

5. Quantum Operations Formalism

All knowledge degenerates into probability; and this probability is greater or less, according to our experience of the veracity or deceitfulness of our understanding, and according to the simplicity or intricacy of the question.

– David Hume, in “Of scepticism with regard to reason”.

The correctness of the theory is judged by the degree of agreement between the conclusions of the theory and human experience. This experience, which alone enables us to make inferences about reality, in physics takes the form of experiment and measurement.

– Einstein, Podolsky and Rosen, in their seminal paper “Can quantum-mechanical description of physical reality be considered complete?”.

In reality, all systems will interact with their environments and there are no perfectly closed systems. As a result, unwanted interactions with the external world show up as noise in quantum systems.

The mathematical formalism of quantum operations is an important and widely applicable tool to describe the dynamics of open quantum systems, either weakly or strongly coupled to their environments. Also called Kraus operator formalism and operator-sum representation, it is very useful in applications to quantum computation and quantum information, to describe discrete state changes *without* explicit time-dependence. This discrete-time analysis is rather different from the usual differential equation description of state dynamics, such as ‘master equations’ which we will study in the next chapter.

5.1 Noise and Markov Processes in Classical Systems

Suppose we have a classical bit, either 0 or 1. The physical realization of a bit of information could be aligned magnetic domains in a hard disk, or electrical charge held in a transistor of a solid state drive. These physical systems interact with the environment, where stray magnetic fields could cause misalignments of a magnetic domain and stray electric fields could cause electrical charge movement. The unwanted physical interactions could cause a bit to flip from a

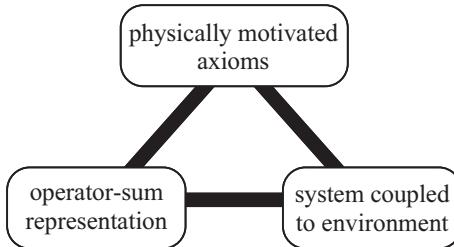


Figure 5.1: Three approaches to quantum operations which are equivalent, but offer different advantages depending upon the intended application. System-environment coupling: studying dynamics as the result of an interaction between a system and an environment is close to the way quantum physics is presented, in terms of Hamiltonians and unitary evolution. It is concrete but not mathematically convenient, especially when the environment is very large. You have seen problems of this nature already, but for tiny environments like a single qubit, or a single harmonic oscillator. In this chapter, we study the operator-sum representation approach, which is completely equivalent to the first approach. It provides a powerful but abstract mathematical representation that is convenient to work with for both numerical and analytical calculations. The third and also equivalent approach to studying open quantum system dynamics is via a set of physically motivated axioms that we would expect a dynamical map in quantum mechanics to satisfy. The advantage of this approach is that it is exceedingly general – it shows that quantum dynamics will be described by quantum operations under an amazingly wide range of circumstances. However, it does not offer the calculational convenience of the second approach, nor the concrete nature of the first. Figure from Nielsen and Chuang.

0 to 1 and vice versa in a hard disk and solid state drive, respectively.

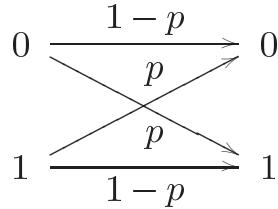


Figure 5.2: Bit flip with probability p . Figure from Nielsen and Chuang.

Usually, the exact physical model of the system-environment interaction is unknown, difficult to establish, or not of interest. In such cases, noise can still be studied by modelling its effects. In the case of the classical bit flips, *transition probabilities* can succinctly summarize the changes in the system.

Let us take the following probabilities:

$$\begin{cases} p_0 : \text{ initial probability that the bit is in the state 0.} \\ p_1 : \text{ initial probability that the bit is in the state 1.} \end{cases} \quad (5.1)$$

$$\begin{cases} q_0 : \text{ final probability that the bit is in the state 0, after noise has occurred.} \\ q_1 : \text{ final probability that the bit is in the state 1, after noise has occurred.} \end{cases} \quad (5.2)$$

Let X be the initial state of the bit, and Y the final state of the bit. Then the law of total probability states that

$$P(Y = y) = \sum_x P(Y = y|X = x)P(X = x). \quad (5.3)$$

(Capital letters denote the state; small letters denote the random variable belonging to that state.) The first term on the RHS is a conditional probability, which reads “the probability

that the final state of the bit is equal to y , given that the initial state is x ." These conditional probabilities are called *transition probabilities*. Suppose $y = 0$ or 1 , Eq. (5.3) can be written as

$$\underbrace{P(Y=0)}_{q_0} = \underbrace{P(Y=0|X=0)P(X=0)}_{(1-p)p_0} + \underbrace{P(Y=0|X=1)P(X=1)}_{(p)p_1}, \quad (5.4)$$

$$\underbrace{P(Y=1)}_{q_1} = \underbrace{P(Y=1|X=0)P(X=0)}_{(p)p_0} + \underbrace{P(Y=1|X=1)P(X=1)}_{(1-p)p_1}, \quad (5.5)$$

or in matrix form,

$$\begin{pmatrix} q_0 \\ q_1 \end{pmatrix} = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}, \quad (5.6)$$

where the 2×2 matrix is the transition probability matrix.

In the above example, we have assumed a simple error model where an initial state X goes to a final state Y . We can extend this idea to ask about the states after that, i.e.

$$X \rightarrow Y \rightarrow Z \dots$$

If the process $Y \rightarrow Z$ only depends on the state Y and is independent of the earlier states, e.g. X and earlier, then we call this a **Markov process**. This assumption that consecutive noise processes act independently is a physically reasonable assumption to make in many situations. Sometimes it is described as being a *memory-less* process.

Noise in classical systems can be described using the theory of stochastic processes, and often with good reason to model noise as Markovian. For a single stage process, the output probabilities \vec{q} are related to the input probabilities \vec{p} by E , the transition probabilities matrix, linearly relating the initial and final states. This is the extension of Eq. (5.6),

$$\vec{q} = E\vec{p}. \quad (5.7)$$

Important considerations for this description of how probabilities change are the following. We require that if \vec{p} is a valid probability distribution, then $E\vec{p}$ must also be a valid probability distribution. Satisfying this condition turns out to be equivalent to two conditions on E :

1. **Positivity requirement:** all the entries of E must be non-negative.
If not, it would be possible to obtain negative probabilities in $E\vec{p}$.
2. **Completeness requirement:** all the columns of E must sum to one.
If not, it would be possible to obtain invalid probability distributions in $E\vec{p}$.

■ **Example 5.1** Consider the completeness requirement. Suppose the transition probability matrix is given by

$$E = \begin{pmatrix} 1-p & p \\ p' & 1-p \end{pmatrix}, \quad (5.8)$$

where p' is not necessarily equal to p . Compute $E\vec{p}$ and explain its implications. ■

For our description of quantum noise, we replace probability distributions with density matrices, and ask, what properties the transition matrix should possess. We answer that in the next section.

5.2 Operator-Sum Representation

Consider a system \mathcal{S} and its environment \mathcal{E} , initially in a separable, uncorrelated state,¹

$$\hat{\rho}(0) = \hat{\rho}_S(0) \otimes \hat{\rho}_E(0). \quad (5.9)$$

¹In general this is not true. Quantum systems interact constantly with their environments and build up correlations, e.g. by exchanging heat with the environment. However, in cases of practical interest, it can be reasonable to assume no correlations between the system and its environment, e.g. when a quantum system is prepared in a specified state in an experiment, prior correlations between that system and the environment are completely ‘reset’, leaving the system in a pure state.

We can then write down the diagonal decomposition of the initial density matrix of the environment which, to be general, may be mixed,

$$\hat{\rho}_E(0) = \sum_i p_i |e_i\rangle\langle e_i|. \quad (5.10)$$

Here $\sum_i p_i = 1$, and environment states $|e_i\rangle$ form an orthonormal basis set of the Hilbert space \mathcal{H}_E of the environment. Given the total Hamiltonian \hat{H} (assumed for simplicity to be time-independent here), the time evolution of the combined system-environment state is given by $\hat{U} = e^{-i\hat{H}t/\hbar}$. The system density matrix at any time is the reduced trace of the combined density matrix at that time, given as

$$\begin{aligned} \hat{\rho}_S(t) &= \text{Tr}_E \left\{ \hat{U}(t) \hat{\rho}(0) \hat{U}^\dagger(t) \right\} \\ &= \text{Tr}_E \left\{ \hat{U}(t) \left[\hat{\rho}_S(0) \otimes \left(\sum_i p_i |e_i\rangle\langle e_i| \right) \right] \hat{U}^\dagger(t) \right\}. \end{aligned} \quad (5.11)$$

Now, taking the trace in the orthonormal basis of environment states $|e_i\rangle$, we get the operator-sum representation,

$$\hat{\rho}_S(t) = \sum_j \langle e_j | \hat{U}(t) \left[\hat{\rho}_S(0) \otimes \left(\sum_i p_i |e_i\rangle\langle e_i| \right) \right] \hat{U}^\dagger(t) |e_j\rangle \quad (5.12)$$

$$\begin{aligned} &= \sum_{i,j} \underbrace{\sqrt{p_i} \langle e_j | \hat{U}(t) | e_i \rangle}_{\equiv \hat{E}_{ij}} \hat{\rho}_S(0) \underbrace{\sqrt{p_i} \langle e_i | \hat{U}^\dagger(t) | e_j \rangle}_{\equiv \hat{E}_{ij}^\dagger} \\ &= \sum_{i,j} \hat{E}_{ij} \hat{\rho}_S(0) \hat{E}_{ij}^\dagger, \end{aligned} \quad (5.13)$$

where the Kraus operators have been defined in the last line for a mixed environment initial state,

$$\hat{E}_{ij} \equiv \sqrt{p_i} \langle e_j | \hat{U}(t) | e_i \rangle. \quad (5.14)$$

The Kraus operators act on the Hilbert space \mathcal{H}_S of the system.² The Kraus operators contain all the available information about the initial state of the environment and about the dynamics of the combined system-environment state. Thus they entirely encapsulate the effect of the environment on the reduced density matrix of the system.

The Kraus operators are **not unique**; they depend on the particular choice of the basis $\{|e_i\rangle\}$ used in performing the trace. Apart from the choice of basis, they are uniquely determined by the initial state of the environment \mathcal{E} and the Hamiltonian of the joint \mathcal{SE} system.

The evolution of the joint \mathcal{SE} system is unitary and the Kraus operators satisfy the *completeness* constraint,

$$\sum_{i,j} \hat{E}_{ij}^\dagger \hat{E}_{ij} = \hat{\mathbb{I}}_S, \quad (5.15)$$

where the subscript S on the right-hand-side denote that the identity is over the system Hilbert space.

■ **Example 5.2** Show that Eq. (5.15) is true. ■

²Note that $\hat{U}(t)$ is a unitary operator acting on the combined system-environment space. After sandwiching between the bra-kets of the environment, it becomes \hat{E}_{ij} which is still an operator but acting on the system subspace. The subscripts ij should be thought of as the index of the operator, rather than the matrix element of the system operator. Also, how can \hat{U} be specified if the environment has nearly infinite degrees of freedom? It turns out that in order for this model to properly describe any possible transformation $\hat{\rho} \rightarrow S(\hat{\rho})$, if the principal system has a Hilbert space of d dimensions, then it suffices to model the environment as being in a Hilbert space of no more than d^2 dimensions. It also turns out not to be necessary for the environment to start out in a mixed state; a pure state will do.

5.2.1 Pure initial environment state

To write down the quantum operation in a closed form, without loss of generality, we *assume the initial environment state to be in a pure state*,

$$\hat{\rho}_E = |e_0\rangle\langle e_0|. \quad (5.16)$$

Even when $\hat{\rho}_E$ is in a mixed state, we are always able to complement \mathcal{H}_E by enlarging the Hilbert space, so that the environment state in the enlarged Hilbert space is pure. This procedure is known as *purification*³, which we will come to later. With this assumption, we have

Kraus Operator

$$\hat{E}_\alpha : \mathcal{H}_S \longrightarrow \mathcal{H}_S, \quad \hat{E}_\alpha := \langle e_\alpha | \hat{U} | e_0 \rangle. \quad (5.17)$$

We may then write the quantum operation \mathcal{E} that is a linear map which maps a density matrix to a density matrix in the same space:

Kraus Operator-Sum Representation

$$\mathcal{E}(\hat{\rho}_S) = \hat{\rho}'_S = \sum_{\alpha} \hat{E}_{\alpha}(t) \hat{\rho}_S \hat{E}_{\alpha}^{\dagger}(t). \quad (5.18)$$

The properties of the system density matrix are preserved by the map:

1. $\hat{\rho}_S(t)$ is **Hermitian**:

$$\hat{\rho}_S^{\dagger}(t) = \left(\sum_{\alpha} \hat{E}_{\alpha} \hat{\rho}_S(0) \hat{E}_{\alpha}^{\dagger} \right)^{\dagger} = \sum_{\alpha} \hat{E}_{\alpha} \hat{\rho}_S^{\dagger}(0) \hat{E}_{\alpha}^{\dagger} = \hat{\rho}_S(t).$$

2. $\hat{\rho}_S(t)$ has **unit trace**:

$$\begin{aligned} \text{Tr} \hat{\rho}_S(t) &= \text{Tr} \left(\sum_{\alpha} \hat{E}_{\alpha} \hat{\rho}_S(0) \hat{E}_{\alpha}^{\dagger} \right) = \sum_{\alpha} \text{Tr} \left(\hat{E}_{\alpha} \hat{\rho}_S(0) \hat{E}_{\alpha}^{\dagger} \right) = \sum_{\alpha} \text{Tr} \left(\hat{\rho}_S(0) \hat{E}_{\alpha}^{\dagger} \hat{E}_{\alpha} \right) \\ &= \text{Tr} \left(\hat{\rho}_S(0) \sum_{\alpha} \hat{E}_{\alpha}^{\dagger} \hat{E}_{\alpha} \right) = \text{Tr} (\hat{\rho}_S(0) \mathbb{I}) = 1. \end{aligned}$$

3. $\hat{\rho}_S(t)$ is **non-negative (also termed positive semi-definite)**, i.e. its eigenvalues, which represent probabilities, are zero or positive.

Consider an arbitrary state $|\phi\rangle$. Then

$\langle \phi | \hat{\rho}_S | \phi \rangle = \langle \phi | (\sum_i p_i |\psi_i\rangle\langle\psi_i|) | \phi \rangle = \sum_i p_i \langle \phi | \psi_i \rangle \langle \psi_i | \phi \rangle = \sum_i p_i |\langle \phi | \psi_i \rangle|^2 \geq 0$, where the last inequality is because $p_i \geq 0$. This shows that the original density matrix $\hat{\rho}_S$ is non-negative.

Now the density matrix after transformation,

$\langle \phi | \hat{\rho}_S(t) | \phi \rangle = \langle \phi | (\sum_{\alpha} \hat{E}_{\alpha} \hat{\rho}_S \hat{E}_{\alpha}^{\dagger}) | \phi \rangle = \sum_{\alpha} \langle \phi | \hat{E}_{\alpha} \hat{\rho}_S \hat{E}_{\alpha}^{\dagger} | \phi \rangle = \sum_{\alpha} \langle \phi_{\alpha} | \hat{\rho}_S | \phi_{\alpha} \rangle \geq 0$ for any $|\phi\rangle$, where we defined $|\phi_{\alpha}\rangle = \hat{E}_{\alpha}^{\dagger} |\phi\rangle$.

Finally, if the set $\{\hat{E}_{\alpha}\}$ has only one operator, then the map reduces to the unitary evolution of the density matrix $\hat{\rho} \rightarrow \hat{U} \hat{\rho} \hat{U}^{\dagger}$. In that case, a pure state would remain pure. If that is not the case, then the evolution is not unitary, which means that in the course of the evolution the system and environment became entangled, so the system is in a mixed state after partial trace. Because of the loss of unitarity, the map is in general not invertible and thus there is a specific arrow of time.

5.2.2 Purification

We mentioned in Chapter 5.2.1 that starting from a pure environment state for the operator-sum representation does not lose generality. Here, we give the purification procedure which states

³See e.g. Chap 2.5 of Nielsen and Chuang

that any mixed environment state can always be purified by going into a larger environment space.

Suppose we are given a mixed state $\hat{\rho}_A = \sum_i p_i |i_A\rangle\langle i_A|$ of a quantum system A . It is possible to introduce another system, which we denote as R , and define a pure state for the composite system which we denote by $|AR\rangle \equiv \sum_i \sqrt{p_i} |i_A\rangle \otimes |i_R\rangle$. This decomposition is allowed by the Schmidt decomposition theorem. Then,

$$\begin{aligned}\text{Tr}_R(|AR\rangle\langle AR|) &= \sum_{i,j} \sqrt{p_i p_j} |i_A\rangle\langle j_A| \underbrace{\text{Tr}(|i_R\rangle\langle j_R|)}_{\delta_{ij}} \\ &= \sum_i p_i |i_A\rangle\langle i_A| = \hat{\rho}_A.\end{aligned}\quad (5.19)$$

That is, the pure state $|AR\rangle$ reduces to $\hat{\rho}_A$ when we look at system A alone. This is a purely mathematical procedure, known as purification, which allows us to associate pure states with mixed states. For this reason we call system R a reference system: it is a fictitious system, without a direct physical significance.

5.3 Completely Positive and Trace Preserving (CPTP) Maps

In a more abstract sense, we can define a quantum superoperator⁴ describing the time evolution law for density matrices as a map

$$\mathcal{E} : \hat{\rho} \longrightarrow \hat{\rho}', \quad (5.20)$$

with the following properties.

1. **Linearity.**

Although a non-linear map could also always map a density matrix to another density matrix, if we impose linearity we retain the ensemble interpretation of the density matrix.

2. **Trace preserving.**

The trace of a density matrix represents the sum of the probabilities of all possible states in the ensemble, so it is important that the trace be preserved. An exception can be made for operators that describe measurement (and not time evolution) – these operators do not preserve trace.

3. **Hermiticity preserving.**

Related to trace preservation, the eigenvalues of the density matrix represent probabilities and therefore must be real. If the map preserves the Hermitian property of the density matrix, this is satisfied.

4. **Complete Positivity.**

Positivity means that the map is such that $\mathcal{E}(\hat{\rho})$ is non-negative if $\hat{\rho}$ is non-negative. Although this condition is enough to obtain a valid density matrix, it leads to a contradiction when we consider composite systems. For example, consider the map that is the partial transpose on the system of one qubit: $\hat{\rho} \longrightarrow \hat{\rho}^{PT}$. For a qubit system entangled with an environment of one ancilla qubit, such that the composite system is in the Bell state $(|0_S 0_A\rangle + |1_S 1_A\rangle)/\sqrt{2}$, the density operator is

$$\hat{\rho} = \frac{1}{2}(|0_S 0_A\rangle\langle 0_S 0_A| + |0_S 0_A\rangle\langle 1_S 1_A| + |1_S 1_A\rangle\langle 0_S 0_A| + |1_S 1_A\rangle\langle 1_S 1_A|) \quad (5.21)$$

and the partial transpose on the system qubit gives

$$\hat{\rho}^{PT} = \frac{1}{2}(|0_S 0_A\rangle\langle 0_S 0_A| + |1_S 0_A\rangle\langle 0_S 1_A| + |0_S 1_A\rangle\langle 1_S 0_A| + |1_S 1_A\rangle\langle 1_S 1_A|). \quad (5.22)$$

Although $\hat{\rho}$ has all positive eigenvalues, $\hat{\rho}^{PT}$ has *one negative eigenvalue!* Therefore, the condition has to be strengthened to complete positivity, which is the condition that maps must not only be positive for the system S that they act on, but also positive on larger systems that include S as a subsystem. When this is the case, we call the map completely positive.

■ **Example 5.3** Verify that the density matrix in Eq. (5.22) has one negative eigenvalue $-1/2$ and three degenerate eigenvalues $1/2$. ■

⁴An operator which acts on an operator.

5.4 Kraus Representation Theorem

Kraus Representation Theorem

Any operator $\hat{\rho} \rightarrow \mathcal{E}(\hat{\rho})$ in a space of dimensions d^2 that obeys the properties of linearity, trace preservation, Hermiticity preservation, and complete positivity, can be written in the form

$$\mathcal{E}(\hat{\rho}) = \sum_{\alpha=1}^K \hat{E}_\alpha \hat{\rho} \hat{E}_\alpha^\dagger, \quad \text{with} \quad \sum_{\alpha=1}^K \hat{E}_\alpha^\dagger \hat{E}_\alpha = \mathbb{I}, \quad (5.23)$$

where $K \leq d^2$ is the Kraus number and d is the dimension of the system.

We first started from the evolution of a composite system and environment, then tracing out the environment degrees of freedom, and arrived at a description of the open system evolution via the operator sum representation in Section 5.2. Then in Section 5.3, we wrote an abstract set of properties of the linear map describing the open system evolution in order to arrive at an acceptable, physical evolution (where the density operator still possess the characteristics of a density operator after evolution). The Kraus representation theorem says that these two descriptions are equivalent.

■ **Example 5.4** (From Nielsen and Chuang:) Consider an example where the system and the environment are single qubits. The environment starts off in the state $|0_E\rangle$. The interaction between the system and the environment after some time t , is described by the unitary

$$\hat{U}_{SE} = \frac{1}{\sqrt{2}} (\hat{\sigma}_x \otimes \mathbb{I} + \hat{\sigma}_y \otimes \hat{\sigma}_x). \quad (5.24)$$

Compute the Kraus operators for such a process. ■

5.4.1 Unitary freedom in Kraus operators

The Kraus operators that appear in the operator-sum representation are not unique. Two representations are equivalent if the Kraus operators between the representations are related by a unitary transformation. That is, $\sum_{k=1}^K \hat{A}_k \hat{\rho} \hat{A}_k^\dagger$ representing some process could also be written as $\sum_{k=1}^K \hat{B}_k \hat{\rho} \hat{B}_k^\dagger$ if

$$B_i = \sum_{j=1}^K u_{ij} A_j, \quad (5.25)$$

where u_{ij} are elements of a unitary matrix. We can show the equivalence explicitly.

$$\begin{aligned} \sum_{k=1}^K \hat{B}_k \hat{\rho} \hat{B}_k^\dagger &= \sum_{k=1}^K \left(\sum_{j=1}^K u_{kj} A_j \right) \hat{\rho} \left(\sum_{i=1}^K u_{ik}^* A_i^\dagger \right) \\ &= \sum_{i,j} \underbrace{\left(\sum_k u_{ik}^* u_{kj} \right)}_{\delta_{ij}} A_j \hat{\rho} A_i^\dagger \\ &= \sum_{j=1}^K A_j \hat{\rho} A_j^\dagger. \end{aligned} \quad (5.26)$$

This unitary freedom comes from the trace over the environment in Eq. (5.12). In this step, the trace was taken over an environment $|e_i\rangle$. However, there is nothing that distinguishes this basis from another basis of the environment, say $|f_i\rangle$ that is related to the first by a unitary transformation. The unitary freedom of the environment basis is reason for the unitary freedom of the Kraus operators.

■ **Example 5.5** To illustrate the non-uniqueness of the Kraus operators, let us re-visit Example 5.4. Instead of measuring the environment qubit in the $|0_E\rangle, |1_E\rangle$ basis, let us measure in the $|+_E\rangle, |-_E\rangle$ basis. Complete the calculation of the Kraus operators.

■

5.5 Noisy Quantum Channels

The operator-sum approach neatly represents the effect of the environment as a sequence of (in general non-unitary) transformations of the reduced density matrix generated by the Kraus operators \hat{E}_α . This approach provides a compact and transparent framework for the formal representation of the evolution of the reduced density matrix, either over time or through space such as a communication channel. For the latter reason, the evolution or map of the reduced system density matrix is called a **quantum channel**. We can think of a noisy quantum channel \mathcal{E} as describing the evolution of quantum information (described by a density matrix $\hat{\rho}$) transmitted with some loss of fidelity from a sender to a receiver.

The task of computing the Kraus operators corresponds to diagonalizing the full system-environment Hamiltonian, and is therefore limited to small environments. However, if we can describe the dynamics of the system in terms of the operator sum representation, based on some *effective model* of the environment, then it is possible to find a larger, composite system that evolves unitarily and yields the operator sum upon partial trace. The ancillary system might not have all the characteristics of the unknown physical environment, which is unnecessary, because *what we are looking for is a minimal model for the environment*. This approach turns out to be very useful in quantum information. In this section we will explore simple quantum channels by giving them a physical interpretation.

5.5.1 Phase damping or dephasing channel

Phase damping describes a process where the system interacts with a large environment composed of many small subsystems. The interaction of the system with each of the environment subsystems is weak compared to the system energy, but strong compared to the subsystem energy. Therefore the system is unchanged, while the environment subsystem is changed. Since there will be many of these interactions with the environment subsystem, their combined action does have an effect on the system. However, it will not be enough to change the system energy.

A physical example is the interaction of a dust particle with photons in the environment is such that scattering of the photon does not change the particle state but changes the state of the photon with probability p . Furthermore, the photon acquires more energy in the scattering if the particle was in the excited state compared to the ground state.

Assuming the environment starts in some abstract pure state $|0_E\rangle$, this process is given by

$$|0_S 0_E\rangle \longrightarrow \underbrace{\sqrt{1-p} |0_S 0_E\rangle + \sqrt{p} |0_S 1_E\rangle}_{\equiv |A\rangle}, \quad (5.27)$$

$$|1_S 0_E\rangle \longrightarrow \underbrace{\sqrt{1-p} |1_S 0_E\rangle + \sqrt{p} |1_S 2_E\rangle}_{\equiv |B\rangle}, \quad (5.28)$$

where environment states $|1_E\rangle, |2_E\rangle$ are some abstract states with different energy. The full system-environment unitary is given by

$$\hat{U}_{SE} = |A\rangle \langle 0_S 0_E| + |B\rangle \langle 1_S 0_E|. \quad (5.29)$$

The Kraus operators are given by

$$\begin{cases} \hat{E}_0 &= \langle 0_E | \hat{U}_{SE} | 0_E \rangle = \sqrt{1-p} |0_S\rangle \langle 0_S| + \sqrt{1-p} |1_S\rangle \langle 1_S| = \sqrt{1-p} \mathbb{I}_S \\ \hat{E}_1 &= \langle 1_E | \hat{U}_{SE} | 0_E \rangle = \sqrt{p} |0_S\rangle \langle 0_S| \\ \hat{E}_2 &= \langle 2_E | \hat{U}_{SE} | 0_E \rangle = \sqrt{p} |1_S\rangle \langle 1_S| \end{cases} \quad (5.30)$$

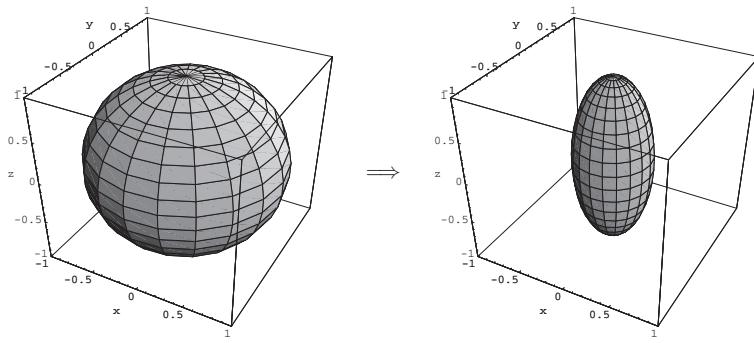


Figure 5.3: Effect of phase damping on the Bloch sphere. Transverse (x, y) components of the Bloch sphere are reduced over time, while the longitudinal (z) component is unaffected. Figure from Nielsen and Chuang.

The state evolution is then

$$\begin{aligned}\mathcal{E}(\hat{\rho}) &= \sum_{\alpha=0}^2 \hat{E}_\alpha \hat{\rho} \hat{E}_\alpha^\dagger \\ &= (1-p)\hat{\rho} + p|0_S\rangle\langle 0_S|\hat{\rho}|0_S\rangle\langle 0_S| + p|1_S\rangle\langle 1_S|\hat{\rho}|1_S\rangle\langle 1_S| \\ &= \begin{pmatrix} \rho_{00} & (1-p)\rho_{01} \\ (1-p)\rho_{10} & \rho_{11} \end{pmatrix}. \end{aligned} \quad (5.31)$$

Consider the Bloch vector: $(n_x, n_y, n_z)^T = ((1-p)n_x, (1-p)n_y, n_z)^T$. Assume $p = \Gamma\Delta t$ is the probability of one such phase damping event during a short time interval Δt . For n events in a time $t = n\Delta t$, then $(1-p)^n = (1 - \Gamma t/n)^n \approx e^{-\Gamma t}$. Therefore, the explicit time dependence in such a map is

$$\mathcal{E}(\hat{\rho}) = \begin{pmatrix} \rho_{00} & \rho_{01}e^{-\Gamma t} \\ \rho_{10}e^{-\Gamma t} & \rho_{11} \end{pmatrix}. \quad (5.32)$$

The phase damping channel leads to the decay of the off-diagonal term in the density matrix and the state progressively evolves to a mixed state. Because of the loss of phase coherence in the off-diagonal term, this process is called dephasing and the time constant $1/\Gamma \equiv T_2$ is called the dephasing time. We use subscript “2” in the dephasing time by convention, in line with nuclear magnetic resonance (NMR) terminology.

One can also check that the completeness relation is satisfied for the phase damping channel,

$$\hat{E}_0^\dagger \hat{E}_0 + \hat{E}_1^\dagger \hat{E}_1 + \hat{E}_2^\dagger \hat{E}_2 = (1-p)\mathbb{I} + p|0\rangle_S\langle 0|_S + p|1\rangle_S\langle 1|_S = \mathbb{I}. \quad (5.33)$$

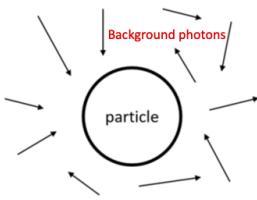


Figure 5.4: An example of a physical process described by phase damping. From Preskill (see Box, next page): A heavy “classical” particle (e.g., an interstellar dust grain) interacting with a background gas of light particles (e.g., the 3K microwave photons). We can imagine that the dust is initially prepared in a superposition of position eigenstates $|\psi\rangle = (|x\rangle + | -x \rangle)/\sqrt{2}$. We might be able to monitor the behaviour of the dust particle, but it is hopeless to keep track of the quantum state of all the photons that scatter from the particle; for our purposes, the quantum state of the particle is described by the density matrix obtained by tracing over the photon degrees of freedom.

Connection to Decoherence (from John Preskill’s notes)

Our analysis of the phase damping channel indicates that if photons are scattered by the dust particle at a rate Γ , then the off-diagonal terms decay as $\exp(-\Gamma t)$ and so become completely negligible for $t \gg \Gamma^{-1}$. At that point, the coherence of the superposition of position eigenstates is completely lost – there is no chance that we can recombine the wavepackets and induce them to interfere. If we attempt to do a double-slit interference experiment with dust grains, we will not see any interference pattern if it takes a time $t \gg \Gamma^{-1}$ for the grain to travel from the source to the screen.

The dust grain is heavy. Because of its large inertia, its state of motion is little affected by the scattered photons. Thus, there are two disparate time scales relevant to its dynamics. On the one hand, there is a damping time scale, the time for a significant amount of the particle’s momentum to be transferred to the photons, which is a long time for such a heavy particle. On the other hand, there is the decoherence time scale. In this model, the time scale for decoherence is of order $1/\Gamma$, the time for a single photon to be scattered by the dust grain, which is far shorter than the damping time scale. **For a macroscopic object, decoherence is fast.**

The preceding discussion tells us that **the phase-damping channel picks out a preferred basis for decoherence**, which in our “interpretation” we have assumed to be the position-eigenstate basis. Physically, decoherence prefers the spatially localized states of the dust grain because the interactions of photons and grains are localized in space. Grains in distinguishable positions tend to scatter the photons of the environment into mutually orthogonal states.

Even if the separation between the “grains” were so small that it could not be resolved very well by the scattered photons (i.e. the environmental photon states scattered off at the $+x$ and $-x$ locations are not fully orthogonal, $\langle E_+ | E_- \rangle = \epsilon \neq 0$), the decoherence process would still work in a similar way but at a slower rate $\Gamma_{\text{decoherence}} = \epsilon \Gamma_{\text{scattering}}$. The intuition we distill from this simple model applies to a wide variety of physical situations. A coherent superposition of macroscopically distinguishable states of a “heavy” object decoheres very rapidly compared to its damping rate. The spatial locality of the interactions of the system with its environment gives rise to a preferred “local” basis for decoherence. The same principle applies to Schrödinger’s unfortunate cat, delicately prepared in a coherent superposition of its dead state and its alive state, two states that are easily distinguished by spatially localized probes. The cat quickly interacts with its environment, which is “scattered” into one of two mutually orthogonal states perfectly correlated with the cat’s state in the $\{| \text{dead} \rangle, | \text{alive} \rangle\}$ basis, thus transforming the cat into an incoherent mixture of those two basis states.

5.5.2 Amplitude damping channel (energy dissipation)

The amplitude damping process or quantum operation is the description of energy dissipation. Here, our model is a two-level atom (system), which is stable in its ground state but decays due to spontaneous emission of a photon to the environment when in its excited state. By detecting the photon (“measuring the environment”), we can get information about the state of the atom.

We denote the atomic ground state by $|0\rangle_S$ and excited state by $|1\rangle_S$. The environment is the electromagnetic field, assumed to be in its vacuum state (0 photons) initially, $|0\rangle_E$. After a fixed interval of time, the probability that the excited state has decayed to the ground state, and a photon emitted, is p . When this happens the environment will be in the state of 1 photon, denoted by $|1\rangle_E$. The probability that the excited atom has not decayed, is $1 - p$. Naturally, if the atom was in the ground state, it remains in the ground state. The process is denoted by

$$|0_S 0_E\rangle \longrightarrow |0_S 0_E\rangle, \quad (5.34)$$

$$|1_S 0_E\rangle \longrightarrow \sqrt{p} |0_S 1_E\rangle + \sqrt{1-p} |1_S 0_E\rangle. \quad (5.35)$$

Therefore, the system-environment unitary dynamics is given by

$$U_{SE} = |0_S 0_E\rangle \langle 0_S 0_E| + \sqrt{p} |0_S 1_E\rangle \langle 1_S 0_E| + \sqrt{1-p} |1_S 0_E\rangle \langle 1_S 0_E|. \quad (5.36)$$

We find the Kraus operators from the definition in Eq. (5.17),

$$\hat{E}_0 \equiv \langle 0_E | U_{SE} | 0_E \rangle = |0_S\rangle \langle 0_S| + \sqrt{1-p} |1_S\rangle \langle 1_S| = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \quad (5.37)$$

$$\hat{E}_1 \equiv \langle 1_E | U_{SE} | 0_E \rangle = \sqrt{p} |0_S\rangle \langle 1_S| = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}. \quad (5.38)$$

In the final equality of each line, we have written the matrix form of the Kraus operators in the atomic basis $\{|0\rangle_S, |1\rangle_S\}$. \hat{E}_0 represents the situation if no decay occurs; \hat{E}_1 represents the situation if a decay from excited to ground state occurs. The density matrix of the system evolves according to Eq. (5.23),

$$\mathcal{E}(\hat{\rho}) = \hat{E}_0 \hat{\rho} \hat{E}_0^\dagger + \hat{E}_1 \hat{\rho} \hat{E}_1^\dagger = \begin{pmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\rho_{01} \\ \sqrt{1-p}\rho_{01} & (1-p)\rho_{11} \end{pmatrix}. \quad (5.39)$$

If the probability of a transition in a time interval Δt is $p = \Gamma \Delta t$, and the channel is applied sequentially $n = t/\Delta t$ times, then

$$\begin{cases} \rho_{11} & \longrightarrow \rho_{11}(1-p)^n = \rho_{11}(1-\Gamma \frac{t}{n})^n \approx \rho_{11} e^{-\Gamma t} \\ \rho_{01} & \longrightarrow \rho_{01}(1-p)^{n/2} = \rho_{01}(1-\Gamma \frac{t}{n})^{n/2} \approx \rho_{01} e^{-\Gamma t/2} \end{cases} \quad (5.40)$$

When $t \rightarrow \infty$,

$$\mathcal{E}(\hat{\rho}) = \begin{pmatrix} \rho_{00} + \rho_{11} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (5.41)$$

which means that the atom eventually decays into the ground state, a reasonable result. It also means that if the atom started in a mixed state, it ends up in a pure state eventually. That is, measurement of the environment prepares a pure state of the atom, in effect preventing the atom from decohering.

Relaxation and dephasing times

- The characteristic decay time $T_1 \equiv 1/\Gamma$ is called the relaxation time.
- For the amplitude-damping channel, the decoherence time T_2 given by the decay of the off-diagonal element, is related and comparable to the relaxation time:

$$T_2 = 2/\Gamma = 2T_1.$$

- Note that this model is not universal – there are many qubit systems for which dephasing times are very much shorter than relaxation times.

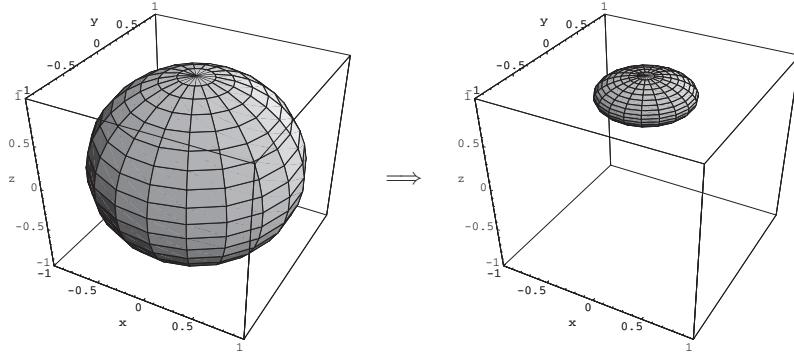


Figure 5.5: Effect of amplitude damping channel. The Bloch sphere shrinks towards the north pole. Figure from Nielsen and Chuang.

Finally, checking the completeness relation, we find that indeed

$$\hat{E}_0^\dagger \hat{E}_0 + \hat{E}_1^\dagger \hat{E}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & p \end{pmatrix} = \mathbb{I}. \quad (5.42)$$

5.5.3 Depolarizing channel

The depolarizing channel is a model of a qubit undergoing decoherence where, with some probability an “error” occurs. The error can be any one of these three types, where each type of error is equally likely. Let us use probability $p/4$ for each of these errors (this is arbitrary, and only for getting a nicer algebraic result at the end).

1. X error (Bit-flip) : $|\psi\rangle_S \rightarrow \hat{\sigma}_x |\psi\rangle_S$, with probability $p/4$.
2. Z error (Phase-flip) : $|\psi\rangle_S \rightarrow \hat{\sigma}_z |\psi\rangle_S$, with probability $p/4$.
3. Y error (Combination of both): $|\psi\rangle_S \rightarrow \hat{\sigma}_y |\psi\rangle_S$, with probability $p/4$.

This is a combination of both bit-flip and phase-flip because $\hat{\sigma}_y = i\hat{\sigma}_x\hat{\sigma}_z$.

The depolarizing channel can be represented by a unitary operator acting on $\mathcal{H}_{SE} = \mathcal{H}_S \otimes \mathcal{H}_E$, where the environment has dimension 4. Assuming the environment to be initially in state $|0\rangle_E$, the unitary operator \hat{U}_{SE} performs the evolution,

$$\begin{aligned} \hat{U}_{SE} |\psi\rangle_S \otimes |0\rangle_E &\rightarrow \sqrt{p/4} \hat{\sigma}_x |\psi\rangle_S \otimes |1\rangle_E \quad (\text{bit-flip}) \\ &+ \sqrt{p/4} \hat{\sigma}_y |\psi\rangle_S \otimes |2\rangle_E \quad (\text{bit-and-phase-flip}) \\ &+ \sqrt{p/4} \hat{\sigma}_z |\psi\rangle_S \otimes |3\rangle_E \quad (\text{phase-flip}) \\ &+ \sqrt{1-3p/4} |\psi\rangle_S \otimes |0\rangle_E \quad (\text{unchanged}) \end{aligned}$$

The environment evolves to one of four mutually orthogonal states that “record” of what error occurred. If we could measure the environment in the basis $\{|\mu\rangle_E, \mu = 0, 1, 2, 3\}$, we would know what kind of error had occurred (and we would be able to intervene and reverse the error).

The Kraus operators of the depolarizing channel are

$$\hat{E}_0 = \sqrt{1-3p/4} \mathbb{I}, \quad \hat{E}_1 = \sqrt{p/4} \hat{\sigma}_x, \quad \hat{E}_2 = \sqrt{p/4} \hat{\sigma}_y, \quad \hat{E}_3 = \sqrt{p/4} \hat{\sigma}_z. \quad (5.43)$$

The qubit density matrix evolves as

$$\mathcal{E}(\hat{\rho}) = (1-3p/4)\hat{\rho} + \frac{p}{4} (\hat{\sigma}_x \hat{\rho} \hat{\sigma}_x + \hat{\sigma}_y \hat{\rho} \hat{\sigma}_y + \hat{\sigma}_z \hat{\rho} \hat{\sigma}_z) \quad (5.44)$$

$$= \begin{pmatrix} \frac{p}{2} + (1-p)\rho_{00} & (1-p)\rho_{01} \\ (1-p)\rho_{10} & 1 - \frac{p}{2} - (1-p)\rho_{00} \end{pmatrix}, \quad (5.45)$$

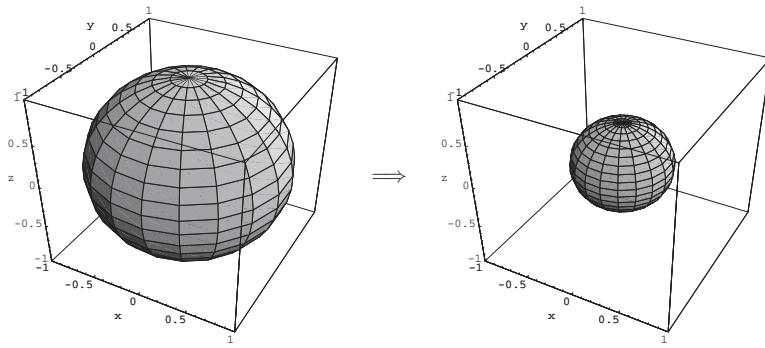


Figure 5.6: Effect of depolarizing channel on the Bloch sphere. Figure from Nielsen and Chuang.

which, after a long time or many state changes becomes a fully mixed state,

$$\hat{\rho} \rightarrow \begin{pmatrix} \frac{p}{2} & 0 \\ 0 & 1 - \frac{p}{2} \end{pmatrix}. \quad (5.46)$$

■ **Example 5.6** For the depolarizing channel, derive the length of the Bloch vector components from the density matrix in Eq. (5.44), in terms of p . Suggest why the “uniform deflation” picture of the Bloch sphere is reasonable.

■

5.6 Quantum Measurement as a Quantum Operation

Recall the measurement postulate of quantum mechanics (Chapter 3.5), which states that a quantum state $\hat{\rho}$ after a measurement by a measurement operator \hat{M}_k , gets mapped as

$$\hat{\rho} \rightarrow \frac{\hat{M}_k \hat{\rho} \hat{M}_k^\dagger}{p_k}, \quad \text{where } p_k = \text{Tr}(\hat{\rho} \hat{M}_k^\dagger \hat{M}_k). \quad (5.47)$$

This implies that a measurement is also a quantum operation, where the density operator after measurement by operator \hat{M}_k is $\mathcal{E}_k(\hat{\rho}) \equiv \hat{M}_k \hat{\rho} \hat{M}_k^\dagger$, so

$$\hat{\rho} \rightarrow \frac{\mathcal{E}_k(\hat{\rho})}{\text{Tr}(\mathcal{E}_k(\hat{\rho}))}. \quad (5.48)$$

■ **Example 5.7** A quantum measurement with outcomes labelled by k is described by a set of measurement operators \hat{M}_k . Show that measurement operators are complete, i.e.

$$\sum_k \hat{M}_k^\dagger \hat{M}_k = \hat{\mathbb{I}}. \quad (5.49)$$

■

The case of projective measurements, presented in Chapter 1.2.1, corresponds to taking $\hat{M}_k = |k\rangle \langle k|$ to be projection operators onto some basis $\{|k\rangle\}$. In such a case, Eqs. (5.49), (5.48) then reduce to

$$\sum_k |k\rangle \langle k| = \hat{\mathbb{I}}, \quad (5.50)$$

$$p_k = \text{Tr}(|k\rangle \langle k| \hat{\rho} |k\rangle \langle k|) = \langle k | \hat{\rho} | k \rangle, \quad (5.51)$$

$$\hat{\rho} \rightarrow \frac{|k\rangle \langle k| \hat{\rho} |k\rangle \langle k|}{\langle k | \hat{\rho} | k \rangle} = |k\rangle \langle k|, \quad (5.52)$$

which is the usual projective measurement or “full collapse” scenario. But now, suppose one the measurement is performed but the outcome is not recorded. The state has changed, but we don’t

PH3406 (2024) T.S.KOH

know exactly which k -th outcome occurred. The best guess we can make to the new state of the system is *a statistical mixture of the possible outcomes, each weighted with probability p_k* , giving us

$$\hat{\rho}' = \sum_k p_k \frac{\hat{M}_k \hat{\rho} \hat{M}_k^\dagger}{p_k} = \sum_k \hat{M}_k \hat{\rho} \hat{M}_k^\dagger. \quad (5.53)$$

Physical Interpretation of Operator Sum Representation

This gives a physical interpretation of quantum operations: they are statistical mixtures of the results of measurements made on a quantum state but not knowing what the measurement outcomes were, i.e. Eq. (5.53).

Seen in this way, noisy quantum channels make sense: the quantum state is measured by the environment but the outcome of the environmental monitoring is unknown. As such, we obtain statistical mixtures of the quantum state, given by the operator-sum representation of Eq. (5.18).

5.7 Quantum State Tomography

Quantum operations provide a wonderful mathematical model for open quantum systems, and are conveniently visualized for qubits. In this section, we relate the formalism to experimentally measurable quantities and describe the measurements an experimentalist has to do in order to characterize the dynamics of a quantum system. There are two steps: first, the state of the quantum system has to be characterized, a procedure called quantum state tomography. Following this, characterizing the dynamics of the quantum system, called quantum process tomography, can be done.

5.7.1 Single Qubit State Tomography

State tomography is the procedure of experimentally determining an unknown quantum state $\hat{\rho}$. If we are given just a single copy of $\hat{\rho}$ then it turns out to be impossible to characterize the state. The basic problem is that there is no quantum measurement which can distinguish non-orthogonal quantum states like $|0\rangle$ and $|+\rangle$ with certainty. However, it is possible to estimate $\hat{\rho}$ if we have a large number of copies of $\hat{\rho}$. For instance, if $\hat{\rho}$ is the quantum state produced by some experiment, then we simply repeat the experiment many times to produce many copies of the state $\hat{\rho}$.

In this section, we will discuss state tomography of a single qubit. We recall that the components of the Bloch vector \vec{v} are

$$v_i = \langle \sigma_i \rangle = \text{Tr}(\hat{\rho} \hat{\sigma}_i). \quad (5.54)$$

Then, when we have many copies of a single qubit density matrix, we can reconstruct the density matrix from measurement statistics,

$$\hat{\rho} = \frac{1}{2} \left(\text{Tr}(\hat{\rho}) \hat{\mathbb{I}} + \text{Tr}(\hat{\rho} \hat{\sigma}_x) \hat{\sigma}_x + \text{Tr}(\hat{\rho} \hat{\sigma}_y) \hat{\sigma}_y + \text{Tr}(\hat{\rho} \hat{\sigma}_z) \hat{\sigma}_z \right). \quad (5.55)$$

Each of the expressions like $\text{Tr}(\hat{\rho} \hat{\sigma}_i)$ are the average value of observable $\hat{\sigma}_i$. In each experiment, when $\hat{\sigma}_i$ is measured, the values are either $+1$ or -1 . The average over a series of N experiments is the sample mean, which provides an estimate for the true value of $\text{Tr}(\hat{\rho} \hat{\sigma}_i)$.

More generally, a projective measurement along an axis m is represented by the observable $\hat{\sigma}_m = \vec{\sigma} \cdot \vec{m}$, and has an expectation value

$$\langle \hat{\sigma}_m \rangle = \text{Tr}(\hat{\sigma}_m \hat{\rho}) = \vec{m} \cdot \vec{v}. \quad (5.56)$$

The probabilities of measuring the qubit in the “+” and “-” states on the \vec{m} axis are

$$p_{\vec{m}+} = \text{Tr}(|\vec{m}+\rangle \langle \vec{m}+| \hat{\rho}) = \frac{1}{2}(1 + \vec{m} \cdot \vec{v}), \quad (5.57)$$

$$p_{\vec{m}-} = \text{Tr}(|\vec{m}-\rangle \langle \vec{m}-| \hat{\rho}) = \frac{1}{2}(1 - \vec{m} \cdot \vec{v}) = 1 - p_{\vec{m}+}. \quad (5.58)$$

If we identically prepare N_m qubits and measure the observable $\hat{\sigma}_m$ on each one, we will find N_{m+} qubits in the $|\vec{m}\rangle$ state and N_{m-} qubits in the $|\vec{m}\rangle$ state, giving an estimate of the expectation value (sample mean)

$$\langle\langle\hat{\sigma}_m\rangle\rangle = \frac{N_{m+} - N_{m-}}{N_{m+} + N_{m-}} = \frac{N_{m+} - N_{m-}}{N_m}. \quad (5.59)$$

A statistical estimate of the error of this expectation value is given by the width of a binomial distribution with the same expectation value,

$$\Delta\langle\langle\hat{\sigma}_m\rangle\rangle = \frac{2\sqrt{N_{m+}N_{m-}}}{N_m^{3/2}}. \quad (5.60)$$

Suppose we perform N_x measurements along the x -axis, N_y along the y -axis and N_z along the z -axis, the probability of getting a certain set of results given the state $\hat{\rho}$ is

$$\begin{aligned} & \mathcal{P}(N_{x+}, N_{x-}, N_{y+}, N_{y-}, N_{z+}, N_{z-} | \hat{\rho}) \\ &= \binom{N_x}{N_{x+}} (p_{x+})^{N_{x+}} (p_{x-})^{N_{x-}} \times \binom{N_y}{N_{y+}} (p_{y+})^{N_{y+}} (p_{y-})^{N_{y-}} \times \binom{N_z}{N_{z+}} (p_{z+})^{N_{z+}} (p_{z-})^{N_{z-}} \\ &= \binom{N_x}{N_{x+}} \left(\frac{1+v_x}{2}\right)^{N_{x+}} \left(\frac{1-v_x}{2}\right)^{N_{x-}} \\ &\quad \times \binom{N_y}{N_{y+}} \left(\frac{1+v_y}{2}\right)^{N_{y+}} \left(\frac{1-v_y}{2}\right)^{N_{y-}} \\ &\quad \times \binom{N_z}{N_{z+}} \left(\frac{1+v_z}{2}\right)^{N_{z+}} \left(\frac{1-v_z}{2}\right)^{N_{z-}} \end{aligned} \quad (5.61)$$

where $N_x \equiv N_{x+} + N_{x-}$ and similar for N_y and N_z . When the number of measurements N becomes large, this calculation can become unwieldy. In the limit of large N (see Fig. 5.7 when distribution becomes approximately Gaussian,⁵ the probability of getting a result of r number of “+” results in N measurements is

$$\mathcal{P}(r) = \underbrace{\binom{N}{r} \times p^r (1-p)^{N-r}}_{r \sim \mathcal{B}(N,p)} \approx \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(r-\mu)^2}{2\sigma^2}\right\}}_{r \sim \mathcal{N}(\mu, \sigma^2)} \quad (5.62)$$

where the random variable r is approximated by the normal mean $\mu = Np$ and variance $\sigma^2 = Np(1-p)$, where p is given by Eq. (5.57) in this example. Therefore, Eq. (5.61) can be approximated as

$$\begin{aligned} & \mathcal{P}(N_{x+}, N_{x-}, N_{y+}, N_{y-}, N_{z+}, N_{z-} | \hat{\rho}) \\ & \propto \exp\left\{-\frac{(N_{x+} - \mu_{x+})^2}{2\sigma_{x+}^2}\right\} \times \exp\left\{-\frac{(N_{y+} - \mu_{y+})^2}{2\sigma_{y+}^2}\right\} \times \exp\left\{-\frac{(N_{z+} - \mu_{z+})^2}{2\sigma_{z+}^2}\right\}, \end{aligned} \quad (5.63)$$

where $\mu_{x+} = N_x \frac{(1+v_x)}{2}$ and $\sigma_{x+}^2 = N_x \frac{(1+v_x)(1-v_x)}{2}$, and similarly for the corresponding y and z quantities. This is an easier expression to compute.

The problem of quantum state tomography is then to invert Eq. (5.61): given a set of experimental results, what can we say about the qubit’s density matrix that has given rise to these results?

There are many statistical methods that can be used for quantum state tomography. Each of these methods has its own advantages and disadvantages, e.g. reconstruction quality, ease of implementation, or computational benefits. In this course, we shall just outline a conceptually simple method called direct inversion tomography, with a simple correction for when direct inversion produces an unphysical result.

⁵There are two approximations to binomial probabilities: the Poisson distribution when N is large and p is small, and the normal or Gaussian distribution when $Np(1-p)$ is large.

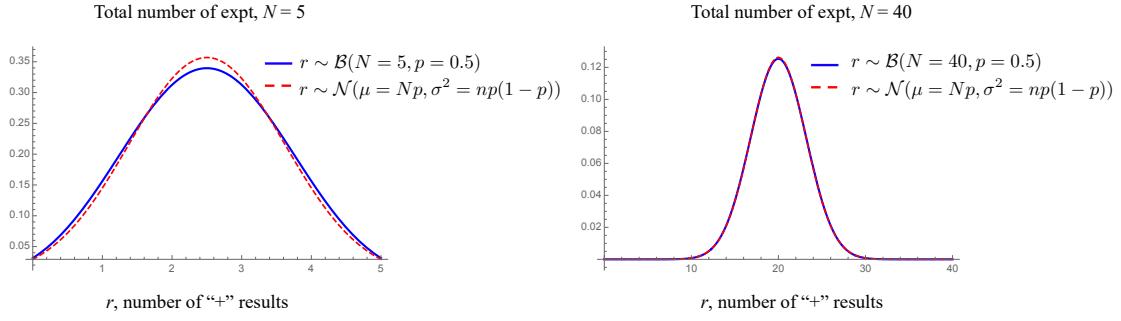


Figure 5.7: Binomial distribution (blue) approximates Gaussian distribution (red, dashed) as N increases. Plots shown for $p = 0.5$.

Direct Inversion

The simplest tomographic method, called direct inversion, assumes that the sample mean $\langle\langle\hat{\sigma}_m\rangle\rangle$ is a good and unbiased estimate of the population mean $\langle\hat{\sigma}_m\rangle$. Combining Eqs. (5.59), (5.56) along the three Cartesian axes fully defines an estimate of the qubit's Bloch vector,

$$\vec{v}_{\text{DI}} = \left(\frac{N_{x+} - N_{x-}}{N_x}, \frac{N_{y+} - N_{y-}}{N_y}, \frac{N_{z+} - N_{z-}}{N_z} \right)^T. \quad (5.64)$$

We note that \vec{v}_{DI} is the global maximum of Eq. (5.61), which is a definition of \vec{v}_{DI} that is readily extensible to different measurement schemes.

However, due to experimental errors and limitation of small sample size, direct inversion has a finite chance of generating an unphysical density matrix where $\|\vec{v}_{\text{DI}}\| > 1$. This can happen even though Eq. (5.64) is an estimation of three individual parameters – the three Cartesian components of the Bloch vector. Each parameter estimate is unconstrained on its own, but the three estimates must satisfy the joint constraint $v_{\text{DI},x}^2 + v_{\text{DI},y}^2 + v_{\text{DI},z}^2 \leq 1$, which is not enforced.

Maximum Likelihood Estimation

Bayes' theorem states that if we are given measurement datum $(N_{x+}, N_{x-}, N_{y+}, N_{y-}, N_{z+}, N_{z-})$ from an experiment, the likelihood that a certain density matrix $\hat{\rho}$ was at the source of these data is

$$\mathcal{L}(\hat{\rho}|N_{x+}, N_{x-}, N_{y+}, N_{y-}, N_{z+}, N_{z-}) \propto \mathcal{M}(\hat{\rho}) \mathcal{P}(N_{x+}, N_{x-}, N_{y+}, N_{y-}, N_{z+}, N_{z-}|\hat{\rho}), \quad (5.65)$$

where $\mathcal{M}(\hat{\rho})$ is a probability distribution called the prior, corresponding to our prior knowledge. Choosing a prior can be a matter of taste or actual prior knowledge; in this course, we only use the simplest prior,

$$\mathcal{M}(\hat{\rho}) = \begin{cases} \text{const.} & \text{if } \|\vec{v}\| \leq 1, \\ 0 & \text{if } \|\vec{v}\| > 1. \end{cases} \quad (5.66)$$

In practice, it is more convenient to work with the *log*-likelihood $\log \mathcal{L}(\hat{\rho})$ than the likelihood $\mathcal{L}(\hat{\rho})$ function. The log-likelihood function reduces the multiplicative factors in Eq. (5.65) to a sum. A sum is easier to work with than a product since its dynamic range is much smaller, and usually avoids the numerical overflow or underflow errors common in computing exponentials in a computer.

Since the global maximum of the probability \mathcal{P} , Eq. (5.61), is at \vec{v}_{DI} , we see that whenever $\|\vec{v}_{\text{DI}}\| \leq 1$ the maximum-likelihood estimate (MLE) of the Bloch vector is simply $\vec{v}_{\text{MLE}} = \vec{v}_{\text{DI}}$.

On the other hand, if $\|\vec{v}_{\text{DI}}\| > 1$, we define a distance measure between \vec{v}_{DI} and another

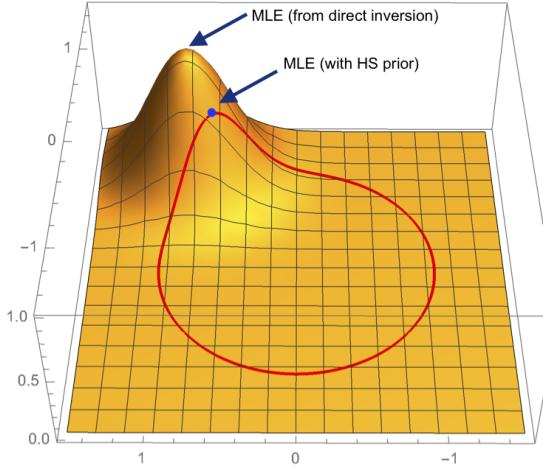


Figure 5.8: An example of the likelihood function Eq. (5.65) with the Gaussian approximation of Eq. (5.63), plotted against two parameters, say v_x and v_y for some v_z . Red line indicates the boundary of physically valid Bloch vectors. For $v_z = 0$, the red line represents the equator of the Bloch sphere. The peak (maximum likelihood) of the Gaussian approximation lies outside the physically valid region. With the Hilbert-Schmidt prior, all points outside the red line are reduced to zero likelihood, giving the maximum likelihood at the blue point. There are many methods to find the “blue point”; we outline one method in Fig. 5.9

Bloch vector \vec{v} which we call the scaled log-likelihood or Kullback-Leibler divergence,

$$\begin{aligned} \mathcal{D}_{\text{KL}}(\vec{v}_{\text{DI}}, \vec{v}) &= \ln \left(\frac{\mathcal{L}(\vec{v}_{\text{DI}})}{\mathcal{L}(\vec{v})} \right) = \ln \left(\frac{\mathcal{M}(\rho)\mathcal{P}(\vec{v}_{\text{DI}})}{\mathcal{M}(\rho)\mathcal{P}(\vec{v})} \right) \\ &= N_{x+} \ln \left(\frac{1 + v_{\text{DI},x}}{1 + v_x} \right) + N_{x-} \ln \left(\frac{1 - v_{\text{DI},x}}{1 - v_x} \right) \\ &\quad + N_{y+} \ln \left(\frac{1 + v_{\text{DI},y}}{1 + v_y} \right) + N_{y-} \ln \left(\frac{1 - v_{\text{DI},y}}{1 - v_y} \right) \\ &\quad + N_{z+} \ln \left(\frac{1 + v_{\text{DI},z}}{1 + v_z} \right) + N_{z-} \ln \left(\frac{1 - v_{\text{DI},z}}{1 - v_z} \right). \end{aligned} \quad (5.67)$$

We minimize this distance measure over the space of physically valid Bloch vectors $\|\vec{v}\| \leq 1$, in order to find a Bloch vector \vec{v}_{MLE} that is closest in distance, and hence of maximum likelihood to be the physically valid result. This is shown schematically in Fig. 5.9. Note that this is just one method of performing MLE; there are many other methods. Maximizing the likelihood, Eq. (5.65), is equivalent to maximizing \mathcal{P} . Since the logarithm is monotonic, this is equivalent to minimizing \mathcal{D}_{KL} .

One can also use the analytical expression for the likelihood function of Eq. (5.65) with or without the Gaussian approximation, once the sample mean and standard deviation from direct inversion is known. Together with the Hilbert-Schmidt prior \mathcal{C}_{HS} , this allows use to find the Bloch vector that gives the maximum likelihood. One such plot for (v_x, v_y) is shown in Fig. 5.8.

With the Hilbert-Schmidt prior \mathcal{C}_{HS} , this allows use to find the Bloch vector that gives the maximum likelihood. An example plot for (v_x, v_y) is shown in Fig. 5.8.

5.7.2 Multi-qubit State Tomography

Similar to the single qubit case, an arbitrary density matrix on n qubits can be expanded as

$$\hat{\rho} = \frac{1}{2} \sum_{v_1 \dots v_n = 0,1,2,3} \text{Tr}(\hat{\sigma}_{v_1} \otimes \hat{\sigma}_{v_2} \otimes \dots \otimes \hat{\sigma}_{v_n} \hat{\rho}) \hat{\sigma}_{v_1} \otimes \hat{\sigma}_{v_2} \otimes \dots \otimes \hat{\sigma}_{v_n}. \quad (5.68)$$

By performing measurements of observables which are products of Pauli matrices we can estimate each term in this sum, and thus obtain an estimate for $\hat{\rho}$. In practice, however, performing

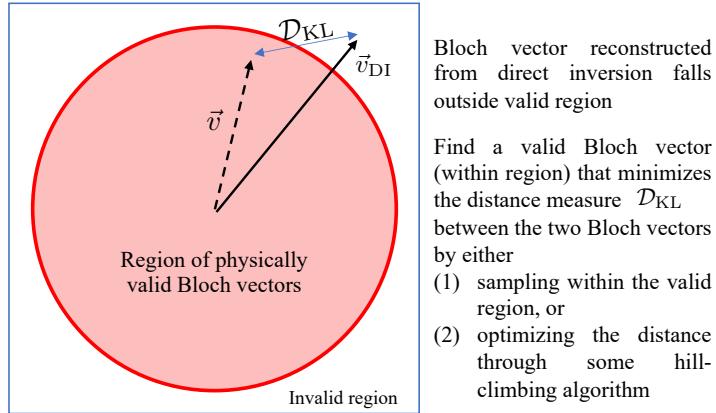


Figure 5.9: Quantum state tomography.

quantum state tomography for multiple qubits gets very complicated easily. We shall limit our study to single qubit cases in this course.

5.8 Quantum Process Tomography

Now let us outline the experimental procedure for quantum process tomography. Given that a system has d dimensions, e.g. $d = 2$ for a single qubit. Choose d^2 possible pure states, so that the set of density matrices $\{\hat{\rho}_1 = |\psi_1\rangle\langle\psi_1|, \hat{\rho}_2 = |\psi_2\rangle\langle\psi_2|, \dots, \hat{\rho}_{d^2} = |\psi_{d^2}\rangle\langle\psi_{d^2}|\}$ form a basis set for the space of matrices.

For each state, $|\psi_i\rangle\langle\psi_i|$, the system is prepared in that state and subjected to the process \mathcal{E} we wish to perform the process tomography on. After the process is completed, we perform quantum state tomography on the new state $\mathcal{E}(|\psi_i\rangle\langle\psi_i|)$. In principle, that is all that is required. However, in practice, we would like to have a way of determining a useful representation of \mathcal{E} from experimentally available data. Our goal is to determine a set of operation elements $\{\hat{E}_i\}$ for \mathcal{E} ,

$$\mathcal{E}(\hat{\rho}) = \sum_i \hat{E}_i \hat{\rho} \hat{E}_i^\dagger. \quad (5.69)$$

However, experimental results involve numbers, not operators, which is a theoretical concept. To determine the Kraus operators \hat{E}_i from measurable parameters, it is convenient to consider an equivalent description of \mathcal{E} using a fixed set of operators \hat{A}_i , which form a basis for the set of operators on the state space. For the single qubit case, $\{\hat{A}_i\}$, correspond to the Pauli matrices and identity, $\{\mathbb{I}, \hat{X}, \hat{Y}, \hat{Z}\}$. (Here, we have used $\hat{X}, \hat{Y}, \hat{Z}$ in place of $\hat{\sigma}_{x,y,z}$ for easy notation.) So,

$$\hat{E}_i = \sum_m a_{im} \hat{A}_m \quad (5.70)$$

for some set of complex numbers a_{im} . Eq. (5.69) may be rewritten as

$$\begin{aligned} \mathcal{E}(\hat{\rho}) &= \sum_i \left(\sum_m a_{im} \hat{A}_m \right) \hat{\rho} \left(\sum_n a_{in}^* \hat{A}_n^\dagger \right) \\ &= \sum_m \sum_n \hat{A}_m \hat{\rho} \hat{A}_n^\dagger \underbrace{\left(\sum_i a_{im} a_{in}^* \right)}_{\chi_{mn}} \\ &= \sum_m \sum_n \hat{A}_m \hat{\rho} \hat{A}_n^\dagger \chi_{mn}. \end{aligned} \quad (5.71)$$

Here, $\chi_{mn} \equiv \sum_i a_{im} a_{in}^*$, are the entries of a complex-valued matrix χ which is Hermitian⁶ and

⁶ $\chi_{mn} = \sum_i a_{im} a_{in}^* = \chi_{nm}^*$, which is Hermitian.

positive semidefinite⁷. This expression, known as the *chi*-matrix representation, shows that \mathcal{E} can be completely described by a complex number matrix χ , once the set of operators \hat{A}_m has been fixed.

What have we done here? The point is that the Kraus operators \hat{E}_i which determine the process \mathcal{E} are unknown. We rewrite them as linear combinations of known operators \hat{A}_m that we chose with unknown coefficients a_{im} . The problem of determining the unknown \hat{E}_i is then transferred to the problem of determining the unknown coefficients a_{im} .

In general, χ is a $d^2 \times d^2$ complex-valued matrix. It has $d^4 - d^2$ independent real parameters. This is because a $d^2 \times d^2$ complex-valued matrix has d^4 independent parameters, but χ is constrained by the completeness relation, which adds d^2 constraints. The completeness relation in terms of χ is

$$\mathbb{I} = \sum_i \hat{E}_i^\dagger \hat{E}_i = \sum_{i,j,k} a_{ik}^* \hat{A}_k a_{ij} \hat{A}_j = \sum_{j,k} \left(\sum_i a_{ij} a_{ik}^* \right) \hat{A}_k \hat{A}_j = \sum_{j,k} \chi_{jk} \hat{A}_k \hat{A}_j. \quad (5.72)$$

Determining χ

Next, we show how to determine χ experimentally. Given the basis states $\{\hat{\rho}_k\}$, we can express the tomographically reconstructed output state $\mathcal{E}(\hat{\rho}_j)$ in terms of linear combinations of the basis states,

$$\mathcal{E}(\hat{\rho}_j) = \sum_k \lambda_{jk} \hat{\rho}_k. \quad (5.73)$$

Coefficients λ_{jk} can be determined by standard linear algebra techniques. Next, we can also write

$$\hat{A}_m \hat{\rho}_j \hat{A}_n^\dagger = \sum_k \beta_{jk}^{mn} \hat{\rho}_k \quad (5.74)$$

where β_{jk}^{mn} are complex numbers which can be determined by standard linear algebra techniques, since the \hat{A} and $\hat{\rho}$ operators are known. Combining the last two equations and Eq. (5.71), we get

$$\begin{aligned} \mathcal{E}(\hat{\rho}_j) &= \sum_m \sum_n \hat{A}_m \hat{\rho}_j \hat{A}_n^\dagger \chi_{mn} = \sum_k \lambda_{jk} \hat{\rho}_k \\ \implies \sum_m \sum_n \sum_k \beta_{jk}^{mn} \chi_{mn} \hat{\rho}_k &= \sum_k \lambda_{jk} \hat{\rho}_k \\ \therefore \sum_m \sum_n \beta_{jk}^{mn} \chi_{mn} &= \lambda_{jk}. \end{aligned} \quad (5.75)$$

The last line can be written due to the linear independence of $\hat{\rho}_k$. We note that λ_{jk} and β_{jk}^{mn} are known. Also, we can think of χ and λ as vectors with composite indices, and β as a $d^4 \times d^4$ matrix with columns indexed by mn , and rows by jk . Then, an inverse matrix $\kappa = \beta^{-1}$ can be found such that the last equation can be inverted to get the χ vector,

$$\chi_{mn} \equiv \sum_j \sum_k \kappa_{jk}^{mn} \lambda_{jk}. \quad (5.76)$$

Determining \hat{E}_i

Having determined χ , one immediately obtains the operator-sum representation for \mathcal{E} in the following manner. Let the unitary matrix \hat{U}^\dagger diagonalize χ , which is possible since χ is hermitian. Then,

$$\chi_{mn} = \sum_{i,j} U_{mi} d_i \delta_{ij} U_{nj}^* \quad (5.77)$$

$$\sum_i a_{im} a_{in}^* = \sum_i U_{mi} d_i U_{ni}^* \quad (5.78)$$

$$\implies a_{im} a_{in}^* = U_{mi} \sqrt{d_i} U_{ni}^* \sqrt{d_i} \quad (5.79)$$

$$\therefore \hat{E}_i = \sqrt{d_i} \sum_m U_{mi} \hat{A}_m. \quad (5.80)$$

⁷For any \vec{v} which is a d^2 -dimensional complex vector, we have $\vec{v}^\dagger \chi \vec{v} = \sum_{mn} v_m^* \chi_{mn} v_n = \sum_{imn} (a_{im} v_m^*)(a_{in}^* v_n) = \sum_i |\sum_m a_{im} v_m^*|^2 \geq 0$, which is positive semi-definite.

■ **Example 5.8 — Depolarizing channel.** The depolarizing channel in Section 5.5.3 is given by Eq. (5.44). Show that the following process matrix corresponding to the depolarizing channel is

$$\chi = \begin{pmatrix} 1 - 3p/4 & 0 & 0 & 0 \\ 0 & p/4 & 0 & 0 \\ 0 & 0 & p/4 & 0 \\ 0 & 0 & 0 & p/4 \end{pmatrix} \quad (5.81)$$

5.9 References

1. Operator Sum Representation, Quantum State and Process Tomography:
Nielsen and Chuang, Chapter 8.
2. Quantum Channels:
John Preskill's Lecture Notes, Chapter 3.
<http://theory.caltech.edu/~preskill/ph229/>
3. Quantum State Tomography:
Roman Schmied, Quantum state tomography of a single qubit: comparison of methods,
Journal of Modern Optics, 2016 VOL. 63, NO. 18, 1744–1758.
<http://dx.doi.org/10.1080/09500340.2016.1142018>

A recipe for single qubit quantum process tomography

■ **Example 5.9** Consider the quantum process on a single qubit ($d = 2$):

$$\mathcal{E}_1 : \hat{\rho} \longrightarrow \hat{\rho}' = \mathcal{E}_1(\hat{\rho}) = \sum_i \hat{E}_i \hat{\rho} \hat{E}_i^\dagger. \quad (5.82)$$

1. Choose a fixed set of operators (there should be $d^2 = 4$):

$$\hat{A}_0 = \mathbb{I}, \hat{A}_1 = X, \hat{A}_2 = -iY, \hat{A}_3 = Z.$$

Here, we have used X, Y, Z in place of $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ for notational ease.

2. Choose a fixed set of basis states (there should be $d^2 = 4$):

$$\{\hat{\rho}_j\} = \{\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3, \hat{\rho}_4\} = \{|0\rangle\langle 0|, |0\rangle\langle 1|, |1\rangle\langle 0|, |1\rangle\langle 1|\}.$$

Here, we note that $\hat{\rho}_2 = \hat{\rho}_1 X$, $\hat{\rho}_3 = X \hat{\rho}_1$, $\hat{\rho}_4 = X \hat{\rho}_1 X$, i.e. all basis states can be written in terms of $\hat{\rho}_1$ and an X operator.

3. Prepare the input states:

$$\{|0\rangle, |1\rangle, |+x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |+y\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)\}.$$

4. Perform the experiment, i.e. allow the quantum process to act on the input states.

For each input state, determine the output state via quantum state tomography:
 $\mathcal{E}_1(|0\rangle\langle 0|)$, $\mathcal{E}_1(|1\rangle\langle 1|)$, $\mathcal{E}_1(|+x\rangle\langle +x|)$, $\mathcal{E}_1(|+y\rangle\langle +y|)$.

5. With the data, use Eq. (5.73) to write the output states in terms of the basis states. This will give us λ_{jk} . Here, we note that

$$|0\rangle\langle 1| = |+x\rangle\langle +x| + i|+y\rangle\langle +y| - \frac{1+i}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|), \text{ and}$$

$$|1\rangle\langle 0| = |+x\rangle\langle +x| - i|+y\rangle\langle +y| - \frac{1-i}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|).$$

Therefore,

$$\hat{\rho}'_1 = \mathcal{E}_1(|0\rangle\langle 0|)$$

$$\hat{\rho}'_2 = \mathcal{E}_1(|0\rangle\langle 1|) = \mathcal{E}_1(|+x\rangle\langle +x|) + i\mathcal{E}_1(|+y\rangle\langle +y|) - \frac{1+i}{2}(\mathcal{E}_1(|0\rangle\langle 0|) + \mathcal{E}_1(|1\rangle\langle 1|))$$

$$\hat{\rho}'_3 = \mathcal{E}_1(|1\rangle\langle 0|) = \mathcal{E}_1(|+x\rangle\langle +x|) - i\mathcal{E}_1(|+y\rangle\langle +y|) - \frac{1-i}{2}(\mathcal{E}_1(|0\rangle\langle 0|) + \mathcal{E}_1(|1\rangle\langle 1|))$$

$$\hat{\rho}'_4 = \mathcal{E}_1(|1\rangle\langle 1|)$$

and

$$\hat{\rho}'_1 = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{13} & \lambda_{14} \end{pmatrix}; \quad \hat{\rho}'_2 = \begin{pmatrix} \lambda_{21} & \lambda_{22} \\ \lambda_{23} & \lambda_{24} \end{pmatrix}; \quad \hat{\rho}'_3 = \begin{pmatrix} \lambda_{31} & \lambda_{32} \\ \lambda_{33} & \lambda_{34} \end{pmatrix}; \quad \hat{\rho}'_4 = \begin{pmatrix} \lambda_{41} & \lambda_{42} \\ \lambda_{43} & \lambda_{44} \end{pmatrix}$$

6. Use Eq. (5.74) and our choices of \hat{A}_m and $\hat{\rho}_j$ to compute each value of β_{jk}^{mn} . Depending on how you wish to perform indexing, you could take β^{mn} as a 4×4 matrix for each pair of (m, n) . Then there are 16 β^{mn} matrices. These matrices can be “assembled” into a $d^4 \times d^4$ matrix afterward.

7. Once β is found, obtain the inverse matrix κ (usually from a computer package) to express the χ matrix using Eq. (5.76).

8. Once the χ matrix is found, it can be diagonalized and the Kraus operators found, from Eqs. (5.79), (5.80). ■

PH3406 (2024) T.S.KOH

PH3406 (2024) T.S.KOH

5.10 Discussion Set

Problem 5.1 — Unitary freedom in operator-sum representation. Two quantum operations \mathcal{E} and \mathcal{F} act on a single qubit with the operator-sum representations $\mathcal{E}(\hat{\rho}) = \sum_k \hat{E}_k \hat{\rho} \hat{E}_k^\dagger$ and $\mathcal{F}(\hat{\rho}) = \sum_k \hat{F}_k \hat{\rho} \hat{F}_k^\dagger$. If the operation elements of \mathcal{E} are

$$\hat{E}_1 = \mathbb{I}/\sqrt{2}, \quad \text{and} \quad \hat{E}_2 = \hat{\sigma}_z/\sqrt{2}, \quad (5.83)$$

construct operation elements of \mathcal{F} so that the two quantum operations produce the same dynamics.

Problem 5.2 — Spontaneous emission as amplitude damping. A two-level atom is initially in a superposition of ground $|0\rangle$ and excited $|1\rangle$ states, $a|0\rangle + b|1\rangle$, where the decay of the excited state due to spontaneous emission of a photon is of probability p in a time interval Δt .

- (i) What is the probability p_0 of the atom being in the ground state initially?
- (ii) Determine the density matrix of the atom after time Δt .
- (iii) Show that if no photon is detected, then the probability of the atom being in the ground state is $p_0 > |a|^2$. Explain why this result makes sense.

Problem 5.3 — Depolarizing channel. The depolarizing channel is an important type of quantum noise. Imagine we take a single qubit, and with probability p that qubit is depolarized. That is, it is replaced by the completely mixed state, $\mathbb{I}/2$. With probability $1 - p$ the qubit is left untouched. The quantum operation is

$$\mathcal{E}(\hat{\rho}) = \frac{p\mathbb{I}}{2} + (1-p)\hat{\rho}. \quad (5.84)$$

However, this is not in the operator-sum representation. We note that

$$\mathcal{F}(A) = \frac{A + XAX + YAY + ZAZ}{4}, \quad (5.85)$$

gives the results $\mathcal{F}(I) = I$, $\mathcal{F}(X) = \mathcal{F}(Y) = \mathcal{F}(Z) = 0$. Here the Pauli matrices are denoted using X, Y, Z . Using this information, and the Bloch sphere representation for single qubit density matrices, rewrite Eq. (5.84) in the operator-sum representation. How does your result compare with Eq. (5.44)?

Problem 5.4 — Amplitude damping of a harmonic oscillator. Suppose that our principal system, a harmonic oscillator, interacts with an environment, modeled as another harmonic oscillator, through the Hamiltonian

$$\hat{H} = \hbar\chi(\hat{a}^\dagger\hat{b} + \hat{a}\hat{b}^\dagger), \quad (5.86)$$

where \hat{a}^\dagger, \hat{a} and \hat{b}^\dagger, \hat{b} are the creation and annihilation operators for the respective harmonic oscillators.

- (i) Using $\hat{U} = \exp(-i\hat{H}\Delta t/\hbar)$, denoting the eigenstates of $\hat{b}^\dagger\hat{b}$ as $|k_b\rangle$, and selecting the vacuum state $|0_b\rangle$ as the initial state of the environment, show that the quantum operation elements $\hat{E}_k = \langle k_b|\hat{U}|0_b\rangle$ are

$$\hat{E}_k = \sum_n \sqrt{\binom{n}{k}} \sqrt{(1-\gamma)^{n-k}\gamma^k} |n-k\rangle \langle n|. \quad (5.87)$$

where $\gamma = 1 - \cos^2(\chi\Delta t)$ is the probability of losing a single quantum of energy, and states such as $|n\rangle$ are eigenstates of $\hat{a}^\dagger\hat{a}$.

Hint: You will have to evaluate $\hat{U}\hat{a}^\dagger\hat{U}^\dagger = e^{s\hat{A}}\hat{B}e^{-s\hat{A}}$, where $\hat{A} \equiv \hat{a}^\dagger\hat{b} + \hat{a}\hat{b}^\dagger$, and $\hat{B} \equiv \hat{a}^\dagger$, $s \equiv -i\chi\Delta t$. The BCH formula is useful here, and you should be able to find patterns in the infinite series.

- (ii) Show that the operation elements \hat{E}_k define a trace-preserving quantum operation.

Problem 5.5 — Simultaneous amplitude and phase damping. The characteristic phase coherence time T_2 is related to the exponential decay of the *off-diagonal* elements in the qubit density matrix, while the characteristic relaxation time T_1 refers to the exponential decay of the *diagonal* elements. Amplitude damping has both non-zero T_1 and T_2 characteristic times where $T_2 = T_1/2$ (see Eqs. (5.39)), while phase damping alone has a T_2 coherence time only (see Eq. (5.32)).

Show that if amplitude and phase damping are simultaneously applied, then $T_2 \leq T_1/2$.

Problem 5.6 — Quantum process tomography. For the following two quantum channels, show that the process matrices can be written as shown.

- (i) **Dephasing.** Eq. (5.31) gives the evolution of the density matrix for the dephasing channel. Verify that it can also be written as $\mathcal{E}(\hat{\rho}) = (1 - p/2)\hat{\rho} + (p/2)\hat{Z}\hat{\rho}\hat{Z}$, and show that the χ matrix is

$$\chi = \begin{pmatrix} 1 - p/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p/2 \end{pmatrix}. \quad (5.88)$$

- (ii) **Rotation.** A rotation by angle θ about an axis \hat{n} applied to a state $\hat{\rho}$ is given by $\mathcal{E}(\hat{\rho}) = \exp[-i(\theta/2)\hat{n} \cdot \vec{\sigma}]\hat{\rho}\exp[i(\theta/2)\hat{n} \cdot \vec{\sigma}]$. Show that, for a rotation of angle θ about the z -axis, the χ matrix is

$$\chi = \frac{1}{2} \begin{pmatrix} 1 + \cos\theta & 0 & 0 & i \sin\theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i \sin\theta & 0 & 0 & 1 - \cos\theta \end{pmatrix}. \quad (5.89)$$

Problem 5.7 — Quantum process tomography. Consider a black box containing a single qubit of unknown dynamics \mathcal{E}_1 . Suppose that the following four density matrices are obtained from experimental measurements, performed according to the rules of Example 5.9.

$$\hat{\rho}'_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (5.90)$$

$$\hat{\rho}'_2 = \begin{pmatrix} 0 & \sqrt{1-\gamma} \\ 0 & 0 \end{pmatrix} \quad (5.91)$$

$$\hat{\rho}'_3 = \begin{pmatrix} 0 & 0 \\ \sqrt{1-\gamma} & 0 \end{pmatrix} \quad (5.92)$$

$$\hat{\rho}'_4 = \begin{pmatrix} \gamma & 0 \\ 0 & 1-\gamma \end{pmatrix} \quad (5.93)$$

where $0 \leq \gamma < 1$ is a real parameter, and $\{\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3, \hat{\rho}_4\} = \{|0\rangle\langle 0|, |0\rangle\langle 1|, |1\rangle\langle 0|, |1\rangle\langle 1|\}$. Take basis operations as $\hat{A}_0 = \mathbb{I}$, $\hat{A}_1 = X$, $\hat{A}_2 = -iY$, $\hat{A}_3 = Z$.

From the input-output relations, one could make several important observations: the ground state $|0\rangle$ is left invariant by \mathcal{E}_1 , the excited state $|1\rangle$ partially decays to the ground state, and superposition states are damped.

Determine the χ matrix for this process.

