Quantum Error Correction - Notes

Ben Karsberg

2021-22

1 Quantum Mechanics: Conventions and Notation

- We take the three postulates of QM from Nielsen and Chuang:

  1. Associated to any isolated physical system is a Hilbert space $\mathcal{H}$, known as the *state space* of the system, and the system is described by its *state vector* (or just *state*), which is a **unit** vector in $\mathcal{H}$
  2. The evolution of a **closed** quantum system is described by a *unitary transformation*
  3. *Quantum measurements* are specified by a collection $\{M_m\}$ of *measurement operators* acting on $\mathcal{H}$. The index m refers to the measurement outcomes, and if the system is in state $|\psi\rangle$ initially, the probability m occurs is:

     $$\mathbb{P}(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$

     and the state collapses to

     $$|\psi\rangle \to \frac{M_m|\psi\rangle}{\sqrt{\mathbb{P}(m)}}$$

1.1 Projective Measurements

- A *projective measurement* is described by an *observable*/Hermitian operator M acting on $\mathcal{H}$
- M has a spectral decomposition:

  $$M = \sum_m mP_m \tag{1.1}$$

  where m are the (real) eigenvalues of M and P_m is the projector onto the corresponding eigenspace
- Note $\sum_m P_m = 1$
- This means the outcomes are indexed by m, and the info in postulate (3) reduces to

  $$\mathbb{P}(m) = \langle\psi|P_m|\psi\rangle \quad \text{and} \quad |\psi\rangle \to \frac{P_m|\psi\rangle}{\sqrt{\mathbb{P}(m)}}$$

- The *expectation* of M on state $|\psi\rangle$ in a probabilistic sense reduces to

  $$\mathbb{E}(M) = \langle M\rangle_\psi = \langle\psi|M|\psi\rangle \tag{1.2}$$

- **Notation:** usually just list the projectors as $\{P_m\}$ so $\sum_m P_m = 1$ and $P_mP_n = \delta_{mn}P_m$

- Also sometimes say 'measure in basis $|m\rangle$', which just means 'perform projective measurement with projectors $P_m = |m\rangle\langle m|$

- Often refer to *positive operator-valued measurements (POVMs)*, which we define here

- Suppose we do a measurement with operators $\{M_m\}$ on system in state $|\psi\rangle$, so $\mathbb{P}(m) = |\psi|M_m^\dagger M_m|\psi\rangle$

- Define *POVM elements* $E_m = M_m^\dagger M_m$; then $\sum_m E_m = 1$ and $\mathbb{P}(m) = \langle\psi|E_m|\psi\rangle$

- Since E_m are sufficient to describe the outcomes, we call the set $\{E_m\}$ a POVM

1.2 Density Operator Formalism

- **Motivation:** nice and easy to talk about subsystems of a composite system, systems where we don't know the state precisely, statistical mechanics

- Consider a system which is in a state from the set $\{|\psi_i\rangle\}$, with the probability of being in state i given by p_i

- The set $\{|\psi_i\rangle, p_i\}$ is called an *ensemble of pure states*

- The *density operator/matrix* for this system is defined as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{1.3}$$

- Clearly the states $|\psi_i\rangle$ uniquely define ρ and vice-versa, so postulate (1) is unchanged

- For postulate 2, we see that if U governs evolution of states, ρ evolves as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \xrightarrow{U} \sum_i p_i U|\psi_i\rangle\langle\psi_i| U^\dagger = U\rho U^\dagger \tag{1.4}$$

- For describing measurements, suppose we have measurement given by measurement operators $\{M_m\}$

- If we were initially in $|\psi_i\rangle$, probabilities are

$$\mathbb{P}(m|i) = \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = \operatorname{Tr}\left(M_m^\dagger M_m |\psi_i\rangle\langle\psi_i|\right)$$

- Therefore, the probability of obtaining m outright on ρ is

$$\mathbb{P}(m) = \sum_i \mathbb{P}(m|i)p_i = \operatorname{Tr}(M_m^\dagger M_m \rho) \tag{1.5}$$

  and after a bit of algebra, we find ρ collapses to

$$\rho \to \rho_m = \frac{M_m \rho M_m^\dagger}{\operatorname{Tr}(M_m^\dagger M_m \rho)}$$

- If we know the state of a system is certainly $|\psi\rangle$, the system is in a *pure state* and $\rho = |\psi\rangle\langle\psi|$; otherwise we have a *mixed state*

- We have criteria $\operatorname{Tr}(\rho^2) = 1$ for pure states and $\operatorname{Tr}(\rho^2) < 1$ for mixed states

- **Theorem:** an operator ρ is a density operator for some system iff

1. $\mathrm{Tr}(\rho) = 1$
2. ρ is a positive operator

- **Theorem:** ensembles $\{|\psi\rangle\}$ and $\{|\phi_i\rangle\}$ generate the same density operator iff

$$|\psi_i\rangle = \sum_j U_{ij} |\phi_j\rangle$$

  where U is a unitary operator

2 Noise and Error Correction

2.1 Markov Processes and Classical Noise

- Consider the classical bit-flip process; the environment contains magnetic fields which can cause a bit to 'flip' with probability p
- To figure out p, we need a model both for the distribution of magnetic fields in the environment, and one for how they interact with bits
- This is a common thing: need a model for both the environment and for the system-environment interaction
- To be concrete, suppose p_0 and p_1 are the initial probabilities for the bit to be a 0 and a 1, and q_0 and q_1 are the probabilities after flipping
- Denote the initial state of the bit as A and the final state as B; then

$$\mathbb{P}(B = b) = \sum_a \mathbb{P}(B = b \,|\, A = a)\mathbb{P}(A = a)$$

- The probabilities $\mathbb{P}(B = b \,|\, A = a)$ are called *transition probabilities*; the above equation can also be written

$$\begin{pmatrix} q_0 \\ q_1 \end{pmatrix} = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}$$

- If some further noise occurs after the initial bit-flip, we can model this as being **independent** of the first flip - this is called *Markovicity*, and we can model the total noise process as a *Markov process*
- For a single-state process, we therefore have that output probabilities $\mathbf{q}$ are related to input probabilities $\mathbf{p}$ by

$$\mathbf{q} = E\mathbf{p} \tag{2.1}$$

  where E is a matrix of transition probabilities called the *evolution matrix*

- This means the final state of the system is **linearly** related to the initial state
- We need $E\mathbf{p}$ to be a valid probability distribution, so this gives us some conditions on E:
  - All entries of E are non-negative, so E is positive
  - All columns of E sum to 1, called *completeness*
- This is all classical, but there are quantum analogues

2.2 Quantum Operations

- **Not to be confused with operators!!**
- Similar to (2.1) for evolution of classical states, quantum states transform as

$$\rho' = \mathcal{E}(\rho) \tag{2.2}$$

  where $\mathcal{E}$ is called a *quantum operation*
- Simple examples: for time-evolution governed by U, $\mathcal{E}(\rho) = U\rho U^\dagger$, for measurements $\mathcal{E}_m(\rho) = M_m \rho M_m^\dagger$

2.2.1 Closed and Open Systems

- *Closed* quantum systems can be thought of as being isolated, and not interacting with an environment; evolution is governed by a unitary operation
- *Open* systems are thought of as arising from an interaction between a system of interest and an environment, together forming a closed system
- For an open system, this means the final state $\mathcal{E}(\rho)$ may not be unitarily related to the initial state
- Suppose the system-environment is initially in $\rho \otimes \rho_{\text{env}}$; then, the operation for evolving ρ alone is found by taking a partial trace:

$$\mathcal{E}(\rho) = \mathrm{Tr}_{\text{env}}\left[U(\rho \otimes \rho_{\text{env}})U^\dagger\right] \tag{2.3}$$

2.2.2 Operator-Sum Representation

- **Motivation:** restate (2.3) in terms of operators on the principal Hilbert space only
- Suppose $|e_k\rangle$ is an orthonormal basis for $\mathcal{H}_{\text{env}}$, and initially $\rho_{\text{env}} = |e_0\rangle\langle e_0|$
- Then:

$$\mathcal{E}(\rho) = \sum_k \langle e_k| U \left[\rho \otimes |e_0\rangle\langle e_0|\right] U^\dagger |e_k\rangle = \sum_k E_k \rho E_k^\dagger \tag{2.4}$$

  where $E_k = \langle e_K|U|e_0\rangle$ are operators on $\mathcal{H}$ only
- This defines the *operator-sum representation* of $\mathcal{E}$, and $\{E_k\}$ are the *operation elements* of $\mathcal{E}$
- Since $\mathrm{Tr}(\mathcal{E}(\rho)) = 1$, it follows that

$$\sum_k E_k^\dagger E_k = 1$$

  which is satisfied by *trace-preserving operation*
- This is **very** useful, since we can talk about the operation without referring to properties of the environment
- This representation is **not** unique
- **Theorem: (Unitary Freedom)** Suppose $\{E_1, \ldots, E_m\}$ and $\{F_1, \ldots, F_n\}$ are operation elements of operations $\mathcal{E}$ and $\mathcal{F}$ respectively. Append zero operators to the shorter list so $m = n$. Then, $\mathcal{E} = \mathcal{F}$ iff $E_i = \sum_j u_{ij} F_j$ where u is a $m \times m$ unitary matrix

2.2.3 Axiomatisation and Properties

- Quantum operations can be defined axiomatically
- A quantum operation $\mathcal{E}$ is a map from the set of density operators on an input space $\mathcal{H}_1$ to an output $\mathcal{H}_2$, satisfying the following properties:

 1. $\text{Tr}(\mathcal{E}(\rho))$ is the probability the process $\mathcal{E}$ occurs when ρ is the initial state. This means $0 \leq \text{Tr}(\mathcal{E}(\rho)) \leq 1 \ \forall \rho$
 2. $\mathcal{E}$ is convex-linear; that is, for real numbers/probabilities $\{p_i\}$

$$\mathcal{E}\left(\sum_i p_i \rho_i\right) = \sum_i p_i \mathcal{E}(\rho_i) \tag{2.5}$$

 3. $\mathcal{E}$ is completely positive; so for any positive operator A, $\mathcal{E}(A)$ must also be positive. More generally, if we introduce auxillary system R, $(\mathcal{I} \otimes \mathcal{E})(A)$ is positive for any positive A on combined system $R \otimes \mathcal{H}_1$

- We also have the following theorem:
- **Theorem:** $\mathcal{E}$ satisfies the above axioms iff

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger \tag{2.6}$$

 for some set of operators $\{E_i\}$ mapping $\mathcal{H}_1 \to \mathcal{H}_2$, and $\sum_i E_i^\dagger E_i \leq 1$

2.3 Error Correction

2.3.1 Examples

- Two main examples: *bit-flip correction* and *Shor code*
- **Bit-flip:**
 - Suppose Alice sends qubit $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ to Bob across a noisy channel, which has probability p of 'flipping' each qubit; that is

$$|0\rangle \mapsto |1\rangle \quad \text{and} \quad |1\rangle \mapsto |0\rangle$$

 - Equivalently, the channel has probability p to send $X\,|\psi\rangle$
 - However, we can protect against this noise by sending three qubits instead, *encoding* as

$$\begin{aligned} |0\rangle &\to |0_L\rangle = |000\rangle \\ |1\rangle &\to |1_L\rangle = |111\rangle \end{aligned} \tag{2.7}$$

 - Bob then receives the three-qubit state, with some qubits possibly flipped
 - He then has to perform *error-detection/syndrome diagnosis*: a measurement to detect if an error occurred and if so, on which qubit it occurred to
 - The relevant 4 projectors for this measurement are:

$$\begin{aligned} P_0 &= |000\rangle\langle 000| + |111\rangle\langle 111| \qquad &&\text{(no error)} \\ P_1 &= |100\rangle\langle 100| + |011\rangle\langle 011| \qquad &&\text{(qubit 1 flipped)} \\ P_2 &= |010\rangle\langle 010| + |101\rangle\langle 101| \qquad &&\text{(qubit 2 flipped)} \\ P_3 &= |001\rangle\langle 001| + |110\rangle\langle 110| \qquad &&\text{(qubit 3 flipped)} \end{aligned} \tag{2.8}$$

  - Note that the outcome of the relevant measurements is 1 if the corresponding error condition is met, and the measurement also does not change the state of the three qubits
  - Upon getting our measurement result, we just apply X to the relevant qubit if an error occurred to get back our original state
  - This error-correction procedure works perfectly so long as one or fewer errors occur: this happens with probability $(1-p)^3+3p(1-p)^2 = 1-3p^2+2p^3$, and the probability we cannot correct the error is therefore $3p^2 - 2p^3$
  - For $p < 1/2$ we therefore have increased reliability
  - In the operation language above, the action of the bit-flip channel can be written

$$\mathcal{E}(\rho) = (1-p)\rho + pX\rho X \tag{2.9}$$

- Before looking at the Shor code, we note another important basic error: the *phase-flip*
- This is similar to the bit-flip, except the relative phase of $|0\rangle$ and $|1\rangle$ are flipped; that is, with probability p we have

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \mapsto a\,|0\rangle - b\,|1\rangle \tag{2.10}$$

- This is easy to correct for: we just move to the $|\pm\rangle$ basis and perform the same procedure as in the bit-flip case
- The action of the phase-flip channel is:

$$\mathcal{E}(\rho) = (1-p)\rho + pZ\rho Z \tag{2.11}$$

- **The Shor code:**
  - The initial motivation is to correct for both phase and bit-flips
  - To do this, we first encode our qubit via the phase-flip machinery: $|0\rangle \to |+++\rangle$ and $|1\rangle \to |---\rangle$
  - Then, we encode each of our new three qubits via the bit-flip code: $|+\rangle \to \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ and $|-\rangle \to \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)$
  - The final result is a nine qubit code, with logical qubits

$$\begin{aligned} |0\rangle \to |0_L\rangle &= \frac{1}{2\sqrt{2}}\left(|000\rangle + |111\rangle\right)\left(|000\rangle + |111\rangle\right)\left(|000\rangle + |111\rangle\right) \\ |1\rangle \to |1_L\rangle &= \frac{1}{2\sqrt{2}}\left(|000\rangle - |111\rangle\right)\left(|000\rangle - |111\rangle\right)\left(|000\rangle - |111\rangle\right) \end{aligned} \tag{2.12}$$

  - This hierarchical way of encoding qubits is called *concatenation*
  - It turns out that the Shor code can correct **arbitrary** single-qubit errors
  - To see why, suppose the encoded qubit is initially $|\psi\rangle = \alpha\,|0_L\rangle + \beta\,|1_L\rangle$, and consider the operator-sum representation of an arbitrary error $\mathcal{E}$ with elements $\{E_i\}$
  - After the noise acts, the state is now

$$\mathcal{E}(|\psi\rangle\langle\psi|) = \sum_i E_i\,|\psi\rangle\langle\psi|\,E_i^\dagger$$

  - Let's focus on the single term $E_i\,|\psi\rangle\langle\psi|\,E_i^\dagger$

- Any single-qubit operator is a 2×2 Hermitian matrix, so as an operator on qubit 1 only, it can be expanded in the basis of Hermitian matrices $\{I, X, Z, Y = iXZ\}$:

$$E_i = e_{i0}I + e_{i1}X_1 + e_{i2}Z_2 + e_{i3}X_1Z_1 \tag{2.13}$$

- Therefore, the unnormalised state $E_i |\psi\rangle$ is a superposition of $|\psi\rangle, X_1 |\psi\rangle, Z_1 |\psi\rangle, X_1Z_1 |\psi\rangle$
- Measuring the error syndrome collapses this superposition into one of the 4 states, and thus we can reverse the error by performing the appropriate inversion
- The same is true for all other errors, so error correction results in $|\psi\rangle$ being recovered, even though the first qubit error was arbitrary
- This is **very** powerful: we can correct a continuous spectrum of errors by just correcting the bit-flip, phase-flip, and combined bit/phase-flip

2.3.2 Generalities

- General procedure: states $|\psi\rangle \in \mathcal{H}$ encoded by unitary operation into a *quantum error-correcting code*, which is a subspace $\mathcal{H}_{\text{code}}$ of a larger Hilbert space
- Often refer to the *projector into the code space* P_{code} or P_c
- After encoding, the state is subjected to noise, a *syndrome measurement* is performed to diagnose what type of error occurred, and then a *recovery operation* is performed to obtain the original state
- Different errors correspond to **orthogonal** subspaces of the full Hilbert space so they can be distinguished
- In the general theory, we make no assumptions about a two-stage detection-recovery method: just assume noise is given by operation $\mathcal{E}$ and error-correction done by operation $\mathcal{R}$
- For error-correction to be successful, we require that for any state ρ supported in $\mathcal{H}_{\text{code}}$:

$$(\mathcal{R} \circ \mathcal{E})(\rho) \propto \rho \tag{2.14}$$

- The proportionality rather than equals means that we include non-trace-preserving operations in $\mathcal{E}$
- **Theorem: (Quantum Error-Correction Conditions)** Let $\mathcal{H}_{\text{code}}$ be a quantum code, and P_c be the projector onto it. Suppose $\mathcal{E}$ is a quantum operation with elements $\{E_i\}$; an error-correction operation $\mathcal{R}$ correcting $\mathcal{E}$ on $\mathcal{H}_{\text{code}}$ exists iff

$$P_c E_i^\dagger E_j P_c = \alpha_{ij} P_c \tag{2.15}$$

  where α is a complex Hermitian matrix
- The set $\{E_i\}$ are called the *errors*, and if $\mathcal{R}$ exists we say that they are a *correctable set of errors*
- In general, we don't know the form of the noise, but we can adapt the quantum error correction conditions to find a whole class of noises which a code $\mathcal{H}_c$ and correction operation $\mathcal{R}$ can correct for
- **Theorem:** Suppose $\mathcal{H}_c$ is a quantum code, and $\mathcal{R}$ is the full error-correction operation correcting $\mathcal{E}$ with elements $\{E_i\}$. Then, $\mathcal{R}$ also corrects for $\mathcal{F}$ with elements $\{F_j\} = \{\sum_i m_{ji}E_i\}$ for complex matrix m on $\mathcal{H}_c$

- This theorem means we can talk about a class of errors $\{E_i\}$ which are *correctable* rather than a class of error processes $\mathcal{E}$
- This is useful: if (for example) we can find a process satisfying

$$P_c \sigma_i^1 \sigma_j^1 P_c = \alpha_{ij} P_c$$

  for the Pauli matrices, then we can correct for arbitrary single-qubit errors since any single qubit operation can be described by having operation elements given by the Pauli matrices

- The Shor code can do this!

2.4 Quantum Erasure

- One class of error which is particularly relevant to Harlow is *quantum erasure correction*
- This is defined as the channel 'erasing' a subsystem of the transmitted state, i.e. losing access to a **known** subsystem
- To formalise this, suppose $\mathcal{H}_{\text{code}} \subseteq \mathcal{H}$ where $\mathcal{H}$ has a tensor product structure $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$
- Then, the operation

$$\mathcal{E}(\rho) = (1-p)\rho + p\mathrm{Tr}_{\overline{A}}(\rho) \otimes |e\rangle\langle e| \tag{2.16}$$

  where $|e\rangle$ is a state orthogonal to $\mathcal{H}$ living in the Hilbert space $\mathcal{H}_E = \mathcal{H} \otimes \mathrm{span}(|e\rangle)$ - think of this as being a 'flag' for erasure

- If we set $\{|a\rangle\}$ and $\{|\overline{a}\rangle\}$ to be orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_{\overline{A}}$ respectively, we can rewrite this explicitly as

$$\mathcal{E}(\rho) = (1-p)\rho + p\sum_{\overline{a}} |e\rangle \langle \overline{a}|\rho|\overline{a}\rangle \langle e| \tag{2.17}$$

  from which we can read off operation elements

$$\begin{aligned} E_0 &= \sqrt{1-p}\left(\sum_a |a\rangle_E \langle a|_A + \sum_{\overline{a}} |\overline{a}\rangle_E \langle \overline{a}|_{\overline{A}}\right) \\ E_{\overline{a}} &= \sqrt{p}\,|e\rangle_E \langle \overline{a}|_{\overline{A}} \end{aligned} \tag{2.18}$$

- Note that $\mathcal{E} : \mathcal{H} \to \mathcal{H}_E$, so the output space has one extra dimension
- We now want to see what the error correction conditions (2.15) reduce to; the first step to do this is calculating:

$$\begin{aligned} E_0^\dagger E_0 &= (1-p) I_{\mathcal{H}} \\ E_{\overline{a}}^\dagger E_{\overline{b}} &= p\,|\overline{a}\rangle_{\overline{A}} \langle \overline{b}|_{\overline{A}} \\ E_0^\dagger E_{\overline{a}} &= E_{\overline{a}}^\dagger E_0 = 0 \end{aligned} \tag{2.19}$$

- (2.15) then becomes trivial for $E_0^\dagger E_0$ and 'cross-terms' $E_0^\dagger E_{\overline{a}}$ etc., and the only non-trivial condition is

$$pP_{\text{code}} |\overline{a}\rangle \langle \overline{b}| P_{\text{code}} = \alpha_{\overline{a}\overline{b}} P_{\text{code}} \tag{2.20}$$

  or, more simply

$$P_{\text{code}} |\overline{a}\rangle \langle \overline{b}| P_{\text{code}} \propto P_{\text{code}} \tag{2.21}$$

- This has a very relevant implication to Harlow: any operator $X_{\overline{A}}$ on $\mathcal{H}_{\overline{A}}$ can be decomposed as $X_{\overline{A}} = \sum_{\overline{a},\bar{b}} x_{\overline{a},\bar{b}} \left|\overline{a}\right\rangle \left\langle\overline{b}\right|$, so if (2.21) holds, then for any such $X_{\overline{A}}$ we must have
$$P_{\text{code}} X_{\overline{A}} P_{\text{code}} \propto P_{\text{code}} \tag{2.22}$$

- **This is exactly (3.3) of Harlow**, establishing its necessity for the correctability of erasure of $\overline{A}$