Ph 220: Quantum Learning Theory
Lecture Note 1: From Quantum Sensing to Learning

Hsin-Yuan Huang (Robert)

Caltech

Introduction

This lecture introduces the fundamentals of quantum sensing by focusing on a canonical problem: the optimal measurement of a magnetic field. This will serve as a foundation for understanding broader concepts in quantum learning.

Q: How can we optimally sense or learn a magnetic field?

Problem Setup

We model the magnetic field, $\vec{B}$, as pointing in the $\hat{z}$ direction, such that $\vec{B} = B\hat{z}$. The sensor is a qubit whose spin rotates under the influence of this field. The dynamics of this qubit sensor are described by the Hamiltonian:

$$H = B \cdot Z$$

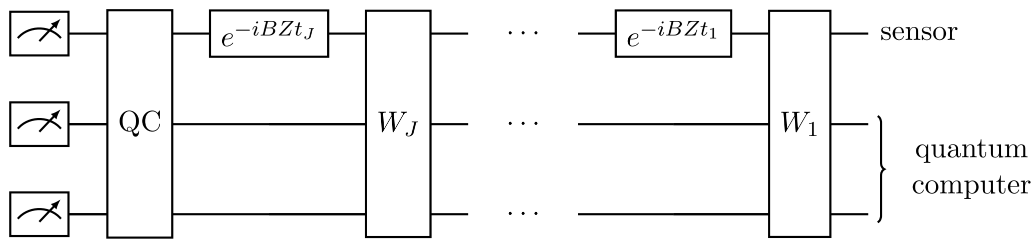
where Z is the Pauli-Z operator. For simplicity, we assume the field strength is bounded, $|B| \leq 1$. To determine the optimal way to learn the magnetic field B, we must first define the general class of protocols available to us.

Q: What is the ultimate family of protocols for learning the value of B?

A general sensing protocol consists of several steps:

1. Prepare an initial quantum state, typically involving a sensor qubit and a multi-qubit quantum computer.
2. Apply a unitary operation W_1 to the joint system.
3. Let the sensor evolve under the Hamiltonian for a time t_1.
4. Repeat steps 2 and 3 for subsequent operations $W_2, \ldots, W_J$ and evolution times $t_2, \ldots, t_J$.
5. Perform a final quantum computation (QC) and measure to produce an estimate $\hat{B}$.

This process is visualized in the quantum circuit below, drawn from right to left. The protocol begins with an initial state (far right), alternates between unitary operations W_j and free evolution, and concludes with a final computation (QC) and measurement (far left) to yield a classical estimate $\hat{B}$.



The final state of the $n+1$ qubits before the QC and measurement is given by:

$$|\psi_B\rangle = \left[\prod_{j=1}^{J} \left(e^{-iBZt_j} \otimes I\right) W_j\right] |0^{n+1}\rangle$$

From the measurement outcome, we compute an estimate, $\hat{B}$, that approximates the true value of B. The total **sensing time** is the sum of all evolution periods: $T = \sum_{j=1}^{J} t_j$. With this framework, we can pose a more precise question.

Q: What is the minimum sensing time required to learn B such that the error $|\hat{B} - B| < \varepsilon$ with high probability?

Our roadmap to answer this question is:

1. Design an effective protocol that achieves a good scaling.
2. Prove a fundamental lower bound (LB) on the required time for any protocol.

1 Protocol Design: Standard Quantum Limit (SQL)

We begin with a simple protocol consisting of independent measurements:

1. Prepare the sensor qubit in the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.
2. Let it evolve under $H = BZ$ for a time t. The state becomes:

$$|\psi_t\rangle = e^{-iBZt}|+\rangle = \frac{1}{\sqrt{2}}(e^{-iBt}|0\rangle + e^{iBt}|1\rangle)$$

3. Measure the qubit in the Y-basis, composed of states $|y+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ and $|y-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$.

The probabilities of the measurement outcomes are:

$$P(y+) = |\langle y+|\psi_t\rangle|^2 = \frac{1}{2}(1 + \sin(2Bt))$$

$$P(y-) = |\langle y-|\psi_t\rangle|^2 = \frac{1}{2}(1 - \sin(2Bt))$$

For simplicity, we fix the sensing time $t = 1/2$, so the probability simplifies to $P(y+) = \frac{1}{2}(1+\sin(B))$. We repeat this process N times. The empirical probability, $\hat{P}(y+)$, is estimated by dividing the number of $y+$ outcomes by N.

Analysis with Hoeffding's Inequality

Hoeffding's inequality for N i.i.d. random variables $X_i \in [0,1]$ states:

$$\Pr\left(\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \mathbb{E}[X]\right| \geq \varepsilon'\right) \leq 2e^{-2N(\varepsilon')^2}$$

This bounds the error in our probability estimate. To ensure $|\hat{P}(y+) - P(y+)| < \varepsilon'$ with failure probability at most δ, we need $N = O\left(\frac{\log(1/\delta)}{(\varepsilon')^2}\right)$. Our estimate for B is found by inverting the probability: $\hat{B} = \arcsin(2\hat{P}(y+) - 1)$. Using the Lipschitz continuity of $\arcsin(x)$ for x not close to ± 1, an error ε' in the probability corresponds to a field error $\varepsilon = O(\varepsilon')$. To achieve a final error $|\hat{B} - B| < \varepsilon$ with probability at least $1-\delta$, the number of measurements must be:

$$N = O\left(\frac{\log(1/\delta)}{\varepsilon^2}\right)$$

This result is the **Standard Quantum Limit (SQL)**. Since each measurement uses sensing time t, the total sensing time is $T = N \cdot t = O(1/\varepsilon^2)$ for a fixed success probability.

2 Protocol Design: Heisenberg Limit (HL)

To achieve a better sensing time of $O(1/\varepsilon)$, we use an iterative, multi-stage approach analogous to phase estimation.

Level 1: Run the SQL protocol to get a rough estimate. With precision $\varepsilon_1 = 0.1$ and confidence $1-\delta$, we obtain $\hat{B}^{(1)}$ where $\Pr[|\hat{B}^{(1)} - B| < 0.1] > 1-\delta$. The sensing time is $T_1 = C \cdot \log(1/\delta)$.

Level 2: Using this initial estimate, we amplify the phase to learn the next digit of B. We effectively evolve the system under a new Hamiltonian $H' = 10(B - \hat{B}^{(1)})Z$. This is achieved by applying the forward evolution $e^{-iB(10t)}$ followed by a corrective inverse evolution $e^{i\hat{B}^{(1)}(10t)}$. Running the SQL protocol on this task gives an estimate $\hat{b}^{(2)}$ for $10(B - \hat{B}^{(1)})$. Our new estimate for B is $\hat{B}^{(2)}_{\text{new}} = \hat{B}^{(1)} + \frac{\hat{b}^{(2)}}{10}$ with error less than 0.01. The sensing time for this stage is $T_2 = C \cdot 10 \cdot \log(1/\delta)$.

Level k: This process is repeated. At level k, we run SQL to estimate $10^{k-1}(B - \hat{B}^{(k-1)}_{\text{new}})$, with a sensing time of $T_k = C \cdot 10^{k-1} \cdot \log(1/\delta)$.

After $L = \lceil \log_{10}(1/\varepsilon) \rceil$ levels, the final estimate is $\hat{B} = \hat{B}^{(1)} + \frac{\hat{b}^{(2)}}{10} + \frac{\hat{b}^{(3)}}{100} + \dots$. The final error is bounded by $|\hat{B} - B| < 10^{-L} \approx \varepsilon$, and the cumulative failure probability is bounded by $L \cdot \delta$.

The total sensing time is the sum of the times from all levels:

$$T = \sum_{k=1}^{L} T_k = C \cdot \log(1/\delta) \cdot \sum_{k=0}^{L-1} 10^k = O\left(10^L \cdot \log(1/\delta)\right) = O\left(\frac{\log(1/\delta)}{\varepsilon}\right)$$

This improved scaling is the **Heisenberg Limit (HL)**.

3 Lower Bound

To establish a fundamental limit, we consider the task of distinguishing two nearby hypotheses: $B = +\varepsilon$ versus $B = -\varepsilon$. Any protocol that learns B with precision ε must be able to distinguish these two scenarios. The distinguishability of the measurement outcome distributions is limited by the distance between the corresponding quantum states, $|\psi_{+\varepsilon}\rangle$ and $|\psi_{-\varepsilon}\rangle$. The total variation distance (TVD) is bounded by the trace distance, which in turn is bounded by the Euclidean norm of the state difference:

$$\mathrm{TVD}(\Pr(x|B=+\varepsilon), \Pr(x|B=-\varepsilon)) \le \frac{1}{2}\| |\psi_{+\varepsilon}\rangle\langle\psi_{+\varepsilon}| - |\psi_{-\varepsilon}\rangle\langle\psi_{-\varepsilon}| \|_1 \le \| |\psi_{+\varepsilon}\rangle - |\psi_{-\varepsilon}\rangle \|_2$$

Using a hybrid argument (telescoping sum) and the fact that unitaries preserve norms, we can bound this difference:

$$\begin{aligned} \| |\psi_{+\varepsilon}\rangle - |\psi_{-\varepsilon}\rangle \|_2 &= \left\| \left(\prod_{j=1}^{J} (e^{-i\varepsilon Z t_j} \otimes I) W_j - \prod_{j=1}^{J} (e^{+i\varepsilon Z t_j} \otimes I) W_j \right) |0^{n+1}\rangle \right\|_2 \\ &\le \sum_{j=1}^{J} \left\| e^{-i\varepsilon Z t_j} - e^{+i\varepsilon Z t_j} \right\|_\infty = \sum_{j=1}^{J} \| -2i \sin(\varepsilon Z t_j) \|_\infty \\ &\le \sum_{j=1}^{J} 2 |\sin(\varepsilon t_j)| \le \sum_{j=1}^{J} 2\varepsilon t_j = 2\varepsilon T \end{aligned}$$

where $T = \sum_j t_j$ is the total sensing time. For the measurement outcomes to be reliably distinguishable, the TVD must be a constant value bounded away from zero. This requires:

$$\text{const} \le \mathrm{TVD} \le 2\varepsilon T \implies T = \Omega(1/\varepsilon).$$

This proves that any protocol, even with access to a quantum computer, must use a total sensing time of at least $\Omega(1/\varepsilon)$.

4 Conclusion

In summary:

(A) We constructed a protocol (phase estimation) with total sensing time $T = O(1/\varepsilon)$.

(B) We proved that any protocol achieving ε error requires a sensing time of $T = \Omega(1/\varepsilon)$.

Together, these results establish that the minimum sensing time to measure a magnetic field B to error ε is $\Theta(1/\varepsilon)$. At this point, we can say that we have fully understood how to optimally sense a magnetic field in an idealized setting. I hope this gives a flavor of how learning theory works.