Solutions Problem Sheet 8: Shot noise and convergence inequalities

Class problems

1. Suppose you want to compute the Hilbert-Schmidt norm $||\rho-\sigma||_2^2 = \mathrm{Tr}[\rho^2+\sigma^2-2\rho\sigma]$ between two mixed states ρ and σ.

   a) Show that $\mathrm{Tr}[\rho\sigma]$ can be measured on a quantum computer via performing a SWAP measurement. (Assume both ρ and σ are given to you.)

   b) Show that $\mathrm{Tr}[\rho\sigma]$ can also be measured using a generalization of the Loschmidt echo test. (Assume ρ is given to you but you know a circuit to prepare σ - note there are different ways in which the circuit to prepare σ could be given to you).

   c) Which measurement method converges more efficiently?

   Answer: Everything is well explained in the *Quantum Mixed State Compiling* paper (https://arxiv.org/abs/2209.00528). You can consider the eigen-decomposition of both ρ and σ and compute all fidelity terms between the eigenstates of σ and the eigenstates of ρ (pure states). For n qubits states, there are 2^n eigenstates for both (if full rank) so in total there are 2^{2n} fidelity to compute. Then, the answer is the same as in the standard setting with pure states discussed in lecture, but the number of shots required is multiplied by 2^{2n} (which is bad).

2. (Old Exam Problem) This question will compare the effect of shot noise for three approaches to estimating $\mathrm{Tr}[\rho H]$, i.e., the energy of a state ρ with Hamiltonian $H=\sum_{i=1}^{M}\alpha_i P_i$ where the P_i are Pauli operators.

   (a) First suppose you measure each Pauli term individually using N/M shots for each Pauli term. Write down an estimator for this measurement strategy. Show that the estimator is unbiased and compute its variance. Find an upper bound for the variance of the estimator in terms of a norm of the vector $\boldsymbol{\alpha}$. Carefully list any assumptions about N and M needed in this calculation.

   (5 marks)

   Estimator: $\hat{X}=\frac{1}{\tilde{N}}\sum_{j=1}^{\tilde{N}}\sum_{i=1}^{M}\alpha_i\lambda_i^{(j)}$ where $\tilde{N}=N/M$.

   (1 mark)

   To show that this is unbiased we compute

$$\langle\hat{X}\rangle=\frac{1}{\tilde{N}}\sum_{j=1}^{\tilde{N}}\sum_{i=1}^{M}\alpha_i\langle\lambda_i^{(j)}\rangle=\frac{1}{\tilde{N}}\sum_{j=1}^{\tilde{N}}\sum_{i=1}^{M}\alpha_i\langle P_i\rangle=\frac{1}{\tilde{N}}\sum_{j=1}^{\tilde{N}}\langle H\rangle=\langle H\rangle.$$

   (1 mark)

   To compute the variance we use

$$\mathrm{Var}(\hat{X})=\mathrm{Var}\left(\sum_i\alpha_i\hat{X}_i\right)=\sum_i|\alpha_i|^2\mathrm{Var}(\hat{X}_i)$$

   where we have defined the estimator $\hat{X}_i=\frac{1}{\tilde{N}}\sum_{j=1}^{\tilde{N}}\lambda_i^{(j)}$ and used the fact that each of these estimators is itself a random variable and independent from the others. The rest of the calculation is now straightforward:

$$\begin{aligned}\mathrm{Var}(\hat{X})&=\sum_i|\alpha_i|^2\mathrm{Var}(\hat{X}_i)=\sum_i|\alpha_i|^2\frac{\mathrm{Var}(P_i)}{\tilde{N}}\\&=\sum_i|\alpha_i|^2\frac{\langle P_i^2\rangle-\langle P_i\rangle^2}{\tilde{N}}=\sum_i|\alpha_i|^2\frac{1-\langle P_i\rangle^2}{\tilde{N}}\leqslant\sum_i|\alpha_i|^2\frac{1}{\tilde{N}}=\frac{M||\boldsymbol{\alpha}||_2^2}{N}\end{aligned}$$

   (3 marks)

(b) Next suppose you instead use $N_i = p_i N$ shots for each Pauli term where $p_i = |\alpha_i| / \sum_i |\alpha_i|$. Why intuitively might this be a better strategy than the previous one?

Because we use more shots to measure the higher weight terms which contribute more to the term we are trying to compute. (1 mark)

Write down an estimator for this measurement strategy. Show that the estimator is unbiased and compute its variance. Find an upper bound for the variance of the estimator in terms of a norm of the vector $\boldsymbol{\alpha}$. Carefully list any assumptions about N needed in this calculation.

(5 marks)

Intuition: Because we use more shots to measure the higher weight terms which contribute more to the term we are trying to compute. (1 mark)

Estimator: $\hat{X} = \sum_{i=1}^{M} \alpha_i \hat{X}_i$ where $\hat{X}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \lambda_i^{(j)}$ with $N_i = p_i N$. (1 mark)

Unbiased: (Same as above but $\tilde{N}$ replaced with N. (1 mark)

Variance: As the $\hat{X}_i$ are again independent random variables we can follow the same steps as above to get:

$$\mathrm{Var}(\hat{X}) = \sum_i |\alpha_i|^2 \frac{\mathrm{Var}(P_i)}{N_i} \leqslant \sum_i |\alpha_i|^2 \frac{1}{N_i} = \sum_i |\alpha_i|^2 \frac{||\alpha||_1}{\alpha_i N} = \frac{||\alpha||_1^2}{N} \tag{1}$$

(2 marks)

(c) Discuss which of these strategies is best for estimating $\mathrm{Tr}[\rho H]$ if only taking into account shot noise? Construct some example Hamiltonians to illustrate your arguments.

(5 marks)

Although $||\boldsymbol{\alpha}||_1 \geqslant ||\boldsymbol{\alpha}||_2$ the extra factor of M for method 1 means that the convergence is generally worse for method 1. (1 mark)

As a quick sanity check the methods should agree when the α_is are all equal. For example, consider $H = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_x)$. The upper bound on the variance for both methods is $2/N$.

But generally for uneven terms 2 will have a lower upper bound on the variance than method 1. For example, say $\alpha_0 = \frac{1}{\sqrt{10}}$ and $\alpha_1 = \frac{\sqrt{9}}{\sqrt{10}}$ then for method 1 the variance is $2/N$ but for method 2 the variance is roughly $1.6/N$.

(4 marks)

(d) Finally, suppose you instead have a source of classical randomness and with probability $p_i = |\alpha_i| / \sum_i |\alpha_i|$ measure Pauli P_i.

Again, write down an estimator for this measurement strategy. Show that the estimator is unbiased and compute its variance. And find an upper bound for the estimator in terms of a norm of the vector $\boldsymbol{\alpha}$.

(5 marks)

Estimator: We measure P_i with probability p_i to give $\lambda_i^{(j)}$. This leads to $\hat{X} = \frac{1}{N} \sum_{j=1}^{N} \hat{x}_j$ where $\hat{x}_j = \frac{\alpha_i \lambda_i^{(j)}}{p_i}$.
(1 mark)

Unbiased: The average of this estimator is

$$\langle \hat{X} \rangle = \frac{1}{N} \sum_j \langle x_j \rangle = \langle x_j \rangle = \sum_i p_i \frac{\alpha_i \langle \lambda_i^{(j)} \rangle}{p_i} = \sum_i \alpha_i \langle P_i \rangle = \langle H \rangle \,, \tag{2}$$

where in a slight abuse of notation we use $\langle ... \rangle$ to first denote the average over both the sampling of the Pauli to measure and the measurement of the Pauli, and later to just measure the Pauli. (1 mark)

Variance:

$$\mathrm{Var}\left(\hat{X}\right) \leqslant \frac{1}{N} \left\langle \frac{{\alpha_i}^2 {\lambda_i^{(j)}}^2}{{p_i}^2} \right\rangle = \frac{1}{N} \sum_i p_i \frac{\alpha_i^2 \langle {P_i}^2 \rangle}{{p_i}^2} = \frac{1}{N} ||\boldsymbol{\alpha}||_1 \sum_i \frac{{\alpha_i}^2}{|\alpha_i|} = \frac{||\boldsymbol{\alpha}||_1^2}{N} \tag{3}$$

(3 marks)

Discuss how this method compares to the others? When might you choose to use it?

(2 marks)

The scaling (at least of the upper bound) is the same as method 2. You might choose to use it if $N_i = p_i N$ is not an integer and so will induce rounding errors.

(e) Suggest two other methods you might use to estimate $\mathrm{Tr}[\rho H]$. What are the advantages / disadvantages of these methods? (3 marks)

1) Find a circuit to rotate into the eigenbasis of the Hamiltonian and measure in that basis (not possible in general!).

2) Group together terms that commute and measure those simultaneously.