Note on POVM and Control

Vincent Văn Dương
vv2102@nyu.edu
New York University

1 Introduction

In this note, we provide numerical evidence that quantum sensors are only efficient in a certain window of time. In particular, we sensed an external magnetic field $\boldsymbol{B}$ coupled to the quantum Ising model. Our sensor achieves a quantum advantage within a certain window of time; outside of this window, the sensor can not extract additional information about the external field parameters.

1.1 Baye's theorem

We will study how to construct a quantum measurement that allows parameters to be estimated with minimum variance. The approach is Bayesian, allowing us to account for prior knowledge of the parameters and the uncertainty in our observations. (This approach draws strong parallels to a partially observable Markov decision process (POMPD) [1, 2]. The POMPD framework could construct the optimal sequence of measurements that sharpens the posterior probability distribution.)

There are two sources of randomness in our problem: quantum and classical. Measuring a quantum state ρ leads to distinct outcomes with varying probabilities; and on the other hand, the parameter's prior distribution contributes to the posterior probability distribution.

Consider a quantum state ρ with POVM M_μ acting on the Hilbert space. (Each μ labels a different measurement outcome.) In a typical system, the parameters are encoded in its state: $\rho = \rho(\phi)$ and have a prior probability distribution ϕ. Following Baye's theorem, we update our estimates after measuring the system that minimises the variance (or increases the information gained). For example, after observing outcome μ, we can arbitrarily estimate the parameter to be ξ_μ. The square loss (or) is

$$\int \mathrm{d}\phi p(\phi|\mu)\varepsilon(\phi, \xi_\mu), \tag{1}$$

where $\varepsilon(\phi, \xi_\mu) = (\phi - \xi_\mu)^2$ is the square loss.

The cost function $\mathcal{C}$, following the work of [3], is the total variance of the parameters ϕ:

$$\mathcal{C} = \sum_\mu p(\mu) \int \mathrm{d}\phi p(\mu|\phi)p(\phi)\varepsilon(\phi, \xi_\mu). \tag{2}$$

The variable μ labels the possible outcomes from measuring the quantum system ρ. The prior probability density of the parameters is $p(\phi)$

1.2 Simple spin

We study a single spin-1/2 particle in a random magnetic field. We show that fluctuations in the random magnetic field lead to a long-time behaviour that is non-trivial. The Hamiltonian of the system is

$$H = \boldsymbol{B} \cdot \boldsymbol{L}, \tag{3}$$

where $\boldsymbol{B}$ is the external magnetic field, and $\boldsymbol{L}$ is the total angular momentum of the particle.

There are six kinds of measurements that we can make:

$$\begin{aligned}
M_1 &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\
M_2 &= \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\
M_3 &= \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \\
M_4 &= \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \\
M_5 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\
M_6 &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned} \tag{4}$$

Turning the crank, we find that

$$\int \mathrm{d}\phi p(\mu|\phi)p(\phi) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \frac{1}{3} - \frac{e^{-2t^2\sigma^2}}{3}(1 - 4t^2\sigma^2) \\ \frac{2}{3} + \frac{e^{-2t^2\sigma^2}}{3}(1 - 4t^2\sigma^2) \end{pmatrix}, \tag{5}$$

and

$$\int \mathrm{d}\phi p(\mu|\phi)p(\phi)\varepsilon(\phi, 0) = 3\sigma^2 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \frac{1}{3} - \frac{e^{-2t^2\sigma^2}}{9}(3 - 24t^2\sigma^2 + 16t^4\sigma^4) \\ \frac{2}{3} + \frac{e^{-2t^2\sigma^2}}{9}(3 - 24t^2\sigma^2 + 16t^4\sigma^4) \end{pmatrix} \tag{6}$$

When measurements are projectors of the initial conditions, the probability distribution varies in time due to fluctuations. This is an artefact that reappears in more complicated dynamics.

2 Optimisation Framework

The goal of numerics is to learn the POVM and estimates $\Theta = \left\{\widehat{M}_\mu, \boldsymbol{\xi}_\mu\right\}$ that minimise the cost function:

$$\begin{aligned} L(\Theta) = \sum_\mu \int d\boldsymbol{\theta} P(\boldsymbol{\theta}) \ \mathrm{Tr}\left[\widehat{M}_\mu e^{-i\widehat{H}(\boldsymbol{\theta})t} \hat{\rho} e^{+i\widehat{H}(\boldsymbol{\theta})t}\right] \left[\boldsymbol{\xi}_\mu - \boldsymbol{\theta}\right]^2 \\ \text{subject to} \\ \widehat{M}_\mu \succcurlyeq 0 \\ \sum_\mu \widehat{M}_\mu = \mathbb{I}. \end{aligned} \tag{7}$$

To minimise the cost function, we will use the typical Newton-Raphson Method, treating Θ as a one-dimensional vector parameterising all $\widehat{M}_\mu$ and $\boldsymbol{\xi}_\mu$. Expanding the loss function simplifies the problem:

$$L(\Theta) = \mathrm{Tr}\left[\sum_\mu \widehat{M}_\mu \left(\boldsymbol{\xi}_\mu^2 \hat{F}[1] - 2\boldsymbol{\xi}_\mu \cdot \hat{F}[\boldsymbol{\theta}] + \hat{F}[\boldsymbol{\theta}^2]\right)\right], \tag{8}$$

where the matrix-valued function $\hat{F}[...]$ is defined via integration

$$\hat{F}[g(\boldsymbol{\theta})] = \int d\boldsymbol{\theta} P(\boldsymbol{\theta}) e^{-i\widehat{H}(\boldsymbol{\theta})t} \hat{\rho} e^{+i\widehat{H}(\boldsymbol{\theta})t} g(\boldsymbol{\theta}). \tag{9}$$

The three matrices $\hat{F}[1], \hat{F}[\boldsymbol{\theta}], \hat{F}[\boldsymbol{\theta}^2]$ are evaluated to arbitrary accuracy via Monte Carlo integration. This is straightforward from sampling the prior:

$$\begin{aligned} \hat{F}[g(\boldsymbol{\theta})] &= \mathbb{E}_{\boldsymbol{\theta} \sim P(\boldsymbol{\theta})}\left[e^{-i\widehat{H}(\boldsymbol{\theta})t} \hat{\rho} e^{+i\widehat{H}(\boldsymbol{\theta})t} g(\boldsymbol{\theta})\right] \\ &\approx \frac{1}{N} \sum_{i=1}^N \left[e^{-i\widehat{H}(\boldsymbol{\theta}_i)t} \hat{\rho} e^{+i\widehat{H}(\boldsymbol{\theta}_i)t} g(\boldsymbol{\theta}_i)\right], \boldsymbol{\theta} \sim P(\boldsymbol{\theta}). \end{aligned} \tag{10}$$

The time-evolved density matrices

$$e^{-i\widehat{H}(\boldsymbol{\theta}_i)t} \hat{\rho} e^{+i\widehat{H}(\boldsymbol{\theta}_i)t} \tag{11}$$

can be calculated using a variety of techniques. The one used in this paper makes use of the singular value decomposition of $\hat{\rho}$:

$$\hat{\rho} = \hat{U}\hat{S}\hat{U}^\dagger. \tag{12}$$

Therefore,

$$e^{-i\widehat{H}(\boldsymbol{\theta}_i)t} \hat{\rho} e^{+i\widehat{H}(\boldsymbol{\theta}_i)t} = \left[e^{-i\widehat{H}(\boldsymbol{\theta}_i)t} \hat{U} \sqrt{\hat{S}}\right] \left[e^{-i\widehat{H}(\boldsymbol{\theta}_i)t} \hat{U} \sqrt{\hat{S}}\right]^\dagger. \tag{13}$$

In this form, the density matrix is the square of one matrix.

The loss function from Equation 7 is constrained: POVM must satisfy the two conditions. Positive semidefiniteness is simple to enforce since every positive semidefinite matrix admits a unique Cholesky decomposition:

$$\widehat{M}_\mu = \hat{L}_\mu \hat{L}_\mu^\dagger, \tag{14}$$

where $\hat{L}_\mu$ is a lower triangular complex matrix. That is, we re-parameterise the problem with these lower triangular matrices. The remaining constraint is non-trivial to enforce. To enforce the constraint, we relax the loss function with a term that penalises deviations from the constraint:

$$L(\Theta) \to L(\Theta) + \lambda \left\| \mathbb{I} - \sum_\mu M_\mu \right\|_2^2. \tag{15}$$

The upshot is that we must increase λ during the optimisation to enforce the constraint.

We may perform Newton-Raphson minimisation since the problem is amenable to auto-differentiation. Updating the parameters allow us to find a minima:

$$\Theta_a{}' = \Theta_a - \sum_b \left[H^{-1}\right]_{ab} D_b, \tag{16}$$

where

$$\begin{aligned} D_a &= \frac{\partial L}{\partial \Theta_a} \\ H_{ab} &= \frac{\partial^2 L}{\partial \Theta_a \partial \Theta_b}. \end{aligned} \tag{17}$$

3 Numerical Results

In this section, we provide evidence that optimal sensing can only be performed within a certain time interval. Inside the time interval, we can reduce the total uncertainty on the external magnetic field by measuring the system with a POVM. I.e., any POVM measurement outside this window yields no added benefit.

3.1 Set-up

The first step is to provide a Hamiltonian $\widehat{H}$ that describes our sensor. The Hamiltonian $\widehat{H}$ that we considered is the Ising model coupled to an unknown magnetic field $\boldsymbol{B}$:

$$\widehat{H} = J \sum_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z + \boldsymbol{B} \cdot \hat{\boldsymbol{L}}, \tag{18}$$

where $\boldsymbol{B}$ is the unknown external field and $\hat{\boldsymbol{L}}$ is the total angular momentum operator. In our analyses, the Ising coupling J is known. (We studied whether large or small J increase or reduce sensing abailities.)

The second step is to include our prior knowledge of the field that we want to sense. The only information available about the external field $\boldsymbol{B}$ is its prior $P(\boldsymbol{B})$, which we take to be a spherically symmetric Gaussian probability distribution:

$$P(\boldsymbol{B}) = \left(2\pi\sigma^2\right)^{-\frac{3}{2}} e^{-\frac{1}{2\sigma^2}\boldsymbol{B}^2}. \tag{19}$$

In the following analyses, we fix the width of the distribution: $\sigma = 0.1$.

The third step is to intialise the state of the system $\hat{\rho}$. For simplicity, we initialise each spin in the $|-z\rangle$ state. Though this corresponds to a pure state $\hat{\rho} = |\psi\rangle\langle\psi|$, we can (and should) consider mixed states.

The fourth step is to let the system evolve for a duration t:

$$\hat{\rho}(t) = e^{-i\widehat{H}t}\hat{\rho}e^{+i\widehat{H}t}. \tag{20}$$

In our analyses, the evolution time t is fixed. (We will find that there is an interval where sensing can be made optimal.)

The fifth, and final step is to make a judicious choice of POVM $\left\{\widehat{M}_\mu\right\}$ that provides the most information about the magnetic field $\boldsymbol{B}$. Put another way, we want to minimise the total uncertainty on the external field $\boldsymbol{B}$. This is the most challenging step because finding the optimal POVM is non-trivial. The technique used in this paper is outlined in the Numerics section.

In the following sub-sections, we chart the regions where there is a quantum advantage. These are regions where the cost (or loss) function can be reduced via a judicious choice of POVM and estimates. The regions where there is a quantum advantage take place when the evolution time t and Ising strength J take particular values – one should tune t and J to guarantee optimal performance.

3.2 Numerics

We studied the fully connected transverse-field Ising model. The initial state consisted of all $|-z\rangle$ states. The numerics were performed with 2, 3, and 4 qubits. All simulations show that a reduction of $0.014 - 0.040$ on the base value of 0.30. Achieving optimality requires a time t and Ising interaction J that is not too large. There is a strip of constant J that makes optimality robust in time. This is reminiscent of measuring the system near criticality as mentioned in

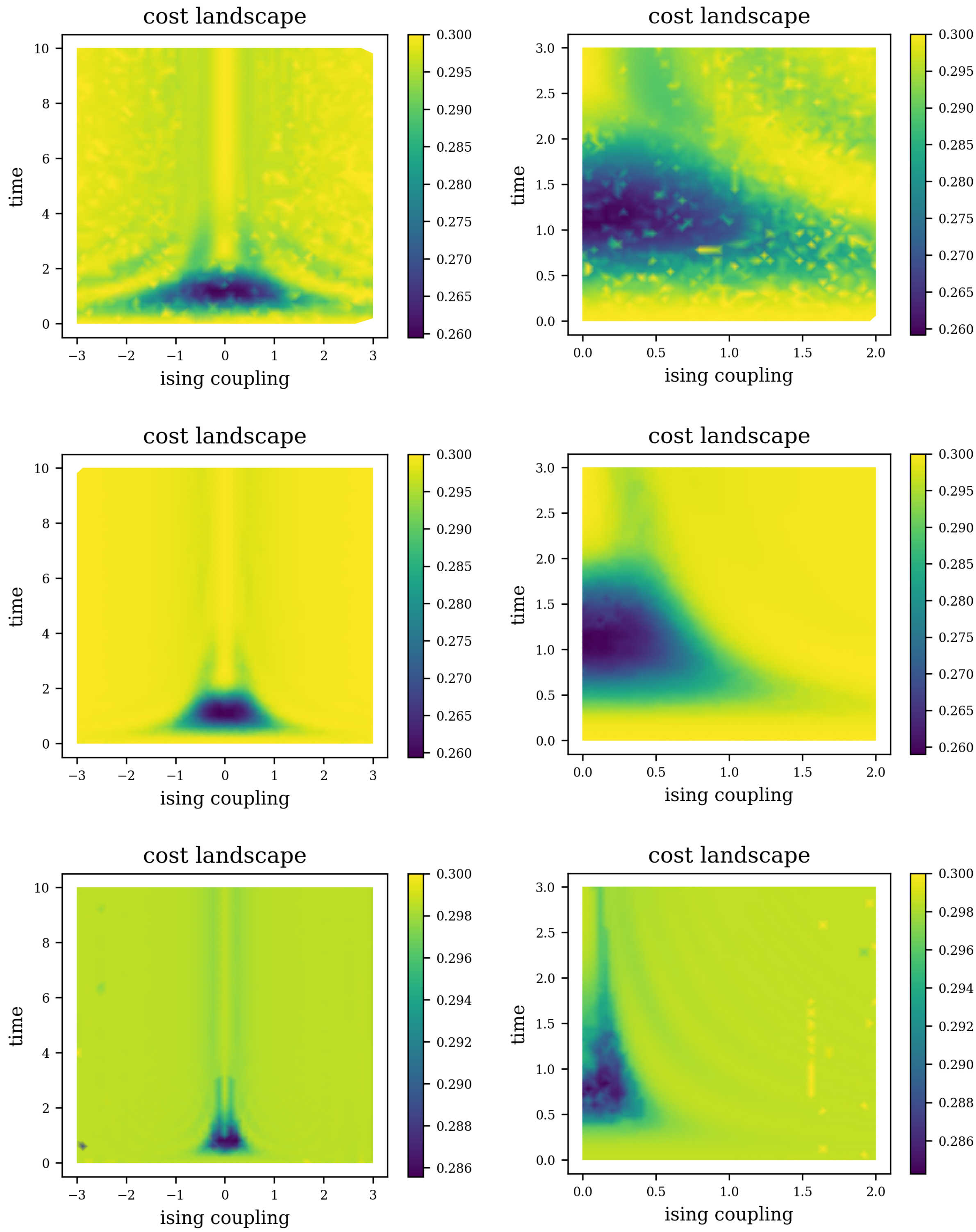

Figure 1: The cost landscape for two, three, and four qubit sensors. Reductions are only obtained for certain parameters of \$J\$ and \$t\$.

4 Comments

5 Conclusion

sec_conclusion.typ

Bibliography

[1] Åström, Karl Johan, “Optimal Control of Markov Processes with Incomplete State Information I,” vol. 10, pp. 174–205, 1965, doi: 10.1016/0022-247X(65)90154-X. [Online]. Available: https://lup.lub.lu.se/search/files/5323668/8867085.pdf

[2] E. Sondik, “The optimal control of partially observable markov process over the infinite horizon: discounted costs,” *Operations Res.*, vol. 26, pp. 282–304, 1978, doi: 10.1287/opre.26.2.282.

[3] R. Kaubruegger, A. Shankar, D. V. Vasilyev, and P. Zoller, “Optimal and variational multiparameter quantum metrology and vector-field sensing,” *PRX Quantum*, vol. 4, no. 2, p. 20333, Jun. 2023, doi: 10.1103/PRXQuantum.4.020333. [Online]. Available: https://link.aps.org/doi/10.1103/PRXQuantum.4.020333