AN ADAPTIVE BAYESIAN ALGORITHM FOR OPTIMAL QUBIT MEASUREMENT

KEVIN SHAIN

Department of Physics, Harvard University

1. Introduction

The premise of quantum computing is that one can initialize, operate upon, and read-out the quantum state of a qubit while using entanglement between qubits to enhance the scaling of information that can be processed. However, the energy scales needed to maintain quantum-ness are extremely small so small environmental fluctuations cause decoherence of the quantum state which means information is lost. The system explored in this paper is a single singlet-triplet spin qubit created by two gate-defined quantum dots in a GaAs/AlGaAs heterostructure. The spatial confinement of the electrons by the gates leads to atom-like energy levels and the two quantum dots are coupled electrostatically. In the spin qubit, one significant source of decoherence is ΔB_z, which represents how the magnetic field from the GaAs nuclei couples via the hyperfine interaction to the qubit. ΔB_z, along with electric field bias, J, allows manipulation around the Bloch sphere of possible states. However, unknown ΔB_z from the environment causes unknown changes in the state and thus loss of information. It is not critical to stay fixed at the same point on the Bloch sphere; one can avoid information loss by just monitoring how the environment will change the state. Thus, accurately estimating environmental sources of magnetic field, principally the nuclear spins in the GaAs, would allow the experimenter to keep track of changes in the state and allay one major source of decoherence.

Since the time scale for nuclear fluctuations is slow compared to operations on the qubit, one can use the qubit itself to estimate ΔB_z before running a quantum algorithm. This is done by initializing the qubit to a known state, letting it evolve for some time, and then measuring the state. Despite the quantum state existing anywhere on the Bloch sphere, each data point is a single measurement that can only take values of either $|S\rangle$ or $|T_0\rangle$. Each measurement is therefore a single Bernoulli trial with the probability of measuring a $|S\rangle$ or $|T_0\rangle$ depending on the evolution of the qubit on the Bloch sphere.

Though the value of ΔB_z varies slowly, assuming that it is completely stationary would limit the accuracy of its inferred value. The inference in this paper will allow for some deterministic drift of ΔB_z from its initial value as well as random diffusion of ΔB_z. The accuracy of the final estimated ΔB_z will then be limited by when the reduction in variance from an additional measurement is less than the increase in variance that occurs between measurements. Therefore, it is paramount that each measurement best reduce the variance in the estimate of ΔB_z. Current measurement schemes naively take measurements at linearly increasing evolution times. However, evolution time can be optimized based on the current state of knowledge. Therefore, the method described in this paper will infer a

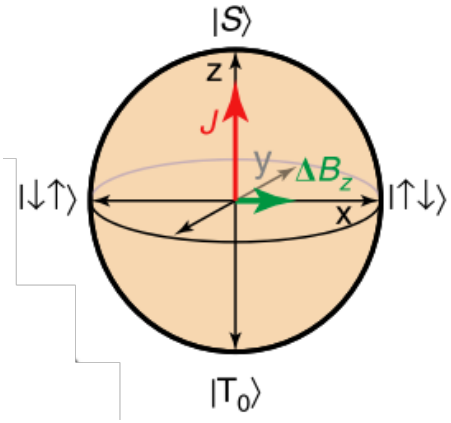

FIGURE 1. The Bloch sphere for a singlet-triplet spin qubit. The J and ΔB_z vectors are terms in the Hamiltonian that cause rotations of the qubit state about their respective axes. J originates from electric potential differences between the two quantum dots. ΔB_z originates from hyperfine coupling to nuclear spins.

distribution of the parameter ΔB_z after each measurement and then estimate the evolution time for the next measurement that will most reduce the variance of the estimated ΔB_z.

2. Statistical Model

The data from each experiment is very manageable. Each experiment has a measurement record of around 100 points of which the independent variable is the evolution time of the qubit between preparation and measurement and the dependent variable is a single-shot measurement outcome. The data is small, but computational efficiency is still extremely important because an estimate of ΔB_z is only useful if it can be computed before the value of ΔB_z changes.

The underlying model of how the qubit state evolves and is measured is very simple. If the singlet state is called $|+\rangle$ and the triplet state is called $|-\rangle$, then the wavefunction is,

$$|\psi(t\rangle = \cos(2\pi\Delta B_z t/2)\,|+\rangle - i\sin(2\pi\Delta B_z t/2)\,|-\rangle \tag{1}$$

such that,

$$P(+|\Delta B_z) = \cos^2(2\pi\Delta B_z t/2) \tag{2}$$

$$P(-|\Delta B_z) = \sin^2(2\pi\Delta B_z t/2) \tag{3}$$

Thus, if the measurement result, m, is denoted +1 if $|+\rangle$ is measured or -1 if $|-\rangle$ is measured, then,

$$P(m|\Delta B_z) = \frac{1}{2}\left(1 + m\cos(2\pi\Delta B_z t)\right) \tag{4}$$

This simple model of the probabilities doesn't consider any errors, but it is important to note that it will still generate an unknown measurement outcome with defined probabilities so no deterministic model will fit the data exactly, even if ΔB_z is known exactly.

One source of error is measurement error. Essentially, the readout of the qubit is done by measuring voltage across another quantum dot. This gives well differentiated Gaussian peaks of voltage depending on the qubit state. These peaks are then thresholded so that the singlet(triplet) peak is marked +(-). This thresholding process is imperfect as there is

some small chance that a tail from the singlet Gaussian will leak into the triplet region or vice versa which creates measurement error. With measurement error, the probability of measuring a + is the probability of measuring + correctly plus the probability of measuring - incorrectly. If the probability of incorrectly measuring +(-) is defined as $\eta_+(\eta_-)$, this gives,

$$P(+|\Delta B_z) = (1-\eta_+)\cos^2(\pi\Delta B_z t) + \eta_- \sin^2(\pi\Delta B_z t) \tag{5}$$

$$P(-|\Delta B_z) = \eta_+ \cos^2(\pi\Delta B_z t) + (1-\eta_-)\sin^2(\pi\Delta B_z t) \tag{6}$$

State preparation error can be accounted for in a similar way. The initial state would ideally be purely singlet every time, but it actually is prepared in the triplet state some small percentage of the time. This means that the initial density matrix is,

$$\begin{pmatrix} 1-\epsilon & 0 \\ 0 & \epsilon \end{pmatrix}$$

Accounting for just state preparation error gives,

$$P(+|\Delta B_z) = (1-\epsilon)\cos^2(\pi\Delta B_z t) + \epsilon \sin^2(\pi\Delta B_z t) \tag{7}$$

$$P(-|\Delta B_z) = \epsilon \cos^2(\pi\Delta B_z t) + (1-\epsilon)\sin^2(\pi\Delta B_z t) \tag{8}$$

Finally, there is axis of rotation error. This accounts for small differences in the electric potential between the two quantum dots that make up the qubit. This shifts the axis of rotation on the Bloch sphere about which the state vector rotates. Defining $\delta \equiv \cos^2(\theta)$ with θ being the angle of ΔB_z away from the x-axis, this error alone leads to,

$$P(+|\Delta B_z) = \cos^2(\pi\Delta B_z t) + \delta \sin^2(\pi\Delta B_z t) \tag{9}$$

$$P(-|\Delta B_z) = (1-\delta)\sin^2(\pi\Delta B_z t) \tag{10}$$

Putting all of the errors together and using $\cos^2(x) - \sin^2(x) = \cos(2x)$ yields,

$$P(+|\Delta B_z) = \eta_- + \frac{1}{2}(1-\eta_+-\eta_-)[1+(1-2\epsilon)(\delta+(1-\delta)\cos(2\pi\Delta B_z t))] \tag{11}$$

$$P(-|\Delta B_z) = \eta_+ + \frac{1}{2}(1-\eta_+-\eta_-)[1-(1-2\epsilon)(\delta+(1-\delta)\cos(2\pi\Delta B_z t))] \tag{12}$$

Combining these probabilities into one using the fact that the result of any measurement is $m_k = \pm 1$,

$$P(m_k|\Delta B_z) = \eta_- + \frac{1}{2}(1-\eta_+-\eta_-)[1+m_k(1-\epsilon)(\delta+(1-\delta)\cos(2\pi\Delta B_z t))] \tag{13}$$

Reparameterizing the long time-scale errors using,

$$\alpha = \eta_- - \eta_+ + (1-\eta_- - \eta_+)(\delta - 2\epsilon\delta) \tag{14}$$

$$\beta = (1-\eta_- - \eta_+)(1-\delta)(1-2\epsilon) \tag{15}$$

results in,

$$P(m_k|\Delta B_z) = \frac{1}{2}[1+m_k(\alpha+\beta\cos(2\pi\Delta B_z t))] \tag{16}$$

Though it seems that reparameterizing the errors sacrifices clarity, it turns out that α and β are more easily extracted than the errors individually since repeatedly measuring immediately after preparation yields $p(+) = \frac{1}{2}[1+\alpha+\beta]$ and making many measurements at random times and averaging gives $p(+) = \frac{1}{2}[1+\alpha]$

3. Inference Framework

The goal of choosing an evolution time for the next measurement that reduces the variance of the estimate of ΔB_z is most naturally handled in the Bayesian framework as the variance can be extracted from the posterior distribution for ΔB_z. One can imagine the posterior being brute-force evaluated at a range of values for ΔB_z and then a variance being calculated. Though brute-force evaluation of the posterior is made more reasonable due to the one-dimensional posterior, (α, β, and the drift and diffusion coefficients being device, not time dependent can be inferred over many experiments), it would still be a lengthy computation. This is made more computationally expensive by the fact that expected variance after the next measurement requires simulating measurements at many evolution times and evaluating the posterior for each simulated next measurement. Therefore, any sensible implementation of adaptive measurement requires calculating the variance of the posterior without its explicit discretization and evaluation.

Both a prior and a likelihood function are needed to express the posterior. Based on previous experiments, a conservative prior can set to be uniform in the range of $50MHz < \Delta B_z < 100MHz$ and zero everywhere else. For the likelihood, the reparameterized single measurement likelihoods can be multiplied to get a likelihood function for an entire measurement series. This is somewhat more complicated than just taking a product because of the drift and diffusion of ΔB_z between measurements. With a drift coefficient, v, a diffusion coefficient, D, and the time between measurements, Δt, the likelihood of a series is,

$$\begin{aligned} P(\{m\}|\Delta B_{z_0}, v, D) = &\prod_k \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi Dk}} e^{-[\Delta B_{z_k} - (\Delta B_{z_0} + vk\Delta t)]^2/4Dk} \\ &\times \frac{1}{2}[1 + m_k(\alpha + \beta\cos(2\pi(\Delta B_{z_0} + vk\Delta t)t)]d\Delta B_{z_k} \end{aligned} \tag{17}$$

with the integral coming from the distribution of ΔB_{z_k} values caused by diffusion, given by,

$$\Delta B_{z_k} = \frac{1}{\sqrt{4\pi Dk}} e^{-[\Delta B_{z_k} - (\Delta B_{z_0} + vk\Delta t)]^2/4Dk} \tag{18}$$

To actually implement this, the Gaussian integral for $P(m_k|\Delta B_{z_0}, v, D)$ can be evaluated by decomposing the cosine term into exponentials and then integrating the complex Gaussian analytically. The result for a single measurement is,

$$\begin{aligned} P(m_k|\Delta B_{z_0}, v, D) = &\frac{1}{2}\sqrt{2Dk} e^{-(2\pi t)^2(2Dk)/2} \beta m_k \cos(2\pi(\Delta B_{z_0} + vk\Delta t)t) \\ &+ \frac{1}{2}(1 + \alpha * m_k) \end{aligned} \tag{19}$$

Neglecting drift and diffusion, though this is not necessary, the likelihood has only cosine terms. Since the prior is uniform, the posterior is also just a function of cosines. This allows the product of cosines to be written as a sum of cosines by repeated use of the identity, $\cos(a)\cos(b) = \frac{1}{2}[\cos(a+b) + \cos(a-b)]$ and the coefficients of the cosine terms can be interpreted as Fourier coefficients. Since there are a finite number of terms in the product, the posterior can be represented exactly by a finite Fourier series. Representing the posterior as a Fourier series will eventually allow the calculation of $\langle \Delta B_z \rangle$ and $\langle \Delta B_z^2 \rangle$ just from the Fourier coefficients using Parseval's theorem. This allows the posterior to be

written as,

$$P_k(\Delta B_z|\{m\}) = A\, P_0(\Delta B_z) \prod_{i=1}^{k} P_i(m_i|\Delta B_z) = A \sum_{q=-K}^{K} c_k(q) \cos(2\pi\tau q \Delta B_z) \tag{20}$$

where A is just a normalization constant, $P_0(\Delta B_z)$ is the uniform prior, τ is the minimum measurement time, and $K = \sum_{i=1}^{k} m_i$ represents the maximum number of terms necessary to represent the posterior exactly. In this experiment, a reasonable value for τ is 10 ns which gives a bandwidth of 50 MHz. The posterior can easily be normalized in the Fourier picture by dividing through by $\frac{c_k(0)}{2\tau}$ to yield,

$$P_k(\Delta B_z|\{m\}) = \sum_{q=-K}^{K} \frac{c_k(q)}{2\tau c_k(0)} \cos(2\pi\tau q \Delta B_z) \tag{21}$$

The Fourier coefficients are calculated using an update procedure after each measurement. Based on the $\cos(a+b)+\cos(a-b)$ that comes from the multiplication of cosines, each coefficient is calculated from a combination of its previous value and coefficients shifted by $\pm n_{k+1}$ with n_{k+1} determined by $t_{k+1} = n_{k+1}\tau$. The drift and diffusion can be handled naturally in the Fourier domain by taking the Fourier transform of the drift and diffusion operators. Thus, for the drift operator, $\Delta B_z \to \Delta B_z + v\Delta t$ acting on the Fourier coefficients is,

$$c(q) \to e^{-i(2\pi\tau q)v\Delta t} c(q) \tag{22}$$

The diffusion operator for the Fourier coefficients is,

$$c(q) \to e^{-D(2\pi\tau q)^2 \Delta t} c(q) \tag{23}$$

Therefore, the total coefficient update equation is,

$$c_{k+1}(q) = e^{-D(2\pi\tau q)^2\Delta t - i(2\pi\tau q)v\Delta t}\{(1 + m_{k+1}\alpha)c_k(q) + m_{k+1}\beta[c_k(q + n_{k+1}) + c_k(q - n_{k+1})]\} \tag{24}$$

where k is the measurement index and q is the Fourier series index. Since the drift introduces a complex exponential, the Fourier coefficients are now explicitly complex which necessitates using cosine and sine terms in the Fourier series. Thus, the posterior is,

$$P_k(\Delta B_z|\{m\}) = \mathrm{Re}\left[\sum_{q=-K}^{K} \frac{c_k(q)}{2\tau c_k(0)} e^{i2\pi\tau q \Delta B_z}\right] \tag{25}$$

Evaluating expectation values from the posterior is now the convolution of the posterior with the quantity to be averaged. For example,

$$\langle \Delta B_z \rangle = \int_{\frac{1}{2\tau}}^{\frac{2}{2\tau}} \Delta B_z P_k(\Delta B_z|\{m\}) d\Delta B_z \tag{26}$$

where the limits of integration are designed to span the bandwidth of the Fourier series, 50 MHz, with bounds determined by the prior which is uniform and nonzero from 50 MHz to 100 MHz. Using Parseval's theorem, the convolution integral can be expressed as the sum of the product of ΔB_z and $P_k(\Delta B_z|\{m\})$ Fourier coefficients. Using the fact that the

Fourier coefficients of ΔB_z are $c_{\Delta B_z}(q) = \frac{(-1)^q - 1}{(\pi q)^2}$ and the boundary conditions gives,

$$\begin{aligned}\langle \Delta B_z \rangle =& \frac{\left(\frac{2}{2\tau}\right)^2 - \left(\frac{1}{2\tau}\right)^2}{\left(2 * \frac{1}{2\tau}\right)} - \sum_{q=-K,\, q\neq 0}^{K} \frac{c_k(q)}{2\tau c_k(0)} \frac{(-1)^q - 1}{(\pi q)^2} \\ =& \frac{3}{4\tau} - \sum_{q=-K,\, q\neq 0}^{K} \frac{c_k(q)}{2\tau c_k(0)} \frac{(-1)^q - 1}{(\pi q)^2}\end{aligned} \tag{27}$$

A very similar technique can be used to calculate $\langle \Delta B_z^2 \rangle$ from just the Fourier coefficients using $c_{\Delta B_z^2}(q) = \frac{(-1)^q}{(\pi q)^2}$. The quantity of interest, variance in ΔB_z, can then be written as,

$$\begin{aligned}\sigma^2_{\Delta B_z} &= \langle \Delta B_z^2 \rangle - \langle \Delta B_z \rangle^2 \\ &= \left[\frac{1}{3 * (2\tau)^2} + \sum_{q=-K,\, q\neq 0}^{K} \frac{c_k(q)}{2\tau^2 c_k(0)} \frac{(-1)^q}{(q\pi)^2}\right] - \left[\frac{1}{2 * 2\tau} + \sum_{q=-K,\, q\neq 0}^{K} \frac{c_k(q)}{2\tau c_k(0)} \frac{[(-1)^q - 1]}{(q\pi)^2}\right]^2\end{aligned} \tag{28}$$

4. Adaptive Evolution Time Algorithm

Because of the probabilistic nature of the measurement, it is impossible to know the variance that will be obtained after a particular measurement. However, for any evolution time, the variance in either measurement outcome can be calculated and the expected value of the variance can be written as the weighted sum,

$$\langle \sigma^2 \rangle_{n_k} = \frac{c_{n_k+}(0)\sigma^2_{n_k+} + c_{n_k-}(0)\sigma^2_{n_k-}}{c_{n_k+}(0) + c_{n_k-}(0)} \tag{29}$$

where $c_{n_k+}(0)$ is the zeroth Fourier coefficient after a measurement of + at time $n_k\tau$ and $\sigma^2_{n_k+}$ is the variance after that hypothetical measurement. Before each measurement, the algorithm iterates through the possible values of n_k, calculates an expected variance for each, and selects the evolution time with the minimum expected variance. While this does involve discretization, one only has to consider approximately 100 values of n_k before the evolution time gets so long that its timescale is similar to that of the ΔB_z fluctuations. This is a much less costly discretization than evaluating the posterior at various ΔB_z. All the work representing the posterior as a Fourier series is useful because it allows calculating the variance in ΔB_z without ever explicitly evaluating the posterior. In practice, calculating the expected variance at 100 values of n_k takes on the order of 1 second using Python. Since it is naturally parallelizable, this could be executed in a fraction of a second and possibly faster than the 4 μs between measurements with an optimized implementation. It is important to note that the adaptive evolution time algorithm will not necessarily outperform a predetermined evolution time for any given measurement. However, it should on average reduce the variance more quickly as a function of the number of measurements and the variance should saturate at a lesser value where the reduction in variance from an optimized measurement equals the increase in variance from diffusion.

5. Results and Discussion

The most important figure of merit for a measurement scheme is the mean squared error between the estimated value of ΔB_z and the actual ΔB_z. The most robust way to measure the error would be to implement the scheme and determine if the estimated ΔB_z

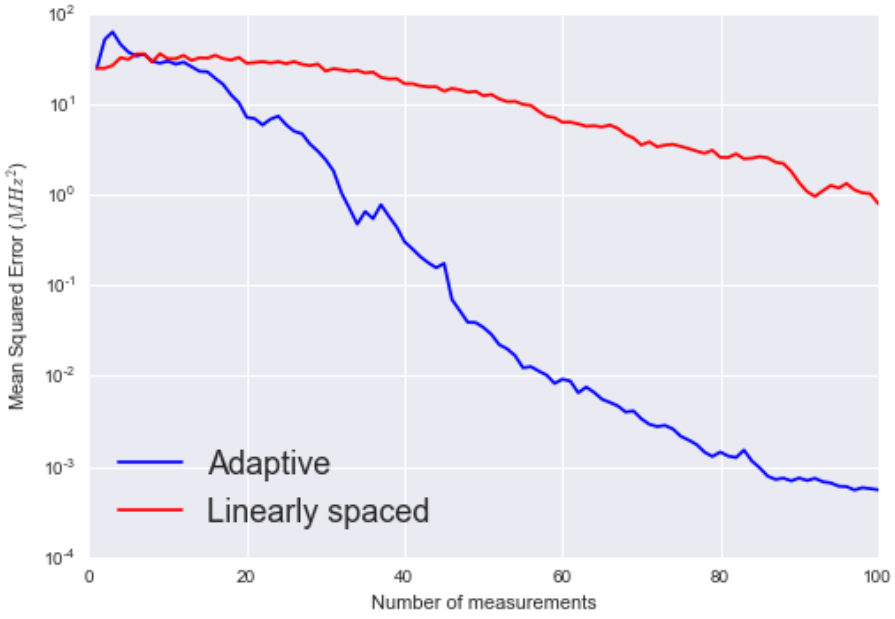

FIGURE 2. Plot of the mean squared error in estimates of ΔB_z with adaptive evolution times and linearly spaced evolution times. This data is averaged over 100 adaptive experiments and 1000 linearly spaced experiments.

is more accurate based on whether it increases the coherence time of the qubit. However, any implementation would have to be programmed on a FPGA if there is any hope of it running within the measurement time of the qubit, approximately 4 μs. Short of a way to prove the efficacy using real data, one must rely on simulated data to effectively slow down time and allow an unoptimized algorithm to run between each measurement. To this end, a time series of simulated ΔB_z values is generated using specified drift and diffusion coefficients. Based on this series of simulated ΔB_z, a probabilistic measurement is performed using the current ΔB_z and the evolution time determined by the measurement scheme. Though the accuracy of the estimated ΔB_z after all measurements are made is most important, the error can be calculated after each measurement to monitor how it decreases over time. By repeating the experiment many times, a MSE can determine the performance of each measurement scheme.

As seen in Figure 2, the adaptive measurement algorithm outperforms the currently used linearly spaced evolution times by over three orders of magnitude in MSE after one hundred measurements. The underlying ΔB_z was chosen to have a typical starting frequency of 70 MHz and diffusion constant of 6.7 kHz$^2/\mu$s. There was no drift included to reduce simulation time by making the coefficients purely real. However, this has no effect since drift is deterministic and therefore does not change the variance between measurements. The adaptive algorithm achieves similar accuracy in less than half the measurements, therefore less than half of the time. This could be very important as there are likely unknown sources of noise that will have more of an effect the more measurements must be performed. If it can be implemented within the 4 μs measurement time, the adaptive evolution time algorithm would likely lead to significant increases in qubit coherence time.

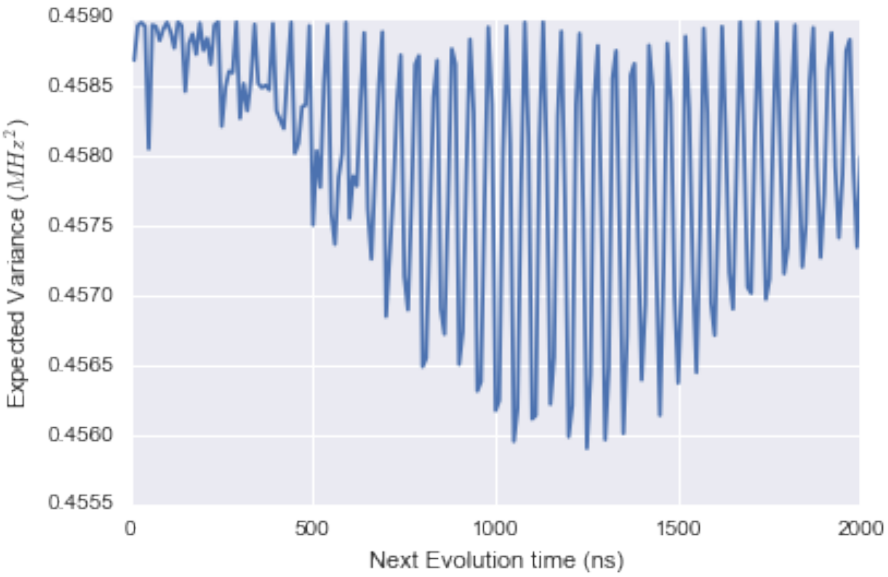

FIGURE 3. Expected variance versus next evolution time after 50 measurements. Oscillations reflect that measuring at different phases of the underlying cosine likelihood function differently constrain the estimate of ΔB_z.

It is worth noting that an evolution time series generated from the adaptive algorithm cannot be used generally for future experiments. The expected variance as a function of next evolution time changes significantly based on the current estimate of ΔB_z. While it seems that a variety of evolution times is best, Figure 3 indicates that the expected variance after a measurement oscillates significantly on the time scale of 10 ns. This means that the optimal evolution times for one experiment are likely to be highly non-optimal for another experiment.

6. CONCLUSIONS

Decoherence due to coupling to nuclear spins has long been an obstacle to quantum information processing with spin qubits. Hamiltonian parameter estimation has recently been used successfully to increase coherence times. However, the lack of data available within the time scale of the experiment severely limits the accuracy of estimates using predetermined linearly increasing evolution times. The adaptive Bayesian algorithm described in this paper allows for more informative measurements by considering the current state of knowledge. Furthermore, the Fourier series representation could be useful in many other quantum computing contexts that have likelihoods made of a product of cosine terms. The current algorithm uses the known structure of posterior to calculate expectation values from the posterior exactly using a finite number of terms. This enables computation that scales with the number of data points, not the desired precision when discretizing the posterior, which could make it valuable to many problems of estimation with small datasets.

References

[1] Alexandr Sergeevich, Anushya Chandran, Joshua Combes, Stephen D. Bartlett, and Howard M. Wiseman, *Characterization of a qubit Hamiltonian using adaptive measurements in a fixed basis*, Physical Review A - Atomic, Molecular, and Optical Physics **84** (2011), no. 5, 1–5, available at 1102.3700.

[2] M D Shulman, S P Harvey, J M Nichol, S D Bartlett, A C Doherty, V Umansky, and A Yacoby, *Suppressing qubit dephasing using real-time Hamiltonian estimation.*, Nature communications **5** (2014), no. May, 5156, available at arXiv:1405.0485v1.