Hamiltonian Learning for Parameter Estimation

Tejas Naik

Supervisor: Prof. David Cory

ABSTRACT

In this project, we study how Hamiltonian learning based on Bayesian Inference can be used to estimate unknown parameters of a Quantum system. We use a numerical method to perform the Bayesian Inference known as Sequential Monte Carlo Algorithm. We study two problems – one, the estimation of Larmor frequency of a spin and the second estimating the dipolar coupling strength between protons in crystalline Gypsum. Finally, we see how the above algorithm could help in the molecular simulations of protein-protein complexes.

I. INTRODUCTION

One of the most important tasks in Quantum Computing is to characterize the dynamics of a physical Quantum system. An efficient and reliable characterization of the process gives information about how well the system performs and what the errors are in the system. The most common technique used to perform this characterization is Quantum Process Tomography. With Quantum Process Tomography, we essentially assume that the Process into consideration is a “Black Box”, meaning we have no information about the dynamics of the system. The information about the dynamics is extracted by evolving a set of initial states which spans the whole Hilbert space of the system in consideration and perform a set of tomographically complete set of measurements. These measurements when processed with methods

such as Linear Inversion, Maximum Likelihood Estimation gives information about the system dynamics in terms of what are known as ‘Chi’ matrices.

Even though Quantum Process Tomography gives a complete characterization of the process, the number of measurements required to characterize a system, scales exponentially with the number of qubits. Further, even though the assumption of the process being a Black Box is a very conservative one, in reality we have some insights about the dynamics of the system. This insight is typically a Hamiltonian which we expect to explain the dynamics we observe in the physical system. Thus, the problem of characterization boils down to finding the optimal parameters of the Hamiltonian such that the Hamiltonian best describes a set of observed measurement results. This process of Hamiltonian learning is also sometimes called as partial process estimation or partial tomography.

We perform the Hamiltonian learning using a technique from Statistics – Bayesian Inference and use the Sequential Monte Carlo Algorithm for estimation of parameter. For the project we look at estimation of a single unknown parameter in the Hamiltonian. In section (II) we start by describing the protocol of Bayesian Inference and give the algorithm for Sequential Monte Carlo (SMC) Algorithm. In section (III), we implement the SMC for estimating the Larmor frequency of spin using a spin-locking pulse experiment. In section (IV), we implement the SMC for estimating the dipolar coupling present between protons on a water molecule in Gypsum. In section (V), we give an outlook on the SMC algorithm and compare the advantages and disadvantages with other methods. Finally, we look at the relevance of using SMC for the dipolar coupling estimation and conclude the report in section (VII).

II. BAYESIAN INFERENCE

Bayesian Inference is inferring the probability distribution of a random variable using Bayes' Theorem. This is a commonly used technique in Statistics and Probability Theory. Here, we will use it for inferring the probability distribution of unknown Hamiltonian Parameter. Let's say we performed some experiment on the system dynamics to be explored and recorded the data D(C). Here C are control parameters. These include frequencies of drive pulses, Rabi strengths, time etc. Thus, D would essentially be a list of data points $\{d_i\}$ recorded at control parameters $\{c_i\}$. For the experiments we consider, the data recorded is the free induction decay (f.i.d) spectra, and thus the only variable control parameter for the data would be time, $\{t_i\}$. Let, $f(\theta, t)$ be the model Hamiltonian for our system. Here, θ is the unknown parameter we want to estimate. Let, I be the prior information we have about the parameter. This information could be the bounds on the parameter or any other information we have about the parameter beforehand. Thus, our task is essentially to find out the probability distribution $P(\theta|D, I)$, i.e. the probability of the parameter given the recorded data and the Prior information. This probability is known as the posterior distribution and can be find out using Bayes' theorem –

$$P(\theta|D, I) = \frac{P(D|\theta)P(\theta|I)}{P(D|I)} \tag{1}$$

The first term on the right-hand side, $P(D|\theta)$ is the likelihood of the data. It essentially gives what's the probability that for the parameter value θ, our model Hamiltonian produces the data D. The second term, $P(\theta|I)$ is the distribution of the unknown parameter θ using the prior information I. The denominator, $P(D|I)$ is just a normalization term and can be ignored for our analysis. We will now derive a form for

calculating the likelihood as per the reference [4]. If we model the noise in the system as $\{n_i\}$, the recorded data can be written as,

$$d_i = f(\theta, t_i) + n_i \tag{2}$$

Given a standard deviation σ for the distribution of noise parameters n_i, we can assume that the noise is a Gaussian distribution with mean 0 and standard deviation σ,

$$P(n|\sigma) = \prod_{i=1}^{N} (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{n_i^2}{2\sigma^2}\right) \tag{3}$$

Here, N is the size of data D. Thus, given the noise parameters n and σ, either the model plus the errors equals the data or it does not equal to the data. Thus, the likelihood given the error parameters becomes a product of delta functions,

$$P(D|\theta, I, n, \sigma) = \prod_{i=1}^{N} \delta(d_i - f(\theta, t_i) - n_i) \tag{4}$$

In order to remove the conditional dependence on n and σ, we can use the sum rule, where we integrate the probabilities over the parameter we want to eliminate,

$$P(D|\theta, I) = \int dn d\sigma P(n|\sigma) P(\sigma|I) P(D|\theta, I, n, \sigma) \tag{5}$$

For the prior distribution of σ, we assume Jeffrey's Prior[4],

$$P(\sigma|I) \propto \frac{1}{\sigma} \tag{6}$$

As we have delta functions in the conditional likelihood, the integral over n gives,

$$P(D|\theta,I) \propto \int d\sigma(2\pi\sigma^2)^{-1/2}\sigma^{-1}\exp\left(-\frac{Q}{2\sigma^2}\right)$$

$$P(D|\theta,I) \propto \int d\sigma\sigma^{-N-1}\exp\left(-\frac{Q}{2\sigma^2}\right) \quad (7)$$

Where,

$$Q = \sum_{i=1}^{N}(d_i - f(\theta,t_i))^2 \quad (8)$$

For a given standard deviation of error, we see that the likelihood is a Gaussian distribution. Integrating equation (7) over σ can be solved using Gamma functions and gives,

$$P(D|\theta,I) \propto \left(\frac{Q}{2}\right)^{-\frac{N}{2}} \quad (9)$$

This distribution is also known as the student's T distribution. Now, we look at a numerical method, known as Sequential Monte Carlo Algorithm, to perform the above stated Bayesian Inference and estimate the unknown parameter distribution.

Sequential Monte Carlo Algorithm:

We start by considering the prior distribution $P(\theta|I)$. Let, $\theta_1 < \theta < \theta_2$ be the prior information we have about the unknown parameter. We need to construct uninformative prior distributions using this prior information. Here, we consider two prior distributions,

1. Uniform Distribution between (θ_1, θ_2).
2. Flat Gaussian Distribution with mean $(\theta_1 + \theta_2)/2$ and a large standard deviation $(\theta_2 - \theta_1)/3$.

Using these priors, we would sample n points. In literature these points are called as particles and we will store the location of these particles

(the value of the parameters) in an array X. Without loss of generality, we would assume that the respective probability for each particle is equal to $1/n$ and would store these weights in an array Ω. Once, we calculate the likelihood of data using our model Hamiltonian, $P(D|\theta, I)$, we will use the Bayes' formula to update the weights of the particles,

$$\Omega_f(X_i) = P(D|X_i, I)\Omega_I(X_i) \tag{10}$$

where, $\Omega_f(X_i)$ is the updated weight of particle at location X_i and $\Omega_I(X_i)$ is the initial weight of particle at location X_i, which is $1/n$. We iteratively update the weights, by taking the updated weights as initial weight for our next iteration until we the weight starts to saturate. Also, with each update as the likelihood probabilities are not normalized, we will need to normalize the weights after each iteration. The estimate of the parameter then becomes the weighted average of the particles with the weights at that iteration,

$$Estimate = \mu = \sum_i \Omega_f(X_i)X_i \tag{11}$$

Similarly, the variance becomes,

$$\Sigma = \left(\sum_i \Omega_f(X_i){X_i}^2\right) - \mu^2 \tag{12}$$

Thus, using the variance if the final distribution shows Gaussian-like distribution, we can also calculate confidence intervals. There is one issue which limits the numerical precision of the algorithm, which is the issue of zero weights. As we are using a finite number of points to predict a continuous distribution, the accuracy of our distribution is highly dependent on the number of points in our analysis. Within, a few iterations of Bayesian updates, the particles which are far from the original distribution, would get zero or near-zero weights and thus

would essentially be knocked out of the algorithm. The effective number of particles would thus decrease with each iteration and would affect the numerical precision of the estimation. In order to circumvent this problem, we do resampling of the particles. When the effective number of particles, which can be written in terms of weights as,

$$n_{eff} = \frac{1}{\sum_i (\Omega_f(X_i))^2} \tag{13}$$

becomes less than some threshold $n_{threshold}$, we resample the particles. The resampling is done through the probability distribution given by weighted sum of Gaussians,

$$p(x') = \sum_i \Omega_f(X_i) \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x' - \mu_i\ )^2}{\sigma}\right) \tag{14}$$

Where the means μ_i is given by a linear combination of particle position $X_i$and the mean μ,

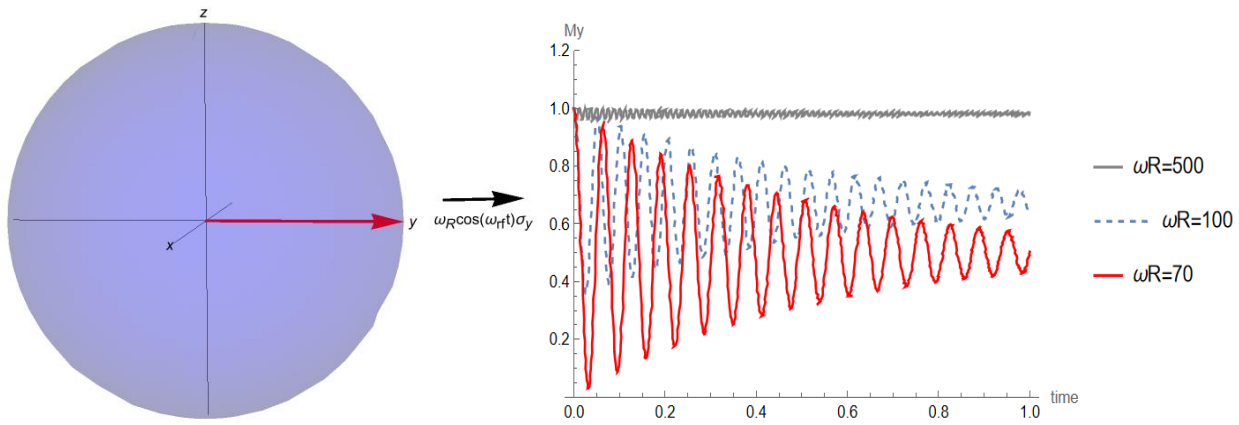
$$\mu_i = aX_i + (1-a)\mu \tag{15}$$

and variance σ are given by,

$$\sigma = (1-a^2)\Sigma \tag{16}$$

Here a is a hyperparameter whose value we will use as 0.98 as used in reference [2]. What we essentially are doing is considering a Gaussian distribution around each particle position with mean μ_i and variance σ, and each of these Gaussian distributions are weighted by the weights of the particle positions. Once we resample all the particles, the weight vector is again reset to 1/n, and the algorithm is resumed.

In the next section we implement the SMC algorithm for estimation of Larmor frequency.



(a) (b)

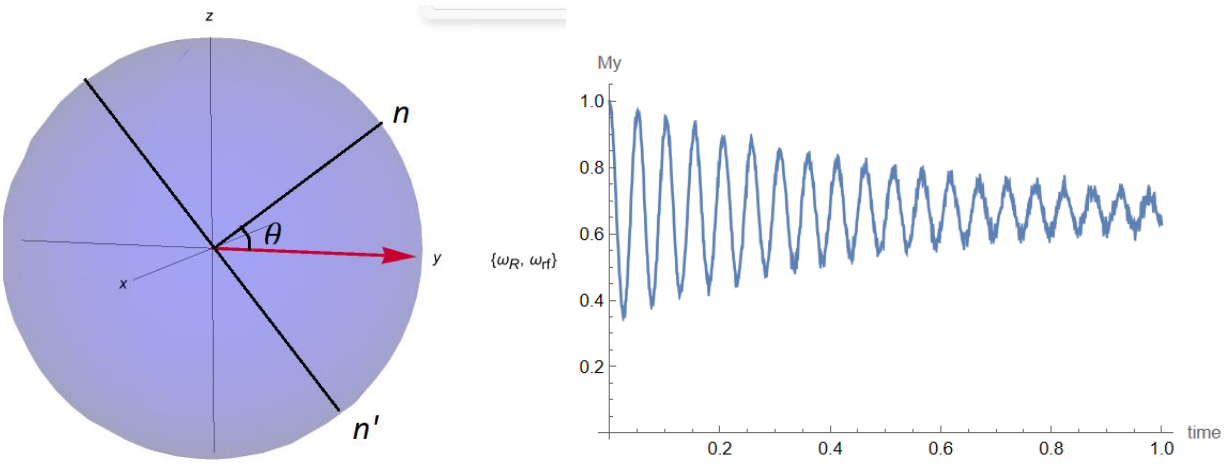
Fig.1. (a) Bloch sphere representation of initial spin state along y-direction and spin-locking pulse along y-direction. (b) Magnetization measured for different ω_R, keeping $\Delta\omega = 70\ s^{-1}$.

III. LARMOR PRECESSION

In this section we estimate the Larmor frequency of a spin, by collecting the data using a spin-locking experiment and then using the SMC. In a spin-locking experiment, a strong RF-field is applied along a certain direction in the transverse plane, which locks the spin in that direction in the rotating frame of reference, depending on the strength of the pulse with respect to the detuning. This can be understood by looking at the Hamiltonian in the rotating frame of reference,

$$H = -\frac{(\omega_0 - \omega_{rf})}{2}\sigma_z - \frac{\omega_R}{2}\sigma_y \qquad (17)$$

Here, ω_R and ω_{rf} are the Rabi strengths and the frequency of the RF field applied in the y-direction. ω_0 is the Larmor frequency we need to estimate. Applying the RF field at frequency ω_{rf} would thus create a detuning of $\Delta\omega = \omega_0 - \omega_{rf}$. Now, let's say we start with the initial state σ_y and apply the RF field in the y-direction, then the net

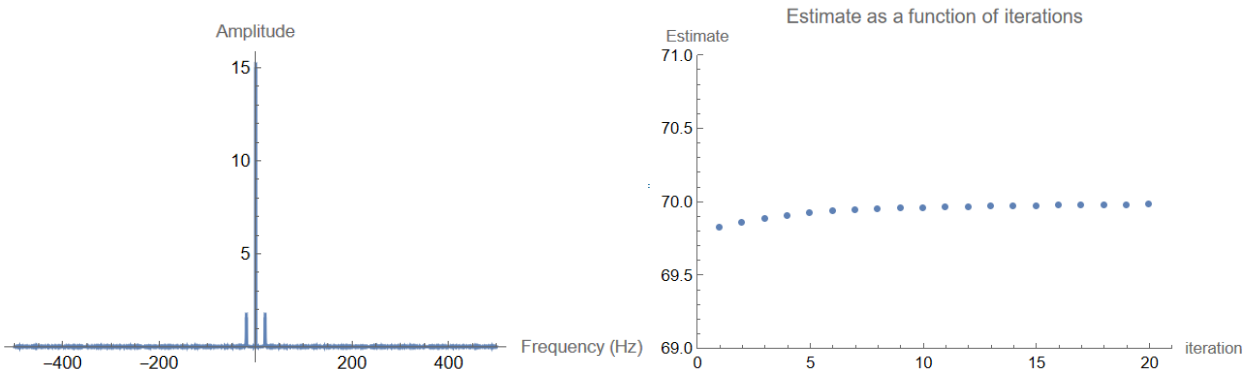


(a) (b)

Fig.2. (a) Decomposition of magnetization along y in n and n' components. (b) Experimental data simulated with $\Delta\omega = 70\ s^{-1}$ and $\omega_R = 100\ s^{-1}$

Hamiltonian acting would be given by equation (17). In the rotating frame of reference, if we see the free induction decay by measuring the magnetization $M_y(t)$ for a fixed $\Delta\omega$ and vary the Rabi strength ω_R, for $\omega_R \gg \Delta\omega$, the spin would be essentially locked in the y-direction, and we wouldn't observe any oscillations in the f.i.d. This is shown in figure (1b) by the gray line. The data recorded at this point isn't very informative as it doesn't capture the oscillatory dynamics of $\Delta\omega$. Now, if the ω_R is reduced such that $\omega_R \sim \Delta\omega$, we would start to see a combination of the locking and the oscillatory dynamics. Thus, we would look at this regime when $\omega_R \sim \Delta\omega$, and the prior information we have is that $\omega_1 < \omega_0 < \omega_2$. With this prior information and the recorded data, we need to estimate the value of $\Delta\omega$.

For this problem, there exists a simple analytic expression for the magnetization $M_y(t)$. In the interaction frame, we basically have two magnetic field, one in direction z with strength $\Delta\omega$ and one in direction y with strength ω_R. Thus, the net magnetic field can be thought of in



(a) (b)

Fig. 3. (a) Fourier Transform of F.I.D. from fig.(2b). (b) Estimate as a function of iterations.

direction $n = (0, \frac{\omega_R}{\omega}, \frac{\Delta\omega}{\omega})$, as seen in figure (2a). Here, $\omega = \sqrt{{\omega_R}^2 + \Delta\omega^2}$ is the precession frequency for spins along the direction n'. The magnetization along the n direction would remain constant with time while the magnetization along the n' direction would precess around the n axis with frequency ω while decaying because of the spin-spin relaxation. Thus, the magnetization along the y-direction would have a constant component along n and an oscillatory decaying component along n'. With the initial condition $M_y(0)$,

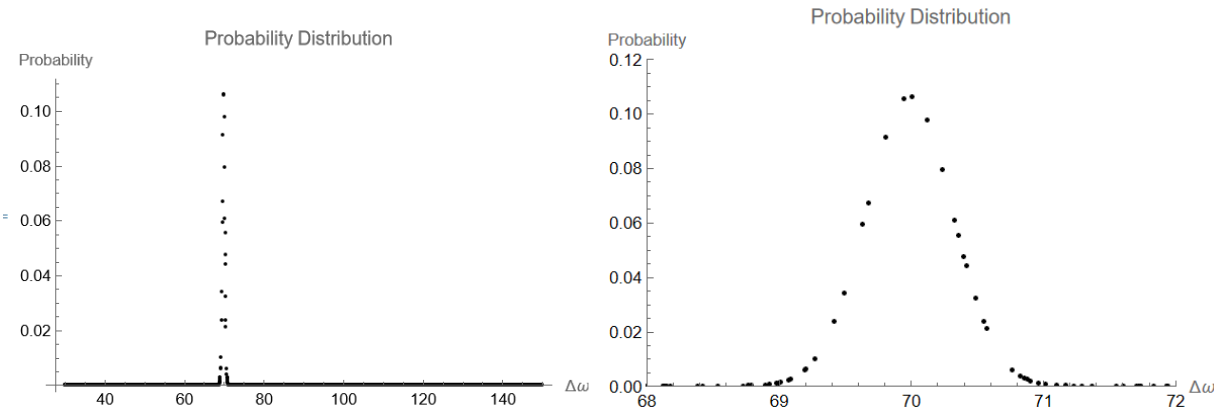
$$\mathrm{My(t)} = \cos(\theta)\cos(\theta) + \sin(\theta)\sin(\theta)\cos(\omega t)e^{-\frac{t}{T2}}$$

$$\mathrm{My(t)} = cos^2(\theta) + sin^2(\theta)\cos(\omega t)e^{-\frac{t}{T2}} \qquad (18)$$

Where, θ is the angle the n axis makes with y-axis and is given by,

$$\theta = cos^{-1}\left(\frac{\omega_R}{\omega}\right) \qquad (19)$$

Now with this model, we simulate the experimental result. In order to mimic the experimental errors, we added a Gaussian white noise of mean 0 and standard deviation 0.01. Taking $\Delta\omega = 70$ and $\omega_R = 100$ and $T_2 = 0.5\ seconds$, the free induction decay curve is given in figure (2b). One way of estimating the $\Delta\omega$ is by looking at the Fourier



(a) (b)

Fig. 4. (a) Probability Distribution for Larmor frequency. (b) Zoomed out version to see variation near the peak.

transform of the data. From equation (18), we can see that the free induction decay contains a constant term $cos^2(\theta)$ which would give a peak at the centre and due to the cosine variation of the second term, we would get two peaks at frequencies $+\omega/2\pi$ Hz and $-\omega/2\pi$ Hz. As ω depends on $\Delta\omega$, by finding at what frequency the peaks occur we can estimate $\Delta\omega$. But we can see in figure (3a) that the amplitude of peaks is too small compared to the middle peak. If the signal to noise ratio is worse than it would be difficult to precisely locate the peak. Moreover, the broadening caused by T2 relaxation would also create a broadening of the peak.

For SMC Implementation, the graph of estimate obtained as a function of iterations is given in figure (3b). For this implementation when the likelihood probability as derived in equation (9) was used, the probability near the actual value blows up giving indeterminate values and thus becomes unnormalizable. Thus, a Gaussian-like likelihood was used for calculating the likelihood,

$$P(D|\Delta\omega) = \exp\left(-\left(\sum_i D_i - f(\Delta\omega, t_i)\right)^2\right) \tag{20}$$

For this implementation, 2000 particles were sampled from a uniform distribution between (30,150). The threshold for sampling was set at 0.6. We see that as the number of iterations increase, the estimate saturates around the true value. After 20 iterations, the estimated value was 70.0451 s^{-1} which gives an absolute percentage error of 0.0645% with the actual value of $\Delta\omega = 70s^{-1}$. The distribution of the parameter from the particles and weight is given in figure (4). In figure (4a), we can see that we get a sharp peak near 70 and 0 everywhere else. Figure (4b) is a zoomed-out version where we can see that the distribution of parameter, we get is Gaussian-like. Obtaining such a distribution is helpful in field like Optimal Control theory, where a region estimation is helpful to describe the range of dynamics experienced by a quantum system.

In reality, the estimation of the Larmor frequency is a much simpler problem and we do not need to use the spin-locking experiment or Hamiltonian learning for the estimation. Moreover, In order to initialize our system to state σ_y, we need to apply a $\pi/2$ pulse on resonance and thus we would need to know ω_0 beforehand. But this problem serves as a example to study how SMC can estimate parameters efficiently and give insights about the distribution. With this problem in our mind, we move to the main problem of estimating the dipolar coupling strength in Gypsum.

IV. DIPOLAR DYNAMICS

The dipolar magnetic coupling between spins is the dominant multi-spin interaction in dielectric solids. This is because unlike liquid

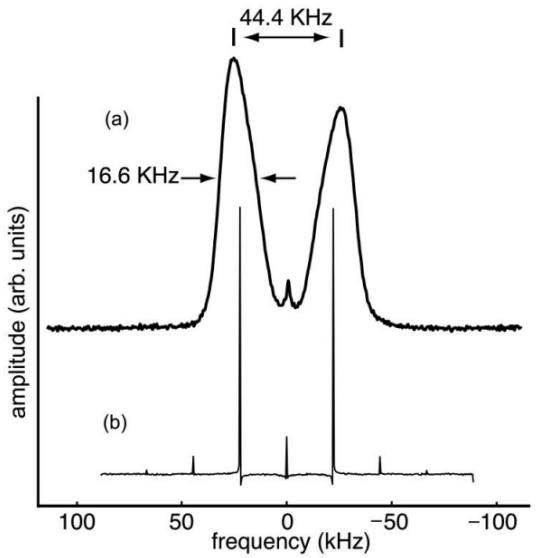
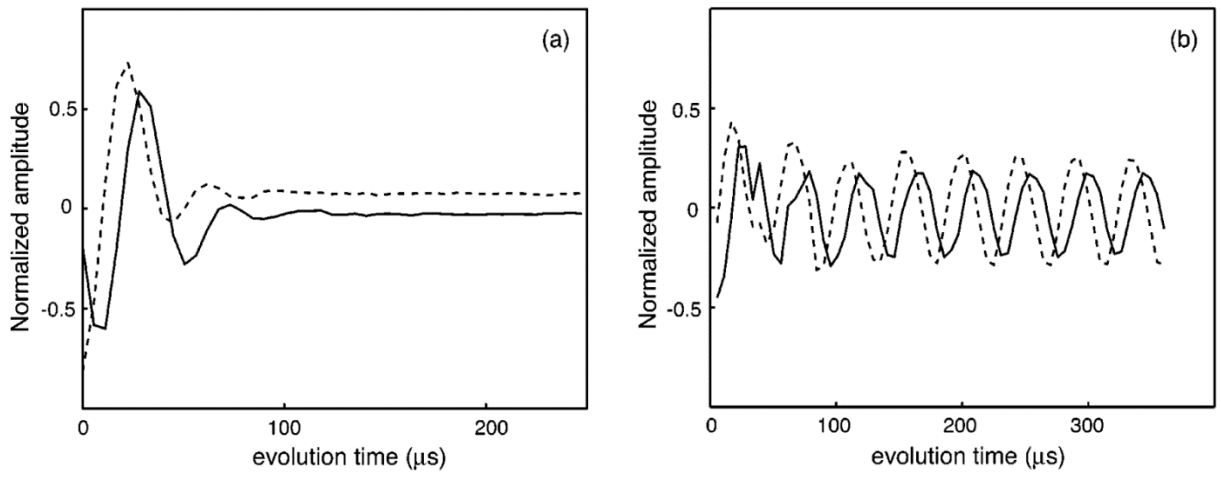


Fig 5. (a) Pake Doublet from one pulse experiment. (b) Pake Doublet with AM RF field. Figure adapted from [1].

state NMR, the spins in solids doesn't undergo constant motion and thus the interaction doesn't average to 0. Thus, the effect of this interaction would be observable in the NMR spectra of solids. G. Pake in the year 1948 [5], saw that when the proton spectra of crystalline Gypsum were observed under a one-pulse experiment, we observe a splitting of peaks and the distance between the peaks is proportional to dipolar coupling strength between the proton spins. In order to understand this splitting, we look at the mean field felt by a spin. Let α be the strength of dipolar coupling between two spins. This strength is inversely proportional to r^3, where r is the distance between the two spins. In absence of this interaction, if we apply a magnetic field H_0 in the z-direction causing the Zeeman energy splitting, the spins in transverse plane would precess with the Larmor frequency dependent on the magnitude of H_0 and we would observe a single peak in the spectrum corresponding to this frequency. Now considering the dipolar interaction, if the angle between the line joining the two spins and H_0, then the net field felt by each spin for spin-1/2 particles is –



(a) (b)

Fig. 6. Solid lines represent measurement signal of $\sigma_x^1 + \sigma_x^2$ while dotted line represents measurement signal of $\sigma_z^1\sigma_y^2 + \sigma_z^2\sigma_y^1$. (a) Free Induction Decay from one-pulse experiment. (b) Free Induction decay when AM RF field is applied. Figure adapted from [1].

$$H_{eff} = H_0 \pm \alpha(3\,cos^2(\theta) - 1) \qquad (21)$$

The positive and negative sign corresponds to the neighboring spin being aligned or anti-aligned with the spin we are observing. Thus, now if we observe the spectrum, we will see peaks at two frequencies one corresponding to net field $H_0 + \alpha(3\,cos^2(\theta) - 1)$, and the other corresponding to $H_0 - \alpha(3\,cos^2(\theta) - 1)$. In the rotating frame of reference of the Larmor frequency, these peaks would be equidistant on either side of 0 frequency and the splitting between the two peaks would be $2\alpha(3\,cos^2(\theta) - 1)$. This splitting of peak due to the dipolar coupling interaction is known as Pake Doublet.

This particular phenomenon can be observed by looking at the protons on water of crystallization of crystalline Gypsum. Gypsum, $Ca_2SO_4.2H_2O$ has two molecules of water of crystallization and in its crystalline form, the distance between any two water molecules is much larger than the distance between protons on a single molecule.

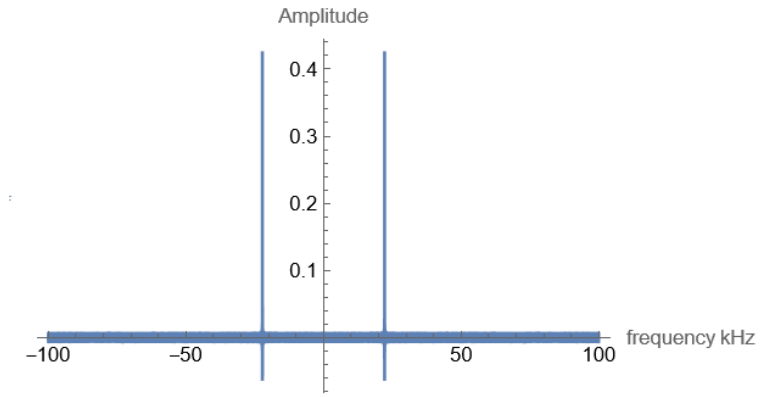
Thus, we can classify the interaction strengths felt by these protons into two categories namely – (i) a strong dipolar interaction ω_s, between protons on the same water molecule. This interaction is the same for every molecule as the distance between the protons on a molecule is the same. ii) weak dipolar interaction ω_w, between protons on different water molecules. Depending on the distance the molecules, these interaction strengths vary for every pair of molecules considered and is significantly smaller than the stronger coupling. The interesting parameter for us is the strong dipolar coupling strength as that can be used to estimate the bond length between protons in H_2O in Gypsum. The secularly approximated dipolar Hamiltonian for Gypsum can thus be written as,

$$H_d = \frac{\omega_S}{4}\sum_i h_{12}^{ii} + \frac{1}{4}\sum_{\alpha,\beta}\sum_{i\neq j}\omega_{ij}^{\alpha\beta}h_{\alpha\beta}^{ij} \qquad (22)$$

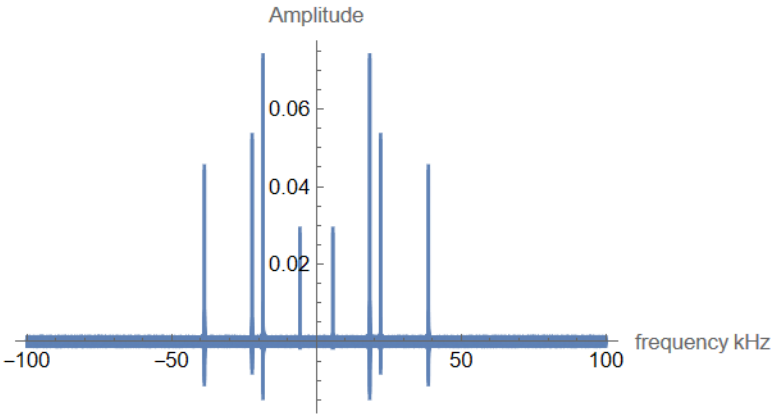
Where, $h_{\alpha\beta}^{ij} = 2\sigma_{zi}^{\alpha}\sigma_{zj}^{\beta} - \sigma_{xi}^{\alpha}\sigma_{xj}^{\beta} - \sigma_{xi}^{\alpha}\sigma_{xj}^{\beta}$ and i and j represent the molecule and α, β represents which proton on the molecule. If the weaker interactions were absent, evolution of σ_x takes a very nice form. Taking a spin-pair in the state $\sigma_x^1 + \sigma_x^2$, evolution under the strong-coupling interaction gives,

$$\sigma_x^1 + \sigma_x^2 \rightarrow (\sigma_x^1 + \sigma_x^2)\cos\left(\frac{3\omega_S}{2}t\right) + \left(\sigma_z^1\sigma_y^2 + \sigma_z^2\sigma_y^1\right)\sin\left(\frac{3\omega_S}{2}t\right)$$

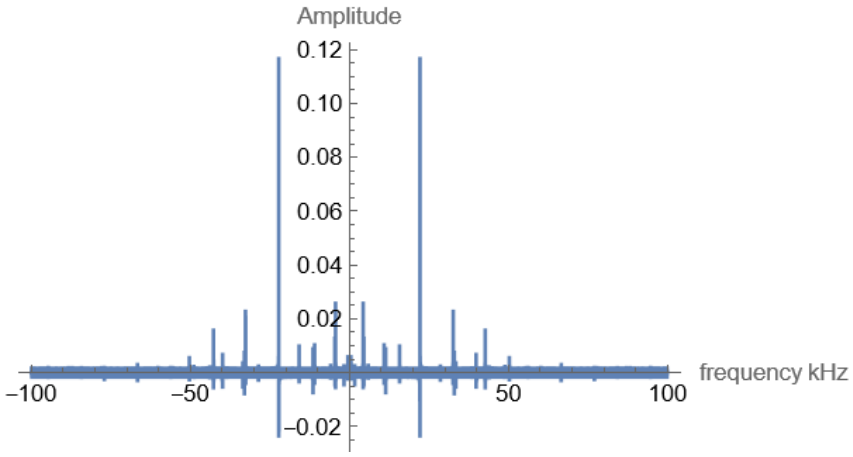
Thus, we see that if we measure the magnetization in the x-direction, it will oscillate with frequency $\frac{3\omega_S}{2}$, the Fourier analysis of which will give two sharp peaks centered at $\frac{3\omega_S}{4\pi}$ Hz and $\frac{3\omega_S}{4\pi}$ Hz. We can see this in the simulated spectra in figure (7a), where we observe sharp peaks at frequencies $\pm$22.22 kHz. Unfortunately, if the factor the weaker interactions in these peaks are broadened. This can be seen in figure



(a)



(b)



(c)

Fig. 7. (a) Simulated spectrum for isolated pair of protons on water molecule in Gypsum. (b) 4-spins simulation of 2 molecules interacting with each other weakly. (c) 4-spins simulation with application of AM RF field to suppress weak inter-molecular interactions.

(5a). The FWHM of the broadening can be seen as 16.6 kHz which can be is $3\omega_{avgw}/2\pi$, where ω_{avgw} is the average weak interaction strength. This figure is used from [1], where they applied the external magnetic field along the [010] orientation which causes the splitting of molecules in inequivalent sites of crystalline Gypsum to coincide causing Pake Doublet. Due to this broadening its difficult to find the exact splitting between the peak and thus its difficult to estimate the strong coupling strength. Also, if one looks at the free induction, we see that the signal dies off very quickly due to the weaker interactions. Thus, the net coherence time is very low for such systems. The authors of [1], used Amplitude Modulated Pulse to decouple these weak interactions while keeping the strong coupling intact. This was done by the application of the following AM RF pulse,

$$H_{mod}(t) = \frac{\omega_1}{2}\cos(\omega_m t)\, e^{i\left(\frac{\omega_0 t}{2}\right)\sum_i \sigma_{zi}} \left(\sum_i \sigma_{xi}\right) e^{-i\left(\frac{\omega_0 t}{2}\right)\sum_i \sigma_{zi}} \quad (23)$$

Under the conditions that the pulse strength $\omega_{avgw} \ll \omega_1 \ll \omega_S$ and the frequency $\omega_m = 3\omega_S/2$, the weaker interactions are decoupled. The free induction decay with the application of AM RF in the regime described above is given in figure (6b), where we see that the signal oscillates without any significant attenuation upto 360 μs which is much larger than that in figure (6a). As the time dependent AM RF field is periodic in the rotating frame of reference, the condition for decoupling of weaker interactions can be proved using Average Hamiltonian Theory as shown in reference [1]. Because of the decoupling of weaker interactions, we would get sharp peaks at $\pm 3\omega_S/4\pi$ as expected with just the strong interaction and can be seen in figure (5b). Figure (7b) shows a 4-spin simulation of evolution under the dipolar coupling where spins 1-2 and 3-4 are strongly coupled with strength $\omega_S = 92991.1425\ s^{-1}$ and spins 1-3,1-4,2-3 and 2-4 are weakly coupled with

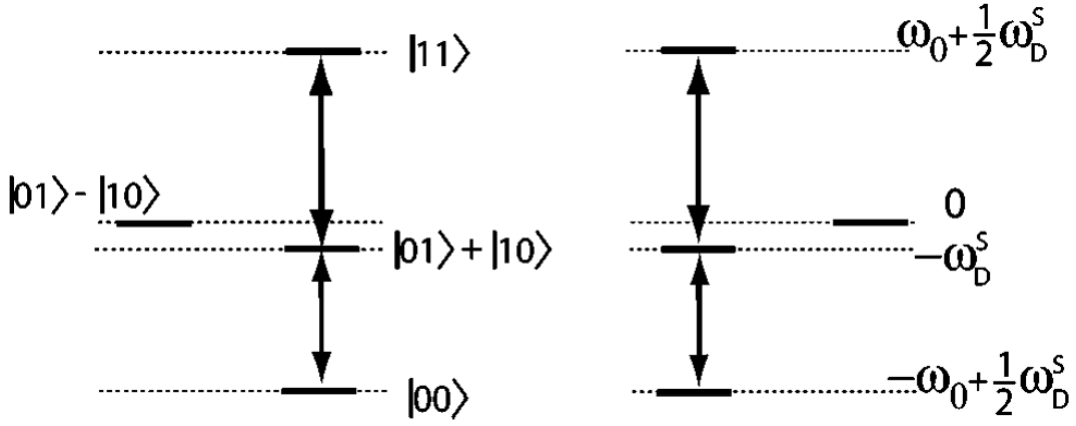


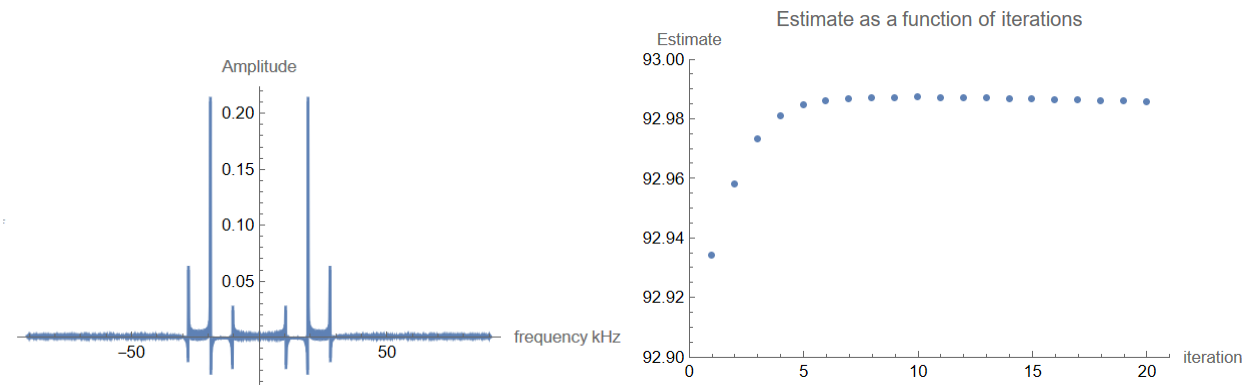
Fig. 8. Energy level structure of pair of protons on a water molecule. Each molecule forms a triplet and singlet. Weaker inter-molecular couplings drive transitions between singlet and triplet states. Figure adapted from [1].

interaction strength $\omega_{avgw} = 34557.5192\ s^{-1}$. As we are taking only 4-spins for the simulation, we wouldn't see a smooth broadening as per figure (5a). But we can see that instead of getting two peaks at $\pm 3\omega_S/4\pi$ Hz, we see multiple peaks near it. Figure (7c), shows the simulated spectra obtained when an AM RF pulse (Appendix) is applied in the 4-spin simulation with $\omega_1 = 3\omega_S/4$ and $\omega_m = 3\omega_S/2$. We can see that the peaks due to the weaker interactions have been suppressed and we get sharp peaks at $\pm 3\omega_S/4\pi$ Hz broadened by only the T2 relaxation.

An alternative way of proving this is by looking at the energy level structure. If we consider a single isolated molecule, as we have two spin-1/2 protons, they form triplet states acting as a spin-1 particle and a singlet acting as spin-0. These triplet states and the singlet and their energy levels are given in figure (8). The transitions between triplet and singlet state are forbidden for an isolated molecule. Now if we take two such molecules who are weakly interacting, the weak interaction preserves the energy level structure of single molecules but the transitions between singlet and triplet and singlet states in a molecule

are no longer forbidden. These transitions are driven by the weak interaction strength. Now, consider if we drive the system with pulse of frequency $\omega_m = 3\omega_S/2$, in the rotating frame of reference, it will simultaneously drive transitions between the states $|00\rangle \leftrightarrow |01\rangle + |10\rangle$ and $|01\rangle + |10\rangle \leftrightarrow |11\rangle$. The rate of these transitions would depend on the Rabi strength of the field ω_1. Thus, if this strength is much higher than the average weak interaction strength, we would effectively only have transitions within the triplet states of the molecule and the weaker interactions would be decoupled. But if $\omega_1 \gg \omega_S$, in that case the triplet-singlet energy structure would break down and the spins within a molecule would also get decoupled. Thus, for preserving the strong interactions but decoupling the weaker interactions, apply a pulse with $\omega_{avgw} \ll \omega_1 \ll \omega_S$ and $\omega_m = 3\omega_S/2$.

As we want to estimate the strong coupling, we saw that using just a one-pulse experiment gives broadened peaks from which it is difficult to precisely locate the splitting. We saw that by using the AM Modulated pulse, we can decouple the weak interactions and get sharp peaks but even for that we need to know what the value of ω_S is beforehand as the Rabi strength ω_1 and frequency ω_m depend on ω_S. Considering that we have prior information as bounds on the value of ω_S, we can choose ω_1 and ω_S randomly in that range and as discussed earlier if this value is $\omega_{avgw} \ll \omega_1$, the weaker interactions would be decoupled. Thus, we would see two distinct peaks but the splitting between these peaks would depend on the value of ω_m and would no longer be $3\omega_S/2\pi$ Hz. This can be seen in figure (9a), where the splitting between peaks in no longer $\frac{3\omega_S}{2\pi} = 44.44$ kHz, when simulated with AM RF pulse with $\omega_1 = 50000\ s^{-1}$ and $\omega_m = 120{,}000\ s^{-1}$ was applied. We see that the value of ω_m is not equal to $\frac{3\omega_S}{2} = 139486.714\ s^{-1}$. Thus, given this spectrum and the corresponding free induction decay, we use the SMC algorithm



(a) (b)

Fig. 9. (a) Fourier Spectra with $\omega_1 = 50000\ s^{-1}$ and $\omega_m = 120{,}000\ s^{-1}$. In this case peaks are not separated by $\frac{3\omega_S}{2\pi}$ Hz. (b) Estimate vs. Iterations for strong coupling problem. Here the scale on x-axis is ms^{-1}.

to estimate the value of strong coupling. The model Hamiltonian we consider is,

$$H(\omega_s) = \frac{\omega_s}{4}(2\sigma_z^1\sigma_z^2 - \sigma_x^1\sigma_x^2 - \sigma_x^1\sigma_x^2) + \frac{\omega_1}{2}\cos(\omega_m t)\,(\sigma_x^1 + \sigma_x^2) \qquad (24)$$

For SMC Implementation, the graph of estimate obtained as a function of iterations is given in figure (9b). For this implementation, 2000 particles were sampled from a uniform distribution between (30,000,150,000). The threshold for sampling was set at 0.6. We see that as the number of iterations increase, the estimate saturates around the true value. After 20 iterations, the estimated value was $92985.7\ s^{-1}$ with a standard deviation $(36.0992)\ s^{-1}$of which gives an absolute percentage error of 0.0059% with the actual value of $\omega_S = 92991.1425\ \ s^{-1}$. We see that using SMC algorithm, we can estimate the value of unknown parameter upto a very good precision and we also get uncertainty estimates in our estimation. We can use this uncertainty estimate to find confidence intervals and region estimation.

V. OUTLOOK

Sequential Monte Carlo has many advantages over methods for parameter estimation such as Maximum Likelihood Estimation and gradient-based techniques. These include –

i) Robustness to noisy or incomplete data. In contrast, gradient based methods and MLE are sensitive to noise and outliers and may produce inaccurate estimates in such cases.
ii) SMC would converge to true posterior distribution even when the likelihood distribution is multimodal. In contrary, gradient based and MLE methods may get stuck in local optima and fail to converge to true global optima.
iii) As SMC is based on Bayesian Inference, we can get uncertainty estimates and model comparison, which is not possible in the case of gradient based and MLE methods.

Apart from these advantages, SMC also suffers from some of the disadvantages such as-

i) High Computational Cost, as the SMC perform Monte Carlo Simulations multiple times which involves evaluating the likelihood function many times and is computationally expensive.
ii) The precision of the estimates depends on the values of hyperparameters such as number of sampled points, resampling threshold. Thus, hyperparameter tuning is needed to get accurate estimates.
iii) For our experiments we found that using a uniform distribution or a flat Gaussian for prior distribution gave similar results. But this may not always be the case and for complex systems, the accuracy might depend on the choice of the Prior distribution.

Similar works in the Quantum Technology field using Bayesian Inference and SMC for different purposes include –

i) In the paper "Robust Online Hamiltonian Learning" [2], the authors presented SMC for parameter estimation which can work in online mode i.e., the algorithm is ran in real-time as the data is collected and they gave a method of selecting the best controls for future data collection based on the present data collected such that the information gained is maximum through subsequent measurements. The work presented in this project is based on this paper and uses on offline version of this method where complete data has been collected and the algorithm is an applied as a post-processing step.

ii) In the paper "Exponential Parameter Estimation (in NMR) using Bayesian Probability Theory" [4] used Bayesian Inference for estimating the exponential decay parameters in the spectra of water relaxation experiments of Human Brain gray-matter in vivo. Estimation of these parameters is important for the understanding of water content and structure of gray-matter in human brain. Exponential parameter estimation of a sum of exponentials is a difficult task and by using Markov Chain Monte Carlo Simulation with stimulated annealing for sampling, the authors of the paper used it to estimate the exponential parameters using Bayesian Inference.

iii) In the paper "Bayesian Inference for NMR Spectroscopy with Applications to Chemical Quantification"[6], the authors used Bayesian Inference for estimating the relative chemical composition of chemicals in a mixture. Traditionally this can be done by looking at the Fourier Transform of the free induction decay, where different chemicals would have different Larmor frequencies and thus show peaks at different frequencies.

The above papers are just some of the works in the Quantum Technologies field which have used Bayesian Inference for different goals. In the next section, we discuss why our problem of estimating dipolar coupling through SMC is important and explain its relevance in today's time.

VI. RELEVANCE AND FUTURE DIRECTION

Protein-Protein complex mapping is a technique used to determine interactions between two or more proteins in a complex. Using this information, researchers can obtain insights about biological processes such as enzyme catalysis, protein folding etc. In happens that proteins with very large molecular weights are inaccessible through NMR. This is because beyond a certain size of protein complexes, the spectral quality and resolution obtained through NMR drops and thus it becomes difficult to obtain any meaningful information from the spectra. Thus, the technique used to probe the interactions between Protein-Protein complexes is known as Electron Paramagnetic Resonance (EPR). High Q, a Quantum Technology based startup is using EPR for Protein Dynamics [7]. This technique is similar to NMR with the only difference that instead of addressing the nuclear spins, we address the electrons spins. Thus, similar to NMR, EPR is used to measure distances between the protein-protein complexes by measuring the magnetic interaction strength between unpaired electrons on protein complexes. Dipolar coupling between two electrons would lead to splitting of the EPR spectrum, which can be analyzed to determine the dipolar coupling constant and hence the distance between the paramagnetic centres. These distance act as constraints on the Molecular dynamics simulations of proteins which is used by researchers to get insights in drug-design. Thus, the work we did in this

project can be extended to finding the coupling strength of electrons from the EPR spectra.

As the experimental data was generated by simulations with an added white noise, the next step is to collect experimental data by performing the AM RF pulse experiment on crystalline Gypsum and then use SMC for parameter estimation. We studied the algorithm which can be used for prediction of single unknown parameter. Thus, the next step is to think how to extend this for more than one unknown parameters.

VII. CONCLUSION

In this project we studied how Bayesian Inference equipped with Sequential Monte Carlo Simulation can be used for parameter estimation of Quantum system. We used SMC for estimating Larmor frequency using spin-locking pulse experiment and for estimating the dipolar coupling strength between protons on a water molecule in Gypsum. We saw why such estimating the strength using this method is important and how it can help in enhancing the protein dynamics simulations.

VIII. Appendix

As the Amplitude modulated RF field is time-dependent, we need to use numerical approximation in order to find its unitary evolution. The net Hamiltonian in the rotating frame of reference is,

$$H = H_d + H_1(t) \tag{25}$$

Where, H_d is the dipolar Hamiltonian and $H_1(t)$ is the time dependent AM RF pulse Hamiltonian. To perform numerical approximation, we divide the time in m discrete intervals of duration dt. The value of dt should be small enough such that the value of $H_1(dt)$ is reasonably

constant in each interval. In this case the unitary evolution at time T, can be calculated as,

$$U_T = U_m U_{m-1} \ldots . U_1 \tag{26}$$

Where,

$$U_k = \exp\left(-i\left(H_d + H_1(kt)\right)dt\right) \tag{27}$$

IX. REFERENCES

[1]. “Selective Coherence Transfers in Homonuclear Dipolar Coupled Spin Systems”; Chandrasekhar Ramanathan et. al. Physical Review A 71, 020203(R) 2005.

[2]. “Robust Online Hamiltonian Learning”, Christopher E. Granade et. al.

[3]. “Hamiltonian Learning and Certification using Quantum Resources”, Nathan Wiebe et. al.

[4]. “Exponential Parameter Estimation in NMR using Bayesian Probablity Theory”, G Larry Bretthorst et. al.

[5]. G. Pake, J. Chem. Phys. 16, 327 s1948d.

[6]. “Bayesian Inference for NMR Spectroscopy with applications to Chemical Quantification” (2014) by Andrew Wilson et. al.

[7]. https://highqtechnologies.com/products/