

A Bayesian approach to quantifying uncertainty from experimental noise in DEER spectroscopy



Thomas H. Edwards, Stefan Stoll*

Department of Chemistry, University of Washington, Seattle, WA 98103, United States

ARTICLE INFO

Article history:

Received 25 May 2016

Revised 29 June 2016

Accepted 30 June 2016

Available online 2 July 2016

Keywords:

Inverse problem

Tikhonov regularization

MCMC

Statistical inference

ABSTRACT

Double Electron-Electron Resonance (DEER) spectroscopy is a solid-state pulse Electron Paramagnetic Resonance (EPR) experiment that measures distances between unpaired electrons, most commonly between protein-bound spin labels separated by 1.5–8 nm. From the experimental data, a distance distribution $P(r)$ is extracted using Tikhonov regularization. The disadvantage of this method is that it does not directly provide error bars for the resulting $P(r)$, rendering correct interpretation difficult. Here we introduce a Bayesian statistical approach that quantifies uncertainty in $P(r)$ arising from noise and numerical regularization. This method provides credible intervals (error bars) of $P(r)$ at each r . This allows practitioners to answer whether or not small features are significant, whether or not apparent shoulders are significant, and whether or not two distance distributions are significantly different from each other. In addition, the method quantifies uncertainty in the regularization parameter.

© 2016 Published by Elsevier Inc.

1. Introduction

Double Electron-Electron Resonance (DEER), also called Pulse Electron-Electron Double Resonance (PELDOR), is a solid-state pulse Electron Paramagnetic Resonance (EPR) experiment that measures long-range (1.5–8 nm) distances, r , between paramagnetic species to sub-ångström resolution [1–3]. It is commonly applied to proteins and other biomacromolecules immobilized in frozen glassy solutions, and has been used to investigate systems that can be challenging to NMR and X-ray crystallography, such as membrane proteins [4–10]. DEER provides complete distance distributions, $P(r)$, by measuring the through-space magnetic dipole-dipole interaction between two stable paramagnetic spin labels.

DEER data consist of an oscillatory time-domain signal that can be transformed via regularized numerical inversion into an interspin distance distribution function $P(r)$. Since this inversion is mathematically ill-conditioned, numerical regularization is necessary. Tikhonov regularization is the most widely used regularization procedure. It is a least-squares method that imposes a degree of smoothness in $P(r)$, reflecting the inherent width in the distribution owing to spin label and protein mobility [11–13]. The magnitude of the roughness penalty is scaled with a regular-

ization parameter, α . This regularization scheme complicates error analysis and uncertainty quantification [13].

The distance distribution $P(r)$ is, in principle, rich in structural information, but in practice its usable information is often limited to the dominant spin-spin distance. Peak widths and shapes, which describe structural heterogeneity, are more uncertain as they are highly sensitive to noise levels, degree of regularization, time-domain truncation, background correction, etc. There does not currently exist a general and robust method for fully quantifying uncertainty in the Tikhonov estimate of $P(r)$ [4,12–14]. As a result, it can be difficult to draw robust conclusions from DEER distance distributions. The lack of uncertainty quantification prevents access to important structural information.

Concerningly, it is very uncommon for any error or uncertainty estimates to be displayed in $P(r)$. After surveying over 100 recently published papers containing DEER distance distributions, we have found that nearly 90% lacked error estimates on $P(r)$ (see SI). This state of affairs may hinder more widespread adoption of DEER spectroscopy. Enabling robust statistical analysis of DEER data will significantly expand the reliability, and therefore the utility, of DEER.

In this paper, we present a Bayesian statistical method to quantify the uncertainty arising from measurement noise and regularization in the DEER distance distribution estimated with Tikhonov regularization. The effects of other sources of uncertainty, such as background correction, can be assessed by extending the method. When evaluating a given $P(r)$, DEER practitioners

* Corresponding author.

E-mail addresses: edwardst@uw.edu (T.H. Edwards), stst@uw.edu (S. Stoll).

often request to see the time-domain trace in order to qualitatively assess uncertainty in the $P(r)$. There is an empirical understanding that data with a higher signal-to-noise ratio (SNR) yield more reliable distance distributions, particularly in the case of multi-modal distributions, but quantifying the reliability of $P(r)$ remains an open issue. Further, the selection of the value for the regularization parameter α , which is required for Tikhonov regularization, is a source of bias and uncertainty.

Bayesian methods have been successfully implemented in many applications involving the solution of ill-conditioned inverse problems, including medical imaging, underground resource exploration, image reconstruction, and astronomical spectroscopy [15–18]. Recently, the utility of Bayesian statistics has been recognized in the estimation and analysis of model parameters related to biophysical measurements [19–21].

The next section summarizes the basics of DEER and Tikhonov regularization. Then, we introduce a new Bayesian approach that provides a full probabilistic analysis of $P(r)$ using a Markov-Chain Monte Carlo (MCMC) algorithm. After that, we illustrate its performance using model distributions. Finally, we discuss how the method can be applied to answer interpretational questions about $P(r)$ and how it relates to other forms of error assessment.

2. DEER and Tikhonov regularization

2.1. The DEER signal

The time-domain DEER signal, $V(t)$, is the integrated spin echo measured as a function of the timing of the pump pulse relative to the probe pulses within the DEER pulse sequence. $V(t)$ is modulated by the dipolar interaction between spin labels. It is a combination of the desired modulation due to intramolecular spin-spin dipolar interactions, $F(t)$, and an undesired background contribution due to intermolecular interactions, $B(t)$ (Fig. 1). In dilute solutions, $V(t)$ can be expressed as a product [1,2]

$$V(t) = V_0 \cdot F(t) \cdot B(t) \quad (1)$$

$$F(t) = (1 - \lambda) + \lambda S(t) \quad (2)$$

$$B(t) = \exp(-(kt)^{D/3}) \quad (3)$$

where V_0 is the unmodulated integrated echo intensity, λ is the modulation depth (a measure of the efficiency of the DEER pump

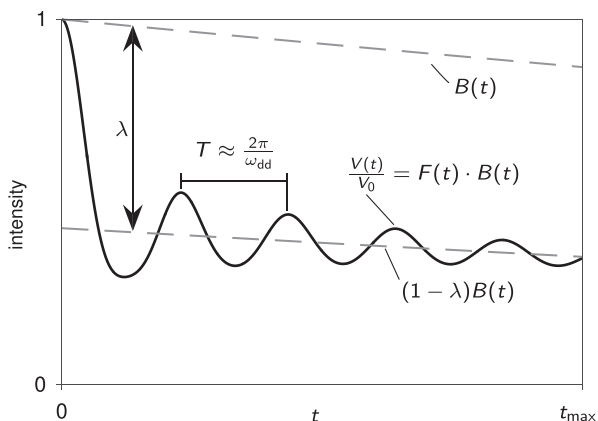


Fig. 1. The normalized time-domain DEER signal, $V(t)/V_0$ (solid black), is the product of two signals: the oscillatory intramolecular modulation, $F(t)$, that provides information about spin-spin distances in biradicals, and an averaged intermolecular signal in the form of a decaying exponential background, $B(t)$ (dashed gray). The background is removed and $F(t)$ is shifted and rescaled by λ , the modulation depth, to obtain the dipolar modulation function, $S(t)$.

pulse, $0 \leq \lambda \leq 1$), and $S(t)$ is the intramolecular dipolar modulation function. Eq. (2) is based upon the assumption that the sample contains only A-B spin pairs, and is not valid for the case where an A-A pair is excited by the observer sequence. This case is unavoidable in homogenous labeling schemes, as in the common case of two nitroxide radical labels on one protein. A more complete expression contains additional terms, as explained in [22,23]. For the purposes of this work, the specific forms of $B(t)$ and $F(t)$ are irrelevant; the analysis is only concerned with $S(t)$.

The background $B(t)$ is a decaying exponential, where D is the dimensionality of the system (3 for soluble proteins). The decay constant k depends upon the spin concentration and inversion efficiency. For coupled pairs of spin-1/2 in an amorphous glass or powder, and in the absence of orientation selection, $S(t)$ is given by the integral

$$S(t) = \int_0^\infty dr K(t,r) P(r) \quad (4)$$

where the integral kernel $K(t,r)$ describes the powder average of the dipolar interaction and $P(r)$ is the distance distribution probability density function (PDF) ($\int dr P(r) = 1$) that describes the conformational heterogeneity of the ensemble of spin-labeled protein molecules in the sample. The dipolar kernel is [1,2,24]

$$K(t,r) = \int_0^{\pi/2} d\theta \sin \theta \cos [(1 - 3 \cos^2 \theta) \cdot \omega_\perp t] \quad (5)$$

$$= \frac{C(z)}{z} \cdot \cos(\omega_\perp t) + \frac{S(z)}{z} \cdot \sin(\omega_\perp t) \quad (6)$$

where $\omega_\perp = \frac{\mu_0}{4\pi} \frac{\mu_B^2}{h} \frac{g_A g_B}{r^3}$ is the perpendicular dipolar frequency, $z = \sqrt{6\omega_\perp t/\pi}$, and $C(z)$ and $S(z)$ are the cosine and sine Fresnel integrals, respectively [25,26]. The frequency of oscillation in the time domain is proportional to the inverse cube of the spin-spin distance. With this kernel and the normalized $P(r)$, $S(t)$ is normalized such that $S(0) = 1$ in the absence of noise.

Experimentally, $S(t)$ is measured at n discrete time points t_i . For the data analysis, P is also discretized, typically at n points r_j , although a different number of points is possible. The discretized form of Eq. (4) is

$$S = KP \quad (7)$$

where K is here an $n \times n$ matrix, and S and P are $n \times 1$ vectors, with elements:

$$K_{ij} = K(t_i, r_j) \cdot \Delta r \quad (8)$$

$$S_i = S(t_i) \quad (9)$$

$$P_j = P(r_j) \quad (10)$$

The step size in the distance domain, Δr , has been incorporated into K in order to simplify notation. This discretization preserves normalization of S and P . After background correction, the DEER experiment yields the noisy signal, S . However, the practitioner seeks the distance distribution, P .

2.2. Tikhonov regularization

In order to extract P from S , it is necessary to solve Eq. (7) for P . This would require inverting the matrix K . Unfortunately, K has a large condition number, and so the computation of its inverse is prone to large numerical errors. Thus, Eq. (7) is an ill-conditioned inverse problem [27]. To obtain a solution, numerical regularization is required. The regularization method preferred in the DEER literature is Tikhonov regularization [11–13,27], which solves the multidimensional least-squares minimization problem:

$$P_\alpha = \min_{P \geq 0} (\|S - KP\|^2 + \alpha^2 \|LP\|^2) = \min_{P \geq 0} (\rho + \alpha^2 \eta) \quad (11)$$

where L is commonly the $((n-2) \times n)$ -dimensional second-derivative operator matrix

$$L = \begin{pmatrix} 1 & -2 & 1 & & 0 \\ & 1 & -2 & 1 & \\ & & \ddots & \ddots & \ddots \\ 0 & & & 1 & -2 & 1 \end{pmatrix} \quad (12)$$

α is the regularization parameter [11–13], and $\|\dots\|$ is the Euclidean norm (ℓ^2 -norm).

The first term in Eq. (11), ρ , is the residual sum of squares and captures how well the solution P_α fits the data in the time domain (time-domain misfit, or prediction error). The second term, η , is a measure of the overall roughness of the solution. The regularization parameter α determines the importance of the roughness relative to the prediction error in the fitting.

Since P is a probability density, it is constrained to be non-negative. Therefore, the simple analytical solution to the unconstrained problem, which is

$$P_\alpha = (K^T K + \alpha^2 L^T L)^{-1} K^T S = Q_\alpha K^T S \quad (13)$$

cannot be used. In order to solve the constrained problem, numerical iterative algorithms have to be employed [27,28].

2.3. Regularization parameter selection

The regularization term in Eq. (11) represents an assumption about the distance distribution: it should be smooth. This is a reasonable assumption as it is difficult to imagine a biologically relevant, or microscopically realistic, scenario where the distance between two intrinsically mobile spin labels varies in a very rough and discontinuous fashion on a sub-ångström scale. The degree of regularization is determined by the magnitude of the scalar parameter α , which is a nuisance parameter that must be separately determined. For very small α , roughness is not penalized, the solution is extremely spiky and over-fits the noise. For large α , roughness is over-penalized, P_α will be over-smoothed, and the fit to the data will be poor.

Since the value of α significantly influences the form of the solution, it is important to choose an optimal value for α . One commonly employed method is the L-curve criterion [12,13,29]. The L-curve is a parametric curve in the log-log plot of η against ρ . This curve typically adopts a characteristic “L” shape. The L-curve criterion considers the optimal value for α to be found at the corner of the L-curve [29]. In practice, the corner is often determined by eye. Alternatively, the point of maximum curvature can be found automatically [30]. This removes some practitioner bias from the data processing procedure and increases reproducibility.

Leave-One-Out Cross Validation (LOO-CV) is another approach [30]. Generally, CV approaches withhold a subset of the data from the Tikhonov analysis and then iteratively adjust α to reduce the error between the fit and the withheld data [30]. The LOO variant of CV excludes each data point, one at a time, from the analysis and then minimizes the total prediction error. This can be expressed as the minimization problem [31]

$$\alpha_{\text{LOOCV}} = \min_\alpha \left(\sum_{i=1}^n \frac{|K(:,i)P_\alpha - S(i)|^2}{1 - h_\alpha(i,i)} \right) \quad (14)$$

where h_α is the α -dependent matrix $KQ_\alpha K^T$ and P_α is the Tikhonov solution for a particular α . In our hands, LOO-CV outperforms the L-curve criterion. However, it is still a heuristic method and does not provide uncertainty estimates for α .

3. Bayesian uncertainty quantification

In this section, we present an MCMC method based on Bayesian statistics for quantifying uncertainty in P and α . It utilizes $S(t)$ and assumes that background correction has already been performed. We introduce the basic probabilistic equations, show the connection to Tikhonov regularization, and describe the MCMC algorithm, generate representative samples of P and α , and analyze them statistically.

3.1. The posterior distribution

We start by considering what is observable and given, and what is unknown, uncertain, and wanted. Here, we consider the intramolecular dipolar modulation function $S(t)$ observable and given. $S(t)$ is obtained from the experimental data $V(t)$ after background correction. It is an inseparable sum of the true signal and corrupting noise. A second useful quantity that can be obtained from experiment is the noise variance. We obtain it experimentally by acquiring and storing each phase-cycled DEER trace separately and then statistically analyzing the distribution of values for each t_i across the set of acquired traces. As shown in the SI, this noise is Gaussian with amplitude roughly independent of t_i . The noise deviations at adjacent time points are not correlated. Therefore, we can characterize the noise by a single noise variance, σ_S^2 . (If recording separate traces is not possible on a particular spectrometer, one may use the root mean square difference between a Tikhonov fit and the data as a stand-in. We do not recommend this, as errors in the fit will propagate to the final results.) For notational convenience, we will henceforth use its inverse, called the precision

$$\tau = 1/\sigma_S^2. \quad (15)$$

Next, we consider the unknown and uncertain quantities. These include the desired underlying distance distribution P . However, since the problem is not solvable without regularization, a regularization parameter is required as well. There is no absolute certainty about its optimal value. Here, we will use the symbol δ and show later that δ is proportional to α^2 .

To summarize, S and τ are observable and fixed, and P and δ are unknown and desired. Our goal is to fully characterize the conditional probability distribution for P and δ , given the measured S and τ : $\pi(P, \delta|S, \tau)$. First, we construct an expression for $\pi(P, \delta|S, \tau)$. Like any conditional probability, it is proportional to the full joint probability

$$\pi(P, \delta|S, \tau) = \frac{\pi(P, \delta, S, \tau)}{\pi(S, \tau)} \propto \pi(P, \delta, S, \tau) = \pi(S, P, \delta, \tau). \quad (16)$$

In the second step we omitted the marginal probability $\pi(S, \tau)$, since it is a constant independent of P and δ . In the last step, we reordered the variables. This joint probability can be written as a chain product of four probabilities

$$\pi(S, P, \delta, \tau) = \pi(S|P, \delta, \tau) \cdot \pi(P|\delta, \tau) \cdot \pi(\delta|\tau) \cdot \pi(\tau). \quad (17)$$

Next, we introduce specific expressions for the factors on the right-hand side. The first term indicates the conditional probability of a certain S given P , δ and τ . For this, we can use a Gaussian with the weighted sum-of-squares-deviation in the exponent

$$\pi(S|P, \delta, \tau) = \pi(S|P, \tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left(-\frac{\tau}{2}\|S - KP\|^2\right). \quad (18)$$

(We omit δ in the notation, since the expression does not directly depend on it). This expression captures the essence of least-squares fitting: if P is such that KP fits S well, the magnitude of the exponent is small and the probability large. On the other hand,

if the fit is poor, the exponent is large in magnitude, and the probability is small (Fig. 2).

The second term in Eq. (17) is the conditional probability of a certain P given values for δ and τ . For this, we choose a Gaussian smoothness requirement:

$$\pi(P|\delta, \tau) = \pi(P|\delta) = \left(\frac{\delta}{2\pi}\right)^{n/2} \exp\left(-\frac{\delta}{2}\|LP\|^2\right) \quad (19)$$

where L is the same operator as in Tikhonov regularization (commonly the second derivative) and δ determines the roll-off rate of the probability as a function of the roughness $\|LP\|^2$. This expression assigns higher probabilities to smoother P , as illustrated in Fig. 2.

The third term in Eq. (17) is the probability of a given δ . Since we are uncertain about the value of δ , we use a very broad distribution. For mathematical convenience and because of its highly tunable shape, we choose a gamma distribution [17]

$$\pi(\delta|\tau) = \pi(\delta|a, b) = \frac{b^a}{\Gamma(a)} \delta^{a-1} \exp(-b\delta) \quad (20)$$

where we omitted τ because it is independent of δ . The fixed parameters a and b determine the shape and width of the distribution such that its mean is a/b and its variance is a/b^2 . We use $a = 1$, in which case the expression simplifies to a decaying exponential, $b \exp(-b\delta)$ (see Fig. 2). We discuss the selection of b in Section 3.3. It is not currently clear if our choice of Eq. (20) is optimal; other superior choices could exist. The gamma distribution is, however, commonly used in conjunction with Gaussian priors.

Finally, the last term in Eq. (17), $\pi(\tau)$, is a constant, since τ is known from experiment. We can therefore drop it and use the experimental τ value.

To summarize, our target expression for the conditional probability for P and δ , given S and τ , is

$$\pi(P, \delta|S, \tau) \propto \pi(S|P, \tau) \cdot \pi(P|\delta) \cdot \pi(\delta|a, b) \quad (21)$$

with the right-hand side quantities given by Eqs. (18)–(20) and illustrated in Fig. 2. This expression is a form of Bayes' law and constitutes a Bayesian hierarchical model. $\pi(P, \delta|S, \tau)$ is called the posterior PDF, or simply the posterior. $\pi(S|P, \tau)$ is called the likelihood and describes how likely it is to measure a certain S given a particular P and τ . $\pi(P|\delta)$ is called the prior PDF, or simply the prior. It describes our knowledge about P prior to the acquisition of the data

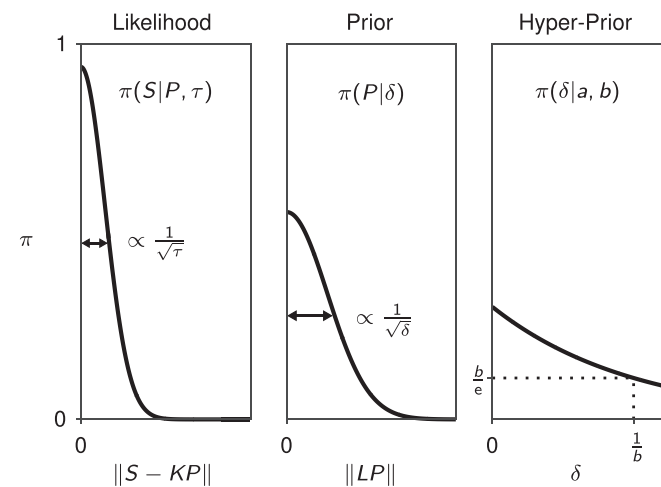


Fig. 2. Illustration of the factors composing the posterior distribution: the likelihood, prior, and hyperprior. The likelihood function (left) is a Gaussian in $\|S - KP\|$, the fit error. Its width is determined by τ . The prior (middle) is also a Gaussian, this time in $\|LP\|$, the roughness. It biases towards zero roughness with width determined by δ . The hyper-prior (right) is a gamma distribution with the form of a slowly decaying exponential, given $a = 1$ and small b .

(i.e. P is not rough). $\pi(\delta|a, b)$ is a hyper-prior which describes prior knowledge about δ . The reason we use a gamma distribution in δ is that this makes the total expression a gamma function in δ . Eq. (20) is therefore called a conjugate hyper-prior [15]. Similarly, the Gaussian prior in P is conjugate with the likelihood, since the total expression is also Gaussian in P .

3.2. Connection to Tikhonov regularization

The particular P that maximizes the posterior probability is the most likely P and is called the maximum a posteriori (MAP) solution: $P_{\text{MAP}} = \max_{P \geq 0} \pi(P, \delta|S, \tau)$. The same P can be obtained by minimizing the negative logarithm of that probability: $\max_{P \geq 0} \pi(P, \delta|S, \tau) = \min_{P \geq 0} (-\log \pi(P, \delta|S, \tau))$. Dropping all P -independent prefactors in $\pi(P, \delta|S, \tau)$, taking the negative logarithm, and dividing by $\tau/2$ gives

$$-\log \pi(P, \delta|S, \tau) \propto \|S - KP\|^2 + \frac{\delta}{\tau} \|LP\|^2. \quad (22)$$

This is identical to the Tikhonov functional in Eq. (11) if [15,32,33]

$$\alpha^2 = \frac{\delta}{\tau} \quad (23)$$

Consequently, the P_α obtained by Tikhonov regularization is equal to P_{MAP} in the special case where Eq. (23) is true. Eq. (23) is not a general constraint on δ or α : τ and δ are independent variables in the Bayesian model.

3.3. Gibbs sampling

In order to obtain visualizable and interpretable results, we need to compute the average and other measures of P over the distribution given in Eq. (21). Due to the multivariate and constrained nature of the expression, this is not possible analytically. Therefore, we need to resort to numerical sampling techniques to generate a set of P and δ such that any expectation (e.g. mean) over this set approximates well the corresponding analytical expectation value over the entire posterior distribution, $\pi(P, \delta|S, \tau)$.

For this, we use the Gibbs sampler, which is a Markov-Chain Monte-Carlo (MCMC) algorithm for drawing numerical samples from a multivariate probability distribution [34]. The Gibbs sampler generates samples of P and δ from $\pi(P, \delta|S, \tau)$ by successively sampling from the posterior of each variable (P and δ) in turn. In our case, the necessary posteriors are $\pi(P|\delta, S, \tau)$ and $\pi(\delta|P, S, \tau)$.

To get $\pi(P|\delta, S, \tau)$, we take Eq. (21) and omit terms that do not contain P . As shown in the SI, the resulting expression gives a multivariate normal distribution in P

$$\pi(P|\delta, S, \tau) \propto \exp\left[-\frac{1}{2}(P - P_{\text{MAP}})^T \Sigma^{-1} (P - P_{\text{MAP}})\right] \quad (24)$$

with mean $P_{\text{MAP}} = \tau \Sigma K^T S$ and covariance matrix $\Sigma = (\tau K^T K + \delta L^T L)^{-1}$. Drawing a random sample of P from this multivariate normal distribution under the nonnegativity constraint $P \geq 0$ can be done by solving the following randomized minimization problem ([33], see SI):

$$\min_{P \geq 0} \left[P^T \Sigma^{-1} P + P^T \left(\tau K^T S + (C_L^{-1})^T v \right) \right] \quad (25)$$

where v is a random vector sampled from the normal distribution $\text{Normal}(0, I)$ with an identity covariance matrix, and C_L is the lower triangular Cholesky factor of Σ such that $C_L C_L^T = \Sigma$ (see SI). To solve Eq. (25), we use the Fast Non-Negative Least Squares (FNNLS) algorithm [28] (see Section 7).

To get $\pi(\delta|P, S, \tau)$, we again take Eq. (21) and omit terms independent of δ . This results in a gamma distribution with parameters $\tilde{a} = a + \frac{n}{2}$ and $\tilde{b} = b + \frac{1}{2}\|LP\|^2$:

$$\pi(\delta|P, S, \tau) \propto \text{gamma}(\delta; \tilde{a}, \tilde{b}). \quad (26)$$

One common approach with gamma hyper-priors is to use small values for a and b that satisfy $a \ll n/2$ and $b \ll \|LP\|^2/2$. With these values, a and b have little influence on \tilde{a} and \tilde{b} and on the posterior of δ in Eq. (26), therefore introducing minimal bias [33]. Initially, to achieve such an uninformative flat distribution, we chose b values between 10^{-6} and 10^{-3} , which render the decay rate of the exponential very slow [33], as in Fig. 2. We found that b had to be 10^{-5} or smaller to avoid influencing the results. This choice resulted in estimates for α that were unreasonably large, so we instead used the LOO-CV estimate of α to select b . Recalling the relationship between δ and α from Eq. (23), we set the mean of Eq. (26) (\tilde{a}/\tilde{b}) equal to $\alpha_{\text{LOOCV}}^2 \tau$ and solve for b to obtain:

$$b = \frac{\tilde{a}}{\alpha_{\text{LOOCV}}^2 \tau} - \frac{1}{2} \|LP_{\text{init}}\|^2 \quad (27)$$

where P_{init} is calculated via Eq. (25) using $\delta = \alpha_{\text{LOOCV}}^2 \tau$. To draw random samples from the gamma distribution, we use the algorithm of Marsaglia and Tsang [35].

The Gibbs sampler explores the full (P, δ) space by taking random steps (Monte Carlo) along each parameter dimension, using the last value sampled for the other parameter (Markov Chain [15,36]):

```

Set  $\delta_0 = \tau \alpha_{\text{LOOCV}}^2$ 
Set  $P_0 = P_{\alpha_{\text{LOOCV}}}$ 
Set  $t = 1$ 
while not converged do
   $\delta_t \sim \pi(\delta|P_{t-1}, S, \tau)$ 
   $P_t \sim \pi(P|\delta_t, S, \tau)$ 
   $t = t + 1$ 
end while

```

The symbol \sim denotes drawing a random sample from the distribution to its right, and subscripted t refers to the sample index. Note that δ_t , the sampled value of δ , is used immediately in the generation of a sample of P , P_t .

The Gibbs sampler produces a chain of P and δ values that is guaranteed to converge to the true distribution [15,17,34]. Upon convergence, the algorithm yields a large set of (P_t, δ_t) samples, distributed over parameter space such that their density is proportional to $\pi(P, \delta|S, \tau)$. The initial part of the chain (called the burn-in) is usually discarded, since initially the distribution is not yet stationary and depends on the particular choice of the starting value.

Assessment of convergence needs to be done carefully. We consider the sampling converged when the 2nd, 25th, 50th, 75th, and 98th percentiles for each point in P change by less than 10^{-3} nm^{-1} . For more easily computed values such as α and the mean of P , we use a multi-chain convergence test [33,37]. This method requires that multiple chains are run from different starting points. At each convergence check, the intra- and inter-chain variances of some test statistic are compared. When the inter- and intra-chain variances coincide, the sampler is considered converged. The details are discussed in the SI.

3.4. Analyzing the Gibbs samples

Next, we calculate the median and spread estimates for the set of P vectors and δ values from the Gibbs sampler to obtain measures of uncertainty for P and δ . For each r_j , the values of $P_t(r_j)$

are sorted into quantiles, using method 8 from [38]. Specifically, we calculate the 2nd, 25th, 50th, 75th, and 98th percentiles. We generate distance distribution plots, where the median P (50th percentile) is shown as a line, and the other quantiles are used to draw error bands. The 2nd and 98th percentiles encapsulate the 96% Bayesian credible interval (BCI). This interval indicates that the parameter in question has a 96% probability of falling within the interval [17]. Likewise, the 25th and 75th percentiles bound the 50% credible interval.

We analyze the distribution of δ , and consequently α , in a similar fashion. From the set of δ samples, we obtain a mean and the root-mean-square deviation. The method therefore automatically determines α , starting from the LOO-CV estimate, without the need of user choices and the danger of user bias and inconsistency. However, any systematic error in the LOO-CV estimate will affect the final results. We next show that the Bayesian method described above captures well the effect of noise on P .

4. Results

4.1. Unimodal and bimodal distributions

4.1.1. Credible intervals

Fig. 3 shows DEER time-domain traces generated from two model distance distributions (one unimodal and one bimodal) and three different levels of noise. For these data, six independent chains per data set converged within 11,500–20,500 iterations. The sets of P and δ from the resulting chains represent the posterior distribution and are illustrated in Fig. 4. The results are plotted in terms of α , rather than δ . This scatter plot of the value of P at $r = 3 \text{ nm}$ vs. α for each iteration of two chains of the Gibbs sampler reveals that most of the samples fall near to the model value for P and the α that best recovered the model through Tikhonov regularization. As the distance from the maximum of the posterior increases, the density of points decreases. To illustrate the Markov chains that built this particular posterior, the initial values and the first 25 iterations for two chains are highlighted. The sampler alternates between steps in the δ and P dimensions. The initial value starts the Markov Chain in a region of low probability, and then quickly finds the region of high probability, sampling around the posterior. Though highly informative, this view of the posterior would be difficult to visualize for all $n = 400$ points in each P considered, so we examine the results in terms of quantiles.

The resulting median and the 50% and 96% credible intervals obtained from the Gibbs sampler for simulated data are shown in Fig. 5. The width of the credible intervals increases with the noise level and decreases with increasing δ : variance is reduced at the cost of increased bias (see SI), but τ dominates. High noise levels may lead to much more drastic deviations in the distance distribution. Bimodal distributions display wider credible intervals than unimodal distributions for a given noise level, again matching our expectations. The average and standard deviations of α , obtained from δ via Eq. (23), are given in Fig. 5.

Small features adjacent to the primary peaks are likely to be artifactual. For example, the small feature at 4.8 nm in the unimodal $P(r)$ s in Fig. 3 is shown in Fig. 5 to be indistinguishable from zero and therefore artifactual at all noise levels considered.

The credible intervals in Fig. 5 show the persistence of a second peak at 3.9 nm for bimodal $P(r)$ s in the mid- and high-SNR cases. In the low-SNR case, where Tikhonov regularization is unable to recover a bimodal distribution due to poor data quality, the credible intervals reflect that the second peak cannot be resolved. In some cases, as in Fig. 5 top right panel, the BCIs do not encapsulate all of the original model distribution. This is likely due to a systematic error introduced by the LOOCV method that is unable to deter-

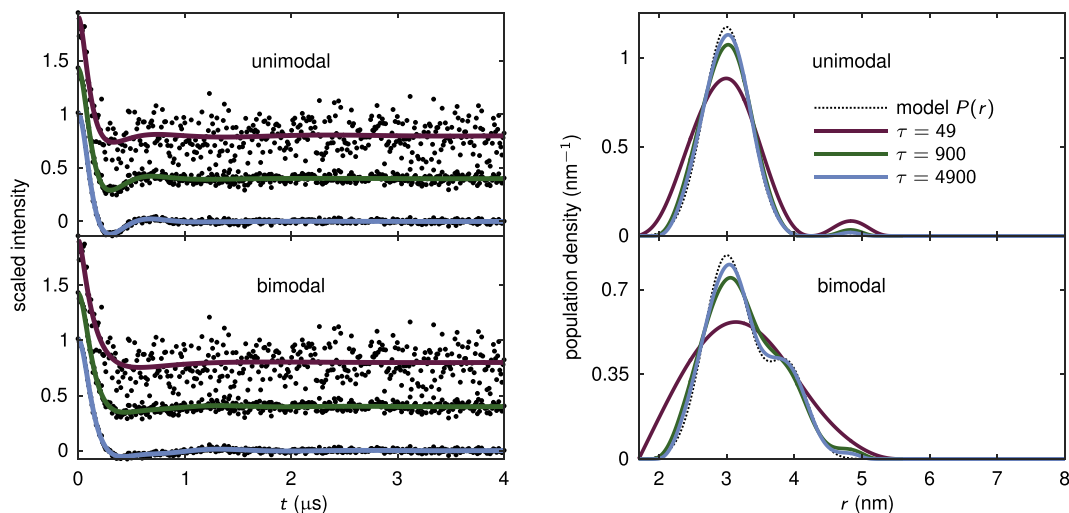


Fig. 3. Simulated model data. Left panel: DEER time traces with varied SNR generated from unimodal and bimodal distributions are shown as black dots in the top and bottom plots, respectively. The unimodal distribution is a normalized Gaussian centered at 3 nm with a FWHM (full-width at half maximum) of 0.8 nm. The bimodal distribution is a normalized sum of two Gaussians with centers of 3 and 3.9 nm, FWHM of 0.8 and 0.7 nm, and weights of 1 and 0.4, respectively. For all S and P , $n = 400$. Right panel: the corresponding unimodal and bimodal $P(r)$ vectors obtained from Tikhonov regularization are displayed in the top and bottom plots, respectively. The model $P(r)$ is shown as a dotted line and the low-, medium-, and high-SNR solutions are shown as magenta, green, and blue curves, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

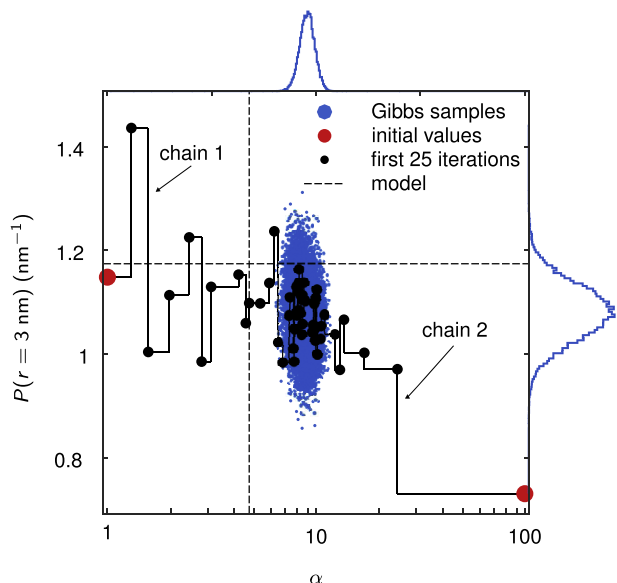


Fig. 4. Illustrative plot of the result of a Gibbs sampler, showing $P_t(r_i)$ vs. α , from the unimodal $\tau = 900$ (medium noise) case (Fig. 3), including marginal histograms for both variables. The scatter plot shows two chains of Gibbs sampled values for $P(r_i = 3 \text{ nm})$ vs. the sampled δ values (converted to α through Eq. (23)). This distance corresponds to the peak maximum of the model P . Chain 1 begins at $\alpha = 1$ and chain 2 begins at $\alpha = 100$. The dashed line in the α dimension corresponds to the α that minimizes the error between the Tikhonov solution and the model P . After a brief burn-in for both chains, the algorithm finds and samples the posterior distribution, visiting regions of high probability more often than those of low probability. The offset from the model values is due to noise and the systematic bias introduced by the LOO-CV estimate of α .

mine the optimal regularization parameter. This error affects the MCMC procedure because the LOOCV estimate is used to determine the hyperprior.

4.1.2. Histograms

By generating histograms of single-point estimators such as the mean, the mode, or the variance for each P_t , we can gain additional insight. Fig. 6 shows an example for the model data from Fig. 3. The

histograms for the high-noise cases are larger than those for the low- and medium-noise cases because the Gibbs sampler took many more iterations to converge. In both the unimodal and bimodal cases, the location of the most populated distance does not significantly change with increasing SNR, though it does become more precise. This stability of the mode with respect to noise has been previously observed [11]. In contrast, the peak width is statistically significantly different at different noise levels, and unrelated to the precision of the mode. However, the confounding effect of overestimation of δ due to the systematic error in the LOOCV estimator, which would tend to increase peak width via oversmoothing, cannot be ruled out. The mean likewise shifts with increasing noise.

4.1.3. Regularization parameter

In Tikhonov regularization, error assessment for α is not performed. One distinct advantage of the method described here is that it generates a distribution of δ and thereby α , allowing statistical analysis of this parameter. Fig. 7 shows an example. For the high- and medium-SNR cases from Fig. 3, the Bayesian MCMC approach generates a narrowly distributed set of α s close to the value that minimizes error between P_α and the model P . The discrepancy is greater in the low-SNR case. LOO-CV outperforms the ‘L’-curve, but both heuristic methods overestimate α . This systematic overestimation may result in an underestimation of the credible intervals (see Fig. S2).

4.2. Mixed-width distribution

One example that is frequently brought up to describe the limitations of Tikhonov regularization is a bimodal $P(r)$ consisting of a wide and a narrow Gaussian peak of roughly equal area (Fig. 8) [11,14]. From the noisy time-domain signal corresponding to this $P(r)$, Tikhonov regularization predicts the narrow peak well, but represents the wide peak as a combination of features that have widths similar to that of the narrow peak. The smoothness bias in Tikhonov regularization applies approximately equally over the entire range of r and, in this case, this bias prevents faithful recovery of the original mixed-width bimodal $P(r)$. The Bayesian method outlined here provides error bands that visually indicate

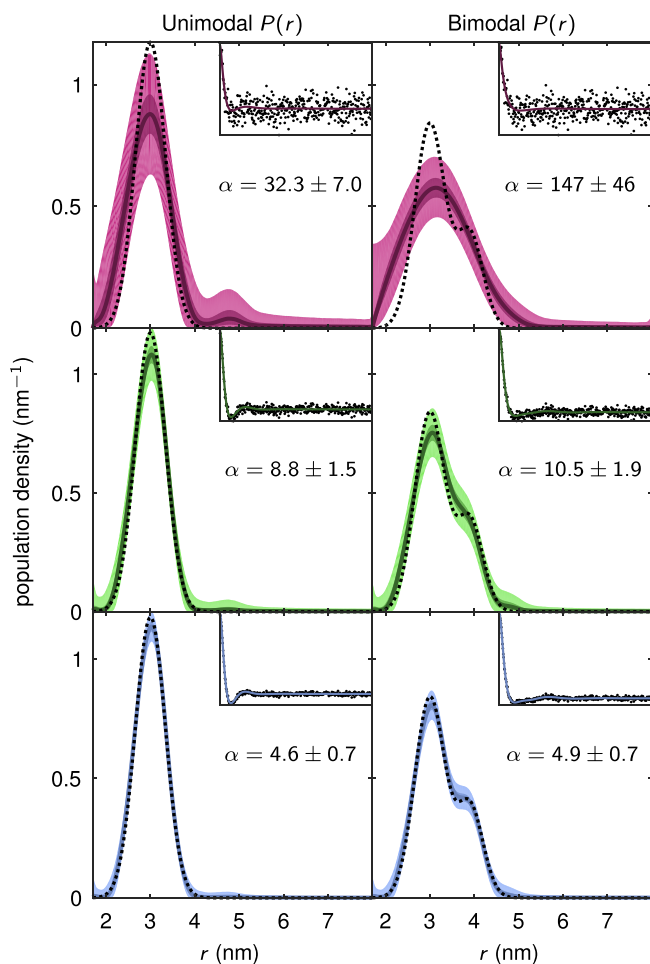


Fig. 5. Bayesian credible intervals on $P(r)$ from model data (Fig. 3). Median P solutions and Bayesian credible intervals for the unimodal and bimodal models are displayed in the left and right columns, respectively. Low-, medium-, and high-SNR cases are in the top, middle, and bottom rows, respectively. The 96% and 50% credible intervals are light and dark bands, respectively. The model P s are shown as black dashed lines. Model S traces with variable SNR (black dots) and time-domain fits from the median P s are shown as insets. α is reported as the mean plus or minus 2σ .

that the shape of $P(r)$ is quite uncertain in the area of the wide feature. This shows that the wide peak cannot be recovered as faithfully as the narrow feature.

5. Discussion

Although not the only source of uncertainty in DEER distance distributions, noise in the time-domain signal contributes spurious features and partially determines the degree of regularization required. Researchers often go to great lengths to minimize the presence of noise. As a first step towards full uncertainty quantification in DEER analysis, the present work successfully quantifies the uncertainty in $P(r)$ stemming from experimental noise. This will make the interpretation of DEER data easier and more robust, as more quantitative arguments about the significance or insignificance of a given feature can be made.

5.1. Interpretation guidelines

Bayesian credible intervals are a natural way to visualize the uncertainty in $P(r)$ given a certain noise level in the data. Features lying within a given credible interval are statistically indistinguish-

able at that level of credibility. The credible intervals can therefore guide interpretation of DEER distance distributions and help guard against overinterpretation. Next, we consider the following typical interpretative questions: (a) Significance of small peaks, (b) presence of shoulders, and (c) relevance of the difference between two distributions.

- Small side peaks are not uncommon in experimental DEER distance distributions. With the credible intervals, it is possible to assess whether they are significant or not. If the credible interval of a small peak includes the baseline, then the small peak is statistically indistinguishable from zero at that level of credibility and must be considered insignificant given the data. For example, in the top left panel of Fig. 5, the small feature at about 4.7 nm can be considered insignificant. Likewise, the small feature about 1.7 nm in Fig. 8 can be discounted.
- Shoulders in distributions can indicate conformational subpopulations. If a straight reference line (or a spline) connecting two points on the median distribution to the left and to the right of the shoulder fully falls within the credible intervals, then the shoulder is not significant. In such a case, a P with a shoulder is statistically indistinguishable from a P without a shoulder. For example, in the bottom right panel of Fig. 5, a reference line connecting the median P from the left and right of the higher distance feature about 3.9 nm runs outside of the credible interval, implying that the shoulder is statistically significant. In Fig. 8, the shoulder feature at about 2.6 nm can be considered insignificant, as a line connecting the median P local maximum at about 3.2 nm to the same curve at about 2.2 nm falls entirely within the BCI. Overall, this indicates that both the Tikhonov solution and the P median show detail between 2.5 and 3.5 nm that is not significant.
- When comparing two distributions, the credible intervals allows the practitioner to quickly identify whether the distributions are significantly different. If the credible intervals of two P do not overlap, then they are different. If there is overlap in the credible intervals, then significance testing can be employed to answer basic questions such as: are the mean spin-spin distances significantly different between these two samples? For example, the BCIs in Fig. 5 for the unimodal and bimodal high- or mid-SNR cases are clearly distinct, revealing a statistically significant difference. Additionally, one can construct histograms as in Fig. 6 and perform a Student's t -test. Fig. 6 demonstrates that $P(r)$ peak width can be significantly affected by experimental noise. This implies that quantifying uncertainty in the width of $P(r)$ is crucial in studies that compare the relative widths of distributions.

5.2. Comparison with other methods

In the literature, the quality of a given $P(r)$ is most often assessed by examining the time-domain data, considering the magnitude of the regularization parameter, and then making a qualitative judgement call (see SI). Several attempts have been made to quantify uncertainty in DEER. Those quantitative estimates of error from signal noise that the authors have encountered come from one of three methods: the noise feature of the Validation Tool in the software DeerAnalysis [13]; error propagation [13,27]; and various model-based approaches [14,24,39–45].

5.2.1. DeerAnalysis

DeerAnalysis [13] is the most commonly used software for DEER data processing, and it uses Tikhonov regularization to calcu-

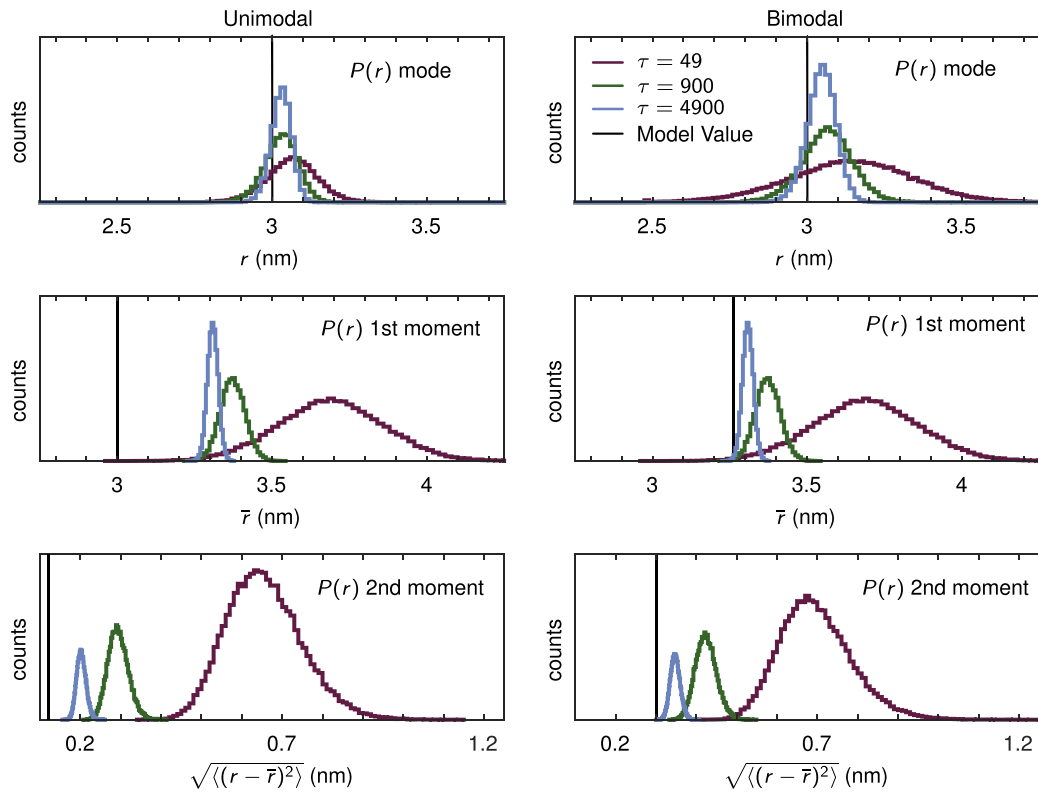


Fig. 6. Analysis of the Gibbs sampler output. Left column: unimodal case. Right column: bimodal case. Top row: $P(r)$ mode. Middle row: $P(r)$ 1st moment. Bottom row: $P(r)$ 2nd moment. The mode of the $P(r)$ is largely invariant to noise, but its precision does correlate with SNR. Added noise does cause statistically significant shifts in the 1st moment and especially in the 2nd moment. Such analysis can be used to demonstrate the statistical difference or similarity of two separate $P(r)$ s.

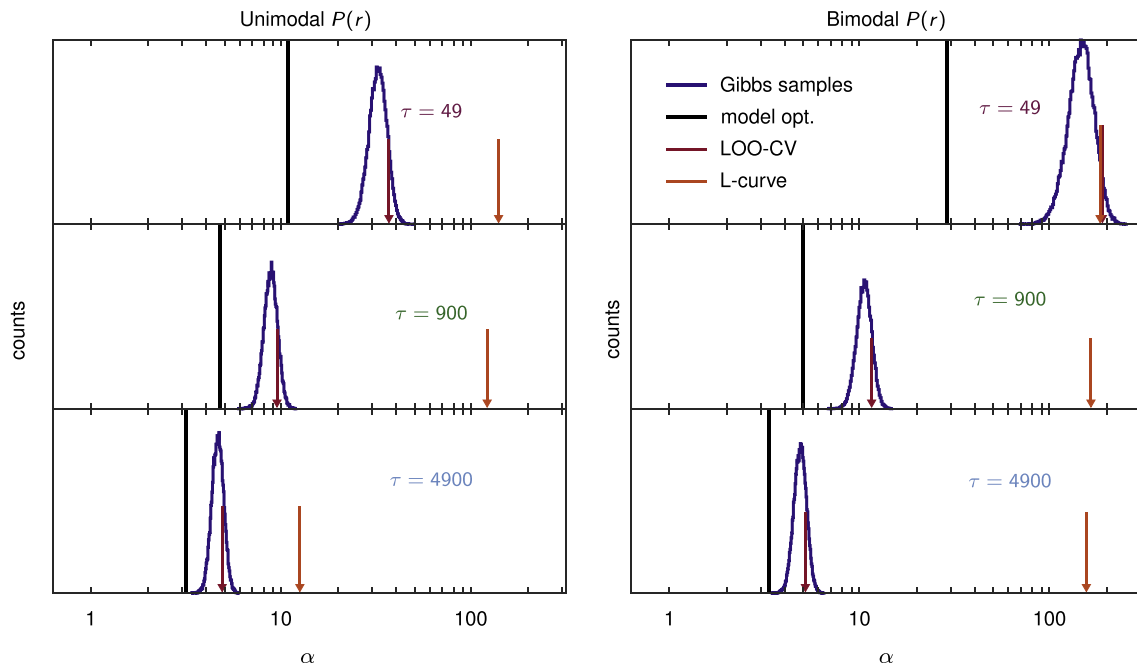


Fig. 7. Comparison of α values for the model data in Fig. 3 determined from the maximum curvature of the L-curve (orange), LOO-CV (dark red), and Gibbs sampling (dark blue), with the value that best recovers the model using Tikhonov regularization (black). Left column: unimodal case. Right column: bimodal case. Top row: low SNR. Middle row: medium SNR. Bottom row: high SNR. The Gibbs sampler out-performs point estimates for α in all cases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

late $P(r)$. The method proposed here quantifies uncertainty in the Tikhonov solution for $P(r)$, so DeerAnalysis is the most comparable extant method. DeerAnalysis provides the capability to assess error

from time-domain noise by repeatedly adding additional normally distributed synthetic noise to the data V/V_0 and then examining the resulting changes in the P vectors. This approach falls under

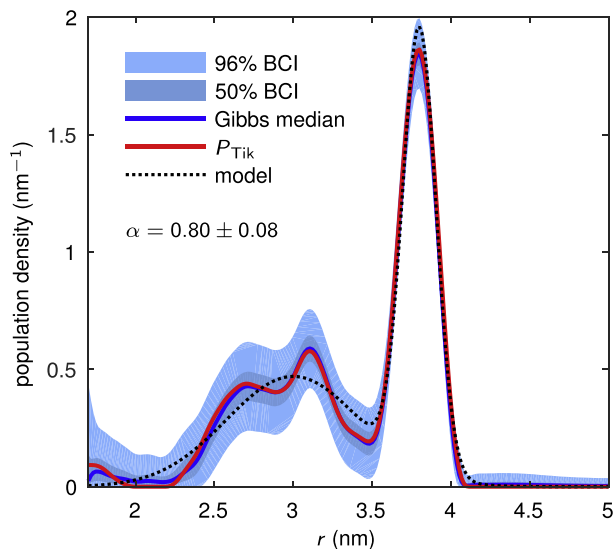


Fig. 8. Bayesian analysis applied to a bimodal mixed-width distribution. The model was generated with two equally weighted Gaussian peaks centered at 3 and 3.8 nm with FWHM of 1 and 0.25 nm, respectively. Normally distributed noise with $\tau = 2500$ (SNR = 50) was added to the corresponding S . Both the Tikhonov solution and the Gibbs median P fail to recapitulate the model distribution. The Bayesian credible intervals indicate that there is increased uncertainty in the regions where the Tikhonov solution is the most incorrect.

the category of bootstrapping methods. The user sets several parameters, including the magnitude of the added synthetic noise and the number of iterations. The noise is therefore not guaranteed to be equal in magnitude to the experimental noise level, compromising accuracy. Convergence is likewise not guaranteed. The error bars displayed are the $\pm 2\sigma$ bands of the set of calculated P vectors. This visualization is useful if the population is normally, or close to normally, distributed. This is not guaranteed to be the case, especially where P is zero or close to zero. The confidence intervals often include regions of negative probability, which is not physically valid, so negative valued indices of the $\pm 2\sigma$ bands are set to 0. No error analysis regarding the uncertainty in the choice of α can be performed.

Compared to the validation tool in DeerAnalysis, the present method is more time consuming because of the large number of required iterations. DeerAnalysis also considers additional sources of uncertainty, such as background correction and pulse bandwidth. However, the Bayesian method features convergence criteria, utilizes real noise information, reduces user bias, and provides uncertainty estimates for α .

5.2.2. Propagation

The next method is more limited in scope than either DeerAnalysis or the proposed Bayesian approach. The variance of the noise in $S(t)$, σ_s^2 , can be propagated from the time domain to the distance domain using $\sigma_p^2 = \sigma_s^2 \text{diag}(Q_z' Q_z'^T)$ where the primes indicate that only rows and columns in Q_z (as defined in Eq. (13)) are retained that correspond to the indices where the constrained Tikhonov solution is non-zero. Commercial EPR spectrometers store acquired data only after signal averaging, so noise statistics are usually unavailable. In this case, the fit RMSD (root mean square difference) between the Tikhonov fit KP and the data S is often used as a stand-in for σ_s . As a result, any error in the fit will corrupt the calculation. This approach gives no error estimate for regions where the Tikhonov solution is zero. This method is used in the program FTIKREG [27], which was used by DeerAnalysis until 2016. We have used this method in the past [46,47].

Unlike the propagation method, for which true noise statistics are usually unavailable, the proposed Bayesian method uses observed noise information. It also provides estimates for the full r domain, including regions where $P(r)$ has been constrained to non-negative values. Finally, the propagation method considers α to be given and fixed, and so provides no uncertainty estimates for the level of regularization applied.

5.2.3. Model fitting

In addition to the Bayesian approach, it is possible to use classical frequentist methods. One such approach to quantify uncertainty is to assume a particular model for $P(r)$, fit the model to the data, and then statistically evaluate the error in the various fit parameters, usually with some variant of the χ^2 cost function. This method has been employed in LongDistances [45], DeerSim [40–43], and DEERconstruct [44]. Alternatively, DEFit fits Gaussian distributions directly to the time-domain data [39]. Hustedt and coworkers have implemented a model-based approach that simultaneously fits the background function and $P(r)$ to time-domain data, thereby uncovering potential correlations between background parameters and features in $P(r)$ [14,24].

The Bayesian method proposed here quantifies uncertainty arising from time-domain noise and regularization, ameliorating one of the disadvantages associated with Tikhonov regularization compared to model-based approaches. The Bayesian approach so far does not take into account background correction. As with Tikhonov regularization, the proposed method is model-free with respect to $P(r)$. The credible intervals produced by the present method do not include any negative values.

With the Bayesian credible intervals in hand, it is now possible to quantitatively compare model-based and model-free methods. This will require an extensive benchmark set of distributions that are likely to be encountered in practice.

6. Conclusions

We introduce a Bayesian statistical method for the analysis of DEER spectroscopic data. The method extends the existing method of Tikhonov regularization: First, it not only determines $P(r)$, but it also fully quantifies the associated uncertainty (error) due to noise in the data. Second, it automatically determines the regularization parameter α and provides an associated error estimate. The uncertainty regarding α is included in the overall error estimates for P . The statistically determined uncertainty estimates are visualized using quantiles. This is an important improvement to the application of DEER spectroscopy as it guides interpretation, reduces practitioner bias, increases reliability, and helps reproducibility.

Although it outperforms current methods, the LOOCV method, used to set the hyperparameter b in the δ hyperprior (Eq. (20)), introduces a systematic overestimation of the level of regularization. This error is propagated into the final results, where it may have the effect of underestimating the credible intervals, and should be kept in mind. We hope that a better hyperprior will be found in the future to eliminate this bias.

The Bayesian treatment is far from complete. There are still many other sources of uncertainty affecting the DEER distance distribution that need to be addressed. The error estimates determined here do not capture uncertainty arising from background correction, time-domain truncation, and discretization. Uncertainty in the background correction, in particular, can give rise to very significant uncertainty in the distance distribution [11,13]. The proposed method may be combined sequentially with the validation tool of DeerAnalysis to assure that a good choice for the background function has been used. The Bayesian statistical approach outlined here is flexible, general and robust enough to

be extended to capture the influences of, and correlations between, these other sources of uncertainty.

7. Methods

All computations were carried out using MATLAB R2016a [48]. The dipolar kernel matrix K was calculated using Fresnel integrals [49]. Non-negativity constrained minimization for Tikhonov regularization and for Gibbs sampling was done using the Fast Non-Negative Least Squares (FNNLS) algorithm [28]. Each minimization took between 0.01 and 0.1 s on a standard laptop computer. Uniform and normally distributed pseudo-random numbers were generated with the Mersenne Twister pseudo-random number generator (PRNG) as implemented in MATLAB R2016a using a seed value of 19,560 to generate model data and 12,345 for Gibbs sampler runs. The PRNG was reset before each MCMC run, but not during a run, i.e. not between chains. The gamma-distributed pseudo-random numbers were generated using the algorithm of Marsaglia and Tsang [35,50].

Acknowledgements

We thank Jonathan M. Bardsley (University of Montana) [33] for a very helpful discussion over breakfast, Francois Monard (University of Washington) for useful literature recommendations, and Niket Thakkar (University of Washington) for valuable advice. This work was supported by the National Science Foundation (NSF) grant CHE-1452967, the National Institutes of Health (NIH) grant R01 EY010329 (S.S.), and the National Institute of General Medical Sciences (NIGMS) of the NIH under Award Number T32GM008268 (T.H.E.).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jmr.2016.06.021>.

References

- [1] A.D. Milov, A.B. Ponomarev, Yu.D. Tsvetkov, Electron-electron double resonance in electron spin echo: model biradical systems and the sensitized photolysis of decalin, *Chem. Phys. Lett.* 110 (1984) 67–72.
- [2] R.G. Larsen, D.J. Singel, Double electron-electron resonance spin-echo modulation: spectroscopic measurement of electron spin pair separations in orientationally disordered solids, *J. Chem. Phys.* 98 (1993) 5134–5146.
- [3] R.E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich, H.W. Spiess, Determination of end-to-end distances in a series of TEMPO diradicals of up to 2.8 nm length with a new four-pulse double electron electron resonance experiment, *Angew. Chem. Int. Ed.* 37 (1998) 2834–2837.
- [4] G. Jeschke, DEER distance measurements on proteins, *Annu. Rev. Phys. Chem.* 63 (2012) 419–446.
- [5] O. Dalmás, H.C. Hyde, R.E. Hulse, E. Perozo, Symmetry-constrained analysis of pulsed double electron-electron resonance (DEER) spectroscopy reveals the dynamic nature of the KcsA activation gate, *J. Am. Chem. Soc.* 134 (2012) 16360–16369.
- [6] I.D. Sahu, R.M. McCarrick, K.R. Troxel, R. Zhang, H.J. Smith, M.M. Dunagan, M.S. Swartz, P.V. Rajan, B.M. Kroncke, C.R. Sanders, G.A. Lorigan, DEER EPR measurements for membrane protein structures via bifunctional spin labels and lipidic nanoparticles, *Biochemistry* 52 (2013) 6627–6632.
- [7] J.N. Rao, C.C. Jao, B.G. Hegde, R. Langen, T.S. Ulmer, A combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins, *J. Am. Chem. Soc.* 132 (2010) 8657–8668.
- [8] R. Guo, K. Gaffney, Z. Yang, M. Kim, S. Sungsuwan, X. Huang, W.L. Hubbell, H. Hong, Steric trapping reveals a cooperativity network in the intramembrane protease GlpG, *Nat. Chem. Biol.* 12 (2016) 353–360.
- [9] I. Kaminker, M. Bye, N. Mendelman, K. Gislason, S.Th. Sigurdsson, D. Goldfarb, Distance measurements between manganese(II) and nitroxide spin-labels by DEER determine a binding site of Mn²⁺ in the HP92 loop of ribosomal RNA, *Phys. Chem. Chem. Phys.* 17 (2015) 15098–15102.
- [10] P. Kumar, I.M.J. Segers-Nolten, N. Schilderink, V. Subramaniam, M. Huber, Parkinson's protein α -Synuclein binds efficiently and with a novel conformation to two natural membrane mimics, *PLOS ONE* 10 (2015) 1–11.
- [11] G. Jeschke, G. Panek, A. Godt, A. Bender, H. Paulsen, Data analysis procedures for pulse ELDOR measurements of broad distance distributions, *Appl. Magn. Reson.* 26 (2004) 223–244.
- [12] Y. Chiang, P.P. Borbat, J.H. Freed, The determination of pair distance distributions by pulsed ESR using Tikhonov regularization, *J. Magn. Reson.* 172 (2005) 279–295.
- [13] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C.R. Timmel, D. Hilger, H. Jung, DeerAnalysis2006: a comprehensive software package for analyzing pulsed ELDOR data, *Appl. Magn. Reson.* 30 (2006) 473–498.
- [14] R.A. Stein, A.H. Beth, E.J. Hustedt, Inter-spin label distance measurements: analysis considerations and approaches, in: P.Z. Qin, K. Warncke (Eds.), *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Part A*, vol. 563, Elsevier, 2015, pp. 531–564 (Chapter 20).
- [15] J. Kaipio, E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2005.
- [16] L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. Marzouk, L. Tenorio, B. van Bloemen Waanders, K. Willcox, *Large-Scale Inverse Problems and Quantification of Uncertainty*, John Wiley & Sons, Ltd, 2011.
- [17] R.C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM Computational Science and Engineering, 2014.
- [18] J.F. Giovannelli, J. Idier, *Regularization and Bayesian Methods for Inverse Problems in Signal and Image Processing*, Wiley, 2015.
- [19] G. Zhou, V.A. Voelz, Bayesian inference of conformational state populations from computational models and sparse experimental observables, *J. Comput. Chem.* 35 (2014) 2215–2224.
- [20] K.E. Hines, A primer on Bayesian inference for biophysical systems, *Biophys. J.* 108 (2015) 2103–2113.
- [21] J.L. MacCallum, A. Perez, K.A. Dill, Determining protein structures by combining semireliable data with atomistic physical models by Bayesian inference, *Proc. Natl. Acad. Sci. USA* 112 (2015) 6985–6990.
- [22] K.M. Salikhov, I.T. Khairuzhdinov, R.B. Zaripov, Three-pulse ELDOR theory revisited, *Appl. Magn. Reson.* 45 (2014) 573–619.
- [23] K.M. Salikhov, I.T. Khairuzhdinov, Four-pulse ELDOR theory of the spin 1/2 label pairs extended to overlapping EPR spectra and to overlapping pump and observer excitation bands, *Appl. Magn. Reson.* 46 (2015) 67–83.
- [24] S. Brandon, A.H. Beth, E.J. Hustedt, The global analysis of DEER data, *J. Magn. Reson.* 218 (2012) 93–104.
- [25] NIST Digital Library of Mathematical Functions. <<http://dlmf.nist.gov/>> (Release 1.0.10 of 2015-08-07. Online companion to [26]).
- [26] F.W.J. Olver, D.W. Lozier, R.F. Boisvert, C.W. Clark (Eds.), *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, 2010 (Print companion to [25]).
- [27] J. Weese, A reliable and fast method for the solution of Fredholm integral equations of the first kind based on Tikhonov regularization, *Comput. Phys. Commun.* 69 (1992) 99–111.
- [28] R. Bro, S. Jong, A fast non-negativity-constrained least squares algorithm, *J. Chemometr.* 11 (1997) 393–401.
- [29] P.C. Hansen, The L-curve and its use in the numerical treatment of inverse problems, in: P. Johnston (Ed.), *Advances in Computational Bioengineering: Computational Inverse Problems in Electrocardiology*, WIT Press, 2000, pp. 119–142.
- [30] P.C. Hansen, *Discrete Inverse Problems: Insight and Algorithms*, SIAM, 2010.
- [31] G. Wahba, *Spline Models for Observational Data*, SIAM, 2007.
- [32] V.A. Voelz, V.S. Pande, Calculation of rate spectra from noisy time series data, *Proteins* 80 (2012) 342–351.
- [33] J.M. Bardsley, C. Fox, An MCMC method for uncertainty quantification in nonnegativity constrained inverse problems, *Inverse Probl. Sci. Eng.* 20 (2012) 477–498.
- [34] R.M. Neal, *Probabilistic Inference Using Markov Chain Monte Carlo Methods*, Technical Report CRG-TR-93-1, Department of Computer Science, University of Toronto, 1993.
- [35] G. Marsaglia, W.W. Tsang, A simple method for generating gamma variables, *ACM Trans. Math. Softw.* 26 (2000) 363–372.
- [36] J.J. France, Y. Gur, R.M. Kirby, C.R. Johnson, A Bayesian approach to quantifying uncertainty in Tikhonov solutions for the inverse problem of electrocardiography, *Comput. Card. 41* (2014) 529–532.
- [37] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin, *Bayesian Data Analysis*, third ed., Taylor & Francis Group LLC, 2014.
- [38] R.J. Hyndman, F. Yanan, Sample quantiles in statistical packages, *Am. Stat.* 50 (1996) 361–365.
- [39] K.I. Sen, T.M. Logan, P.G. Fajer, Protein dynamics and monomer-monomer interactions in AnR activation by electron paramagnetic resonance and double electron-electron resonance, *Biochemistry* 46 (2007) 11639–11649.
- [40] L. Galiano, F. Ding, A.M. Veloro, M.E. Blackburn, C. Simmerling, G.E. Fanucci, Drug pressure selected mutations in HIV-1 protease alter flap conformations, *J. Am. Chem. Soc.* 131 (2009) 430–431.
- [41] M.E. Blackburn, A.M. Veloro, G.E. Fanucci, Monitoring inhibitor-induced conformational population shifts in HIV-1 protease by pulsed EPR spectroscopy, *Biochemistry* 48 (2009) 8765–8767.
- [42] I.M.S. de Vera, A.N. Smith, M.C.A. Dancel, X. Huang, B.M. Dunn, G.E. Fanucci, Elucidating a relationship between conformational sampling and drug resistance in HIV-1 protease, *Biochemistry* 52 (2013) 3278–3288.
- [43] X. Huang, M.D. Britto, J.L. Kear-Scott, C.D. Boone, J.R. Rocca, C. Simmerling, R. McKenna, M. Bieri, P.R. Gooley, B.M. Dunn, G.E. Fanucci, The role of select

- subtype polymorphisms on HIV-1 protease conformational sampling and dynamics, *J. Biol. Chem.* 289 (2014) 17203–17214.
- [44] T. Casey, DEERconstruct, MATLAB Central File Exchange, 2015 (accessed and downloaded May 19, 2016).
- [45] C.J. Lopez, Z. Yang, C. Altenbach, W.L. Hubbell, Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants, *Proc. Natl. Acad. Sci. USA* 110 (2013) E4306–E4315.
- [46] M.C. Puljung, H.A. DeBerg, W.N. Zagotta, S. Stoll, Double electron-electron resonance reveals cAMP-induced conformational change in HCN channels, *Proc. Natl. Acad. Sci. USA* 111 (2014) 9816–9821.
- [47] H.A. DeBerg, J.R. Bankston, J.C. Rosenbaum, P.S. Brzovic, W.N. Zagotta, S. Stoll, Structural mechanism for the regulation of HCN ion channels by the accessory protein TRIP8b, *Structure* 23 (2015) 734–744.
- [48] The MathWorks Inc., MATLAB Release 2016a, 2016.
- [49] J. D'Errico, FresnelS and FresnelC, MATLAB Central File Exchange, 2010 (accessed and downloaded November 25, 2013).
- [50] A. Bar-Guy, RANDRAW, MATLAB Central File Exchange, 2005 (accessed and downloaded May 25, 2015).