Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 177 (2005) 184–196
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

# Maximum entropy: A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR

Yun-Wei Chiang, Peter P. Borbat, Jack H. Freed *

*Baker Laboratory of Chemistry and Chemical Biology, and National Biomedical ACERT Center for Advanced ESR Technology, Cornell University, Ithaca, NY 14853-1301, USA*

Received 23 May 2005; revised 14 July 2005
Available online 30 August 2005

**Abstract**

Tikhonov regularization (TIKR) has been demonstrated as a powerful and valuable method for the determination of distance distributions of spin-pairs in bi-labeled biomolecules directly from pulsed ESR signals. TIKR is a direct method, which requires no iteration, and, therefore, provides a rapid and unique solution. However, the distribution obtained tends to exhibit oscillatory excursions with negative portions in the presence of finite noise, especially in the peripheral regions of the distribution. The Shannon–Jaynes entropy of a probability distribution provides an intrinsic non-negativity constraint on the probability distribution and an unbiased way of obtaining information from incomplete data. We describe how the maximum entropy regularization method (MEM) may be applied to solve the ill-posed nature of the dipolar signal in pulsed ESR. We make use of it to suppress the negative excursions of the distance distribution and to increase the tolerance to noise in the dipolar signal. Model studies and experimental data are investigated, and they show that, with the initial or "seed" probability distribution that is required for MEM taken as the TIKR result, then MEM is able to provide a regularized solution, subject to the non-negativity constraint, and it is effective in dealing with noise that is problematic for TIKR. In addition we have incorporated into our MEM method the ability to extract the intermolecular dipolar component, which is embedded in the raw experimental data. We find that MEM minimization, which is implemented iteratively, is greatly accelerated using the TIKR result as the seed, and it converges more successfully. Thus we regard the MEM method as a complement to TIKR by securing a positive pair distance distribution and enhancing the accuracy of TIKR.


*Keywords:* Maximum entropy; Double-quantum coherence ESR; Double electron–electron resonance; Cytochrome *c*; T4 lysozyme

## 1. Introduction

Tikhonov regularization (TIKR) [1] has recently been introduced [2–4] to extract distance distributions between spin-pairs from intramolecular dipole–dipole interactions measured by pulsed electron spin resonance (ESR) techniques [5], in particular double quantum coherence (DQC) ESR and double electron–electron resonance (DEER). The ability to obtain such distributions, for distances as great as 7 nm [6], significantly enhances the application of ESR techniques in studying protein structure [7–9] in large biomolecules. A particularly significant application is in providing distance distributions of intermediate folding states of proteins [2,10,11].

The integral equation [2] that connects the intramolecular dipole–dipole interactions, which are measured in the ESR experiment, with the distance distributions between spin-pairs is in the form of a Fredholm equation of the first kind [1], that can be represented in discrete form as a linear equation $K(r,t)P(r) = S_0(t)$,

* Corresponding author. Fax: +1 607 255 6969.
*E-mail address:* jhf@ccmr.cornell.edu (J.H. Freed).

1090-7807/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2005.07.021

where $K$ is a kernel matrix representing the shape of a Pake doublet dipolar signal in the time domain for a given radial separation, $r$, while $P$ is the vector representing the pair distance distribution, and $S_0$ is the vector representing the intramolecular dipolar signal from pulsed ESR experiments, which includes experimental noise [2]. This discrete form, in the presence of finite noise, is ill-posed, so it cannot usually be solved in a straightforward manner by singular value decomposition (SVD) or other least squares methods, and, therefore, requires regularization methods. The difficulties connected with solving this ill-posed problem to extract $P(r)$ from pulsed ESR experiments have been previously discussed using SVD and regularization methods [2–4].

Among the regularization methods, TIKR has been generally accepted as an important one. It has been demonstrated in model and experimental studies that TIKR with a regularization parameter determined by the L-curve criterion provides a mathematically reliable estimate for the distance distribution for reasonable noise levels (i.e., SNR of 30–500 [2]). Another virtue of this TIKR method is that it is easy to compute numerically and it leads to the unique solution for a given regularization parameter, $\lambda$. That is, the formal minimization of the TIKR functional given by

$$\Phi_{\mathrm{TIKR}}[P] \equiv \|KP - S_0\|^2 + \lambda^2 \|P\|^2 \tag{1}$$

yields the simple matrix equation [12]

$$P_\lambda = \left(K^{\mathrm{T}}K + \lambda^2\right)^{-1} K^{\mathrm{T}} S_0. \tag{2a}$$

This equation for the Tikhonov functional minimizer is readily solved by SVD to yield the form given by [12]

$$P_\lambda = (\Sigma^2 + \lambda I)^{-1} V \Sigma U^{\mathrm{T}} S_0 = \sum_{i=1}^{\mathrm{Rank}(K)} f_i \frac{u_i^{\mathrm{T}} S_0}{\sigma_i} v_i, \text{ where } f_i \equiv \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \tag{2b}$$

to obtain the unique $P$ for the given value of $\lambda$. In Eq. (2b), the kernel matrix $K$ (of dimension $M \times N$) is decomposed by SVD into two orthonormal matrices, $V$ (whose columns are composed of the $N$-dimensional vectors $v_i$) and $U$ (whose columns are composed of the $M$-dimensional vectors $u_i$), and a diagonal matrix $\Sigma$ whose elements are the singular values, $\sigma_i$. It can be rewritten (as shown by the second equality on the right-hand side) as the standard SVD solution, but where the term for the $i$th singular value is multiplied by a filter function $f_i$, which acts to screen out undesired small singular values, $\sigma_i$ for which $\sigma_i \ll \lambda$. The filter, shown in Eq. (2b), is specifically for the functional in Eq. (1). The difference between various regularization methods lies in the way that the filters are defined. The analysis given by Eqs. (2a) and (2b) is easily repeated for many values of $\lambda$, from which the L-curve criterion is readily applied to yield the optimum $\lambda$.

However, a possible criticism of this approach is that the $P(r)$ obtained tends to exhibit oscillatory excursions with negative portions in the presence of finite noise, especially in the peripheral regions of the regularized distance distributions. A physically meaningful $P(r)$ must, of course be non-negative, but corrupting noise has no such requirement. In model tests it was shown that the oscillations do not significantly affect the main distribution. It nevertheless is desirable to constrain the $P(r)$ to be positive, so the TIKR method of Eq. (1) has been adapted to include this constraint, and this has been applied to pulsed ESR experiments [3,4]. The resulting computational algorithm, using a self-consistent (SC) approach to determine the optimum $\lambda$ and subjecting to the non-negativity constraint, requires iterative methods [13] that may be initiated with a good starting guess of $P(r)$ [14]. (These authors do supply a sophisticated deterministic annealing method when good initial values are not available [14].) Thus the introduction of the non-negativity constraint leads to a computationally more cumbersome algorithm.

In the present paper, we discuss the use of a different approach than Eq. (1) for the regularization in pulsed ESR. It is based on maximizing the entropy (ME) function associated with $P(r)$. This implicitly restricts $P(r)$ to be positive, but also at the expense of computational ease, requiring a numerical non-linear minimization procedure. In one study [15] the ME approach was found to be computationally significantly faster than constrained TIKR. Our main objective is to determine whether the inclusion of this information-theoretic ME principle would lead to improved estimates of $P(r)$ under conditions of lower SNR typically encountered in real experiments. Furthermore, we use the proposed method to determine $P(r)$ and to simultaneously extract information about the baseline in the experimental data. The ME minimization is implemented using a conjugate-gradient (CG) algorithm, which provides rapid local convergence. A good initial guess of $P(r)$ to seed the iterative minimization is crucial to obtaining useful results by this method. Thus we also address this issue and we report an approach that we find to work quite well, that is based on using the result of Eqs. (2a) and (2b) in conjunction with the L-curve criterion, as the seed to initiate the ME regularization approach.

Maximum entropy (ME) has been widely used as a general and powerful method for reconstructing results from noisy and incomplete data in various fields (e.g., image reconstruction [16] such as radio astronomy, medical tomography, and X-ray imaging; and fluorescence spectroscopy [17]). The idea is that by maximizing the Shannon–Jaynes entropy

$$E = -\int \alpha(\tau) \ln \alpha(\tau) \mathrm{d}\tau, \tag{3}$$

where $\alpha(\tau)$ is the function to be determined, the solution is determined/selected from the many sets of functions that can fit the data. This underlying idea has been well explained in information theory and statistical mechanics by taking entropy as a probabilistic concept [18,19], i.e., the entropy of a probability distribution can be considered as a measure of the uncertainty of the experimental outcomes. The virtues of using the maximum entropy concept are that: (i) it provides an unbiased way of obtaining information from incomplete data; (ii) it implicitly possesses the non-negativity constraint to the probability distribution. Application of ME in the context of image reconstruction was originally introduced by Frieden [20] and later, was found useful in this field. Many efforts have been made using ME to reconstruct an image from incomplete/noisy data [16,21,22], and even to increase SNR in NMR spectroscopy [23,24].

In this report, we replace the penalty term of the Tikhonov regularization functional, i.e., the second term on the right-hand side of Eq. (1) with the entropy Eq. (3) to solve the ill-posed inverse problem that we encounter in determining pair distance distributions from pulsed ESR experiments. In the past, the convergence behavior of the functional has been investigated and tested with numerical experiments. A transformed form of the entropy, which retains the advantages of using the Shannon–Jaynes entropy, is suggested in the literature and used in the analysis to provide better convergence for the functional minimization [25,26]. Details about this transformation are given in Section 2 and discussed in later sections. The model distributions and experimental data we use are the same as those used previously [2]. We find that ME regularization is best utilized as a complement to TIKR for the determination of pair distance distributions. In fact, in the early stages of this work we found that ME regularization used independently of TIKR failed to recover satisfactory results (as compared to the TIKR results) in all tested cases. However, when the result of the TIKR was used to initiate the ME regularization it does become a viable method.

The dipolar time evolution signal obtained from pulse ESR experiments includes both intra- and intermolecular interactions, where the former is what we desire and the latter usually is unwanted. The intermolecular contribution manifests itself as a modification of a DEER signal envelope, thus producing a large decaying "baseline"; whereas in DQC it leads to a small offset [7], whose slope increases with concentration. It also leads to a damping of the intramolecular signal. Analysis of the dipolar signal often becomes difficult when the spin concentration is increased, which makes these concentration effects stronger. The removal (typically by subtraction) of the baseline from raw experimental data has been performed independently of the determination of distance distributions of spin-pairs [2,7,27–29]. Such a procedure is also commonly used in the spectral analysis of nuclear ESEEM.

One proposed method was that intramolecular interactions could be separated from intermolecular interactions by studying the concentration dependence of the pulsed ESR signals. More typically, the baseline function contributed from intermolecular interactions is generally obtained by fitting a low-degree polynomial to the raw experimental data. This method, in some cases of DEER experiments, usually utilizes weighted fitting in certain regions (e.g., weighted fitting of the last part of the raw data) where intermolecular interactions dominate the overall decay. In other words, the baseline subtraction requires some prior insight. In this report we propose a more unified way, which is based on the ME regularization, to simultaneously extract the intermolecular contribution in the raw experimental data and determine the spin-pair distance distribution from the intramolecular dipolar interactions. A virtue of such an approach is that all the data are utilized to fit the baseline as well as the $P(r)$, thus, this approach is more likely to avoid bias from, e.g., artifacts in local regions of the signal.

Three models, bimodal, box-like, and broad trimodal distributions, are tested with the levels of SNR $\sim 500$ and 30. The model study is first performed without considering any baseline in the time evolution data and then tested with a model baseline in the analysis. The dipolar time evolution signals of the proteins T4 lysozyme and cytochrome *c*, respectively measured by DQC-ESR and DEER techniques, are analyzed using the ME regularization method. The improvements include successful suppression of the peripheral noise-like oscillations of the Tikhonov results, the better tolerance to the noise levels of SNR $< 30$, and simultaneous determination of the intermolecular interaction and the spin-pair distance distribution. Based on these improvements, ME regularization is proposed as a complement to the Tikhonov regularization using the L-curve criterion, and we suggest that a similar strategy might be useful for the self-consistent (SC) method of TIKR.

## 2. Method

### 2.1. MEM functional and its minimization

The maximum entropy regularization functional is given below in Eq. (4). One readily recognizes that the introduction of ME to the least squares problem is equivalent to the method of Lagrange multipliers. ME was first rigorously studied as a regularization method by Klaus and Smith [30]. It uses the entropy as the penalty term and is what we call MEM in this report. (Note that MEM is referred to maximizing the

Shannon–Jaynes entropy instead of a regularization method in some literature references.) There are two terms in the MEM functional: the first is the residual norm, which approaches zero as the regularized solution $P$ resembles the exact solution and is the same as in TIKR; the second is negative of the entropy, $-E$, multiplied by the square of regularization parameter, $\lambda$. Thus ME regularization involves minimizing the following functional for a given $\lambda$:

$$\Phi_{\mathrm{ME}}[P] \equiv \|KP - S_0\|^2 + \lambda^2 E = \|KP(r) - S_0\|^2 + \lambda^2 \int P(r) \ln P(r) \mathrm{d}r, \quad (4)$$

whereas this ME functional is a natural and appropriate way to introduce the regularization, from a practical viewpoint it requires an iterative approach to minimization that exhibits slow convergence (partly due to the complicated non-linearity of the functional) [26,31,32]. Instead of using Eq. (4) directly, it has been suggested that one use a modified form of the entropy in the functional to be minimized, [25,26,33].

$$\Phi_{\mathrm{ME}}[P] = \|KP(r) - S_0\|^2 + \lambda^2 \times \int \left[ P(r) \ln \frac{P(r)}{P_0(r)} + \frac{P_0(r)}{e} \right] \mathrm{d}r \Rightarrow \min . \quad (5)$$

In this modified entropy approach, the use of an appropriately chosen prior probability function, $P_0$ suppresses the difficulty in the numerical algorithm that one encounters when simultaneously minimizing the residual term to zero and the penalty to minus infinity. It has been shown that the modified form of the maximum entropy functional (cf. Eq. (5)) is equivalent to Eq. (4), and it is also analogous to the Tikhonov functional in its properties [25,26]. This transformed functional not only possesses the non-negativity constraint on the regularized solution for $P(r)$ by virtue of the natural logarithm but also is proven to inherit the stability and the convergence rate from the Tikhonov functional, although the actual numerical solution lacks the simplicity of that for Eqs. (2a) and (2b). The modified ME functional has been rigorously examined mathematically and tested for the case of analyzing fluorescence data [26]. No marked difference between the results obtained with Tikhonov and maximum entropy regularization, respectively, was found [26,34].

The functional minimization of Eq. (5) for the determination of the solution vector $P$ is performed using a conjugate-gradient (CG) method [35], which is generally considered as the appropriate method for large problems; i.e., problems with hundreds (or thousands) of elements in solution vector. Two versions of CG, the Fletcher–Reeves (FR) and the Polak–Ribiere (PR), have been tested for our purposes. We found that the latter is better than the former in most of the cases studied for two reasons: (i) PR converges much more quickly than FR; (ii) PR always provides a better or at least a similar solution than FR in the tested models. (The methods only differ by one line in the numerical code, and the algorithm can be found in the literature [35].) The program we developed is performed with Matlab version 7.0, and naturally uses Matlab functions for matrix and vector calculations.

The regularization parameter of the TIKR functional is determined using the L-curve criterion incorporated in the Regularization toolbox [36–38]. The subroutines we have used have been modified from those in the Regularization toolbox. The basis for the use of the L-curve criterion in TIKR is that the residual norm and the norm of the penalty term are monotonic functions of the regularization parameter. This has been proven to be true for TIKR [12], but this matter has never previously been studied for MEM. In this report, the regularization parameter of the MEM functional (cf. Eq. (5))is determined in a similar way using the maximum curvature criterion. However, the $y$-axis of the L-curve plot is now the modified entropy term rather than the solution norm. A typical L-curve obtained in the MEM analysis is shown in Fig. 1. If a very small regularization parameter is used, the curve becomes discontinuous and oscillatory (cf. the upper left corner in Fig. 1),

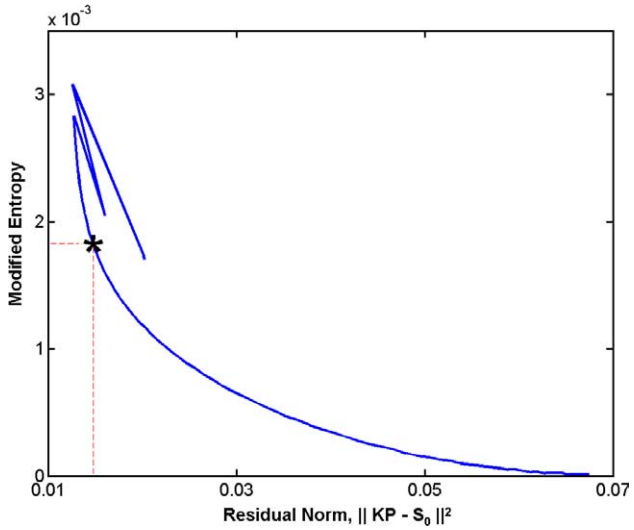


Fig. 1. A typical L-curve obtained in MEM analysis. The ordinate is given by the modified negative of entropy, i.e., by the second term on the right-hand side of Eq. (5) with $\lambda = 1$. When a small regularization parameter is used, the curve becomes discontinuous and oscillatory (the upper left corner in the plot). This is due to the fact that the residual norm oscillates back-and-forth unstably as the penalty term fails to stabilize the regularization minimization due a small regularization parameter. In such an L-curve plot the search range for an optimal regularization parameter must be restricted to the smooth region to avoid obtaining a noise-like solution.

indicating that the penalty term is too small to stabilize the solution. Thus, we ignore the range of $\lambda$ yielding such unstable behavior. Based on the model studies shown in Section 3, we suggest the following approach for selecting an optimal regularization parameter for the MEM functional: (i) perform an extensive, but coarse-grained search over values of the regularization parameter to determine the region in the L-curve plot which excludes the unstable behavior. (ii) Repeat the regularization within the selected region with finer resolution in $\lambda$ and determine the optimal regularization parameter, $\lambda_{opt}$ according to the maximum curvature criterion. The $\lambda_{opt}$ is selected and marked as asterisk in Fig. 1. This strategy is tested in Section 4 and then used in analyzing experimental results in this report.

### 2.2. MEM functional that includes both intra- and intermolecular interactions

The results from pulsed ESR will in general include both intra- and intermolecular dipolar interactions. The former from the spin label pairs and the latter from spin labels on other nearby molecules. It is best to work with low enough concentrations to minimize the intermolecular contributions. There are currently standard approaches for subtracting or removing such contributions prior to the analysis of the intramolecular component of the time-domain signal [7,27,29]. Because the exact form of the baseline is not known, a judicious removal is prone to a potentially biased judgment. However, both $P(r)$ and the "baseline," can be fit simultaneously where the latter may be taken as one of several convenient functional forms and parametrized as needed. This offers the prospect of a more reliable estimation of the intra molecular component to the signal. We consider two examples of a simple form for the baseline. It is straightforward to include them directly into the MEM functional of Eq. (5). Let us refer to the raw experimental data as $S$. Then in DQC experiments on dilute systems one usually approximates $S$ as a sum of intramolecular ($S_0$) and intermolecular terms ($B$) asz

$$S_0 = S - B, \tag{6a}$$

which was found to be satisfactory in empirical studies [7].

A multiplicative factor is usually used for DEER, as given in Eq. (6b) [27].

$$S_0 = \frac{S}{B} - 1. \tag{6b}$$

(Note another baseline form has recently been recommended [3]. It can also be implemented in our MEM approach, but we have used Eq. (6b) in the present study.)

In either case of Eqs. (6a) or (6b), $B$ can be approximated as a low-degree polynomial [27]:

$$B = \sum_{i=0}^{\text{integer}} a_i t^i. \tag{6c}$$

Then with the replacement of $S_0$ by $S$ into Eq. (5), where S is given by either Eqs. (6a) or (6b), we can simultaneously determine the distance distribution and the unknown intermolecular contribution (i.e., baseline). That is we utilize the iterative CG algorithm to determine both $P(r)$ and the coefficients $a_i$. More refined formulations of the intermolecular contribution can be incorporated into the analysis in a similar fashion, but we confine ourselves in this study to the simple, but commonly used, approaches given by Eqs. (6a)–(6c).

In the model (and experimental) cases we studied, we have found that this method works very well. In fact it does not even require good initial seed parameters for the baseline function in order to reproduce $B$ accurately. However, for high concentrations, for which the baseline dominates over the intramolecular dipolar signal, a reasonable seed for the baseline function is recommended to avoid undesired minima and expedite the convergence. Details are given in the following sections.

## 3. Results for the model data

Three model distributions[1] were first tested under the condition of high SNR (~500) using MEM, and then for the case of low SNR ~30, which we previously found challenging for TIKR [2]. We found that a good approximation for $P(r)$ to initialize the MEM regularization is necessary. Given that both the residual and entropy terms are non-linear and complicated one may expect there are many local minima. Thus a good initial estimate is needed. That minimizing the MEM functional can be challenging is well known [12]. More details about the minimizations of MEM and TIKR will be given in Section 5. Here we note that the $P_o(r)$ in Eq. (5) and the initial or seed value of $P(r)$ for the algorithm is always taken as the optimum distribution obtained by TIKR. It is necessary to first modify the negative and zero values of $P(r)$ obtained by TIKR to be positive and nearly zero in order to insert the TIKR $P(r)$ into the argument of the natural logarithm in Eq. (5). (In cases of high SNR, $P_o(r)$ can be set to be unity, but using a good estimate for $P_o(r)$ significantly reduces the computation time.)

[1] We give the parameters used for the model distributions. *Bimodal distribution*: a sum of two equally weighted Gaussian peaks at 3.6 and 3.9 nm with respective standard deviations of 0.1 and 0.077 nm. *Box-like distribution*: the two maxima are located at 3.2 and 4.3 nm and the curve between the peaks decreases as $r^{-1}$ from each maximum such that it is symmetric about the center at 3.75 nm. *Trimodal distribution*: a sum of three equally weighted Gaussian peaks at 3, 4, and 5 nm with respective standard deviations of 0.4, 0.5, and 0.2 nm.

Fig. 2 shows the $P(r)$ obtained by MEM for the (A) bimodal and (B) box-like distributions for a SNR $\sim$ 500. The original distributions, estimates by MEM, and the estimates by TIKR are shown. Both TIKR and MEM recover very good estimates of $P(r)$, especially for the dominant portion of the distribution. However, the oscillations about zero in the peripheral regions of $P(r)$ seen in the TIKR result are significantly reduced and non-negative in the MEM result. For the box-like distribution in Fig. 2B, we do find some high-frequency noise superposed on the main portion of the MEM distribution. These results are consistent with a previous study, where it was found that, where $P$ is small MEM was often superior to Tikhonov regularization, but was comparable to or slightly worse than TIKR where $P$ was not small [15].

The optimal regularization parameters for both TIKR and MEM regularizations were determined by the L-curve criterion [36,37] (cf. Fig. 1 and above discussion). Our model results appear to indicate the appropriateness of the L-curve criterion for determining an optimal regularization parameter for the MEM regularization of dipolar signals. However, this might not necessarily be the case for other kernel functions.

Fig. 3 shows the results for the (A) bimodal and (B) box-like distributions for SNR $\sim$30. We found that the MEM result shows some improvements over the TIKR result. It suppresses the undesirable oscillations in the wings of the distribution. Also the bimodal distribution (in Fig. 3A) is more apparent in the MEM result. A similar observation may be made for the box-like model (Fig. 3B) but at the expense of more noise in the main region of the distribution. The sharp edges of the box are much better represented by MEM, and the overall shape within the box is better if one smoothes over the noisier central region. Comparing the results in

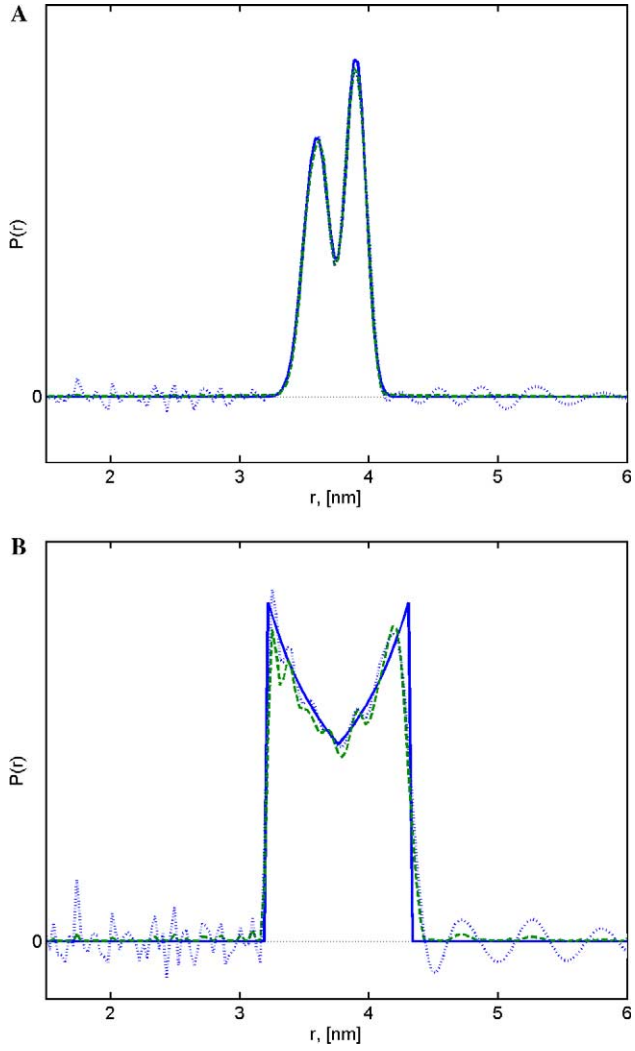


Fig. 2. The model study results of the (A) bimodal and (B) box-like distributions for SNR $\sim$ 500. The original, MEM regularized, and TIKR regularized solutions are plotted by solid, dashed, and dotted lines, respectively.

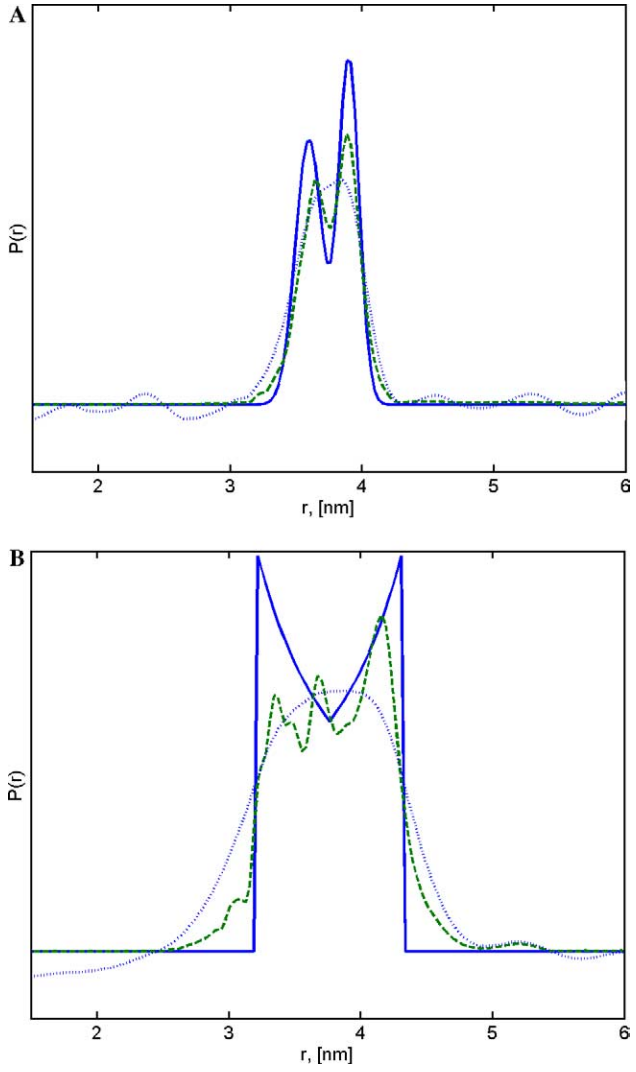


Fig. 3. The model study results of the (A) bimodal and (B) box-like distributions for SNR $\sim$ 30 in the simulated dipolar signals. The original, MEM regularized, and TIKR regularized solutions are plotted by solid, dashed, and dotted lines, respectively.

Fig. 3B with Fig. 2B, it appears that the artifacts in the main distribution are primarily caused by the lower SNR. These results indicate the limitations of both regularization methods for recovering distributions from pulsed ESR experiments. However, the average distances determined by TIKR and MEM are both close to that of the original model. One of the most important observations from these model tests is that both TIKR and MEM are able to provide an average distance for the distribution with satisfactory accuracy under substantial noise conditions, i.e., SNR $\leqslant$30, which is typical in DQC and DEER experiments. Our results also indicate that MEM must be coupled with TIKR. The MEM can be used to improve on the TIKR result only if the TIKR result is used as the seed for the MEM procedure.

Fig. 4 shows the results for the broad tri-modal distribution with values of SNR $\sim$ (A) 500, (B) 30, and (C) 10 using both TIKR and MEM. The distribution is well recovered for a SNR $\sim$ 500 using both methods (Fig. 4A). For a SNR $\sim$ 30 (Fig. 4B) the MEM and TIKR results are very similar and are in good, but not precise, agreement with the original distribution. However, when the SNR was decreased to 10, the MEM failed to provide a good estimate even though the TIKR result, which is not so bad, was seeded as the initial distribution. This observation, that the minimization converges to a poorer solution for this poor SNR, is referred to as semiconvergence [12]. This means that the solution vector initially approaches a regularized solution vector and then, in later stages of the iteration, converges to a undesired vector. Some undesirable oscillations appear in the main distribution of the MEM result (cf. Fig. 4C) for this low SNR case.

Next we used MEM to extract the distance distribution from a simulated dipolar signal, which includes both intra- and intermolecular contributions (cf. Eqs. (6a)–(6c)). We added a second-degree polynomial (cf. Eq. (6c))to the simulated data and found that the MEM method is able to provide good estimates for all the distributions, including the bimodal, boxlike, and

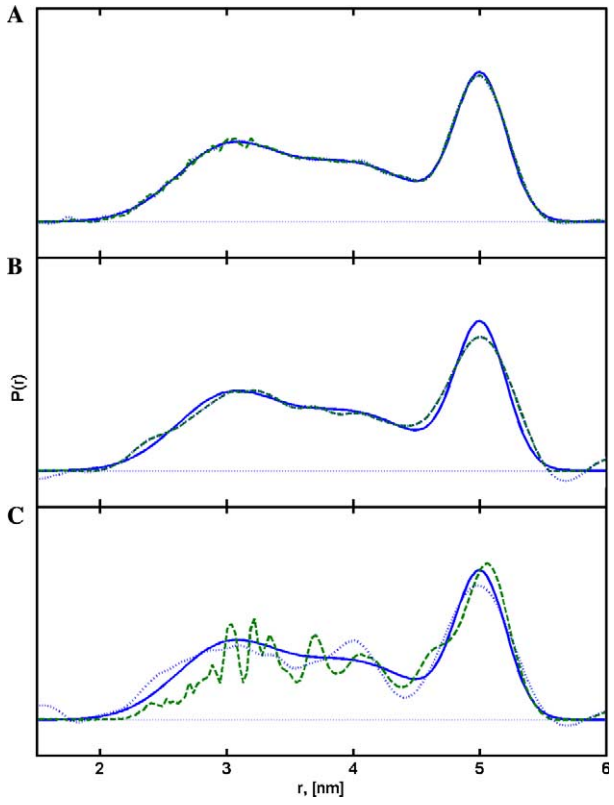


Fig. 4. The model study results for a broad tri-modal distribution for SNR $\sim$ (A) 500, (B) 30, and (C) 10 using TIKR and MEM. The original distribution, MEM, and TIKR regularized solutions are plotted by solid, dashed, and dotted lines, respectively.

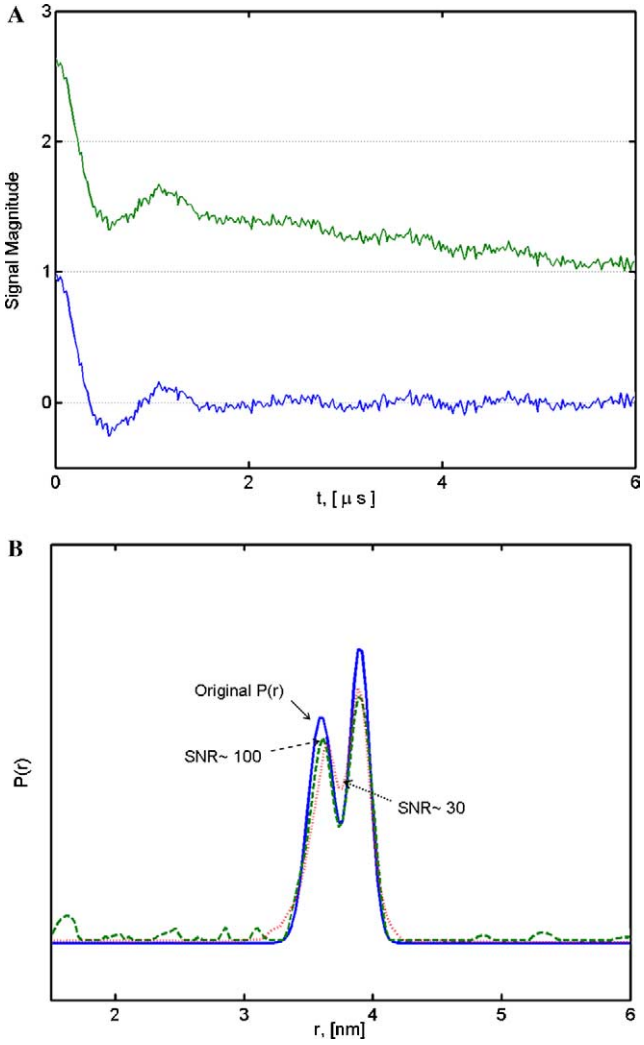


Fig. 5. (A) The time evolution data of the bimodal distribution for SNR $\sim$ 30 before (lower trace) and after (upper trace) a second-order polynomial baseline, $(a_0, a_1, a_2) = (1.65, -0.13, 0.004)$, is added using Eq. (6a). (B) The distributions recovered from the baseline-added time evolution data by minimizing Eq. (5). The original and the estimates for SNR $\sim$ 100 and 30 are, respectively, plotted by solid, dashed, and dotted lines. The added baseline and the model distribution are simultaneously obtained using the proposed method.

trimodal distributions. Fig. 5 shows one of the cases studied, which assumes the condition of Eq. (6a). Fig. 5A displays the simulated dipolar signals with SNR of 30 before (lower trace; i.e., the intramolecular signal) and after a second-degree polynomial baseline is added (upper trace). The original $P(r)$ for the intramolecular spin pairs is shown in Fig. 5B as well as the MEM estimates obtained from the baseline-added signals for the cases of SNR $\sim$ 100 and 30. The recovered distribution for a SNR $\sim$ 100 is very good. With the initial values seeded with the TIKR results, the program returns good approximations for both the inserted baseline and the original distribution. The recovered distribution for the case of a SNR $\sim$30 is quite good. It is almost the same as the case obtained in the absence of any baseline, (cf. the dashed line in Fig. 3A). This indicates that the proposed method, i.e., to minimize Eq. (5) including the presence of a baseline (e.g. Eqs. (6a)–(6c)), is a viable approach.

The computation times required for the TIKR and MEM calculations are quite different. Typically, the determination of a distribution, whose solution vector length is 150 and with SNR $\sim$ 30 in the time-domain signal, requires more than an hour using MEM and less than a minute using TIKR to complete the calculations for 200 $\lambda$ values. It requires at least double this time for the MEM if fitting the baseline is included in the calculation. Actual computation time varies with the baseline function. The computations were performed on a Linux-operating PC (CPU: Intel Pentium IV 2.0 GHz; RAM: 2 GB).

## 4. Results for the experimental data

The experimental data (i.e., the manually pre-processed baseline-subtracted data) that we previously analyzed using TIKR was used without any modifications in this report on the MEM analysis. Hereafter, we call the experimental signals before and after baseline subtraction as "*raw experimental signals*" and "*intramolecular dipolar signals*," respectively. The latter is the "baseline-subtracted" signal obtained using Eqs. (6a)–(6c) to manually remove the intermolecular component. Two proteins were studied, T4-lysozyme (T4L) and iso-cytochrome *c* (iso-1-cyt c). They were performed using the DQC and DEER techniques, respectively. The residues of interest were substituted by cysteines and labeled with the nitroxide spin label, (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl) methanethiosulfonate (MTSSL). Further details about the materials used and the six-pulse DQC and four-pulse DEER experiments at 17.3 GHz are given elsewhere [2,7].

The distance distributions of spin-pairs recovered by TIKR and MEM from the "intramolecular DQC dipolar" signals of the T4L 65/135 and the T4L 61/80 mutants are shown in Figs. 6 and 7, respectively, by

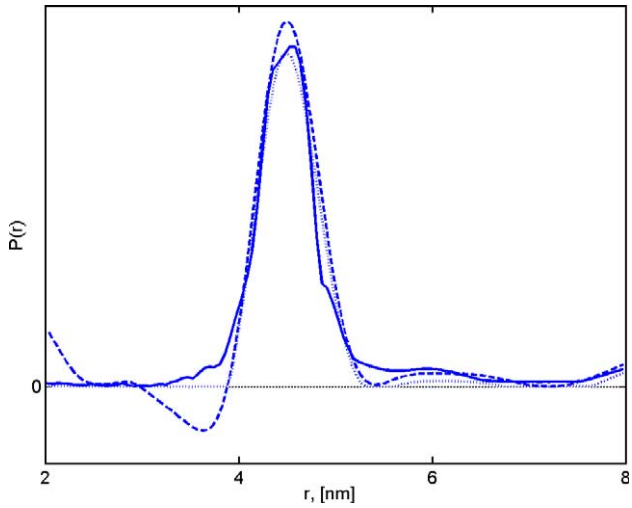


Fig. 6. The regularized distributions recovered from the dipolar signals of the T4L 65/135 mutant obtained from six-pulse DQC-ESR at 17.3 GHz. The distributions obtained using TIKR (Eqs. (2a) and (2b)) and MEM (Eq. (5)) are plotted by dashed and solid lines, respectively. The distribution obtained from minimizing Eq. (5) with consideration of baseline presence (i.e., Eq. (6a)) is plotted by the dotted line. The coefficients for the extracted baseline are $(a_0, a_1) = (-0.01, -0.001)$; this is to be compared to an initial signal [i.e., $S(t = 0)$] magnitude of 0.36.

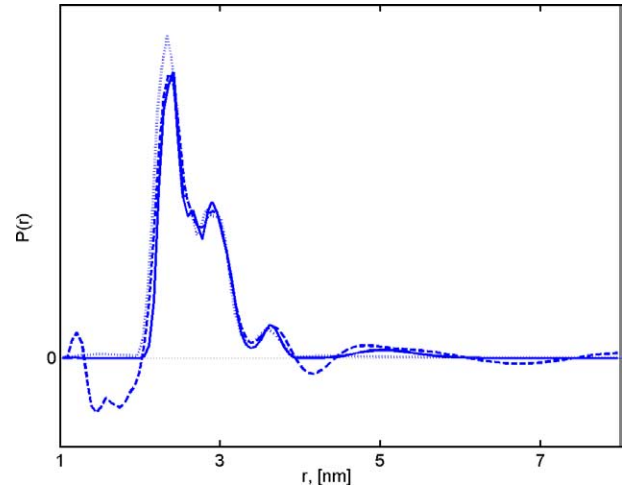


Fig. 7. The regularized distributions recovered from the dipolar signals of the T4L 61/80 mutant obtained from six-pulse DQC-ESR at 17.3 GHz. The distributions obtained from TIKR and MEM (Eq. (5)) are plotted by dashed and solid lines, respectively. The distribution obtained from minimizing the MEM functional that includes Eq. (6a) is plotted by the dotted line. The coefficients for the extracted baseline (cf. Eq. (6a)) are $(a_0, a_1) = (0.19, -0.09)$; this is to be compared to an initial signal [i.e., $S(t = 0)$] magnitude of 1.55.

dashed and solid lines. The $P(r)$ obtained directly from the "raw experimental" signal by minimizing the MEM functional Eq. (5) that includes Eq. (6a) is plotted as a dotted line. In Fig. 7 the main part of the $P(r)$ is consistently recognized as bimodal from the three results. The non-negative $P(r)$'s that are obtained by MEM are more

appropriate than those obtained previously by TIKR [2]. The main portion of the distributions obtained from the MEM and TIKR regularizations are otherwise nearly the same. This supports our previous contention that the oscillations in the peripheral region of the TIKR distribution do not significantly affect the main part of the distribution, so they can be disregarded in the TIKR results [2]. The MEM results from the raw experimental signals are seen to suppress the oscillations in the wings more effectively than those from the intramolecular dipolar signals. The parameters from fitting the baseline function, Eq. (6a), are given in the captions of Figs. 6 and 7 and are found to be very small, as expected for low spin concentrations. In fact, they are of the order of estimated experimental uncertainties.

The results from the DEER signals for the $P(r)$ of the iso-1-cytochrome $c$ doubly labeled at S47C and K79C are shown in Fig. 8. They provide insight into the conformations of possible folding intermediates. Fig. 8 includes results for three different conformational states of the iso-1-cyt c: (a) the completely folded state, (b) partially unfolded state, and (c) a more unfolded state [2].

Overall, the $P(r)$ recovered by MEM are consistent with our previous results obtained by TIKR (cf. Fig. 8). Before discussing the $P(r)$ recovered from the raw experimental data, we first focus on the minor differences between the results of the two regularization methods (MEM vs. TIKR) applied to the intramolecular dipolar data and summarize them as follows: (i) the undesirable oscillations around zero in the wings

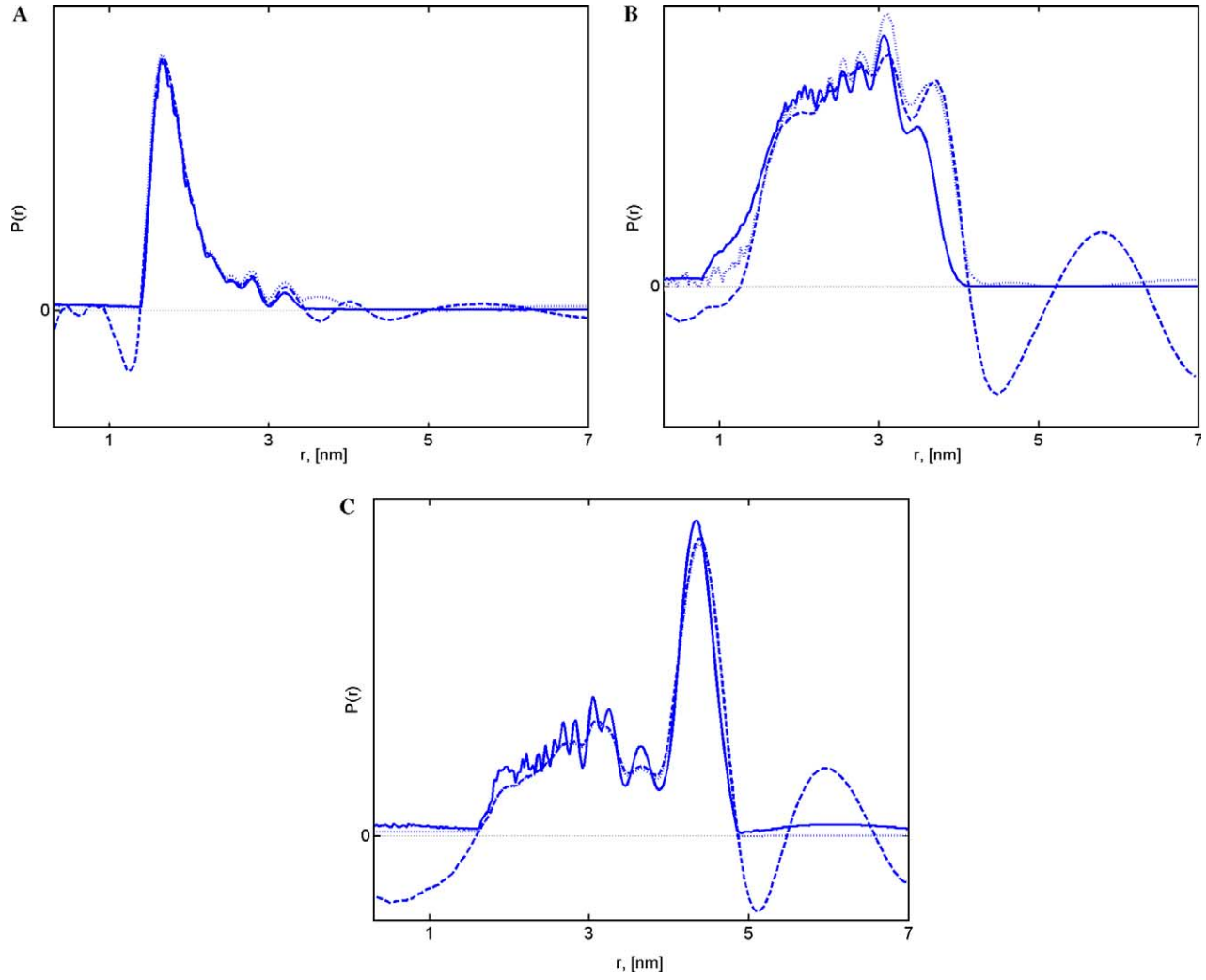


Fig. 8. The pair distance distributions for three different folding states recovered from the dipolar signals of the iso-1-cyt c S47C/K79C obtained from four-pulse DEER experiments at 17.3 GHz. The intermediate states of the protein are maintained by guanidinium hydrochloride concentrations, [GdnHCL] = (A) 0, (B) 0.7, and (C) 1.5 M, respectively. The iso-1-cyt c is fully folded and unfolded at 0 and 2 M GdnHCL. The distributions obtained from TIKR and MEM (Eq. (5)) are plotted by dashed and solid lines, respectively. The distribution obtained from minimizing the MEM functional that includes Eq. (6b) is plotted by the dotted line. The oscillations around zero, which appear in the TIKR results, are all corrected in the MEM results. The coefficients for the respective extracted baselines, from (A to C), are (cf. Eq. (6b)) $(a_0, a_1, a_2) = (2.29, -0.25, -0.06)$, $(2.52, -0.36, 0.03)$, and $(2.43, -0.51, 0.05)$. The respective magnitudes of the initial signal [i.e., $S(t = 0)$] are 2.41, 2.61, and 2.50, respectively.

of $P(r)$ are completely suppressed in the MEM results, as we expected; this is very desirable here, since they are very substantial in the TIKR result. (ii) There are some high-frequency oscillations that appear in the main distribution of the MEM results (the solid lines in Figs. 8B and C), which are not seen in the TIKR analysis. According to the model test results for low SNR (cf. Figs. 3B and 4C), we attribute such oscillation as mainly due to the substantial noise level in the data and partially due to the inaccuracy of the baseline subtraction (discussed later). This may indicate a limitation of MEM for dealing with cases of more severe noise. In the results for T4L in Figs. 6 and 7 the SNR was higher than those for Fig. 8 (cf. Figs. 7–11 in [2] for the time evolution signals). This may be why there are no such high-frequency oscillations in Figs. 6 and 7. To summarize, we suggest that the superposition of the MEM results with those from TIKR, as in Fig. 8, provides useful insight into the $P(r)$ where one should emphasize the respective strengths (and de-emphasize the weaknesses) of the two methods.

The $P(r)$ recovered from the raw experimental data by minimizing the MEM functional Eq. (5) with Eqs. (6a)–(6c) are shown by dotted lines in Fig. 8. In this study we found that the baseline contributed from the intermolecular interactions is very significant because of the high protein concentration. The concentration of cytochrome $c$ was five times that of T4 lysozyme in the DQC experiment. This substantially increases the signal intensity from the intermolecular interactions, and also reduces the intramolecular dipolar oscillation, thereby raising a large challenge for the CG algorithm to search for the global minimum. When we initially used a zero vector as the initial condition for the baseline we obtained noise-like distributions for $P(r)$. This indicated that the seed is far from the global minimum. A physically meaningful distribution (i.e., not noise-like) was not obtained until we used reasonable seed values for the baseline, which were obtained by fitting a second-degree polynomial to the last three quarters of the time-domain raw data [27].

Overall, the averages of the distributions obtained by each of the three methods in Fig. 8 agree within estimated error (ca. $\pm 0.2$ nm). The high-frequency oscillation in the main distribution observed in the MEM result from the intramolecular dipolar data is absent in the MEM result from the raw experimental data for the cases of Figs. 8A and C. These latter results (plotted as dotted lines) converge more nearly to the initial distributions, i.e., the TIKR results in the main portion of $P(r)$ with more nearly zero values for $P(r)$ in the wings of the distribution. To check this convergence, we repeated the fitting to the raw experimental data using the MEM results (i.e., the solid lines) as the seed distributions, and we found that they still converge more closely to the TIKR distributions (i.e., the dashed lines) in the central region of $P(r)$. It is likely that oscillations in the main part of $P(r)$ are suppressed because the intermolecular interaction is more properly fit using the raw data and Eq. (5) with Eqs. (6a)–(6c). As for Fig. 8B, the artifactual oscillations remain in the main part of the distribution, obtained from the raw data and MEM, whether we seeded the calculations with either of the two other distributions. Perhaps a higher order polynomial approximation (cf. Eq. (6c))for the baseline is required for this case. Overall these results indicate that incorporating an approximate function for the baseline into the MEM functional is a valuable asset to the analysis.

## 5. Discussion

### 5.1. MEM used as a complement to Tikhonov regularization

The maximum entropy method of regularization has been rigorously examined for its stability as well as its convergence [25,26], and was previously mainly applied to analyses of sums of exponentials [39]. The previous study which most relates to our present study is that of Amato and Hughes [15]. They demonstrated the convergence of the solution of the minimization problem (i.e., Eq. (5)) by showing the procedure leads to a valid regularization method. In addition, they compare the performance of MEM vs. TIKR by several numerical experiments. Their results indicate that MEM is inferior to TIKR for a one-dimensional Fredholm equation of the first kind when $P$ (cf. Eq. (5)) is not close to zero, but better when $P$ is close to zero. Another study, however, concluded that there are no clear criteria when to prefer one over the other [33]. Performance may vary with the kernel functions. Our report may be considered as proposing a regularization scheme for the determination of $P(r)$ which employs both TIKR and MEM. This scheme benefits from the virtues of TIKR, which provides a much faster and direct regularization calculation (cf. Eq. (2b)) as compared to MEM which requires an iterative CG algorithm with good seed values. On the other hand MEM constrains the regularized solution to be non-negative, it suppresses undesirable noise in the wings of $P(r)$, and the intermolecular "baseline" can be fit without further complicating the algorithm.

Our method using MEM as a complement to TIKR is developed in this context. As we have discussed before [2], the most important advantage of using TIKR with the regularization parameter determined by the L-curve criterion is that it provides a fast regularization algorithm and a graphical presentation (i.e., L-curve) for examining how much the penalty term is weighted in the calculation. Because the TIKR regularized solution is in a quadratic form (cf. Eqs. (1), and (2a), (2b)) and readily solved by SVD, it is uniquely determined for

each given regularization parameter. The regularized solution of TIKR is not affected by other minimization parameters, such as the convergence and iteration step parameters required for performing CG algorithm in MEM regularization or other algorithms.

In our model tests MEM is found more useful for recovering original distributions for the bimodal and box-like models under the condition of SNR $\sim 30$ than TIKR. However, this was the case only if the TIKR result was used as the initial distribution for MEM regularization. Our regularization procedure is summarized in the following.

(1) Obtain an approximation for the intermolecular contribution to the dipolar signal using manual baseline subtraction methods [7,27]. Then perform TIKR to obtain a solution for $P(r)$ with an optimal regularization parameter determined by the L-curve criterion.

(2) Use the $P(r)$ obtained in step 1 as the initial condition and the prior distribution, $P_o$ for MEM regularization. Note that we first set the negative values of $P(r)$ in the TIKR result to be positive but close to zero before substituting into the MEM regularization, since negative values are not allowed as arguments of the logarithm in Eq. (5). One has two choices in employing Eq. (5). Either manually preprocess the raw experimental signal to remove baseline and then apply Eq. (5), or else use Eq. (5) with Eqs. (6a)–(6c) to allow the baseline to be fit as part of the MEM regularization. In the latter case it is helpful to use the manual estimate of the baseline to seed the MEM. Our results indicate a preference for the latter method.

(3) The optimal regularization parameter $\lambda_{\text{opt}}$ for the MEM functional (Eq. (5)) may also be estimated by application of the L-curve criterion, i.e., it is taken as the value for $\lambda$ corresponding to the point of the maximum curvature of the L-curve whose $x$- and $y$-axis are the residual norm and the modified entropy term, respectively. Care must be taken to avoid unstable regions in the L-curve plot. We must caution the reader that the mathematical validity for the L-curve criterion in MEM has never been rigorously studied. However, we have found from our model studies that, while the L-curve plot no longer has a distinct L-shaped elbow in most cases, the maximum curvature criterion is still robust and useful for selecting an optimal regularization parameter. This is different from a report that an L-shaped curve may not appear in some numerical studies [40]. The study of other methods for selecting the $\lambda_{\text{opt}}$ in the MEM functional is very limited. Generalized cross validation (GCV; [41]) has been employed [15,42], as has manual assignment [25,26].

MEM has recently been used to extract the distribution of first order rate constants from fluorescence resonance energy transfer (FRET) experiments from which distances between donor and acceptor labeled residues in proteins were estimated [43,44]. That MEM application is different from what we have discussed in this work in the following respects: (i) the kernel functions for the FRET (a sum of exponential decays) and for pulsed ESR (Pake doublets) are different, which likely implies different ill-posed behavior for the respective problems; (ii) the ill-posed problem of FRET was solved by considering two solutions for the distance distributions, i.e., one narrow $P(r)$ obtained from non-linear least squares fits subject to a non-negativity constraint (cf. the LSQNONNEG subroutine in Matlab), and one broad $P(r)$ determined by minimizing a MEM functional; (iii) the resolution for the estimate distribution $P(r)$ (i.e., dimension of the solution vector) shown in [43,44] is much reduced (by a factor of about 10–100) than we obtain (cf. Section 4). (Note that the instability of an ill-posed problem can be subdued by reducing the size of the regularized solution vector, which has the same effect as preventing small singular values from contributing to the solution). It is not possible to compare further the two MEM regularizations since several key issues were neither shown nor discussed in those applications [43,44], (e.g., the MEM functional, the method used for the functional minimization, the initial seed for the minimization algorithm, and the definition of the L-curve axes), and, most important, they were used to solve different problems. However, the maximum entropy application to the analysis of FRET supports our work in the following two aspects: (i) the L-curve criterion seems to be a useful criterion in finding a balance between the least squares fits and the entropy; (ii) the use of entropy provides a natural and unbiased way of constraining the regularization solution to be non-negative.

### 5.2. Implications for other approaches to the regularization

In addition to the L-curve method used in [2], the self-consistent (SC) method has been used for selecting an optimal regularization parameter in TIKR in the determination of spin-pair distance distributions from pulsed ESR experiments [3]. The SC-method [13] relies on an iterative approach to determine $\lambda_{\text{opt}}$ in a self-consistent manner; whereas, the L-curve method [37] finds $\lambda_{\text{opt}}$ by making use of a posteriori regularization information, i.e., the residual and solution norms for all the tested regularization parameters plotted on the L-curve. Using the models of bimodal and box-like distributions, we found it is not possible to draw a clear conclusion as to which is better for the tested noise conditions, SNR $\sim 500$, 100, and 30. However, we did find that the SC-method with non-negativity constraint, (let us call this the modified SC-method; [13]) was not as successful as the L-curve criterion, especially for the broad trimodal distribution with SNR $\sim 30$ as noted in [2] (cf. Fig. 4B for the L-curve result and Fig. 4b in [3] for the

modified SC result). The functional to be minimized in the modified SC-method is a combination of the TIKR functional and an additional constraint term to ensure a non-negative solution. This may have interfered with the objective of TIKR of introducing the penalty term by means of a regularization parameter to stabilize the solution, because the contribution from the extra term to the functional is not controlled by any regularization parameter. As for the minimization algorithms, the estimated $P(r)$ from the modified SC-method is obtained by a functional minimization procedure [13,14] where the SC-method result is used as the initial condition. In other words, the idea of minimizing the MEM functional using TIKR result as initial condition that we proposed here is similar to that aspect of the modified SC-method, but the two functionals are different. Both methods require good seeds to avoid undesired local minima. From these considerations, one might prefer to use MEM when a non-negative structure on the regularized solution is required. A method that performs MEM regularization using the result of the SC-method as the initial condition might be useful.

## 6. Summary

The maximum entropy regularization method for the determination of distance distributions of spin-pairs directly from pulse ESR dipolar signals has been introduced as a complement to Tikhonov regularization in this study. From the model distributions investigated, it is clear that the Tikhonov regularization must first be used. The MEM regularization, which implicitly guarantees a non-negative solution, can then be used to refine the TIKR result that implicitly guarantees a non-negative solution. Its main improvement to the TIKR distribution is the removal of negative and positive oscillations in the wings of the distance distribution. For SNR less than or comparable to 30, MEM regularization is often useful to refine the TIKR result to provide a better resolution. In addition, MEM regularization, which is implemented iteratively, can be readily modified to incorporate effects of intermolecular dipolar interactions that are embedded in the raw experimental data as a "baseline." The proposed method, tested on model and experimental data, was found useful for determining both the baseline and pair distance distribution functions simultaneously. This report, for the first time, provides an iterative regularization method for a direct determination of the two functions. The method for determining an optimal regularization parameter for the MEM functional is crucial to the MEM regularization. This topic has never been extensively investigated using existing methods, such as L-curve, GVC, SC, etc. The reason probably is the TIKR regularization is sufficient for solving ill-posed problems in general applications. In this report we have studied this topic and derived an empirical rule, which is based on the L-curve criterion, for selecting an optimal regularization parameter of the MEM functional according to the numerical experiments performed.

It is suggested that in the post-proteomic era, which focuses on interactions between various structural domains, one would employ TIKR to quickly obtain an estimate for the pair distance distribution for all the bilabeled protein mutants and then refine those distributions, which are of most interest, using MEM with a greater expense in computational time.

## Acknowledgments

This work was supported by grants from NIH/NCRR, NIH/NIBIB, and NSF/Chemistry. It made use of the computer facility of the Cornell Center for Materials Research (CCMR).

## References

[1] C.W. Groetsch, The theory of Tikhonov regularization for Fredholm equations of the first kind, Pitman Advanced Pub. Program, Boston, 1984.
[2] Y.W. Chiang, P.P. Borbat, J.H. Freed, The determination of pair distance distributions by pulsed ESR using Tikhonov regularization, J. Mag. Reson. 172 (2005) 279–295.
[3] 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.
[4] M.K. Bowman, A.G. Maryasov, N. Kim, V.J. DeRose, Visualization of distance distribution from pulsed double electron–electron resonance data, Appl. Magn. Reson. 26 (2004) 23–39.
[5] L.J. Berliner, G.R. Eaton, S.S. Eaton, Biological magnetic resonance: distance measurements in biological systems by EPR, Kluwer Academic/Plenum Publishers, New York, 2000.
[6] P.P. Borbat, J.H. Davis, S.E. Butcher, J.H. Freed, Measurement of large distances in biomolecules using double-quantum filtered refocused electron spin-echoes, J. Am. Chem. Soc. 126 (2004) 7746–7747.
[7] P.P. Borbat, H.S. Mchaourab, J.H. Freed, Protein structure determination using long-distance constraints from double-quantum coherence ESR: study of T4 lysozyme, J. Am. Chem. Soc. 124 (2002) 5304–5314.
[8] B.G. Dzikovski, P.P. Borbat, J.H. Freed, Spin-labeled gramicidin A: channel formation and dissociation, Biophys. J. 87 (2004) 3504–3517.
[9] G. Jeschke, C. Wegener, M. Nietschke, H. Jung, H.-J. Steinhoff, Interresidual distance determination by four-pulse double electron–electron resonance in an integral membrane protein: the Na+/proline transporter PutP of *Escherichia coli*, Biophys. J. 86 (2004) 2551–2557.
[10] V.M. Grigoryants, K.A. DeWeerd, C.P. Scholes, Method of rapid mix EPR applied to the folding of bi-spin-labeled protein as a probe for the dynamic onset of interaction between sequentially distant side chains, J. Phys. Chem. B. 108 (2004) 9463–9468.
[11] K. DeWeerd, V. Grigoryants, Y. Sun, J.S. Fetrow, C.P. Scholes, EPR-detected folding kinetics of externally located cysteine-

directed spin-labeled mutants of iso-1-cytochrome *c*, Biochemistry 40 (2001) 15846–15855.
[12] P.C. Hansen, Rank-deficient and discrete ill-posed problems : numerical aspects of linear inversion, SIAM, Philadelphia, 1997.
[13] J. Weese, A reliable and fast method for the solution of Fredholm integral-equations of the 1st kind based on Tikhonov regularization, Comput. Phys. Commun. 69 (1992) 99–111.
[14] T. Roths, M. Marth, J. Weese, J. Honerkamp, A generalized regularization method for nonlinear ill-posed problems enhanced for nonlinear regularization terms, Comput. Phys. Commun. 139 (2001) 279–296.
[15] U. Amato, W. Hughes, Maximum-entropy regularization of fredholm integral-equations of the 1st kind, Inverse Probl. 7 (1991) 793–808.
[16] J. Skilling, R.K. Bryan, Maximum-entropy image-reconstruction—general algorithm, Monthly Notices R. Astron. Soc. 211 (1984) 111–124.
[17] A.K. Livesey, J.C. Brochon, Analyzing the distribution of decay constants in pulse-fluorometry using the maximum-entropy method, Biophys. J. 52 (1987) 693–706.
[18] E.T. Jaynes, Information theory and statistical mechanics, Phys. Rev. 106 (1957) 620–630.
[19] E.T. Jaynes, Information theory and statistical mechanics. 2, Phys. Rev. 108 (1957) 171–190.
[20] B.R. Frieden, Restoring with maximum likelihood and maximum entropy, J. Opt. Soc. Am. 62 (1972) 511–518.
[21] A.K. Livesey, J. Skilling, Maximum-entropy theory, Acta Crystallogr. A 41 (1985) 113–122.
[22] S.F. Gull, G.J. Daniell, Image-reconstruction from incomplete and noisy data, Nature 272 (1978) 686–690.
[23] E.D. Laue, J. Skilling, J. Staunton, S. Sibisi, R.G. Brereton, Maximum-entropy method in nuclear magnetic-resonance spectroscopy, J. Mag. Reson. 62 (1985) 437–452.
[24] E.D. Laue, J. Skilling, J. Staunton, Maximum-entropy reconstruction of spectra containing antiphase peaks, J. Mag. Reson. 63 (1985) 418–424.
[25] P.P.B. Eggermont, Maximum-entropy regularization for fredholm integral-equations of the 1st kind, SIAM J. Math. Anal. 24 (1993) 1557–1576.
[26] H.W. Engl, G. Landl, Convergence-rates for maximum-entropy regularization, SIAM J. Numer. Anal. 30 (1993) 1509–1536.
[27] G. Jeschke, A. Koch, U. Jonas, A. Godt, Direct conversion of EPR dipolar time evolution data to distance distributions, J. Magn. Reson. 155 (2002) 72–82.
[28] A.D. Milov, Y.D. Tsvetkov, F. Formaggio, S. Oancea, C. Toniolo, J. Raap, Solvent effect on the distance distribution between spin labels in aggregated spin labeled trichogin GA IV dimer peptides as studied by pulsed electron–electron double resonance, Phys. Chem. Chem. Phys. 6 (2004) 3596–3603.
[29] A.D. Milov, A.G. Maryasov, Y.D. Tsvetkov, Pulsed electron double resonance (PELDOR) and its applications in free-radicals research, Appl. Magn. Reson. 15 (1998) 107–143.
[30] M. Klaus, R.T. Smith, A hilbert-space approach to maximum-entropy reconstruction, Math. Methods Appl. Sci. 10 (1988) 397–406.
[31] G. Landl, R.S. Anderssen, Non-negative differentially constrained entropy-like regularization, Inverse Probl. 12 (1996) 35–53.
[32] T. Elfving, An algorithm for maximum entropy image reconstruction from noisy data, Math. Comput. Model. 12 (1989) 729–745.
[33] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer Academic Publishers, Dordrecht; Boston, 1996.
[34] G. Landl, T. Langthaler, H.W. Englt, H.F. Kauffmann, Distribution of event times in time-resolved fluorescence: the exponential series approach–algorithm, regularization, analysis, J. Computat. Phys. 95 (1991) 1–28.
[35] R. Fletcher, Practical methods of optimization, Wiley, Chichester; New York, 1987.
[36] P.C. Hansen, Regularization tools Version 3.0 for Matlab 5.2, Numer. Algorithms 20 (1999) 195–196.
[37] P.C. Hansen, D.P. Oleary, The use of the l-curve in the regularization of discrete Ill-posed problems, SIAM J. Sci. Comput. 14 (1993) 1487–1503.
[38] P.C. Hansen, in: P. Johnston (Ed.), Computational Inverse Problems in Electrocardiology, WIT Press, 2000, pp. 119–142.
[39] A.A. Istratov, O.F. Vyvenko, Exponential analysis in physical phenomena, Rev. Sci. Instrum. 70 (1999) 1233–1257.
[40] T. Reginska, A regularization parameter in discrete ill-posed problems, SIAM J. Sci. Comput. 17 (1996) 740–749.
[41] G. Wahba, Practical approximate solutions to linear operator equations when data are noisy, SIAM J. Numer. Anal. 14 (1977) 651–667.
[42] A. Isayama, N. Iwama, T. Showa, Y. Hosoda, N. Isei, S. Ishida, M. Sato, Maximum entropy estimation of electron cyclotron emission spectra from incomplete interferograms in ELMy H-mode tokamak experiment, Jpn. J. Appl. Phys. Part 1 42 (2003) 5787–5796.
[43] J.G. Lyubovitsky, H.B. Gray, J.R. Winkler, Mapping the cytochrome *c* folding landscape, J. Am. Chem. Soc. 124 (2002) 5481–5485.
[44] E.V. Pletneva, H.B. Gray, J.R. Winkler, Many faces of the unfolded state: conformational heterogeneity in denatured yeast cytochrome *c*, J. Mol. Biol. 345 (2005) 855–867.