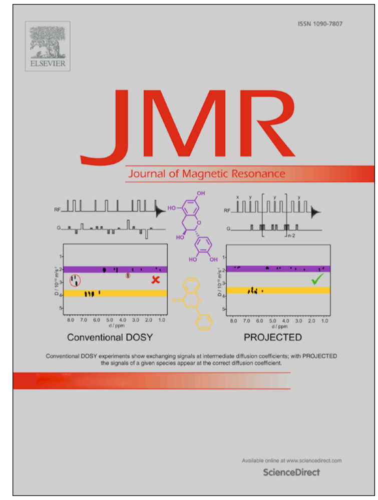
Accepted Manuscript



Optimal Tikhonov Regularization for DEER Spectroscopy

Thomas H. Edwards, Stefan Stoll

PII:	S1090-7807(18)30044-2
DOI:	https://doi.org/10.1016/j.jmr.2018.01.021
Reference:	YJMRE 6244
To appear in:	*Journal of Magnetic Resonance*
Received Date:	23 December 2017
Accepted Date:	29 January 2018

Please cite this article as: T.H. Edwards, S. Stoll, Optimal Tikhonov Regularization for DEER Spectroscopy, *Journal of Magnetic Resonance* (2018), doi: https://doi.org/10.1016/j.jmr.2018.01.021

Optimal Tikhonov Regularization for DEER Spectroscopy

Thomas H. Edwards, Stefan Stoll*

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

Abstract

Tikhonov regularization is the most commonly used method for extracting distance distributions from experimental double electron-electron resonance (DEER) spectroscopy data. This method requires the selection of a regularization parameter, α, and a regularization operator, L. We analyze the performance of a large set of α selection methods and several regularization operators, using a test set of over half a million synthetic noisy DEER traces. These are generated from distance distributions obtained from *in silico* double labeling of a protein crystal structure of T4 lysozyme with the spin label MTSSL. We compare the methods and operators based on their ability to recover the model distance distributions from the noisy time traces. The results indicate that several α selection methods perform quite well, among them the Akaike information criterion and the generalized cross validation with either the first- or second-derivative operator. They perform significantly better than currently utilized L-curve methods.

Keywords: inverse problem; Tikhonov regularization; DEER; PELDOR; penalized least-squares

1. Introduction

Double electron-electron resonance (DEER) spectroscopy, also called pulsed electron-electron double resonance (PELDOR) spectroscopy, measures the magnetic dipolar coupling between two or more paramagnetic centers, such as spin labels attached to proteins [1–3]. DEER data analysis usually involves the removal of a background signal followed by a transformation of the oscillatory time-domain signal into a distance-domain probability distribution function describing the distances between nearby paramagnetic centers (1.5-10 nm).

There exist several different approaches for extracting distance distributions from DEER data: Tikhonov regularization [4–7], Gaussian mixture models [8–10], Tikhonov regularization post-processed with Gaussians [11, 12], Tikhonov regularization combined with maximum entropy [13], Bayesian inference (based upon Tikhonov regularization) [14], regularization by limiting the number of points in the distance domain [15], wavelet denoising [16], truncated singular-value decomposition [6, 17], and neural networks [18]. Among them, Tikhonov regularization is the most widely employed method.

In this paper, we are concerned with the determination of optimal settings for Tikhonov regularization. This involves the choice of a regularization operator L and of a value for the regularization parameter α. An optimal choice of L and α ensures good distance distribution recovery and prevents overfitting the data; a bad choice causes poor recovery and either under- or overfitting to the data. There are several operators to choose from, and many methods are available for selecting α, each based on a defensible rationale. However, they vary greatly both in terms of theoretical justification and empirical track record. Therefore, the selection of the method/operator combination ought to be based on a thorough comparison of their performance for a practically relevant benchmark set of data analysis problems.

Tikhonov regularization was introduced to NMR for de-Pake-ing [19], for the extraction of internuclear distances from dipolar time-domain signals such as those from REDOR [20], for the determination of orientational distributions from ^{2}H NMR data [21, 22], and relaxation rate distributions [23, 24]. These approaches used the self-consistent method for selecting α, as introduced and implemented in the program FTIKREG [25, 26]. In the context of extracting distance distributions from DEER data, Tikhonov regularization was initally mentioned in 2002 [27, 28], and first applications appeared in 2004 [4, 5]. In these papers, the regularization parameter was selected manually or using FTIKREG. A thorough paper examining Tikhonov regularization and introducing the use of the L-curve maximum-curvature criterion appeared in 2005 [6]. A different L-curve method, the minimum-radius criterion, was introduced in 2006 in the program DeerAnalysis [7] and is used in its current release (2016).

Despite the long history and the widespread use of

*Corresponding author

Email addresses: edwardst@uw.edu (Thomas H. Edwards), stst@uw.edu (Stefan Stoll)