

# Estimation of The Discrete Spectrum of Relaxations For Electromagnetic Induction Responses Using $\ell_p$ -Regularized Least Squares for $0 \leq p \leq 1$

Mu-Hsin Wei, *Student Member, IEEE*, James H. McClellan., *Fellow, IEEE*, and Waymond R. Scott, Jr, *Fellow, IEEE*

**Abstract**—The EMI response of a target can be accurately modeled by a sum of real exponentials. However, in practice it is difficult to obtain the model parameters from measurements. We previously proposed a constrained linear method that can robustly estimate the model parameters when they are nonnegative. In this paper, we present a modified  $\ell_p$ -regularized least squares algorithm, for  $0 \leq p \leq 1$ , that eliminates the nonnegative constraint. An empirical method for choosing the regularization parameter is also studied. Using tests on synthetic data and laboratory measurements, the proposed method is shown to provide robust estimates of the model parameters in practice.

**Index Terms**—Electromagnetic induction (EMI), discrete spectrum of relaxation frequencies (DSRF), sum of exponentials,  $\ell_1$  minimization.

## I. INTRODUCTION

Recent research has shown that broadband EMI sensors are capable of discriminating between certain types of targets [1], [2]. The EMI frequency response  $H(\omega)$  of a metallic target can be expressed as [3]:

$$H(\omega) = c_0 + \sum_{k=1}^K \frac{c_k}{1 + j\omega/\zeta_k} \quad (1)$$

where  $c_0$  is the shift,  $K$  the model order,  $c_k$  the real spectral amplitudes, and  $\zeta_k$  the relaxation frequencies. The parameter set  $S = \{(\zeta_k, c_k) : k = 1 \dots K\}$  is called the Discrete Spectrum of Relaxation Frequencies (DSRF) or simply the spectrum; each pair  $(\zeta_k, c_k)$  is one relaxation.

It is advantageous to model the EMI signal with (1) because the relaxation frequencies are invariant to target orientation, and this is valuable in target detection. However, it is difficult in practice to obtain the model parameters in (1) from a small number of measurements. For most existing estimation methods, a good guess of the model order  $K$  is required for the fitting process to converge. Prior knowledge of  $K$ , however, is usually unavailable. The highly correlated summands in (1) and the nonlinear relation between  $H(\omega)$  and  $\zeta_k$  also make

estimation difficult. Most existing methods often give sub-optimal solutions that are far from the truth, or return complex parameters that sometimes lack physical meaning [4].

In [5] we proposed a constraint linear method that can robustly estimate the DSRFs without prior knowledge of the model order. The method however presumes nonnegative spectra for the targets which is valid for most targets using our system [6]; however, some targets can have a spectrum with positive and negative relaxations. In this paper, we propose an estimation method using  $\ell_p$ -regularized least squares ( $0 \leq p \leq 1$ ) that removes the nonnegative constraint. Since we ultimately want to solve for the  $\ell_0$ -regularized problem, more accurate estimates may be obtained using  $p < 1$  than  $p = 1$  when approximating, as argued in [7]. As with the previously proposed constrained optimization method, the  $\ell_p$  method always returns real model parameters and is stable under noise. While the proposed method is presented here in a frequency-domain application, the method can also be extended to time-domain applications in a straightforward manner.

## II. ESTIMATION METHOD

We formulate the DSRF estimation problem as a set of linear equations as described in Section II of [5]. When the target response is measured at  $N$  distinct frequencies ( $\omega_{\min} = \omega_1 < \omega_2 < \dots < \omega_N = \omega_{\max}$ ), the problem can be written in matrix form:

$$\begin{bmatrix} H(\omega_1) \\ H(\omega_2) \\ \vdots \\ H(\omega_N) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \frac{1}{1+j\omega_1/\zeta_1} & \frac{1}{1+j\omega_1/\zeta_2} & \cdots & \frac{1}{1+j\omega_1/\zeta_M} \\ 1 & \frac{1}{1+j\omega_2/\zeta_1} & \frac{1}{1+j\omega_2/\zeta_2} & \cdots & \frac{1}{1+j\omega_2/\zeta_M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \frac{1}{1+j\omega_N/\zeta_1} & \frac{1}{1+j\omega_N/\zeta_2} & \cdots & \frac{1}{1+j\omega_N/\zeta_M} \end{bmatrix}}_{\tilde{\mathbf{Z}}} \begin{bmatrix} \tilde{c}_0 \\ \tilde{c}_1 \\ \tilde{c}_2 \\ \vdots \\ \tilde{c}_K \end{bmatrix} + error \quad (2)$$

where  $\tilde{\zeta}_m$  are the *sampled* relaxation frequencies,  $\tilde{c}_m$  are the corresponding spectral amplitude estimates,  $\mathbf{h}$  the observation vector, and  $\tilde{\mathbf{Z}}$  the overcomplete dictionary. The sampled  $\tilde{\zeta}_m$  are generated by enumerating a large set of  $M$  possible relaxation frequencies uniformly distributed in the log- $\zeta$  space ( $M \gg K$ ).  $\tilde{\mathbf{c}}$  is the weighted selector vector containing the shift estimator  $\tilde{c}_0$  followed by the spectral amplitude

This work is supported in part by the US Army REDCOM CERDEC Night Vision and Electronic Sensors Directorate, Science and Technology Division, Countermeasures Branch and in part by the U. S. Army Research Office under Contract Number W911NF-05-1-0257.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA. (e-mail: m.wei@gatech.edu).

estimators. We expect the solution for  $\tilde{c}$  to have many zero elements because  $M \gg K$ , i.e.,  $\tilde{c}$  will be sparse.

We then utilize the  $\ell_p$ -regularized least squares technique, for  $0 \leq p \leq 1$ , because it promotes sparse solutions [8].

$$\arg \min_{\tilde{c}} \|\tilde{Z}'\tilde{c} - \mathbf{h}'\|_2^2 + \lambda \|\tilde{c}\|_p^p, \quad 0 \leq p \leq 1 \quad (3)$$

$$\text{where } \tilde{Z}' = \begin{bmatrix} \Re(\tilde{Z}) \\ \Im(\tilde{Z}) \end{bmatrix} \text{ and } \mathbf{h}' = \begin{bmatrix} \Re(\mathbf{h}) \\ \Im(\mathbf{h}) \end{bmatrix}$$

where  $\lambda$  is the regularization parameter. Separating the real and imaginary parts in  $\tilde{Z}$  makes the whole system real. Ideally, in the optimal  $\tilde{c}$ , only those  $\tilde{c}_m$  with corresponding  $\tilde{\zeta}_m$  that are near a true  $\zeta_k$  will be nonzero, and they will take on the correct spectral amplitudes  $c_k$ . It follows that a DSRF can then be deduced from the nonzero estimated  $\tilde{c}_m$  and their corresponding  $\tilde{\zeta}_m$ .

The  $\ell_p$ -regularized least squares solution for  $p < 1$  can be approximated by the iteratively reweighted  $\ell_1$  algorithm proposed by Candès *et al.* [8]. The weights are updated as suggested in [9]. We also adopt the  $\epsilon$ -regularization technique used in the same paper. In summary, (3) is approximated by (see also [10]):

---

**Algorithm 1:** Approximated  $\ell_p$ -regularized least squares

---

**Input:**  $\tilde{Z}', \mathbf{h}', p, \lambda, \tilde{c}^0$

- 1  $\tilde{c}^n \leftarrow \tilde{c}^0$
- 2 **for**  $k \leftarrow 0$  **to**  $-8$  **step**  $-1$  **do**
- 3      $\epsilon \leftarrow 10^k$
- 4     **repeat**
- 5          $\tilde{c}^{n-1} \leftarrow \tilde{c}^n$
- 6          $w_i^n \leftarrow (|\tilde{c}_i^{n-1}| + \epsilon)^{p-1}$
- 7          $\tilde{c}^n \leftarrow \arg \min \|\tilde{Z}'\tilde{c} - \mathbf{h}'\|_2^2 + \lambda \sum_{i=1}^{M+1} w_i^n |\tilde{c}_i|$
- 8     **until**  $\|\tilde{c}^n - \tilde{c}^{n-1}\|_2 < \sqrt{\epsilon}/100$
- 9 **return**  $\tilde{c}^n$

---

The  $\ell_1$  minimization problem in step 7 is solved by **11\_ls**, a MATLAB optimizer proposed by Kim *et al.* [11]. We have also found that normalizing the input data  $\mathbf{h}$ , as well as the columns of  $\tilde{Z}'$ , to have unit  $\ell_2$  norm increases the accuracy of estimation. While it is often suggested to initialize  $\tilde{c}^0$  using the least-squares solution [9], we observe that setting entries of  $\tilde{c}^0$  to all ones also seems to be effective and converges faster. The nonzero entries of  $\tilde{c}$  selected by (3) along with the corresponding  $\tilde{\zeta}_m$  are the relaxations needed in the estimated DSRF,  $\hat{S} = \{(\hat{\zeta}_l, \hat{c}_l) : l = 1 \dots L\}$ .

### III. ESTIMATION RESULTS

The proposed estimation method is tested against synthetic and laboratory data to show its functionality, accuracy, and stability. The hardware system used is a wideband EMI sensor operating at 21 frequencies approximately logarithmically distributed over the range 300 Hz–90 kHz (2.5 decades) [6]. The synthetic data is generated in accordance with the hardware specification. The range of  $\zeta$  for estimation is chosen such that  $\log(\tilde{\zeta}_{\min})$  and  $\log(\tilde{\zeta}_{\max})$  are 2.45 and 6.62, respectively, i.e.,

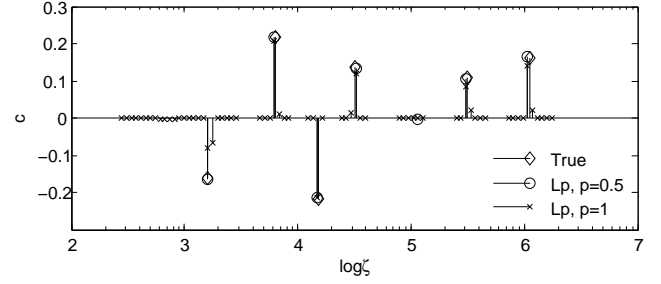


Fig. 1. Estimation of a synthetic six-relaxation DSRF.

4.17 decades. All estimations are performed with  $M = 100$ , and all presented spectra are normalized such that  $\sum |c_i| = 1$ . Spectral amplitudes less than  $10^{-5}$  are not displayed. All frequency responses are normalized such that  $\|\mathbf{h}\|_2 = 1$ . Unless specified,  $p = 0.5$  is chosen as a representative case. The regularization parameter  $\lambda$  is chosen based on the method described in Section IV. Results presented in this section may achieve higher accuracy with a more sophisticated  $\lambda$  selection rule. Here we demonstrate the usability of the proposed algorithm with a simple  $\lambda$  selection rule. See Section IV for more discussion on the choice of  $\lambda$ .

**Notation:**  $\zeta$  and  $c$  are the true/theoretical relaxation frequencies and spectral amplitudes;  $\hat{\zeta}$  and  $\hat{c}$  are the estimates.

#### A. Dissimilarity Measure Between Two DSRFs

In order to evaluate performance between the estimated DSRF and the true spectrum, we need to define a measure of dissimilarity that is appropriate for sparse spectra with multiple peaks. We use the Earth Mover’s Distance (EMD) [12] which quantifies the “amount of work” to morph one spectrum into the other. Strictly speaking the EMD is only defined for positive spectra, but we can account for negative spectral amplitudes by defining the distance function between two relaxations  $(\zeta_i, c_i)$  and  $(\hat{\zeta}_j, \hat{c}_j)$  to be:

$$d_{ij} = \begin{cases} |\log \zeta_i - \log \hat{\zeta}_j| & , c_i \hat{c}_j \geq 0 \\ 1 + |\log \zeta_i - \log \hat{\zeta}_j| & , c_i \hat{c}_j < 0 \end{cases}$$

which penalizes relaxations with different signs. Spectra are made nonnegative and normalized prior to the EMD computation. Finally, notice that the EMD is measured in decades because it is examined in  $\log$ - $\zeta$  space.

The EMD compares two spectra as a whole, so the effect of very small amplitude relaxations is tiny in the EMD, and neglecting these small components amounts to assuming they are near the noise level of the measured frequency response.

#### B. Synthetic Six-relaxation DSRF

We test our method (using  $p = 0.5$ ) on a six-relaxation DSRF synthesized at 70 dB SNR with AWGN (Fig. 1). This is a case that cannot be handled by traditional nonlinear parameter optimization [4], or the nonnegative linear method [5]. All six relaxation frequencies are recovered by using **Algorithm 1**. The estimation is nearly perfect, because the estimated model parameters are real, and the deviation from truth is

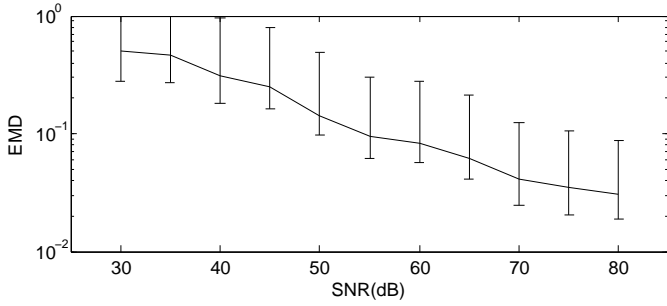


Fig. 2. Monte Carlo simulation on goodness of estimation vs. SNR performed on a four-relaxation DSRF. Sample size is 100 per SNR.

small. The EMD between the estimated and the true DSRF is 0.01 decades. There is one extra relaxation near  $\log\zeta=5$  in the estimated spectrum of Fig. 1, but it has very small amplitude and can be safely neglected.

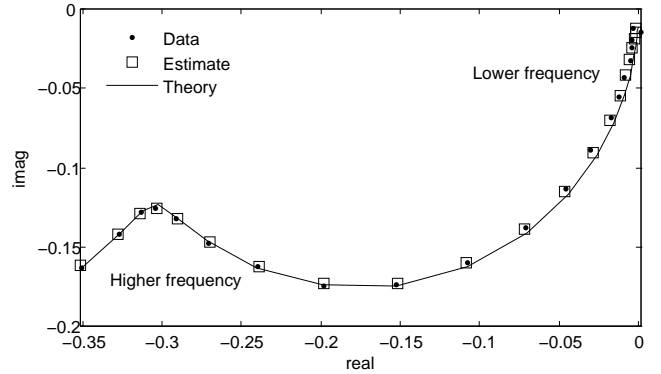
This spectrum is also estimated with  $p = 1$  using `ll_ls`. In this case, many extra relaxations are introduced by the fitting process (Fig. 1); the EMD is 0.03. Real targets are not likely to have a spectrum with many small relaxations around a strong relaxation. In fact, Baum argues that physical relaxation frequencies are discrete [3]. However, the small relaxations introduced by  $p = 1$  seem to give a continuous spectrum of relaxation frequencies. In this sense,  $p < 1$  gives a sparser solution that more accurately resembles a physical spectrum even though this may not always be reflected in the EMD measure.

C. Signal to Noise Ratio

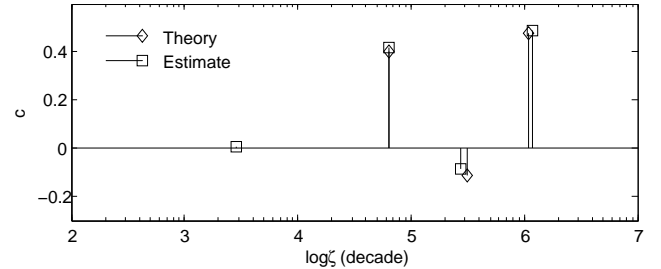
To see how the proposed method performs in noise, a Monte Carlo simulation versus SNR is performed. The true spectrum is from a target with a four-relaxation DSRF including negative relaxations. The simulation result, shown in Fig. 2, shows the robustness of the estimation method at different signal-to-noise ratios. The EMD between the estimate and the truth increases as the SNR decreases. This suggests that the proposed method is usable in a range of SNR where the EMD is below some threshold. This threshold, however, depends on the application of the estimated spectrum. For example, if the DSRF produces features for classification, a more robust classifier may tolerate worse estimations and, therefore, allow lower SNR. For our purpose, spectra with an EMD below 0.1 are considered visually similar, those with an EMD above 0.2 exhibit visual differences, but may still resemble each other. In our laboratory measurements, a typical SNR for loop targets is 70 dB.

D. Laboratory Data

We verify the functionality of the estimation method on laboratory data where we know the theoretical DSRF [13]. We examine a target that consists of three mutually orthogonal copper loops. The loop diameters and thickness are 3/20, 4/30, and 5/36, respectively in cm/AWG<sup>1</sup>. We pick a specific orientation and position relative to the EMI sensor that best



(a)



(b)

Fig. 3. (a) Frequency response of three mutually orthogonal copper loops. (b) Theoretical and estimated DSRF of the response in (a).

shows the bipolar relaxation amplitudes in a spectrum. The target frequency response of this configuration is shown in Fig. 3(a), the SNR is 38 dB, and its estimated DSRF in Fig. 3(b). Theoretical data are also displayed. We see that the estimate and theory agree well, and the EMD between the theoretical and estimated DSRF is 0.10 decades.

Next we examine the changes in the DSRF as the target moves relative to the EMI sensor. The same target configured at a fixed orientation is displaced at different positions along a horizontal axis, which we will call  $x$ . The vertical distance between the target and sensor is 6 cm. The EMI sensor is located at  $x = 0$ . Samples of the measured target responses are shown in Fig. 4(a); their corresponding spectra are in Fig. 4(b). Theoretical results are also shown. Overall, the theory and measurement agree. The disagreement at  $x = -0.5$  may be because of approximations in the model and/or inaccuracies in the positions measured in the experiment.

As expected from the theory, while the frequency response changes dramatically as the target moves along the  $x$  axis, the corresponding change in the spectral domain only occurs in the spectral amplitudes. The three dominant relaxation frequencies remain unchanged. The proposed method successfully estimates the spectra that agree with this phenomenon. All three relaxation frequencies are consistently estimated. The extra relaxations all have small amplitudes that can be safely ignored. This invariant property of the relaxation frequencies makes the DSRF valuable especially for target discrimination.

IV. CHOOSING  $\lambda$

In this section, we first examine the behavior of the proposed method in relation to the regularization parameter  $\lambda$ , and then

<sup>1</sup>American Wire Gauge

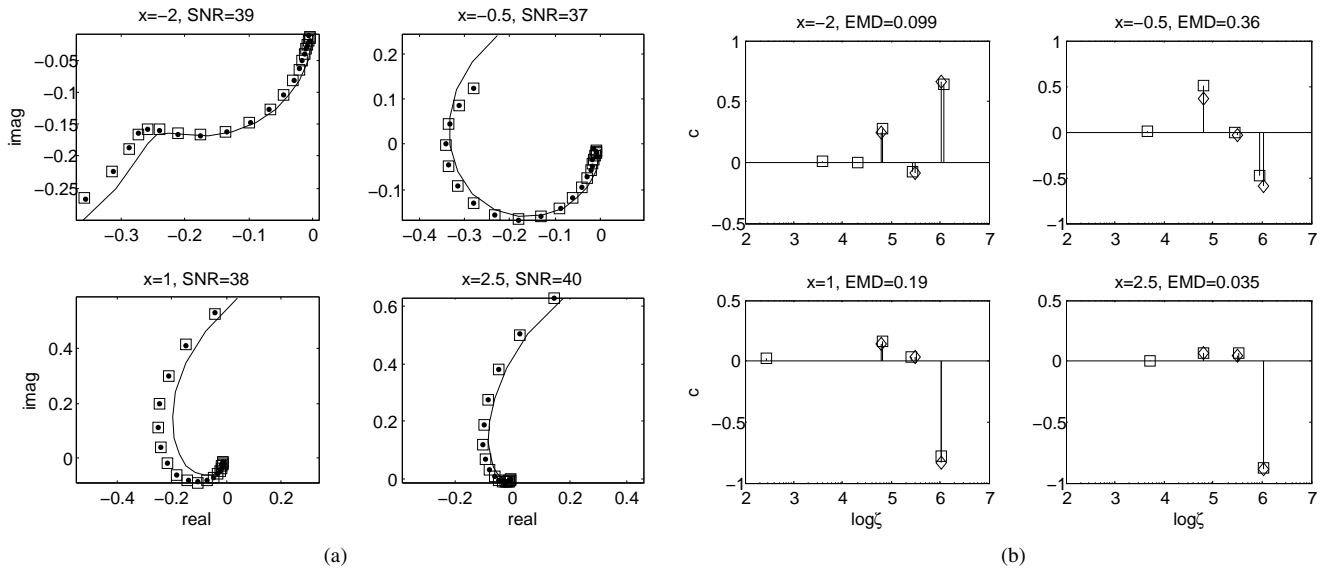


Fig. 4. The plots share the same annotation as Fig. 3. (a) Frequency responses of the three mutually orthogonal copper loops at different locations. (b) Theoretical and estimated DSRF of the corresponding responses in (a). The SNR is measured in dB,  $x$  positions in cm, and EMD in decades.

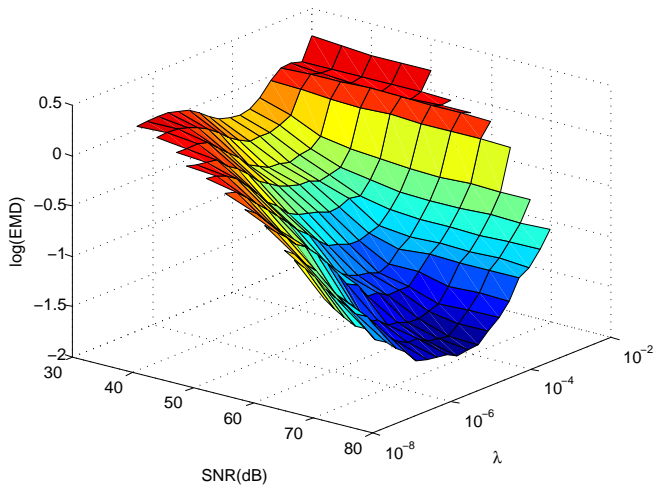


Fig. 5. Monte Carlo simulation of the goodness of estimation (EMD) of a four-relaxation spectrum at different SNR's and  $\lambda$ 's.

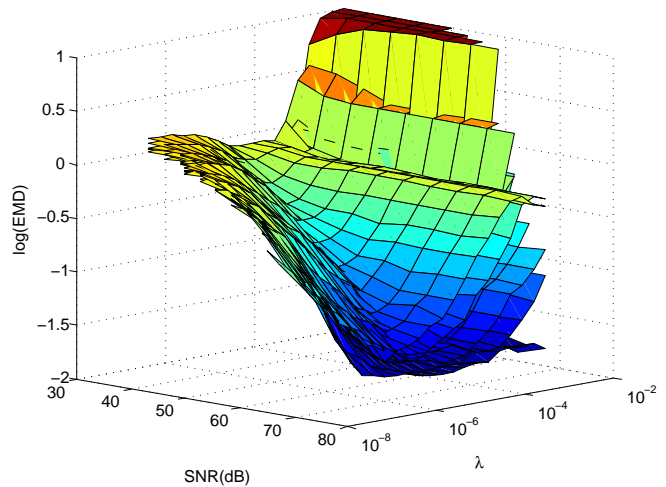


Fig. 6. Same simulation as in Fig. 5, but for spectra of different model orders (1 to 6). Each spectrum constitutes one surface in the figure. All surfaces have their minimum in the same SNR- $\lambda$  region.

we propose a simple  $\lambda$  selection rule exploiting the observed properties of  $\lambda$ . All discussions and figures presented here assume  $p = 0.5$  unless otherwise specified.

To understand how the goodness of fit changes with  $\lambda$  and SNR, we conduct a cross-validation-like simulation. First, we build a collection of synthetic spectra with different model orders and a variety of distributions of relaxations. For each spectrum at a fixed SNR, the spectrum is estimated 100 times for each  $\lambda$  within a range, and the average goodness of fit, measured by the EMD between the available truth and the estimate, is recorded. This is done for a range of SNRs. The simulation result for a four-relaxation spectrum, as an example, is shown in Fig. 5. We see that not only is the EMD surface well-behaved (i.e., smooth) with respect to the SNR and  $\lambda$ , but more importantly the surface itself is convex-shaped. Thus, at each SNR, the minimum EMD is achievable by a unique  $\lambda$ .

The wide valley of the surface also shows that the goodness of fit is not very sensitive near the optimal  $\lambda$  that gives the minimum EMD per SNR.

Simulations of spectra for other model orders and distributions also exhibit the same property (Fig. 6). Moreover, the valleys of the EMD surface all occur in nearly the same SNR- $\lambda$  region. In other words, the  $\lambda$  that produces the minimum EMD at a given SNR is quasi-independent of the model order. Figure 7 shows the averaged EMD of different model orders in Fig. 6. The resulting surface still exhibits the properties described above. This allows us to pick a near-optimal  $\lambda$  based solely on the SNR.

From the data in Fig. 7, we can find the optimal  $\lambda$  at each SNR which is also plotted in the same figure. Using the wide-valley property, we can achieve the near-minimum EMD by choosing  $\lambda$ 's that are near the optimal  $\lambda$  in the valley. Here we

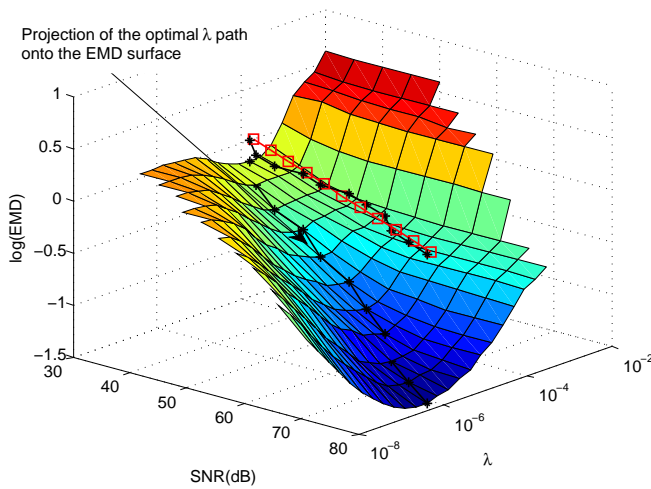


Fig. 7. Average of EMD surfaces in Fig. 6. The curve with asterisk markers traces out the optimal  $\lambda$ 's. The line with square markers is the approximated optimal  $\lambda$  curve used to select  $\lambda$  in practical estimation.

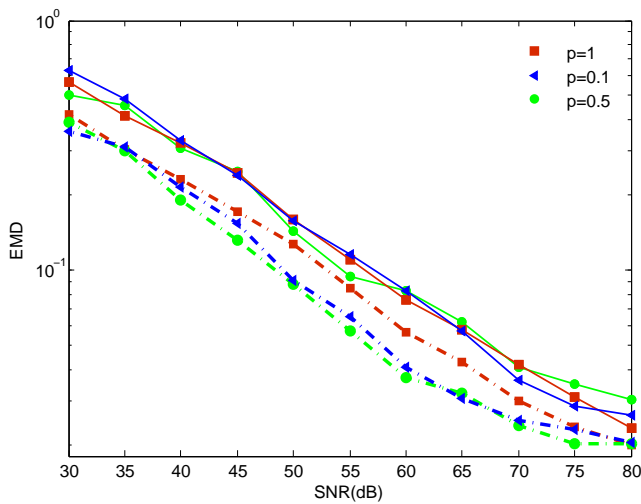


Fig. 8. Goodness of fit using the linear  $\log\text{-}\lambda$  selection rule for several  $p$ 's. The true spectrum is the same as in Fig. 5. Dash-dot curves denote the optimal  $\lambda$ , solid lines denote the linear  $\log\text{-}\lambda$  selection rule.

approximate the optimal  $\lambda$  with a semilog function of SNR. This is done by fitting the optimal  $\log\text{-}\lambda$  curve with a linear function. Weights may be added to promote certain SNR's that are more important. For our problem setup, the  $\lambda$  is chosen by (also shown in Fig. 7)

$$\log \lambda = -0.05 \cdot \text{SNR} - 2.2 \quad (4)$$

In practice, this  $\log\text{-}\lambda$  selection rule that is linear in SNR allows the regularization parameter to be determined with negligible computation time. When processing the laboratory data, we use (4) along with an estimate of the SNR to determine  $\lambda$  for use in **Algorithm 1**. The same empirical method can be repeated for other  $p$ 's, and the result is also a linear relationship between  $\log \lambda$  and SNR.

Figure 8 compares the goodness of estimation of a four-relaxation target using the linear  $\log\text{-}\lambda$  selection rule and the optimal  $\lambda$  which uses the true spectrum that is not available in practice. We see a slight increase in the EMD when the linear

$\log\text{-}\lambda$  selection rule is used, which is reasonable and expected. The increase is acceptable and hence the linear  $\log\text{-}\lambda$  selection rule is an appropriate  $\lambda$  selector.

Also shown in Fig. 8 are the performances of other  $p$  values. It is seen that  $p < 1$  gives more accurate results than  $p = 1$  when the optimal  $\lambda$  is used, but this advantage is significantly diminished when the linear  $\log\text{-}\lambda$  selection rule is used. While this lessens the advantages of using  $p < 1$ , we emphasize that  $p = 1$  tends to give estimates with many relaxation frequencies while  $p < 1$  gives sparser spectra which are more physically accurate (see Section III-B). It is possible that both the accuracy and the sparsity advantages for  $p < 1$  could be obtained with a better  $\lambda$  selection rule. Lastly, since the performance of a certain  $p$  value is dependent on the  $\lambda$  selection rule used, different optimum  $p$  values would be determined if the  $\lambda$  selection rule is changed.

## REFERENCES

- [1] P. Gao, L. Collins, P. M. Garber, N. Geng, and L. Carin, "Classification of landmine-like metal targets using wideband electromagnetic induction," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1352–1361, May 2000.
- [2] E. B. Fails, P. A. Torrione, W. R. Scott, Jr., and L. M. Collins, "Performance of a four parameter model for modeling landmine signatures in frequency domain wideband electromagnetic induction detection systems," in *Proc. SPIE*, vol. 6553, Orlando, FL, Apr. 2007, p. 65530D.
- [3] C. E. Baum, "On the singularity expansion method for the solution of electromagnetic interaction problems," Air Force Weapons Laboratory, Interaction Notes 88, 1971.
- [4] Y. Das and J. E. McFee, "Limitations in identifying objects from their time-domain electromagnetic induction response," in *Proc. SPIE*, vol. 4742, Orlando, FL, Apr. 2002, pp. 776–788.
- [5] M. Wei, W. R. Scott, Jr., and J. H. McClellan, "Robust estimation of the discrete spectrum of relaxations for electromagnetic induction responses," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1169–1179, Mar. 2010.
- [6] W. R. Scott, Jr., "Broadband array of electromagnetic induction sensors for detecting buried landmines," in *Proc. IGARSS*, Boston, MA, Jul. 2008.
- [7] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, 2007.
- [8] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted  $\ell_1$  minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, 2008.
- [9] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *ICASSP*, Las Vegas, NV, Mar. 2008, pp. 3869–3872.
- [10] M. A. T. Figueiredo and R. D. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *ICIP*, vol. 2, Genoa, Italy, 2005, pp. 782–785.
- [11] S. J. Kim, K. Koh, M. Lustig, and S. Boyd, "An efficient method for compressed sensing," in *ICIP*, vol. 3, San Antonio, TX, 2007, pp. 117–120.
- [12] Y. Rubner, C. Tomasi, and L. J. Guibas, "A metric for distributions with applications to image databases," in *Proc. ICCV*, Bombay, India, Jan. 1998, pp. 59–66.
- [13] G. D. Larson and W. R. Scott, Jr., "Automated, non-metallic measurement facility for testing and development of electromagnetic induction sensors for landmine detection," in *Proc. SPIE*, vol. 7303, Orlando, FL, Apr. 2009, p. 73030X.