HIGH RESOLUTION SPARSE ESTIMATION OF EXPONENTIALLY DECAYING TWO-DIMENSIONAL SIGNALS

Stefan Ingi Adalbjörnsson, Johan Swärd, and Andreas Jakobsson

Dept. of Mathematical Statistics, Lund University, Sweden

ABSTRACT

In this work, we consider the problem of high-resolution estimation of the parameters detailing a two-dimensional (2-D) signal consisting of an unknown number of exponentially decaying sinusoidal components. Interpreting the estimation problem as a block (or group) sparse representation problem allows the decoupling of the 2-D data structure into a sum of outer-products of 1-D damped sinusoidal signals with unknown damping and frequency. The resulting non-zero blocks will represent each of the 1-D damped sinusoids, which may then be used as non-parametric estimates of the corresponding 1-D signals; this implies that the sought 2-D modes may be estimated using a sequence of 1-D optimization problems. The resulting sparse representation problem is solved using an iterative ADMM-based algorithm, after which the damping and frequency parameter can be estimated by a sequence of simple 1-D optimization problems.

***Index Terms*—** Sparse signal modeling, Spectral analysis, Sparse reconstruction, Parameter estimation, ADMM.

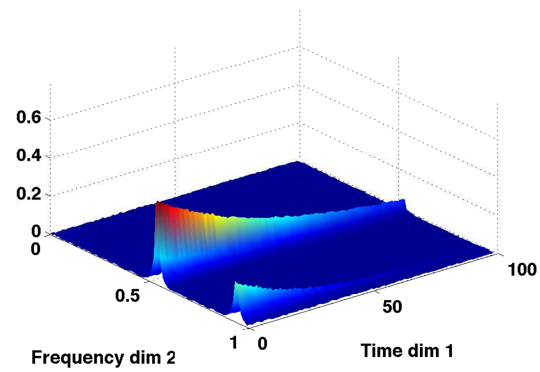

Fig. 1. An estimate of $\mathbf{Z}$ for data containing two damped sinusoids. Estimating the frequency and damping coefficient for these damped sinusoids yields the estimates of $f_{1,1}$, $f_{1,2}$, $\beta_{1,1}$, and $\beta_{1,2}$ for the first time dimension.

1. INTRODUCTION

High-dimensional decaying sinusoidal signals occur in a wide variety of fields, such as spectroscopy, geology, sonar, and radar, and given the importance of such signals in a variety of applications, the topic has attracted notable attention in the recent literature (see, e.g. [1–7]). Common solutions include subspace-based algorithms [1–3, 5–7], typically making strong model assumptions, or the use of high-dimensional representations necessitating an iterative zooming procedure over multiple dimensions, such as the technique introduced in [4]. Such approaches often suffer from high complexity and sub-optimal performance, typically requiring an accurate initialization or model order information to yield reliable results, information which is commonly not available in many of the discussed applications. Often, the measurements are also assumed to be uniformly sampled, which may well be undesired in applications such as, for instance, spectroscopy. In this work, we formulate a sparse representation separating the frequency and damping dimensions in order to facilitate multiple low-dimensional searches over the separate dimensions, without imposing the need for detailed initialization or assuming any *a priori* model order information. By separating the two-dimensional (2-D) search into two *linked* 1-D searches, the 2-D estimation problem may be decoupled allowing for a both computationally and memory efficient estimation procedure. The work allows for non-uniformly sampled data and is an extension of our recent technique for accurate estimation of 1-D decaying sinusoidal signals [8], which is here extended to allow for 2-D signals. In order to reduce complexity, we herein further propose a computationally efficient implementation based on the concept of the alternating direction method of multipliers (ADMM) [9]. The remainder of the paper is organized as follows: in the next section, we introduce the considered data model. Then, in section 3, we introduce the idea behind decoupling the search dimensions. Section 4 introduce the ADMM formulation of the estimator, and Section 5 illustrates the performance of the proposed estimator using numerical simulations and using measured 2-D nuclear magnetic resonance (NMR) data. Finally, Section 6 contains our conclusions.

This work was supported in part by the Swedish Research Council and the Crafoord's and Carl Trygger's foundations.

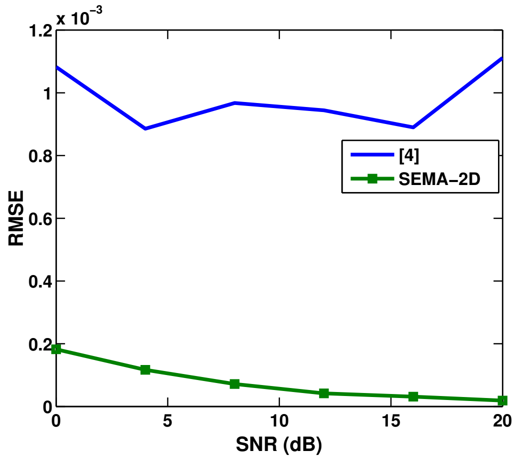

Fig. 2. The RMSE of $f_{1,1}$ as a function of the SNR. The signal contains two modes.

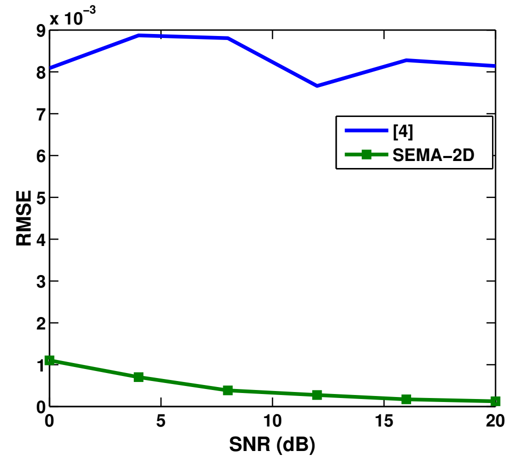

Fig. 3. The RMSE of $\beta_{1,1}$ as a function of the SNR. The signal contains two modes.

2. DATA MODEL

Let $y(\tau_1, \tau_2)$ denote a (possibly non-uniformly sampled) signal of interest that may be well modeled as a sum of 2-D exponentially decaying sinusoids corrupted by an additive noise, i.e.,

$$y(\tau_1, \tau_2) = \sum_{k=1}^{K} \alpha_k z_{1,k}^{\tau_1} z_{2,k}^{\tau_2} + e(\tau_1, \tau_2) \quad (1)$$

where $\tau_l = t_{\ell,1}, \ldots, t_{\ell,N_\ell}$, for $\ell = 1$ and 2,

$$z_{\ell,k} = e^{2i\pi f_{\ell,k} - \beta_{\ell,k}} \quad (2)$$

and with K denoting the number of exponentially decaying sinusoids, α_k the complex amplitude of the k:th mode, and $f_{\ell,k}$ and $\beta_{\ell,k}$ the k:th frequency and damping coefficient for dimension ℓ, respectively. Here, the unknown parameters to be estimated are not only the complex amplitudes, the frequency, and damping, but also the number of modes K. The additive noise $e(\tau_1, \tau_2)$ is assumed to be well modeled as an uncorrelated 2-D Gaussian process. Let

$$\mathbf{a}_{\ell,k} = \begin{bmatrix} z_{\ell,k}^{t_{\ell,1}} & \ldots & z_{\ell,k}^{t_{\ell,N_\ell}} \end{bmatrix}^T \quad (3)$$

where $(\cdot)^T$ denotes the transpose. Then, (1) may be expressed concisely over the $N_1 \times N_2$ time samples as

$$\mathbf{Y} = \tilde{\mathbf{A}}_1 \mathbf{B} \tilde{\mathbf{A}}_2^T + \mathbf{E} \quad (4)$$

where $\tilde{\mathbf{A}}_\ell$ is an $N_\ell \times K$ dimensional matrix formed as

$$\tilde{\mathbf{A}}_\ell = \begin{bmatrix} \mathbf{a}_{\ell,1} & \ldots & \mathbf{a}_{\ell,K} \end{bmatrix} \quad (5)$$

for $\ell = 1$ and 2, and

$$\mathbf{B} = \text{diag}\left\{ \begin{bmatrix} \alpha_1 & \ldots & \alpha_K \end{bmatrix} \right\} \quad (6)$$

with diag $\{\mathbf{x}\}$ denoting the diagonal matrix formed with the vector $\mathbf{x}$ along its diagonal, and where $\mathbf{E}$ is formed similarly to $\mathbf{Y}$.

3. DECOUPLING THE SEARCH DIMENSIONS

Exploiting the sparsity of the representation, one may form a large dictionary matrix, say $\mathbf{U}$, containing all considered combinations of frequencies and dampings, and then solve

$$\underset{\mathbf{x}}{\text{minimize}} \quad \left|\left|\text{vec}(\mathbf{Y}) - \mathbf{U}\mathbf{x}\right|\right|_2^2 + \lambda \left|\left|\mathbf{x}\right|\right|_1 \quad (7)$$

where $||\cdot||_F$ denotes the Frobenius norm and vec$\{\cdot\}$ the vectorization operator. The estimated (sparse) vector $\mathbf{x}$ would then contain the dictionary elements corresponding to the desired modes. However, such a solution would be practically impossible due to the enormous dimension of the resulting dictionary; even for a grid containing only 100 grid points in each dimension, this would require 10^8 dictionary elements. As an alternative, one may form the minimization

$$\underset{\mathbf{B}}{\text{minimize}} \quad \left|\left|\mathbf{Y} - \mathbf{A}_1 \mathbf{B} \mathbf{A}_2^T\right|\right|_F^2 + \lambda \left|\left|\mathbf{B}\right|\right|_1 \quad (8)$$

which instead aims at fitting a (sparse) diagonal matrix $\mathbf{B}$. In order to reduce the dimensionality, one may note that the k:th mode in (1) may be expressed as the outer product of $\mathbf{a}_{1,k}$ and $\mathbf{a}_{2,k}$. Thus, by introducing $\mathbf{x} = \alpha_k \mathbf{a}_{2,k}$ and $\mathbf{z} = \alpha_k \mathbf{a}_{1,k}$, the minimization in (8) may be replaced by

$$\underset{\mathbf{X}}{\text{minimize}} \quad \left|\left|\mathbf{Y} - \mathbf{A}_1 \mathbf{X}\right|\right|_F^2 + \lambda_1 \sum_{m=1}^{M_1} \left|\left| \left[\mathbf{X}\right]_m \right|\right|_2 \quad (9)$$

$$\underset{\mathbf{Z}}{\text{minimize}} \quad \left|\left|\mathbf{Y} - \mathbf{Z}\mathbf{A}_2\right|\right|_F^2 + \lambda_2 \sum_{m=1}^{M_2} \left|\left| \left[\mathbf{Z}^T\right]_m \right|\right|_2 \quad (10)$$

where $\left[\mathbf{X}\right]_m$ denotes the m:th row of $\mathbf{X}$, which efficiently decouples the problem into two sub-problems of dimension $N_1 M_1 B_1 + N_2 M_2 B_2$, where M_1, M_2, B_1, and B_2 denote the number of grid points in the frequency and damping grids,

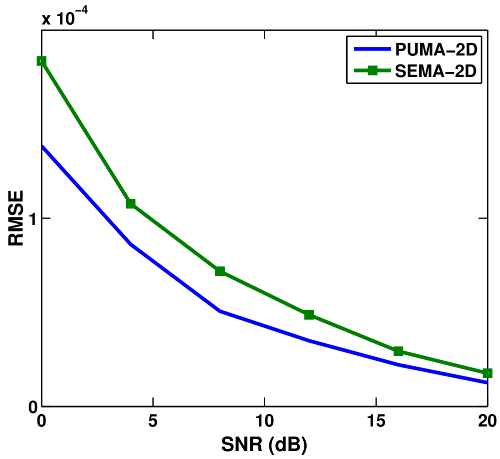

Fig. 4. The RMSE of the $f_{1,1}$ as a function of SNR. The signal contains two modes.

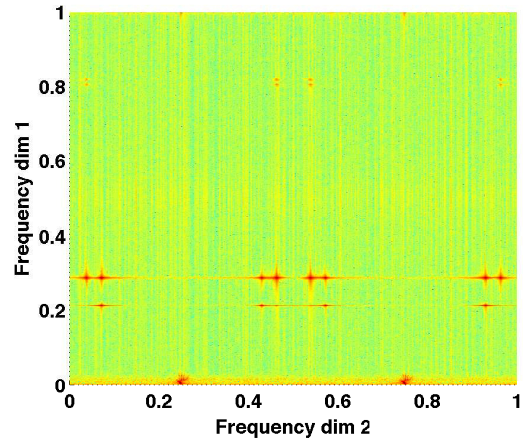

Fig. 5. The logarithmic periodogram estimate of the examined 2-D NMR data.

respectively. Reminiscent to the development in [8], one may then proceed to initially form the dictionary matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ by setting all the damping parameters to zero, i.e., $\mathbf{A}_1$ and $\mathbf{A}_2$ are initially only Fourier matrices. Solving (9) and (10) with these matrices yield the sparse estimates $\hat{\mathbf{X}}$ and $\hat{\mathbf{Z}}$. Each non-zero row, k, in $\hat{\mathbf{X}}$ corresponds to a damped sinusoid in the second dimension as well as the on-grid frequency estimate $\hat{f}_{1,k}$ in the first dimension. Similarly, each non-zeros column, j, in $\hat{\mathbf{Z}}$ corresponds to a damped sinusoid in the first dimension and a frequency $\hat{g}_{2,j}$ in the second dimension, as is illustrated in Figure 1. The figure illustrates how $\hat{\mathbf{Z}}$ captures the two modes which then decay over the first time dimension, τ_1. Utilizing this, one may then find the damping and frequency parameters for the second dimension by solving a non-linear least squares (NLS) problem over the frequency and damping for each non-zero row in $\mathbf{X}$, i.e., say that the m:th column of $\mathbf{X}$ contains non-zero elements. Then, one may find the frequency and damping parameters for the second time dimension as

$$\{\hat{f}, \hat{\beta}\} = \underset{f,\beta}{\text{minimize}} \quad \left|\left| [\mathbf{X}]_m - \hat{\alpha}\underline{\mathbf{a}} \right|\right|_2^2 \tag{11}$$

where (see also [10])

$$\hat{\alpha} = \frac{[\mathbf{X}]_m \underline{\mathbf{a}}}{\underline{\mathbf{a}}^H \underline{\mathbf{a}}} = \frac{1}{N_1} [\mathbf{X}]_m \underline{\mathbf{a}} \tag{12}$$

with $(\cdot)^H$ denoting the conjugate transpose, and $\underline{\mathbf{a}}$ is a vector formed similar to $\mathbf{a}_{\ell,k}$, defined as in (3), but for a generic f and β. Typically, (11) is solved by evaluating the cost function for a grid of different f and β, where then $\underline{\mathbf{a}}$ is created using those parameters according to (3) . Similarly, the frequency and damping parameters for the first time dimension may be estimated by minimizing

$$\{f^*, \beta^*\} = \underset{f,\beta}{\text{minimize}} \quad \left|\left| \left[\mathbf{Z}^T\right]_m - \hat{\alpha}\underline{\mathbf{a}}^T \right|\right|_2^2 \tag{13}$$

where $\hat{\alpha}$ is estimated as in (12). One may thus by iteratively solving (11) and (13), using (12), find the frequency and damping coefficients in the two dimensions without having to solve the full 2-D minimization directly. We thus obtain a set of estimates, such that from $\mathbf{X}$, we will, in the first dimension, obtain a grid-restricted estimate of $\hat{f}_{1,k}$, as well as $\hat{f}_{2,k}$ and $\hat{\beta}_{2,k}$, in the second dimension, which will, due to the search, be unrestricted by any grid. Similarly, from $\mathbf{Z}$, in the first dimension, we obtain $\hat{g}_{1,k}$ and $\hat{\beta}_{1,k}$, which will then be unrestricted by any grid, as well as the grid-restricted $\hat{g}_{2,k}$ in the second dimension. As a final step, we pair together the obtained unrestricted frequency estimates using the minimum frequency distance

$$\sum_{l=1}^{\hat{K}} d(\hat{f}_{1,l}, \tilde{g}_{1,l}) + d(\hat{f}_{2,l}, \tilde{g}_{2,l}) \tag{14}$$

where the distance function is defined as

$$d(a,b) = \min(|b-a|, |b-(1+a)|, |1+b-a|)$$

and $\{\tilde{g}_{1,j}, \tilde{g}_{2,j}\}$ is formed from all possible permutations of $\{\hat{g}_{1,j}, \hat{g}_{2,j}\}$, for $j = 1, \ldots, \hat{K}$. The unrestricted damping estimates are paired accordingly with the resulting permutation. Alternatively, for signals with many modes, the pairings may be found by ordering both sets in order of increasing frequency in the first dimension. It should be stressed that the decoupling of the problem retains the 2-D structure and dependencies, such that the two minimizations combined correspond to the actual modes, having the benefit that the parameters estimate need not be restricted to a fixed grid of values. We note that the resulting algorithm may also be extended to enable a similar decoupling of higher dimensional data; we are currently exploring such an extension.

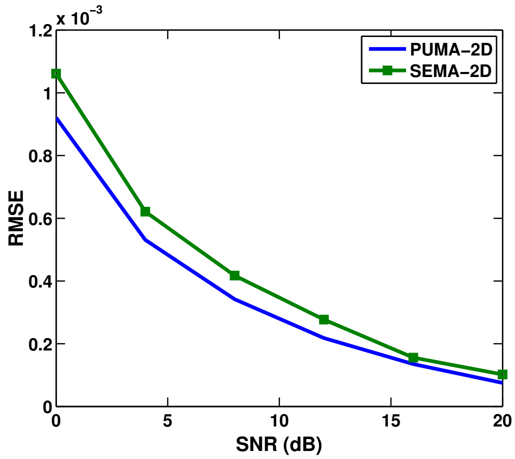

Fig. 6. The RMSE of $\beta_{1,1}$ as a function of SNR. The signal contains two modes.

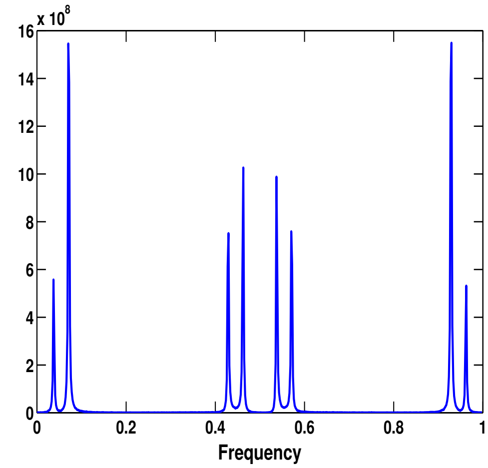

Fig. 7. The DFT of the time series for the second time dimension found in the row in $\mathbf{X}$ that corresponds to $f_1 = 0.2881$.

4. IMPLEMENTATION USING ADMM

In order to reduce the complexity of the minimization of solving (9) and (10), we proceed to present an ADMM-based formulation of the joint minimization. In the interest of brevity, we only show the derivation of (9) since the one for (10) follows analogically if the problem is transposed. Define

$$f(\mathbf{X}) = \frac{1}{2}||\mathbf{Y} - \mathbf{A}\mathbf{X}||_F^2 \tag{15}$$

$$g(\mathbf{V}) = \lambda \sum_{m=1}^{M_1} ||\mathbf{V}_m||_2 \tag{16}$$

The minimization in (9) may then be expressed as

$$\begin{aligned} \underset{\mathbf{X}}{\text{minimize}} \quad & f(\mathbf{X}) + g(\mathbf{V}) \\ \text{subject to} \quad & \mathbf{X} - \mathbf{V} = \mathbf{0} \end{aligned} \tag{17}$$

The augmented Lagrangian of (17) is

$$L(\mathbf{X}, \mathbf{V}, \mathbf{D}) = f(\mathbf{X}) + g(\mathbf{V}) + \frac{\mu}{2}||\mathbf{X} - \mathbf{V} + \mathbf{D}||_F^2 \tag{18}$$

where μ is the step size and $\mathbf{D}$ is the scaled dual variable. The steps in the ADMM, for iteration $k+1$, then becomes

$$\begin{aligned} \mathbf{X}^{k+1} =& \arg\min_{\mathbf{X}} \quad L(\mathbf{X}, \mathbf{V}^k, \mathbf{D}^k) \\ =& \left(\mathbf{A}^H\mathbf{A} + \mu\mathbf{I}\right)^{-1}\left(\mathbf{A}^H\mathbf{Y} + \mathbf{V}^k - \mathbf{D}^k\right) \end{aligned} \tag{19}$$

$$\mathbf{V}^{k+1} = \arg\min_{\mathbf{V}} \quad L(\mathbf{X}^{k+1}, \mathbf{V}, \mathbf{D}^k) \tag{20}$$

$$\mathbf{D}^{k+1} = \mathbf{X}^{k+1} - \mathbf{V}^{k+1} + \mathbf{D}^k \tag{21}$$

where (20) may be solved using a soft threshold, i.e., for all m,

$$[\mathbf{V}]_m^{k+1} = \max\left(0, 1 - \frac{\kappa}{||\mathbf{x}||_2}\right)\mathbf{x} \tag{22}$$

with $\mathbf{x} = [\mathbf{X}]_m^k - [\mathbf{V}]_m^k$ and $\kappa = \lambda/\mu$. The choice of the step size $\mu > 0$ will not affect whether the algorithm converges or not, but will influence the convergence time [11]. In all simulations in this paper, μ was set to 1. We coin the resulting algorithm the Sparse Exponential Mode Analysis for 2-D data (SEMA-2D).

5. NUMERICAL EXAMPLES

In order to examine the performance of the proposed algorithm, we illustrate the use of the algorithm for the estimation of the frequency and damping parameters of a signal mimicking 2-D NMR data. Initially the method is compared to the recent non-parametric estimator proposed in [4], which utilize a sparse estimation formulation combined with a zooming procedure to reduce the dimensionality of the dictionary. Figures 2 and 3 illustrate the total RMSE of all the unknown parameters, defined as

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\sum_{k=1}^{K}\left(\theta_{m,k} - \hat{\theta}_{m,k}\right)^2} \tag{23}$$

where $\hat{\theta}_{m,k}$ denotes the estimate of the parameter $\theta_{m,k}$, where $\theta_{m,k}$ denotes either the frequencies or dampings, M the number of Monte-Carlo simulations, for the first frequency and damping coefficients, $f_{1,1}$ and $\beta_{1,1}$ (the performance along the second dimension is similar). As is clear from the figures, the SEMA-2D estimator drastically outperforms the zooming-based method presented in [4]. The shown results have been obtained using 100 Monte-Carlo simulations, using frequencies selected randomly over all frequencies, whereas the damping were (arbitrarily) selected fixed to $\beta_1 = \left[\begin{array}{cc} 0.007, & 0.02 \end{array}\right]$ and $\beta_2 = \left[\begin{array}{cc} 0.008, & 0.01 \end{array}\right]$. The data set contains 100×100 uniformly sampled data points, and the algorithm in [4] has been allowed an initial grid of 100 grid points in each search dimension, and 50 levels of refinement. We proceed to examine how the proposed (non-parametric) SEMA 2-D algorithm compares to the PUMA 2-D algorithm presented in [6], which is a statistically efficient *parametric* estimator, here allowed perfect model order information. Figures 4 and 6 show the performance of the computationally efficient parametric PUMA 2-D estimator as compared to the proposed algorithm for $f_{1,1}$ and $\beta_{1,1}$ (the performance of the second mode is similar), clearly indicating that the proposed non-parametric estimator achieves almost the same performance as the parametric PUMA 2-D. Here, PUMA 2-D has been allowed 10 iterations. Finally, we examine the performance of the proposed algorithm on measured 2-D NMR data from a ^{15}N-HSQC experiment on a Histidine sample acquired at 600 Mhz. Figure 5 illustrates the 2-D periodogram estimate of the examined data set, which consists of 1024×512 uniformly sampled data points, containing, seemingly, 22 modes. Typically, the most interesting aspect of this form of estimates is to estimate the damping coefficient of the weak modes. Figure 7 shows the DFT of the rows of $\mathbf{X}$ corresponding to finding $f_1 = 0.2881$, which can be seen to match the corresponding row in Figure 5, illustrating how the SEMA 2-D algorithm has been able to effectively decouple the search dimensions. Throughout these simulations, λ has been selected such as being the average of the dominant amplitudes and that of the remaining spectrum, where the former is found as the mean of the amplitude of the k_0 largest peaks of the periodogram, the latter is computed as the mean of the remaining periodogram estimate, and with k_0 denoting the number of considered peaks, being equivalent to the model order given to the competing algorithms. It should be stressed that selecting λ in this way does not imply using k_0 as the assumed model order, but is rather only acting as a guideline for how this weighting might be selected. Another approach to find a suitable λ would, e.g., be to use cross-validation [12].

6. CONCLUSION

In this work, we have introduced a sparse decoupling framework for two-dimensional exponentially decaying sinusoids. The method allows the modes to be estimated using a sequence of 1-D searches, while still yielding the same optimal minimum as a full 2-D search would. Furthermore, we propose a computationally efficient ADMM based implementation, drastically reducing the required complexity.

7. ACKNOWLEDGMENT

The authors would like to thank the authors of [6] for providing their implementation of the PUMA 2-D algorithm, as well as Prof. Mikael Akke for providing the NMR data.

8. REFERENCES

[1] J. Liu and X. Liu, "An Eigenvector-Based Approach for Multidimensional Frequency Estimation With Improved Identifiability," *IEEE Transactions on Signal Processing*, vol. 54, pp. 4543–4556, 2006.

[2] Y. Hua, "Estimating Two-Dimensional Frequencies by Matrix Enhancement and Matrix Pencil," *IEEE Transactions on Signal Processing*, vol. 40, no. 9, pp. 2267–2280, September 1992.

[3] J. Sacchini, W. Steedly, and R. Moses, "Two-dimensional Prony modeling and parameter estimation," *IEEE Transactions on Signal Processing*, vol. 41, no. 11, pp. 3127–3137, November 1993.

[4] S. Sahnoun, E. H. Djermoune, and D. Brie, "Sparse Modal Estimation of 2-D NMR Signals," in *38th IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing*, Vancouver, Canada, May 26-31 2013.

[5] S. Rouquette and M. Najim, "Estimation of Frequencies and Damping Factors by Two-Dimensional ESPRIT Type Methods," *IEEE Transactions on Signal Processing*, vol. 49, no. 49, pp. 237–245, January 2001.

[6] F. K. W. Chan, H. C. So, and W. Sun, "Subspace approach for two-dimensional parameter estimation of multiple damped sinusoids," *Signal Process.*, vol. 92, pp. 2172 – 2179, 2012.

[7] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-Order SVD-Based Subspace Estimation to Improve the Parameter Estimation Accuracy in Multidimensional Harmonic Retrieval Problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3198–3213, July 2008.

[8] J. Swärd, S. Adalbjörnsson, and A. Jakobsson, "High Resolution Sparse Estimation of Exponentially Decaying Signals," in *Proceedings of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing*, 2014.

[9] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[10] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Prentice Hall, Upper Saddle River, N.J., 2005.

[11] J. Eckstein and D.P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, pp. 293–318, April 1992.

[12] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society B*, vol. 58, no. 1, pp. 267–288, 1996.