Fast Gridless Estimation of Damped Modes*

Maria Juhlin, Filip Elvander, Johan Swärd, and Andreas Jakobsson
Mathematical Statistics, Lund University, Sweden

***Abstract*—Estimating the parameters of damped modes is generally computationally cumbersome, especially in cases when the model order of the signal is not known *a priori*. In this work, we propose a computationally efficient grid-less sparse reconstruction technique based on the recent SURE estimator developed for sinusoidal signals, extending it as to also allow for damped modes. In contrast to other estimators the proposed method allows for a non-uniform sampling scheme, thus presenting superior performance in applications where uniform sampling is suboptimal, hence broadening the application range. The performance of the proposed method is illustrated using both simulated and real spectroscopic data, clearly showing the improved performance as compared to previous techniques.**

***Index Terms*—Sparse signal analysis, grid-less estimation, damped sinusoids, NMR spectroscopy**

I. Introduction

Signals that may be well modeled as superpositions of exponentially decaying sinusoidal components occur naturally in a wide variety of fields, such as, e.g, nuclear magnetic resonance (NMR) spectroscopy, seismology, and various forms of audio signals, and the problem of estimating the parameters detailing such modes has as a result attracted significant interest in the recent literature (see, e.g., [1]–[9]). Many of the earlier works have focused on parametric techniques that typically impose strong *a priori* assumptions on the structure of the signal, such as assuming a known model order and a uniform sampling scheme [5], [6]. In [10], a subspace technique for magnetic resonance spectroscopy relying on prior information about the signal components was introduced. Although yielding promising results, the technique is computationally cumbersome, and, as with other parametric estimators, the imposition of a certain model structure and *a priori* model order information severely restricts the field of application. A subspace-based method aiming at circumventing the problem of requiring model order information was presented in [9], estimating both the signal parameters and the model orders simultaneously. As such, the estimator is comprised of three sequential steps, resulting in a computationally cumbersome framework. Moreover, since no prior information is required, some thresholding has to be introduced, thus impacting estimation results for lower signal to noise ratios (SNRs).

As assumptions on the expected signal structure are often too restrictive for practical signals, much attention has also focused on developing non-parametric techniques, imposing no or only weak assumptions on the observed signal (see, e.g., [1]–[3]). One such method was proposed in [11], where a computationally efficient estimator based on interpolation of the coefficient of the discrete-time Fourier transform (DFT) was introduced. The suggested method is computationally fast, but only considers a single signal component, restricting the usability of the method. Furthermore, the optimal amount of samples needed will depend heavily on the signal components to be estimated. Similar estimators have been developed in [11], [12], although both methods are restricted to signals containing only a single sinusoid.

Recently, efforts have been made to develop semi-parametric techniques, by imposing a detailed model structure for the decaying modes constituting the signal, but without assuming knowledge of the model order, other than that the signal being sparse in the sense of being expected to contain only a few modes [7], [8], [13], [14]. However, dictionary-based techniques with a fixed grid of signal candidate in general require a fine gridding of the parameter space, in this case frequency and damping, in order to allow for an accurate and sparse signal representation. Thus, in order to avoid off-grid effects such as omitting relevant signal components, a large dictionary is in general required, causing an increase in computational complexity [15]–[17].

Often, there is also a need for allowing for non-uniformly sampled signals, and several recent works have focused on this case (see, e.g., [18]–[20]). Such techniques are relevant in cases when a reduction in the number of samples implies a benefit in terms of time and/or cost, or in applications where acquiring uniformly sampled measurements may be prohibitively costly or even unfeasible. One such application is NMR spectroscopy. In such cases, it is highly beneficial to construct non-uniform sampling patterns that allow for considerable decreases in the sample size, while mitigating the loss in estimation accuracy [18], [19]. However, in order to exploit such sampling schemes, the used estimator needs to be able to make efficient use of the available samples, even if these are non-uniformly spaced.

The SEMA estimator is one example of a semi-parametric method that allows for both non-uniform sampling and that restricts the computational complexity by iteratively refining the estimated frequency and decay coefficients for each active candidate mode [8], [14]. Another promising technique is the so-called SURE estimator that has been developed for sinusoidal data [21], [22]. This estimator also allows for non-uniform sampling and offers a grid-less formulation, wherein the dictionary elements are updated iteratively such that redundant dictionary elements are also removed during the process, speeding up the estimator significantly. In this paper, we extend the SURE estimator to allow for decaying modes,

*This work was supported in part by the Crafoord foundation (Grant no. 20180641), eSSENCE (Grant no 2017-4:2), and the Swedish Research Council (Grant no. 2015-04148).

comparing the performance of the resulting estimator for both real and simulated data to both SEMA and the Cramér-Rao lower bound (CRLB), indicating the optimal performance for any unbiased estimator.

II. Signal model and prior research

Consider a signal $y(t_n) \in \mathbb{C}$ sampled at (possibly non-uniformly spaced) times t_n, for $n = 1, 2, \ldots, N$, that may be well modeled as a superposition of K exponentially decaying complex sinusoidals (cisoids), i.e.,

$$y(t_n) = \sum_{k=1}^{K} \alpha_k e^{2i\pi f_k t_n - \beta_k t_n} + \epsilon(t_n), \tag{1}$$

where α_k, f_k, and β_k denote the complex amplitude, frequency, and damping factor for the kth component, respectively, and where $\epsilon(t_n)$ denotes a circularly symmetric white Gaussian noise. For later reference in the numerical section, we denote the variance of the noise component by σ^2. Several statistically efficient methods have been proposed for estimating the parameters of this model, such as HTLS [1], and extensions of sub-space methods such as MUSIC, ESPRIT, and PUMA [6], [23]. However, although sporting low computational complexity, these methods suffer from the drawback of requiring the model order, i.e., the number of signal components K, to be known, typically requiring the use of some model order selection criteria [23], [24] in order to operate. Also, more fundamentally, these methods are restricted to signals that have been uniformly sampled. This prevents them from being used together with design methods such as [19] and in other applications where signals are sampled non-uniformly.

In order to adress the problem of the unknown model order K, as well as allowing for arbitrarily sampled data, one may pose the estimation of the damped modes as a sparse reconstruction problem. Specifically, by defining a set of signal candidates

$$a(t; \boldsymbol{\theta}_p) = e^{2i\pi f_p t - \beta_p t}, \tag{2}$$

for $p = 1, \ldots, P$, with $\boldsymbol{\theta}_p = \left[\begin{array}{cc} f_p & \beta_p \end{array} \right]^T$, the problem may be relaxed to a linear problem of amplitude estimation, where the amplitude coefficients of the vast majority of the candidate components $a(t; \boldsymbol{\theta}_p)$ are assumed to be zero. By defining

$$\mathbf{y} = \left[\begin{array}{ccc} y(t_1) & \cdots & y(t_n) \end{array} \right]^T \tag{3}$$

$$\boldsymbol{\theta} = \left[\begin{array}{ccc} \boldsymbol{\theta}_1 & \cdots & \boldsymbol{\theta}_p \end{array} \right]^T \tag{4}$$

$$\mathbf{A}_{\boldsymbol{\theta}} = \left[\begin{array}{ccc} \mathbf{a}(\boldsymbol{\theta}_1) & \cdots & \mathbf{a}(\boldsymbol{\theta}_p) \end{array} \right] \tag{5}$$

$$\mathbf{a}(\boldsymbol{\theta}_p) = \left[\begin{array}{ccc} a(t_1; \boldsymbol{\theta}_p) & \cdots & a(t_n; \boldsymbol{\theta}_p) \end{array} \right]^T, \tag{6}$$

this may be posed as

$$\underset{\boldsymbol{\alpha} \in \mathbb{C}^P}{\text{minimize}} \; ||\mathbf{y} - \mathbf{A}_{\boldsymbol{\theta}} \boldsymbol{\alpha}||_2^2 + \lambda ||\boldsymbol{\alpha}||_1, \tag{7}$$

where $\boldsymbol{\alpha}$ denotes the vector of fitted complex amplitudes, and with $\lambda > 0$ denoting a user-defined regularization parameter. The used ℓ_1-norm is included in the minimization in order to encourage sparse solutions. This formulation has been successfully exploited in [8]. However, due to the discretization of the parameter space, the set of frequencies $\{f_p\}_{p=1}^P$ and damping parameters $\{\beta_p\}_{p=1}^P$ must be selected as to yield a fine grid in order to allow for both an accurate and sparse signal representation (see, e.g., [15] for a discussion on this topic). This implies, in the absence of accurate prior knowledge of the values of the signal parameters, that $P \gg K$, thus resulting in potentially high computational load as common solvers, such as, e.g., ADMM [25], have a complexity of $\mathcal{O}(P^3)$.

Striving to alleviate this problem, the use of integrated dictionary elements have been introduced [14], [17], using dictionary elements with frequency domain support of non-zero measure, realized as integrals of elements $a(t; \boldsymbol{\theta}_p)$, together with an iterative dictionary refinement, drastically decreasing the number of dictionary elements needed. As an alternative to such a solution, we will in this paper introduce an iterative estimation algorithm, based on the sinusoidal parameter estimator presented in [21], that throughout the iterations utilize signal candidates of the relevant signal class, i.e., elements $a(t; \boldsymbol{\theta})$, without having to employ a large set of candidates.

III. Proposed Estimator

Modifying the formulation in (7), we will herein consider the minimization problem (see also [21])

$$\min_{\boldsymbol{\alpha}, \boldsymbol{\theta}} \|\mathbf{y} - \mathbf{A}_{\boldsymbol{\theta}} \boldsymbol{\alpha}\|_2^2 + \lambda T(\boldsymbol{\alpha}), \tag{8}$$

where the sparsity inducing function $T(\boldsymbol{\alpha})$ is defined as

$$T(\boldsymbol{\alpha}) = \sum_{m=1}^{M} \log(|\alpha_m|^2 + \eta), \tag{9}$$

and where η is a small positive number, added to ensure well-posedness as well as numerical stability. It may be noted that the non-linear signal parameter vector, $\boldsymbol{\theta}$, appears explicitly in the optimization problem, allowing for the interpretation of an initial set of M signal candidates that are allowed to vary over the optimization procedure. This particular feature allows for choosing $M \ll P$ (although larger than K), while still allowing for a sparse and accurate signal reconstruction. Here, the sparsity inducing ℓ_1-norm has been replaced by a non-convex function lower-bounding it. This may be compared to the so-called reweighting schemes often applied to problems of the form (7) as to yield sparser solutions than those induced by the ℓ_1-norm alone [26]. Being non-convex, the criterion in (8) may not be readily minimized. However, local minimima may be found by applying a majorization-minimization scheme, where the objective function in (8) at iteration j is majorized by the surrogate function (see also [21], [22])

$$\begin{aligned} S\left(\boldsymbol{\alpha}, \boldsymbol{\theta} \mid \boldsymbol{\alpha}^{(j-1)}\right) = \lambda \left(Q(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}^{(j-1)}) + T(\boldsymbol{\alpha}^{(j-1)}) \right) \\ + \|\mathbf{y} - \mathbf{A}_{\boldsymbol{\theta}} \boldsymbol{\alpha}\|_2^2 , \end{aligned} \tag{10}$$

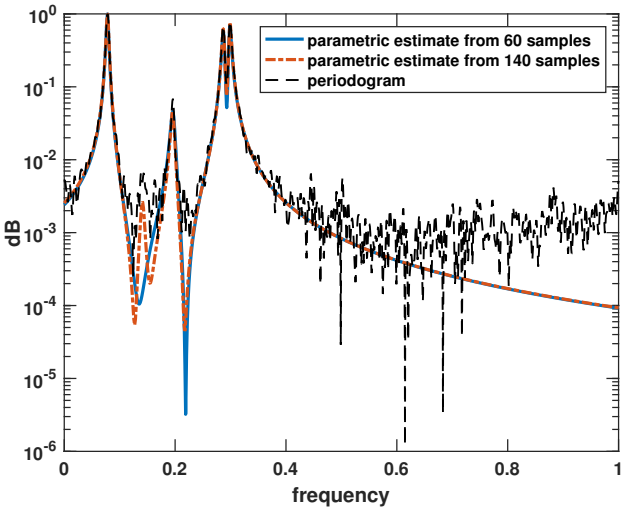

Fig. 1. Squared magnitude of the DFT for a measured NMR signal consisting of four signal components, computed from 220 uniformly spaced samples. Superpositioned is the squared magnitude of the theoretical Fourier transform, computed from parameter estimates obtained using the proposed method utilizing 60 and 140 non-uniformly spaced signal samples. The samples are selected as to minimize the Cramér-Rao lower bound for the frequency and damping parameters.

with

$$Q(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}^{(j-1)}) = -M + \sum_{m=1}^{M} \frac{|\alpha_m|^2 + \eta}{\left|\alpha_m^{(j-1)}\right|^2 + \eta}, \qquad (11)$$

where it may be noted that

$$Q(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}^{(j-1)}) + T(\boldsymbol{\alpha}^{(j-1)}) \geq T(\boldsymbol{\alpha}) \qquad (12)$$

and $Q(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}) \equiv 0$. In fact (for details, see [21]), performing the iterations

$$\boldsymbol{\theta}^{(j)} = \arg\min_{\boldsymbol{\theta}} g(\boldsymbol{\theta}; \mathbf{D}^{(j-1)}), \qquad (13)$$

where

$$g(\boldsymbol{\theta}; \mathbf{D}^{(j-1)}) = -\mathbf{y}^H \mathbf{A}_{\boldsymbol{\theta}} \left(\mathbf{A}_{\boldsymbol{\theta}}^H \mathbf{A}_{\boldsymbol{\theta}} + \lambda \mathbf{D}^{(j-1)} \right)^{-1} \mathbf{A}_{\boldsymbol{\theta}}^H \mathbf{y} \qquad (14)$$

$$\mathbf{D}^{(j-1)} = \mathrm{diag}\left(\frac{1}{\left|\alpha_1^{(j-1)}\right|^2 + \eta} \quad \cdots \quad \frac{1}{\left|\alpha_M^{(j-1)}\right|^2 + \eta} \right), \qquad (15)$$

produces a sequence $\left\{\boldsymbol{\theta}^{(j)}\right\}_j$ yielding a non-increasing sequence of values of the objective function in (8). It may be noted that the complex amplitudes may be computed as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}^{(j)}) = \left(\mathbf{A}_{\boldsymbol{\theta}^{(j)}}^H \mathbf{A}_{\boldsymbol{\theta}^{(j)}} + \lambda \mathbf{D}^{(j-1)} \right)^{-1} \mathbf{A}_{\boldsymbol{\theta}^{(j)}}^H \mathbf{y}. \qquad (16)$$

As the function g is differentiable, one may perform the update in (13) using block-wise coordinate descent via gradient descent, i.e., performing minimization with respect to $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M$ in sequence. The computationally dominating step in computing (13) is computing the gradient $\nabla_{\boldsymbol{\theta}} g$, requiring $\mathcal{O}(M^3)$ operations. However, as the presented framework allows for having M small, this is not a cumbersome operation. Also, in order to speed up the computations, one may sequentially identify and discard components whose amplitude coefficients tend to zero, i.e., $\lim_{j\to\infty} \alpha_m^{(j)} = 0$. Although one could envision implementing an online identification scheme along the lines in [27], [28] by considering the geometry of a corresponding dual problem, we here propose to simply discard a component whenever $\left|\alpha_m^{(j)}\right| < \xi$ for a non-negative threshold ξ.

As noted, sparse, grid-based methods in general require a fine grid in order for an accurate signal representation to be attainable. Also, when decreasing the number of components used to grid the parameter space, one may also be exposed to the risk of not being able to detect signal components. Considering the continuous-time representation of the signal components $a(t; \boldsymbol{\theta})$, we note that for $\beta > 0$, these signals are in the Hilbert space L_2^+, i.e., the space of square integrable functions on the positive real line. Thus, the maximal inner product for a signal component $a(t; \boldsymbol{\theta})$ with any dictionary element $a(t; \boldsymbol{\theta}_m)$ is

$$\begin{aligned} h(\boldsymbol{\theta}) &\triangleq \max_m \left| \langle a(t; \boldsymbol{\theta}_m), a(t; \boldsymbol{\theta}) \rangle_{L_2^+} \right| \\ &= \max_m \left| \int_0^\infty e^{-i2\pi f_m t - \beta_m t} e^{i2\pi f - \beta t} dt \right| \\ &= \max_m \frac{1}{\sqrt{4\pi^2 (f - f_m)^2 + (\beta + \beta_m)^2}}. \end{aligned} \qquad (17)$$

As a criteria including the ℓ_2-norm, the discrete time analog to L_2, yield optimality criteria related to the inner product in ℓ_2, it may be noted that narrow-band dictionary components may yield a risk of not detecting all signal components of multi-component signals. Specifically, consider the quantity

$$\sup_{\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}} \frac{h(\boldsymbol{\theta})}{h(\tilde{\boldsymbol{\theta}})}, \qquad (18)$$

i.e., the ratio of the maximal inner products corresponding to two signal components $a(t; \boldsymbol{\theta})$ and $a(t; \tilde{\boldsymbol{\theta}})$. Then, if both these signal components have damping parameters close to zero and the dictionary is initialized with $\beta_m = 0$, for $m = 1, \ldots, M$, one suffers a high risk of not detecting the component with parameter $\tilde{\boldsymbol{\theta}}$ should this have a frequency far from any of the candidate frequencies f_m if, at the same time, the frequency component of the parameter $\boldsymbol{\theta}$ is close to some grid point. This is caused by the quantity in (18) tending to infinity. To safe-guard against such a problem, while still allowing for using few candidate components, we propose initializing the dictionary components with $\beta_m > 0$ as to somewhat equalize the magnitudes of the correlation between possible signal components and the dictionary elements. This also adds robustness for the gradient descent scheme; terms on the form $\mathbf{A}_{\boldsymbol{\theta}}^H \mathbf{y}$ appear in the expression of $\nabla_{\boldsymbol{\theta}} g$. Thus, initializing $\beta_m > 0$ prevents the gradient practically vanishing in some directions due to coarse gridding in the frequency dimension.

As the regularization term T may induce bias for the estimates of the damping parameters β_m, we, in order to

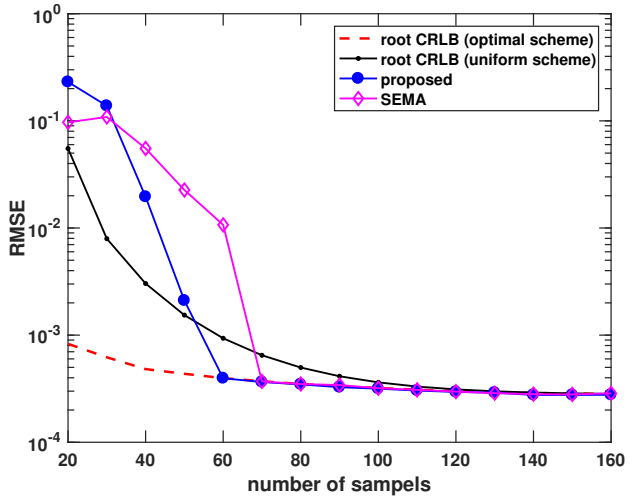

Fig. 2. Root mean squared error for the frequency parameters f for the signal displayed in Figure 1, aggregated over the four signal components, for the proposed estimator, as well as the SEMA estimator, when varying the number of signal samples included in the estimation. The samples are selected as to minimize the Cramér-Rao lower bound for the frequency and damping parameters.

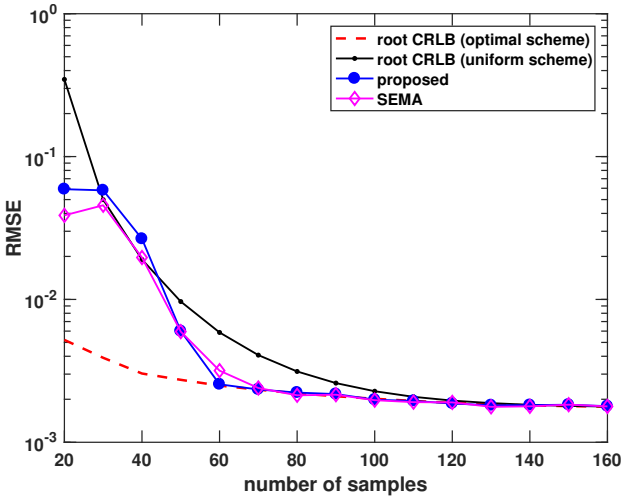

Fig. 3. Root mean squared error for the damping parameters β for the signal displayed in Figure 1, aggregated over the four signal components, for the proposed estimator, as well as the SEMA estimator, when varying the number of signal samples included in the estimation. The samples are selected as to minimize the Cramér-Rao lower bound for the frequency and damping parameters.

yield statistically efficient estimates, propose to perform a non-linear least squares (NLS) search in a small neighborhood of the obtained damping estimates. That is, given an estimate $\hat{\boldsymbol{\theta}}$, with corresponding vectors of estimated frequencies $\hat{\mathbf{f}}$ and dampings $\hat{\boldsymbol{\beta}}$, the damping parameters are updated as the solution to

$$\underset{\boldsymbol{\alpha}\,,\,\boldsymbol{\theta}\in\Theta(\hat{\boldsymbol{\theta}})}{\text{minimize}} \ \|\mathbf{y}-\mathbf{A}_{\boldsymbol{\theta}}\boldsymbol{\alpha}\|_2^2 \tag{19}$$

where $\Theta(\hat{\boldsymbol{\theta}})$ is a neighborhood of $\hat{\boldsymbol{\theta}}$ on the form

$$\Theta(\hat{\boldsymbol{\theta}}) = \left\{\boldsymbol{\theta} \mid \mathbf{f}=\hat{\mathbf{f}}, \left\|\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right\|_\infty \leq \Delta\right\}, \tag{20}$$

i.e., the frequency parameters are fixed to those corresponding to $\hat{\boldsymbol{\theta}}$, whereas any damping parameter is allowed to deviate at maximum Δ from the initial damping estimates, where $\Delta \ll 1$. The proposed estimator is summarized in Algorithm **??**[1].

IV. Numerical examples

In this section, we demonstrate the proposed method's ability to produce statistically efficient estimates in realistic estimation scenarios, also leveraging the possibility to use non-uniformly sampled data. Specifically, we consider a measured NMR signal consisting of four damped modes, where the data consists of 220 uniformly sampled measurements. The squared magnitude of the DFT, i.e., the periodogram, of the entire signal is displayed in Figure 1. As may be noted, two of the signal components are closely spaced in frequency. Superimposed in the figure is the corresponding theoretical squared magnitude of the Fourier transform, computed from the estimated parameters yielded by the proposed method when using only 60 and 140 out of the 220 available signal samples. These samples were selected using the method presented in [19], specifically optimizing the sampling scheme to optimize the estimation accuracy for the frequency and damping parameters given the restricted number of samples. As is clear from the figure, the proposed method is able to accurately estimate the spectrum of the signal even with as few as 60 samples. The frequency and damping parameters, when using a normalized sampling frequency $f_s = 1$, are $(f_1, f_2, f_3, f_4) = (0.0781, 0.1960, 0.2873, 0.2987)$ and $(\beta_1, \beta_2, \beta_3, \beta_4) = (1.53, 2.36, 1.73, 1.70)\times 10^{-2}$, respectively. Furthermore, the complex amplitudes and noise variance are

$$(|\alpha_1|\,,|\alpha_2|\,,|\alpha_3|\,,|\alpha_4|) = (7.09, 2.31, 5.98, 6.33)\times 10^4 \tag{21}$$

$$(\angle\alpha_1, \angle\alpha_2, \angle\alpha_3, \angle\alpha_4) = -(0.102, 0.283, 0.173, 0.149) \tag{22}$$

$$\sigma^2 = 2.55\times 10^7, \tag{23}$$

yielding an amplitude-to-noise ratio

$$\psi \triangleq 10\log_{10}\left(\frac{\sum_{k=1}^4 |\alpha_k|^2}{\sigma^2}\right) = 27.12 \text{ dB}. \tag{24}$$

Throughout these examples, we use $M = 10$ initial signal candidate components, initialized with $\beta_m = 0.005$, for $m = 1, \ldots, M$. Also, we set $\Delta = 10^{-2}$ and $\xi = 10^{-2}\cdot \max_k |\alpha_k|$.

A. Varying the number of signal samples

Using the obtained signal parameters, we perform Monte Carlo simulations, where in each simulation, the signal is generated using the frequencies, dampings, and complex amplitudes, obtained from the real data as detailed above, with each simulation being corrupted by a realization of a circularly symmetric white Gaussian noise with variance, i.e., σ^2, equal to the variance estimated from the measured data.

[1] An implementation of the algorithm will be provided online on the authors' website upon publication.

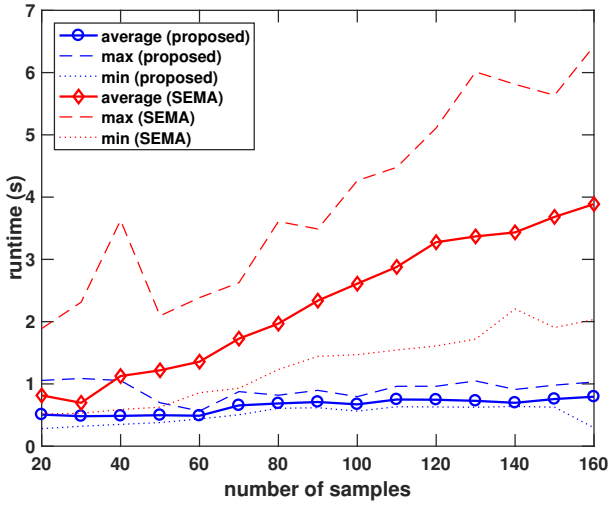

Fig. 4. Average, maxmimal, and minimal runtimes for the proposed estimator, as well as the SEMA estimator for the estimation scenario corresponding to Figures 2 and 3.

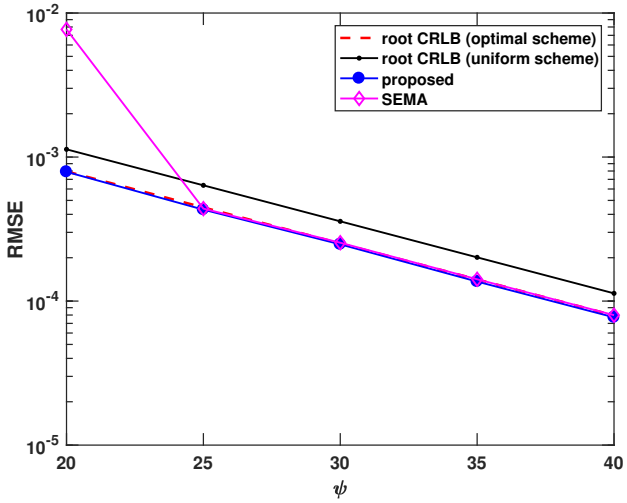

Fig. 5. Root mean squared error for the frequency parameters f for the signal displayed in Figure 1, aggregated over the four signal components, for the proposed estimator, as well as the SEMA estimator, when varying the amplitude-to-noise ratio, ψ. For both estimators, 80 non-uniformly spaced signal samples are used. The samples are selected as to minimize the Cramér-Rao lower bound for the frequency and damping parameters.

To demonstrate the usefulness of allowing for non-uniformly sampled data, we in the simulations vary the number of signal samples used by the proposed estimator. These samples are selected as detailed in [19], where the estimation accuracy for the frequency and damping parameters are again prioritized. The resulting root mean squared errors (RMSE) are presented in Figures 2 and 3 for, respectively, the frequency and the damping parameters. It may here be noted that, as to increase readability, the mean squared errors for all parameters have been summed. For reference, the (summed) root of the CRLB for the corresponding parameters are also plotted. As can be seen, the proposed method indeed attains the CRLB when

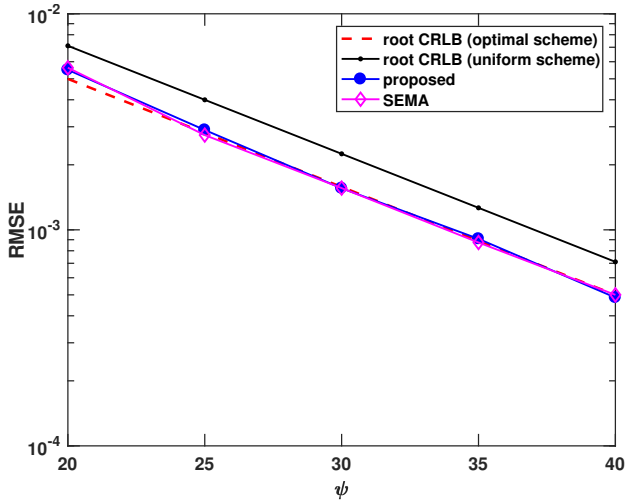

Fig. 6. Root mean squared error for the damping parameters β for the signal displayed in Figure 1, aggregated over the four signal components, for the proposed estimator, as well as the SEMA estimator, when varying the amplitude-to-noise ratio, ψ. For both estimators, 80 non-uniformly spaced signal samples are used. The samples are selected as to minimize the Cramér-Rao lower bound for the frequency and damping parameters.

given 60 or more signal samples. As a comparison, the figures also include the root CRLB corresponding to the estimation scenario where one, instead of selecting the samples in a theoretically optimal fashion, instead use the corresponding number of uniform samples selected in the beginning of the signal. It may be noted that the selection method in [19] yields a considerably lower CRLB than that given by uniform sampling, and that the proposed method is able to exploit this improvement in achieving the theoretical estimation accuracy.

Also presented are corresponding results from performing estimation using the SEMA method [8], which, as may be noted also attains the CRLB, although for 70 samples or more. However, the computation time for SEMA was considerably higher. Figure 4 displays the average, maximal, and minimal computation time for the proposed method, as well as SEMA, when running this experiment. As may be noted, the computation time for SEMA increases as more samples are included in the estimation. This is caused by SEMA relying on a complete NLS search for finding the damping parameters, resulting in the computational complexity scaling significantly with the number of signal samples. In contrast, the proposed method yields quite accurate estimates of the damping parameters already in the gradient ascent step, requiring only a small NLS search to yield statistically efficient estimates.

B. *Varying* ψ

Further, we consider the performance of the proposed estimator when applied to signals with varying levels of noise. Again, we perform Monte Carlo simulations in the same manner as described above using the same signal parameters. In contrast with the previous experiment, we here fix the number of samples to 80, selected as before using the methodology in [19], and instead vary the noise variance σ^2. This variation

is presented in terms of ψ as defined in (24); specifically, ψ is varied between 20 and 40 dB. It may here be noted that selecting ψ below 20 dB will result in the noise floor overpowering the weakest of the four signal components. Indeed, neither the proposed estimator nor SEMA was able to correctly recover the signal components for $\psi < 20$ dB. For ψ in the range 20 to 40 dB, the obtained RMSE for the frequency and damping are shown in Figures 5 and 6, respectively. As may be seen, both estimators are able to attain the CRLB also in this scenario. It may also be noted that both estimators are able to improve on the theoretically best performance obtainable when instead using the first 80 samples of the signal.

V. Conclusions

In this paper, we have proposed a grid-less estimator of the parameters detailing the type of exponentially decaying modes commonly appearing in spectroscopy applications. The proposed method allows for arbitrarily sampled data, thus making it applicable to application scenarios wherein uniform sampling is costly, or even unfeasible. As demonstrated in the numerical examples, the proposed estimator is able to produce statistically efficient parameter estimates in realistic spectroscopy scenarios, also displaying a competitive decrease in computation time compared to state-of-the-art sparse estimators.

References

[1] L. Vanhamme, T. Sundin, P. van Hecke, and S. van Huffel, "MR spectroscopy quantitation: a review of time-domain methods," *NMR in Biomedicine*, vol. 14, no. 4, pp. 233–246, 2001.

[2] P. Stoica and T. Sundin, "Nonparametric NMR Spectroscopy," *J. Magn. Reson.*, vol. 152, pp. 57–69, 2001.

[3] E. Gudmundson, P. Stoica, J. Li, A. Jakobsson, M. D. Rowe, J. A. S. Smith, and J. Ling, "Spectral Estimation of Irregularly Sampled Exponentially Decaying Signals with Applications to RF Spectroscopy," *J. Magn. Reson.*, vol. 203, no. 1, pp. 167–176, March 2010.

[4] Y. Li, J. Razavilar, and K. J. R. Liu, "A High-Resolution Technique for Multidimensional NMR Spectroscopy," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 1, pp. 78–86, 1998.

[5] 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.

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

[8] J. Swärd, S. I. Adalbjörnsson, and A. Jakobsson, "High Resolution Sparse Estimation of Exponentially Decaying N-dimensional Signals," *Elsevier Signal Processing*, vol. 128, pp. 309–317, Nov 2016.

[9] Z. Zhou, H. C. So, and M. G. Christensen, "Parametric Modeling for Damped Sinusoids from Multiple Channels," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 3895–3907, Aug. 2013.

[10] P. Stoica, Y. Selén, N. Sandgren, and S. Van Huffel, "Using prior knowledge in SVD-based parameter estimation for magnetic resonance spectroscopy—the ATP example," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1568–1578, September 2004.

[11] E. Aboutanios, "Estimation of the Frequency and Decay Factor of a Decaying Exponential in Noise," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 501–509, Feb 2010.

[12] B. G. Quinn, "Estimating parameters in noisy low frequency exponentially damped sinusoids and exponentials," in *41st IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Shanghai, China, March 2016, pp. 4298–4302.

[13] J. J. Fuchs, "On the Use of Sparse Representations in the Identification of Line Spectra," in *17th World Congress IFAC*, Seoul, July 2008, pp. 10225–10229.

[14] M. Jälmby, J. Swärd, F. Elvander, and A. Jakobsson, "Computationally Efficient Estimation of Multi-dimensional Damped Modes Using Sparse Wideband Dictionaries," in *26th European Signal Processing Conference*, Rome, Italy, 2018, Accepted for publication.

[15] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to Basis Mismatch in Compressed Sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182 –2195, May 2011.

[16] P. Stoica and P. Babu, "Sparse Estimation of Spectral Lines: Grid Selection Problems and Their Solutions," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 962–967, Feb. 2012.

[17] M. Butsenko, J. Swärd, and A. Jakobsson, "Estimating Sparse Signals Using Integrated Wideband Dictionaries," *IEEE Trans. Signal. Process*, 2018, Accepted for publication.

[18] S. P. Chepuri and G. Leus, "Sparsity-Promoting Sensor Selection for Non-Linear Measurement Models," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 684–698, February 2015.

[19] J. Swärd, F. Elvander, and A. Jakobsson, "Designing sampling schemes for multi-dimensional data," *Elsevier Signal Processing*, vol. 150, pp. 1–10, Sep. 2018.

[20] J. Swärd, F. Elvander, and A. Jakobsson, "Designing Optimal Sampling Schemes," in *25th European Signal Processing Conference*, Aug 28 - Sep 2 2017.

[21] J. Fang, F. Wang, Y. Shen, H. Li, and R. S. Blum, "Super-Resolution Compressed Sensing for Line Spectral Estimation: An Iterative Reweighted Approach," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4649–4662, September 2016.

[22] J. Swärd, H. Li, and A. Jakobsson, "Off-grid fundamental frequency estimation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 26, no. 2, pp. 296–303, 2018.

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

[24] P. Stoica and Y. Selén, "Model-order Selection — A Review of Information Criterion Rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, July 2004.

[25] 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.

[26] E. J. Candès, M. B. Wakin, and S. Boyd, "Enhancing Sparsity by Reweighted l_1 Minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, Dec. 2008.

[27] O. Fercoq, A. Gramfort, and J. Salmon, "Mind the duality gap: safer rules for the Lasso," in *32nd International Conference on Machine Learning*, Lille, France, 2015.

[28] A. Bonnefoy, V. Emiya, L. Ralaivola, and R. Gribonval, "Dynamic Screening: Accelerating First-Order Algorithms for the Lasso and Group-Lasso," *IEEE Trans. Signal. Process*, vol. 63, no. 19, pp. 5121–5132, Oct. 2015.