2018 26th European Signal Processing Conference (EUSIPCO)

Least Squares and Maximum Likelihood Estimation of Mixed Spectra

Johan Brynolfsson, Johan Swärd, Andreas Jakobsson, and Maria Sandsten
Mathematical Statistics, Lund University, Sweden

***Abstract*—In this paper, we propose a novel 1-D spectral estimator for signals with mixed spectra. The proposed method is partly based on the recently introduced smooth spectral estimator LIMES, in which the smoothness is accounted for by assuming linearity within predefined segments of the spectrum. The proposed method utilizes this formulation but also allows segments to change size to better estimate the spectrum, thereby allowing for the estimation of spectra that are neither completely smooth or sparse in frequency, but rather contains a mixture of such components. Using simulated data, we illustrate the performance of the proposed estimator, comparing to other recent spectral estimation techniques.**

***Index Terms*—spectral estimation, time-series, LIMES, covariance-fitting**

I. Introduction

Spectral estimation has long been an important tool in both the analysis and classification of a wide variety of signals. Classical techniques based on the Fourier transform are popular and widely used, often in combination with various windowing techniques in order to improve the estimates. Such methods often suffer from poor resolution and/or high variance [1]. As an alternative, one may use a variety of parametric estimation techniques, generally imposing strong model assumptions on the signal. Such estimators may render highly accurate and reliable estimates, if both the model structure and model order of the signal are well known, but may also fail to provide meaningful estimates if this is not the case. Recently, significant efforts have been made to develop methods that aim at exploiting the strengths of the parametric estimators, but without requiring a priori knowledge of the model order of the signal. This is often done by assuming that the signals of interest are sparse in frequency (see, e.g., [2]–[7]). Similar efforts can be made for signals with smooth spectra, although this problem has attracted less attention in the literature, especially for the case when the signal is irregularly sampled.

A notable exception is the maximum-likelihood based estimator for smooth spectra, denoted LIMES, introduced in [8], [9]. This estimator models the spectral smoothness using a piece-wise linear model, and proceeds to formulate a sparse estimation technique to model such signals. The LIMES algorithm is iterative and requires a reasonably accurate initial estimate in order to render good results, which in [8] was provided by the Daniell's method (DAM) [10]. The LIMES estimator is formed using a set of grid points, usually equidistant, in the frequency domain, wherein the spectra is assumed to be well modelled as being piece-wise linear. Based on these points, a transformation from the frequency domain to the covariance domain is performed, from which a covariance fitting problem may be formulated and solved to yield the smooth spectral estimate. As shown in [8], the method is able to accurately model smooth spectra, and the technique was later extended to allow for 2-D signals and for time varying signals with smooth spectra in [11], [12], where also a fast approximative solver based on LIMES was introduced.

In this paper, we further extend upon these ideas, examining signals with mixed spectra, containing both smooth components and strong peaks, such as signals containing a mixture of sinusoidal components and ARMA-processes. To this end, we generalize the LIMES estimator such that the estimator determines an optimal set of grid points used to form the piecewise linear segments in the frequency domain. To allow for a computationally efficient implementation, an initial estimate of the spectrum is formed by solving a least-squares (LS) optimization problem using the same piece-wise linearity structure assumed in [8], but including the set of grid points in the optimization problem as variables. Solving the LS problem yields initial estimates of the grid points and the corresponding spectrum. These initial estimates are then refined using a generalized version of LIMES that also allows for the grid points to be part of the optimization. This allows the proposed method to better capture non-smooth parts of the spectrum and correctly represent a larger range of spectra. We evaluate the method on simulated data, comparing it to the original LIMES estimator as well as other state-of-the-art spectral estimators. The proposed method is shown to outperform earlier related estimators when the signal contains a mixture of smooth and non-smooth spectral components.

II. Spectral estimation

We consider stationary signals with unknown spectral densities, thus without any prior knowledge on the shape of the spectra. To derive our method, we start with forming the transformation matrix $\mathcal{C}$ for smooth spectra and then show how one may use this formulation to also allow for non-smooth spectral content.

A. Piece-wise linear spectrum

Consider a stationary signal $y(t)$, for $t = t_0, \ldots, t_{N-1}$, that belongs to a second order stationary process with a

*This work was supported in part by the Swedish Research Council and the Swedish strategic research programme eSSENCE.

smooth spectrum that is band-limited to a band $|B| \leq \pi$. In order to form the smooth spectral estimate, we make use of the transformation scheme proposed in [8], wherein it was suggested that the smoothness of the spectrum could be accounted for by assuming that the spectrum is approximately linear over a user-defined set of segments. In particular, we assume that the spectral content between frequencies $-B$ and B may be divided into M segments, and that any point inside such a segment, say segment k, may be well modeled as

$$\phi(\omega) = \frac{\omega - \omega_k}{\Delta_k}\phi_{k+1} + \frac{\omega_{k+1} - \omega}{\Delta_k}\phi_k \tag{1}$$

for $\omega \in [\omega_k, \omega_{k+1})$, where $\Delta_k = \omega_{k+1} - \omega_k$ is the length of the kth segment. For any second order stationary process, it holds that the covariance function is the Fourier transform of the spectral function, i.e.,

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(\omega) e^{i\omega\tau} d\omega \tag{2}$$

where τ is the time lag. Including the assumed bandlimitation and by defining $\boldsymbol{\omega} = \begin{bmatrix} \omega_1 & \ldots & \omega_{M+1} \end{bmatrix}^T$, as a vector containing the frequency vertexes, as well as the corresponding vector $\Phi = \begin{bmatrix} \phi_1 & \ldots & \phi_{M+1} \end{bmatrix}^T$, containing the spectral densities at the vertexes, and by inserting (1) in (2), one obtains

$$R(\tau, \Phi, \boldsymbol{\omega}) = \frac{1}{2\pi}\int_{-B}^{B} \sum_{k=1}^{M} \mathbf{1}_{\Omega_k}(\omega) \left(\frac{\omega - \omega_k}{\Delta_k}\phi_{k+1} + \frac{\omega_{k+1} - \omega}{\Delta_k}\phi_k \right) e^{i\omega\tau} d\omega \tag{3}$$

where $\mathbf{1}_{\Omega_k}(\omega)$ denotes the indicator function of that ω is in $\Omega_k = [\omega_k, \omega_{k+1})$. As all sets Ω_k are assumed to be disjoint, the integral in (3) may be expressed as a sum of integrals

$$R(\tau, \Phi, \boldsymbol{\omega}) = \frac{1}{2\pi}\sum_{k=1}^{M} \int_{\omega_k}^{\omega_{k+1}} \left(\frac{\omega - \omega_k}{\Delta_k}\phi_{k+1} + \frac{\omega_{k+1} - \omega}{\Delta_k}\phi_k \right) e^{i\omega\tau} d\omega \tag{4}$$

By introducing

$$F_{k+1}(\tau, \boldsymbol{\omega}) = \frac{1}{2\pi}\int_{\omega_k}^{\omega_{k+1}} \frac{\omega - \omega_k}{\Delta_k} e^{i\omega\tau} d\omega \tag{5}$$

$$G_k(\tau, \boldsymbol{\omega}) = \frac{1}{2\pi}\int_{\omega_k}^{\omega_{k+1}} \frac{\omega_{k+1} - \omega}{\Delta_k} e^{i\omega\tau} d\omega \tag{6}$$

one may express (4) as

$$R(\tau, \Phi, \boldsymbol{\omega}) = \sum_{k=1}^{M} F_{k+1}(\tau, \boldsymbol{\omega})\phi_{k+1} + G_k(\tau, \boldsymbol{\omega})\phi_k \tag{7}$$

The functions $F_k(\tau, \boldsymbol{\omega})$ and $G_k(\tau, \boldsymbol{\omega})$ may then be calculated as (see [8] for details)

$$F_k(\tau, \boldsymbol{\omega}) = \begin{cases} \Delta_{k-1}/4\pi & \text{if } \tau = 0 \\ \frac{e^{i\omega_k\tau}}{2\pi}\left(\frac{1}{i\tau} + \frac{1}{\tau^2\Delta_{k-1}}\left(1 - e^{-i\Delta_{k-1}\tau}\right)\right) & \text{if } \tau \neq 0 \end{cases} \tag{8}$$

$$G_k(\tau, \boldsymbol{\omega}) = \begin{cases} \Delta_k/4\pi & \text{if } \tau = 0 \\ \frac{e^{i\omega_k\tau}}{2\pi}\left(-\frac{1}{i\tau} - \frac{1}{\tau^2\Delta_k}\left(e^{i\Delta_k\tau} - 1\right)\right) & \text{if } \tau \neq 0 \end{cases} \tag{9}$$

Further, by defining

$$C_k(\tau, \boldsymbol{\omega}) = \begin{cases} G_1(\tau, \boldsymbol{\omega}) & \text{if } k = 1 \\ F_k(\tau, \boldsymbol{\omega}) + G_k(\tau, \boldsymbol{\omega}) & \text{if } k = 2, .., M \\ F_{k+1}(\tau, \boldsymbol{\omega}) & \text{if } k = M+1 \end{cases} \tag{10}$$

a piece-wise linear version of the Fourier transform may be expressed as

$$R(\tau, \Phi, \boldsymbol{\omega}) = \sum_{k=1}^{M+1} C_k(\tau, \boldsymbol{\omega})\phi_k \tag{11}$$

If the observed signal has N samples, R and C_k, with $k = 1...M+1$, will be $N \times N$ matrices. To simplify the notation, we henceforth drop the τ.

III. Proposed spectral estimation

Using the piece-wise linear formulation, multiple different spectral estimators may be formed. In this paper, we will introduce two different estimators, one being based on Least squares (LS) and the other on a maximum likelihood (ML) formulation. The former estimator is computationally cheaper than the latter, although it is not as accurate. As a result, it may be used to form an initial estimator for the ML approach, which may then be used to refine the quality of the estimate.

A. Least squares estimator

An LS estimate of both the spectral density as well as the position of the frequency vertexes may be formulated as

$$\left\{\hat{\Phi}_{LS}, \hat{\boldsymbol{\omega}}\right\} = \arg\min_{\Phi, \boldsymbol{\omega}} L(\Phi, \boldsymbol{\omega}) \tag{12}$$

where

$$L(\Phi, \boldsymbol{\omega}) = \left\| \hat{R} - \sum_{k=1}^{M+1} C_k(\boldsymbol{\omega})\phi_k \right\|_F^2 \tag{13}$$

$$= \left\| \hat{R}_{vec} - \boldsymbol{\mathcal{C}}(\boldsymbol{\omega})\Phi \right\|_2^2 \tag{14}$$

with $\|\cdot\|_F$ denoting the Frobenius norm, and $\hat{R}$ the sample covariance matrix. Further, $\hat{R}_{vec} = \text{vec}(\hat{R})$, where $\text{vec}(\cdot)$ denotes the vectorization operator, and $\boldsymbol{\mathcal{C}} = [\text{vec}(C_1), \ldots, \text{vec}(C_{M+1})]$. For a given frequency vector $\boldsymbol{\omega}$, the LS estimate is thus

$$\hat{\Phi}_{LS} = [\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})^* \boldsymbol{\mathcal{C}}(\boldsymbol{\omega})]^{-1} \boldsymbol{\mathcal{C}}(\boldsymbol{\omega})^* \hat{R}_{vec} \tag{15}$$

Given an estimate $\hat{\Phi}_{LS}$, the gradients of the LS cost function may then be calculated with respect to $\boldsymbol{\omega}$ as shown in the Appendix. These are then used to update the frequency points by taking incremental steps in the gradient direction. The step size taken may be chosen in multiple ways. Here, we begin with an empirically chosen starting value $s_i = 0.01/\frac{\partial L(\Phi_i, \boldsymbol{\omega})}{\partial \boldsymbol{\omega}}$ and increase or decrease this depending on if the previous step generated a better or worse cost function. Between each gradient step, the spectral estimate is updated using the new frequency evaluation points. The resulting algorithm, termed CASED (Covariance fitting Approach for Spectral Estimation with gradient Descent), is presented in Algorithm 1.

ISBN 978-90-827970-1-5 © EURASIP 2018

Algorithm 1 CASED

$\boldsymbol{\omega}_0 = \text{linspace}(-B, B, M+1)$
$\hat{R}(n,m) = \frac{1}{N}\sum_{k=n-m}^{N} y(k)y^*(k-n+m)$
$\boldsymbol{\mathcal{C}}_0 \leftarrow [\text{vec}\,(C_1(\boldsymbol{\omega}_0)), \ldots, \text{vec}\,(C_{M+1}(\boldsymbol{\omega}_0))]$
for i = 1, ..., until convergence **do**
  $\hat{\Phi}_i \leftarrow (\boldsymbol{\mathcal{C}}(\boldsymbol{\omega}_{i-1})\boldsymbol{\mathcal{C}}(\boldsymbol{\omega}_{i-1}))^{-1}\boldsymbol{\mathcal{C}}(\boldsymbol{\omega}_{i-1})\hat{R}_{vec}$
  $\boldsymbol{\omega}_i \leftarrow \boldsymbol{\omega}_{i-1} + s_i\frac{\partial L(\Phi_i,\boldsymbol{\omega})}{\partial\boldsymbol{\omega}}$
  $\boldsymbol{\mathcal{C}}_i \leftarrow [\text{vec}\,(C_1(\boldsymbol{\omega}_i)), \ldots, \text{vec}\,(C_{M+1}(\boldsymbol{\omega}_i))]$
  if $L(\Phi_i, \boldsymbol{\omega}_i) < L(\Phi_i, \boldsymbol{\omega}_{i-1})$ **then**
    $s_{i+1} \leftarrow s_i/1.5$
  else
    $s_{i+1} \leftarrow 1.25 s_i$
    $\boldsymbol{\omega}_i \leftarrow \boldsymbol{\omega}_{i-1}$
  end if
end for

B. Maximum likelihood estimator

After finding suitable initial values for Φ and $\boldsymbol{\omega}$ using the LS method in Algorithm 1, these may be used as starting values for finding a solution to the ML problem formulation. Defining the negative log likelihood as

$$f(\Phi,\boldsymbol{\omega}) = \log\left(\det|R(\Phi,\boldsymbol{\omega})|\right) + \boldsymbol{y}^* R(\Phi,\boldsymbol{\omega})^{-1}\boldsymbol{y}$$

where $\boldsymbol{y}$ is the observed signal, yields the ML estimator

$$\left\{\hat{\Phi}_{ML}, \hat{\boldsymbol{\omega}}\right\} = \arg\min_{\Phi,\boldsymbol{\omega}} f(\Phi,\boldsymbol{\omega}) \tag{16}$$

where, given a fixed frequency grid $\boldsymbol{\omega}$, the spectral density Φ may be estimated using the minimization-majorizaion method introduced in [8]. The frequency grid may then be updates in the same manner as the CASED. As the resulting estimator is computationally cumbersome, especially when also updating the evaluation points, one should initialize the method using a reasonably accurate initial estimate, which is here done using the CASED algorithm. The resulting algorithm, termed LIMESD (LIkelihood-based Method for Estimation of Spectra with gradient Descent), is summarized in Algorithm 2.

IV. Numerical examples

To exemplify the performance of the proposed algorithm, we generate observations of a sinusoidal signal with a normalized frequency of $\pi/2$ rad/s and an amplitude of 0.5. The signal is observed with an additive ARMA noise with MA coefficients $[1,\ 0.2s,\ 0.9s^2]$ and AR coefficients $[1,\ 0.3s]$ and where the process is driven by a white Gaussian noise with unit standard deviation. We generate 250 realizations of the signal, each with $N = 128$ samples. The resulting spectral density is shown in Fig. 1, as compared to the presented CASED and LIMESD algorithms (Fig. 1a and b, respectively), the periodogram using a rectangular window (Fig. 1c), the LASSO estimator [13] (Fig. 1d), the Iterative Adaptive Approach (IAA) estimator [4] (Fig. 1e), and the LIMES estimator (Fig. 1f). One may note that the Lasso estimator and IAA accurately identifies the sinusoidal frequency, but as expected perform poorly on the ARMA part. Similarly, LIMES accurately estimates the ARMA spectrum, as compared to the other estimators, but by design fails to represent the sinusoidal component accurately. Both CASED and LIMESD outperform the other estimators, accurately estimating both the sinusoidal and the smooth ARMA part. As can be seen from the figure, LIMESD produces slightly more precise estimates, although this comes at the cost of higher computational complexity. The average run times for the methods were $1.5{\cdot}10^{-4}$s for the periodogram, $1.4 \cdot 10^{-2}$s for the LASSO, 0.78s for IAA, 0.29s for LIMES, whereas CASED had an average running time of 12.80s and LIMESD 46.19s.

Algorithm 2 LIMESD

$\boldsymbol{\omega}_0 = \boldsymbol{\omega}_{CASED}$
$\Phi_0 = \Phi_{CASED}$
for i = 1, ..., until convergence **do**
  $\hat{\Phi}_i \leftarrow$ LIMES (see [8])
  $\boldsymbol{\omega}_i \leftarrow \boldsymbol{\omega}_{i-1} + s_i\frac{\partial f(\Phi_i,\boldsymbol{\omega})}{\partial\boldsymbol{\omega}}$
  if $f(\Phi_i, \boldsymbol{\omega}_i) < f(\Phi_i, \boldsymbol{\omega}_{i-1})$ **then**
    $s_{i+1} \leftarrow s_i/1.5$
  else
    $s_{i+1} \leftarrow 1.25 s_i$
    $\boldsymbol{\omega}_i \leftarrow \boldsymbol{\omega}_{i-1}$
  end if
end for

V. Conclusions and future work

In this work, we have presented methods for estimating mixed spectra using the piece-wise linear model, wherein we have also allowed the vertexes to shift, to better model the spectra. We aim to refine the method by allowing on-line addition or subtraction of frequency vertexes. We will also focus on reducing the complexity of the proposed algorithms, introducing better schemes for selecting appropriate step-sizes.

VI. Appendix

In this appendix, we present the derivatives of the cost-functions used to update the frequency evaluation points.

A. Least-Squares derivatives

Define the least-squares cost function as

$$\begin{aligned} L(\Phi) = &\left\|\hat{R}_{vec} - \boldsymbol{\mathcal{C}}(\boldsymbol{\omega})\Phi\right\|_2^2 \\ =&\hat{R}_{vec}^*\hat{R}_{vec} - \hat{R}_{vec}^*\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})\Phi - \Phi^T\overline{\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}^T\hat{R}_{vec} \\ &+ \Phi^T\overline{\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}^T\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})\Phi \end{aligned}$$

where $\overline{x}$ denotes complex conjugation of x. As $\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})$ is a complex-valued matrix, the gradient is found using Wirtinger calculus [14]

$$\begin{aligned} \frac{\partial L(\Phi)}{\partial\omega_k} =&\text{tr}\left(\left(\frac{\partial L(\Phi)}{\partial\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}\right)^T\frac{\partial\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}{\partial\omega_k}\right) \\ &+ \text{tr}\left(\left(\frac{\partial L(\Phi)}{\partial\overline{\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}}\right)^T\frac{\partial\overline{\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}}{\partial\omega_k}\right) \end{aligned} \tag{17}$$

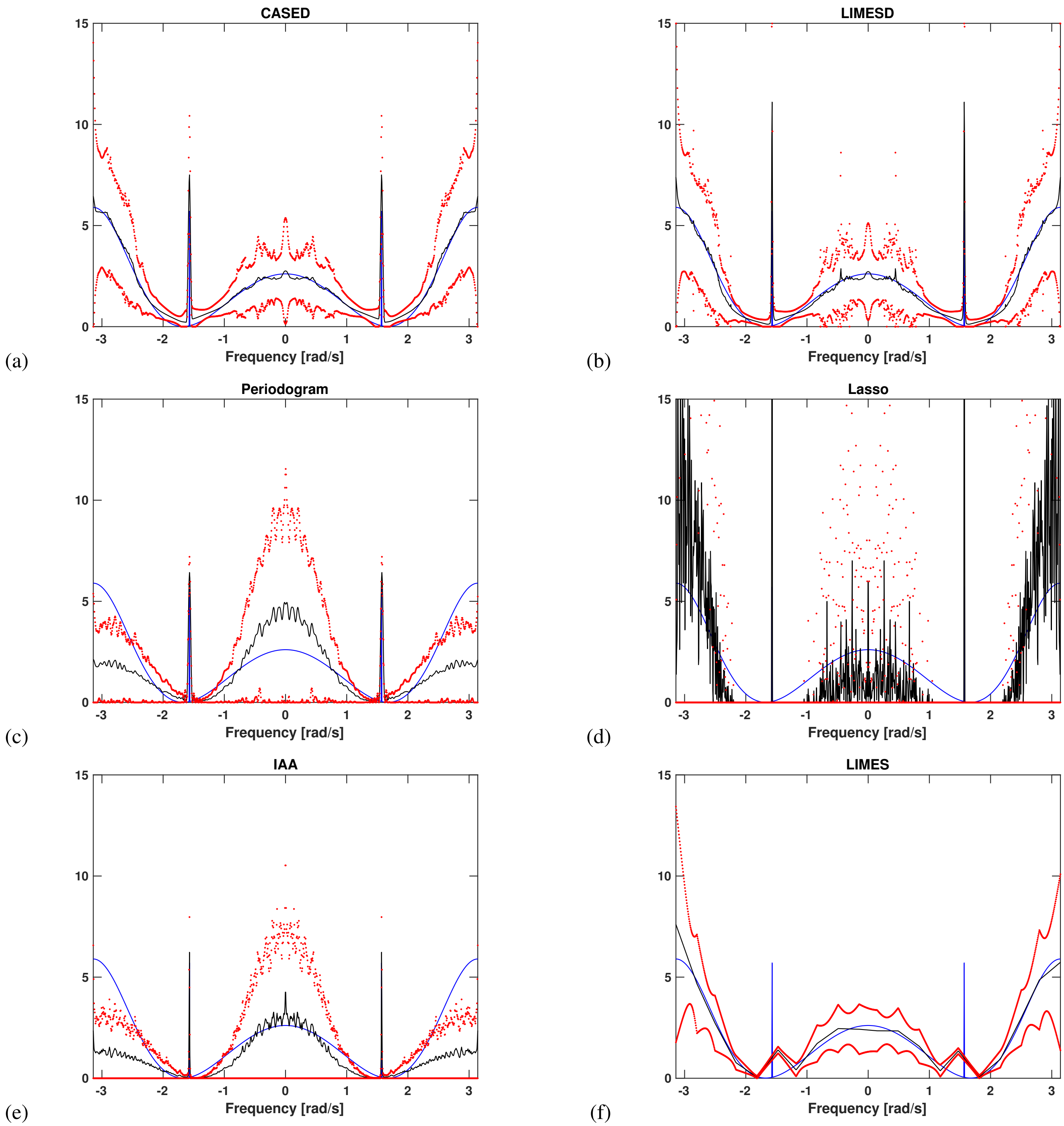

Fig. 1: The figures a-f shows the performance, as evaluated on 250 realizations of the example signal in Sec. IV. The blue lines represent the true spectral content. The black lines are the average spectral estimate, and the red dots are the mean $\pm$ one standard deviation, for each method respectively.

ISBN 978-90-827970-1-5

The two parts of the Wirtinger derivative are found as

$$\frac{\partial L(\Phi)}{\partial \boldsymbol{\mathcal{C}}(\boldsymbol{\omega})} = -\hat{R}_{vec}\Phi^T + \overline{\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}\Phi\Phi^T \tag{18}$$

$$\frac{\partial L(\Phi)}{\partial \overline{\boldsymbol{\mathcal{C}}(\boldsymbol{\omega})}} = -\hat{R}_{vec}\Phi^T + \boldsymbol{\mathcal{C}}(\boldsymbol{\omega})\Phi\Phi^T \tag{19}$$

The gradient of the matrices $C_k,\ \forall k$ are common for both the LS cost function and the ML cost function and are presented in Subsection VI-C.

B. Maximum-Likelihood derivatives

Define the negative log-likelihood as

$$f(\Phi,\boldsymbol{\omega}) = \ln|\det(R(\Phi,\boldsymbol{\omega}))| + y^* R(\Phi,\boldsymbol{\omega})^{-1} y \tag{20}$$

where $R(\Phi,\boldsymbol{\omega}) = \sum_{k=1}^{M+1}\phi_k C_k$. Even though C_k is complex valued, R is real-valued and the derivative of $f(\Phi)$ is found as

$$\frac{\partial f(\Phi,\boldsymbol{\omega})}{\partial \omega_k} = \operatorname{tr}\left(\left(\frac{\partial f(\Phi,\boldsymbol{\omega})}{\partial R(\Phi,\boldsymbol{\omega})}\right)^T \frac{\partial R(\Phi,\boldsymbol{\omega})}{\partial \omega_k}\right) \tag{21}$$

where

$$\frac{\partial f(\Phi,\boldsymbol{\omega})}{\partial R(\Phi,\boldsymbol{\omega})} = R(\Phi,\boldsymbol{\omega})^{-T} - R(\Phi,\boldsymbol{\omega})^{-T} y y^* R(\Phi,\boldsymbol{\omega})^{-T} \tag{22}$$

and

$$\frac{\partial R(\Phi,\boldsymbol{\omega})}{\partial \omega_k} = \sum_{j=1}^{M+1}\phi_j \frac{\partial C_j(\boldsymbol{\omega})}{\partial \omega_k} \tag{23}$$

where the derivatives of $C_k(\boldsymbol{\omega})$ is found in subsection VI-C

C. Linear Fourier transform derivative

The derivative of the piece-wise linear Fourier transformation with respect to frequency vertex ω_k is found as

$$\begin{aligned}
\frac{\partial R(\Phi,\boldsymbol{\omega})}{\partial \omega_k} &= \sum_{j=1}^{M+1}\phi_j \frac{\partial C_j(\boldsymbol{\omega})}{\partial \omega_k} \\
&= \phi_{k-1}\frac{\partial C_{k-1}(\boldsymbol{\omega})}{\partial \omega_k} + \phi_k \frac{\partial C_k(\boldsymbol{\omega})}{\partial \omega_k} + \phi_{k+1}\frac{\partial C_{k+1}(\boldsymbol{\omega})}{\partial \omega_k} \\
&= \phi_{k-1}\frac{\partial G_{k-1}(\boldsymbol{\omega})}{\partial \omega_k} + \phi_k\left(\frac{\partial G_k(\boldsymbol{\omega})}{\partial \omega_k} + \frac{\partial F_k(\boldsymbol{\omega})}{\partial \omega_k}\right) \\
&\quad + \phi_{k+1}\frac{\partial F_{k+1}(\boldsymbol{\omega})}{\partial \omega_k}
\end{aligned}$$

The partial derivatives of matrices F_k and G_k are presented below. Assuming $\tau = 0$ yields

$$\frac{\partial F_{k+1}(\boldsymbol{\omega})}{\partial \omega_k} = \frac{\partial G_k(\boldsymbol{\omega})}{\partial \omega_k} = \frac{-1}{4\pi} \tag{24}$$

and

$$\frac{\partial F_k(\boldsymbol{\omega})}{\partial \omega_k} = \frac{\partial G_{k-1}(\boldsymbol{\omega})}{\partial \omega_k} = \frac{1}{4\pi} \tag{25}$$

Further, assuming $\tau \neq 0$ implies that

$$\begin{aligned}
\frac{\partial F_k(\boldsymbol{\omega})}{\partial \omega_k} =& \frac{e^{i\omega_k\tau}}{2\pi}\Bigg(1 + \frac{i}{\tau(\omega_k-\omega_{k-1})} \\
&- \frac{1}{\tau^2(\omega_k-\omega_{k-1})^2}\left(1 - e^{-i(\omega_k-\omega_{k-1})\tau}\right)\Bigg)
\end{aligned} \tag{26}$$

$$\begin{aligned}
\frac{\partial F_{k+1}(\boldsymbol{\omega})}{\partial \omega_k} =& \frac{e^{i\omega_{k+1}\tau}}{2\pi}\Bigg(\frac{1 - e^{-i(\omega_{k+1}-\omega_k)\tau}}{\tau^2(\omega_{k+1}-\omega_k)^2} \\
&- \frac{i}{\tau(\omega_{k+1}-\omega_k)}e^{-i(\omega_{k+1}-\omega_k)\tau}\Bigg)
\end{aligned} \tag{27}$$

$$\begin{aligned}
\frac{\partial G_k(\boldsymbol{\omega})}{\partial \omega_k} =& \frac{e^{i\omega_k\tau}}{2\pi}\Bigg(-1 + \frac{i}{\tau(\omega_{k+1}-\omega_k)} \\
&- \frac{1}{\tau^2(\omega_{k+1}-\omega_k)^2}\left(e^{i(\omega_{k+1}-\omega_k)\tau} - 1\right)\Bigg)
\end{aligned} \tag{28}$$

$$\begin{aligned}
\frac{\partial G_{k-1}(\boldsymbol{\omega})}{\partial \omega_k} =& \frac{e^{i\omega_{k-1}\tau}}{2\pi}\Bigg(\frac{e^{i(\omega_k-\omega_{k-1})\tau} - 1}{\tau^2(\omega_k-\omega_{k-1})^2} \\
&- \frac{i}{\tau(\omega_k-\omega_{k-1})}e^{i(\omega_k-\omega_{k-1})\tau}\Bigg)
\end{aligned} \tag{29}$$

REFERENCES

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

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

[3] I. F. Gorodnitsky and B. D. Rao, "Sparse Signal Reconstruction from Limited Data Using FOCUSS: A Re-weighted Minimum Norm Algorithm," *IEEE Transactions on Signal Processing*, vol. 45, pp. 600–616, March 1997.

[4] T. Yardibi, J. Li, P. Stoica, M. Xue, and A. B. Baggeroer, "Source Localization and Sensing: A Nonparametric Iterative Approach Based on Weighted Least Squares," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, pp. 425–443, January 2010.

[5] G.-O. Glentis and A. Jakobsson, "Efficient Implementation of Iterative Adaptive Approach Spectral Estimation Techniques," *IEEE Transactions on Signal Processing*, vol. 59, pp. 4154–4167, Sept. 2011.

[6] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Transactions on Signal Processing*, vol. 59, pp. 35–47, Jan 2011.

[7] J. Swärd, S. I. Adalbjörnsson, and A. Jakobsson, "Generalized Sparse Covariance-based Estimation," *Elsevier Signal Processing*, vol. 143, pp. 311–319, February 2018.

[8] P. Stoica and P. Babu, "Maximum-Likelihood Nonparametric Estimation of Smooth Spectra From Irregularly Sampled Data," *IEEE Transactions on Signal Processing*, vol. 59, pp. 5746–5758, Dec. 2011.

[9] P. Stoica and P. Babu, "On the LIMES approach to spectral analysis mof irregularly sampled data," *Electronics Letters*, vol. 48, pp. 218–220, February 2012.

[10] P. Daniell, "On the Theoretical Specification and Sampling Properties of Autocorrelated Time-Series," *J. Roy. Stat. Ser. B*, vol. 8, pp. 88–90, 1946.

[11] J. Swärd, J. Brynolfsson, A. Jakobsson, and M. Sandsten, "Smooth 2-D Frequency Estimation using Covariance Fitting," in *22nd European Signal Processing Conference*, (Lisbon, Portugal), 1-5 Sep 2014.

[12] J. Brynolfsson, J. Swärd, A. Jakobsson, and M. Sandsten, "Smooth Time-Frequency Estimation Using Covariance Fitting," in *39th IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, (Florence, Italy), 4-8 May 2014.

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

[14] A. Hjørungnes, *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge University Press, 2011.

ISBN 978-90-827970-1-5 © EURASIP 2018