Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN), Vol. 1 (2012) pp. 1-12.

FREQUENCY EXTRAPOLATION THROUGH SPARSE SUMS OF LORENTZIANS

FREDRIK ANDERSSON*, MARCUS CARLSSON†, AND MAARTEN V. DE HOOP‡

Abstract. Sparse sums of Lorentzians can give good approximations to functions consisting of linear combination of piecewise continuous functions. To each Lorentzian, two parameters are assigned: translation and scale. These parameters can be found by using a method for complex frequency detection in the frequency domain. This method is based on an alternating projection scheme between Hankel matrices and finite rank operators, and have the advantage that it can done in weighted spaces. The weighted spaces can be used to partially revoke the effect of finite band-width filters. Apart from frequency extrapolation the method provides a way of estimating discontinuity locations.

1. Introduction. We study the frequency extrapolation of seismic records with a view to recovering information of the wavefield at frequencies away from the bandwidth in which observations are made. On the low frequency side, this extrapolation is important for applications in full waveform inversion.

Frequency extrapolation to enhance deconvolution of seismic records has been a recurring subject of study. We discuss a few different approaches developed since the 1970s. Dasgupta and Nowack [10] consider an iterative two-stage autoregressive (AR) extrapolation (EARD) approach, based on the work of Claerbout [8, 9]. In the first stage they apply a spectral division with a water table following the approach of Clayton and Wiggins. Clayton and Wiggins [16] used a maximum entropy method to extend the deconvolution bandwidth. Escalante *et al.* [12] invoke a basic sparsity constraint directly in the time domain. Extended time-domain deconvolution (ETDD) with frequency extrapolation was analyzed by Ligoria and Ammon [14]. A sequence of wavelets is iteratively applied to predict the observed data using an estimated source-time function. This results in a series of weighted and delayed pulses which is then bandpass filtered. By increasing the frequency bandwidth of the filter, an extended bandwidth for the deconvolution can be obtained. Chartrand, Sidky and Pan [7] develop a frequency extrapolation method by nonconvex compressed sensing. Frequency extrapolation has been a long standing problem studied in sampling theory; see, for example, Papoulis [15].

We will use the following model: Let f be a linear combination of piecewise smooth functions, and suppose that we measure

$$g(t) = f * \varphi(t) + n(t), \tag{1.1}$$

where $*$ denotes convolution. The band-limitation wavelet φ is assumed to be known. For the numerical simulations conducted in the paper, we will use normally distributed noise n. In [6] convincing examples are presented to motivate why many piecewise continuous functions are well approximated by sparse sums using scaled and translated sums using functons of the type

$$\frac{1}{1+t^2}, \quad \frac{t}{1+t^2},$$

i.e., of the form

$$f(t) \approx \sum_{j=1}^{J} c_j \frac{1}{a_j^2 + (t-\tau_j)^2} + d_j \frac{(t-\tau_j)}{a_j^2 + (t-\tau_j)^2}. \tag{1.2}$$

Functions of the form $(a^2 + (t-\tau)^2)^{-1}$ are commonly referred to as Lorentzians. For convenience of reference, we will use the term also to denote the second type of rational functions above.

*Centre for Mathematical Sciences, Lund University, Box 118, SE-22100 Lund Sweden, (fa@maths.lth.se).
†Centre for Mathematical Sciences, Lund University, Box 118, SE-22100 Lund Sweden, (mc@maths.lth.se).
‡Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA (mdehoop@purdue.edu).

In [6] approximations of type above is obtained by finding approximations of

$$\widehat{f}(\omega), \quad \omega > 0,$$

by using sparse sums of exponentials. The tool to do this is based on using the structure of Hankel matrices generated by functions on an interval.

In order to follow the same approach for the model (1.1), it would be necessary to divide $\widehat{g}$ by $\widehat{\varphi}$ in the frequency domain. Without the presence of noise, this procedure would likely be successful (as long as $\widehat{\varphi}$ is non-vanishing). However, the contribution of the noise term n will be severely boosted for typical choices of φ (i.e. such that get close to zero or vanish at subsets of frequencies of interest).

However, we will resolve this problem by using an approach which allows use to make approximations by sums of exponentials of functions on an interval, but in a *weighted* norm. In this way we can allow for the division by $\widehat{\varphi}$ – and the boost of the noise level associated with that –provided that we construct a weighted norm that counteracts the effect of this division. The methodology we will use is based on recent results using alternating projections for complex frequency estimation [3, 2].

2. Sparse approximation in the frequency domain. We will use the following unitary definition of the Fourier transform

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(t)\mathrm{e}^{-2\pi \mathrm{i}t\omega}\, dt.$$

Following the presentation of BM2009, we note that if p is a piecewise polynomial with compact support then

$$\widehat{p}(\omega) = \sum_{j=1}^{J} c_j |\omega|^{-\alpha_j} \mathrm{e}^{-2\pi \mathrm{i}\tau_j \omega}, \quad \alpha_j > 0, c_j \in \mathbb{C}, \tau_j \in \mathbb{R}.$$

Due to the compact support of p, $\widehat{p}$ is analytic which implies a strong interrelation constraint between the parameters α_j, τ_j and c_j.

It is shown in BM2005 that functions of the type $\omega^{-\alpha}$, $\alpha > 0$ are well approximated by short sums of exponential on finite intervals

$$\Omega = [\Omega_1, \Omega_2], \quad \Omega_1 \leq 0, \Omega_2 < \infty. \tag{2.1}$$

Hence, it seems plausible that compactly supported piecewise polynomials (and indirectly functions that are well approximated by such) are well approximated by sums of the form

$$\widehat{f_{\mathrm{app}}}(\omega) = \sum_{j=1}^{J} c_j \mathrm{e}^{-2\pi(a_j|\omega| + \mathrm{i}\tau_j\omega)}. \tag{2.2}$$

Note that f is real valued since $\widehat{f}(-\omega) = \overline{\widehat{f}(\omega)}$. Moreover, representations of the above form transforms into simple rational functions of Lorentzian type under the Fourier transform. This is straightforwardly derived as

$$\begin{aligned} \mathcal{F}^{-1}\left(c\mathrm{e}^{-2\pi(a|\cdot| + \mathrm{i}\tau\cdot)}\right)(t) &= 2\Re\int_0^{\infty} c\mathrm{e}^{-2\pi(a+\mathrm{i}(\tau-t))\omega}\, d\omega \\ &= 2\Re\frac{c}{a - \mathrm{i}(t-\tau)} = \frac{\Re(c)}{\pi}\frac{1}{(t-\tau)^2 + a^2} - \frac{\Im(c)}{\pi}\frac{(t-\tau)}{(t-\tau)^2 + a^2}. \end{aligned} \tag{2.3}$$

These functions are plotted in the two top panels of Figure 1 for $a = 1$ and $\tau = 0$. In this way, we

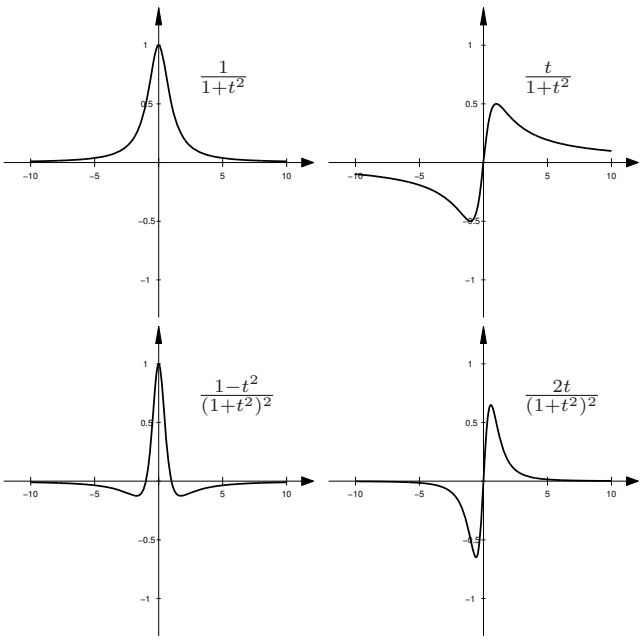

FIG. 1. *Lorentzian building blocks. The top two plots correspond to (2.3), and the two bottom ones to (2.5) and (2.6) for* $q = 1$.

obtain a appealing transition from the decay (a) and frequency (τ) parameters in the representation (2.2) to parameters of scale (a) and localization (τ).

These two functions decay as $|t-\tau|^{-1}$ and $|t-\tau|^{-2}$, respectively, as $|t-\tau| \to \infty$. For some cases, this decay rate may seem a bit to slow. A remedy could be to instead consider representations of the form

$$\widehat{f_{\text{app}}}(\omega) = \sum_{j=1}^{J} c_j |\omega|^q \mathrm{e}^{-2\pi(a_j|\omega| + \mathrm{i}\tau_j\omega)}, \tag{2.4}$$

since

$$\begin{aligned}\mathcal{F}^{-1}\left(c|\omega|^q \mathrm{e}^{-2\pi(a|\cdot| + \mathrm{i}\tau\cdot)}\right)(t) &= \frac{q!}{\pi(2\pi)^q} \Re \frac{c}{(a - \mathrm{i}(t-\tau))^{q+1}} \\ &= \frac{q!}{\pi(2\pi)^q}\left(\Re c L^{\text{even}}_{q,a,\tau}(t) - \Im c L^{\text{odd}}_{q,a,\tau}(t)\right),\end{aligned}$$

where

$$L^{\text{even}}_{q,a,\tau}(t) = \frac{1}{(a^2 + (t-\tau)^2)^{q+1}} \sum_{k=0}^{\lceil \frac{q}{2} \rceil} \binom{q+1}{2k} (-1)^k (t-\tau)^{2k} a^{n+1-2k}, \tag{2.5}$$

$$L^{\text{odd}}_{q,a,\tau}(t) = \frac{1}{(a^2 + (t-\tau)^2)^{q+1}} \sum_{k=0}^{\lfloor \frac{q}{2} \rfloor} \binom{q+1}{2k+1} (-1)^k (t-\tau)^{2k+1} a^{n-2k}. \tag{2.6}$$

The functions $L^{\text{even}}_{q,a,\tau}(t)$ and $L^{\text{odd}}_{q,a,\tau}(t)$ decay as $|t-\tau|^{2(q+1-\lceil \frac{q}{2} \rceil)}$ and $|t-\tau|^{2(q-\lfloor \frac{q}{2} \rfloor)+1}$, respectively, as $|t-\tau| \to \infty$.

A main motivation for the approximation using (2.2) is that there are constructive methods for computing such approximations. The approach used in [6] relies on the technique developed in [6].

The main principle to approximate a function $\widehat{f}$ by (2.2) on an interval Ω is to sample f at equally spaced points $\omega_k = k(\Omega_2 - \Omega_1)/(2N)$, $k = 0, \dots, 2N$, and construct a $N+1 \times N+1$- Hankel matrix

$$H_f(k,l) = \widehat{f}(\omega_{k+l}), \quad 0 \leq k, l \leq N.$$

The Hankel matrix has the complex symmetric property $H_f = H_f^T$ (which is different from usual Hermitian symmetry), and it therefore has a Takagi factorization $H_f = U\Sigma U^T$, where U is unitary and Σ is a diagonal matrix with positive entries σ_j. The columns of u_l of U are *con-eigenvectors* of H_f, and σ_l are the corresponding *con-eigenvalues*, i.e., $H_f u_l = \sigma_l \overline{u_l}$, $0 \leq l \leq N$. Typically, the con-eigenvalues σ_l decay rapidly.

In the next step, the J-th con-eigenvector u_J is used to construct a polynomial

$$p_J(z) = \sum_{k=0}^{N} u_J(k) z^k,$$

and let $z_n \in \mathbb{C}$ be the roots of p_J. Out of these N roots, it is *often* (but not always) possible to select J so that (after renumbering)

$$\sum_{j=1}^{J} \tilde{c}_j z_j^k \approx \widehat{f}(\omega_k),$$

in such a way that the error is *exactly* equal to σ_J when measured in a certain norm [4]. The result above can be expressed in the form (2.2) by a change of variables.

These results are very useful, and for functions whose Hankel matrices have rapidly decaying con-eigenvalues it yields good approximations using only few terms. However, the approximation is done using a slightly unusual norm, and the fact that the number of exponentials needed to achieve an approximation with an error of σ_J is not always true. For an extensive review of these properties, see [4].

The norm induced by this algorithm (the operator norm associated with the formation of Hankel matrices from sampled data) is not suitable for the applications we have in mind. While it is true that there is an L^∞ interpretation of the approximation error in the case of infinite sampling on a half axis [1], it is not at all as transparent the usual L^2 or a weighted L^2 version. Interpretations and estimates can be found in [4].

Instead we will employ the method alternating directions developed in [2] for computing approximations using sums of exponentials. A main advantage with this approach is that it allows for approximations in weighted L^2 spaces, for a certain class of weight functions. This ability will be crucial for making complex frequency detection when the relationship between measured data and the underlying function is perturbed by bandlimitation (wavelet filter) and when the presence of noise prevents deconvolution by means of simply division in the frequency domain.

For instance, one way of obtaining approximations using the higher order Lorentzians (2.5) and (2.6), is to use approximations

$$\frac{\widehat{f}(\omega)}{\omega^q} \approx \sum_{j=0}^{J} c_j \mathrm{e}^{-2\pi(a+\mathrm{i}\tau_j)\omega}, \quad a_j > 0, c_j \in \mathbb{C}, \tau_j \in \mathbb{R}, \text{for } 0 < \Omega_1 \leq \omega \leq \Omega_2.$$

The division by ω^q will then weight low-frequency data much higher than high-frequency data, and this may be undesirable. The usage of weights in the method of [2] can, however, be used to revoke some of this effect.

We need to review some results about Hankel matrices before presenting the suggested approximation method. For a more detailed overview, cf [2].

3. Hankel matrices and Takagi factorizations. A *Hankel* matrix H_f has constant entries on the anti-diagonals, and can thus be generated from some vector $f = (f_j)_{j=2}^{2N}$ by

$$Hf(j,l) := f(j+l), \quad 1 \le j,l \le N. \tag{3.1}$$

An orthonormal basis for the Hankel matrices in $\mathbb{M}_{N,N}$ is given by

$$e_m(j,l) = \begin{cases} \frac{1}{\sqrt{N-|N+1-m|}}, & \text{if } j+l=m; \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

for $2 \le m \le 2N$, where the normalization factor originates from the number of elements along anti-diagonal m.

Hankel matrices are a special case of the class of complex symmetric matrices, i.e., matrices that satisfy the symmetry condition $A = A^T$ in contrast to the usual (Hermitian) self-adjointness condition $A = A^*$. Similarly to real symmetric matrices, which are always diagonalizable, complex symmetric matrices can be decomposed as

$$A = \sum_{m=1}^{N} s_m \overline{u_m} u_m^*, \quad s_m \ \ge 0+, \quad u_m \in \mathbb{C}^N,$$

where the vectors $\{u_m\}_m$ are mutually orthogonal. This decomposition of A is called a *Takagi* factorization. The vectors u_m also satisfy the relation $Au_m = s_m\overline{u_m}$. In [13] the terminology *con-eigenvectors* and *con-eigenvalues* are used for u_m and σ_m, respectively. There are different ways to compute Takagi factorizations, out of one is easily derived from the following result, which proof is given as an exercise in [13].

PROPOSITION 3.1. *Let A and B be real symmetric $(N \times N)$-matrices and let*

$$W = \begin{pmatrix} A & -B \\ -B & -A \end{pmatrix}.$$

Let $d_1 \ge d_2 \ge \ldots \ge d_{2N}$ be the eigenvalues of W. Then $d_j + d_{2N+1-j} = 0$ for $j = 1, 2, \ldots, 2N$, and an orthonormal basis of eigenvectors can be chosen as

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}, \ldots, \begin{pmatrix} X_{2n} \\ Y_{2n} \end{pmatrix},$$

where $X_j, Y_j \in \mathbb{R}^N$, $\quad X_{2n+1-j} = -Y_j$ and $Y_{2n+1-j} = X_j$ for $j = 1, 2, \ldots, N$.

A finite dimensional version of Kroneckers theorem [2] say that in the generic case, if a Hankel matrix has rank J, then its generating function has the form

$$f(l) = \sum_{j=1}^{J} c_j e^{\zeta_j l}, \quad c_p, \zeta_p \in \mathbb{C}. \tag{3.3}$$

Given a Hankel matrix of rank J, how can the coefficients c_j and nodes ζ_j above be found? It turns out that if u_k are con-eigenvalues of H_f and we define

$$P_{u_k}(x) = \sum_{n=1}^{N} u_k(n) x^n,$$

then e^{ζ_j} will be zeros to each of the polynomials $P_{u_k}(x)$, $k > J$, cf. [5]. The set of roots of any of the of the polynomials P_{u_k} with corresponding zero con-eigenvalue will thus contain the nodes sought for. Thus, the problem of nodes ζ_j can be solved by root finding of a polynomial. The (linear) problem of finding the coefficients c_j is then easily solved. For a stable way of computing ζ_j, see [2].

Crucial for our approach to recovering functions from the measurement of their convolution with a wavelet function φ, is the usage of the weighted matrix approximations. Given a positive weight $w \in \mathbb{R}^N$, we denote by $\mathbb{M}^w_{N,N}$ the Hilbert space of matrices with the weighted Frobenius norm, given by

$$\|A\|_w^2 = \sum_{j,k=1}^{N} w(j)|A(j,k)|^2 w(k) = \|\mathrm{diag}(\sqrt{w})\, A\, \mathrm{diag}(\sqrt{w})\|^2.$$

THEOREM 3.2. *Let $A \in \mathbb{M}^w_{N,N}$, and let s_m and q_m denote con-eigenvalues and con-eigenvectors of $B = \mathrm{diag}(\sqrt{w})\ A\ \mathrm{diag}(\sqrt{w})$. Then the best rank k-approximation of A (in $\mathbb{M}^w_{N,N}$) is given by*

$$\sum_{m=1}^{k} s_m \overline{u_m} u_m^*,$$

where $u_m(l) = q_m(l)/\sqrt{w(l)}$, $1 \le l \le N$. The proof of the theorem is based on the Eckart-Young theorem (see e.g. [13, p 205], [11]).

When applying the weights above to Hankel matrices, a normalization with respect to the induced (matrix) weights along the anti-diagonal is needed, and we associate the weights

$$\omega(m) = \sum_{\substack{j+l=m \\ 1\le j,l\le N}} w(j)w(l), \quad 2 \le m \le 2N, \tag{3.4}$$

to w. A basis for Hankel matrices in the weighted space $\mathbb{M}^w_{N,N}$ is then given by

$$e^w_m(j,l) = \begin{cases} \frac{1}{\sqrt{\omega(m)}}, & \text{if } j+l=m; \\ 0, & \text{otherwise.} \end{cases}$$

for $2 \le m \le 2N$. Note that in the case $w = 1$ we get the "triangle weight" which appeared in (3.2). We let ℓ^ω_{2N-1} be the space of complex sequences $f = (f_j)_{j=2}^{2N}$, equipped with the norm defined by

$$\|f\|_\omega^2 = \sum_j |f_j|^2 \omega(j).$$

The mapping H (given by (3.1)) will in the sequel be considered as a mapping from ℓ^ω_{2N-1} to $\langle \mathbb{M}_{N,N}, \|\cdot\|_w \rangle$. It is a unitary map (isometric isomorphism), whose adjoint is the weighted averaging operator

$$H^*A(m) = \frac{1}{\omega(m)} \sum_{j+l=m} w(j)A(j,l)w(l), \tag{3.5}$$

and $H^*H = I$. The following proposition is now immediate.

PROPOSITION 3.3. *Let $w \in \mathbb{R}^N_+$ be given and let ω be the associated weight defined by (3.4). Let $f = (f_j)_{j=2}^{2N}$ and let $\mathcal{S}$ be any set of Hankel matrices. Then the problem*

$$\operatorname*{argmin}_{\tilde{H}\in\mathcal{S}} \|Hf - \tilde{H}\|_w$$

is equivalent to the problem

$$\operatorname*{argmin}_{g\in H^*(\mathcal{S})} \|f - g\|_{\ell^\omega_{2N-1}}.$$

The solutions are related by $Hg = \tilde{H}$.

4. Alternating projections. Let $\mathcal{H}_N$ the set of $N \times N$ Hankel matrices and $\mathcal{R}_{N,J}$ will denote the set of $N \times N$ matrices of rank at most k.

From Proposition 3.3 we know that the problem of finding the best approximation to $f \in \ell^\omega_{2N-1}$ of the form $f_{opt}(l) = \sum_{j=1}^J c_j e^{\zeta_j l}$ is equivalent to solving

$$\operatorname*{argmin}_{Hg \in \mathcal{R}_{N,J} \cap \mathcal{H}_N} \|Hf - Hg\|_w. \tag{4.1}$$

By starting with $Hf_0 = Hf$ and alternatively projecting onto the subsets $\mathcal{R}_{N,J}$ and $\mathcal{H}_N$, the idea is that the so arising sequence Hf_m will converge to an intersection point $Hf_\infty \in \mathcal{H}_{N,J} = \mathcal{R}_{N,J} \cap \mathcal{H}_N$, and moreover that Hf_∞ is in fact close to the optimal one, Hf_{opt}. This idea was investigated in a general framework in [3]. The main result of [3] roughly says that the above scheme indeed works if we start not too far away from $\mathcal{H}_N$.

Let $P_{\mathcal{R}_{N,J}}$ and $P_{\mathcal{H}_N}$ denote the maps taking a given matrix B onto (potentially one of) the closest point(s) in the respective manifolds. The alternating projection scheme then reads.

ALGORITHM 1.

1. *Let $f_0 = f$, $l = 0$*
2. *(Application of $P_{\mathcal{R}_{N,J}}$) Compute the first k con-eigenvalues s_m and the corresponding con-eigenvectors u_m of Hf_l using Theorem 3.2. The projection $P_{\mathcal{R}_{N,J}} Hf_l$ is then given*

$$\sum_{m=1}^k s_m \overline{u_m} u_m^*. \tag{4.2}$$

3. *(Application of $P_{\mathcal{H}_N}$) Compute*

$$f_{l+1} = H^* \left(\sum_{m=1}^k s_m \overline{u_m} u_m^* \right)$$

4. *Increase l and repeat from (2).*

It is crucial for practical signal processing that there are fast algorithms for computing both of the projection operators. A key components for both $P_{\mathcal{R}_{N,J}}$ and $P_{\mathcal{H}_N}$ is the utilization of FFT. For details, we refer to [2].

The following theorem is proven in [3, Theorem 6.1].

THEOREM 4.1. *For all $A \in \mathcal{H}_N$ outside a thin subset, the following is true. Given any $\epsilon > 0$, there exists an $s > 0$ such that, for all Hf with $\|Hf - A\| \leq s$, the sequence of alternating projections given by $B_0 = Hf$ and*

$$B_{j+1} = \begin{cases} P_{\mathcal{R}_{N,J}}(B_j) & j \text{ is even} \\ P_{\mathcal{H}}(B_j) & j \text{ is odd} \end{cases} \tag{4.3}$$

(i) converges to a point $Hf_\infty \in \mathcal{H}^n_{N,J}$
(ii) $\|Hf_\infty - Hf_{opt}\|_\omega \leq \epsilon \|Hf - Hf_{opt}\|_\omega$

In our setting, the condition (i) says that the algorithm produces an approximation of f of the form

$$f_\infty(l) = \sum_{j=1}^J c_j e^{\zeta_j l}, \tag{4.4}$$

is obtained as output of the of the algorithm. The condition (ii) says that the error $\|f_\infty - f_{opt}\|_{\ \omega}$ can be made arbitrarily small relative to the distance $\|f - f_{opt}\|_\omega$.

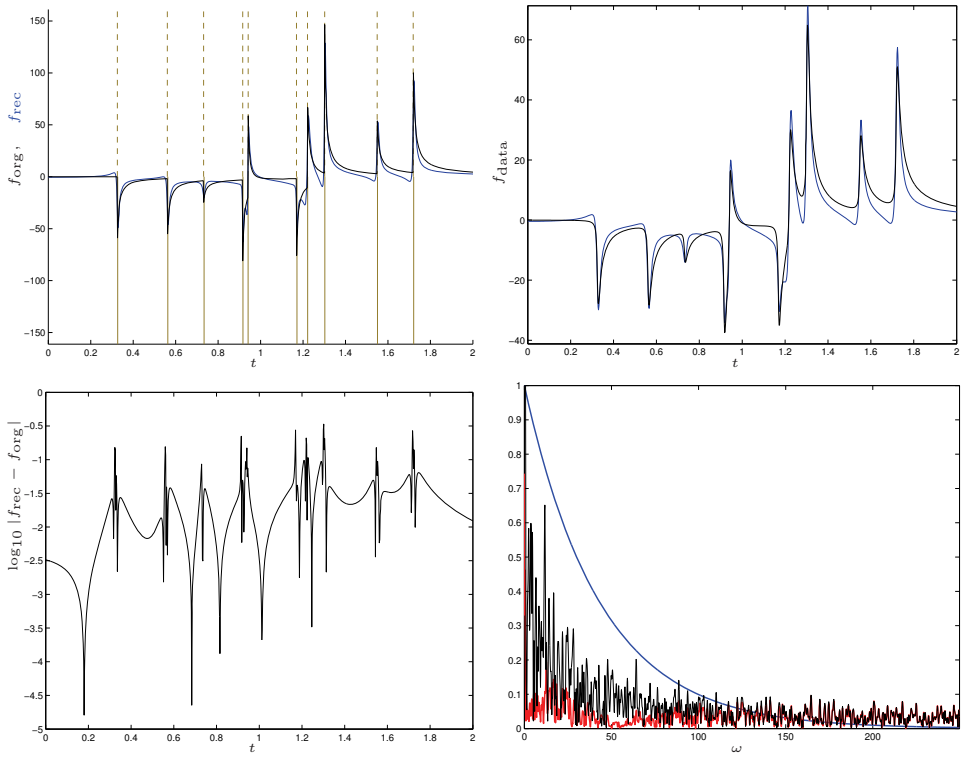

FIG. 2. *The top left panel shows the original function in black and the recovered approximation in blue. The dashed gold lines show the location of true singularites x_l in (5.1) and the solid gold lines the translations τ_j from (1.2). The right top panel shows in black and blue the functions of the left top panel, but convolved with φ from (1.1). In the lower left panel the approximation error is shown in $\log_{10}$-scale. The right lower panel shows Fourier transform of the approximation in black, the approximation error in red, and $\widehat{\varphi}$ in blue.*

Having the tool to find good approximations with sums of exponentials in weighted spaces, allows the construction of approximations of the form

$$f(t) \approx \sum_{j=1}^{J} c_j L^{\text{even}}_{q,a,\tau}(t) + d_j L^{\text{odd}}_{q,a,\tau}(t),$$

which takes the form (1.2) for the case where $q = 0$.

5. Numerical Experiments. In [6] a sum of piecewise polynomials with different degrees of discontinuities was used as a benchmark example to test the approximation capability of sparse sums of Lorentzians. The highest degree of the discontinuity of the function ranged from discontinuities in the function itself up to its third derivative.

As a representative example we will instead use function consisting of a sum of the form

$$\sum_{l=1}^{10} c_l \theta(t - x_l) \frac{1}{a_l(t - x_l) + 0.01}, \tag{5.1}$$

where θ denotes the Heaviside function.

The top left panel of Figure 2 shows one such function in black. The discontinuities x_l are shown in by the golden dashed lines. These dashed lines are only drawn in the upper plane (positive y-axis). In the lower plane there are solid golden lines, and these illustrate the detected locations of τ_j in the approximation (1.2). We see that the detected points are accurately τ_j aligned with

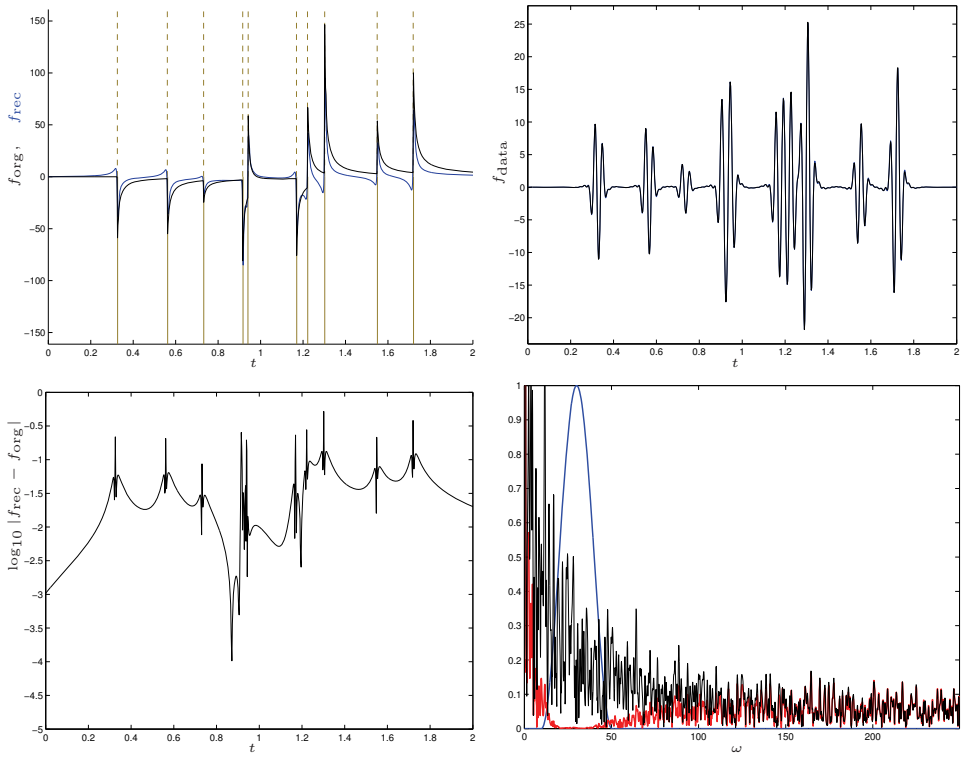

FIG. 3. *Illustrations showing counterparts to the plot in Figure 2 for the deconvolution problem (1.1) with* $\widehat{\varphi}$ *that has support in* $[10, 50]$ *Hz.*

the discontinuity locations x_l in (5.1), although the function types are substantially different as (for instance (1.2) is continuous whie (5.1)) is not.

The number of components J was determined by the size of the con-eigenvalues of the Hankel matrix H_f. For the simulations done with (5.1), we have chosen J so that $\sigma_{J-1}(H_f) > 0.1\sigma_0(H_f)$ and $\sigma_J(H_f) \leq 0.1\sigma_0(H_f)$, where we use the notation $\sigma_k(H_f)$ to denote the con-eigenvectors of the Hankel matrix H_f. As more terms in (1.2) are used, the approximations should improve, but the locations will not necessary align with the singularities of (5.1). However, they will have a tendency to cluster close to the singularities. See [6] for illustrations of this. Also, the recovery of more components will be more instable than the recovery of few. This corresponds to a loosening of the recovery (sparsity) constraint.

The right panel of Figure 2 shows results after the wavelet filter φ of 1.1 has been applied. Again the black curve shows the filtered version of the original function while the blue curve shows the filtered version of the reconstruction. For this example, a filter that preserves low-frequency information has been used. The frequency content of the filter ($\widehat{\varphi}$) is shown in blue in the bottom right panel of Figure 2. In this panel, the frequency content (amplitude) of the reconstruction is shown in black, and the error in red. The spatial error is also shown in the lower left panel, where it is depicted in log_{10}-scale that is normalized to the maximum value of the original function. We can see see error peaks at the singularity points, which is expected since we are trying to approximate discontinuous function by means of continuous functions.

The results in Figure 2 are expected from the results presented in [6]. Next, we consider the case were low-frequency information is not available. In Figure 3 results are shown where $\widehat{\varphi}$ has the shape of a bell ($\cos^2$) and is supported on an interval $[10, 50]$ Hz. The actual shape of $\widehat{\varphi}$ is shown in blue in the lower right panel. The panels of Figure 3 shows the counterpart of the plots of Figure 2. We can see that the performance in this case is comparable of the case where low frequency

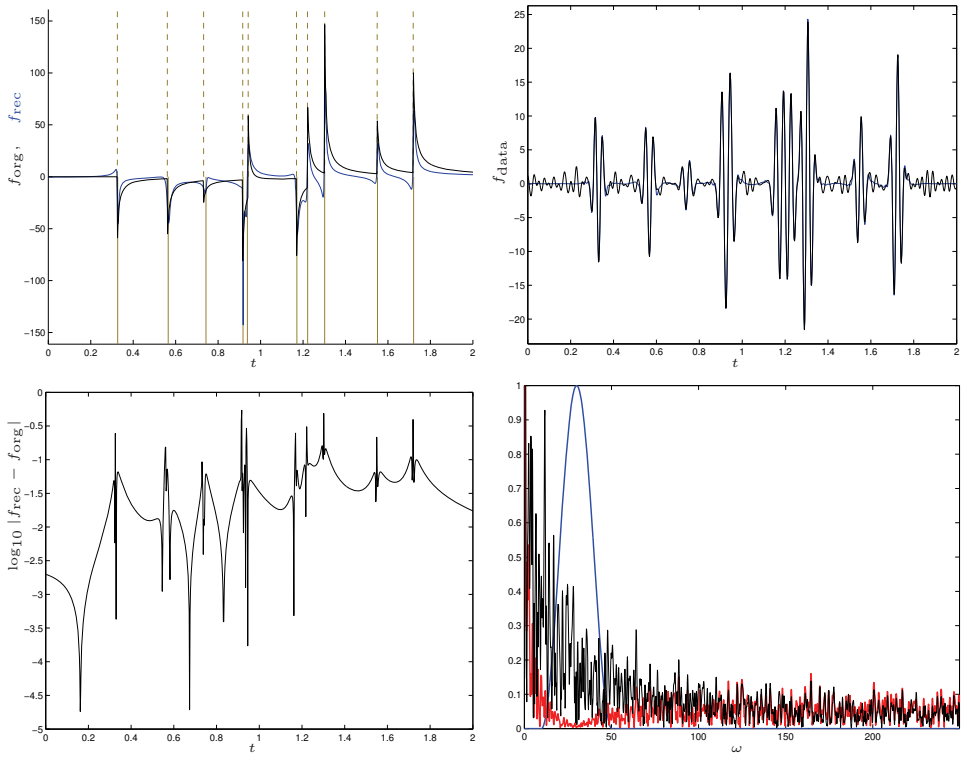

FIG. 4. *Illustrations showing counterparts to the plot in Figure 3 for the deconvolution problem (1.1) with added noise with a maximum amplitude that is* 10% *of the maximum filtered signal.*

information is available, although the plots in the upper right panels of Figure 2 and Figure 3 are rather different.

However, in the absence of noise, it should be rather easy to recover the same results in the two Figures, since one can simply divide by $\widehat{\varphi}$, use complex frequency estimation on the support of $\widehat{\varphi}$. For a more realistic comparison, we investigate how stable the method is to noise perturbations. First we need to choose a fair way of how to add noise. Since we assume that $\widehat{\varphi}$ is known, we will use add a filter satisfying

$$\widehat{\chi}(\omega) = \begin{cases} 1 & \text{if } \omega \in \operatorname{supp}(\widehat{\varphi}) \\ 0 & \text{otherwise.} \end{cases}$$

and add normally distributed random noise that is filtered by χ. We will show simulations using two noise levels.

Figures 4 and 5 shows results with added noise. The noise have been normalized so that its maximum amplitude is 10% of the maximum amplitude of f in (5.1) for Figure 4, while the maximum amplitude has been normalized to 20% in Figure 5. The noise levels are hence not discardable. In Figure 4 the discontinuity locations are essentially found, but the exact shape of the recovered function is note perfect. The results in Figure 5 are a bit less good. In particular, one of the singularity locations are noticeably incorrect. Still the method is able to determine most of the locations despite the rather high noise present.

6. Discussion. We have presented a new method for frequency extrapolation and deconvolution using a sparsity constraint. The sparsity constraint is that the function should be possible to represent as a sum of parametrized Lorentzians. The parameters consists apart from amplitude of translation and scale. We have assumed that measured data can be modeled as a convolution with

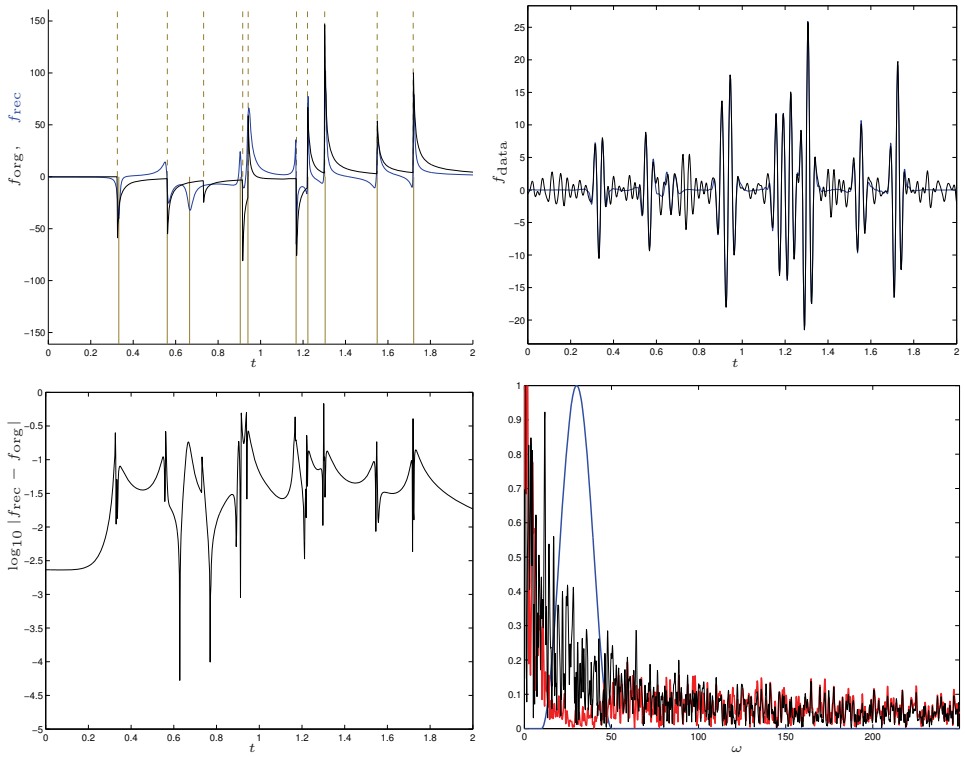

FIG. 5. *Illustrations showing counterparts to the plot in Figure 3 for the deconvolution problem (1.1) with added noise with a maximum amplitude that is* 20% *of the maximum filtered signal.*

a known (possibly band-limited filter which suppresses low frequency information) filter, along with additive noise.

The estimation is done by means of complex frequency detection in the frequency domain, meaning that both oscillations (imaginary part) and decay factor (real part) as parameters of exponentials are detected. The detection is made by using an alternate projection scheme between Hankel matrices and finite rank matrices. A major advantage of the approach is that it allows for approximations in weighted spaces which can be used to revoke the convolution effect of in band-limited measurements appearing e.g. in seismic measurements.

REFERENCES

[1] V. M. ADAMJAN, D. Z. AROV, AND M. G. KREĬN, *Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz*, Funkcional. Anal. i Priložen., 2 (1968), pp. 1–19.

[2] FREDRIK ANDERSSON AND MARCUS CARLSSON, *'a fast alternating projection method for complex frequency estimation*, Proceedings of the GMIG, 11 (2011), pp. 21—36.

[3] ———, *Alternating projections on low-dimensional manifolds*, Proceedings of the GMIG, 11 (2011), pp. 37–56.

[4] FREDRIK ANDERSSON, MARCUS CARLSSON, AND MAARTEN V DE HOOP, *Sparse approximation of functions using sums of exponentials and aak theory*, Journal of Approximation Theory, 163 (2011), pp. 213–248.

[5] GREGORY BEYLKIN AND LUCAS MONZÓN, *On approximation of functions by exponential sums*, Appl. Comput. Harmon. Anal., 19 (2005), pp. 17–48.

[6] GREGORY BEYLKIN AND LUCAS MONZÓN, *Nonlinear inversion of a band-limited fourier transform*, Applied and Computational Harmonic Analysis, 27 (2009), pp. 351–366.

[7] RICK CHARTRAND, EMIL Y. SIDKY, AND XIAOCHUAN PAN, *Frequency extrapolation by nonconvex compressive sensing*, in IEEE International Symposium on Biomedical Imaging (ISBI), 2011.

[8] JON CLAERBOUT, *Earth Soundings Analysis: Processing versus Inversion*, Blackwell Scientific Publications, Cambridge, MA:, 1992.

[9] ———, *Multidimensional recursive filters via a helix*, Geophysics, 63 (1998), pp. 1532–1541.

[10] Saptarshi Dasgupta and Robert L Nowack, *Frequency extrapolation to enhance the deconvolution of transmitted seismic waves*, Journal of Geophysics and Engineering, 5 (2008), pp. 118–127.
[11] Carl Eckart and Gale Young, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.
[12] Christian Escalante, Yu J. Gu, and Mauricio Sacchi, *Simultaneous iterative time-domain sparse deconvolution to teleseismic receiver functions*, Geophysical Journal International, 171 (2007), pp. 316–325.
[13] Roger A. Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994.
[14] Juan Pablo Ligorria and Charles J. Ammon, *Iterative deconvolution and receiver-function estimation*, Bulletin of the Seismological Society of America, 89 (1999), pp. 1395–1400.
[15] A. Papoulis, *A new algorithm in spectral analysis and band-limited extrapolation*, Circuits and Systems, IEEE Transactions on, 22 (1975), pp. 735 – 742.
[16] Rob W. Clayton Wiggins and Ralph A., *Source shape estimation and deconvolution of teleseismic bodywaves*, Geophysical Journal of the Royal Astronomical Society, 47 (1976), pp. 151–177.