A Minimax Dictionary Expansion for Sparse Continuous Reconstruction

Thiago A. R. Passarin, Daniel R. Pipa and Marcelo V. W. Zibetti
Federal University of Technology – Paraná (UTFPR)
Curitiba, Brazil
passarin@utfpr.edu.br, danielpipa@utfpr.edu.br, marcelozibetti@utfpr.edu.br

***Abstract*—The concept of dictionary expansion has been applied in inverse problems as a means to overcome a problem known as off-grid deviation. Within this framework and under the assumption that the off-grid deviations obey an uniform distribution, we propose a minimax error criterion to build expanded dictionaries. To this end, we formulate the problem as a polynomial regression and cast it as a second-order cone program. A robust method for the recovery of continuous time shifts and amplitudes from reconstructed expanded coefficients is also presented. Empirical results with a greedy algorithm and a convex optimization algorithm, both conceived to work with expanded dictionaries, show that the proposed expanded basis provides accurate reconstruction of continuous-time located events in the presence of noise.**

Keywords*—Inverse problems, dictionary expansion, manifold, optimization, sparse reconstruction.

I. Introduction

Consider a linear time-invariant (LTI) system whose impulse response is the signal $h(t)$ in Fig. 1a. The signal $h(t-\Delta)$ is the system response for an impulse with a delay of Δ. The problem of deconvolving an arbitrary signal that has been outputted by the system consists in finding a set of N impulses at the input, each having amplitude v_n and delay $n\Delta$, such that the output $c(t)$ can be represented as a finite combination of impulse responses:

$$c(t) = \sum_{n=1}^{N} v_n h(t - n\Delta) + e(t) \tag{1}$$

where $e(t)$ represents additive Gaussian measurement noise with zero mean and variance σ^2. This formulation is found in many applications of inverse problems such as neuron spike detection [1] and ultrasound nondestructive testing (NDT) [2], [3], where the signal to be reconstructed is modelled as a limited sum of shifted impulses with arbitrary amplitudes.

The use of a finite number of possible values for the delay $n\Delta$ implies a discretization (sampling) of the manifold formed in the space of acquired data by the continuous variation of the delay parameter Δ, as explained in [4]. This sampling (along with the signal time sampling) generates the discrete dictionary $\mathbf{H} \in \mathbb{R}^{M\times N}$, where each column $\mathbf{h}_n \in \mathbb{R}^M$ contains the delayed response $h(t - n\Delta)$. Considering that, Eq. (1) can be represented as

$$\mathbf{c} = \mathbf{H}\mathbf{v} + \mathbf{e} \tag{2}$$

where $\mathbf{c} \in \mathbb{R}^M$ is the vector of acquired data, $\mathbf{v} \in \mathbb{R}^N$ is the vector of amplitudes of each delayed impulse response, and $\mathbf{e} \in \mathbb{R}^M$ is the noise vector.

The problem of finding a sparse $\mathbf{v}$ that explains the data $\mathbf{c}$ through the model $\mathbf{H}$ is often formulated as

$$\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{v}\|_0 \ \text{ s.t. } \ \|\mathbf{c} - \mathbf{H}\mathbf{v}\|_2^2 \leq \epsilon \tag{3}$$

where $\|\cdot\|_0$ is the ℓ_0 pseudonorm and ϵ defines the upper bound for the reconstruction residual.

While the problem (3) is suitable for greedy algorithms such as [5], a convex relaxation can be obtained substituting the ℓ_0 pseudonorm by a ℓ_1 norm, which yields similar results in most cases while making the problem solvable by many convex optimization algorithms [6]. A well-known relaxed formulation is the LASSO [7] or Basis Pursuit (BP) [8]:

$$\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \frac{1}{2\sigma^2}\|\mathbf{c} - \mathbf{H}\mathbf{v}\|_2^2 + \lambda\|\mathbf{v}\|_1 \tag{4}$$

where λ is a parameter that controls the trade-off between the sparsity of the solution and the fidelity of the solution to the acquired data.

The problem known as off-grid deviation arises when some component of the data vector $\mathbf{c}$ corresponds to a response $h(t-\tau)$ with a value for τ not contemplated by the model $\mathbf{H}$, i.e., not a multiple of Δ. To cope with that phenomenon, the acquisition model (1) can be adapted to:

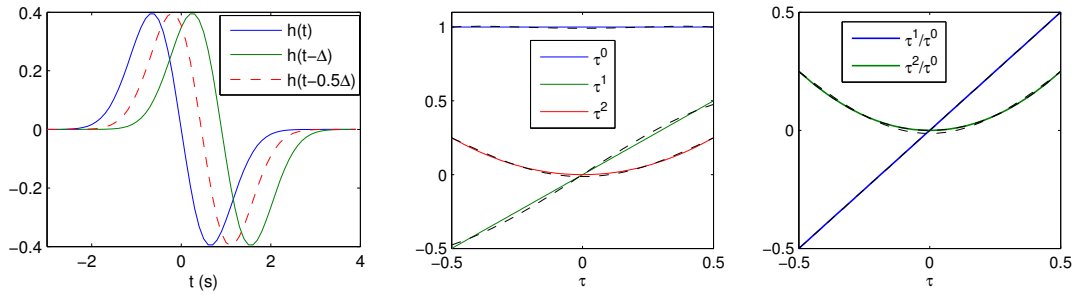
$$c(t) = \sum_{n=1}^{N} v_n h(t - (n + \tau_n)\Delta) + e(t) \tag{5}$$

where τ_n can vary continuously within the interval $[-0.5, 0.5]$. Rigorously, a continuous variation would require infinitesimal sampling on the manifold, which is not practically feasible. To overcome this problem, dictionary expansion has been proposed in [4]. The model matrix $\mathbf{H}$ is replaced by a set of K matrices $\{\mathbf{D}^{(k)}\}_{1\leq k\leq K}$, with the same dimensions as $\mathbf{H}$, which combined span an approximation of the continuous manifold formed by the variation of τ. (For simplicity, we will represent sets $\{x^{(k)}\}_{1\leq k\leq K}$ as $\{x^{(k)}\}$.) Eq. (2) is then adapted to the expanded acquisition model

$$\mathbf{c} \approx \sum_{k=1}^{K} \mathbf{D}^{(k)}\mathbf{x}^{(k)} + \mathbf{e} \tag{6}$$

where $\{\mathbf{x}^{(k)}\}$ are the K expanded vectors of coefficients used with $\{\mathbf{D}^{(k)}\}$, while the reconstruction problem (3) is general-

ISBN 978-0-9928626-7-1 © EURASIP 2017



(a) (b) (c)

Fig. 1. (a) Waveforms $h(t)$, $h(t-\Delta)$ and off-grid manifold sample $h(t-0.5\Delta)$. (b) Curves τ^0, τ^1 and τ^2 for $\tau \in [-0.5, 0.5]$. The dashed lines show the values of the coefficients $x_1^{(1)}$, $x_1^{(2)}$ and $x_1^{(3)}$ resulting from an unconstrained reconstruction with a single waveform $h(t-(1+\tau)\Delta)$. (c) Ratios τ^1/τ^0 and τ^2/τ^0 to be used in the formulation of the linear constraint set. The dashed lines, whose maxima and minima are the actual values used in the constraint set $\mathcal{C}$, are the ratios between the dashed lines of Fig. 1b.

ized to

$$\{\hat{\mathbf{x}}^{(k)}\} = \arg\min_{\{\mathbf{x}^{(k)}\}} \|\mathbf{x}^{(1)}\|_0$$
$$\text{s.t. } \|\mathbf{c} - \sum_{k=1}^{K} \mathbf{D}^{(k)}\mathbf{x}^{(k)}\|_2^2 \leq \epsilon, \{\mathbf{x}^{(k)}\} \in \mathcal{C} \tag{7}$$

where $\mathcal{C}$ is a convex constraint set relating each K-tuple $\{x_n^{(k)}\}$. The formulation of $\mathcal{C}$ depends on the type of expansion used. Mapping back from coefficients $\{x_n^{(k)}\}$ to impulses of amplitude v_n and time shift τ_n is possible and also depends on the type of expansion [4], [9].

A. Previous expansion approaches

In [4], the BP formulation (4) is adapted for expanded dictionaries and called Continuous Basis Pursuit (CBP):

$$\{\hat{\mathbf{x}}^{(k)}\} = \arg\min_{\{\mathbf{x}^{(k)}\}} \frac{1}{2\sigma^2}\|\mathbf{c} - \sum_{k=1}^{K} \mathbf{D}^{(k)}\mathbf{x}^{(k)}\|_2^2 + \lambda\|\mathbf{x}^{(1)}\|_1$$
$$\text{s.t. } \{\mathbf{x}^{(k)}\} \in \mathcal{C} \tag{8}$$

Two expansion schemes (Taylor and Polar) are proposed. The Taylor expansion is motivated by the fact that time shifts $h(t-\tau)$ on a sufficiently smooth function $h(t)$ can be approximated by a linear combination of $h(t)$ and its derivatives, while the Polar expansion is motivated by the fact that the manifold formed by the variation of τ on an LTI system must have both constant ℓ_2 norm and curvature. Both geometrically conceived approximations allow the recovered coefficients to be directly used on the reconstruction of continuous amplitudes and time shifts.

In [9], the Orthogonal Matching Pursuit (OMP) algorithm is adapted to solve problem (7) and called Continuous Orthogonal Matching Pursuit (COMP). An expansion scheme is also proposed, where the continuous manifold is finely sampled, generating a dense dictionary which then undergoes a singular value decomposition (SVD). The resulting K principal left singular vectors are then used as the expanded atom. Although the SVD expansion is not motivated by a geometric model, the recovered coefficients from the reconstructions using the expanded dictionary are indirectly used on the recovery of continuous amplitudes and time shifts.

B. Motivation

We assume that the off-grid deviation τ follows a uniform distribution. In the case of ultrasound imaging, for instance, that means that the probability of a scatterer to be located exactly on a location corresponding to a discrete manifold sample $h(t-n\Delta)$ is the same as on any neighbouring off-grid location corresponding to $h(t-(n+\tau)\Delta)$ with $\tau \in [-0.5, 0.5]$. An equivalent interpretation holds for the case of neuron spike detection: we expect the probability of a spike to occur exactly on a modeled instant of time to be the same as in any other time. This assumption motivates an expanded dictionary that aims not to privilege any specific point on the continuous manifold. Such criterion opposes, for example, the Taylor expansion, which yields exact representations for responses modeled by the original dictionary $\mathbf{H}$ at the expense of large error in the representation of off-grid deviated responses. Instead, our criterion ignores how the representation error evolves as τ_n varies in (5) and focuses on minimizing the maximum error.

II. Minimax Expansion

A. Problem formulation

Let $\mathbf{h}_n(\tau) \in \mathbb{R}^M$ be the vector containing the response $h(t-(n+\tau)\Delta)$, $\tau \in [-0.5, 0.5]$ with appropriate sampling rate. In this case, $\mathbf{h}_n(0)$ corresponds to the n-th atom $\mathbf{h}_n$ of the original dictionary.

Our formulation is based on a polynomial regression, with τ as the independent variable. We want to approximate deviations within the n-th time bin $\mathbf{h}_n(\tau)$ as a polynomial composition of $\{\mathbf{d}_n^{(k)}\}$:

$$\mathbf{h}_n(\tau) \approx \sum_{k=1}^{K} \tau^{k-1}\mathbf{d}_n^{(k)} \tag{9}$$

where $\mathbf{d}_n^{(k)}$ is the n-th column of $\mathbf{D}^{(k)}$. For instance, for $K=3$, we have a second degree polynomial:

$$\mathbf{h}_n(\tau) \approx \mathbf{d}_n^{(1)} + \tau\mathbf{d}_n^{(2)} + \tau^2\mathbf{d}_n^{(3)}.$$

In order to cast the regression problem, a fine sampling of the off-grid deviation τ is performed uniformly within the interval $[-0.5, 0.5]$, yielding a set of T different deviations

ISBN 978-0-9928626-7-1 © EURASIP 2017

$\{\tau_i\}_{1\le i\le T}$. For any arbitrary expanded set $\{\mathbf{d}_n^{(k)}\}$, each deviated waveform $\mathbf{h}_n(\tau_i)$ is approximated by (9) with the residual $\mathbf{r}_{n,i}$:

$$\mathbf{r}_{n,i} = \mathbf{h}_n(\tau_i) - \sum_{k=1}^{K} \tau_i^{k-1}\mathbf{d}_n^{(k)}. \tag{10}$$

Our goal now is the definition of a set $\{\mathbf{d}_n^{(k)}\}$ that minimizes the residual norms $\|\mathbf{r}_{n,i}\|$ for every τ_i drawn from our oversampling of the manifold. One approach would be solving a simple Least Squares (LS) problem:

$$\{\hat{\mathbf{d}}_n^{(k)}\} = \arg\min_{\{\mathbf{d}_n^{(k)}\}} \sqrt{\sum_{i=1}^{T} \|\mathbf{r}_{n,i}\|_2^2}$$
$$\text{s.t. } \mathbf{r}_{n,i} \text{ as in (10) } \forall i \in \{1,\dots,T\}. \tag{11}$$

However, recalling Section I-B, we aim to minimize the *maximum* residual norm among every $\|\mathbf{r}_{n,i}\|$, regardless of the values obtained for the other, nonmaximal i-th residuals. Such minimize-maximum (minimax) problem is formalized in Eq.(12):

$$\{\hat{\mathbf{d}}_n^{(k)}\} = \arg\min_{\{\mathbf{d}_n^{(k)}\}} \max_i \|\mathbf{r}_{n,i}\|$$
$$\text{s.t. } \mathbf{r}_{n,i} \text{ as in (10) } \forall i \in \{1,2,\dots,T\} \tag{12}$$

We cast the problem (12) as a Second-Order Cone Program (SOCP) [10] with the introduction of a slack variable r_{max}, whose value is constrained to be equal to or greater than every $\|\mathbf{r}_{n,i}\|$. The resulting optimization problem is defined in Eq. (13):

$$\{\hat{\mathbf{d}}_n^{(k)}\} = \arg\min_{\{\mathbf{d}_n^{(k)}\}} r_{max}$$
$$\text{s.t. } \left\{\begin{matrix}\|\mathbf{r}_{n,i}\| \le r_{max} \\ \mathbf{r}_{n,i} \text{ as in (10)}\end{matrix}\right\} \forall i \in \{1,2,\dots,T\} \tag{13}$$

We solve the SOCP (13) with CVX package for MATLAB [11].

Fig. 2 shows the plots of $\|\mathbf{r}_{1,i}\|$ (expansion of the first atom) obtained from the solution of the LS (11) and the Minimax (13) optimization problems for the waveform $h(t)$ of Fig. 1a, with $T = 21$ manifold samples. The LS solution has maximum residual norm $\|\mathbf{r}_{1,21}\| = 0.093$, while for the Minimax solution the maximum value is $\|\mathbf{r}_{1,16}\| = 0.068$ (27% smaller).

The expansion procedure described in this section has to be performed for every n-th atom $\mathbf{h}_n$. For LTI systems with proper sampling rate, the set $\{\mathbf{d}_1^{(k)}\}$ resulting from the expansion of the first atom $\mathbf{h}_1$ can be time-shifted in order to obtain all other sets $\{\mathbf{d}_n^{(k)}\}$.

B. Constraint set

The model (9) induces the ideal constraint set

$$x_n^{(k)} = \tau^{k-1} x_n^{(1)} \ \forall k \in \{1,\dots,K\}, \ \forall n \in \{1,\dots,N\}. \tag{14}$$

For instance, with $K = 3$ that means that any reconstructed coefficients set $\{x_n^{(k)}\}$ should follow the relation

$$(x_n^{(1)}, x_n^{(2)}, x_n^{(3)}) = v_n(\tau^0, \tau^1, \tau^2)$$

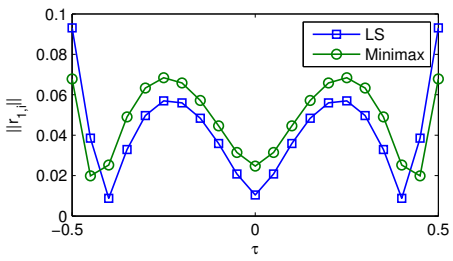


Fig. 2. Residual norms $\|\mathbf{r}_{1,i}\|$ from the expansion of the first atom with LS (11) and Minimax (13) for the waveform $h(t)$ of Fig. 1a, with $T = 21$ manifold samples.

where v_n is the scaling factor from (5). The values for the triplet (τ^0, τ^1, τ^2) for $\tau \in [-0.5, 0.5]$ are shown in Fig. 1b. The black dashed lines in Fig. 1b show the values of the reconstructed coefficients $x_1^{(1)}$, $x_1^{(2)}$ and $x_1^{(3)}$ from the unconstrained reconstruction of the single impulse $\mathbf{h}_{1,\tau}$ of Fig. 1a using only the first expanded set $\{\mathbf{d}_1^k\}$ obtained from (12). Note that, for this controlled case, the recovered coefficients nearly follow the ideal constraint set (14).

The ideal constraint set (14) is non-convex and nonlinear. A simple linear, convex relaxation is obtained if we constrain every $x_n^{(1)}$ to be positive and limit minimum and maximum ratios between $x_n^{(1)}$ and $x_n^{(k)}$ for $2 \le k \le K$:

$$\left\{\begin{matrix} x_n^{(1)} \ge 0 \\ \min\{0, (-0.5)^{k-1}\} \le \frac{x_n^{(k)}}{x_n^{(1)}} \le 0.5^{k-1} \forall k \in \{2,\dots,K\} \end{matrix}\right\}$$
$$\forall n \in \{1,\dots,N\}. \tag{15}$$

For $K = 3$, this yields $x_n^{(1)} \ge 0$, $-0.5 \le \frac{x_n^{(2)}}{x_n^{(1)}} \le 0.5$ and $0 \le \frac{x_n^{(3)}}{x_n^{(1)}} \le 0.25$. The constraint set (15) uses minima and maxima of the ratios represented in Fig. 1c, where the dashed curves represent the practical ratios obtained using the dashed curves of Fig. 1b. Our constraint set $\mathcal{C}$ uses those practical maxima and minima, which allows the expanded dictionary to represent the shifted waveforms with smaller error than with (15).

C. Recovery of continuous time shifts and amplitudes

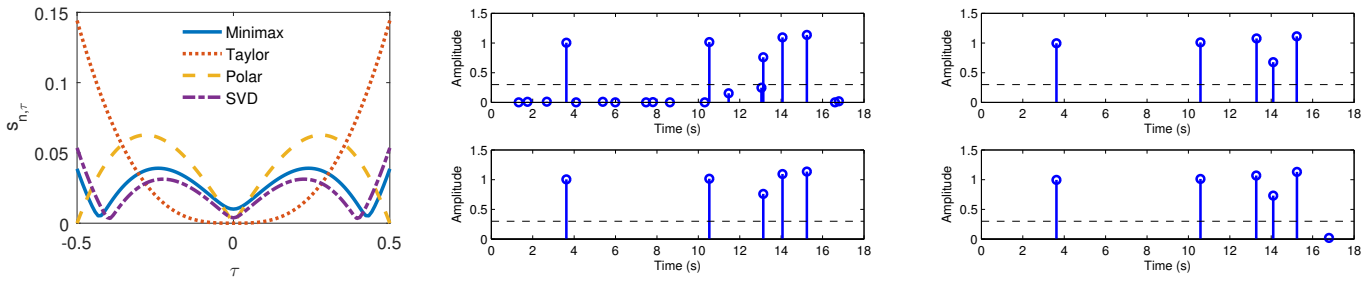
After solving (7) or (8) with a proper reconstruction algorithm, we translate from each n-th recovered expanded coefficients $\{x_n^{(k)}\}$ to time shift τ_n. We search for the value of τ_n that maximizes the dot product of Eq. (16).

$$\hat{\tau}_n = \arg\max_{\tau} \left(\sum_{k=1}^{K} x_n^{(k)} \mathbf{d}_n^{(k)}\right)^H \mathbf{h}_n(\tau) \tag{16}$$

where $(\cdot)^H$ represents the Hermitian operator. Since the dot product is a simple and fast operation, the solution $\hat{\tau}_n$ for (16) is found from brute force testing over a collection of responses obtained from a fine sampling of $\tau \in [-0.5, 0.5]$.

The recovered coefficients are also used in Eq. (17) to determine v_n from the ratio between the norm of the synthesized

ISBN 978-0-9928626-7-1 © EURASIP 2017



(a) (b) (c)

Fig. 3. (a) Distance $s_{1,\tau}$ with $\{\mathbf{d}_1^{(k)}\}$ as obtained with Taylor, Polar, SVD and Minimax expansion methods. (b) Signal recovered with CBP before (top) and after (bottom) the thresholding step described in Section III-B. The dashed lines represent the threshold of 0.3. (c) Signal recovered at the 5th (top) and 6th (bottom) iterations of COMP. The event added to the solution at the 6th iteration has amplitude 0.02, which is below the threshold of 0.3 (represented in the dashed lines), causing the algorithm to stop and chose the solution of the 5th iteration as final solution.

recovered response and the template response $\mathbf{h}_n$.

$$\hat{v}_n = \frac{\| \sum_{k=1}^{K} x_n^{(k)} \mathbf{d}_n^{(k)} \|}{\|\mathbf{h}_n\|} \tag{17}$$

A simpler approach motivated by the model (9) and the ideal constraint set (14) is possible making $\hat{v}_n = x_n^{(1)}$ and $\hat{\tau}_n = x_n^{(2)}/x_n^{(1)}$, but we found that (16) and (17) consistently provide smaller errors in our simulations. A possible explanation is the fact that the ideal constraint set is relaxed, as described in Section II-B, allowing solutions that do not strictly follow (14).

III. Preliminary Results

The following tests were performed with the waveform $h(t) \propto e^{-t^2}\sin(-t)$ shown in Fig. 1a with $\Delta = 0.9s$, a time sampling ratio of 10 samples per second, a dictionary size of $N = 20$ and $T = 21$ manifold samples per time bin. We choose $K = 3$ for all expansion schemes to allow for comparison with the Polar basis, for which $K = 3$ is the only possible order. For Taylor expansion, that implies using the original waveform and its first and second derivatives. For all alogirthms, the translation from reconstructed coefficients to continuous amplitudes and time shifts were performed according to each basis: see [4], Table 1 for Taylor and Polar, [9], Table 1 for SVD and Eqs. (16) and (17) for Minimax. All the quadratic optimization problems within CBP and COMP were solved using the CVX package for Matlab [11].

A. Distance to continuous manifold

We define $s_{n,\tau}$ as the minimum distance between the continuous manifold and the subspace spanned by $\{\mathbf{d}_n^{(k)}\}$. It is the norm of the difference between a shifted waveform $\mathbf{h}_n(\tau)$ and the reconstruction using only the corresponding n-th set $\{\mathbf{d}_n^{(k)}\}$. Fig. 3a shows $s_{n,\tau}$ for $n = 1$ as a function of τ for the set $\{\mathbf{d}_1^{(k)}\}$ as obtained with Taylor, Polar, SVD and Minimax bases. Note that the Taylor expansion yields an exact representation of the on-grid waveform (when $\tau_n = 0$) and privileges cases where τ_n is very small; the Polar expansion represents exactly the cases where $\tau = 0$ or $\tau = \pm 0.5$, privileging the neighbourings of those cases; and the SVD expansion yields a better average error along all the interval $\tau = [-0.5, 0.5]$ but still results in a large discrepancy between minimum and maximum distances. The Minimax expanded set yields a curve similar to that of SVD but with minimum maximum distance, which provides a more accurate representations near the extremes $\tau = \pm 0.5$.

B. Misses and false positives

We simulated the acquisitions of 100 signals, each consisting of 5 shifted copies of $h(t)$ with continuous time shifts drawn from a uniform distribution from $0s$ to $17s$ and unitary amplitudes. White Gaussian noise was added with signal-to-noise ratios (SNR) from 30dB to 10dB with a step of 5dB. The reconstructions from all simulated signals were performed with CBP and COMP using Taylor, Polar, SVD and Minimax bases.

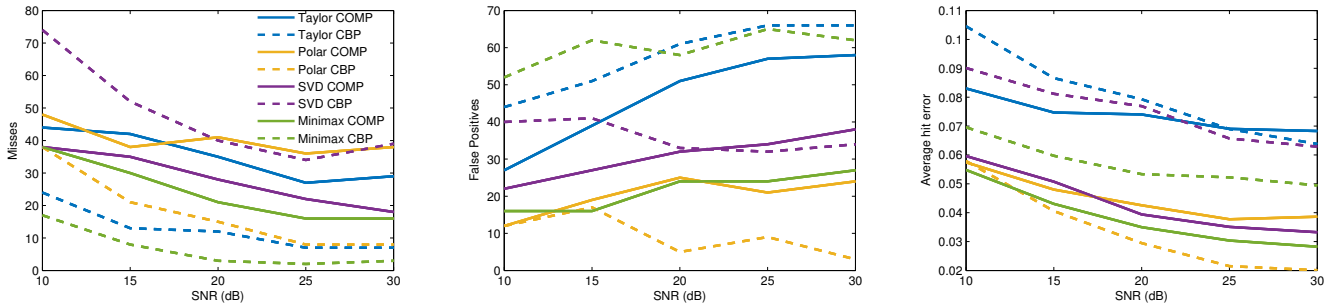
CBP formulation (8) uses a ℓ_1 norm penalization which, differently from the ℓ_0 pseudonorm of COMP, does not force solutions to be sparse in a strict sense, i.e., does not limit the number of recovered events, allowing some events of very small amplitude to participate in the solution. For this reason, we follow [9] and impose a thresholding step after the recovery with CBP, discarding events with amplitude less than 0.3. Fig. 3b shows an example of a recovered signal before and after the thresholding.

Following the same principles of OMP [5], COMP greedily adds expanded sets to the dictionary, yielding a solution with j events at each j-th iteration, until the stop criterion is met. From the prior knowledge on the fact that all original events have unitary amplitude, the stop criterion we use with COMP is based on the amplitude attributed to the newly added event at each iteration: when this amplitude is below a threshold of 0.3, the algorithm stops iterating and the result of the previous iteration is considered as the final solution. Fig. 3c shows an example where the event added at the 6th iteration has amplitude below the threshold, causing the algorithm to stop and chose the solution of the 5th iteration.

Following [4], a recovered event is matched to an original event if the shift between them is not greater than a limit of 0.5Δ and no other recovered event is at a shorter shift from the original event. In this case, a hit is computed. Unmatched original events are computed as misses, while unmatched recovered events are computed as false positives.

Fig. 4a shows a plot of misses computed from the 100 reconstructions (500 simulated events) using Taylor, Polar, SVD

ISBN 978-0-9928626-7-1 © EURASIP 2017



(a) (b) (c)

Fig. 4. Results from reconstructions of 100 signals, each consisting of 5 shifted copies of the template response $h(t) \propto e^{-t^2}\sin(-t)$ shown in Fig. 1a, using CBP and COMP with Taylor, Polar, SVD and Minimax bases. (a) Total number of misses, out of 500 original events. (b) Total number of false positives. (c) Average error between time shifts of original and recovered events for computed hits. A time bin of $\Delta = 0.45$s was used.

and Minimax bases with COMP and CBP. The Minimax basis provided the smallest number of misses for both algorithms and the best global result was obtained with CBP.

Fig. 4b shows a plot of false positives computed from the same set of reconstructions. For COMP, the Polar and Minimax bases have shown similar results, alternating as providers of the best result for this criterion. However, a sensible increase in the number of false positives is obtained with Minimax when we change from COMP to CBP.

C. Continuous time shift error

From the same reconstruction results of Section III-B we computed the time shifts between the original and the reconstructed events continuous locations for every computed hit. The averaged values for the reconstructions with Taylor, Polar, SVD and Minimax bases using both CBP and COMP are shown in Fig. 4c. The Minimax basis provided the smallest errors with COMP and was the second best with CBP.

IV. Discussion

We presented a Minimax formulation for the creation of expanded dictionaries in inverse problems. The criterion is motivated by the assumption that the off-grid deviations follow a uniform distribution, meaning that we search for the "fairest" distribution of the representation error along the manifold formed by the variation of the delay parameter τ. We also presented a SOCP formulation to reach that criterion, making the problem solvable by off-the-shelf convex optimization algorithms, as well as a robust method for the recovery of continuous time shifts and amplitudes from reconstructed expanded coefficients.

Our first empirical results show that the Minimax expansion scheme is competitive with earlier proposed Polar, Taylor and SVD schemes, especially within the greedy algorithm COMP, which is a generalization of the OMP algorithm for expanded dictionaries. This encourages further investigations on practical problems where the problem of off-grid deviation is present. For instance, the Taylor and the Polar bases have been used, respectively, in neuron spike identification [1] and compressive sensing of frequency-sparse signals [12].

Although the expansion process of Section II was derived for the LTI case, it is also suitable for time-variant (or, more generally, shift-variant) systems as long the direct model $h(t, \tau)$ is known. A typical application where shift-variant models arise is ultrasound NDT, where shift-variance is introduced by dispersion and attenuation. Recent efforts on the construction of dictionaries that contemplate those phenomena have been recently reported, for instance [2] and [3].

References

[1] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "A unified framework and method for automatic neural spike identification," *J. Neuroscience Methods*, vol. 222, pp. 47–55, Jan 2014.

[2] E. Carcreff, S. Bourguignon, J. Idier, and L. Simon, "A linear model approach for ultrasonic inverse problems with attenuation and dispersion," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no. 7, pp. 1191–1203, July 2014.

[3] E. Mor, A. Azoulay, and M. Aladjem, "A matching pursuit method for approximating overlapping ultrasonic echoes," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, no. 9, pp. 1996–2004, September 2010.

[4] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4735–4744, Oct 2011.

[5] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec 2007.

[6] J. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," *Information Theory, IEEE Transactions on*, vol. 52, no. 3, pp. 1030 –1051, march 2006.

[7] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. pp. 267–288, 1996.

[8] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[9] K. C. Knudson, J. Yates, A. Huk, and J. W. Pillow, "Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit," in *Advances in Neural Information Processing Systems 27*, 2014, pp. 1215–1223.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[11] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[12] K. Fyhn, H. Dadkhahi, and M. F. Duarte, "Spectral compressive sensing with polar interpolation," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 6225–6229.

ISBN 978-0-9928626-7-1 © EURASIP 2017