SPARSE REGULARIZED TOTAL LEAST SQUARES FOR SENSING APPLICATIONS

Hao Zhu†, Geert Leus, and Georgios B. Giannakis†*

†Dept. of ECE, University of Minnesota, 200 Union St. SE, Minneapolis, MN 55455, USA
*Delft University of Technology, Fac. EEMCS, Mekelweg 4, 2628CD, Delft, Netherlands

ABSTRACT

This paper focuses on solving sparse reconstruction problems where we have noise in both the observations and the dictionary. Such problems appear for instance in compressive sampling applications where the compression matrix is not exactly known due to hardware non-idealities. But it also has merits in sensing applications, where the atoms of the dictionary are used to describe a continuous field (frequency, space, angle, ...). Since there are only a finite number of atoms, they can only approximately represent the field, unless we allow the atoms to move, which can be done by modeling them as noisy. In most works on sparse reconstruction, only the observations are considered noisy, leading to problems of the least squares (LS) type with some kind of sparse regularization. In this paper, we also assume a noisy dictionary and we try to combat both noise terms by casting the problem into a sparse regularized total least squares (SRTLS) framework. To solve it, we derive an alternating descent algorithm that converges to a stationary point at least. Our algorithm is tested on some illustrative sensing problems.

Index Terms— Total least squares (TLS), sparsity, spectrum sensing, direction-of-arrival estimation.

1. INTRODUCTION

Although sparse reconstruction algorithms have been around for a while, there has been a renewed interest in this field inspired by the results obtained in compressive sampling [1, 2]. The basic problem in sparse reconstruction is to model an observation using only a few atoms of a large dictionary. A dictionary can be represented by a matrix, whose columns are the atoms of the dictionary. In compressive sampling applications for instance, the dictionary matrix is the product of a compression matrix with a basis matrix, whereas in sensing applications the atoms of the dictionary basically represent a dense grid of points in some continuous field such as frequency, space, angle, ..., or any combination thereof. In most sparse reconstruction applications, such as the ones discussed above, the dictionary matrix can be considered fat, making the reconstruction problem underdetermined. However, the knowledge that the solution is sparse can help us out by including a sparse regularization.

Generally, only the observations are considered noisy. The sparse reconstruction problem is then typically solved by formulating it as a least squares (LS) problem with sparse regularization, such as the least-absolute shrinkage and selection operator (LASSO) [3]. Little efforts have been made to tackle the case when both the observations and the dictionary are noisy. A noisy dictionary can occur in compressive sampling, where the compression actually takes place in the analog domain, and as a result the compression matrix is only approximately known due to non-idealities of the analog components. In this paper, on the other hand, we mainly address sensing applications where the atoms of the dictionary can be viewed as candidate grid points where sources could be active. Since in practice the actual sources will not be exactly on the grid, we can take that into account by modeling the grid points as noisy. We consider two sensing applications in this paper: spectrum sensing for cognitive radio and direction estimation in array processing. In spectrum sensing, every grid point corresponds to a specific location where a specific waveform (basis function) is transmitted [4, 5, 6], whereas in direction estimation, we use an angular grid, where every grid point represents a different direction-of-arrival (DoA) [7, 8]. An important benefit of assuming noisy grid points is that the estimated disturbance of these points provides a correction of the grid, leading to an improved accuracy of the model fitting.

In [9], it has been shown how the solution of an LS problem with sparse regularization is affected by noise in the dictionary, but no methods have been devised that take this noise into account. In this paper we develop an approach that is robust against noise in both the observations and the dictionary. We formulate this problem as a sparse regularized total least squares (SRTLS) problem, and we develop an alternating descent algorithm to solve this problem, where we alternate between the unknown coefficients and the unknown error on the dictionary. The problem can be shown to converge to a stationary point at least. We apply this algorithm to both sensing applications discussed earlier, illustrating the benefits of SRTLS over LASSO. Finally, note that the setup considered here is different from the one in [10], where a seemingly related problem was considered but for dictionary learning instead of sparse reconstruction.

Notations: Upper (lower) bold face letters will be used for matrices (column vectors); $(\cdot)^{\mathbf{T}}$ denotes transposition; $(\cdot)^{\mathbf{T}}$ Hermitian transpose; $\mathbf{1}_{m\times n}$ the $m \times n$ matrix of all ones; $\|\cdot\|_F$ the Frobenius norm; and $\|\cdot\|_p$ the vector $p-$norm for $p \geq 1$.

2. DATA MODEL AND PROBLEM STATEMENT

Consider an underdetermined linear system for sensing an unknown sparse signal vector $\mathbf{x}$ of length n, and the observation $\mathbf{y} \approx \mathbf{A}\mathbf{x}$ is a vector of length m. Under the assumption that only $\mathbf{y}$ is corrupted by an additive noise $\mathbf{e}$ in the form of

$$\mathbf{y} + \mathbf{e} = \mathbf{A}\mathbf{x}, \tag{1}$$

it is well known that a sparse solution can be obtained by solving an LS problem with an ℓ_1 norm regularization, which is also known as

This work is supported by the NSF grants CCF-0830480 and ECCS-0824007; Geert Leus is supported in part by NWO-STW under the VICI program (project 10382).

the LASSO [3]

$$\hat{\mathbf{x}}_{LASSO} := \arg\min_{\mathbf{x},\mathbf{e}} \|\mathbf{e}\|_2^2 + \lambda\|\mathbf{x}\|_1 \quad \text{s.t. } \mathbf{y}+\mathbf{e} = \mathbf{A}\mathbf{x}, \tag{2}$$

for some $\lambda > 0$. However, provided that both the observation $\mathbf{y}$ and the dictionary $\mathbf{A}$ are given data, it is more reasonable to treat them symmetrically by assuming that $\mathbf{A}$ is also corrupted by an additive noise term $\mathbf{E}$. Then the system model becomes

$$\mathbf{y}+\mathbf{e} = (\mathbf{A}+\mathbf{E})\mathbf{x}. \tag{3}$$

The above model takes into account the (possible) noise, or system mismatch, in the dictionary matrix $\mathbf{A}$, and has been considered extensively as the total least squares (TLS) problem, with applications as broad as image reconstruction, speech and audio processing, modal and spectral analysis, system identification, and so on; see e.g., [11]. However, as far as we know, TLS has not yet been studied in combination with ℓ_1 regularization. In addition, TLS modeling of (3) as well as all the TLS solvers so far, including the ℓ_2 regularized ones, have not yet been applied for an underdetermined system.

Therefore, compared to the TLS problem, we consider a system where the number of observations is less than the number of unknown coefficients, i.e., $m < n$ in (3). Then the objective becomes to obtain a parsimonious estimate of the signal $\mathbf{x}$ given $\mathbf{y}$ and $\mathbf{A}$, both noise-corrupted. Assuming the noise terms $\mathbf{E}$ and $\mathbf{e}$ in (3) are uncorrelated with each other and across the entries, we are interested in solving the sparsity regularized TLS (SRTLS) problem formulated as

$$\hat{\mathbf{x}}_{SRTLS} := \arg\min_{\mathbf{x},\mathbf{e},\mathbf{E}} \| [\mathbf{E},\ \mathbf{e}] \|_F^2 + \lambda\|\mathbf{x}\|_1 \quad \text{s.t. } \mathbf{y}+\mathbf{e} = (\mathbf{A}+\mathbf{E})\mathbf{x}, \tag{4}$$

for some $\lambda > 0$. Different from (2), the problem (4) looks for minimal (in the Frobenius norm sense) corrections $\mathbf{E}$ and $\mathbf{e}$ on both the given data $\mathbf{A}$ and $\mathbf{y}$ that allows the corrected system to afford the solution $\hat{\mathbf{x}}_{SRTLS}$ with a minimal ℓ_1 norm. Notice that the formulation (4) treats the noise variance for each entry of $[\mathbf{E},\ \mathbf{e}]$ as equal, yet disproportional or even structured noise models can also be allowed but are omitted here for space consideration. A close look at the optimization problem (4) reveals that it involves the product of the optimization variables $\mathbf{E}$ and $\mathbf{x}$. Thus it is generally a non-convex problem and we will seek an iterative alternating-descent type algorithm which converges to a stationary point at least in Section 3. Before discussing the algorithmic implementation, it is useful to understand the importance of the noise term $\mathbf{E}$. Next, we will look at two sparse signal reconstruction applications, where the noise $\mathbf{E}$ originates from errors due to the assumptions that are made about the dictionary matrices. Including this noise term $\mathbf{E}$ in the data model and the sparse system solver, a better system performance can be obtained.

2.1. Spectrum Sensing in Cognitive Radio Networks

Spectrum sensing is a critical prerequisite in envisioned applications of wireless cognitive radio (CR) networks which promise to resolve the perceived bandwidth scarcity versus under-utilization dilemma. The task is to estimate the transmitting source locations and identify its (un)used frequency bands; see e.g., [4]. Specifically, consider N_s sources (transmitters) located at position vectors $\{\boldsymbol{x}_s\}_{s=1}^{N_s}$ and N_r CRs at $\{\boldsymbol{x}_r\}_{r=1}^{N_r}$. Relying on a virtual grid of *candidate* source locations depicted in Fig. 1, vectors $\boldsymbol{x}_s$ no longer describe the actual positions of e.g., primary users but grid points with *known* spatial coordinates where transmitting radios could be present. Under this scenario, the power spectral density (PSD) $\Phi_s(f)$ is nonzero if there is a transmitter present at $\boldsymbol{x}_s$, and zero otherwise.

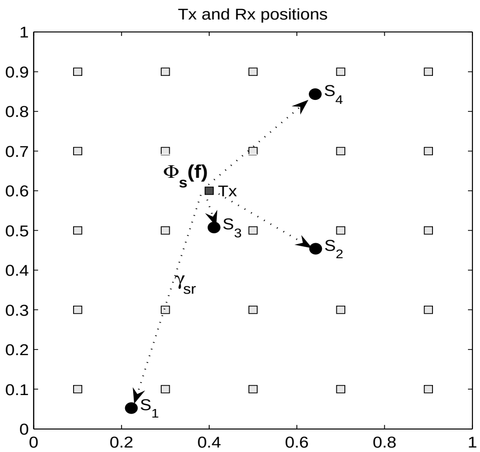


Fig. 1. Virtual CR network grid with $N_s = 25$ candidate locations, 1 transmitting source, and $N_r = 4$ receiving CRs.

For a sufficiently large number of bases N_b, the PSD of each source can be well approximated by the basis expansion model

$$\Phi_s(f) = \sum_{\nu=1}^{N_b} \theta_{s\nu} b_\nu(f), \quad s = 1, 2, \ldots, N_s, \tag{5}$$

where $\{b_\nu(f)\}_{\nu=1}^{N_b}$ is a collection of known bases, and $\{\theta_{s\nu}\}$ denote the expansion coefficients to be estimated. Furthermore, the channel gain γ_{sr} from the source at $\boldsymbol{x}_s$ to the CR at $\boldsymbol{x}_r$ is assumed to follow a known function of the source-receiver distance; e.g., the exponential pathloss model where $\gamma_{sr} = e^{-||\boldsymbol{x}_r - \boldsymbol{x}_s||^2/\delta^2}$, with δ a known constant. Together with some other system specifications as detailed in [4], the received PSD can be approximated as the linear combination of the PSDs of all the sources and noise

$$\begin{aligned} \Phi_r(f) &\approx \sum_{s=1}^{N_s} \gamma_{sr}\Phi_s(f) + \sigma_r^2 \\ &\approx \sum_{s=1}^{N_s} \gamma_{sr} \sum_{\nu=1}^{N_b} \theta_{s\nu} b_\nu(f) + \sigma_r^2 \\ &\approx \boldsymbol{b}_r^T(f)\boldsymbol{\theta} + \sigma_r^2, \end{aligned} \tag{6}$$

where σ_r^2 stands for the noise power and the $P := N_b N_s \times 1$ vector $\boldsymbol{\theta}$ is formed by stacking the columns of the matrix with entries $\theta_{s\nu}$, and $\boldsymbol{b}_r(f)$ by concatenating the columns of the matrix with entries $\gamma_{sr} b_\nu(f)$.

Stacking (6) across receivers and sampling frequencies leads to a linear model for this specific spectrum sensing application, which fits the SRTLS one (3) in the following three ways. Firstly, sparsity in $\boldsymbol{\theta}$ is manifested since the model (6) is parsimonious both in frequency as well as in space. Secondly, the noise term $\mathbf{e}$ in (3) can account for the imperfection of estimating the receiver PSD due to the limited memory in practice. Last but most importantly, the disturbance $\mathbf{E}$ in (3) is well motivated as (6) assumes that the sources are located exactly *on the grid*. Notice that in (6) only the model parameters corresponding to the candidate grid points are stacked in the vector $\boldsymbol{\theta}$ to be estimated. However, most likely in a real implementation, the actual source locations differ from any of the grid points, as shown in Fig. 1. This results in a mismatch between $\boldsymbol{b}_r(f)$ in (6) which corresponds to only the grid points and the one to the actual transmitter locations. Moreover, the estimated error $\mathbf{E}$ that is obtained by solving the problem (4) can further be used to improve the source location estimation, thereby increasing the spectrum sensing accuracy over the spatial field, as explained in more details in Section 4. Note that the error $\mathbf{E}$ could also be used to recover from possible inaccuracies in the bases $b_\nu(f)$, but we will not consider this here.

2.2. Direction Estimation in Array Processing

The goal of array source localization is to find the direction-of-arrival (DoA) of sources of wavefields that impinge on an array consisting of a number of antennas. For purposes of exposition, we only focus on the narrowband scenario with a single snapshot. Similar to the spectrum sensing problem where the parameters for all grid points are to be estimated, under this setup we consider all possible source DoAs $\{\theta_s\}_{s=1}^{N_s}$, usually taken from a uniform sampling. We represent the signal field by a vector $\boldsymbol{z}$ of length N_s, where the s-th entry z_s is nonzero and equal to its transmitting signal amplitude if there exists a source from the angle θ_s and zero otherwise.

To this end, define the $N_r \times 1$ array steering vector of a uniform linear array (ULA) of N_r antennas corresponding to the source from direction θ_s as

$$\boldsymbol{\psi}(\theta_s) := [1 \; e^{-j\alpha_s} \; e^{-j\alpha_s(N_r-1)}]^T, \tag{7}$$

where the array phase shift $\alpha_s := 2\pi(d/\lambda_s)\sin(\theta_s)$ with d the distance between neighboring antennas and λ_s the signal wavelength of the source from θ_s. Then the received signal at the ULA can be written as an $N_r \times 1$ vector $\boldsymbol{\varphi}$ of the form

$$\boldsymbol{\varphi} = \boldsymbol{\Psi}\mathbf{z} + \mathbf{w} \tag{8}$$

where $\mathbf{w}$ is the receiver additive noise and $\boldsymbol{\Psi}$ is the angle scanning matrix taking the form $\boldsymbol{\Psi} = [\boldsymbol{\psi}(\theta_1) \; \ldots \; \boldsymbol{\psi}(\theta_{N_s})]$. Similar to the spectrum sensing example, this overcomplete representation in (8) allows us to exchange the problem of parameter estimation of the source signal for the problem of sparse estimation of the source scanning vector $\mathbf{z}$. However, the same issue also appears since the finite sampling will fail to cover the sources from the angles other than $\{\theta_s\}_{s=1}^{N_s}$, and equivalently an error in the angle scanning matrix $\boldsymbol{\Psi}$ may result. Therefore, applying the SRTLS problem formulation (4), which takes into account the error in $\boldsymbol{\Psi}$, has the potential to recover a more accurate estimate of $\mathbf{z}$. What is more, the estimate of the error in $\boldsymbol{\Psi}$ is also useful to improve the source angle inference, as explained in Section 4.

In the next section, we will develop the alternating descent algorithm used for solving the SRTLS problem (4).

3. SRTLS ALTERNATING DESCENT ALGORITHM

As mentioned earlier, the SRTLS problem (4) is non-convex, thus no efficient convex solver can guarantee to achieve the global optimum. To this end, we adopt an iterative approach where all the variables are updated in turn per iteration. First notice that $\mathbf{e}$ follows a fixed relationship with $\mathbf{E}$ and $\mathbf{x}$ from the constraint in (4). For a given $\mathbf{E}$ matrix, the SRTLS problem (4) reduces to the LASSO formulation in (2) so $\mathbf{x}$ can be solved efficiently; also by fixing $\mathbf{x}$, the optimal $\mathbf{E}$ can be obtained as the solution of a constrained LS problem. Therefore, our alternating descent algorithm only needs to iterate between $\mathbf{E}$ and $\mathbf{x}$ in turn per iteration.

Specifically, let $\mathbf{E}(k)$ denote the iterate for $\mathbf{E}$ at iteration k; and likewise for $\mathbf{x}(k)$. After initializing the algorithm at iteration $k = 0$ with $\mathbf{E}(0) = \mathbf{0}_{m\times n}$, the update iterations follow. For any iteration $k \geq 0$, the signal iterate $\mathbf{x}(k)$ is obtained as the optimum of the problem

$$\mathbf{x}(k) = \arg\min_{\mathbf{x},\mathbf{e}} \|\mathbf{e}\|_2^2 + \lambda\|\mathbf{x}\|_1 \quad \text{s.t. } \mathbf{y} + \mathbf{e} = [\mathbf{A} + \mathbf{E}(k)]\mathbf{x}, \tag{9}$$

using some convex optimization solver; e.g., the interior point solver SeDuMi [12]. Substituting the constraint of (4) back into its cost function for given $\mathbf{x}(k)$, the noise matrix $\mathbf{E}(k+1)$ is further estimated as

$$\mathbf{E}(k+1) = \arg\min_{\mathbf{E}} \|\mathbf{E}\|_F^2 + \|\mathbf{y} - \mathbf{A}\mathbf{x}(k) - \mathbf{E}\mathbf{x}(k)\|_2^2 \tag{10}$$

By setting the first-order derivative of its cost function to zero, the optimal solution to the quadratic problem (10) is found in closed form as

$$\mathbf{E}(k+1) = [\mathbf{y} - \mathbf{A}\mathbf{x}(k)]\mathbf{x}^T(k)[\mathbf{I} + \mathbf{x}(k)\mathbf{x}^T(k)]^{-1}. \tag{11}$$

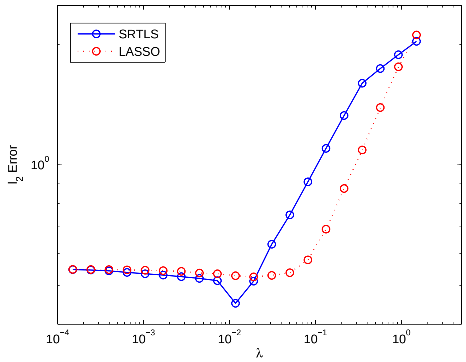
This completes one update, and the algorithm will terminate once the difference between two consecutive iterations becomes smaller than a given threshold.

In this iterative algorithm, both updates (9) or (10) at iteration k may either improve or maintain, but cannot worsen, the SRTLS cost function. Thus, monotonous convergence of the (bounded, non-negative) cost function is established. Also the alternating descent algorithm will at least converge to a stationary point for the SRTLS problem (4), while the limit point most likely depends on the initialization. By setting $\mathbf{E}(0) = \mathbf{0}_{m\times n}$, $\mathbf{x}(0)$ is equivalent to the LASSO solution. This is a good starting point, since it was shown in [9] that even with errors in both $\mathbf{A}$ and $\mathbf{y}$, the stability of the recovered signal of the regular compressive sensing techniques is still guaranteed.

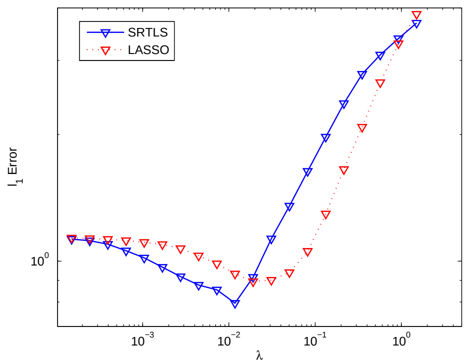
4. NUMERICAL EXAMPLES

In this section, three simulated tests are presented to illustrate the merits of the SRTLS approach. First, a general model is considered, and then the specific sensing examples of Section 2 are studied.

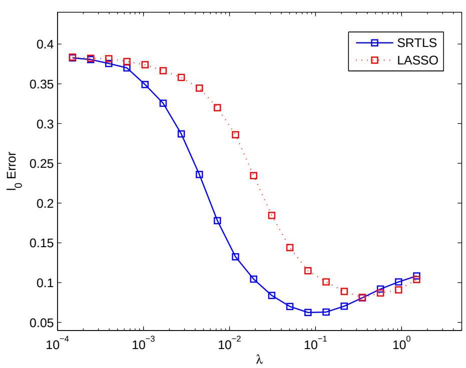
Test Case 1: (General Model.) This test is based on a general setup where in each trial a new matrix $\mathbf{A}$ of size 20×40 was randomly generated with independent normally distributed entries of variance $1/20$ (so that the expected ℓ_2 norm of each column was unity). The entries of the noise terms $\mathbf{E}$ and $\mathbf{e}$ in (3) were also generated using a Gaussian distribution with a standard deviation that is 5% of the one of $\mathbf{A}$, thus the equivalent entry-wise signal-to-noise ration (SNR) is 26dB. A vector $\mathbf{x}$ with 5 nonzero entries was then randomly generated with nonzero entries taken from a unit Gaussian distribution. We compare the SRTLS solution with the LASSO one for 20 values of λ uniformly distributed in the log-scale. The comparison is done in terms of the ℓ_2, ℓ_1, and ℓ_0 errors with respect to (w.r.t.) the actual $\mathbf{x}$ value. (The ℓ_0 error is calculated by the percentage of entries where the support of the two vectors is different.) There are 50 trials in total.



(a)

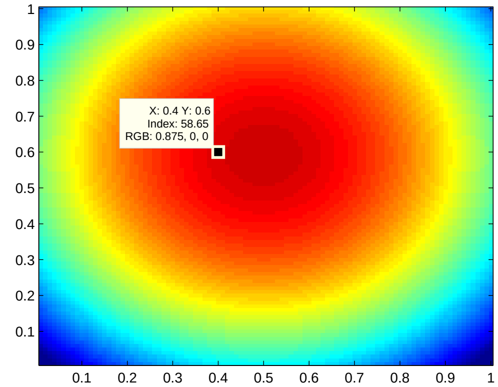


(b)

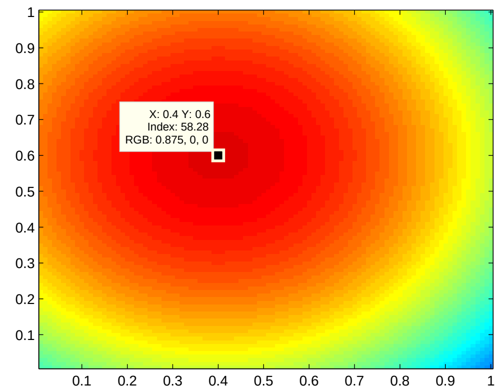


(c)

Fig. 2. Comparison between SRTLS and LASSO in terms of (a) ℓ_2 norm, (b) ℓ_1 norm, and (c) ℓ_0 norm of the estimation errors.



(a)



(b)

Fig. 3. Comparison between PSD maps estimated by (a) LASSO, and (b) SRTLS for the CR network in Fig. 1.

Fig. 2 depicts the improvement of the SRTLS results over the LASSO ones, especially in the ℓ_0 norm. With a moderate λ value range, the SRTLS solution is consistently better than the LASSO one in recovering the right support for $\mathbf{x}$, as shown in Fig. 2(c). Nevertheless, when λ gets large, both estimates tend to prefer the all-zero vector, so that the ℓ_0 norm becomes more or less the same, although the LASSO solution has a smaller error in ℓ_2 and ℓ_1. However, for both these error norms, the SRTLS is still slightly preferred with a moderate choice of λ.

Test Case 2: (Spectrum Sensing.) This simulation is performed with reference to the CR network in the region $[0\ 1] \times [0\ 1]$ as depicted in Fig. 1. The setup includes $N_r = 4$ CRs to estimate the PSD map in space and frequency, generated by one source located at $[0.4\ 0.6]$, in the center of four neighboring candidate locations on the grid. The CRs can scan 128 frequencies from 15MHz to 30MHz, and adopt the basis expansion model in (5) over the considered band comprising $N_b = 16$ rectangles of 1MHz width as frequency bases, and the single source only transmits over the 6th band. The average gains of the network links obey an exponential decaying model for γ_{sr} with $\delta = 1/2$. The received data are generated using the transmit PSD described earlier, a regular Rayleigh fading channel model with 6 taps, and additive white Gaussian receiver noise with the SNR at 0dB. The receiver PSD is obtained using the exponentially weighted moving average (EWMA) with exponent 0.99 of the periodogram estimate across 1000 coherent blocks.

We apply both the LASSO and SRTLS algorithms to solve the sparse reconstruction problem for the linear model (6). The penalty parameter λ is chosen following [1], since here the atoms are not orthogonal. Both algorithms identify the frequency band correctly. The LASSO algorithm detects two transmitting sources at the positions $[0.5\ 0.5]$ and $[0.5\ 0.7]$, which are among the four nearest neighboring grid points of the actual source at $[0.4\ 0.6]$. The proposed SRTLS further refines the source fitting by iteratively

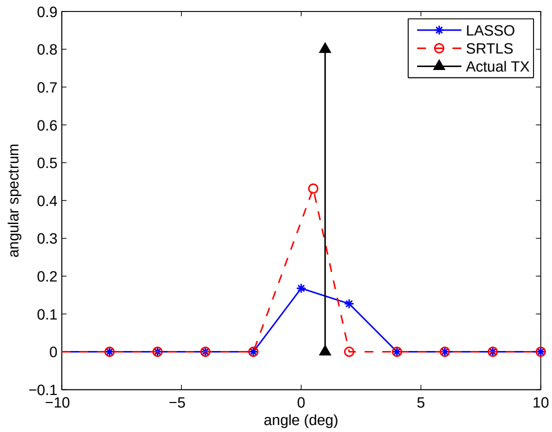


Fig. 4. Angular spectrum of the LASSO algorithm and the proposed SRTLS one, compared to the actual transmission pattern.

searching for a better sensing model and stacked source parameters. This algorithm terminates with only one source at the location $[0.5\ 0.5]$ and also a refined estimate of the corresponding atom, which is matched with all the atoms corresponding to 25 uniformly spread candidate points in the region $[0.3\ 0.7] \times [0.3\ 0.7]$. It is shown that the one corresponding to the point $[0.4\ 0.6]$ renders the best match w.r.t. the refined estimate, and with this information this algorithm recovers the source location exactly. In order to illustrate this difference, the estimated maps of the spatial PSDs over the 6th frequency band are plotted. This is done for the LASSO (Fig. 3 (a)), and the SRTLS with the re-calibrated source location (Fig. 3 (b)), where in both maps the marked point indicates the actual source location $[0.4\ 0.6]$. Relative to the LASSO map, the SRTLS map is more accurate and matches with the position of the source exactly.

Test Case 3: (DoA Estimation.) The ULA has $N_r = 8$ antenna elements and the number of scanning angles is $N_s = 90$, searching from -90° to 90° w.r.t. the array boresight. There is one single source of unit amplitude with its DoA $\theta = 1^\circ$, which is the mean of the two scanning angles $\theta_{45} = 0^\circ$ and $\theta_{46} = 2^\circ$. The noise power for each antenna is set to 0.01; i.e., the SNR is 20dB. The distance between two neighboring antennas is set to $d = (1/2)\lambda_s$.

Selecting λ as done in Test Case 2, the LASSO solution yields two nonzero entries at both $\theta_{45} = 0^\circ$ and $\theta_{46} = 2^\circ$, while the SRTLS gives one nonzero entry at $\theta_{45} = 0^\circ$. We further match the estimated array steering vector $\hat{\psi}(\theta_{45})$ to the ones corresponding to 20 uniformly sampled angles in the region $[-2^\circ\ 2^\circ]$, and this refines the source DoA to be $\hat{\theta} = 0.5^\circ$. We compare the estimated angle spectrum using the LASSO algorithm and the SRTLS algorithm with refinement in Fig. 4. The black arrow denotes the actual source location from the direction 1° and serves as a benchmark to the true angular spectrum. The proposed SRTLS solver can clearly identify the single source and provides a good approximation to the true source angle, while the LASSO solution achieves the peak at two angles near the actual one.

5. CONCLUSIONS AND CURRENT RESEARCH

In this paper, we have looked at a problem that as far as we know has not yet been considered in the sparse reconstruction literature. More specifically, we have studied sparse reconstruction for applications where both the observations and the dictionary are noisy. A local solution to this problem has been obtained by formulating it as an SRTLS problem and applying an alternating descent algorithm. First of all, we have shown by simulations that if the dictionary is noisy, SRTLS is capable of reaching a better solution than LASSO, especially in terms of the ℓ_1 and ℓ_0 norm. We have next applied our SRTLS algorithm to some illustrative spectrum sensing and direction estimation problems, leading to an improved spectrum (frequency or angular) estimate compared to LASSO. The main reason for this is that the SRTLS approach is capable of shifting the grid points to the actual locations (in any domain) of the sources.

Currently, our research is aimed at analyzing the stationary point that is reached by the proposed SRTLS algorithm, and comparing it to the LASSO results. Further, we are studying methods to obtain the global solution to the SRTLS problem.

6. REFERENCES

[1] D. L. Donoho S. S. Chen and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. on Scientific Computing*, vol. 20, pp. 33–61, Jan. 1998.

[2] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Information Theory*, vol. 52, pp. 5406–5425, Dec. 2006.

[3] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal. Statist. Soc. B*, vol. 58, pp. 267–288, 1996.

[4] J.-A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Trans. on Signal Process.*, Mar. 2010, to appear.

[5] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proc. CROWNCOM*, Mykonos Island, Greece, June 2006, pp. 1–5.

[6] Z. Quan, S. Cui, A. Sayed, and H. V. Poor, "Spatial-spectral joint detection for wideband spectrum sensing in cognitive radio networks," in *Proc. IEEE 56th Intl. Conf. Acoustics, Speech, Signal Processing*, Las Vegas, NV, Mar. 2008, pp. 2793–2796.

[7] D. M. Malioutov, M. Cetin, and A.S. Willskys, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. on Signal Process.*, vol. 53, pp. 3010–3022, Aug. 2005.

[8] J. Zheng, M. Kaveh, and H. Tsuji, "Sparse spectral fitting for direction of arrival and power estimation," in *Proc. IEEE 15th Workshop on Statistical Signal Processing*, Cardiff, Wales, UK., Aug. 2009, pp. 429–432.

[9] M. A. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressive sensing," *IEEE J. of Selected Topics in Signal Process.*, Apr. 2010, to appear.

[10] S.F. Cotter and B.D. Rao, "Application of total least squares (TLS) to the design of sparse signal representation dictionaries," in *Proc. Asilomar Conference on Signals, systems and Computers*, Pacific Grove, CA, USA, Nov. 2002, pp. 963–966.

[11] S. Van Huffel and J. Vandewalle, *The total least squares problem: Computational aspects and analysis*, vol. 9 of *Frontier in Applied Mathematics*, SIAM, Philadelphia, 1991.

[12] J. F. Sturm, "Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones," *Optimization Methods Software*.