Signal Processing 130 (2017) 204–216

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

c–LASSO and its dual for sparse signal estimation from array data

Christoph F. Mecklenbräuker [a,c,*], Peter Gerstoft [b], Erich Zöchmann [a,c]

[a] Institute of Telecommunications, TU Wien, Gusshausstr. 25/389, A-1040 Wien, Austria
[b] University of California San Diego, La Jolla, USA
[c] Christian Doppler Laboratory for Dependable Connectivity for the Society in Motion, Austria

ARTICLE INFO

Article history:
Received 11 February 2016
Received in revised form
27 June 2016
Accepted 28 June 2016
Available online 1 July 2016

Keywords:
Sparsity
c-LASSO
Duality theory
Homotopy

ABSTRACT

We treat the estimation of a sparse set of sources emitting plane waves observed by a sensor array as a complex-valued LASSO (c–LASSO) problem where the usual ℓ_1-norm constraint is replaced by the ℓ_1-norm of a matrix $\boldsymbol{D}$ times the solution vector. When the sparsity order is given, algorithmically selecting a suitable value for the c–LASSO regularization parameter remains a challenging task. The corresponding dual problem is formulated and it is shown that the dual solution is useful for selecting the regularization parameter of the c-LASSO. The solution path of the c-LASSO is analyzed and this motivates an order-recursive algorithm for the selection of the regularization parameter and a faster iterative algorithm that is based on a further approximation. This greatly facilitates computation of the c-LASSO-path as we can predict the changes in the active indices as the regularization parameter is reduced. Using this regularization parameter, the directions of arrival for all sources are estimated.

1. Introduction

This paper proposes the complex-valued LASSO[1] (c-LASSO [1]) to solve the sparse signal estimation problem from sensor array data. This work proposes an array data model selection procedure and high-resolution signal estimator by selecting a suitable value for the c-LASSO regularization parameter. The key results in this paper motivate an order-recursive algorithm for its solution and two faster iterative algorithms.

Sparse signal estimation techniques retrieve a signal vector from an undercomplete set of noisy measurements when the signal vector is assumed to have only few nonzero components at unknown positions. Research in this area was spawned by the Least Absolute Shrinkage and Selection Operator (LASSO) [2]. In the related field of compressed sensing, this sparse signal reconstruction problem is known as the atomic decomposition problem [3]. The early results for sparse signals [4–6] have been extended to compressible (approximately sparse) signals and sparse signals buried in noise [7–11] which enables application to problems in array processing.

We aim at selecting the best fitting configuration of plane wave sources which satisfactorily explain the observed array data set when the number of sources is constrained. Similar selection and fitting problems in array processing can be treated by, e.g., model order selection, multiple hypotheses testing, and cross-validation [12].

Here, the real-valued approach in [13] is applied to the complex-valued LASSO (c-LASSO, [1]) and the resulting fast LASSO-Path solvers [14] are shown to solve the signal estimation problem from complex-valued array data. The sparsity of the complex-valued solution requires simultaneous sparsity of the real and imaginary parts with identical support [1,15,16]. This relates the c-LASSO to real-valued block-sparse compressive sensing.

It is shown here that the corresponding dual vector is interpretable as the output of a weighted matched filter acting on the residuals of the linear observation model, cf. [17].

The regularization parameter μ in LASSO defines the trade-off between the model fit and the estimated sparsity order K given by the number of estimated nonzero signal components. When the sparsity order K_0 is given, choosing a suitable value for the LASSO regularization parameter μ remains a challenging task. The homotopy techniques [18,14,19] provide an approach to sweep over a range of μ values to select the signal estimate with the given K_0.

The maximum magnitudes of the dual vector can be used for selecting the regularization parameter of the c-LASSO. This is the basis for an order-recursive algorithm to solve the sparse signal reconstruction problem [19–21] for the given K_0. In this work, a fast and efficient choice of μ is proposed for Direction of Arrival (DOA) estimation from array data. The choice exploits the sidelobe levels of the array's beam pattern. We motivate this choice after proving several relations between the regularization parameter μ, the c-LASSO residuals, and the c-LASSO's dual solution.

* Corresponding author.
E-mail addresses: cfm@nt.tuwien.ac.at (C.F. Mecklenbräuker), gerstoft@ucsd.edu (P. Gerstoft), ezoechma@nt.tuwien.ac.at (E. Zöchmann).

[1] Abbreviation: c-LASSO = complex-valued Least Absolute Shrinkage and Selection Operator.

http://dx.doi.org/10.1016/j.sigpro.2016.06.029
0165-1684/

The main achievements of this work are summarized as follows: We extend the convex duality theory [13] from the real-valued to the complex-valued case and formulate the corresponding dual problem to the c-LASSO. We show that the dual solution is useful for selecting the regularization parameter of the c-LASSO. Three signal processing algorithms are formulated and evaluated to support our theoretical results and claims.

The developed methods have been applied to acoustic imaging both with single and multiple data snapshots [22]. Its root mean square error performance is superior to beamforming, minimum variance distortion-free response (MVDR), and multiple signal classification (MUSIC). Experiments with measured acoustic array data indicate that the developed algorithms resolve multiple coherent waves [22].

1.1. Notation

Matrices $\boldsymbol{A}$, $\boldsymbol{B}$, … and vectors $\boldsymbol{a}$, $\boldsymbol{b}$, … are complex-valued and denoted by boldface letters. The zero vector is $\boldsymbol{0}$. The Hermitian transpose, inverse, and Moore–Penrose pseudo inverse are denoted as $\boldsymbol{X}^H$, $\boldsymbol{X}^{-1}$, $\boldsymbol{X}^+$ respectively. We abbreviate $\boldsymbol{X}^{-H} = (\boldsymbol{X}^H)^{-1}$. The complex vector space of dimension N is written as $\mathbb{C}^N$. $\mathcal{N}(\boldsymbol{A})$ is the null space of $\boldsymbol{A}$ and $\mathrm{span}(\boldsymbol{A})$ denotes the linear hull of $\boldsymbol{A}$. The projection onto $\mathrm{span}(\boldsymbol{A})$ is $\boldsymbol{P}_{\boldsymbol{A}}$. The ℓ_p-norm is written as $\|\cdot\|_p$. For a vector $\boldsymbol{x} \in \mathbb{C}^M$, $\|\boldsymbol{x}\|_\infty = \max_{1\le m\le M}|x_m|$, for a matrix $\boldsymbol{X} \in \mathbb{C}^{N\times M}$, we define $\|\boldsymbol{X}\|_{\max} = \max_{1\le n\le N}\max_{1\le m\le M}|X_{nm}|$.

2. Problem formulation

We start from the following problem formulation: Let the complex-valued array data $\boldsymbol{y} \in \mathbb{C}^N$ and array steering matrix $\boldsymbol{A} \in \mathbb{C}^{N\times M}$ be given. Find the sparse signal $\boldsymbol{x}_{\ell_0} \in \mathbb{C}^M$ for given sparsity order $K_0 \in \mathbb{N}$ such that the squared data residuals are minimal,

$$\boldsymbol{x}_{\ell_0} = \underset{\boldsymbol{x}}{\mathrm{argmin}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{x}\|_0 \le K_0, \tag{P0}$$

where $\|\cdot\|_p$ denotes the ℓ_p-norm. The problem (P0) is known as complex-valued ℓ_0-reconstruction. It is *non-convex* and hard to solve [23]. Therefore, the ℓ_0-constraint in (P0) is commonly relaxed to an ℓ_1 constraint which renders the problem (P1) to be *convex*. Further, a matrix $\boldsymbol{D}$ is introduced in the formulation of the constraint which gives flexibility in the problem definition. Let the number of rows of $\boldsymbol{D}$ be arbitrary at first. In Section 3 suitable restrictions on $\boldsymbol{D}$ are imposed where needed. Several real-valued cases are discussed in [13]. Here, we generalize the c-LASSO [1] problem by introducing the $\boldsymbol{D}$ matrix in the constraint,[2]

$$\boldsymbol{x}_{\ell_1} = \underset{\boldsymbol{x}}{\mathrm{argmin}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{D}\boldsymbol{x}\|_1 \le \varepsilon. \tag{P1}$$

The sparsity of the complex-valued signal $\boldsymbol{x}$ implies joint sparsity of its real and imaginary parts with identical support [15,1]. This relates (P1) to real-valued block-sparse compressive sensing. Incorporating the ℓ_1 norm constraint into the objective function results in the equivalent formulation (P1′),

$$\boldsymbol{x}_{\ell_1} = \underset{\boldsymbol{x}}{\mathrm{argmin}} \left(\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu \|\boldsymbol{D}\boldsymbol{x}\|_1 \right). \tag{P1′}$$

The equivalence of (P0) and (P1′) requires suitable conditions to be satisfied such as the restricted isometry property (RIP) condition or mutual coherence condition imposed on $\boldsymbol{A}$, cf. [24,25,5]. Under such condition, the problems (P0) and (P1′) yield the same sparsity order, $K_0 = K$ with $K = \|\boldsymbol{x}_{\ell_1}\|_0$, if the regularization parameter μ in (P1′) is suitably chosen. The algorithms of Section 6 calculate suitable regularization parameters in this sense.

[2] If $\boldsymbol{D}$ is invertible then this is equivalent to the c-LASSO [1].

3. Dual problem to the c-LASSO

The c-LASSO problem [13,1] is written in constraint form, all vectors and matrices are assumed to be complex-valued. The following discussion is valid for arbitrary $N, M \in \mathbb{N}$: both the over-determined and the under-determined cases are included. Following [26,27], a vector $\boldsymbol{z} \in \mathbb{C}^M$ and an equality constraint $\boldsymbol{z} = \boldsymbol{D}\boldsymbol{x}$ are introduced to obtain the equivalent problem

$$\min_{\boldsymbol{x},\boldsymbol{z}} \left(\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu \|\boldsymbol{z}\|_1 \right) \quad \text{subject to} \quad \boldsymbol{z} = \boldsymbol{D}\boldsymbol{x}. \tag{1}$$

The complex-valued dual vector $\boldsymbol{u} = (u_1, \ldots, u_M)^T$ is introduced and associated with this equality constraint. The corresponding Lagrangian is

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu \|\boldsymbol{z}\|_1 + \mathrm{Re}\left[\boldsymbol{u}^H (\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z}) \right] \tag{2}$$

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}) = \mathcal{L}_1(\boldsymbol{x}, \boldsymbol{u}) + \mathcal{L}_2(\boldsymbol{z}, \boldsymbol{u}). \tag{3}$$

To derive the dual problem, the Lagrangian is minimized over $\boldsymbol{x}$ and $\boldsymbol{z}$. The terms involving $\boldsymbol{x}$ are

$$\mathcal{L}_1(\boldsymbol{x}, \boldsymbol{u}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mathrm{Re}\left(\boldsymbol{u}^H \boldsymbol{D}\boldsymbol{x} \right). \tag{4}$$

The terms in (2) involving $\boldsymbol{z}$ are

$$\mathcal{L}_2(\boldsymbol{z}, \boldsymbol{u}) = \mu \|\boldsymbol{z}\|_1 - \mathrm{Re}(\boldsymbol{u}^H \boldsymbol{z}). \tag{5}$$

The value $\hat{\boldsymbol{x}}$ minimizing (4) is found by differentiation, $\partial\mathcal{L}_1/\partial\boldsymbol{x} = 0$. This gives

$$\boldsymbol{D}^H \boldsymbol{u} = 2\boldsymbol{A}^H \left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}} \right) \tag{6}$$

whereby

$$\boldsymbol{A}^H \boldsymbol{A}\hat{\boldsymbol{x}} = \boldsymbol{A}^H \boldsymbol{y} - \frac{1}{2}\boldsymbol{D}^H \boldsymbol{u}. \tag{7}$$

If $\boldsymbol{D}^H\boldsymbol{u} \in \mathrm{span}(\boldsymbol{A}^H)$ the solution to (7) is expressible as a sum of three terms,

$$\hat{\boldsymbol{x}} = \underbrace{\boldsymbol{A}^+\boldsymbol{y} + \boldsymbol{\xi}}_{\hat{\boldsymbol{x}}_{\mathrm{LS}}} - \frac{1}{2}(\boldsymbol{A}^H\boldsymbol{A})^+ \boldsymbol{D}^H \boldsymbol{u}, \tag{8}$$

where $(\cdot)^+$ denotes the Moore–Penrose pseudoinverse. The Moore–Penrose pseudoinverse $\boldsymbol{X}^+$ is defined and unique for all matrices $\boldsymbol{X}$. The sole purpose of (8) is to provide a geometrical interpretation of the three terms as illustrated in Fig. 1 and explained below. In the following, we assume that $\boldsymbol{A}$ has full row-rank and $\boldsymbol{A}^+ = \boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{A}^H)^{-1}$ is a right-inverse [28]. Here, $\boldsymbol{\xi} \in \mathcal{N}(\boldsymbol{A})$ is any vector in the nullspace $\mathcal{N}(\boldsymbol{A}) = \{\boldsymbol{\xi} \in \mathbb{C}^M | \boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{0}\}$. By setting $\boldsymbol{\xi} = \boldsymbol{\xi}_{\ell_1} = \boldsymbol{x}_{\ell_1} - \boldsymbol{A}^+\boldsymbol{y} + \frac{1}{2}(\boldsymbol{A}^H\boldsymbol{A})^+\boldsymbol{D}^H\boldsymbol{u} \in \mathcal{N}(\boldsymbol{A})$, we specialize (8) to the solution of (P1′),

$$\boldsymbol{x}_{\ell_1} = \boldsymbol{A}^+\boldsymbol{y} + \boldsymbol{\xi}_{\ell_1} - \tfrac{1}{2}(\boldsymbol{A}^H\boldsymbol{A})^+\boldsymbol{D}^H\boldsymbol{u}. \tag{9}$$

Thus, the solution (9) to the c-LASSO problem (P1′) consists of three terms, as illustrated in Fig. 1. The first two terms are the least norm solution $\boldsymbol{A}^+\boldsymbol{y}$ and the nullspace solution $\boldsymbol{\xi}_{\ell_1}$ which together form the unconstrained least squares (LS) solution $\hat{\boldsymbol{x}}_{\mathrm{LS}}$. The third term in (9) is associated with the dual solution and lies in the same subspace as $\boldsymbol{x}_{\text{least norm}}$, namely $\mathrm{span}(\boldsymbol{A}^H)$, i.e., the row space of $\boldsymbol{A}$. Fig. 2 shows the three terms of (9) individually for a simple array-processing scenario. The continuous angle φ is discretized uniformly in $[-90, 90]^\circ$ using 361 samples and the wavefield is

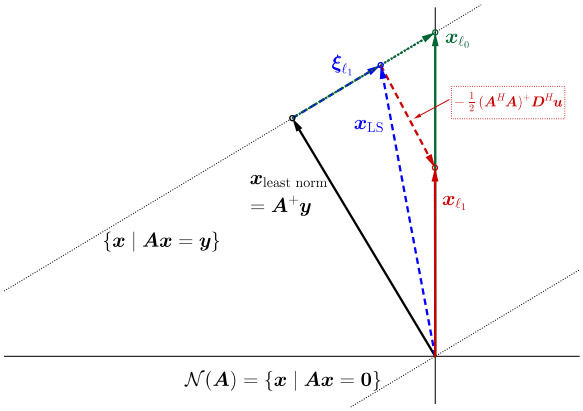

Fig. 1. Sketch of the relations between the primal solution and the terms in (9): least norm solution $\boldsymbol{x}_{\text{least norm}}$, least squares solution $\boldsymbol{x}_{\text{LS}}$, and the sparse solutions $\boldsymbol{x}_{\ell_0}$, $\boldsymbol{x}_{\ell_1}$. The nullspace term $\boldsymbol{\xi}_{\ell_1}$ is any vector along the line perpendicular to $\text{span}(\boldsymbol{A}^+) = \text{span}\left(\boldsymbol{A}^H\right)$. The red dashed arrow represents the last term in (9) which is perpendicular to $\boldsymbol{\xi}_{\ell_1}$, but need not be parallel to $\boldsymbol{x}_{\text{least norm}}$. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

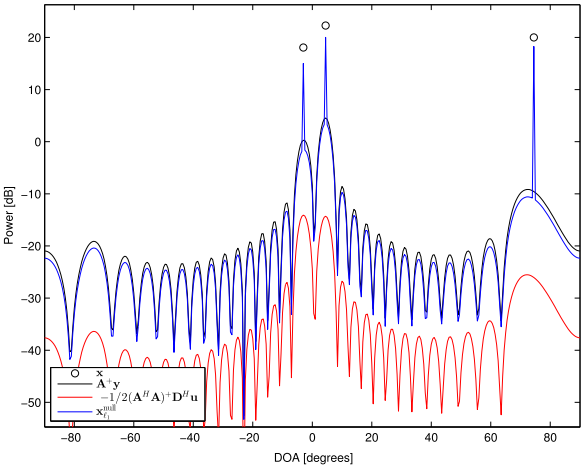

Fig. 2. Numerical example of the solution terms in Eq. (9) versus Direction of Arrival (DOA).

observed by 30 sensors resulting in a complex-valued 30×361 $\boldsymbol{A}$ matrix (see Section 4.1). At those indices m which correspond to DOAs at -3°, 4. 5° and 74. 5° in Fig. 2, the three terms in (9) sum constructively giving a non-zero x_m (we say: "the mth source position is active"), while for all other entries they interfere destructively. Constructive interference is illustrated in Fig. 1 which is in contrast to the destructive interference when the three terms in (9) sum to zero. This is formulated rigorously in Corollary 1.

We evaluate (4) at the minimizing solution $\hat{\boldsymbol{x}}$ and express the result solely by the dual $\boldsymbol{u}$. Firstly, we expand

$$\|\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}\|_2^2 = \|\boldsymbol{y}\|_2^2 + \|\boldsymbol{A}\hat{\boldsymbol{x}}\|_2^2 - 2\,\text{Re}\{\boldsymbol{y}^H\boldsymbol{A}\hat{\boldsymbol{x}}\} \tag{10}$$

Secondly using (6),

$$\begin{aligned}\boldsymbol{u}^H\boldsymbol{D}\hat{\boldsymbol{x}} &= (\boldsymbol{D}^H\boldsymbol{u})^H\hat{\boldsymbol{x}} = 2(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}})^H\boldsymbol{A}\hat{\boldsymbol{x}} \\ &= 2\boldsymbol{y}^H\boldsymbol{A}\hat{\boldsymbol{x}} - 2\,\|\boldsymbol{A}\hat{\boldsymbol{x}}\|_2^2\end{aligned} \tag{11}$$

Adding Eq. (10) and the real part of (11) gives

$$\begin{aligned}\mathcal{L}_1(\hat{\boldsymbol{x}}, \boldsymbol{u}) &= \|\boldsymbol{y}\|_2^2 - \|\boldsymbol{A}\hat{\boldsymbol{x}}\|_2^2 \\ &= \boldsymbol{y}^H\boldsymbol{y} - \left\|\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{D}}^H\boldsymbol{u}\right\|_2^2,\end{aligned} \tag{12}$$

where we used (8) which assumes $\boldsymbol{D}^H\boldsymbol{u} \in \text{span}(\boldsymbol{A}^H)$ and introduced the abbreviations

$$\tilde{\boldsymbol{D}} = \tfrac{1}{2}\boldsymbol{D}\boldsymbol{A}^+, \tag{13}$$

$$\tilde{\boldsymbol{y}} = \boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{y}, \quad \text{with } \boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{A}^+ \tag{14}$$

The statement $\boldsymbol{v} \in \text{span}(\boldsymbol{A}^H)$ is equivalent to $\boldsymbol{U}^H\boldsymbol{v} = \boldsymbol{0}$, where $\boldsymbol{U}$ is a unitary basis of the null space $\mathcal{N}(\boldsymbol{A})$. With $\boldsymbol{v} = \boldsymbol{D}^H\boldsymbol{u}$, this becomes $(\boldsymbol{D}\boldsymbol{U})^H\boldsymbol{u} = \boldsymbol{0}$, resulting in

$$\inf_{\boldsymbol{x}} \mathcal{L}_1(\boldsymbol{x}, \boldsymbol{u}) = \begin{cases} \boldsymbol{y}^H\boldsymbol{y} - \left\|\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{D}}^H\boldsymbol{u}\right\|_2^2, & \text{if } (\boldsymbol{D}\boldsymbol{U})^H\boldsymbol{u} = \boldsymbol{0}, \\ -\infty, & \text{otherwise.} \end{cases} \tag{15}$$

Next (5) is minimized with respect to $\boldsymbol{z}$, see Appendix A,

$$\inf_{\boldsymbol{z}} \mathcal{L}_2(\boldsymbol{z}, \boldsymbol{u}) = \begin{cases} 0, & \text{if } \|\,\boldsymbol{u}\,\|_\infty \le \mu \\ -\infty, & \text{otherwise.} \end{cases} \tag{16}$$

Combining the results and conditions of (15) and (16), we obtain the dual objective function and the dual constraints, i.e., the formulation of the dual problem to (P1),

$$\max_{\boldsymbol{u} \in \mathbb{C}^M} \left(\boldsymbol{y}^H\boldsymbol{y} - \left\|\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{D}}^H\boldsymbol{u}\right\|_2^2 \right) \tag{17a}$$

$$\text{subject to} \quad \|\,\boldsymbol{u}\,\|_\infty \le \mu, \tag{17b}$$

$$(\boldsymbol{D}\boldsymbol{U})^H\boldsymbol{u} = \boldsymbol{0}. \tag{17c}$$

Eq. (6) is solvable for $\boldsymbol{u}$ if the row space constraint (17c) is fulfilled. In this case, solving (6) directly gives

Result 1. If $\boldsymbol{D}$ is non-singular, the dual vector $\boldsymbol{u}$ is the output of a weighted matched filter acting on the vector of residuals, i.e.,

$$\boldsymbol{u} = 2\boldsymbol{D}^{-H}\boldsymbol{A}^H(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\ell_1}), \tag{18}$$

where $\boldsymbol{x}_{\ell_1}$ is the c-LASSO solution (P1′).

The dual vector $\boldsymbol{u}$ gives an indication of the sensitivity of the primal solution to small changes in the constraints of the primal problem (cf. [26]: Section 5.6). For the real-valued case the solution to (P1′) is more easily constructed and better understood via the dual problem [13]. Result 1 asserts a linear one-to-one relation between the corresponding dual and primal solution vectors also in the complex-valued case. Thus, any results formulated in the primal domain are readily applicable in the dual domain. This allows a more fundamental interpretation of sequential Bayesian approaches to density evolution for sparse source reconstruction [20,21]: they can be rewritten in a form that shows that they solve a c-LASSO problem and its dual. It turns out that the posterior probability density is strongly related to the dual solution [21,29].

The following corollaries clarify useful element-wise relations between the primal and dual solutions: Corollary 1 relates the *magnitudes* of the corresponding primal and dual variables. Further, Corollary 2 certifies what conditions on $\boldsymbol{D}$ are sufficient for guaranteeing that the *phase angles* of the corresponding primal and dual variables are equal. Finally, Corollary 3 states that *both* the primal and the dual solutions to (P1′) are piecewise linear in the regularization parameter μ.

Corollary 1. *For a diagonal matrix $\boldsymbol{D}$ with real-valued positive diagonal entries, we conclude: If the mth primal coordinate is active, i.e., $x_{\ell_1,m} \neq 0$ then the box constraint (17b) is tight in the mth dual coordinate. Formally,*

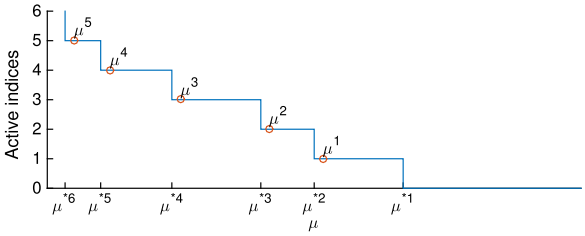

Fig. 3. Illustration of the c-LASSO path: Number of active indices versus the regularization parameter μ. Increments in the active set occur at μ^{*p}.

$$x_{\ell_1,m} \neq 0 \quad \Rightarrow \quad |u_m| = \mu, \qquad (m = 1, \ldots, M). \tag{19}$$

The proof is given in Appendix B.

Thus, the mth dual coordinate hits the boundary as the mth primal coordinate becomes active. Conversely, when the bound on $|u_m|$ is loose (i.e., the constraint on u_m is *inactive*), the corresponding primal variable x_m is zero (the mth primal coordinate is *inactive*). The active set $\mathcal{M}$ is

$$\mathcal{M} = \left\{ m | x_{\ell_1,m} \neq 0 \right\} \subseteq \left\{ m | |u_m| = \mu \right\} = \mathcal{U}. \tag{20}$$

Here, we have also defined the dual active set $\mathcal{U}$ which is a superset of $\mathcal{M}$ in general. This is due to Corollary 1 which states an implication in (19) only, but not an equivalence. The active set $\mathcal{M}$ implicitly depends on the choice of μ in problem (P1′). Let $\mathcal{M}$ contain exactly K indices,

$$\mathcal{M} = \{m_1, \ m_2, \ \ldots, \ m_K\}. \tag{21}$$

The number of active indices versus μ is illustrated in Fig. 3. Starting from a large choice of regularization parameter μ and then decreasing, we observe incremental changes in the active set $\mathcal{M}$ at specific values μ^{*p} of the regularization parameter, i.e., the candidate points of the c-LASSO path [14]. The active set remains constant within the interval $\mu^{*p} > \mu > \mu^{*p+1}$. By decreasing μ, we tend to enlarge the sets $\mathcal{M}$ and $\mathcal{U}$ [13,14]. By Eq. (20), we see that $\mathcal{U}$ may serve as a relaxation of the set of active indices $\mathcal{M}$.

Corollary 2. *If matrix $\boldsymbol{D}$ is diagonal with real-valued positive diagonal entries, then the phase angles of the corresponding entries of the dual and primal solution vectors are equal.*

$$\arg(u_m) = \arg(x_{\ell_1,m}) = \theta_m, \quad \forall\, m \in \mathcal{M} \tag{22}$$

Corollary 3. *The primal and the dual solutions to the c-LASSO problem* (P1′) *are continuous and approximately piecewise linear in the regularization parameter $\mu > 0$. The changes in slope occur at those values for μ where the set of active indices $\mathcal{M}$ changes.*

The proofs for these corollaries are given in Appendix B.

3.1. Relation to the ℓ_0 solution

It is now assumed that $\mathcal{M}$ defines the indices of the K non-zero elements of the corresponding ℓ_0 solution. In other words: the ℓ_1 and ℓ_0 solutions share the same sparsity pattern. The ℓ_0 solution with sparsity order K is then obtained by regressing the K *active* columns of $\boldsymbol{A}$ to the data $\boldsymbol{y}$ in the least-squares sense. Let

$$\boldsymbol{A}_{\mathcal{M}} = [\boldsymbol{a}_{m_1}, \ \boldsymbol{a}_{m_2}, \ \ldots, \boldsymbol{a}_{m_K}], \tag{23}$$

where $\boldsymbol{a}_m$ denotes the mth column of $\boldsymbol{A}$. The ℓ_0 solution becomes (cf. Appendix C)

$$\boldsymbol{x}_{\ell_0,\mathcal{M}} = \boldsymbol{A}_{\mathcal{M}}^{+}\boldsymbol{y}. \tag{24}$$

Here, $\boldsymbol{A}_{\mathcal{M}}^{+} = (\boldsymbol{A}_{\mathcal{M}}^H\boldsymbol{A}_{\mathcal{M}})^{-1}\boldsymbol{A}_{\mathcal{M}}^H$ is the left inverse of $\boldsymbol{A}_{\mathcal{M}}$. By subtracting (9) from (24) and restricting the solutions to the contracted basis $\boldsymbol{A}_{\mathcal{M}}$ yields

$$\boldsymbol{A}_{\mathcal{M}}(\boldsymbol{x}_{\ell_0,\mathcal{M}} - \boldsymbol{x}_{\ell_1,\mathcal{M}}) = \tfrac{1}{2}\boldsymbol{A}_{\mathcal{M}}\left(\boldsymbol{A}_{\mathcal{M}}^H\boldsymbol{A}_{\mathcal{M}}\right)^{+}\boldsymbol{D}_{\mathcal{M}}^H\boldsymbol{u}_{\mathcal{M}} = \tilde{\boldsymbol{D}}_{\mathcal{M}}^H\mu e^{j\theta_{\mathcal{M}}}, \tag{25}$$

where (13) was used. For the diagonal matrix $\boldsymbol{D}$, we conclude that the ℓ_0-solution (P0) and the c-LASSO solution (P1′) coincide in the image of $\boldsymbol{A}_{\mathcal{M}}$, if the c-LASSO problem is pre-informed (prior knowledge) by setting the mth column of $\boldsymbol{D}$, $m \in \mathcal{M}$ to zero. Such prior knowledge is obtainable by iterative re-weighting [30] or by a sequential algorithm on stationary sources [21,29].

4. Direction of arrival estimation

For the numerical examples, we model a uniform linear array (ULA), which is described with its steering vectors representing the incident wave for each array element.

4.1. Array data model

Let $\boldsymbol{x} = (x_1, \ldots, x_M)^{\mathrm{T}}$ be a vector of complex-valued source amplitudes. We observe time-sampled waveforms on an array of N sensors which are stacked in the vector $\boldsymbol{y}$. The following linear model for the narrowband sensor array data $\boldsymbol{y}$ at frequency ω is assumed,

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}. \tag{26}$$

The mth column of the transfer matrix $\boldsymbol{A}$ is the array steering vector $\boldsymbol{a}_m$ for hypothetical waves from DOA φ_{m}. To simplify the analysis all columns are normalized such that their ℓ_2 norm is one. The transfer matrix $\boldsymbol{A}$ is constructed by sampling all possible DOAs, but only few are active. Therefore, the dimension of $\boldsymbol{A}$ is $N \times M$ with $N \ll M$ and $\boldsymbol{x}$ is sparse. The linear model (26) is underdetermined.

The nmth element of $\boldsymbol{A}$ is modeled by

$$A_{nm} = \frac{1}{\sqrt{N}} \exp\left[\mathrm{j}(n-1)\pi \sin \varphi_m\right]. \tag{27}$$

Here $\varphi_m = \frac{(m-1)180^\circ}{M} - 90^\circ$ is the DOA of the mth hypothetical DOA to the nth array element.

The additive noise vector $\boldsymbol{n}$ is assumed spatially uncorrelated and follows a zero-mean complex normal distribution with diagonal covariance matrix $\sigma^2\boldsymbol{I}$.

Following a sparse signal reconstruction approach [13,21], this leads to minimizing the c-LASSO's Lagrangian

$$\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu\|\boldsymbol{D}\boldsymbol{x}\|_1, \tag{28}$$

where the weighting matrix $\boldsymbol{D}$ gives flexibility in the formulation of the penalization term in (28). Prior knowledge about the source vector leads to various forms of $\boldsymbol{D}$. This provides a Bayesian framework for sequential sparse signal trackers [20,21,29]. Specific choices of $\boldsymbol{D}$ encourage both sparsity of the source vector and sparsity of their successive differences which is a means to express that the source vector is locally constant versus DOA [31]. The minimization of (28) constitutes a convex optimization problem. Minimizing the c-LASSO's Lagrangian (28) with respect to $\boldsymbol{x}$ for given μ, gives a sparse source estimate $\boldsymbol{x}_{\ell_1}$. If $\mathrm{rank}(\boldsymbol{A}) < N$, (28) is no longer strictly convex and may not have a unique solution, cf. [13].

Earlier approaches formulated this as a (ordinary) c-LASSO problem [2,7,8] which is equivalent to (28) when specializing to $\boldsymbol{D} = \boldsymbol{I}$.

4.2. Basis coherence

The following examples feature different levels of basis coherence in order to examine the solution's behavior. As described

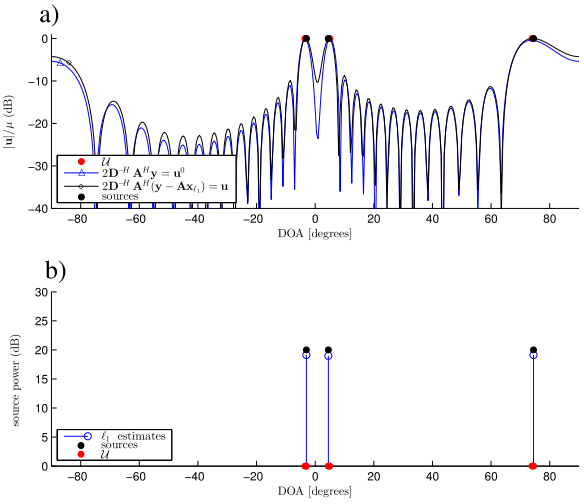

Fig. 4. Dual (a) and primal (b) solutions for 3 well separated sources at DOA -3°, 4.5°, 84.5° with low basis coherence. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

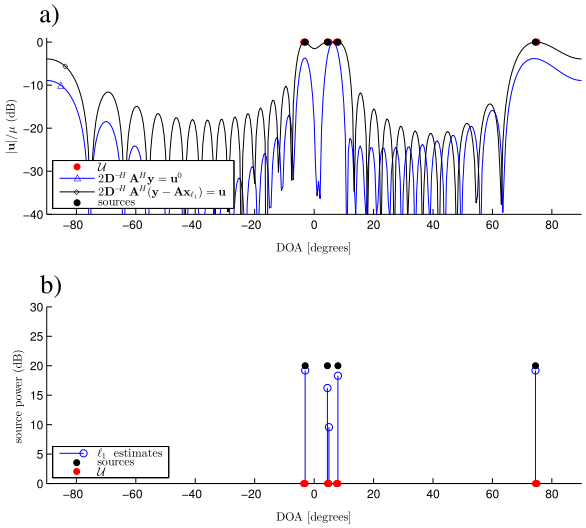

Fig. 5. Dual (a) and primal (b) solutions for 4 sources at DOA -3°, 4.5°, 8°, 84.5° with higher basis coherence.

in [32], the *basis coherence* is a measure of correlation between two steering vectors and defined as the inner product between atoms, i.e., the columns of $\boldsymbol{A}$. The maximum of these inner products is called mutual coherence and is customarily used for performance guarantees of recovery algorithms. To state the difference formally:

$$\mathrm{coh}(\boldsymbol{a_i}, \boldsymbol{a_j}) = \boldsymbol{a_i}^H \boldsymbol{a_j} \tag{29}$$

$$\text{mutual coh}(\boldsymbol{A}) = \|\boldsymbol{A}^H\boldsymbol{A} - \boldsymbol{I}\|_{\max} \tag{30}$$

The mutual coherence is bounded between 0 and 1

The following noiseless example in Figs. 4 and 5 demonstrates the dual solution for two different values of μ with $N=30$ and $M=361$. In Fig. 4, the c-LASSO with $\mu = 1$ is solved for a scenario with three sources at DOA -3°, 4.5°, 84.5° and all sources have same power level and are in-phase (see Fig. 4b). The dual solution $\boldsymbol{u}$ is shown for $\mu = 1$ (black, '$\diamond$') compared with the dual solution $\boldsymbol{u}^0$ (blue, '$\triangle$') for a very large choice of $\mu \geq 2 \| \boldsymbol{D}^{-H}\boldsymbol{A}^H\boldsymbol{y} \|_\infty$ which gives the trivial primal solution $\boldsymbol{x}_{\ell_1} = \boldsymbol{0}$. Fig. 5 shows the dual solutions for a scenario with an additional fourth source at 8°.

4.2.1. Low basis coherence

Fig. 4 shows the performance when the steering vectors of the active sources have small basis coherence. The basis of source 1 is weakly coherent with source 2, coh ≈ 0.02 using (29).

Fig. 4 a shows the normalized magnitude of the dual vector for a very large $\mu \geq 2 \| \boldsymbol{D}^{-H}\boldsymbol{A}^H\boldsymbol{y} \|_\infty$ (blue, '$\triangle$') and the normalized magnitude of the dual vector for $\mu = 1$ (black, '$\diamond$'). The dual active set $\mathcal{U}$ defined in (20) is depicted in red color. This figure shows that the true source parameters (DOA and power) are well estimated. It is also seen here that the behavior of the weighted matched filter closely resembles the magnitude of the dual vector and the weighted matched filter may be used as an approximation of the dual vector. This idea is further explored in Section 6.

4.2.2. High basis coherence

Fig. 5 a shows that the sources are not separable with the weighted matched filter, because the steering vectors belonging to sources 2 and 3 are coherent, coh = 0.61 using (29). The c-LASSO approach is still capable of resolving all four sources. The DOA region defined by $\mathcal{U}$ is much broader around the nearby sources, allowing for spurious peaks close to the true DOA. Fig. 5b shows that the true source locations (DOA) are still well estimated, but for the 2nd source from left, the power is split into two bins, causing a poor source estimate.

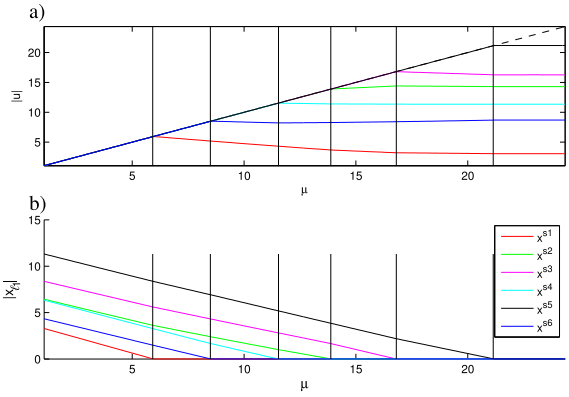

Fig. 6. Magnitudes of the solution paths versus μ for the simulation parameters in Table 1 and SNR = 40 dB: (a) dual, and (b) primal vectors for the case of the *complete basis*. The path is obtained by solving (P1') and (18) repeatedly for many values of μ.

5. Solution path

The c-LASSO solution path [13,14,1] gives the primal and dual solution vector versus the regularization parameter μ. The primal and dual trajectories are piece-wise smooth and related according to Corollaries 1–3. Figs. 6–9 show results from individual c-LASSO runs by varying μ.

The problem (P1) is complex-valued and the corresponding solution paths behave differently from the real-valued case [13]. Here, the behaviors of magnitude and phase for the primal and dual solution paths are given in Corollaries 1 and 2, whereas in the real-valued setting this reduces to results for the sign of the solutions. In the following figures, only the magnitudes of the primal variables containing a source and the corresponding dual solutions are illustrated. Note that Corollary 2 guarantees that the phase angle of each active primary solution its duals are identical and independent of μ.

Based on the observed solution paths, we notice that the hitting times (cf. [13], i.e., the μ-values when $|u_m| = \mu$ for an additional

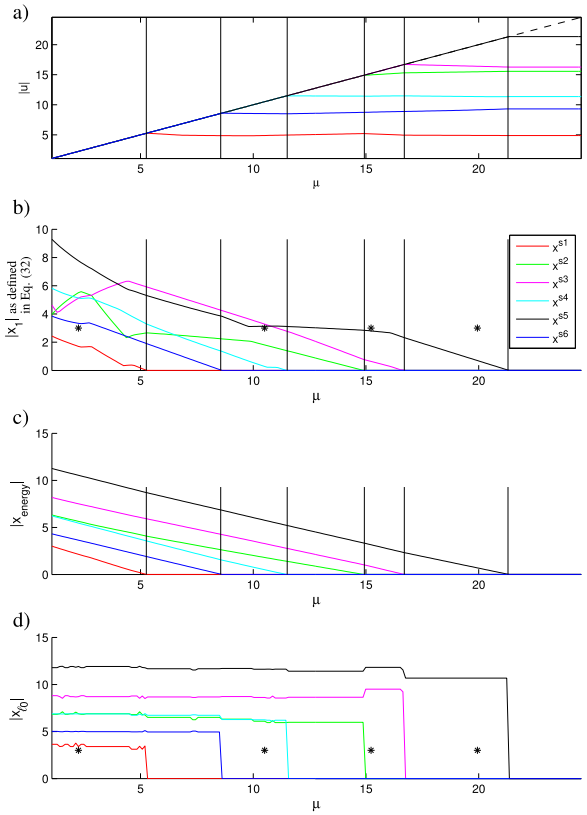

Fig. 7. Magnitudes of the solution paths versus μ for the simulation parameters in Table 1 and SNR = 40 dB: (a) dual, and (b, c and d) primal vectors for the case of an 80-vector *overcomplete basis*. For the primal solution, the peak within ± 2 bins from the true bin is tracked based on (b) maximum (c) energy. The magnitudes of the corresponding elements of $\boldsymbol{x}_{\ell_0}$ are shown in (d). The figure is produced by solving (P1') and (18) repeatedly for many values of μ.

m) of the dual variables (at lower μ) are well predictable from the solution at higher μ.

For the following simulations and Figs. 6–9, the signal to noise ratio (SNR) is defined as

$$\mathrm{SNR} = 10 \log_{10}\left(\mathrm{E}\, \|\boldsymbol{A}\boldsymbol{x}\|_2^2 / \mathrm{E}\, \|\boldsymbol{n}\|_2^2 \right) \mathrm{dB}. \tag{31}$$

Results for high SNR = 40 dB are shown in Figs. 6–8, 9a, c, whereas lower SNR = 20 dB is in Figs. 9b and d.

5.1. Complete basis

First (Fig. 6) discusses the dual and primal solution for a complete basis with $M=6$, sparsity order $K=6$, and $N=30$ sensors linearly spaced with half wavelength spacing. This simulation scenario is not sparse and all steering vectors $\boldsymbol{a}_m$ for $1 \le m \le M$ will eventually be used to reconstruct the data for small μ. The source parameters that are used in the simulation scenario are given in Table 1.

We discuss the solution paths in Figs. 6–9 from right ($\mu = \infty$) to left ($\mu = 0$). Initially all dual solution paths are horizontal (slope=0), since the primal solution $\boldsymbol{x}_{\ell_1} = \boldsymbol{0}$ for $\mu > 2 \| \boldsymbol{D}^{-H}\boldsymbol{A}^H\boldsymbol{y} \|_\infty$. In this strongly penalized regime, the dual vector is $\boldsymbol{u} = 2\boldsymbol{D}^{-H}\boldsymbol{A}^H\boldsymbol{y}$ which does not depend on μ.

At the point $\mu^1 = 2 \| \boldsymbol{D}^{-H}\boldsymbol{A}^H\boldsymbol{y} \|_\infty$ the first dual coordinate hits the boundary (17b). This occurs at $\mu^1 = 21$ in Fig. 6a and the corresponding primal coordinate becomes active. As long the active set $\mathcal{M}$ does not change, the magnitude of the corresponding dual coordinate is μ, due to Corollary 1. The remaining dual variables change slope relative to the basis coherence level of this index and the active set.

As μ decreases, the source magnitudes at the primal active indices increase since the ℓ_1-constraint in (P1′) becomes less important, see Fig. 6b. The second source will become active when the next dual coordinate hits the boundary (at $\mu^1 = 17$ in Fig. 6).

When the active set is constant, the primary and dual solution is piecewise linear with μ, as proved in Corollary 3. The changes in slope are quite gentle, as shown for the example in Fig. 6. Finally, at $\mu = 0$ the problem (P1′) degenerates to an unconstrained (underdetermined) least squares problem. Its primal solution $\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_{\mathrm{LS}}$, see (8), is not unique and the dual vector is trivial, $\boldsymbol{u} = \boldsymbol{0}$.

5.2. Overcomplete basis

We now enlarge the basis to $M=81$ with hypothetical source locations $\varphi_m \in [-20°, 20°]$ with $0.5°$ spacing, and all other parameters as before. The solution is thus sparse.

The c-LASSO path [14] is illustrated in Fig. 7 where we expect the source location estimate within ± 2 bins from the true source location. The dual Fig. 7a appears to be quite similar to Fig. 6a.

To obtain a simple figure, we plot the maximum magnitude of five adjacent bins near each true source, i.e., we define

$$|x_1|_m = \max(|x_{m-2}|, \ldots, |x_{m+2}|). \tag{32}$$

This shows that changes in slope at points not predicted by Corollary 3 (i.e. $\mu = 10$ in Fig. 7b).

Corollary 3 gives that the primary solution should change linearly, as demonstrated for the complete basis in Fig. 6b. Here we explain why this is not the case for the overcomplete basis primary solution in Fig. 7b. This is understood by examining the full solution at selected values of μ (asterisk (*) in Fig. 7). At $\mu = 20$ just one solution is active, only the black source (source 5) is active though one bin to the left, as shown in Fig. 8a2. The dual vector in Fig. 8a1–d1, has a broad maximum, explaining the sensitivity to offsets around the true DOA. The shape of this maximum is imposed by the dictionary; the more coherent the dictionary, the broader the maximum. Between $\mu = 16$ and $\mu = 11$, the black source appears constant, this is because at large values the source is initially located in a neighboring bin. As μ decreases, the true bin receives more power, see Fig. 8b2 and Fig. 8c2 for $\mu = 15$ and $\mu = 10$, respectively. When the bin is stronger than the neighboring bin at $\mu \le 11$, see Fig. 8d2, this source power starts increasing again. This trading in source power causes the fluctuations in Fig. 7b.

One way to correct for this fluctuation is to sum the coherent energy for all bins near a source, i.e., multiplying the source vector with the corresponding neighbor columns of $\boldsymbol{A}$, which also touch the boundary (marked region in Fig. 8) and then compute the energy based on the average received power at each sensor. This gives a steady rise in contributed power to the array data. This is observed in Fig. 7c which shows the steady increase in $|\mathbf{x}_{\mathrm{energy}}|_m$ defined by

$$|\mathbf{x}_{\mathrm{energy}}|_m = \left\| \sum_{j=-2}^{2} x_{m-j} \boldsymbol{a}_{m-j} \right\|_2 \tag{33}$$

for all $m \in \mathcal{M}$ when μ decreases.

We motivated solving (P1′) as a substitute for ℓ_0-reconstruction (P0)—finding the active indexes of the ℓ_1 solution, see Fig. 7d. The ℓ_0 primal can be found with the restricted basis and the value of the ℓ_1 primal from (8), which depends on μ, or by just solving (24).

To investigate the sensitivity to noise, 10 c-LASSO paths are simulated for 10 noise realizations for both SNR = 40 dB (Figs. 9a and c) and SNR = 20 dB (Figs. 9b and d). The primal $|\mathbf{x}_{\ell_1}|$ (Figs. 9a and b) show quite large variation with noise. This is because the

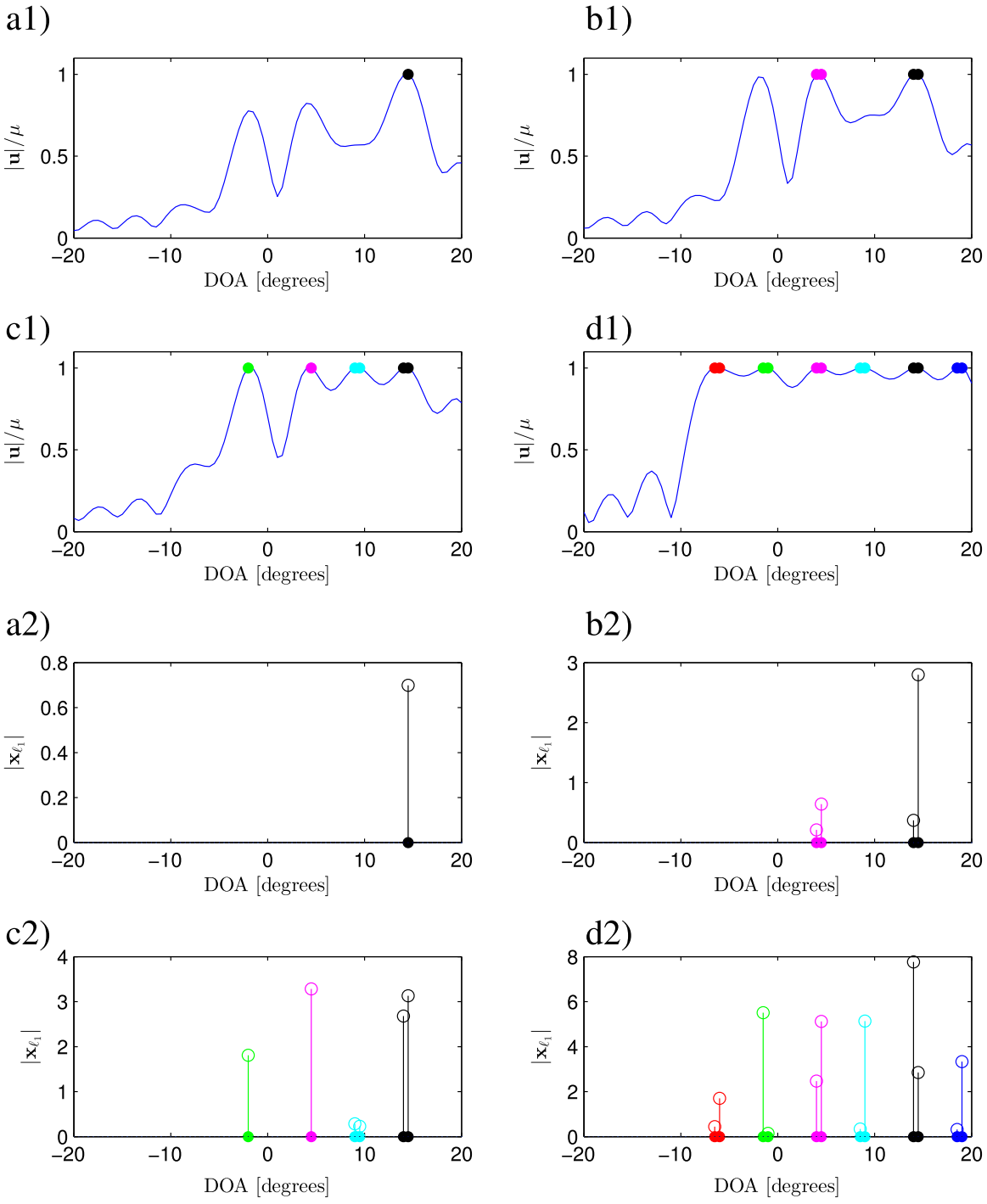

Fig. 8. Dual and primal solutions at selected values of μ for 81-vector overcomplete basis for SNR = 40 dB.

noise causes the active indexes to shift and thus the magnitude to vary. The mapping to energy $|\mathbf{x}_{\text{energy}}|_m$ defined in (33) is shown in Figs. 9c and d.

6. Solution algorithms

Motivated by Result 1 and Corollary 1, we propose the order-recursive algorithm in Table 2 for approximately solving problem (P0) by selecting a suitable regularization parameter μ in problem (P1′), a faster iterative algorithm in Table 3, and a dual-based iterative algorithm in Table 4.

As shown by Result 1, the dual vector is evaluated by a weighted matched filter acting on the c-LASSO residuals. The components of the dual vector which hit the boundary, i.e., $|u_m| = \mu$, correspond to the primal variables containing a source $|x_m| > 0$. As $|u_m| = \mu$ constitutes a necessary condition, this condition is at least $|\mathcal{M}|$ times fulfilled. Informally, we express this as: "The dual vector must have $|\mathcal{M}|$ peaks of height μ, where the shaping is defined by the dictionary $\boldsymbol{A}$ and the weighting matrix $\boldsymbol{D}$."

The key observation is the reverse relation. By knowing the peak magnitudes of the dual vector, one estimates the appropriate μ-value to make i peaks hit the boundary. We denote this regularization parameter value as $\mu^{(i)}$. This is a necessary condition to obtain i active sources.

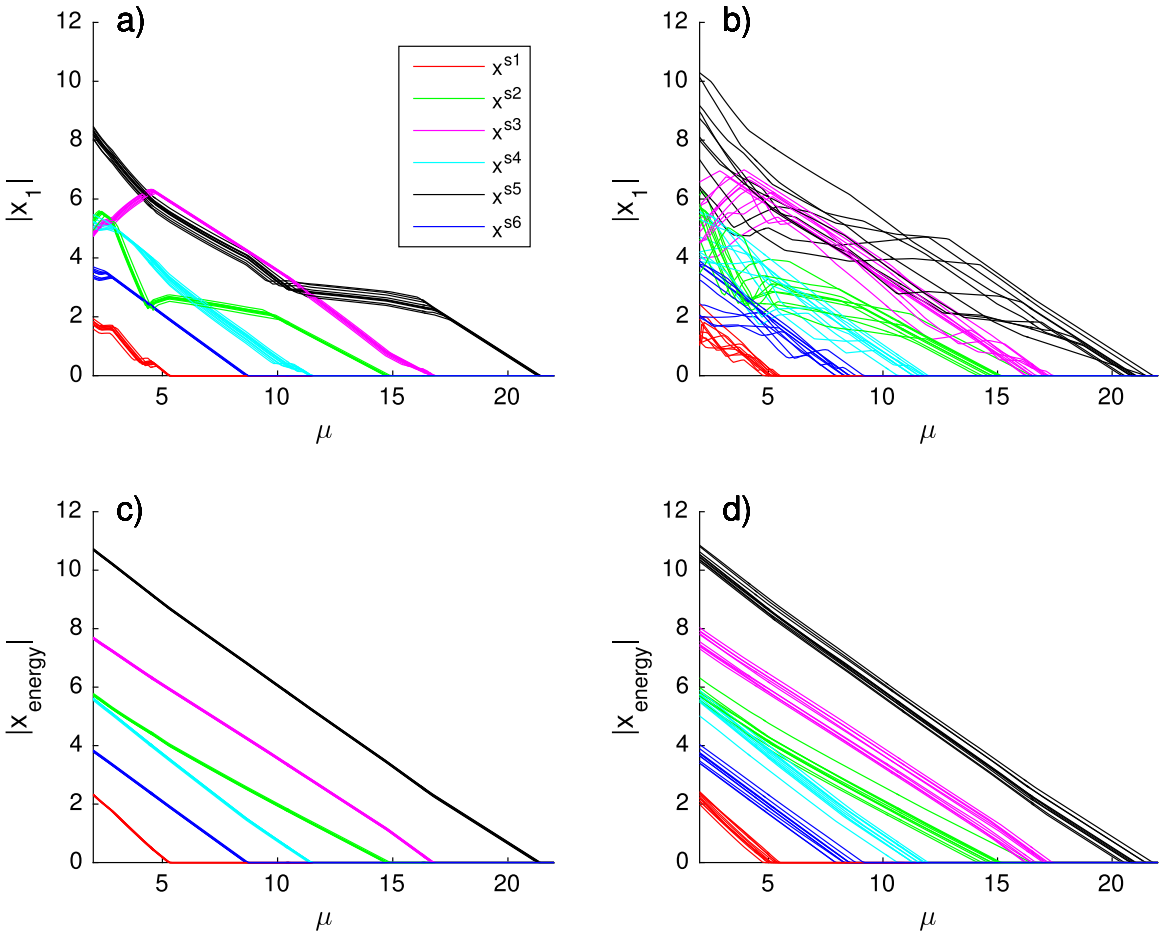

Fig. 9. For 10 noise realizations, magnitudes of the primal solution path versus μ are shown for the settings in Table 1 and an overcomplete basis, $M=81$. The peak within ± 2 bins from the true bin is tracked based on the maximum (top: (a) and (b)) or energy (bottom: (c) and (d)). Left ((a) and (c)): SNR=40 dB, right ((b) and (d)): SNR=20 dB.

Table 1
Source parameters for simulation scenario.

No.	DOA (°)	Power (lin.)
1	−6.0	4.0
2	−1.0	7.0
3	4.0	9.0
4	9.0	7.0
5	14.0	12.0
6	19.0	5.0

Table 2
Order-recursive algorithm to select μ for given sparsity order K_0.

Line	Given: $\boldsymbol{A} \in \mathbb{C}^{N\times M}$, $\boldsymbol{D} \in \mathrm{diag}\mathbb{R}^M$, $\boldsymbol{y} \in \mathbb{C}^N$ Given: $K_0 \in \mathbb{N}$, $F \in]0, 1[$, $\boldsymbol{x}_{\ell_1}$.
1	$\mathcal{M} = \{m \vert\vert x_{\ell_1,m}\vert > \delta_i\}$, $\delta_i = \epsilon \parallel \boldsymbol{x}^i_{\ell_1} \parallel_\infty$
2	$\boldsymbol{u}^{i-1} = 2\boldsymbol{D}^{-H}\boldsymbol{A}^H\left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\ell_1}\right)$
3	if $\vert\mathcal{M}\vert < K_0$
4	$\mathcal{U} = \left\{ m \middle\vert 1 - \frac{\vert u_m\vert}{\mu} < \epsilon_\mu \right\}$
5	$i = \vert\mathcal{U}\vert + 1$
6	$\mu = (1-F)\mathrm{peak}\left(\boldsymbol{u}^{i-1}, i\right) + F\mathrm{peak}\left(\boldsymbol{u}^{i-1}, i+1\right)$
7	else if $\vert\mathcal{M}\vert > K_0$
8	bisecting between μ^{i-1} and μ^i defined in (34)
9	end
10	Output: μ

Table 3
Iterative primal based algorithm to select μ for given sparsity order K_0.

Line	Given: $\boldsymbol{A} \in \mathbb{C}^{N\times M}$, $\boldsymbol{D} \in \mathrm{diag}\mathbb{R}^M$, $\boldsymbol{y} \in \mathbb{C}^N$ Given: $K_0 \in \mathbb{N}$, $F \in]0, 1[$, $\boldsymbol{x}_{\ell_1}$.
1	$\mathcal{M} = \{m \vert\vert x_{\ell_1,m}\vert > \delta_i\}$, $\delta_i = \epsilon \parallel \boldsymbol{x}^i_{\ell_1} \parallel_\infty$
2	$\boldsymbol{u}^{i-1} = 2\boldsymbol{D}^{-H}\boldsymbol{A}^H\left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\ell_1}\right)$
3	if $\vert\mathcal{M}\vert < K_0$
4	$\mathcal{U} = \left\{ m \middle\vert 1 - \frac{\vert u_m\vert}{\mu} < \epsilon_\mu \right\}$
5	$i = \vert\mathcal{U}\vert + 1$
6	$\mu = (1-F)\mathrm{peak}\left(\boldsymbol{u}^{i-1}, K\right) + F\mathrm{peak}\left(\boldsymbol{u}^{i-1}, K_0+1\right)$
7	else if $\vert\mathcal{M}\vert > K_0$
8	bisecting between μ^{i-1} and μ^i defined in (34)
9	end
10	Output: μ

We define the peak($\boldsymbol{u}$, i)–function which returns the ith largest local peak in magnitude of the vector $\boldsymbol{u}$. A local peak is defined as an element which is larger than its adjacent elements. The peak function can degenerate to a simple sorting function giving the ith largest value, this will cause slower convergence in the algorithms below.

Proposition 1. *Assuming all sources to be separated such that there is at least a single bin in between, the* peak *function relates the regularization parameter to the dual vector via*

Table 4
Iterative dual based algorithm to select μ for given sparsity order K_0.

Line	Given: $\boldsymbol{A} \in \mathbb{C}^{N\times M}$, $\boldsymbol{D} \in \mathrm{diag}\mathbb{R}^{M}$, $\boldsymbol{y} \in \mathbb{C}^{N}$ Given: $K_0 \in \mathbb{N}$, $F \in]0,1[$, $\boldsymbol{u}$.
1	$\mathcal{U} = \left\{ m \middle\vert 1 - \frac{\lvert u_m \rvert}{\mu} < \epsilon_\mu \right\}$
2	$\boldsymbol{x}_{\ell_0} = \boldsymbol{A}_{\mathcal{U}}^{+}\boldsymbol{y}$
3	$\mathcal{M} = \{m \lvert x_{\ell_0,m}\rvert > \delta\}$, $\delta = \epsilon \parallel \boldsymbol{x}_{\ell_0} \parallel_\infty$
4	if $\lvert\mathcal{M}\rvert < K_0$
5	$i = \lvert\mathcal{U}\rvert + 1$
6	$\mu = (1-F)\mathrm{peak}\left(\boldsymbol{u}^{i-1}, K\right) + F\mathrm{peak}\left(\boldsymbol{u}^{i-1}, K_0+1\right)$
7	Else if $\lvert\mathcal{M}\rvert > K_0$
8	Bisecting between μ^{i-1} and μ^{i}defined in (34)
9	End
10	Output: μ

$$\mu^i = \mathrm{peak}\left(\left\lvert \boldsymbol{u}(\mu^i)\right\rvert, i\right) = \mathrm{peak}\left(\boldsymbol{u}(\mu^i), i\right). \tag{34}$$

Eq. (34) is a fixed-point equation for μ^i which is demanding to solve. Therefore we approximate (34) with previously obtained dual vectors.[3] At a potential new source position n, the dual vector is expanded as

$$u_n(\mu^i) = \frac{2}{D^*_{n,n}}\boldsymbol{a}_n^H\left(\boldsymbol{y} - \sum_{m\in\mathcal{M}_i}\boldsymbol{a}_m x_{\ell_1,m}(\mu^i)\right) \tag{35}$$

$$u_n(\mu^i) \approx \frac{2}{D^*_{n,n}}\boldsymbol{a}_n^H\left(\boldsymbol{y} - \sum_{m\in\mathcal{M}_{i-1}}\boldsymbol{a}_m x_{\ell_1,m}(\mu^{i-1})\right) \tag{36}$$

$$u_n(\mu^i) \approx \frac{2}{D^*_{n,n}}\boldsymbol{a}_n^H\left(\boldsymbol{y} - \sum_{m\in\mathcal{M}_{i-2}}\boldsymbol{a}_m x_{\ell_1,m}(\mu^{i-2})\right) \tag{37}$$

$$\vdots$$

$$u_n(\mu^i) \approx \frac{2}{D^*_{n,n}}\boldsymbol{a}_n^H\boldsymbol{y} \tag{38}$$

The approximations used in (36)–(38) are progressive. These approximations are good if the steering vectors associated with the active set are sufficiently incoherent: $\lvert\boldsymbol{a}_n^H\boldsymbol{a}_m\rvert \approx 0$ for $n, m \in \mathcal{M}$. Eq. (38) corresponds to the conventional beamformer $\boldsymbol{A}^H\boldsymbol{y}$ for a single snapshot. In the solution algorithms, the approximations (36)–(38) are used for the (fast) selection of the regularization parameter μ only. Even if the approximations (36)–(38) do not hold, the bisection algorithm in Tables 2, 3, and 4 will guarantee the correct size of the active set. Thus, the accuracy of the approximations (36)–(38) affects the computation time only, but not the numerical accuracy of the $\boldsymbol{x}_{\ell_1}$ solution.

Our simulations have shown that a significant speed-up achievable, so we named it fast-iterative algorithm, cf Section 6.2. From the box constraint (17b), the magnitude of the ith peak in $\boldsymbol{u}$ does not change much during the iteration over i: It is bounded by the difference in regularization parameter. For any $\mu^i < \mu^{i-1}$, we conclude from (19) and (36) that

$$\underbrace{\mathrm{peak}(\boldsymbol{u}(\mu^{i-1}), i)}_{\le\mu^{i-1}} - \underbrace{\mathrm{peak}(\boldsymbol{u}(\mu^{i}), i)}_{=\mu^{i}} \le \mu^{i-1} - \mu^i. \tag{39}$$

Thus, the magnitude of the ith peak cannot change more than the corresponding change in the regularization parameter. The left hand side of (39) is interpretable as the prediction error of the regularization parameter and this shows that the prediction error is bounded.

Assuming our candidate point estimates (μ^{*1}, μ^{*2}, …) are correct, we follow a path of regularization parameters μ^1, μ^2, … where μ^p is slightly higher than the lower end μ^{*p+1} of the regularization interval. Specifically, $\mu^p = (1-F)\mu^{*p} + F\mu^{*p+1}$ with $F < 1$. For the numerical examples F=0.9 is used. This F is chosen because the primal solution $\boldsymbol{x}_{\ell_1}$ is closest to $\boldsymbol{x}_{\ell_0}$ at the lower end of the interval.

In the following we focus on the order recursive algorithm, and indicate the differences to the other approaches.

6.1. Recursive-in-order algorithm

The recursive-in-order algorithm in Table 2 finds one source at a time as μ is lowered. For this purpose it employs an approximation of the height of the ith local peak given a solution with $(i-1)$ peaks. The underlying assumption is that the next source will become active at the location corresponding to the dual coordinate of the next peak. Eq. (36) allows to approximate

$$\mu^i = \mathrm{peak}(\boldsymbol{u}(\mu^i), i) \approx \mathrm{peak}(\boldsymbol{u}(\mu^{i-1}), i). \tag{40}$$

This assumption is not universally valid as it may happen that the coordinate corresponding to the $(i+1)$th peak becomes active first, although $\mathrm{peak}(\boldsymbol{u}^{i-1}, i) > \mathrm{peak}(\boldsymbol{u}^{i-1}, i+1)$. In this case, two sources become active as the regularization parameter is chosen too low. This exception can be handled by, e.g., bisection in μ.

The recursive-in-order algorithm provided in Table 2 takes as input the dictionary $\boldsymbol{A}$, the generalization matrix $\boldsymbol{D}$, the measurement vector $\boldsymbol{y}$, the given sparsity order K_0 and the previous order c-LASSO solution $\boldsymbol{x}_{\ell_1}$. In line 1 the actual active set is determined by thresholding and line 2 produces the dual vector by Result 1. Line 2 can be omitted, if the c-LASSO solver makes the dual solution available, e.g., through primal–dual interior point methods or alternating direction method of multipliers. If the size of the active set of the previous c-LASSO solution is less than the given sparsity order K_0, the algorithm determines the dual active set $\mathcal{U}$ in line 4, cf. Eq. (20). The incremented cardinality of $\mathcal{U}$ is the new requested number of hitting peaks in the dual vector, cf. [13]. Finally, line 6 calculates μ based on the candidate point estimate (40).

6.2. Fast-iterative algorithm

The approximation in Eq. (40) is not limited to a single iteration. From (37)–(38) with (19), we extend (40) to

$$\begin{aligned}\mu^i &\approx \mathrm{peak}(\boldsymbol{u}(\mu^{i-1}), i)\\ &\approx \mathrm{peak}(\boldsymbol{u}(\mu^{i-2}), i)\\ &\approx \cdots\\ &\approx \mathrm{peak}(\boldsymbol{u}(\mu^{0}), i) = \mathrm{peak}(2\boldsymbol{D}^{-H}\boldsymbol{A}^H\boldsymbol{y}, i).\end{aligned} \tag{41}$$

This observation motivates the iterative algorithm in Table 3. The main difference to the recursive-in-order algorithm is found in line 6. The peakfinder estimates the maximum of the Kth peak. This leads to a significant speed-up, if sources are well separated and

[3] For the first step, we define $\boldsymbol{u}_0 = 2\boldsymbol{D}^{-H}\boldsymbol{A}^H\boldsymbol{y}$.

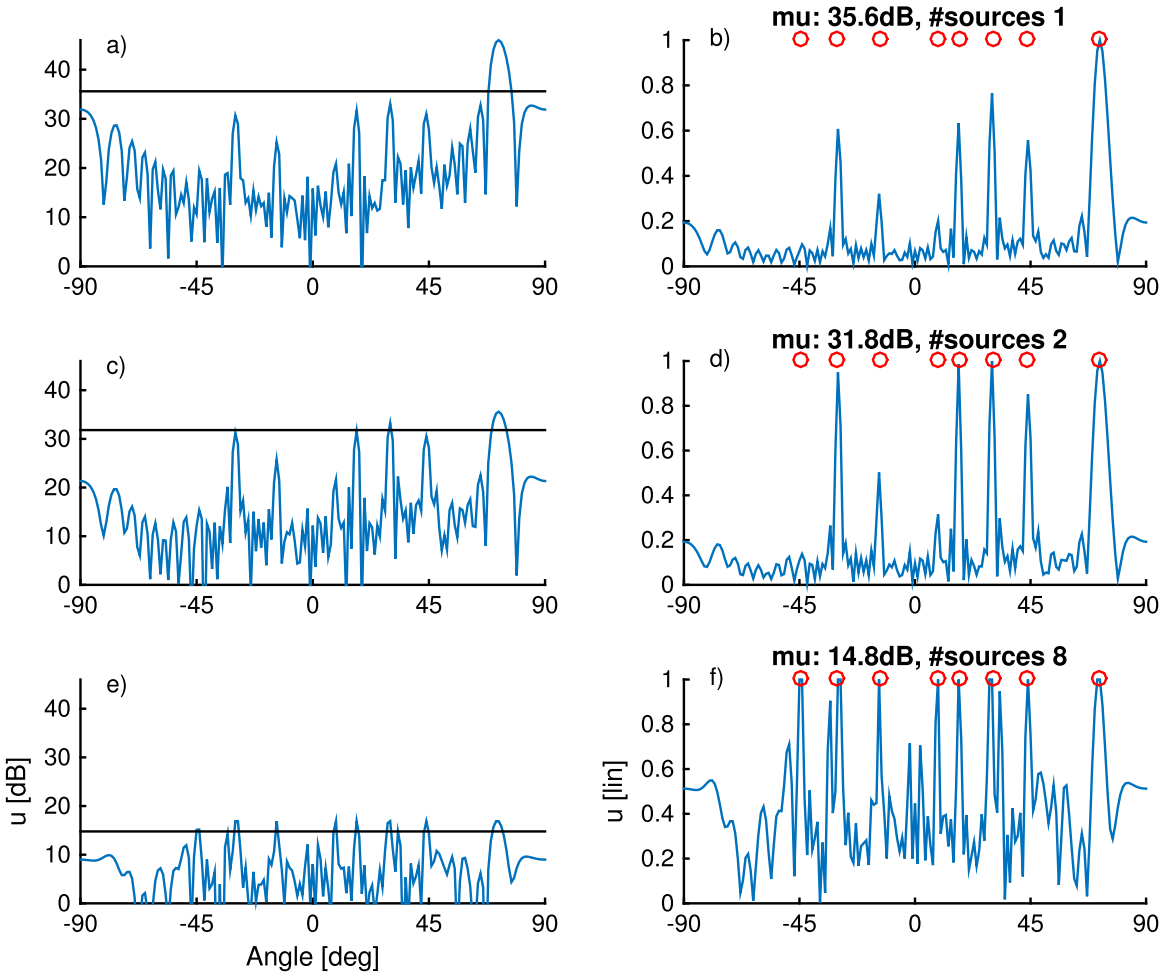

Fig. 10. Dual solution for order-recursive approach corresponding to step $i=1$ (a and b), $i=2$ (c and d), and $i=8$ (e and f). Left column: Dual (dB) for the previous step which is used for selecting μ (horizontal line). Right column: Dual (lin) normalized with μ (maximum is 1), the true source locations are marked with ○, and the actual value of μ and number of sources found is also indicated.

their basis coherence is low.

6.3. Detection in the dual domain

As a demonstrative example, we provide the fast iterative algorithm formulated *solely* in the dual domain in Table 4. Note that the gird-free atomic norm solutions [33–37] follow a similar approach.

As asserted by (20), searching for active indices in the dual domain is effectively a form of relaxation of the primal problem (P1′). This amounts to peak finding in the output of a weighted matched filter acting on the residuals, cf. Result 1. In line 1, the active set $\mathcal{M}$ is effectively approximated by the relaxed set $\mathcal{U}$. Therefore, the ℓ_0 solution is determined by regression on the relaxed set in line 2 and the primal active set is found by thresholding this solution in line 3. The remainder of the algorithm equals the primal based ones.

7. Simulation

In this section, the performance of the proposed dual estimation algorithms is evaluated based on numerical simulation. We use synthetic data from a uniform linear array with $N=64$ elements with half-wavelength spacing. The DOA domain is discretized by $\varphi_m = (m-1)\frac{180^\circ}{M} - 90^\circ$ with $m=1,\ldots,M$ and $M=180$. The simulation scenario has $K_0 = 8$ far-field plane-waves sources (26). The uncorrelated noise $\boldsymbol{n}$ is zero-mean complex-valued circularly symmetric normally distributed $\sim\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$, i.e., 0 dB power. Eight sources are stationary at $\varphi^{\mathrm{T}} = [-45, -30, -14, 9, 17, 30, 44, 72]$ degrees relative to broadside with constant power level (PL) $[-5, 10, 5, 0, 11, 12, 9, 25]$ dB [21].

The dual solution for the order-recursive approach, Table 2, corresponds to the results shown in Fig. 10. The faster iterative approach, Table 3, yields the results in Fig. 11. The dual solution using the primal solution from the previous iteration is interpreted as a weighted matched filter and used for the selection of μ (left column). Next, the convex optimization is carried out for that value of μ giving the dual solution. We plot the dual solution on a linear scale and normalized to a maximum value of 1 which is customary in implementations of the dual for compressed sensing [33–35]. The number of active sources (see the right column in Figs. 10 and 11) are determined according to line 1 in Tables 2 and 3.

For the order-recursive approach step 1, Fig. 10a, the μ is selected based on the main peak $\varphi = 72^\circ$ and a large side lobe at $\varphi = 80^\circ$. Once the solution for that μ is obtained it turns out that there is no an active source in the sidelobe.The solution progresses steadily down the c-LASSO path. Fig. 11 shows the faster iterative approach in Table 3 for the 8-source problem. In the first iteration we use a μ between the 8th and 9th peaks based on the weighted matched filter solution (Fig. 11a). There are many sidelobes

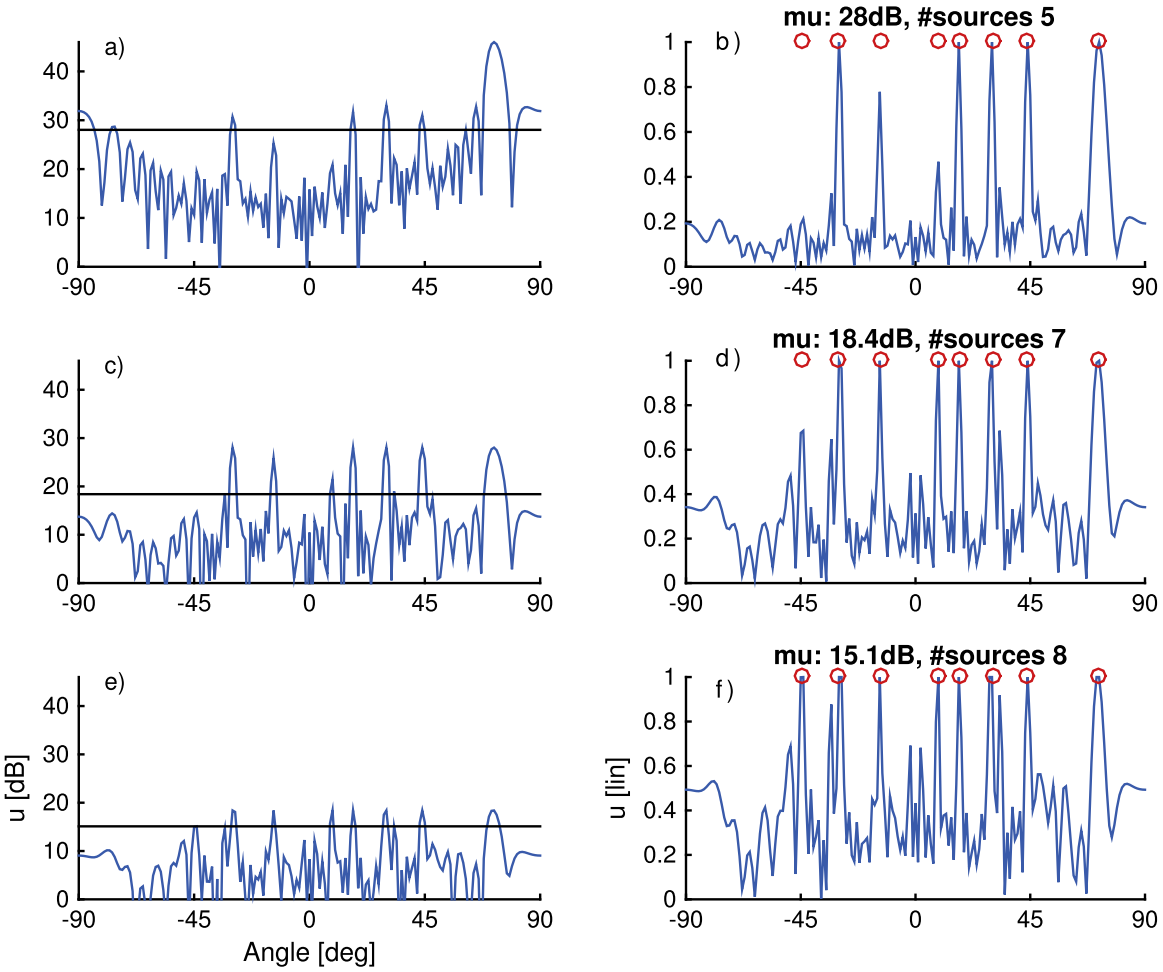

Fig. 11. Dual solution for iterative approach corresponding for localizing $K_0 = 8$ sources for step $i=1$ (a and b), $i=2$ (c and d), and $i=3$ (e and f). Left column: Dual (dB) for the previous step which is used for selecting μ (horizontal line). Right column: Dual (lin) normalized with μ (maximum is 1), the true source locations are marked with $\circ$, and the actual value of μ and number of sources found is also indicated.

associated with the source at $\varphi = 72°$. As soon as the dominant source is determined, the sidelobes in the residuals are reduced and only five sources are observed. After two more iterations, all eight sources are found at their correct locations.

For both algorithms, the main CPU time is used in solving the convex optimization problem. Thus the iterative algorithm is a factor 8/3 faster in this case than the straightforward approach which strictly follows the c-LASSO path. The approach described in Table 2 has approximately the same CPU time usage as the approach in Ref. [21], but it is conceptually simpler and provides deeper physical insight into the problem.

8. Conclusion

The c-LASSO problem is convex. The corresponding dual problem is interpretable as a weighted matched filter acting on the residuals of the c-LASSO. There is a linear one-to-one relation between the dual and primal vectors. Any results formulated for the primal problem are readily extendable to the dual problem. Thus, the sensitivity of the primal solution to small changes in the constraints can be easily assessed. Further, the difference between the solutions $\boldsymbol{x}_{\ell_0}$ and the $\boldsymbol{x}_{\ell_1}$ is characterized via the dual vector.

Based on mathematical and physical insight, an order-recursive and a faster iterative c-LASSO-based algorithm are proposed and evaluated. These algorithms use the dual variable of the c-LASSO for regularization parameter selection. This greatly facilitates computation of the c-LASSO-path as we can predict the changes in the active indexes as the regularization parameter is reduced. Further, a dual-based algorithm is formulated which solves only the dual problem. The examples demonstrate the algorithms, confirming that the dual and primal solutions are piecewise linear in the regularization parameter μ.

Acknowledgments

This work has been funded by the Christian Doppler Laboratory for Wireless Technologies for Sustainable Mobility, the Christian Doppler Laboratory for Dependable Connectivity for the Society in Motion, the Office of Naval Research, Grant Nos. N00014-11-1-0439 and N00014-13-1-0632 (MURI), as well as the Telecommunications Research Center Vienna (FTW) through the project Compressed channel state information feedback for time-variant MIMO channels. The financial support by the Austrian Federal Ministry of Economy, Family and Youth, the Austrian Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development is gratefully acknowledged.

Appendix A. Proof of (16)

Set $\boldsymbol{u} = (u_1, \ldots, u_M)^T \in \mathbb{C}^M$. From (5),

$$\mu \parallel \boldsymbol{z} \parallel_1 - \mathrm{Re}(\boldsymbol{u}^H \boldsymbol{z}) = \sum_{m=1}^{M} \left(\mu |z_m| - \mathrm{Re}(u_m^* z_m) \right) \tag{A1}$$

$$\mu \parallel \boldsymbol{z} \parallel_1 - \mathrm{Re}(\boldsymbol{u}^H \boldsymbol{z}) = \sum_{m=1}^{M} \underbrace{(\mu - |u_m| \cos \phi_{mm})}_{=\tilde{\mu}_m} |z_m|, \tag{A2}$$

where we set $u_m^* z_m = |u_m||z_m| e^{j\phi_{mm}}$. The phase difference ϕ_{mm} depends on both u_m and z_m. If all coefficients $\tilde{\mu}_m$ in (A2) are non-negative, $\tilde{\mu}_m \geq 0$, for all $z_m \in \mathbb{C}$, then

$$\min_{\boldsymbol{z}} \left(\mu \parallel \boldsymbol{z} \parallel_1 - \mathrm{Re}(\boldsymbol{u}^H \boldsymbol{z}) \right) = 0, \tag{A3}$$

otherwise there is no lower bound on the minimum. Therefore, all $|u_m|$ must be bounded, i.e., $|u_m| \leq \mu \ \forall \ m = 1, \ldots, M$ to ensure that all $\tilde{\mu}_m \geq 0$ for all possible phase differences $-1 \leq \cos \phi_{mm} \leq 1$. Finally, we note that $\parallel \boldsymbol{u} \parallel_\infty = \max_m |u_m|$.

Appendix B. Proofs of Corollaries 1, 2, and 3

Proof of Corollary 1

Let the objective function of the c-LASSO problem (P1′) be

$$\mathcal{L} = \| \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} \|_2^2 + \mu \parallel \boldsymbol{D}\boldsymbol{x} \parallel_1. \tag{B1}$$

In the following, we evaluate the subderivative $\partial \mathcal{L}$ [38] as the set of all complex subgradients as introduced in [39]. First, we observe

$$\partial \mathcal{L} = -2\boldsymbol{A}^H (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}) + \mu \partial \parallel \boldsymbol{D}\boldsymbol{x} \parallel_1. \tag{B2}$$

Next, it is assumed that $\boldsymbol{D}$ is a diagonal matrix with positive real-valued diagonal entries. Then the subderivate $\partial \parallel \boldsymbol{D}\boldsymbol{x} \parallel_1$ evaluates to

$$\partial \parallel \boldsymbol{D}\boldsymbol{x} \parallel_1 = \begin{cases} \dfrac{D_{mm} x_m}{|x_m|} & \text{for } x_m \neq 0 \\ \{ z \in \mathbb{C}, |z| \leq 1 \} & \text{for } x_m = 0. \end{cases} \tag{B3}$$

The minimality condition for $\mathcal{L}$ is equivalent to setting (B2) to zero. For all m with $x_m \neq 0$ and with (18), this gives

$$D_{mm} u_m = \mu \frac{D_{mm} x_m}{|x_m|}. \tag{B4}$$

It readily follows that $|u_m| = \mu$ for $x_m \neq 0$ and $D_{mm} \neq 0$.

Proof of Corollary 2

Starting from Eq. (B4), dividing by μ and invoking Corollary 1, we conclude for matrices $\boldsymbol{D}$ with positive diagonal entries and for $m \in \mathcal{M}$,

$$\mu e^{j \arg(x_m)} = \frac{2}{D_{mm}} \boldsymbol{e}_m^H \boldsymbol{A}^H (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}) = u_m, \tag{B5}$$

where $\boldsymbol{e}_m$ is the mth standard basis vector. This concludes the proof of Corollary 2.

Proof of Corollary 3

For the primal vector, this was shown in the real-valued case by Tibshirani [2] and for the complex-valued case, this is a direct consequence of Appendix B, Eq. (55) in [21] under the assumption of a piecewise constant phase angle of the primal solution, cf. [20].

For the dual vector, this was shown in the real-valued case by Tibshirani [13] and for the complex-valued case, this readily follows from Result 1: If the primal vector $\boldsymbol{x}_{\ell_1}$ depends linearly on μ in (18) then so does the dual vector $\boldsymbol{u}$.

Appendix C. ℓ_0 Solution

The gradient (cf. Appendix B) of the data objective function is

$$\nabla \| \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} \|_2^2 = -2\boldsymbol{A}^H \left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} \right) \tag{C1}$$

For the active source components, x_m with $m \in \mathcal{M}$, the ℓ_0-constraint of (P0) is without effect and the solution results from setting the gradient to zero, i.e., solving the normal equations.

$$\boldsymbol{A}_{\mathcal{M}}^H \boldsymbol{y} = \boldsymbol{A}_{\mathcal{M}}^H \boldsymbol{A}_{\mathcal{M}} \boldsymbol{x}_{\ell_0, \mathcal{M}} \quad \Rightarrow \quad \boldsymbol{x}_{\ell_0, \mathcal{M}} = \boldsymbol{A}_{\mathcal{M}}^{\dagger} \boldsymbol{y} \tag{C2}$$

We set

$$\boldsymbol{x}_{\ell_0, \mathcal{M}} = \boldsymbol{x}_{\ell_1, \mathcal{M}} + \boldsymbol{\Delta}_{\mathcal{M}}. \tag{C3}$$

This is inserted into (C1),

$$\nabla \left\| \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\ell_1, \mathcal{M}} \right\|_2^2 = -2\boldsymbol{A}^H \left(\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{x}_{\ell_0, \mathcal{M}} - \boldsymbol{\Delta}_{\mathcal{M}}) \right). \tag{C4}$$

Using (6) gives

$$\boldsymbol{D}_{\mathcal{M}}^H \boldsymbol{u}_{\mathcal{M}} = 2\boldsymbol{A}_{\mathcal{M}}^H \left(\boldsymbol{y} - \boldsymbol{A}_{\mathcal{M}} \boldsymbol{x}_{\ell_1, \mathcal{M}} \right) \tag{C5}$$

$$\boldsymbol{D}_{\mathcal{M}}^H \mu e^{j\theta_{\mathcal{M}}} = 2\boldsymbol{A}_{\mathcal{M}}^H \left(\boldsymbol{y} - \boldsymbol{A}_{\mathcal{M}} \left(\boldsymbol{x}_{\ell_0, \mathcal{M}} - \boldsymbol{\Delta}_{\mathcal{M}} \right) \right) \tag{C6}$$

$$\mu \boldsymbol{D}_{\mathcal{M}}^H e^{j\theta_{\mathcal{M}}} = 2\boldsymbol{A}_{\mathcal{M}}^H \boldsymbol{A}_{\mathcal{M}} \boldsymbol{\Delta}_{\mathcal{M}} \tag{C7}$$

This results in

$$\boldsymbol{\Delta}_{\mathcal{M}} = \frac{\mu}{2} \left(\boldsymbol{A}_{\mathcal{M}}^H \boldsymbol{A}_{\mathcal{M}} \right)^+ \boldsymbol{D}_{\mathcal{M}}^H e^{j\theta_{\mathcal{M}}} \tag{C8}$$

which depends on μ both explicitly and implicitly through $\mathcal{M}$. If the set of nonzero elements of (P0) is equal to the active set of (P1′), the solutions of (P0) and (P1′) differ by (C8).

References

[1] A. Maleki, L. Anitori, Z. Yang, R.G. Baraniuk, Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP), IEEE Trans. Inf. Theory 59 (7) (2013) 4290–4308.

[2] R. Tibshirani, Regression shrinkage and selection via the LASSO, J. R. Stat. Soc. Ser. B (Methodological) (1996) 267–288.

[3] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (1) (2001) 129–159.

[4] I.F. Gorodnitsky, B.D. Rao, Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm, IEEE Trans. Signal Process. 45 (3) (1997) 600–616.

[5] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory 52 (2) (2006) 489–509.

[6] D.L. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (4) (2006) 1289–1306.

[7] J.J. Fuchs, Recovery of exact sparse representations in the presence of bounded noise, IEEE Trans. Inf. Theory 51 (10) (2005) 3601–3608.

[8] D. Malioutov, M. Çetin, A.S. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Trans. Signal Process. 53 (8) (2005) 3010–3022.

[9] D.L. Donoho, M. Elad, V.N. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, IEEE Trans. Inf. Theory 52 (1) (2006) 6–18.

[10] J.A. Tropp, Just relax: convex programming methods for identifying sparse signals in noise, IEEE Trans. Inf. Theory 52 (3) (2006) 1030–1051.

[11] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer, 2010.

[12] H. Van Trees, Optimum Array Processing, Wiley-Interscience, New York, 2002, Ch. 1–10.

[13] R.J. Tibshirani, J. Taylor, The solution path of the generalized LASSO, Ann. Stat. 39 (3) (2011) 1335–1371.

[14] A. Panahi, M. Viberg, Fast candidate points selection in the LASSO path, IEEE Signal Process. Lett. 19 (2) (2012) 79–82.

[15] Z. Yang, L. Xie, On phase transition of compressed sensing in the complex domain, IEEE Signal Process. Lett. 19 (1) (2012) 47–50.

[16] J.F. de Andrade Jr, M.L. de Campos, J.A. Apolinário Jr, A complex version of the LASSO algorithm and its application to beamforming, in: Proceedings of the 7th International Telecommunications Symposium (ITS 2010), Manaus, Brazil, 2010.
[17] D.L. Donoho, Y. Tsaig, I. Drori, J.-L. Starck, Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit, IEEE Trans. Inf. Theory 58 (2) (2012) 1094–1121.
[18] M.R. Osborne, B. Presnell, B.A. Turlach, A new approach to variable selection in least squares problems, IMA J. Numer. Anal. 20 (3) (2000) 389–403.
[19] Z. Koldovsky, P. Tichavsky, A homotopy recursive-in-model-order algorithm for weighted LASSO, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2014, pp. 4151–4155.
[20] A. Panahi, M. Viberg, Fast LASSO based DOA tracking, in: 4th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), IEEE, 2011, pp. 397–400.
[21] C.F. Mecklenbräuker, P. Gerstoft, A. Panahi, M. Viberg, Sequential bayesian sparse signal reconstruction using array data, IEEE Trans. Signal Process. 61 (24) (2013) 6344–6354.
[22] P. Gerstoft, A. Xenaki, C. Mecklenbräuker, Multiple and single snapshot compressive beamforming, J. Acoust. Soc. Am. 138 (4) (2015) 2003–2014.
[23] D. Needell, J.A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, Appl. Comput. Harmonic Anal. 26 (3) (2009) 301–321.
[24] D.L. Donoho, M. Elad, Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ_1 minimization, Proc. Natl. Acad. Sci. USA 100 (5) (2003) 2197–2202.
[25] E.J. Candes, T. Tao, Decoding by linear programming, IEEE Trans. Inf. Theory 51 (12) (2005) 4203–4215.
[26] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.
[27] S.-J. Kim, K. Koh, S. Boyd, D. Gorinevsky, ℓ_1 trend filtering, SIAM Rev. 51 (2) (2009) 339–360.
[28] G.H. Golub, C.F. Van Loan, Matrix Computations, Vol. 3, JHU Press, 2012.
[29] E. Zöchmann, P. Gerstoft, C.F. Mecklenbräuker, Density evolution of sparse source signals, in: 3rd International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa), IEEE, 2015, pp. 124–128.
[30] E.J. Candes, M.B. Wakin, S.P. Boyd, Enhancing sparsity by reweighted ℓ_1 minimization, J. Fourier Anal. Appl. 14 (5–6) (2008) 877–905.
[31] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, K. Knight, Sparsity and smoothness via the fused LASSO, J. R. Stat. Soc. Ser. B (Stat. Methodol.) 67 (1) (2005) 91–108.
[32] A. Xenaki, P. Gerstoft, K. Mosegaard, Compressive beamforming, J. Acoust. Soc. Am. 136 (1) (2014) 260–271.
[33] G. Tang, B.N. Bhaskar, P. Shah, B. Recht, Compressed sensing off the grid, IEEE Trans. Inf. Theory 59 (11) (2013) 7465–7490.
[34] E.J. Candès, C. Fernandez-Granda, Super-resolution from noisy data, J. Fourier Anal. Appl. 19 (6) (2013) 1229–1254.
[35] E.J. Candès, C. Fernandez-Granda, Towards a mathematical theory of super-resolution, Commun. Pure Appl. Math. 67 (6) (2014) 906–956.
[36] A. Panahi, M. Viberg, Gridless compressive sensing, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2014, pp. 3385–3389.
[37] A. Xenaki, P. Gerstoft, Grid-free compressive beamforming, J. Acoust. Soc. Am. 137 (4) (2015) 1923–1935.
[38] D.P. Bertsekas, Nonlinear Programming, Athena Scientific, 1999.
[39] P. Bouboulis, K. Slavakis, S. Theodoridis, Adaptive learning in complex reproducing kernel Hilbert spaces employing Wirtinger's subgradients, IEEE Trans. Neural Netw. Learn. Syst. 23 (3) (2012) 425–438.