Direction Finding with L1-norm Subspaces

P. P. Markopoulos, a N. Tsagkarakis, a D. A. Pados, a and G. N. Karystinos b

a Department of Electrical Engineering, The State University of New York at Buffalo, Buffalo, NY 14260, USA;
b School of Electronic and Computer Engineering, Technical University of Crete, Chania, 73100, Greece.

ABSTRACT

Conventional subspace-based signal direction-of-arrival estimation methods rely on the familiar L2-norm-derived principal components (singular vectors) of the observed sensor-array data matrix. In this paper, for the first time in the literature, we find the L1-norm maximum projection components of the observed data and search in their subspace for signal presence. We demonstrate that L1-subspace direction-of-arrival estimation exhibits (i) similar performance to L2 (usual singular-value/eigen-vector decomposition) direction-of-arrival estimation under normal nominal-data system operation and (ii) significant resistance to sporadic/occasional directional jamming and/or faulty measurements.

Keywords: Dimensionality reduction, direction-of-arrival estimation, erroneous data, faulty measurements, jamming, L1 norm, L2 norm, principal-component analysis, outlier resistance, subspace signal processing.

1. INTRODUCTION

Direction-of-arrival (DoA) estimation is a fundamental problem in signal processing with important applications in radar, sonar, source localization, wireless communications, and much more.1–6 Existing DoA estimation techniques may be broadly categorized into spectral estimation methods,7,8 likelihood maximization methods,9–14 and subspace-based methods.15–18 Subspace-based methods enjoy great popularity, mostly due to their favorable trade-off between target-angle resolution ability and computational simplicity in implementation.

Conventional subspace-based DoA-estimation methods, such as the celebrated MUltiple SIgnal Classification (MUSIC) procedure,15 rely on the L2-norm principal subspaces of the recorded snapshots. The L2-norm principal subspaces are obtained either by means of singular-value-decomposition (SVD) of the sensor-array data matrix or eigen-vector decomposition (EVD) of the estimated received-signal autocorrelation matrix.19 In unobstructed system operation and additive white Gaussian noise, such methods are well known to offer unbiased, asymptotically consistent DoA estimates20–22 with high target-angle resolution (“super-resolution” methods).

In several real-world applications, however, recorded snapshots may be corrupted by faulty/erroneous measurements, impulsive additive noise,23–25 and/or intermittent (sporadic) directional interference. Such interference may appear either as an endogenous communication-system characteristic, as for example in frequency-hopped spread-spectrum (FHSS) systems,26 or due to exogenous factors such as intentional jamming. In such cases, L2-subspace-based methods are known to suffer from significant performance degradation.27 Indeed, L2-principal subspaces—essentially the squared-norm minimizers—naturally respond strongly to corrupted snapshots that appear in the sensor-array data matrix as points that lie far from the nominal-signal subspace.28 DoA estimators that rely on such subspaces are inevitably misled. In this paper, we focus on designing a subspace DoA estimator that (i) is capable of attaining performance similar to that of L2-subspace methods (e.g., MUSIC) in unobstructed system operation and (ii) the DoA estimator, at the same time, exhibits inherently significant resistance against data record contamination.

Further author information: (Send correspondence to D.A.P.)
P.P.M: E-mail: pmarkopo@buffalo.edu, Telephone: +1 716 645 1207
N.T.: E-mail: ntsagkar@buffalo.com, Telephone: +1 716 645 1207
D.A.P.: E-mail: pados@buffalo.edu, Telephone: +1 716 645 1150
G.N.K.: E-mail: karystinos@telecom.tuc.gr, Telephone: +30 28210 37343
System developers have long been aware that absolute-error minimizers place much less emphasis than square-
error minimizers on individual data points that diverge from the nominal data subspace. Along this line, there
has been in the past few years an extensive documented effort in the signal processing and pattern recognition
literature to define and calculate $L_1$-norm-optimal subspaces of recorded data (under various forms of optimality)$^{29-36}$
that offer resistance against erroneous data in training. Markopoulos et al.$^{37,38}$ calculated optimally for
the first time the maximum-projection $L_1$-principal subspaces of real-valued data, for which only suboptimal
approximations were known in the literature.$^{34,35}$ Experimental studies accompanying the algebraic developments
illustrated sturdy resistance of the $L_1$-optimal subspaces against data set corruption.$^{37}$

In this work, we consider system operation in the presence of intermittent directional interference (jamming)
and, for the first time in the literature, present a complete DoA-estimation methodology based on optimally
calculated $L_1$-principal subspaces of the recorded snapshots. Specifically, we define the real $L_1$-principal data
subspace of the array-collected data and employ for its calculation the fastest-known optimal algorithm.$^{37}$ Then,
we search in the calculated subspace for signal presence by an appropriately defined MUSIC-analogous power
spectrum. Simulation studies presented herein demonstrate that the proposed $L_1$-based DoA-estimation method
attains performance similar to conventional MUSIC ($L_2$) under unobstructed system operation, while it offers
greatly superior signal-subspace approximation and DoA-estimation performance in the presence of intermittent
directional interference. The add-on cost associated with the proposed $L_1$ DoA estimator versus standard $L_2$ is
increased computational complexity.

2. SIGNAL MODEL

Consider a uniform linear array (ULA) of $D$ elements. The length-$D$ response vector to a far-field signal that
impinges on the array with angle of arrival $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ with respect to broadside is denoted by

$$
\mathbf{s}(\theta) \triangleq \left[ 1, e^{-j2\pi f_c d \sin(\theta)}, \ldots, e^{-j(D-1)2\pi f_c d \sin(\theta)} \right]^T
$$

(1)

where $f_c$ is the carrier frequency, $c$ is the signal propagation speed, and $d$ is the fixed inter-element spacing
of the array. To avoid unnecessary mathematical complications, we consider the uniform inter-element spacing $d$
to be no greater than half the carrier wavelength (i.e., $d \leq \frac{c}{2f_c}$ per Nyquist spatial sampling theorem). Then, for
any two distinct angles of arrival $\theta, \theta' \in (-\frac{\pi}{2}, \frac{\pi}{2}]$, $\mathbf{s}(\theta)$ and $\mathbf{s}(\theta')$ are linearly independent.

The array collects $N$ narrowband-signal snapshots from $K$ sources of interest (targets) with distinct angles
of arrival $\theta_1, \theta_2, \ldots, \theta_K \in (-\frac{\pi}{2}, \frac{\pi}{2})$, $K < D \leq N$. The angles of arrival are to be estimated. In addition
to the signals of interest, however, we experience intermittently co-channel presence of $L$ independent directional
interfering sources (jammers) at angles $\theta'_1, \theta'_2, \ldots, \theta'_L \in (-\frac{\pi}{2}, \frac{\pi}{2}]$. Each jammer may corrupt a snapshot with
a fixed, small, unknown to the receiver probability $\rho_l$, $l = 1, \ldots, L$. Thus, the $n$th signal observation vector
(snapshot), $n = 1, 2, \ldots, N$, is of the form

$$
\mathbf{y}_n = \sum_{k=1}^{K} x_{n,k} \mathbf{s}(\theta_k) + \sum_{l=1}^{L} \gamma_{n,l} x'_{n,l} \mathbf{s}(\theta'_l) + \mathbf{n}_n \in \mathbb{C}^{D \times 1}
$$

(2)

where $x_{n,k}$, $x'_{n,l} \in \mathbb{C}$ denote the random statistically independent values of target $k$ and jammer $l$, respectively,
in snapshot $n$, and $\gamma_{n,l}$ is the activity indicator random variable for jammer $l$ modeled as $\{0, 1\}$-Bernoulli with
activation probability $\rho_l$. Finally, $\mathbf{n}_n \in \mathbb{C}^{D \times 1}$ accounts for additive zero-mean, white circularly-symmetric
complex Gaussian noise with per-element variance $\sigma_n^2$, i.e., $\mathbf{n}_n \sim \mathcal{CN}(0_D, \sigma_n^2 I_D)$. From this point on, we refer
to the case of $\rho_l = 0$ for every $l = 1, \ldots, L$ as normal system operation (the array collects target-only data with
probability one).

*The real and imaginary parts of each element of $\mathbf{n}_n$, $n = 1, 2, \ldots, N$, are independently drawn form $\mathcal{N}(0, \frac{\sigma^2}{2})$.  

Proc. of SPIE Vol. 9109 91090J-2
3. REAL-FIELD $L_2$-PRINCIPAL-SUBSPACE DIRECTION FINDING

In this section, we set the foundation to define the real-field $L_1$-optimal subspace of interest by showing how familiar MUSIIC-type DoA estimation can be carried out involving only real-field operations.

To begin, we rewrite the two sums in (2) in equivalent matrix/vector form as follows

$$y_n = S(\Theta)x_n + S(\Theta')\Gamma_n x'_n + n_n \in \mathbb{C}^{D \times 1}$$

(3)

where $x_n \triangleq [x_{n,1}, x_{n,2}, \ldots, x_{n,K}]^T$, $x'_n \triangleq [x'_{n,1}, x'_{n,2}, \ldots, x'_{n,L}]^T$, $\Gamma_n \triangleq \text{diag}(\gamma_n)$, $\gamma_n \triangleq [\gamma_1, \gamma_2, \ldots, \gamma_L]^T$, $\Theta \triangleq \{\theta_1, \theta_2, \ldots, \theta_K\}$, and $\Theta' \triangleq \{\theta'_1, \theta'_2, \ldots, \theta'_L\}$. For any size-$m$ distinct-DoA set $\Phi = \{\phi_1, \phi_2, \ldots, \phi_m\}$ with $m \leq D$,

$$S(\Phi) \triangleq [s(\phi_1), s(\phi_2), \ldots, s(\phi_m)] \in \mathbb{C}^{D \times m}$$

(4)

denotes the corresponding rank-$m$ array steering matrix.\(^1\)

Next, we define the real-valued representation $\overline{A} \in \mathbb{R}^{2m \times 2n}$ of any complex-valued matrix $A \in \mathbb{C}^{m \times n}$ by concatenating real and imaginary parts as follows

$$\overline{A} \triangleq \begin{bmatrix} \Re\{A\} & -\Im\{A\} \\ \Im\{A\} & \Re\{A\} \end{bmatrix} \in \mathbb{R}^{2m \times 2n}.$$  

(5)

The transition from $A \in \mathbb{C}^{m \times n}$ to $\overline{A} \in \mathbb{R}^{2m \times 2n}$ is based on what is commonly referred to as complex-number realification in representation theory.\(^4\) Realification allows for any complex system of equations to be converted into (and solved through) a real system.\(^3\) Before we proceed, we review briefly in the following two Lemmas below fundamental properties of matrix realification.

Lemma 1. For any $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{n \times q}$, $(A + B) = \overline{A} + \overline{B}$, $(AB) = \overline{A} \overline{B}$, and $(A^H) = \overline{A}^T$. \hspace{1cm} $\blacksquare$

Lemma 2. For any $A \in \mathbb{C}^{m \times n}$, rank$(\overline{A}) = 2 \times \text{rank}(A)$. In particular, each singular value of $A$ will appear twice among the singular values of $\overline{A}$. \hspace{1cm} $\blacksquare$

By (5) and Lemma 1, we define the $n$th real-valued snapshot $\overline{y}_n$ by

$$\overline{y}_n = \overline{S}(\Theta)x_n + \overline{S}(\Theta')\overline{\Gamma}_n x'_n + \overline{n}_n \in \mathbb{R}^{2D \times 2}.$$  

(6)

In the following, DoA estimation will be carried out operating directly on $\overline{y}_n \in \mathbb{R}^{2D \times 1}$, $n = 1, \ldots, N$.

By (6) and Lemma 2, $\overline{S}(\Theta)$ defines a $2K$-dimensional subspace $S$ wherein the $K$ signal components of interest with angles of arrival $\theta_1, \theta_2, \ldots, \theta_K$ lie. The following proposition shows how this property of $\overline{S}(\Theta)$ can be utilized to estimate the target DoAs.

Proposition 1. For any $\phi \in (-\pi, \pi]$, span$(\overline{S}(\phi)) \subseteq S = \text{span} (\overline{S}(\Theta))$ if and only if $\phi \in \Theta$. Set equality holds only if $K = 1$. \hspace{1cm} $\blacksquare$

By Proposition 1, the two columns of $\overline{S}(\phi)$ have zero projection onto $S^\perp \triangleq \text{null}(\overline{S}(\Theta)^T)$ if and only if $\phi \in \Theta$. Therefore, knowledge of $S$ (thus, $S^\perp$) suffices to accurately decide whether some $\phi \in (-\pi, \pi]$ is a target DoA or not. In the sequel, we elaborate on obtaining (estimating) $S$ from the collected snapshots.

First, we define and calculate the first and second-order signal and activity-indicator statistics. For simplicity in presentation and without loss of generality, we assume that all activation probabilities are equal to each other, $\rho_1 = \rho_2 = \cdots = \rho_L = \rho$. Then, per (3), we set $\mu_x \triangleq \mathbb{E}[x]$, $\mu_{x'} \triangleq \mathbb{E}[x']$, $\mu_{\gamma} \triangleq \mathbb{E}[\gamma_n] = \rho \mathbf{1}_L$, $R_x \triangleq \mathbb{E}[xx^H]$, $R_{xx'} \triangleq \mathbb{E}[x'x'^H]$, and $R_{\gamma} \triangleq \mathbb{E}[\gamma_n \gamma_n^T] = \rho^2 \mathbf{1}_{L \times L} + (1-\rho)\mathbf{I}_L$. For any random matrix $A \in \mathbb{C}^{m \times n}$, $\mathbb{E}[\overline{A}] = \mathbb{E}[A]$ and $\mathbb{E}[\overline{AA}^H] = \mathbb{E}[(\overline{A})^T \overline{A}^T]$. Then, by Lemma 1 and (6) the autocorrelation matrix of the real-valued version of the observations is

$$R_y \triangleq \mathbb{E}\{\overline{y}_n \overline{y}_n^T\} = [\overline{S}(\Theta), \overline{S}(\Theta')]^T \begin{bmatrix} R_x & \rho \overline{P}_x \overline{P}_x^T \\ \rho \overline{P}_x \overline{P}_x^T & R_{xx'} \odot (\mathbf{1}_{2 \times 2} \otimes R_{\gamma}) \end{bmatrix} \begin{bmatrix} \overline{S}(\Theta)^T \\ \overline{S}(\Theta')^T \end{bmatrix} + \sigma^2 \mathbf{I}_{2D}.$$  

(7)

\(^1\) $\overline{S}(\Phi) \in \mathbb{C}^{D \times m}$ is built by columns of a transposed Vandermonde matrix\(^3\) and has rank $m$ if $|\Phi| = m \leq D$. 

---

Proc. of SPIE Vol. 9109 91090J-3
Since \( \text{rank}(R_x) = K \), Lemma 2 implies that \( \text{rank}(R_x) = 2K \) and the \( 2K \)-dimensional column-span of \( \mathbf{S}(\Theta)R_x\mathbf{S}(\Theta)^T \) that appears in (7) coincides with that of \( \mathbf{S}(\Theta) \). Therefore, given the eigenvalue-decomposition (EVD)

\[
\mathbf{S}(\Theta)R_x\mathbf{S}(\Theta)^T = \mathbf{U} \mathbf{A} \mathbf{U}^T
\]

(8)

where \( \mathbf{A} = \text{diag}(\{\lambda_1, \lambda_2, \ldots, \lambda_{2D}\}) \) is the diagonal matrix containing the sorted eigenvalues \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{2K} > \lambda_{2K+1} = \cdots = \lambda_{2D} = 0 \) and \( \mathbf{U} \) is the corresponding \( 2D \times 2D \) orthonormal eigenmatrix, the \( 2K \) leftmost columns of \( \mathbf{U} \) collected in \( \mathbf{U}(\Theta) \triangleq [\mathbf{U}_{1:2K}]_N \) constitute a complete orthonormal basis for \( \mathbf{S} \).

For regular system operation \( (\rho = 0) \), \( \mathbf{R}_\mathbf{\hat{Y}} \) admits EVD

\[
\mathbf{R}_\mathbf{\hat{Y}} \overset{\text{e.v.d.}}{=} \mathbf{U}(\Lambda + \sigma^2\mathbf{I}_{2D})\mathbf{U}^T.
\]

(9)

Then, the solution \( \mathbf{Q}_{\text{opt}} \in \mathbb{R}^{2D \times 2K} \) to the \( L_2 \)-principal subspace problem\(^\text{19} \)

\[
\max_{\mathbf{Q} \in \mathbb{R}^{2D \times 2K}, \mathbf{Q}^T \mathbf{Q} = \mathbf{I}_{2K}} \text{Tr} \left( \mathbf{Q}^T \mathbf{R}_\mathbf{\hat{Y}} \mathbf{Q} \right)
\]

(10)

coincides with \( \mathbf{U}(\Theta) \) and offers a complete characterization of \( \mathbf{S} \). By Proposition 1, any \( \phi \in (\frac{\pi}{2}, \frac{\pi}{2}] \) can be optimally examined for target presence by

\[
\phi \in \Theta \iff (\mathbf{I}_{2D} - \mathbf{Q}_{\text{opt}}\mathbf{Q}_{\text{opt}}^T)\mathbf{s}(\phi) = 0_{2D \times 2}.
\]

(11)

In field operations, however, perfect knowledge of \( \mathbf{R}_\mathbf{\hat{Y}} \) cannot be assumed. Instead, \( \mathbf{R}_\mathbf{\hat{Y}} \) may be sample-average estimated over the \( N \) real-valued snapshot matrices as

\[
\mathbf{\hat{R}}_\mathbf{\hat{Y}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{\overline{y}}_n \mathbf{\overline{y}}_n^T = \frac{1}{N} \mathbf{\overline{Y}} \mathbf{\overline{Y}}^T
\]

(12)

where \( \mathbf{\overline{Y}} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N] \in \mathbb{C}^{D \times N} \) is the size-\( N \) complex snapshot record.\(^{\text{4}} \) The maximization argument in (10) is accordingly calculated by \( \frac{1}{N} \text{Tr} \left( \mathbf{Q}^T \mathbf{\overline{Y}} \mathbf{\overline{Y}}^T \mathbf{Q} \right) = \frac{1}{N} \| \mathbf{Q}^T \mathbf{\overline{Y}} \|^2_2 \) and \( \mathbf{S} \) estimated by \( \hat{\mathbf{S}}_{L_2} \triangleq \text{span}(\mathbf{Q}_{L_2}) \) where \( \mathbf{Q}_{L_2} \in \mathbb{R}^{2D \times 2K} \) is the solution to

\[
\max_{\mathbf{Q} \in \mathbb{R}^{2D \times 2K}, \mathbf{Q}^T \mathbf{Q} = \mathbf{I}_{2K}} \| \mathbf{Q}^T \mathbf{\overline{Y}} \|^2_2.
\]

(13)

The maximizer \( \mathbf{Q}_{L_2} \) of (13) is obtained by singular-value-decomposition of the real-valued data matrix \( \mathbf{\overline{Y}} \).\(^\text{19} \) Under normal system operation, \( \mathbf{\hat{R}}_\mathbf{\hat{Y}} \) converges to \( \mathbf{R}_\mathbf{\hat{Y}} \) (a.s.) and \( \hat{\mathbf{S}}_{L_2} \) converges to \( \mathbf{S} \) (a.s.) as the number of snapshots \( N \) increases to infinity. In view of (11), directions of arrival are estimated as the arguments that correspond to the \( K \) highest distinct local maxima (peaks) of the power spectrum\(^3 \)

\[
P_{L_2}(\phi) \triangleq 2\| (\mathbf{I}_{2D} - \mathbf{Q}_{L_2}\mathbf{Q}_{L_2}^T)\mathbf{s}(\phi) \|^2_2.
\]

(14)

Proposition 2 below establishes the equivalence of complex-MUSIC with the presented real-field \( L_2 \)-subspace DoA estimation method.

**Proposition 2.** For every solution \( \mathbf{Q}_{L_2} \in \mathbb{C}^{D \times K} \) of (13), there exists a solution \( \mathbf{Q}_{CL_{L_2}} \) of

\[
\max_{\mathbf{Q} \in \mathbb{C}^{D \times K}, \mathbf{Q}^T \mathbf{Q} = \mathbf{I}_K} \| \mathbf{Q}^H \mathbf{\overline{Y}} \|^2_2
\]

(15)

\(^1\)If the real-valued snapshots are elliptically contoured, \( \mathbf{\hat{R}}_\mathbf{\hat{Y}} \) is a maximum-likelihood, unbiased, and asymptotically consistent estimator for \( \mathbf{R}_\mathbf{\hat{Y}} \).

\(^2\)For finite data-record size \( N \), \( \text{Pr}\{\text{span} (\mathbf{s}(\phi)) \not\subset \hat{\mathbf{S}}_{L_2}\} = 1 \) for all \( \phi \in (-\frac{\pi}{2}, \frac{\pi}{2}] \) and (14) is defined w.p.1.
with \(\|Q^T_{L_2} \mathbf{Y}\|_2 = \sqrt{2}\|Q^H_{CL_2} \mathbf{Y}\|_2\) and \(Q_{L_2} Q^T_{L_2} = (Q_{CL_2} Q^H_{CL_2}) = Q_{CL_2} Q^T_{CL_2}\). ■

By Proposition 2, Lemma 1, and the fact that \(\|\mathbf{X}\|_2 = \sqrt{2}\|A\|_2\), for every complex matrix \(A\), \(P_{L_2}(\phi)\) in (14) calculated solely by means of the real-field operations (13) and (14) coincides with the MUSIC spectrum:

\[
\| (I_D - Q_{CL_2} Q^H_{CL_2}) \mathbf{s}(\phi) \|_2^{-2}
\]

for every \(\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]\). Thus, standard (complex) \(L_2\)-subspace DoA estimation can be carried out equivalently via real \(L_2\)-principal subspaces of the realified snapshots of twice the original dimension.

4. NOVEL \(L_1\)-PRINCIPAL-SUBSPACE DIRECTION FINDING

4.1 Directional Interference and \(L_1\)-norm Direction Finding

We steer now our focus to the scenario of interest where each snapshot may be corrupted by one up to \(L\) directional jammers with non-zero probability. Following Sections 2 and 3 notation, the probability that a interfering source is \(p = 1 - (1 - \rho)^L > 0\), \(n = 1, 2, \ldots, N\). In this case, signal and disturbance (interference plus noise) subspaces are not SVD/EVD-separable any more. The \(2K\) \(L_2\)-principal components of \(R_F\) define now a \(2K\)-dimensional subspace, which is a subset of span \((\mathbf{S}(\theta), \mathbf{S}(\theta'))\) that does not necessarily coincide with \(S\).

At the same time, recent studies in the emerging field of \(L_1\)-principal component analysis have exhibited that \(L_1\)-principal subspaces are far more resistant to data contamination than \(L_2\)-principal subspaces.\(^{29-38}\) Therefore, in the DoA estimation problem of interest herein, we are highly motivated to substitute the \(L_2\)-principal subspace derived through (15) by a contamination resistant \(L_1\)-norm optimal subspace. Specifically, we choose to change the norm from \(L_2\) to \(L_1\) in (15) and focus on the \(L_1\)-optimal subspace \(\hat{S}_{L_1}\) that is described by the orthonormal matrix \(Q_{L_1}\) that solves

\[
\text{maximize}_{Q \in \mathbb{R}^{2D \times 2K}, \; Q^T Q = I_{2K}} \|Q^T \mathbf{Y}\|_1.
\]

Thereafter, directions of arrival \(\theta_1, \theta_2, \ldots, \theta_K\) will be estimated as the \(K\) highest distinct local maxima (peaks) of the power spectrum

\[
P_{L_1}(\phi) \overset{\Delta}{=} 2 \| (I_{2D} - Q_{L_1} Q^T_{L_1}) \mathbf{s}(\phi) \|_2^{-2}.
\]

Although \(L_1\)-principal subspace finding in the form of (17) is not a new problem in the literature,\(^{34,35}\) the optimal solution was unknown until recently.\(^{37}\) In the sequel, we present the optimal solution, discuss the general complexity, and provide the fastest known algorithm.\(^{37,38}\)

4.2 The Optimal \(L_1\)-Principal Subspace

Consider the \(L_1\)-principal subspace finding problem in (17). By the definition of the entry-wise \(L_1\) norm, for any matrix \(A \in \mathbb{R}^{m \times n}\), \(\|A\|_1 = \sum_{i=1}^m \sum_{j=1}^n |A_{i,j}| = \text{Tr}(A^T \text{sgn}(A))\) where \(\text{sgn}(A) = \text{arg max}_{B \in \{\pm 1\}^{m \times n}} \text{Tr}(A^T B)\). Therefore, the \(L_1\)-principal subspace finding problem in (17) can be rewritten as

\[
\text{maximize}_{Q \in \mathbb{R}^{2D \times 2K}, \; Q^T Q = I_{2K}} \left\{ \max_{B \in \{\pm 1\}^{2N \times 2K}} \text{Tr}(Q^T YB) \right\}.
\]

By the Cauchy-Schwarz inequality for matrices,\(^{19}\) for any \(Q \in \mathbb{R}^{m \times n}\) such that \(Q^T Q = I_n\) and \(A \in \mathbb{R}^{m \times n}\) that admits compact SVD \(A \overset{\text{comp svd}}{=} \mathbf{U}_{m \times K} \Sigma V^T\),

\[
\text{Tr}(Q^T A) = \text{Tr}(Q^T U \Sigma V^T) = \text{Tr}(U \Sigma^2 V^T Q^T) \\
\leq \|U \Sigma^2\|_2 \|\Sigma V^T Q^T\|_2 = \|\Sigma^2\|_2^2 \\
= \text{Tr}(\Sigma) = \|A\|_*
\]

\(^*\)rank \((\mathbf{S}(\theta), \mathbf{S}(\theta'))\) = min\{2(K + L), 2D\} ≥ 2K.

Proc. of SPIE Vol. 9109 91090J-5
where \( \|A\|_* \) denotes the nuclear-norm (also known as trace-norm) of \( A \) and equals the sum of its singular values. Equality in (20) holds for \( (U \Sigma^2) \Sigma^2 = \Sigma^2 V^T Q^T \), which is satisfied by \( Q = UV^T \). Therefore, for any fixed \( B \in \{\pm1\}^{2N \times 2K} \) in the inner maximization in (19),

\[
\max_{Q \in \mathbb{R}^{2D \times 2K}, Q^T Q = I_{2K}} \text{Tr} (Q^T YB) = \|YB\|_* .
\]

(21)

In view of (20) and (21), Proposition 3 below provides the optimal solution to (17).

**Proposition 3.** If \( B_{\text{nuc}} \) is a solution to

\[
\max_{B \in \{\pm1\}^N} \|YB\|_*
\]

and \( YB_{\text{nuc}} = U \Sigma V^T \), then

\[
Q_{L_1} = U_{:,1:2K} V^T
\]

is a solution to (17). Moreover, \( \|Q_{L_1}^T Y\|_1 = \|YB_{\text{nuc}}\|_* \). □

By Proposition 3, if we have the solution to (22) the \( L_1 \)-optimal subspace is obtained by a simple SVD. Thus, we have cast the \( L_1 \)-subspace finding problem of interest as a combinatorial optimization problem. A natural approach to solve (22) would be exhaustive search over all points \( B \in \{\pm1\}^{2N \times 2K} \) with complexity \( O(2^{4NK}) \). Certainly, exhaustive search becomes soon practically infeasible as the data-record size \( N \) increases. Instead, Markopoulos et al.\(^{37} \) presented recently an optimal algorithm that finds \( B_{\text{nuc}} \) with complexity \( O(4^{DK - 2K}) \), that is, in polynomial time with respect to \( N \) for any fixed dimension (antenna-array length) \( D < N \). For completeness purposes, we give the optimal algorithm in Fig. 1.

### 5. EXPERIMENTAL STUDIES

In this section, we carry out simulation studies to illustrate the theoretical developments presented in this paper.

In Fig. 2, we present a realization of the DoA-estimation spectra \( P_{L_1} \) and \( P_{L_2} \), as defined in (14) and (18), respectively, that offers a first hint on the effectiveness of the proposed methodology as compared with conventional \( L_2 \)-subspace-based DoA estimation (MUSIC). The receiver antenna-array is equipped with \( D = 6 \) elements and collects \( N = 8 \) snapshots. All snapshots contain the target signal of interest (\( K = 1 \)). Two of the eight snapshots are corrupted by two interfering sources (\( L = 2 \)). The signal-to-noise ratio (SNR) is set to 10dB for the target source and to 20dB for each of the two interferers. MUSIC (\( L_2 \)) clearly fails to identify the DoA of interest and is misled by the two jammer-corrupted measurements. Interestingly, the \( L_1 \)-subspace manages to separate the target signal location successfully.

In Fig. 3, we generalize our study to include probabilistic presence of an interferer/jammer. Specifically, we set \( D = 3 \) antenna elements, \( N = 4 \) snapshots, \( K = 1 \) target signal at \( \theta = -40^\circ \) with SNR 10dB, and \( L = 1 \) jammer at \( \theta' = 32^\circ \) with activation probability \( \rho = 0, .1, .25, .5, .75 \), or 1. We plot the root-mean-square-error (RMSE)\(^{\dagger} \) of both \( L_1 \) and \( L_2 \) methods averaged over 1 000 independent realizations as a function of the SNR of the jammer. As documented, in normal system operation (\( \rho = 0 \)) the RMSE of both estimators is extremely close to each other and low, with slight superiority of the \( L_2 \)-designed method. Quite interestingly, for any non-zero jammer activation probability and over the whole range of jammer SNR, the \( L_1 \) RMSE is lower than the \( L_2 \).\(^{\ddagger} \) It is tempting to consider that, if the computational cost can be afforded, \( L_1 \)-subspace based DoA estimation may be the overall preferred method due to the exhibited resistance to faulty data and relatively similar performance to \( L_2 \) under clean data.

\(^{\dagger} \)Denoting by \( \hat{\theta} \) the DoA estimate obtained by either method, RMSE is defined as the square root of the average value of \( |\hat{\theta} - \theta| \).

\(^{\ddagger} \)At high jammer SNR values and \( \rho = 1 \), both methods will fail reaching the horizontal line in the plot that indicates the number of degrees that separate the two sources (i.e., both methods detect the jammer instead of the target).
Algorithm: Polynomial-time calculation of the $2K$-dimensional $L_1$-principal subspace

**Input:**
1) Real data matrix $X_{2D \times 2N}$ with rank($X$) = $2D$ and 2) subspace dimensionality $2K$

1: $B_{N \times M} \leftarrow \text{compute\_candidates}(X)$, $M \leftarrow \{1, 2, \ldots, M\}$
2: $z_{\text{opt}} \leftarrow \arg \max_{z \in A^{2K}} \|XB_z\|$
3: $B_{\text{nuc}} \leftarrow B_{z_{\text{opt}}}$
4: $(U_{2D \times 2K}, \Sigma_{2K \times 2K}, V_{2K \times 2K}) \leftarrow \text{svd}(XB_{\text{nuc}})$

**Output:** $Q_{1L} \leftarrow UV^T$

**Function** compute\_candidates

**Input:** Full-rank $Z_{n \times m}$ with $n > m$
1: if $m > 2$ then $i \leftarrow 0$
2: for $I \subset \{1, 2, \ldots, n\}$ s.t. $|I| = m − 1$, $i \leftarrow i + 1$
3: $Z_{(m−1) \times m} \leftarrow Z_{I}$
4: $c_{m \times 1} \leftarrow \text{null}(Z)$, $c \leftarrow \text{sgn}(c_{m})c$
5: $B_{i,i} \leftarrow \text{sgn}(Zc)$
6: for $j = 1 : m − 1$
7: $c_{(m−1) \times 1} \leftarrow \text{null}(Z_{/j,m−1})$, $c \leftarrow \text{sgn}(c_{m−1})c$
8: $B_{Z(j,i)} \leftarrow \text{sgn}(Z_{/1:m−1}c)$
9: $B \leftarrow [B, \text{compute\_candidates}(Z_{/1:m−2})]$
10: else $m = 2$, for $i = 1, 2, \ldots, n$
11: $c_{2 \times 1} \leftarrow \text{null}(Z_{i})$, $c \leftarrow \text{sgn}(c_{2})c$
12: $B_{i,i} \leftarrow \text{sgn}(Zc)$, $B_{i,i} \leftarrow \text{sgn}(Z_{i,1})$
13: else $B \leftarrow \text{sgn}(Z)$

**Output:** $B$

Figure 1. Optimal $O(N^{4DK+2K+1})$ computation of the $2K$-dimensional $L_1$-principal subspace of a rank-$2D$ data matrix $Y_{2D \times 2N}$.

Another metric of interest specifically designed to capture the resistance to jammed measurements by the employed subspace estimators is the subspace representation ratio defined as the average value of

$$\text{srr}(Q) \triangleq \frac{\|Q^{T}\bar{s}_0\|^2}{\|Q^{T}\bar{s}_0\|^2 + \|Q^{T}s_0\|^2}$$

(24)

where $Q$ represents the principal orthonormal basis for the estimated signal subspace given by the solution to (13) for MUSIC ($L_2$) and the solution to (17) for the proposed method ($L_1$). In Fig. 4, we plot for the set-up of Fig. 3 srr($Q$) as a function of the jammer SNR for activation probabilities $\rho = 0, .1, .25, .5, .75$, or 1. Similar to Fig. 3, in normal system operation ($\rho = 0$) the performance of both methods is high-quality and very close to each other ($L_2$-MUSIC being negligibly superior). In the presence of corrupted measurements, $L_1$-principal subspace estimation outperforms $L_2$ universally in the study. In constant jammer presence ($\rho = 1$), both methods capture the jammer signal subspace as the jammer SNR increases relative to the fixed target signal SNR.

6. SUMMARY

We revisited the problem of subspace-based DoA estimation in the presence of (intermittent) directional interference. We presented for the first time in the literature a complete DoA-estimation methodology based on optimally calculated $L_1$-principal subspaces of the recorded snapshots. First, we demonstrated how conventional $L_2$-principal subspace DoA estimation can be carried out using only real-field operations. Then, in view of recent developments in the subspace signal processing literature, we proposed $L_1$-principal-components calculation of
Figure 2. DoA-estimation spectra $P_{L_1}$ and $P_{L_2}$ (target signal location $\Delta$; interfering sources locations $\times$).

Figure 3. Root-mean-squared-error (RMSE) of $L_2$ (MUSIC) and proposed $L_1$-subspace DoA-estimation method (target-from-jammer degrees of separation 72°).
the target-signal subspace of interest. We showed how the $L_1$-subspace finding problem can be cast as a combinatorial optimization problem and presented the fastest known optimal algorithm for its calculation. Simulation studies demonstrated that the proposed $L_1$ DoA-estimation procedure attains performance similar to familiar $L_2$ (MUSIC) under normal system operation, while it offers significant DoA-estimation enhancement when the data record is corrupted by faulty measurements due to jamming/interference (or sensor malfunction). One may conclude that, when the additional computational cost can be afforded, $L_1$-subspace DoA estimation is to be preferred over $L_2$.

REFERENCES


