Open Access

Realified L1-PCA for direction-of-arrival estimation: theory and algorithms

Panos P. Markopoulos[1], Nicholas Tsagkarakis[2], Dimitris A. Pados[3*] and George N. Karystinos[4]

Abstract

Subspace-based direction-of-arrival (DoA) estimation commonly relies on the Principal-Component Analysis (PCA) of the sensor-array recorded snapshots. Therefore, it naturally inherits the sensitivity of PCA against outliers that may exist among the collected snapshots (e.g., due to unexpected directional jamming). In this work, we present DoA-estimation based on outlier-resistant L1-norm principal component analysis (L1-PCA) of the realified snapshots and a complete algorithmic/theoretical framework for L1-PCA of complex data through realification. Our numerical studies illustrate that the proposed DoA estimation method exhibits (i) similar performance to the conventional L2-PCA-based method, when the processed snapshots are nominal/clean, and (ii) significantly superior performance when the snapshots are faulty/corrupted.

Keywords: Data contamination, Direction-of-arrival estimation, Faulty measurements, L1 norm, L2 norm, Multiple signal classification, Principal-component analysis, Outlier resistance, Singular-value decomposition, Subspace data processing

1 Introduction

Direction-of-arrival (DoA) estimation is a fundamental problem in signal processing theory with important applications in localization, navigation, and wireless communications [1–6]. Existing DoA-estimation methods can be broadly categorized as (i) likelihood maximization methods [7–13], (ii) spectral estimation methods, as in the early works of [14, 15], and (iii) subspace-based methods [16–19]. Subspace-based methods have enjoyed great popularity in applications, mostly due to their favorable trade-off between angle estimation quality and computational simplicity in implementation.

In their most common form, subspace-based DoA estimation methods rely on the L2-norm principal components (L2-PCs) of the recorded snapshots, which can be simply obtained by means of singular-value decomposition (SVD) of the sensor-array data matrix, or by eigenvalue decomposition (EVD) of the received-signal autocorrelation matrix [20]. Importantly, under nominal system operation (i.e., no faulty measurements or unexpected jamming/interfering sources), in additive white

Gaussian noise (AWGN) environment, such methods are known to offer unbiased, asymptotically consistent DoA estimates [21–23] and exhibit high target-angle resolution ("super-resolution" methods).

However, in many real-world applications, the collected snapshot record may be unexpectedly corrupted by faulty measurements, impulsive additive noise [24–26], and/or intermittent directional interference. Such interference may appear either as an endogenous characteristic of the underlying communication system, as for example in frequency-hopped spread-spectrum systems [27], or as an exogenous factor (e.g., jamming). In cases of such snapshot corruption, L2-PC-based methods are well known to suffer from significant performance degradation [28–30]. The reason is that, as squared error-fitting minimizers, L2-PCs respond strongly to corrupted snapshots that appear in the processed data matrix as points that lie far from the nominal signal subspace [29]. Accordingly, DoA estimators that rely upon the L2-PCs are inevitably misled.

At the same time, research in signal processing and data analysis has shown that absolute error-fitting minimizers place much less emphasis on individual data points that diverge from the nominal signal subspace than square-fitting-error minimizers. Based on this observation, in

*Correspondence: dpados@fau.edu
[3]I-SENSE and Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, 33431 USA
Full list of author information is available at the end of the article

the past few years, there have been extended documented research efforts toward defining and calculating L1-norm principal components (L1-PCs) of data under various forms of L1-norm optimality, including absolute-error minimization and projection maximization [31–46]. Recently, Markopoulos et al. [47, 48] calculated optimally the maximum-projection L1-PCs of real-valued data, for which up to that point only suboptimal approximations were known [36–38]. Experimental studies in [47–53] demonstrated the sturdy resistance of optimal L1-norm principal-component analysis (L1-PCA) against outliers, in various signal processing applications. Recently, [43, 45] introduced a heuristic algorithm for L1-PCA that was shown to attain state-of-the-art performance/cost trade-off. Another popular approach for outlier-resistant PCA is "Robust PCA" (RPCA), as introduced in [29] and further developed in [54, 55].

In this work, we consider system operation in the presence of unexpected, intermittent directional interference and propose a new method for DoA-estimation that relies on the L1-PCA of the recorded complex snapshots. Importantly, this work introduces a complete paradigm on how L1-PCA, defined and solved over the real field [47, 48], can be used for processing complex data, through a simple "realification" step. An alternative approach for L1-PCA of complex-valued data was presented in [46], where the authors reformulated complex L1-PCA into unimodular nuclear-norm maximization (UNM) and estimated its solution through a sequence of converging iterations. It is noteworthy that for the UNM introduced in [46], no general exact solver exists to date.

Our numerical studies show that the proposed L1-PCA-based DoA-estimation method attains performance similar to the conventional L2-PCA-based one (i.e., MUSIC [16]) in the absence of jamming sources, while it offers significantly superior performance in the case of unexpected, sporadic contamination of the snapshot record.

Preliminary results were presented in [56]. The present paper is significantly expanded to include (i) an Appendix section with all necessary technical proofs, (ii) important new theoretical findings (Proposition 3 on page 7), (iii) new algorithmic solutions (Section 3.5), and (iv) extensive numerical studies (Section 4).

The rest of the paper is organized as follows. In Section 2, we present the system model and offer a preliminary discussion on subspace-based DoA estimation. In Section 3, we describe in detail the proposed L1-PCA-based DoA-estimation method and present three algorithms for L1-PCA of the snapshot record. Section 4 presents our numerical studies on the performance of the proposed DoA estimation method. Finally, Section 5 holds some concluding remarks.

1.1 Notation

We denote by $\mathbb{R}$ and $\mathbb{C}$ the set of real and complex numbers, respectively, and by j the imaginary unit (i.e., $j^2 = -1$). $\Re\{(\cdot)\}, \Im\{(\cdot)\}, (\cdot)^*, (\cdot)^\top$, and $(\cdot)^H$ denote the real part, imaginary part, complex conjugate, transpose, and conjugate transpose (Hermitian) of the argument, respectively. Bold lowercase letters represent vectors and bold uppercase letters represent matrices. $\mathbf{diag}(\cdot)$ is the diagonal matrix formed by the entries of the vector argument. For any $\mathbf{A} \in \mathbb{C}^{m \times n}, [\mathbf{A}]_{i,q}$ denotes its (i, q)th entry, $[\mathbf{A}]_{:,q}$ its qth column, and $[\mathbf{A}]_{i,:}$ its ith row; $\|\mathbf{A}\|_p \triangleq \left(\sum_{i=1}^{m} \sum_{q=1}^{n} |[\mathbf{A}]_{i,q}|^p\right)^{\frac{1}{p}}$ is the pth entry-wise norm of $\mathbf{A}, \|\mathbf{A}\|_*$ is the nuclear norm of $\mathbf{A}$ (sum of singular values), span($\mathbf{A}$) represents the vector subspace spanned by the columns of $\mathbf{A}$, rank($\mathbf{A}$) is the dimension of span($\mathbf{A}$), and null($\mathbf{A}^\top$) is the kernel of span($\mathbf{A}$) (i.e., the nullspace of $\mathbf{A}^\top$). For any square matrix $\mathbf{A} \in \mathbb{C}^{m \times m}$, det($\mathbf{A}$) denotes its determinant, equal to the product of its eigenvalues. $\otimes$ and $\odot$ are the Kronecker and entry-wise (Hadamard) product operators [57], respectively. $\mathbf{0}_{m \times n}, \mathbf{1}_{m \times n}$, and $\mathbf{I}_m$ are the $m \times n$ all-zero, $m \times n$ all-one, and size-m identity matrices, respectively. Also, $\mathbf{E}_m \triangleq \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \otimes \mathbf{I}_m$, for $m \in \mathbb{N}_{\geq 1}$, and $\mathbf{e}_{i,m}$ is the ith column of $\mathbf{I}_m$. Finally, E$\{\cdot\}$ is the statistical-expectation operator.

2 System model and preliminaries

We consider a uniform linear antenna 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 (w.r.t.) the broadside is defined as

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

where f_c is the carrier frequency, c is the signal propagation speed, and d is the fixed inter-element spacing of the array. We consider that the uniform inter-element spacing d is no greater than half the carrier wavelength, adhering to the Nyquist spatial sampling theorem; i.e., $d \leq \frac{c}{2f_c}$. Accordingly, for any two distinct angles of arrival $\theta, \theta' \in (-\frac{\pi}{2}, \frac{\pi}{2}]$, the corresponding array response vectors $\mathbf{s}(\theta)$ and $\mathbf{s}(\theta')$ are linearly independent.

The ULA collects N narrowband snapshots from K sources of interest (targets) arriving from distinct DoAs $\theta_1, \theta_2, \ldots, \theta_K \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right], K < D \leq N$. We assume that the system may also experience intermittent directional interference from L independent sources (jammers), at angles $\theta_1', \theta_2', \ldots, \theta_L' \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. A schematic illustration of the targets and jammers is given in Fig. 1. We assume that $\theta_i \neq \theta_q'$, for any $i \in \{1, 2, \ldots, K\}$ and $q \in \{1, 2, \ldots, L\}$. For any $l \in \{1, 2, \ldots, L\}$, the l-th jammer may be active during any of the N snapshots with some fixed and unknown
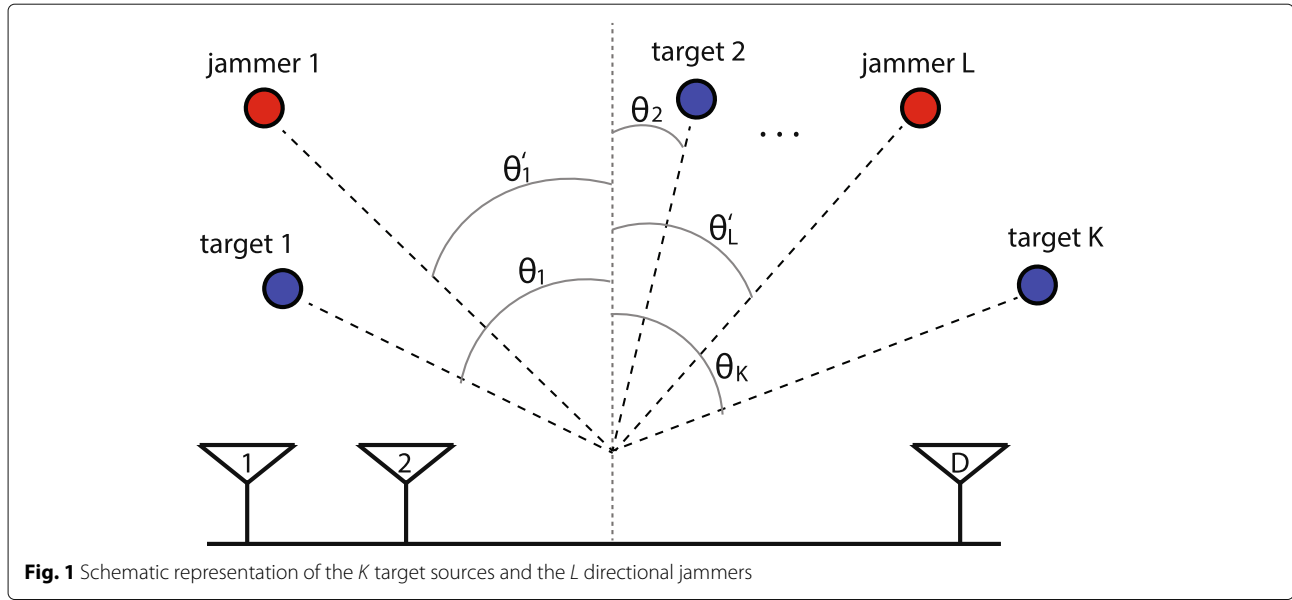
Fig. 1 Schematic representation of the K target sources and the L directional jammers

to the receiver probability p_l. Accordingly, the n-th down-converted received data vector is of the form

$$\mathbf{y}_n = \sum_{k=1}^{K} x_{n,k}\mathbf{s}(\theta_k) + \sum_{l=1}^{L} \gamma_{n,l}x'_{n,l}\mathbf{s}\left(\theta'_l\right) + \mathbf{n}_n \in \mathbb{C}^{D\times 1}, \quad (2)$$
$$n = 1, 2, \ldots, N,$$

where, $x_{n,k}$ and $x'_{n,l} \in \mathbb{C}$ denote the statistically independent signal values of target k and jammer l, respectively, comprising power-scaled information symbols and flat-fading channel coefficients, and $\gamma_{n,l}$ is the activity indicator for jammer l, modeled as a $\{0,1\}$-Bernoulli random variable with activation probability p_l. $\mathbf{n}_n \in \mathbb{C}^{D\times 1}$ accounts for additive white Gaussian noise (AWGN) with mean equal to zero and per-element variance σ^2; i.e., $\mathbf{n}_n \sim \mathcal{CN}\left(\mathbf{0}_D, \sigma^2\mathbf{I}_D\right)$. Henceforth, we refer to the case of target-only presence in the collected snapshots (i.e., $\gamma_{n,l} = 0$ for every $n = 1, 2, \ldots, N$ and every $l = 1, 2, \ldots, L$) as *normal system operation*.

Defining $\mathbf{x}_n \triangleq [x_{n,1}, x_{n,2}, \ldots, x_{n,K}]^\top$, $\mathbf{x}'_n \triangleq [x'_{n,1}, x'_{n,2}, \ldots, x'_{n,L}]^\top$, $\Gamma_n \triangleq \mathbf{diag}\left([\gamma_{n,1}, \gamma_{n,2}, \ldots, \gamma_{n,L}]^\top\right)$, and $\mathbf{S}_\Phi \triangleq [\mathbf{s}(\phi_1), \mathbf{s}(\phi_2), \ldots, \mathbf{s}(\phi_m)] \in \mathbb{C}^{D\times m}$ for any size-m set of angles $\Phi \triangleq \{\phi_1, \phi_2, \ldots, \phi_m\} \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]^{m}$,[1] (2) can be rewritten as

$$\mathbf{y}_n = \mathbf{S}_\Theta\mathbf{x}_n + \mathbf{S}_{\Theta'}\Gamma_n\mathbf{x}'_n + \mathbf{n}_n \in \mathbb{C}^{D\times 1}, \quad n = 1, 2, \ldots, N, \quad (3)$$

for $\Theta \triangleq \{\theta_1, \theta_2, \ldots, \theta_K\}$ and $\Theta' \triangleq \{\theta'_1, \theta'_2, \ldots, \theta'_L\}$. The goal of a DoA estimator is to identify correctly all angles

in the DoA set Θ. Importantly, by the Vandermonde structure of $\mathbf{S}_\Theta$, it holds that

$$\mathbf{s}(\phi) \subseteq \mathrm{span}(\mathbf{S}_\Theta) \Leftrightarrow \phi \in \Theta, \quad (4)$$

for any $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ [16]. That is, given $\mathcal{S} \triangleq \mathrm{span}(\mathbf{S}_\Theta)$, the receiver can decide accurately for any candidate angle $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ whether it is a DoA in Θ, or not.

2.1 DoA estimation under normal system operation.
Considering for a moment $p_l = 0$ for every $l \in \{1, 2, \ldots, L\}$, (2) becomes

$$\mathbf{y}_n = \mathbf{S}_\Theta\mathbf{x}_n + \mathbf{n}_n \in \mathbb{C}^{D\times 1}, \quad n = 1, 2, \ldots, N \quad (5)$$

with autocorrelation matrix $\mathbf{R} \triangleq E\left\{\mathbf{y}_n\mathbf{y}_n^H\right\} = \mathbf{S}_\Theta E\left\{\mathbf{x}_n\mathbf{x}_n^H\right\} \mathbf{S}_\Theta^H + \sigma^2\mathbf{I}_D$. Certainly, $\mathcal{S} = \mathrm{span}(\mathbf{S}_\Theta)$ coincides with the K-dimensional principal subspace of $\mathbf{R}$, spanned by its K highest-eigenvalue eigenvectors [5]. Therefore, being aware of $\mathbf{R}$, the receiver could obtain $\mathcal{S}$ through standard EVD and then conduct accurate DoA estimation by means of (4). However, in practice, the nominal received-signal autocorrelation matrix $\mathbf{R}$ is unknown to the receiver and sample-average estimated as $\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}_n\mathbf{y}_n^H$ [5, 16]. Accordingly, $\mathcal{S}$ is estimated by the span of the K highest-eigenvalue eigenvectors of $\hat{\mathbf{R}}$, which coincide with the K highest-singular-value left singular-vectors of $\mathbf{Y} \triangleq [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N]$. The eigenvectors of $\hat{\mathbf{R}}$, or left singular-vectors of $\mathbf{Y}$, are also commonly referred to as the L2-PCs of $\mathbf{Y}$, since they constitute a solution to the L2-PCA problem

$$\mathbf{Q}_{L2} = \underset{\mathbf{Q}\in\mathbb{C}^{D\times K}, \, \mathbf{Q}^H\mathbf{Q}=\mathbf{I}_K}{\mathrm{argmax}} \left\|\mathbf{Q}^H\mathbf{Y}\right\|_2^2. \quad (6)$$

In accordance to (4), the DoA set Θ is estimated by the arguments that yield the K local maxima (peaks) of the familiar MUSIC [16] spectrum

$$P(\phi) = \left\| \left(\mathbf{I}_D - \mathbf{Q}_{L2}\mathbf{Q}_{L2}^H\right) \mathbf{s}(\phi) \right\|_2^{-2}, \quad \phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad (7)$$

which clarifies why MUSIC is, in fact, an L2-PCA-based DoA estimation method. Certainly, as N increases asymptotically, $\hat{\mathbf{R}}$ tends to $\mathbf{R}$, $\mathbf{Q}_{L2}$ tends to span $\mathbf{S}_\Theta$, and $P(\phi)$ goes to infinity for every $\phi \in \Theta$ and finding its peaks becomes a criterion equivalent to (4). Therefore, for sufficient N, L2-PCA-based MUSIC is well-known to attain high performance in normal system operation.

2.2 Complications in the presence of unexpected jamming

In this work, we focus on the case where $p_l > 0$ for all l, so that some snapshots in $\mathbf{Y}$ are corrupted by unexpected, unknown, directional interference, as modeled in (2). In this case, the K eigenvectors of $\mathbf{R} = E\left\{\mathbf{y}_n\mathbf{y}_n^H\right\}$ do not span $\mathcal{S}$ any more. Thus, the K eigenvectors of $\hat{\mathbf{R}}$ or singular-vectors of $\mathbf{Y}$ would be of no use, even for very high sample-support N. In fact, interference-corrupted snapshots in $\mathbf{Y}$ may constitute *outliers* with respect to $\mathcal{S}$. Accordingly, due to the well documented high responsiveness of L2-PCA in (6) to outlying data, $\mathbf{Q}_{L2}$ may diverge significantly from $\mathcal{S}$ [29, 48], rendering DoA estimation by means of (7) highly inaccurate. Below, we introduce a novel method that exploits the outlier-resistance of L1-PCA [36, 47, 48] to offer improved DoA estimates.

3 Proposed DoA estimation method
3.1 Operation on realified snapshots

In order to employ L1-PCA algorithms that are defined for the processing of real-valued data, the proposed DoA estimation method operates on real-valued representations of the recorded complex snapshots in (2), similar to a number of previous works in the field [58–60]. In particular, we define the real-valued representation of any complex-valued matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$, by concatenating its real and imaginary parts, as

$$\overline{\mathbf{A}} \triangleq \begin{bmatrix} \Re\{\mathbf{A}\}, & -\Im\{\mathbf{A}\} \\ \Im\{\mathbf{A}\}, & \Re\{\mathbf{A}\} \end{bmatrix} \in \mathbb{R}^{2m \times 2n}. \quad (8)$$

In Lie algebras and representation theory, this transition from $C^{m \times n}$ to $\mathbb{R}^{2m \times 2n}$ is commonly referred to as complex-number *realification* [61, 62] and is a method that allows for any complex system of equations to be converted into (and solved through) a corresponding real system [63]. Lemmas 1, 2, and 3 presented in the Appendix provide three important properties of realification. By (8)

and Lemma 1, the nth complex snapshot $\mathbf{y}_n$ in (3) can be *realified* as

$$\overline{\mathbf{y}}_n = \overline{\mathbf{S}}_\Theta \overline{\mathbf{x}}_n + \overline{\mathbf{S}}_{\Theta'} \overline{\Gamma}_n \overline{\mathbf{x}'}_n + \overline{\mathbf{n}}_n \in \mathbb{R}^{2D \times 2}. \quad (9)$$

In accordance with Lemma 2, the rank of $\overline{\mathbf{S}}_\Theta$ is $2K$ and, hence, $\mathcal{S}_R \triangleq \text{span}\left(\overline{\mathbf{S}}_\Theta\right)$ is a $2K$-dimensional subspace wherein the K realified signal components of interest with angles of arrival in Θ lie. The following Proposition, deriving straightforwardly from (4) by means of Lemma 1 and Lemma 2, highlights the utility of $\mathcal{S}_R$ for estimating the target DoAs.

Proposition 1 *For any $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$, it holds that*

$$\text{span}(\overline{\mathbf{s}}(\phi)) \subseteq \mathcal{S}_R \iff \phi \in \Theta. \quad (10)$$

Set equality may hold only if $K = 1$. ∎

By Proposition 1, given an orthonormal basis $\mathbf{Q}_R \in \mathbb{R}^{2D \times 2K}$ that spans $\mathcal{S}_R$, the receiver can decide accurately whether some $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is a target DoA, or not, by means of the criterion

$$\left(\mathbf{I}_{2D} - \mathbf{Q}_R\mathbf{Q}_R^\top\right)\overline{\mathbf{s}}(\phi) = \mathbf{0}_{2D \times 2} \iff \phi \in \Theta. \quad (11)$$

Similar to the complex-data case presented above, in normal system operation, $\mathcal{S}_R$ coincides with the span of the K dominant eigenvectors of $\mathbf{R}_R \triangleq E\left\{\overline{\mathbf{y}}_n\overline{\mathbf{y}}_n^\top\right\}$. When the receiver, instead of $\mathbf{R}_R$, possesses only the realified snapshot record $\overline{\mathbf{Y}}$, $\mathcal{S}_R$ can be estimated as the span of

$$\mathbf{Q}_{R,L2} = \underset{\mathbf{Q} \in \mathbb{R}^{2D \times 2K},\ \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_{2K}}{\text{argmax}} \left\| \mathbf{Q}^\top \overline{\mathbf{Y}} \right\|_2^2. \quad (12)$$

Then, in accordance with (11), the target DoAs can be estimated as the arguments that yield the K highest peaks of the spectrum

$$P_R(\phi;\ \mathbf{Q}_{R,L2}) \triangleq 2 \left\| \left(\mathbf{I}_{2D} - \mathbf{Q}_{R,L2}\mathbf{Q}_{R,L2}^\top\right) \overline{\mathbf{s}}(\phi) \right\|_2^{-2},$$
$$\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]. \quad (13)$$

Similar to (6), the solution to (12) can be obtained by singular-value decomposition (SVD) of $\overline{\mathbf{Y}}$. Interestingly, the L2-PCA-based DoA estimator of (13) is equivalent to the complex-field MUSIC estimator presented in Section 2. In fact, as we prove in the Appendix,

$$P_R(\phi; \mathbf{Q}_{R,L2}) = P(\phi) \quad \forall \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right). \quad (14)$$

Hence, exhibiting performance identical to that of MUSIC, (12) can offer highly accurate estimates of the target DoAs under normal system operation. However, when $\mathbf{Y}$ contains corrupted snapshots, the L2-PCA-calculated span $(\mathbf{Q}_{R,L2})$ is a poor approximation to $\mathcal{S}_R$ and DoA estimation by means of $P_R(\phi; \mathbf{Q}_{R,L2})$ tends to be highly inaccurate. In the following subsection, we present

an alternative, L1-PCA-based method for obtaining an outlier-resistant estimate of Θ.

3.2 DoA estimation by realified L1-PCA

Over the past few years, L1-PCA has been shown to be far more resistant than L2-PCA against outliers in the data matrix [31–40, 47, 48]. In this work, we propose the use of a DoA-estimation spectrum analogous to that in (13) that is formed by the L1-PCs of $\overline{\mathbf{Y}}$. Specifically, the proposed method has two steps. First, we obtain the L1-PCs of $\overline{\mathbf{Y}}$, solving the L1-PCA problem

$$\mathbf{Q}_{R,L1} = \underset{\mathbf{Q}\in\mathbb{R}^{2D\times 2K},\ \mathbf{Q}^\top\mathbf{Q}=\mathbf{I}_{2K}}{\text{argmax}} \sum_{n=1}^{N} \left\| \mathbf{Q}^\top \overline{\mathbf{y}}_n \right\|_1. \qquad (15)$$

That is, (15) searches for the subspace that maximizes data presence, quantified as the aggregate L1-norm of the projected points.

Then, similarly to MUSIC, we estimate the target angles in Θ by the K highest peaks of the L1-PCA-based spectrum

$$P_R(\phi; \mathbf{Q}_{R,L1}) = 2 \left\| \left(\mathbf{I}_{2D} - \mathbf{Q}_{R,L1}\mathbf{Q}_{R,L1}^\top \right) \overline{\mathbf{s}}(\phi) \right\|_2^{-2}, \qquad (16)$$
$$\phi \in (-\frac{\pi}{2}, \frac{\pi}{2}].$$

In accordance to standard practice, to find the K highest peaks of (16), we examine every angle in $\left\{ \phi = -\frac{\pi}{2} + k\Delta\phi : k \in \left\{ 1, 2, \ldots, \left\lfloor \frac{\pi}{\Delta\phi} \right\rfloor \right\} \right\}$, for some small scanning step $\Delta\phi > 0$. Next, we place our focus on solving the L1-PCA in (15).

3.3 Principles of realified L1-PCA

Although L1-PCA is not a new problem in the literature (see, e.g., [36–38]), its exact optimal solution was unknown until the recent work in [48], where the authors proved that (15) is formally NP-hard and offered the first two exact algorithms for solving it. Proposition 2 below, originally presented in [48] for real-valued data matrices of general structure (i.e., not having necessarily the realified structure of $\overline{\mathbf{Y}}$) translates L1-PCA in (15) to a nuclear-norm maximization problem over the binary field.

Proposition 2 *If* $\mathbf{B}_{opt}$ *is a solution to*

$$\underset{\mathbf{B}\in\{\pm 1\}^{2N\times 2K}}{\text{maximize}} \left\| \overline{\mathbf{Y}}\mathbf{B} \right\|_*^2 \qquad (17)$$

and $\overline{\mathbf{Y}}\mathbf{B}_{opt}$ *admits SVD* $\overline{\mathbf{Y}}\mathbf{B}_{opt} = \mathbf{U}\Sigma_{2K\times 2K}\mathbf{V}^\top$, *then*

$$\mathbf{Q}_{R,L1} = \mathbf{U}\mathbf{V}^\top \qquad (18)$$

is a solution to (15). *Moreover,* $\left\| \mathbf{Q}_{R,L1}^\top \overline{\mathbf{Y}} \right\|_1 = \left\| \overline{\mathbf{Y}}\mathbf{B}_{opt} \right\|_*$. ∎

Since $\mathbf{Q}_{R,L1}$ can be obtained by $\mathbf{B}_{opt}$ via standard SVD, L1-PCA is in practice equivalent to a combinatorial optimization problem over the $4NK$ binary variables in $\mathbf{B}$. The authors in [48] presented two algorithms for exact solution of (17), defined upon real-valued data matrices of general structure.

In this work, for the first time, we simplify the solutions of [48] in view of the special, realified structure of $\overline{\mathbf{Y}}$. Specifically, in the following Proposition 3, we show that for $K = 1$ we can exploit the special structure of $\overline{\mathbf{Y}}$ and reduce (17) to a binary quadratic-form maximization problem over half the number of binary variables (i.e., $2N$ instead of $4N$). A proof for Proposition 3 is provided in the Appendix.

Proposition 3 *If* $\mathbf{b}_{opt}$ *is a solution to*

$$\underset{\mathbf{b}\in\{\pm 1\}^{2N\times 1}}{\text{maximize}} \left\| \overline{\mathbf{Y}}\mathbf{b} \right\|_2^2, \qquad (19)$$

then $[\mathbf{b}_{opt}, \mathbf{E}_N\mathbf{b}_{opt}]$ *is a solution to*

$$\underset{\mathbf{B}\in\{\pm 1\}^{2N\times 2}}{\text{maximize}} \left\| \overline{\mathbf{Y}}\mathbf{B} \right\|_*^2 \qquad (20)$$

with $\left\| \overline{\mathbf{Y}}[\mathbf{b}_{opt}, \mathbf{E}_N\mathbf{b}_{opt}] \right\|_*^2 = 4 \left\| \overline{\mathbf{Y}}\mathbf{b}_{opt} \right\|_2^2$. ∎

In view of Propositions 2 and 3, $\mathbf{Q}_{R,L1}$ derives easily from the solution of

$$\underset{\mathbf{B}\in\{\pm 1\}^{2N\times m}}{\text{maximize}} \left\| \overline{\mathbf{Y}}\mathbf{B} \right\|_* , \qquad (21)$$

for $m = 1$, if $K = 1$, or $m = 2K$, if $K > 1$.

Since (21) is a combinatorial problem, the conceptually simplest approach for solving it is an exhaustive search (possibly in parallel fashion) over all elements of its feasibility set $\{\pm 1\}^{2N\times m}$. By means of this method, one should conduct 2^{2Nm} nuclear norm evaluations (e.g., by means of SVD of $\overline{\mathbf{Y}}\mathbf{B}$) to identify the optimum argument in the feasibility set; thus, the asymptotic complexity of this method is $\mathcal{O}\left(2^{2Nm}\right)$. Exploiting the well-known nuclear-norm properties of column-permutation and column-negation invariance, we can expedite practically the exhaustive procedure by searching for a solution to (21) in the set of all binary matrices that are column-wise built by the elements of a size-m multiset[2] of $\{\mathbf{b} \in \{\pm 1\}^{2N} : [\mathbf{b}]_1 = 1\}$. By this modification, the exact number of binary matrices examined (thus, the number of nuclear-norm evaluations) decreases from 2^{2Nm} to $\binom{2^{2N-1}+2K-1}{m}$. Of course, exhaustive-search approaches, being of exponential complexity in N, become impractical as the number of snapshots increases. For completeness, in Fig. 2, we provide a pseudocode for the exhaustive-search algorithm presented above.

For the case of engineering interest where $N > D$ and D is a constant, the authors in [48] presented a polynomial-cost algorithm that solves (21) with complexity $\mathcal{O}(N^{2Dm})$.

Algorithm 1: Exponential-cost, optimal L1-PCA of $\overline{\mathbf{Y}}$

Input: i) $\overline{\mathbf{Y}}_{2D \times 2N}$ and ii) K

1: $\mathbf{B}_{2N \times 2^{2N-1}} \leftarrow$ concatenation of all vectors in $\{\mathbf{b} \in \{\pm 1\}^{2N} : [\mathbf{b}]_1 = 1\}$
2: If $K = 1$, $m \leftarrow 1$; else, $m \leftarrow 2K$
3: $\mathbf{z}_{\text{opt}} \leftarrow \arg\max_{\mathbf{z} \in \{1, 2, \dots, 2^{2N-1}\}^m, z_1 \leq z_2 \leq \dots \leq z_m} \|\mathbf{X}\mathbf{B}_{:,\mathbf{z}}\|_*$
4: If $m = 1$, $\mathbf{B}_{\text{opt}} \leftarrow [\mathbf{B}_{:,\mathbf{z}_{\text{opt}}}, \mathbf{E}_N \mathbf{B}_{:,\mathbf{z}_{\text{opt}}}]$; else, $\mathbf{B}_{\text{opt}} \leftarrow \mathbf{B}_{:,\mathbf{z}_{\text{opt}}}$
5: $(\mathbf{U}_{2D \times 2K}, \mathbf{\Sigma}_{2K \times 2K}, \mathbf{V}_{2K \times 2K}) \leftarrow \text{svd}(\mathbf{X}\mathbf{B}_{\text{opt}})$

Output: $\mathbf{Q}_{R,L1} \leftarrow \mathbf{U}\mathbf{V}^\top$

Fig. 2 Algorithm for optimal computation of the $2K$ L1-PCs of rank-$2D$ data matrix $\overline{\mathbf{Y}}_{2D \times 2N}$ with exponential (w.r.t. N) asymptotic complexity $\mathcal{O}\left(2^{2Nm}\right)$ ($m = 1$, for $K = 1$; $m = 2K$, for $K > 1$)

In the following subsection, we exploit further the structure of $\overline{\mathbf{Y}}$ and reduce significantly the computational cost of this algorithm.

3.4 Polynomial-cost realified L1-PCA

The authors in [48] showed that, according to Proposition 2, a solution to (21) can be found among the binary matrices that draw columns from

$$\mathcal{B} \stackrel{\triangle}{=} \left\{ \text{sgn}\left(\overline{\mathbf{Y}}^\top \mathbf{a}\right) : \mathbf{a} \in \Omega_{2D} \right\} \subseteq \{\pm 1\}^{2N \times 1} \qquad (22)$$

where $\Omega_{2D} \stackrel{\triangle}{=} \left\{ \mathbf{a} \in \mathbb{R}^{2D \times 1} : \|\mathbf{a}\|_2 = 1, [\mathbf{a}]_{2D} > 0 \right\}$ —with the positivity constraint in the last entry of $\mathbf{a}$ deriving from the invariance of the nuclear norm to column negations of its matrix argument. That is, a solution to (21) belongs to the mth Cartesian power of $\mathcal{B}$, $\mathcal{B}^m \subseteq \{\pm 1\}^{2N \times m}$.

In addition, [48] pointed out that, since the nuclear-norm maximization is also invariant to column permutations of the argument, we can maintain problem equivalence while further narrowing down our search to the elements of a set $\tilde{\mathcal{B}}$, subset of $\mathcal{B}^m$, that contains the $\binom{|\mathcal{B}|+m-1}{m}$ binary matrices that are built by the elements of all size-m multisets of $\mathcal{B}$. That is, we can obtain a solution to (21) by solving instead

$$\underset{\mathbf{B} \in \tilde{\mathcal{B}}}{\text{maximize}} \left\|\overline{\mathbf{Y}}\mathbf{B}\right\|_*^2. \qquad (23)$$

Importantly, $|\tilde{\mathcal{B}}| = \binom{|\mathcal{B}|+m-1}{m} < |\mathcal{B}|^m = |\mathcal{B}^m|$. The exact multiset-extraction procedure for obtaining $\tilde{\mathcal{B}}$ from $\mathcal{B}$ follows.

Calculation of $\tilde{\mathcal{B}}$ from $\mathcal{B}$ [48]. For every $i \in \left\{1, 2, \dots, \binom{|\mathcal{B}|+m-1}{m}\right\}$, we define a distinct indicator function $f_i : \mathcal{B} \mapsto \{0, 1, \dots, m\}$ that assigns to every $\mathbf{b} \in \mathcal{B}$ a natural number $f_i(\mathbf{b}) \leq m$, such that $\sum_{\mathbf{b} \in \mathcal{B}} f_i(\mathbf{b}) = m$. Then, for every $i \in \left\{1, 2, \dots, \binom{|\mathcal{B}|+m-1}{m}\right\}$, we define a unique binary matrix $\mathbf{B}_i \in \{\pm 1\}^{2N \times m}$ such that every $\mathbf{b} \in \mathcal{B}$ appears exactly $f_i(\mathbf{b})$ times among the columns of $\mathbf{B}_i$. Finally, we define the sought-after set as $\tilde{\mathcal{B}} \stackrel{\triangle}{=} \left\{\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_{\binom{|\mathcal{B}|+m-1}{m}}\right\}$.

Evidently, the cost to solve (23), and thus (21), amounts to the cost of constructing the feasibility set $\tilde{\mathcal{B}}$ added to the cost of conducting nuclear-norm evaluations (through SVD) over all its elements. Therefore, the cost to solve (23) depends on the construction cost and cardinality of $\tilde{\mathcal{B}}$. As seen above, $|\tilde{\mathcal{B}}| = \binom{|\mathcal{B}|+m-1}{m}$ and $\tilde{\mathcal{B}}$ can be constructed online, by multiset selection on $\mathcal{B}$, with negligible computational cost. Therefore, for determining the cardinality and construction cost of $\tilde{\mathcal{B}}$, we have to find the cardinality and construction cost of $\mathcal{B}$.

Next, we present a novel method to construct $\mathcal{B}$, different than the one in [48], that exploits the realified structure of $\overline{\mathbf{Y}}$ to achieve lower computational cost.

Construction of $\mathcal{B}$, in view of the structure of $\overline{\mathbf{Y}}$.

Considering that any group of $m \leq 2D$ columns of $\overline{\mathbf{Y}}$ spans a m-dimensional subspace, for each index set $\mathcal{X} \subseteq \{1, 2, \dots, 2N\}$ —elements in ascending order (e.a.o.)— of cardinality $|\mathcal{X}| = 2D - 1$, we denote by $\mathbf{z}(\mathcal{X})$ the unique left-singular vector of $[\overline{\mathbf{Y}}]_{:,\mathcal{X}}$ that corresponds to zero singular value. Calculation of $\mathbf{z}(\mathcal{X})$ can be achieved either by means of SVD or by simple Gram-Schmidt Orthonormalization (GMO) of $[\overline{\mathbf{Y}}]_{:,\mathcal{X}}$ —both SVD and GMO are of constant cost with respect to N. Accordingly, we define

$$\mathbf{c}(\mathcal{X}) \stackrel{\triangle}{=} \text{sgn}([\mathbf{z}(\mathcal{X})]_{2D})\mathbf{z}(\mathcal{X}) \in \Omega_{2D}. \qquad (24)$$

Being a scaled version of $\mathbf{z}(\mathcal{X})$, $\mathbf{c}(\mathcal{X})$ also belongs to $\text{null}\left([\overline{\mathbf{Y}}]_{:,\mathcal{X}}^\top\right)$, satisfying $[\overline{\mathbf{Y}}]_{:,\mathcal{X}}^\top \mathbf{c}(\mathcal{X}) = \mathbf{0}_{2D-1}$. Next, we define the set of binary vectors

$$\mathcal{B}(\mathcal{X}) \stackrel{\triangle}{=} \left\{\mathbf{b} \in \{\pm 1\}^{2N \times 1} : [\mathbf{b}]_{\mathcal{X}^c} = \text{sgn}\left([\overline{\mathbf{Y}}]_{:,\mathcal{X}^c}^\top \mathbf{c}(\mathcal{X})\right)\right\} \qquad (25)$$

of cardinality $|\mathcal{B}(\mathcal{X})| = 2^{2D-1}$, where $\mathcal{X}^c \stackrel{\triangle}{=} \{1, 2, \dots, 2N\} \setminus \mathcal{X}$ (e.a.o.) is the complement of $\mathcal{X}$. In [48], the authors showed that

$$\mathcal{B} = \bigcup_{\substack{\mathcal{X} \subseteq \{1,2,\ldots,2N\} \\ |\mathcal{X}| = 2D-1}} \mathcal{B}(\mathcal{X}). \qquad (26)$$

Since $\mathcal{X}$ can take $\binom{2N}{2D-1}$ different values, $\mathcal{B}$ can be built by (26) through $\binom{2N}{2D-1}$ nullspace calculations in the form of (24), with cost $\binom{2N}{2D-1}D^3 \in \mathcal{O}\left(N^{2D}\right)$. Accordingly, $\mathcal{B}$ consists of

$$|\mathcal{B}| \le 2^{2D-1}\binom{2N}{2D-1} \in \mathcal{O}\left(N^{2D-1}\right) \qquad (27)$$

elements. In fact, in view of [64], the exact cardinality of $\mathcal{B}$ is

$$|\mathcal{B}| = \sum_{d=0}^{2D-1}\binom{2N-1}{d} \in \mathcal{O}\left(N^{2D-1}\right). \qquad (28)$$

Next, we show for the first time how we can reduce the cost of calculating $\mathcal{B}$, exploiting the realified structure of $\overline{\mathbf{Y}}$.

Consider $\mathcal{X}_1 \subseteq \{1,2,\ldots,N\}$ (e.a.o.), $\mathcal{X}_2 \subseteq \{N+1,N+2,\ldots,2N\}$ (e.a.o.), and their union $\mathcal{X}_A = \{\mathcal{X}_1,\mathcal{X}_2\}$ (e.a.o.), such that $|\mathcal{X}_1| < D$ and $|\mathcal{X}_A| = |\mathcal{X}_1| + |\mathcal{X}_2| = 2D-1$. Define also the set of indices $\mathcal{X}_B = \{\mathcal{X}_1 + N, \mathcal{X}_2 - N\}$ (e.a.o.) with $|\mathcal{X}_B| = 2D-1$. By the structure of $\overline{\mathbf{Y}}$, it is straightforward that

$$\mathbf{c}(\mathcal{X}_B) = \mathbf{E}_D\mathbf{c}(\mathcal{X}_A)\mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D). \qquad (29)$$

In turn, by the definition in (25) and (29), it holds that

$$\mathcal{B}(\mathcal{X}_B) = \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{E}_N\mathcal{B}(\mathcal{X}_A). \qquad (30)$$

The proof of (29) and (30) is offered in the Appendix. Notice now that, for every $\mathcal{X} \subset \{1,2,\ldots,2N\}$ with $|\mathcal{X}| = 2D-1$, there exist $\mathcal{X}_1 \subset \{1,2,\ldots,N\}$ and $\mathcal{X}_2 \subset \{N+1,N+2,\ldots,2N\}$, satisfying $|\mathcal{X}_1| < D$ and $|\mathcal{X}_1| + |\mathcal{X}_2| = 2D-1$, such that

$$\text{either } \mathcal{X} = \{\mathcal{X}_1,\mathcal{X}_2\}, \text{ or } \mathcal{X} = \{\mathcal{X}_1+N,\mathcal{X}_2-N\}. \qquad (31)$$

Thus, by (26) and (31), $\mathcal{B}$ can constructed as

$$\mathcal{B} = \bigcup_{d=0}^{D-1} \bigcup_{\substack{\mathcal{X}_1 \subset \{1,2,\ldots,N\},\ |\mathcal{X}_1|=d \\ \mathcal{X}_2 \subset \{\{1,2,\ldots,N\}+N\},\ |\mathcal{X}_2|=2D-1-d}} \{\mathcal{B}(\{\mathcal{X}_1,\mathcal{X}_2\}), \mathcal{B}(\{\mathcal{X}_1+N,\mathcal{X}_2-N\})\}. \qquad (32)$$

In view of (30), $\mathcal{B}(\{\mathcal{X}_1+N,\mathcal{X}_2-N\})$ can be directly constructed from $\mathcal{B}(\{\mathcal{X}_1,\mathcal{X}_2\})$ with negligible computational overhead. In addition, for $|\mathcal{X}_1| < D$ and $|\mathcal{X}_1| + |\mathcal{X}_2| = 2D-1$, by the Chu-Vandermonde binomial-coefficient property [65], $\{\mathcal{X}_1,\mathcal{X}_2\}$ can take

$$\sum_{d=0}^{D-1}\binom{N}{d}\binom{N}{2D-1-d} = \frac{1}{2}\binom{2N}{2D-1} \qquad (33)$$

values. Therefore, exploiting the structure of $\overline{\mathbf{Y}}$, the proposed algorithm constructs $\mathcal{B}$ by (32), avoiding half the

nullspace calculations needed in the generic method of [48], presented in (26).

In view of (28), the feasibility set of (23), $\tilde{\mathcal{B}}$, consists of exactly

$$|\tilde{\mathcal{B}}| = \binom{|\mathcal{B}| + m - 1}{m} = \binom{\sum_{d=0}^{2D-1}\binom{2N-1}{d} + m - 1}{m}$$
$$\in \mathcal{O}\left(N^{2Dm-m}\right) \qquad (34)$$

elements. Thus, $\mathcal{O}(N^{2Dm-m})$ nuclear-norm evaluations suffice to obtain a solution to (17). The asymptotic complexity for solving (15) by the presented algorithm is then $\mathcal{O}\left(N^{2Dm-m}\right)$. The described polynomial-time algorithm is presented in detail in Fig. 3, including element-by-element construction of $\mathcal{B}$.

3.5 Iterative realified L1-PCA

For large problem instances (large N, D), the above presented optimal L1-PCA calculators could be computationally impractical. Therefore, at this point we present and employ the bit-flipping-based iterative L1-PCA calculator, originally introduced in [43] for processing general real-valued data matrices. Given $\overline{\mathbf{Y}} \in \mathbb{R}^{2D \times 2N}$, and some $m < 2\min(D,N)$, the algorithm presented below attempts to solve (21), by conducting a converging sequence of optimal single-bit flips.

Specifically, the algorithm is initializes at a $2N \times m$ binary matrix $\mathbf{B}^{(1)}$ and conducts optimal single-bit flipping iterations. Specifically, at the tth iteration step ($t > 1$), the algorithm generates the new matrix $\mathbf{B}^{(t)}$ so that (i) $\mathbf{B}^{(t)}$ differs from $\mathbf{B}^{(t-1)}$ in exactly one entry (bit flipping) and (ii) $\|\overline{\mathbf{Y}}\mathbf{B}^{(t)}\|_* > \|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)}\|_*$. Mathematically, we notice that if we flip at the tth iteration the (n,k)th bit of $\mathbf{B}^{(t-1)}$ setting $\mathbf{B}^{(t)} = \mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{n,k}\mathbf{e}_{n,2N}\mathbf{e}_{k,m}^\top$, it holds that

$$\overline{\mathbf{Y}}\mathbf{B}^{(t)} = \overline{\mathbf{Y}}\mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{n,k}[\overline{\mathbf{Y}}]_{:,n}\mathbf{e}_{k,m}^\top. \qquad (35)$$

Therefore, at step t, the presented algorithm searches for a solution (n,k) to

$$\max_{\substack{(j,l) \in \{1,2,\ldots,2N\} \times \{1,2,\ldots,m\} \\ (l-1)2N+j \in \mathcal{L}^{(t)}}} \left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{j,l}[\overline{\mathbf{Y}}]_{:,j}\mathbf{e}_l^{m\top}\right\|_*. \qquad (36)$$

The constraint set $\mathcal{L}^{(t)} \subseteq \{1,2,\ldots,2Nm\}$, employed to restrain the greediness of the presented iterations, contains the indices[3] of bits that have not been flipped before and, thus, is initialized as $\mathcal{L}^{(1)} = \{1,2,\ldots,2Nm\}$. Having obtained the solution to (36), (n,k), the algorithm proceeds as follows. If $\left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{n,k}[\overline{\mathbf{Y}}]_{:,n}\mathbf{e}_{k,m}^\top\right\|_* > \left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)}\right\|_*$, the algorithm generates $\mathbf{B}^{(t)} = \mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{n,k}\mathbf{e}_{n,2N}\mathbf{e}_{k,m}^\top$

Algorithm 2: Polynomial-cost, optimal L1-PCA of $\overline{\mathbf{Y}}$

Input: i) $\overline{\mathbf{Y}}_{2D \times 2N}$ and ii) K

1:   If $K = 1$, $m \leftarrow 1$; else, $m \leftarrow 2K$
2:   $\mathbf{B}_{2N \times M} \leftarrow$ compute_candidates($\overline{\mathbf{Y}}^\top$), $\mathcal{M} \leftarrow \{1, 2, \ldots, M\}$
3:   $\mathbf{z}_{\mathrm{opt}} \leftarrow \arg\max_{\mathbf{z} \in \mathcal{M}^m, z_1 \leq z_2 \leq \ldots \leq z_m} \|\mathbf{X}\mathbf{B}_{:,\mathbf{z}}\|_*$
4:   If $m = 1$, $\mathbf{B}_{\mathrm{opt}} \leftarrow [\mathbf{B}_{:,\mathbf{z}_{\mathrm{opt}}}, \overline{\mathbf{E}}_N \mathbf{B}_{:,\mathbf{z}_{\mathrm{opt}}}]$; else, $\mathbf{B}_{\mathrm{opt}} \leftarrow \mathbf{B}_{:,\mathbf{z}_{\mathrm{opt}}}$
5:   $(\mathbf{U}_{2D \times 2K}, \boldsymbol{\Sigma}_{2K \times 2K}, \mathbf{V}_{2K \times 2K}) \leftarrow$ svd($\mathbf{X}\mathbf{B}_{\mathrm{opt}}$)

Output: $\mathbf{Q}_{R,L1} \leftarrow \mathbf{U}\mathbf{V}^\top$

Function *compute_candidates* $(\mathbf{Z}_{n \times m})$ –with $n > m$

1:   for $d = 0 : D - 1$
2:    for all $X_1 \subset \{1, \ldots, \frac{n}{2}\}$ s.t. $|X_1| = d$
3:     for all $X_2 \subset \{\frac{n}{2} + 1, \ldots, n\}$ s.t. $|X_2| = m - 1 - d$
4:      $\mathcal{I} = \{\mathcal{X}_1, \mathcal{X}_2\}$, $i \leftarrow i + 1$,
5:      $\mathbf{c}_{m \times 1} \leftarrow \mathrm{null}([\mathbf{Z}]_{\mathcal{I},:})$, $\mathbf{c} \leftarrow \mathrm{sgn}(c_m)\mathbf{c}$
6:      $\mathbf{F} \leftarrow \mathrm{sgn}(\mathbf{Z}\mathbf{c})\mathbf{1}_{2^m}^\top$
7:      $[\mathbf{F}]_{\mathcal{I},:} \leftarrow$ all 2^m binary vectors of length m
8:      $\mathbf{B}_{\mathcal{I},(i-1)2^m+1:i2^m} \leftarrow \mathbf{F}$, $i \leftarrow i + 1$
9:      $\mathbf{B}_{\mathcal{I},(i-1)2^m+1:i2^m} \leftarrow \mathrm{sgn}(c_{\frac{m}{2}})\mathbf{E}_N \mathbf{F}$
10:   Return $\mathbf{B}$

Fig. 3 Algorithm for optimal computation of the $2K$ L1-PCs of the rank-$2D$ data matrix $\overline{\mathbf{Y}}_{2D \times 2N}$ with polynomial (w.r.t. N) asymptotic complexity $\mathcal{O}\left(N^{4DK-m+1}\right)$ ($m = 1$ for $K = 1$; $m = 2K$, for $K > 1$)

and updates $\mathcal{L}^{(t+1)}$ to $\mathcal{L}^{(t)} \setminus \{(k-1)2N + n\}$. If, otherwise, $\left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{n,k}[\overline{\mathbf{Y}}]_{:,n}\mathbf{e}_{k,m}^\top\right\|_* \leq \left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)}\right\|_*$, the algorithm obtains a new solution (n, k) to (36) after resetting $\mathcal{L}^{(t)}$ to $\{1, 2, \ldots, 2Nm\}$. If this new (n, k) is such that $\left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{n,k}[\overline{\mathbf{Y}}]_{:,n}\mathbf{e}_{k,m}^\top\right\|_* > \left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)}\right\|_*$, then the algorithm sets $\mathbf{B}^{(t)} = \mathbf{B}^{(t-1)} - 2[\mathbf{B}^{(t-1)}]_{n,k}\mathbf{e}_{n,2N}\mathbf{e}_{k,m}^\top$ and updates $\mathcal{L}^{(t+1)} = \mathcal{L}^{(t)} \setminus \{(k-1)2N + n\}$. Otherwise, the iterations terminate and the algorithm returns $\mathbf{B}^{(t)}$ as a heuristic solution to (21). Notice that since at each iteration, the optimization metric increases. At the same time, the metric is certainly upper-bounded by $\|\overline{\mathbf{Y}}\mathbf{B}_{opt}\|_*$. Therefore, the iterations are guaranteed to terminate in a finite number of steps, for any initialization $\mathbf{B}^{(1)}$. Our studies have shown that, in fact, the iterations terminate for $t < 2Nm$, with very high frequency of occurrence.

For solving (36), one has to calculate $\left\|\overline{\mathbf{Y}}\mathbf{B}^{(t-1)} - 2\left[\mathbf{B}^{(t-1)}\right]_{j,l}[\overline{\mathbf{Y}}]_{:,j}\mathbf{e}_{l,m}^\top\right\|_*$, for all $(j, l) \in \{1, 2, \ldots, 2N\} \times \{1, 2, \ldots, m\}$ such that $(l-1)2N + j \in \mathcal{L}$. At worst case, $\mathcal{L} = \{1, 2, \ldots, 2Nm\}$ and this demands $2Nm$ independent singular-value/nuclear-norm calculations. Therefore, the total cost for solving (36) is $\mathcal{O}\left(N^2m^3\right)$. If we limit the number of iterations to $2Nm$, for the sake of practicality, then the total cost for obtaining a heuristic solution to (21) is $\mathcal{O}\left(N^3m^4\right)$ –significantly lower than the cost of the polynomial-time optimal algorithm presented

above, $\mathcal{O}\left(N^{2Dm-m}\right)$. When the iterations terminate, the algorithm returns the bit-flipping-derived L1-PC matrix $\mathbf{Q}_{R,BF} \triangleq \mathbf{U}\mathbf{V}^\top$, where $\mathbf{U}\boldsymbol{\Sigma}_{2K \times 2K}\mathbf{V}^\top \stackrel{\mathrm{svd}}{=} \overline{\mathbf{Y}}\mathbf{B}_{opt}$. Formal performance guarantees for the presented bit-flipping procedure were offered in [43], for general real-valued matrices and $K = 1$.

A pseudocode of the presented algorithm for the calculation of the $2K$ L1-PCs $\overline{\mathbf{Y}}_{2D \times 2N}$ is presented in Fig. 4.

4 Numerical results and discussion

We present numerical studies to evaluate the DoA estimation performance of realified L1-PCA, compared to other PCA calculation counterparts. Our focus lies on cases where a nominal source (the DoA of which we are looking for) operates in the intermittent presence of a jammer located at a different angle. Ideally, we would like the DoA estimator to be able to identify successfully the DoA of the source of interest, despite the unexpected directional interference.

To offer a first insight into the performance of the proposed method, in Fig. 5 we present a realization of the DoA-estimation spectra $P_R(\phi; \mathbf{Q}_{R,L1})$ and $P_R(\phi; \mathbf{Q}_{R,L2})$, as defined in (14) and (16), respectively. In this study, we calculate the exact L1-PCs of $\overline{\mathbf{Y}}$, using the polynomial-cost optimal algorithm of Fig. 3. The receiver antenna-array is equipped with $D = 3$ elements and collects $N = 8$ snapshots. All snapshots contain a signal from the single source of interest ($K = 1$) impinging on the array with DoA $- 20°$. One out of the eight snapshots is corrupted

---

Algorithm 3: Cubic-cost L1-PCA of $\overline{\mathbf{Y}}$ (heuristic)

Input:		i) $\overline{\mathbf{Y}}_{2D \times 2N}$ and ii) K
	1:	If $K = 1$, $m \leftarrow 1$; else, $m \leftarrow 2K$
	2:	$\mathbf{B}_{2N \times m} \leftarrow \mathrm{bfk}(\overline{\mathbf{Y}}, \mathbf{B}^{(1)}, m)$
	3:	If $m = 1$, $\mathbf{B} \leftarrow [\mathbf{B}, \mathbf{E}_N \mathbf{B}]$
	4:	$(\mathbf{U}_{2D \times 2K}, \mathbf{\Sigma}_{2K \times 2K}, \mathbf{V}_{2K \times 2K}) \leftarrow \mathrm{svd}(\overline{\mathbf{Y}}\mathbf{B})$
Output:		$\mathbf{Q}_{R,BF} \leftarrow \mathbf{U}\mathbf{V}^\top$ (bit-flipping derived estimate to $\mathbf{Q}_{R,L1}$)

Function $\mathrm{bfK}(\overline{\mathbf{Y}}_{2D \times 2N}, \mathbf{B}_{2N \times m}, m \leq \min\{2N, 2D\})$

1:	$\omega \leftarrow \|\overline{\mathbf{Y}}\mathbf{B}\|_*$
2:	$\mathcal{L} \leftarrow \{1, 2, \ldots, 2Nm\}$
3:	while true (or terminate at $2Nm$ bit-flippings)
4:	for $x \in \mathcal{L}$, $l \leftarrow \lceil \frac{x}{2N} \rceil$, $v \leftarrow x - 2N(l-1)$
5:	$[\sim, \mathbf{\Sigma}, \sim] \leftarrow \mathrm{svd}(\overline{\mathbf{Y}}(\mathbf{B} - 2[\mathbf{B}]_{v,l}\mathbf{e}_{v,2N}\mathbf{e}_{m,l}^\top))$, $a_{v,l} \leftarrow \|\mathbf{\Sigma}\|_1$
6:	$(y, k) \leftarrow \arg\max_{\{(v,l):\, 2N(l-1)+v \in \mathcal{L}\}} a_{v,l}$,
7:	if $a_{y,k} > \omega$,
8:	$[\mathbf{B}]_{y,k} \leftarrow -[\mathbf{B}]_{y,k}$, $\omega \leftarrow a_{y,k}$, $\mathcal{L} \leftarrow \mathcal{L} \setminus \{2N(l-1) + v\}$
9:	elseif $\omega \geq a_{y,k}$ and $|\mathcal{L}| < 2Nm$, $\mathcal{L} \leftarrow \{1, 2, \ldots, 2Nm\}$
10:	else, break
11:	Return $\mathbf{B}$

Fig. 4 Algorithm for estimation of the $2K$ L1-PCs of rank-$2D$ data matrix $\overline{\mathbf{Y}}_{2D \times 2N}$ with cubic (w.r.t. *N*) asymptotic complexity $\mathcal{O}\left(N^3 m^4\right)$ ($m = 1$ for $K = 1$; $m = 2K$, for $K > 1$)

by two jamming sources with DoAs 31° and 54°. The signal-to-noise ratio (SNR) is set to 2 dB for the target source and to 5 dB for each of the jammers. We observe that standard MUSIC (L2-PCA) is clearly misled by the two jammer-corrupted measurement. Interestingly, the

proposed L1-PCA-based method manages to identify the target location successfully.

Next, we generalize our study to include probabilistic presence of an jammer. Specifically, we keep $D = 3$ and $N = 8$ and consider $K = 1$ target at $\theta = -41°$ with
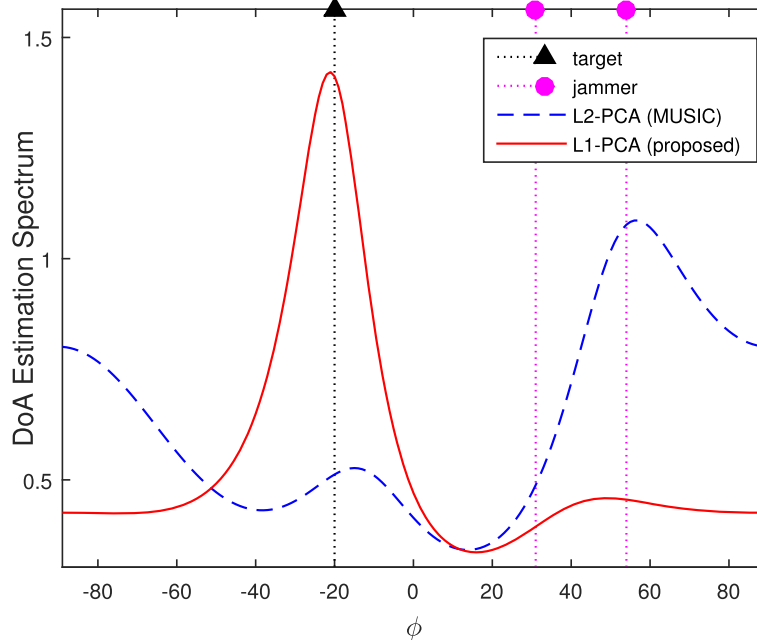

Fig. 5 DoA-estimation spectra $P_R(\phi; \mathbf{Q}_{R,L2})$ (MUSIC) and $P_R(\phi; \mathbf{Q}_{R,L2})$ (proposed); one target and two jamming signals with angles of arrival marked by ▲ and ●, ●, respectively

SNR 2 dB, and $L = 1$ jammer at $\theta' = 24°$ with activation probability p taking values in $\{0, .1, .2, .3, .4, .5\}$.

In Fig. 6, we plot the root-mean-square-error (RMSE)[4], calculated over 5000 independent realizations, vs. jammer SNR, for three DoA estimators: (a) the standard L2-PCA-based one (MUSIC), (b) the proposed L1-PCA DoA estimator with the L1-PCs calculated optimally by means of the polynomial-cost algorithm of Fig. 3, and (c) the proposed L1-PCA estimator with the L1-PCs found by means of the algorithm of Fig. 4. For all three methods, we plot the performance attained for each value of $p \in \{0, .1, .2, .3, .4, .5\}$. Our first observation is that the two L1-PCA-based estimators exhibit almost identical performance for every value of p and jammer SNR. Then, we notice that, in normal system operation ($p = 0$) the RMSEs of the L2-PCA-based and L1-PCA-based estimators are extremely close to each other and low, with slight (almost negligible) superiority of the L2-PCA-based method. Quite interestingly, for any non-zero jammer activation probability p and over the entire range of jammer SNR values, the RMSE attained by the proposed L1-PCA-based methods is lower than that attained by the L2-PCA-based one. For instance, for jammer SNR 12 dB and $p = .1$, the proposed methods offer 8° smaller RMSE than MUSIC. Of course, at high jammer SNR values and $p = .5$ the RMSE of both methods approaches 65°, which is the angular distance of the target and the jammer; i.e. both methods tend to peak the significantly (18 dB) stronger jammer present in half the snapshots.

In Fig. 7, we change the metric and study the more general Subspace Representation Ratio (SRR), attained by L2-PCA and L1-PCA. For any orthonormal basis $\mathbf{Q} \in \mathbb{R}^{2D \times 2K}$, SRR is defined as

$$\text{SRR}(\mathbf{Q}) \triangleq \frac{\left\| \mathbf{Q}^\top \bar{\mathbf{s}}(\theta) \right\|_2^2}{\left\| \mathbf{Q}^\top \bar{\mathbf{s}}(\theta') \right\|_2^2 + \left\| \mathbf{Q}^\top \bar{\mathbf{s}}(\theta) \right\|_2^2}. \qquad (37)$$

In Fig. 7, we plot $\text{SRR}(\mathbf{Q}_{R,L2})$ (L2-PCA), $\text{SRR}(\mathbf{Q}_{R,L1})$ (optimal L1-PCA), and $\text{SRR}(\mathbf{Q}_{R,BF})$ averaged over 5000 realizations, for multiple values of p, versus the jammer SNR. We observe that, again, the performance of the optimal and heuristic L1-PCA calculators almost coincides for every value of p and jammer SNR. Also, we notice that under normal system operation ($p = 0$) the spans of $\mathbf{Q}_{R,L2}$, $\mathbf{Q}_{R,L1}$, and $\mathbf{Q}_{R,BF}$ are equally good approximations to $\mathcal{S}_R$ and their respective SRR curves lie close (as close as the target SNR and the number of snapshots allow) to the benchmark of $\text{SRR}(\mathbf{U})$, where $\mathbf{U}$ is an orthonormal basis for the exact $\mathcal{S}_R$. On the other hand, when half the snapshots are jammer corrupted ($p = .5$) both methods capture more of the interference. Similar to Fig. 6, for any jammer activation probability and over the entire range of jammer SNR values, the SRR attained by L1-PCA (both algorithms) is superior to that attained by conventional L2-PCA.

Next, we set $D = 4, N = 10, \theta = -20°$ and $\theta' = 50°$. The source SNR is set to 5 dB. In Fig. 8, we plot the RMSE vs. jamming SNR performance attained by L2-PCA, RPCA[5] (algorithm of [29]), and L1-PCA (proposed –computed by efficient Algorithm 3), all computed on the realified snapshots. We observe that for $p = 0$ (i.e., no jamming corruption) all methods perform well; in particular, L2-PCA and L1-PCA demonstrate almost identical performance of about 3° RMSE. For jammer operation probability $p > 0$, we observe that the proposed L1-PCA method outperforms clearly all counterparts, exhibiting
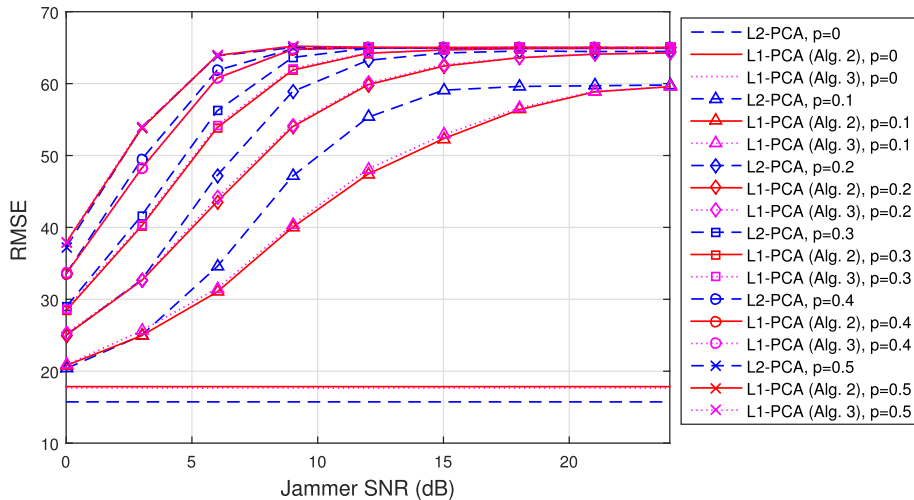

Fig. 6 Root-mean-squared-error (RMSE) vs. jammer SNR, for: L2-PCA (MUSIC), optimal L1-PCA, calculated by means of Algorithm 2 in Fig. 3, and L1-PCA by means of Algorithm 3 in Fig. 4. For each estimator, we present the RMSE curves for $p = 0, .1, .2, .3, .4, .5. N = 8, D = 3, \theta = -41°, \theta' = 24°$, source SNR 2 dB
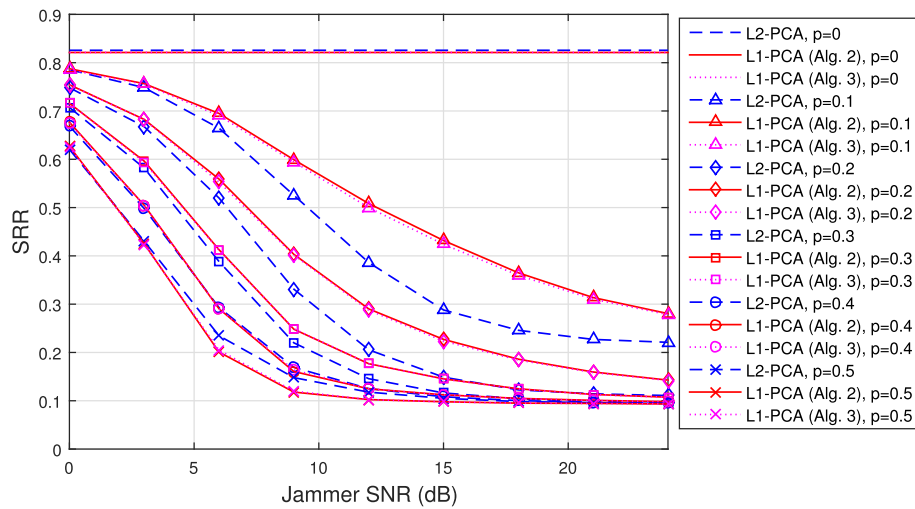
Fig. 7 Average SRR vs. jammer SNR, for: L2-PCA (MUSIC), optimal L1-PCA, calculated by means of Algorithm 2 in Fig. 3, and L1-PCA by means of Algorithm 3 in Fig. 4. For each estimator, we present the RMSE curves for $p = 0, .1, .2, .3, .4, .5$. $N = 8, D = 3, \theta = -41°, \theta' = 24°$, source SNR 2 dB

from 5° (for jammer SNR 6 dB) to 20° (for jammer SNR 11 dB) lower RMSE.

In Fig. 9. we plot the RMSE attained by the three counterparts, this time fixing jamming SNR to 10 dB and varying the snapshot corruption probability $p \in \{0, .1, .2, .3, .4, .5, .6\}$. Once again, we observe that, for $p = 0$ (no jamming activity), all methods perform well. For $p > 0$, L1-PCA outperforms both counterparts across the board.

Finally, in the study of Fig. 9, we measure the computation time expended by the three PCA methods, for $p = 0$ and $p = 0.5$. We observe that standard PCA, implemented by SVD, is the fastest method, with average computation time about $4 \cdot 10^{-5}$ s, for both values of p. The computation time of RPCA is $1.5 \cdot 10^{-2}$ s for $p = 0$ and $1.9 \cdot 10^{-2}$ s for $p = 0.5$. L1-PCA (Algorithm 3) computation takes, on average, $4.3 \cdot 10^{-2}$ s for both values of p, comparable to RPCA.[6]
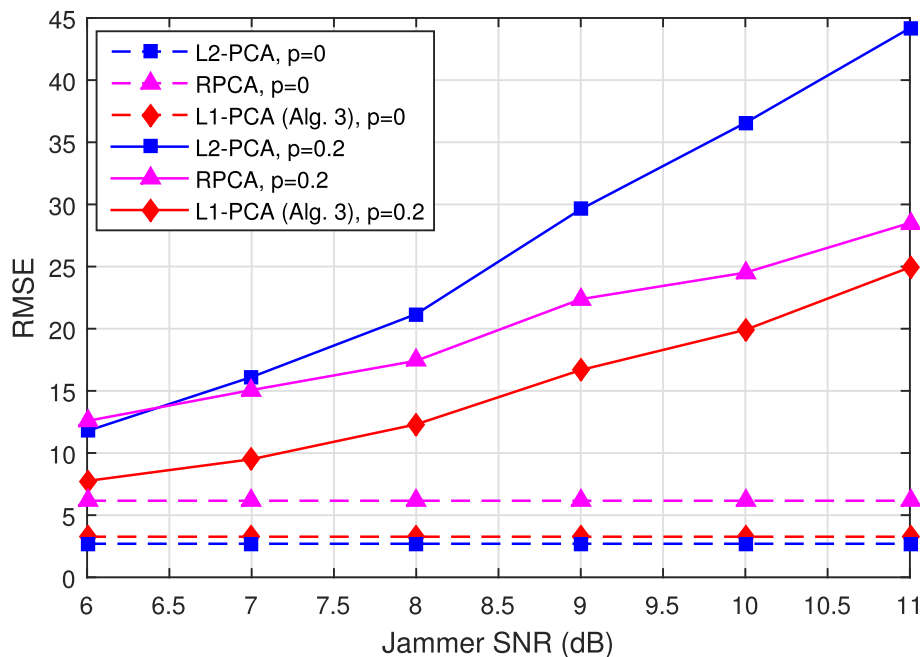

Fig. 8 RMSE vs. jammer SNR, for: L2-PCA (MUSIC), RPCA [29], and L1-PCA by means of Algorithm 3 in Fig. (4). For each estimator, we present the RMSE curves for $p = 0, .2$. $N = 10, D = 4, \theta = -20°, \theta' = 50°$, source SNR 5 dB
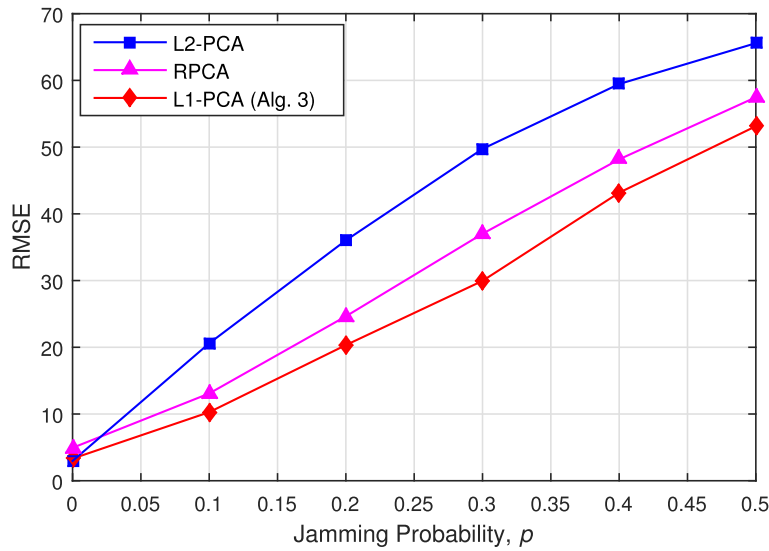
Fig. 9 RMSE vs. jammer operation probability p, for: L2-PCA (MUSIC), RPCA [29], and L1-PCA by means of Algorithm 3 in Fig. 4. $N = 10, D = 4, \theta = -20°, \theta' = 50°$, source SNR 5dB, jammer SNR 10 dB

5 Conclusions

We considered the problem of DoA estimation in the possible presence of unexpected, intermittent directional interference and presented a new method that relies on the L1-PCA of the recorded snapshots. Accordingly, we presented three algorithms (two optimal ones and one iterative/heuristic) for realified L1-PCA; i.e., L1-PCA of realified complex data matrices. Our numerical studies showed that the proposed method attains performance similar to conventional L2-PCA-based DoA estimation (MUSIC) in normal system operation (absence of jammers), while it attains significantly superior performance in the case of unexpected, sporadic corruption of the snapshots.

Endnotes

[1] $\mathbf{S}_\Phi \in \mathbb{C}^{D \times m}$ is a transposed Vandermonde matrix [66] and has rank m if $|\Phi| = m < D$.

[2] For any underlying set of elements $\mathcal{A}$ of cardinality n, a size m multiset may be defined as a pair $(\mathcal{A}, f)$ where $f : \mathcal{A} \to \mathbb{N}_{\geq 1}$ is a function from $\mathcal{A}$ to $\mathbb{N}_{\geq 1}$ such that $\sum_{a \in \mathcal{A}} f(a) = m$; the number of all distinct size m multisets defined upon $\mathcal{A}$ is $\binom{n+m-1}{m}$ [67].

[3] The (n, k)th bit of the binary matrix argument in (21) has the corresponding single-integer index $(k-1)2N+n \in \{1, 2, \ldots, 2Nm\}$.

[4] Denoting by θ the DoA and $\hat{\theta}$ the DoA estimate, RMSE is defined as the square root of the average value of $|\hat{\theta} - \theta|^2$.

[5] Given $\overline{\mathbf{Y}}$, RPCA minimizes $\|\mathbf{L}\|_* + \lambda \|\mathbf{O}\|_1$, over $\mathbf{L}$ and $\mathbf{O}$, subject to $\overline{\mathbf{Y}} = \mathbf{L} + \mathbf{O}$ and $\lambda = \sqrt{\max\{2D, 2N\}}^{-1}$

[29]. Then, it returns the $2K$ L2-PCs of $\mathbf{L}$ (computed by SVD).

[6] Reported computation times are measured in MAT-LAB R2017a, run on a computer equipped with Intel(R) core(TM) i7-6700 processor 3.40 GHz and 32 GB RAM. The MATLAB code of L1-PCA (Algorithm 3) can be found at [68]. The MATLAB code for RPCA was provided at [69].

Appendix
Useful properties of realification

Lemma 1 below follows straightforwardly from the definition in (8).

Lemma 1 *For any* $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$, *it holds that* $\overline{(\mathbf{A} + \mathbf{B})} = \overline{\mathbf{A}} + \overline{\mathbf{B}}$. *For any* $\mathbf{A} \in \mathbb{C}^{m \times n}$ *and* $\mathbf{B} \in \mathbb{C}^{n \times q}$, *it holds that* $\overline{(\mathbf{AB})} = \overline{\mathbf{A}}\,\overline{\mathbf{B}}$ *and* $\overline{(\mathbf{A}^H)} = \overline{\mathbf{A}}^\top$.   ∎

Lemma 2 below was discussed in [70] and [20], in the form of problem 8.6.4. Here, we also provide a proof, for the sake of completeness.

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

Proof Consider a complex matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ of rank $k \leq \min\{m, n\}$ and its singular value decomposition $\mathbf{A} \overset{\text{SVD}}{=} \mathbf{U}_{m \times m} \Sigma_{m \times n} \mathbf{V}_{n \times n}^H$, where

$$\Sigma = \begin{bmatrix} \mathbf{diag}(\sigma) & \mathbf{0}_{k \times (n-k)} \\ \mathbf{0}_{(m-k) \times k} & \mathbf{0}_{(m-k) \times (n-k)} \end{bmatrix} \quad (38)$$

and $\boldsymbol{\sigma} \triangleq [\sigma_1, \sigma_2, \ldots, \sigma_k]^\top \in \mathbb{R}_+^k$ is the length k vector containing (in descending order) the positive singular values of $\mathbf{A}$. By Lemma 1,

$$\overline{\mathbf{A}} = \overline{\mathbf{U}}\ \overline{\boldsymbol{\Sigma}}\ \overline{\mathbf{V}}^\top \tag{39}$$

with $\overline{\mathbf{U}}^\top \overline{\mathbf{U}} = \overline{\mathbf{U}}\ \overline{\mathbf{U}}^\top = \mathbf{I}_{2m}$ and $\overline{\mathbf{V}}^\top \overline{\mathbf{V}} = \overline{\mathbf{V}}\ \overline{\mathbf{V}}^\top = \mathbf{I}_{2n}$. Define now, for every $a, b \in \mathbb{N}_{\geq 1}$, the $ab \times ab$ permutation matrix

$$\mathbf{Z}_{a,b} \triangleq \left[\mathbf{I}_a \otimes \mathbf{e}_1^b,\ \mathbf{I}_a \otimes \mathbf{e}_2^b,\ \ldots,\ \mathbf{I}_a \otimes \mathbf{e}_b^b \right]^\top \tag{40}$$

where $\mathbf{e}_i^b \triangleq [\mathbf{I}_b]_{:,i}$, for every $i \in \{1, 2, \ldots, b\}$. Then,

$$\mathbf{Z}_{a,b}^\top \mathbf{Z}_{a,b} = \sum_{i=1}^b \left(\mathbf{I}_a \otimes \left(\mathbf{e}_i^b \right)^\top \right)^\top \left(\mathbf{I}_a \otimes \left(\mathbf{e}_i^b \right)^\top \right)$$

$$= \mathbf{I}_a \otimes \left(\sum_{i=1}^b \mathbf{e}_i^b \left(\mathbf{e}_i^b \right)^\top \right) = \mathbf{I}_a \otimes \mathbf{I}_b = \mathbf{I}_{ab}. \tag{41}$$

By (39),

$$\overline{\mathbf{A}} = \overline{\mathbf{U}}\mathbf{Z}_{2,m}^\top \mathbf{Z}_{2,m}\overline{\boldsymbol{\Sigma}}\mathbf{Z}_{2,n}^\top \mathbf{Z}_{2,n}\overline{\mathbf{V}}^\top$$

$$= \left(\overline{\mathbf{U}}\mathbf{Z}_{2,m}^\top \right) \left(\mathbf{Z}_{2,m}(\mathbf{I}_2 \otimes \boldsymbol{\Sigma})\mathbf{Z}_{2,n}^\top \right) \left(\overline{\mathbf{V}}\mathbf{Z}_{2,n}^\top \right)^\top$$

$$= \check{\mathbf{U}}\check{\boldsymbol{\Sigma}}\check{\mathbf{V}}^\top \tag{42}$$

where $\check{\mathbf{U}} \triangleq \overline{\mathbf{U}}\mathbf{Z}_{2,m}^\top$, $\check{\mathbf{V}} \triangleq \overline{\mathbf{V}}\mathbf{Z}_{2,n}^\top$, and $\check{\boldsymbol{\Sigma}} \triangleq \mathbf{Z}_{2,m}(\mathbf{I}_2 \otimes \boldsymbol{\Sigma})\mathbf{Z}_{2,n}^\top$. It is easy to show that $\check{\mathbf{U}}^\top \check{\mathbf{U}} = \check{\mathbf{U}}\check{\mathbf{U}}^\top = \mathbf{I}_{2m}$, $\check{\mathbf{V}}^\top \check{\mathbf{V}} = \check{\mathbf{V}}\check{\mathbf{V}}^\top = \mathbf{I}_{2n}$, and $\check{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma} \otimes \mathbf{I}_2$. Therefore, (42) constitutes the standard (sorted singular values) SVD of $\overline{\mathbf{A}}$ and $\text{rank}(\overline{\mathbf{A}}) = 2k$. □

Lemma 3 below follows from Lemmas 1 and 2.

Lemma 3 *For any* $\mathbf{A} \in \mathbb{C}^{m \times n}$, $\|\mathbf{A}\|_2^2 = \frac{1}{2}\|\overline{\mathbf{A}}\|_2^2$. ∎

Proof of (14)
We commence our proof with the following auxiliary Lemma 4.

Lemma 4 *For any matrix* $\mathbf{A} \in \mathbb{C}^{m \times n}$, *if* $\mathbf{Q}_{real} \in \mathbb{R}^{2m \times 2l}$, $l < m \leq n$ *is a solution to*

$$\underset{\mathbf{Q} \in \mathbb{R}^{2m \times 2l},\ \mathbf{Q}^\top \mathbf{Q} = \mathbf{I}_{2l}}{\text{maximize}} \|\overline{\mathbf{A}}^\top \mathbf{Q}\|_2. \tag{43}$$

and $\mathbf{Q}_{comp.} \in \mathbb{C}^{m \times l}$ *is a solution to*

$$\underset{\mathbf{Q} \in \mathbb{C}^{m \times l},\ \mathbf{Q}^H \mathbf{Q} = \mathbf{I}_l}{\text{maximize}} \|\mathbf{A}^H \mathbf{Q}\|_2, \tag{44}$$

then $\mathbf{Q}_{real}\mathbf{Q}_{real}^\top = \overline{\left(\mathbf{Q}_{comp.}\mathbf{Q}_{comp.}^H\right)} = \overline{\mathbf{Q}}_{comp.}\overline{\mathbf{Q}}_{comp.}^\top$. ∎

Proof Orthonormal basis $[\check{\mathbf{U}}]_{:,1:2l}$, defined in the proof of Lemma 2 above, contains the $2l$ highest-singular-value

left-singular vectors of $\overline{\mathbf{A}}$, and, thus, solves (43) [20]. Since the objective value in (43) is invariant to column permutations of the argument $\mathbf{Q}$, any column permutation of $[\check{\mathbf{U}}]_{:,1:2l}$ is still a solution to (43). Next, we define the permutation matrix $\mathbf{W}_{m,l} \triangleq [\mathbf{I}_l,\ \mathbf{0}_{l \times (m-l)}]^\top$ and notice that $[\check{\mathbf{U}}]_{:,1:2l} = \overline{\mathbf{U}}[\mathbf{I}_2 \otimes \mathbf{e}_1^m, \mathbf{I}_2 \otimes \mathbf{e}_2^m, \ldots, \mathbf{I}_2 \otimes \mathbf{e}_l^m]$ is a column permutation of $\overline{\mathbf{U}}(\mathbf{I}_2 \otimes \mathbf{W}_{m,l}) = \overline{\mathbf{U}}\ \overline{\mathbf{W}}_{m,l}$, which, by Lemma 1, equals $\overline{(\mathbf{UW}_{m,l})} = \overline{[\mathbf{U}]_{:,1:l}}$. Thus, $\overline{[\mathbf{U}]_{:,1:l}}$ solves (43) too. At the same time, by (39), $[\mathbf{U}]_{:,1:l}$ contains the l highest-singular-value left-singular vectors of $\mathbf{A}$ and solves (44) [20]. By the above, we conclude that a realification per (8) of any solution to (44) constitutes a solution to (43) and, thus, $\mathbf{Q}_{real}\mathbf{Q}_{real}^\top = \overline{\left(\mathbf{Q}_{comp.}\mathbf{Q}_{comp.}^H\right)} = \overline{\mathbf{Q}}_{comp.}\overline{\mathbf{Q}}_{comp.}^\top$. □

By Lemmas 1, 3, and 4, (14) holds true.

Proof of (29)
We commence our proof by defining $d = |\mathcal{X}_1|$ and the sets $\mathcal{X}_A^c \triangleq \{1, 2, \ldots, 2N\} \setminus \mathcal{X}_A$ (e.a.o.) and $\mathcal{X}_B^c \triangleq \{1, 2, \ldots, 2N\} \setminus \mathcal{X}_B$ (e.a.o.) Then, we notice that

$$[\mathbf{I}_{2N}]_{:,\mathcal{X}_B} = \mathbf{E}_N [\mathbf{I}_{2N}]_{:,\mathcal{X}_A}\ \mathbf{P}, \tag{45}$$

where $\mathbf{P} \triangleq \left[-[\mathbf{I}_{2D-1}]_{:,d+1:2D-1},\ [\mathbf{I}_{2D-1}]_{:,1:d} \right]$. Similarly,

$$[\mathbf{I}_{2N}]_{:,\mathcal{X}_B^c} = \mathbf{E}_N [\mathbf{I}_{2N}]_{:,\mathcal{X}_A^c}\ \mathbf{P}_c, \tag{46}$$

where $\mathbf{P}_c \triangleq \left[-[\mathbf{I}_{2N-2D+1}]_{:,N-d+1:2N-2D+1},\ [\mathbf{I}_{2N-2D+1}]_{:,1:N-d} \right]$. Then,

$$[\overline{\mathbf{Y}}]_{:,\mathcal{X}_B} = \overline{\mathbf{Y}}[\mathbf{I}_{2N}]_{:,\mathcal{X}_B} = \overline{\mathbf{Y}}\mathbf{E}_N [\mathbf{I}_{2N}]_{:,\mathcal{X}_A}\ \mathbf{P}$$

$$= \mathbf{E}_D \overline{\mathbf{Y}}[\mathbf{I}_{2N}]_{:,\mathcal{X}_A}\ \mathbf{P} = \mathbf{E}_D [\overline{\mathbf{Y}}]_{:,\mathcal{X}_A}\ \mathbf{P}. \tag{47}$$

Consider now $\mathbf{z} = \text{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{E}_D \mathbf{c}(\mathcal{X}_A)$. It holds that $[\mathbf{z}]_{2D} > 0$ and

$$[\overline{\mathbf{Y}}]_{:,\mathcal{X}_B}^T \mathbf{z} = \text{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}^\top [\overline{\mathbf{Y}}]_{:,\mathcal{X}_A}^\top \mathbf{E}_D^\top \mathbf{E}_D \mathbf{c}(\mathcal{X}_A)$$

$$= \text{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}^\top [\overline{\mathbf{Y}}]_{:,\mathcal{X}_A}^\top \mathbf{c}(\mathcal{X}_A) = \mathbf{0}_{2D}. \tag{48}$$

Therefore, $\mathbf{z} = \mathbf{c}(\mathcal{B}) = \in \text{null}\left([\overline{\mathbf{Y}}]_{:,\mathcal{X}_B}^\top\right) \cap \Omega_{2D}$ and, hence, (29) holds true.

Proof of Prop. 3
We begin by rewriting the maximization argument of (20) as

$$\|\overline{\mathbf{Y}}\mathbf{B}\|_*^2 = \|\overline{\mathbf{Y}}\mathbf{B}\|_2^2 + 2\sqrt{\det\left(\mathbf{B}^\top \overline{\mathbf{Y}}^\top \overline{\mathbf{Y}}\mathbf{B}\right)}$$

$$= \|\overline{\mathbf{Y}}\mathbf{b}_1\|_2^2 + \|\overline{\mathbf{Y}}\mathbf{b}_2\|_2^2$$

$$+ 2\sqrt{\|\overline{\mathbf{Y}}\mathbf{b}_1\|_2^2 \|\overline{\mathbf{Y}}\mathbf{b}_2\|_2^2 - \left(\mathbf{b}_1^\top \overline{\mathbf{Y}}^\top \overline{\mathbf{Y}}\mathbf{b}_2\right)^2}, \tag{49}$$

where $\mathbf{b}_1$ and $\mathbf{b}_2$ are the first and second columns of $\mathbf{B}$, respectively. Evidently, the maximum value attained at (17) is upper bounded as

$$
\begin{aligned}
\max_{\mathbf{B}\in\{\pm 1\}^{2N\times 2}} \|\overline{\mathbf{Y}}\mathbf{B}\|_*^2 \leq & \max_{\mathbf{b}_1\in\{\pm 1\}^{2N},\mathbf{b}_2\in\{\pm 1\}^{2N}} \|\overline{\mathbf{Y}}\mathbf{b}_1\|_2^2 \\
& + \|\overline{\mathbf{Y}}\mathbf{b}_1\|_2^2 + 2\|\overline{\mathbf{Y}}\mathbf{b}_1\|_2\|\overline{\mathbf{Y}}\mathbf{b}_2\|_2 \\
& = 4 \max_{\mathbf{b}\in\{\pm 1\}^{2N}} \|\overline{\mathbf{Y}}\mathbf{b}\|_2^2. \quad (50)
\end{aligned}
$$

Considering now a solution $\mathbf{b}_{\mathrm{opt}}$ to maximize$_{\mathbf{b}\in\{\pm 1\}^{2N\times 1}}$ $\|\overline{\mathbf{Y}}\mathbf{b}\|_2^2$, and defining $\mathbf{b}'_{\mathrm{opt}} = \mathbf{E}_N\mathbf{b}_{\mathrm{opt}}$, we notice that $\|\overline{\mathbf{Y}}\mathbf{b}'_{\mathrm{opt}}\|_2^2 = \|\overline{\mathbf{Y}}\mathbf{E}_N\mathbf{b}_{\mathrm{opt}}\|_2^2 = \|\mathbf{E}_D\overline{\mathbf{Y}}\mathbf{b}_{\mathrm{opt}}\|_2^2 = \|\overline{\mathbf{Y}}\mathbf{b}_{\mathrm{opt}}\|_2^2$ and $\mathbf{b}_{\mathrm{opt}}^\top\overline{\mathbf{Y}}^\top\overline{\mathbf{Y}}\mathbf{b}'_{\mathrm{opt}} = \mathbf{b}_{\mathrm{opt}}^\top\overline{\mathbf{Y}}^\top\overline{\mathbf{Y}}\mathbf{E}_N\mathbf{b}_{\mathrm{opt}} = \mathbf{b}_{\mathrm{opt}}^\top\overline{\mathbf{Y}}^\top\mathbf{E}_D\overline{\mathbf{Y}}\mathbf{b}_{\mathrm{opt}} = 0$. Therefore, $\|\overline{\mathbf{Y}}\left[\mathbf{b}_{\mathrm{opt}},\mathbf{b}'_{\mathrm{opt}}\right]\|_*^2 = 4 \|\overline{\mathbf{Y}}\mathbf{b}_{\mathrm{opt}}\|_2^2$ and, in view of (50), $\left[\mathbf{b}_{\mathrm{opt}}, \mathbf{E}_N\mathbf{b}_{\mathrm{opt}}\right]$ is a solution to (20).

Proof of (30)

By (29) and (46), it holds that

$$
\begin{aligned}
[\overline{\mathbf{Y}}]_{:,\mathcal{X}_B^c}^\top \mathbf{c}(\mathcal{X}_B) &= \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}_c^\top[\overline{\mathbf{Y}}]_{:,\mathcal{X}_A^c}^\top \mathbf{E}_D^\top\mathbf{E}_D\mathbf{c}(\mathcal{X}_A) \\
&= \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}_c^\top[\overline{\mathbf{Y}}]_{:,\mathcal{X}_A^c}^\top \mathbf{c}(\mathcal{X}_A). \quad (51)
\end{aligned}
$$

Consider now some $\mathbf{b} \in \mathcal{B}(\mathcal{X}_A)$ and define $\mathbf{b}' \overset{\triangle}{=} \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{E}_N\mathbf{b}$. By (46), (51), and the definition in (25), it holds that

$$
\begin{aligned}
[\mathbf{b}']_{\mathcal{X}_B^c} &= [\mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{E}_N\mathbf{b}]_{\mathcal{X}_B^c} = \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)[\mathbf{I}_{2N}]_{:,\mathcal{X}_B^c}^\top \mathbf{E}_N\mathbf{b} \\
&= \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)(\mathbf{E}_N[\mathbf{I}_{2N}]_{:,\mathcal{X}_A}\mathbf{P}_c)^\top\mathbf{E}_N\mathbf{b} \\
&= \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}_c^\top[\mathbf{I}_{2N}]_{:,\mathcal{X}_A}^\top \mathbf{b} = \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}_c^\top[\mathbf{b}]_{:,\mathcal{X}_A} \\
&= \mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}_c^\top\mathrm{sgn}([\overline{\mathbf{Y}}]_{:,\mathcal{X}_A}^\top \mathbf{c}(\mathcal{A})) \\
&= \mathrm{sgn}(\mathrm{sgn}([\mathbf{c}(\mathcal{X}_A)]_D)\mathbf{P}_c^\top[\overline{\mathbf{Y}}]_{:,\mathcal{X}_A}^\top \mathbf{c}(\mathcal{A})) \\
&= \mathrm{sgn}([\overline{\mathbf{Y}}]_{:,\mathcal{X}_B^c}^\top \mathbf{c}(\mathcal{X}_B)).
\end{aligned}
$$

$$(52)$$

Hence, $\mathbf{b}'$ belongs to $\mathcal{B}(\mathcal{X}_B)$ and (30) holds true.

Abbreviations
AWGN: Additive white Gaussian noise; DoA: Direction-of-arrival; EVD: Eigenvalue decomposition; L1-PCA: L1-norm principal-component analysis; L1-PCs: L1-norm principal components; L2-PCs: L2-norm principal components; MUSIC: Multiple signal classification; PCA: Principal-component analysis; RMSE: Root-mean-squared-error; RPCA: Robust PCA; SNR: Signal-to-noise ratio; SRR: Subspace-representation ratio; SVD: Singular-value decomposition; ULA: Uniform linear array; UNM: Unimodular nuclear-norm maximization

Authors' contributions
All the authors have contributed to the presented theoretical, algorithmic, and numerical results, as well as the drafting of the manuscript. All the authors read and approved the final manuscript.

Authors' information
Panos P. Markopoulos (pxmeee@rit.edu) is an Assistant Professor with the Dept. of Electrical and Microelectronic Engineering, Rochester Institute of Technology, Rochester, NY, USA. Nicholas Tsagkarakis (nikolaos.tsagkarakis@ericsson.com) is a System Engineer with Ericsson, Gothenburg, Sweden. Dimitris A. Pados (dpados@fau.edu) is an I-SENSE Fellow and Professor with the Dept. of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA. George N. Karystinos (karystinos@telecom.tuc.gr) is a Professor with the School of Electrical and Computer Engineering, Technical University of Crete, Chania, Greece.

Availability of data and materials
MATLAB scripts for L1-PCA of real data can be found here [68]. All numerical studies in this paper involve synthetic data, generated in MATLAB exactly as presented in Section 4; thus, no other datasets are made available.

Ethics approval and consent to participate
Not applicable. The manuscript does report any studies involving human participants, human data, or human tissue.

Consent for publication
Not applicable. The manuscript does not contain any individual person's data.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Electrical and Microelectronic Engineering, Rochester Institute of Technology, 14623 Rochester, NY, USA. [2]Ericsson AB, 417 56 Göteborg, Sweden. [3]I-SENSE and Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, 33431 USA. [4]School of Electrical and Computer Engineering, Technical University of Crete, 73100 Chania, Crete, Greece.

References
1. S. A. Zekavat, M. Buehrer, *Handbook of Position Location: Theory, Practice and Advances*. (Wiley-IEEE Press, New York, 2011)
2. W.-J. Zeng, X.-L. Li, High-resolution multiple wideband and non-stationary source localization with unknown number of sources. IEEE Trans. Signal Process. **58**, 3125–3136 (2010)
3. M. G. Amin, W. Sun, A novel interference suppression scheme for global navigation satellite systems using antenna array. IEEE J. Select Areas Commun. **23**, 999–1012 (2005)
4. S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, Z. Sahinoglu, Localization via ultra-wideband radios. IEEE Signal Process. Mag. **22**, 70–84 (2005)
5. L. C. Godara, Application of antenna arrays to mobile communications, part ii: Beam-forming and direction-of-arrival considerations. Proc. IEEE. **85**, 1195–1245 (1997)
6. H. Krim, M. Viberg, Two decades of array signal processing research. IEEE Signal Process. Mag., 67–94 (1996)
7. J. M. Kantor, C. D. Richmond, D. W. Bliss, B. C. Jr., Mean-squared-error prediction for bayesian direction-of-arrival estimation. IEEE Trans. Signal Process. **61**, 4729–4739 (2013)
8. P. Stoica, A. B. Gershman, Maximum-likelihood doa estimation by data-supported grid search. IEEE Signal Process. Lett. **6**, 273–275 (1999)
9. P. Stoica, K. C. Sharman, Maximum likelihood method for direction of arrival estimation. IEEE Trans. Acoust. Speech Signal Process., 1132–1143 (1990)

10. J. Sheinvald, M. Wax, A. J. Weiss, On maximum-likelihood localization of coherent signals. IEEE Trans. Signal Process. **44**, 2475–2482 (1996)

11. B. Ottersten, M. Viberg, P. Stoica, A. Nehorai, in *Radar Array Processing*, ed. by S. Haykin, J. Litva, and T. J. Shepherd. Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing (Springer, Berlin, 1993), pp. 99–151

12. M. I. Miller, D. R. Fuhrmann, Maximum likelihood narrow-band direction finding and the em algorithm. IEEE Trans. Acoust., Speech, Signal Process. **38**, 1560–1577 (1990)

13. F. C. Schweppe, Sensor array data processing for multiple-signal sources. IEEE Trans. Inf. Theory. **14**, 294–305 (1968)

14. D. H. Johnson, The application of spectral estimation methods to bearing estimation problems. Proc. IEEE. **70**, 1018–1028 (1982)

15. R. T. Lacoss, Data adaptive spectral analysis method. Geophysics. **36**, 661–675 (1971)

16. R. O. Schmidt, Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**, 276–280 (1986)

17. S. Haykin, *Advances in Spectrum Analysis and Array Processing*. (Prentice-Hall, Englewood Cliffs, 1995)

18. H. L. V. Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*. (Wiley, New York, 2002)

19. R. Grover, D. A. Pados, M. J. Medley, Subspace direction finding with an auxiliary-vector basis. IEEE Trans. Signal Process. **55**, 758–763 (2007)

20. G. H. Golub, C. F. V. Loan, *Matrix Computations*, 4th edn. (The Johns Hopkins Univ. Press, Baltimore, 2012)

21. P. Stoica, A. Nehorai, Music, maximum likelihood, and cramer-rao bound. IEEE Trans. Acoust. Speech Signal Process. **37**, 720–741 (1989)

22. P. Stoica, A. Nehorai, Music, maximum likelihood, and cramer-rao bound: further results and comparisons. IEEE Trans. Acoust. Speech Signal Process. **38**, 2140–2150 (1990)

23. H. Abeida, J.-P. Delmas, Efficiency of subspace-based doa estimators. Signal Process. (Elsevier). **87**, 2075–2084 (2007)

24. K. L. Blackard, T. S. Rappaport, C. W. Bostian, Measurements and models of radio frequency impulsive noise for indoor wireless communications. IEEE J. Select Areas Commun. **11**, 991–1001 (1993)

25. J. B. Billingsley, Ground clutter measurements for surface-sited radar. Massachusetts Inst. Technol. Cambridge, MA, Tech. Rep. 780 (1993). https://apps.dtic.mil/docs/citations/ADA262472

26. F. Pascal, P. Forster, J. P. Ovarlez, P. Larzabal, Performance analysis of covariance matrix estimates in impulsive noise. IEEE Trans. Signal Process. **56**, 2206–2217 (2008)

27. R. L. Peterson, R. E. Ziemer, D. E. Borth, *Introduction to Spread Spectrum Communications*. (Prentice Hall, Englewood Cliffs, 1995)

28. O. Besson, P. Stoica, Y. Kamiya, Direction finding in the presence of an intermittent interference. IEEE Trans. Signal Process. **50**, 1554–1564 (2002)

29. E. J. Candes, X. Li, Y. Ma, J. Wright, Robust principal component analysis?. J. ACM. **58**(11), 1–37 (2011)

30. F. De la Torre, M. J. Black, in *Proc. Int. Conf. Computer Vision (ICCV)*. Robust principal component analysis for computer vision, (Vancouver, 2001), pp. 1–8

31. Q. Ke, T. Kanade, Robust subspace computation using L1 norm, . Internal Technical Report, Computer Science Department, Carnegie Mellon University, CMU-CS-03–172 (2003). http://www-2.cs.cmu.edu/~ke/publications/CMU-CS-03-172.pdf

32. Q. Ke, T. Kanade, in *Proc. IEEE Conf. Comput. Vision Pattern Recog. (CVPR)*. Robust l_1 norm factorization in the presence of outliers and missing data by alternative convex programming (SPIE, San Diego, 2005), pp. 739–746

33. L. Yu, M. Zhang, C. Ding, in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*. An efficient algorithm for l_1-norm principal component analysis (IEEE, Kyoto, 2012), pp. 1377–1380

34. J. P. Brooks, J. H. Dulá, The l_1-norm best-fit hyperplane problem. Appl. Math. Lett. **26**, 51–55 (2013)

35. J. P. Brooks, J. H. Dulá, E. L. Boone, A pure l_1-norm principal component analysis. J. Comput. Stat. Data Anal. **61**, 83–98 (2013)

36. N. Kwak, Principal component analysis based on l_1-norm maximization. IEEE Trans. Pattern Anal. Mach. Intell. **30**, 1672–1680 (2008)

37. F. Nie, H. Huang, C. Ding, D. Luo, H. Wang, in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*. Robust principal component analysis with non-greedy l_1-norm maximization (IJCAI, 2011), pp. 1433–1438

38. M. McCoy, J. A. Tropp, Electron JS. Two proposals for robust PCA using semidefinite programming. **5**, 1123–1160 (2011)

39. C. Ding, D. Zhou, X. He, H. Zha, in *Proc. Int. Conf. Mach. Learn. (ICML)*. r_1-PCA: Rotational invariant l_1-norm principal component analysis for robust subspace factorization (Proc. ICML, Pittsburgh, 2006), pp. 281–288

40. X. Li, Y. Pang, Y. Yuan, l_1-norm-based 2DPCA. IEEE Trans. Syst. Man. Cybern. B. Cybern. **40**, 1170–1175 (2009)

41. Y. Liu, D. A. Pados, Compress-sensed-domain L1-PCA video surveillance. IEEE Trans. Mult. **18**, 351–363 (2016)

42. P. P. Markopoulos, D. A. Pados, G. N. Karystinos, M. Langberg, in *Proc. SPIE Compressive Sensing Conference, Defense and Commercial Sensing (SPIE DCS 2017)*. L1-norm principal-component analysis in l2-norm-reduced-rank data subspaces, (Anaheim, 2017), pp. 1021104–1102110410

43. P. P. Markopoulos, S. Kundu, S. Chamadia, D. A. Pados, Efficient L1-norm principal-component analysis via bit flipping. IEEE Trans. Signal Process. **62**, 4252–4264 (2017)

44. N. Tsagkarakis, P. P. Markopoulos, D. A. Pados, in *Proc. IEEE International Conference on Machine Learning and Applications (IEEE ICMLA 2016)*. On the L1-norm approximation of a matrix by another of lower rank (IEEE, Anaheim, 2016), pp. 768–773

45. P. P. Markopoulos, S. Kundu, S. Chamadia, D. A. Pados, in *Proc. IEEE International Conference on Machine Learning and Applications (IEEE ICMLA 2016)*. L1-norm principal-component analysis via bit flipping (IEEE, Anaheim, 2016), pp. 326–332

46. N. Tsagkarakis, P. P. Markopoulos, G. Sklivanitis, D. A. Pados, L1-norm principal-component analysis of complex data. IEEE Trans. Signal Process. **66**, 3256–3267 (2018)

47. P. P. Markopoulos, G. N. Karystinos, D. A. Pados, in *Proc. 10th Int. Symp. on Wireless Commun. Syst. (ISWCS)*. Some options for l_1-subspace signal processing (IEEE, Ilmenau, 2013), pp. 622–626

48. P. P. Markopoulos, G. N. Karystinos, D. A. Pados, Optimal algorithms for l_1-subspace signal processing. IEEE Trans. Signal Process. **62**, 5046–5058 (2014)

49. P. P. Markopoulos, S. Kundu, D. A. Pados, in *Proc. IEEE Int. Conf. Image Process. (ICIP)*. L1-fusion: Robust linear-time image recovery from few severely corrupted copies (IEEE, Quebec City, 2015), pp. 1225–1229

50. P. P. Markopoulos, F. Ahmad, in *Proc. IEEE Radar Conference (IEEE Radarcon 2017)*. Indoor human motion classification by L1-norm subspaces of micro-doppler signatures (IEEE, Seattle, 2017)

51. D. G. Chachlakis, P. P. Markopoulos, R. J. Muchhala, A. Savakis, in *Proc. SPIE Compressive Sensing Conference, Defense and Commercial Sensing (SPIE DCS 2017)*. Visual tracking with L1-grassmann manifold modeling (SPIE, Anaheim, 2017), pp. 1021102–1102110210

52. P. P. Markopoulos, F. Ahmad, in *Int. Microwave Biomed. Conf. (IEEE IMBioC 2018), IEEE*. Robust radar-based human motion recognition with L1-norm linear discriminant analysis (IEEE, Philadelphia, 2018), pp. 145–147

53. A. Gannon, G. Sklivanitis, P. P. Markopoulos, D. A. Pados, S. N. Batalama, *Semi-blind signal recovery in impulsive noise with L1-norm PCA* (IEEE, Pacific Grove, 2018). to appear

54. X. Ding, L. He, L. Carin, Bayesian robust principal component analysis. IEEE Trans. Image Process. **20**(12), 3419–3430 (2011)

55. J. Wright, A. Ganesh, S. Rao, Y. Peng, Y. Ma, in *Adv. Neural Info. Process. Syst. (NIPS)*. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization, (2009), pp. 2080–2088

56. P. P. Markopoulos, N. Tsagkarakis, D. A. Pados, G. N. Karystinos, in *Proc. IEEE Phased Array Systems and Technology (PAST 2016)*. Direction-of-arrival estimation by L1-norm principal components, (Waltham, 2016), pp. 1–6

57. A. Graham, *Kronecker Products and Matrix Calculus with Applications*. (Ellis Horwood, Chichester, 1981)

58. M. Haardt, J. Nossek, Unitary esprit: How to obtain increased estimation accuracy with a reduced computational burden. IEEE Trans. Signal Process. **43**, 1232–1242 (1995)

59. M. Pesavento, A. B. Gershman, M. Haardt, Unitary root-music with a real-valued eigendecomposition: A theoretical and experimental performance study. IEEE Trans. Signal Process. **48**, 1306–1314 (2000)

60. S. A. Vorobyov, A. B. Gershman, Z.-Q. Luo, Robust adaptive beamforming using worst-case performance optimization. IEEE Trans. Signal Process. **51**, 313–324 (2003)

61. J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*. (Springer, New York, 1972)

62. A. W. Knapp, *Lie Groups Beyond an Introduction vol. 140*. (Birkhauser, Boston, 2002)
63. L. W. Ehrlich, Complex matrix inversion versus real. Commun. ACM. **13**, 561–562 (1970)
64. G. N. Karystinos, A. P. Liavas, Efficient computation of the binary vector that maximizes a rank-deficient quadratic form. IEEE Trans. Inf. Theory. **56**, 3581–3593 (2010)
65. E. S. Andersen, Two summation formulae for product sums of binomial coefficients. Math. Scand. **1**, 261–262 (1953)
66. C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. (SIAM, Philadelphia, 2001)
67. R. P. Stanley, *Enumerative Combinatorics vol. 1*, 2nd edn. (Cambridge University Press, New York, 2012)
68. P. P. Markopoulos, L1-PCA Toolbox. https://www.mathworks.com/matlabcentral/fileexchange/64855-l1-PCA-toolbox. Accessed 18 July 2019
69. D. Laptev, RobustPCA. https://github.com/dlaptev. Accessed 18 July 2019
70. D. Day, M. A. Heroux, Solving complex-valued linear systems via equivalent real formulations. SIAM J. Sci. Comput. **23**(2), 480–498 (2001)

Publisher's Note