Chapter 9
Source Localization and Tracking: A Sparsity-Exploiting Maximum a Posteriori Based Approach

Md Mashud Hyder and Kaushik Mahata

Abstract In this work, we explore the potential of sparse recovery algorithms for localization and tracking the direction-of-arrivals (DOA) of multiple targets using a limited number of noisy time samples collected from a small number of sensors. In target tracking problems, the targets are assumed to be moving with a small random angular acceleration. We show that the target tracking problem can be posed as a problem of recursively reconstructing a sequence of sparse signals where the support of the signals changing slowly with time. Here, one can use the support of last signal as a priori information to estimate the behavior of current signal. In particular, we propose a maximum a posteriori (MAP)-based approach to deal with the sparse recovery problem arising in tracking and detection of DOAs. We consider both narrowband and broadband scenarios. Numerical simulations demonstrate the effectiveness of the proposed algorithm. We found that the proposed algorithm can resolve and track closely spaced DOAs with a small number of sensors.

9.1 Introduction

Direction of arrival (DOA) estimation using sensor array has been an active research area [20, 25, 37], playing an important role in smart antennas, next generation mobile communication systems, various type of imaging systems and target tracking applications. Many algorithms have been developed, see [37] and references therein. The algorithms like Capon [6], pose the DOA estimation problem as a beamforming

Md. M. Hyder (✉) · K. Mahata
The University of Newcastle, Callaghan, NSW 2308, Australia
e-mail: mashud_buet@yahoo.com; md.hyder@uon.edu.au

K. Mahata
e-mail: Kaushik.Mahata@newcastle.edu.au

problem. Here one designs adaptive filterbanks to obtain nonparametric estimate of the spatial spectrum. The popular alternative to this is the subspace algorithms like MUSIC [35], ESPRIT [33] or weighted subspace fitting [36, 41]. The subspace algorithms which exploit the low-rank structure of the noise-free signal. The maximum-likelihood estimation [25] is another efficient technique, but requires accurate initialization to ensure global convergence. All these methods rely on the statistical properties of the data, and hence requires a large number of time samples.

Conventional DOA estimation techniques cannot exploit the target moving statistics into their formulation and hence their performance degrade when a large number of DOAs moving in a field of interest [45]. Recently, several approaches have been developed for tracking targets [4, 8, 30, 31, 44, 45]. The maximum likelihood (ML) methods [30, 31, 45] have good statistical properties and are robust in target tracking with a relatively small number of samples. The works in [30, 31] incorporate the target motion dynamics in ML estimation and computes the the DOA parameters at each time subinterval and refine the ML estimates through Kalman filtering. The concept of Multiple Target States (MTS) has been introduced in [45] to describe the target motion. The DOA tracking is implemented through updating the MTS by maximizing the likelihood function of the array output. However, ML-based algorithms have high computational cost in general [8]. A recursive expectation and maximization (EM) [12] algorithm has been used in [8] to improve computational efficiency of ML algorithms. Cyclostationarity property of the moving targets has been exploited in [44]. In target tracking, the change in DOA from last time frame to the current time frame is computed by exploiting the difference of the averaged cyclic cross correlation of array output. Target tracking in clutter environment is also addressed in [4].

Sparse signal representation has been applied for spectral analysis [10, 14–16, 26, 34]. In [34], a Cauchy-prior is used to enforce sparsity in a temporal spectrum. A recursive weighted least-squares algorithm called FOCUSS has been developed in [16] for source localization. Fuchs [14, 15] formulates the source localization problem as a sparse recovery problem in the beam-space domain. DOA estimation has been posed into joint-sparse recovery problem in [10, 19, 26]. ℓ_1-SVD [26] combines the singular value decomposition (SVD) step of the subspace algorithms with a sparse recovery method based on ℓ_1-norm minimization. The ℓ_1-SVD algorithm can handle closely spaced correlated sources if the number of sources is known. An alternative strategy called joint $\ell_{2,0}$ approximation of DOA (JLZA-DOA) is proposed in [19]. The algorithm represents the snapshots of sensors as some jointly sparse [18] linear combinations of the columns of a manifold matrix. The resulting optimization problem of JLZA-DOA has been solved using a convex–concave procedure. JLZA is further extended to deal with the joint sparse problem with *multiple* measurement matrices arising in broadband DOA estimation, where the manifold matrices for different frequency bands are different. This allows a sensor spacing larger than the smallest half-wavelength of the broadband signal, which in turn results in a significant improvement in the DOA resolution performance. However, these algorithms have been designed for stationary DOA estimation.

There are two aims of the work: (i) develop an efficient algorithm for stationary DOA estimation and, (ii) adopt the algorithm for multiple target tracking. We pose target tracking as a problem of tracking a sparse signal $\mathbf{x}(t)$. Here the field of interest is discretized into a fine grid consisting of a large number (n) of potential DOAs, and $\mathbf{x}(t)$ is an n dimensional complex valued vector, where the ith component of $\mathbf{x}(t)$ is essentially the signal received from the ith point on the DOA-grid at time t. Since there are only a small number of targets at any time t, the vector $\mathbf{x}(t)$ is sparse, and the support of $\mathbf{x}(t)$ gives the locations of the targets. As the targets move, the support of $\mathbf{x}(t)$ changes with time t. If this change is slow enough, then from the estimate of $\mathbf{x}(t-1)$ we can make a fairly accurate prediction of the support of $\mathbf{x}(t)$. Recent papers [24, 39, 40] in compressive sensing (CS) with partially known support demonstrate that such prior knowledge about the support can be used to significantly lower the number of data samples needed for reconstruction.

While the algorithms for CS with partially known support work well for slowly varying support, these methods cannot be used if the target speed is above a particular threshold. To alleviate this problem we propose a MAP estimation approach. At time t we use the past estimates of $\mathbf{x}(\tau)$, $\tau < t$, to construct a a priori predictive probability density function of $\mathbf{x}(t)$. This prior is such that the components of $\mathbf{x}(t)$ which are close to the predicted future location of DOA, will have large magnitude with very high probability. On the other hand, a component of $\mathbf{x}(t)$ which are far from the predicted location is of large magnitude with a very small probability. Subsequently, we use this prior and the array measurements at time t to derive a MAP estimate of $\mathbf{x}(t)$.

We demonstrate the performance of our method on minimum-redundancy linear arrays [27]. In a minimum-redundancy array, the inter-element spacing is not necessarily required to maintain the half wavelength of the receiving narrowband signal. We can resolve relatively large number of DOAs with small sensors and time samples. Moreover, such array arrangement leads to an increase in the resolution of DOA estimation.

Similar to [19], we enforce joint sparsity in DOA estimation for narrowband and broadband signals. In broadband case, this joint sparsity allows a sensor spacing larger than the smallest half-wavelength of the signal, which in turn results in a significant improvement in the DOA resolution performance. This is possible because the spatial aliasing effect can be suppressed by enforcing the joint sparsity of the recovered spatial spectrum across the whole frequency range under consideration.

The chapter is structured as follows. In Sect. 9.2, narrowband stationary DOA estimation is considered. The DOA estimation problem is set as an underdetermined sparse recovery problem. Some state-of-the-art sparsity-based DOA estimation techniques have been discussed, and a MAP-based DOA estimation framework has been developed. Section 9.3 considers the narrowband DOA tracking problem. We formulate MAP framework in DOA tracking. We also demonstrate a possible way to adopt a conventional sparsity-based DOA estimation technique in tracking problem. The proposed MAP approach has been extended for broadband DOA estimation in Sect. 9.4. Finally, in Sect. 9.5 we present some simulation results.

9.2 Stationary Narrowband DOA Estimation

9.2.1 Background

Consider k narrow-band signals $\{s_j(t)\}_{j=1}^k$ incident on a sensor array, consisting of m omnidirectional sensors. Let

$$\mathbf{y}(t) = [\mathbf{y}_1(t) \cdots \mathbf{y}_m(t)]',$$

where $\mathbf{y}_j(t)$ is the signal recorded after demodulation by the jth sensor and $\mathbf{y}'$ denotes the transpose of $\mathbf{y}$. Defining

$$\mathbf{s}(t) = [s_1(t) \cdots s_k(t)]',$$

and using the narrowband observation model [20, 25], we have

$$\mathbf{y}(t) = A(\theta)\mathbf{s}(t) + e(t). \tag{9.1}$$

Here $A(\theta)$ is the manifold matrix, θ is the DOA vector containing the directions of arrival of individual signals, i.e., the jth component θ_j of θ gives the DOA of the signal $s_j(t)$, and $e(t)$ denotes the measurement noise. The manifold matrix consists of the steering vectors $\{a(\theta_j)\}_{j=1}^k$:

$$A(\theta) = [a(\theta_1) \cdots a(\theta_k)].$$

The mapping $a(\theta)$ depends on the array geometry and the wave velocity, which are assumed to be known for any given θ. The problem is to find θ and k from $\{\mathbf{y}_j(t)\}_{j=1}^m$.

9.2.2 Connection to the Blind Source Separation

Blind source separation (BSS) involves recovering unobserved signals from a set of their mixtures [1, 21, 32]. For instance, the signal received by an antenna is a superposition of signals emitted by all the sources which are in its receptive filed. Let us consider k signals $\{s_j(t)\}_{j=1}^k$ incident on a sensor array, consisting of k sensors. Let

$$\hat{\mathbf{y}}(t) = [\hat{\mathbf{y}}_1(t) \cdots \hat{\mathbf{y}}_k(t)]',$$

where $\hat{\mathbf{y}}_j(t)$ is the signal recorded by the jth sensor. The model of sensor output be [1, 21]:

$$\hat{\mathbf{y}}(t) = \hat{A}\mathbf{s}(t) + \hat{e}(t). \tag{9.2}$$

where $\hat{A} \in \mathbb{R}^{k \times k}$ is an unknown nonsingular mixing matrix. Without knowing the properties of the source signals and the mixing matrix, we want to estimate the source signals from the observations $\hat{\mathbf{y}}(t)$ via some linear transformation of the form [1]

$$\hat{\mathbf{s}}(t) = B\hat{\mathbf{y}}(t) \tag{9.3}$$

where $\hat{\mathbf{s}}(t) = [\hat{s}_1(t) \cdots \hat{s}_k(t)]'$, and $B \in \mathbb{R}^{k \times k}$ is a de-mixing matrix. However, without any information about $\hat{A}$ or $\mathbf{s}(t)$, it is impossible to estimate $\hat{\mathbf{s}}(t)$. In BSS, different assumptions have been made on $\mathbf{s}(t)$. For example, independent component analysis (ICA)-based approach assumes that the sources $\mathbf{s}(t)$ are statistically independent [21]. The goal of ICA is to find the transformation matrix B such that the random variables $\hat{s}_1(t), \ldots, \hat{s}_k(t)$ are as independent as possible [32].

There are interesting similarities between the BSS model in (9.2) and the DOA estimation model in (9.1). In the DOA estimation problem (see (9.1)), we have to estimate the source signals, while the matrix $A(\theta)$ and $\mathbf{s}(t)$ are unknown. Unlike BSS, DOA estimation model assumes that the construction of matrix $A(\theta)$ depends on the sensor array geometry and DOA of source signals. Since, the sensor array geometry is known, if one can estimate the DOA of sources, it is possible to construct $A(\theta)$, and hence estimation of $\mathbf{s}(t)$. It is not necessary the sources to be statistically independent [43]. Hence, DOA estimation can be viewed as a semi-BSS problem. Recently, BSS techniques have been applied for DOA estimation [9, 22, 29].

9.2.3 DOA Estimation as a Joint-Sparse Recovery Problem

We divide the whole area of interest into some discrete set of "potential locations". In this work, we consider the far-field scenario and hence the discrete set be a grid of directions-of-arrival angles. Let the set of all potential DOAs be $\mathbb{G} = \{\bar{\theta}_1, \ldots, \bar{\theta}_n\}$, where typically $n \gg k$. The choice of $\mathbb{G}$ is similar to that used in the Capon or MUSIC algorithms. Collect the steering vectors for each element of $\mathbb{G}$ in

$$\Phi = [a(\bar{\theta}_1) \cdots a(\bar{\theta}_n)].$$

Since $\mathbb{G}$ is known, Φ is known and is independent of θ. Now, represent the signal field at time t by $\mathbf{x}(t) \in \mathbb{C}^n$, where the jth component $\mathbf{x}_j(t)$ of $\mathbf{x}(t)$ is nonzero only if $\bar{\theta}_j = \theta_\ell$ for some ℓ, and in that case $\mathbf{x}_j(t) = \mathbf{s}_\ell(t)$. Then one has a model

$$\mathbf{y}(t) = \Phi\mathbf{x}(t) + \bar{e}(t), \tag{9.4}$$

where $\bar{e}(t)$ is the residual due to measurement noise and model-errors. Since $k \ll n$, $\mathbf{x}(t)$ is sparse. Note that the equality $\bar{\theta}_j = \theta_\ell$ may not hold exactly for any $\ell \in \{1, 2, \ldots, k\}$ in practice. Nevertheless, by making $\mathbb{G}$ dense enough, one can ensure $\bar{\theta}_j \approx \theta_\ell$ closely, and the remaining modeling error is absorbed in the residual term

$\bar{e}(t)$. We model the elements of $\bar{e}(t)$ as mutually independent, and identically complex Gaussian distributed random variables with zero mean and variance $1/\lambda$, where λ is a positive real number.

The model, (9.4) lets us pose the problem of estimating k and θ as that of estimating a sparse $\mathbf{x}(t)$ which can be solved using CS [13] framework. If there is a reliable algorithm to recover the sparse $\mathbf{x}(t)$ from $\mathbf{y}(t)$ using (9.4), then all but a few components of the final solution $\mathbf{x}(t)$ will have very small magnitudes. Thus, if the jth component $x_j(t)$ is a dominant component in the recovered $\mathbf{x}(t)$, then we infer that at time t, there is a source with DOA $\bar{\theta}_j$, with an associated signal $x_j(t)$. Finally, the number of these dominant spikes gives k.

A k sparse $\mathbf{x}(t)$ can be recovered uniquely from (9.4) if $k \leq m/2$, and every m columns of Φ form a basis of $\mathbb{C}^m$. The latter is called the *unique representation property*, and is closely connected to the concept of an *un-ambiguous* array. Apart from the limit on k, the single snapshot setting in (9.4) is sensitive to noise. Since noise is ubiquitous in practical problems, we turn to the so-called joint sparse formulation [18]. In practice, we have several snapshots $\{\mathbf{y}(t)\}_{t=1}^{N}$. Using (9.4), we can write

$$Y := [\mathbf{y}(1) \cdots \mathbf{y}(N)] = \Phi X + E, \tag{9.5}$$

where $X = [\mathbf{x}(1) \cdots \mathbf{x}(N)]$ is a sparse matrix, and $E = [\bar{e}_1 \cdots \bar{e}_N]$. If the DOA vector θ is time-invariant over the period of measurement, then for all t the nonzero dominant peaks in $\mathbf{x}(t)$ occur at the same locations corresponding to the actual DOAs. In other words, only k rows of X are nonzero. Such a matrix is called jointly k-sparse. Hence the DOA estimation problem can be posed as the multiple measurement vectors (MMV) problem [18] of finding a jointly sparse X from Y.

9.2.4 Results on the Joint-Sparse Recovery Problem

Assume that $E = 0$, and Y, X are real-valued.[1] The conditions for the existence of a unique solution to the MMV problem is demonstrated by the following lemma [11].

Lemma 1 *Let rank* $(Y) = r \leq m$, *and every m columns of Φ form a basis of* $\mathbb{R}^m$. *Then a solution to* (9.5) *with k nonzero rows is unique provided that $k \leq \lceil (m+r)/2 \rceil - 1$, where $\lceil . \rceil$ denotes the ceiling operation.*

We pose DOA estimation as a MMV problem. Assume that $N > m$, and the matrix

$$X = [\mathbf{x}(1) \cdots \mathbf{x}(N)]$$

---

[1] In the real DOA estimation problem, $E \neq 0$ and X, Y are complex valued. We deal with these issues in the next section.

has rank k. Then according to Lemma 1, the DOAs can be estimated uniquely using the joint sparse framework if $k \leq m - 1$. It is interesting that all the subspace algorithms, e.g., MUSIC or ESPRIT, have the same limitation.

As described in [19], an $\ell_{2,0}$-norm-based minimization approach can be used to solve the MMV problem arising in DOA estimation. However, since zero norm leads to an NP-hard problem, different relaxations used in literature.

9.2.5 DOA Estimation Using ℓ_1 Optimization

ℓ_1-SVD [26] is an efficient algorithm that uses $\ell_{2,1}$-based optimization to solve the DOA estimation problem. In presence of noise, ℓ_1-SVD algorithm considers the following way to solve X given Y in (9.5)

$$\min_X \|Y - \Phi X\|_F^2 + \varsigma \|X\|_{2,1}, \tag{9.6}$$

where $||X||_{2,1}$ is the mixed norm

$$||X||_{2,1} = \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{N} |x_i(j)|^2} = \sum_{i=1}^{n} ||X(i,:)||_2, \tag{9.7}$$

and $\varsigma > 0$ is a tuning factor whose value depends on noise present in the signal. Note that we use the Matlab notation $X(i,:)$ to denote the ith row of X.

The computation time needed to optimization in (9.6) increases with increasing N. To reduce both the computational complexity and sensitivity to noise, ℓ_1-SVD uses SVD of the data matrix $Y \in \mathbb{C}^{m \times N}$. Similar to other subspace algorithms (i.e., MUSIC) the ℓ_1-SVD keeps the signal subspace.

9.2.6 MAP Approach for DOA Localization

In this section we develop a maximum a posteriori (MAP) estimation approach for stationary DOA estimation. Recall that individual columns of E are modeled as mutually independent and identically complex Gaussian distributed random vectors with zero mean and covariance matrix $1/\lambda I$. Using (9.5) the conditional density of Y given X is given by

$$p(Y|X) = \left(\frac{\lambda}{2\pi}\right)^{mN} \exp\{-\lambda||Y - \Phi X||_F^2/2\}. \tag{9.8}$$

Now suppose we know a priori density $p(X)$ of X. Then MAP proposes to estimate X by maximizing the conditional density $p(X|Y)$ given the observed data Y with respect to X. This is same as maximizing the joint log-likelihood function [23]

$$\log p(Y|X) + \log p(X). \tag{9.9}$$

Next we propose a suitable candidate for $p(X)$. Recall that X is row sparse. Furthermore it is reasonable to postulate the rows of X are mutually independent, because the locations of individual targets are independent. In practice, it is common to assume that the elements in a row $X(i,:)$ are independent and identically distributed. The independence follows from the choice of the sampling frequency used in practical arrays. The identical distribution follows because the energy of a target signal remains the same over N snapshots. Now for a given i, we have two possibilities:

- With a high probability $1-q$ there is no target at $\bar{\theta}_i$, and so the elements of $X(i,:)$ are basically of very small energy (contributed by noise and model errors), say μ. We model $X(i,:)$ in this case as a complex Gaussian distributed random vector with mean zero and covariance matrix $\mu^2 I$.
- Otherwise (with a low probability q), there is a target at $\bar{\theta}_i$ so that the elements of $X(i,:)$ have relatively large energy $\rho \gg \mu$ contributed by the target signal. We model $X(i,:)$ in this case as a complex Gaussian distributed random vector with mean zero and covariance matrix $\rho^2 I$.

Consequently, $p(X)$ is a product of Gaussian mixture densities

$$p(X) = \prod_{i=1}^{n} \left\{ \frac{q}{(2\pi\rho)^N} \exp\left(\frac{-||X(i,:)||_2^2}{2\rho^2} \right) + \frac{1-q}{(2\pi\mu)^N} \exp\left(\frac{-||X(i,:)||_2^2}{2\mu^2} \right) \right\}. \tag{9.10}$$

In this work we set $q = k/n$. Such a Gaussian mixture model has been used in simulations [28] and performance analysis [17] in CS literature. Using (9.10) we can write

$$-\ln[p(X)] = \sum_{i=1}^{n} \left(\frac{-||X(i,:)||_2^2}{2\rho^2} - \ln\left[1 + r\exp\left\{ -\frac{||X(i,:)||_2^2}{2\sigma^2} \right\} \right] \right) + \text{constant} \tag{9.11}$$

where the "constant" absorbs the terms independent of X, and

$$r = \frac{(1-q)\rho^N}{q\mu^N}, \qquad \frac{1}{\sigma^2} = \frac{1}{\mu^2} - \frac{1}{\rho^2}. \tag{9.12}$$

Combining (9.11) with (9.8) and (9.9), we can write the criterion function

$$\wp(X) = \sum_{i=1}^{n}\left(\frac{||X(i,:)||_2^2}{2\rho^2} - \ln\left[1 + r_i \exp\left\{-\frac{||X(i,:)||_2^2}{2\sigma^2}\right\}\right]\right) + \frac{\lambda}{2}||Y - \Phi X||_F^2,$$

(9.13)

which we need to minimize with respect to X. The reader may have noticed that while moving from (9.11) to (9.13) we have replaced r by r_i. Indeed for stationary DOA estimation case $r_i = r$, $\forall i$. Having different r_i for different values of i will be useful in tracking moving targets, where it will suffice to minimize the same cost function (9.13). Considering the more general case at this stage allows us to use the results in the next section in the tracking problem.

9.2.7 Solution Strategy

Like many optimization problems encountered in CS literature [7, 38], minimizing (9.13) is a nonconvex problem. To deal with the nonconvex optimization problem we use the concept of graduated nonconvexity (GNC) [2, 3]. GNC is a deterministic annealing method for approximating the global solution for nonconvex minimization problems. Here we construct a sequence of functions $\wp_j(X)$, $j = 0, 1, 2, \ldots, w$ such that

- $\wp_0(X)$ is quadratic;
- $|\wp_{j+1}(X) - \wp_j(X)|$ is small in the neighborhood of the minimizer $X^{(j)}$ of $\wp_j(X)$;
- $\wp_w(X) = \wp(X)$ for a user chosen integer w.

Because $\wp_0(X)$ is quadratic, we can compute $X^{(0)}$ using the standard analytical expression. Then as $|\wp_1(X) - \wp_0(X)|$ is small in the neighborhood of $X^{(0)}$, by initializing a numerical algorithm to minimize $\wp_1(X)$ at $X^{(0)}$ one has a high probability of converging to $X^{(1)}$. If we continue this process of initializing the numerical algorithm to optimize $\wp_{j+1}(X)$ at our estimate of $X^{(j)}$ obtained by numerically optimizing $\wp_j(X)$, then one can expect that $X^{(w)}$ is likely to be the minimizer of $\wp_w(X)$.

The sequence of functions $\wp_j(X)$, $j = 0, 1, 2, \ldots, w$ are constructed as follows. We choose an appropriate real number σ_1 (more details on how the choices are made will follow shortly), and define

$$\wp_0(X) = \sum_i \frac{||X(i,:)||_2^2}{2\rho^2} + \frac{\lambda}{2}||Y - \Phi X||_F^2$$

(9.14)

$$\wp_j(X) = \wp_0(X) - \sum_i \ln\left[1 + r_i \exp\left\{-\frac{||X(i,:)||_2^2}{2\sigma_j^2}\right\}\right] \quad j = 1, 2, 3, \ldots, w;$$

(9.15)

where

$$\sigma_j = (\sigma/\sigma_1)^{j/w} \sigma_1,$$

and w is a user chosen integer. The parameter σ_j controls the degree of nonconvexity in $\wp_j$. As we increase the value of j form 0 to w, we gradually transform $\wp_j$ from a convex function $\wp_0$ to our desired likelihood function $\wp_w$. If w is sufficiently large, then the change from $\wp_{j-1}$ to $\wp_j$ is small, and so is the change from $X^{(j-1)}$ to $X^{(j)}$.

Next we derive an expression for $X^{(0)}$. Let R be a diagonal matrix such that $R_{ii} = \rho^2$. Recall that $X^{(0)}$ is the minimum point of $\wp_0$. Hence, if we differentiate (9.14) with respect to X and evaluate at $X^{(0)}$, we must get zero. Hence

$$X^{(0)} = (R + \lambda\Phi^*\Phi)^{-1}(\lambda\Phi^*Y). \tag{9.16}$$

Note that we denote the conjugate transpose of Φ by Φ^*.

We can reduce the cost of computing $X^{(0)}$ if we use an alternative expression for $X^{(0)}$, which is obtained by applying matrix inversion lemma in the right hand side of (9.16)

$$X^{(0)} = R^{-1}\Phi^*(I/\lambda + \Phi R^{-1}\Phi^*)^{-1}Y. \tag{9.17}$$

where I is a $m \times m$ identity matrix. Computing $X^{(0)}$ via (9.16) requires inverting an $m \times m$ matrix. On the other hand we must invert an $n \times n$ matrix if we compute $X^{(0)}$ via (9.17).

The parameter σ_1 controls the degree of nonconvexity in $\wp_1$. If we take $\sigma_1 \to \infty$, then the logarithmic term in (9.15) tends to $\ln(1+r)$, making $\wp_1$ a quadratic function. In practice, we take

$$\sigma_1 \geq 5 \max_i \|X(i,:)^{(0)}\|_2,$$

This ensures $\exp\{-\|X(i,:)^{(0)}\|_2^2/(2\sigma_1^2)\} \geq 0.99$ for all i. Consequently, $\exp\{-\|X(i,:)^{(0)}\|_2^2/(2\sigma_1^2)\} \approx 1$ for all X satisfying $\|X - X^{(0)}\|_F < \|X^{(1)} - X^{(0)}\|_F$ [28].

9.2.8 Minimizing $\wp_j$

In this section we explore some properties of $X^{(j)}$, and develop a numerical algorithm to compute it. Define

$$\xi_j(X(i,:)) = \frac{1}{\rho^2} + \frac{r_i \exp\left(-\frac{\|X(i,:)\|_2^2}{2\sigma_j^2}\right)}{\sigma_j^2\left[1 + r_i \exp\left(-\frac{\|X(i,:)\|_2^2}{2\sigma_j^2}\right)\right]}, \tag{9.18}$$

and an $n \times n$ diagonal matrix $W_j(X)$ as

$$W_j(X) = \text{diag}\{ \, \xi_j(X(1,:)) \; \xi_j(X(2,:)) \cdots \xi_j(X(m,:)) \, \}.$$

From (9.18) it is readily verified that $\xi_j\{X(i,:)\} > 0$ for all i. Hence $W_j(X)$ is a positive definite matrix.

Now we can verify that

$$\frac{\partial \wp_j(X)}{\partial X} = W_j(X)X - \lambda \Phi^*(Y - \Phi X). \tag{9.19}$$

Since $X^{(j)}$ is the minimum point of $\wp_j$, setting $X = X^{(j)}$ in (9.19) we get

$$X^{(j)} = g_j(X^{(j)}) \tag{9.20}$$

where

$$g_j(X) := \{W_j(X) + \lambda \Phi^* \Phi\}^{-1} \{\lambda \Phi^* Y\}. \tag{9.21}$$

Also, a calculation similar to (9.17) gives

$$g_j(X) := W_j^{-1}(X) \Phi^* \left[I/\lambda + \Phi W_j^{-1}(X) \Phi^* \right]^{-1} Y. \tag{9.22}$$

The equation $X = g_j(X)$ is nonlinear, and cannot be solved analytically. One possibility is to use a fixed point iteration. However, the convergence of the fixed point iteration is not guaranteed. Nevertheless, using (9.19) and (9.21) we have

$$g_j(X) - X = -\{W_j(X^{(j)}) + \lambda \Phi^* \Phi\}^{-1} \frac{\partial \wp_j(X)}{\partial X}. \tag{9.23}$$

Since W_j is a positive definite matrix, and $\lambda > 0$, the matrix $W_j(X^{(j)}) + \lambda \Phi^* \Phi$ is positive definite. Hence (9.23) implies that $\wp_j(X)$ is decreasing along the vector $g_j(X) - X$. In fact moving to $g_j(X)$ from X is same as taking the Newton step associated with some convex–concave procedure to minimize $\wp_j(X)$ [19].

9.2.9 Numerical Algorithm for Solving MAP Optimization

The MAP optimization strategy is given in Table 9.1. We assume that the values of ρ, and μ are known. In fact, simulation results demonstrate that the accurate values of ρ, and μ are not necessary. Instead, an approximation of these values are sufficient [17]. Using initial $X^{(0)}$ we calculate σ_1, and in Step 3 some parameters including w are set. In each iteration, we find a step-length κ along the decent-direction $g_j(X) - X$ using the standard backtracking strategy (step 4–5) [5]. We set $\beta = 0.5$, which

Table 9.1 MAP for narrowband DOA estimation

1. Set $X = X^{(0)}$
2. Set $\sigma_1 = 5 \max_i \|X(i, :)^{(0)}\|_2$
3. Set $j = 1$, choose $\eta \in (0, 1)$, $\beta = 0.5$ and $w \geq 15$
do {
  4. Set $\kappa = 1$
  5. **while** $\wp_j(\kappa g_j(X) + (1 - \kappa)X]) > \wp_j(X)$ {
    $\kappa = \beta\kappa$
  } **end**
  6. $X_o = X$, and $X = \kappa g_j(X) + (1 - \kappa)X$
  7. If $\frac{\|X - X_o\|_2}{\|X_o\|_2} < \eta$ then $j = j + 1$
} **while** $j \leq w$

is very common [5]. The inner-iteration for updating X for a given j terminates when the relative change in the magnitude of X is below η, see step 7. Hence for a smaller η more accurate solutions are sought in expense of higher computation time. According to our experimental study, having $\eta = 0.02$ makes a good tradeoff. Upon convergence of each inner iteration, we increment j (step 7). Note that choosing a larger w helps the optimization problem in (9.15) to move slowly from convex to its desire nonconvex form. Thus we have lower probability to get trapped in local minimum. However, a larger w increases the number of outer iterations (step 4–7) of the algorithm, and hence the computation time. Our experimental study suggests that choosing $w = 20$ makes a good tradeoff between solution accuracy and computation time. Upon convergence for $j = w$, PMAP stops its iterations. The value of λ depends on noise variance. In our simulation, we set $\lambda = 5$ [19].

9.2.10 Acceleration via QR Factorization

Typically, the matrix $X \in \mathbb{C}^{n \times N}$ in (9.5) is large, as n is a large number (we need $n = 360$ to achieve $0.5°$ spatial resolution). If the number of data samples N is large, the algorithm may become very slow. To accelerate the algorithm, we use the QR factorization $Y/\sqrt{N} = \bar{R}Q$, where $\bar{R} \in \mathbb{C}^{m \times m}$ is a nonsingular upper triangular matrix, and $Q \in \mathbb{C}^{m \times N}$ is such that $QQ^* = I$. When $E = 0$, then

$$\text{row span}\{X\} \subset \text{row span}\{Q\}. \tag{9.24}$$

Consequently, $\|Y - \Phi X\|_F^2 = \|\bar{R} - \Phi\bar{X}\|_F^2$, where $\bar{X} = XQ^* \in \mathbb{C}^{n \times m}$ must be jointly row-sparse, and is of significantly smaller size than X. Hence, it is more efficient to estimate $\bar{X}$ via minimizing

$$\wp(\bar{X}) = \sum_{i=1}^{n} \left(\frac{\|\bar{X}(i,:)\|_2^2}{2\rho^2} - \ln\left[1 + r_i \exp\left\{-\frac{\|\bar{X}(i,:)\|_2^2}{2\sigma^2}\right\}\right]\right) + \frac{\lambda}{2}\|\bar{R} - \Phi\bar{X}\|_F^2.$$

$$(9.25)$$

Following (9.10), it is readily verified that $\bar{X}$ and X have identical a priori density function.

9.3 Narrowband Target Tracking

9.3.1 Problem Formulation

Suppose k number of targets are moving in a plane. We wish to

- Detect the targets; and
- Track the DOAs of the targets with a time resolution τ, so that the algorithm will yield estimates DOAs at time instant $0, \tau, 2\tau, 3\tau, \ldots$.

Let f_c be the sampling frequency at the sensors. We assume that over each interval $[\ell\tau, \ell\tau + \frac{N}{f_c})$, the change of $\theta_i(t)$ is negligible, i.e.,

$$\theta(t) \approx \theta(\ell\tau); \quad t \in \left[\ell\tau, \ell\tau + \frac{N}{f_c}\right) \qquad (9.26)$$

where N is the number of snapshots used to detect and estimate the DOAs at time $\ell\tau$. The above is a common assumption made by many state-of-the-art target tracking algorithms [4, 30, 31, 44]. Hence under this assumption, the N snapshots of sensor data in (9.1) can be expressed as

$$y(t) \approx A(\theta(\ell\tau))s(t) + e(t), \quad t \in \left[\ell\tau, \ell\tau + \frac{N}{f_c}\right). \qquad (9.27)$$

Just as in (9.5) we can now formulate a stationary target tracking problem at time instant $\ell\tau$, where we need to recover a joint sparse matrix

$$X_\ell = [\mathbf{x}(\ell\tau) \ \mathbf{x}(\ell\tau + 1/f_c) \ \mathbf{x}\{\ell\tau + (N-1)/f_c\}]$$

given the snapshot matrix

$$Y_\ell = [\mathbf{y}(\ell\tau) \ \mathbf{y}(\ell\tau + 1/f_c) \ \mathbf{y}\{\ell\tau + (N-1)/f_c\}]$$

such that $Y_\ell = \Phi X_\ell + E_\ell$ holds. To solve this problem we can always use the stationary DOA estimation methods discussed before. However, when we track the targets the algorithm is inherently recursive. This means while estimating X_ℓ, we already

know estimates of $X_{\ell-1}$, $X_{\ell-2}$, …. If we can use this prior knowledge efficiently, we can work with a significantly smaller N making the assumption (9.26) a feasible one. In addition it is possible to work with larger number of targets. In order to exploit the prior information available in the estimates of $X_{\ell-1}$, $X_{\ell-2}$, …, we need to consider a dynamic of model for target motion. This is discussed next.

9.3.2 Dynamic Model for the Target Motion

In this contribution, we stick to the most commonly used 'small acceleration' model used in the target tracking literature. For any target i, we assume that its angular acceleration $\ddot{\theta}_i(t)$ is a Wiener process with a small incremental variance. Note that we denote the first order derivative of $\theta(t)$ with respect to t by $\dot{\theta}(t)$, and the second order derivative is denoted by $\ddot{\theta}(t)$.

For a generic DOA $\theta(t)$, it is straightforward to write down the state space equations for $\theta(t)$ in terms of the state

$$\mathbf{s}(t) = [\theta(t) \ \dot{\theta}(t)]'. \tag{9.28}$$

We have

$$\dot{\mathbf{s}}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{s}(t) + \begin{bmatrix} 0 \\ w(t) \end{bmatrix}, \qquad \theta(t) = [1 \ 0] \, \mathbf{s}(t), \tag{9.29}$$

where $w(t)$ denotes the Wiener process with an incremental variance γ. Now we can discretize (9.29) with a time resolution τ, and it is wellknown that the equivalent discrete-time model is given by

$$\mathbf{s}_1(\ell + 1) = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \mathbf{s}_1(\ell) + \mathbf{w}_1(\ell), \qquad \theta(\ell\tau) = [1 \ 0] \, \mathbf{s}_1(\ell), \tag{9.30}$$

where we write $\mathbf{s}_1(\ell) := \mathbf{s}(\ell\tau)$ for short; $\mathbf{w}_1(\ell)$ is a discrete-time, zero mean white noise sequence such that

$$\mathsf{E}\{\mathbf{w}_1\mathbf{w}_1'\} = \gamma \begin{bmatrix} \tau^3/3 & \tau^2/2 \\ \tau^2/2 & \tau \end{bmatrix}. \tag{9.31}$$

Using (9.30), (9.31) and after a few steps of algebra one can show that

$$\theta(\ell\tau) = 2\theta(\ell\tau - \tau) - \theta(\ell\tau - 2\tau) + w_2(\ell), \tag{9.32}$$

where w_2 is a scalar valued discrete-time first order moving average process with zero mean and

$$\mathsf{E}\{w_2\} = 0, \qquad \mathsf{E}\{w_2^2(\ell)\} = \frac{2\gamma\tau^3}{3}, \qquad \mathsf{E}\{w_2(\ell)w_2(\ell - 1)\} = \frac{\gamma\tau^3}{6}. \tag{9.33}$$

9.3.3 Extension of ℓ_1-SVD for DOA Tracking

Using (9.32) we can use the estimates of $X_{\ell-1}$, $X_{\ell-2}$, ... to make predictions about X_ℓ. Then we wish to incorporate this prediction in the MAP framework described above. However, before doing so, we consider how we could naturally extend ℓ_1-SVD method for tracking by using an approach called CS with partially known support [24, 39, 40]. The idea is to use the past estimates $X_{\ell-1}$, $X_{\ell-2}$, ... to estimate the support of X_ℓ, and use that information to estimate a joint sparse X_ℓ.

Suppose that until time $t = (\ell-1)\tau$ the tracking algorithm has detected k targets. The estimated DOAs for the ith target at time$(\ell - j)\tau$ is denoted by $\hat{\theta}_i(\ell - j)$. From (9.32) we know

$$\mathsf{E}\{\theta_i(\ell\tau)\} = 2\hat{\theta}_i(\ell - 1) - \hat{\theta}_i(\ell - 2). \tag{9.34}$$

At this stage we neglect the second order statistics of $w_2(\ell)$, because the standard methods for CS with partially known support does not have any provisions to do so. Nevertheless, while discussing MAP approach in the next section, we will use the second order statistics of $w_2(\ell)$.

Define $\mathbb{I} := \{1, 2, \ldots, n\}$. Recall that n is the number of points on the DOA-grid $\mathbb{G} = \{\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_n\}$. CS with partially known support requires us to predict the support $T(\ell)$ of X_ℓ defined as

$$T_\ell := \{i \in \mathbb{I} : ||X_\ell(i, :)||_2 \neq 0\}.$$

We do so as follows. For each i we identify the point of $\mathbb{G}$ which is the nearest to $\mathsf{E}\{\theta_i(\ell\tau)\}$ and denotes the associated index by ι_i:

$$\iota_i = \arg\min_{j \in \mathbb{I}} |\bar{\theta}_j - \mathsf{E}\{\theta_i(\ell\tau)\}|.$$

Then form

$$T_\ell = \{\iota_1, \iota_2, \ldots, \iota_k\}.$$

Different CS-based algorithms have been developed to exploit the support information in sparse recovery process [24, 39, 40]. Least squares CS-residual (LS-CS) [39] is a two step procedure. First, a least squares (LS) estimate $\bar{X}_\ell$ of X_ℓ is computed assuming that the support of X_ℓ is T_ℓ. To explain the details, let

$$\Phi_+ = [\Phi(:, \iota_1) \ \Phi(:, \iota_2) \cdots (:, \iota_k)]$$

and

$$X_+ = [\Phi_+^* \Phi_+]^{-1} \Phi_+^* Y.$$

Note that X_+ is a $k \times N$ matrix, and we form $\bar{X}_\ell$ as follows:

$$\bar{X}_\ell(i, :) = \begin{cases} X_+(j, :), & \text{if } i = \iota_j \text{ for some } j \in \{1, \ldots, k\}, \\ 0, & \text{otherwise.} \end{cases} \quad (9.35)$$

Next calculate the associated residual

$$\bar{Y} = Y - \Phi_{T_\ell} \bar{X}_\ell.$$

In the subsequent step, LS-CS uses a CS algorithm to find a sparse solution $\hat{X}_\ell$ such that $\bar{Y}_\ell = \Phi \hat{X}_\ell$. The final estimate is $\bar{X}_\ell + \hat{X}_\ell$. Adapting recently proposed modified-CS [40] to our problem, this step requires us to solve

$$\hat{X}_\ell = \arg \min_X \ \|Y - \Phi X\|_F^2 + \varsigma \sum_{i \in \mathbb{I} \setminus T_\ell} \|X(i, :)\|_2. \quad (9.36)$$

We refer the modification of ℓ_1-SVD as ℓ_1-SVD-MCS. It might be worthwhile to note the difference between (9.6) and (9.36), and see how easily ℓ_1-SVD in (9.6) is adapted to the framework of CS with partially known support.

9.3.4 MAP for Tracking

The MAP algorithm can be used for tracking problem with a small modification in the expression for $p(X)$ given in (9.10). Here we have the option to use the estimates of $X_{\ell-1}, X_{\ell-2}, \ldots$ to obtain a better prior density $p(X)$.

As before, suppose that until time $t = (\ell-1)\tau$, the tracking algorithm has detected k targets. Then according to (9.32) the conditional density of $\theta_i(\ell\tau)$ evaluated at θ is proportional to

$$\eta_i(\theta) = \exp \left\{ -\frac{\left[\theta - 2\hat{\theta}_i(\ell - 1) + \hat{\theta}_i(\ell - 2) \right]^2}{4\gamma\tau^3/3} \right\}.$$

It is natural to use this conditional density as a measure of the probability that $\theta_i(\ell\tau)$ is close to a grid point $\bar{\theta}_j$. In particular, if the ith target was the only target detected, then the probability q_j that we will find that target at the grid point $\bar{\theta}_j$ is evaluated as

$$\frac{\eta_i(\bar{\theta}_j)}{\sum_{j \in \mathbb{I}} \eta_i(\bar{\theta}_j)}.$$

When we have k targets in the field, then the probability q_j of finding a target at grid point $\bar{\theta}_j$ is given by

$$q_j = 1 - \prod_{i=1}^{k} \left\{ 1 - \frac{\eta_i(\bar{\theta}_j)}{\sum_{j \in \mathbb{I}} \eta_i(\bar{\theta}_j).} \right\}, \tag{9.37}$$

The above expression for (9.37) works when no new target can appear in the field, and none of the existing targets can disappear. Nevertheless, we can generalize (9.37) to relax these requirements. Let

- α be the probability that an existing target disappears; and
- β be the probability that a new target appears in the field at a grid point.

Now we modify (9.37) to accommodate the possibility that a new target may appear in the field and an existing target may disappear. The event that "the ith target is not presesent at $\bar{\theta}_j$" is the union of two mutually exclusive events:

1. The target has actually disappeared from the field (with probability α); and
2. The target is still there in the field (with probability $1 - \alpha$), but it is not at $\bar{\theta}_j$. The probability of this event is

$$(1 - \alpha) \left\{ 1 - \frac{\eta_i(\bar{\theta}_j)}{\sum_{j \in \mathbb{I}} \eta_i(\bar{\theta}_j)} \right\}.$$

Now combining the probabilities of (1) and (2), the resultant probability that "the ith target is not present at $\bar{\theta}_j$" is

$$\alpha + (1 - \alpha) \left\{ 1 - \frac{\eta_i(\bar{\theta}_j)}{\sum_{j \in \mathbb{I}} \eta_i(\bar{\theta}_j)} \right\}.$$

Then the probability that none of the existing targets is present at $\bar{\theta}_j$ and no new target appears at $\bar{\theta}_j$ is

$$(1 - \beta) \prod_{i=1}^{k} \left[\alpha + (1 - \alpha) \left\{ 1 - \frac{\eta_i(\bar{\theta}_j)}{\sum_{j \in \mathbb{I}} \eta_i(\bar{\theta}_j)} \right\} \right].$$

Thus, to accommodate the possibility that a new target may appear in the field and an existing target may disappear (9.37) is modified accordingly to

$$q_j = 1 - (1 - \beta) \prod_{i=1}^{k} \left[\alpha + (1 - \alpha) \left\{ 1 - \frac{\eta_i(\bar{\theta}_j)}{\sum_{j \in \mathbb{I}} \eta_i(\bar{\theta}_j)} \right\} \right], \quad j \in \mathbb{I} \tag{9.38}$$

Once we know q_i, $\forall i \in \mathbb{I}$, we can replace q by q_i in (9.10) to get

$$p(X) = \prod_{i=1}^{n} \left\{ \frac{q_i}{(2\pi\rho)^N} \exp\left(\frac{-||X(i,:)||_2^2}{2\rho^2} \right) + \frac{1-q_i}{(2\pi\mu)^N} \exp\left(\frac{-||X(i,:)||_2^2}{(2\mu^2)} \right) \right\},$$
(9.39)

which in turn gives

$$-\ln[p(X)] = \sum_{i=1}^{n} \left(\frac{-||X(i,:)||_2^2}{2\rho^2} - \ln\left[1 + r_i \exp\left\{ -\frac{||X(i,:)||_2^2}{2\sigma^2} \right\} \right] \right) + \text{constant}$$
(9.40)

where

$$r_i = \frac{(1-q_i)\rho^N}{q_i \mu^N}, \qquad \frac{1}{\sigma^2} = \frac{1}{\mu^2} - \frac{1}{\rho^2}.$$
(9.41)

Combining (9.40) with (9.8) and (9.9), we again arrive at the criterion function (9.13). However, unlike the stationary case, now all r_i are different from each other. Nevertheless, we can follow the procedure in Sect. 9.2.7 to develop a minimization problem and use the algorithm in Table 9.1 for estimation of $X(\ell)$. In this work, we set $\alpha = 0.01$, $\beta = 0.01$.

9.4 Extension of MAP Framework for Broadband DOA Estimation

We consider the procedure proposed in [19] for broadband DOA estimation. The broadband signal has been splitted into several narrowband signals by using a bank of narrowband filters. Subsequently, the narrowband model (9.5) is applied to each narrowband filter output. Suppose that we have narrowband data at frequencies $\{\omega_i\}_{i=1}^{K}$, and let Φ_i be the "over-complete" manifold matrix at frequency ω_i. Then, the narrowband model at frequency ω_i is of the form

$$Y_i = \Phi_i X_i + E_i, \quad i \in \{1, 2, \ldots, K\}.$$

Here, E_i is the additive noise at frequency ω_i and X_i is the jointly row-sparse signal matrix at frequency ω_i. Now,

$$\mathbf{X} := [X_1 \ X_2 \cdots X_K]$$

is jointly row-sparse. This is because if $X_i(\ell, :)$ is nonzero for some ℓ, then there is source signal at frequency ω_i at direction $\bar{\theta}_\ell$. Therefore, we would expect signals at other frequencies from the direction $\bar{\theta}_\ell$ as well, making $X_i(\ell, :)$ nonzero for all i.

Now to resolve broadband DOA in MAP framework we need to develop a priori density $p(\mathbf{X})$ of $\mathbf{X}$. We assume that the energy of a target at all frequency bands $\{\omega_j\}_{j=1}^{K}$ is almost similar. Nevertheless, the following approach can be extended easily when signal energy is different at different frequency band. Then as described in Sect. 9.2.6, we have two probabilities for every index i: (i) with very small probability q_i, there is a target at $\bar{\theta}_i$, and hence the elements of $\mathbf{X}(i, :)$ have relatively large energy ρ. We model $\mathbf{X}(i, :)$ as a complex Gaussian distributed random vector with zero mean and covariance matrix $\rho^2 I$; (ii) with probability $1 - q_i$, the elements of $\mathbf{X}(i, :)$ have small energy $\mu \ll \rho$. Hence, $p(\mathbf{X})$ is a product of Gaussian mixture densities

$$p(\mathbf{X}) = \prod_{i=1}^{n} \left\{ \frac{q_i}{(2\pi\rho^2)^{KN}} \exp\left(-\frac{\|\mathbf{X}(i, :)\|_2^2}{2\rho^2} \right) + \frac{1 - q_i}{(2\pi\mu^2)^{KN}} \exp\left(-\frac{\|\mathbf{X}(i, :)\|_2^2}{2\mu^2} \right) \right\}. \tag{9.42}$$

Then following (9.8)–(9.13), we can end up the the criterion function

$$\wp(\mathbf{X}) = \sum_{i=1}^{n} \left(\frac{\|\mathbf{X}(i, :)\|_2^2}{2\rho^2} - \ln\left[1 + r_i \exp\left\{ -\frac{\|\mathbf{X}(i, :)\|_2^2}{2\sigma^2} \right\} \right] \right) + \frac{\lambda}{2} \sum_{i=1}^{K} \|Y_i - \Phi_i X_i\|_F^2 \tag{9.43}$$

where, $r_i = \dfrac{(1 - q_i)\rho^{KN}}{q_i \mu^{KN}}$,

which need to minimize with respect to $\mathbf{X}$, and σ are defined in (9.12).

For minimizing (9.43), we follow the GNC procedure of Sect. 9.2.7 and generate w number of suboptimization problem $\{\wp_j(\mathbf{X})\}_{j=1}^{w}$. Then following the calculation (9.18)–(9.23), it can be shown that $\wp_j(\mathbf{X})$ is decreasing along $\mathbf{g}_j(\mathbf{X}) - \mathbf{X}$, where

$$\mathbf{g}_j(\mathbf{X}) = [g_j^{(1)}(\mathbf{X})\, g_j^{(2)}(\mathbf{X})\, \cdots\, g_j^{(K)}(\mathbf{X})], \tag{9.44}$$

$$g_j^{(i)}(\mathbf{X}) = W_j^{-1}(\mathbf{X})\Phi_i^* \left[I/\lambda + \Phi_i W_j^{-1}(\mathbf{X})\Phi_i^* \right]^{-1} Y_i. \tag{9.45}$$

Using the direction we can develop a broadband DOA estimation algorithm. The final algorithm is given in Table 9.2.

9.4.1 Jamming Signal Mitigation

The noise term $e(t)$ in (9.1) is the residual noise due to measurement noise and model error. In general, it is assumed that the noise has uniform distribution and smaller

Table 9.2 MAP for broadband DOA estimation

---

1. Set $X_i^{(0)} = R^{-1}\Phi_i^*(I/\lambda + \Phi_i R^{-1}\Phi_i^*)^{-1}Y_i$, $i = 1, \ldots, K$.

   Form $\mathbf{X}^{(0)} = [\ X_1^{(0)}\ X_2^{(0)} \cdots X_K^{(0)}\ ]$

2. Set $\sigma_1 = 5 \max_i \|\mathbf{X}(i, :)^{(0)}\|_2$

3. Set $j = 1$, choose $\eta \in (0, 1)$, $\beta = 0.5$ and $w \geq 15$

do {

  4. Set $\kappa = 1$

  5. **while** $\wp_j(\kappa \mathbf{g}_j(\mathbf{X}) + (1 - \kappa)\mathbf{X}]) > \wp_j(\mathbf{X})$ {

      $\kappa = \beta\kappa$

  } **end**

  6. $\mathbf{X}_o = \mathbf{X}$, and $\mathbf{X} = \kappa \mathbf{g}_j(\mathbf{X}) + (1 - \kappa)\mathbf{X}$

  7. If $\frac{\|\mathbf{X}-\mathbf{X}_o\|_F}{\|\mathbf{X}_o\|_F} < \eta$ then $j = j + 1$

} **while** $j \leq w$

---

magnitude than the source signal. However, there exists some other types of noisy signals; like, jamming signal. Jamming and deception is the intentional emission of radio frequency signals to interfere with the operation of a radar by saturating its receiver with noise or false information. In general, jamming signals come from fixed directions and have magnitude many times larger than actual source signal [42]. Hence, jamming signals hinder actual source. However, due to larger magnitude, it is easier to know the direction of jamming signal in priori. To mitigate from jamming, we will use the crude estimate of the direction of the jamming signal as a 'partially known support' of the sparse signal $\mathbf{X}$. Let the jammer direction be supported on T_j. Then the value of q_i in (9.42) will be very high if $q_i \in T_j$. We then search any target in rest of the support of $\mathbf{X}$. In our experiments we set $q_i = 0.99$ when $q_i \in T_j$.

9.5 Simulation Results

We compare the performance of MAP-based approach with ℓ_1-SVD [26], Capon's method [6], and ℓ_1-SVD-MCS (see (9.36)). We use four sensors and follow the procedure of minimum-redundancy array [27] for the linear array arrangement. The interelement spacings are d, $3d$ and $2d$, respectively. For narrowband DOA estimation, the value of d is equal to the half wavelength of receiving narrowband signal. For broadband signal we consider two types of the value of d for two classes of algorithms. Similar to [19], MAP algorithm can allow a sensor spacing larger than the half wavelength associated with the highest frequency in the broadband signal. Hence, we set the value of d to be 1.5 times the smallest wavelength in the broadband signal. However, ℓ_1-SVD and Capon cannot allow larger d. Hence we set the value of d equals to 0.5 times the smallest wavelength in the broadband signal. The algorithms starts with a uniform grid with $0.5°$ resolution, i.e., $n = 360$, and $\Phi \in \mathbb{C}^{4 \times 360}$.
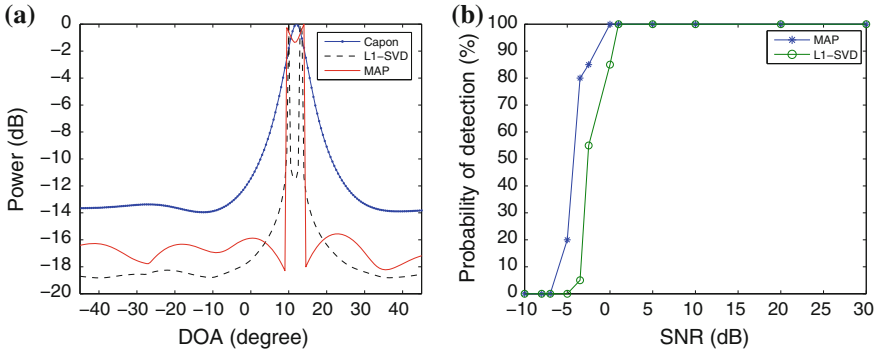
Fig. 9.1 Separating two uncorrelated sources at 10° and 15° by different algorithms. **a** Spatial spectrum obtained by different algorithms. Signal SNR = 2 dB, N = 50. **b** Frequency of separating sources versus SNR

In DOA tracking simulation, we select starting DOA locations first. The future DOA locations are generated using (9.30). The starting $\dot{\theta}(t) = 0.5^0$ and $\gamma = 0.05$. The simulations are performed using MATLAB7.

9.5.1 Narrowband DOA Estimation and Tracking

Each narrowband signal is generated from a zero mean Gaussian distribution. The measurements are corrupted by temporally and spatially uncorrelated zero-mean noise sequence. At first we consider stationary DOA estimation. We assume the number of DOAs k is unknown.

Figure 9.1a shows the spatial spectrum plots of the different algorithms when two uncorrelated sources are placed at 10° and 15°. We take $N = 50$ snapshots and SNR = 2 dB. Capon algorithm cannot separate sources. It indicates one source at 12°. ℓ_1-SVD can separate two sources at 10° and 13°. Hence detection bias is 2°. MAP algorithm can separate sources at 9.5° and 15°. Hence bias in only 0.5°. Next, we investigate the impact of the noise power on the performance of algorithms. Here, we simulate two uncorrelated sources at 10° and 15°. We keep the value of N fixed at 50 and vary the noise power. For each SNR, we carry out 100 independent simulations, and the results are shown in Fig. 9.1b, where we plot the frequency at which the different algorithms separate the sources against noise. Note that MAP outperforms the ℓ_1-SVD. The plots for Capon is not shown, as they are unable to resolve the sources when $N < 140$.

Figure 9.2 shows the results when we simulate two strongly correlated sources at 10° and 15° with a correlation coefficient 0.99. SNR = 6 dB and $N = 50$. Note that MAP can locate the sources clearly, while other methods generates single peak and failed separating sources.
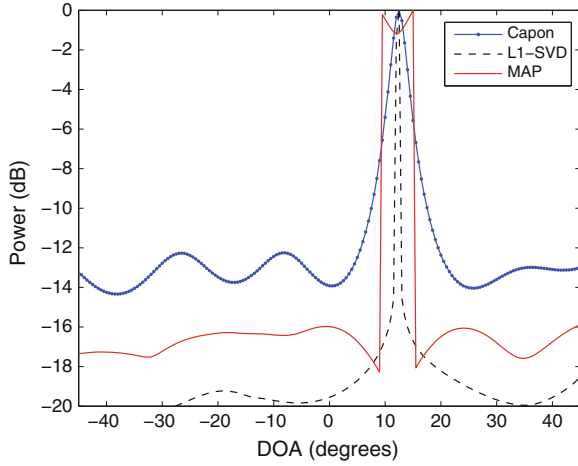
Fig. 9.2 Separating two correlated sources at 10° and 15° by different algorithms. SNR = 6 dB, $N = 50$
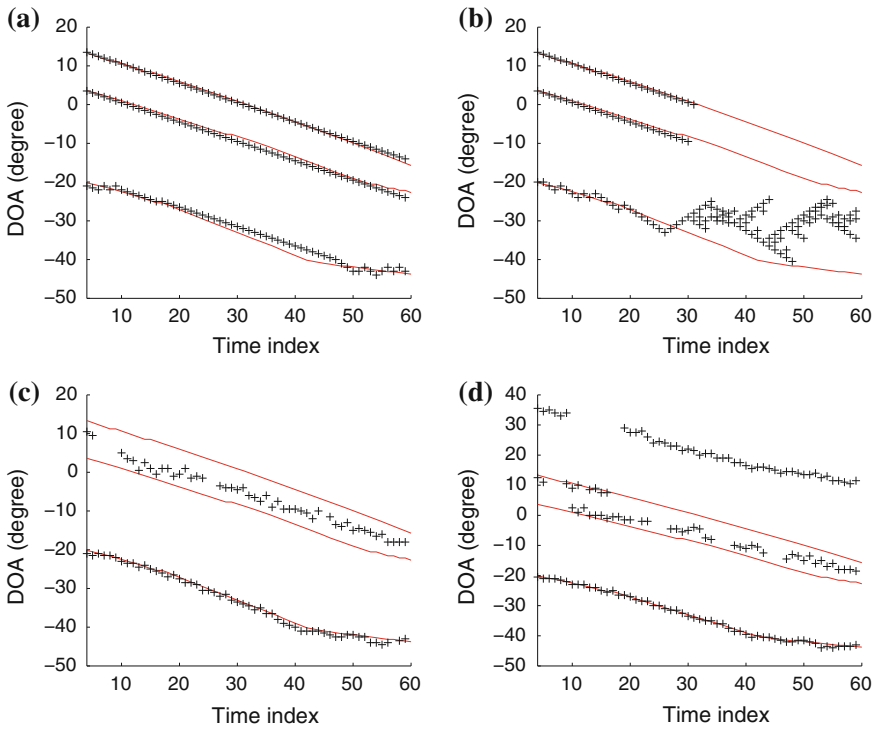

Fig. 9.3 DOA tracking for three targets. SNR = 4 dB, $N = 50$. '−' actual track, '+' estimated track. **a** MAP, **b** ℓ_1-SVD-MCS, **c** ℓ_1-SVD, **d** Capon
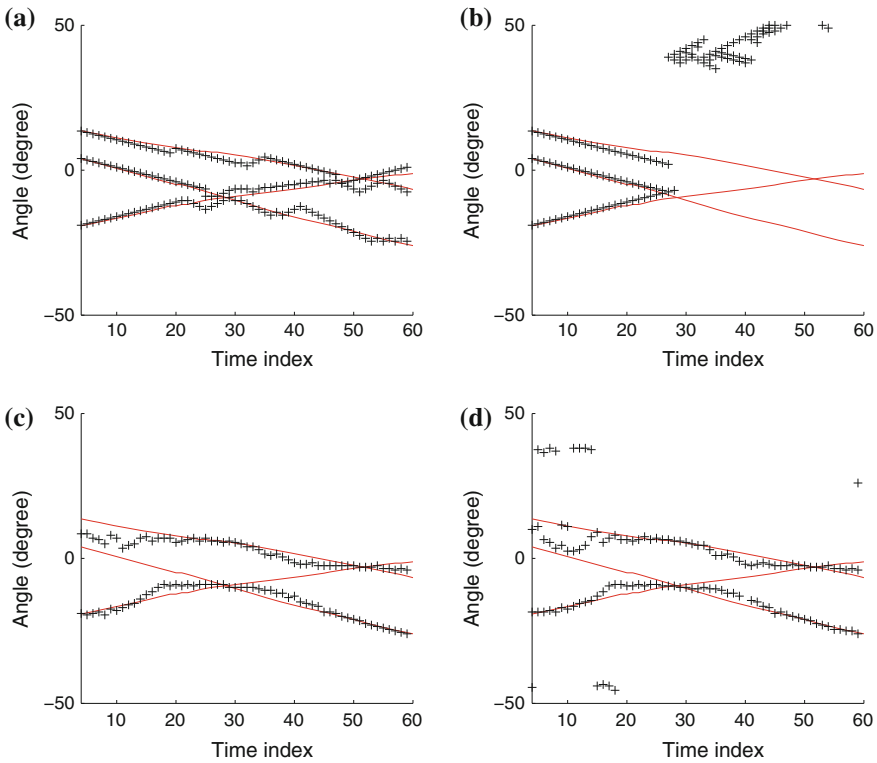
Fig. 9.4 DOA tracking for three targets. SNR = 4 dB, $N = 50$. '$-$' actual track, '$+$' estimated track. **a** MAP, **b** ℓ_1-SVD-MCS, **c** ℓ_1-SVD, **d** Capon

Narrowband DOA tracking results are shown in Fig. 9.3. We consider three uncorrelated moving sources. The starting location of sources are $-20°$, $5°$ and $10°$ respectively. The average SNR = 4 dB and $N = 50$. As can be seen in Fig. 9.3a MAP can track the sources almost accurately. There is a little error tracking the first source between time index 35 and 45. ℓ_1-SVD-MCS can track sources until time index 25. The interesting observation is that once ℓ_1-SVD-MCS loses track of DOAs, it cannot back to the track again. As illustrated in Fig. 9.3b, after time index 25, ℓ_1-SVD-MCS cannot track second and third DOAs anymore. Instead, it generates some random walk around the track of first source. ℓ_1-SVD can track the first source only. Capon also tracks first source. However, it generates another fictitious path from 40° to 20°.

We consider another route in Fig. 9.4. The starting DOA locations are $-20°$, $5°$ and $15°$, respectively. In this scenario the first source crosses the route of other two sources. It is difficult to keep track of sources when they cross each other. Figure 9.4a illustrates that MAP is able to keep tracking the sources. In some cases, the estimated route of some sources are displaced from the actual route, however, MAP algorithm can back to the actual route of sources immediately. As before, ℓ_1-SVD-MCS can
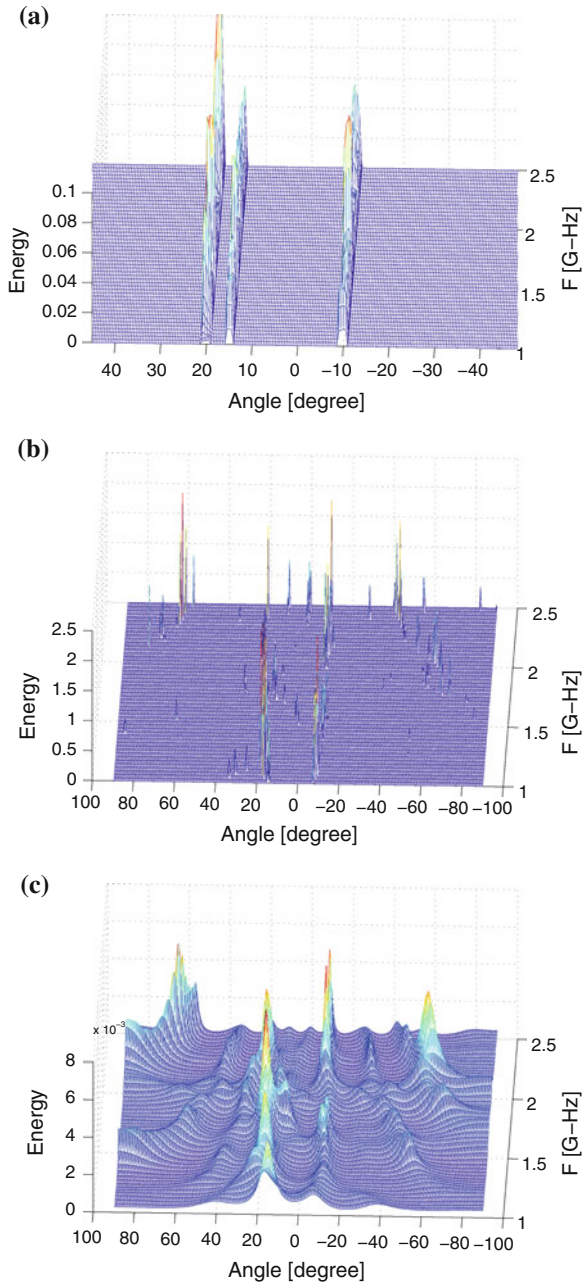
(a)

(b)

(c)

Fig. 9.5 Broadband DOA estimation results for three sources are located at $-10°$, $15°$ and $20°$. SNR $= 6$ dB, $N = 100$. **a** MAP, **b** ℓ_1-SVD, **c** Capon
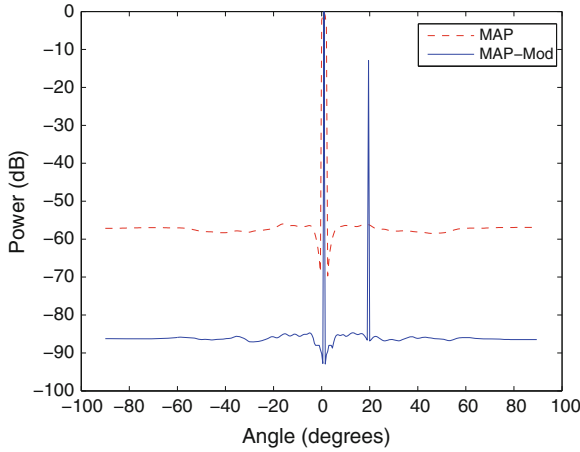
Fig. 9.6 Separating a broadband DOA in presence of jamming signal. The jamming signal comes from 1^0 and actual source is at 20^0. $N = 100$

track the sources until time index 25. The algorithm loses the track of sources when the first and second sources cross each other. ℓ_1-SVD and Capon failed tracking sources. Until time index 25, both algorithms track the first sources, however, they start tracking the second source after time index 25.

### 9.5.2 Broadband DOA Estimation and Tracking

Broadband sources are generated using the procedure [19]. Each source consists of 10 sinusoids with frequencies randomly chosen from the interval [1, 2.5] GHz. The received signal is sampled at 7.5 GHz. The sampled data is filtered through a bank of first-order bandpass filters of the form

$$H_\omega(z) = \frac{(1-r)e^{i\omega}}{z - re^{i\omega}} \tag{9.46}$$

It can be shown that H_ω is a narrow-band filter centered at digital frequency ω (which is related to the analog frequency via the standard relationship). The bandwidth of the filter is controlled by r, where $0 < r < 1$. Taking $r \to 1$ makes the bandwidth smaller, but makes the filter less stable. We take $r = 0.99$. The filterbank consists of 50 filters with center frequencies uniformly distributed over the interval [1, 2.5] GHz.

We simulate three broadband sources at $-10°$, $15°$ and $20°$ in Fig. 9.5. The SNR = 6 dB and $N = 100$. As can be seen in Fig. 9.5, MAP algorithm separate three sources fairly accurately. The detected peaks are sharp and clear. ℓ_1-SVD generates

Fig. 9.7 Broadband DOA
tracking for four targets.
SNR = 4 dB, $N = 100$. '−'
actual track, '+' estimated
track. **a** MAP, **b** ℓ_1-SVD-
MCS, **c** Capon
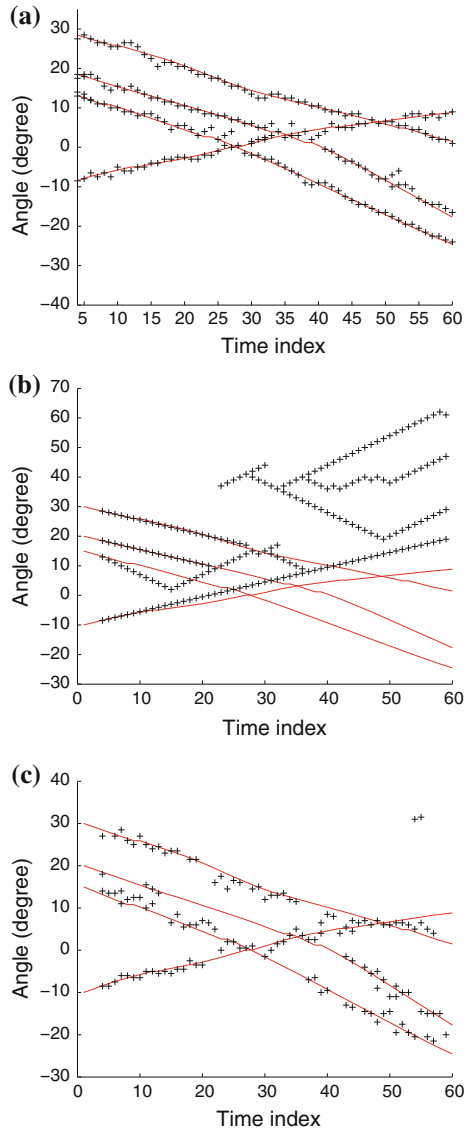
many spurious peaks and failed to locate the actual DOA locations. Capon roughly
generates two peaks at −9° and 19°. However, It generates many other random peaks.

Figure 9.6 shows the source detection performance in presence of jamming signal.
In this setup, a jammer sending signal at an angle 1^0 with a SNR = 40 dB. The
actual source located at 20^0 transmitting signal with SNR = 2 dB. As can be seen,
the proposed modification of MAP for jamming signal (MAP-Mod) in Sect. 9.4.1

can detect the actual source. However, when we are applying the conventional MAP, it detects jamming signal only.

Broadband DOA tracking results are shown in Fig. 9.7. We consider four moving sources. The starting location of sources are $-10°$, $15°$, $20°$ and $30°$, respectively. The SNR $= 4$ dB and $N = 50$. As can be seen in Fig. 9.7a MAP can track the sources with reasonable accuracy.

9.6 Conclusion

A sparse signal reconstruction perspective for source localization and tracking has been proposed. We started with a scheme for localizing narrowband sources and developed a tractable MAP-based optimization approach which can exploit the joint-sparsity arises in the source localization problem. The scheme has been extend for wideband source localization. However, the resulting optimization was nonconvex. Hence, we propose an approach similar to the concept of GNC to cope with the issue. We described how to efficiently mitigate the local minima of the nonconvex optimization through an automatic method by choosing the regularization parameter. We then adopt the MAP formulation for narrowband and wideband source tracking scenario. In source tracking formulation, we utilize the information of current location and moving direction of DOA to estimate its future location. We modify the proposed MAP formulation so that it can use the information efficiently. Finally, we examined various aspects of our approach by using numerical simulations. Several advantages over existing source localization methods were identified, including increased resolution, no need for accurate initialization, and improved robustness to noise.

Some of the interesting questions for future research include an investigation of the applicability of GNC-based sparse recovery algorithms, which have a lower computational cost, to blind source localization. A theoretical study for determining the sequence $\wp_j(X)$ in (9.15) so that the algorithm can avoid local minima will be helpful.

References

1. Amari, S., Cichocki, A., Yang, H.H.: A new learning algorithm for blind signal separation. In: Advances in Neural Information Processing Systems, pp. 757–763. MIT Press, Cambridge (1996)
2. Blake, A., Zisserman, A.: Visual Reconstruction. MIT Press, Cambridge (1987)
3. Blake, A.: Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. IEEE Trans. Pattern Anal. Mach. Intell. **11**(1), 2–12 (1989)
4. Blanding, W., Willett, P., Bar-Shalom, Y.: Multiple target tracking using maximum likelihood probabilistic data association. In: Aerospace Conference, 2007 IEEE, pp. 1–12 (2007). doi:10.1109/AERO.2007.353035

5. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)

6. Capon, J.: High-resolution frequency-wavenumber spectrum analysis. Proc. IEEE **57**(8), 1408–1418 (1969)

7. Chartrand, R.: Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Process. Lett. **14**(10), 707–710 (2007)

8. Chung, P.J., Bohme, J.F., Hero, A.O.: Tracking of multiple moving sources using recursive em algorithm. EURASIP J. Adv. Signal Process. **2005**(1), 534, 685 (2005). doi:10.1155/ASP.2005.50. http://asp.eurasipjournals.com/content/2005/1/534685

9. Chun-yu, K., Wen-Tao, F., Xin-hua, Z., Jun, L.: A kind of method for direction of arrival estimation based on blind source separation demixing matrix. In: Eighth International Conference on Natural Computation (ICNC), pp. 134–137 (2012)

10. Cotter, S.: Multiple snapshot matching pursuit for direction of arrival (doa) estimation. In: European Signal Processing Conference, Poznan, Poland, pp. 247–251 (2007)

11. Cotter, S., Rao, B., Engan, K., Kreutz-Delgado, K.: Sparse solutions to linear inverse problems with multiple measurement vectors. IEEE Trans. Signal Process. **53**(7), 2477–2488 (2005)

12. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc B (Methodol.) **39**(1), 1–38 (1977). http://www.jstor.org/stable/2984875

13. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**, 1289–1306 (2006)

14. Fuchs, J.J.: Linear programming in spectral estimation: application to array processing. In: Proceedings of the IEEE International Conference ICASSP-96 on Acoustics, Speech, and Signal Processing 1996, vol. 6, pp. 3161–3164 (1996). doi:10.1109/ICASSP.1996.550547

15. Fuchs, J.J.: On the application of the global matched filter to doa estimation with uniform circular arrays. IEEE Trans. Signal Process. **49**(4), 702–709 (2001)

16. Gorodnitsky, I, Rao, B.: Sparse signal reconstruction from limited data using focuss: a re-weighted minimum norm algorithm. IEEE Trans. Signal Process. **45**(3), 600–616 (1997)

17. Hyder, M., Mahata, K.: Maximum a posteriori estimation approach to sparse recovery. In: Seventeenth International Conference on Digital Signal Processing (DSP), pp. 1–6 (2011). doi:10.1109/ICDSP.2011.6004892

18. Hyder, M., Mahata, K.: A robust algorithm for joint-sparse recovery. IEEE Signal Process. Lett. **16**(12), 1091–1094 (2009). doi:10.1109/LSP.2009.2028107

19. Hyder, M., Mahata, K.: Direction-of-arrival estimation using a mixed norm approximation. IEEE Trans. Signal Process. **58**(9), 4646–4655 (2010). doi:10.1109/TSP.2010.2050477

20. Johnson, D.H., Dudgeon, D.E.: Array Signal Processing: Concepts and Techniques. Prentice-Hall, Englewood Cliffs (1993)

21. Jutten, C., Herault, J.: Blind separation of sources, Part I: an adaptive algorithm based on neuromimetic architecture. Signal Process. **24**(1), 1–10 (1991)

22. Kang, C., Zhang, X., Han, D.: A new kind of method for DOA estimation based on blind source separation and mvdr beamforming. In: Fifth International Conference on Natural Computation, ICNC '09., vol. 2, pp. 486–490 (2009)

23. Kay, S.M.: Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice-Hall Inc, Upper Saddle River (1993)

24. Khajehnejad, M., Xu, W., Avestimehr, A., Hassibi, B.: Analyzing weighted ℓ_1 minimization for sparse recovery with nonuniform sparse models. IEEE Trans. Signal Process. **59**(5), 1985–2001 (2011)

25. Krim, H., Viberg, M.: Two decades of array signal processing research: the parametric approach. IEEE Signal Process. Mag. **13**(4), 67–94 (1996)

26. Malioutov, D., Cetin, M., Willsky, A.: A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans. Signal Process. **53**(8), 3010–3022 (2005)

27. Moffet, A.: Minimum-redundancy linear arrays. IEEE Trans. Antennas Propag. **16**(2), 172–175 (1968). doi:10.1109/TAP.1968.1139138

28. Mohimani, H., Babaie-Zadeh, M., Jutten, C.: A fast approach for overcomplete sparse decomposition based on smoothed ℓ^0 norm. IEEE Trans. Signal Process. **57**(1), 289–301 (2009)

29. Mukai, R., Sawada, H., Araki, S., Makino, S.: Blind source separation and DOA estimation using small 3-D microphone array. In Proc. HSCMA 2005, pp. d.9–10 (2005)
30. Park, S., Ryu, C., Lee, K.: Multiple target angle tracking algorithm using predicted angles. IEEE Trans. Aerosp. Electron. Syst. **30**(2), 643–648 (1994)
31. Rao, C., Sastry, C.R., Zhou, B.: Tracking the direction of arrival of multiple moving targets. IEEE Trans. Signal Process. **42**(5), 1133–1144 (1994). doi:10.1109/78.295205
32. Rojas, F., Puntonet, C., Rodriguez-Alvarez, M., Rojas, I., Martin-Clemente, R.: Blind source separation in post-nonlinear mixtures using competitive learning, simulated annealing, and a genetic algorithm. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **34**(4), 407–416 (2004)
33. Roy, R., Kailath, T.: ESPRIT—estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. **37**(7), 984–995 (1989)
34. Sacchi, M., Ulrych, T., Walker, C.: Interpolation and extrapolation using a high-resolution discrete fourier transform. IEEE Trans. Signal Process. **46**(1), 31–38 (1998)
35. Schmidt, R.: Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**(3), 276–280 (1986)
36. Stoica, P., Sharman, K.C.: Maximum likelihood methods for direction-of-arrival estimation. IEEE Trans. Acoust. Speech Signal Process. **38**(7), 1132–1143 (1990)
37. Stoica, P., Moses, R.: Introduction to Spectral Analysis, 2nd edn. Prentice-Hall, Upper Saddle River (2004)
38. Trzasko, J., Manduca, A.: Relaxed conditions for sparse signal recovery with general concave priors. IEEE Trans. Signal Process. **57**(11), 4347–4354 (2009)
39. Vaswani, N.: LS-CS-residual (LS-CS): compressive sensing on least squares residual. IEEE Trans. Signal Process. **58**(8), 4108–4120 (2010)
40. Vaswani, N., Lu, W.: Modified-CS: modifying compressive sensing for problems with partially known support. IEEE Trans. Signal Process. **58**(9), 4595–4607 (2010)
41. Viberg, M., Ottersten, B.: Sensor array processing based on subspace fitting. IEEE Trans. Signal Process. **39**(5), 1110–1121 (1991)
42. Xu, L., Li, J., Stoica, P.: Target detection and parameter estimation for mimo radar systems. IEEE Trans. Aerosp. Electron. Syst. **44**(3), 927–939 (2008)
43. Xu, X., Ye, Z., Zhang, Y.: Doa estimation for mixed signals in the presence of mutual coupling. IEEE Trans. Signal Process. **57**(9), 3523–3532 (2009)
44. Yan, H., Fan, H.H.: An algorithm for tracking multiple wideband cyclostationary sources. In: Thirteenth IEEE/SP Workshop on Statistical Signal Processing, pp. 497–502 (2005). doi:10.1109/SSP.2005.1628646
45. Zhou, Y., Yip, P., Leung, H.: Tracking the direction-of-arrival of multiple moving targets by passive arrays: algorithm. IEEE Trans. Signal Process. **47**(10), 2655–2666 (1999). doi:10.1109/78.790648