# 8 Adaptive Beamforming via Sparsity-Based Reconstruction of Covariance Matrix

Yujie Gu, Nathan A. Goodman, and Yimin D. Zhang

Traditional adaptive beamformers are very sensitive to model mismatch, especially when the training samples for adaptive beamformer design are contaminated by the desired signal. In this chapter, we reconstruct a signal-free interference-plus-noise covariance matrix for adaptive beamformer design. Exploiting the sparsity of sources, the interference covariance matrix can be reconstructed as a weighted sum of the outer products of the interference steering vectors, and the corresponding parameters can be estimated from a sparsity-constrained covariance matrix fitting problem. In contrast to classical compressive sensing and sparse reconstruction techniques, the sparsity-constrained covariance matrix fitting problem can be effectively solved as a modified least squares solution by using the *a priori* information on the array structure. Extensive simulation results demonstrate that the proposed adaptive beamformer almost always provides the near-optimal output performance regardless of the input signal power.

## 8.1 Introduction

Adaptive beamforming is an effective spatial filtering technique that adjusts the beamforming weight vector to increase the strength of the signal of interest while suppressing interference and noise. As a ubiquitous task in array signal processing, adaptive beamforming has been widely used in radar, sonar, wireless communications, radio astronomy, seismology, speech processing, medical imaging, and many other areas (see, for example, [1–5] and the references therein). Unlike conventional data-independent beamformers (e.g., fixed or switched beamformers), adaptive beamformers depend on the array received data and hence are expected to provide better capability for interference suppression and signal enhancement. Nevertheless, it is also well known that adaptive beamformers are extremely sensitive to model mismatch, especially when the training samples used for the calculation of the beamforming weight are *contaminated* by the desired signal. In practice, such model mismatch commonly occurs. For example, the data covariance matrix cannot be accurately estimated due to the limited number of training samples, and the steering vector of the desired signal may also be imprecise or even unknown due to look direction error, imperfect calibration, and

other effects. Whenever model mismatches exist, classical adaptive beamformers (e.g., Capon beamformer [6]) will suffer severe performance degradation. To this end, adaptive beamformer design with robustness against model mismatch has been an intensive research topic in the past decades, and various robust adaptive beamforming techniques have been proposed (see, for example, [4, 7, 8] and the references therein). Based on the principle of adaptive beamforming, these robust adaptive beamformers can be classified into two major categories.

In the first category, robust adaptive beamforming techniques process the sample covariance matrix, because the exact interference-plus-noise covariance matrix is usually unavailable in practical applications. The sample covariance matrix is a maximum likelihood estimate of the data covariance matrix, and thus leads to the optimal output of the resulting adaptive beamformer when the sample size tends to infinity. Unfortunately, the sample size is often limited in practice, thus resulting in significant performance degradation, especially when the desired signal is present in the training samples [9, 10]. The most popular robust beamforming technique in this category is the diagonal loading technique [9, 11–13], which adds a scaled identity matrix to the sample covariance matrix to reduce the conditional number. A major problem with diagonal loading is that there is no clear rule to choose the optimal diagonal loading factor in different scenarios. In order to adaptively choose the diagonal loading factor rather than in an *ad hoc* way, several user parameter-free adaptive beamforming algorithms were proposed (see, for example, [14] and the references therein). The shrinkage estimation approach [15] in the sense of minimizing mean square error (MSE) can automatically compute the diagonal loading levels without the need to specify any user parameters. However, this approach leads to an estimate of the statistical covariance matrix of the array received data rather than the required interference-plus-noise covariance matrix. In such a case, the performance degradation becomes severe with the increase of the desired signal power, even when the desired signal steering vector is exactly known. The eigenspace decomposition technique [16, 17] is another popular approach for robust adaptive beamforming applicable to arbitrary steering vector mismatch case. The key idea of this technique is to use the projection of the presumed steering vector onto the sample signal-plus-interference subspace. This approach requires the knowledge of the dimension of the signal-plus-interference subspace. It is known that this approach suffers severe performance degradation from the *subspace swap*[a] when the signal-to-noise ratio (SNR) is low [18, 19]. It also suffers from the signal self-nulling problem, especially at high SNR levels. A sparsity-based iterative adaptive approach (IAA) [20] can iteratively update the spatial power estimates in the whole observation field and subsequently update the covariance matrix used for adaptive beamformer design. Although it does improve the power estimate, the IAA beamforming algorithm is not robust against direction-of-arrival

[a] A subspace swap occurs when the measured data is better approximated by some components of the noise subspace than by some components of the signal subspace, i.e., there is a switch of vectors between the estimated signal and noise subspaces.

(DOA) mismatch because its weight is simply that of the scanning grid point corresponding to the assumed DOA of the desired signal.

In the second category, robust adaptive beamforming techniques process the presumed desired signal steering vector, because the exact knowledge of the steering vector is not easy to obtain in practice. In practical situations, steering vector mismatch can easily occur due to look direction errors [21, 22] or imperfect array calibration and distorted antenna shape [23]. Besides these, other common causes leading to steering vector mismatches include array manifold mismodeling because of source wavefront distortions resulting from environmental inhomogeneities [24, 25], near-far problem [26], source spreading and local scattering [27–30], as well as other effects [10]. In this category, the linear constrained minimum variance (LCMV) beamformer [31] is most commonly used. It provides robustness against uncertainty in the signal look direction by broadening the main lobe of the beampattern. However, the additional imposed constraints reduce the degrees of freedom (DOFs) of the resulting adaptive beamformer. More importantly, the LCMV beamformer becomes less robust when any other types of steering vector mismatch beyond the look direction errors become dominant. To improve the robustness of adaptive beamformers against arbitrary unknown steering vector mismatches, the worst-case performance optimization-based technique [32–34] makes explicit use of an uncertainty set of the signal steering vector. This method requires that the upper bound of the norm of the mismatch vector is *a priori* unknown. Moreover, the worst operating conditions may not always occur. Hence, this adaptive beamforming technique is also an *ad hoc* approach and will suffer from performance degradation whenever the upper bound of the norm of the mismatch vector is either overestimated or underestimated. Another representative technique in this category is to estimate the desired signal steering vector by maximizing the beamformer output, under the constraint that the convergence of the steering vector estimate to any interference steering vector or their combinations is prohibited [35, 36]. However, the imposed norm constraint on the steering vector is too strict to be satisfied, particularly when there exist local scattering encountered in, e.g., mobile communications and indoor speech signal processing. In such cases, gain perturbations in different sensors cannot be ignored, and then the norm constraint no longer holds.

As mentioned above, these two categories of adaptive beamforming techniques were developed almost independently in the past decades. Obviously, these adaptive beamforming techniques are not optimal, because they respectively assume that either the desired signal steering vector or the interference-plus-noise covariance matrix is exactly known. Since the pioneering work of Vorobyov et al. [37], adaptive beamforming has been required to be jointly robust against covariance matrix uncertainty and steering vector mismatch [38–42]. It is worth noting that, in [40], the interference-plus-noise covariance matrix is reconstructed by integrating the outer products of interference steering vectors weighted by the Capon spatial spectrum over a region separated from the desired signal direction, thus removing the desired signal component from the covariance matrix

used for adaptive beamformer design. The reconstructed interference-plus-noise covariance matrix is then used to correct the presumed signal steering vector in order to maximize the beamformer output power under the only constraint that the corrected steering vector does not converge to any interference steering vector or their combinations. Based on the reconstructed interference-plus-noise covariance matrix and the estimated desired signal steering vector, the resulting adaptive beamformer provides a near-optimal output performance with a fast convergence rate. However, the computational complexity of covariance matrix reconstruction is high due to the integral operation. In addition, there is a certain level of performance loss when the number of training samples is small, because the source power obtained from the Capon spatial spectrum is underestimated and, as a result, the estimated interference-plus-noise covariance matrix is inaccurate.

In this chapter, we will elaborate adaptive beamforming via sparsity-based reconstruction of the interference-plus-noise covariance matrix [43]. By exploiting the sparsity of sources in the observed spatial domain, the interference covariance matrix is reconstructed as a linear combination of the outer products of the interference steering vectors weighted by their individual power, which can be estimated from a sparsity-constrained covariance matrix fitting problem. As such, the proposed technique provides a signal-free interference-plus-noise covariance matrix to enable robust adaptive beamformer design that avoids the signal self-nulling problem. It requires a low computational complexity as there is no matrix inversion or eigen-decomposition involved in the sparsity-constrained covariance matrix fitting problem. Hence, the proposed adaptive beamforming technique is suitable for arbitrary number of training samples [44]. When the number of training samples is larger than the number of array sensors, the formulated sparsity-constrained covariance matrix fitting problem can be effectively solved by using the known array structure, i.e., estimate the directions of sources and their power in turn. The proposed adaptive beamformer is compared to existing state-of-the-art adaptive beamformers in terms of computational complexity, output signal-to-interference-plus-noise ratio (SINR) performance, and convergence rate. Numerical simulations clearly demonstrate the near-optimal output performance and faster convergence rate of the proposed adaptive beamforming algorithm exploiting the sparsity of sources in the spatial domain.

## 8.2 Adaptive beamforming criterion

In this section, we first build the narrowband array signal model, then briefly review adaptive beamforming criteria and classical adaptive beamformers.

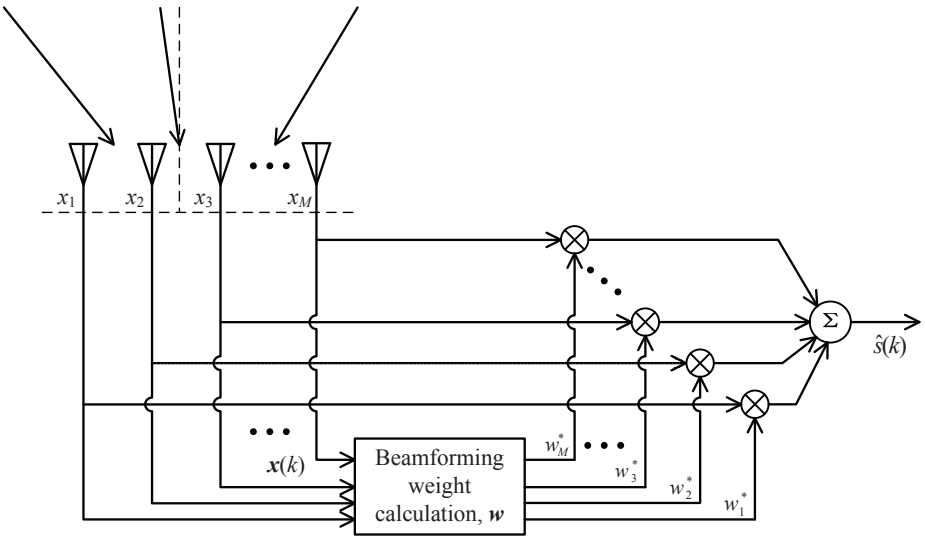



**Figure 8.1** System block diagram of the adaptive beamformer.

### 8.2.1 Array signal model

Consider a narrowband array consisting of $M$ omni-directional sensors depicted in Figure 8.1. The baseband received signal of the array at the time instant $k$, $\boldsymbol{x}(k) = [x_1(k), \cdots, x_M(k)]^T \in \mathbb{C}^M$, can be represented as

$$\boldsymbol{x}(k) = \boldsymbol{x}_s(k) + \boldsymbol{x}_i(k) + \boldsymbol{n}(k), \tag{8.1}$$

where $\boldsymbol{x}_s(k)$, $\boldsymbol{x}_i(k)$, and $\boldsymbol{n}(k)$ are statistically independent components of the desired signal, interference, and noise, respectively. Here, $(\,\cdot\,)^T$ denotes the transpose operator. Among them, the desired signal vector $\boldsymbol{x}_s(k)$ is expressed as

$$\boldsymbol{x}_s(k) = \boldsymbol{a}_s s(k), \tag{8.2}$$

where $s(k)$ is the desired signal waveform, and $\boldsymbol{a}_s \in \mathbb{C}^M$ is the corresponding signal steering vector. Ideally, the steering vector is a function depending on the array geometry as well as source direction, e.g., $\boldsymbol{a}_s \triangleq \boldsymbol{a}(\theta_s)$, where $\theta_s$ is the direction of the desired signal impinging on the array. For example, the ideal steering vector of a uniform linear array (ULA) has the form of

$$\boldsymbol{a}(\theta) = \left[1, e^{-j\frac{2\pi}{\lambda}d\sin\theta}, \cdots, e^{-j\frac{2\pi}{\lambda}(M-1)d\sin\theta}\right]^T, \tag{8.3}$$

where $\theta$ is the DOA of the source, $j = \sqrt{-1}$ is the imaginary unit, $\lambda$ is the wavelength of the narrowband signal, and $d = \lambda/2$ is the inter-element spacing of the array. Similarly, the steering vector of the interference has the similar form with a different source direction. In contrast, there is no such form for the additive noise because noise does not have a fixed direction.

### 8.2.2 Adaptive beamforming criteria

The objective of adaptive beamforming is to design a data-dependent beamforming weight vector $\boldsymbol{w} = [w_1, \cdots, w_M]^T \in \mathbb{C}^M$, such that the beamformer output

$$y(k) = \boldsymbol{w}^H \boldsymbol{x}(k) \tag{8.4}$$

is the best estimate of the desired signal waveform $s(k)$, where $(\cdot)^H$ denotes the Hermitian transpose. To this end, a number of adaptive beamforming criteria have been developed in the past decades. Among them, maximum SINR [6] is the most popular one. Other feasible adaptive beamforming criteria include minimum MSE [3], minimum least-squares error [45], and minimum mutual information [46]. The interested readers are referred to references [3, 4, 47] for the detailed performance tradeoffs among different adaptive beamforming criteria. In this chapter, the maximum SINR criterion will be mainly considered for adaptive beamformer design.

The beamformer output for the SINR maximization problem, defined as

$$\begin{aligned} \max_{\boldsymbol{w}} \ \text{SINR} &\triangleq \frac{\mathbb{E}\big[\, \big|\boldsymbol{w}^H \boldsymbol{x}_s(k)\big|^2 \,\big]}{\mathbb{E}\big[\, \big|\boldsymbol{w}^H \left(\boldsymbol{x}_i(k) + \boldsymbol{n}(k)\right)\big|^2 \,\big]} \\ &= \frac{\sigma_s^2 \big|\boldsymbol{w}^H \boldsymbol{a}_s\big|^2}{\boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w}}, \end{aligned} \tag{8.5}$$

is mathematically equivalent to the minimum variance distortionless response (MVDR) problem [6] as

$$\min_{\boldsymbol{w}} \boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w} \quad \text{s.t.} \quad \boldsymbol{w}^H \boldsymbol{a}_s = 1, \tag{8.6}$$

where $\sigma_s^2 \triangleq \mathbb{E}\big[\, |s(k)|^2 \,\big]$ is the desired signal power, and

$$\boldsymbol{R}_{i+n} \triangleq \mathbb{E}\big[\, \left(\boldsymbol{x}_i(k) + \boldsymbol{n}(k)\right) \left(\boldsymbol{x}_i(k) + \boldsymbol{n}(k)\right)^H \,\big] \in \mathbb{H}^M \tag{8.7}$$

is the interference-plus-noise covariance matrix. Here, $\mathbb{E}\left[\,\cdot\,\right]$ denotes the statistical expectation, and $\mathbb{H}^M$ denotes the $M \times M$ Hermitian matrix. Using the Lagrange multiplier method, the solution of the MVDR problem

$$\boldsymbol{w}_{\text{MVDR}} = \frac{\boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}_s}{\boldsymbol{a}_s^H \boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}_s}, \tag{8.8}$$

is easily obtained. The MVDR beamformer is sometimes referred to as the Capon beamformer, which maximizes the output SINR.

Substituting the data covariance matrix

$$\begin{aligned} \boldsymbol{R} &= \mathbb{E}\left[\boldsymbol{x}(k) \boldsymbol{x}^H(k)\right] \\ &= \sigma_s^2 \boldsymbol{a}_s \boldsymbol{a}_s^H + \boldsymbol{R}_{i+n}, \end{aligned} \tag{8.9}$$

into (8.6) in lieu of the generally unavailable interference-plus-noise covariance

matrix $\boldsymbol{R}_{i+n}$, the corresponding solution,

$$\boldsymbol{w}_{\text{MPDR}} = \frac{\boldsymbol{R}^{-1}\boldsymbol{a}_s}{\boldsymbol{a}_s^H \boldsymbol{R}^{-1}\boldsymbol{a}_s}, \tag{8.10}$$

is referred to as the minimum power distortionless response (MPDR) beamformer. Using the matrix inversion lemma, the MPDR beamformer is proven to be equivalent to the MVDR beamformer as

$$\begin{aligned}\boldsymbol{w}_{\text{MPDR}} &= \frac{1}{\boldsymbol{a}_s^H \boldsymbol{R}^{-1}\boldsymbol{a}_s}\left(\boldsymbol{R}_{i+n} + \sigma_s^2 \boldsymbol{a}_s \boldsymbol{a}_s^H\right)^{-1}\boldsymbol{a}_s \\ &= \frac{1}{\boldsymbol{a}_s^H \boldsymbol{R}^{-1}\boldsymbol{a}_s}\left(\boldsymbol{R}_{i+n}^{-1} - \frac{\boldsymbol{R}_{i+n}^{-1}\boldsymbol{a}_s \boldsymbol{a}_s^H \boldsymbol{R}_{i+n}^{-1}}{\sigma_s^{-2} + \boldsymbol{a}_s^H \boldsymbol{R}_{i+n}^{-1}\boldsymbol{a}_s}\right)\boldsymbol{a}_s \\ &= \alpha \boldsymbol{w}_{\text{MVDR}},\end{aligned} \tag{8.11}$$

where the scalar coefficient $\alpha = \frac{\boldsymbol{a}_s^H \boldsymbol{R}_{i+n}^{-1}\boldsymbol{a}_s}{\boldsymbol{a}_s^H \boldsymbol{R}^{-1}\boldsymbol{a}_s}\left(1 - \frac{\boldsymbol{a}_s^H \boldsymbol{R}_{i+n}^{-1}\boldsymbol{a}_s}{\sigma_s^{-2} + \boldsymbol{a}_s^H \boldsymbol{R}_{i+n}^{-1}\boldsymbol{a}_s}\right)$ does not affect the adaptive beamformer performance in terms of the output SINR. Hence, the MPDR beamformer is also referred to as an MVDR beamformer in the majority of the early literature.

In practical array applications including radar, however, the data covariance matrix cannot be accurately estimated due to the limited training samples, and the signal steering vector may not be precisely known because of the imperfect knowledge of the source location, propagation environment and/or array calibration. In such cases, the MPDR beamformer suffers severe performance degradation, which becomes obvious with the increase of input signal power. Hence, the MPDR beamformer underperforms the MVDR beamformer in practical applications.

### 8.2.3 Adaptive beamformer design

Limited by the size of training samples, the exact data covariance matrix $\boldsymbol{R}$ is not easy available in practical applications, not to mention the signal-free interference-plus-noise covariance matrix $\boldsymbol{R}_{i+n}$. It is usually replaced by the sample covariance matrix

$$\hat{\boldsymbol{R}} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{x}(k)\boldsymbol{x}^H(k), \tag{8.12}$$

where $K$ is the number of snapshots (i.e., training samples). The resulting adaptive beamformer,

$$\boldsymbol{w}_{\text{SMI}} = \frac{\hat{\boldsymbol{R}}^{-1}\bar{\boldsymbol{a}}_s}{\bar{\boldsymbol{a}}_s^H \hat{\boldsymbol{R}}^{-1}\bar{\boldsymbol{a}}_s}, \tag{8.13}$$

is called the sample matrix inversion (SMI) beamformer [48], where $\bar{\boldsymbol{a}}_s = \boldsymbol{a}(\theta_s)$ is the presumed signal steering vector. Whenever there exists a desired signal in the array received signal $\boldsymbol{x}(k)$, the SMI beamformer is in essence an MPDR

beamformer (8.10) rather than an MVDR beamformer (8.8). As $K \to \infty$, $\hat{\boldsymbol{R}}$ will converge to $\boldsymbol{R}$, and the corresponding output SINR will approach the optimal value under stationary and ergodic assumptions. However, when $K$ is small, the large gap between $\hat{\boldsymbol{R}}$ and $\boldsymbol{R}$ is known to dramatically affect the output performance of the SMI beamformer, especially when there is a desired signal in the training samples [9,10].

In order to reduce the sensitivity of the SMI beamformer to model mismatches, many different beamforming algorithms have been developed in the past decades and successfully applied in a wide range of areas (see, for example, [3,4,7,14] and the references therein). In the following, several classical adaptive beamforming algorithms are briefly reviewed.

#### 8.2.3.1 Diagonal loading beamforming

Diagonal loading is the most popular adaptive beamforming approach that is robust to the data uncertainty [9,13]. Replacing the sample covariance matrix $\hat{\boldsymbol{R}}$ in the SMI beamformer (8.13) by a diagonally loaded sample covariance matrix $\hat{\boldsymbol{R}} + \xi \boldsymbol{I}$, the resulting beamformer,

$$\boldsymbol{w}_{\text{DL-SMI}} = \frac{\left(\hat{\boldsymbol{R}} + \xi \boldsymbol{I}\right)^{-1} \bar{\boldsymbol{a}}_s}{\bar{\boldsymbol{a}}_s^H \left(\hat{\boldsymbol{R}} + \xi \boldsymbol{I}\right)^{-1} \bar{\boldsymbol{a}}_s}, \tag{8.14}$$

is referred to as the diagonal loading SMI (DL-SMI) beamformer, where $\xi$ is a diagonal loading factor, and $\boldsymbol{I}$ is an identity matrix.

The performance of the DL-SMI beamformer depends on the diagonal loading factor $\xi$. It is usually chosen in an *ad hoc* way, typically about ten times the noise power, i.e., $\xi = 10\sigma_n^2$, where the noise power $\sigma_n^2$ is assumed to be known [9]. Obviously, it is not optimal, not to mention that the instantaneous noise power is not easy to know. In order to adaptively choose the loading factor, several user parameter-free approaches have been proposed for adaptive beamforming [14]. However, this method leads to an estimate of the data covariance matrix rather than that of the interference-plus-noise covariance matrix. Regardless of the value of the chosen diagonal loading factor, the performance loss of the DL-SMI beamformer is inevitable, and this degradation becomes more severe with the increase of the desired signal power [49]. The main reason is that the desired signal component is always active in any kind of diagonal loading beamformer and its effect becomes more pronounced with the increase of input SNR [40].

#### 8.2.3.2 Eigenspace decomposition beamforming

Motivated by the success of DOA estimation [50], the idea of eigen-decomposition has also been introduced for adaptive beamformer design [16,17]. Replacing the presumed steering vector $\bar{\boldsymbol{a}}_s$ in (8.13) by the projection of $\bar{\boldsymbol{a}}_s$ onto the sample signal-plus-interference subspace, the resulting eigenspace beamformer is given by

$$\boldsymbol{w}_{\text{EIG}} = \hat{\boldsymbol{R}}^{-1} \boldsymbol{P}_{\boldsymbol{E}} \bar{\boldsymbol{a}}_s = \boldsymbol{E} \boldsymbol{\Lambda}^{-1} \boldsymbol{E}^H \bar{\boldsymbol{a}}_s, \tag{8.15}$$

where $\boldsymbol{P}_{\boldsymbol{E}} = \boldsymbol{E}\boldsymbol{E}^H$ is the orthogonal projection matrix onto the signal-plus-interference subspace. Here, the matrix $\boldsymbol{E}$ contains the signal-plus-interference subspace eigenvectors of $\hat{\boldsymbol{R}}$, and the diagonal matrix $\boldsymbol{\Lambda}$ contains the corresponding eigenvalues.

The eigenspace beamformer is robust to arbitrary steering vector mismatch. However, this approach does not work well at low SNR as well as at high signal-to-interference ratio (SIR) cases. In the former case, the estimation of the projection matrix onto the signal-plus-interference subspace breaks down because of the high probability of subspace swaps. In the latter case, the desired signal component denominates the sample covariance matrix, thus degrading the performance of the adaptive beamformer. Furthermore, the eigenspace beamformer also does not work well when the dimension of the signal-plus-interference subspace is high and/or difficult to determine.

#### 8.2.3.3 Worst-case beamforming

The worst-case performance optimization-based adaptive beamforming [32] guarantees a distortionless response for all possible steering vectors in a predetermined set. The worst-case adaptive beamforming problem can be formulated as

$$\min_{\boldsymbol{w}} \boldsymbol{w}^H \hat{\boldsymbol{R}} \boldsymbol{w} \quad \text{s.t.} \quad \max_{\|\boldsymbol{e}_s\|_2 \leq \varepsilon} |\boldsymbol{w}^H (\bar{\boldsymbol{a}}_s + \boldsymbol{e}_s)| \geq 1, \tag{8.16}$$

where $\boldsymbol{e}_s = \boldsymbol{a}_s - \bar{\boldsymbol{a}}_s$ denotes the mismatch vector between the actual signal steering vector $\boldsymbol{a}_s$ and the presumed signal steering vector $\bar{\boldsymbol{a}}_s$, and $\varepsilon$ is the upper bound of the norm of the mismatch vector $\boldsymbol{e}_s$. Here, $\|\cdot\|_2$ denotes the $\ell_2$-norm, also called the Euclidean norm. Because the constraint condition is nonlinear and nonconvex, the worst-case adaptive beamforming problem (8.16) is a semi-infinite nonconvex quadratic program, and is NP-hard[b]. By using the special structure of the objective function and the constraints, the nonconvex optimization problem can be reformulated as a second-order cone programming (SOCP) problem

$$\begin{aligned} \min_{\boldsymbol{w}} \boldsymbol{w}^H \hat{\boldsymbol{R}} \boldsymbol{w} \quad \text{s.t.} \quad & \boldsymbol{w}^H \bar{\boldsymbol{a}}_s \geq \varepsilon \|\boldsymbol{w}\|_2 + 1, \\ & \operatorname{Im}\left(\boldsymbol{w}^H \bar{\boldsymbol{a}}_s\right) = 0, \end{aligned} \tag{8.17}$$

which is convex and can be efficiently solved in polynomial time using the well-established interior point methods. Here, $\operatorname{Im}(\cdot)$ denotes the imaginary part of a complex number.

The worst-case beamformer is robust to arbitrary unknown signal steering vector mismatch with an upper-bounded norm. However, in practical applications, neither the mismatch vector nor its upper bound is *a priori* known. Either overestimation or underestimation of the upper bound of the norm of the steering vector mismatch will degrade the performance of the worst-case beamformer. In addition, the worst-case beamformer also suffers the signal self-nulling problem

[b] In optimization theory, NP-hard problems represent a class of extremely difficult problems that cannot be solved in polynomial time.

because it uses the sample covariance matrix $\hat{\boldsymbol{R}}$ rather than the interference-plus-noise covariance matrix $\boldsymbol{R}_{i+n}$.

#### 8.2.3.4 Iterative adaptive beamforming

The IAA algorithm [20] is a kind of sparse approach to beamforming by iteratively updating the spatial spectrum estimation and beamforming weighting vectors based on a weighted least squares approach. Considering that the IAA depends on the unknown spatial spectrum distribution, it must be implemented in an iterative way. The initialization is done by a delay-and-sum (DAS) beamformer, which is a spatial matched filter with a data-independent weight vector $\boldsymbol{w}_{\mathrm{DAS}} = \frac{\boldsymbol{a}(\theta)}{M}$, as $\hat{s}_l(k) = \boldsymbol{a}^H(\theta_l)\boldsymbol{x}(k)/M,\ l = 1, \cdots, L,\ k = 1, \cdots, K$, from which the power estimates are given by $\hat{p}_l = \frac{1}{K}\sum_{k=1}^{K}|\hat{s}_l(k)|^2, l = 1, \cdots, L$. Here, $L$ is the number of potential source locations in the observed field (or the number of scanning points), which is usually much larger than the true number of sources. Then, the IAA algorithm repeats the following iterative process

$$\begin{aligned}
&\bar{\boldsymbol{R}} = \boldsymbol{A}(\boldsymbol{\theta})\mathrm{diag}(\hat{\boldsymbol{p}})\boldsymbol{A}^H(\boldsymbol{\theta}) \\
&\text{for } l = 1, \cdots, L \\
&\quad \boldsymbol{w}_l = \frac{\bar{\boldsymbol{R}}^{-1}\boldsymbol{a}(\theta_l)}{\boldsymbol{a}^H(\theta_l)\bar{\boldsymbol{R}}^{-1}\boldsymbol{a}(\theta_l)} \\
&\quad \hat{p}_l = \boldsymbol{w}_l^H\bar{\boldsymbol{R}}\boldsymbol{w}_l \\
&\text{end for}
\end{aligned} \tag{8.18}$$

to converge, where $\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_L)] \in \mathbb{C}^{M\times L}$ is the array steering matrix, $\hat{\boldsymbol{p}} = [\hat{p}_1, \hat{p}_2, \cdots, \hat{p}_L]^T \in \mathbb{R}_+^L$ is the estimated spatial spectrum. Here, $\mathbb{R}_+^L$ denotes the set of $L$-dimensional vectors of nonnegative real numbers.

As such, the IAA algorithm can achieve the signal waveform (and hence signal power) estimation in a way of sparse signal representation. It does perform well when there is no model mismatch. However, when there is a slight model mismatch on the signal steering vector, e.g., signal look direction mismatch, performance degradation would occur, and the degradation becomes severe with the increase of the input SNR.

## 8.3 Covariance matrix reconstruction-based adaptive beamforming

In order to avoid, or at least mitigate, the signal self-nulling phenomenon prevalent in adaptive beamformers, in this chapter, we will elaborate a covariance matrix sparse reconstruction method to provide an estimate of the signal-free interference-plus-noise covariance matrix for adaptive beamformer design. In such a case, the performance of the resulting adaptive beamformer will always approach the optimal value in terms of the output SINR. Moreover, the proposed adaptive beamformer has a faster convergence rate than classical adaptive beamformers.

### 8.3.1 Interference-plus-noise covariance matrix reconstruction

Similar to the signal covariance matrix in (8.9), i.e., $\boldsymbol{R}_s = \sigma_s^2 \boldsymbol{a}(\theta_s)\boldsymbol{a}^H(\theta_s)$, the interference-plus-noise covariance matrix has the form of

$$\boldsymbol{R}_{i+n} = \sum_{q=1}^{Q} \sigma_{i_q}^2 \boldsymbol{a}(\theta_{i_q})\boldsymbol{a}^H(\theta_{i_q}) + \sigma_n^2 \boldsymbol{I}, \tag{8.19}$$

where $Q$ is the number of interferers, $\boldsymbol{a}(\theta_{i_q})$ is the steering vector of the $q$-th interference impinging from the DOA $\theta_{i_q}$ and $\sigma_{i_q}^2$ is the corresponding interference power. Hence, in order to have an accurate estimate of the interference-plus-noise covariance matrix $\boldsymbol{R}_{i+n}$, we need to know the steering vectors of all interferers via DOAs and their individual power, together with the noise power. When these pieces of information are unavailable, the interference-plus-noise covariance matrix can be reconstructed as [40]

$$\hat{\boldsymbol{R}}_{i+n} = \int_{\bar{\Theta}} p_{\text{Capon}}(\theta)\boldsymbol{a}(\theta)\boldsymbol{a}^H(\theta)d\theta, \tag{8.20}$$

where $\boldsymbol{a}(\theta)$ is the steering vector associated with a hypothetical direction $\theta$,

$$p_{\text{Capon}}(\theta) = \frac{1}{\boldsymbol{a}^H(\theta)\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}(\theta)} \tag{8.21}$$

is the Capon spatial spectrum estimator, and $\bar{\Theta}$ is the complement sector of $\Theta$. Here, $\Theta$ is a known or estimated angular sector in which the desired signal is located. Hence, the covariance matrix estimator $\hat{\boldsymbol{R}}_{i+n}$ collects all interference and noise in the out-of-sector $\bar{\Theta}$, which effectively excludes the desired signal component.

Correspondingly, the adaptive beamformer based on interference-plus-noise covariance matrix reconstruction

$$\boldsymbol{w}_{\text{Recon}} = \frac{\hat{\boldsymbol{R}}_{i+n}^{-1}\bar{\boldsymbol{a}}_s}{\bar{\boldsymbol{a}}_s^H\hat{\boldsymbol{R}}_{i+n}^{-1}\bar{\boldsymbol{a}}_s}, \tag{8.22}$$

can dramatically improve the performance regardless of the desired signal power (see [40] and accompanying simulations). Nevertheless, the estimation accuracy of $\hat{\boldsymbol{R}}_{i+n}$ in (8.20) is poor because the Capon estimator (8.21) underestimates the true power, especially when the number of snapshots is limited. On the other hand, the computational complexity is high because the covariance matrix reconstruction process introduces the unnecessary integral operation, where the number of interferers is actually countable.

### 8.3.2 Sparsity-based interference-plus-noise covariance matrix reconstruction

Because of the DOF requirement, the number of array sensors is typically larger than the true number of sources. Hence, besides the low-rank characteristic of the array covariance matrix, the target sources in the observed field have the

sparse nature. In such a case, this sparsity can be leveraged to reconstruct the interference-plus-noise covariance matrix $\boldsymbol{R}_{i+n}$, which will provide better estimation accuracy and simplify the integral operation of (8.20) over the entire complement sector $\bar{\Theta}$.

According to (8.19), the interference-plus-noise covariance matrix is a function of the directions and power of interferers, as well as the noise power. The estimation accuracy of these parameters will affect the performance of the adaptive beamformer via the reconstructed interference-plus-noise covariance matrix. To estimate the parameters of both the desired signal and interferers, we formulate a sparsity-constrained covariance matrix fitting problem according to (8.9) as

$$\begin{aligned} \min_{\boldsymbol{p},\sigma_n^2} \quad & \left\| \hat{\boldsymbol{R}} - \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H - \sigma_n^2 \boldsymbol{I} \right\|_F \quad \text{s.t.} \quad \|\boldsymbol{p}\|_0 = Q+1, \\ & \boldsymbol{p} \geq \boldsymbol{0}, \\ & \sigma_n^2 > 0, \end{aligned} \tag{8.23}$$

where $\boldsymbol{p} \in \mathbb{R}_+^L$ is the spatial spectrum distribution on the sample grids of the observed spatial domain (e.g., $\{\theta_1, \theta_2, \cdots, \theta_L\} \in \Theta \cup \bar{\Theta}$), $\boldsymbol{P} = \text{diag}(\boldsymbol{p}) \in \mathbb{R}_+^{L \times L}$ is the corresponding diagonal matrix, $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_L)] \in \mathbb{C}^{M \times L}$ is the array manifold matrix, and $\|\cdot\|_F$ and $\|\cdot\|_0$, respectively, denote the Frobenius norm of a matrix and the $\ell_0$ "norm" of a vector. Note that, although it does not satisfy the positive homogeneity, the $\ell_0$ "norm", which counts the number of non-zero elements in a vector, is an ideal measure of sparsity. According to the sparse observation, the number of potential sources is much larger than the true number of sources, i.e., $L \gg Q+1$. The idea behind (8.23) is intuitive in the sense that it tries to find the sparsest spatial spectrum distribution $\boldsymbol{p}$ and the noise power $\sigma_n^2$ such that the difference between the resulting covariance matrix $\boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma_n^2 \boldsymbol{I}$ and the sample covariance matrix $\hat{\boldsymbol{R}}$ is minimized.

However, the true number of sources is *a priori* unknown. Even if known, it is understood that (8.23) is a difficult combinatorial optimization problem due to the nonconvex $\ell_0$ "norm" constraint, and is intractable even for moderately sized problems. In the past decades, many approximation methods have been proposed to solve this nonconvex optimization problem, such as greedy approximations [51, 52] and $l_p$ ($p \leq 1$) convex relaxations [53, 54]. When the solution $\boldsymbol{p}$ is sufficiently sparse, the $\ell_0$ "norm" can be approximately replaced by the $\ell_1$-norm. By introducing the $\ell_1$-norm convex relaxation, the nonconvex optimization problem (8.23) can be formulated as a convex one

$$\begin{aligned} \min_{\boldsymbol{p},\sigma_n^2} \left\| \hat{\boldsymbol{R}} - \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H - \sigma_n^2 \boldsymbol{I} \right\|_F \quad \text{s.t.} \quad & \|\boldsymbol{p}\|_1 \leq \sigma_s^2 + \sum_{k=1}^{K} \sigma_{i_k}^2 + \sigma_n^2 + \delta, \\ & \boldsymbol{p} \geq \boldsymbol{0}, \\ & \sigma_n^2 > 0, \end{aligned} \tag{8.24}$$

where the $\ell_1$-norm of $\boldsymbol{p}$ equals the power sum of all sources (i.e., $\sigma_s^2 + \sum_{k=1}^{K} \sigma_{i_k}^2 + \sigma_n^2$), and a small number $\delta > 0$ is added to the power constraint in order to allow

a space for the optimization algorithm to search for $\boldsymbol{p}$. However, in practical applications, the true number of sources is not easy to know, not to mention their power.

Alternatively, the above convex optimization problem can be reformulated as a basis pursuit denoising (BPDN) problem [55] as

$$\min_{\boldsymbol{p},\sigma_n^2} \left\| \hat{\boldsymbol{R}} - \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H - \sigma_n^2 \boldsymbol{I} \right\|_F + \gamma \|\boldsymbol{p}\|_1 \quad \text{s.t.} \quad \boldsymbol{p} \geq \boldsymbol{0}, \; \sigma_n^2 > 0, \tag{8.25}$$

where $\gamma$ is a regularization parameter controlling the tradeoff between the sparsity of the spatial spectrum and the residual norm of covariance matrix fitting. The optimization problem is convex and can be solved using standard and highly efficient interior point methods. Besides the BPDN, the least absolute shrinkage and selection operator (LASSO) [56] is another popular formulation based on the $\ell_1$-norm relaxation. Note that, there is no matrix inversion or eigendecomposition required in the proposed sparsity-constrained covariance matrix fitting problem. Hence, it is suitable for arbitrary number of snapshots from one to infinity. However, the obtained solution is not absolutely sparse because of the $\ell_1$-norm relaxation. In addition, the regularization parameter $\gamma$ is difficult to determine in different scenarios. Either overestimation or underestimation will sacrifice the balance between data-fidelity and sparsity, which subsequently leads to performance degradation of the resulting adaptive beamformer.

When the number of snapshots is larger than the number of array sensors, we can decompose the sparsity-constrained covariance matrix fitting problem into two associated sub-problems: 1) a source localization problem to find the DOA support of sources; and 2) a power estimation problem operating on the DOAs estimated in the first sub-problem. The combination of these two sub-problems represents an approximation to the solution of the sparsity-constrained covariance matrix fitting problem.

Compared to adaptive beamforming, DOA estimation is a more mature array processing technique, and there are many sophisticated methods available (see, for example, [3, 53] and the references therein). In general, the DOAs are estimated either from a spectral search algorithm (e.g., for example, [50, 57]) or from a search-free polynomial rooting algorithm (e.g., for example, [58] and the references therein). For the convenience of explanation, here we simply use the classical Capon spatial spectrum $\boldsymbol{p}_{\text{Capon}}(\boldsymbol{\theta})$ in (8.21) to estimate the DOAs of sources. The estimated DOAs provide the support of the sparse vector defined in the proposed sparsity-constrained covariance matrix fitting problem.

Let $\Theta_p$ denote the set of directions corresponding to the peaks of $\boldsymbol{p}_{\text{Capon}}$ on the entire observed spatial domain (i.e., $\Theta \cup \bar{\Theta}$), for which the cardinality is usually greater than the true number of sources because of the spurious peaks (i.e., $|\Theta_p| = \|\boldsymbol{p}\|_0 > Q+1$). Here, $|\cdot|$ denotes the cardinality of a set. In order to minimize the $\ell_0$ "norm" $\|\boldsymbol{p}\|_0$ to find the sparsest solution of (8.23), a common method is to remove the spurious peaks by setting a threshold, such as the noise

power, which can be approximately estimated as the minimum eigenvalue of the sample covariance matrix (i.e., $\hat{\sigma}_n^2 = \lambda_{\min}(\hat{\boldsymbol{R}})$) [59], where $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of a matrix. In theory, there are $M-Q-1$ eigenvalues that equal the actual noise power $\sigma_n^2$. However, in practical applications with a limited number of snapshots, the minimum eigenvalue of the sample covariance matrix is always smaller than the noise power. Hence, if the value of a peak in the Capon spatial spectrum $\boldsymbol{p}_{\text{Capon}}$ is lower than the threshold, it will be regarded as a spurious peak and its corresponding direction will be removed from the set $\Theta_p$. After removing all the spurious peaks, the residual set is denoted as $\tilde{\Theta}_p = \{\tilde{\theta}_{p,1}, \cdots, \tilde{\theta}_{p,\tilde{Q}}\}$ with cardinality $|\tilde{\Theta}_p| = \tilde{Q} \leq |\Theta_p|$. In such a case, $\|\boldsymbol{p}\|_0 = \tilde{Q} \geq Q+1$.

After finding the DOA support $\tilde{\boldsymbol{\theta}}_p = [\tilde{\theta}_{p,1}, \cdots, \tilde{\theta}_{p,\tilde{Q}}]^T$, the sparsity-constrained covariance matrix fitting problem (8.23) degenerates into an inequality-constrained least squares problem:

$$\min_{\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p), \sigma_n^2} \left\| \hat{\boldsymbol{R}} - \boldsymbol{A}(\tilde{\boldsymbol{\theta}}_p)\boldsymbol{P}(\tilde{\boldsymbol{\theta}}_p)\boldsymbol{A}^H(\tilde{\boldsymbol{\theta}}_p) - \sigma_n^2 \boldsymbol{I} \right\|_F \quad \text{s.t.} \quad \boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p) > \boldsymbol{0}, \; \sigma_n^2 > 0, \tag{8.26}$$

where $\boldsymbol{P}(\tilde{\boldsymbol{\theta}}_p) = \text{diag}(\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p))$ is a diagonal matrix with the power distribution $\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p) \in \mathbb{R}_{++}^{\tilde{Q}}$ on the DOA support $\tilde{\boldsymbol{\theta}}_p$, and $\boldsymbol{A}(\tilde{\boldsymbol{\theta}}_p) = [\boldsymbol{a}(\tilde{\theta}_{p,1}), \cdots, \boldsymbol{a}(\tilde{\theta}_{p,\tilde{Q}})] \in \mathbb{C}^{M \times \tilde{Q}}$ is the corresponding array manifold matrix. Here, $\mathbb{R}_{++}^{\tilde{Q}}$ denotes the set of $\tilde{Q}$-dimensional vectors of positive real numbers. The strict inequality constraint enforced here indicates that the signal power on the found DOA support $\tilde{\boldsymbol{\theta}}_p$ are always positive. The above inequality-constrained least squares problem is convex and can be solved using highly efficient interior point methods.

It is noted that covariance matrix reconstruction-based adaptive beamformers are not very sensitive to the estimation error in the noise power. Therefore, for the sake of simplicity, the optimization variable of noise power $\sigma_n^2$ is taken to be the minimum eigenvalue of $\hat{\boldsymbol{R}}$, which leads to a simplified inequality-constrained least squares problem:

$$\min_{\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p)} \left\| \hat{\boldsymbol{R}} - \lambda_{\min}(\hat{\boldsymbol{R}})\boldsymbol{I} - \boldsymbol{A}(\tilde{\boldsymbol{\theta}}_p)\boldsymbol{P}(\tilde{\boldsymbol{\theta}}_p)\boldsymbol{A}^H(\tilde{\boldsymbol{\theta}}_p) \right\|_F \quad \text{s.t.} \quad \boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p) > \boldsymbol{0}. \tag{8.27}$$

Using the vectorization property, the above optimization problem can be further simplified as

$$\min_{\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p)} \left\| \text{vec}\left(\hat{\boldsymbol{R}} - \lambda_{\min}(\hat{\boldsymbol{R}})\boldsymbol{I}\right) - \left(\boldsymbol{A}(\tilde{\boldsymbol{\theta}}_p) \odot \boldsymbol{A}(\tilde{\boldsymbol{\theta}}_p)\right)\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p) \right\|_2 \quad \text{s.t.} \quad \boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p) > \boldsymbol{0}, \tag{8.28}$$

where $\text{vec}(\cdot)$ denotes the vectorization operator, and $\odot$ denotes the Khatri–Rao product. Without the inequality constraint, the closed-form solution to (8.27) is given by

$$\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p) = \left[\boldsymbol{G}^H\boldsymbol{G}\right]^{-1}\boldsymbol{G}^H\boldsymbol{r}, \tag{8.29}$$

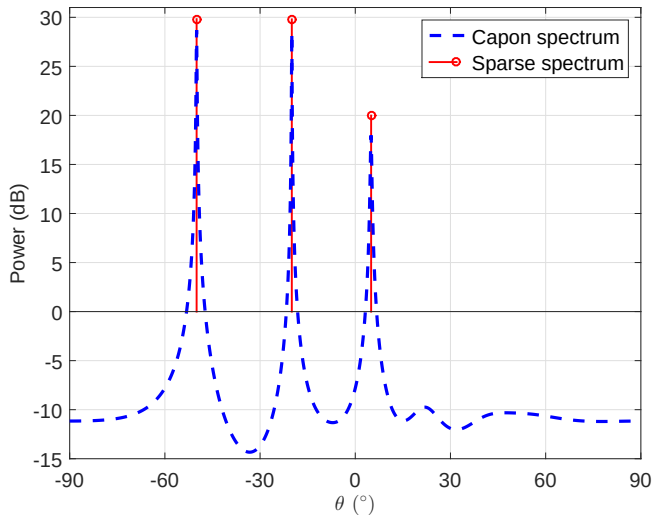



**Figure 8.2** Spatial spectrum comparison.

where $\boldsymbol{G} = \boldsymbol{A}(\tilde{\boldsymbol{\theta}}_p) \odot \boldsymbol{A}(\tilde{\boldsymbol{\theta}}_p) = \left[\boldsymbol{a}(\tilde{\theta}_{p,1}) \otimes \boldsymbol{a}(\tilde{\theta}_{p,1}), \cdots, \boldsymbol{a}(\tilde{\theta}_{p,\tilde{Q}}) \otimes \boldsymbol{a}(\tilde{\theta}_{p,\tilde{Q}})\right] = \left[\mathrm{vec}\left(\boldsymbol{a}(\tilde{\theta}_{p,1})\boldsymbol{a}^H(\tilde{\theta}_{p,1})\right), \cdots, \mathrm{vec}\left(\boldsymbol{a}(\tilde{\theta}_{p,\tilde{Q}})\boldsymbol{a}^H(\tilde{\theta}_{p,\tilde{Q}})\right)\right] \in \mathbb{C}^{M^2 \times \tilde{Q}}$ is obtained by stacking the outer products of the sources steering vectors, and $\boldsymbol{r} = \mathrm{vec}\left(\hat{\boldsymbol{R}} - \lambda_{\min}(\hat{\boldsymbol{R}})\boldsymbol{I}\right) \in \mathbb{C}^{M^2}$ is vectorized from the sample covariance matrix subtracted by an estimated noise covariance matrix. Here, $\otimes$ denotes the Kronecker product. Then, the estimated spatial spectrum of (8.23) is $\tilde{Q}$-sparse and is expressed as

$$\boldsymbol{p}(\theta) = \begin{cases} \boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p) & \theta \in \tilde{\Theta}_p, \\ 0 & \theta \notin \tilde{\Theta}_p. \end{cases} \tag{8.30}$$

Namely, only $\tilde{Q}$ entries of the estimated spatial spectrum $\boldsymbol{p}(\boldsymbol{\theta})$ are nonzero and all other $L - \tilde{Q}$ entries are zero. An example of spatial spectrum comparison between the proposed sparse spectrum (8.30) and the Capon spectrum (8.21) is illustrated in Figure 8.2, where three sources impinge from DOAs of $-50°$, $-20°$ and $5°$ with the SNR of 30 dB, 30 dB and 20 dB, respectively. It is clear that the proposed method achieves a more accurate estimate of the signal power.

However, when some sources are very weak, there may be negative entries in $\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p)$ (8.29), which is obtained by discarding the inequality constraint in (8.27). Without loss of generality, assume that the $\tilde{q}$-th entry of $\boldsymbol{p}(\tilde{\boldsymbol{\theta}}_p)$ is negative, i.e., $p(\tilde{\theta}_{p,\tilde{q}}) < 0$. In such a case, the inequality constraint in (8.26) will not be satisfied, and the closed-form solution in (8.29) should be modified. A simple method is to force $p(\tilde{\theta}_{p,\tilde{q}})$ to be a small positive value $\delta > 0$ (for example, $\delta = 10^{-5}$ is used in our simulations), and the power estimation of other sources $\bar{\boldsymbol{\theta}}_p = \left[\tilde{\theta}_{p,1}, \cdots, \tilde{\theta}_{p,\tilde{q}-1}, \tilde{\theta}_{p,\tilde{q}+1}, \cdots, \tilde{\theta}_{p,\tilde{Q}}\right]^T \in \mathbb{R}^{\tilde{Q}-1}$, will be modified as

$$\bar{\boldsymbol{p}}(\bar{\boldsymbol{\theta}}_p) = \left[\bar{\boldsymbol{G}}^H \bar{\boldsymbol{G}}\right]^{-1} \bar{\boldsymbol{G}}^H \bar{\boldsymbol{r}}, \tag{8.31}$$

where $\bar{\boldsymbol{G}} = \left[\mathrm{vec}\big(\boldsymbol{a}(\tilde{\theta}_{p,1})\boldsymbol{a}^H(\tilde{\theta}_{p,1})\big), \cdots, \mathrm{vec}\big(\boldsymbol{a}(\tilde{\theta}_{p,\tilde{q}-1})\boldsymbol{a}^H(\tilde{\theta}_{p,\tilde{q}-1})\big),\right.$
$\left.\mathrm{vec}\big(\boldsymbol{a}(\tilde{\theta}_{p,\tilde{q}+1})\boldsymbol{a}^H(\tilde{\theta}_{p,\tilde{q}+1})\big), \cdots, \mathrm{vec}\big(\boldsymbol{a}(\tilde{\theta}_{p,\tilde{Q}})\boldsymbol{a}^H(\tilde{\theta}_{p,\tilde{Q}})\big)\right] \in \mathbb{C}^{M^2\times(\tilde{Q}-1)}$, and $\bar{\boldsymbol{r}} = \mathrm{vec}\big(\hat{\boldsymbol{R}} - \lambda_{\min}(\hat{\boldsymbol{R}})\boldsymbol{I} - \delta\boldsymbol{a}(\tilde{\theta}_{p,\tilde{q}})\boldsymbol{a}^H(\tilde{\theta}_{p,\tilde{q}})\big) \in \mathbb{C}^{M^2}$. In other words, we re-calculate the source powers after fixing the power of weak sources, thus resulting in a modified spatial spectrum as

$$\boldsymbol{p}(\theta) = \begin{cases} \bar{\boldsymbol{p}}(\bar{\boldsymbol{\theta}}_p) & \theta \in \bar{\boldsymbol{\theta}}_p, \\ \delta & \theta = \tilde{\theta}_{p,\tilde{q}}, \\ 0 & \theta \notin \tilde{\boldsymbol{\theta}}_p. \end{cases} \tag{8.32}$$

Using the $\tilde{Q}$-sparse spatial spectrum $\boldsymbol{p}(\theta)$, the interference-plus-noise covariance matrix can be sparsely reconstructed as

$$\hat{\boldsymbol{R}}_{i+n} = \sum_{\theta_{i_q}\in\bar{\Theta}\cap\tilde{\Theta}_p} p(\theta_{i_q})\boldsymbol{a}(\theta_{i_q})\boldsymbol{a}^H(\theta_{i_q}) + \hat{\sigma}_n^2\boldsymbol{I}, \tag{8.33}$$

where $\boldsymbol{a}(\theta_{i_q})\boldsymbol{a}^H(\theta_{i_q})$ is the outer product of the $q$-th interference steering vector $\boldsymbol{a}(\theta_{i_q})$. Because there are at most $\tilde{Q}$ elements in the set $\bar{\Theta}\cap\tilde{\Theta}_p$, the integral operation in (8.20) is effectively simplified to be a summation operation (8.33) by using the sparse characteristics of sources in the observed spatial domain. Note that there is no desired signal component in the reconstructed interference-plus-noise covariance matrix.

Considering the possible look direction mismatch, the DOA of the desired signal can be located by searching for the peak of $\boldsymbol{p}_{\mathrm{Capon}}$ in $\Theta$, i.e., $\tilde{\theta}_s = \arg\max_{\theta\in\Theta} p_{\mathrm{Capon}}(\theta)$, and the corresponding steering steering vector is denoted as $\tilde{\boldsymbol{a}}_s = \boldsymbol{a}(\tilde{\theta}_s)$. When $\Theta\cap\tilde{\Theta}_p$ is empty, which is common at low SNRs, we simply use the presumed signal steering vector for adaptive beamformer design, i.e., $\tilde{\boldsymbol{a}}_s = \bar{\boldsymbol{a}}_s$, even if there is a look direction mismatch.

Substituting the reconstructed interference-plus-noise covariance matrix $\hat{\boldsymbol{R}}_{i+n}$ and the estimated signal steering vector $\tilde{\boldsymbol{a}}_s$ into the MVDR beamformer (8.8) together, we can propose the adaptive beamformer as

$$\boldsymbol{w} = \frac{\hat{\boldsymbol{R}}_{i+n}^{-1}\tilde{\boldsymbol{a}}_s}{\tilde{\boldsymbol{a}}_s^H\hat{\boldsymbol{R}}_{i+n}^{-1}\tilde{\boldsymbol{a}}_s}. \tag{8.34}$$

The proposed adaptive beamforming algorithm based on sparse reconstruction of the interference-plus-noise covariance matrix is summarized in Table 8.1.

The computational complexity of the proposed adaptive beamforming algorithm is $\mathcal{O}\left(LM^2\right)$ with $L \gg M$, which is mainly dominated by the spectral search. If a *search-free* DOA estimation technique [58] is adopted, the computational complexity can be further decreased to $\mathcal{O}\left(\max\left(M^3, \tilde{Q}^2M^2\right)\right)$, where $\mathcal{O}\left(M^3\right)$ is the complexity of DOA estimation and $\mathcal{O}\left(\tilde{Q}^2M^2\right)$ is the complexity of power estimation. Therefore, the proposed adaptive beamforming algorithm has complexity slightly higher than the DOA estimation algorithm. Meanwhile, the

**Table 8.1** Adaptive beamforming algorithm based on sparse reconstruction of interference-plus-noise covariance matrix

| | |
|---|---|
| **Step 1**: | Estimate the DOAs of the sources by, e.g., searching for the peaks of the Capon spatial spectrum $\boldsymbol{p}_{\text{Capon}}(\boldsymbol{\theta})$ (8.21); |
| **Step 2**: | Solve the least squares problem (8.27) to obtain the $\tilde{Q}$-sparse spatial spectrum $\boldsymbol{p}(\theta)$ (8.30) or (8.32); |
| **Step 3**: | Reconstruct the interference-plus-noise covariance matrix $\hat{\boldsymbol{R}}_{i+n}$ (8.33) and estimate the signal steering vector $\tilde{\boldsymbol{a}}_s$; |
| **Step 4**: | Compute the proposed adaptive beamformer $\boldsymbol{w}$ (8.34). |

computational complexity of the covariance matrix reconstruction-based adaptive beamforming algorithm [40] is $\mathcal{O}\left(\frac{|\bar{\Theta}|}{|\bar{\Theta}\cup\Theta|}LM^2\right)$. Note however that, if the spatial estimate of the sources in the entire region is desired, the SMI beamformer has the complexity of $\mathcal{O}\left(LM^2\right)$ as well.

## 8.4 Simulation results

In our simulations, a ULA with $M = 10$ omni-directional sensors spaced half wavelength apart is considered. It is assumed that there is one desired signal from the presumed direction $\bar{\theta}_s = 5^\circ$ and two uncorrelated interferers from $-50^\circ$ and $-20^\circ$, respectively. The interference-to-noise ratio (INR) at each sensor is equal to 30 dB. The additive noise is modeled as a complex circularly symmetric Gaussian zero-mean spatially and temporally white process. When comparing the performance of the adaptive beamforming algorithms with respect to the input SNR, the number of snapshots is fixed to be $K = 30$. In the performance comparison of mean output SINR versus the number of snapshots, the SNR in each sensor is set to be fixed at 20 dB. For each data point (SNR or number of snapshots), 500 Monte Carlo trials are performed.

The proposed interference-plus-noise covariance matrix *sparse* reconstruction-based beamformer (8.34) is compared to the SMI beamformer [48], the DL-SMI beamformer [9], the eigenspace decomposition-based beamformer [16], the worst-case performance optimization-based beamformer [32], the IAA beamformer [20], and the interference-plus-noise covariance matrix reconstruction-based beamformer [40]. All the tested beamformers are adaptive beamformers, i.e., their weight vectors depend on the received array data. The diagonal loading factor $\xi$ in the DL-SMI beamformer (8.14) is assumed to be ten times the noise power, where the noise power is regarded as *a priori* known. The eigenspace-based beamformer (8.15) is assumed to know the exact number of interference sources. In the worst-case beamformer (8.16), the upper bound of the mismatched vector is *ad hoc* chosen to be $\varepsilon = 0.3M$ as suggested in [32]. Without loss of generality, the angular sector covering the direction of the desired signal in the reconstruction-

based beamformers, is set to be $\Theta = \left[\bar{\theta}_s - 5^\circ, \bar{\theta}_s + 5^\circ\right]$ (namely, $[0^\circ, 10^\circ]$), and the corresponding out-of-sector is $\bar{\Theta} = \left[-90^\circ, \bar{\theta}_s - 5^\circ\right) \cup \left(\bar{\theta}_s + 5^\circ, 90^\circ\right]$ (namely, $[-90^\circ, 0^\circ) \cup (10^\circ, 90^\circ]$). The sampling grid is uniform in $\Theta \cup \bar{\Theta}$ with $0.1^\circ$ increment between adjacent grid points. As a benchmark, the optimal SINR (8.5) is also shown in all figures, which is calculated from the exact interference-plus-noise covariance matrix and the actual desired signal steering vector. Considering that the output performance of the interference-plus-noise covariance matrix (sparse) reconstruction-based beamformers is very close to the optimal SINR regardless of the input signal power [40, 43], we also compare the output performance in terms of deviation from the optimal SINR. For fair comparison, the actual steering vector of the desired signal is normalized so that $\|\boldsymbol{a}\|_2^2 = M(= 10)$ [32, 33]. The CVX software [60] is used to solve the related convex optimization problems.

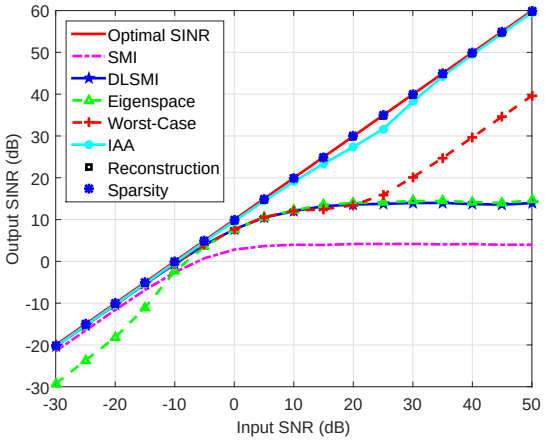
### 8.4.1 Example 1: Exactly known signal steering vector

In the first example, we consider an ideal scenario where the steering vectors of both the desired signal and the interferers are exactly known. Namely, there is no steering vector mismatch. Note that, even in this ideal case, the presence of the desired signal in the training samples may still substantially degrade the output performance of adaptive beamformers as compared with the signal-free training case [3, 32, 40]. However, it can be seen from Figure 8.3(a) that the output performance of adaptive beamformers based on interference-plus-noise covariance matrix reconstruction is almost always equal to the optimal SINR for all SNR values between $-30$ dB and 50 dB (i.e., SIR ranges from $-60$ dB to 20 dB), which illustrates the high dynamic range. In particularly, the output SINR of the proposed interference-plus-noise covariance matrix sparse reconstruction-based adaptive beamformer is approximated as

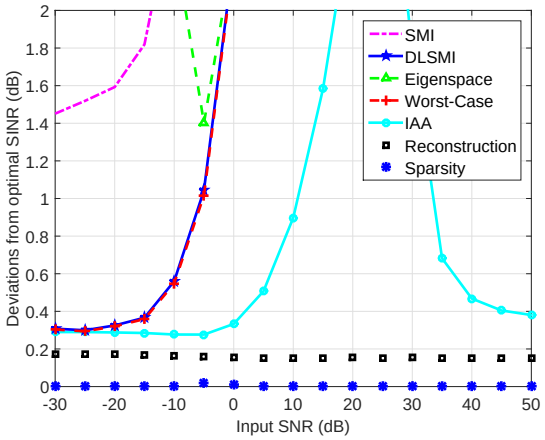
$$\text{SINR} \approx \|\boldsymbol{a}\|_2^2 \,\text{SNR} = M \times \text{SNR}, \tag{8.35}$$

which achieves the design goal of the adaptive beamformer and outperforms the other tested beamformers. From Figure 8.3(b), the average SINR performance loss of the proposed adaptive beamformer is about 0.002 dB. In contrast, there is an average performance loss of 0.158 dB for the interference-plus-noise covariance matrix reconstruction-based adaptive beamformer [40], which is because the Capon spatial spectrum estimator underestimates the interferences power and reduces the estimation accuracy of the interference-plus-noise covariance matrix. It should be noted that the signal power is 100 times higher than the interference power in the case of SNR = 50 dB, which can be used to illustrate the situation when the SIR approximately approaches to infinity. Figure 8.3(c) shows the convergence rates of the tested adaptive beamformers versus the number of snapshots $K$. It is clear that the adaptive beamformer based on interference-plus-noise covariance matrix (sparse) reconstruction converges much faster than the other tested adaptive beamformers.

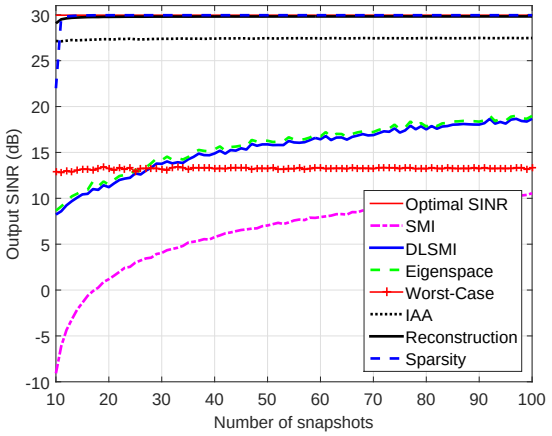
In Figure 8.4, we compare the beampattern of the proposed beamformer with




(a) Output SINR versus input SNR




(b) Deviations from optimal SINR versus input SNR




(c) Output SINR versus number of snapshots

**Figure 8.3** First example: Exactly known steering vectors.

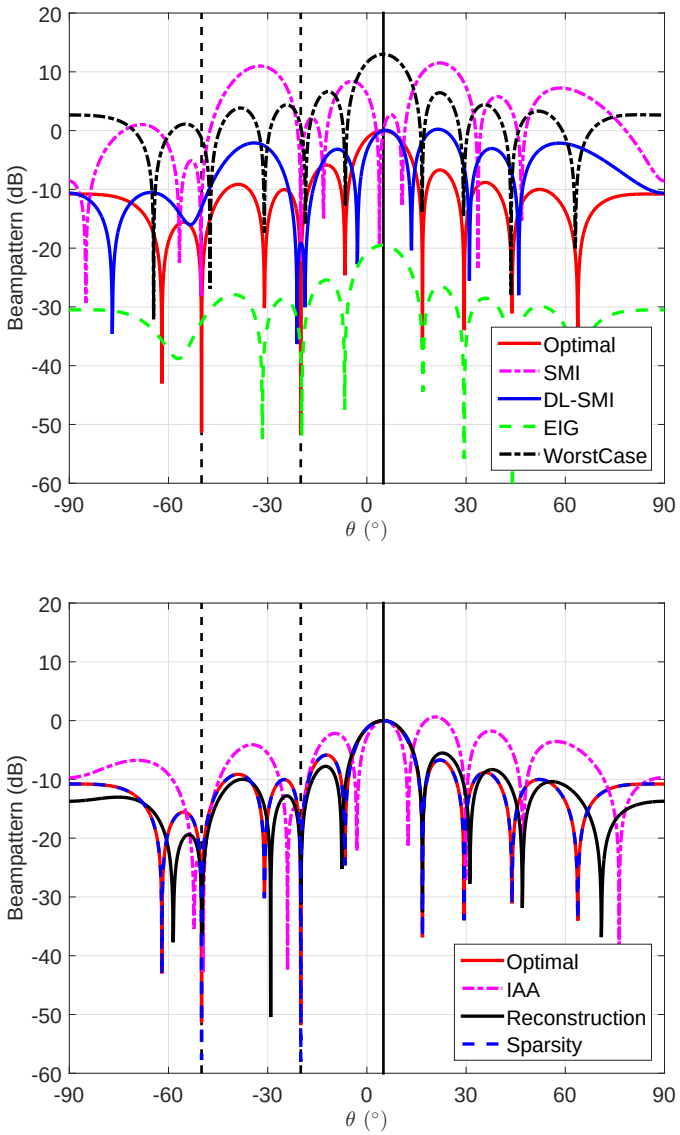



**Figure 8.4** First example: Beampattern comparison.

those of the other tested beamformers for $K = 30$ and SNR $= 20$ dB, where the vertical solid line denotes the direction of the desired signal and the vertical dashed lines denote the directions of interference. It is evident that the beampattern of the proposed adaptive beamformer almost exactly matches that of the optimal one.

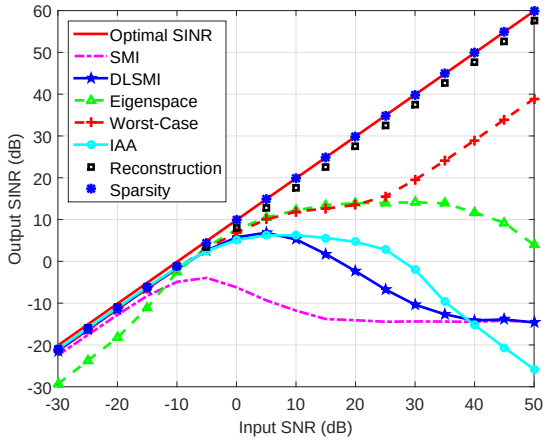
### 8.4.2 Example 2: Fixed signal DOA mismatch

In the second example, a scenario with fixed signal DOA mismatch is considered. We assume that the actual DOA of the desired signal is 8° while the assumed one is 5°. Correspondingly, there is a fixed DOA mismatch of 3° for the desired signal. By comparing Figure 8.5(a) with Figure 8.3(a), we can see that, when the input SNR is 20 dB, there is about 18 dB performance loss for both the SMI beamformer and the DL-SMI beamformer. There is no obvious performance change for the worst-case beamformer, while the eigenspace-based beamformer suffers clear performance loss when the signal power is higher than the interference power. Compared with the performance loss (about 2 dB) of the interference-plus-noise covariance matrix reconstruction-based beamformer, there is almost no performance loss for the proposed adaptive beamformer based on covariance matrix sparse reconstruction when the input SNR is above 0 dB. When the SNR is lower than 0 dB, the performance loss is mainly because of the DOA estimation accuracy of the Capon spatial spectrum. The output performance of the proposed beamformer can be further improved by introducing more sophisticated DOA estimation methods especially for low SNR cases. Similarly, as shown in Figure 8.5(c), the output performance of the proposed adaptive beamformer is close to the optimal SINR when the number of snapshots is larger than the number of array sensors.

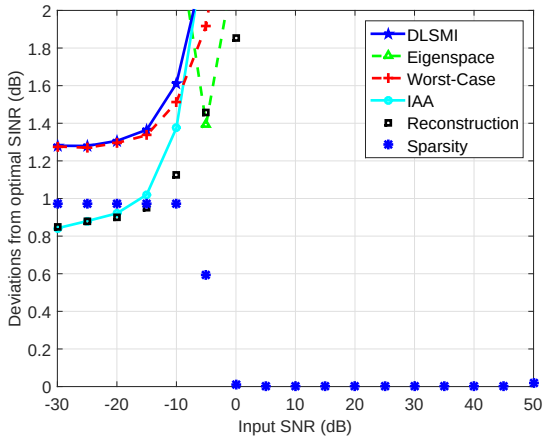
### 8.4.3 Example 3: Random sources DOA mismatch

In the third example, a more practical scenario with random DOA mismatches is considered. More specifically, random DOA mismatches of both the desired signal and the interferers are assumed to be uniformly distributed in $[-4°, 4°]$. That is to say, the actual DOA of the desired signal is uniformly distributed as $\mathcal{U}\left[\bar{\theta}_s - 4°, \bar{\theta}_s + 4°\right]$ (i.e., $\mathcal{U}\left[1°, 9°\right]$), and the DOAs of the interferers are uniformly distributed as $\mathcal{U}\left[-54°, -46°\right]$ and $\mathcal{U}\left[-24°, -16°\right]$, respectively. Note that, random DOAs of the signal and interferers change from trial to trial but remain fixed from snapshot to snapshot.

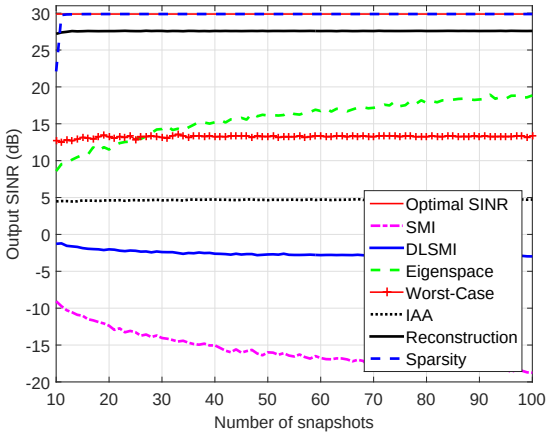
It can be seen from Figure 8.6(a) that the output performance of the proposed beamformer is much closer to the optimal SINR than other tested beamformers. When the input SNR is less than $-10$ dB, there is an approximately 0.6 dB performance loss because there may be no peak in the angular sector $\Theta$ for the Capon spectrum or the peak's value is less than the threshold; therefore, there presents a random DOA mismatch for the desired signal by using the presumed DOA $\bar{\theta}_s$, the center of the desired signal sector $\Theta$. In addition, due to the limited sampling grid, the performance of the proposed beamformer does not exactly converge to the optimal one when the input SNR is higher than 0 dB. In detail, the maximum estimation error of source DOAs is 0.05°, which is half of the grid increment of 0.1°. Such DOA estimation error will degrade the output performance of the proposed beamformer because both the reconstructed




(a) Output SINR versus input SNR

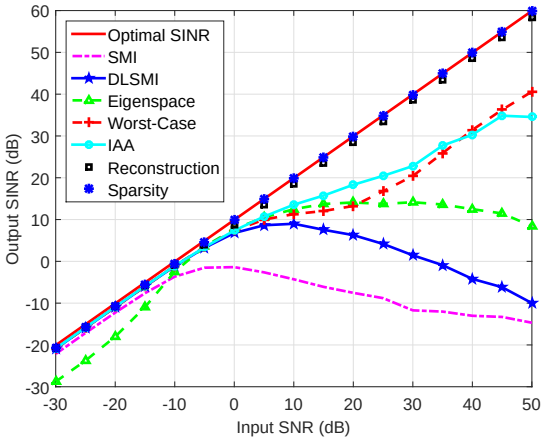



(b) Deviations from optimal SINR versus input SNR

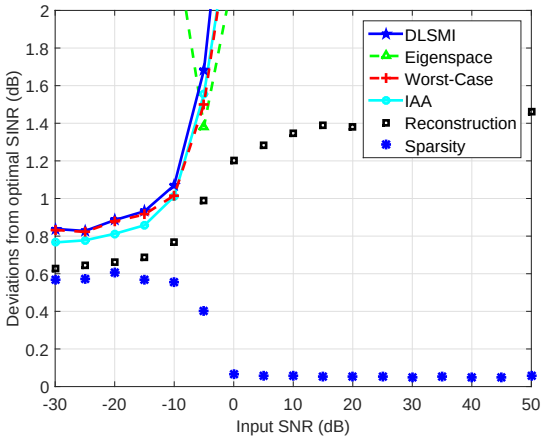



(c) Output SINR versus number of snapshots

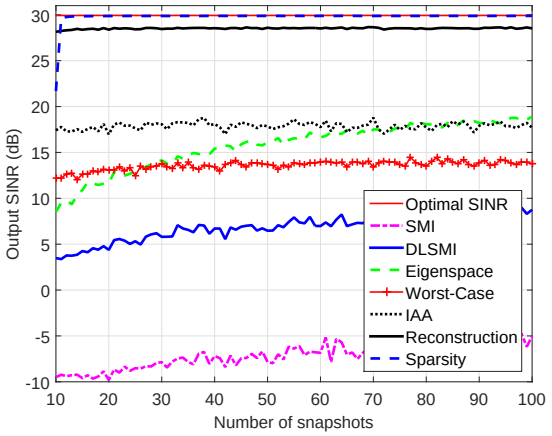
**Figure 8.5** Second example: Fixed signal DOA mismatch.




(a) Output SINR versus input SNR




(b) Deviations from optimal SINR versus input SNR




(c) Output SINR versus number of snapshots

**Figure 8.6** Third example: Random sources look direction mismatch.

interference-plus-noise covariance matrix $\hat{\boldsymbol{R}}_{i+n}$ and the modified signal steering vector $\tilde{\boldsymbol{a}}_s$ depend on the DOA estimation. The possible solutions to mitigate the effect of grid limitation include the grid refinement method [53] and the off-grid direction estimation method [61, 62]. In addition to achieving a faster convergence rate than the interference-plus-noise covariance matrix reconstruction-based beamformer [40], the proposed interference-plus-noise covariance matrix sparse reconstruction-based beamformer offers a stable output performance with the increase of SNR while others do not, as shown in Figure 8.6(c).

### 8.4.4 Example 4: Coherent local scattering

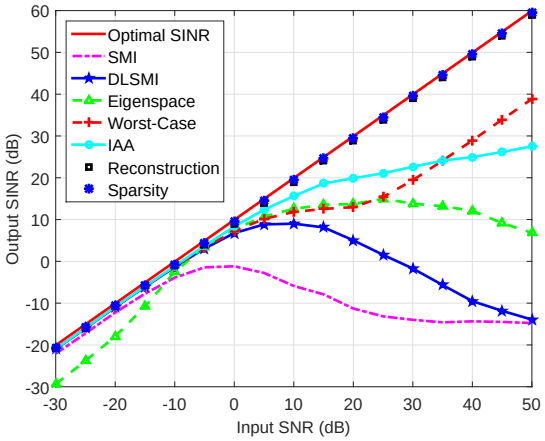
In the fourth example, we consider a scenario where the spatial signature of the desired signal is distorted by local scattering effects. Specifically, the desired signal is assumed to be a plane wave with the presumed steering vector $\bar{\boldsymbol{a}}_s$, whereas the actual steering vector $\boldsymbol{a}_s$ is formed as the superposition of five signal paths including four coherent scattered paths as

$$\boldsymbol{a}_s = \bar{\boldsymbol{a}}_s + \sum_{t=1}^{4} e^{j\psi_t} \boldsymbol{a}(\theta_t), \tag{8.36}$$

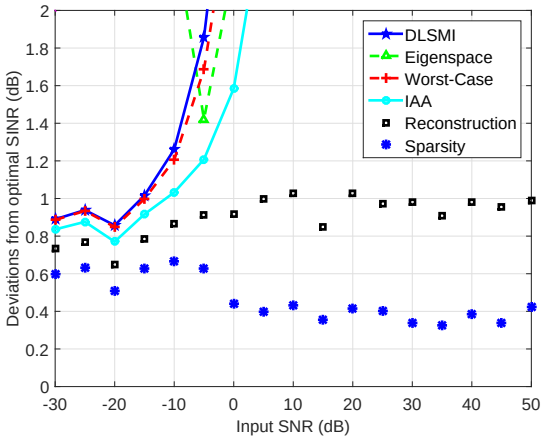
where $\boldsymbol{a}(\theta_t), t = 1, 2, 3, 4$, correspond to coherently scattered paths. The steering vector of the $t$-th path, $\boldsymbol{a}(\theta_t)$, can be modeled as a plane wave from the direction of $\theta_t$. The DOAs of scattered paths follow independent normal distribution $\theta_t \sim \mathcal{N}\left(\bar{\theta}_s, 4^\circ\right), t = 1, 2, 3, 4$, and the phases of scattered paths follow independent uniform distribution $\psi_t \sim \mathcal{U}\left[0, 2\pi\right), t = 1, 2, 3, 4$. Note that the tested adaptive beamformers are implemented in a block adaptive manner, which means that both $\theta_t$ and $\psi_t, t = 1, 2, 3, 4$ change from run to run but do not change from snapshot to snapshot. From Figure 8.7, the output performance loss of the proposed beamformer is less than 0.7 dB, which is much smaller than that suffered by the other tested beamformers.

### 8.4.5 Example 5: Wavefront distortion

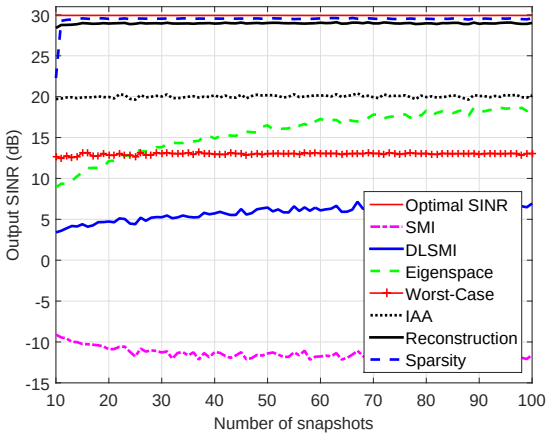
In the fifth example, we consider the situation where the desired signal spatial signature is distorted by wave propagation effects in an inhomogeneous medium. We assume independent-increment phase distortions of the desired signal wavefront [25, 63]. In each Monte Carlo run, each of these phase distortions is independently drawn from a Gaussian random generator $\mathcal{N}(0, 0.04)$, which remains fixed from snapshot to snapshot. From Figure 8.8, it is clear that the proposed beamformer provides more stable and near-optimal output performance than the other tested beamformers regardless of the input signal power or the number of snapshots.




(a) Deviations from optimal SINR versus SNR

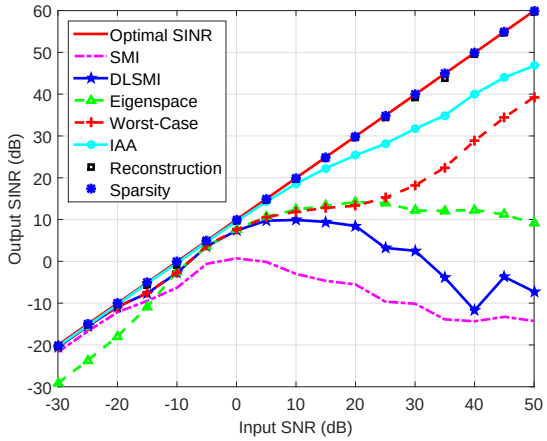



(b) Deviations from optimal SINR versus SNR

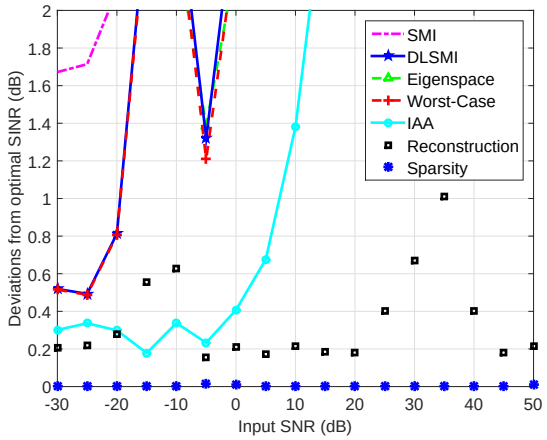



(c) Output SINR versus number of snapshots

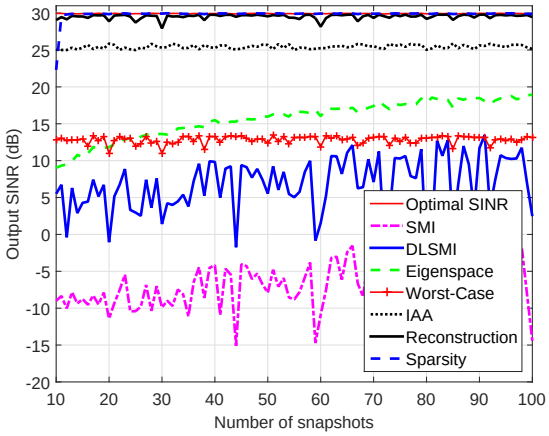
**Figure 8.7** Fourth example: Coherent local scattering.




(a) Deviations from optimal SINR versus SNR




(b) Deviations from optimal SINR versus SNR




(c) Output SINR versus number of snapshots

**Figure 8.8** Fifth example: Wavefront distortion.

### 8.4.6 Example 6: Incoherent local scattering

In the sixth example, we assume incoherent local scattering of the desired signal, which is common in array applications due to the multipath scattering effects caused by the presence of local scatters. In such a case, the desired signal is assumed to have a time-varying spatial signature as [32, 40]

$$\boldsymbol{a}_s(k) = s_0(k)\bar{\boldsymbol{a}}_s + \sum_{t=1}^{4} s_t(k)\boldsymbol{a}(\theta_t), \tag{8.37}$$

where $s_t(k) \sim \mathcal{N}(0,1), t = 0,1,2,3,4$, are independent and identically distributed (i.i.d.) zero-mean complex Gaussian random variables changing from snapshot to snapshot, $\theta_t \sim \mathcal{N}(\bar{\theta}_s, 4^\circ), t = 1,2,3,4$, are the random DOAs changing from run to run while remaining fixed from snapshot to snapshot. This corresponds to the case of incoherent local scattering [30], where the rank of the signal covariance matrix $\boldsymbol{R}_s$ is higher than one. In the general-rank case, the output SINR should be rewritten as [10]

$$\text{SINR} = \frac{\boldsymbol{w}^H \boldsymbol{R}_s \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w}}, \tag{8.38}$$

which is maximized by [10]

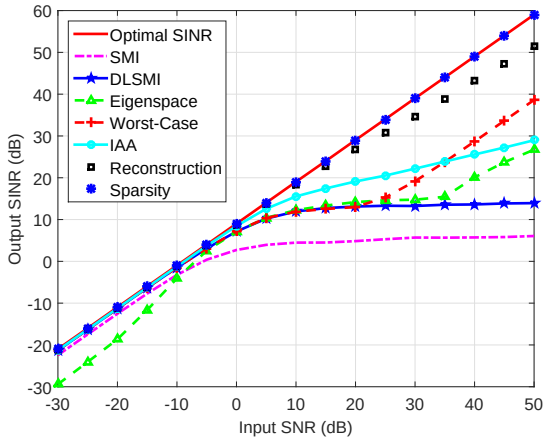
$$\boldsymbol{w} = \mathcal{P}\{\boldsymbol{R}_{i+n}^{-1} \boldsymbol{R}_s\}, \tag{8.39}$$

where $\mathcal{P}\{\cdot\}$ stands for the principal eigenvector of a matrix.

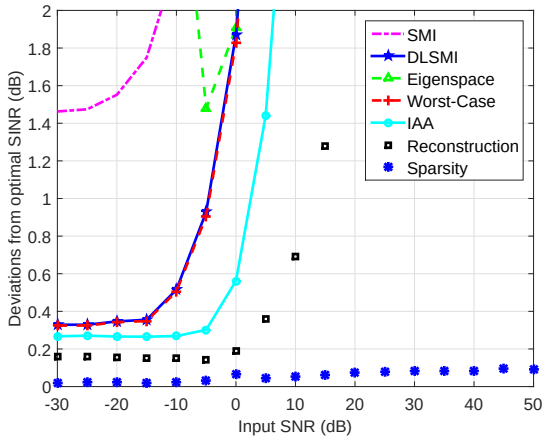
It can be seen from Figure 8.9(a) that the proposed beamformer outperforms all other tested beamformers especially at high SNR. The performance loss of the proposed beamformer is less than 0.1 dB. In contrast, there is about 7.5 dB performance loss for the interference-plus-noise covariance matrix reconstruction-based beamformer. The main reason is that the signal-of-interest leaks into the out-of-sector $\bar{\Theta}$ due to the incoherent local scattering, and then the reconstructed interference-plus-noise covariance matrix $\hat{\boldsymbol{R}}_{i+n}$ (8.20) is contaminated by the leaked desired signal component.

### 8.4.7 Discussion

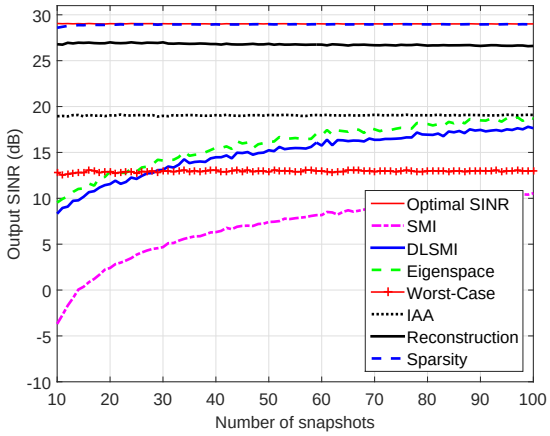
From the above extensive simulation results illustrated for different scenarios, it is clear that the proposed adaptive beamformer based on interference-plus-noise covariance matrix *sparse* reconstruction consistently enjoys the best performance as compared to other tested beamformers. More specifically, benefitted from the interference covariance matrix reconstruction which excludes the desired signal component, the output performance of the proposed beamformer is always close to or equal to the optimal SINR regardless of the input SNR. In contrast, there is a slight output performance degradation for the interference-plus-noise covariance matrix reconstruction-based beamformer because the Capon spatial spectrum estimator underestimates the power of interferers and thus decreases




(a) Deviations from optimal SINR versus SNR




(b) Deviations from optimal SINR versus SNR




(c) Output SINR versus number of snapshots

**Figure 8.9** Sixth example: Incoherent local scattering.

the estimation accuracy of the interference-plus-noise covariance matrix. On the other hand, other tested beamformers use the signal-contaminated covariance matrix, and thus degrade the output performance particularly when the input SNR is high. Another significant advantage of the proposed adaptive beamformer is that it achieves faster convergence, and only requires the number of snapshots to be slightly higher than the number of array sensors in order to approach the optimal SINR. In contrast, if an average performance loss of less than 3 dB is required for the SMI beamformer, the number of signal-free snapshots must be at least twice the number of array sensors [48]. However, in the underlying problem where the snapshots are contaminated by the desired signal, the convergence becomes much slower, i.e., a much higher number of snapshots are required [17].

Note that, although the ULA is adopted in the above simulations, the idea of covariance matrix sparse reconstruction proposed in this chapter for the adaptive beamformer design can be generalized to arbitrary arrays, e.g., the coprime array [64–67]. Benefitted from the larger array aperture due to the sparse deployment, the coprime array adaptive beamforming algorithm proposed in [67] achieves high robustness against model mismatches with a significant reduction in the number of antennas and the associated radio frequency chains.

It is also worth noting that the extension of the proposed narrowband adaptive beamforming technique to the broadband adaptive beamforming is straightforward [3]. For example, applying fast Fourier transform (FFT) to the broadband signal yields narrowband components, which can then be independently processed using the proposed (narrowband) covariance matrix sparse reconstruction-based adaptive beamforming technique. Subsequently, the time-domain broadband beamformer output is obtained by applying an inverse FFT to the output of the individual narrowband beamformers.

## 8.5 Conclusion

In this chapter, we proposed a simple, effective robust adaptive beamforming algorithm based on interference-plus-noise covariance matrix sparse reconstruction. Specifically, by exploiting the sparsity of sources distributed in the observed spatial domain, accurate interference-plus-noise covariance matrix reconstruction can be achieved by estimating the sparse spatial spectrum distribution from a sparsity-constrained covariance matrix fitting problem, which provides a signal-free interference-plus-noise covariance matrix for the beamformer design. The formulated sparsity-constrained covariance matrix fitting problem can be effectively solved with *a priori* information of the estimated source DOAs rather than $\ell_1$-norm relaxation-type approximations. Simulation results evidently demonstrate the effectiveness of the proposed algorithm. Compared to the existing techniques, the performance of the proposed method is nearly optimal over a wide range of input SNR and various error conditions. In addition, the proposed technique also has low computational complexity.

# References


[1]B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[2]H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, July 1996.

[3]H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY: John Wiley & Sons, 2002.

[4]J. Li and P. Stoica, Eds., *Robust Adaptive Beamforming*. New York, NY: John Wiley & Sons, 2005.

[5]W. Liu and S. Weiss, *Wideband Beamforming: Concepts and Techniques*, Chichester, U.K.: Wiley, 2010.

[6]J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[7]S. A. Vorobyov, "Principles of minimum variance robust adaptive beamforming design," *Signal Process.*, vol. 93, no. 12, pp. 3264–3277, Dec. 2013.

[8]K. Yang, T. Ohira, Y. Zhang, C.-Y. Chi, "Super-exponential blind adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1549–1563, June 2004.

[9]H. Cox, R. M. Zeskind, and M. H. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 35, no. 10, pp. 1365–1376, Oct. 1987.

[10]A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. J. Electron. Commun.*, vol. 53, no. 6, pp. 305–314, Dec. 1999.

[11]W. F. Gabriel, "Spectral analysis and adaptive array superresolution techniques," *Proc. IEEE*, vol. 68, no. 6, pp. 654–666, June 1980.

[12]Y. I. Abramovich, "Controlled method for adaptive optimization of filters using the criterion of maximum SNR," *Radio Eng. Electron. Phys.*, vol. 26, no. 3, pp. 87–95, Mar. 1981.

[13]B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, July 1988.

[14]L. Du, T. Yardibi, J. Li, and P. Stoica, "Review of user parameter-free robust adaptive beamforming algorithms," *Digital Signal Process.*, vol. 19, no. 4, pp. 567–582, July 2009.

[15]P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," *IEEE Trans. Signal Process.*, vol. 56, no.6, pp. 2598–2602, June 2008.

[16]L. Chang and C.-C. Yeh, "Performance of DMI and eigenspace-based beamformers," *IEEE Trans. Antennas Propagat.*, vol. 40, no. 11, pp. 1336–1347, Nov. 1992.

[17]D. D. Feldman and L. J. Griffiths, "A projection approach for robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 867–876, Apr. 1994.

[18]J. K. Thomas, L. L. Scharf, and D. W. Tufts, "The probability of a subspace swap in the SVD," *IEEE Trans. Signal Process.*, vol. 43, no. 3, pp. 730–736, Mar. 1995.

[19]M. Hawkes, A. Nehorai, and P. Stoica, "Performance breakdown of subspace-based methods: Prediction and cure," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Salt Lake City, UT, May 2001, pp. 4005–4008.

[20]T. Yardibi, J. Li, P. Stoica, M. Xue, and A. B. Baggeroer, "Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 425–443, Jan. 2010.

[21]L. C. Godara, "The effect of phase-shift errors on the performance of an antenna-array beamformer," *IEEE J. Ocean. Eng.*, vol. 10, no. 3, pp. 278–284, July 1985.

[22]J. W. Kim and C. K. Un, "An adaptive array robust to beam pointing error," *IEEE Trans. Signal Process.*, vol. 40, no. 6, pp. 1582–1584, June 1992.

[23]N. K. Jablon, "Adaptive beamforming with the generalized sidelobe canceller in the presence of array imperfections," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 8, pp. 996–1012, Aug. 1986.

[24]A. B. Gershman, V. I. Turchin, and V. A. Zverev, "Experimental resultsof localization of moving underwater signal by adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2249–2257, Oct. 1995.

[25]J. Ringelstein, A. B. Gershman, and J. F. Böhme, "Direction finding in random inhomogeneous media in the presence of multiplicative noise," *IEEE Signal Process. Lett.*, vol. 7, no. 10, pp. 269–272, Oct. 2000.

[26]Y. J. Hong, C.-C. Yeh, and D. R. Ucci, "The effect of a finite-distance signal source on a far-field steering applebaum array—two dimensional array case," *IEEE Trans. Antennas Propagat.*, vol. 36, no. 4, pp. 468–475, Apr. 1988.

[27]K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments," *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, pp. 437–447, Mar. 2000.

[28]J. Goldberg and H. Messer, "Inherent limitations in the localization of a coherently scattered source," *IEEE Trans. Signal Process.*, vol. 46, no. 12, pp. 3441–3444, Dec. 1998.

[29]D. Astely and B. Ottersten, "The effects of local scattering on direction of arrival estimation with MUSIC," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3220–3234, Dec. 1999.

[30]O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 1872–1882, July 2000.

[31]O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[32]S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[33]J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, July 2003.

[34]R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

[35]A. Hassanien, S. A. Vorobyov, and K. M. Wong, "Robust adaptive beamforming using sequential quadratic programming: An iterative solution to the mismatch problem," *IEEE Signal Process. Lett.*, vol. 15, pp. 733–736, Nov. 2008.

[36]A. Khabbazibasmenj, S. A. Vorobyov, and A. Hassanien, "Robust adaptive beamforming based on steering vector estimation with as little as possible prior information," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2974–2987, June 2012.

[37]S. A. Vorobyov, A. B. Gershman, Z.-Q. Luo, and N. Ma, "Adaptive beamforming with joint robustness against mismatched signal steering vector and interference nonstationarity," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 108–111, Feb. 2004.

[38]Y. J. Gu, W.-P. Zhu, and M.N.S. Swamy, "Adaptive beamforming with joint robustness against covariance matrix uncertainty and signal steering vector mismatch," *Electronics Lett.*, vol. 46, no. 1, pp. 86–88, Jan. 2010.

[39]Y. Gu and A. Leshem, "Robust adaptive beamforming based on jointly estimating covariance matrix and steering vector," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Prague, Czech Republic, May 2011, pp. 2640–2643.

[40]Y. Gu and A. Leshem, "Robust adaptive beamforming based on interference covariance matrix reconstruction and steering vector estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3881–3885, July 2012.

[41]L. Huang, J. Zhang, X. Xu, and Z. Ye, "Robust adaptive beamforming with a novel interference-plus-noise covariance matrix reconstruction method," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1643–1650, Apr. 2015.

[42]J. Yang, G. Liao, J. Li, Y. Lei, and X. Wang, "Robust beamforming with imprecise array geometry using steering vector estimation and interference covariance matrix reconstruction," *Multidim. Syst. Signal Process.*, vol. 28, no. 2, pp 451–469, Apr. 2017.

[43]Y. Gu, N. A. Goodman, S. Hong, and Y. Li, "Robust adaptive beamforming based on interference covariance matrix sparse reconstruction," *Signal Process.*, vol. 96, Part B, pp. 375–381, Mar. 2014.

[44]Y. Gu and Y. D. Zhang, "Single-snapshot adaptive beamforming," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Sheffield, UK, July 2018.

[45]Y. C. Eldar, A. Nehorai, and P. S. La Rosa, "An expected least-squares beamforming approach to signal estimation with steering vector uncertainties," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 288–291, May 2006.

[46]K. Kumatani, T. Gehrig, U. Mayer, E. Stoimenov, J. McDonough, and M. Wölfel, "Adaptive beamforming with a minimum mutual information criterion," *IEEE Trans. Audio Speech Lang. Process.*, vol. 15, no. 8, pp. 2527–2541, Nov. 2007.

[47]Y. Rong, Y. C. Eldar, and A. B. Gershman, "Performance tradeoffs among adaptive beamforming criteria," *IEEE J. Sel. Top. Signal Process.*, vol. 1, no. 4, pp. 651–659, Dec. 2007.

[48]I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 10, no. 6, pp. 853–863, Nov. 1974.

[49]L. Du, J. Li, and P. Stoica, "Fully automatic computation of diagonal loading levels for robust adaptive beamforming," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 449–458, Jan. 2010.

[50]R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[51]J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[52]A. J. Miller, *Subset Selection in Regression*. London, U.K.: Chapman and Hall, 2002.

[53]D. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[54]Y. C. Eldar and G. Kutyniok, Eds., *Compressed Sensing: Theory and Applications*. U.K., Cambridge: Cambridge Univ. Press, 2012.

[55]S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[56]R. Tibshirani. "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. B*, vol. 58, no. 1, pp. 267–288, 1996.

[57]R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.

[58]A. B. Gershman, M. Rubsamen, and M. Pesavento, "One- and two-dimensional direction-of-arrival estimation: An overview of search-free techniques," *Signal Process.*, vol. 90, no. 5, pp. 1338–1349, May 2010.

[59]K. Harmanci, J. Tabrikian, and J. L. Krolik, "Relationships between adaptive minimum variance beamforming and optimal source localization," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 1–12, Jan. 2000.

[60]M. Grant, S. Boyd, and Y. Y. Ye, "CVX: MATLAB software for disciplined convex programming," Dec. 2017. Available: http://cvxr.com/cvx/

[61]G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[62]Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1257–1269, Mar. 2016.

[63]O. Besson, F. Vincent, P. Stoica, and A. B. Gershman, "Approximate maximum likelihood estimators for array processing in multiplicative noise environments," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2506–2518, Sept. 2000.

[64]Y. Gu, C. Zhou, N. A. Goodman, W.-Z. Song, and Z. Shi, "Coprime array adaptive beamforming based on compressive sensing virtual array signal," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Shanghai, China, Mar. 2016, pp. 2981–2985.

[65]C. Zhou, Y. Gu, W.-Z. Song, Y. Xie, and Z. Shi, "Robust adaptive beamforming based on DOA support using decomposed coprime subarrays," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Shanghai, China, Mar. 2016, pp. 2986–2990.

[66]C. Zhou, Z. Shi, and Y. Gu, "Coprime array adaptive beamforming with enhanced degrees-of-freedom capability," in *Proc. IEEE Radar Conf.*, Seattle, WA, May 2017, pp. 1357–1361.

[67]C. Zhou, Y. Gu, S. He, and Z. Shi, "A robust and efficient algorithm for coprime array adaptive beamforming," *IEEE Trans. Veh. Tech.*, vol. 67, no. 2, pp. 1099–1112, Feb. 2018.