Journal of Astronomical Instrumentation

Robust Astronomical Imaging in the Presence of Radio Frequency Interference [1]

Shuimei Zhang, Yujie Gu, Yimin D. Zhang [2]
Department of Electrical and Computer Engineering
Temple University, Philadelphia, PA 19122, USA

Received (to be inserted by publisher); Revised (to be inserted by publisher); Accepted (to be inserted by publisher);

Radio astronomical observations are increasingly contaminated by radio frequency interference (RFI), rendering the development of effective RFI suppression techniques a pressing task. In practice, the existence of model mismatch makes the observing environment more challenging. In this paper, we develop a robust astronomical imaging method in the presence of RFI and model mismatch. The key contribution of the proposed method is the accurate estimation of the actual signal steering vector by maximizing the beamformer output power subject to a constraint that prevents the estimated steering vector from converging to the interference steering vectors. The proposed method is formulated as a quadratically constrained quadratic programming problem that can be solved using efficient numerical approaches. Simulation results demonstrate the effectiveness of the proposed method.

Keywords: Robust adaptive beamforming, astronomical imaging, radio interference, synthetic aperture.

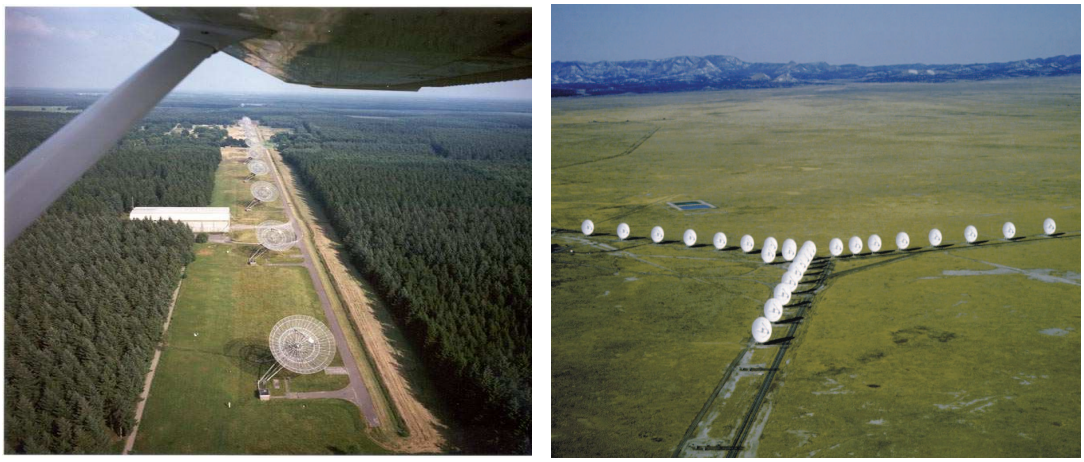
1. Introduction

Radio astronomy discoveries are important for humankind to understand the origin of the universe (Hall, 2013). Modern radio astronomical telescopes commonly use an array consisting of multiple antennas to construct astronomical images based on the principle of radio interferometry (Ryle, 1952). The Westerbork Synthesis Radio Telescope (WSRT) (Hogbom & Brouw, 1974) in the Netherlands (shown in Fig. 1(a)) and the Very Large Array (VLA) (Napier *et al.*, 1983) in New Mexico (shown in Fig. 1(b)) are two examples of such kind of arrays. As the Earth rotates, the array aperture observed at different time epochs can be used for aperture synthesis, thereby enabling high-resolution astronomical imaging. High-resolution imaging of radio astronomy signals plays a critical role in radio astronomy.

Unfortunately, radio astronomy observations are increasingly contaminated by radio frequency interference (RFI). With the development of technological innovation, the soaring demand for wireless broadband communications and other applications would require more spectrum resources. Moreover, various wireless communication devices employ different technologies and collectively span an extremely broad bandwidth. Certain protected frequency bands are slated for exclusive use by radio astronomy. However, radio telescopes are required to provide effective suppression of various RFI signals in the unprotected bands. In addition, capabilities of effective interference cancellation may enable sharing the protected bands slated for radio astronomy with wireless communications and broadcast services in the future.

Compared with active communication services, astronomical signals are extremely weak, typically tens of decibels below the noise floor. As such, astronomical signals are highly vulnerable to RFI. Protection of radio astronomy observations against RFI is typically offered in three different layers. The first layer is called pre-observation, which performs, for example, spectrum management and coordination with active sources (van Driel, 2009). The second layer is during observation, in which the interference may be detected and removed from the data in real time (Raza *et al.*, 2002). The third layer is after observation, which

[1]This work is supported in part by the National Science Foundation (NSF) under grant AST-1547420.
[2]Corresponding author.



(a) WSRT (b) VLA

Fig. 1. The WSRT and the VLA.

mainly concerns the image formation process (Leshem & Van der Veen, 2000; Zhang *et al.*, 2017). In this paper, we focus on the third layer.

Radio astronomical images are traditionally derived from a delay-and-sum (DAS) beamformer, which is data-independent and thus lacks the effective interference suppression capability. To improve the capability to suppress the interference, adaptive beamforming techniques have been adopted for radio astronomical imaging. Among them, minimum variance distortionless response (MVDR) (Van der Veen *et al.*, 2004) and adaptive angular response (AAR) (Ben-David & Leshem, 2008) are two commonly used techniques in recent radio astronomical imaging. However, these adaptive beamforming techniques are sensitive to model mismatch, thereby degrading interference suppression capability and imaging quality in the presence of model impairments.

In practice, while radio telescope systems are carefully designed and operated, they inevitably suffer from model mismatch, particularly as the antennas are continuously adjusted with the rotation of the Earth to steer their beams towards the astronomical sources of interest. Hence, each antenna has its own directional response. Moreover, the propagation of astronomical signals through the atmosphere and ionosphere causes additional amplitude gain error and phase delays to the array with a very long baseline (Wijnholds *et al.*, 2010). With the development of modern astronomy, a lot of new generation radio astronomy telescopes are being deployed, such as the Low Frequency Array (LOFAR) and the Square Kilometer Array (SKA). They will have a much higher sensitivity than before, implying that they would receive more interfering signals, and be more severely affected by ionospheric and tropospheric turbulence properly accounting for low-frequency measurements. Such model mismatch, if uncompensated, would degrade the imaging resolution and quality. In such a case, it is more difficult to suppress the RFI. Therefore, array calibration and robust beamforming against model mismatch in radio astronomy have been important objectives in the past decades (Van der Tol & van der Veen, 2005; Ollier *et al.*, 2017).

More specifically, a number of robust adaptive beamforming methods have been proposed to address the model mismatch problem. For example, diagonal loading (Cox *et al.*, 1987; Carlson, 1988; Li *et al.*, 2003) is a widely used robust adaptive beamforming technique, which adds a scaled identity matrix to the sample correlation matrix to reduce the conditional number. However, it is difficult to choose an optimal loading factor in different scenarios. Eigenspace-based beamformers (Chang & Yeh, 1992; Feldman & Griffiths, 1994) may cope with arbitrary steering vector mismatch. The key idea behind this technique is to project 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. In addition, when the signal-to-noise ratio (SNR) is low, the approach suffers severe performance degradation due to the subspace swap, i.e., the data can be better represented by some components of the noise subspace than by some components of the signal subspace. The worst-case performance based optimization technique (Vorobyov *et al.*, 2003; Li *et al.*,

2004; Lorenz & Boyd, 2005) makes use of the uncertain set of the signal steering vector. In this method, the upper bound of the norm of the mismatch vector is the required *a priori*, and the determination of this upper error bound has to be made in an *ad hoc* manner. Either overestimation or underestimation of the upper error bound would compromise the performance.

In order to address the problems encountered by these existing techniques as described above, in this paper, we propose a new method that provides robust radio astronomical imaging in the presence of RFI and model mismatch. Unlike the existing techniques in which beamforming is performed to achieve a desirable output signal-to-interference-plus-noise ratio (SINR), this paper considers beamforming from imaging perspective. For radio astronomical imaging, the image intensity is the output power of the beamformer. To obtain this image, adaptive beamforming is performed point by point. Inspired by the adaptive beamforming method in Gu's paper (Gu & Leshem, 2012), we formulate the robust beamforming in the presence of model mismatch as a quadratically constrained quadratic programming (QCQP) problem. Accurate estimation of the actual signal steering vectors corresponding to directions within the region of interest is achieved by maximizing the beamformer output power subject to a constraint that prevents the estimated steering vector from converging to the interference steering vectors. We derive an effective numerical solution based on the Karush-Kuhn-Tucker (KKT) conditions, which reduces the computational complexity and provides more insights into the problem. Simulation results evidently demonstrate the effectiveness of the proposed robust astronomical imaging technique compared to the state-of-the-art astronomical imaging techniques.

The rest of this paper is organized as follows. In Section 2, the signal model is formulated. In Section 3, after presenting a brief overview of the existing approaches to astronomical imaging, we describe the proposed robust astronomical imaging method and the numerical solution. Simulation results are presented in Section 4 and conclusions are drawn in Section 5.

Notations: Lower-case (upper-case) bold characters are used to denote vectors (matrices). $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix. $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote the transpose and the Hermitian transpose, respectively. Diag$(\cdot)$ denotes a diagonal matrix with the elements of a vector constituting the diagonal entries, whereas diag$(\cdot)$ denotes a vector consisting of the diagonal elements of a general matrix. $\odot$ is the Hadamard (or element-wise) product.

2. Signal Model

2.1. *Background*

In radio astronomy, as shown in Fig. 2, two coordinate systems are utilized for the telescope array and the astronomical objects, respectively. In the coordinate system described by (u, v, w), the vector baseline $\boldsymbol{d}_\lambda$ connects a pair of antennas and is measured in wavelength. The component w is measured in the direction of $\boldsymbol{s}_0$, which serves as the phase reference position. Components u and v are measured in a plane that is normal to the direction of the phase reference position. The other coordinate system is described by (ℓ, m) and corresponds to the projection of the celestial sphere onto a plane that is a tangent at the field center and is measured in radians (Thompson *et al.*, 2017). It is clear that, for an array with P sensors, there are $P(P-1)/2$ independent baselines. Note that the baselines change over time as the Earth rotates.

By discretizing the specific observation area of interest and assuming D astronomical point sources therein, the source intensity is expressed as

$$I(\ell, m) = \sum_{d=1}^{D} I(\ell_d, m_d)\delta(\ell - \ell_d)\delta(m - m_d), \tag{1}$$

where $I(\ell_d, m_d), d = 1, \cdots, D$, denotes the source intensity at position (ℓ_d, m_d), and $\delta(\cdot)$ is the Dirac delta function.

2.2. *Visibility matrix formulation*

According to the principle of radio interferometry, radio astronomical imaging is based on the "visibility" $V(u_{ij}^k, v_{ij}^k)$, where u_{ij}^k and v_{ij}^k stand for the two components of the baseline in the coordinate system

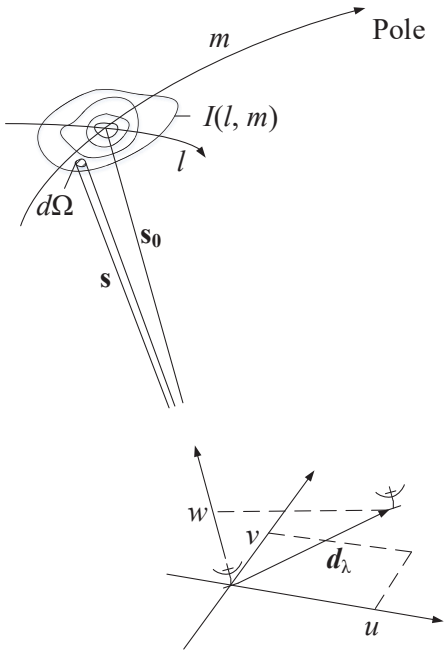


Fig. 2. Geometric relationship between a source under observation and an interferometer (one antenna pair of an array).

corresponding to antenna pairs i and j at time epoch t_k. The geometric delay is assumed to be compensated by an adjustable instrumental delay, so that the reference source position appear as if it were at the North Pole. As a result of such compensation, the baseline components vary over time as the Earth rotates.

After compensation, at the reference source position, the propagation path difference is 0. For a wavefront incident from direction (ℓ_d, m_d) at time epoch t_k, there is a difference of $(u_{ij}^k \ell_d + v_{ij}^k m_d)$ wavelengths in the path length corresponding to antenna pair i and j, yielding a phase difference in the correlator output (Thompson *et al.*, 2017). Under certain standard assumptions, such as planar arrays, small field of view imaging, and uncorrelated source points, the visibility and the astronomical source intensity are associated with the two-dimensional Fourier transform as (Levanda & Leshem, 2010)

$$V(u_{ij}^k, v_{ij}^k) = \sum_{d=1}^{D} I(\ell_d, m_d) e^{-\jmath 2\pi (u_{ij}^k \ell_d + v_{ij}^k m_d)}, \tag{2}$$

where $\jmath = \sqrt{-1}$ is the imaginary unit. Selecting one antenna as the reference point and denoting its coordinate as (u_0^k, v_0^k), we have

$$V(u_{ij}^k, v_{ij}^k) = \sum_{d=1}^{D} e^{-\jmath 2\pi (u_{i,0}^k \ell_d + v_{i,0}^k m_d)} I(\ell_d, m_d) e^{\jmath 2\pi (u_{j,0}^k \ell_d + v_{j,0}^k m_d)}, \tag{3}$$

where $u_{i,0}^k - u_{j,0}^k = u_{ij}^k$ and $v_{i,0}^k - v_{j,0}^k = v_{ij}^k$. The visibility depicts the correlation between different telescope antenna pairs. Define the (i,j)-th entry of the measurement correlation matrix at time epoch t_k as

$$(\boldsymbol{R}_k)_{ij} = V(u_{ij}^k, v_{ij}^k). \tag{4}$$

Then, from (3), the correlation matrix becomes

$$\boldsymbol{R}_k = \overline{\boldsymbol{A}}_k \boldsymbol{B} \overline{\boldsymbol{A}}_k^{\mathrm{H}}, \tag{5}$$

where $\overline{\boldsymbol{A}}_k = [\bar{\boldsymbol{a}}_k(\ell_1, m_1), \cdots, \bar{\boldsymbol{a}}_k(\ell_D, m_D)]$ denotes the telescope array manifold matrix at time epoch t_k due to the geometrical delays associated with the array configuration and the D source positions. Denoting P as the number of antennas in the telescope array, the d-th column of matrix $\overline{\boldsymbol{A}}_k$ is the array steering vector associated with the d-th astronomical source, located at (ℓ_d, m_d), and is expressed as

$$\bar{\boldsymbol{a}}_k(\ell_d, m_d) = \left[e^{-\jmath 2\pi (u_1^k \ell_d + v_1^k m_d)}, \cdots, e^{-\jmath 2\pi (u_P^k \ell_d + v_P^k m_d)}\right]^{\mathrm{T}}. \tag{6}$$

In addition, $\boldsymbol{B} = \text{Diag}\left[I(\ell_1, m_1), \cdots, I(\ell_D, m_D)\right]$ is a diagonal matrix representing the intensity of all point sources. Note that $\boldsymbol{B}$ can be treated as time-invariant because the effect of their variations is insignificant during an observation (Zhang *et al.*, 2018a,b).

For presentational convenience, our discussion will focus on the narrowband signal model. Wideband signals are typically decomposed into multiple frequency bins. In this case, the correlation matrices can be computed for each frequency bin separately (Levanda & Leshem, 2010).

2.3. *Problem formulation in the presence of interference and model mismatch*

Considering Q interferers, the correlation matrix of the interference signals at time epoch t_k is expressed as

$$\boldsymbol{R}_k^i = \boldsymbol{A}_k^i \boldsymbol{B}_k^i (\boldsymbol{A}_k^i)^{\mathrm{H}}, \tag{7}$$

where the superscript i represents the interference, $\boldsymbol{B}_k^i$ is a matrix representing Q interfering powers on the diagonal, and $\boldsymbol{A}_k^i = \left[\boldsymbol{a}_{1k}^i, \cdots, \boldsymbol{a}_{Qk}^i\right]$ is the array response matrix corresponding with $\boldsymbol{a}_{qk}^i$. Here, $\boldsymbol{a}_{qk}^i$ is the spatial signature of the q-th interferer at time epoch t_k, $q = 1, \cdots, Q$ and can be formulated as

$$\boldsymbol{a}_{qk}^i = \begin{bmatrix} \beta_{q1} e^{-\jmath 2\pi f_q \tau_{q1}} \\ \vdots \\ \beta_{qP} e^{-\jmath 2\pi f_q \tau_{qP}} \end{bmatrix}. \tag{8}$$

where β_{qp} denotes the p-th antenna gain in the direction of the q-th interfer. Unlike $\boldsymbol{a}_k(l_d, m_d)$ which has a structured array manifold and is known, $\boldsymbol{a}_{qk}^i$ is not a structured function of the interferer direction (Leshem *et al.*, 2000). Because the interferers are usually located in the near-field and interfere the telescope antennas mainly through the array sidelobes. The array responses to such interference signals are generally time-varying because of the motion of the interfering sources, changes in the propagation environment, and the time-varying geometrical delay compensations applied to the telescope antennas. As the Earth rotates, the sidelobe levels may vary over time as the antennas adjust their main beams towards the region of interest. Also, it is noted that RFI cannot be averaged out during the correlation process since it has a certain directivity (Leshem & Van der Veen, 2000).

In addition to RFI, model mismatch is another type of source that leads to distortions in the array response over different time epochs. While model mismatch alone may compromise the imaging performance, its effect is more pronounced in the presence of RFI because existing adaptive beamformers are sensitive to such model mismatch.

Generally speaking, there are two kinds of model mismatch. The first kind of model mismatch results from the instrumental effects. Each antenna has its own directional response and the primary beam of the individual antennas typically lacks the adequate level of accuracy. In the receiver response, the amplitude and phase can be affected by environmental conditions of the receiver chain. The other kind of model mismatch is caused by propagation effects. Ionospheric and tropospheric turbulence cause time-varying refraction and diffraction, which has profound effects on the propagation of radio waves.

For the first kind of model mismatch, the instrumental effects are characterized by an unknown complex gain matrix $\boldsymbol{G}_k^a = \text{Diag}[g_{1,k}^a, \cdots, g_{P,k}^a]$. For the d-th source at time epoch t_k, the actual array response becomes $\boldsymbol{G}_k^a \bar{\mathbf{a}}_k(\ell_d, m_d)$, where $\bar{\mathbf{a}}_k(\ell_d, m_d)$ is the corresponding presumed steering vector. The difference between the presumed signal steering vector and the actual signal steering vector reflects the signal model mismatch due to imperfect calibration of the array response and the propagation channel conditions. In this case, taking both steering vector mismatch and interference into account, the measured visibility matrix becomes (Leshem *et al.*, 2000; Wijnholds *et al.*, 2010)

$$\boldsymbol{R}_k = \left(\boldsymbol{G}_k^a \overline{\boldsymbol{A}}_k\right) \boldsymbol{B} \left(\boldsymbol{G}_k^a \overline{\boldsymbol{A}}_k\right)^{\mathrm{H}} + \check{\boldsymbol{R}}_k^i, \tag{9}$$

where $\overline{A}_k = [\overline{a}_k(\ell_1, m_1), \cdots, \bar{\boldsymbol{a}}_k(\ell_D, m_D)]$ is the presumed array response matrix, and $\check{\boldsymbol{R}}_k^i$ denotes the correlation matrix due to interference and model mismatch.

For the second kind of model mismatch, we define an unknown complex gain matrix $\boldsymbol{G}_k^s = \left[\boldsymbol{g}_{1,k}^s, \cdots, \boldsymbol{g}_{D,k}^s\right]$ to characterize the propagation effects. Unlike $\boldsymbol{G}_k^a$, which is diagonal, $\boldsymbol{G}_k^s$ is a full matrix. Correspondingly, for the d-th source at time epoch t_k, the actual array response becomes $\boldsymbol{g}_{d,k}^s \odot \bar{\mathbf{a}}_k(\ell_d, m_d)$. The measured visibility matrix can be expressed by

$$\boldsymbol{R}_k = \left(\boldsymbol{G}_k^s \odot \overline{\boldsymbol{A}}_k\right) \boldsymbol{B} \left(\boldsymbol{G}_k^s \odot \overline{\boldsymbol{A}}_k\right)^{\mathrm{H}} + \check{\boldsymbol{R}}_k^i. \tag{10}$$

It is clear that (9) can be seen as a special case of (10) when $\boldsymbol{g}_{1,k}^s = \boldsymbol{g}_{2,k}^s = \cdots = \boldsymbol{g}_{D,k}^s$. Therefore, the actual measured visibility matrix can be formulated as the following universal model,

$$\boldsymbol{R}_k = \left(\boldsymbol{G}_k \odot \overline{\boldsymbol{A}}_k\right) \boldsymbol{B} \left(\boldsymbol{G}_k \odot \overline{\boldsymbol{A}}_k\right)^{\mathrm{H}} + \check{\boldsymbol{R}}_k^i, \tag{11}$$

where $\boldsymbol{G}_k$ is an unknown full matrix which represents the instrumental effects, propagation effects or their combination. The actual array manifold matrix $\boldsymbol{G}_k \odot \overline{\boldsymbol{A}}_k$ can be decomposed into two terms, the presumed manifold matrix, $\overline{\boldsymbol{A}}_k$, and the model mismatch error in the manifold matrix, $\boldsymbol{E}_k$, i.e.,

$$\boldsymbol{R}_k = \left(\overline{\boldsymbol{A}}_k + \boldsymbol{E}_k\right) \boldsymbol{B} \left(\overline{\boldsymbol{A}}_k + \boldsymbol{E}_k\right)^{\mathrm{H}} + \check{\boldsymbol{R}}_k^i. \tag{12}$$

We denote $\boldsymbol{E}_k = [\boldsymbol{e}_{1,k}, \cdots, \boldsymbol{e}_{D,k}]$ with $\boldsymbol{e}_{d,k} = \boldsymbol{a}_{d,k} - \bar{\boldsymbol{a}}_{d,k}$ denoting the mismatch vector for the d-th source at time epoch t_k.

In the real word scenario, the array received signals are contaminated by the additive system noise. Assume that the received signals are corrupted with independent and identically distributed (i.i.d.) zero-mean additive white Gaussian noise which is introduced after the varying receiver gains (Leshem *et al.*, 2000). The visibility matrix at time epoch t_k becomes

$$\boldsymbol{R}_k = \left(\overline{\boldsymbol{A}}_k + \boldsymbol{E}_k\right) \boldsymbol{B} \left(\overline{\boldsymbol{A}}_k + \boldsymbol{E}_k\right)^{\mathrm{H}} + \check{\boldsymbol{R}}_k^i + \sigma_n^2 \boldsymbol{I}_P, \tag{13}$$

where σ_n^2 is the variance of the noise at each antenna.

In practice, the true correlation matrix $\boldsymbol{R}_k$ is not available and is estimated based on the finite samples of the data vector. Let $\boldsymbol{x}_k(n)$ denote the telescope's output signal at time epoch t_k. Then the covariance matrix $\boldsymbol{R}_k$ can be estimated by its maximum likelihood estimate, i.e., the sampled correlation matrix $\hat{\boldsymbol{R}}_k$, which is calculated as

$$\hat{\boldsymbol{R}}_k = \frac{1}{N} \sum_{n=(k-1)N}^{kN-1} \boldsymbol{x}_k(n) \boldsymbol{x}_k^{\mathrm{H}}(n), \tag{14}$$

where N is the number of samples used in a short-term integration.

Based on the imaging technique being used, model mismatch generally yields blurring in the obtained astronomical image. In the presence of RFI, such model mismatch would degrade the capability of effective RFI suppression, thus may generate false images or even obscure the true astronomical sources in the obtained images. Therefore, it is necessary to develop robust astronomical imaging methods to effectively mitigate RFI in the presence of model mismatch.

3. Robust Astronomical Imaging

In this section, we first briefly review the classic radio astronomical imaging methods. Then, the proposed method that provides robust RFI suppression and astronomical imaging in the presence of model mismatch is presented.

3.1. *Classical image formation methods*

The DAS beamformer is a classical beamformer, which estimates the source intensity at all positions within the interested region Ω as (Levanda & Leshem, 2010)

$$I^{\mathrm{DAS}}(\ell, m) = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{a}_k^{\mathrm{H}}(\ell, m) \hat{\boldsymbol{R}}_k \boldsymbol{a}_k(\ell, m), \forall (\ell, m) \in \Omega, \tag{15}$$

where K denotes the number of time epochs. For the DAS beamformer, its weight vector matches the signal steering vector $\boldsymbol{a}_k(\ell, m)$ and thus only depends on the array response matrix. That is, it is a data-independent beamformer, whose weights do not change as the array received data varies. Clearly, the DAS beamformer does not have the interference cancellation capability and its performance significantly degrades in the presence of RFI.

In order to improve the imaging quality and the robustness against the interference, several adaptive beamforming techniques have been developed for radio astronomical imaging. Among them, the MVDR beamformer (Van der Veen *et al.*, 2004) and AAR beamfomer (Ben-David & Leshem, 2008) are popular. The MVDR beamfomer is given as

$$I^{\mathrm{MVDR}}(\ell, m) = \sum_{k=1}^{K} \frac{1}{\boldsymbol{a}_k^{\mathrm{H}}(\ell, m)\hat{\boldsymbol{R}}_k^{-1}\boldsymbol{a}_k(\ell, m)}, \tag{16}$$

which is also known as the Capon beamformer. Compared to DAS, MVDR is a data-dependent beamformer, which utilizes the data to form the beam. The signals from directions other than that of the targeted astronomical signal are suppressed as much as possible. However, in the MVDR dirty image, the beam is spatially varying. In other words, the output noise power is not spatially uniform. To address this issue, the AAR beformer

$$I^{\mathrm{AAR}}(\ell, m) = \frac{\sum_{k=1}^{K} \boldsymbol{a}_k^{\mathrm{H}}(\ell, m)\hat{\boldsymbol{R}}_k^{-1}\boldsymbol{a}_k(\ell, m)}{\sum_{k=1}^{K} \boldsymbol{a}_k^{\mathrm{H}}(\ell, m)\hat{\boldsymbol{R}}_k^{-2}\boldsymbol{a}_k(\ell, m)}, \tag{17}$$

is proposed as a variant of MVDR beamformer. Compared to MVDR, AAR has isotropic white noise response, and provides optimal suppression of interference under the white noise constraint. Note that, DAS, MVDR and AAR all suffer performance degradation in the presence of model mismatch.

3.2. *Robust image formation method*

In Section 2, we learned that the model mismatch due to antenna perturbations will cause errors in the array response matrix. Such errors, if uncompensated, will result in ineffective jammer suppression, self-nulling of the astronomical signal, and image blurring, as will be demonstrated using simulation examples in Section 4.

In this subsection, we develop a robust beamforming technique by estimating the actual array response in the presence of model mismatch. This approach is applied to both instrumental effects and propagation effects.

Consider the source $(\ell, m) \in \Omega$ at time epoch t_k. The intensity of source located at (ℓ, m) is the output power of the beamformer. Because beamforming is performed for each point and at each time epoch, (ℓ, m) and k are omitted in the sequel for notational simplicity. The actual signal steering vector $\boldsymbol{a}$ can be estimated by maximizing the MVDR beamformer output power as

$$\max_{\boldsymbol{a}} \quad \frac{1}{\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}}, \tag{18}$$

which is equivalent to

$$\min_{\boldsymbol{a}} \quad \boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}. \tag{19}$$

Without additional restrictions, $\boldsymbol{a} = \boldsymbol{0}$ becomes the apparent best undesired solution, because $\hat{\boldsymbol{R}} \succeq 0$. To avoid this trivial solution, the presumed signal steering vector $\bar{\boldsymbol{a}}$ is utilized. Recall that in eq. (12), we define the mismatch vector as $\boldsymbol{e} = \boldsymbol{a} - \bar{\boldsymbol{a}}$, and $\boldsymbol{a} = \bar{\boldsymbol{a}} + \boldsymbol{e}$. Note that the desired signal steering vector $\boldsymbol{a}$ should be prevented from getting close to any interference region (Gu & Leshem, 2012). Here, for source

(ℓ, m), other points of the image are also considered as the interference. Then, the optimization problem of estimating the actual steering vector $\boldsymbol{a}$ can be transformed into estimating the mismatch vector $\boldsymbol{e}$ as follows:

$$\begin{aligned} \min_{\boldsymbol{e}} \ & (\bar{\boldsymbol{a}}+\boldsymbol{e})^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}(\bar{\boldsymbol{a}}+\boldsymbol{e}) \\ \text{s.t.} \ & (\bar{\boldsymbol{a}}+\boldsymbol{e})^{\mathrm{H}}\hat{\boldsymbol{R}}_{j+n}(\bar{\boldsymbol{a}}+\boldsymbol{e}) \leq \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}_{j+n}\bar{\boldsymbol{a}}, \end{aligned} \tag{20}$$

where the inequality constraint is introduced to ensure that the output interference-plus-noise power does not increase as a result of the mismatch vector adjustment. As such, it guarantee that the estimated signal steering vector does not converge to any interferer steering vector (Gu & Leshem, 2012). However, because the desired astronomical signal is very weak as compared to the interference, we can use $\hat{\boldsymbol{R}}$ to approximate $\hat{\boldsymbol{R}}_{j+n}$ without performance loss. In this case, (20) becomes:

$$\begin{aligned} \min_{\boldsymbol{e}} \ & (\bar{\boldsymbol{a}}+\boldsymbol{e})^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}(\bar{\boldsymbol{a}}+\boldsymbol{e}) \\ \text{s.t.} \ & (\bar{\boldsymbol{a}}+\boldsymbol{e})^{\mathrm{H}}\hat{\boldsymbol{R}}(\bar{\boldsymbol{a}}+\boldsymbol{e}) \leq \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}}. \end{aligned} \tag{21}$$

In order to exclude the trivial solution $\boldsymbol{e} = -\bar{\boldsymbol{a}}$, the mismatch vector $\boldsymbol{e}$ can be decomposed into $\boldsymbol{e}_{\parallel}$ and $\boldsymbol{e}_{\perp}$, which respectively represent the orthogonal component and parallel component of the presumed steering vector $\bar{\boldsymbol{a}}$. Since $\boldsymbol{e}_{\parallel}$ is a scaled copy of the presumed steering vector $\bar{\boldsymbol{a}}$, it will not affect the output SINR. Hence, by estimating only the orthogonal component $\boldsymbol{e}_{\perp}$, we can retain the optimal beamforming performance without falling into the undesired trivial solution. In this case, instead of estimating $\boldsymbol{e}$, the optimization problem (21) can be reformulated as follows:

$$\begin{aligned} \min_{\boldsymbol{e}_{\perp}} \ & (\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp})^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}(\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp}) \\ \text{s.t.} \ & \bar{\boldsymbol{a}}^{\mathrm{H}}\boldsymbol{e}_{\perp} = 0, \\ & (\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp})^{\mathrm{H}}\hat{\boldsymbol{R}}(\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp}) \leq \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}}, \end{aligned} \tag{22}$$

which is a QCQP problem. It can be readily solved by using convex optimization software, e.g., CVX (Grant *et al.*, 2008), which requires $\mathcal{O}(P^{3.5})$ flops (Gu & Leshem, 2012). However, the solution we will present below only requires $\mathcal{O}(P^3)$ flops. Also, more insights into choosing the parameters of the adaptive beamformer will be provided.

Based on the KKT conditions, the Lagrange function of (22) can be written as

$$L(\boldsymbol{e}_{\perp}, \nu, \mu) = (\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp})^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}(\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp}) + \nu\bar{\boldsymbol{a}}^{\mathrm{H}}\boldsymbol{e}_{\perp} + \mu\left[(\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp})^{\mathrm{H}}\hat{\boldsymbol{R}}(\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp}) - \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}}\right], \tag{23}$$

where ν and μ are real-valued Lagrange multipliers. ν is arbitrary and $\mu \geqslant 0$. As a consequence, $\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}} \succeq 0$. Evidently we have $L(\boldsymbol{e}_{\perp}, \nu, \mu) \leq (\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp})^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}(\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp})$ for any $\boldsymbol{e}_{\perp} \in \bar{S}$ with equality on the boundary of $\bar{S}$, where the set $\bar{S}$ is defined by the inequality constraint in (22).

Differentiating (23) with respect to $\boldsymbol{e}_{\perp}$ and equating the result to zero, we obtain the following equation:

$$\frac{\partial L}{\partial \boldsymbol{e}_{\perp}} = 2\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)(\bar{\boldsymbol{a}}+\boldsymbol{e}_{\perp}) + \nu\bar{\boldsymbol{a}} = 0. \tag{24}$$

Then, the orthogonal component of the mismatched vector, $\boldsymbol{e}_{\perp}$, is obtained as

$$\boldsymbol{e}_{\perp} = -\bar{\boldsymbol{a}} - \frac{\nu}{2}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}. \tag{25}$$

Hence, the steering vector is estimated by

$$\boldsymbol{a} = \bar{\boldsymbol{a}} + \boldsymbol{e}_{\perp} = -\frac{\nu}{2}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}. \tag{26}$$

Note that the two Lagrange multipliers ν and μ are still unknown.

Substituting (25) into the derivative of (23) with respect to ν, and equating the result to zero, we obtain the following equation:

$$\frac{\partial L}{\partial \nu} = \bar{\boldsymbol{a}}^{\mathrm{H}}\boldsymbol{e}_{\perp} = -\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}} - \frac{\nu}{2}\bar{\boldsymbol{a}}^{\mathrm{H}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}} = 0. \tag{27}$$

The Lagrange multiplier ν is obtained as

$$\nu = -2\frac{\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}}}{\bar{\boldsymbol{a}}^{\mathrm{H}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}}. \tag{28}$$

Substituting ν (28) back into (25) and (26), we have

$$\boldsymbol{e}_{\perp} = -\bar{\boldsymbol{a}} + \frac{\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}}}{\bar{\boldsymbol{a}}^{\mathrm{H}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}, \tag{29}$$

and

$$\boldsymbol{a} = \bar{\boldsymbol{a}} + \boldsymbol{e}_{\perp} = \frac{\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}}}{\bar{\boldsymbol{a}}^{\mathrm{H}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}. \tag{30}$$

Thus, the Lagrange multiplier μ is still unknown.

With the estimated steering vector $\bar{\boldsymbol{a}} + \boldsymbol{e}_{\perp}$ (30), we further differentiate (23) with respect to μ and equate the result to zero, resulting in

$$\begin{aligned}\frac{\partial L}{\partial \mu} &= (\bar{\boldsymbol{a}} + \boldsymbol{e}_{\perp})^{\mathrm{H}}\hat{\boldsymbol{R}}(\bar{\boldsymbol{a}} + \boldsymbol{e}_{\perp}) - \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}} \\ &= \left[\frac{\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}}}{\bar{\boldsymbol{a}}^{\mathrm{H}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}}}\right]^2 \bar{\boldsymbol{a}}^{\mathrm{H}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\hat{\boldsymbol{R}}\left(\hat{\boldsymbol{R}}^{-1} + \mu\hat{\boldsymbol{R}}\right)^{-1}\bar{\boldsymbol{a}} - \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}} \\ &= 0.\end{aligned} \tag{31}$$

Perform the eigen-decomposition on $\hat{\boldsymbol{R}}^{-2}$ as

$$\hat{\boldsymbol{R}}^{-2} = \boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{U}^{\mathrm{H}}, \tag{32}$$

where the columns of $P \times P$ matrix $\boldsymbol{U}$ contain the eigenvectors and the elements of the diagonal matrix $\boldsymbol{\Gamma}, \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_P$, are the corresponding eigenvalues. Substituting (32) into (31), we obtain the function $f(\mu)$ of the Lagrange multiplier μ as

$$\begin{aligned}f(\mu) &= \frac{\partial L}{\partial \mu} \\ &= \left[\frac{\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}}}{\bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{U}(\boldsymbol{\Gamma} + \mu\boldsymbol{I})^{-1}\boldsymbol{U}^{\mathrm{H}}\bar{\boldsymbol{a}}}\right]^2 \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{U}(\boldsymbol{\Gamma} + \mu\boldsymbol{I})^{-2}\boldsymbol{U}^{\mathrm{H}}\bar{\boldsymbol{a}} - \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}} \\ &= 0,\end{aligned} \tag{33}$$

where $\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}} = P$.

Lemma: There is a unique real-valued and positive solution $\mu > 0$ to (33).

Proof: See Appendix.

Let μ be the unique positive root of $f(\mu) = 0$ and introduce two auxiliary vectors

$$\boldsymbol{\rho} = \boldsymbol{U}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\bar{\boldsymbol{a}} = [\rho_1, \cdots, \rho_P]^{\mathrm{T}} \in \mathbb{C}^P, \tag{34}$$

$$\boldsymbol{\tau} = \boldsymbol{U}^{\mathrm{H}}\bar{\boldsymbol{a}} = [\tau_1, \cdots, \tau_P]^{\mathrm{T}} \in \mathbb{C}^P. \tag{35}$$

Substituting (34) and (35) into (33), we obtain

$$f(\mu) = P^2 \frac{\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{(\gamma_p+\mu)^2}}{\left(\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{\gamma_p+\mu}\right)^2} - \bar{\boldsymbol{a}}^{\mathrm{H}} \hat{\boldsymbol{R}} \bar{\boldsymbol{a}} = 0. \tag{36}$$

Therefore,

$$\begin{aligned} \bar{\boldsymbol{a}}^{\mathrm{H}} \hat{\boldsymbol{R}} \bar{\boldsymbol{a}} &= P^2 \frac{\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{(\gamma_p+\mu)^2}}{\left(\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{\gamma_p+\mu}\right)^2} \\ &\le P^2 \frac{\frac{1}{(\gamma_P+\mu)^2} \sum_{p=1}^{P} \rho_p^* \tau_p}{\frac{1}{(\gamma_1+\mu)^2} \left(\sum_{p=1}^{P} \rho_p^* \tau_p\right)^2} \\ &= \frac{(\gamma_1+\mu)^2}{(\gamma_P+\mu)^2} \frac{P^2}{\bar{\boldsymbol{a}}^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \bar{\boldsymbol{a}}}, \end{aligned} \tag{37}$$

which gives the following upper bound on μ:

$$\mu \le \frac{\gamma_1 - \zeta \gamma_P}{\zeta - 1}, \tag{38}$$

with

$$\zeta = \frac{1}{P} \sqrt{\bar{\boldsymbol{a}}^{\mathrm{H}} \hat{\boldsymbol{R}} \bar{\boldsymbol{a}} \bar{\boldsymbol{a}}^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \bar{\boldsymbol{a}}}. \tag{39}$$

Hence, the solution $\mu > 0$ to (33) is unique and it falls in the range of $(0, (\gamma_1 - \zeta\gamma_P)/(\zeta - 1))$. Using the Newton's method or the bisection method, equation (33) can be efficiently solved and the solution can be denoted as μ_0. Substituting μ_0 back into (26), the actual steering vector can be estimated as

$$\boldsymbol{a} = \frac{P}{\bar{\boldsymbol{a}}^{\mathrm{H}} \left(\hat{\boldsymbol{R}}^{-1} + \mu_0 \hat{\boldsymbol{R}}\right)^{-1} \bar{\boldsymbol{a}}} \left(\hat{\boldsymbol{R}}^{-1} + \mu_0 \hat{\boldsymbol{R}}\right)^{-1} \bar{\boldsymbol{a}}. \tag{40}$$

Using the estimated actual steering vector in the MVDR or AAR will result in robust adaptive beamforming capability for astronomical imaging. In summary, the proposed robust astronomical imaging method is achieved by performing the following steps for each source in the region of interest and at time epoch t_k:

Step 1: Estimate the actual steering vector $\boldsymbol{a}_k(\ell, m)$ by maximizing the beamformer output power under the constraint that prevents the estimated steering vector getting close to the interference steering vector;

Step 2: Generate the MVDR or AAR dirty image with the estimated steering vector, as described in (16) and (17), respectively.

4. Simulation Results

In this section, we provide the simulation results to demonstrate the imaging performance of the proposed astronomical imaging method. Throughout this section, the VLA telescope with $P = 27$ antennas is taken as an example of the array configuration. An 8-hour observation time period with 695 time epochs is applied, where each time epoch spans a time period of approximately 41.4 seconds. The simulated sky region, shown in Fig. 3, contains four elliptical Gaussian resources. At each time epoch, a sampled covariance matrix $\hat{\boldsymbol{R}}_k$ is generated from a simulated image based on $N = 10^5$ data samples. The input SNR is defined as the ratio of the average signal power and and the average noise power at each data sample, and is assumed

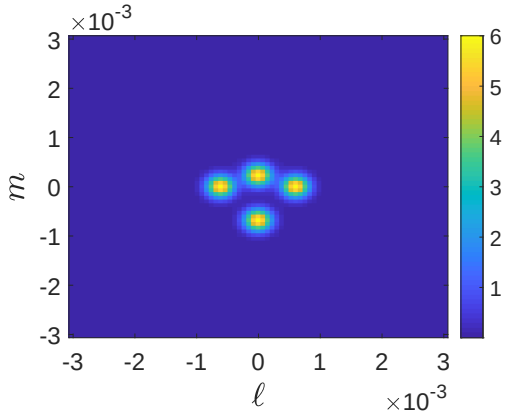


Fig. 3. The simulated image.

to be -20 dB for each antenna unless otherwise specified as in Fig. 7. The mismatch vector $\boldsymbol{e}_{d,k}$ follows a Gaussian distribution with $\boldsymbol{e}_{d,k} \sim \mathcal{CN}(\boldsymbol{0}, 0.25\boldsymbol{I})$. Because the array model mismatch may change the norm the steering vector, the actual steering vector for each point source is normalized as $\boldsymbol{a} = \sqrt{P}\boldsymbol{a}/\|\boldsymbol{a}\|$. In the worst-case beamformer, the upper bound of the mismatched vector is set up as $0.3P$ as suggested in the paper (Vorobyov *et al.*, 2003).

4.1. *The effect of model mismatch*

In this subsection, to better investigate the effect of model mismatch, i.e., instrumental effects and propagation effects, we design an experiment in the absence of the RFI. The result is shown in Fig. 4. The three plots in the first row show results without any array model mismatch, whereas the other nine plots are obtained with the array model mismatch. Compared with the DAS dirty image shown in Fig. 4(a) and the MVDR dirty image shown in Fig. 4(b), we can see that the AAR dirty image shown in Fig. 4(c) provides the highest image resolution.

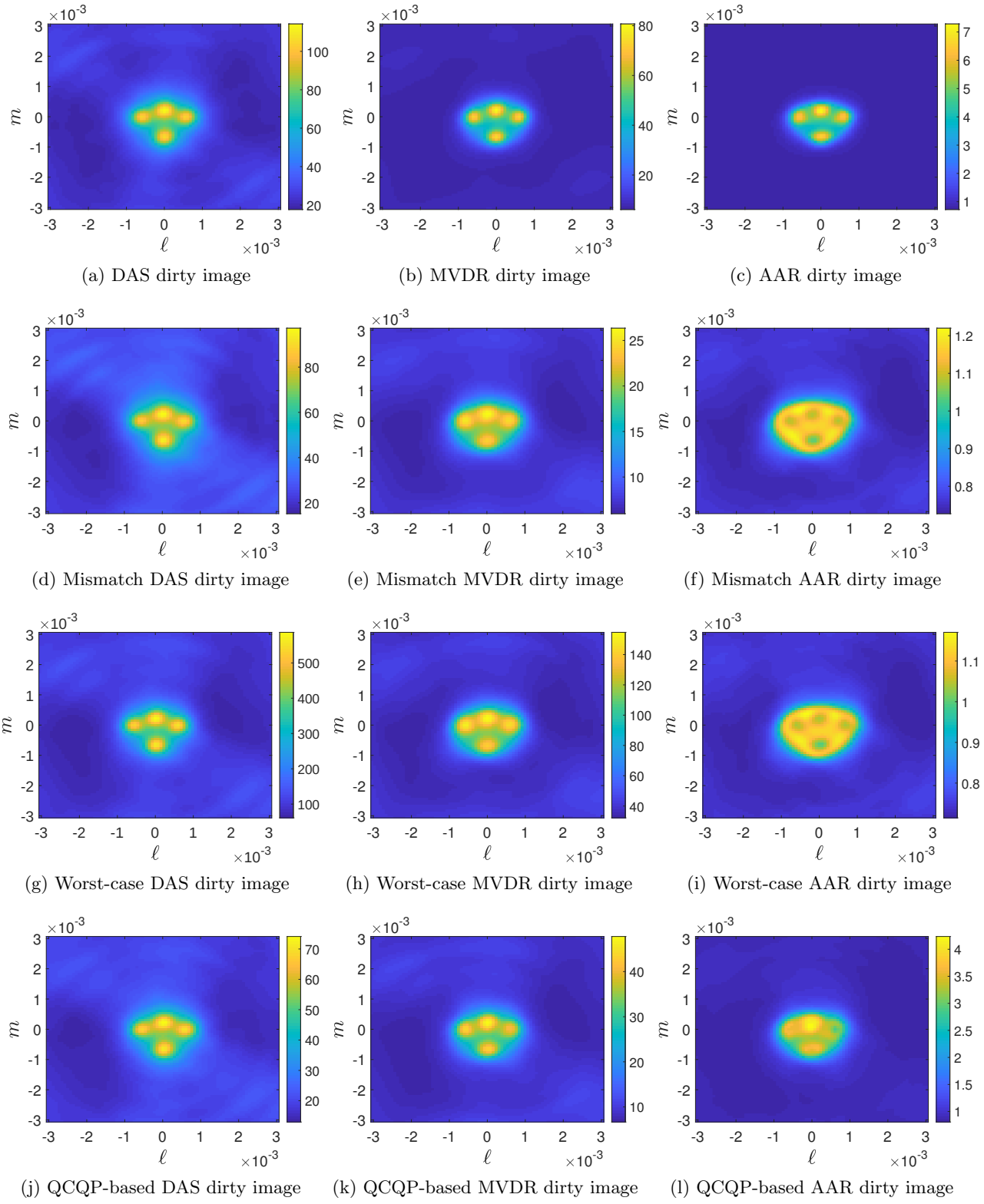
The second row in Fig. 4 shows the effect of model mismatch on the DAS, MVDR, and AAR imaging techniques. The imaging performance of the AAR suffers the most, and then followed by the MVDR and the DAS. The dirty image obtained from the AAR method is highly distorted and, as a result, the four Gaussian sources are unresolved. The third row in Fig. 4 shows the results of worst-case based methods (Vorobyov *et al.*, 2003) in the presence of model mismatch. We notice that the worst-case method hardly improves the imaging performance due to the large error ranges considered in this paper.

The fourth row in Fig. 4 shows the QCQP-based imaging techniques in the presence of model mismatch. We notice that the proposed method provides a certain level of improvement. For the AAR-based method, the image shows a better resolution of the four Gaussian sources. The dynamic image range is increased for the proposed QCQP-based imaging technique as compared with the mismatched cases shown in the second row. For MVDR, the dynamic range is increased from 20.4 to 41.1 whereas for AAR, the dynamic range is increased from 0.5 to 3.4.

Also, an experiment is designed in the absence of steering vector mismatch, and the cross sections at the center ($m = 0$) of the dirty images for the proposed QCQP-based methods are shown in Fig. 5. In the absence of model mismatch, the proposed method yields insignificant steering vector adjustment. As a result, the deviation between proposed QCQP-based dirty images and classical dirty images is negligible.

4.2. *The effect of model mismatch and RFI*

In this subsection, we consider the presence of both model mismatch and RFI. Consider two moving ground-based non-line-of-sight (NLOS) interferers. Due to fast fading effects, we model the corresponding spatial signature vectors $\boldsymbol{a}_{qk}^i$ as a collection of i.i.d. random vectors. The interference steering vectors are assumed to vary independently over time epochs as $\boldsymbol{a}_{qk}^i \sim \mathcal{CN}(0, \boldsymbol{I})$ and normalized $\boldsymbol{a}_{qk}^i = \sqrt{P}\boldsymbol{a}_{qk}^i/\|\boldsymbol{a}_{qk}^i\|$.



(a) DAS dirty image (b) MVDR dirty image (c) AAR dirty image

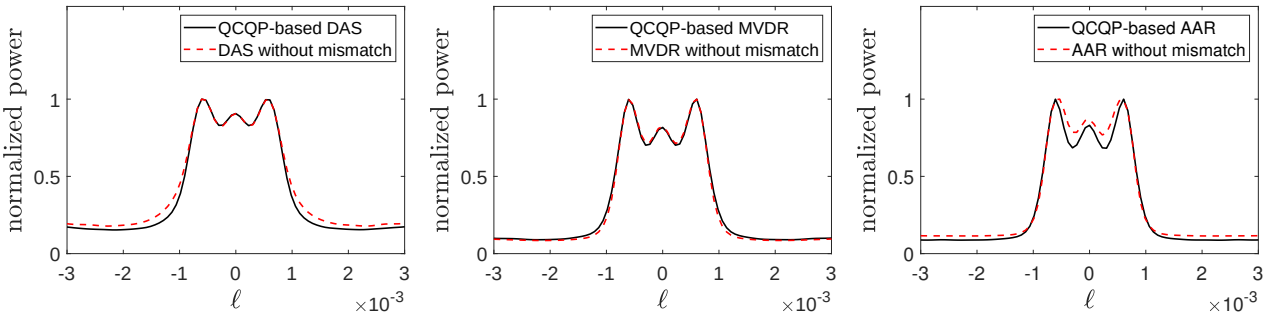
(d) Mismatch DAS dirty image (e) Mismatch MVDR dirty image (f) Mismatch AAR dirty image

(g) Worst-case DAS dirty image (h) Worst-case MVDR dirty image (i) Worst-case AAR dirty image

(j) QCQP-based DAS dirty image (k) QCQP-based MVDR dirty image (l) QCQP-based AAR dirty image

Fig. 4. Comparison of dirty images generated by different beamforming methods. Plots (a)–(c) are generated without steering vector mismatch; Plots (d)–(l) are generated in the presence of steering vector mismatch but without interference.

Unless otherwise specified, the input interference-to-noise ratio (INR) of the two interferers is respectively set to 15 dB and 20 dB in each antenna.



(a) QCQP-based DAS dirty image (b) QCQP-based MVDR dirty image (c) QCQP-based AAR dirty image

Fig. 5. The cross sections of the dirty images in the absence of steering vector mismatch.

In Fig. 6, the imaging performance is compared in the presence of RFI and model mismatch. Fig. 6(a) shows the dirty image generated by the classical DAS method. It is clear that, because DAS cannot suppress interference signals, the interference signals dominate the resulting image and the astronomical sources are completely obscured. The effect of interference can be mitigated by adaptive beamforming methods. We first consider the MVDR and AAR method in the presence of RFI but without model mismatch and the results are respectively shown in Fig. 6(b) and Fig. 6(c). Both achieve good interference suppression capability and the sources are clearly revealed.

The second row in Fig. 6 shows the results of the DAS, MVDR, and AAR imaging techniques in the presence of model mismatch and strong interference. With the existence of model mismatch, dirty images generated by both MVDR and AAR exhibit degradations, as shown in Fig. 6(e) and Fig. 6(f). The performance degradations of these methods are mainly caused due to model mismatch. The third row in Fig. 6 shows the results with the estimated array response by worst-case beamforming. Similar to the third row in Fig. 4, we can notice that the performances have a slight improvement, especially for DAS.

The fourth row in Fig. 6 shows the results that exploit the estimated array response obtained using the proposed method. Similar to the results depicted in Fig. 6, the proposed method outperforms the worst-case method, and maintains the adaptive interference suppression capability in the presence of model mismatch. As for DAS, although it is a data-independent method, the proposed method resolve the four Gaussian sources because the estimated steering vector is inferred based on the MVDR spatial spectrum. Note also that, compared with the results in the second row, the dynamic range is increased in the fourth row with the use of the proposed QCQP-based imaging technique. The proposed method increases the dynamic range from 21.8 to 80.3 for the MVDR and from 0.5 to 10.6 for the AAR.

To better demonstrate the effectiveness of the proposed method, we show the normalized power distribution respectively with the input SNR, input INR, and mismatch variance. Fig. 7 shows the cross sections at the center ($m = 0$) of the dirty images, indicating the performance of the tested methods versus the input SNR. For a better comparison, the cross sections have been normalized by its corresponding maximum value. We can see that the proposed method remains effective in all the tested SNR values. Fig. 8 shows the normalized output power for different values of the mismatch variance. Compared with the mismatched MVDR and AAR, the proposed QCQP-based imaging technique provides significant improvement, particularly when the variance is large. Fig. 9 shows the performance of the proposed method with different RFI levels. We notice that the proposed QCQP-based imaging technique improves the imaging performance, especially with the existence of the strong interference signals.

It is worthy to mention that the effect of the dirty beam due to finite and non-uniform spatial sampling can be mitigated to obtain a high-resolution clean image through various deconvolution methods, such as the CLEAN algorithm (Högbom, 1974), the maximum entropy method (MEM) algorithm (Jaynes, 1957) and the Least Squares Minimum Variance Imaging (LS-MVI) algorithm (Ben-David & Leshem, 2008).

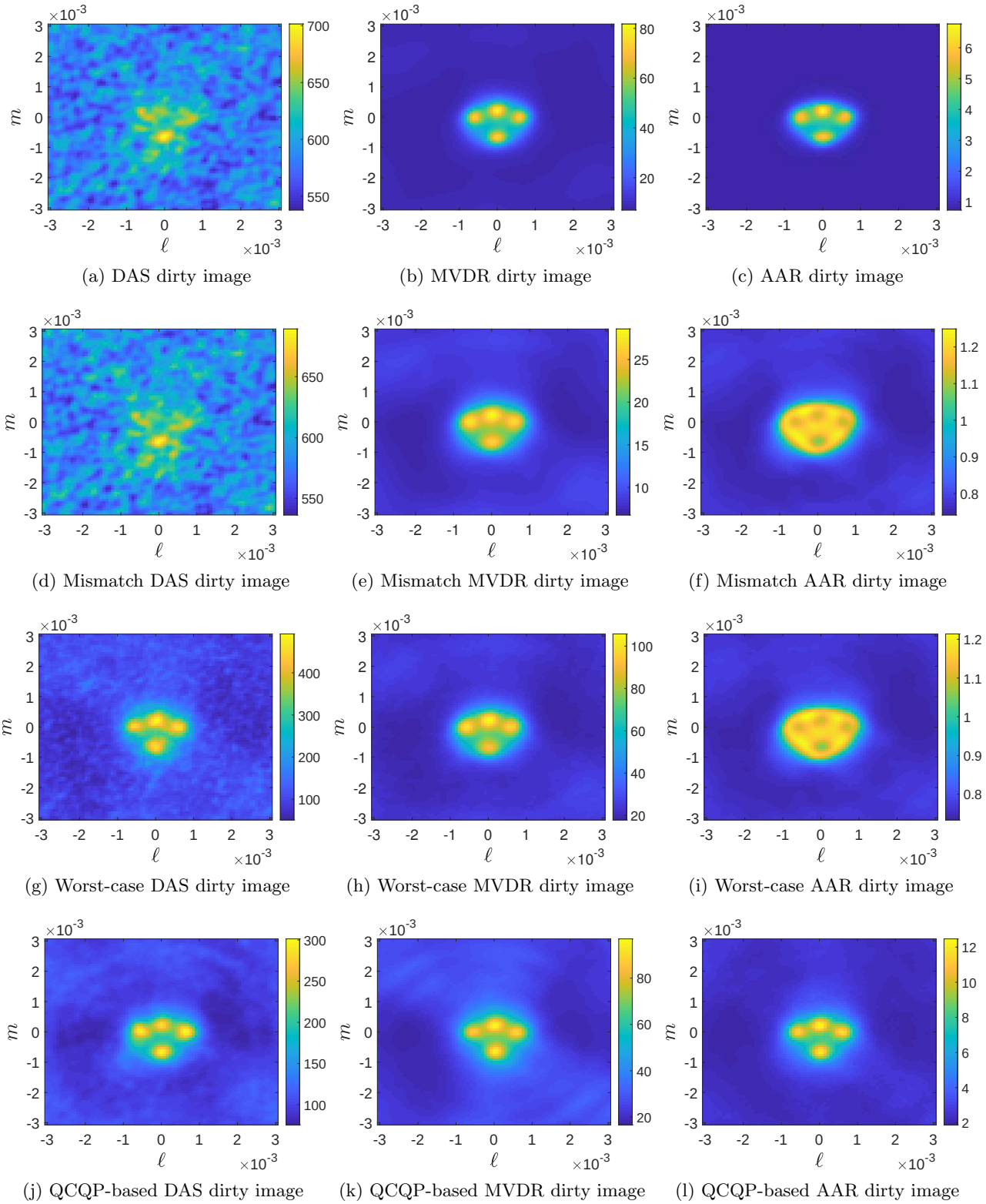
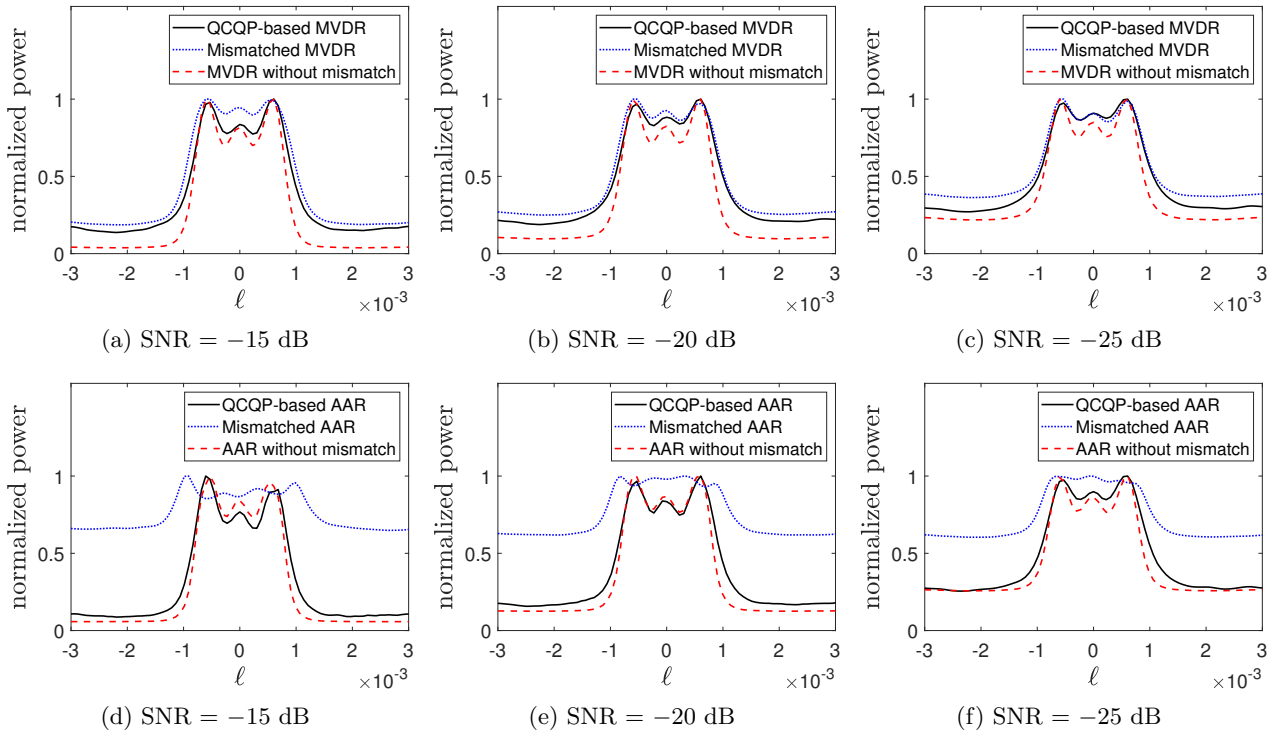


Fig. 6. Comparison of dirty images generated by different beamforming methods. Plots (a)–(c) are generated without steering vector mismatch; Plots (d)–(l) are generated in the presence of steering vector mismatch and strong interference.

5. Conclusion

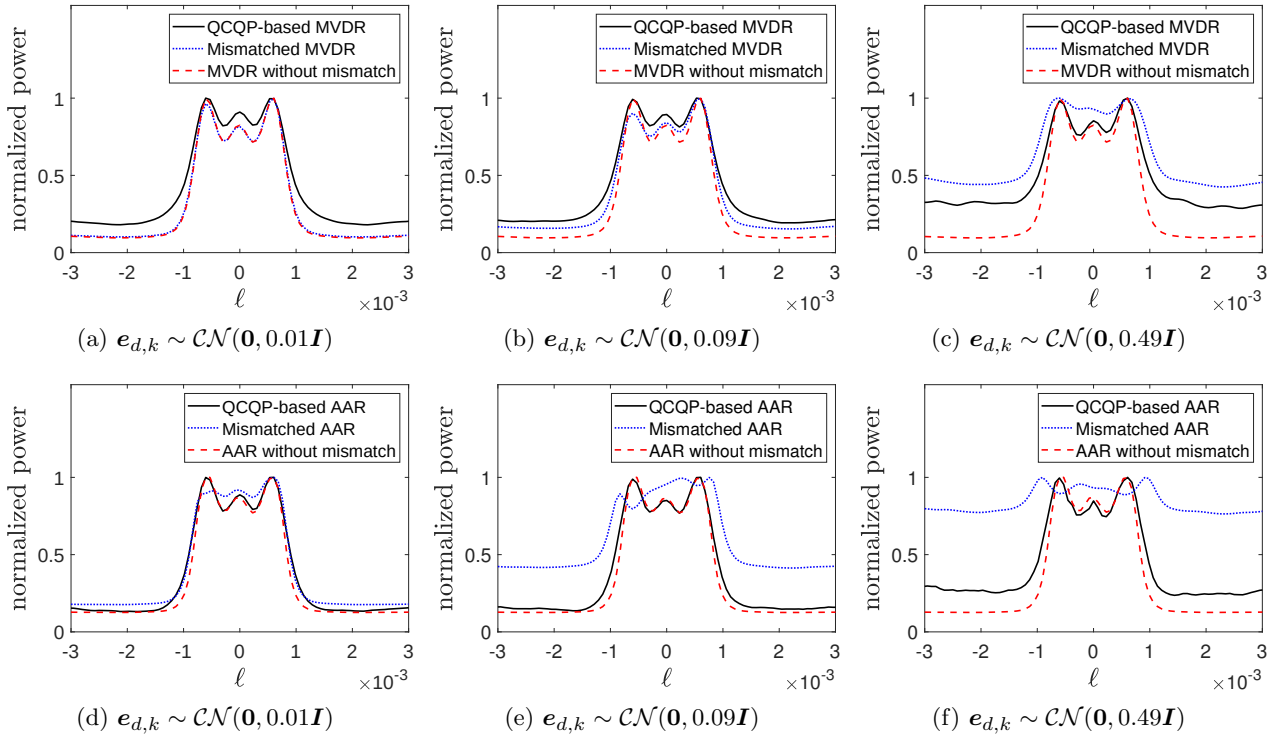
In this paper, we proposed a robust astronomical imaging method which maintains the effective interference suppression capability in the presence of model mismatch. The proposed method is based on the estimation



(a) SNR $= -15$ dB (b) SNR $= -20$ dB (c) SNR $= -25$ dB

(d) SNR $= -15$ dB (e) SNR $= -20$ dB (f) SNR $= -25$ dB

Fig. 7. The cross sections of the dirty images in different SNR conditions in the presence of RFI.

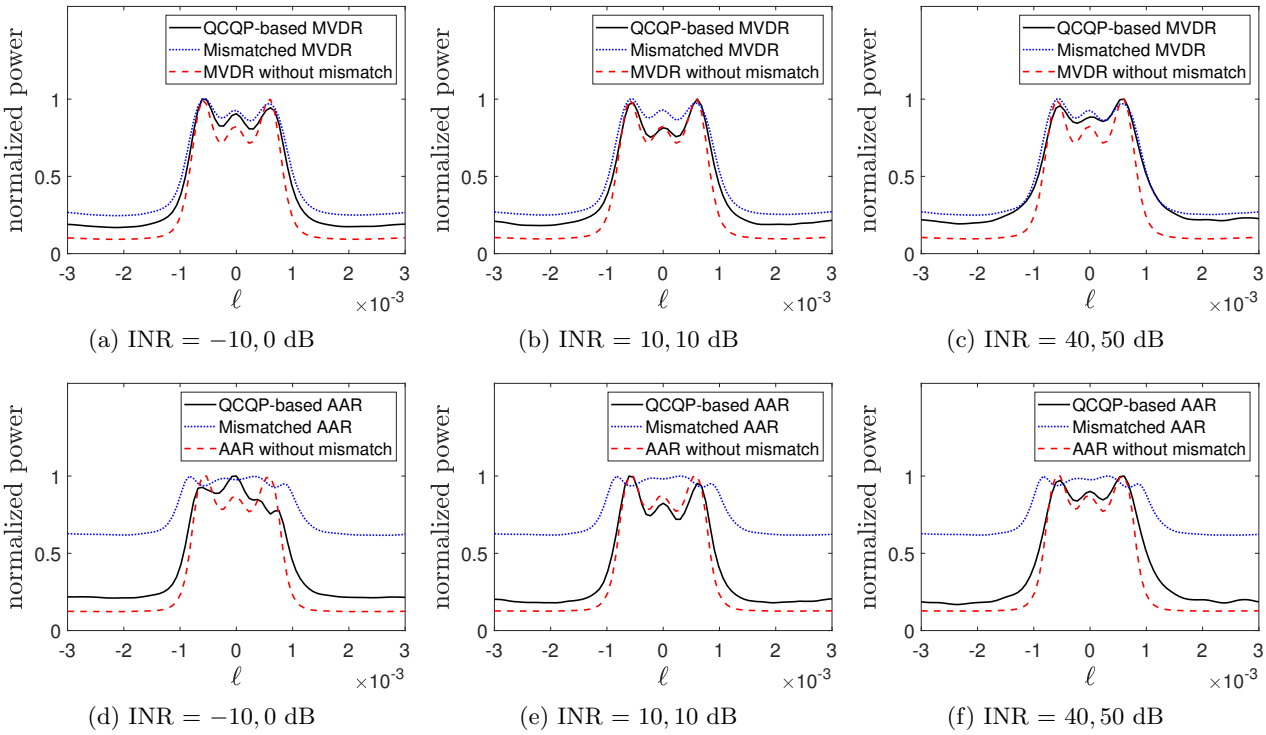


(a) $\boldsymbol{e}_{d,k} \sim \mathcal{CN}(\boldsymbol{0}, 0.01\boldsymbol{I})$ (b) $\boldsymbol{e}_{d,k} \sim \mathcal{CN}(\boldsymbol{0}, 0.09\boldsymbol{I})$ (c) $\boldsymbol{e}_{d,k} \sim \mathcal{CN}(\boldsymbol{0}, 0.49\boldsymbol{I})$

(d) $\boldsymbol{e}_{d,k} \sim \mathcal{CN}(\boldsymbol{0}, 0.01\boldsymbol{I})$ (e) $\boldsymbol{e}_{d,k} \sim \mathcal{CN}(\boldsymbol{0}, 0.09\boldsymbol{I})$ (f) $\boldsymbol{e}_{d,k} \sim \mathcal{CN}(\boldsymbol{0}, 0.49\boldsymbol{I})$

Fig. 8. The cross sections of the dirty images with different mismatch variances in the presence of RFI.

of the actual array responses by solving a QCQP problem, which maximizes the beamformer output power under the constraint that the estimated array response does not converge to that of the interferer. The numerical solution of the proposed steering vector estimation problem based on the KKT conditions is



(a) INR $= -10, 0$ dB (b) INR $= 10, 10$ dB (c) INR $= 40, 50$ dB

(d) INR $= -10, 0$ dB (e) INR $= 10, 10$ dB (f) INR $= 40, 50$ dB

Fig. 9. The cross sections of the dirty images in the presence of RFI with difference interference levels.

provided. Then, the estimated array responses are used in the adaptive beamforming methods for radio astronomical imaging. Simulation results show that the proposed methods, especially the QCQP-based AAR beamformer, achieve a good imaging performance in the presence of RFI and model mismatch.

Appendix: Proof of Lemma

The existence of a unique real-valued and positive solution $\mu \geq 0$ of (33) is proved by showing

(i) $f(\mu)$ is a monotonically decreasing function
(ii) the solution μ of $f(\mu) = 0$ is positive.

First, in order to prove the monotonicity of function $f(\mu)$, we need to obtain the differentiation of function $f(\mu)$ with respect to μ.

$$\frac{\partial f(\mu)}{\partial \mu} = 2P^2 \frac{\left[\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{(\gamma_p+\mu)^2}\right]^2 - \left[\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{(\gamma_p+\mu)^3}\right]\left[\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{(\gamma_p+\mu)}\right]}{\left[\sum_{p=1}^{P} \frac{\rho_p^* \tau_p}{(\gamma_p+\mu)}\right]^3} \leq 0, \tag{41}$$

where the inequality $\partial f(\mu)/\partial \mu \leq 0$ is derived from the Cauchy-Schwarz inequality and, hence, the monotonicity of function $f(\mu)$ on variable $\mu \geq 0$ is proved.

Next, we prove that the solution of $f(\mu) = 0$ is positive.

We note that

$$\lim_{\mu \to \infty} f(\mu) = \frac{(\bar{\boldsymbol{a}}^{\mathrm{H}} \bar{\boldsymbol{a}})^2}{\bar{\boldsymbol{a}}^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \bar{\boldsymbol{a}}} - \bar{\boldsymbol{a}}^{\mathrm{H}} \hat{\boldsymbol{R}} \bar{\boldsymbol{a}} \leq 0, \tag{42}$$

and

$$f(0) = \left(\frac{\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}}}{\bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}}}\right)^2 \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}^3\bar{\boldsymbol{a}} - \bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}} \geq 0. \tag{43}$$

To prove (42), let us use the eigen-decomposition of $\hat{\boldsymbol{R}}$,

$$\hat{\boldsymbol{R}} = \boldsymbol{Q}\Xi\boldsymbol{Q}^{\mathrm{H}}, \tag{44}$$

where $\boldsymbol{Q} \in \mathbb{C}^{P\times P}$ is the unitary matrix whose columns are the eigenvectors of $\hat{\boldsymbol{R}}$, and $\Xi = \mathrm{Diag}\{\xi_1, \cdots, \xi_P\}$ is the diagonal matrix of eigenvalues of $\hat{\boldsymbol{R}}$.

Introducing a new auxiliary vector η as

$$\boldsymbol{\eta} = \boldsymbol{Q}^{\mathrm{H}}\bar{\boldsymbol{a}} = [\eta_1, \cdots, \eta_P] \in \mathbb{C}^P, \tag{45}$$

we have

$$\begin{aligned}
\bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^{\mathrm{H}}\hat{\boldsymbol{R}}\bar{\boldsymbol{a}} &= \bar{\boldsymbol{a}}^{\mathrm{H}}\boldsymbol{Q}\Xi^{-1}\boldsymbol{Q}^{\mathrm{H}}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^{\mathrm{H}}\boldsymbol{Q}\Xi\boldsymbol{Q}^{\mathrm{H}}\bar{\boldsymbol{a}} \\
&= \boldsymbol{\eta}^{\mathrm{H}}\Xi^{-1}\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{H}}\Xi\boldsymbol{\eta} \\
&= \left(\sum_{p=1}^{P}\frac{|\eta_p|^2}{\xi_p}\right)\left(\sum_{p=1}^{P}\xi_p|\eta_p|^2\right) \\
&\geq \left(\sum_{p=1}^{P}|\eta_p|^2\right)^2 \\
&= \left(\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{\eta}\right)^2 \\
&= \left(\bar{\boldsymbol{a}}^{\mathrm{H}}\bar{\boldsymbol{a}}\right)^2,
\end{aligned} \tag{46}$$

where the inequality is from the Cauchy-Schwarz inequality, and the last equality is based on the fact that $\boldsymbol{Q}$ is a unitary matrix. Hence, inequality (42) is proved. Similarly, inequality (43) can also be proved using the Cauchy-Schwarz inequality.

Since $f(0) > 0, f(\infty) < 0$ and $f(\mu)$ is a continuous differentiable monotonic function for positive values of μ, it has a unique root in the interval $(0, +\infty)$. This completes the proof.

References

Ben-David, C. & Leshem, A. [2008] *IEEE J. Sel. Topics Signal Process.* **2**, 670.

Carlson, B. D. [1988] *IEEE Trans. Aerosp. Electron. Syst.* **24**, 397.

Chang, L. & Yeh, C.-C. [1992] *IEEE Trans. on Antennas Propag.* **40**, 1336.

Cox, H., Zeskind, R. & Owen, M. [1987] *IEEE Trans. Acoust., Speech, Signal Process.* **35**, 1365.

Feldman, D. D. & Griffiths, L. J. [1994] *IEEE Trans. Signal Process.* **42**, 867.

Grant, M., Boyd, S. & Ye, Y. [2008] "CVX: Matlab software for disciplined convex programming," http://cvxr.com/cvx/.

Gu, Y. & Leshem, A. [2012] *IEEE Trans. Signal Process.* **60**, 3881.

Hall, S. [2013] "How astronomy benefits society and humankind," https://www.universetoday.com/106302/how-astronomy-benefits-society-and-humankind/.

Högbom, J. [1974] *Astron. Astrophys. Suppl.,* **15**, 417.

Hogbom, J. & Brouw, W. [1974] *Astron. Astrophys.* **33**, 289.

Jaynes, E. T. [1957] *Phys. Rev.* **106**, 620.

Leshem, A. & Van der Veen, A.-J. [2000] *IEEE Trans. Inf. Theory* **46**, 1730.

Leshem, A., van der Veen, A.-J. & Boonstra, A.-J. [2000] *Astron. Astrophys. Suppl.* **131**, 355.

Levanda, R. & Leshem, A. [2010] *IEEE Signal Process. Mag.* **27**.

Li, J., Stoica, P. & Wang, Z. [2003] *IEEE Trans. Signal Process.* **51**, 1702.

Li, J., Stoica, P. & Wang, Z. [2004] *IEEE Trans. Signal Process.* **52**, 2407.

Lorenz, R. G. & Boyd, S. P. [2005] *IEEE Trans. Signal Process.* **53**, 1684.
Napier, P. J., Thompson, A. R. & Ekers, R. D. [1983] *Proc. IEEE* **71**, 1295.
Ollier, V., El Korso, M. N., Boyer, R., Larzabal, P. & Pesavento, M. [2017] *IEEE Trans. Signal Process.* **65**, 5649.
Raza, J., Boonstra, A.-J. & Van der Veen, A.-J. [2002] *IEEE Signal Process. Lett.* **9**, 64.
Ryle, M. [1952] "A new radio interferometer and its application to the observation of weak radio stars," *Proc. R. Soc. Lond. A*, p. 351.
Thompson, A. R., Moran, J. M., Swenson, G. W. *et al.* [2017] *Interferometry and Synthesis in Radio Astronomy* (Springer).
Van der Tol, S. & van der Veen, A.-J. [2005] "Application of robust capon beamforming to radio astronomical imaging," *Proc. IEEE ICASSP* (IEEE), p. iv.
Van der Veen, A.-J., Leshem, A. & Boonstra, A.-J. [2004] "Signal processing for radio astronomical arrays," *Proc. IEEE Sensor Array and Multichannel Signal Process. Workshop*, p. 1.
van Driel, W. [2009] *Proc. IAU Symposium* **5**, 457.
Vorobyov, S. A., Gershman, A. B. & Luo, Z.-Q. [2003] *IEEE Trans. Signal Process.* **51**, 313.
Wijnholds, S. J., Van der Tol, S., Nijboer, R. & Van der Veen, A.-J. [2010] *IEEE Signal Process. Mag.* **27**.
Zhang, S., Gu, Y., Barott, W. C. & Zhang, Y. D. [2018a] "Improved radio astronomical imaging based on sparse reconstruction," *Proc. SPIE*, p. 106580O.
Zhang, S., Gu, Y., Wang, B. & Zhang, Y. D. [2017] "Robust astronomical imaging under coexistence with wireless communications," *Proc. Asilomar Conf. Signals, Systems, and Computers.*
Zhang, S., Gu, Y., Won, C.-H. & Zhang, Y. D. [2018b] "Dimension-reduced radio astronomical imaging based on sparse reconstruction," *Proc. IEEE Sensor Array and Multichannel Signal Process. Workshop.*