Norm-Constrained Capon Beamforming Using Multirank Signal Models With Kalman Filter Implementation

Jwu-Sheng Hu, *Member, IEEE*, and Ming-Tang Lee

***Abstract*—The performance of the Capon beamforming is known to degrade dramatically in the presence of model mismatches, especially when the desired signal is present in the training data. To improve the robustness, the diagonal loading technique was introduced. However, the major drawback is the selection of the diagonal loading level, which is related to the unknown signal powers. Another way to alleviate performance degradation is to use a more accurate signal model of the array response. In this paper, a norm-constrained Capon beamforming using multirank signal models is proposed. Based on the pseudo-observation method, the quadratic constraints can be easily constructed. The problem is solved by a nonlinear Kalman filter which can be implemented on-line. The simulation results show that the design of the norm-constraint value is less sensitive to the signal powers, small angle mismatches, and number of sensors with a standard linear array. Further, it is shown that the use of a multirank signal model and Kalman filter technique result in less self-cancellation and performance degradation than that of the rank-1 signal model and the estimation of sample matrix.**

***Index Terms*—Capon beamforming, Kalman filter, multirank signal model, norm constraint.**

I. Introduction

CAPON beamforming [or minimum variance distortionless response beamforming, (MVDR) beamforming] aims to minimize variances of the interferences and noise while maintaining the desired array response. It is known to degrade dramatically due to even small mismatches of the desired signal model, especially when the desired signal is present in the training data. The research on robust Capon beamforming has focused on maintaining the output signal-to-interference-and-noise-ratio (SINR) performance against several array or propagation uncertainties. In the real world environment, the spatial correlation is typically multirank due to local scattering, wavefront fluctuation, or reverberation [1]–[5]. It requires a multirank signal model to provide a more accurate representation of the wave propagation of the desired source to the sensors. In this case, if the array response formulated by the multirank signal model is known exactly, performance degradation as described above can be alleviated. To further reduce the sensitivity to arbitrary kinds of mismatches, the diagonal loading (DL) [6]–[10] technique has been widely used to improve the robustness of the Capon beamformer. The major drawback of DL is that it is hard to choose the diagonal loading level since it is related to the unknown signal power. Norm-constrained Capon beamforming (NCCB) is known to be equivalent to DL [7], [10]. However, knowledge of how to choose the norm-constraint value has not been completely studied. In this paper, we will demonstrate the superiority of using the norm constraint to the original DL formulation (Section V).

To facilitate flexible implementation, the NCCB is solved via the constrained Kalman filtering technique. Constrained Kalman filtering [11]–[15] has been widely investigated in the last decade. The approaches mainly fall into one of three categories: pseudo-observation methods (or penalty methods), projection methods, and dimension reduction methods. Among these methods, the pseudo-observation method is the most intuitive way in merging the constraints into the state space of the Kalman filtering by considering the constraints as additional measurement equations. In this way, several existing nonlinear Kalman filtering algorithms can be directly applied. Chen and Chiang [16] were the first to introduce the penalty method [14], [15] into the traditional Capon beamforming problem. El-Keyi *et al.* used the penalty method in robust adaptive beamforming based on worst-case performance optimization [17]. In this paper, we also use the penalty method for robust adaptive beamforming with multirank signal models and an additional norm constraint. The settings of the initial conditions and parameter matrices are studied to achieve good performance and prevent the ill-conditioning problem. Compared to the previous work [1], the computation of the principal eigenvector can be avoided in the proposed method.

Since the NCCB problem with multirank signal models is quadratic, the associated nonlinear Kalman filtering can be solved directly by the extended Kalman filter (EKF) [18]–[20]. Another popular method is the unscented Kalman filter (UKF) [20]–[22]. The EKF approximates the Jacobian and Hessian matrices (in the first- and second-order approximations) of the nonlinear functions, while the UKF approximates the probability distribution of the nonlinear transformation using sigma points. Theoretically, the second-order extended Kalman filter (SOEKF) gives the best approximation in the mean square error (MSE) sense. However, due to the approximation of the second-order errors, the SOEKF is more sensitive to improper initial conditions and parameter matrices. The comparison of the above nonlinear Kalman filters will be discussed in Section VI.

The remainder of this paper is organized as follows. In Section II, we briefly review the Capon beamforming problem

Manuscript received October 29, 2013; revised May 21, 2014; accepted June 08, 2014. Date of publication June 30, 2014; date of current version September 01, 2014.

The authors are with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: jshu@cn.nctu.edu.tw; lhoney.ece97g@g2.nctu.edu.tw).

Digital Object Identifier 10.1109/TAP.2014.2333051

with multirank signal models. Section III introduces the NCCB problem with multirank signal models. Section IV gives the state-space model of the NCCB problem based on the penalty method and presents the solutions using the EKFs and the UKF. Section V shows the superiority of design of the norm-constraint value and Section VI compares the beamformers. Finally, the conclusions are drawn in Section VII.

II. Problem Formulation

Considering an array with M sensors, the output of a narrowband beamformer is given by

$$y(k) = \mathbf{w}^H \mathbf{x}(k) \tag{1}$$

where k is the time index, $\mathbf{x}(k)$ is the $M \times 1$ complex observation vector, and $\mathbf{w}$ is the $M \times 1$ complex beamformer weight vector. The observation vector is composed of statistically independent components as

$$\mathbf{x}(k) = \mathbf{s}(k) + \mathbf{i}(k) + \mathbf{n}(k) \tag{2}$$

where $\mathbf{s}(k), \mathbf{i}(k)$, and $\mathbf{n}(k)$ are the desired signal, interference, and sensor noise, respectively. In the general case of the Capon beamforming problem, the weight vector is designed to minimize the output power subject to a distortionless constraint as [1]

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{\Phi}_x \mathbf{w} \quad \text{subject to } \mathbf{w}^H \mathbf{\Phi}_s \mathbf{w} = 1 \tag{3}$$

where

$$\mathbf{\Phi}_s = E[\mathbf{s}(k)\mathbf{s}(k)^H] \tag{4}$$

$$\mathbf{\Phi}_x = E[\mathbf{x}(k)\mathbf{x}(k)^H] = \mathbf{\Phi}_s + \mathbf{\Phi}_{i+n} \tag{5}$$

are the signal and interference-plus-noise covariance matrices, respectively. This problem is equivalent to maximizing the SINR [1]

$$\text{SINR} = \frac{\mathbf{w}^H \mathbf{\Phi}_s \mathbf{w}}{\mathbf{w}^H \mathbf{\Phi}_{i+n} \mathbf{w}} \tag{6}$$

The signal covariance matrix $\mathbf{\Phi}_s$ describes the relationship between the desired source and the sensors. One commonly used model of the signal field is a point source in a homogeneous field [23], where $\mathbf{\Phi}_s$ is measured as

$$\mathbf{\Phi}_s = \phi_s \mathbf{a}_s \mathbf{a}_s^H \tag{7}$$

where ϕ_s and $\mathbf{a}_s$ are the variance and steering vector of the desired signal. In practical environments, $\mathbf{\Phi}_s$ is multirank due to the local scattering, wavefront fluctuation, and reverberation. In the case of an incoherently scattered source, $\mathbf{\Phi}_s$ can be expressed by

$$\mathbf{\Phi}_s = \phi_s \int_{-\pi/2}^{\pi/2} \rho(\theta)\mathbf{a}(\theta)\mathbf{a}^H(\theta)d\theta \tag{8}$$

where $\rho(\theta)$ is the normalized angular power density function ($\int_{-\pi/2}^{\pi/2} \rho(\theta)d\theta = 1$), and $\mathbf{a}(\theta)$ is the steering vector at direction θ.

In practical applications, $\mathbf{\Phi}_x$ is estimated by the input sample covariance matrix as

$$\hat{\mathbf{\Phi}}_x = \frac{1}{N} \sum_{k=1}^{N} \mathbf{x}(k)\mathbf{x}^H(k) \tag{9}$$

where N is the training size. An analytical solution to the Capon beamforming problem in (3) is given by

$$\hat{\mathbf{w}}_{\text{MRSMI}} = \mathcal{P}\left\{\hat{\mathbf{\Phi}}_x^{-1} \mathbf{\Phi}_s\right\} \tag{10}$$

where $\mathcal{P}\{\cdot\}$ denotes the operator that yields the *normalized principal eigenvector* of a matrix where the normalization is applied to satisfy the distortionless constraint in (3). This solution is denoted as the multirank sample matrix inverse (MRSMI) beamformer [1]. However, when the desired signal exists in the training data, Capon beamforming is known to degrade dramatically due to the mismatches between the presumed and actual array responses to the desired signal [1]. This is the so-called *self-cancellation* phenomenon. To improve the robustness of Capon beamforming against mismatches, one of the most popular approaches is DL method. It is equivalent to imposing an additive noise on the input covariance matrix [6], [7], and the solution is modified as

$$\hat{\mathbf{w}}_{\text{MRLSMI}} = \mathcal{P}\{(\hat{\mathbf{\Phi}}_x + \lambda \mathbf{I})^{-1} \mathbf{\Phi}_s\} \tag{11}$$

where λ is the DL level to be determined. This solution is referred to as the multirank loaded SMI (MRLSMI) beamformer. The major drawback of MRLSMI is that it is not clear how to choose the best DL level λ since the optimal choice depends on the unknown source power. In this paper, we will show the superiority of the selection of the norm-constraint value compared to the DL level. Further, the benefits of using multirank signal models and the Kalman filter are also studied.

III. NCCB With Multirank Signal Models

Cox *et al.* [7] have shown that the DL problem is equivalent to the NCCB problem. For the multirank case, the NCCB can be expressed by

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{\Phi}_x \mathbf{w} \quad \text{subject to } \begin{cases} \mathbf{w}^H \breve{\mathbf{\Phi}}_s \mathbf{w} = 1 \\ \|\mathbf{w}\|^2 = T \end{cases} \tag{12}$$

where

$$\breve{\mathbf{\Phi}}_s = \frac{\mathbf{\Phi}_s}{\text{tr}\{\mathbf{\Phi}_s\}/M} \tag{13}$$

denotes the normalized matrix of $\mathbf{\Phi}_s$; $\text{tr}\{\cdot\}$ denotes the trace operator; and T is the designed constraint value of the squared weight vector norm. The normalization of (13) removes the effect of ϕ_s and leaves spatial characteristics only. This helps to design the threshold T, which will be discussed later.

The Lagrangian function of (12) can be defined by

$$J(\mathbf{w}, \lambda, \mu) = \mathbf{w}^H \mathbf{\Phi}_x \mathbf{w} + \lambda(\|\mathbf{w}\|^2 - T) + \mu(1 - \mathbf{w}^H \breve{\mathbf{\Phi}}_s \mathbf{w}) \tag{14}$$

By letting the derivatives with respect to $\mathbf{w}$, λ, and μ be equal to zero, an explicit solution to the Lagrangian function (13) has the same form as (11)

$$\mathbf{w}_{\mathrm{DL}} = \mathcal{P}\{(\mathbf{\Phi}_x + \lambda \mathbf{I})^{-1} \breve{\mathbf{\Phi}}_s\} \tag{15}$$

where the normalized principle eigenvector $\mathbf{w}_{\mathrm{DL}}$ has to satisfy both the constraints in (12). According to the norm constraint $\|\mathbf{w}_{\mathrm{DL}}\|^2 = T$ and the solution given by (15), the boundaries of the norm-constraint threshold T can be obtained by choosing $\lambda = 0$ and $\lambda \to \infty$

$$\|\mathbf{w}_{\mathrm{matched}}\|^2 \leq T \leq \|\mathbf{w}_{\mathrm{Capon}}\|^2 \tag{16}$$

where

$$\mathbf{w}_{\mathrm{Capon}} = \mathcal{P}\{\mathbf{\Phi}_x^{-1} \breve{\mathbf{\Phi}}_s\} \tag{17}$$

$$\mathbf{w}_{\mathrm{matched}} = \mathcal{P}\{\breve{\mathbf{\Phi}}_s\} \tag{18}$$

$\mathbf{w}_{\mathrm{Capon}}$ denotes the standard Capon beamformer and $\mathbf{w}_{\mathrm{matched}}$ denotes the matched filter for the multirank case. It is worth noting that the squared norm of $\mathbf{w}_{\mathrm{matched}}$ that satisfies the distortionless constraint $\mathbf{w}^H \breve{\mathbf{\Phi}}_s \mathbf{w} = 1$ can be written as

$$\|\mathbf{w}_{\mathrm{matched}}\|^2 = \frac{\mathcal{P}^H\{\breve{\mathbf{\Phi}}_s\}\mathcal{P}\{\breve{\mathbf{\Phi}}_s\}}{\mathcal{P}^H\{\breve{\mathbf{\Phi}}_s\}\breve{\mathbf{\Phi}}_s\mathcal{P}\{\breve{\mathbf{\Phi}}_s\}} = \frac{1}{\sigma_{\max}\{\breve{\mathbf{\Phi}}_s\}} \tag{19}$$

where $\sigma_{\max}\{\breve{\mathbf{\Phi}}_s\}$ denotes the maximum eigenvalue of $\breve{\mathbf{\Phi}}_s$. The matched filter is known to be optimal in the incoherent noise field (for rank-1 signal model, it is the well-known delay-and-sum filter).

For the case of the rank-1 signal model, the weight vector can be decomposed into the subspace of the presumed steering vector and its null space based on the concept of the generalized sidelobe canceller (GSC) [6] as

$$\mathbf{w} = \frac{\mathbf{a}_s}{M} + \mathbf{a}_s^{\perp} \quad \text{where } \mathbf{a}_s^H \mathbf{a}_s^{\perp} = 0 \tag{20}$$

where $\mathbf{a}_s^{\perp}$ is the null space and the weight vector satisfies the distortionless constraint $\mathbf{w}^H \mathbf{a}_s = 1$. In this case, the norm of the weight vector can be expressed as

$$\|\mathbf{w}\|^2 = \frac{1}{M} + \|\mathbf{a}_s^{\perp}\|^2. \tag{21}$$

It can be observed that the norm constraint restricts the norm of the vector $\mathbf{a}_s^{\perp}$ that does not belong to the feasible set of the distortionless constraint. In the case of the multirank signal models, the weight vector can be decomposed in a similar way

$$\mathbf{w} = \mathbf{w}_{\mathrm{matched}} + \mathbf{w}^{\perp}. \tag{22}$$

If the signal model described by $\mathbf{\Phi}_s$ is full-rank, its null space does not exist. As a result, the norm constraint is used to penalize the additional term $\mathbf{w}^{\perp}$ that does not satisfy the distortionless constraint. In order to express the effect of $\mathbf{w}^{\perp}$, we decompose the threshold T into

$$T = \frac{1}{\sigma_{\max}\{\breve{\mathbf{\Phi}}_s\}} + \gamma, \quad \text{where } \gamma \geq 0. \tag{23}$$

By expressing T as (23), we found that the best selection of γ is between 0.005 and 0.02 which only depends on the number of sensors. It is useful that the best selected γ does not change with different signal mismatches or noise conditions. In comparison, the selection of the DL level λ is hard and sensitive to the desired signal power. In Section V, we will demonstrate the superiority of tolerance to the signal power and small angle mismatches using the best selected γ.

Note that the norm constraint on the Capon beamforming problem is useful only when there is a contradiction between the minimization of the output power and the distortionless constraint. In detail, when the presumed and actual signal models $\breve{\mathbf{\Phi}}_s$ are mismatched, the input desired signal may be considered as interference and a sharp transition can occur to satisfy the minimization and distortionless objectives (similar principle in [10]). Such transition increases the Euclidean norm of the weight vector, thus the imposed norm constraint is helpful in alleviating the *self-cancellation*. Once the model mismatch is large, the desired signal is wholly considered as interference and the contradiction is not activated. In this case, the norm constraint can no longer protect the desired signal from *self-cancellation*. On the other hand, the DL works better in large model mismatches with carefully designed DL levels.

As a result, the design with the norm constraint is expected to be useful without choosing γ and is expected to be more robust under moderate model mismatch conditions. In large model mismatch conditions (which more easily occur with small angular spreading and large angle mismatch), the norm constraint may fail to protect the desired signal.

IV. Solutions Using the Kalman Filter

The NCCB problem is formulated into a state-space model by using the penalty method. Based on the state-space model, nonlinear Kalman filtering approaches can be used to solve the quadratic problem. Subsequently, the settings of the initial conditions and parameter matrices are discussed.

A. State-Space Formulation Using the Penalty Method

The pseudo-observation method (or penalty method) treats the set of constraint equations as additional observations without measurement noise [14], [15]. In this case, the constraint equations are called *perfect measurements*, and the constraints are considered as "hard constraints". However, it is known that perfect measurements result in a singular error covariance matrix, which will lead to the ill-conditioning problem in the Kalman filter. Thus, small variances of the constraint equations are added instead which gives the "soft-constrained" solutions.

Considering the NCCB problem in (12), the state-space model is given as (24)–(28),

State-Space Formulation of the NCCB Problem

State equation

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{v}_s(k). \tag{24}$$

Measurement equation

$$\begin{bmatrix} 0 \\ 1 \\ T \end{bmatrix} = \begin{bmatrix} \mathbf{x}^H(k)\mathbf{w}(k) \\ f_1(\mathbf{w}(k)) \\ f_2(\mathbf{w}(k)) \end{bmatrix} + \begin{bmatrix} v_{m1}(k) \\ v_{m2}(k) \\ v_{m3}(k) \end{bmatrix} \tag{25}$$

where

$$f_1(\mathbf{w}(k)) = \mathbf{w}^H(k)\breve{\mathbf{\Phi}}_s\mathbf{w}(k) \tag{26}$$

$$f_2(\mathbf{w}(k)) = \mathbf{w}^H(k)\mathbf{w}(k). \tag{27}$$

A vector form of (25) is expressed as

$$\mathbf{z} = \mathbf{f}(\mathbf{w}(k)) + \mathbf{v}_m(k) \tag{28}$$

where $\mathbf{v}_s(k)$ and $\mathbf{v}_m(k)$ are the process and measurement noises respectively. Typically, the noise processes $\mathbf{v}_s(k)$ and $\mathbf{v}_m(k)$ are assumed to be zero-mean and mutually uncorrelated with the covariance matrices

$$\mathbf{Q} = E\left[\mathbf{v}_s(k)\mathbf{v}_s^H(k)\right] = \sigma_0^2\mathbf{I} \tag{29}$$

$$\mathbf{R} = E\left[\mathbf{v}_m(k)\mathbf{v}_m^H(k)\right] = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}. \tag{30}$$

The only real measurement in (25) is the input vector $\mathbf{x}(k)$ in the first equation given by the objective of minimizing the filtered output power in the MSE sense, i.e., $E[|0 - \mathbf{x}^H(k)\mathbf{w}(k)|^2]$.

Considering the measurement update in the Kalman filter

$$\hat{\mathbf{w}}(k) = \hat{\mathbf{w}}(k-1) + \mathbf{K}(k)(\mathbf{z} - \mathbf{f}(\mathbf{w}(k-1))). \tag{31}$$

It can be shown that the estimated state $\hat{\mathbf{w}}(k)$ is the solution to the optimization problem [15]

$$\begin{aligned} &\hat{\mathbf{w}}_{\mathrm{SCPO}}(k) \\ &\quad = \arg\min_{\mathbf{w}} \Big\{ (\mathbf{w} - \hat{\mathbf{w}}(k-1))^H \\ &\qquad \times [\mathbf{P}^-(k-1)]^{-1}(\mathbf{w} - \hat{\mathbf{w}}(k-1)) \\ &\qquad + \frac{1}{\sigma_1^2}\|0 - \mathbf{x}^H(k)\mathbf{w}\|^2 + \frac{1}{\sigma_2^2}\|1 - \mathbf{w}^H\breve{\mathbf{\Phi}}_s\mathbf{w}\|^2 \\ &\qquad + \frac{1}{\sigma_3^2}\|T - \mathbf{w}^H\mathbf{w}\|^2 \Big\} \end{aligned} \tag{32}$$

where $\mathbf{K}(k)$ is the Kalman gain and $\mathbf{P}^-(k)$ is the *a priori* state error covariance matrix at time index k. Now the constraint parameters σ_2^2 and σ_3^2 act as penalty terms. When the constraint parameters approach zero, the weightings of the constraint costs are increased, and the solutions that do not satisfy the constraint are more penalized. The solution of the penalty should approach the solution of the NCCB problem in (12) if the constraint parameters σ_2^2 and σ_3^2 are much smaller than σ_1^2. This means that the approximation of the nonlinear functions is adequate. To avoid numerical problems, typically the constraint parameters are set to be nonzero values. Therefore, the penalty method does not satisfy the constraints strictly, but provides a flexible approach to incorporate different equality constraints.

B. Solutions Using EKFs

The measurement equation described by (28) can be approximated by the Taylor series expansion to the second-order, around an estimate $\hat{\mathbf{w}}(k)$, as

$$\begin{aligned} \mathbf{f}(\mathbf{w}(k)) &= \mathbf{f}(\hat{\mathbf{w}}(k)) + \mathbf{F}_{\mathbf{w}}(\mathbf{w}(k) - \hat{\mathbf{w}}(k)) \\ &\quad + \frac{1}{2}\sum_{i=1}^{P} \boldsymbol{\phi}_i(\mathbf{w}(k) - \hat{\mathbf{w}}(k))^H\mathbf{F}_{\mathbf{ww}}^{(i)}(\mathbf{w}(k) - \hat{\mathbf{w}}(k)) \end{aligned} \tag{33}$$

where $\boldsymbol{\phi}_i = [\overbrace{0, \cdots, 1, \cdots, 0}^{i-\text{th}}]^T$ and P is the number of measurement equations. $\mathbf{F}_{\mathbf{w}}$ denotes the Jacobian matrix of the nonlinear function $\mathbf{f}(\mathbf{w}(k))$, and $\mathbf{F}_{\mathbf{ww}}^{(i)}$ denotes the Hessian matrix of the i-th measurement equation in $\mathbf{f}(\mathbf{w}(k))$. Two major algorithms of the nonlinear Kalman filtering are the EKF [18]–[20] and the UKF [20]–[22]. The EKF directly estimates the Jacobian and Hessian matrices. On the other hand, the UKF implicitly estimates the first- and second-order terms in the nonlinear transformation in (33) using *sigma points*. The Jacobian and Hessian matrices of the nonlinear function in (28) can be computed as

$$\begin{aligned} \mathbf{F}_{\mathbf{w}}(k) &= [\nabla_{\mathbf{w}}\mathbf{f}^T(\mathbf{w}(k))]^T \\ &= [\,\mathbf{x}^H(k) \quad \mathbf{w}^H(k)\breve{\mathbf{\Phi}}_s \quad \mathbf{w}^H(k)\,]^T \end{aligned} \tag{34}$$

$$\mathbf{F}_{\mathbf{ww}}^{(1)}(k) = \nabla_{\mathbf{w}}\nabla_{\mathbf{w}}^H\{\mathbf{x}^H(k)\mathbf{w}(k)\} = \mathbf{0} \tag{35}$$

$$\mathbf{F}_{\mathbf{ww}}^{(2)}(k) = \nabla_{\mathbf{w}}\nabla_{\mathbf{w}}^H\{f_1(\mathbf{w}(k))\} = \breve{\mathbf{\Phi}}_s \tag{36}$$

$$\mathbf{F}_{\mathbf{ww}}^{(3)}(k) = \nabla_{\mathbf{w}}\nabla_{\mathbf{w}}^H\{f_2(\mathbf{w}(k))\} = \mathbf{I}. \tag{37}$$

The first- and second-order extended Kalman filters (FOEKF and SOEKF) can be summarized as (38)–(43) [20].

Multirank NCCB Using the FOEKF and SOEKF

$$\mathbf{P}^-(k) = \mathbf{P}^+(k-1) + \mathbf{Q} \tag{38}$$

$$\mathbf{e}(k) = \mathbf{z} - \mathbf{f}(\hat{\mathbf{w}}(k-1)) - \boldsymbol{\pi}(k) \tag{39}$$

$$\mathbf{S}(k) = \mathbf{F}_{\mathbf{w}}(k)\mathbf{P}^-(k)\mathbf{F}_{\mathbf{w}}^H(k) + \mathbf{R} + \mathbf{\Lambda}(k) \tag{40}$$

$$\mathbf{K}(k) = \mathbf{P}^-(k)\mathbf{F}_{\mathbf{w}}^H(k)\mathbf{S}^{-1}(k) \tag{41}$$

$$\hat{\mathbf{w}}(k) = \hat{\mathbf{w}}(k-1) + \mathbf{K}(k)\mathbf{e}(k) \tag{42}$$

$$\mathbf{P}^+(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{F}_{\mathbf{w}}(k)]\mathbf{P}^-(k) \tag{43}$$

where $\mathbf{e}(k)$ and $\mathbf{S}(k)$ are the innovation vector and its covariance matrix. In the SOEKF, the Hessian matrices lead to the additional terms in the innovation $\pi(k)$ and its covariance matrix $\Lambda(k)$ under the MSE sense. The bias terms $\pi(k)$ and $\Lambda(k)$ in our problem can be expressed as

$$\boldsymbol{\pi}(k) = \frac{1}{2}\begin{bmatrix} 0 \\ \mathrm{tr}\{\breve{\mathbf{\Phi}}_s\mathbf{P}^-(k)\} \\ \mathrm{tr}\{\mathbf{P}^-(k)\} \end{bmatrix} \tag{44}$$

$$\begin{aligned} \mathbf{\Lambda}(k) &= \frac{1}{2} \\ &\quad \times \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathrm{tr}\{\breve{\mathbf{\Phi}}_s\mathbf{P}^-(k)\breve{\mathbf{\Phi}}_s\mathbf{P}^-(k)\} & \mathrm{tr}\{\breve{\mathbf{\Phi}}_s\mathbf{P}^-(k)\mathbf{P}^-(k)\} \\ 0 & \mathrm{tr}\{\breve{\mathbf{\Phi}}_s\mathbf{P}^-(k)\mathbf{P}^-(k)\} & \mathrm{tr}\{\mathbf{P}^-(k)\mathbf{P}^-(k)\} \end{bmatrix}. \end{aligned} \tag{45}$$

C. Solutions Using UKFs

The UKF uses *sigma points* to approximate the first- and second-order moments of the nonlinear transformation. There are different ways to set the *sigma points* and the weightings [20]–[22]. In this paper, we choose the method given by [20] since it gives positive weightings. The $(2M+1)$ *sigma points* for the approximation of the nonlinear measurement equations are generated by (46), shown at the bottom of the following page, and the transformed *sigma points* are given by

$$[\hat{\mathbf{Z}}_\sigma(k)]_i = \mathbf{f}([\hat{\mathbf{W}}_\sigma(k)]_i), \quad i = 0, \ldots, 2M \tag{47}$$

where $\mathbf{1}$ denotes the M-by-1 all-one vector. The UKF is summarized as (48)–(56) [20].

Multirank NCCB Using the UKF

$$\mathbf{P}^{-}(k) = \mathbf{P}^{+}(k-1) + \mathbf{Q} \tag{48}$$

$$\bar{\mathbf{z}}(k) = \frac{1}{2M+1}\sum_{i=0}^{2M}[\hat{\mathbf{Z}}_{\sigma}(k)]_i \tag{49}$$

$$\mathbf{P}_{\mathbf{wz}}(k) = \frac{1}{2M+1} \times \sum_{i=0}^{2M}([\hat{\mathbf{W}}_{\sigma}(k)]_i - \hat{\mathbf{w}}(k-1))([\hat{\mathbf{Z}}_{\sigma}(k)]_i - \bar{\mathbf{z}}(k))^H \tag{50}$$

$$\mathbf{P}_{\mathbf{zz}}(k) = \frac{1}{2M+1} \times \sum_{i=0}^{2M}([\hat{\mathbf{Z}}_{\sigma}(k)]_i - \bar{\mathbf{z}}(k))([\hat{\mathbf{Z}}_{\sigma}(k)]_i - \bar{\mathbf{z}}(k))^H \tag{51}$$

$$\mathbf{e}(k) = \mathbf{z} - \bar{\mathbf{z}}(k) \tag{52}$$

$$\mathbf{S}(k) = \mathbf{P}_{\mathbf{zz}}(k) + \mathbf{R} \tag{53}$$

$$\mathbf{K}(k) = \mathbf{P}_{\mathbf{wz}}(k)\mathbf{S}^{-1}(k) \tag{54}$$

$$\hat{\mathbf{w}}(k) = \hat{\mathbf{w}}(k-1) + \mathbf{K}(k)\mathbf{e}(k) \tag{55}$$

$$\mathbf{P}^{+}(k) = \mathbf{P}^{-}(k) - \mathbf{K}(k)\mathbf{S}(k)\mathbf{K}^{H}(k). \tag{56}$$

D. Initial Conditions

The setting of the initial conditions is important for constrained Kalman filtering problems. The initial conditions should satisfy the constraints or at least be close to the feasible sets of the constraint in the order of the chosen variance parameter in matrix $\mathbf{R}$. Improper settings can dramatically degrade the performance under nonlinear constraints. In our application, the distortionless constraint should be satisfied to maintain the array response. According to the decomposition of $\mathbf{w}$ in (22), it is suggested that the initial state be set as

$$\hat{\mathbf{w}}(0) = \mathbf{w}_{\text{matched}} \tag{57}$$

which strictly satisfies the distortionless constraint. On the other hand, for the initialization of the *a posteriori* state error covariance matrix $\mathbf{P}^{+}(k)$, we can refer to the update in the projection method for the rank-1 case [11], which gives a projection matrix of the null space of the steering vector

$$\mathbf{P}^{+}(0) = \mathbf{I} - \mathbf{a}_s\left[\mathbf{a}_s^H\mathbf{a}_s\right]^{-1}\mathbf{a}_s^H \tag{58}$$

Likewise, for the case with multirank signal models, $\mathbf{P}^{+}(k)$ can be initialized as

$$\mathbf{P}^{+}(0) = \mathbf{I} - \frac{\mathbf{w}_{\text{matched}}\mathbf{w}_{\text{matched}}^H}{\|\mathbf{w}_{\text{matched}}\|^2}. \tag{59}$$

E. Estimation of Parameter Matrices

In the update equations of the proposed Kalman filters, there are two parameter matrices to be determined: $\mathbf{Q}$ and $\mathbf{R}$. In general, $\mathbf{Q}$ stands for the *random walk* during the state update, which is typically assumed as stochastically white. For stationary environments, $\sigma_0^2 = 0$ can be chosen. The larger the parameter σ_0^2 chosen, the larger the allowed *random walk* of the state. i.e., the state variation can be large to track the nonstationary environmental changes. Second, $\mathbf{R}$ corresponds to the error variances of each measurement. In (30), σ_1^2 corresponds to the average output power. It is suggested to be the same order of the optimal output power of the array; however, the variance of the desired signal is not known *a priori* and it is related to the norm of the state $\mathbf{w}$. σ_2^2 and σ_3^2 correspond to the augmented distortionless and norm constraints. σ_2^2 controls the fitness of the distortionless constraint in (12). When σ_2^2 approaches zero, the beamformer approaches the matched filter in (18), which is distortionless to the presumed signal model but fails to reject the interferences. σ_3^2 controls the fitness of the norm constraint. The norm constraint controls the sidelobe level of the beamformer. When the norm-constraint value γ is small and σ_3^2 approaches zero, the beamformer emphasizes reducing the sidelobe level instead of rejecting interferences.

Since there is only one "true measurement" in the measurement equations, we propose only to estimate the parameter σ_1^2 and consider σ_2^2 and σ_3^2 as adjustable parameters to control the tradeoff between signal distortion and interference rejection. Contrary to the adaptive beamforming problems [16], [17], which suggested setting the variance parameters corresponding to the constraints as a very small value 10^{-10}, it is suggested in this dissertation to set a "properly" small value if there are some tradeoffs between the constraint sets. For such a case, if both σ_2^2 and σ_3^2 are set very small, the problem may not be feasible. For the uniform linear array (ULA) with ten elements and half-wavelength spacing, 10^{-5} is a good choice for the tradeoff between the distortionless and norm constraints. Further, to avoid the ill-conditioning problem, the condition number of the parameter matrix $\mathbf{R}$ should be controlled. In this case, we propose to use the following form:

$$\hat{\mathbf{R}} = \hat{\sigma}_1^2\begin{bmatrix}1 & 0 & 0\\ 0 & \kappa & 0\\ 0 & 0 & \kappa\end{bmatrix} \tag{60}$$

where $10^{-10} < \kappa \leq 1$ is the small variance set for the constraints. In this case, the condition number of $\hat{\mathbf{R}}$ is guaranteed to be smaller than 10^{10}.

The parameter σ_1^2 is estimated in a recursive way as

$$\hat{\sigma}_1^2(k) = \alpha\cdot\hat{\sigma}_1^2(k-1) + (1-\alpha)\cdot e_1(k)e_1^*(k) \tag{61}$$

where α is the forgetting factor close to unity and $e_1(k)$ is the first element of the innovation vector $\mathbf{e}(k)$. The recursive approach is similar to the work in [24].

$$\hat{\mathbf{W}}_{\sigma}(k) = \left[\underbrace{\hat{\mathbf{w}}(k-1)}_{M\times 1},\ \underbrace{\hat{\mathbf{w}}(k-1)\cdot\mathbf{1}^T + \sqrt{M\mathbf{P}^{-}(k)}}_{M\times M},\ \underbrace{\hat{\mathbf{w}}(k-1)\cdot\mathbf{1}^T - \sqrt{M\mathbf{P}^{-}(k)}}_{M\times M}\right] \tag{46}$$

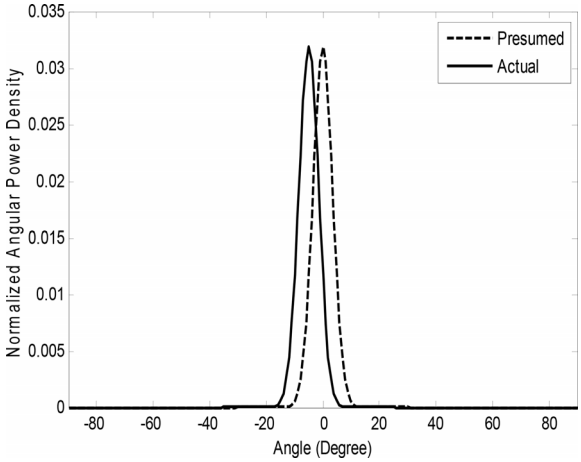

Fig. 1. Normalized angular distribution function $\rho(\theta)$ of the scattered source.

Compared to the variance estimation which used the *a priori* state error covariance matrix $\mathbf{P}^{-}(k)$ [25], [26], the recursive estimator in (61) depends on only the innovation and is not affected by the initial conditions. This guarantees the stability of the whole system. Typically, the parameter σ_1^2 is initialized as zero. For the simulations, $\alpha = 0.9$ is chosen.

V. Analysis of the Norm-Constraint Value

In this section, we show that the best selected norm-constraint value $\gamma = 0.0074$ is less sensitive to the input signal powers, central angle mismatches, and number of sensors than the selection of the DL level λ with a *standard linear array*. For all the analyses, the theoretical covariance matrices of the signals were used. The equivalent λ of a chosen γ can be derived from the equation $\|\mathbf{w}_{\mathrm{DL}}\|^2 = T$.

For Simulation 1, the sensitivity of λ and γ to the input signal powers are studied. A ULA of $M = 10$ sensors with half-wavelength spacing was used. The multirank signal model shown in Fig. 1 was utilized with presumed central angles equal to 0°. The scattered desired source and interference impinged into the array from the central angles -5° and 45°, respectively. Thus, a 5° central angle mismatch was considered. The sensor noise power was set to 1, and the interference-to-noise ratio (INR) was 30 dB. The signal-to-noise ratio (SNR) varies from -20 dB to 30 dB. Fig. 2 shows the output SINRs versus input SNRs for different selections of λ and γ. It can be seen that $\lambda = 10^2$ (note that the $\text{sensor noise power} = 1$) and $\gamma = 0.0074$ are good choices considering all input SNR conditions. These values will be used as the best choices for the rest of the simulations. For the selection of λ, higher values give more penalties on the spatially white noise (or incoherent noise), which leads to the matched filter being $\lambda \to \infty$. The matched filter does not form nulls in the directions of interferences; hence, it may have poor output SINR performance. As $\lambda \to 0$, this becomes the MVDR solution without norm constraint, which is sensitive to the array mismatches and has severe *self-cancellation* at high input SNRs. For the selection of γ, it can be expected that the variation of output SINR with different γ is relatively smaller than that with different λ due to normalization of the eigenvector. When the

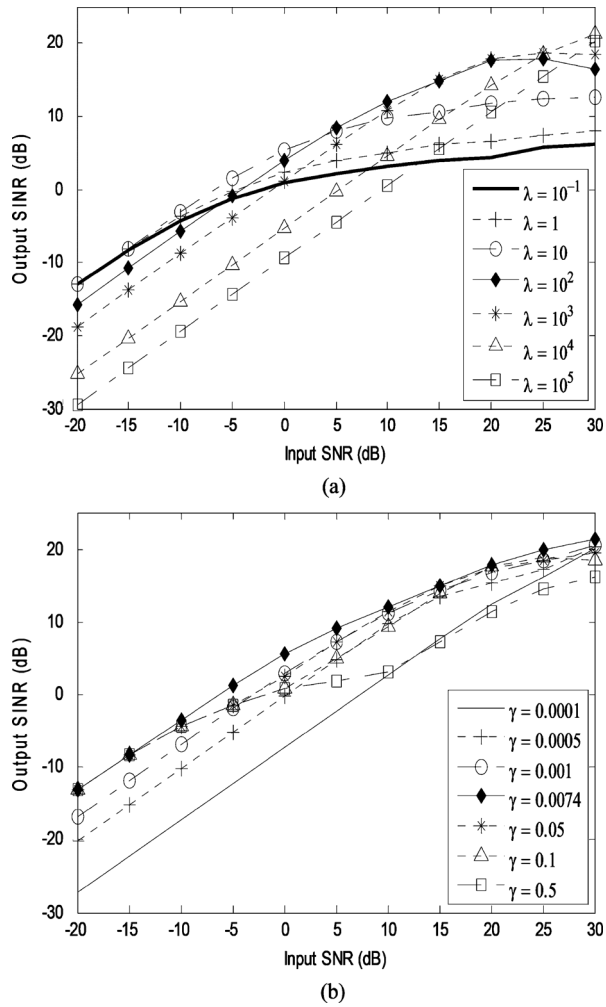

Fig. 2. Output SINR versus input SNR. (a) Different DL levels λ. (b) Different norm-constraint values γ.

norm-constraint value $\gamma \to 0$, the weight vector becomes the matched filter. When γ approaches the upper bound given in (18), it results in the standard Capon solution.

From Figs. 3 and 4, the optimal selections of λ and γ for different input SNR conditions are illustrated. In the figures, the star symbols mark the optimal selections for each case. It is obvious that the optimal selection of λ strongly depends on the input signal power, while the optimal selection of γ is less sensitive to the input signal power. Further, we found that the best selection of γ only depends on the number of sensors and does not change with different signal mismatch or noise conditions. Since the powers of the desired signals are unknown *a priori* and even time varying, the proposed robust beamformer provides consistent SINR performances due to the insensitivity of γ to the signal powers. Fig. 5 compares the best selected λ and γ for different input SNR conditions. Since the optimal choice of λ depends on the desired signal powers, the proposed norm-constrained robust beamformer with the optimal γ gives better output SINRs for most SNR conditions.

In Simulation 2, the comparison between the best selected λ and γ for different central angle mismatches is analyzed. For Simulations 2 and 3, the input SNR was set as 30 dB. It can be

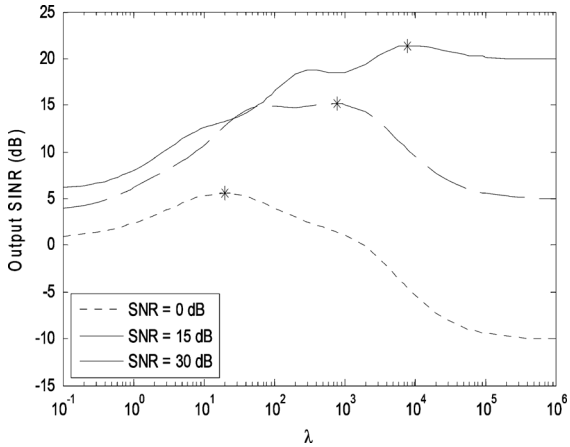

Fig. 3. Output SINR versus λ for different input SNRs.

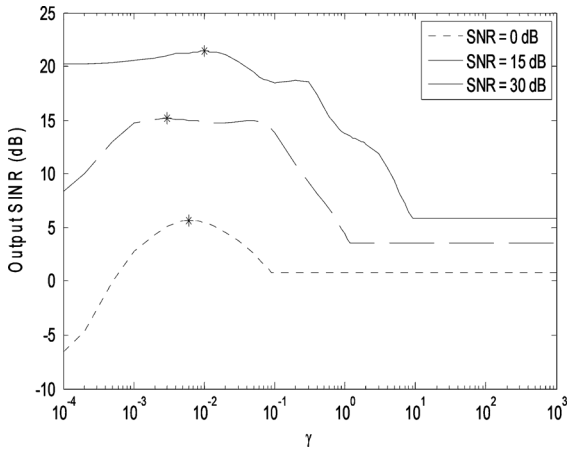

Fig. 4. Output SINR versus γ for different input SNRs.

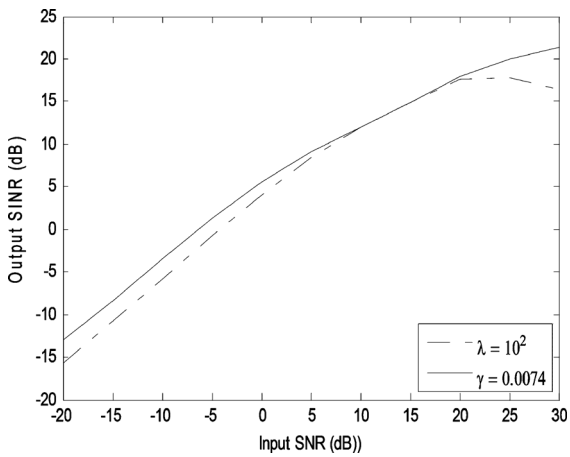

Fig. 5. Comparison between the best selected λ and γ for different input SNR conditions.

seen from Fig. 6 that for small angular mismatches, the proposed norm-constrained robust beamformer with the best selected γ performs better than the best selected λ. As the input SNR increases, the performance superiority becomes more obvious when a small angular mismatch exists. This is because the diagonal level becomes relatively small to the signal power and is insufficient to maintain the robustness. It is also worth noting that for low and moderate input SNR conditions, the method

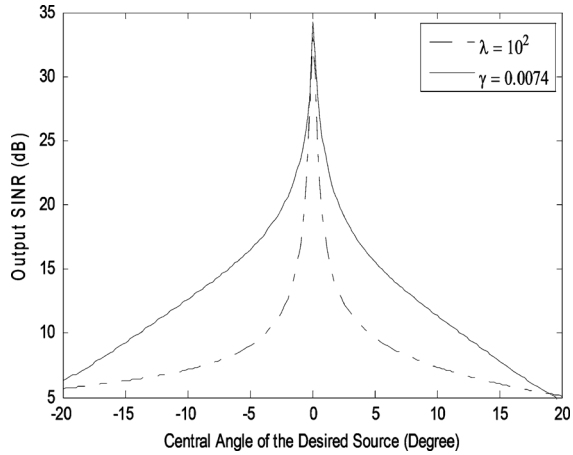

Fig. 6. Comparison between the best selected λ and γ for different central angle mismatches.

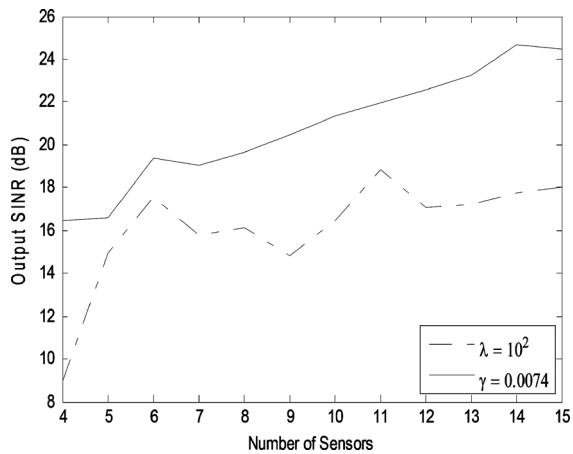

Fig. 7. Comparison between the best selected λ and γ for different numbers of sensors.

using norm constraint outperforms that using DL even when there is no angular mismatch. In this case, the improper large DL reduces the eigenvalue spread of the sample matrix that is not needed when there is no mismatch and hence degrades the output SINR performance.

In Simulation 3, the comparison between the best selected λ and γ for different numbers of sensors is investigated. Again, the superiority of using the norm-constrained robust beamformer is demonstrated in Fig. 7.

VI. Simulation Results

In this section, the comparisons of the beamformers are studied. The simulation condition is the same as in Simulation 1, except that the generated simulated data is used. For each scenario, the average of 100 simulation runs is used to obtain each data point. The detailed parameter settings and abbreviations of the algorithms are listed below:

1) MRSMI: Multirank sample matrix inverse [1]. The algorithm is implemented by (10).
2) MRLSMI: Multirank loaded sample matrix inverse [1]. The algorithm is implemented by (11), where the DL level is chosen as $\lambda = 10^2$ (note that the $\text{sensor noise power} = 1$).

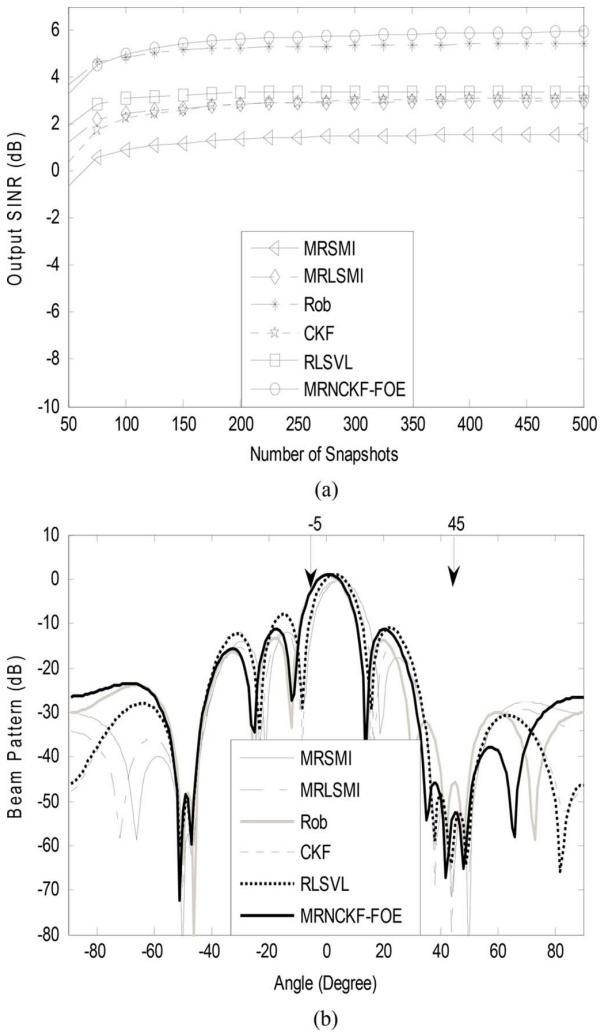

Fig. 8. Comparisons of the beamformers at 0 dB input SNR. (a) Output SINR versus training size. (b) Beam pattern.

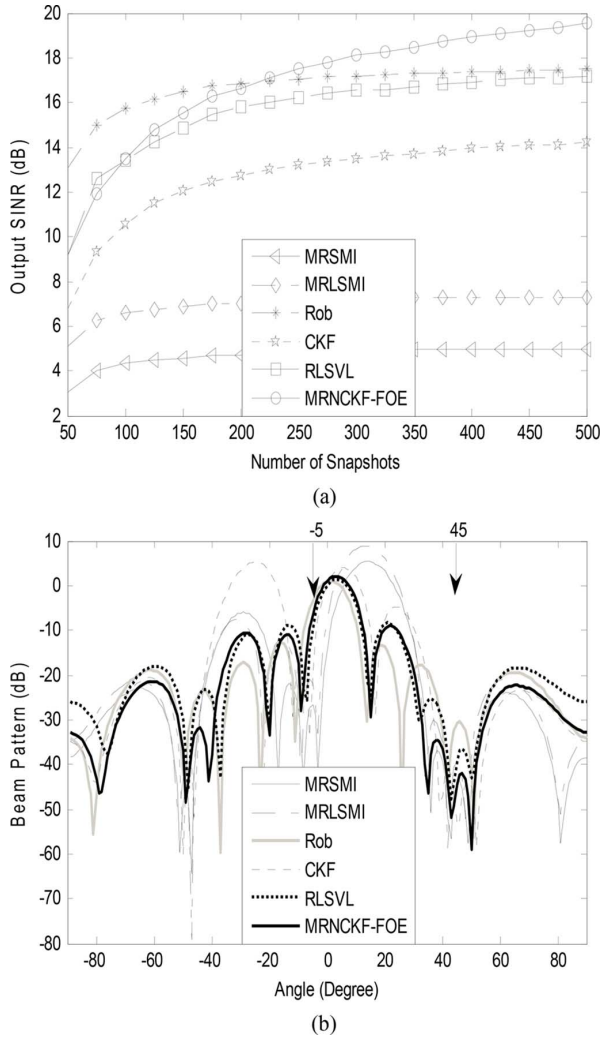

Fig. 9. Comparisons of the beamformers at 20 dB input SNR. (a) Output SINR versus training size. (b) Beam pattern.

3) Rob: Robust adaptive filter given by (53) in [1]. The positive and negative DLs are chosen as $\lambda = 10^2$ and $\varepsilon = 6$. Note that for the selection of ε, the multirank signal model is normalized as (13). The selection of ε is smaller than the maximal eigenvalue of the normalized signal model to guarantee the feasibility of the robust beamformer [1].
4) CKF: Constrained Kalman filter [16]. The CKF uses the rank-1 model without the norm constraint and constructs the state space using the penalty method. The parameter matrices $\mathbf{Q}$ is set to a zero matrix, and the diagonal terms of $\mathbf{R}$ are set as $\hat{\sigma}_1^2$ and $10^{-5} \cdot \hat{\sigma}_1^2$, where $\hat{\sigma}_1^2$ is estimated using (61) with the forgetting factor $\alpha = 0.9$.
5) RLSVL: Recursive least square with variable loading [27]. The RLSVL uses the rank-1 model and the GSC structure. A norm constraint is imposed on the nulled vector to improve the robustness. The forgetting factor and the norm constraint are set as 0.999 and 0.2 as suggested in paper [27].
6) MRNCKF-FOE: Multirank norm-constrained first-order extended Kalman filter. The proposed beamformer is implemented using (38)–(43) by letting both $\pi(k)$ and $\Lambda(k)$ be zero. The parameter matrix $\mathbf{Q}$ is set to a zero matrix, and $\mathbf{R}$ is set as in (60) with $\kappa = 10^{-5}$ for all the proposed Kalman filters, where $\hat{\sigma}_1^2$ is estimated using (61) with the forgetting factor $\alpha = 0.9$. The norm constraint value $\gamma = 0.0074$ is chosen.
7) MRNCKF-SOE: Multirank norm-constrained second-order extended Kalman filter. The proposed beamformer is implemented using (38)–(43) with the second-order terms $\pi(k)$ and $\Lambda(k)$.
8) MRNCKF-U: Multirank norm-constrained unscented Kalman filter. The proposed beamformer is implemented using (46)–(56).

For the first case, the convergences and beam patterns at $\mathrm{SNR} = 0$ dB are studied. Fig. 8(a) shows the output SINR performance versus the training size. The presence of the desired signal causes the SINR performance to deteriorate due to the *self-cancellation* phenomenon, which can be observed in Fig. 8(b) around the central angle of the desired source $-5°$. Considering the performance of the pairs (MRSMI, MRLSMI) and (CKF, RLSVL), it can be observed that the norm constraint improves the SINR performance. From the comparison of the MRLSMI and Rob beamformer, the negative loading in Rob clearly improves the SINR performance. In this simulation,

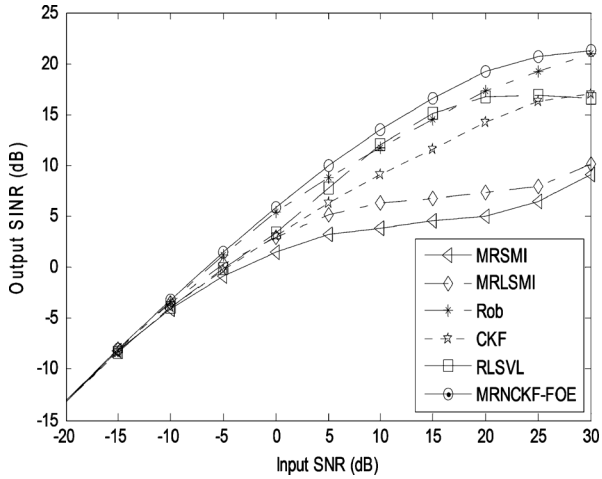

Fig. 10. Output SINR versus input SNR for different beamformers.

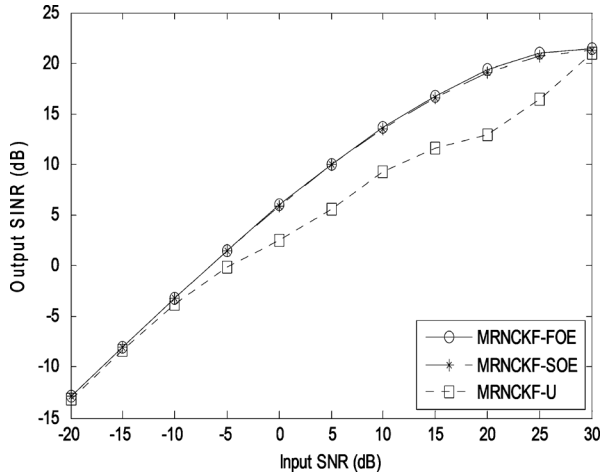

Fig. 11. Output SINR versus input SNR for proposed Kalman filters.

the proposed robust beamformer, MRNCKF-FOE, gives the best performance among all the algorithms. It is worth noting that the Kalman filter solutions seem to be more robust to the steering mismatches than the beamformers using the estimation of sample matrix. The Kalman filter is a close-loop system which constrains the weight vector to the desired array response in each iteration. On the other hand, the sample matrix inverse method is an open-loop system which constrains the weight vector after the sample matrix is estimated. Therefore, the latter can be easily affected by the contaminated sample matrix since the importance of the training data is the same when estimating the sample matrix. Compared with the Rob [1], the proposed method provides a more flexible structure so that additional constraints can be easily imposed for other problems. Besides, the eigenvalue decomposition is not needed. From the beam pattern, it is shown that the proposed MRNCKF-FOE, gives the best output SINR since it has the smallest signal distortion at -5° while keeping the same order of noise rejection at 45°.

For the second case, the convergences and beam patterns at $\mathrm{SNR} = 20$ dB are studied. In Fig. 9(a), the large signal power slows down the convergence of the algorithms. The strong signal power leads to larger *self-cancellation* for the MRSMI and MRLSMI beamformers. Despite the difference between the Kalman filter and the one using the sample matrix, it also reveals that the chosen DL level of the MRLSMI beamformer is not appropriate under this SNR condition (see Fig. 3). This demonstrates the advantage of using the norm constraint with a more robust selection of the norm-constraint value. The negative loading skill used in the Rob beamformer greatly compensates for the problem of selection of λ. However, the performance of the proposed beamformer outruns the Rob beamformer as the number of iterations grows.

For the third case, the output SINR of the beamformers versus the input SNR is illustrated in Fig. 10. The training size of this simulation is $N = 500$. When the input SNR is small, all the beamformers converge to the optimal MVDR solution. As the SNR increases, the differences between the algorithms become obvious. It is shown that the proposed beamformer has the best performance through different SNR conditions.

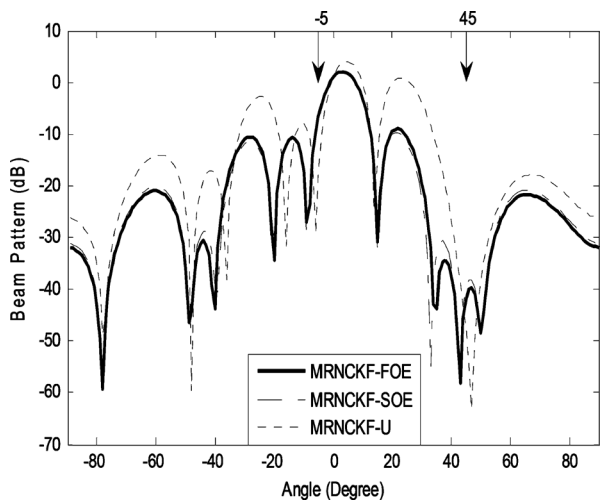

Fig. 12. Beam patterns of the proposed Kalman filters at 20-dB input SNR.

For the last case, the output SINR of the proposed Kalman filter solutions versus the input SNR is illustrated in Fig. 11. It can be seen that the performances of the FOEKF and SOEKF are almost the same. Compared to the extended solutions, the UKF with the *sigma points* suggested in [20] has a worse SINR performance. The UKF implicitly estimates the first- and second-order approximation terms of the Taylor expansion using *sigma points*. The *sigma points* were spread based on $\sqrt{M}$ times eigenvectors of the error covariance $\mathbf{P}^{-}(k)$. An issue for the spreading of *sigma points* is invoked when some error dominates the covariance $\mathbf{P}^{-}(k)$. In this case, some *sigma points* are spread far away from the constraint sets and the neighborhood of the current state estimate, which can induce improper nonlinear transformations that degrade the performance of the UKF. Tuning the sigma points for UKF could possibly give rise to a better performance, but it may require new algorithms that are beyond the scope of this paper. Besides, the results indicate that the first-order extended solution is good enough for our problem. In Fig. 12, it can be seen that the large error of interference rejection enforces noise reduction at both 45° and -5°, which results in *self-cancellation*.

VII. Conclusion

In this paper, we proposed a robust beamformer with multirank signal models and the Kalman filter. Compared to the use of the DL form, we showed that designing a standard linear array with the norm-constraint value is less sensitive to unknown signal powers, small angle mismatches, and the number of sensors. Further, in comparison with the selection of the DL level, the best selection of the norm-constraint value γ only depends on the number of sensors and is robust to different signal mismatch or noise conditions. Besides, the simulation results indicate that the designs using the Kalman filter have less *self-cancellation* than the estimations of the sample matrix.

References

[1] S. Shahbazpanahi, A. B. Gershman, Z. Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2257–2269, Sep. 2003.

[2] A. Pezeshki, B. D. Van Veen, L. L. Scharf, H. Cox, and M. L. Nordenvaad, "Eigenvalue beamforming using a multirank MVDR beamformer and subspace selection," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1954–1967, May 2008.

[3] L. Zhang and W. Liu, "Robust forward backward based beamformer for a general-rank signal model with real-valued implementation," *Signal Process.*, vol. 92, no. 1, pp. 163–169, Jan. 2012.

[4] S. Valaee, B. Champagne, and P. Kabal, "Parametric localization of distributed sources," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2144–2153, Sep. 1995.

[5] A. Zoubir, Y. Wang, and P. Charge, "Efficient subspace-based estimator for localization of multiple incoherently distributed sources," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 532–542, Feb. 2008.

[6] H. L. Van Trees, *Optimum Array Processing*. Hoboken, NJ, USA: Wiley, 2002.

[7] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1365–1375, Oct. 1987.

[8] J. E. Hudson, *Adaptive Array Principles*. London, U.K.: Peter Peregrinus, 1981.

[9] 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, Jul. 2003.

[10] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust Capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, Sep. 2004.

[11] D. Simon and T. L. Chia, "Kalman filtering with state equality constraints," *IEEE Trans. Aerospace Electron. Syst.*, vol. 38, no. 1, pp. 128–136, Jan. 2002.

[12] S. J. Julier and J. J. LaViola, "On Kalman filtering with nonlinear equality constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2774–2784, Jun. 2007.

[13] C. Yang and E. Blasch, "Kalman filtering with nonlinear state constraints," *IEEE Trans. Aerospace Electron. Syst.*, vol. 45, no. 1, pp. 70–84, Jan. 2009.

[14] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory Appl.*, vol. 4, no. 8, pp. 1303–1318, Aug. 2010.

[15] R. J. Hewett, M. T. Heath, M. D. Butala, and F. Kamalabadi, "A robust null space method for linear equality constrained state estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 3961–3971, Aug. 2010.

[16] Y. H. Chen and C. T. Chiang, "Adaptive beamforming using the constrained Kalman filter," *IEEE Trans. Antennas Propag.*, vol. 41, no. 11, pp. 1576–1580, Nov. 1993.

[17] A. El-Keyi, T. Kirubarajan, and A. B. Gershman, "Robust adaptive beamforming based on the Kalman filter," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3032–3041, Aug. 2005.

[18] S. Schmidt, "Application of state-space methods to navigation problems," *Adv. Control Syst.*, pp. 293–340, 1966.

[19] M. Athans, R. P. Wishner, and A. Bertolini, "Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements," *IEEE Trans. Autom. Control*, vol. 13, no. 5, pp. 504–514, Oct. 1968.

[20] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, NJ, USA: Wiley, 2006.

[21] S. J. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proc. AeroSense: 11th Int. Symp. Aerospace/Defense Sensing, Simulation and Controls*, 1997, pp. 182–193.

[22] F. Gustafsson and G. Hendeby, "Some relations between extended and unscented Kalman filters," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 545–555, Feb. 2012.

[23] I. Cohen, "Multichannel post-filtering in nonstationary noise environments," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1149–1160, May 2004.

[24] P. K. Dash, S. Hasan, and B. K. Panigrahi, "Adaptive complex unscented Kalman filter for frequency estimation of time-varying signals," *IET Science, Meas. Technol.*, vol. 4, no. 2, pp. 93–103, 2010.

[25] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering," *IEEE Trans. Autom. Control*, vol. 15, no. 2, pp. 175–184, Apr. 1970.

[26] K. Myers and B. D. Tapley, "Adaptive sequential estimation with unknown noise statistics," *IEEE Trans. Autom. Control*, vol. 21, no. 4, pp. 520–523, Aug. 1976.

[27] Z. Tian, K. L. Bell, and H. L. Van Trees, "A recursive least squares implementation for LCMP beamforming under quadratic constraint," *IEEE Trans. Signal Process.*, vol. 49, no. 6, pp. 1138–1145, Jun. 2001.

Jwu-Sheng Hu (M'94) received the B.S. degree from the Department of Mechanical Engineering, National Taiwan University, Taiwan, R.O.C., in 1984, and the M.S. and Ph.D. degrees from the Department of Mechanical Engineering, University of California, Berkeley, CA, USA, in 1988 and 1990, respectively.

From 1991–1993, he was an Assistant Professor in the Department of Mechanical Engineering, Wayne State University, Detroit, MI, USA, where he received the Research Initiation Award from National Science Foundation. Since 1993, he joined the Department of Electrical Engineering, National Chiao Tung University, Taiwan, R.O.C., and became a full Professor in 1998. He became Vice-Chairman of the Department in 2006. Since 2008, he has worked in-part at the Industrial Technology Research Institute (ITRI) of Taiwan where now he serves as the Technical Director of the Mechanical and System Lab as well as the Intelligent Robotics Division Director. He also served as a research faculty at the National Chip Implementation Center (CIC) of Taiwan for embedded system design applications. His current research interests include robotics, microphone array, mechatronics, and embedded systems.

Dr. Hu is currently an associate editor of the *ASME Journal of Dynamic Systems, Measurement and Control*. He also served as the Chairman of the Taipei Chapter, IEEE RA Society.

Ming-Tang Lee was born in Taoyuan, Taiwan, R.O.C., in 1984. He received the B.S., M.S., and Ph.D. degrees from the Department of Electrical and Control Engineering, National Chiao Tung University, Taiwan, R.O.C., in 2007, 2008, and 2013, respectively.

He is currently with Realtek Inc., Taiwan, R.O.C., as a Design Engineer for speech signal processing. His research interests include array signal processing, adaptive signal processing, detection and estimation, and speech enhancement.