Iterative Sparse Asymptotic Minimum Variance Based Approaches for Array Processing*

Habti Abeida, Qilin Zhang, and Jian Li, Fellow, IEEE

Abstract—This paper presents a series of user parameter-free iterative Sparse Asymptotic Minimum Variance (SAMV) approaches for array processing applications based on the asymptotically minimum variance (AMV) criterion. With the assumption of abundant snapshots in the direction-of-arrival (DOA) estimation problem, the signal powers and noise variance are jointly estimated by the proposed iterative AMV approach, which is later proved to coincide with the Maximum Likelihood (ML) estimator. We then propose a series of power-based iterative SAMV approaches, which are robust against insufficient snapshots, coherent sources and arbitrary array geometries. Moreover, to overcome the direction grid limitation on the estimation accuracy, the SAMV-Stochastic ML (SAMV-SML) approaches are derived by explicitly minimizing a closed form stochastic ML cost function with respect to one scalar parameter, eliminating the need of any additional grid refinement techniques. To assist the performance evaluation, approximate solutions to the SAMV approaches are also provided for high signal-to-noise ratio (SNR) and low SNR scenarios. Finally, numerical examples are generated to compare the performances of the proposed approaches with those of the existing ones.

Index Terms—Array Processing, Asymptotically Minimum Variance Estimator, Direction-Of-Arrival (DOA) Estimation, Sparse AMV Estimation.

I. INTRODUCTION

Sparse signal representation has attracted a lot of attention in recent years and it has been successfully used for solving inverse problems in various applications such as channel equalization (e.g., [1–4]), source localization (e.g., [5], [6]) and radar imaging (e.g., [7]–[9]). In its basic form, it attempts to find the sparsest signal vector satisfying the constraint $y = Ax + e$ where $A \in \mathbb{C}^{M \times K}$ is an overcomplete basis (i.e., $K > M$), $y$ is the observation data, and $e$ is the noise term. Theoretically, this problem is underdetermined and has multiple solutions. However, the additional constraint that $x$ should be sparse allows one to eliminate the ill-posedness problem (e.g., [10], [11]). In recent years, a number of practical algorithms such as $\ell_1$ norm minimization (e.g., [12], [13]) and focal underdetermined system solution (FOCUSS) (e.g., [14], [15]) have been proposed to approximate the sparse solution.

Conventional subspace-based source localization algorithms such as multiple signal classification (MUSIC) and estimation of signal parameters via a rotational invariance technique (ESPRIT) [16], [17] are only applicable when $M > K$, and they require sufficient snapshots and high signal-to-noise ratio (SNR) to achieve high spatial resolution. However, it is often impractical to collect a large number of snapshots, especially in fast time-varying environment, which deteriorates the construction accuracy of the subspaces and degrades the localization performance. In addition, even with appropriate array calibration, subspace-based methods are incapable of handling the source coherence problem due to their sensitivity to subspace orthogonality (e.g., [16]).

Recently, a user parameter-free non-parametric algorithm, referred to as the iterative adaptive approach (IAA), has been proposed in [6] and employed in various applications (e.g., [8], [9]). It is demonstrated in these works that the weighted least-squares fitting-based IAA algorithm provides accurate DOA and signal power estimates, and it is insensitive to practical impairments such as few (even one) snapshots, arbitrary array geometries and coherent sources. The iterative steps are based on the IAA covariance matrix, which could be singular in the noise-free scenarios when only a few components of the signal power vector are non-zero. To mitigate this problem, a regularized version of the IAA algorithm (IAA-R) is later proposed in [9] for single-snapshot and nonuniform white noise cases. Stoica et al. have recently proposed a user parameter-free SParse Iterative Covariance-based Estimation (SPICE) approach in [18], [19] based on minimizing a covariance matrix fitting criterion. We note that the source localization performance of the power-based algorithms is mostly limited by the fineness of the direction grid size [5].

In this paper, we propose a series of iterative Sparse Asymptotic Minimum Variance (SAMV) approaches based on the asymptotically minimum variance (AMV) approach (also called asymptotically best consistent (ABC) estimators in [20]), which is initially proposed for DOA estimation in [21], [22]. After presenting the sparse signal representation data model for the DOA estimation problem in Section II, we first propose an iterative AMV approach in Section III, which is later proven to be identical to the stochastic Maximum Likelihood (ML) estimator. Based on this approach, we then propose the user parameter-free iterative SAMV approaches that can handle arbitrary number of snapshots ($N < M$ or $N > M$), and only a few non-zero components in the signal power vector in Section IV. In addition, a series of SAMV-Stochastic ML (SAMV-SML) approaches are proposed in Section V to alleviate the direction grid limitation and enhance...
the performance of the power-based SAMV approaches. In Section VI, we derive approximate expressions for the SAMV power-iteration formulas for both high and low SNR scenarios. In Section VII, numerical examples are generated to compare the performances of the proposed approaches with those of the existing ones. Finally, conclusions are given in Section VIII.

The following notations are used throughout the paper. Matrices and vectors are represented by bold upper case and bold lower case characters, respectively. Vectors are by default in column orientation, while $T$, $H$, and $*$ stand for transpose, conjugate transpose, and complex conjugate, respectively. $E(\cdot)$, $\text{Tr}(\cdot)$ and $\text{det}(\cdot)$ are the expectation, trace and determinant operators, respectively. $\text{vec}(\cdot)$ is the “vectorization” operator that turns a matrix into a vector by stacking all columns on top of one another, $\otimes$ denotes the Kronecker product, $I$ is the identity matrix of appropriate dimension, and $u_m$ denotes the $m$th column of $I$.

II. PROBLEM FORMATION AND DATA MODEL

Consider an array of $M$ omnidirectional sensors receiving $\hat{K}$ narrowband signals impinging from the sources located at $\theta_{k} \triangleq (\theta_{k1}, \ldots, \theta_{K})$ where $\theta_{k}$ denotes the location parameter of the $k$th signal, $k = 1, \ldots, \hat{K}$. The $M \times 1$ array snapshot vectors can be modeled as (see e.g., [6], [19])

$$y(n) = Ax(n) + e(n), \quad n = 1, \ldots, N,$$

where $A \triangleq [a(\theta_1), \ldots, a(\theta_{\hat{K}})]$ is the steering matrix with each column being a steering vector $a_k \triangleq a(\theta_k)$, a known function of $\theta_k$. The vector $x(n) \triangleq [x_1(n), \ldots, x_{\hat{K}}(n)]^T$ contains the source waveforms, and $e(n)$ is the noise term. Assume that $E(e(n)e^H(n)) = \sigma I_M \delta_n \delta_k$, where $\delta_n \delta_k$ is the Kronecker delta and it equals to 1 only if $n = n$ and $k = k$ otherwise. We also assume first that $e(n)$ and $x(n)$ are independent of each other, and that $E(x(n)x^H(n)) = P \delta_n \delta_k$, where $P \triangleq \text{Diag}(p_1, \ldots, p_{\hat{K}})$. Let $p$ be a vector containing the unknown signal powers and noise variance, $p_0 \triangleq [p_1, \ldots, p_{\hat{K}}, \sigma]^T$.

The covariance matrix of $y(n)$ that conveys information about $p$ is given by

$$R \triangleq APA^H + \sigma I.$$

This covariance matrix is traditionally estimated by the sample covariance matrix $R_N \triangleq YY^H/N$ where $Y \triangleq [y(1), \ldots, y(N)]$. After applying the vectorization operator to the matrix $R$, the so-obtained vector $r(p) \triangleq \text{vec}(R)$ is linearly related to the unknown parameter $p$ as

$$r(p) \triangleq \text{vec}(R) = Sp,$$

where $S \triangleq [S_1, \bar{a}_{\hat{K}+1}]$, $S_1 = [\bar{a}_1, \ldots, \bar{a}_{\hat{K}}]$, $\bar{a}_k \triangleq a_k \otimes a_k$, $k = 1, \ldots, \hat{K}$, and $\bar{a}_{\hat{K}+1} \triangleq \text{vec}(I)$.

We note that the Gaussian circular asymptotic covariance matrix $r_N \triangleq \text{vec}(R_N)$ is given by [23, Appendix B], [22]

$$C_r = R^* \otimes R.$$

The number of sources, $\hat{K}$, is usually unknown. The power-based algorithms, such as the proposed SAMV approaches, use a predefined scanning direction grid $\{\theta_k\}_{k=1}^{\hat{K}}$ to cover the entire region-of-interest $\Omega$, and every point in this grid is considered as a potential source whose power is to be estimated. Consequently, $\hat{K}$ is the number of points in the grid and it is usually much larger than the actual number of sources present, and only a few components of $p$ will be non-zero. This is the main reason why sparse algorithms can be used in array processing applications [18], [19].

To estimate the parameter $p$ from the statistic $r_N$, we develop a series of iterative SAMV approaches based on the AMV approach introduced by Porat and Fridelander in [24], Stoica et al. in [20] with their asymptotically best consistent (ABC) estimator, and Delmas and Abeida in [21], [22].

III. THE ASYMMETRICALLY MINIMUM VARIANCE APPROACH

In this section, we develop a recursive approach to estimate the signal powers and noise variance (i.e., $p$) based on the AMV criterion using the statistic $r_N$. We assume that $p$ is identifiable from $r(p)$. Exploiting the similarities to the works in [21], [22], it is straightforward to prove that the covariance matrix $\text{Cov}_p^{\text{Alg}}$ of an arbitrary consistent estimator of $p$ based on the second-order statistic $r_N$ is bounded below by the following real-valued symmetric positive definite matrix:

$$\text{Cov}_p^{\text{Alg}} \geq [S_d C_r^{-1} S_d]^T,$$

where $S_d \triangleq \text{dr}(p)/\text{dp}$. In addition, this lower bound is attained by the covariance matrix of the asymptotic distribution of $\hat{p}$ obtained by minimizing the following AMV criterion:

$$\hat{p} = \arg \min_p f(p),$$

where

$$f(p) \triangleq \text{vec}(r_N - r(p))^H C_r^{-1} [r_N - r(p)].$$

From (3) and by using (2), the estimate of $p$ is given by the following results proved in Appendix A:

**Result 1.** The $\{\hat{p}_k\}_{k=1}^{\hat{K}}$ and $\hat{\sigma}$ that minimize (3) can be computed iteratively. Assume $\hat{p}_k^{(i)}$ and $\hat{\sigma}_i^{(i)}$ have been obtained in the $i$th iteration, they can be updated at the $(i+1)$th iteration as:

$$\hat{p}_{k}^{(i+1)} = \frac{a_k^H R^{-1(i)} r_N R^{-1(i)} a_k}{a_k^H R^{-1(i)} a_k} + \hat{p}_k^{(i)} - \frac{1}{a_k^H R^{-1(i)} a_k}, \quad k = 1, \ldots, \hat{K},$$

$$\hat{\sigma}_i^{(i+1)} = \left( \text{Tr}(R^{-2(i)} r_N) + \hat{\sigma}_i^{(i)} \text{Tr}(R^{-2(i)} - \text{Tr}(R^{-1(i)})) \right) / \text{Tr}(R^{-2(i)}),$$

where the estimate of $R$ at the $i$th iteration is given by $R(i) = A P(i) A^H + \hat{\sigma}_i(i) I$ with $P(i) = \text{Diag}(\hat{p}_1^{(i)}, \ldots, \hat{p}_{\hat{K}}^{(i)})$.

Assume that $x(n)$ and $e(n)$ are both circularly Gaussian distributed, $y(n)$ also has a circular Gaussian distribution with zero-mean and covariance matrix $R$. The stochastic negative log-likelihood function of $\{y(n)\}_{n=1}^N$ can be expressed as (see, e.g., [6], [25])

$$L(p) = \ln(\text{det}(R)) + \text{Tr}(R^{-1} r_N).$$
In lieu of the cost function (3) that depends linearly on \( p \) (see (2)), this ML cost-function (6) depends non-linearly on the signal powers and noise variance embedded in \( R \). Despite this difficulty and reminiscent of [6], we prove in Appendix B that the following result holds:

**Result 2.** The estimates given by (4) and (5) are identical to the ML estimates.

Consequently, there always exist approaches that give the same performance as the ML estimator which is asymptotically efficient.

Returning to the Result 1, we notice that the expression (4) remains valid regardless of \( K > M \) or \( K < M \). In the scenario where \( K > M \), we observe from numerical calculations that the \( \hat{p}_k \) and \( \hat{\sigma} \) given by (4) and (5) may be negative; therefore, the nonnegativity of the power estimates can be enforced similar to [6, Eq. (30)],

\[
\hat{p}_{k}^{(i+1)} = \max \left( 0, \frac{a_k^H R^{-1}(i) R_N R^{-1}(i) a_k}{(a_k^H R^{-1}(i) a_k)^2} \right) + \hat{p}_k^{(i)} - \frac{1}{a_k^H R^{-1}(i) a_k}, \quad k = 1, \ldots, K,\]

\[
\hat{\sigma}_{(i+1)} = \max \left( 0, \frac{\text{Tr}(R^{-2}(i) R_N) + \hat{\sigma}_{(i)} \text{Tr}(R^{-2}(i)) - \text{Tr}(R^{-2}(i))}{\text{Tr}(R^{-2}(i))} \right),
\]

The above updating formulas of \( \hat{p}_k \) and \( \hat{\sigma} \) at the \((i+1)\)th iteration require knowledge of \( \hat{p}_k \) and \( \hat{\sigma} \) at the \( i \)th iteration, hence this algorithm must be implemented iteratively. The initialization of \( \hat{p}_k \) can be done with the periodogram (PER) power estimates (see, e.g., [26]),

\[
\hat{p}_{k,\text{PER}}^{(0)} = \frac{a_k^H R_N a_k}{\|a_k\|^2}, \quad \text{(8)}
\]

The noise variance estimator \( \hat{\sigma} \) can be initialized as, for instance,

\[
\hat{\sigma} = \frac{1}{MN} \sum_{n=1}^{N} \| y(n) \|^2. \quad \text{(9)}
\]

**Remark 1.** In the classical scenario where there are more sensors than sources (i.e., \( K \leq M \)), closed form approximate ML estimates of a single source power and noise variance are derived in [27] and [28] assuming uniform white noise and nonuniform white noise, respectively. However, these approximate expressions are derived at high and low SNR regimes separately\(^2\), compared to the unified expressions (4) and (5) regardless of SNR or the number of sources.

**Remark 2.** Result 1 can be extended easily to the nonuniform white Gaussian noise case where the covariance matrix is given by

\[
E\left( ee^H(n)\right) = \text{Diag}(\sigma_1, \ldots, \sigma_M) = \sum_{m=1}^{M} \sigma_m a_m a_m^H, \quad \text{(10)}
\]

where \( a_m = u_m, \ m = 1, \ldots, M, \) denote the canonical vectors. Under these assumptions and from Result 1, the estimate of \( p \) at the \((i+1)\)th iteration is given by

\[
\hat{p}_k^{(i+1)} = \frac{a_k^H R^{-1}(i) R_N R^{-1}(i) a_k}{(a_k^H R^{-1}(i) a_k)^2} + \hat{p}_k^{(i)} - \frac{1}{a_k^H R^{-1}(i) a_k}, \quad k = 1, \ldots, K, \quad \text{(11)}
\]

where \( R^{(i)} = \mathbf{A} P^{(i)} \mathbf{A}^H + \sum_{m=1}^{M} \hat{\sigma}_m a_m a_m^T + \sigma^2 \mathbf{I} \), \( \mathbf{P}^{(i)} = \text{Diag}(\hat{p}_1^{(i)}, \ldots, \hat{p}_K^{(i)}) \) and \( \hat{\sigma}_m = \hat{p}_k^{(i)} \), \( m = 1, \ldots, M \).

As mentioned before, the \( \hat{p}_k \) may be negative when \( K > M \), therefore, the power estimates can be iterated similar to (7) by forcing the negative values to zero.

**IV. The Sparse Asymptotic Minimum Variance Approaches**

In this section, we propose the iterative SAMV approaches to estimate \( p \) even when the number of grid points \( K \) exceeds the number of sources (i.e., when the steering matrix \( \mathbf{A} \) can be viewed as an overcomplete basis for \( y(n) \)) and only a few non-zero components are present in \( p \). This is the common case encountered in many spectral analysis applications, where only the estimation of \( p \) is deemed relevant (e.g., [18], [19]).

As mentioned in Result 2, the estimates given by (4) and (5) may give iterational negative values due to the presence of the non-zero terms \( p_k - 1/(a_k^H a_k) \) and \( \sigma_{(i)} \text{Tr}(R^{-1}(i)/\text{Tr}(R^{-2})) \). To resolve this difficulty, let\(^3\) \( p_k = 1/(a_k^H a_k) \) and \( \sigma = \text{Tr}(R^{-1})/\text{Tr}(R^{-2}) \), and propose the following SAMV approaches based on Result 1:

**SAMV-0 approach:**

The estimates of \( p_k \) and \( \sigma \) are updated at the \((i+1)\)th iteration as:

\[
\hat{p}_k^{(i+1)} = \frac{a_k^H R^{-1}(i) R_N R^{-1}(i) a_k}{(a_k^H R^{-1}(i) a_k)^2}, \quad k = 1, \ldots, K, \quad \text{(12)}
\]

\[
\hat{\sigma}_{(i+1)} = \frac{\text{Tr}(R^{-2}(i) R_N)}{\text{Tr}(R^{-2}(i))}, \quad \text{(13)}
\]

**SAMV-1 approach:**

The estimates of \( p_k \) and \( \sigma \) are updated at the \((i+1)\)th iteration as:

\[
\hat{p}_k^{(i+1)} = \frac{a_k^H R^{-1}(i) R_N R^{-1}(i) a_k}{(a_k^H R^{-1}(i) a_k)^2}, \quad k = 1, \ldots, K, \quad \text{(14)}
\]

\[
\hat{\sigma}_{(i+1)} = \frac{\text{Tr}(R^{-2}(i) R_N)}{\text{Tr}(R^{-2}(i))}. \quad \text{(15)}
\]

**SAMV-2 approach:**

The estimates of \( p_k \) and \( \sigma \) are updated at the \((i+1)\)th iteration as:

\[
\hat{p}_k^{(i+1)} = \frac{a_k^H R^{-1}(i) R_N R^{-1}(i) a_k}{(a_k^H R^{-1}(i) a_k)^2}, \quad k = 1, \ldots, K, \quad \text{(16)}
\]

\[
\hat{\sigma}_{(i+1)} = \frac{\text{Tr}(R^{-2}(i) R_N)}{\text{Tr}(R^{-2}(i))}. \quad \text{(17)}
\]

In the case of nonuniform white Gaussian noise with covariance matrix given in Remark 2, the SAMV noise power estimates can be updated alternatively as

\[
\hat{\sigma}_{m} = \frac{u_m^H R^{-1}(i) R_N R^{-1}(i) u_m (u_m^H R^{-1}(i) u_m)^2}, \quad m = 1, \ldots, M, \quad \text{(17)}
\]

\(^2\)For high and low SNR scenarios, the ML function (6) is linearized by different approximations in [27] and [28].

\(^3\)Different substitutions are applied in the SAMV algorithms resulting different levels of sparsity and robustness, which is further demonstrated and analysed in Section VII.
where $R^{(i)} = AP^{(i)}A^H + \sum_{m=1}^M \sigma_m^{(i)} u_m u_m^T$, $P^{(i)} = \text{Diag}(\hat{p}_1^{(i)}, \ldots, \hat{p}_K^{(i)})$ and $u_m$ are the canonical vectors, $m = 1, \ldots, M$.

In the following Result 3 proved in Appendix C, we show that the SAMV-1 signal power and noise variance updating formulas given by (14) and (15) can also be obtained by minimizing a weighted least square (WLS) cost function.

**Result 3.** The SAMV-1 estimate is also the minimizer of the following WLS cost function:

$$\hat{p}_k = \arg\min_{p_k} g(p_k),$$

where

$$g(p_k) = \arg\min_{p_k} \|r_N - p_k \hat{a}_k\|^2,\quad (18)$$

and $C_k = C_r - p_k^2 \hat{a}_k \hat{a}_k^H$, $k = 1, \ldots, K + 1$.

The implementation steps of these SAMV approaches are summarized in Table 1.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The SAMV approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initialization:</strong> ${p_k^{(0)}}_{k=1}^K$ and $\hat{\sigma}^{(0)}$ using, e.g., (8) and (9).</td>
<td></td>
</tr>
<tr>
<td><strong>repeat</strong></td>
<td></td>
</tr>
<tr>
<td>• Update $R^{(i)} = AP^{(i)}A^H + \sigma^{(i)}I$</td>
<td></td>
</tr>
<tr>
<td>• Update $p_k^{(i+1)}$ using SAMV formulas (12) or (14) or (16),</td>
<td></td>
</tr>
<tr>
<td>• Update $\hat{\sigma}^{(i+1)}$ using (15).</td>
<td></td>
</tr>
</tbody>
</table>

**Remark 3.** Since $R_N = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(n)\mathbf{y}^H(n)$, the SAMV-1 source power updating formula (14) becomes

$$p_k^{(i+1)} = \frac{1}{N(a_k^H R^{-1}(i) a_k)^2} \sum_{n=1}^N |a_k^H R^{-1}(i) \mathbf{y}(n)|^2.\quad (19)$$

Comparing this expression with its IAA counterpart [6, Table II], we can see that the difference is that the IAA power estimate is obtained by adding up the signal magnitude estimates $\{x_k(n)\}_{k=1}^K$. The matrix $R$ in the IAA approach is obtained as $\mathbf{A}\mathbf{P}\mathbf{A}^H$, where $\mathbf{P} = \text{Diag}(p_1, \ldots, p_K)$. This $R$ can suffer from matrix singularity problem when only a few elements of $\{p_k\}_{k=1}^K$ are non-zero (i.e., the noise-free case).

**Remark 4.** The main difference among the three versions of the SAMV algorithms lies in the formation of the power matrix $\mathbf{P}$. SAMV-0 provides highly sparse solution because its power estimates are proportional to the square of the power estimates at the previous iteration, whereas the SAMV-1 power estimates are proportional to the square of the Capon power estimates $\hat{p}_k^{(i)}\text{Capon}$. When comparing SAMV-1 with SAMV-2, we note that the SAMV-2 power estimates are proportional to $p_k^{(i)}\hat{p}_k\text{Capon}$, a combination of power estimates at previous iteration and the standard Capon power estimates. Due to the factor $p_k^{(i)}\hat{p}_k\text{Capon}$, SAMV-2 achieves better resolution than SAMV-1.

V. **DOA estimation: The Sparse Asymptotic Minimum Variance-Stochastic Maximum Likelihood approaches**

It has been noticed in [5] that the resolution of most power-based sparse source localization techniques is limited by the fineness of the direction grid that covers the location parameter space. In the sparse signal recovery model, the sparsity of the truth is actually dependent on the distance between the adjacent element in the overcomplete dictionary. Therefore, the difficulty of choosing the optimum overcomplete dictionary (i.e., particularly, the DOA scanning direction grid) arises. Since the computational complexity is proportional to the fineness of the direction grid, a highly dense grid is not computationally practical. To overcome this resolution limitation imposed by the grid size, we propose the grid-free SAMV-SML approaches, which refine the location estimates $\theta = (\theta_1, \ldots, \theta_K)^T$ by iteratively minimizing a stochastic ML cost function with respect to a single scalar parameter $\theta_k$.

By using (47), the ML objective function can be decomposed into $\mathcal{L}(p_{-k})$ and $l(p_k)$. Assuming that the parameters $\{p_k\}_{k=1}^K$ and $\sigma$ are estimated using the SAMV approaches, these two terms are simplified, respectively, as $\mathcal{L}(\theta_{-k})$, the marginal likelihood function with parameter $\theta_k$ excluded, and $l(\theta_k)$ with terms concerning $\theta_k$:

$$l(\theta_k) = \log \left( \frac{1}{1 + p_k \alpha_{1,k}(\theta_k)} + \frac{\alpha_{N,k}(\theta_k)}{1 + p_k \alpha_{N,k}(\theta_k)} \right),\quad (20)$$

where the $\alpha_{1,k}(\theta_k)$ and $\alpha_{N,k}(\theta_k)$ are defined in Appendix B. Therefore, the estimate of $\theta_k$ can be obtained by minimizing (20) with respect to the scalar parameter $\theta_k$.

The classical stochastic ML estimates are obtained by minimizing the cost function with respect to a multi-dimensional vector $(\theta_k)_{k=1}^K$, (see e.g., [25, Appendix B, Eq. (B.1)]). The computational complexity of the multi-dimensional optimization is so high that the classical stochastic ML estimation problem is usually unsolvable. On the contrary, the proposed SAMV-SML algorithms only require minimizing (20) with respect to a scalar $\theta_k$, which can be efficiently implemented using derivative-free uphill search methods such as the Nelder-Mead algorithm [29].

The SAMV-SML approaches are summarized in Table 2.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>The SAMV-SML approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initialization:</strong> ${p_k^{(0)}}<em>{k=1}^K$, $\hat{\sigma}^{(0)}$ and ${\theta_k^{(0)}}</em>{k=1}^K$ based on the result of SAMV-0, e.g., SAMV-2 estimates, (16) and (15).</td>
<td></td>
</tr>
<tr>
<td><strong>repeat</strong></td>
<td></td>
</tr>
<tr>
<td>• Compute $R^{(i)}$ and $Q_k^{(i)}$ given by (43).</td>
<td></td>
</tr>
<tr>
<td>• Update $p_k$ using (4) or (14) or (16), update $\sigma$ using (5) or (15),</td>
<td></td>
</tr>
<tr>
<td>• Minimizing (20) with respect to $\theta_k$ to obtain the stochastic ML estimate</td>
<td></td>
</tr>
</tbody>
</table>

VI. **HIGH AND LOW SNR APPROXIMATIONS**

To get more insights into the SAMV approaches, we derive the following approximate expressions for the SAMV approaches for high and low SNR cases, respectively.

4SAMV-SML variants use different $p_k$ and $\sigma$ estimates: AMV-SML: (4) and (5), SAMV1-SML: (14) and (15), SAMV2-SML: (16) and (15).

5The Nelder-Mead algorithm has already been incorporated in the function “fminsearch” in MATLAB®.
1) Zero-Order Low SNR Approximation: Note that the inverse of the matrix $\mathbf{R}$ can be written as:

$$
\mathbf{R}^{-1} = (\mathbf{R} + \sigma \mathbf{I})^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1} \left( \frac{1}{\sigma} \mathbf{I} + \mathbf{R}^{-1} \right)^{-1} \mathbf{R}^{-1},
$$

(21)

At low SNR (i.e., $\frac{\sigma}{\mathbf{R}} \ll 1$), from (21), we obtain $\mathbf{R}^{-1} \approx \frac{1}{\sigma} \mathbf{I}$. Thus,

$$
a^H_k \mathbf{R}^{-1} \mathbf{R} \mathbf{R}^{-1} a_k \approx \frac{1}{\sigma} \sigma_k^H \mathbf{R} \mathbf{R}^{-1} a_k,
$$

(22)

$$
a^H_k \mathbf{R}^{-1} a_k \approx \frac{M^2}{\sigma} \sigma_k^H a_k,
$$

(23)

$$
\text{Tr}(\mathbf{R}^{-2} \mathbf{R}^N) \approx \frac{1}{N \sigma^2} \sum_{n=1}^{N} \|y(n)\|^2,
$$

(24)

$$
\text{Tr}(\mathbf{R}^{-2(i)}) \approx \frac{M}{\sigma^2}.
$$

(25)

Substituting (22) and (23) into the SAMV updating formulas (12)-(16), we obtain

$$
p_{k, \text{SAMV-0}}^{(i+1)} = \frac{M}{\sigma} \sigma_k^H \hat{p}_{k, \text{SAMV-0}} \hat{p}_{k, \text{PER}},
$$

(26)

$$
p_{k, \text{SAMV-1}}^{(i+1)} = \hat{p}_{k, \text{PER}},
$$

(27)

$$
p_{k, \text{SAMV-2}}^{(i+1)} = \frac{M}{\sigma} \sigma_k^H \hat{p}_{k, \text{SAMV-2}} \hat{p}_{k, \text{PER}},
$$

(28)

where $\hat{p}_{k, \text{PER}}$ is given by (8). Using (24) and (25), the common SAMV noise updating equation (15) is approximated as

$$
\hat{\sigma} = \frac{1}{MN} \sum_{n=1}^{N} \|y(n)\|^2.
$$

(29)

From (27), we comment that the SAMV-1 approach is equivalent to the PER method at low SNR. In addition, we remark that at very low SNR, the SAMV-0 and SAMV-2 power estimates given by (26) and (28) are scaled versions of the PER estimate $\hat{p}_{k, \text{PER}}$, provided that they are both initialized by PER.

2) Zero-Order High SNR Approximation: At high SNR (i.e., $\frac{\sigma}{\mathbf{R}} \gg 1$), from (21), we obtain $\mathbf{R}^{-1} \approx \mathbf{R}^{-1}$. Thus,

$$
a^H_k \mathbf{R}^{-1} \mathbf{R} \mathbf{R}^{-1} a_k \approx a^H_k \mathbf{R}^{-1} \mathbf{R} a_k,
$$

(30)

Substituting (29) and (30) into the SAMV formulas (12)-(16) yields:

$$
p_{k, \text{SAMV-0}}^{(i+1)} = \sigma_k^H \sigma_k^{-1} \sigma_k^{-1} a_k,
$$

(31)

$$
p_{k, \text{SAMV-1}}^{(i+1)} = \frac{a^H_k \mathbf{R}^{-1(i)} \mathbf{R} \mathbf{R}^{-1(i)} a_k}{(a^H_k \mathbf{R}^{-1(i)} a_k)^2},
$$

(32)

$$
p_{k, \text{SAMV-2}}^{(i+1)} = \frac{a^H_k \mathbf{R}^{-1(i)} \mathbf{R} \mathbf{R}^{-1(i)} a_k}{a^H_k \mathbf{R}^{-1(i)} a_k}.\n$$

(33)

From (32), we get

$$
p_k^{(i+1)} = \frac{a^H_k \mathbf{R}^{-1(i)} \mathbf{R} \mathbf{R}^{-1(i)} a_k}{(a^H_k \mathbf{R}^{-1(i)} a_k)^2} = \frac{1}{N} \frac{\sum_{n=1}^{N} [y_k^{(i)}(n)]^2}{\sum_{n=1}^{N} \|y_k^{(i)}(n)\|^2},
$$

(34)

where $y_k^{(i)}(n) \triangleq \frac{a^H_k \mathbf{R}^{-1(i)} y(n)}{a^H_k \mathbf{R}^{-1(i)} a_k}$ is the the signal waveform estimate at the direction $\theta_k$ and the nth snapshot [6, Eq. (7)]. From (34), we comment that SAMV-1 and IAA are equivalent at high SNR.

VII. SIMULATION RESULTS

A. Source Localization

This subsection focuses on evaluating the performances of the proposed SAMV and SAMV-SML algorithms using an $M = 12$ element uniform linear array (ULA) with half-wavelength inter-element spacing, since the application of the proposed algorithms to arbitrary arrays is straightforward. For all the considered power-based approaches, the scanning direction grid $\{\theta_k\}_{k=1}^{K}$ is chosen to uniformly cover the entire region-of-interest $\Omega = [0^\circ 180^\circ]$ with the step size of $0.2^\circ$. The various SNR values are achieved by adjusting the noise variance $\sigma$, and the SNR is defined as:

$$
\text{SNR} \triangleq 10 \log_{10} \left( \frac{p_{\text{avg}}}{\sigma} \right) \text{[dB]},
$$

(35)

where $p_{\text{avg}}$ denotes the average power of all sources. For $K$ sources, $p_{\text{avg}} \triangleq \frac{1}{K} \sum_{k=1}^{K} p_k$.

First, the DOA estimation results using a 12 element ULA and $N = 120$ snapshots of both independent and coherent sources are given in Figure 1 and Figure 2, respectively. Three sources with 5 dB, 3 dB and 4 dB power at locations $\theta_1 = 35.11^\circ$, $\theta_2 = 50.15^\circ$ and $\theta_3 = 55.05^\circ$ are present in the region-of-interest. For the coherent source case in Figure 2, the sources at $\theta_1$ and $\theta_3$ share the same phases but the source at $\theta_2$ is independent of them. The true source locations and powers are represented by the circles and vertical dashed lines that align with these circles. In each plot, the estimation results of 10 Monte Carlo trials for each of the algorithms are shown together.

Due to the strong smearing effects and limited resolution, the PER approach fails to correctly separate the closely spaced sources at $\theta_2$ and $\theta_3$ (Figure 1(a) and Figure 2(a)). The IAA algorithm has reduced the smearing effects significantly, resulting in much lower sidelobe levels in Figure 1(b) and Figure 2(b). However, the resolution provided by IAA is still not high enough to separate the two closely spaced sources at $\theta_2$ and $\theta_3$.

In the scenario with independent sources, the eigen-analysis based MUSIC algorithm and existing sparse methods such as the SPICE+ algorithm, are capable of resolving all three sources in Figures 1(c)–(d), thanks to their superior resolution. However, the source coherence degrades their performances dramatically in Figures 2(c)–(d). In contrast, the proposed SAMV algorithms, depicted in Figure 2(e)–(g), are much more robust against signal coherence. We observe in Figures 1–2 that the SAMV-1 approach generally provides almost identical spatial estimates to its IAA counterpart, which verifies the comments in Section VI. In Figures 1–2, the SAMV-0 and SAMV-2 algorithms generate high resolution sparse spatial estimates for both the independent and coherent sources. However, we notice in our simulations that the sparsest SAMV-0 algorithm requires a high SNR to work properly. Therefore, SAMV-0 is not included when comparing mean-square-errors of angle estimation over a wide range of SNR in Figures 3–4. From Figures 1–2, we comment that the SAMV-SML algorithms (AMV-SML, SAMV1-SML and SAMV2-SML) provide the most accurate estimates of the source locations and powers simultaneously.
Fig. 1. Source localization with a ULA of $M = 12$ sensors and $N = 120$ snapshots, SNR = 25 dB. Three uncorrelated sources are at $35.11^\circ$, $50.15^\circ$ and $55.05^\circ$, as represented by the red circles and vertical dashed lines in each plot. 10 Monte Carlo trials are shown in each plot. Spatial estimates are shown with (a) Periodogram (PER), (b) IAA, (c) SPICE+, (d) MUSIC, (e) SAMV-0, (f) SAMV-1, (g) SAMV-2, (h) AMV-SML, (i) SAMV1-SML and (j) SAMV2-SML.

Fig. 2. Source localization with a ULA of $M = 12$ sensors and $N = 120$ snapshots, SNR = 25 dB. Three sources are at $35.11^\circ$, $50.15^\circ$ and $55.05^\circ$. The first and the last sources are coherent with each other. These sources are represented by the red circles and vertical dashed lines in each plot. 10 Monte Carlo trials are shown in each plot. Spatial estimates are shown with (a) Periodogram (PER), (b) IAA, (c) SPICE+, (d) MUSIC, (e) SAMV-0, (f) SAMV-1, (g) SAMV-2, (h) AMV-SML, (i) SAMV1-SML and (j) SAMV2-SML.

Next, Figures 3–4 compare the total angle mean-square-error (MSE)$^6$ of each algorithm with respect to varying SNR values for both independent and coherent sources. These DOA localization results are obtained using a 12 element ULA and $N = 16$ or 120 snapshots. Two sources with 5 dB and 3 dB power at locations $\theta_1 = 35.11^\circ$ and $\theta_2 = 50.15^\circ$ are present$^7$.

While calculating the MSEs for the power-based grid-dependent algorithms$^8$, only the highest two peaks in $\{\hat{\theta}_k\}_{k=1}^K$ are selected as the estimates of the source locations. The grid-independent SAMV-SML algorithms (AMV-SML, SAMV1-SML and SAMV2-SML) are all initialized by the SAMV-2 algorithm. Each point in Figures 3–4 is the average of 1000 Monte Carlo trials.

Due to the severe smearing effects (already shown in Figures 1–2), the PER approach gives high total angle MSEs in Figures 3–4. The IAA algorithm provides almost equivalent MSE performances to the SAMV-1 approach for a wide range of the SNR values (see comments in Section VI), and thus only the MSE curves of the SAMV-1 algorithm are plotted in Figures 3–4. In contrast, the SAMV-2 approach offers lower total angle estimation MSEs, especially for the coherent source case, and this is also the main reason why we initialize the SAMV-SML approaches with the SAMV-2 result. Note that in Figure 3, the SAMV-1 and SAMV-2 provide similar MSEs at very low SNR, which has already been investigated in Section VI. The zero-order low SNR approximation shows that the SAMV-1 and SAMV-2 estimates are equivalent to the PER result or a scaled version of it.

We also observe that there exist the plateau effects for the power-based grid-dependent algorithms (SAMV-1 and SAMV-2) in Figures 3–4 when the SNR is sufficiently high. These phenomena reflect the resolution limitation imposed by the direction grid size as detailed in Section V. Since the power-based grid-dependent algorithms estimate each source location $\hat{\theta}_{source}$ by selecting one element from a fixed set of discrete values (i.e., the direction grid points, $\{\theta_k\}_{k=1}^K$), there always exists an estimation bias provided that the sources are not located precisely on the direction grid. Theoretically, this bias can be reduced if the adjacent distance between the grid points is reduced. However, a uniformly fine direction grid with large $K$ values incurs prohibitive computational costs and is not applicable for practical applications. In lieu of

$^6$Defined as the summation of the angle MSE for each source.

$^7$These DOA true values are selected so that neither of them is on the direction grid.

$^8$Include the IAA, SAMV-1 and SAMV-2 algorithms.
increasing the value of $K$, some adaptive grid refinement postprocessing techniques have been developed (e.g., [5]) by refining this grid locally. To combat the resolution limitation without relying on additional grid refinement postprocessing, the SAMV-SML approaches employ a grid-independent one-dimensional minimization scheme, and the resulted angle estimation MSEs are significantly reduced at high SNR compared to the SAMV approaches in Figures 3–4. In our simulations, the MSE performances of the AMV-SML and SAMV1-SML approaches are identical to their SAMV2-SML counterpart, which verifies that the SAMV signal powers and noise variance updating formulas (Eq. 14–16) are good approximations to the ML estimates (Eq. 4–5). In the independent source scenario in Figure 3, the MSE curves of the SAMV-SML approaches agree well with the stochastic Cramér-Rao lower bound (CRB, see, e.g., [25]) over most of the indicated range of SNR. Even with coherent sources, these SAMV-SML approaches are still asymptotically efficient and they provide lower angle estimation MSEs than competing algorithms over a wide range of SNR.

**B. Active Sensing: Range-Doppler Imaging Examples**

This subsection focuses on numerical examples for the SISO radar/sonar Range-Doppler imaging problem. Since this imaging problem is essentially a single-snapshot application, only algorithms that work with single snapshot are included in this comparison, namely, Matched Filter (MF, another alias of the periodogram approach), IAA, SAMV-0, SAMV-1 and SAMV-2. First, we follow the same simulation conditions as in [6]. A 30-element P3 code is employed as the transmitted pulse, and a total of nine moving targets are simulated. Of all the moving targets, three are of 5 dB power and the rest six are of 25 dB power, as depicted in Figure 5(a). The received signals are assumed to be contaminated with uniform white Gaussian noise of 0 dB power. Figure 5 shows the comparison of the imaging results produced by the aforementioned algorithms.

The Matched Filter (MF) result in Figure 5(b) suffers from severe smearing and leakage effects both in the Doppler and range domain, making it impossible to distinguish the 5 dB targets. In contrast, the IAA algorithm in Figure 5(c) and SAMV-1 in Figure 5(e) offer similar and greatly enhanced imaging results with observable target range estimates and Doppler frequencies. The SAMV-0 approach provides highly sparse result and eliminates the smearing effects completely, but it misses the weak 5 dB targets in Figure 5(d), which agrees well with our previous comment on its sensitivity to SNR. In Figure 5(f), the smearing effects (especially in the Doppler domain) are further attenuated by SAMV-2, compared with the IAA/SAMV-1 results. We comment that among all the competing algorithms, the SAMV-2 approach provides the best balanced result, providing sufficiently sparse images without
missing weak targets.

In Figure 5(d), the three 5 dB sources are not resolved by the SAMV-0 approach due to the excessive low SNR. After increasing the power levels of these sources to 15 dB (the rest conditions are kept the same as in Figure 5), all the sources can be accurately resolved by the SAMV-0 approach in Figure 6(d). We comment that the SAMV-0 approach provides the most accurate imaging result provided that all sources have adequately high SNR.

VIII. CONCLUSIONS

We have presented a series of user parameter-free array processing algorithms, the iterative SAMV algorithms, based on the AMV criterion. It has been shown that these algorithms have superior resolution and sidelobe suppression ability and are robust to practical difficulties such as insufficient snapshots, coherent source signals, without the need of any decoration preprocessing. Moreover, a series of grid-independent SAMV-SML approaches are proposed to combat the limitation of the direction grid size problem. It is shown that these approaches provide grid-independent asymptotically efficient estimates without any additional grid refinement postprocessing.

APPENDIX A

PROOF OF RESULT 1

Given \( \{ \hat{p}_k \}_{i=1}^K \) and \( \hat{\sigma} \), which are the estimates of the first \( K \) components and the last element of \( p \) at the \( i \)th iteration, the matrix \( R(i) = A \hat{p}_k ^{(i)} A^H + \sigma(i) I \) is known, and thus the matrix \( C_r(i) = R(i) \otimes R(i) \) is also known. For notational simplicity, we omit the iteration index and use \( C_r \) instead in this section.

Define the vectorized covariance matrix of the interference and noise as

\[
r'_k \overset{\text{def}}{=} r - p_k a_k, \quad k = 1, \ldots, K.
\]

Assume that \( r'_k \) is known and substitute \( r'_k + p_k a_k \) for \( r \) in (3). Then minimizing (3) is equivalent to minimizing the following cost function:

\[
f(p_k) = [r_N - p_k a_k]^H C_r^{-1} [r_N - p_k a_k] - [r_N - p_k a_k]^H C_r^{-1} r'_k - r'_k^H C_r^{-1} [r_N - p_k a_k]. \tag{36}\]

Note that \( r'_k \) does not depend on \( p_k \). Differentiating (3) with respect to \( p_k \) and setting the results to zero, we get

\[
\hat{p}_k = \frac{1}{a_k^H C_r^{-1} a_k} (a_k^H C_r^{-1} r_N - a_k^H C_r^{-1} r'_k), \quad k = 1, \ldots, K+1. \tag{37}\]

Replacing \( r'_k \) with its definition in (37) yields

\[
\hat{p}_k = \frac{1}{a_k^H C_r^{-1} a_k} (a_k^H C_r^{-1} r_N + p_k a_k^H C_r^{-1} a_k - a_k^H C_r^{-1} r). \tag{38}\]

Using the following identities (see, e.g., [2, Th. 7.7, 7.16]),

\[
\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B), \tag{39}\]

\[
 (A \otimes B) \otimes (C \otimes D) = AC \otimes BD, \tag{40}\]

Eq. (38) can be simplified as

\[
\hat{p}_k = \frac{a_k^H R^{-1} R_k^{-1} a_k + p_k - \frac{1}{a_k^H R^{-1} a_k}}{1 + \text{Tr}(R^{-2})} ( \text{Tr}(R^{-2} R_N) + \sigma \text{Tr}(R^{-2}) - \text{Tr}(R^{-1}) ) \tag{41}\]

Computing \( \hat{p}_k \) and \( \hat{\sigma} \) requires the knowledge of \( p_k, \sigma, \) and \( R \). Therefore, this algorithm must be implemented iteratively as detailed in Table 1.

APPENDIX B

PROOF OF RESULT 2

Define the covariance matrix of the interference and noise as

\[
Q_k \overset{\text{def}}{=} R - p_k a_k a_k^H, \quad k = 1, \ldots, K. \tag{43}\]

Applying the matrix inversion lemma to (43) yields

\[
R^{-1} = Q_k^{-1} - p_k \beta_k b_k b_k^H, \quad k = 1, \ldots, K, \tag{44}\]

where \( b_k \overset{\text{def}}{=} Q_k^{-1} a_k \) and \( \beta_k \overset{\text{def}}{=} (1 + p_k a_k^H Q_k^{-1} a_k)^{-1} \). Since

\[
\text{Tr}(R^{-1} R_N) = \text{Tr}(Q_k^{-1} R_N) - p_k \beta_k b_k^H R_N b_k, \tag{45}\]

and using the algebraic identity \( \det(I + AB) = \det(I + BA) \), we obtain

\[
\ln(\det(R)) = \ln(\det(Q_k + p_k a_k a_k^H)) = \ln \left(1 + p_k a_k^H Q_k^{-1} a_k\right) \det(C_r) = \ln(\det(Q_k)) - \ln(\beta_k). \tag{42}\]
Substituting (45) and (46) into the ML function (6) yields
\[
\mathcal{L}(p) = \ln(\det(Q_k)) + \text{Tr}(Q_k^{-1}R_N - (\ln(\beta_k) + p_k\beta_k (b_k^H R_N b_k)) = \mathcal{L}(p_{k-1}) - l(p_k),
\]
with
\[
l(p_k) \equiv \ln\left(1 + \frac{1}{1 + p_k\alpha_{k,1}(\theta_k)}\right) + p_k \frac{\alpha_{k,1}(\theta_k)}{1 + p_k \alpha_{k,1}(\theta_k)}, \tag{48}
\]
where \(\alpha_{k,1}(\theta_k) \equiv (a_k^H Q_k^{-1} a_k)^{-1}\) and \(\alpha_{k,2}(\theta_k) \equiv (a_k^H Q_k^{-1} R_N Q_k^{-1} a_k)^{-1}\). The objective function has now been decomposed into \(\mathcal{L}(p_{k-1})\), the marginal likelihood with \(p_k\) excluded, and \(l(p_k)\), where terms concerning \(p_k\) are conveniently isolated. Consequently, minimizing (6) with respect to \(p_k\) is equivalent to minimizing the function (48) with respect to the parameter \(p_k\).

It has been proved in [6, Appendix, Eqs. (27) and (28)] that the unique minimizer of the cost function (48) is
\[
\hat{p}_k = \frac{a_k^H Q_k^{-1} (R_N - Q_k) Q_k^{-1} a_k}{(a_k^H Q_k^{-1} a_k)^2}, \quad k = 1, \ldots, K. \tag{49}
\]
We note that \(\hat{p}\) is strictly positive if \(a_k^H Q_k^{-1} R_N Q_k^{-1} a_k > a_k^H Q_k^{-1} a_k\). Using (44), we have
\[
a_k^H Q_k^{-1} a_k = \gamma_k (a_k^H R_k^{-1} a_k), \tag{50}
a_k^H Q_k^{-1} R_N Q_k^{-1} a_k = \gamma_k^2 (a_k^H R_k^{-1} R_N R_k^{-1} a_k), \tag{51}
\]
where \(\gamma_k \equiv 1 + p_k a_k^H Q_k^{-1} a_k\). Substituting (50) and (51) into (49), we obtain the desired expression
\[
\hat{p}_k = \frac{a_k^H R_k^{-1} (R_N - R_k) R_k^{-1} a_k}{(a_k^H R_k^{-1} a_k)^2} \cdot p_k = \frac{a_k^H R_k^{-1} R_N R_k^{-1} a_k}{(a_k^H R_k^{-1} a_k)^2} + p_k \tag{52}
\]
Differentiating (6) with respect to \(\sigma\) and setting the result to zero, we obtain
\[
\hat{\sigma} = \frac{\text{Tr}(R_k^{-1}(R_N - \bar{R}) R_k^{-1})}{\text{Tr}(R_k^{-2})}, \tag{53}
\]
and after substituting \(\bar{R} - \sigma I\) for \(\bar{R}\) in the above equation,
\[
\hat{\sigma} = \frac{\text{Tr}(R_k^{-1}(R_N - R_k) R_k^{-1})}{\text{Tr}(R_k^{-2})} + \sigma = \text{Tr}(R_k^{-2} R_N)/\text{Tr}(R_k^{-2}) + \sigma, \tag{54}
\]
Computing \(\hat{p}_k\) and \(\hat{\sigma}\) requires the knowledge of \(p_k, \sigma,\) and \(\bar{R}\). Therefore, the algorithm must be implemented iteratively as detailed in Result 1.

**APPENDIX C**

**PROOF OF RESULT 3**

Differentiating (18) with respect to \(p_k\) and setting the result to zero, we get
\[
\hat{p}_{k+1} = \frac{a_k^H C_k^{-1} r_N}{a_k^H C_k^{-1} a_k}. \tag{55}
\]
Applying the matrix inversion lemma to \(C_k^{-1}\), the numerator and denominator of Eq. (55) can be expressed respectively, as
\[
a_k^H C_k^{-1} r_N = w_k(a_k^H C_k^{-1} r_N), \tag{47}
\]
and
\[
\hat{a}_k^H C_k^{-1} \hat{a}_k = w_k(a_k^H C_k^{-1} \hat{a}_k), \tag{48}
\]
Thus,
\[
\hat{p}_{k+1} = \frac{\hat{a}_k^H C_k^{-1} r_N}{\hat{a}_k^H C_k^{-1} \hat{a}_k}, \quad k = 1, \ldots, K. \tag{56}
\]
Using the Kronecker product properties and the identities (39) and (40), with \(A = B = R\) and \(C = R_N\), the numerator and denominator of Eq. (56) can be expressed, respectively, as
\[
a_k^H C_k^{-1} r_N = a_k^H R^{-1} R_N R^{-1} a_k, \quad k = 1, \ldots, K, \tag{57}
a_k^H C_k^{-1} \hat{a}_k = (a_k^H R^{-1} a_k)^2, \quad k = 1, \ldots, K, \tag{58}
\]
and
\[
\hat{a}_k^H C_k^{-1} \hat{a}_k = a_k^H C_k^{-1} \hat{a}_k, \quad k = 1, \ldots, K, \tag{59}
\]
Therefore, dividing (57) by (58) gives (14), and dividing (59) by (60) yields (15).

**REFERENCES**


