Persymmetric Parametric Adaptive Matched Filter for Multichannel Adaptive Signal Detection

Pu Wang, Zafer Sahinoglu, Man-On Pun, and Hongbin Li

***Abstract*—This correspondence considers a *parametric* approach for multichannel adaptive signal detection in Gaussian disturbance which can be modeled as a multichannel autoregressive (AR) process and, moreover, possesses a persymmetric structure induced by a symmetric antenna geometry. By introducing the persymmetric AR (PAR) modeling for the disturbance, a persymmetric parametric adaptive matched filter (Per-PAMF) is proposed. The developed Per-PAMF extends the classical PAMF by exploiting the underlying persymmetric properties and, hence, improves the detection performance in training-limited scenarios. The performance of the proposed Per-PAMF is examined by the Monte Carlo simulations and simulation results demonstrate the effectiveness of the Per-PAMF compared with the conventional PAMF and nonparametric detectors.**

***Index Terms*—Multichannel adaptive signal detection, maximum likelihood estimation, multichannel autoregressive process, parametric approach, persymmetry.**

I. Introduction

Multichannel adaptive signal detection against strong spatially and temporally colored disturbances has been encountered in many applications, e.g., wireless communications, hyperspectral imaging, and medical imaging [1]–[3]. Traditional techniques are limited for practical applications due to their excessive training requirement and high computational complexity. For example, the covariance-matrix-based detectors, e.g., Kelly's generalized likelihood ratio test (GLRT) [4], the adaptive matched filter (AMF) [5] and the recent Rao test [6], need $K \geq JN$ training signals to ensure a full-rank estimate of the disturbance covariance matrix and have to invert the $JN \times JN$ covariance matrix, where J denotes the number of antennas and N denotes the number of pulses.

Among other techniques, a class of parametric detectors provide an efficient way to simultaneously mitigate the training requirement and reduce the computational complexity [7]–[11] (and reference therein). By modeling the disturbance as a multichannel autoregressive (AR) process, the parametric detectors decompose the jointly spatio-temporal whitening of the covariance-matrix-based detectors into successive temporal whitening and spatial whitening. As a well-known parametric detector, the parametric AMF (PAMF) is simple to implement. Using measured datasets [7], [8], the PAMF was found to yield better performance with significantly reduced computational complexity than the nonparametric counterpart, i.e., the AMF, especially when $K \ll JN$.

In this correspondence, extending the multichannel AR process based parametric detector, we exploit additional structure of the disturbance covariance matrix, i.e., the persymmetric property [12], [13]. This enables an improved parametric detector with better training-signal efficiency. The utilization of persymmetry for applications in communications and radar can be traced back to [12] and [13] and has been proved to be an efficient way to mitigate the demanding requirement of homogeneous training signals. In [13], Nitzberg shows that the efficiency of usage of training signals is improved by up to a factor of two by utilizing persymmetry. Following [13], several adaptive detection schemes explicitly taking into account the persymmetry have been proposed in [14] and, more recently, [15]–[19]. Specifically, [14] proposed a GLRT to detect a multi-band signal from the disturbance with the persymmetric property. Extension to the compound-Gaussian environment was made in [15] and has been further verified in [16] with experimentally measured datasets. The results show that exploiting the persymmetric property significantly improves the robustness of the adaptive detection algorithms in terms of the constant false alarm rate (CFAR). Meanwhile, [17] proposed a persymmetric GLRT for a partially-homogeneous environment.

Our incorporation of the persymmetric constraint in the PAMF leads to a new persymmetric PAMF (Per-PAMF), which, while maintaining the simple implementation as the PAMF, further improves the robustness in training-limited scenarios. The Per-PAMF is developed in a two-step procedure. In detail, a nonadaptive parametric matched filter (PMF) is first introduced by assuming the knowledge of the nuisance parameters and, then, the Per-PAMF is developed from the PMF by replacing the nuisance parameters by their maximum likelihood (ML) estimates from training signals under the persymmetric constraint. The performance of the Per-PAMF detector is examined by the Monte-Carlo simulations and the simulation results show that the Per-PAMF has slightly better performance than the PAMF when the number of training signals is sufficient, while it significantly outperforms the PAMF in cases with extremely limited training signals.

The remainder of the correspondence is organized as follows. Section II contains the signal model and introduces the persymmetric AR modeling for the disturbance. The Per-PAMF detector is derived in Section III. Numerical results with two distinct datasets are provided in Section IV. The conclusion is finally drawn in Section V.

II. Signal Model

The problem of interest is to decide which of the following two hypotheses is true [4], [5], [7], [9], [10]:

$$\begin{aligned} H_0 &: \mathbf{x}_0 = \mathbf{d}_0 \\ H_1 &: \mathbf{x}_0 = \alpha \mathbf{s} + \mathbf{d}_0 \end{aligned} \tag{1}$$

where $\mathbf{x}_0$ is the $JN \times 1$ test signal, $\mathbf{s}$ is the *known* space-time steering vector which is a Kronecker product between the temporal $(\mathbf{s}_d)$ and spatial $(\mathbf{s}_s)$ steering vectors, i.e., $\mathbf{s} = \mathbf{s}_d \otimes \mathbf{s}_s$, α is an *unknown* complex-valued amplitude, and $\mathbf{d}_0$ is the disturbance signal (e.g., clutter and noise) which is modeled as a complex Gaussian vector with zero-mean and unknown covariance matrix $\mathbf{R}$, i.e., $\mathbf{d}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$. Aside from the test signal, there are K target-free independent and identically distributed (i.i.d.) training signals $\mathbf{x}_k = \mathbf{d}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$, $k = 1, \ldots, K$, which are also independent of the test signal.

Moreover, by following the parametric approach introduced in [7], the disturbance signals $\mathbf{d}_k$ are modeled as a multichannel AR process. Specifically, let $\mathbf{d}_k(n) \in \mathbb{C}^{J \times 1}, n = 0, 1, \ldots, N-1$, denote the N nonoverlapping temporal segments of $\mathbf{d}_k$, i.e., $\mathbf{d}_k \triangleq [\mathbf{d}_k^T(0), \mathbf{d}_k^T(1), \ldots, \mathbf{d}_k^T(N-1)]^T$. The multichannel AR process of the disturbance signal is described as [7], [9], [10]

$$\mathbf{d}_k(n) = -\sum_{p=1}^{P} \mathbf{A}^H(p) \mathbf{d}_k(n-p) + \varepsilon_k(n) \tag{2}$$

Manuscript received October 06, 2011; revised February 02, 2012; accepted February 10, 2012. Date of publication March 08, 2012; date of current version May 11, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Milica Stojanovic.

P. Wang and H. Li are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: pwang4@stevens.edu; Hongbin.Li@stevens.edu).

Z. Sahinoglu and M.-O. Pun are with Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139 USA (e-mail: zafer@merl.com; mpun@merl.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2012.2190411

where $\varepsilon_k(n) \sim \mathcal{CN}(0, \mathbf{Q})$ is the J-channel temporally white but spatially colored Gaussian driving noise with $\mathbf{Q}$ denoting the *unknown* $J \times J$ spatial covariance matrix, and $\{\mathbf{A}(p)\}_{p=1}^{P}$ denote the *unknown* $J \times J$ AR coefficient matrices.

In this correspondence, we consider a case frequently encountered in practice, where the systems use a symmetric antenna configuration (symmetrical with respect to its phase center) and transmit a set of pulses of equal duration [13]–[17], [19]. For example, the widely used uniform linear array with a constant pulse repetition frequency (PRF) is such a system. The structured antenna array configurations and constant PRF cause the spatio-temporal covariance matrix $\mathbf{R}$ to be *persymmetric-block-Toeplitz*, as shown in [20]. Specifically, the return from a discrete disturbance source has a similar form as a target echo. Unlike a target, the disturbance is distributed in both range and azimuth. As an approximation to a continuous field, the disturbance return from a specific range is modeled as the superposition of a large number N_c of independent disturbance sources in azimuth. Assuming that returns from different disturbance sources are uncorrelated, the disturbance covariance matrix can be calculated as [20, Sec. 2.6.1]

$$\mathbf{R} = \sum_{i=1}^{N_c} \epsilon_i \left(\mathbf{s}_{d,i}\mathbf{s}_{d,i}^H\right) \otimes \left(\mathbf{s}_{s,i}\mathbf{s}_{s,i}^H\right) \tag{3}$$

where ϵ_i is proportional to the radar cross section (RCS) for the ith disturbance source, and $\mathbf{s}_{d,i}$ and $\mathbf{s}_{s,i}$ denote the temporal (Doppler) and spatial steering vectors, respectively. Due to the configurations in both spatial and temporal domains, the temporal and spatial steering vectors of the disturbance source satisfy the following properties, i.e., $\mathbf{s}_{d,i}$ is a Vandermonde vector and $\mathbf{s}_{s,i} = \mathbf{E}\mathbf{s}_{s,i}^*$ is persymmetric, where $\mathbf{E}$ denotes the exchange matrix with unit anti-diagonal elements and zeros elsewhere,

$$\mathbf{E} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{4}$$

Therefore, $\mathbf{s}_{d,i}\mathbf{s}_{d,i}^H$ is a Toeplitz matrix, while $\mathbf{s}_{s,i}\mathbf{s}_{s,i}^H$ is a persymmetric matrix. In addition to the Kronecker operation in (3), it is straightforward to show that the overall covariance matrix $\mathbf{R}$ to be a persymmetric-block-Toeplitz matrix.

To exploit this structure information of the covariance matrix $\mathbf{R}$, a new multichannel AR model by incorporating the persymmetric property is introduced below.

- **AS1—Persymmetric Spatial Covariance Matrix**: Following the traditional AR process of (2), the spatial covariance matrix is further assumed to be persymmetric as

$$\mathbf{Q} = \mathbf{E}\mathbf{Q}^*\mathbf{E} \tag{5}$$

  where $[\,\cdot\,]^*$ denotes the complex conjugate and $\mathbf{E}$ denotes the exchange matrix of (4).
- **AS2—Persymmetric AR Coefficient Matrices**: In addition to the persymmetric $\mathbf{Q}$, we assume that the AR coefficient matrices satisfy the following property:

$$\mathbf{A}(p) = \mathbf{E}\mathbf{A}^*(p)\mathbf{E}. \tag{6}$$

As shown in the Appendix, we prove that the proposed persymmetric AR model provides a parametric approach to approximate the persymmetric-block-Toeplitz structure of the covariance matrix $\mathbf{R}$. In summary, the goal is to develop a decision rule for the problem in (1) together with the assumptions AS1 and AS2.

III. Persymmetric Parametric Adaptive Matched Filter

The Per-PAMF is developed in a two-step approach: 1) find the GLRT when the nuisance parameters $\mathbf{A}$ and $\mathbf{Q}$ are assumed both known and 2) replace $\mathbf{A}$ and $\mathbf{Q}$ by their ML estimates from training signals subject to the persymmetric constraints.

A. PMF—The GLRT With Known $\mathbf{A}$ and $\mathbf{Q}$

When $\mathbf{A}$ and $\mathbf{Q}$ are both known, the GLRT has the form as

$$T = \frac{\max\limits_{\alpha} p_1(\mathbf{x}_0; \alpha, \mathbf{A}, \mathbf{Q})}{p_0(\mathbf{x}_0; \mathbf{A}, \mathbf{Q})} \tag{7}$$

where $\{p_i\}_{i=0,1}$ are the likelihood functions under H_0 and H_1. It can easily be shown that the GLRT reduces to the nonadaptive PMF [7]

$$T_{\mathrm{PMF}} = \frac{\left|\sum\limits_{n=P}^{N-1} \tilde{\mathbf{s}}^H(n)\mathbf{Q}^{-1}\tilde{\mathbf{x}}_0(n)\right|^2}{\sum\limits_{n=P}^{N-1} \tilde{\mathbf{s}}^H(n)\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n)} \tag{8}$$

where $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{x}}_0$ are, respectively, the temporally whitened steering vector and test signal obtained with the *true* temporal correlation matrices $\mathbf{A}(p), p = 1, \ldots, P$,

$$\tilde{\mathbf{s}}(n) = \mathbf{s}(n) + \sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{s}(n-p) \tag{9}$$

$$\tilde{\mathbf{x}}_0(n) = \mathbf{x}_0(n) + \sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{x}_0(n-p). \tag{10}$$

In the case of unknown $\mathbf{A}$ and $\mathbf{Q}$, the above PMF cannot be implemented and, therefore, we need to replace $\mathbf{A}$ and $\mathbf{Q}$ with their ML estimates under the persymmetric constraints of (5) and (6).

B. Persymmetric ML Estimate of $\mathbf{Q}$

The persymmetric ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ are obtained from training signals only. According to the signal model, the joint likelihood function of training signals can be written as

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_K; \mathbf{A}, \mathbf{Q}) = \left[\frac{1}{\pi^J|\mathbf{Q}|} e^{-\mathrm{tr}(\mathbf{Q}^{-1}\boldsymbol{\Gamma}_0)}\right]^{K(N-P)}$$

where

$$K(N-P)\boldsymbol{\Gamma}_0 = \sum_{k=1}^{K}\sum_{n=P}^{N-1} \varepsilon_k(n)\varepsilon_k^H(n) \tag{11}$$

with definitions

$$\varepsilon_k(n) = \mathbf{x}_k(n) + \sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{x}_k(n-p). \tag{12}$$

Alternatively, $K(N-P)\boldsymbol{\Gamma}_0$ can be rewritten as

$$K(N-P)\boldsymbol{\Gamma}_0 = \hat{\mathbf{R}}_{xx} + \mathbf{A}^H\hat{\mathbf{R}}_{yx} + \hat{\mathbf{R}}_{yx}^H\mathbf{A} + \mathbf{A}^H\hat{\mathbf{R}}_{yy}\mathbf{A} \tag{13}$$

where

$$\mathbf{A} \triangleq \left[\mathbf{A}^H(1), \mathbf{A}^H(2), \ldots, \mathbf{A}^H(P)\right]^H \tag{14}$$

$$\hat{\mathbf{R}}_{xx} = \sum_{k=1}^{K}\sum_{n=P}^{N-1} \mathbf{x}_k(n)\mathbf{x}_k^H(n) \tag{15}$$

$$\hat{\mathbf{R}}_{yy} = \sum_{k=1}^{K}\sum_{n=P}^{N-1} \mathbf{y}_k(n)\mathbf{y}_k^H(n) \tag{16}$$

$$\hat{\mathbf{R}}_{yx} = \sum_{k=1}^{K}\sum_{n=P}^{N-1} \mathbf{y}_k(n)\mathbf{x}_k^H(n) \tag{17}$$

and $\mathbf{y}_k(n)$ is a regression vector of $\mathbf{x}_k(n)$:

$$\mathbf{y}_k(n) \triangleq [\mathbf{x}_k^T(n-1), \mathbf{x}_k^T(n-2), \ldots, \mathbf{x}_k^T(n-P)]^T$$

. By exploiting the persymmetric property of $\mathbf{Q}$, i.e., (5), we have

$$\mathrm{tr}(\mathbf{Q}^{-1}\mathbf{\Gamma}_0) = \mathrm{tr}\left(\mathbf{Q}^{-1}\frac{\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E}}{2}\right) \tag{18}$$

which leads to

$$\ln p \propto -\ln|\mathbf{Q}| - \frac{1}{2}\mathrm{tr}(\mathbf{Q}^{-1}[\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E}]).$$

Taking the derivative of $\ln p$ with respect to $\mathbf{Q}$ and equating the results to zero lead to the persymmetric ML estimate of $\mathbf{Q}$

$$\hat{\mathbf{Q}}_{\mathrm{PML}} = \frac{1}{2}(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E}). \tag{19}$$

As a result, $\ln p \propto -\ln\left|\frac{1}{2}(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E})\right|$. Therefore, the persymmetric ML estimate of $\mathbf{A}$ is equivalent to minimizing the determinant of $K(N-P)(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E})/2$.

C. Persymmetric ML Estimate of $\mathbf{A}$

Recall that $\mathbf{A}$ of (14) is formed by stacking $\mathbf{A}(p)$ column-wise. From (6), $\mathbf{A}$ has the following property:

$$\mathbf{A} = \mathbf{E}_B\mathbf{A}^*\mathbf{E} \tag{20}$$

where $\mathbf{E}_B = \mathbf{I}_P \otimes \mathbf{E}$ with $\mathbf{I}_P$ denoting a $P \times P$ identity matrix.

From (13), $K(N-P)(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E})$ can be expressed as

$$\begin{aligned}
&K(N-P)\left(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E}\right)\\
&\quad\overset{(a)}{=} \left[\hat{\mathbf{R}}_{xx} + \mathbf{E}\hat{\mathbf{R}}_{xx}^*\mathbf{E}\right]\\
&\qquad + \left[\mathbf{A}^H\hat{\mathbf{R}}_{yx} + \mathbf{E}(\mathbf{A}^H\mathbf{E}_B\mathbf{E}_B\hat{\mathbf{R}}_{yx})^*\mathbf{E}\right]\\
&\qquad + \left[\hat{\mathbf{R}}_{yx}^H\mathbf{A} + \mathbf{E}\left(\hat{\mathbf{R}}_{yx}^H\mathbf{E}_B\mathbf{E}_B\mathbf{A}\right)^*\mathbf{E}\right]\\
&\qquad + \left[\mathbf{A}^H\hat{\mathbf{R}}_{yy}\mathbf{A} + \mathbf{E}(\mathbf{A}^H\mathbf{E}_B\mathbf{E}_B\hat{\mathbf{R}}_{yy}\mathbf{E}_B\mathbf{E}_B\mathbf{A})^*\mathbf{E}\right]\\
&\quad\overset{(b)}{=} \left[\hat{\mathbf{R}}_{xx} + \mathbf{E}\hat{\mathbf{R}}_{xx}^*\mathbf{E}\right] + \mathbf{A}^H\left[\hat{\mathbf{R}}_{yx} + \mathbf{E}_B\hat{\mathbf{R}}_{yx}^*\mathbf{E}\right]\\
&\qquad + \left[\hat{\mathbf{R}}_{yx}^H + \mathbf{E}\left(\hat{\mathbf{R}}_{yx}^H\right)^*\mathbf{E}_B\right]\mathbf{A}\\
&\qquad + \mathbf{A}^H\left[\hat{\mathbf{R}}_{yy} + \mathbf{E}_B\hat{\mathbf{R}}_{yy}^*\mathbf{E}_B\right]\mathbf{A}
\end{aligned}$$

where (a) has used the fact that $\mathbf{E}_B\mathbf{E}_B = \mathbf{I}_{JP}$, and (b) is due to (20) and $\mathbf{E}_B^H = \mathbf{E}_B$. Denote the persymmetric matrices:

$$\hat{\mathbf{R}}_{xx,P} = \left(\hat{\mathbf{R}}_{xx} + \mathbf{E}\hat{\mathbf{R}}_{xx}^*\mathbf{E}\right)/2 \tag{21}$$

$$\hat{\mathbf{R}}_{yx,P} = \left(\hat{\mathbf{R}}_{yx} + \mathbf{E}_B\hat{\mathbf{R}}_{yx}^*\mathbf{E}\right)/2 \tag{22}$$

$$\hat{\mathbf{R}}_{yy,P} = \left(\hat{\mathbf{R}}_{yy} + \mathbf{E}_B\hat{\mathbf{R}}_{yy}^*\mathbf{E}_B\right)/2. \tag{23}$$

As a result, $K(N-P)(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E})/2$ can be rewritten as

$$\begin{aligned}
&K(N-P)\left(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E}\right)/2\\
&\quad= \hat{\mathbf{R}}_{xx,P} + \mathbf{A}^H\hat{\mathbf{R}}_{yx,P} + \hat{\mathbf{R}}_{yx,P}^H\mathbf{A} + \mathbf{A}^H\hat{\mathbf{R}}_{yy,P}\mathbf{A}\\
&\quad= \left(\mathbf{A}^H + \hat{\mathbf{R}}_{yx,P}^H\hat{\mathbf{R}}_{yy,P}^{-1}\right)\hat{\mathbf{R}}_{yy,P}\left(\mathbf{A}^H + \hat{\mathbf{R}}_{yx,P}^H\hat{\mathbf{R}}_{yy,P}^{-1}\right)^H\\
&\qquad + \left(\hat{\mathbf{R}}_{xx,P} - \hat{\mathbf{R}}_{yx,P}^H\hat{\mathbf{R}}_{yy,P}^{-1}\hat{\mathbf{R}}_{yx,P}\right).
\end{aligned} \tag{24}$$

Since $\hat{\mathbf{R}}_{yy,P}$ is nonnegative definite and the remaining term $\hat{\mathbf{R}}_{xx,P} - \hat{\mathbf{R}}_{yx,P}^H\hat{\mathbf{R}}_{yy,P}^{-1}\hat{\mathbf{R}}_{yx,P}$ does not depend on $\mathbf{A}$, it follows that

$$\begin{aligned}
\left(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E}\right)/2 &\geq \left(\mathbf{\Gamma}_0 + \mathbf{E}\mathbf{\Gamma}_0^*\mathbf{E}\right)/2\big|_{\mathbf{A}=-\hat{\mathbf{R}}_{yx,P}^H\hat{\mathbf{R}}_{yy,P}^{-1}}\\
&= \frac{\hat{\mathbf{R}}_{xx,P} - \hat{\mathbf{R}}_{yx,P}^H\hat{\mathbf{R}}_{yy,P}^{-1}\hat{\mathbf{R}}_{yx,P}}{K(N-P)}
\end{aligned} \tag{25}$$

which implies that

$$\hat{\mathbf{A}}_{\mathrm{PML}} = -\hat{\mathbf{R}}_{yy,P}^{-1}\hat{\mathbf{R}}_{yx,P} \tag{26}$$

and the persymmetric ML estimate of $\mathbf{Q}$ of (19) reduces to

$$\hat{\mathbf{Q}}_{\mathrm{PML}} = \frac{\hat{\mathbf{R}}_{xx,P} - \hat{\mathbf{R}}_{yx,P}^H\hat{\mathbf{R}}_{yy,P}^{-1}\hat{\mathbf{R}}_{yx,P}}{K(N-P)}. \tag{27}$$

D. Per-PAMF

By replacing $\mathbf{A}$ and $\mathbf{Q}$ with the persymmetric ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ in (8), we obtained the Per-PAMF as

$$T_{\mathrm{Per-PAMF}} = \frac{\left|\sum_{n=P}^{N-1}\hat{\tilde{\mathbf{s}}}_P^H(n)\hat{\mathbf{Q}}_{\mathrm{PML}}^{-1}\hat{\tilde{\mathbf{x}}}_{0,P}(n)\right|^2}{\sum_{n=P}^{N-1}\hat{\tilde{\mathbf{s}}}_P^H(n)\hat{\mathbf{Q}}_{\mathrm{PML}}^{-1}\hat{\tilde{\mathbf{s}}}_P(n)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\mathrm{Per-PAMF}} \tag{28}$$

where $\gamma_{\mathrm{Per-PAMF}}$ is a threshold subject to a preset probability of false alarm, and the temporally whitened steering vector $\hat{\tilde{\mathbf{s}}}_P$ and test signal $\hat{\tilde{\mathbf{x}}}_{0,P}$ are *adaptively* obtained, respectively, with the persymmetric ML estimate $\hat{\mathbf{A}}_{\mathrm{PML}}$ of (26):

$$\hat{\tilde{\mathbf{s}}}_P(n) = \mathbf{s}(n) + \sum_{p=1}^{P}\hat{\mathbf{A}}_{\mathrm{PML}}^H(p)\mathbf{s}(n-p) \tag{29}$$

$$\hat{\tilde{\mathbf{x}}}_{0,P}(n) = \mathbf{x}_0(n) + \sum_{p=1}^{P}\hat{\mathbf{A}}_{\mathrm{PML}}^H(p)\mathbf{x}_0(n-p). \tag{30}$$

From (28), it is seen that the Per-PAMF inherits the reduced computational complexity of the conventional PAMF by performing successively a temporal whitening followed by a spatial whitening, in contrast to the computationally intensive joint spatio-temporal whitening of the covariance matrix based approach (e.g., the AMF [5]). On the other hand, it further improves robustness of the parameter estimation by exploiting the underlying structure of the disturbance covariance matrix via the persymmetric ML estimates of the AR coefficient matrices $\mathbf{A}$ and the spatial covariance matrix $\mathbf{Q}$.

E. Comparison With the Conventional PAMF

Compared with the conventional PAMF [7], the proposed Per-PAMF shares the same test variable as shown in (28), except the underlying estimates of the unknown parameters $\mathbf{A}$ and $\mathbf{Q}$. The conventional PAMF uses the unconstrained ML estimate from training signals:

$$\hat{\mathbf{A}}_{\mathrm{ML}} = -\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{yx}, \tag{31}$$

$$\hat{\mathbf{Q}}_{\mathrm{ML}} = \frac{\hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{yx}^H\hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{yx}}{K(N-P)}. \tag{32}$$

Compared with the above ML estimates, it is noted that the persymmetric ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ in (26) and (27) explicitly utilize the persymmetric properties through the operations of (21), (22), and (23).

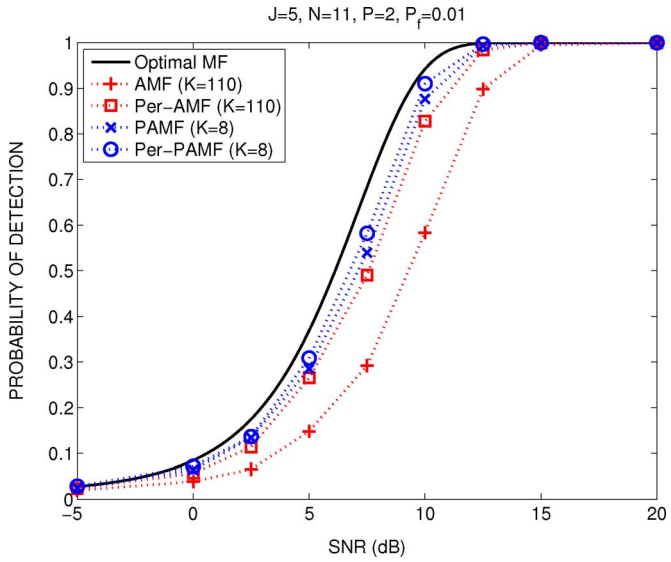

Fig. 1. Probability of detection versus SINR for $K = 8$ when $J = 5, N = 11$, $P = 2$, and $P_f = 0.01$.

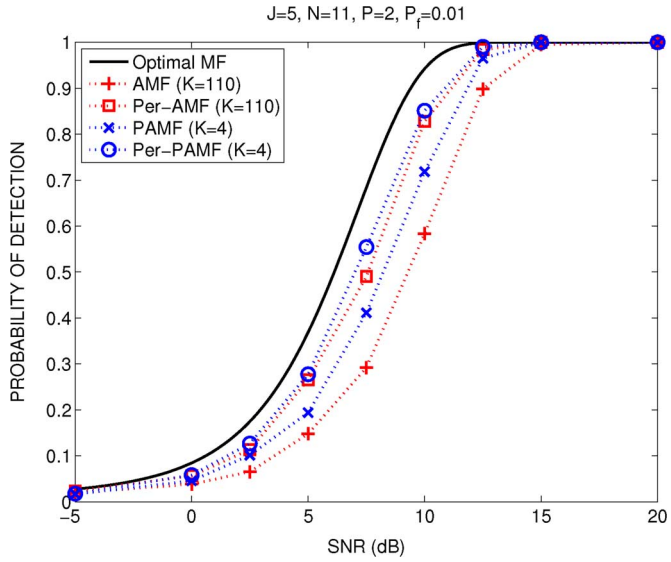

Fig. 2. Probability of detection versus SINR for $K = 4$ when $J = 5, N = 11$, $P = 2$, and $P_f = 0.01$.

IV. Performance Evaluation

In this section, simulation results are provided to demonstrate the efficiency of the proposed Per-PAMF in the training-limited case, e.g., $K \ll JN$. The performance is evaluated with two datasets: 1) a synthesized AR dataset, in which the disturbance signal $\mathbf{d}_k$ is generated as a multichannel second-order AR process $(P = 2)$ with a AR coefficient $\mathbf{A}$ and a spatial covariance matrix $\mathbf{Q}$ satisfying (6) and (5); and 2) a physical clutter dataset, in which the disturbance is generated according to the clutter model used in [20]. The signal-to-interference-plus-noise ratio (SINR) is defined as

$$\mathrm{SINR} = |\alpha|^2 \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s} \tag{33}$$

where $\mathbf{R}$ is the spatial-temporal covariance matrix. Regarding the synthesized dataset, $\mathbf{R}$ corresponds to the selected $\mathbf{A}$ and $\mathbf{Q}$. The simulated performance is obtained by using at least 10 000 Monte Carlo trials for the probability of false alarm $P_f = 0.01$. Performance comparisons are made among the nonparametric AMF [5], the nonparametric persymmetric AMF (Per-AMF) [18], the conventional PAMF [7], and the clairvoyant matched filter (MF) [5], [20]. Particularly, the simulated scenario uses $J = 5$ antenna elements and $N = 11$ pulses, while the number of training signals are, respectively, $K = 2$, $K = 4$, and $K = 8$.

A. Synthesized AR Dataset

In this case, the steering vector $\mathbf{s}$ is generated with a normalized spatial frequency $f_s = 0.2$ and a normalized Doppler frequency is $f_d = 0.2$, respectively. Fig. 1 shows the probability of detection versus the SINR with comparably sufficient training signals. In this case, the performance gain of the Per-PAMF over the conventional PAMF is marginal since both detectors have enough training signals to obtain good estimates of the unknown parameters. Meanwhile, both parametric detectors, i.e., the PAMF and Per-PAMF with $K = 8$ training signals, show better detection performance than the nonparametric covariance matrix based AMF and Per-AMF with $K = 2JN = 110$ training signals.

In Fig. 2, the number of training signals is reduced to $K = 4$. As shown in Fig. 2, the performance gap between the Per-PAMF and the PAMF is about 1.5 dB when $P_d = 0.8$, while the Per-PAMF with $K = 4$ training signals is slightly better than its nonparametric counterpart Per-AMF with $K = 110$ training signals. The most challenging case is the third scenario where only $K = 2$ training signals are available. As shown in Fig. 3, the conventional PAMF gives much worse performance than the Per-PAMF. In other words, with only $K = 2$ training signals, the conventional PAMF cannot obtain reliable estimates of unknown parameters, e.g., $\mathbf{A}$ and $\mathbf{Q}$, which leads to performance degradation, while the Per-PAMF has better efficiency of using training signals for unknown parameter estimation and thus maintains its performance even with only $K = 2$ training signals. The performance gain for the Per-PAMF over the conventional PAMF is about 5 dB when $P_d = 0.8$. On the other hand, the AMF cannot work functionally with $K = JN = 55$ training signals and the Per-AMF with $K = 55$ training signals gives similar detection performance with the Per-PAMF with $K = 2$ training signals.

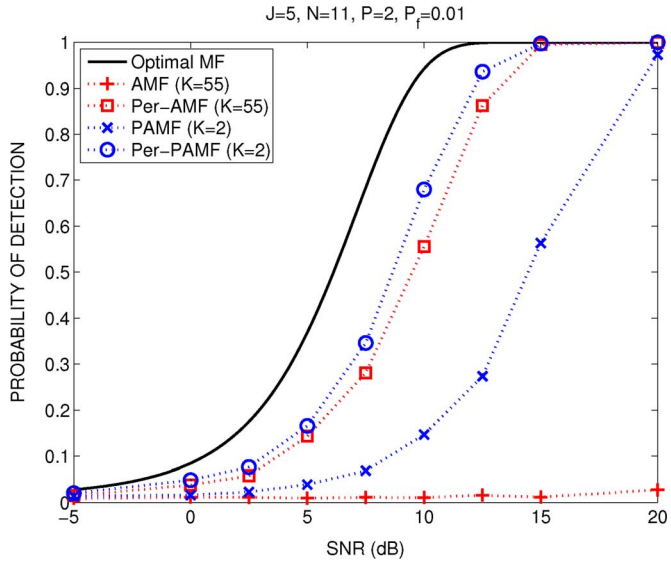

Fig. 3. Probability of detection versus SINR for $K = 2$ when $J = 5, N = 11$, $P = 2$, and $P_f = 0.01$.

From all three scenarios, it is clearly seen that, when there is a sufficient number of training signals, the performance gain is marginal and thus one may not need to utilize the prior persymmetric information.

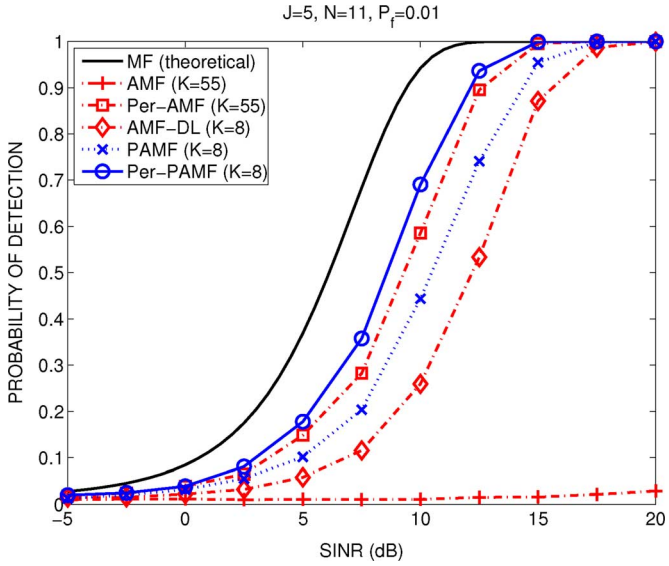

Fig. 4. Probability of detection versus SINR for $K = 8$ when $J = 5$, $N = 11$, and $P_f = 0.01$.

In contrast, the advantage of the proposed Per-PAMF becomes evident when the training signals are not enough for the PAMF to reliably estimate the unknown parameters and to maintain a reasonable detection performance. Therefore, in the training-limited scenarios, it is crucial to make use of the prior structure knowledge to improve the robustness of the parametric detectors. Here, we have a brief note on the CFAR property of the proposed Per-PAMF. Similar to the conventional PAMF, the CFAR property is usually achieved in the asymptotic case with a sufficiently large N (the number of pulses) or at the sufficient training scenario where $K \gg JN$, but may vanish when $J \approx N$ or when the disturbance does not follow the AR process; see [7, Sec. VII.B].

B. Physical Clutter Dataset

Unlike the previous section which used synthesized AR dataset, this section provides simulation results in a different environment, where the disturbance is generated from a physical clutter model described in [20], rather than a multichannel AR process. The platform is at altitude of 9 km with the velocity of 50 m/s, and the range of interest is 130 km. The clutter-to-noise ratio (CNR) is 35 dB. The clutter is divided among 360 clutter patches equally distributed in the azimuth about the platform and the RCS for each patch is weighted by a transmit pattern with a backlobe level of -30 dB; see [20, (80) and Figs. 9 and 10]. Moreover, we assume a linear symmetric array and a set of pulses with a constant PRF of 300 Hz, which results in a persymmetric spatial steering vector and a Vandermonde temporal steering vector:

$$\mathbf{s}_s = \left[e^{j2\pi 3f_s}, e^{j2\pi f_s}, 1, e^{-j\pi f_s}, e^{-j2\pi 3f_s}\right]^T$$
$$\mathbf{s}_t = \left[1, e^{j2\pi f_d}, e^{j2\pi 2f_d}, \ldots, e^{j2\pi 9f_d}, e^{j2\pi 10f_d}\right]^T.$$

The steering vectors for the clutter sources are similarly generated.

We consider the case with $K = 8$ training signals. Since there is no knowledge about the AR order, we try a variety of AR orders, i.e., $P = \{1, 2, 3\}$, and select the one giving the best performance. In contrary, we use $K = 55$ training signals for the nonparametric detectors, e.g., the conventional AMF and Per-AMF. Alternatively, the AR order can be adaptively determined from the training signals with conventional order determination techniques, e.g., the Akaike information criterion (AIC) and the minimum description length (MDL). Specifically, a low-complexity, joint model-order selection and parametric detection procedure was proposed in [21]. The simulated results are shown in Fig. 4. It is seen that the proposed Per-PAMF is better than the conventional PAMF with a performance gain of 2 dB at $P_d = 0.9$. Moreover, it is seen that the proposed Per-PAMF detector needs much less training signals ($K = 8$ versus $K = 55$) than the nonparametric Per-AMF detector to achieve similar detection performance, and is significantly better than the conventional AMF detector with $K = 55$ training signals. We also note that, compared with the same case of $K = 8$ but with the synthesized AR dataset in Fig. 1, the performance gap between the PAMF/Per-PAMF and the optimal matched filter is increased, which shows that the parametric PAMF/Per-PAMF detectors suffer from the model mismatch. Finally, we also compare the proposed Per-PAMF with the diagonal-loaded AMF [22], [23] with the same amount of training signals, i.e., $K = 8$. It is seen from Fig. 4 that, with a choice of 10-dB loading factor (ten times the white noise level), the diagonal-loaded AMF (denoted as the AMF-DL) gives a detection performance, much better than that of the conventional AMF detector but still worse than that of the proposed Per-PAMF detector.

V. Conclusion

This correspondence extends the conventional PAMF by exploiting the structure properties of the disturbance covariance matrix, for widely used systems with a symmetric antennas geometry and pulses with a constant PRF. The developed Per-PAMF shares the same detection statistics as the conventional PAMF but utilizes the structure information through the estimation of the unknown AR coefficient matrices $\mathbf{A}$ and spatial covariance matrix $\mathbf{Q}$. Numerical results have verified that the proposed Per-PAMF gives better detection performance than the conventional PAMF as well as the covariance matrix based detectors when training signals are limited.

Appendix

In the following, we show that the persymmetric AR based disturbance in Section II has a *persymmetric-block-Toeplitz* space-time covariance matrix. First, the space-time covariance matrix of an arbitrary multichannel AR process $\mathbf{x}(n)$ is defined by

$$\mathbf{R} = E\left\{\begin{bmatrix}\mathbf{x}(0)\\ \mathbf{x}(1)\\ \vdots\\ \mathbf{x}(N-1)\end{bmatrix}\begin{bmatrix}\mathbf{x}^H(0) & \mathbf{x}^H(1) & \cdots & \mathbf{x}^H(N-1)\end{bmatrix}\right\}$$
$$= \begin{bmatrix}\mathbf{R}(0) & \mathbf{R}(-1) & \cdots & \mathbf{R}(1-N)\\ \mathbf{R}(1) & \mathbf{R}(0) & \cdots & \mathbf{R}(2-N)\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{R}(N-1) & \mathbf{R}(N-2) & \cdots & \mathbf{R}(0)\end{bmatrix} \quad (34)$$

where $\mathbf{R}(m) = E[\mathbf{x}(n)\mathbf{x}^H(n-m)]$ which implies that $\mathbf{R}(-m) = \mathbf{R}^H(m)$. It is easily observed that $\mathbf{R}$ is a block-Toeplitz matrix.

To further show $\mathbf{R}$ is persymmetric-block-Toeplitz, we need to prove that the sub-block matrix $\mathbf{R}(m)$ is a persymmetric matrix. To this end, we consider an iterative procedure to find an explicit expression of $\mathbf{R}(m)$ as a function of the AR coefficient matrices $\mathbf{A}(p)$ and spatial covariance matrix $\mathbf{Q}$ and, then, prove the persymmetric property of the sub-block matrices $\mathbf{R}(m)$. First, we treat the AR process as a causal filter and $\mathbf{x}(n)$ is therefore the output of the causal filter

$$\mathbf{x}(n) = \sum_{m=0}^{\infty} \mathbf{H}(m)\boldsymbol{\varepsilon}(n-m). \quad (35)$$

Given $\mathbf{H}(m)$, we have, for $m > 0$

$$\begin{aligned}\mathbf{R}(m) &= E[\mathbf{x}(n)\mathbf{x}^H(n-m)] \\ &= E\left[\sum_{l_1=0}^{\infty}\mathbf{H}(l_1)\varepsilon(n-l_1)\sum_{l_2=0}^{\infty}\varepsilon^H(n-l_2)\mathbf{H}^H(l_2)\right] \\ &= \sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\mathbf{H}(l_1)E[\varepsilon(n-l_1)\varepsilon^H(n-l_2)]\mathbf{H}^H(l_2) \\ &\overset{(a)}{=} \sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\mathbf{H}(l_1)\mathbf{Q}\delta(m+l_2-l_1)\mathbf{H}^H(l_2) \\ &= \sum_{l_1=0}^{\infty}\mathbf{H}(l_1)\mathbf{Q}\mathbf{H}^H(l_1-m) \end{aligned} \tag{36}$$

where (a) holds since the driving noise $\varepsilon(n)$ is temporally white across n. From (35), it is seen that the output $\mathbf{x}(n)$ with respect to impulse input is

$$\begin{aligned}\mathbf{x}(n) &= \sum_{m=0}^{\infty}\mathbf{H}(m)\boldsymbol{\delta}(n-m) \\ &= \mathbf{H}(0)\boldsymbol{\delta}(n) + \mathbf{H}(1)\boldsymbol{\delta}(n-1) + \mathbf{H}(2)\boldsymbol{\delta}(n-2) + \cdots \end{aligned} \tag{37}$$

and we can enumerate

$$\begin{aligned}\mathbf{x}(0) &= \mathbf{H}(0)\boldsymbol{\delta}(0) \\ \mathbf{x}(1) &= \mathbf{H}(1)\boldsymbol{\delta}(0) \\ \mathbf{x}(2) &= \mathbf{H}(2)\boldsymbol{\delta}(0) \end{aligned} \tag{38}$$

and so on. Meanwhile, since $\mathbf{x}(n)$ is a multichannel AR process, we have

$$\mathbf{x}(n) = -\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{x}(n-p) + \varepsilon(n) \tag{39}$$

and, therefore, the output $\mathbf{x}(n)$ with respect to impulse input can be alternatively obtained as

$$\begin{aligned}\mathbf{x}(0) &= \boldsymbol{\delta}(0) \\ \mathbf{x}(1) &= -\mathbf{A}^H(1)\mathbf{x}(0) = -\mathbf{A}^H(1)\boldsymbol{\delta}(0) \\ \mathbf{x}(2) &= -\mathbf{A}^H(1)\mathbf{x}(1) - \mathbf{A}^H(2)\mathbf{x}(0) \\ &= \{(-\mathbf{A}^H(1))^2 - \mathbf{A}^H(2)\}\boldsymbol{\delta}(0) \end{aligned} \tag{40}$$

and so on. By comparing (38) with (40), we have

$$\begin{aligned}&\mathbf{H}(0) = \mathbf{I}, \\ &\mathbf{H}(1) = -\mathbf{A}^H(1), \\ &\mathbf{H}(2) = -\mathbf{A}^H(1)\mathbf{H}(1) - \mathbf{A}^H(2)\mathbf{H}(0), \\ &\mathbf{H}(3) = -\mathbf{A}^H(1)\mathbf{H}(2) - \mathbf{A}^H(2)\mathbf{H}(1) - \mathbf{A}^H(3)\mathbf{H}(0) \end{aligned} \tag{41}$$

and so on. As a result, the impulse response matrix $\mathbf{H}(m)$ can be determined in an iterative way as

$$\mathbf{H}(m) = \begin{cases} -\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{H}(m-p), & m \geq P \\ -\sum_{p=1}^{m}\mathbf{A}^H(p)\mathbf{H}(m-p), & 0 < m < P \\ \mathbf{I}, & m = 0. \end{cases} \tag{42}$$

Since $\mathbf{A}(p), p = 1, \ldots, P$, satisfy the persymmetric property as shown in (6), we can conclude from (42) that $\mathbf{H}(m)$ is also subject to the persymmetric constraint as

$$\mathbf{H}(m) = \mathbf{E}\mathbf{H}^*(m)\mathbf{E}. \tag{43}$$

In addition to the spatial persymmetry of $\mathbf{Q}$ in (5), we can rewrite the block matrices $\mathbf{R}(m)$ in (36) as

$$\begin{aligned}\mathbf{E}\mathbf{R}^*(m)\mathbf{E} &= \sum_{l_1=0}^{\infty}\mathbf{E}\mathbf{H}^*(l_1)\mathbf{Q}^*\mathbf{H}^T(l_1-m)\mathbf{E} \\ &\overset{(a)}{=} \sum_{l_1=0}^{\infty}\{\mathbf{E}\mathbf{H}^*(l_1)\mathbf{E}\}\{\mathbf{E}\mathbf{Q}^*\mathbf{E}\}\{\mathbf{E}\mathbf{H}^T(l_1-m)\mathbf{E}\} \\ &\overset{(b)}{=} \sum_{l_1=0}^{\infty}\mathbf{H}(l_1)\mathbf{Q}\mathbf{H}^H(l_1-m) \\ &= \mathbf{R}(m) \end{aligned} \tag{44}$$

where (a) used the fact that $\mathbf{E}\mathbf{E} = \mathbf{I}$ and (b) holds due to the persymmetric properties of the $\mathbf{H}(m)$ and $\mathbf{Q}$. Combining (34) and (44), the overall space-time covariance matrix $\mathbf{R}$ of a persymmetric AR process is proved to be a persymmetric-block-Toeplitz matrix.

REFERENCES

[1] A. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Process. Mag.*, vol. 14, pp. 49–83, Nov. 1997.

[2] R. Klemm, *Principles of Space-Time Adaptive Processing*. London, U.K.: IET, 2002.

[3] L. Huang, E. A. Thompson, V. Schmithorst, S. K. Holland, and T. M. Talavage, "Partially adaptive STAP algorithm approaches to functional MRI," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 2, pp. 518–521, Feb. 2009.

[4] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, pp. 115–127, Mar. 1986.

[5] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[6] A. De Maio, "Rao test for adaptive detection in Gaussian interference with unknown covariance matrix," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3577–3584, Jul. 2007.

[7] J. R. Román, M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 677–692, Apr. 2000.

[8] J. H. Michels, M. Rangaswamy, and B. Himed, "Performance of parametric and covariance based STAP tests in compound-Gaussian clutter," *Digit. Signal Process.*, vol. 12, no. 2, 3, pp. 307–328, Apr./Jul. 2002.

[9] K. J. Sohn, H. Li, and B. Himed, "Parametric GLRT for multichannel adaptive signal detection," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5351–5360, Nov. 2007.

[10] P. Wang, H. Li, and B. Himed, "A new parametric GLRT for multichannel adaptive signal detection," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 317–325, Jan. 2010.

[11] P. Wang, H. Li, and B. Himed, "A Bayesian parametric test for multichannel adaptive signal detection in non-homogeneous environments," *IEEE Signal Process. Lett.*, vol. 17, no. 4, pp. 351–354, Apr. 2010.

[12] A. Cantoni and P. Butler, "Properties of the eigenvectors of persymmetric matrices with applications to communication theory," *IEEE Trans. Commun.*, vol. 24, no. 8, pp. 804–809, Aug. 1976.

[13] R. Nitzberg, "Application of maximum likelihood estimation of persymmetric covariance matrices to adaptive processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 16, no. 1, pp. 124–127, Jan. 1980.

[14] L. Cai and H. Wang, "A persymmetric multiband GLR algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 806–816, July 1992.

[15] E. Conte and A. De Maio, "Exploiting persymmetry for CFAR detection in compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 2, pp. 719–724, Apr. 2003.

[16] A. De Maio, G. Foglia, E. Conte, and A. Farina, "CFAR behavior of adaptive detectors: An experimental analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 233–251, Jan. 2005.

[17] M. Casillo, A. De Maio, S. Iommelli, and L. Landi, "A persymmetric GLRT for adaptive detection in partially-homogeneous environment," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 1016–1019, Dec. 2007.

[18] G. Pailloux, P. Forster, J. P. Ovarlez, and F. Pascal, "On persymmetric covariance matrices in adaptive detection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Las Vegas, NV, Mar./Apr. 2008, pp. 2305–2308.

[19] M. Jansson, P. Wirfalt, K. Werner, and B. Ottersten, "ML estimation of covariance matrices with Kronecker and persymmetric structure," in *Proc. IEEE DSP/SPE Workshop*, Marco Island, FL, Jan. 2009, pp. 298–301.
[20] J. Ward, "Space-time adaptive processing for airborne radar," Lincoln Lab., MIT, Cambridge, MA, Tech. Rep. 1015, Dec. 1994.
[21] K. J. Sohn, H. Li, and B. Himed, "Recursive parametric tests for multichannel adaptive signal detection," *IET Radar, Sonar, Navig.*, vol. 2, no. 1, pp. 63–70, Feb. 2008.
[22] Y. I. Abramovich, "Controlled method for adaptive optimization of filters using the criterion of maximum SNR," *Radio Eng. Electron. Phys.*, vol. 26, pp. 87–95, Mar. 1981.
[23] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.

Single-Carrier Systems With MMSE Linear Equalizers: Performance Degradation due to Channel and CFO Estimation Errors

Athanasios P. Liavas and Despoina Tsipouridou

***Abstract*—We assess the impact of the channel and the carrier frequency offset (CFO) estimation errors on the performance of single-carrier systems with MMSE linear equalizers. Performance degradation is caused by the fact that a *mismatched* MMSE linear equalizer is applied to channel output samples with *imperfectly canceled* CFO. Assuming a single-block training, we develop an asymptotic expression for the excess mean square error (EMSE) induced by the channel and CFO estimation errors and derive a simple EMSE approximation which reveals the following: 1) performance degradation is mainly caused by the imperfectly canceled CFO and 2) the EMSE is approximately proportional to the CFO estimation error variance, with the factor of proportionality being independent of the training sequence. We also highlight the fact that the placement of the single-block training at the middle of the packet is a good practice.**

***Index Terms*—Joint channel and CFO estimation, linear MMSE equalization.**

I. Introduction

A problem that frequently arises in packet-based wireless communication systems is the *joint* estimation of the frequency selective channel and the CFO [1], [2]. Optimal training sequence (TS) design for this problem has been considered in [2], where the optimized cost function was the *asymptotic* Cramér-Rao bound (CRB). However, in [2], the channel and CFO estimation errors were assigned equal weight which might be *suboptimal* since "*... presumably channel estimation errors will have a different impact, e.g., on bit-error rate, than frequency estimation errors*" [2].

Manuscript received June 20, 2011; revised October 31, 2011 and January 19, 2012; accepted February 13, 2012. Date of publication March 05, 2012; date of current version May 11, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yimin D. Zhang.

The authors are with the Department of Electronic and Computer Engineering, Technical University of Crete, Greece (e-mail: liavas@telecom.tuc.gr; despoina@telecom.tuc.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2012.2189858

It seems that the *unequal weighting* problem cannot be resolved unless we consider specific communication systems. An important structure to study is a single-carrier system with an MMSE linear equalizer. Performance degradation is caused by the fact that a *mismatched* MMSE linear equalizer is applied to channel output samples with *imperfectly canceled* CFO. This system has been considered in [3], where the authors derived an expression for the EMSE and designed optimal TSs.

We consider the same system as in [3] but our aim is different. More specifically, we assume a *single-block* training and, as done in [3], we develop an asymptotic expression for the induced EMSE which, however, is difficult to interpret. Our main contribution lies in the fact that, assuming *small ideal MMSE*, we derive a simple and informative EMSE approximation, which reveals the following:

1) the dominant error source is the imperfectly canceled CFO;
2) the EMSE is approximately proportional to the CFO estimation error variance, with the factor of proportionality being *independent* of the TS.[1]

We also highlight the fact that the placement of the single-block TS at the middle of the transmitted packet is a good practice.

Notation: Superscripts ^{T}, ^{H}, and * denote transpose, conjugate transpose, and elementwise conjugation, respectively. $\mathrm{tr}(\cdot)$ denotes the trace operator, $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number, and $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. $\sigma_{\max}(\cdot)$, $\sigma_{\min}(\cdot)$, $\|\cdot\|_2$, $\|\cdot\|_F$, and $k_2(\cdot)$ denote, respectively, the maximum singular value, the minimum singular value, the spectral norm, the Frobenius norm, and the condition number, with respect to the spectral norm, of the matrix argument. $\mathcal{E}[X]$ denotes the expected value of X. $\mathbf{P}_{\mathcal{R}(\mathbf{A})}$ and $\mathbf{P}^{\perp}_{\mathcal{R}(\mathbf{A})}$ denote, respectively, the orthogonal projector onto the column space of matrix $\mathbf{A}$ and its orthogonal complement.

II. Channel and CFO Estimation

A. The Channel Model

We consider a packet-based single-carrier system with input packet length N. We assume that the baseband-equivalent frequency-selective channel has impulse response $\mathbf{h} \triangleq [h_0 \cdots h_L]^T$, angular CFO ω, and phase ϕ. The output at time instant n, for $n = 1, \ldots, N+L$, is

$$r_n = \mathrm{e}^{j(\omega n+\phi)} \sum_{l=0}^{L} h_l a_{n-l} + w_n \tag{1}$$

where $\{a_n\}_{n=1}^{N}$ and $\{w_n\}_{n=1}^{N+L-1}$ denote the channel input and additive channel noise, respectively. The input symbols are i.i.d. unit variance circular. The noise samples are i.i.d. circular Gaussian with variance σ_w^2. In the sequel, we absorb term $e^{j\phi}$ into channel $\mathbf{h}$.

The channel output vector $\mathbf{r}_{n:n-M} \triangleq [r_n \ \cdots \ r_{n-M}]^T$ can be expressed as

$$\mathbf{r}_{n:n-M} = \mathbf{\Gamma}_{n:n-M}(\omega)\mathbf{H}\mathbf{a}_{n:n-L-M} + \mathbf{w}_{n:n-M} \tag{2}$$

where

$$\mathbf{\Gamma}_{n:n-M}(\omega) \triangleq \mathrm{diag}(\mathrm{e}^{j\omega n}, \ldots, \mathrm{e}^{j\omega(n-M)}) \tag{3}$$

and $\mathbf{H}$ is the $(M+1)\times(M+L+1)$ Toeplitz filtering matrix constructed by $\mathbf{h}$.

[1]Thus, optimal TS design for CFO estimation is also highly relevant for *joint* channel and CFO estimation.