Adaptive Widely Linear Reduced-Rank Interference Suppression Based on the Multistage Wiener Filter

Nuan Song, *Student Member, IEEE*, Rodrigo C. de Lamare, *Senior Member, IEEE*, Martin Haardt, *Senior Member, IEEE*, and Mike Wolf

***Abstract*—We propose a widely linear multistage Wiener filter (WL-MSWF) receiver to suppress inter-/intra-symbol interference, multiuser interference, and narrowband interference in a high data rate direct-sequence ultra wideband (DS-UWB) system. The proposed WL receiver fully exploits the second-order statistics of the received signal, yielding a smaller Minimum Mean Square Error (MMSE) than the linear receiver. The WL-MSWF receiver mainly consists of a low-rank transformation and an adaptive reduced-rank filter. The rank-reduction is achieved via a transformation matrix. Based on the linear MSWF concept, two constructions of this rank-reduction matrix, namely total WL (TWL) and quasi WL (QWL), are proposed. We develop stochastic gradient (SG) and recursive least squares (RLS) adaptive versions of the proposed TWL/QWL-MSWF and theoretically analyze their convergence behavior. The comparison of the proposed TWL/QWL-MSWF and the existing algorithms is carried out in terms of the computational complexity and the resulting MMSE performance. Extensive simulation results show that the proposed TWL/QWL-MSWF schemes outperform the existing schemes in both convergence and steady-state performance under various conditions.**

***Index Terms*—Direct-sequence ultrawideband, multistage Wiener filter, narrowband interference, noncircular, reduced-rank, widely linear.**

I. Introduction

COMPLEX-VALUED signals have been widely used in various fields such as mobile communications, smart antennas, radar, biomedicine, optics and seismics, etc. Complex-domain representations are quite convenient to physically characterize the signals in practice [1]–[3]. Most parameter estimation and filtering techniques for complex-valued signals, whose samples are often organized in a vector $\boldsymbol{r}$, are based on their second-order statistics. It is often assumed that the signal $\boldsymbol{r}$ is second-order circular (or proper). As a result, only the covariance matrix $\boldsymbol{R} = \mathbb{E}\{\boldsymbol{r}\boldsymbol{r}^H\}$ is utilized for signal processing. However, it is shown that in many applications when $\boldsymbol{r}$ is noncircular or improper, the second-order behavior should be described by both the covariance matrix $\boldsymbol{R}$ and the pseudo-covariance (also called complementary covariance in [2], [4]) matrix $\check{\boldsymbol{R}} = \mathbb{E}\{\boldsymbol{r}\boldsymbol{r}^T\}$, where $\check{\boldsymbol{R}}$ is not vanishing [5]. The improperness may arise from modulations which employ improper signal constellations such as binary phase shift keying (BPSK), amplitude shift keying (ASK), biorthogonal keying (BOK), or the ones that can be interpreted as a real constellation after reformulation such as offset quadrature phase shift keying (OQPSK), minimum shift keying (MSK), or Gaussian MSK (GMSK) [6].

Widely linear (WL) processing, which fully exploits the second-order statistics ($\boldsymbol{R}$ and $\check{\boldsymbol{R}}$) of improper signals, can significantly improve the estimation performance [4], [5], [7], [8]. The WL filtering techniques have gained a great popularity in the applications of interference suppression, equalization, and synchronization. Data-aided and blind adaptive WL minimum mean square error (MMSE) receivers based on recursive least squares (RLS) [9] and stochastic gradient (SG) [10] techniques are proposed to achieve interference suppression in BPSK-based direct sequence code division multiple sccess (DS-CDMA) systems. Different equalization strategies based on WL processing have been developed for DS-CDMA [11] and DS ultra wideband (DS-UWB) [12]. The authors of [13] provide new insights into the optimum WL array receivers for their applications to single antenna interference cancellation techniques [14] as well as to synchronization schemes [15] for GSM systems, considering BPSK, MSK, and GMSK signals in the presence of noncircular interferences. Compared to the linear processing, these WL receivers exhibit an increased robustness against interference, and the related adaptive algorithms are able to provide a better convergence performance. One important property is that the WL estimate of the real-valued data from a sequence of complex and improper observations results in a real-valued estimate. This not only produces a smaller estimation error than the linear estimate but may also reduce the receiver complexity since only the real-valued signal is processed [9], [11].

In many situations, the observation data used for parameter estimation has a large size due to a high processing gain, a large number of antennas, or numerous multipath components, which requires a long receive filter. However, a filter with a large number of taps requires substantial training, which considerably slows down the convergence speed, and becomes highly sensitive to interference. Thereby, in order to decrease the number of estimated parameters (e.g., filter coefficients), reduced-rank processing can be applied such that the received vector is transformed into a lower dimensional subspace and the

Manuscript received November 01, 2010; revised July 21, 2011 and April 04, 2012; accepted April 13, 2012. Date of publication May 03, 2012; date of current version July 10, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ut-Va Koc. This work was supported in part by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under Contract WO 1442/1-2. This paper was presented in part at the *The Seventh International Symposium on Wireless Communication Systems (ISWCS 2010)*, York, U.K., September, 2010.

N. Song, M. Haardt, and M. Wolf are with the Communications Research Laboratory, Ilmenau University of Technology, D-98684 Ilmenau, Germany (e-mail: nuan.song@tu-ilmenau.de; martin.haardt@tu-ilmenau.de; mike.wolf@tu-ilmenau.de).

R. C. de Lamare is with the Communications Research Group, Department of Electronics, University of York, Heslington, North Yorkshire, York Y010 5DD, U.K. (e-mail: rcdl500@ohm.york.ac.uk).

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.2012.2197747

filtering optimization is carried out within this subspace. Compared to the full-rank techniques, the reduced-rank methods are able to achieve a faster convergence, an increased robustness against interference, and a lower complexity by estimating a reduced number of parameters. There have been several reduced-rank techniques proposed for interference suppression. Some well-known approaches, namely the "principal components" (PC) [16], [17] and the "cross spectral" metric [18], exclusively rely on the eigen-decomposition for estimating the signal subspace. This demands huge computational efforts and an often large rank to reach a satisfactory performance [18]. A more effective method called multistage Wiener filter (MSWF) was proposed in [19] and [20]. In contrast to the eigen-decomposition algorithms, the MSWF does not require the knowledge of the signal subspace but utilizes a successive orthogonal decomposition for parameter estimation. It is capable of attaining an improved convergence with a filter rank which is much less than the dimension of the signal subspace [21]. Another reduced-rank approach is called auxiliary vector filtering (AVF), which iteratively updates the filter weights according to a sequential and conditioned optimization of auxiliary vectors [22]. Both the MSWF and the AVF estimators can be combined with different design criteria such as MMSE [23], constrained minimum variance (CMV) [24], or constrained constant modulus (CCM) [24], [25]. The AVF outperforms the MSWF but has a higher complexity. In the WL case, both the original received signal $\boldsymbol{r}$ and its complex conjugate $\boldsymbol{r}^*$ have to be considered, which further increases the filter length and thus decelerates the convergence [3], [26]. Reduced-rank techniques are thus more attractive and efficient in WL signal processing. So far, most of the reduced-rank algorithms are based on linear processing [24], [27]–[29]. One of the few algorithms that combine both is the WL reduced-rank Wiener filter investigated in [4], where the computationally expensive eigen-decomposition is employed. This reduced-rank WL estimator usually requires twice the rank of its linear counterpart.

Wireless communication systems can substantially benefit from the use of UWB signals. However, in high data rate DS-UWB applications [30], the system performance may be deteriorated by inter-/intra-symbol interference (ISI), multiuser interference (MUI), or even by the interference from other non-UWB systems operating in the same bandwidth. The emissions of the IEEE 802.11a wireless local area network (WLAN) in the range of 5.2 GHz [31], for example, occur in a frequency band which is permitted for UWB operations in the US [32]. The IEEE 802.11a WLAN signal may exhibit a much higher power than the UWB signal and is treated as narrowband interference (NBI). The large bandwidth requires a high sampling rate and leads to a received vector with a large size. The reduced-rank techniques are thus very promising for interference suppression in DS-UWB systems [33]. One mandatory modulation scheme for DS-UWB systems is the noncircular BPSK modulation [30]. Therefore, the combination of the robust MSWF method and the WL processing is motivated to ensure a faster convergence and a lower complexity than the full-rank and/or the linear counterparts.

In this paper, we propose a WL-MSWF receiver for interference suppression in DS-UWB systems. The proposed receiver consists of a bijective transformation to form an augmented observation vector, a rank-reduction block to perform the low-rank transformation, and an adaptive reduced-rank filter. In contrast to the WL reduced-rank Wiener filter based on PC [4], the proposed receiver applies the linear MSWF concept in the WL case. It does not require the eigen-decomposition and thus its computational complexity is considerably reduced. Combining the WL processing with the MSWF not only achieves a lower MMSE than that of the linear case, but also has a better convergence performance compared to the full-rank techniques.

The main contributions of our work are summarized as follows:

1) We derive the WL-MSWF and characterize some key properties. Two constructions of the rank-reduction matrix are introduced, namely the total WL (TWL) and the quasi WL (QWL) designs.
2) For both low-rank WL projection methods (TWL and QWL), we develop the SG and the RLS adaptive algorithms to compute the WL-MSWF.
3) We analyze the statistical performance in terms of mean square error (MSE) for the adaptive SG and RLS algorithms, including the stability and the convergence performance.
4) We estimate and compare the computational complexity of the proposed and the existing schemes in terms of real additions and multiplications.
5) The proposed TWL/QWL-MSWF schemes are examined for interference suppression in a DS-UWB system under realistic scenarios and compared with the linear MSWF counterparts, linear/WL full-rank schemes, as well as the linear/WL PC-based methods. We mainly focus on the scenario when both the signal and the interference (MUI and NBI) are noncircular. We also show the suitability of the proposed methods applied in the case when the desired signal is strictly circular but the interference (MUI or NBI) is noncircular.

Section II introduces the data model for the DS-UWB system. Section III reviews the linear reduced-rank Wiener filter according to the MSWF design. The WL-MSWF receiver is presented along with its key properties in Section IV. Section V details the SG/RLS adaptive algorithms for the WL-MSWF and analyzes the corresponding convergence and transient behavior. The computational complexity of all the studied algorithms is evaluated in Section VI. Section VII provides extensive simulation results of the proposed TWL/QWL-MSWF algorithms and compares them to the existing schemes.

Notation: The superscripts T, H, and $*$ stand for transpose, conjugate transpose, and complex conjugation, respectively. We use a as the subscript to denote the associated augmented quantities. The reduced-rank quantities are symbolized with a "bar". The Hadamard (element-wise) product is denoted by $\odot$. The expectation and the trace operations are expressed by $\mathbb{E}\{\cdot\}$ and $\mathrm{tr}\{\cdot\}$. The floor/ceiling operator $\lfloor x \rfloor / \lceil x \rceil$ rounds the argument x down/up to the closest integer that is less/greater than or equal to x. The operation $\Re\{\cdot\}$ is to take the real part of a variable. We use the bold capital letters to represent matrices and the bold small letters for vectors.

II. System Model

We consider the uplink of a BPSK DS-UWB system with N_u asynchronous users in the presence of NBI. In the complex baseband, the transmitted signal for the kth user is given by

$$s_k(t) = \sum_{i=-\infty}^{\infty} b_k(i) \sum_{n=0}^{N-1} \sqrt{E_k} c_k(n) g(t - iT_b - nT_c) \quad (1)$$

where $b_k(i) \in \{\pm 1\}$ is the ith BPSK symbol for the user k with unit variance $\sigma_b^2 = \mathbb{E}\{|b_1(i)|^2\} = 1$, T_b is the bit duration, E_k and $c_k(n) \in \left\{\pm 1/\sqrt{N}\right\}$ denote the corresponding energy per bit and the multiple access code with chip interval T_c. The baseband reference pulse $g(t)$ is the impulse response of a Root Raised Cosine (RRC) low pass filter with 30% excess bandwidth, i.e., the roll-off factor is $\beta = 0.3$. For both the low and high frequency bands, the filter cutoff frequency (-3 dB point) is $\frac{1}{2T_c}$ [30]. The processing gain N is equal to T_b/T_c.

Since the signal bandwidth is constrained to $B = (\beta+1)B_3$, the complex-valued impulse response of the multipath UWB channel can be fully described by the discrete response, i.e., tapped-delay line model written as $h_k(t) = \sum_{l=0}^{L-1} \alpha_k(l)\delta(t - l/B)$, where $\alpha_k(l)$ is the lth complex channel tap for the kth user and $\sum_{l=0}^{L-1} |\alpha_k(l)|^2 = 1$. In our case, the channel is assumed to be time-invariant block fading. For UWB communications with $B \geq 500$ MHz, the statistics of the path gains are different from those in narrowband systems. The large bandwidth also results in a significant number of resolvable multipath components and severe ISI.

The received signal at the output of a pulse matched filter with the impulse response $g(T - t)$ can be expressed as

$$y(t) = \sum_{i=-\infty}^{\infty} \sum_{k=1}^{N_u} \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} \sqrt{E_k} b_k \left(i + \left\lfloor \frac{n - D_k}{N} \right\rfloor\right) c_k(n) \alpha_k(l) \hat{g}\left(t - iT_b - nT_c - \frac{l}{B} - \tau_k\right) + J(t) + n(t) \quad (2)$$

where T represents the delay to ensure that the received pulse filter is causal, $\hat{g}(t) = g(t) * g(T - t)$, $J(t) = \hat{J}(t) * g(T - t)$, and $n(t) = \hat{n}(t) * g(T - t)$ are the filtered pulse, NBI, and noise, respectively. The zero-mean, complex additive white Gaussian noise (AWGN) $\hat{n}(t)$ is assumed to have a power spectral density N_0. Asynchronous (but chip synchronous) transmission is assumed, meaning that $\tau_1 - \tau_k = D_k T_c$, where the random variable D_k takes values in $\{0, 1, \ldots, N-1\}$ with equal probability. Without loss of generality, we assume that the delay of the desired user τ_1 is known and $\tau_1 = 0$ is chosen.

The NBI is often modeled as a single tone. It is more realistic to consider the orthogonal frequency division multiplexing (OFDM) signal from the IEEE 802.11a WLAN that overlays the UWB emission spectrum. Such an OFDM signal can be regarded as a sum of multiple single-tone NBIs, given by

$$\hat{J}(t) = \sqrt{\frac{P_J}{N_c}} \sum_{n=0}^{N_c-1} x_n e^{j(2\pi(f_J + n\cdot\Delta f)t + \theta)} \quad (3)$$

where P_J is the NBI power, N_c is the number of subcarriers, $x_n \in \{\pm 1\}$ is a BPSK-modulated symbol, f_J is the frequency difference between the carrier frequencies of the NBI and the UWB signal, Δf denotes the subcarrier frequency spacing, and a random phase θ is uniformly distributed in $[0, \pi)$. The Signal to NBI ratio is computed as $\mathrm{SIR} = E_s/(P_J T_s)$, where E_s is the signal energy per symbol and $E_s = E_b$ for BPSK. Usually in UWB communications, it is assumed that the duration of a NBI T_J is greater than T_b.

At the receiver, by sampling $y(t)$ at a chip rate $\frac{1}{T_c}$, the received signal vector is obtained. For the ith transmitted bit $i = 0, 1, \ldots, N_s - 1$, the corresponding received vector of length $M = N + L - 1$ can be written as

$$\boldsymbol{r}(i) = \sqrt{E_1} b_1(i) \boldsymbol{C}_1 \boldsymbol{h}_1 + \boldsymbol{v}(i) + \boldsymbol{\eta}(i) + \boldsymbol{j}(i) + \boldsymbol{n}(i) \quad (4)$$

including the desired user signal, the MUI part $\boldsymbol{v}(i)$, all the interference from the chips of the current symbols (intrasymbol) as well as from the previous and subsequent symbols (intersymbol) $\boldsymbol{\eta}(i)$, the NBI vector $\boldsymbol{j}(i)$ observed in the ith bit, and the AWGN. The code matrix for the kth user $\boldsymbol{C}_k \in \mathbb{R}^{M \times L}$ is a Toeplitz matrix, which can be expressed as

$$\boldsymbol{C}_k = \begin{bmatrix} c_k(0) & 0 & \cdots & 0 \\ c_k(1) & c_k(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_k(N) & c_k(N-1) & \cdots & 0 \\ 0 & c_k(N) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_k(N) \end{bmatrix}. \quad (5)$$

In what follows, we denote $\boldsymbol{X}(m:n,:)$ as a matrix consisting of the rows in $\boldsymbol{X}$ that are indexed from m to n.

The NBI vector is expressed as

$$\boldsymbol{j}(i) = \sqrt{\frac{P_J}{N_c}} \sum_{n=0}^{N_c-1} \boldsymbol{x}_n(i) \odot e^{j[2\pi(f_J + n\cdot\Delta f)\cdot \boldsymbol{k} T_c + \theta]} \quad (6)$$

where $\boldsymbol{x}_n(i) = \left[x_n\left(\left\lfloor \frac{iNT_c}{T_J} \right\rfloor\right), \ldots, x_n\left(\left\lfloor \frac{(iN+M-1)T_c}{T_J} \right\rfloor\right)\right]^T$, $j = \sqrt{-1}$, and $\boldsymbol{k} = [iN, iN+1, \ldots, iN+M-1]^T$.[1] We represent the asynchronous MUI each with an offset D_k by $\boldsymbol{v}(i) = \sum_{k=2}^{N_u} \sqrt{E_k} b_k(i) \check{\boldsymbol{C}}_k \boldsymbol{h}_k$, where $\check{\boldsymbol{C}}_k \in \mathbb{R}^{M \times L}$ is constructed from a zero matrix and the first $M - D_k$ rows of $\boldsymbol{C}_k$ defined as

$$\check{\boldsymbol{C}}_k = \begin{bmatrix} \boldsymbol{0}_{D_k} \\ \boldsymbol{C}_k(1 : M - D_k, :) \end{bmatrix}. \quad (7)$$

The ISI is expressed as

$$\boldsymbol{\eta}(i) = \sum_{k=1}^{N_u} \sqrt{E_k} \sum_{j=i-\xi}^{i-1} \tilde{\boldsymbol{C}}_k \boldsymbol{h}_k b_k(j) + \sum_{k=1}^{N_u} \sqrt{E_k} \sum_{j=i+1}^{i+\xi} \hat{\boldsymbol{C}}_k \boldsymbol{h}_k b_k(j)$$
$$\xi = \left\lceil \frac{L - 1 + D_k}{N} \right\rceil \quad (8)$$

[1]For a quantity, either a vector $\boldsymbol{x}$ or a matrix $\boldsymbol{X}$, the expression $e^{\boldsymbol{x}}$ or $e^{\boldsymbol{X}}$ returns the exponential for each element in $\boldsymbol{x}$ or $\boldsymbol{X}$ (MATLAB-like notation).

where $\tilde{\boldsymbol{C}}_k$ and $\hat{\boldsymbol{C}}_k \in \mathbb{R}^{M\times L}$ include the last $M-(i-j)N+D_k$ and the first $M-(j-i)N+D_k$ rows of $\boldsymbol{C}_k$, respectively, given by

$$\tilde{\boldsymbol{C}}_k = \begin{bmatrix} \boldsymbol{C}_k(\tilde{\xi}:M,:) \\ \boldsymbol{0} \end{bmatrix}, \text{ with } \tilde{\xi} = (i-j)N - D_k + 1,$$
$$\hat{\boldsymbol{C}}_k = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{C}_k(1:\hat{\xi},:) \end{bmatrix}, \text{ with } \hat{\xi} = M-(j-i)N + D_k. \quad (9)$$

III. Linear Reduced-Rank Wiener Filter

This section recalls the key concept of linear reduced-rank filters as preliminaries and summarizes the major results on the Linear MSWF (L-MSWF) algorithm. The cost function of the linear MMSE filter is given by[2]

$$J = \mathbb{E}\left\{ \left| b_1(i) - \boldsymbol{w}^H \boldsymbol{r}(i) \right|^2 \right\}. \quad (10)$$

The Wiener solution $\boldsymbol{w}_o = \boldsymbol{R}^{-1}\boldsymbol{p}$ with $\boldsymbol{p} = \mathbb{E}\{b_1^*(i)\boldsymbol{r}(i)\}$ can be estimated by adaptive algorithms such as SG and RLS [34]. However, when a large amount of data is processed, the conventional full-rank filter $\boldsymbol{w} \in \mathbb{C}^M$ that has the same length as the received vector $\boldsymbol{r}(i) \in \mathbb{C}^M$ exhibits a slow convergence and a high interference sensitivity. The reduced-rank technique is able to exploit the key features of the data and to reduce the number of adaptive parameters. The rank-reduction is achieved by transforming the received vector $\boldsymbol{r}(i)$ onto a D-dimensional subspace with $D \ll M$. Let us denote the rank-reduction matrix as $\boldsymbol{S}_D \in \mathbb{C}^{M\times D}$ and the reduced-rank vector is given by $\bar{\boldsymbol{r}}(i) = \boldsymbol{S}_D^H \boldsymbol{r}(i) \in \mathbb{C}^D$. The weight vector $\bar{\boldsymbol{w}} \in \mathbb{C}^D$ is estimated based on $\bar{\boldsymbol{r}}(i)$ and the filter length can be significantly reduced. The linear reduced-rank Wiener solution can be obtained as $\bar{\boldsymbol{w}}_o = \bar{\boldsymbol{R}}^{-1}\bar{\boldsymbol{p}}$, where the reduced-rank covariance matrix is $\bar{\boldsymbol{R}} = \mathbb{E}\{\bar{\boldsymbol{r}}(i)\bar{\boldsymbol{r}}^H(i)\} = \boldsymbol{S}_D^H \boldsymbol{R}\boldsymbol{S}_D$ and the reduced-rank cross-correlation vector is $\bar{\boldsymbol{p}} = \mathbb{E}\{b_1^*(i)\bar{\boldsymbol{r}}(i)\} = \boldsymbol{S}_D^H\boldsymbol{p}$. We can then calculate the corresponding MMSE

$$\bar{J}_{\min} = 1 - \bar{\boldsymbol{p}}^H \bar{\boldsymbol{R}}^{-1}\bar{\boldsymbol{p}} \quad (11)$$

and the SINR

$$\bar{\mathrm{SINR}} = \frac{\bar{\boldsymbol{p}}^H \bar{\boldsymbol{R}}^{-1}\bar{\boldsymbol{p}}}{1 - \bar{\boldsymbol{p}}^H \bar{\boldsymbol{R}}^{-1}\bar{\boldsymbol{p}}} = \frac{1}{\bar{J}_{\min}} - 1. \quad (12)$$

A. Linear MSWF

One method to construct the rank-reduction matrix is to apply the L-MSWF [19], [20]. It is shown in [21] that the rank-reduction matrix for the L-MSWF $\boldsymbol{S}_D$ is spanned by D normalized basis vectors $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_D$, where $\boldsymbol{f}_n = \boldsymbol{R}^{n-1}\boldsymbol{p}$ can be chosen. In other words, the linear reduced-rank filter transforms the received signal into the Krylov subspace represented by

$$\boldsymbol{S}_D = \left[\boldsymbol{p}, \boldsymbol{R}\boldsymbol{p}, \ldots, \boldsymbol{R}^{D-1}\boldsymbol{p}\right]. \quad (13)$$

[2]In some cases when the observation data vector $\boldsymbol{r}(i)$ is not stationary, e.g., it contains time-varying interference, the cost function shown in (10) also depends on the time index i [9]. For notational simplicity, we remove the index i in some cases that are related with nonstationary variables such as $\boldsymbol{R}_a$ and $\check{\boldsymbol{R}}$ shown in (15).

The MMSE and the output SINR of the L-MSWF asymptotically converge to the linear full-rank case, i.e., $\bar{J}_{\min} \geq J_{\min}$ and $\bar{\mathrm{SINR}} \leq \mathrm{SINR}$. Another important property is that the rank D required to achieve the full rank performance does not scale significantly with the system size such as the number of users N_u and the length of the received vector M. Generally, $D \leq 8$ can be chosen. The analysis in [19] and [20] also indicates that D can be decreased without considerably increasing the MSE.

The associated adaptive algorithms based on the powers of $\boldsymbol{R}$ given in (13) can be carried out in terms of SG or RLS [24]. Compared to the full-rank adaptive algorithms, the adaptive L-MSWF with a small rank D can provide a faster convergence speed and a better steady state performance for a given data record.

IV. Widely Linear MultiStage Wiener Filter

The main purpose of this section is to investigate the WL-MSWF techniques and compare them to the linear counterpart.

A. Preprocessing: Augmented Vector Formulation

In order to exploit the information contained in both second-order statistics, i.e., $\boldsymbol{R}$ and $\check{\boldsymbol{R}}$, the received signal $\boldsymbol{r}(i)$ and its complex conjugate $\boldsymbol{r}^*(i)$ are formulated into an augmented vector using a bijective transformation $\mathcal{T}$

$$\boldsymbol{r} \xrightarrow{\mathcal{T}} \boldsymbol{r}_a: \qquad \boldsymbol{r}_a = \frac{1}{\sqrt{2}}\left[\boldsymbol{r}^T, \quad \boldsymbol{r}^H\right]^T \in \mathbb{C}^{2M\times 1}. \quad (14)$$

The filter with coefficients $\boldsymbol{w}_a$, which is designed according to the augmented received vector $\boldsymbol{r}_a(i)$, is widely linear with $\boldsymbol{r}(i)$. It is thus named as a WL filter.

For example, the solution for a WL Wiener filter has a similar expression as in the linear case shown in Section III but with a subscript "a," denoting the augmented quantities. Let us then analyze the augmented covariance matrix, which can be represented by the covariance matrix $\boldsymbol{R}$ and the pseudo-covariance matrix $\check{\boldsymbol{R}}$ of $\boldsymbol{r}(i)$ as

$$\boldsymbol{R}_a = \frac{1}{2}\begin{bmatrix} \boldsymbol{R} & \check{\boldsymbol{R}} \\ \check{\boldsymbol{R}}^* & \boldsymbol{R}^* \end{bmatrix} \quad (15)$$

where

$$\boldsymbol{R} = \sum_{k=1}^{N_u} E_k \check{\boldsymbol{C}}_k \boldsymbol{h}_k \boldsymbol{h}_k^H \check{\boldsymbol{C}}_k^H + \boldsymbol{R}_{\eta\eta} + \boldsymbol{R}_{jj} + N_0 \boldsymbol{I}_M$$

and

$$\check{\boldsymbol{R}} = \sum_{k=1}^{N_u} E_k \check{\boldsymbol{C}}_k \boldsymbol{h}_k \boldsymbol{h}_k^T \check{\boldsymbol{C}}_k^T + \check{\boldsymbol{R}}_{\eta\eta} + \check{\boldsymbol{R}}_{jj}(i).$$

The covariance and pseudo-covariance matrices of ISI $\boldsymbol{\eta}(i)$ are denoted by $\boldsymbol{R}_{\eta\eta}$ and $\check{\boldsymbol{R}}_{\eta\eta}$ as

$$\boldsymbol{R}_{\eta\eta} = \sum_{k=1}^{N_u} E_k \left(\sum_{j=i-\xi}^{i-1} \tilde{\boldsymbol{C}}_k \boldsymbol{h}_k \boldsymbol{h}_k^H \tilde{\boldsymbol{C}}_k^H + \sum_{j=i+1}^{i+\xi} \hat{\boldsymbol{C}}_k \boldsymbol{h}_k \boldsymbol{h}_k^H \hat{\boldsymbol{C}}_k^H \right)$$
$$\check{\boldsymbol{R}}_{\eta\eta} = \sum_{k=1}^{N_u} E_k \left(\sum_{j=i-\xi}^{i-1} \tilde{\boldsymbol{C}}_k \boldsymbol{h}_k \boldsymbol{h}_k^T \tilde{\boldsymbol{C}}_k^T + \sum_{j=i+1}^{i+\xi} \hat{\boldsymbol{C}}_k \boldsymbol{h}_k \boldsymbol{h}_k^T \hat{\boldsymbol{C}}_k^T \right).$$

Since the modulated symbols $\boldsymbol{x}_n(i)$ on different subcarriers are uncorrelated, the second-order statistics of the BPSK-modulated OFDM NBI vector $\boldsymbol{j}(i)$ can be expressed as

$$\begin{aligned}\boldsymbol{R}_{jj} &= \frac{P_J}{N_c}\sum_{n=0}^{N_c-1} e^{j2\pi(f_J+n\cdot\Delta f)\cdot\boldsymbol{K}T_c} \odot \mathbb{E}\left\{\boldsymbol{x}_n(i)\boldsymbol{x}_n^H(i)\right\} \\ \check{\boldsymbol{R}}_{jj}(i) &= \frac{P_J}{N_c}\sum_{n=0}^{N_c-1} e^{j[2\pi(f_J+n\cdot\Delta f)\cdot\check{\boldsymbol{K}}(i)T_c+2\theta]} \\ &\quad \odot \mathbb{E}\left\{\boldsymbol{x}_n(i)\boldsymbol{x}_n^T(i)\right\} \end{aligned} \tag{16}$$

where

$$\boldsymbol{K} = \begin{bmatrix} 0 & 1 & \cdots & M-1 \\ -1 & 0 & \cdots & M-2 \\ \vdots & \vdots & \ddots & \vdots \\ -(M-1) & -(M-2) & \cdots & 0 \end{bmatrix}$$

and

$$\check{\boldsymbol{K}}(i) = \begin{bmatrix} 2iN & 2iN+1 & \cdots & 2iN+M-1 \\ 2iN+1 & 2(iN+1) & \cdots & 2iN+M \\ \vdots & \vdots & \ddots & \vdots \\ 2iN+M-1 & 2iN+M & \cdots & 2(iN+M-1) \end{bmatrix}.$$

If $\frac{T_j}{T_c}$ is an integer, $\boldsymbol{R}_{jj}$ does not vary with respect to the time index i. The matrix $\check{\boldsymbol{R}}_{jj}$ is time-varying with respect to i. In our case where BPSK modulated signals are considered, the improperness of $\boldsymbol{r}$ arises from signals of all users, the ISI, and the NBI. Since $\check{\boldsymbol{R}}$ is nonzero, the WL processing is able to take full advantage of this improper nature.

It is shown in [5] and [9] that when the data to be estimated are real, the WL Wiener filter weight vector $\boldsymbol{w}_a$ follows the transformation defined in (14) such that $\boldsymbol{w}_a = [\check{\boldsymbol{w}}^T, \quad \check{\boldsymbol{w}}^H]^T/\sqrt{2}$, where $\check{\boldsymbol{w}} \in \mathbb{C}^{M\times 1}$. Therefore, for the real estimated data, a key property of the WL filtering is conjugate symmetry defined by

$$\boldsymbol{w}_a^H\boldsymbol{r}_a(i) = \boldsymbol{r}_a^T(i)\boldsymbol{w}_a^* = \Re\left\{\check{\boldsymbol{w}}^H\boldsymbol{r}(i)\right\}. \tag{17}$$

In contrast to the conventional linear filter whose estimate is generally complex, the WL procedure exploits the statistics of both the covariance matrix and the pseudo-covariance matrix, yielding a real estimate with a smaller error [5], [9].

B. Widely Linear Reduced-Rank Filter

In the WL case, the augmented vector with twice the size of the received signal has to be considered. This requires a large number of symbols to reach the steady-state performance and imposes an even higher complexity on the receiver. To this end, the reduced-rank signal processing techniques can be combined with the WL filter to achieve a fast convergence, increased robustness to interference, and a lower complexity.

The principle of the proposed WL reduced-rank receiver is shown in Fig. 1, where the reduced-rank signal processing and the adaptive receiver design follow after the bijective transformation $\mathcal{T}$. The augmented received signal $\boldsymbol{r}_a$ of dimension $2M$ is then transformed by a rank-reduction matrix $\boldsymbol{S}_{a,D} \in \mathbb{C}^{2M\times D}$ onto a D-dimensional subspace, yielding a reduced-rank vector $\bar{\boldsymbol{r}}_a(i) = \boldsymbol{S}_{a,D}^H\boldsymbol{r}_a(i) \in \mathbb{C}^D$. The WL reduced-rank Wiener solution is written as $\bar{\boldsymbol{w}}_{a,o} = \bar{\boldsymbol{R}}_a^{-1}\bar{\boldsymbol{p}}_a$. Using augmented notations,

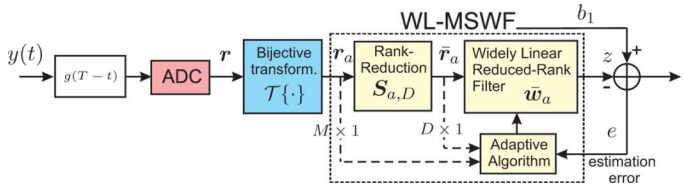


Fig. 1. Block diagram of the WL reduced-rank receiver in the complex baseband.

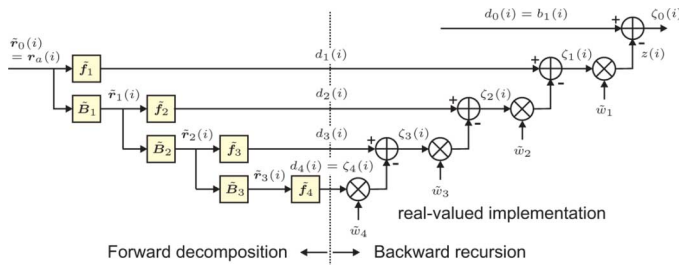


Fig. 2. Structure of 4-stage WL-MSWF.

the resulting MMSE $\bar{J}_{a,\min}$ and the SINR $\bar{\mathrm{SINR}}_a$ can also be represented in the same fashion as (11) and (12), respectively.

It is worth mentioning that if the received signal is circular, the WL solutions become equivalent to the linear case. Therefore, the proposed WL reduced-rank receiver, which additionally requires a bijective transformation before the filtering implementation, can be regarded as a generalized framework.

C. The WL-MSWF Strategies

1) Total-WL Construction (TWL): One way to construct the rank-reduction matrix $\boldsymbol{S}_{a,D}$ is to extend the L-MSWF to the widely linear case. Fig. 2 represents the four-stage MSWF, which consists of several nested filters $\tilde{\boldsymbol{f}}_1, \ldots, \tilde{\boldsymbol{f}}_D \in \mathbb{C}^{2M\times 1}$ and a combining procedure via the weighting coefficients $\tilde{w}_1, \ldots, \tilde{w}_D$. The "observation" data $\tilde{\boldsymbol{r}}_{n-1}(i)$ is successively decomposed by the filters $\tilde{\boldsymbol{f}}_n$ into one direction of the cross-correlation vector and the other subspace orthogonal to this direction by a blocking matrix $\tilde{\boldsymbol{B}}_n$. This matrix satisfies $\tilde{\boldsymbol{B}}_n^H\tilde{\boldsymbol{f}}_n = \boldsymbol{0}$ and can be chosen as the $2M \times 2M$-dimensional matrix $\tilde{\boldsymbol{B}}_n = \boldsymbol{I}_{2M} - \tilde{\boldsymbol{f}}_n\tilde{\boldsymbol{f}}_n^H$. In Fig. 2, $d_n(i)$ denotes the output of the filter $\tilde{\boldsymbol{f}}_n$ and $\tilde{\boldsymbol{r}}_n(i)$ is the output of $\tilde{\boldsymbol{B}}_n$. When $n = 0$, $d_0(i) = b_1(i)$ is the desired signal and $\tilde{\boldsymbol{r}}_0(i) = \boldsymbol{r}_a(i)$ is the augmented vector of the received signal. At the nth stage, the filter $\tilde{\boldsymbol{f}}_n$ is calculated according to the cross-correlation between the "desired" data $d_{n-1}(i)$ and the "observation" data vector $\tilde{\boldsymbol{r}}_{n-1}(i)$ from the previous stage

$$\tilde{\boldsymbol{f}}_n = \mathbb{E}\left\{d_{n-1}^*(i)\tilde{\boldsymbol{r}}_{n-1}(i)\right\},\ \|\tilde{\boldsymbol{f}}_n(i)\| = 1,\ n = 1, \ldots D. \tag{18}$$

Then the forward recursion can be continued by

$$d_n(i) = \tilde{\boldsymbol{f}}_n^H\tilde{\boldsymbol{r}}_{n-1}(i),\ n = 1, \ldots, D, \tag{19}$$

$$\tilde{\boldsymbol{r}}_n(i) = \tilde{\boldsymbol{B}}_n^H\tilde{\boldsymbol{r}}_{n-1}(i),\ n = 1, \ldots, D-1. \tag{20}$$

In the combining phase, the weighting coefficients are designed based on the MMSE criterion, i.e., $\tilde{w}_n$ is chosen so that

$\mathbb{E}\left\{|\zeta_{n-1}(i)|^2\right\}$ is minimized. For $n = D, \ldots, 1$, the backward recursion is completed by

$$\tilde{w}_n = \frac{\mathbb{E}\left\{d^*_{n-1}(i)\zeta_n(i)\right\}}{\mathbb{E}\left\{|\zeta_n(i)|^2\right\}} \tag{21}$$

$$\zeta_{n-1}(i) = d_{n-1}(i) - \tilde{w}^*_n \zeta_n(i). \tag{22}$$

Note that when $n = D$, $\zeta_D(i) = d_D(i)$ and when $n = 1$, $\tilde{w}^*_1\zeta_1(i)$ is the estimate for $d_0(i)$.

Similarly to [21], the rank-reduction matrix $\boldsymbol{S}_{a,D}$ defines the D-dimensional subspace spanned by $\tilde{\boldsymbol{f}}_n$ and can be constructed by the Krylov subspace, i.e.,

$$\boldsymbol{S}_{a,D} = \left[\tilde{\boldsymbol{f}}_1, \tilde{\boldsymbol{f}}_2, \ldots, \tilde{\boldsymbol{f}}_D\right] \tag{23}$$

$$= \left[\boldsymbol{p}_a, \boldsymbol{R}_a\boldsymbol{p}_a, \ldots, \boldsymbol{R}_a^{D-1}\boldsymbol{p}_a\right]. \tag{24}$$

The TWL construction of the rank-reduction matrix fully utilizes the second-order statistics of the observation signal. This scheme is denoted as TWL-MSWF.

2) Quasi-WL Construction (QWL): A simpler way to construct the rank-reduction matrix is based on adopting a transformation $\mathcal{T}$ on $\boldsymbol{S}_D$ using the L-MSWF

$$\boldsymbol{S}_{a,D} = \frac{1}{\sqrt{2}}\left[\boldsymbol{S}_D^T, \boldsymbol{S}_D^H\right]^T = \mathcal{T}\left\{\boldsymbol{S}_D\right\} \tag{25}$$

where $\boldsymbol{S}_D$ represents the Krylov subspace as shown in (13). The reduced-rank vector is thus calculated by $\bar{\boldsymbol{r}}_a(i) = \Re\left\{\boldsymbol{S}_D^H\boldsymbol{r}(i)\right\} = \Re\left\{\bar{\boldsymbol{r}}(i)\right\}$, i.e., by taking the real part of the reduced-rank vector from the L-MSWF algorithm. With the QWL design, the general block diagram shown in Fig. 1 can be simplified to an equivalent model depicted in Fig. 3, where the block "Widely Linear Reduced-Rank Filter" is still preserved. Compared to the TWL method, the only difference lies in how to construct the rank-reduction matrix $\boldsymbol{S}_{a,D}$. Both constructions (24) and (25) can be generalized in the form of $\boldsymbol{S}_{a,D} = \mathcal{T}\left\{\check{\boldsymbol{S}}_D\right\} = \mathcal{T}\left\{\boldsymbol{S}_D + \Delta\boldsymbol{S}_D\right\}$, where $\Delta\boldsymbol{S}_D$ contains the difference between the linear and the widely linear designs, i.e., the first M rows of $\boldsymbol{S}_{a,D} - \mathcal{T}\left\{\boldsymbol{S}_D\right\}$. If $\Delta\boldsymbol{S}_D = \boldsymbol{0}$, we have a QWL construction, which does not exploit the second-order information contained in the pseudo-covariance matrix $\check{\boldsymbol{R}}$. However, the succeeding filter design still takes advantage of the improper signals, providing a better performance than the L-MSWF. The associated filtering method is named QWL-MSWF. When $D = 1$, i.e., $\boldsymbol{S}_{a,D} = \boldsymbol{p}_a = \mathcal{T}\left\{\boldsymbol{p}\right\} = \mathcal{T}\left\{\boldsymbol{S}_D\right\}$, the QWL-MSWF and the TWL-MSWF methods have the same performance. We will show in Sections IV-E and VIII that in most cases for improper signals, the TWL-MSWF outperforms the QWL-MSWF.

D. Comparison With the PC Methods

One of the few WL reduced-rank filters has been proposed in [4] using the PC technique. It is based on the eigen-decomposition of the augmented covariance matrix $\boldsymbol{R}_a = \boldsymbol{V}\Sigma\boldsymbol{V}^H$, where the columns of $\boldsymbol{V} \in \mathbb{C}^{2M\times 2M}$ are the eigenvectors of $\boldsymbol{R}_a$ and $\boldsymbol{\Sigma}$ is a diagonal matrix with the ordered eigenvalues σ_k on its diagonal such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{2M}$. The rank-reduction matrix obtained via PC is $\boldsymbol{S}_{a,D} = \boldsymbol{V}(:, 1:D)$, which contains the first D columns of $\boldsymbol{V}$, corresponding to the D largest eigenvalues with a descending order. A modified PC method introduced in [21] improves the performance. It chooses the eigenvectors associated with the D largest values of $\frac{\left|\boldsymbol{v}_k^H\boldsymbol{p}_a\right|^2}{\sigma_k}$, where $\boldsymbol{v}_k$ is the kth column of $\boldsymbol{V}$. This method selects a set of D eigenvectors to form the rank-reduction matrix that minimizes the MSE.

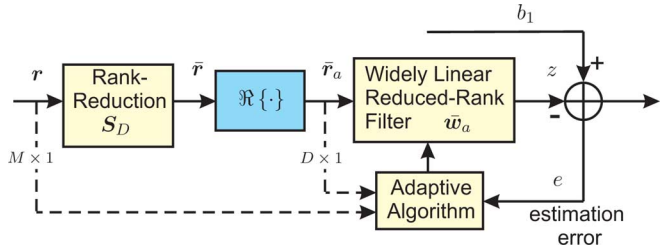


Fig. 3. Receiver structure of QWL-MSWF.

Compared to the proposed TWL/QWL-MSWF, there are some disadvantages of the above WL-PC techniques:

1) a larger rank D is required than that for the MSWF;
2) these methods rely on the eigen-decomposition, which is much more computationally expensive;
3) the WL-PC requires a larger D to achieve a better performance than the linear PC [4].

E. MMSE and SINR Analysis of the WL-MSWF

Let us first consider the L-MSWF described in Section III. The eigenvalue decomposition of the reduced-rank covariance matrix can be obtained by $\bar{\boldsymbol{R}} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^H$, where $\boldsymbol{Q}$ contains the eigenvectors $\boldsymbol{q}_k, k = 1, \ldots, D$ and $\boldsymbol{\Lambda}$ is a diagonal matrix consisting of eigenvalues λ_k in a descending order. Applying (13) to (11), the MMSE of the L-MSWF can thus be expressed as

$$\begin{aligned}\bar{J}_{\min} &= 1 - \boldsymbol{p}^H\boldsymbol{S}_D\boldsymbol{Q}\boldsymbol{\Lambda}^{-1}\boldsymbol{Q}^H\boldsymbol{S}_D^H\boldsymbol{p} \\ &= 1 - \sum_{k=1}^{D}\frac{\left|\boldsymbol{q}_k^H(\boldsymbol{S}_D^H\boldsymbol{p})\right|^2}{\lambda_k}\end{aligned} \tag{26}$$

where it can be easily proven that $\boldsymbol{S}_D^H\boldsymbol{p}$ is real-valued.

Similarly to the linear case, the eigenvalue decomposition of the reduced-rank augmented covariance matrix is computed by $\bar{\boldsymbol{R}}_a = \boldsymbol{Q}_a\boldsymbol{\Lambda}_a\boldsymbol{Q}_a^H$, where the columns of $\boldsymbol{Q}_a$ are the eigenvectors $\boldsymbol{q}_{ak}, k = 1, \ldots, D$ and $\boldsymbol{\Lambda}_a$ contains the eigenvalues λ_{ak} in a descending order on its diagonal. With $\boldsymbol{p}_a = \mathcal{T}\left\{\boldsymbol{p}\right\}$ and $\boldsymbol{S}_{a,D} = \mathcal{T}\left\{\check{\boldsymbol{S}}_D\right\}$, the resulting MMSE of the WL-MSWF can be written by

$$\begin{aligned}\bar{J}_{a,\min} &= 1 - \boldsymbol{p}_a^H\boldsymbol{S}_{a,D}\boldsymbol{Q}_a\boldsymbol{\Lambda}_a^{-1}\boldsymbol{Q}_a^H\boldsymbol{S}_{a,D}^H\boldsymbol{p}_a \\ &= 1 - \sum_{k=1}^{D}\frac{\left|\boldsymbol{q}_{ak}^H(\boldsymbol{S}_{a,D}^H\boldsymbol{p}_a)\right|^2}{\lambda_{ak}} \\ &= 1 - \sum_{k=1}^{D}\frac{\left|\boldsymbol{q}_{ak}^H\left(\check{\boldsymbol{S}}_D^H\boldsymbol{p} + \check{\boldsymbol{S}}_D^T\boldsymbol{p}^*\right)/2\right|^2}{\lambda_{ak}} \\ &= 1 - \sum_{k=1}^{D}\frac{\left|\boldsymbol{q}_{ak}^H\Re\left\{\check{\boldsymbol{S}}_D^H\boldsymbol{p}\right\}\right|^2}{\lambda_{ak}}.\end{aligned} \tag{27}$$

The MMSE is mainly determined by the eigenvalues of $\bar{\boldsymbol{R}}_a$. In Appendix A, we show that $\lambda_{ak} < \lambda_k, k = 1, 2, \ldots, D, D \ll K$

with K being the number of eigenvectors of $\boldsymbol{R}$ (or $\boldsymbol{R}_a$) that correspond to the signal subspace. This applies to both the TWL and the QWL constructions. If the QWL is applied, $\check{\boldsymbol{S}}_D = \boldsymbol{S}_D$ holds and thus (27) is simplified to $\bar{J}_{a,\min} = 1 - \sum_{k=1}^{D} \frac{\left|\boldsymbol{q}_{ak}^H(\boldsymbol{S}_D^H\boldsymbol{p})\right|^2}{\lambda_{a,k}}$. When the TWL is used, more information can be explored, yielding a smaller MMSE than the QWL. Therefore, a comparison of (26) and (27) indicates that even with the same filter length D, the MMSE of the WL-MSWF estimate with both constructions is smaller than that of the L-MSWF, i.e., $\bar{J}_{a,\min} < \bar{J}_{\min}$. Since the SINR has a simple relationship with the MMSE as shown in (12), $\bar{\text{SINR}}_a > \bar{\text{SINR}}$ holds. This will be verified in Section VII-A.

F. Properties of the WL-MSWF

With the real-valued data being estimated, the WL-MSWF has the following key properties:

1) It has been shown in [35] that after the multistage decomposition, the reduced-rank vector $\bar{\boldsymbol{r}}_a(i)$, the filter weight vector $\bar{\boldsymbol{w}}_a(i)$, the decision variable $z(i)$, and the estimation error $e(i)$ are all real-valued.
2) With increasing D, the MMSE and the output SINR of the WL-MSWF converge to the solutions of the WL full-rank Wiener filter.
3) In contrast to the eigen-decomposition methods, the WL-MSWF inherently extracts key characteristics of the processed data and the rank D required to achieve the full-rank performance is much smaller.
4) With the same rank D, the WL-MSWF outperforms the L-MSWF in terms of the MMSE and the maximum SINR.
5) The rank D required to approach the full-rank performance is only slightly affected by the system load such as the number of users N_u, the NBI, as well as the processing gain N and the number of channel taps L, which determine the ISI impact.
6) Compared to the full-rank filters, the complexity is significantly reduced by using the reduced-rank techniques [21], [24]. On one hand, due to the processing on the augmented received vector, the WL forward decomposition has a higher complexity compared to the linear case. On the other hand, it has been shown that the combining phase of the WL-MSWF is carried out on the real-valued data, which alleviates the computational efforts. It is worth mentioning that the QWL-MSWF design simply deals with the real part of the reduced-rank vector from the L-MSWF algorithm. Consequently, it has an even lower complexity than the L-MSWF. The complete computational complexity analysis will be addressed in Section VI.

V. Adaptive Algorithms and Convergence Analysis

In this section we develop two training-based adaptive algorithms, the SG and the RLS, for the proposed WL-MSWF techniques. The convergence performance of the WL adaptive schemes based on the SG has been discussed in [1], [10]. However, it is of prime interest to evaluate the convergence behavior of the adaptive reduced-rank algorithms. In this section, we focus on the convergence analysis of both the SG and the RLS versions of the WL-MSWF as well as the comparison with their linear counterparts.

TABLE I
SG Adaptive Algorithm for WL-MSWF[3]

Initialize the algorithm by setting:
$\boldsymbol{p}_a(0) = \boldsymbol{0}, \boldsymbol{R}_a(0) = \delta\boldsymbol{I}, \bar{\boldsymbol{w}}_a(0) = \boldsymbol{0}$
Choose the rank D and the step size μ
For the time index $i = 1, 2, \cdots, N_s$
The rank-reduction matrix is estimated by TWL or QWL
The reduced-rank vector $\bar{\boldsymbol{r}}_a(i) = \boldsymbol{S}_{a,D}^H(i)\boldsymbol{r}_a(i)$
The estimate of $b_1(i)$ is $z(i) = \bar{\boldsymbol{w}}_a^H(i)\bar{\boldsymbol{r}}_a(i)$
The estimation error $e(i) = b_1(i) - z(i)$
Update WL-MSWF $\bar{\boldsymbol{w}}_a(i+1) = \bar{\boldsymbol{w}}_a(i) + \mu e^*(i)\bar{\boldsymbol{r}}_a(i)$
end

A. SG and RLS Adaptive Algorithms for the WL-MSWF

The rank-reduction matrix $\boldsymbol{S}_{a,D}$ for the TWL is constructed based on estimating the augmented covariance matrix $\boldsymbol{R}_a$ and the augmented cross-correlation vector $\boldsymbol{p}_a$ by

$$\boldsymbol{R}_a(i) = \lambda\boldsymbol{R}_a(i-1) + \boldsymbol{r}_a(i)\boldsymbol{r}_a^H(i) \tag{28}$$

$$\boldsymbol{p}_a(i) = \lambda\boldsymbol{p}_a(i-1) + b_1^*(i)\boldsymbol{r}_a(i) \tag{29}$$

where $0 < \lambda < 1$ is the forgetting factor and $b_1(i)$ is the ith training symbol. Using (24), the rank-reduction matrix at time instant i can thus be calculated by

$$\boldsymbol{S}_{a,D}(i) = \left[\boldsymbol{p}_a(i), \boldsymbol{R}_a(i)\boldsymbol{p}_a(i), \ldots, \boldsymbol{R}_a^{D-1}(i)\boldsymbol{p}_a(i)\right]. \tag{30}$$

The QWL construction $\boldsymbol{S}_{a,D}$ is obtained by (25), where $\boldsymbol{R}(i)$ and $\boldsymbol{p}(i)$ are recursively estimated. Tables I and II show the related SG and RLS algorithms for the WL-MSWF, where δ and $\tilde{\delta}$ are initialization scalars to ensure the numerical stability. In Table II, the reduced-rank augmented covariance matrix is given by $\bar{\boldsymbol{R}}_a(i) = \boldsymbol{S}_{a,D}^H(i)\boldsymbol{R}_a(i)\boldsymbol{S}_{a,D}(i)$ and the RLS scheme estimates its inverse $\bar{\boldsymbol{R}}_a^{-1}(i)$.

B. Convergence Analysis of the WL-MSWF With SG

1) Step Size: As discussed in [34], to ensure the convergence, the step size should be chosen such that

$$0 < \mu < \frac{2}{\max\{\lambda_{ak}\}}, k = 1, \ldots, D. \tag{31}$$

Similarly, the step size of the L-MSWF-SG approach satisfies $0 < \mu < \frac{2}{\max\{\lambda_k\}}$, $k = 1, \ldots, D$. Since for $k = 1, \ldots, D, D < K, \lambda_{ak} < \lambda_k$ is observed, indicating that the step size of the WL-MSWF-SG algorithm can be larger than the L-MSWF-SG.

2) The Mean Square Error Learning Curve: The MSE of the WL-MSWF-SG algorithm at time i can be expressed as [34]

$$\bar{J}_a(i) = \mathbb{E}\left\{|e(i)|^2\right\}. \tag{32}$$

[3]We use this "complex conjugate" to have a general expression, since for linear filtering methods, the estimate z might be complex-valued. The real-valued estimate is observed as one special property of the WL algorithms, when the data to be estimated is real (e.g., BPSK).

TABLE II
RLS ADAPTIVE ALGORITHM FOR WL-MSWF

Initialize the algorithm by setting:
$\boldsymbol{p}_a(0)=\boldsymbol{0}, \boldsymbol{R}_a(0)=\delta\boldsymbol{I}, \bar{\boldsymbol{p}}_a(0)=\boldsymbol{0}, \bar{\boldsymbol{R}}_a^{-1}(0)=\tilde{\delta}^{-1}\boldsymbol{I}, \bar{\boldsymbol{w}}_a(0)=\boldsymbol{0}$

Choose the rank D,
For the time index $i=1,2,\cdots,N_s$
The rank-reduction matrix is estimated by TWL or QWL
The reduced-rank vector $\bar{\boldsymbol{r}}_a(i)=\boldsymbol{S}_{a,D}^H(i)\boldsymbol{r}_a(i)$
The estimate of $b_1(i)$ is $z(i)=\bar{\boldsymbol{w}}_a^H(i)\bar{\boldsymbol{r}}_a(i)$
The recursive calculation:
$\boldsymbol{k}(i)=\bar{\boldsymbol{R}}_a^{-1}(i-1)\bar{\boldsymbol{r}}_a(i)$
$\boldsymbol{g}(i)=\dfrac{\lambda^{-1}\boldsymbol{k}(i)}{1+\lambda^{-1}\bar{\boldsymbol{r}}_a^H(i)\boldsymbol{k}(i)}$
$\bar{\boldsymbol{R}}_a^{-1}(i)=\lambda^{-1}\bar{\boldsymbol{R}}_a^{-1}(i-1)-\lambda^{-1}\boldsymbol{g}(i)\bar{\boldsymbol{r}}_a^H(i)\bar{\boldsymbol{R}}_a^{-1}(i-1)$
$\bar{\boldsymbol{p}}_a(i)=\lambda\bar{\boldsymbol{p}}_a(i-1)+b_1^*(i)\bar{\boldsymbol{r}}_a(i)$
Update WL-MSWF $\bar{\boldsymbol{w}}_a(i)=\bar{\boldsymbol{R}}_a^{-1}(i)\bar{\boldsymbol{p}}_a(i)$
end

Applying the eigen-decomposition of $\bar{\boldsymbol{R}}_a$, when the steady state is achieved, i.e., $i\to\infty$, we get

$$\begin{aligned}\bar{J}_a(\infty)&=\bar{J}_{a,\min}+\mu\bar{J}_{a,\min}\sum_{k=1}^{D}\frac{\lambda_{ak}}{2-\mu\lambda_{ak}}\\&\approx\bar{J}_{a,\min}+\frac{\mu\bar{J}_{a,\min}}{2}\sum_{k=1}^{D}\lambda_{ak},\ \ \mu\ \text{small}\end{aligned}\tag{33}$$

where $\bar{J}_{a,\min}$ is calculated by (11). The excess MSE $\bar{J}_{a,\mathrm{ex}}(i)$ can be represented as

$$\bar{J}_{a,\mathrm{ex}}(i)=\bar{J}_a(i)-\bar{J}_{a,\min}\tag{34}$$

meaning that $\bar{J}_{a,\mathrm{ex}(\infty)}\approx\frac{\mu\bar{J}_{a,\min}}{2}\sum_{k=1}^{D}\lambda_{ak}$. Considering that $\bar{J}_{a,\min}<\bar{J}_{\min},\lambda_{ak}<\lambda_k,k=1,\ldots,D,D<K$ shown in Section IV-E, we can conclude that the steady-state MSE and excess MSE of the WL-MSWF-SG method are both smaller than that of the linear case, i.e., $\bar{J}_a(\infty)<\bar{J}(\infty)$ and $\bar{J}_{a,\mathrm{ex}}(\infty)<\bar{J}_{\mathrm{ex}}(\infty)$.

The transient behavior of the MSE is mainly determined by the excess MSE, consisting of the transient excess MSE $\bar{J}_{a,\mathrm{extrans}}(i)$ and the steady-state excess MSE [34] as

$$\bar{J}_{a,\mathrm{ex}}(i)=\bar{J}_{a,\mathrm{extrans}}(i)+\bar{J}_{a,\mathrm{ex}}(\infty).\tag{35}$$

It will be shown via experiments that the WL-MSWF-SG algorithm has a smaller transient excess MSE than the linear method, showing a superior convergence performance for the WL case even with the same rank D.

C. Convergence Analysis of the WL-MSWF With RLS

1) Weight Error Correlation Matrix: To analyze the RLS implementation of the WL-MSWF receiver shown in Table II, we assume the forgetting factor $\lambda=1$ and obtain the weight error as follows [34]

$$\boldsymbol{\epsilon}_a(i)=\bar{\boldsymbol{w}}_a(i)-\bar{\boldsymbol{w}}_{a,o}=\bar{\boldsymbol{R}}_a^{-1}(i)\sum_{n=1}^{i}\bar{\boldsymbol{r}}_a(n)e_o^*(n)\tag{36}$$

where $e_o(i)=b_1(i)-\bar{\boldsymbol{w}}_{a,o}^H\bar{\boldsymbol{r}}_a(i)$ is the estimation error produced by the optimal solution $\bar{\boldsymbol{w}}_{a,o}$. We assumed $e_o(n)$ to be white with zero-mean and variance σ_e^2, where $\mathbb{E}\{e_o(m)e_o^*(n)\}=\begin{cases}\sigma_e^2=\bar{J}_{a,\min}, & m=n\\ 0, & m\neq n\end{cases}$. The weight error correlation matrix can then be expressed as

$$\begin{aligned}\boldsymbol{K}_a(i)&=\mathbb{E}\left\{\boldsymbol{\epsilon}_a(i)\boldsymbol{\epsilon}_a^H(i)\right\}\\&=\bar{J}_{a,\min}\mathbb{E}\left\{\bar{\boldsymbol{R}}_a^{-1}(i)\sum_{n=1}^{i}\sum_{m=1}^{i}\bar{\boldsymbol{r}}_a(m)\bar{\boldsymbol{r}}_a^H(n)\bar{\boldsymbol{R}}_a^{-1}(i)\right\}\\&=\bar{J}_{a,\min}\mathbb{E}\left\{\bar{\boldsymbol{R}}_a^{-1}(i)\right\}\\&=\frac{\bar{J}_{a,\min}}{i-D-1}\bar{\boldsymbol{R}}_a^{-1},\quad i>D+1.\end{aligned}\tag{37, 38}$$

2) The Learning Curve of A Priori Estimation Error: In RLS algorithms, the a priori estimation error defined by $\xi(i)=b_1(i)-\bar{\boldsymbol{w}}_a^H(i-1)\bar{\boldsymbol{r}}_a(i)$ is chosen to characterize the learning curve [34]. By eliminating $b_1(i)$ based on the expression of $e_o(i)$, we can represent $\xi(i)$ in terms of the weight error $\boldsymbol{\epsilon}_a(i-1)$ as

$$\xi(i)=e_o(i)-\boldsymbol{\epsilon}_a^H(i-1)\bar{\boldsymbol{r}}_a(i).\tag{39}$$

The resulting learning curve is expressed as

$$\begin{aligned}\bar{J}'_a(i)&=\mathbb{E}\left\{|\xi(i)|^2\right\}=\bar{J}_{a,\min}+\mathrm{tr}\left\{\bar{\boldsymbol{R}}_a\boldsymbol{K}_a(i-1)\right\}\\&=\bar{J}_{a,\min}+\frac{D}{i-D-1}\bar{J}_{a,\min},\ \ i>D+1.\end{aligned}\tag{40}$$

Compared to SG in (34) and (35), the learning curve of RLS indicates that the excess MSE $\bar{J}'_{a,\mathrm{ex}}(i)=\frac{D}{i-D-1}\bar{J}_{a,\min}$ vanishes as $i\to\infty$ and does not depend on the eigenvalue spread of $\bar{\boldsymbol{R}}_a$. In the steady state, a zero excess MSE can be reached by the RLS algorithm, exhibiting a faster convergence and a higher robustness than the SG method. Since $\bar{J}_{a,\min}<\bar{J}_{\min}$, the transient excess MSE of the WL-MSWF-RLS approach is smaller than those of the linear counterparts even with the same rank D, i.e., $\bar{J}'_{a,\mathrm{ex}}(i)<\bar{J}'_{\mathrm{ex}}(i)$.

VI. COMPLEXITY ANALYSIS

The computational complexity of the adaptive algorithms is estimated according to the number of real additions and real multiplications per iteration for each received symbol of size M. The estimated computational complexity of the proposed WL-MSWF schemes is summarized in Table III, where we consider the existing algorithms for comparison. Fig. 4 illustrates the total number of real operations (additions and multiplications) per iteration per symbol for each algorithm as a function of M, where the rank of the MSWF $D=4$ is chosen. For all the algorithms, the SG always has a lower complexity than RLS. In the full-rank case, the WL-SG is slightly simpler than the L-SG due to the conjugate symmetric property of the WL approaches, while the multiplication of bigger matrices results in a higher complexity of the WL-RLS than that of the L-RLS. In the MSWF, the construction of the rank-reduction matrix that requires a higher-order matrix multiplication imposes more computational efforts than the full-rank case. A larger D will considerably increase the computational costs. We can observe that the proposed TWL-MSWF SG/RLS methods

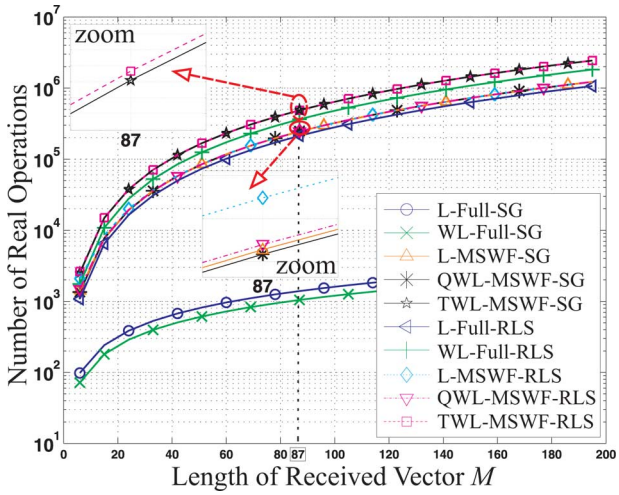


Fig. 4. Computational complexity in terms of real additions and multiplications per iteration per symbol as a function of M. For MSWF schemes, $D = 4$ is chosen. The zoomed-in curves are also shown at $M = 87$.

TABLE III
ESTIMATED COMPUTATIONAL COMPLEXITY ACCORDING TO THE NUMBER OF REAL OPERATIONS

Algorithms	Additions	Multiplications
L-Full-SG	$8M$	$8M+2$
WL-Full-SG	$6M-1$	$6M+1$
L-MSWF-SG	$4(D-1)M^2$	$4(D-1)M^2$
	$+2(D+1)M+6D$	$+4D(M+2)+2$
QWL-MSWF-SG	$4(D-1)M^2+$	$4(D-1)M^2+$
	$2(D+1)M$	$2D(2M+1)+1$
TWL-MSWF-SG	$8(D-1)M^2+$	$8(D-1)M^2+$
	$2(D+1)M$	$2D(2M+1)+1$
L-Full-RLS	$12M^2+2M-1$	$16M^2+10M+2$
WL-Full-RLS	$20M^2+2M-1$	$28M^2+10M+1$
L-MSWF-RLS	$4(D-1)M^2+$	$4(D-1)M^2+$
	$2M(D+1)+12D^2-1$	$2D(2M+5)+16D^2+2$
QWL-MSWF-RLS	$4(D-1)M^2+$	$4(D-1)M^2+$
	$2M(D+1)+(D-1)^2+1$	$4DM+(2D+1)^2$
TWL-MSWF-RLS	$8(D-1)M^2+$	$8(D-1)M^2+$
	$2M(D+1)+(D-1)^2+1$	$4DM+(2D+1)^2$

exhibit the highest complexity. It is worth emphasizing that the proposed QWL-MSWF SG/RLS algorithms are slightly less complex than the L-MSWF counterparts and significantly reduce the complexity compared to the full-rank WL-RLS.

VII. SIMULATION RESULTS

In this section, we evaluate the steady-state, the transient, and the convergence performance of the proposed TWL/QWL-MSWF schemes and compare them with the linear MSWF, the linear/WL full-rank Wiener filters, as well as the linear/WL PC-based reduced-rank methods. The rank-dependent performance along with the adaptive rank selection algorithms are presented. We further analyze the SINR performance of the proposed methods in the case when the desired signal is strictly circular (e.g., QPSK-modulated signal) but the interference (MUI or NBI) is noncircular.

TABLE IV
PARAMETERS FOR IEEE 802.11A OFDM SIGNAL

modulation	f_{OFDM}	N_c	Δf	T_J
BPSK	5.22 GHz	48	312.5 KHz	4 μs

For the multipath propagation channel, we use UWB channels measured in a line-of-sight office of size 5 m × 5 m × 2.6 m. The measurements (including antennas) were carried out by the IMST GmbH [36]. The transfer function of a certain channel realization is firstly transformed from the band-pass to the low-pass range at a center frequency $f_c = 4$ GHz, and afterwards converted into a tapped-delay line model with equally spaced taps. The RRC pulse is chosen with $B_3 = 500$ MHz and $\beta = 0.3$. At the receiver, the sampling rate of the ADC is 1 GHz and thus the channel resolution is 1 ns. The maximum channel delay is 64 ns. We assume that the UWB channel is time-invariant block fading during the estimation. The DS code of length $N = 24$ is generated pseudo-randomly for the DS-UWB system. The dimension of the received vector $\boldsymbol{r}$ is $M = 87$. The parameters of the OFDM interference used for the simulations are shown in Table IV, where the cyclic prefix and the guard interval are not considered for simplicity[4] and the OFDM symbol period T_J is larger than the symbol duration. We consider a scheme in which the proposed adaptive WL-MSWF algorithms are first trained by a pilot sequence of 400 symbols and are then switched to the decision-directed mode.

A. Achievable SINR and Transient Analysis

The simulation results are presented to validate the theoretical analysis in Sections V-B and V-C. We first compare the eigenvalues of the reduced-rank covariance matrix for both linear and WL cases ($\bar{\boldsymbol{R}}$ and $\bar{\boldsymbol{R}}_a$). Fig. 5(a) depicts the eigenvalues using linear, QWL, and TWL reduced-rank matrix constructions for $D = 2, 4, 6$, where the number of users $N_u = 16$, $E_b/N_0 = 15$ dB, and NBI is absent. It is observed that the eigenvalues of using both TWL and QWL constructions are smaller than the linear case, i.e., $\lambda_{ak} < \lambda_k, k = 1, \ldots, D$, meaning that a larger step size for WL-MSWF-SG algorithms can be chosen compared to the L-MSWF-SG [cf. Equation (31)]. When the NBI is present, the eigenvalues are shown in Fig. 5(b) with $D = 4$. With very low SIR, the TWL-MSWF method has larger eigenvalues ($k = 3, 4$) than the L-MSWF due to the "contribution" of the strong NBI. However, the dominant eigenvalues (i.e., $k = 1$, 2) of TWL-MSWF are no greater than L-MSWF at various SIR values. Fig. 5(c) plots the eigenvalues changing with the number of users, which shows the higher values of L-MSWF than those of the TWL/QWL-MSWF algorithms. The SINR values of different schemes as a function of $\frac{E_b}{N_0}$ (dB), SIR, and the number of users, are also illustrated in Fig. 6, where the rank $D = 4$ is chosen. It can be clearly seen that both the TWL-MSWF and QWL-MSWF outperform the L-MSWF in terms of the SINR and the TWL construction which utilizes more second-order information produces a higher SINR than the QWL case. The

[4]The overall spectrum does not change with the cyclic prefix or the guard interval. This implies that the performance of the algorithms will not be affected by adding the guard interval for the OFDM signal. Therefore, we ignore this for simplicity.

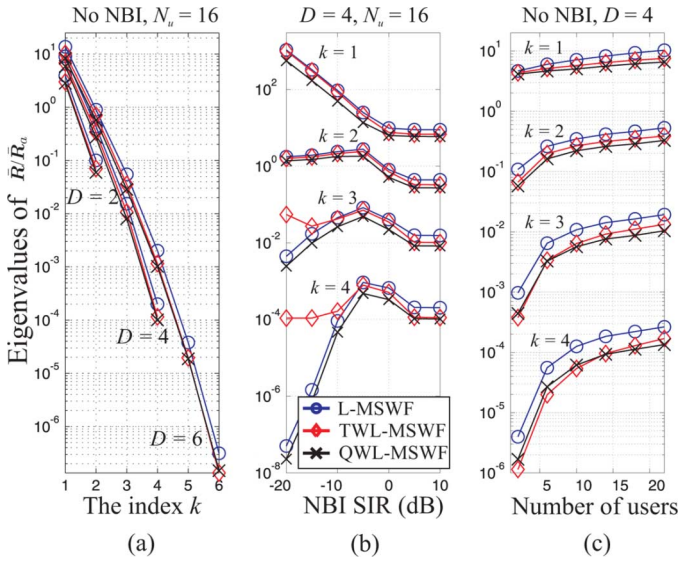


Fig. 5. Eigenvalues of the reduced-rank covariance matrix constructed by L/TWL/QWL-MSWF algorithms with $\frac{E_b}{N_0} = 15$ dB versus (a) the kth stage projection for different D, (b) various SIR in the presence of OFDM NBI, and (c) different number of users.

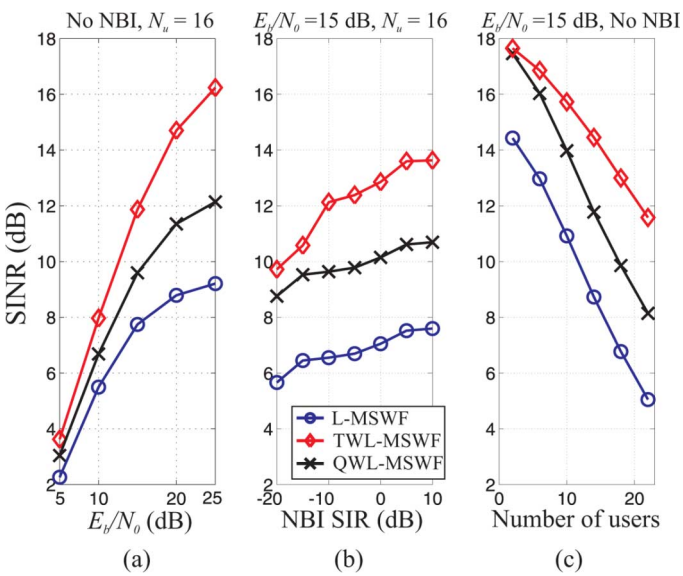


Fig. 6. SINR of L/TWL/QWL-MSWF algorithms versus (a) $\frac{E_b}{N_0}$ (dB), (b) various SIR in the presence of OFDM NBI, and (c) different number of users.

performance gain of the TWL over the QWL increases with the number of users, cf., Fig. 6(c). In summary, this shows that given a value of rank D, the proposed TWL-/QWL-MSWF schemes are more robust to interference and can accommodate more users compared to the L-MSWF.

We assess the SINR of the proposed TWL/QWL-MSWF algorithms as a function of the rank D and compare them to the PC-based reduced-rank filters shown in Fig. 7(a) and (b). The performance of the full-rank linear/WL schemes is shown only for the case, where NBI is present and $N_u = 16$. The conventional PC method that uses the first D eigenvectors of $\boldsymbol{V}$ corresponding to D largest eigenvalues of $\boldsymbol{R}$ or $\boldsymbol{R}_a$ in a descending order is denoted as "PC-conv." The modified PC scheme is called "PC-modi." As D increases, i.e., more signal information

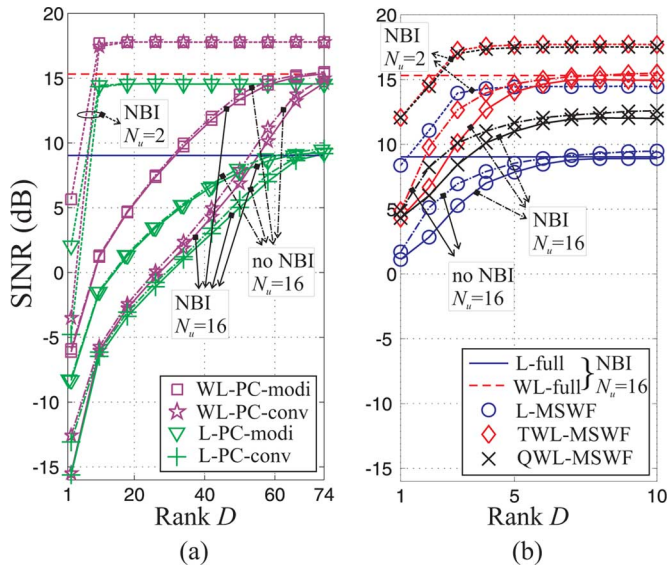


Fig. 7. SINR of the discussed algorithms versus the rank D for (a) the L/WL-PC algorithms and for (b) the L/TWL/QWL-MSWF algorithms. We consider $\frac{E_b}{N_0} = 15$ dB, $N_u = 2$ and 16, OFDM NBI of SIR $= -5$ dB.

is utilized, the SINR increases until it gets close to the full-rank state. The TWL-MSWF only requires the rank $D = 2$ to $D = 6$ to achieve the highest SINR and the selected D is only slightly affected by the number of users and the presence of NBI. For both the PC-conv and the PC-modi, the necessary D to approach the full-rank SINR does not depend on the presence of NBI but is quite sensitive to the number of users, e.g., to obtain the best performance, we need $D = 10$ for the 2-user case and $D > 60$ for $N_u = 16$. The QWL-MSWF cannot reach the WL full-rank SINR but it still outperforms the PC-based methods with a much smaller rank. For $N_u = 16$ with $D < 35$, the advantage of the WL-PC-conv scheme over the L-PC-conv is lost, unless a higher rank is chosen. With the same rank D, the WL-PC-modi method exhibits a higher SINR than the L-PC-modi, since the D eigenvectors are selected to minimize the MSE.

Fig. 8 shows the transient excess MSE of the training-based SG algorithms $\bar{J}_{a,\text{extrans}}(i)$ for the TWL/QWL-MSWF-SG schemes compared to the linear counterpart. It is assumed that the augmented covariance matrix is known and is computed by (15). We consider the step size $\mu = 0.02$ without NBI and $\mu = 0.024$ in the presence of NBI, $N_u = 16$, $\frac{E_b}{N_0} = 15$ dB, and $D = 4$. For each time instant, the excess MSE of the WL methods is smaller than that of the linear case and TWL exhibits a better transient performance than QWL.

B. BER Convergence Performance

We show the bit error rate (BER) performance of the adaptive TWL/QWL-MSWF algorithms and compare it to the existing methods in Fig. 9(a) for SG and in (b) for RLS. The rank $D = 4$ is chosen as a representative value to compare the performance of different schemes. It is obvious that all the RLS algorithms outperform the SG in the convergence and tracking performances. Even with the same D, the TWL-MSWF which fully exploits the second-order behavior of the noncircular signal performs the best. Since the QWL-MSWF constitutes the rank-reduction matrix from the linear estimates and utilizes the complementary covariance statistics only for the weight

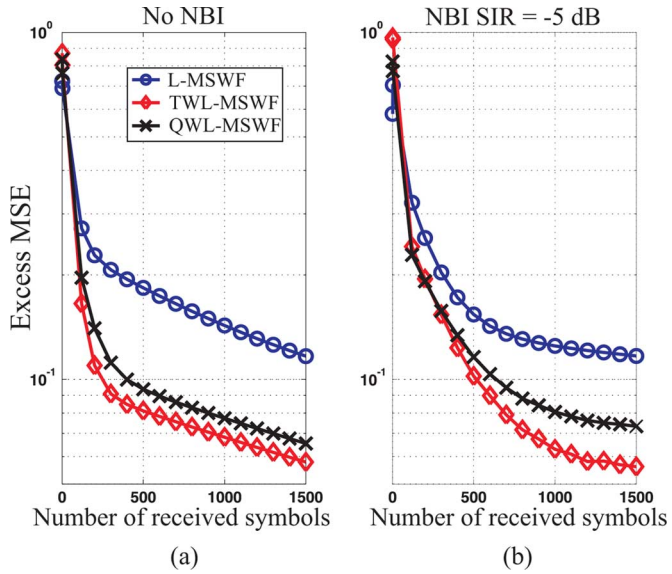


Fig. 8. Transient excess MSE of L/TWL/QWL-MSWF-SG algorithms in the cases when OFDM NBI is absent (a) and present (b). It is chosen that $\frac{E_b}{N_0} = 15$ dB, $D = 4$, and $N_u = 16$.

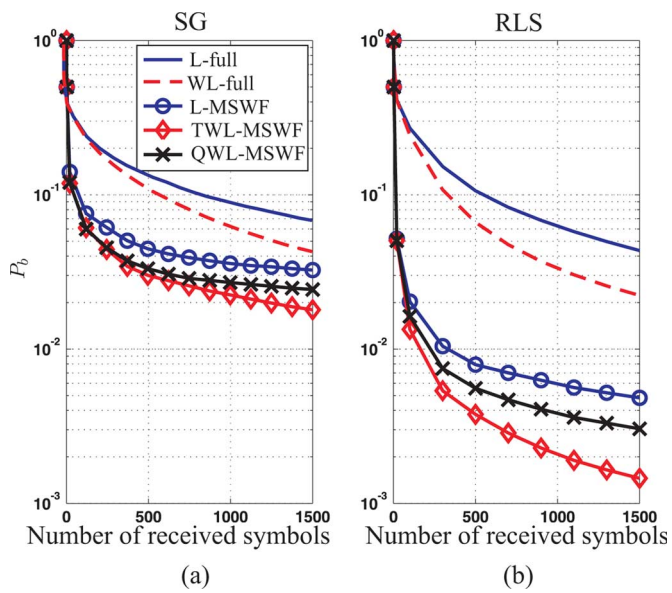


Fig. 9. BER convergence performance of (a) SG and (b) RLS algorithms for $\frac{E_b}{N_0} = 15$ dB, $N_u = 16$, and OFDM-NBI with SIR $= -5$ dB. We consider $D = 4$ for the MSWF techniques.

adaptation, it still exhibits a better convergence performance than the L-MSWF but has a lower complexity. The proposed TWL/QWL-MSWF algorithms show a better BER performance compared to the WL full-rank counterparts. The reason is that after the augmented received signal of a dimension $2M$ is projected onto a Krylov subspace with a much lower dimension D, the estimation of filter weights is only based on a small amount of parameters. This implies a faster convergence to the steady-state performance.

C. Rank-Dependent Performance

The number of parameters for estimating the filter weights, i.e., the rank D, has an influence on the performance of the proposed adaptive algorithms. We first examine the BER performance versus the rank D and then introduce an adaptive rank selection method. Fig. 10 depicts the BER of the TWL/QWL-MSWF algorithms as a function of the rank D, where the performances of the L-MSWF as well as the full-rank counterparts are included for comparison. It can be observed that for both SG and RLS algorithms, $D = 4$ provides the best performance. It is worth remarking that $D = 3$ which performs the same as $D = 4$ is preferred for the SG methods.

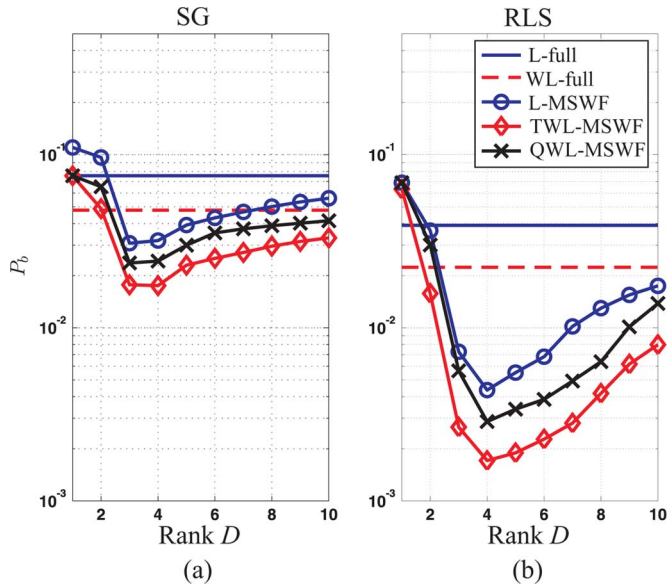


Fig. 10. BER performance of (a) SG and (b) RLS algorithms versus the rank D for $\frac{E_b}{N_0} = 15$ dB, $N_u = 16$, and OFDM-NBI with SIR $= -5$ dB. The number of the received symbols is chosen as 1500.

The performance of the proposed algorithms is rank-dependent. A smaller rank D provides a faster convergence at the beginning of the adaptation and a larger D results in a better steady-state performance (cf. Fig. 11). Thereby, the rank can be adapted to ensure both advantages. We employ an adaptive method proposed in [21] to select the rank D, based on the MSE estimate from *a posteriori* least-squares cost function

$$\mathcal{C}_d(i) = \sum_{m=1}^{i} \lambda^{i-m} \left| b_1(m) - \bar{\boldsymbol{w}}_{a,d}^H(m-1)\boldsymbol{S}_{a,d}^H(m-1)\boldsymbol{r}_a(m) \right|^2 \tag{41}$$

where d represents the rank to be chosen and λ is the exponential weighting factor. For each received symbol, the optimal rank that minimizes the exponentially weighted cost function (41) is selected

$$D_{\text{opt}}(i) = \arg \min_{D_{\min} \le d \le D_{\max}} \mathcal{C}_d(i) \tag{42}$$

$D_{\min}$ and $D_{\max}$ are the minimum and maximum ranks considered. We assess the adaptive rank selection technique for the TWL/QWL-MSWF with both SG and the RLS adaptive algorithms as shown in Fig. 11, where the performance using a fixed rank is also included for comparison. We choose the range of the considered rank is $D_{\min} = 2$ and $D_{\max} = 6$. By adapting the rank at each received symbol, both a fast convergence and a

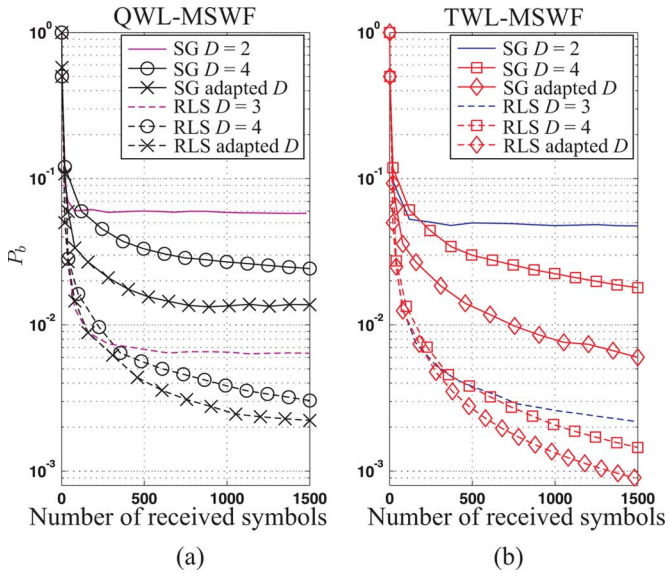


Fig. 11. BER convergence performance of the adaptive rank selection method for (a) the QWL-MSWF and (b) the TWL-MSWF. We choose $\frac{E_b}{N_0} = 15$ dB, $N_u = 16$, and OFDM-NBI with SIR $= -5$ dB.

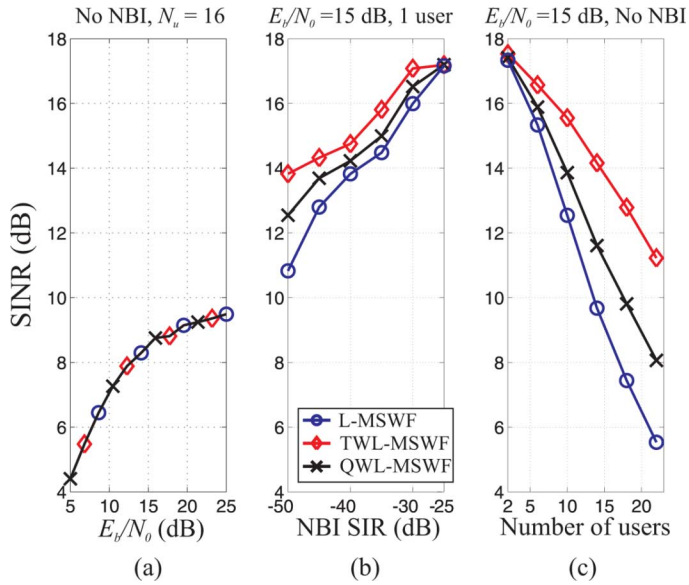


Fig. 12. The SINR of L/TWL/QWL-MSWF algorithms for the QPSK system versus (a) $\frac{E_b}{N_0}$ (dB), (b) various SIR in the presence of BPSK-OFDM NBI, and (c) different number of users (MUI is BPSK modulated with $N = 24$).

better steady-state performance can be attained. The complexity of the adaptive rank selection algorithm lies in the adaptation of the involved quantities for $D_{\min} \leq D \leq D_{\max}$ and the additional calculations of the cost function in (41). The complexity can be reduced by switching off the rank-selection after the steady state is reached.

D. Other Applicable Situations

In the above discussions, we consider the case when both the desired signal and the interference (MUI as well as NBI) are strictly noncircular. In the following, we show the proposed TWL/QWL-MSWF algorithms are still applicable and outperform the L-MSWF in the situation when the desired signal is strictly circular but the interference is noncircular (i.e., the received observation vector $\boldsymbol{r}$ is still noncircular). If $\boldsymbol{r}$ is circular, the performance of the WL algorithms is the same with the linear counterpart. In Fig. 12(a), QPSK is considered for all the users and the same processing gain $N = 24$ is chosen for simplicity. It is obvious that since no advantage can be exploited for the circular observation data (QPSK), the WL methods performs the same as the linear one. Fig. 12(b) and (c) show the case when the desired signal is QPSK modulated (circular) but the interference is noncircular, i.e., MUI is BPSK modulated with $N = 24$ and NBI is the BPSK-OFDM signal. The WL schemes fully exploit the second-order information of the interference, showing a superior performance over the linear scheme.

VIII. Conclusion

To suppress the ISI, the MUI, and the NBI in a high-data-rate DS-UWB system, we propose a WL-MSWF receiver and develop the corresponding adaptive algorithms (i.e., SG and RLS). Based on the linear MSWF concept, two constructions of the rank-reduction matrix (TWL and QWL) are derived. The TWL/QWL-MSWF schemes fully/partially exploit the second-order information of the noncircular signal, yielding a higher SINR than the L-MSWF. Compared to the WL-PC methods, the proposed TWL/QWL-MSWF are simpler and can approach the optimal MMSE with a much smaller rank. We show that the QWL-MSWF can be simplified by taking the real part of the reduced-rank vector after the low-rank transformation in the L-MSWF receiver, indicating a lower complexity. The computational complexity with respect to the number of real additions and multiplications is estimated for the associated SG and RLS adaptive algorithms. The convergence analysis shows that the step size of the WL-MSWF-SG can be larger than that of the L-MSWF-SG. From the MSE point of view, the proposed adaptive algorithms (SG and RLS) exhibit a better transient behavior than the linear counterparts.

Extensive simulation results in terms of the SINR and the BER convergence performance are presented to assist the theoretical analyses. It is shown that the TWL/QWL-MSWF perform better than the existing techniques and the TWL-MSWF provides the best performance. The BER of the WL-MSWF is rank-dependent, where $D = 3$ is desired for the SG algorithm and $D = 4$ for the RLS. Furthermore, we assess an adaptive rank selection method for the WL-MSWF to achieve both a faster convergence and a better steady-state performance. Under the situation when the desired signal is strictly circular but the interference (MUI or NBI) is noncircular, the proposed WL-MSWF outperforms the L-MSWF.

Appendix

Eigenvalue Analysis of $\bar{\boldsymbol{R}}$ and $\bar{\boldsymbol{R}}_a$: We consider the same rank for both the linear and the WL MSWF schemes. Two constructions for the rank-reduction matrix can be represented as $\boldsymbol{S}_{a,D} = \mathcal{T}\{\check{\boldsymbol{S}}_D\} = \mathcal{T}\{\boldsymbol{S}_D\} + \mathcal{T}\{\Delta\boldsymbol{S}_D\}$, where $\Delta\boldsymbol{S}_D = \boldsymbol{0}$ indicates the QWL construction. We define

$\Delta \boldsymbol{S}_{a,D} = \mathcal{T}\{\Delta \boldsymbol{S}_D\}$. The augmented reduced-rank covariance matrix $\bar{\boldsymbol{R}}_a$ can be written as

$$\begin{aligned}\bar{\boldsymbol{R}}_a &= \frac{1}{2}\left(\mathcal{T}\{\boldsymbol{S}_D\}+\Delta \boldsymbol{S}_{a,D}\right)^H \begin{bmatrix} \boldsymbol{R} & \check{\boldsymbol{R}} \\ \check{\boldsymbol{R}}^* & \boldsymbol{R}^* \end{bmatrix} \left(\mathcal{T}\{\boldsymbol{S}_D\}+\Delta \boldsymbol{S}_{a,D}\right) \\ &= \frac{1}{2}\bar{\boldsymbol{R}} + \frac{1}{2}\underbrace{\mathrm{Re}\left\{\boldsymbol{S}_D^H \check{\boldsymbol{R}} \boldsymbol{S}_D^*\right\}}_{\bar{\bar{\boldsymbol{R}}}} \\ &+ \frac{1}{4}\underbrace{\left(\Delta \boldsymbol{S}_{a,D}^H \boldsymbol{R}_a \boldsymbol{S}_{a,D} + \boldsymbol{S}_{a,D}^H \boldsymbol{R}_a \Delta \boldsymbol{S}_{a,D} + \Delta \boldsymbol{S}_{a,D}^H \boldsymbol{R}_a \Delta \boldsymbol{S}_{a,D}\right)}_{\Delta \bar{\boldsymbol{R}}_a} \\ &= \frac{1}{2}\bar{\boldsymbol{R}} + \frac{1}{2}\bar{\bar{\boldsymbol{R}}} + \frac{1}{4}\Delta \bar{\boldsymbol{R}}_a. \end{aligned} \tag{43}$$

Since all the components in (43) are Hermitian matrices, by using the theorem (Weyl) 4.3.1 in [37], we can obtain the kth eigenvalue of the augmented reduced-rank covariance matrix (expressed by $\lambda_k(\cdot), k = 1, \ldots, D, D \ll K$) satisfying

$$\begin{aligned}\lambda_k\left(\bar{\boldsymbol{R}}_a\right) &= \lambda_k\left(\frac{1}{2}\bar{\boldsymbol{R}} + \frac{1}{2}\bar{\bar{\boldsymbol{R}}} + \frac{1}{4}\Delta \bar{\boldsymbol{R}}_a\right) \\ &\le \frac{1}{2}\lambda_k\left(\bar{\boldsymbol{R}}\right) + \frac{1}{2}\lambda_{\max}\left(\bar{\bar{\boldsymbol{R}}}\right) \\ &\quad + \frac{1}{4}\lambda_{\max}\left(\Delta \bar{\boldsymbol{R}}_a\right). \end{aligned} \tag{44}$$

If the QWL is applied, $\Delta \bar{\boldsymbol{R}}_a = \boldsymbol{0}$ and (44) can be simplified as

$$\begin{aligned}\lambda_{ak}^{\mathrm{QWL}} &\le \frac{1}{2}\lambda_k\left(\bar{\boldsymbol{R}}\right) + \frac{1}{2}\lambda_{\max}\left(\bar{\bar{\boldsymbol{R}}}\right) \\ &< \frac{1}{2}\lambda_k\left(\bar{\boldsymbol{R}}\right) + \frac{1}{2}\lambda_k\left(\bar{\boldsymbol{R}}\right) = \lambda_k \end{aligned} \tag{45}$$

where it is given that in general, the singular values of the complementary covariance matrix are smaller than those of the covariance matrix. For the TWL construction, the eigenvalue analysis shows that

$$\lambda_{ak}^{\mathrm{TWL}} < \lambda_k + \frac{1}{4}\lambda_k\left(\Delta \bar{\boldsymbol{R}}_a\right). \tag{46}$$

If $\lambda_k\left(\Delta \bar{\boldsymbol{R}}_a\right)$ is not dominant, $\lambda_{ak}^{\mathrm{TWL}} < \lambda_k$ still holds. However, it is shown in Section VII-A that when there is strong NBI, $\lambda_{ak}^{\mathrm{TWL}} > \lambda_k$ will occur.

REFERENCES

[1] T. Adali and S. Haykin, *Adaptive Signal Processing: Next Generation Solutions*. Hoboken, NJ: Wiley, 2010.

[2] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge, U.K.: Cambridge Univ, Press, 2010.

[3] D. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Hoboken, NJ: Wiley, 2009.

[4] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[5] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[6] H. Trigui and D. Slock, "Performance bounds for cochannel interference cancellation within the current GSM standard," *Signal Process.*, vol. 80, no. 7, pp. 1335–1346, 2000.

[7] F. Römer and M. Haardt, "Deterministic Cramér-Rao bounds for strict sense non-circular sources," in *Proc. Int. ITG/IEEE Workshop Smart Antennas (WSA'07)*, Vienna, Austria, Feb. 2007.

[8] M. Haardt and F. Römer, "Enhancements of unitary ESPRIT for non-circular sources," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Montreal, QC, Canada, 2004, vol. 2, pp. 101–104.

[9] S. Buzzi, M. Lops, and A. M. Tulino, "A new family of MMSE multiuser receivers for interference suppression in DS/CDMA systems employing BPSK modulation," *IEEE Tran. Commun.*, vol. 49, no. 1, pp. 154–167, Jan. 2001.

[10] R. Schober, W. H. Gerstacker, and L. H. J. Lampe, "Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 746–756, Mar. 2004.

[11] A. Mirbagheri, K. N. Plataniotis, and S. Pasupathy, "An enhanced widely linear CDMA receiver with OQPSK modulation," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 261–261, Feb. 2006.

[12] A. Parihar, L. Lampe, R. Schober, and C. Leung, "Equalization for DS-UWB systems—Part I: BPSK modulation," *IEEE Trans. Commun.*, vol. 55, no. 6, pp. 1164–1173, Jun. 2007.

[13] P. Chevalier and F. Pipon, "New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK, and GMSK signals corrupted by noncircular interferences—Application to SAIC," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 870–883, Mar. 2006.

[14] R. Meyer, W. H. Gerstacker, R. Schober, and J. B. Huber, "A single antenna interference cancellation algorithm for increased gsm capacity," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1616–1621, Jul. 2006.

[15] P. Chevalier, F. Pipon, and F. Delaveau, "Second-order optimal array receivers for synchronization of BPSK, MSK, and GMSK signals corrupted by noncircular interferences," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 3, pp. 1–16, 2007.

[16] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 677–690, Feb. 1998.

[17] Y. Song and S. Roy, "Blind adaptive reduced-rank detection for DS-CDMA signals in multipath channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, Nov. 1999.

[18] J. S. Goldstein and I. S. Reed, "Reduced-rank adaptive filtering," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 492–496, Feb. 1997.

[19] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener Filter based on orthogonal projections," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2943–2959, Jul. 1998.

[20] M. L. Honig and W. Xiao, "Performance of reduced-rank linear interference suppression," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1928–1946, May 2001.

[21] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 986–994, Jun. 2002.

[22] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Process.*, vol. 49, no. 2, Feb. 2001.

[23] H. Qian and S. N. Batalama, "Data record-based criteria for the selection of an auxiliary vector estimator of the MMSE/MVDR filter," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1700–1708, Oct. 2003.

[24] R. C. de Lamare, M. Haardt, and R. Sampaio-Neto, "Blind adaptive constrained reduced-rank parameter estimation based on constant modulus design for CDMA interference suppression," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2470–2482, Jun. 2008.

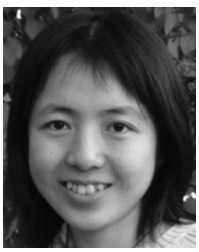
[25] L. Wang and R. C. de Lamare, "Robust auxiliary vector filtering algorithm based on constrained constant modulus design for adaptive beamforming," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, Dallas, TX, 2010, pp. 2530–2533.

[26] S. C. Douglas, "Widely-linear recursive least-squares algorithm for adaptive beamforming," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'09)*, Taipei, Taiwan, 2009, pp. 2041–2044.

[27] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank adaptive filtering based on joint iterative optimization of adaptive filters," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 980–983, Dec. 2007.

[28] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation and filtering," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2503–2514, Jul. 2009.

[29] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank space-time adaptive interference suppression with joint iterative least squares algorithms for spread-spectrum systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1217–1228, Jun. 2010.
[30] R. Fisher, R. Kohno, M. M. Laughlin, and M. Welbourn, "DS-UWB physical layer submission to 802.15 Task Group 3a," *IEEE Standard Proposal IEEE P802.15-04/0137r4*, Jan. 2005.
[31] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications High-Speed Physical Layer in the 5 GHz Band*, IEEE 802.15 TG4a, 1999.
[32] First Report and Order: Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems Federal Communications Commission, 2002, ET Docker 98-153.
[33] Q. Z. Ahmed and L.-L. Yang, "Reduced-rank adaptive multiuser detection in hybrid direct-sequence time-hopping ultra wide bandwidth systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 156–167, Jan. 2010.
[34] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice Hall, 2002.
[35] N. Song, R. de Lamare, M. Wolf, and M. Haardt, "Adaptive reduced-rank interference suppression for DS-UWB systems based on the widely linear multistage Wiener filter," in *Proc. 7th Int. Symp. Wireless Commun. Syst.(ISWCS 2010)*, York, U.K., Sep. 2010.
[36] J. Kunisch and J. Pamp, "Measurement results and modeling aspects for the UWB radio channel," in *Proc. IEEE Conf. Ultra Wide-Band Syst. Technol.*, Baltimore, MD, May 2002.
[37] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.



Nuan Song (S'07) was born in China in December, 1981. She received the B.Sc. degree in electronic and information engineering from the Northwestern Polytechnical University, Xi'an, China, in 2004, and the M.Sc. degree in digital communication systems and technology from Chalmers University of Technology, Gothenburg, Sweden, in 2006. She is currently working toward the Ph.D. at the Communication Research Laboratory at Ilmenau University of Technology, Ilmenau, Germany, where she has worked as a Research Assistant since February, 2006.

Her current research interests include the transmission and receiver techniques for ultrawideband communication systems and signal processing for wireless communications. She was the recipient of the STINT scholarship by the Swedish Foundation for International Cooperation in Research and Higher Education during the Master's study in Chalmers University of Technology.

Rodrigo C. de Lamare (S'99–M'04–SM'10) received the Dipl. degree in electronic engineering from the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, in 1998, and the M.Sc. and Ph.D. degrees from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, both in electrical engineering, in 2001 and 2004, respectively.

Since January 2006, he has been with the Communications Research Group, Department of Electronics, University of York, York, U.K., where he is currently a Senior Lecturer in Communications Engineering. His research interests include communications and signal processing, areas in which he has published about 250 papers in refereed journals and conferences. He has served as the General Chair of the 7th IEEE International Symposium on Wireless Communications Systems (ISWCS), held in York, U.K., in September 2010, and will serve as the technical programme chair of ISWCS 2013 in Ilmenau, Germany.

Dr. de Lamare serves as an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking*.

Martin Haardt (S'90–M'98–SM'99) received the B.Sc. degree in electrical engineering from Purdue University, West Lafayette, IN. He received the M.Sc. degree from the Ruhr-University Bochum, Bochum, Germany, in 1991, and the Ph.D. degree from Munich University of Technology, Munich, Germany, in 1996.

He has been a Full Professor in the Department of Electrical Engineering and Information Technology and Head of the Communications Research Laboratory at Ilmenau University of Technology, Ilmenau, Germany, since 2001. In 1997, he joined Siemens Mobile Networks in Munich, Germany, where he was responsible for strategic research for third generation mobile radio systems. From 1998 to 2001, he was the Director for International Projects and University Cooperations in the mobile infrastructure business of Siemens in Munich, where his work focused on mobile communications beyond the third generation. During his time at Siemens, he also taught in the international Master of Science in Communications Engineering program at Munich University of Technology. In the fall of 2006 and the fall of 2007, he was a visiting professor at the University of Nice in Sophia-Antipolis, France, and at the University of York, U.K., respectively. His research interests include wireless communications, array signal processing, high-resolution parameter estimation, as well as numerical linear and multilinear algebra.

Prof. Haardt has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2002–2006 and since 2011), the IEEE SIGNAL PROCESSING LETTERS (2006–2010), the *Research Letters in Signal Processing* (2007–2009), the *Hindawi Journal of Electrical and Computer Engineering* (since 2009), the *EURASIP Signal Processing Journal* (since 2011), and as a Guest Editor for the *EURASIP Journal on Wireless Communications and Networking*. He has also served as an elected member of the Sensor Array and Multichannel (SAM) technical committee of the IEEE Signal Processing Society (since 2011), as the technical cochair of the IEEE International Symposiums on Personal Indoor and Mobile Radio Communications (PIMRC) 2005 in Berlin, Germany, as the technical program chair of the IEEE International Symposium on Wireless Communication Systems (ISWCS) 2010 in York, U.K., and as the general chair of ISWCS 2013 in Ilmenau, Germany. He received the 2009 Best Paper Award from the IEEE Signal Processing Society, the Vodafone (formerly Mannesmann Mobilfunk) Innovations-Award for outstanding research in mobile communications, the ITG best paper award from the Association of Electrical Engineering, Electronics, and Information Technology (VDE), and the Rohde & Schwarz Outstanding Dissertation Award.

Mike Wolf received the Dipl. and Ph.D. degrees in electrical engineering and information technology from Ilmenau University of Technology, Ilmenau, Germany.

He is currently working as a Lecturer in the Communications Research Laboratory of Ilmenau University of Technology. His research areas include wireless optics and ultrabroadband communications.