Adaptive Frequency-Domain Group-based Shrinkage Estimators (GSE) for UWB Systems

Sheng Li, *IEEE Member*, Rodrigo C. de Lamare, *IEEE Senior Member*, and Martin Haardt, *IEEE Senior Member*

Abstract—In this work, we propose low-complexity adaptive biased estimation algorithms, called group-based shrinkage estimators (GSEs), for parameter estimation and interference suppression scenarios with mechanisms to automatically adjust the shrinkage factors. The proposed estimation algorithms divide the target parameter vector into a number of groups and adaptively calculate one shrinkage factor for each group. GSE schemes improve the performance of the conventional least squares (LS) estimator in terms of the mean-squared error (MSE), while requiring a very modest increase in complexity. An MSE analysis is presented which indicates the lower bounds of the GSE schemes with different group sizes. We prove that our proposed schemes outperform the biased estimation with only one shrinkage factor and the best performance of GSE can be obtained with the maximum number of groups. Then, we consider an application of the proposed algorithms to single-carrier frequency-domain equalization (SC-FDE) of direct-sequence ultra-wideband (DS-UWB) systems, in which the structured channel estimation (SCE) algorithm and the frequency domain receiver employ the GSE. The simulation results show that the proposed algorithms significantly outperform the conventional unbiased estimator in the analyzed scenarios.

Index Terms–DS-UWB systems, parameter estimation, interference suppression, biased estimation, adaptive algorithm.

I. Introduction

In this work, biased estimation algorithms are considered in two common deterministic estimation scenarios in communications engineering, which are parameter estimation and interference suppression [1]-[5]. It is known that under the assumption of AWGN, the least-square (LS) algorithm can provide an efficient solution to these estimation problems and will lead to minimum variance unbiased estimators (MVUE). The unbiasness is usually considered as a good property for an estimator because the expected value of unbiased estimators is the true value of the unknown parameter [6]. However, in some scenarios the LS method is not directly related to the mean square error (MSE) associated with the target parameter vector and it has been found that a lower MSE can be achieved by adding an appropriately chosen bias to the conventional LS estimators [7],[31]. Note that some reduced-rank techniques

Sheng Li is with the College of Information Engineering, Zhejiang University of Technology, China (e-mail: shengli@zjut.edu.cn); Rodrigo C. de Lamare is with the Communications Research Group, Department of Electronics, University of York, UK (e-mail: rcdl500@ohm.york.ac.uk); Martin Haardt is with the Communications Research Laboratory, Ilmenau University of Technology, Germany (e-mail: martin.haardt@tu-ilmenau.de)

also employ a bias to accelerate the convergence speed [32]-[35].

A class of biased estimator that has been studied in recent years is known as the biased estimators with a shrinkage factor [7]-[14]. These biased estimation algorithms have shown their ability to outperform the existing unbiased estimators especially in low signal-to-noise ratios (SNR) scenarios and/or with short data records [9]. For these biased estimators [7]-[14], the complexity is much lower than for MMSE algorithms because the additional number of coefficients to be computed is only one. The motivation for the group-based shrinkage estimator (GSE) is to find a generalized estimator with a number of shrinkage factors that can achieve a better performance and complexity tradeoff than the biased estimator with only one shrinkage factor.

In the parameter estimation scenario,some biased estimators have been proposed to achieve a smaller estimation error than the LS solutions by removing the unbiasedness of the conventional estimators with a shrinkage factor in the parameter estimation scenario. The earliest shrinkage estimators that reduce the MSE over MVUE include the well known James-Stein estimator [10] and the work of Thompson [11]. Some extensions of the James-Stein estimator have been proposed in [12]-[15]. In [16], blind minimax estimation (BME) techniques have been proposed, in which the biased estimators were developed to minimize the worst case MSE among all possible values of the target parameter vector within a parameter set. If a spherical parameter set is assumed, the shrinkage estimator obtained is named spherical BME (SBME) [16].

For the interference suppression scenario, the biased estimators can be employed to achieve a lower estimation error between the estimated filter and the optimal linear LS estimator. The major motivation for adopting the biased algorithms here is to accelerate the convergence rate for the adaptive implementations and provide a better performance with short training data support in long filter scenarios [13].

To the best of our knowledge, biased estimators with shrinkage factors are rarely implemented into real-world signal processing and have not been considered in the frequency domain for communication systems. One possible reason is that some assumptions required for the signal model may not be satisfied. For example, in time-hopping UWB (TH-UWB) systems, the multiple access interference (MAI) cannot be accurately approximated by a Gaussian distribution for some values of the the signal-to-interference-plus-noise ratio (SINR) [17]. Another possible reason is that the existing shrinkage-based estimators usually require some statistical information such as the noise variance and the norm of the actual parameter

vector. In our previous work [13] and [14], adaptive biased estimation algorithms with only one shrinkage factor have been proposed to fulfill the tasks of interference suppression and parameter estimation. In this work, a novel biased estimation technique, named group-based shrinkage estimators (GSE), is proposed. In this algorithm, the target parameter vector is divided into several groups and one shrinkage factor is calculated for each group. Least mean square (LMS)-based adaptive estimation algorithms are then developed to calculate the shrinkage factors. The GSE estimators are able to improve the performance of the recursive least squares (RLS) algorithm that recursively computes the LS estimator. In DS-UWB systems, the estimation tasks are usually very challenging because the environments include dense multipath. In this work, in order to test the proposed algorithms, we consider applications of DS-UWB systems with SC-FDE. Specifically, we concentrate on the channel estimation and interference suppression with the proposed algorithms. The MSE performance of the proposed GSE schemes is then analyzed, a lower bound of the MSE performance is derived and the relationship between the number of groups and the lower bound is set up. Simulations show that with an additional complexity that is only linearly dependent on the size of the parameter vector and the number of groups, the proposed biased GSE algorithms outperform the conventional RLS algorithm in terms of MSE in low SNR scenarios and/or with short data support. It should be noted that the proposed GSE estimator can be employed for applications where a high estimation accuracy is required. These include localization in wireless sensor networks [28] and in dense cluttered environments with UWB technology [29]. The proposed estimators can also be employed into emergent multicast and broadcast systems [5], such as the orthogonal frequency-division multiplexing (OFDM) based multi-user multiple-input multiple-output (MIMO) systems as specified in the IEEE 802.11ac standard and the 3GPP long-term-evolution (LTE) systems.

The main contributions of this work are summarized as follows:

- Novel GSE schemes are proposed to improve the performance of the frequency domain RLS algorithms in the applications of parameter estimation and interference suppression in DS-UWB systems.
- LMS based adaptive algorithms are developed for both scenarios to adjust the shrinkage factors.
- The MSE analysis is carried out which indicates a lower bound of the proposed estimator and the relationship between the lower bound and the number of groups.
- The performance of the proposed biased estimators is examined in multiuser SC-FDE for DS-UWB systems with the IEEE 802.15.4a channel model, convolutional and low-density parity-check (LDPC) codes.

The rest of this paper is structured as follows. In Section II, we first review the LS solution for the parameter estimation scenario and present the structured channel estimation (SCE) problem in SC-FDE of DS-UWB systems. Then, the signal model of the frequency domain receiver design for DS-UWB systems that represents the interference suppression scenario,

is presented. The proposed GSE scheme and its adaptive implementations for the parameter estimation scenario and the interference suppression scenario are developed in Section III and Section IV, respectively. The MSE analysis is shown in Section V. The simulation results are shown in Section VI. Section VII draws the conclusions.

II. SYSTEM MODEL

In this section, we introduce the channel estimation and receiver design tasks in the frequency-domain of DS-UWB systems with SC-FDE that represent the parameter estimation scenario and the interference suppression scenario, respectively.

A. Problem statement for the parameter estimation scenario

The linear model for the parameter estimation scenario can be expressed as:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{h} + \boldsymbol{n}, \tag{1}$$

where the training data matrix $\boldsymbol{X} \in \mathbb{C}^{M \times L}$ and the received signal $\boldsymbol{y} \in \mathbb{C}^{M \times 1}$ are given, $\boldsymbol{n}$ is AWGN with zero mean and variance σ^2. In this scenario, the typical target is to estimate the parameter vector $\boldsymbol{h} \in \mathbb{C}^{L \times 1}$ that leads to the minimum MSE. The MSE consists of the estimation variance and the squared bias and is given by

$$\mathbb{E}\{\|\boldsymbol{h}-\hat{\boldsymbol{h}}\|^2\} = \mathbb{E}\{(\hat{\boldsymbol{h}}-\mathbb{E}\{\hat{\boldsymbol{h}}\})^H(\hat{\boldsymbol{h}}-\mathbb{E}\{\hat{\boldsymbol{h}}\})\}+[\|\mathbb{E}\{\hat{\boldsymbol{h}}\}-\boldsymbol{h}\|^2],$$

where $\mathbb{E}\{\cdot\}$ represents expectation of a random variable.

The conventional LS algorithm estimates the parameter by minimizing

$$J_{\text{LS}}(\boldsymbol{h}) = \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{h}\|^2. \tag{2}$$

Assuming that the matrix $\boldsymbol{X}^H\boldsymbol{X}$ is a full rank matrix, the LS solution is given by

$$\hat{\boldsymbol{h}}_{\text{LS}} = (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H\boldsymbol{y}. \tag{3}$$

Under the assumption of AWGN with zero mean and variance σ^2, the LS estimator is a MVUE that leads to a minimum MSE

$$\mathbb{E}\{\|\boldsymbol{h}-\hat{\boldsymbol{h}}_{\text{LS}}\|^2\} = \mathbb{E}\{(\hat{\boldsymbol{h}}_{\text{LS}}-\boldsymbol{h})^H(\hat{\boldsymbol{h}}_{\text{LS}}-\boldsymbol{h})\} = var(\boldsymbol{h},\hat{\boldsymbol{h}}_{\text{LS}}).$$

we define $v = var(\boldsymbol{h},\hat{\boldsymbol{h}}_{\text{LS}}) = tr\{\sigma^2(\boldsymbol{X}^H\boldsymbol{X})^{-1}\}$, where $tr\{\cdot\}$ represents the trace operator [6].

The objective of introducing the biased estimator in the parameter estimation scenario is to achieve a lower MSE than the unbiased estimator, which can be expressed as

$$\mathbb{E}\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\text{b}}\|^2\} \leq \mathbb{E}\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\text{LS}}\|^2\}. \tag{4}$$

Although the objective shown here is similar to MMSE algorithms that is to achieve an MSE as small as possible. It should be noted that the biased algorithm developed in this work adopts a different strategy from MMSE algorithms.
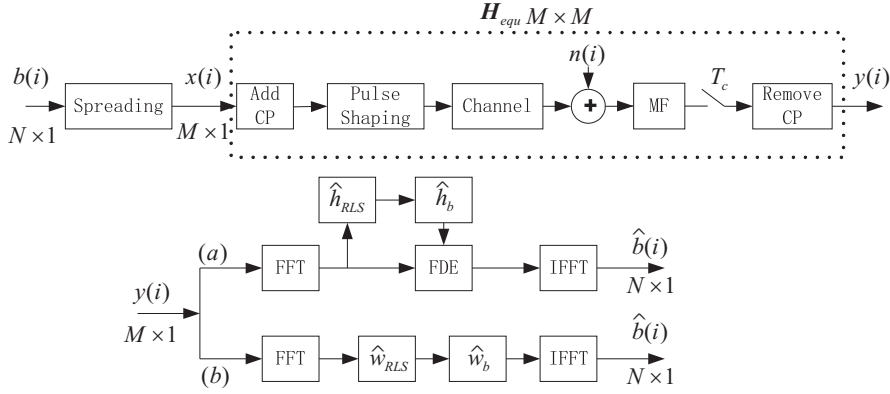
Fig. 1. Block diagram of SC-FDE schemes in DS-UWB systems: (a) Parameter estimation scenario,(b) Interference suppression scenario.

B. System model for the SCE: parameter estimation scenario

Here, we consider the channel estimation problem of a synchronous downlink block-by-block transmission DS-UWB system based on SC-FDE with K users. The block diagram of the parameter estimation scenario is shown as branch (a) in Fig. 1. For notational simplicity, we assume that a N_c-by-1 Walsh spreading code s_k is assigned to the k-th user. The spreading gain is $N_c = T_s/T_c$, where T_s and T_c denote the symbol duration and chip duration, respectively. At each time instant i, a data vector $b_k(i) \in \mathbb{C}^{N \times 1}$ is transmitted by the k-th user. We define the signal after spreading as $x_k(i) = D_k b_k(i)$, where the block diagonal matrix $D_k \in \mathbb{C}^{M \times N}$ ($M = N \cdot N_c$) performs the spreading of the data block and its first column is constructed by the spreading code s_k zero-padded to the length of M. In order to prevent inter block interference (IBI), a cyclic-prefix (CP) is added and the length of the CP is assumed to be larger than the length of the channel impulse response (CIR). With the insertion of the CP at the transmitter and its removal at the receiver, the equivalent channel is denoted as a circulant Toeplitz matrix $H_{equ} \in \mathbb{C}^{M \times M}$, whose first column is composed of a vector h_{equ} zero-padded to length M, where $h_{equ} = [h(0), h(1), \ldots, h(L-1)]$ is the equivalent CIR. At the receiver, a chip matched-filter (CMF) is applied and the received sequence is then sampled at chip-rate and organized in an M-dimensional vector. This signal then goes through the discrete Fourier transform (DFT). The frequency-domain received signal is given by

$$y(i) = F H_{equ} \sum_{k=1}^{K} x_k(i) + F n(i) = \Lambda_H F \sum_{k=1}^{K} x_k(i) + F n(i),$$
(5)

where $n(i)$ represents the AWGN, $\Lambda_H = F H_{equ} F^H$ is a diagonal matrix whose diagonal vector is defined as $\tilde{h}$ and its a-th entry is given by $\tilde{h}_a = \sum_{l=0}^{L-1} h_l \cdot \exp\{-j(2\pi/M)a\,l\}$, where $a \in \{0, 1, 2, \ldots, M-1\}$. $F \in \mathbb{C}^{M \times M}$ represents the DFT matrix and its (a,b)-th entry is $F_{a,b} = (1/\sqrt{M})\exp\{-j(2\pi/M)a\,b\}$, where $a, b \in \{0, 1, 2, \ldots, M-1\}$. By defining a matrix $F_{M,L} \in \mathbb{C}^{M \times L}$ that contains the first L columns of the DFT matrix F, we obtain the following relationship

$$\tilde{h} = \sqrt{M} F_{M,L} h_{equ},$$
(6)

In unstructured channel estimation (UCE), the vector $\tilde{h} \in \mathbb{C}^{M \times 1}$ is directly estimated, while in the structured channel estimation (SCE), the fact that $L < M$ is taken into account and the vector $h_{equ} \in \mathbb{C}^{L \times 1}$ is the parameter vector to be estimated. The concept of SCE was proposed in [20], where the SCE shows a better performance than the UCE. In [21], adaptive MMSE detection schemes for SC-FDE in multi-user DS-UWB systems based on SCE are developed, where the estimated h_{equ} is adaptively calculated based on RLS, least-mean squares (LMS) and the conjugate gradient (CG) algorithm for the detection and the RLS version performs the best. The purpose of developing biased estimation in this scenario is to further improve the performance of the RLS algorithm in terms of the MSE.

We consider user 1 as the desired user and omit the subscript of this user for simplicity. Note that the frequency domain received signal can be expressed as

$$y(i) = \sqrt{M} \Delta(i) F_{M,L} h_{equ} + n_e(i),$$
(7)

where we define a diagonal matrix $\Delta(i) = \text{diag}[Fx(i)]$, the noise and interference vector $n_e(i)$ consists of the MAI and the noise and is assumed to be AWGN. As shown in (7), the SCE problem is an implementation example of the parameter estimation problem where a given matrix $X(i) \in \mathbb{C}^{M \times L}$ is defined as $X(i) = \sqrt{M} \Delta(i) F_{M,L}$. The LS solution of h_{equ} is given by

$$h_{equ,LS}(i) = R_h^{-1}(i) p_h(i)$$
(8)

where $R_h(i) = \sum_{j=1}^{i} \lambda^{i-j} F_{M,L}^H \Delta^H(j) \Delta(j) F_{M,L}$, $p_h(i) = \sum_{j=1}^{i} \lambda^{i-j} F_{M,L}^H \Delta^H(j) z(j)$ and λ is the forgetting factor. Then the LS solution can be computed recursively by the following RLS algorithm [20]

$$\hat{h}_{RLS}(i+1) = \hat{h}_{RLS}(i) + R_h^{-1}(i) F_{M,L}^H \Delta^H(i) e_h(i),$$
(9)

where $e_h(i) = z(i) - \Delta(i) F_{M,L} \hat{h}_{RLS}(i)$ is the M-dimensional error vector.

In Section III, a novel biased estimation algorithm called group-based shrinkage estimator (GSE) is incorporated into the unbiased LS estimator that is able to improve the estimation performance in terms of the MSE.

C. System model for the frequency domain receiver design: interference suppression scenario

The block diagram of the interference suppression scenario is shown as branch (b) in Fig. 1. For each time instant i, an N-dimensional data vector $\boldsymbol{b}_k(i) = [b_k^{(1)}(i), \ldots, b_k^{(N)}(i)]^T$ is transmitted by the k-th user. After the spreading, the M-dimensional transmit signal is given by

$$\boldsymbol{x}_k(i) = \boldsymbol{S}_k \boldsymbol{b}_{k,\mathrm{e}}(i), \qquad (10)$$

where the spreading matrix $\boldsymbol{S}_k \in \mathbb{C}^{M \times M}$, $M = N \times N_c$, is a circulant Toeplitz matrix and its first column consists of the spreading codes and zero-padding [26]. The equivalent M-dimensional expanded data vector is

$$\boldsymbol{b}_{k,\mathrm{e}}(i) = [b_k^{(1)}(i), \boldsymbol{0}_{N_c-1}, b_k^{(2)}(i), \boldsymbol{0}_{N_c-1}, \cdots, b_k^{(N)}(i), \boldsymbol{0}_{N_c-1}]^T,$$

where $(\cdot)^T$ is the transpose. Using this signal expression we can obtain a simplified frequency domain receiver design. At the receiver, a CMF is applied and the received sequence is then sampled at chip-rate and organized in an M-dimensional vector. After the DFT, the received signal is given by

$$\boldsymbol{z}(i) = \boldsymbol{F}\boldsymbol{y}(i) = \sum_{k=1}^K \boldsymbol{F}\boldsymbol{H}_{\mathrm{equ}}\boldsymbol{S}_k \boldsymbol{b}_{k,\mathrm{e}}(i) + \boldsymbol{F}\boldsymbol{n}(i), \qquad (11)$$

where $\boldsymbol{n}(i)$ is the AWGN and $\boldsymbol{F} \in \mathbb{C}^{M \times M}$ represents the DFT matrix. Since both $\boldsymbol{H}_{\mathrm{equ}}$ and $\boldsymbol{S}_k$ are circulant Toeplitz matrices, their product also has the circulant Toeplitz form. This feature makes $\boldsymbol{\Lambda}_k = \boldsymbol{F}\boldsymbol{H}_{\mathrm{equ}}\boldsymbol{S}_k \boldsymbol{F}^H$ a diagonal matrix. Hence, we have

$$\boldsymbol{z}(i) = \sum_{k=1}^K \boldsymbol{\Lambda}_k \boldsymbol{F}\boldsymbol{b}_{k,\mathrm{e}}(i) + \boldsymbol{F}\boldsymbol{n}(i). \qquad (12)$$

We can further expand $\boldsymbol{F}\boldsymbol{b}_{k,\mathrm{e}}(i)$ as [26]

$$\boldsymbol{F}\boldsymbol{b}_{k,\mathrm{e}}(i) = (1/\sqrt{N_c})\boldsymbol{I}_{\mathrm{e}}\boldsymbol{F}_N \boldsymbol{b}_k(i), \qquad (13)$$

where $\boldsymbol{F}_N \in \mathbb{C}^{N \times N}$ denotes the DFT matrix and $\boldsymbol{I}_{\mathrm{e}} \in \mathbb{C}^{M \times N}$ is structured as

$$\boldsymbol{I}_{\mathrm{e}} = [\underbrace{\boldsymbol{I}_N, \cdots, \boldsymbol{I}_N}_{N_c}]^T. \qquad (14)$$

where $\boldsymbol{I}_N$ denotes the N-by-N identity matrix. Finally, the frequency domain received signal $\boldsymbol{z}(i)$ is given by

$$\boldsymbol{z}(i) = \sum_{k=1}^K (1/\sqrt{N_c})\boldsymbol{\Lambda}_k \boldsymbol{I}_{\mathrm{e}}\boldsymbol{F}_N \boldsymbol{b}_k(i) + \boldsymbol{F}\boldsymbol{n}(i). \qquad (15)$$

Note that the expression in (15) is an implementation example of the interference suppression scenario where the unknown matrix for each time instant is given by $\boldsymbol{H}(i) = (1/\sqrt{N_c})\boldsymbol{\Lambda}_k \boldsymbol{I}_{\mathrm{e}}\boldsymbol{F}_N$. To fulfill the interference suppression task, an MMSE filter $\boldsymbol{W}(i) \in \mathbb{C}^{M \times N}$ can be developed via the following cost function:

$$J_{\mathrm{MSE}}(i) = \mathbb{E}\left\{ \left\| \boldsymbol{b}(i) - \boldsymbol{F}_N^H \boldsymbol{W}^H(i)\boldsymbol{z}(i) \right\|^2 \right\}. \qquad (16)$$

The MMSE solution is given by [21]

$$\boldsymbol{W}_{\mathrm{MMSE}} = \boldsymbol{V}\boldsymbol{I}_{\mathrm{e}}, \qquad (17)$$

where the matrix $\boldsymbol{V} \in \mathbb{C}^{M \times M}$ is

$$\boldsymbol{V} = \frac{1}{\sqrt{N_c}}\left(\frac{1}{N_c}\sum_{k=1}^K \boldsymbol{\Lambda}_k \boldsymbol{I}_{\mathrm{e}}\boldsymbol{I}_{\mathrm{e}}^H \boldsymbol{\Lambda}_k^H + \sigma^2 \boldsymbol{I}_M \right)^{-1}\boldsymbol{\Lambda}_k, \quad (18)$$

where $\boldsymbol{I}_M \in \mathbb{C}^{M \times M}$ denotes the identity matrix. Note that the matrix $\boldsymbol{V}$ consists of N_c times N_c diagonal matrices $\boldsymbol{V}_{ij} \in \mathbb{C}^{N \times N}$, where $i,j \in \{1, N_c\}$. Hence, we take a closer look at the product of $\boldsymbol{V}$ and $\boldsymbol{I}_{\mathrm{e}}$:

$$\boldsymbol{V}\boldsymbol{I}_e = \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} & \cdots & \boldsymbol{V}_{1,N_c} \\ \boldsymbol{V}_{2,1} & \boldsymbol{V}_{2,2} & \cdots & \boldsymbol{V}_{2,N_c} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{V}_{N_c,1} & \boldsymbol{V}_{N_c,2} & \cdots & \boldsymbol{V}_{N_c,N_c} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_N \\ \boldsymbol{I}_N \\ \vdots \\ \boldsymbol{I}_N \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^{N_c} \boldsymbol{V}_{1,j} \\ \sum_{j=1}^{N_c} \boldsymbol{V}_{2,j} \\ \vdots \\ \sum_{j=1}^{N_c} \boldsymbol{V}_{N_c,j} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{W}}_1 & & & \\ & \hat{\boldsymbol{W}}_2 & & \\ & & \ddots & \\ & & & \hat{\boldsymbol{W}}_{N_c} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_N \\ \boldsymbol{I}_N \\ \vdots \\ \boldsymbol{I}_N \end{bmatrix} = \hat{\boldsymbol{W}}\boldsymbol{I}_{\mathrm{e}},$$

where $\hat{\boldsymbol{W}}_i = \sum_{j=1}^{N_c} \boldsymbol{V}_{i,j}$, $i = 1, \ldots, N_c$, are diagonal matrices. Hence, the product of $\boldsymbol{V}$ and $\boldsymbol{I}_{\mathrm{e}}$ can be converted into a product of a diagonal matrix $\hat{\boldsymbol{W}} \in \mathbb{C}^{M \times M}$ ($M = N \times N_c$) and $\boldsymbol{I}_{\mathrm{e}}$, where the diagonal entries of $\hat{\boldsymbol{W}}$ are $\hat{w}_l$, $l = 1, \ldots, M$, and equal the sum of all entries in the l-th row of matrix $\boldsymbol{V}$. Finally, we express the MMSE design as

$$\boldsymbol{W}_{\mathrm{MMSE}} = \hat{\boldsymbol{W}}\boldsymbol{I}_{\mathrm{e}} = \mathrm{diag}(\hat{\boldsymbol{w}}_{\mathrm{e}})\boldsymbol{I}_{\mathrm{e}}, \qquad (19)$$

where $\hat{\boldsymbol{w}}_{\mathrm{e}} = (\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_M)$ is an equivalent filter with M taps.

The expression shown in (19) enables us to design an M-dimensional receive filter rather than an M-by-N matrix form receive filter. The estimated data vector can be expressed as

$$\hat{\boldsymbol{b}}(i) = \boldsymbol{F}_N^H \boldsymbol{I}_{\mathrm{e}}^H \hat{\boldsymbol{W}}^H(i)\boldsymbol{z}(i) = \boldsymbol{F}_N^H \boldsymbol{I}_{\mathrm{e}}^H \hat{\boldsymbol{Z}}(i)\hat{\boldsymbol{w}}(i), \qquad (20)$$

where $\hat{\boldsymbol{Z}}(i) = \mathrm{diag}(\boldsymbol{z}(i))$ and $\hat{\boldsymbol{w}}(i) = \hat{\boldsymbol{w}}_{\mathrm{e}}^*(i)$ is the weight vector of the adaptive receiver. Since $\boldsymbol{F}_N$ and $\boldsymbol{I}_{\mathrm{e}}$ are fixed, we consider the equivalent N-by-M received data matrix as $\boldsymbol{Y}(i) = \boldsymbol{F}_N^H \boldsymbol{I}_{\mathrm{e}}^H \hat{\boldsymbol{Z}}(i)$, and express the estimated data vector as $\hat{\boldsymbol{b}}(i) = \boldsymbol{Y}(i)\hat{\boldsymbol{w}}(i)$.

D. LS solution and adaptive RLS algorithm for the interference suppression scenario

Here, we detail the LS and RLS designs for the frequency domain multiuser receiver $\hat{\boldsymbol{w}}$. The cost function for the development of the LS estimation is given by

$$J_{\mathrm{LS}} = \|\boldsymbol{b} - \boldsymbol{Y}\boldsymbol{w}\|^2. \qquad (21)$$

The LS design of the linear receiver can be expressed as

$$\hat{\boldsymbol{w}}_{\mathrm{LS}} = (\boldsymbol{Y}^H \boldsymbol{Y})^{-1}\boldsymbol{Y}^H \boldsymbol{b} = \boldsymbol{R}_{\mathrm{LS}}^{-1}\boldsymbol{p}_{\mathrm{LS}}, \qquad (22)$$

where the matrix $\boldsymbol{R}_{\mathrm{LS}}$ is defined as $\boldsymbol{Y}^H \boldsymbol{Y}$ and $\boldsymbol{p}_{\mathrm{LS}}$ represents the vector $\boldsymbol{Y}^H \boldsymbol{b}$. Note that, the data vector can be expressed as

$$\boldsymbol{b} = \boldsymbol{Y}\boldsymbol{w}_{\mathrm{o}} + \boldsymbol{\epsilon}_{\mathrm{o}}, \qquad (23)$$

where $\boldsymbol{\epsilon}_{\mathrm{o}}$ is the measurement error vector and $\boldsymbol{w}_{\mathrm{o}}$ is the optimum tap-weight vector of the receiver (optimum in the

MSE sense). Assuming that $\boldsymbol{\epsilon}_{\mathrm{o}}$ is white and Gaussian with zero mean and covariance of $\sigma_e^2 \boldsymbol{I}_N$, then the LS solution in (22) is a MVUE [27]. Now, let us have a look at the following MSE:

$$\mathbb{E}\{\|\boldsymbol{w}_{\mathrm{o}}-\hat{\boldsymbol{w}}_{\mathrm{LS}}\|^2\} = \mathbb{E}\{(\boldsymbol{w}_{\mathrm{o}}-\hat{\boldsymbol{w}}_{\mathrm{LS}})^H(\boldsymbol{w}_{\mathrm{o}}-\hat{\boldsymbol{w}}_{\mathrm{LS}})\} = var(\boldsymbol{w}_{\mathrm{o}},\hat{\boldsymbol{w}}_{\mathrm{LS}}) \tag{24}$$

Defining $v_w = var(\boldsymbol{w}_{\mathrm{o}},\hat{\boldsymbol{w}}_{\mathrm{LS}})$, we have [6]

$$v_w = tr\{\sigma_e^2(\boldsymbol{Y}^H\boldsymbol{Y})^{-1}\}, \tag{25}$$

where σ_e^2 is the variance of the measurement error.

In the interference suppression scenario, it is possible to introduce the biased estimation to reduce the MSE between the optimal receive filter $\boldsymbol{w}_{\mathrm{o}}$ and the LS estimator $\hat{\boldsymbol{w}}_{\mathrm{LS}}$. Note that, for the interference suppression scenario, the typical objective is to minimize the overall performance criterion which is determined as $\mathbb{E}\{\|\boldsymbol{b}-\hat{\boldsymbol{b}}\|^2\}$, rather than to minimize $\mathbb{E}\{\|\boldsymbol{w}_{\mathrm{o}}-\hat{\boldsymbol{w}}\|^2\}$. The main motivation to introduce the bias in the interference suppression scenario is to provide an initial improvement for the overall performance when the adaptive filtering techniques are employed and the training data are limited. This can also help with tracking problems and with robustness against interference.

The LS solution of the receiver can be computed recursively by the RLS adaptive algorithm. We employ the RLS update equation that is proposed in [21]

$$\hat{\boldsymbol{w}}_{\mathrm{RLS}}(i+1) = \hat{\boldsymbol{w}}_{\mathrm{RLS}}(i) + \boldsymbol{R}_{\mathrm{w}}^{-1}(i)\boldsymbol{Y}^H(i)\boldsymbol{e}_{\mathrm{aw}}(i), \tag{26}$$

where $\boldsymbol{R}_{\mathrm{w}}(i) = \sum_{j=1}^i \lambda_w^{i-j}\boldsymbol{Y}^H(j)\boldsymbol{Y}(j)$ and $\boldsymbol{e}_{\mathrm{aw}}(i) = \boldsymbol{b}(i) - \boldsymbol{Y}(i)\hat{\boldsymbol{w}}(i)$. Note that $\boldsymbol{R}_{\mathrm{w}}(i)$ is an M-by-M symmetric sparse matrix in which the number of nonzero elements equals MN_c. Hence, the complexity of each adaptation by using this algorithm is $\mathcal{O}(MN_c^2)$.

III. Proposed GSE for parameter estimation scenario

A. Proposed GSE: Optimal Solution

It is known that the biased estimator with a shrinkage factor can be expressed as

$$\hat{\boldsymbol{h}}_{\mathrm{b}} = (1+\alpha)\hat{\boldsymbol{h}}_{\mathrm{LS}}, \tag{27}$$

where $\hat{\boldsymbol{h}}_{\mathrm{LS}} \in \mathbb{C}^{L\times 1}$ is the LS estimator of the parameter vector and $\hat{\boldsymbol{h}}_{\mathrm{b}} \in \mathbb{C}^{L\times 1}$ is the biased estimator with a shrinkage factor, α is a real-valued variable and $(1+\alpha)$ is defined as the real-valued shrinkage factor that is larger than 0 but smaller than 1 (i.e., $-1 < \alpha < 0$).

Actually, for the parameter estimation scenario, the MMSE estimators with the following expression can also be considered as a biased estimator,

$$\hat{\boldsymbol{h}}_{\mathrm{MMSE}} = \boldsymbol{A}\hat{\boldsymbol{h}}_{\mathrm{LS}}, \tag{28}$$

where $\hat{\boldsymbol{h}}_{\mathrm{MMSE}} \in \mathbb{C}^{L\times 1}$ and $\boldsymbol{A} \in \mathbb{C}^{L\times L}$ is a full-rank matrix. As in [22] and [23], such MMSE channel estimators are developed for MIMO and OFDM systems, respectively. Although these MMSE estimators can achieve a much lower MSE than the LS estimator especially in low SNR regime, they experience much higher complexity than the biased estimator with only one shrinkage factor. In [22], the proposed scaled LS channel estimator can be considered as a biased estimator with only one shrinkage factor, which outperforms the conventional LS estimator while it requires a much lower complexity than the MMSE estimator. The basic idea of the following proposed group-based shrinkage estimator (GSE) is to find a solution with a better tradeoff between the complexity and the performance than the MMSE estimator and the biased estimator with only one shrinkage factor.

The proposed GSE can be expressed as follows:

$$\hat{\boldsymbol{h}}_{\mathrm{b}} = \begin{bmatrix} (1+\alpha_1)\hat{h}_{\mathrm{LS}}(1) \\ \vdots \\ (1+\alpha_1)\hat{h}_{\mathrm{LS}}(\frac{L}{S}) \\ (1+\alpha_2)\hat{h}_{\mathrm{LS}}(\frac{L}{S}+1) \\ \vdots \\ (1+\alpha_2)\hat{h}_{\mathrm{LS}}(\frac{2L}{S}) \\ \vdots \\ (1+\alpha_S)\hat{h}_{\mathrm{LS}}(\frac{(S-1)L}{S}+1) \\ \vdots \\ (1+\alpha_S)\hat{h}_{\mathrm{LS}}(L) \end{bmatrix} = \hat{\boldsymbol{H}}_{\mathrm{LS}}(\mathbf{1}_S + \boldsymbol{\alpha}) \tag{29}$$

where $\hat{\boldsymbol{H}}_{\mathrm{LS}} \in \mathbb{C}^{L\times S}$ is a block diagonal matrix that is constructed from the elements of $\hat{\boldsymbol{h}}_{\mathrm{LS}}$ as well as zeros, S is the number of groups, we define the S-dimensional column vectors $\mathbf{1}_S = [1,1,\ldots,1]^T$ and $\boldsymbol{\alpha} = [\alpha_1,\alpha_2,\ldots,\alpha_S]^T$. The scalar α_s is a real-valued variable and $(1+\alpha_s)$ is defined as the shrinkage factor for the s-th group of coefficients that is larger than 0 but smaller than 1, where $s = 1,2,\ldots,S$. Here we propose to use a uniform group size for the L-dimensional parameter vector, hence the size of each group is L/S. If the length of the parameter vector divided by the group size is not an integer, we can perform zero-padding in the parameter estimation vector to fulfill this requirement. If any statistical knowledge of the parameter vector is given, the group size could be different from each other. But this approach will introduce a higher complexity because we need to select the size of each group and choose a suitable one. In this work, for notational simplicity, we will focus on the low complexity uniform group size approach.

The goal is to minimize the MSE defined by

$$\mathbb{E}\{\|\boldsymbol{h}-\hat{\boldsymbol{h}}_{\mathrm{b}}\|^2\} = \mathbb{E}\{(\hat{\boldsymbol{h}}_{\mathrm{b}}-\mathbb{E}\{\hat{\boldsymbol{h}}_{\mathrm{b}}\})^H(\hat{\boldsymbol{h}}_{\mathrm{b}}-\mathbb{E}\{\hat{\boldsymbol{h}}_{\mathrm{b}}\})\} + \|\mathbb{E}\{\hat{\boldsymbol{h}}_{\mathrm{b}}\}-\boldsymbol{h}\|^2. \tag{30}$$

Note that

$$\hat{\boldsymbol{h}}_{\mathrm{b}} - \mathbb{E}\{\hat{\boldsymbol{h}}_{\mathrm{b}}\} = (\hat{\boldsymbol{H}}_{\mathrm{LS}} - \boldsymbol{H})(\mathbf{1}_S + \boldsymbol{\alpha}) \tag{31}$$

and we have $\hat{\boldsymbol{H}}_{\mathrm{LS}} = \boldsymbol{H} + \boldsymbol{N}$ where

$$\boldsymbol{N} = \begin{bmatrix} \tilde{\boldsymbol{n}}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & \tilde{\boldsymbol{n}}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \ldots & \ldots & \tilde{\boldsymbol{n}}_S \end{bmatrix} \tag{32}$$

and $\tilde{\boldsymbol{n}} = (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H\boldsymbol{n} = [\tilde{\boldsymbol{n}}_1,\tilde{\boldsymbol{n}}_2,\ldots,\tilde{\boldsymbol{n}}_S]^T \in \boldsymbol{C}^{L\times 1}$, assuming that all the elements in this equivalent noise vector

are independent and identically distributed (i.i.d.) random variables. Hence, we have

$$\mathbb{E}\{(\hat{h}_{\mathrm{b}} - \mathbb{E}\{\hat{h}_{\mathrm{b}}\})^H (\hat{h}_{\mathrm{b}} - \mathbb{E}\{\hat{h}_{\mathrm{b}}\})\} = \tilde{\sigma}^2 (1_S + \alpha)^H (1_S + \alpha) \tag{33}$$

Note that $\tilde{\sigma}^2$ equals the variance of the equivalent noise $\tilde{n}$ times the length of the group. We also have

$$\|\mathbb{E}\{\hat{h}_{\mathrm{b}}\} - h\|^2 = \alpha^H H^H H \alpha \tag{34}$$

Finally, the optimal solution of the vector α that minimizes (30) is given by

$$\alpha_{\mathrm{opt}} = -\tilde{\sigma}^2 (\tilde{\sigma}^2 I_S + H^H H)^{-1} 1_S. \tag{35}$$

and we have

$$\hat{h}_{\mathrm{b,opt}} = \hat{H}_{\mathrm{LS}} (1_S + \alpha_{\mathrm{opt}}) \tag{36}$$

Note that this equation is a general expression for different numbers of groups. The complexity of this algorithm is very low because the inverse matrix required to calculate the α_{opt} is a diagonal matrix. Hence, this estimator combined with the conventional RLS algorithm will only introduce an additional complexity that is linear in the length of the parameter vector L and the number of groups S. If the group size equals L, then the GSE converges to the biased estimator with only one shrinkage factor. In the following section, adaptive algorithms will be developed to compute the best GSE with a given group size.

B. Proposed GSE: Adaptive Algorithms

It should be noted that the optimal solution of the biased estimator requires some prior knowledge of the system, which is the matrix $H^H H$ and the scalar term $\tilde{\sigma}^2$ for calculation of the vector α. In addition, the LS channel estimator is also required. The LS channel estimator can be recursively calculated by the RLS adaptive algorithm that is detailed in Section II-B. In this work, we propose LMS-based adaptive algorithms that enable us to estimate the vector α without prior knowledge of the channel and the noise variance. Substituting (34) and (33) into (30) and considering the MSE cost function as a function of α, we can obtain a new cost function

$$f(\alpha) = \tilde{\sigma}^2 (1_S + \alpha)^H (1_S + \alpha) + \alpha^H H^H H \alpha. \tag{37}$$

The gradient of $f(\alpha)$ with respect to α is given by

$$g_\alpha = \tilde{\sigma}^2 (1_S + \alpha) + H^H H \alpha. \tag{38}$$

Note that, because α is a real-valued vector, there is a factor of 2 for this gradient. In what follows, this factor is absorbed into the step size of gradient-type recursions. Hence, the LMS-based update equation of the vector α for the $(i+1)$-th time slot can be expressed as

$$\hat{\alpha}(i+1) = \hat{\alpha}(i) - \mu \hat{g}_\alpha(i), \tag{39}$$

where μ is the step size of the LMS algorithm and the estimated gradient vector is given by

$$\hat{g}_\alpha(i) = \hat{\tilde{\sigma}}^2(i)(1_S + \hat{\alpha}(i)) + \hat{P}_m(i)\hat{\alpha}(i). \tag{40}$$

Here, $\hat{\tilde{\sigma}}^2(i)$ is the estimated equivalent noise variance and the diagonal matrix $\hat{P}_m(i)$ is defined as the estimator of the matrix $H^H H$, the main diagonal vector of this matrix is defined as $\mathrm{diag}[\hat{P}_m(i)] = [\hat{P}_{m,1}(i), \ldots, \hat{P}_{m,S}(i)]$. In this work, we adopt the instantaneous estimator as $\hat{\tilde{\sigma}}^2(i) = \|\hat{h}_{\mathrm{RLS}}(i) - \hat{h}(i)\|^2/S$, where $\hat{h}_{\mathrm{RLS}}(i)$ is the RLS channel estimator and $\hat{h}(i) = \frac{1}{i}\sum_{j=1}^i \hat{h}_{\mathrm{RLS}}(j)$ represent the time averaged channel estimator. Note that $H^H H \in \mathbb{C}^{S \times S}$ is a diagonal matrix with its s-th diagonal element equals $h_{\Sigma,s} = \sum_{l=(s-1)L/S+1}^{sL/S} |h(l)|^2$.

Hence, the elements in the optimal solution can be expressed as $\alpha_{\mathrm{opt},s} = -(1 + h_{\Sigma,s}/\tilde{\sigma}^2)^{-1}$. If we use the matrix $\hat{P}_m(i)$ to replace the matrix $H^H H$, the estimated optimal solution becomes

$$\hat{\alpha}_{\mathrm{opt},s} = -\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2 + \hat{P}_{m,s}(i)}. \tag{41}$$

Recall that we assume that the shrinkage factors for each group are larger than zero but smaller than one. It can be found that if $\hat{P}_{m,s}(i) < 0$, this assumption no longer holds. In addition, if $\hat{P}_{m,s}(i) \to \infty$, the shrinkage factor converges to zero, and the biased estimator actually converges to the unbiased estimator. So we should constrain the values of $\hat{P}_{m,s}(i)$ into the range $(0, +\infty)$.

In order to determine the diagonal matrices $\hat{P}_m(i)$ for each time instant, two approaches are developed in this work. In the first approach, which is named estimator based (GSE-EB) method, the matrices $\hat{P}_m(i)$ are replaced by the diagonal matrices $\hat{H}_{\mathrm{RLS}}^H(i)\hat{H}_{\mathrm{RLS}}(i)$, where $\hat{H}_{\mathrm{RLS}}(i)$ is the RLS estimator of H. Note that, when the number of groups is only one, the GSE-EB method will lead to an optimal shrinkage factor α that has the same expression as the SBME that is proposed in [16]. However, the knowledge of the noise variance is not required in our work. In the second approach, which is named automatic tuning (GSE-AT) method, an LMS-based algorithm is proposed to update the diagonal matrices $\hat{P}_m(i)$.

For the GSE-EB method, the estimation of $\hat{P}_m(i)$ is only determined by the RLS estimator. If the effective spreading codes and the channel information is known, the RLS algorithm can be initialized efficiently. Here, we consider a general scenario where all these quantities are unknown and the initialization of the RLS algorithm is an all zero vector, which means the beginning stage of the RLS algorithm is not very accurate. In order to improve the convergence rate of the proposed GSE schemes, we develop the following GSE-AT algorithm. For each time instant, $\hat{P}_m(i)$ is firstly set to $\hat{H}_{\mathrm{RLS}}^H(i)\hat{H}_{\mathrm{RLS}}(i)$ as in the GSE-EB algorithm, then we consider $\hat{P}_{m,s}(i)$ as the variables of the MSE cost function, where $\hat{P}_{m,s}(i)$ are the diagonal elements of the diagonal matrix $\hat{P}_m(i)$ and $s = 1, 2, \ldots, S$. Then we develop an LMS adaptive recursion to further adapt these values and improve the estimation accuracy for each time instant. Let us reexpress the MSE cost function as shown in (30) as follows. Here, we omit the time index i for simplicity

$$f(\hat{P}_m) = \tilde{\sigma}^2 (1_S + \alpha)^H (1_S + \alpha) + \alpha^H \hat{P}_m \alpha, \tag{42}$$

where $\alpha = -\tilde{\sigma}^2 (\tilde{\sigma}^2 I_S + \hat{P}_m)^{-1} 1_S$ is also a function of $\hat{P}_m$.

Expanding this cost function, we have

$$f([\hat{P}_{m,1}, \ldots, \hat{P}_{m,S}]) = \tilde{\sigma}^2 \sum_{s=1}^{S} \left(1 + \alpha_s\right)^2 + \sum_{s=1}^{S} \alpha_s^2 \hat{P}_{m,s}. \quad (43)$$

Hence, for each group, the corresponding $\hat{P}_{m,s}$ can be obtained by the following equation

$$\hat{P}_{m,s}(c+1) = \hat{P}_{m,s}(c) - \mu_p \hat{g}_s, \quad (44)$$

where c is the iteration index and $\hat{g}_s$ is the estimator of the gradient of the function (43) with respect to $\hat{P}_{m,s}$, which is given by

$$\hat{g}_s = 2\hat{\tilde{\sigma}}^2 \left(1 + \hat{\alpha}_s\right) \frac{\hat{\tilde{\sigma}}^2}{\left(\hat{\tilde{\sigma}}^2 + \hat{P}_{m,s}\right)^2} + 2\hat{\alpha}_s \hat{P}_{m,s} \frac{\hat{\tilde{\sigma}}^2}{\left(\hat{\tilde{\sigma}}^2 + \hat{P}_{m,s}\right)^2} + \hat{\alpha}_s^2. \quad (45)$$

In order to obtain a low complexity solution, for the GSE-AT algorithm, we set the iteration index c to 1, which means for each time instant we only update the values of $\hat{P}_{m,s}$ once. As pointed out previously, the values of $\hat{P}_m(i)$ should be constrained in the range $(0, +\infty)$. Hence, for the GSE-AT algorithm, if the updated values are negative, then we set these values to the same values of GSE-EB algorithm. In Table I, the proposed biased estimators using two approaches to calculate $\hat{P}_m(i)$ are summarized.

Note that the GSE-EB approach (where $\hat{P}_m(i) = \hat{H}_{\text{RLS}}^H(i)\hat{H}_{\text{RLS}}(i)$) requires $2L + 3S + 2$ complex multiplications and $3L + S$ complex additions for one update of the shrinkage factors. For the GSE-AT approach, in which $\hat{P}_m(i)$ is updated by using equation (44), the number of complex multiplications required to update the shrinkage factor is $2L + 13S + 2$ and the number of complex additions required is $3L + 7S$. It will be demonstrated by the simulations that the performance of the GSE-AT approach is better than the GSE-EB approach, while the GSE-EB approach has a lower complexity.

IV. PROPOSED GSE FOR INTERFERENCE SUPPRESSION SCENARIO

A. Proposed GSE: Optimal Solution

For the interference suppression scenario, the biased estimator with a shrinkage factor is given by [13]

$$\hat{w}_{\text{b}} = (1 + \alpha)\hat{w}_{\text{LS}}. \quad (46)$$

where $\hat{w}_{\text{LS}} \in \mathbb{C}^{M \times 1}$ is the LS estimator of the receive filter. The proposed GSE can be expressed as follows:

$$\hat{w}_{\text{b}} = \begin{bmatrix} (1 + \alpha_{w,1})\hat{w}_{\text{LS}}(1) \\ \vdots \\ (1 + \alpha_{w,1})\hat{w}_{\text{LS}}(\frac{L}{S}) \\ (1 + \alpha_{w,2})\hat{w}_{\text{LS}}(\frac{L}{S} + 1) \\ \vdots \\ (1 + \alpha_{w,2})\hat{w}_{\text{LS}}(\frac{2L}{S}) \\ \vdots \\ (1 + \alpha_{w,S})\hat{w}_{\text{LS}}(\frac{(S-1)L}{S} + 1) \\ \vdots \\ (1 + \alpha_{w,S})\hat{w}_{\text{LS}}(L) \end{bmatrix} = \hat{W}_{\text{LS}}(\mathbf{1}_S + \boldsymbol{\alpha}_w) \quad (47)$$

where $\hat{W}_{\text{LS}} \in \mathbb{C}^{M \times S}$ is a block diagonal matrix that is constructed by the elements from $\hat{w}_{\text{LS}}$ and zeros, S is the number of groups. Moreover, we define the S-dimensional column vectors $\mathbf{1}_S = [1, 1, \ldots, 1]^T$ and $\boldsymbol{\alpha}_w = [\alpha_{w,1}, \alpha_{w,2}, \ldots, \alpha_{w,S}]^T$. The quantities $\alpha_{w,s}$ are real-valued variables and $(1 + \alpha_{w,s})$ is defined as the shrinkage factor for the s-th group of coefficients that is larger than zero but smaller than one, where $s = 1, 2, \ldots, S$.

The objective of the GSE is to achieve a smaller MSE than the LS algorithm, which can be expressed as

$$\mathbb{E}[\|\boldsymbol{w}_{\text{o}} - \hat{\boldsymbol{w}}_{\text{b}}\|^2] \leq \mathbb{E}[\|\boldsymbol{w}_{\text{o}} - \hat{\boldsymbol{w}}_{\text{LS}}\|^2]. \quad (48)$$

It should be noted that the problem that we want to solve for both parameter estimation and interference suppression scenarios has a similar form. Hence, we can follow the derivation as shown in Section III-A, and the optimal solution of the parameter vector $\boldsymbol{\alpha}_w$ is given by

$$\boldsymbol{\alpha}_{w,\text{opt}} = -\tilde{\sigma}_w^2 \left(\tilde{\sigma}_w^2 \boldsymbol{I}_S + \boldsymbol{W}^H \boldsymbol{W}\right)^{-1} \mathbf{1}_S. \quad (49)$$

where $\boldsymbol{W} = \mathbb{E}[\hat{\boldsymbol{W}}_{\text{LS}}]$, $\tilde{\sigma}_w^2 = v_w / S$ and we have

$$\hat{\boldsymbol{w}}_{\text{b,opt}} = \hat{\boldsymbol{W}}_{\text{LS}}(\mathbf{1}_S + \boldsymbol{\alpha}_{w,\text{opt}}) \quad (50)$$

B. Proposed GSE: Adaptive Algorithms

In this section, the LMS-based adaptive algorithms are proposed to estimate the vector $\boldsymbol{\alpha}_w$. First, we consider the MSE cost function as a function of $\boldsymbol{\alpha}_w$, i.e.,

$$f(\boldsymbol{\alpha}_w) = \tilde{\sigma}_w^2 (\mathbf{1}_S + \boldsymbol{\alpha}_w)^H (\mathbf{1}_S + \boldsymbol{\alpha}_w) + \boldsymbol{\alpha}_w^H \boldsymbol{W}^H \boldsymbol{W} \boldsymbol{\alpha}_w. \quad (51)$$

The gradient of $f(\boldsymbol{\alpha}_w)$ with respect to $\boldsymbol{\alpha}_w^*$ is given by $\boldsymbol{g}_w = \tilde{\sigma}_w^2 (\mathbf{1}_S + \boldsymbol{\alpha}_w) + \boldsymbol{W}^H \boldsymbol{W} \boldsymbol{\alpha}_w$. Hence, the LMS-based update equation of the vector $\boldsymbol{\alpha}_w$ for the $(i+1)$-th time slot can be expressed as

$$\hat{\boldsymbol{\alpha}}_w(i+1) = \hat{\boldsymbol{\alpha}}_w(i) - \mu_w \hat{\boldsymbol{g}}_w(i), \quad (52)$$

where μ_w is the step size of the LMS algorithm and the estimated gradient vector is given by

$$\hat{\boldsymbol{g}}_w(i) = \hat{\tilde{\sigma}}_w^2(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}_w(i)) + \hat{\boldsymbol{P}}_w(i)\hat{\boldsymbol{\alpha}}_w(i), \quad (53)$$

where $\hat{\tilde{\sigma}}_w^2(i)$ is the estimated equivalent noise variance and the diagonal matrix $\hat{\boldsymbol{P}}_w(i) \in \mathbb{C}^{S \times S}$ is defined as the estimator of

TABLE I
GSE FOR SCE-RLS IN SC-FDE DS-UWB SYSTEMS

Proposed GSE-EB	Proposed GSE-AT
1. Initialization: $\hat{\boldsymbol{\alpha}}(1) = \mathbf{0} \in \mathbb{C}^{S \times 1}$ Set value of μ	**1. Initialization:** $\hat{\boldsymbol{\alpha}}(1) = \mathbf{0} \in \mathbb{C}^{S \times 1}$ Set values of μ and μ_p
2. Calculate the biased estimator: For $i = 1, 2, \ldots$ $\hat{\boldsymbol{h}}_{\mathrm{b}}(i) = \hat{\boldsymbol{H}}_{\mathrm{RLS}}(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}(i))$	**2. Calculate the biased estimator:** For $i = 1, 2, \ldots$ $\hat{\boldsymbol{h}}_{\mathrm{b}}(i) = \hat{\boldsymbol{H}}_{\mathrm{RLS}}(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}(i))$
3. Calculate the shrinkage factor: $\hat{\boldsymbol{h}}(i) = \frac{1}{i}\sum_{j=1}^{i}\hat{\boldsymbol{h}}_{\mathrm{RLS}}(j)$ $\hat{\bar{\sigma}}^2(i) = \|\hat{\boldsymbol{h}}_{\mathrm{RLS}}(i) - \hat{\boldsymbol{h}}(i)\|^2/S$ $\hat{\boldsymbol{P}}_m(i) = \hat{\boldsymbol{H}}_{\mathrm{RLS}}^H(i)\hat{\boldsymbol{H}}_{\mathrm{RLS}}(i)$ $\hat{\boldsymbol{g}}_{\boldsymbol{\alpha}}(i) = \hat{\bar{\sigma}}^2(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}(i)) + \hat{\boldsymbol{P}}_m(i)\hat{\boldsymbol{\alpha}}(i),$ $\hat{\boldsymbol{\alpha}}(i+1) = \hat{\boldsymbol{\alpha}}(i) - \mu\hat{\boldsymbol{g}}_{\boldsymbol{\alpha}}(i).$	**3. Calculate the shrinkage factor:** $\hat{\boldsymbol{h}}(i) = \frac{1}{i}\sum_{j=1}^{i}\hat{\boldsymbol{h}}_{\mathrm{RLS}}(j)$ $\hat{\bar{\sigma}}^2(i) = \|\hat{\boldsymbol{h}}_{\mathrm{RLS}}(i) - \hat{\boldsymbol{h}}(i)\|^2/S$ $\hat{\boldsymbol{P}}_m(i) = \hat{\boldsymbol{H}}_{\mathrm{RLS}}^H(i)\hat{\boldsymbol{H}}_{\mathrm{RLS}}(i)$ For $s = 1, 2, \ldots, S$ $\hat{P}_{m,s}(i) = \hat{P}_{m,s}(i) - \mu_p \hat{g}_s(i)$, where $\hat{g}_s(i)$ is given in (45). If $\hat{P}_{m,s}(i) < 0$, set $\hat{\boldsymbol{P}}_m(i) = \hat{\boldsymbol{H}}_{\mathrm{RLS}}^H(i)\hat{\boldsymbol{H}}_{\mathrm{RLS}}(i)$ break; End For. $\hat{\boldsymbol{g}}_{\boldsymbol{\alpha}}(i) = \hat{\bar{\sigma}}^2(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}(i)) + \hat{\boldsymbol{P}}_m(i)\hat{\boldsymbol{\alpha}}(i),$ $\hat{\boldsymbol{\alpha}}(i+1) = \hat{\boldsymbol{\alpha}}(i) - \mu\hat{\boldsymbol{g}}_{\boldsymbol{\alpha}}(i).$

the matrix $\boldsymbol{W}^H\boldsymbol{W}$, the main diagonal vector of this matrix is defined as $\mathrm{diag}[\hat{\boldsymbol{P}}_w(i)] = [\hat{P}_{w,1}(i), \ldots, \hat{P}_{w,S}(i)]$. The instantaneous estimator of $\hat{\tilde{\sigma}}_w^2(i)$ is given by $\hat{\tilde{\sigma}}_w^2(i) = (\hat{\boldsymbol{w}}_{\mathrm{RLS}}(i) - \boldsymbol{w}_{\mathrm{o}}(i))^H(\hat{\boldsymbol{w}}_{\mathrm{RLS}}(i) - \boldsymbol{w}_{\mathrm{o}}(i))$, where $\boldsymbol{w}_{\mathrm{o}}$ is replaced by the time averaged RLS estimator, that is $\boldsymbol{w}_{\mathrm{o}}(i) = \frac{1}{i}\sum_{j=1}^{i}\hat{\boldsymbol{w}}_{\mathrm{RLS}}(j)$.

In order to determine the diagonal matrices $\hat{\boldsymbol{P}}_w(i)$ for each time instant, the GSE-EB method and the GSE-AT method are developed in the interference suppression scenario. In the GSE-EB approach, the matrices $\hat{\boldsymbol{P}}_w(i)$ are replaced by the diagonal matrices $\hat{\boldsymbol{W}}_{\mathrm{RLS}}^H(i)\hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)$. However, because we assume that the initialization of the RLS algorithm is an all zero vector, the beginning stage of the RLS algorithm is not very accurate. Hence, in order to improve the convergence rate of the proposed GSE schemes, we develop the GSE-AT algorithm. For each time instant, the matrix $\hat{\boldsymbol{P}}_w(i)$ is firstly set to $\hat{\boldsymbol{W}}_{\mathrm{RLS}}^H(i)\hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)$ as in the GSE-EB algorithm, then we consider $\hat{P}_{w,s}(i)$ as the variables of the MSE cost function, where $\hat{P}_{w,s}(i)$ are the diagonal elements of the diagonal matrix $\hat{\boldsymbol{P}}_w(i)$ and $s = 1, 2, \ldots, S$. Then we develop an LMS adaptive equation to further adapt these values and improve the estimation accuracy for each time instant. Here, we omit the time index i for simplicity and have

$$f([\hat{P}_{w,1}, \ldots, \hat{P}_{w,S}]) = \tilde{\sigma}_w^2 \sum_{s=1}^{S}(1 + \alpha_{w,s})^2 + \sum_{s=1}^{S}\alpha_{w,s}^2\hat{P}_{w,s}. \tag{54}$$

Hence, for each group, the corresponding $\hat{P}_{w,s}$ can be updated by the following equation

$$\hat{P}_{w,s}(c+1) = \hat{P}_{w,s}(c) - \mu_s \hat{g}_{w,s}, \tag{55}$$

where c is the iteration index and $\hat{g}_{w,s}$ is the estimator of the gradient of the cost function with respect to $\hat{P}_{m,s}$, which is given by

$$\hat{g}_{w,s} = 2\hat{\tilde{\sigma}}^2(1+\hat{\alpha}_s)\frac{\hat{\tilde{\sigma}}^2}{(\hat{\tilde{\sigma}}^2 + \hat{P}_{m,s})^2} + 2\hat{\alpha}_s\hat{P}_{m,s}\frac{\hat{\tilde{\sigma}}^2}{(\hat{\tilde{\sigma}}^2 + \hat{P}_{m,s})^2} + \hat{\alpha}_s^2. \tag{56}$$

In order to obtain a low complexity solution, we set the iteration index c to 1 for the GSE-AT algorithm. Note that if the updated values in the GSE-AT algorithm become negative, then we set these values to the same as obtained in the GSE-EB algorithm. In Table II, the proposed biased estimators with two approaches to calculate $\hat{P}_{\mathrm{w}}(i)$ are summarized.

Note that the GSE-EB approach (where $\hat{P}_{\mathrm{w}}(i) = \hat{\boldsymbol{W}}_{\mathrm{RLS}}^H(i)\hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)$), requires $2M + 3S + 2$ complex multiplications and $3M + S$ complex additions for one update of the shrinkage factors. For the GSE-AT approach, in which $\hat{P}_{\mathrm{w}}(i)$ is updated by using equation (55), the number of complex multiplications required to update the shrinkage factor is $2M + 13S + 2$ and the number of complex additions required is $3M + 7S$. It will be demonstrated by the simulations that the performance of the GSE-AT approach is better than the GSE-EB approach, while the GSE-EB approach has a lower complexity.

V. MSE ANALYSIS

In this section, we will analyze the MSE performance of the proposed GSE. Since the proposed GSE has similar forms in the parameter estimation scenario and the interference suppression scenario, we carried out the following derivations based on the parameter estimation scenario. Firstly, we will prove that the minimum MSE obtained by the GSE schemes will always be smaller or equal to the MSE that can be achieved by minimum variance unbiased estimator (MVUE) such as the LS estimator. Then the MSE lower bounds of the GSE schemes will be derived. In addition, we will prove that when the numbers of groups is larger than or equal to two, the MSE lower bound will always be lower than the biased estimator with only one shrinkage factor (when the number of groups S equals one).

A. MMSE Comparison

Assuming AWGN with zero mean and variance σ_n^2, the LS estimator is a minimum variance unbiased estimator. The MSE

TABLE II
GSE FOR FREQUENCY DOMAIN RECEIVER IN SC-FDE DS-UWB SYSTEMS

Proposed GSE-EB	Proposed GSE-AT
1. Initialization:	**1. Initialization:**
$\hat{\boldsymbol{\alpha}}_w(1) = \mathbf{0} \in \mathbb{C}^{S \times 1}$ Set value of μ_w	$\hat{\boldsymbol{\alpha}}_w(1) = \mathbf{0} \in \mathbb{C}^{S \times 1}$ Set values of μ_w and μ_s
2. Calculate the biased estimator:	**2. Calculate the biased estimator:**
For $i = 1, 2, \ldots$ $\hat{\boldsymbol{w}}_{\mathrm{b}}(i) = \hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}_w(i))$	For $i = 1, 2, \ldots$ $\hat{\boldsymbol{w}}_{\mathrm{b}} = \hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}_w(i))$
3. Calculate the shrinkage factor:	**3. Calculate the shrinkage factor:**
$\hat{\boldsymbol{w}}(i) = \frac{1}{i} \sum_{j=1}^{i} \hat{\boldsymbol{w}}_{\mathrm{RLS}}(j)$ $\hat{\tilde{\sigma}}_w^2(i) = \|\hat{\boldsymbol{w}}_{\mathrm{RLS}}(i) - \hat{\boldsymbol{w}}(i)\|^2/S$ $\hat{\boldsymbol{P}}_w(i) = \hat{\boldsymbol{W}}_{\mathrm{RLS}}^{H}(i)\hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)$ $\hat{\boldsymbol{g}}_w(i) = \hat{\tilde{\sigma}}_w^2(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}_w(i)) + \hat{\boldsymbol{P}}_w(i)\hat{\boldsymbol{\alpha}}_w(i),$ $\hat{\boldsymbol{\alpha}}_w(i+1) = \hat{\boldsymbol{\alpha}}_w(i) - \mu_w \hat{\boldsymbol{g}}_w(i).$	$\hat{\boldsymbol{w}}(i) = \frac{1}{i} \sum_{j=1}^{i} \hat{\boldsymbol{w}}_{\mathrm{RLS}}(j)$ $\hat{\tilde{\sigma}}_w^2(i) = \|\hat{\boldsymbol{w}}_{\mathrm{RLS}}(i) - \hat{\boldsymbol{w}}(i)\|^2/S$ $\hat{\boldsymbol{P}}_w(i) = \hat{\boldsymbol{W}}_{\mathrm{RLS}}^{H}(i)\hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)$ For $s = 1, 2, \ldots, S$ $\hat{P}_{w,s}(i) = \hat{P}_{w,s}(i) - \mu_s \hat{g}_{w,s}(i)$, where $\hat{g}_{w,s}(i)$ is given in (56). If $\hat{P}_{w,s}(i) < 0$, set $\hat{\boldsymbol{P}}_w(i) = \hat{\boldsymbol{W}}_{\mathrm{RLS}}^{H}(i)\hat{\boldsymbol{W}}_{\mathrm{RLS}}(i)$ break; End For. $\hat{\boldsymbol{g}}_w(i) = \hat{\tilde{\sigma}}_w^2(i)(\mathbf{1}_S + \hat{\boldsymbol{\alpha}}_w(i)) + \hat{\boldsymbol{P}}_w(i)\hat{\boldsymbol{\alpha}}_w(i),$ $\hat{\boldsymbol{\alpha}}_w(i+1) = \hat{\boldsymbol{\alpha}}_w(i) - \mu_w \hat{\boldsymbol{g}}_w(i).$

for the LS estimator is

$$\mathbb{E}\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{LS}}\|^2\} = \mathbb{E}\{(\hat{\boldsymbol{h}}_{\mathrm{LS}} - \boldsymbol{h})^H (\hat{\boldsymbol{h}}_{\mathrm{LS}} - \boldsymbol{h})\} = var(\boldsymbol{h}, \hat{\boldsymbol{h}}_{\mathrm{LS}}). \tag{57}$$

Defining $v = var(\boldsymbol{h}, \hat{\boldsymbol{h}}_{\mathrm{LS}})$, we have

$$v = tr\{\sigma_n^2 (\boldsymbol{X}^H \boldsymbol{X})^{-1}\} \tag{58}$$

It should be noted that we can also express the LS estimator as $\hat{\boldsymbol{h}}_{\mathrm{LS}} = \boldsymbol{h} + \tilde{\boldsymbol{n}}$, hence, we have that the variance of the elements in the equivalent noise $\tilde{\boldsymbol{n}}$ is $\tilde{\sigma}_n^2 = v/L$.

Recall that the target of the biased estimation that is to reduce the MSE introduced by $\hat{\boldsymbol{h}}_{\mathrm{LS}}$. The objective is to obtain a biased estimator that results in

$$\mathbb{E}\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{b}}\|^2\} \leq \mathbb{E}\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{LS}}\|^2\}. \tag{59}$$

Recall the equations (33) and (34), the objective becomes

$$\tilde{\sigma}^2 (\mathbf{1}_S + \boldsymbol{\alpha})^H (\mathbf{1}_S + \boldsymbol{\alpha}) + \boldsymbol{\alpha}^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{\alpha} \leq v. \tag{60}$$

Since $\tilde{\sigma}_n^2 = v/L$, we have $\tilde{\sigma}^2 = (L/S)\tilde{\sigma}_n^2 = v/S$. In appendix A, we prove that this objective is always fulfilled with the optimal solution $\boldsymbol{\alpha}_{\mathrm{opt}}$ as shown in (35).

B. MSE Lower Bound and the Effect of the Number of Groups

It should be noted that a lower bound of the MSE performance of the proposed GSE schemes that corresponds to the optimal $\boldsymbol{\alpha}$ can be obtained as

$$\mathbb{E}\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{b}}\|^2\} \geq (S + \sum_{s=1}^{S} \alpha_{\mathrm{opt},s}) \frac{v}{S} = v - \frac{\sum_{s=1}^{S} (\frac{v^2}{v + h_{\Sigma,s}S})}{S}$$
$$= v - \sum_{s=1}^{S} \frac{v^2}{Sv + h_{\Sigma,s}S^2}. \tag{61}$$

Since the second term on the right hand-side is non-negative, it can be concluded that the MSE lower bound will always be smaller than or equal to the unbiased Cramér-Rao Lower

Bound (CRLB), which is expressed as v in this equation. Note that this expression can be considered as the relationship between the MSE lower bound of the GSE and the unbiased CRLB. In the case where the number of groups equals 1, we have $S = 1$, $\tilde{\sigma}^2 = v$ and $\alpha_{\mathrm{opt}} = -v/(v + \|\boldsymbol{h}\|^2)$. The lower bound becomes

$$\mathbb{E}\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{b}}\|^2\} \geq \frac{1}{1 + v/\|\boldsymbol{h}\|^2} v. \tag{62}$$

In this case, the one-group GSE scheme converges to our previously proposed shrinkage factor biased estimator [13], [14].

For the proposed GSE schemes, we prove the following statements:

Statement 1: The MSE lower bound as shown in (61) with $S > 1$ will always be lower than or equal to the lower bound for $S = 1$. This statement indicates that our proposed GSE outperforms the biased estimator with only one shrinkage factor. The proof is detailed in appendix B.

Statement 2: The lowest MSE lower bound as shown in (61) can be obtained in $S = L$ case, where L is the length of the parameter vector to be estimated. This statement indicates that the optimal performance can be obtained with the largest possible group number. The proof is detailed in Appendix B.

The performance of the algorithm depends on the number of groups and the scenario. If some a priori knowledge of the parameter vector to be estimated is available, then a possible extension of the GSE is to develop a method to determine the size of the groups. For example, the knowledge of the expected value of the number of clusters of a UWB channel might enable a more attractive tradeoff between the performance and the complexity. Moreover, the increase in the number of groups can improve the MSE performance. However, this comes with diminishing returns and an increase in the computational complexity.

VI. SIMULATIONS

In this section, the proposed GSE estimators are employed in the SCE and in the design of the frequency domain receiver of a synchronous downlink block-by-block transmission binary phase shift keying (BPSK) DS-UWB system that are detailed in Section II-B and Section II-C, respectively. Their MSE performances are compared with the conventional RLS adaptive algorithms. The pulse shape adopted is the root-raised-cosine (RRC) pulse with the pulse-width $T_c = 0.375$ ns. The length of the data block is set to $N = 32$ symbols. The Walsh spreading code with a spreading gain $N_c = 8$ is generated for the simulations and we assume that the maximum number of active users is 7. The channel has been simulated according to the standard IEEE 802.15.4a channel model for the NLOS indoor environment as shown in [25]. We assume that the channel is constant during the whole transmission and the time domain channel impulse response has 100 taps. The CP guard interval has a length of 35 chips, which has the equivalent length of 105 samples and it is enough to eliminate the IBI. The uncoded data rate of the transmission is approximately 293 Mbps. For all the simulations, the adaptive receivers/estimators are initialized as null vectors. All the curves are obtained by averaging 200 channel realizations. In Fig. 2 - Fig. 5, the performance of the proposed GSE in the parameter estimation scenario (SCE as an example) is shown. In this scenario, the GSE is employed to improve the MSE performance of the channel estimation. In Fig. 6 to Fig. 8, the performance of the proposed GSE in the interference suppression scenario (frequency domain receiver design as an example) is presented. In this scenario, the GSE schemes accelerate the convergence speed of the adaptive algorithm.
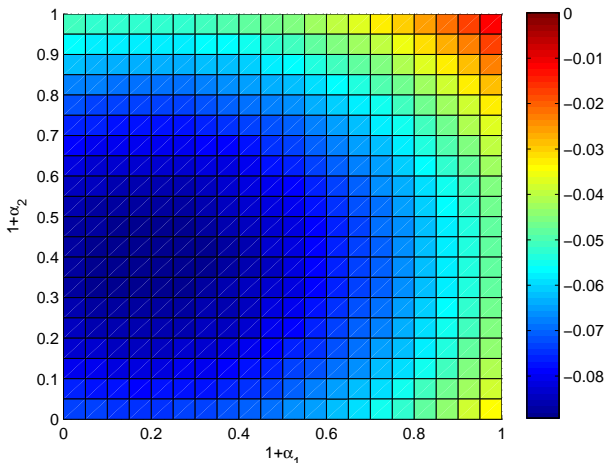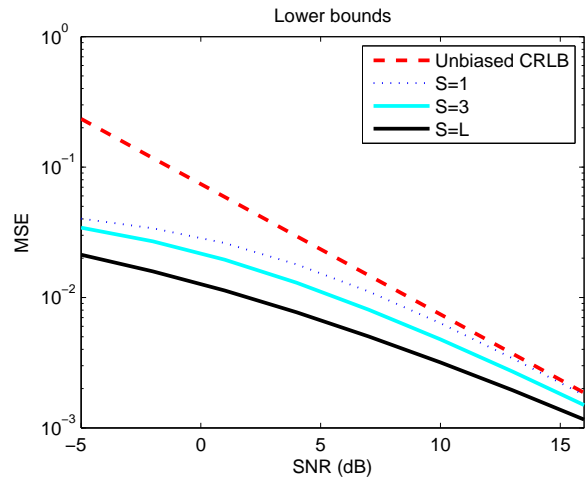

Fig. 3. MSE lower bounds for the proposed GSE with different numbers of groups (SCE).
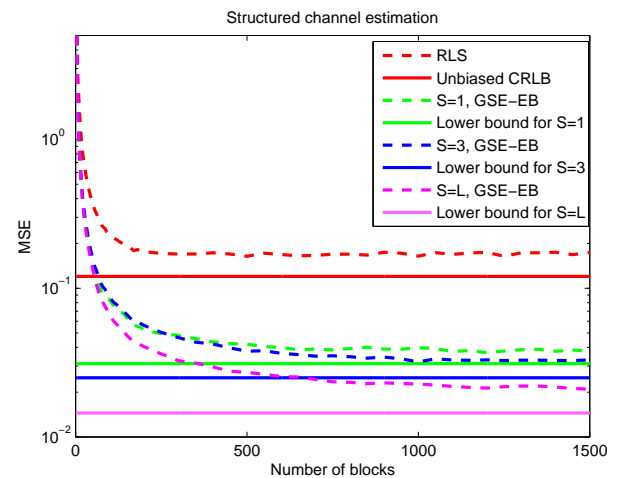

Fig. 4. MSE performance with different numbers of groups (SCE). Parameters used: RLS: $\lambda = 0.998$, $\delta = 10$. Proposed GSE EB: $\mu = 0.075$.

Fig. 2. Surface of the MSE difference in a scenario with $S = 2$ (SCE). The MSE difference is defined as $\delta\mathrm{MSE} = \|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{b}}\|^2 - \|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{RLS}}\|^2$.

First, we examine the MSE difference between the unbiased estimator and the proposed GSE as a function of the shrinkage factors for each group in a single user system with 0 dB SNR. In Fig. 2, the surface defined as $\delta\mathrm{MSE} = \|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{b}}\|^2 - \|\boldsymbol{h} - \hat{\boldsymbol{h}}_{\mathrm{RLS}}\|^2$ in the range of the shrinkage factors between 0 to 1 is shown. Note that, when the bias equals

zero, which corresponding to the point $[1, 1]$ in the figure, the MSE difference equals zero. The optimal solution is located at $[1 + \alpha_1, 1 + \alpha_2] = [0.15, 0.35]$. For this channel realization, the optimal solution that is obtained by our algorithm after transmitting 1000 data blocks is reported as $[1 + \alpha_1, 1 + \alpha_2] = [0.16, 0.35]$, which is very close to the optimal solution shown in this figure.

In Fig. 3, the MSE lower bounds for the proposed GSE schemes with different numbers of groups are shown as a function of SNR. The proposed biased estimators show a better performance than the conventional unbiased estimator. As we have proved, the best performance can be obtained in case of $S = L$. The proposed GSE schemes can maintain the MSE gain for low and medium SNR regimes. In high SNR scenarios with a small noise variance, the biased estimator converges to the unbiased estimator and the gain becomes smaller. Compared with the biased estimators with only one shrinkage factor (which is equivalent to the case $S = 1$), the GSE schemes with a number of groups can maintain the MSE gain until higher SNR regime.
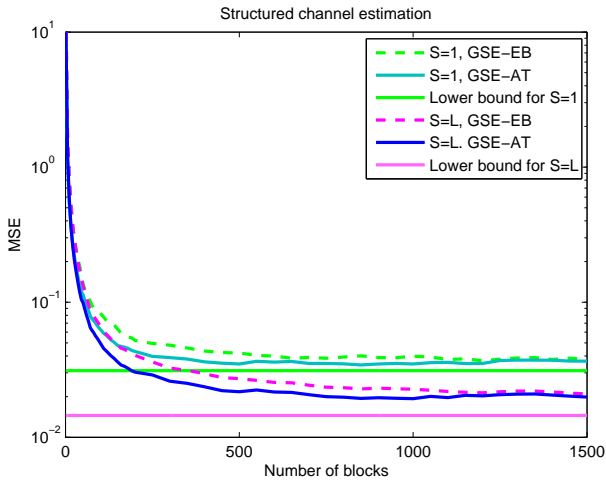
In the third experiment, we examine the proposed GSE

Fig. 5. MSE performance of the proposed GSE-EB and GSE-AT algorithms with $S = 1$ and $S = L$ (SCE). Parameters used: Proposed GSE-EB: $\mu = 0.075$. Proposed GSE-AT: $\mu = 0.075$, $\mu_p = 0.05$.
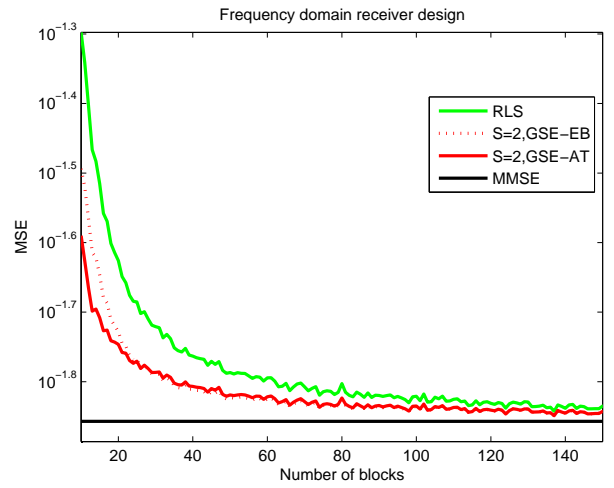

Fig. 7. Normalized MSE performance ($\|\boldsymbol{b} - \hat{\boldsymbol{b}}\|^2 / \|\boldsymbol{b}\|^2$) of the biased estimators with 5 users in 3 dB SNR.
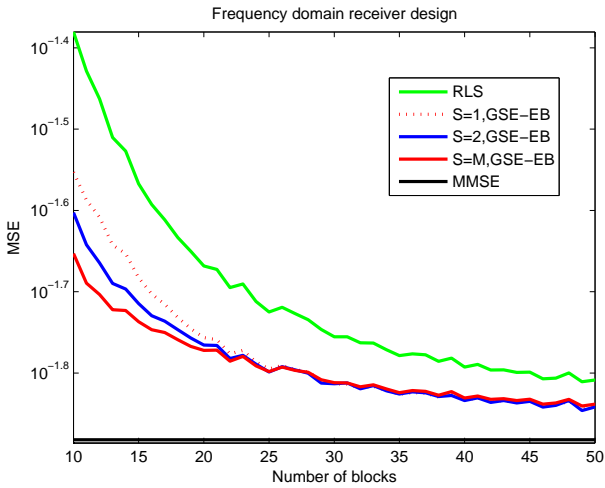
Fig. 6. Normalized MSE performance ($\|\boldsymbol{b} - \hat{\boldsymbol{b}}\|^2 / \|\boldsymbol{b}\|^2$) of the biased estimator with 5 users in 5 dB SNR.

Fig. 8. Coded BER performance of the GSE schemes in interference suppression scenario with 5 users.

schemes with different numbers of groups for the SCE in a single-user system with 0 dB SNR. In Fig. 4, the MSE performance of the channel estimators are compared as a function of the number of blocks transmitted. The RLS algorithm approaches the unbiased CRLB while the proposed biased estimators approach the lower bounds as given in (61). The biased estimators converge faster than the RLS algorithm and the steady-state performance is also improved. Note that the additional complexity to employ the proposed biased estimation techniques increases linearly with the product of the number of groups and the length of the channel.

In Fig. 5, the proposed GSE-EB and GSE-AT algorithms are compared in $S = 1$ and $S = L$ scenarios in a single user system with 0 dB SNR. It can be found that the GSE-AT algorithm can provide a noticeable gain over the GSE-EB algorithm especially at the beginning stage of transmission. This is because the GSE-AT algorithm allows us to further adjust the diagonal matrix $\hat{\boldsymbol{P}}_m(i)$ for each time instant. At
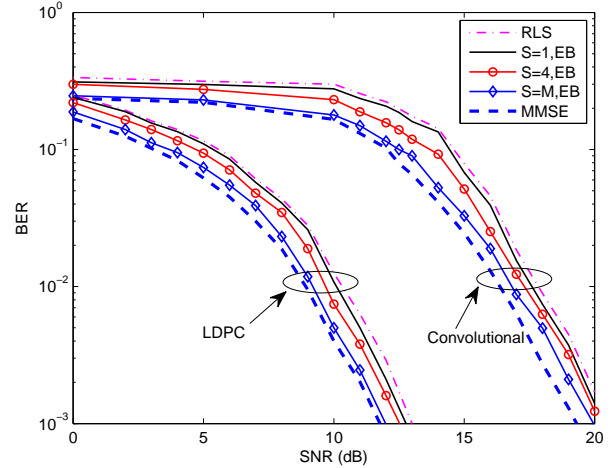
the beginning stage when the RLS algorithm does not very accurately estimate the channel, the GSE-AT algorithm can be used to improve the convergence rate.

In Fig. 6, the performance of the proposed GSE is shown in the interference suppression scenario (frequency domain receiver design as an example) with a short training sequence. In this simulation, 50 training blocks are transmitted in a scenario with 5 users with 5 dB SNR. The proposed GSE-EB algorithm outperforms the RLS algorithm and the best performance is obtained by setting $S = M$. In Fig. 7, we compare the GSE-EB and GSE-AT algorithms in 5-user communications and an SNR of 3 dB. In this experiment, 150 training blocks are transmitted. The GSE-AT algorithm can further accelerate the convergence rate of the GSE-EB algorithm. At the beginning stage, the GSE-AT algorithm introduces the best performance.

In Fig. 8, the bit error rate (BER) performance of the proposed GSE with different numbers of groups are shown in a scenario with 5 users. The coded BER performance is

obtained by adopting a convolutional code and an LDPC code [30] designed according to the PEG approach [30]. For the convolutional code, the constraint length is 5, the rate is 2/3 and the code polynomial is [7,5,5]. For the LDPC code, the rate is 1/2 and the code length is 200 bits. The maximum number of iterations is set to 20. In this experiment, 100 training blocks are transmitted followed by 400 data blocks. With the convolutional code, the proposed GSEs perform better than the RLS algorithm and the maximum gains are obtained in the medium SNR range from 12 dB to 17 dB. By employing the LDPC code, a BER of around 10^{-3} is achieved at 12 dB. The proposed GSEs with LDPC codes also outperform the RLS algorithm and the maximum gains are obtained for an SNR around 7 dB.

VII. CONCLUSIONS

In this work, a novel biased estimation algorithm called group-based shrinkage estimator (GSE) is proposed, which divides the target parameter vector into a number of groups and calculates one shrinkage factor for each group. Adaptive algorithms are developed for the GSE scheme in the parameter estimation and the interference suppression scenarios. The incorporation of the proposed estimators has been considered in the frequency-domain of DS-UWB systems, where structured channel estimation and the receiver designs are considered as examples of the parameter estimation scenario and interference suppression scenario, respectively. An MSE analysis is presented that indicates the lower bound of the proposed GSE schemes. The relationship between the lower bound and the number of groups are also established. It has been proved that the GSE provides a better performance than the biased estimators with only one shrinkage factor. In addition, the lowest MSE lower bound can be obtained in the $S = L$ case. As for future research directions, the GSE scheme can be developed in different systems and scenarios. In addition, if we have some prior knowledge of the target parameter vector, we can then divide it into groups with different sizes and find more attractive tradeoffs between the computational complexity and the performance.

APPENDIX A
MMSE COMPARISON

In order to check if the objective as shown in (60) is fulfilled with $\boldsymbol{a}_{\text{opt}}$, the equation (35) is rearranged by taking the expression of the diagonal matrix $\boldsymbol{H}^H \boldsymbol{H} \in \mathbb{C}^{S \times S}$ into account. We have

$$\boldsymbol{\alpha}_{\text{opt}} = [\alpha_{\text{opt},1}, \ldots, \alpha_{\text{opt,S}}]^T = -\tilde{\sigma}^2 (\tilde{\sigma}^2 \boldsymbol{I} + \boldsymbol{H}^H \boldsymbol{H})^{-1} \boldsymbol{1}_S$$
$$= -[(1 + h_{\Sigma,1}/\tilde{\sigma}^2)^{-1}, (1 + h_{\Sigma,2}/\tilde{\sigma}^2)^{-1}, \ldots, (1 + h_{\Sigma,S}/\tilde{\sigma}^2)^{-1}]^T \tag{63}$$

where $h_{\Sigma,s} = \sum_{i=(s-1)L/S+1}^{sL/S} |h(i)|^2$. Hence, we have the following expressions

$$\tilde{\sigma}^2 (\boldsymbol{1}_S + \boldsymbol{\alpha})^H (\boldsymbol{1}_S + \boldsymbol{\alpha}) = \tilde{\sigma}^2 \sum_{s=1}^{S} (1 + \alpha_{\text{opt,s}})^2, \tag{64}$$

$$\boldsymbol{\alpha}^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{\alpha} = \sum_{s=1}^{S} \alpha_{\text{opt,s}}^2 h_{\Sigma,s}. \tag{65}$$

By substituting (64) and (65) into the MSE expression (60) and bearing in mind that $S\tilde{\sigma}^2 = v$, the inequality that we want to prove becomes

$$\tilde{\sigma}^2 \sum_{s=1}^{S} (1 + \alpha_{\text{opt,s}})^2 + \sum_{s=1}^{S} \alpha_{\text{opt,s}}^2 h_{\Sigma,s} - S\tilde{\sigma}^2 \leq 0. \tag{66}$$

Note that the left hand-side of (66) can be expressed as

$$\sum_{s=1}^{S} \alpha_{\text{opt,s}}^2 + \sum_{s=1}^{S} 2\alpha_{\text{opt,s}} + \sum_{s=1}^{S} \alpha_{\text{opt,s}}^2 h_{\Sigma,s}/\tilde{\sigma}^2$$
$$= \sum_{s=1}^{S} \alpha_{\text{opt,s}}^2 (1 + h_{\Sigma,s}/\tilde{\sigma}^2) + \sum_{s=1}^{S} 2\alpha_{\text{opt,s}}. \tag{67}$$

By recalling that $\alpha_{\text{opt,s}} = -(1 + h_{\Sigma,s}/\tilde{\sigma}^2)^{-1}$ is always a non-positive scalar value, the left-hand side of the equation becomes

$$\sum_{s=1}^{S} (1 + h_{\Sigma,s}/\tilde{\sigma}^2)^{-1} + \sum_{s=1}^{S} 2\alpha_{\text{opt,s}} = \sum_{s=1}^{S} \alpha_{\text{opt,s}}. \tag{68}$$

As $\alpha_{\text{opt,s}}$ are always non-positive, the summation on the right hand-side of (68) will always be smaller or equal to zero, which completes the proof.

APPENDIX B
PROOFS OF THE STATEMENTS

First, we want to prove that the lower bound of MSE as shown in (61) with $S > 1$ will always be lower than or equal to the $S = 1$ case. Based on the equation (61), we can focus on the following function

$$f(S) = \sum_{s=1}^{S} \frac{1}{Sv + h_{\Sigma,s} S^2} = \frac{1}{S^2} \sum_{s=1}^{S} \frac{1}{(v/S) + h_{\Sigma,s}}, \tag{69}$$

where $h_{\Sigma,s} = \sum_{i=(s-1)L/S+1}^{sL/S} |h(i)|^2$ The task now is equivalent to proving that $f(S) \geq f(1)$ for all the possible values of S. Note that $\sum_{s=1}^{S} h_{\Sigma,s} = \|\boldsymbol{h}\|^2$ and

$$f(1) = \frac{1}{v + \|\boldsymbol{h}\|^2} = \frac{1}{v + \sum_{s=1}^{S} h_{\Sigma,s}} = \frac{1}{\sum_{s=1}^{S} (v/S + h_{\Sigma,s})}. \tag{70}$$

Hence, the relation becomes

$$\sum_{s=1}^{S} \frac{1}{(v/S) + h_{\Sigma,s}} \geq \frac{S^2}{\sum_{s=1}^{S} (v/S + h_{\Sigma,s})}. \tag{71}$$

Actually, we can express this problem as the following mathematical problem:

$$\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_S} \geq \frac{S^2}{\sum_{s=1}^{S} a_s}, \tag{72}$$

where S is a positive integer and $a_s = (v/S) + h_{\Sigma,s}$ are all non-negative values. This inequality can be proved by using the mathematical induction as follows:

For $S = 1$, the left-hand side and the right-hand side are both equal to $\frac{1}{a_1}$.

For $S = 2$, the left-hand side equals $\frac{1}{a_1} + \frac{1}{a_2} = \frac{a_1+a_2}{a_1 a_2}$ and the right-hand side equals $\frac{4}{a_1+a_2}$. Since $(a_1 + a_2)^2 - 4a_1 a_2 = (a_1 - a_2)^2 \geq 0$, the inequality holds.

Assuming the inequality holds for $S = n$, where $n \geq 2$, we have

$$\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n} \geq \frac{n^2}{\sum_{s=1}^{n} a_s}. \tag{73}$$

For $S = n + 1$, we first consider the left-hand side as follows

$$\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n} + \frac{1}{a_{n+1}} \geq \frac{n^2}{\sum_{s=1}^{n} a_s} + \frac{1}{a_{n+1}} \\ = \frac{n^2 a_{n+1} + \sum_{s=1}^{n} a_s}{\left(\sum_{s=1}^{n} a_s\right) \cdot a_{n+1}}. \tag{74}$$

and then, the right-hand side is given by

$$\frac{(n+1)^2}{\sum_{s=1}^{n} a_s + a_{n+1}}. \tag{75}$$

Because

$$\left(\sum_{s=1}^{n} a_s + a_{n+1}\right)\left(n^2 a_{n+1} + \sum_{s=1}^{n} a_s\right) - (n+1)^2 \left(\sum_{s=1}^{n} a_s\right) \cdot a_{n+1}$$

$$= \left(\sum_{s=1}^{n} a_s - n \cdot a_{n+1}\right)^2 \geq 0,$$

the inequality also holds for $S = n + 1$. This completes the proof.

Now, we can prove the second statement which points out that the lowest MSE lower bound as shown in (61) can be obtained when the numbers of groups is equal to the length of the parameter vector to be estimated. Following the proof of of $Statement$ 1, we can express the problem that needs to be solved as: prove that $f(L) \geq f(S)$, for any possible values of number of groups S, and mathematically, we need to prove that

$$\frac{1}{L^2} \sum_{i=1}^{L} \frac{1}{(v/L) + |h(i)|^2} \geq \frac{1}{S^2} \sum_{s=1}^{S} \frac{1}{(v/S) + h_{\Sigma,s}}, \tag{76}$$

holds for all the possible values of S. Since the parameter vector is divided into a number of S groups, this inequality holds if the following relationship is fulfilled:

$$\frac{1}{L^2} \sum_{i=(s-1)L/S+1}^{sL/S} \frac{1}{(v/L) + |h(i)|^2} \geq \frac{1}{S^2} \frac{1}{(v/S) + h_{\Sigma,s}}, \tag{77}$$
$$with \ \ s = 1, 2, \ldots, S$$

Actually, the $S = L$ case can be considered as the division of each group (with length L/S) into a number of L/S length-one sub-groups, and the inequality of (77) always holds for each group because of $Statement$ 1. This completes the proof.

REFERENCES

[1] 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, June 2008.

[2] 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, Mar. 2010.

[3] Z. Wang, Y. Xin, G. Mathew, and X. Wang, "A Low-Complexity and Efficient Channel Estimator for Multiband OFDM-UWB Systems," *IEEE Trans. Veh. Technol.*, vol.59, no.3, pp.1355-1366, Mar. 2010.

[4] Q. Z. Ahmed, L. L. Yang, and Sheng Chen, "Reduced-Rank Adaptive Least Bit-Error-Rate Detection in Hybrid Direct-Sequence Time-Hopping Ultrawide Bandwidth System," *IEEE Trans. Veh. Technol.*, vol.60, no.3, pp.849-857, Mar. 2011.

[5] M. Marques da Silva et al.,*Transmission Techniques for Emergent Multicast and Broadcast Systems*, CRC Press Auerbach Publications, 1st Edition, ISBN: 9781439815939, Boca Raton, USA, May 2010.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Upper Saddle River, NJ:Prentice Hall, 1993.

[7] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust Mean-Squared Error Estimation in the Presence of Model Uncertainties," *IEEE Journal on Signal Processing*, vol. 53, no. 1, pp. 168-181, Jan. 2005.

[8] Y. C. Eldar, "Rethinking Biased Estimation: Improving Maximum Likelihood and the Cramèr-Rao Bound," *Foundations and Trends in Signal Processing*, vol. 1, no. 4, pp. 305-449, 2008.

[9] S. Kay and Y. C. Eldar, "Rethinking Biased Estimation," *IEEE Signal Process. Mag.*, pp. 133-136, May. 2008.

[10] W. James and C. Stein, "Estimation with quadratic loss," *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability, Berkeley, CA*, vol. 1, pp.361-379, 1961.

[11] J. R. Thompson, "Some Shrinkage Techniques for Estimating the Mean," *Journal of the American Statistical Association*, vol. 63, no. 321, pp. 113-122, Mar. 1968.

[12] M. E. Bock, "Minimax estimators of the mean of a multivariate normal distribution," *The Annals of Statistics*, vol. 3, no. 1, pp. 209-218, Jan. 1975.

[13] S. Li, R. C. de Lamare, and M. Haardt, "Linear interference suppression in the frequency domain for DS-UWB systems using biased RLS estimation with adaptive shrinkage factors," *in European Wireless (EW)*, Apr. 2011.

[14] S. Li, R. C. de Lamare, and M. Haardt, "Adaptive frequency-domain biased estimation algorithms with automatic adjustment of shrinkage factors," *in Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 4268 - 4271, May 2011.

[15] J. H. Manton, V. Krishnamurthy, and H. V. Poor, "James-Stein State Filtering Algorithms," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2431-2447, Sep. 1998.

[16] Z. Ben-Haim and Y. C. Eldar, "Blind Minimax Estimation," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3145-3157, Sep. 2007.

[17] G. Durisi and S. Benedetto, "Performance evaluation of TH-PPM UWB systems in the presence of multiuser interference," *IEEE Commun. Lett.*, vol. 7, pp. 224-226, May 2003.

[18] R. Fisher, R. Kohno, and M. M. Laughlin et al., "DS-UWB Physical Layer Submission to IEEE 802.15 Task Group 3a (Doc. Number P802.15-04/0137r4)," IEEE P802.15, Jan. 2005.

[19] A. F. Molisch, D. Cassioli, and C. C. Chong et al., "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Trans. Antennas Propag.*, vol. 54, no. 11, pp. 3151-3166, Nov. 2006.

[20] M. Morelli, L. Sanguinetti and U. Mengali, "Channel Estimation for Adaptive Frequency-Domain Equalization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2508-2518, Sep. 2005.

[21] S. Li and R. C. de Lamare, "Frequency Domain Adaptive Detectors for SC-FDE in Multiuser DS-UWB Systems Based on Structured Channel Estimation and Direct Adaptation," *IET Communications*, vol. 4, issue. 13, pp. 1636-1650, 2010.

[22] M. Biguesh and A. B. Gershman, "Training-Based MIMO Channel Estimation: A Study of Estimator Tradeoffs and Optimal Training Signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884-893, Mar. 2006.

[23] J. J. de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On Channel Estimation in OFDM Systems," *in Proc. IEEE VTC*, pp. 815 - 819, Jul. 1995.

[24] O. Edfors, M. Sandell, J. J. de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM Channel Estimation by Singular Value Decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931-939, Sep. 1998.

[25] A. F. Molisch, D. Cassioli, and C. C. Chong et al., "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Trans. Antennas Propagat.*, vol. 54, no. 11, pp. 3151-3166, Nov. 2006.

[26] M. X. Chang and C.C. Yang, "A Novel Architecture of Single-Carrier Block Transmission DS-CDMA," in *IEEE Vehicular Technology Conference (VTC)*, Sep. 2006.

[27] S. Haykin, *Adaptive Filter Theory, 4th Edition*, Pearson Education, 2002.

[28] J. Teng, H Snoussi, C. Richard, and R. Zhou, "Distributed Variational Filtering for Simultaneous Sensor Localization and Target Traching in Wireless Sensor Networks," *IEEE Trans. Veh. Technol.*, vol.61, no.5, pp.2305-2318, Jun. 2012.

[29] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position Error Bound for UWB localization in Dense Cluttered Environments," *IEEE Trans. Aero. Elec. Sys.*, vol. 444, no. 2, pp. 613-628, Apr. 2008.

[30] W. Ryan and S. Lin, *Channel Codes: Classical and Modern*, Cambridge University Press, 2009.

[31] K. Todros and J. Tabrikian, "Uniformly Best Biased Estimators in Non-Bayesian Parameter Estimation," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7635-7647, 2011.

[32] 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.

[33] D. A. Pados, G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Sig. Proc.*, vol. 49, No. 2, February, 2001.

[34] S. Li and R. C. de Lamare, "Reduced-Rank Linear Interference Suppression for DS-UWB Systems Based on Switched Approximations of Adaptive Basis Functions," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 485-497, Feb. 2011.

[35] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation, and Filtering," *IEEE Trans. Sig. Proc.*, vol. 57, no. 7, pp. 2503 - 2514, July 2009.

PLACE PHOTO HERE

Sheng Li (S'08 - M'11) received his Bachelor Degree in Zhejiang University of Technology in China in 2006 and the M.Sc. degree in communications engineering and Ph.D. degree in electronics engineering, both from the University of York, U.K. in 2007 and 2010, respectively. In 2010, he was awarded the K. M. Stott prize for excellence in scientific research. During November 2010 to October 2011, he has carried out a postdoctoral research in the Ilmenau University of Technology, Germany. Since April 2012, he has been with Zhejiang University of Technology, China, where he is currently a lecturer.

PLACE PHOTO HERE

Rodrigo C. de Lamare (S'99 - M'04 - SM'10) received the Diploma in electronic engineering from the Federal University of Rio de Janeiro (UFRJ) in 1998 and the M.Sc. and PhD degrees, both in electrical engineering, from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio) in 2001 and 2004, respectively. Since January 2006, he has been with the Communications Research Group, Department of Electronics, University of York, where he is currently a Reader. Since April 2011, he has also been a Professor with PUC-RIO. His research interests lie in communications and signal processing, areas in which he has published nearly 300 papers in refereed journals and conferences. Dr. de Lamare serves as associate editor for the EURASIP Journal on Wireless Communications and Networking. He is a Senior Member of the IEEE has served as the General Chair of the 7th IEEE International Symposium on Wireless Communications Systems (ISWCS), held in York, UK in September 2010, and will serve as the Technical Programme Chair of ISWCS 2013 in Ilmenau, Germany.

PLACE PHOTO HERE

Martin Haardt (S'90 - M'98 - SM'99) 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, Germany, since 2001. Since 2012, he has also served as an Honorary Visiting Professor in the Department of Electronics at the University of York, UK.

After studying electrical engineering at the Ruhr-University Bochum, Germany, and at Purdue University, USA, he received his Diplom-Ingenieur (M.S.) degree from the Ruhr-University Bochum in 1991 and his Doktor-Ingenieur (Ph.D.) degree from Munich University of Technology in 1996.

In 1997 he joint 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.

Martin Haardt has 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 and Schwarz Outstanding Dissertation Award. 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, UK, respectively. His research interests include wireless communications, array signal processing, high-resolution parameter estimation, as well as numerical linear and multi-linear 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 co-chair 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, UK, as the general chair of ISWCS 2013 in Ilmenau, Germany, and as the general-co chair of the 5-th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) 2013 in Saint Martin, French Caribbean.