Adaptive Distributed Beamforming for Relay Networks based on Local Channel State Information

Lei Zhang, Wei Liu, Atta ul Quddus, Mehrdad Dianati and Rahim Tafazolli

Abstract. Most of the existing distributed beamforming algorithms for relay networks require global channel state information (CSI) at relay nodes and the overall computational complexity is high. In this paper, a new class of adaptive algorithms is proposed which can achieve a globally optimum solution by employing only local CSI. A reference signal based (RSB) scheme is first derived, followed by a constant modulus (CM) based scheme when the reference signal is not available. Considering individual power transmission constraint at each relay node, the corresponding constrained adaptive algorithms are also derived as an extension. An analysis of the overhead and stepsize range for the derived algorithms are then provided and the excess mean square error (EMSE) for the RSB case is studied based on the energy reservation method. As demonstrated by our simulation results, a better performance has been achieved by our proposed algorithms and they have a very low computational complexity and can be implemented on low cost and low processing power devices.

Keywords. Distributed beamforming, relay networks, constant modulus, mean square error, channel state information

I. Introduction

Distributed beamforming is a collaborative communication technique using a relay network consisting of two or more nodes forwarding the message from a transmitter to an intended receiver when there is no direct link between them or the link is so weak that it cannot support the minimum required quality of service (QoS) [1], [2], [3], [4]. It can not only improve the QoS significantly, especially for communications through poor channel, but also provide benefits of increased range, data rate or energy efficiency [3], [5], [6]. There are generally three classes of relay design schemes: amplify-and-forward (AF) [5], [6], [7], [8], [9], decoded-and-forward (DF) [8], [10], and compress-and-forward (CF) [11], [12]. The AF approach is particularly of interest and has been studied extensively in literature due to its simplicity in both algorithm design and implementation aspects. In this paper we only focus on the AF-based scheme.

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 619563. The author would also like to acknowledge the support of the University of Surrey 5GIC (http://www.surrey.ac.uk/5gic) members for this work.

Lei Zhang, Atta ul Quddus, Mehrdad Dianati and Rahim Tafazolli are with the Institute for Communication Systems (ICS), University of Surrey, Guildford, GU2 7XH, UK (Email: {lei.zhang,a.Quddus, m.dianati, r.tafazolli}@surrey.ac.uk).

Wei Liu is with the Communications Research Group, Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, S1 3JD, UK (Email: w.liu@sheffield.ac.uk).

Based on different assumptions on the knowledge of the channel state information (CSI), maximizing output signal-to-noise ratio (SNR) and minimizing mean square error (MMSE) are two commonly used design criteria for AF-based distributed beamformers. For the former one: with the assumption that perfect instantaneous CSI of both links (from source to the relay nodes and from relay nodes to destination) is known, [13] solved the optimal power control problem with each individual relay constraint for both with and without direct link scenarios. The optimal distributed beamforming problem with total and individual power constraints are presented and solved in [14]. Based on the second (2nd) order statistics of the CSI, [15] proposed two algorithms with different design objectives: minimization of the total transmit power, subject to the receiver QoS constraint, and maximizing the receiver SNR subject to two different types of power constraints. The work was then extended to two-way communication systems [16], and multiple peer-to-peer communications based on a common relay network [17]. With partial CSI, [18] proposed a worst-case optimization robust algorithm against the error in relay to destination coefficients with the consideration that the source to relay coefficient could be estimated more accurately. In [19], a robust distributed beamforming algorithm is presented by considering not only the channel estimation error, but also the CSI outdate factor. Another robust scheme was proposed in [20] for source to relay transmission based on a relay network with each node equipped with an antenna array. However, all of the proposed algorithms have a high computational complexity, which may not be practical for a relay network mainly consisting of low processing power mobile handsets. Based on the MMSE criterion, [21] proposed a multi-sensor strategy that achieves the MMSE performance subject to either local or global power constraints, and with some reasonable approximations. Furthermore, the authors proved that as the number of relay sensors (M) increases, the average power usage per node and the total power consumption will drop in the magnitude of $O(M^2)$ and $O(M)$, respectively. When M is large, the proposed distributed beamforming scheme can be performed with local CSI approximately. Based on the same design criterion (MMSE), in [22], [23], a joint channel estimation and distributed beamforming scheme with two suboptimal solutions was proposed with local CSI. A convergence analysis was provided for adaptive relay beamforming schemes that can be reformulated within the random search framework in [24], and two sufficient conditions are derived for a guaranteed convergence of the underlying adaptive algorithms.

In this paper, considering MMSE criterion, a reference signal based (RSB) distributed beamforming scheme with associated adaptive algorithms is proposed with local CSI only by estimating the destination received data through a feedback link to the relay network. The main advantage is that only local CSI is required to achieve a globally optimum solution without any approximation due to the feedback data from the destination contains the global information. Based on the same signal transmission structure, we further study the case where the reference signal is not available and derive a class of constant modulus based (CMB) adaptive beamforming algorithms. Considering individual power transmission constraint

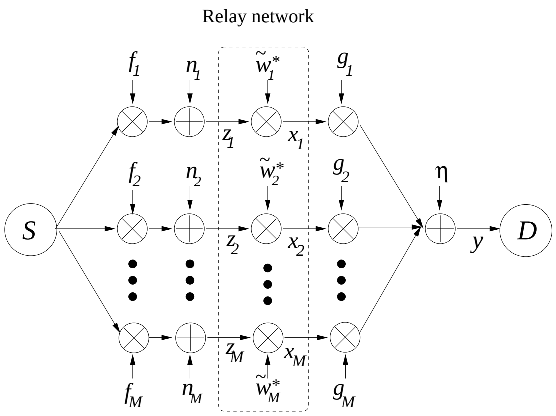


Fig. 1. A distributed beamforming structure for relay networks.

at each relay node, the corresponding constrained adaptive algorithms are then derived as an extension. The overhead for the proposed adaptive algorithms are analyzed and compared with an existing 2nd-order statistics based algorithm. Then the stepsize ranges for the derived algorithms are provided and the excess MSE (EMSE) for the RSB case is studied based on the energy reservation method. All of the proposed algorithms have extremely low computational complexity, and can be implemented on low cost and low processing power devices.

The rest of the paper is organized as follows. In Section II, the MMSE-based distributed beamforming model is introduced and the traditional global CSI based algorithms are provided. The proposed adaptive beamforming algorithms for both unconstrained and constrained approaches are derived in Section III. The overhead, computational complexity, convergence range of stepsize and EMSE of the proposed algorithms are analyzed in Section IV. Simulations are provided in Section V and conclusions are drawn in Section VI.

Notation: Vectors and matrices are denoted by uppercase and lowercase bold letters. $\{\cdot\}^H, \{\cdot\}^T, \{\cdot\}^*$ are Hermitian conjugate, transpose and conjugate operations, respectively. $\mathcal{E}\{\cdot\}$ denotes the expectation operation. $\mathbf{I}$ is the identity matrix and $j = \sqrt{-1}$. $\odot$ denote the element-wise (Schur-Hadamard) multiplication of two vectors. We also define the vector $\mathbf{l}_1 = [1, 1, \cdots, 1]^T$ with proper dimension.

II. System Model and Problem Formulation

Consider a wireless network with a source $\mathcal{S}$, a destination $\mathcal{D}$, and a relay network consisting of M available nodes with each equipped with a single antenna for up- and down-link communications, as shown in Fig. 1. The channel is assumed to be fixed during the relay beamforming process. Due to poor channel condition, there is no direct link between the source and the destination. The message is sent from the source to the destination through the relay network with two steps: first, the source sends the signal to all of the available relay nodes with an attenuation coefficient f_i for the ith relay; then the received signal by each relay node will be weighted by a complex coefficient $\tilde{w}_i^*$ and then forwarded to the destination.

At the n-th snapshot and i-th relay, the received signal is $z_i[n] = f_i s[n] + n_i[n]$ with s and $n_i[n]$ being the transmitted signal and noise at the i-th relay, respectively. The signal power $\mathcal{E}\{|s[n]|^2\} = \sigma_s^2$. To simply the expression in the derived algorithms, we assume noise power at all relay nodes are the same, i.e. $\mathcal{E}\{|n_i[n]|^2\} = \sigma_n^2$. However, the proposed algorithms are not restricted by this assumption. In a vector form, we have

$$\mathbf{z}[n] = \mathbf{f}s[n] + \mathbf{n}[n] \ , \tag{1}$$

where $\mathbf{z} = [z_1[n], \cdots, z_M[n]]^T$, $\mathbf{f} = [f_1, \cdots, f_M]^T$ and $\mathbf{n} = [n_1[n], \cdots, n_M[n]]^T$.

In the AF relay scheme, $z_i[n]$ is weighted by $\tilde{w}_i^*[n]$ before forwarding to the destination, which can be expressed as

$$\mathbf{x}[n] = \tilde{\mathbf{w}}^*[n] \odot (\mathbf{f}s[n] + \mathbf{n}[n]) \ , \tag{2}$$

where $\mathbf{x}[n] = [x_1[n], \cdots, x_M[n]]^T$ and $\tilde{\mathbf{w}}[n] = [w_1[n], \cdots, w_M[n]]^T$. With the relay to destination coefficient $g_i[n]$, the destination received signal $y[n] = \sum_{i=1}^{M} g_i x_i[n] + \eta[n]$ is a linear mixture of $x_i[n]$ with additive noise $\eta \sim \mathcal{CN}(0, \sigma_\eta^2)$, i.e.

$$y[n] \ = \ \mathbf{g}^T(\tilde{\mathbf{w}}^*[n] \odot \mathbf{f})s[n] + \mathbf{g}^T(\tilde{\mathbf{w}}^*[n] \odot \mathbf{n}[n]) + \eta[n] \tag{3}$$

where $\mathbf{g} = [g_1, \cdots, g_M]^T$. Since $\mathbf{g}$ and $\tilde{\mathbf{w}}^*[n]$ are exchangeable, we can rewrite (3) as

$$\begin{aligned} y[n] &= \tilde{\mathbf{w}}^H[n](\mathbf{g} \odot \mathbf{f})s[n] + \tilde{\mathbf{w}}^H[n](\mathbf{g} \odot \mathbf{n}[n]) + \eta[n] \\ &= \tilde{\mathbf{w}}^H[n]\mathbf{t}[n] + \eta[n] \ , \end{aligned} \tag{4}$$

with

$$\mathbf{t}[n] = (\mathbf{g} \odot \mathbf{f})s[n] + (\mathbf{g} \odot \mathbf{n}[n]) = \mathbf{g} \odot \mathbf{x}[n] \ . \tag{5}$$

For the power unconstrained MMSE-based method, the optimum solution can be obtained by solving the following optimization problem:

$$\begin{aligned} &\min_{\mathbf{w}} f_o = \mathcal{E}\|r[n] - y[n]\|^2 \\ =\ &\mathcal{E}\|r[n] - \tilde{\mathbf{w}}^H[n]\mathbf{t}[n] - \eta[n]\|^2 \ , \end{aligned} \tag{6}$$

where $r[n]$ is a reference signal available to all relay nodes, which can be any scaled version of the transmitted signal $s[n]$, i.e., $r[n] = c \cdot s[n]$ with c being a nonzero constant.

A well-known adaptive solution to (6) is given by

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \mu e^*[n]\mathbf{t}[n] \ , \tag{7}$$

with μ being the stepsize and $e[n] = r[n] - y[n]$. One problem with the solution in (7) is that $e[n]$ is unknown in distributed relay networks since the destination received data $y[n]$ is unknown at the nodes.

Alternatively, if the 2nd order statistics of the signals are available (or can be estimated), we can use the standard Wiener solution to solve the problem, and the optimum coefficient vector $\mathbf{w}_o$ is given by [25], [26]

$$\mathbf{w}_o = \mathbf{R}_{tt}^{-1}\mathbf{r}_{tr} \ , \tag{8}$$

where $\mathbf{R}_{tt} = \mathcal{E}\{\mathbf{t}[n]\mathbf{t}^H[n]\}$ and $\mathbf{r}_{tr} = \mathcal{E}\{r^*[n]\mathbf{t}[n]\}$. Using (5), we have $\mathbf{R}_{tt} = \mathcal{E}\{\mathbf{t}[n]\mathbf{t}^H[n]\} = \sigma_s^2\mathbf{G}\mathbf{f}\mathbf{f}^H\mathbf{G}^H + \sigma_v^2\mathbf{G}\mathbf{G}^H$

and $\mathbf{r}_{tr} = \sigma_s^2\mathbf{G}\mathbf{f}$, where $\mathbf{G} = \mathrm{diag}(\mathbf{g})$. Substituting them into (8) leads to the following result

$$\mathbf{w}_o = \sigma_s^2(\sigma_s^2\mathbf{G}\mathbf{f}\mathbf{f}^H\mathbf{G}^H + \sigma_v^2\mathbf{G}\mathbf{G}^H)^{-1}\mathbf{G}\mathbf{f} \ . \tag{9}$$

One clear problem with (9) is that the calculation of the relay coefficients for each node can not be realised in a truly distributed manner since it requires the global channel coefficients (both $\mathbf{f}$ and $\mathbf{g}$) at each relay node. Ideally we need a central node to collect all the CSI and calculate the optimum solution and then send them to each relay node. Assume the destination is selected as the central node. The channels from relay nodes to destination can be estimated at the destination. However, the channel coefficients from source to relay nodes ($\mathbf{f}$) should also be collected, which could be done by each relay node transmitting with a scheme similar to AF[1].

Specifically, f_i will be treated in a similar way as the reference signal transmitted by a communication system for CSI estimation. It is forwarded by the relay node to destination over the noisy channel g_i, which is assumed already perfectly known at the central node. Then the received channel coefficients data can be written as $\hat{f}_i = g_i f_i + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is the additive noise. The estimation of f_i can be achieved using the MMSE receiver as $f_{e,i} = g_i^*\hat{f}_i/(|g_i|^2 + \sigma_\epsilon^2)$, which could be written in a vector form as $\mathbf{f}_e = [f_{e,1}, \cdots, f_{e,M}]^T$. The calculated beamformer at the destination in terms of estimated $\mathbf{f}_e$ can be written as $\mathbf{w}_D = \sigma_s^2(\sigma_s^2\mathbf{G}\mathbf{f}_e\mathbf{f}_e^H\mathbf{G}^H + \sigma_v^2\mathbf{G}\mathbf{G}^H)^{-1}\mathbf{G}\mathbf{f}_e$.

The corresponding elements of $\mathbf{w}_D$ (i.e. $w_{D,i}$) will be fed back from destination to each relay node. Suppose each node receive a noisy coefficient over the channel g_i as $\hat{w}_i = g_i w_{D,i} + \varepsilon_i$ with $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ being the noise. Similarly, the estimation of $w_{D,i}$ at each relay node can be written as $w_{R,i} = g_i^*\hat{w}_i/(|g_i|^2 + \sigma_\varepsilon^2)$, and in a vector form,

$$\mathbf{w}_R = \frac{\mathbf{g}^* \odot \mathbf{w}_D}{(\mathbf{g} \odot \mathbf{g}^* + \sigma_\epsilon^2\mathbf{1}_1)} \ . \tag{10}$$

Thus, in order to achieve the cooperation gain in a distributed relay network, two estimation errors are introduced, which may lead to performance degradation, as will be shown in our simulations later.

III. The Proposed Distributed Beamforming Scheme

To overcome the problem with traditional relay beamforming design, in this section, we propose a new data transmission scheme that employs only the local CSI for each relay node. In this scheme, we estimate the received data at destination and then broadcast the estimation to all of the relay nodes. Depending on the duplex status of the involved relay nodes, the transmission schemes are different. With half-duplex relay nodes, different from the traditional AF distributed beamforming where two time slots are used, at the first and the second stages, the source sends signal $s[n]$ to the relay network and the amplified signal $\mathbf{x}[n]$ is forwarded to the destination; an extra slot is required for the destination to broadcast the last snapshot it received ($y[n]$) back to the relay network, which is the third stage. Certainly this new stage will reduce the overall data rate. However, as discussed later, after the learning process converges, less and less feedback is needed and this third stage could be removed. While in case the relay node can work in a double-duplex mode, the proposed scheme could be realised by a two time slots strategy. In the first stage, the source sends signal $s[n]$ to the relay network. In the second stage, the amplified signal is forwarded to the destination; meanwhile, the destination is broadcasting the last snapshot it received back to the relay network. In either case, we can write the received feedback at relay node as

$$\tilde{\mathbf{y}}[n] = \mathbf{g}y[n] + \mathbf{v}[n] \ . \tag{11}$$

Notice here $\tilde{\mathbf{y}}[n] = [\tilde{y}_1[n], \cdots, \tilde{y}_M[n]]^T$ with $\tilde{y}_i[n]$ being the i-th node received data, and $\mathbf{v}[n]$ is the noise vector with each element $v_i[n] \sim \mathcal{CN}(0, \sigma_v^2)$. An estimate of the true data $y[n]$ by using the known feedback $\tilde{\mathbf{y}}[n]$ and a set of coefficients $\mathbf{u}$ is given by:

$$\bar{\mathbf{y}}[n] = \mathbf{u} \odot \tilde{\mathbf{y}}[n] \ , \tag{12}$$

$\mathbf{u} = [u_1, \cdots, u_M]^T$ and u_i is the weight coefficient of the ith node. Note here the channel estimation errors are not taken into consideration. Two schemes will be proposed in the next to solve the problem: a reference signal based (RSB) scheme and a constant modulus based (CMB) blind scheme.

A. Reference Signal Based (RSB) Scheme

The RSB scheme is characterized by the optimization problem

$$\begin{aligned}\min_{\mathbf{w},\mathbf{u}} f_r &= \frac{1}{M}\mathcal{E}\|r[n]\mathbf{1}_1 - \mathbf{u} \odot \tilde{\mathbf{y}}[n]\|^2 \\ &= \frac{1}{M}\min_{w_i,u_i}\sum_{i=1}^{M}\mathcal{E}|r[n] - u_i\tilde{y}_i[n]|^2 \ . \end{aligned} \tag{13}$$

By substituting (11) and (4) into (13), taking the gradient of f_r with respect to $\mathbf{w}^*$ and setting it to zero, we have

$$\frac{\partial f_r}{\partial \mathbf{w}^*} = \sum_{i=1}^{M}|u_i g_i|^2\mathbf{R}_{tt}\mathbf{w} - \sum_{i=1}^{M}u_i g_i\mathbf{r}_{tr} = 0 \ . \tag{14}$$

Then we can solve $\mathbf{w}$ as a function of u_i

$$\mathbf{w} = \frac{\sum_{i=1}^{M}u_i g_i}{\sum_{i=1}^{M}|u_i g_i|^2}\mathbf{R}_{tt}^{-1}\mathbf{r}_{tr} \ . \tag{15}$$

To solve the unknown u_i in equation (15), we minimize the following cost function $\mathcal{E}|r[n] - y[n]|^2$, which is the same as (6) and to achieve the MMSE at the destination, the solution of (15) must be the same as (8), which implies that

$$\frac{\sum_{i=1}^{M}u_i g_i}{\sum_{i=1}^{M}|u_i g_i|^2} = 1 \ . \tag{16}$$

Since equation (16) is an underdetermined equation, the solutions is not unique. One of particular interest is

$$u_{r,i} = 1/g_i \ , \tag{17}$$

[1]A scheme similar to DF can be applied to the transmission of the channel coefficients $\mathbf{f}$ from relay nodes to destination. However, this option may lead to a larger latency, extra complexity (e.g. quantization, modulation, channel coding, etc.) and overhead (e.g. channel coding). For fair comparison, the same AF process is also adopted by the proposed adaptive algorithms.

where the subscript $\{\}_r$ denote the solution to the RSB algorithm. Substituting (17) into (14), we have the following RSB adaptive distributed beamforming algorithm

$$w_{r,i}[n+1] = w_{r,i}[n] + \mu_r (r[n] - \frac{1}{g_i}\tilde{y}_i[n])^* t_i[n] \ , \quad (18)$$

with μ_r being the stepsize. Note that $\sum_{i=1}^{M} u_i g_i$ and $\sum_{i=1}^{M} |u_i g_i|^2$ are canceled out in the adaptive algorithm, which means only local CSI is needed and for each node, the update of its coefficient does not need CSI related to any other node in the relay. Equation (18) can be expressed in a vector form as

$$\mathbf{w}_r[n+1] = \mathbf{w}_r[n] + \mu_r (r[n] \odot \mathbf{l}_1 - \frac{\mathbf{1}_1}{\mathbf{g}} \odot \tilde{\mathbf{y}}[n])^* \odot \mathbf{t}[n] \quad (19)$$

B. *Constant Modulus Based (CMB) Scheme*

The RSB scheme is based on knowledge of a reference signal ($r[n]$) at each relay node. When $r[n]$ is unavailable, but the transmitted signals have a constant modulus (for example: phase modulated signals) [27], [28], [29], [30], [31], the CMB scheme can be employed for distributed beamforming. Generally, a CMB algorithm is derived by minimizing the following cost function [32], [33]

$$\min_{\mathbf{w}} E\{(h^p - |y[n]|^p)^q\} \ , \quad (20)$$

where h is modulus of the desired signal and can normally be set as 1, and p and q are positive integers and generally chosen to be 1 or 2. For different values of p and q, the performance of the derived algorithms are different. [34] has shown that for the case of $p = q = 2$ the algorithm can achieve the best performance and in this paper, we will only consider this scenario.

Since $y[n]$ is unknown in (20), replacing it by its estimation $u_i\tilde{y}[n]$ with $p = q = 2$, (20) changes to

$$\min_{\mathbf{w},u_i} E\{(1 - |u_i\tilde{y}[n]|^2)^2\} \ . \quad (21)$$

There are no closed-form solutions to the problem and here we provide three adaptive solutions.

1) Separate optimal constant modulus based (SCMB): In this solution, we first obtain the optimal solution for $\mathbf{u}$ and $\bar{\mathbf{y}}[n]$ based on the MMSE criterion, and then replace $y[n]$ by $\bar{y}_i[n]$ at each node to find the final weight vector iteratively.

Suppose that each note i knows its own relay to destination coefficient g_i and the power of the noise σ_v^2. Then, the original $y[n]$ can be estimated at each node by minimizing the following cost function,

$$\min_{u_i} \mathcal{E}|u_i\tilde{y}_i[n] - y[n]|^2 \quad \text{for} \quad i = 1, \cdots, M \ , \quad (22)$$

where u_i is the coefficient at the i-th node to estimate $y[n]$. Taking the gradient of (22) with respect to u_i then setting it to zero, we have

$$\mathcal{E}(u_i\tilde{y}_i[n] - y[n])\tilde{y}_i^*[n] = 0 \ . \quad (23)$$

Substituting equations (4) and (11) into (23) and noticing that

$$\begin{aligned} &\mathcal{E}(\tilde{y}_i[n]\tilde{y}_i[n]^*) = \mathcal{E}|\tilde{y}_i[n]|^2 = |g_i|^2\sigma_y^2 + \sigma_v^2 \\ &\mathcal{E}(y_i[n]\tilde{y}_i^*[n]) = g_i^*\sigma_y^2 \ , \end{aligned} \quad (24)$$

with $\sigma_y^2 = \mathcal{E}|y[n]|^2$, the coefficients can be given as

$$u_{s,i} = \frac{g_i^*\sigma_y^2}{|g_i|^2\sigma_y^2 + \sigma_v^2} \ , \quad (25)$$

where the subscript $\{\}_s$ in $u_{s,i}$ denotes the separate CMB algorithm. Then the estimation $\bar{y}[n]$ of $y[n]$ at each node can be written as

$$\bar{y}_{s,i}[n] = u_{s,i}\tilde{y}_i[n] = \frac{g_i^*\sigma_s^2}{|g_i|^2\sigma_s^2 + \sigma_v^2}\tilde{y}_i[n] \ . \quad (26)$$

Due to presence of noise, the estimations $\bar{y}_{s,i}[n]$ for $i = 1, \cdots, M$ of $y[n]$ have different values at each relay sensors. However, all of the $\bar{y}_{s,i}[n]$ at different nodes have the same second-order parameters. Note that for each node with different $\bar{y}[n]$, we can formulate a similar cost function as (21) and find the corresponding solution $\mathbf{w}^{(i)}[n+1]$ as follows

$$\mathbf{w}_s^{(i)}[n+1] = \mathbf{w}_s^{(i)}[n] + \mu_s b_{s,i}[n]\mathbf{t}[n] \ , \quad (27)$$

where μ_s is the stepsize, the superscript $^{(i)}$ denotes the weight vector calculated at node i, and

$$b_{s,i}[n] = u_i g_i (1 - |\bar{y}_{s,i}[n]|^2)\bar{y}_{s,i}^*[n] \ . \quad (28)$$

Similarly, the weight vectors $\mathbf{w}_s^{(i)}[n+1]$ calculated at different node are different, each corresponding to a set of suboptimal solution. However, we will prove in the next that the differences of weight vectors calculated at each node with different u_i, g_i are very small based on the assumption that the noise power σ_v^2 is small compared with signal power.

Taking the expectation of $b_{s,i}[n]$ and using $\mathcal{E}(v[n]) = 0$, we have

$$\begin{aligned} \mathcal{E}\{b_{s,i}[n]\} &= |u_{s,i}|^2|g_i|^2\mathcal{E}\{y^*[n](1 \\ &- |u_{s,i}|^2|g_i|^2|y[n]|^2 - \sigma_v^2)\} \ . \end{aligned} \quad (29)$$

When $n \to \infty$, we can use the approximation $|y[n]|_{n\to\infty} \approx 1$. Substituting it into (29) and with (25), we have

$$\begin{aligned} &\mathcal{E}\{b_{s,i}[n]\}_{n\to\infty} \approx |u_{s,i}|^2|g_i|^2\mathcal{E}\{y^*[n]\}(1 - |u_{s,i}|^2|g_i|^2 - \sigma_v^2) \\ &= \mathcal{E}\{y^*[n]\}\frac{|g_i|^4}{(|g_i|^2 + \sigma_v^2)^2}(1 - \frac{|g_i|^4}{(|g_i|^2 + \sigma_v^2)^2} - \sigma_v^2) \ . \end{aligned} \quad (30)$$

The converged weight vector at each node will depend on the level of noise power σ_v^2. Larger σ_v^2 may cause each node to converge to a significantly different value. With the assumption of $|g_i|^2 \gg \sigma_v^2$, (30) can be reduced into

$$\mathcal{E}\{b_{s,i}[n]\}_{n\to\infty} \approx -\mathcal{E}\{y^*[n]\}\sigma_v^2 \ , \quad (31)$$

which is not a function of $u_{s,i}$ and g_i anymore. Furthermore, all of the weight vectors approximately converge to the same value.

By taking the corresponding weight coefficient from $\mathbf{w}_s^{(i)}[n]$, we can write the solution of (22) in a vector form $\mathbf{w}_s = [w_{s,1}^{(1)}, \cdots, w_{s,M}^{(M)}]^T$, where $w_{s,i}^{(i)}$ is the i-th element of $\mathbf{w}_s^{(i)}[n]$, which means at node i, only $w_{s,i}^{(i)}$ needs to be calculated. $\mathbf{w}$ can be written as

$$\mathbf{w}_s[n+1] = \mathbf{w}_s[n] + \mu_s\mathbf{b}_s[n] \odot \mathbf{t}[n] \ , \quad (32)$$

where $\mathbf{b}_s[n] = \mathbf{u}_s \odot \mathbf{g} \odot (\mathbf{l}_1 - |\bar{\mathbf{y}}_{\mathbf{s}}[n]|^2) \odot \bar{\mathbf{y}}_{\mathbf{s}}^*[n]$. One advantage of this selection strategy is that only local CSI is needed at relay nodes.

2) Consistent CMB (CCMB) adaptive algorithm: By taking

$$u_{c,i} = 1/g_i \ , \tag{33}$$

(29) will be reduced to

$$\mathcal{E}\{b_{c,i}[n]\} = \mathcal{E}\{y^*[n]\}(1 - \mathcal{E}|y[n]|^2 - \sigma_v^2) \ , \tag{34}$$

which is unrelated with the coefficient g_i, implying that all of the weight vectors calculated by each node are the same, or we can say the algorithm is consistent. In practice, it is not necessary to calculate all of the weights at each node. Instead, node i only needs to calculate $w_i[n+1]$, which can be written as

$$w_{c,i}[n+1] = w_{c,i}[n] + \mu_c(1 - |\bar{y}_{c,i}[n]|^2)\bar{y}^*_{c,i}[n]t_i[n] \ , \tag{35}$$

with $\bar{y}_{c,i}[n] = u_{c,i}\tilde{y}[n]$. Similarly, in vector form, we have

$$\mathbf{w}_c[n+1] = \mathbf{w}_c[n] + \mu_c(\mathbf{l}_1 - |\bar{\mathbf{y}}_c[n]|^2) \odot \bar{\mathbf{y}}^*_c[n] \odot \mathbf{t}[n] \ . \tag{36}$$

3) Iterative CMB (ICMB) adaptive algorithm: In the last two solutions, we used different strategies to decide the estimator u_i and then the weight vector can be calculated consequently. Now consider equation (21), which has no closed-form solution. Taking the gradient of (21) with respect to w_i and u_i, we arrive at the following ICMB adaptive algorithm

$$\begin{aligned} w_{it,i}[n+1] &= w_{it,i}[n] + \mu_{it}u_{it,i}[n]g_i(1 - |\bar{y}_{it,i}[n]|^2)\bar{y}^*_{it,i}[n]t_i[n] \\ u_{it,i}[n+1] &= u_{it,i}[n] + \mu_u(1 - |\bar{y}_{it,i}[n]|^2)\bar{y}_{it,i}[n]\tilde{y}^*[n] \end{aligned} \tag{37}$$

where μ_{it} and μ_u are the stepsizes, and $\bar{y}_{it,i}[n] = u_{it,i}[n]\tilde{y}_i[n]$.

Define $\bar{\mathbf{y}}_{it}[n] = [\bar{y}_{it,1}[n], \cdots, \bar{y}_{it,M}[n]]^T$. Then (37) can be written in a vector form as

$$\begin{aligned} \mathbf{w}_{it}[n+1] &= \mathbf{w}_{it}[n] + \mu_{it}\mathbf{u}_{it}[n] \odot \mathbf{b}_{it}[n] \\ \mathbf{u}_{it}[n+1] &= \mathbf{u}_{it}[n] + \mu_u(1 - |\bar{\mathbf{y}}_{it,i}[n]|^2) \odot \bar{\mathbf{y}}_{it}[n]\tilde{y}^*[n] \end{aligned} \tag{38}$$

where $\mathbf{b}_{it}[n] = \mathbf{g} \odot (\mathbf{l}_1 - |\bar{y}_{it}[n]|^2) \odot \bar{\mathbf{y}}^*_{it}[n]\tilde{y}[n]$.

For distributed beamforming algorithms, assumption of global CSI at each node is not practical. (19) for RSB, (32), (36) (38) for CMB are adaptive solutions for relay networks based on local CSI only. However, as a compromise, the global parameter $y[n]$ has to be broadcasted back to the relay nodes, which incurs extra power consumption. At the beginning of each communication session, a training stage is required. However, with the learning process converging, less and less feedback is needed and some low complexity techniques such as set-membership (SM) [35], [36] can be used to further improve the data rate. Except for the local CSI, the proposed algorithms have extremely low computational complexity, which is especially important for relay nodes comprised of mobile handsets. Moreover, the adaptive algorithms proposed here can also track channel variations very rapidly, which will be shown in the simulations part.

C. Constrained adaptive beamforming with local CSI

In this subsection, we will consider the problem that each individual node is restricted in its transmit power, which is particularly important for battery-powered relay nodes. Taking the RSB adaptive algorithm as an example, with individual node power transmission constraint, we can formulate the following optimization problem

$$\begin{aligned} &\min_{\mathbf{w},\mathbf{u}} f_r = \frac{1}{M}\mathcal{E}\|r[n] - \mathbf{u} \odot \tilde{\mathbf{y}}[n]\|^2 \\ &\mathcal{E}|w_i^* x_i[n]|^2 \le p_i, \quad \text{for } i = 1, \cdots M \ , \end{aligned} \tag{39}$$

where $\mathcal{E}|w_i^* x_i[n]|^2$ and p_i are the average transmission power and the maximum power transmission for node i, respectively. This is a nonconvex optimization problem that can be solved as follows.

First we use the normal adaptive algorithm in (32) to calculate $\tilde{w}_i[n+1]$ and the instantaneous power transmission $|\tilde{w}_i^*[n+1]x_i[n+1]|^2$, and then compare it with p_i. If $|\tilde{w}_i^*[n+1]x_i[n+1]|^2 > p_i$, normalize $|\tilde{w}_i^*[n+1]x_i[n+1]|^2$ by replacing $\tilde{w}_i[n+1]$ by $w_i[n+1] = \frac{\sqrt{|\tilde{w}_i^*[n+1]x_i[n+1]|^2}}{\sqrt{p_i}}\tilde{w}_i[n+1]$. In this case, the modulus normalization of $\tilde{w}_i[n+1]$ will keep the individual power transmission under the allowance. Meanwhile, this processing will affect the convergence of the MSE algorithm and a suboptimal solution may be achieved when some of the sensors power transmission budgets are lower than actually required. However, due to properties of the adaptive algorithm, part of the loss will be compensated by some other nodes with sufficient power budget automatically. Note that the power constraint for the destination feedback signal at the training stage is not considered in the paper.

Similar idea can be used for the CMB adaptive algorithms and it is omitted here.

IV. Performance analysis

A. Overhead analysis

The proposed adaptive algorithms require a training stage at the beginning of a communication session. Suppose this stage lasts for K snapshots before converging to the target performance. Each snapshot includes 3 time slots, i.e. reference signal ($r[n]$) from source to relay node, relay nodes forward $x[n]$ to destination, and the destination feeds back received $y[n]$ to relay nodes. Assume that $r[n]$, $x[n]$ and $y[n]$ occupy one symbol, respectively. In addition, each relay node needs the channel coefficient g_i, which can be estimated when the destination broadcasts the one symbol reference signal back to all relay nodes. In total, the signal overhead will be $3K+1$.

For the 2nd order statistics based algorithm in (10), the estimation of $\mathbf{f}$ at relay nodes need one reference symbol. M symbols are required for $\mathbf{f} = [f_1, \cdots, f_M]^T$ to be transmitted to the central node, where the beamforming coefficients $\mathbf{w}_D = [w_{D,1}, \cdots, w_{D,M}]$ are calculated, which are then fed back to the relay nodes using M symbols. In addition, extra M symbols are required for estimation of $\mathbf{g}$. In total, the overhead for the distributed beamforming implementation will be $3M+1$.

Apparently, whether the proposed algorithms have a larger overhead or not depends on converging speed of the algorithms (i.e. K) and the number of cooperative nodes M. We will show in the simulations that with a moderate number of M, the proposed algorithms can converge to the target MSE in a few tens of snapshots, which leads to a smaller overhead than the solution in (10).

B. Computational complexity

All the proposed algorithms have a extremely low complexity and suitable for relay nodes with limited power and processing capability. For each node and at each snapshot, the unconstrained RSB algorithm in total needs 10 real-valued multiplications and 4 real-valued additions, while the SCMB, CCMB and ICMB require 13, 17, 21 real-valued multiplications and 4, 6, 6 real-valued additions, respectively. For the constrained algorithms, extra operations are needed due to the normalization operation. In the worst case, where each node regularizes its weight at each snapshot, it would need extra 10 real-valued multiplications and 1 real-valued addition.

C. Range of stepsize for convergence

Let us first consider the RSB adaptive algorithm. Define the weight vector error $\mathbf{e}_{r,w}[n] = \mathbf{w}_o - \mathbf{w}_r[n]$. Substituting it into (19) and taking the expectation of both sides, we have

$$\begin{aligned}&\mathcal{E}\{\mathbf{e}_{r,w}[n+1]\} = \mathcal{E}\{\mathbf{e}_{r,w}[n] - \mu_r(r^*[n]\mathbf{t}[n] - \bar{\mathbf{y}}_r^*[n]\odot\mathbf{t}[n])\}\\&= \mathcal{E}\{\mathbf{e}_{r,w}[n] - \mu_r(\mathbf{R}_{tt}\mathbf{w}_o - \mathbf{R}_{tt}\mathbf{w}[n])\}\\&= (\mathbf{I} - \mu_r\mathbf{R}_{tt})\mathbf{e}_{r,w}[n]\ ,\end{aligned} \tag{40}$$

where we have used $\mathcal{E}\{r^*[n]\mathbf{t}[n]\} = \mathbf{r}_{tr} = \mathbf{R}_{tt}\mathbf{w}_o$ and $\mathcal{E}\{\bar{\mathbf{y}}_r^*[n]\odot\mathbf{t}[n]\} = \mathbf{R}_{tt}\mathbf{w}_r[n]$. The stable condition of $\lim_{n\to\infty} E\{\mathbf{w}_r[n+1]\} = \mathbf{w}_{opt}$ or $\lim_{n\to\infty} E\{\mathbf{e}_{r,w}[n+1]\} = \mathbf{0}$ is equivalent to $E\{\Pi_{n=1}^{\infty}(\mathbf{I}-\mu_r\mathbf{R}_{tt})\} = \mathbf{0}$ [37]. A sufficient condition for stability is that the stepsize is limited to the following range [37]

$$0 \le \mu_r \le \min_k \frac{2}{\lambda_k^{\mathbf{R}_{tt}}}\ , \tag{41}$$

where $\lambda_k^{\mathbf{R}_{tt}}$ is the $k-$th eigenvalue of $\mathbf{R}_{tt}$.

For the SCMB adaptive algorithm, defining the matrix $\mathbf{R}_{sx} = \mathbf{I} + \mu_s|\mathbf{u}_s|^2\odot|\mathbf{g}|^2\odot(\mathbf{l}_1 - |\bar{\mathbf{y}}_\mathbf{s}[n]|^2)\odot\mathbf{t}[n]\mathbf{t}^H[n]$ and using (32), we have

$$\begin{aligned}&\mathbf{e}_{s,\mathbf{w}}[n+1] = \mathbf{w}_o - \mathbf{w}_s[n+1] = \mathbf{w}_o - \mathbf{R}_{sx}\mathbf{w}_s[n]\\&= -\mathbf{R}_{sx}(\mathbf{w}_o - \mathbf{e}_{s,w}[n]) + \mathbf{w}_o\\&= \mathbf{R}_{sx}\mathbf{e}_{s,w}[n] + \mu\mathbf{R}_{st}\mathbf{w}_o\ ,\end{aligned} \tag{42}$$

where

$$\mathbf{R}_{st} = \mathcal{E}\{|\mathbf{u}_s|^2\odot|\mathbf{g}|^2\odot(\mathbf{l}_1 - |\bar{\mathbf{y}}_\mathbf{s}[n]|^2)\odot\mathbf{t}[n]\mathbf{t}^H[n]\} \tag{43}$$

and $\mathbf{w}_o$ is the optimal solution as defined in equation (8), which could be achieved by each relay node through (27), i.e, the optimal solution $\mathbf{w}^{(i)}[n]_{n\to\infty}$ calculated at each relay node is optimal and identical. However, due to separate optimisation of u_i and $\mathbf{w}$, the optimal solutions for SCMB at each relay node are different, as shown in (27). However, as proved in (29) and (30), the differences of weight vectors calculated at each node with different u_i, g_i are very small based on the assumption that the noise power σ_v^2 is small compared to signal power σ_s^2. Then we have $\mathbf{R}_{st}\mathbf{w}_o \approx \mathbf{0}$. Substituting it into (42) and then taking the expectation operation, we have

$$\mathcal{E}\{\mathbf{e}_{s,w}[n+1]\} = \mathcal{E}\{\mathbf{R}_{sx}\}\mathcal{E}\{\mathbf{e}_{s,w}[n]\}\ . \tag{44}$$

A sufficient condition for stability is that the stepsize is limited to the following range

$$0 \le \mu_s \le \min_k \frac{2}{\lambda_k^{\mathbf{R}_{st}}}\ , \tag{45}$$

where $\lambda_k^{\mathbf{R}_{st}}$ is the $k-$th eigenvalue of $\mathbf{R}_{st}$. Due to effect of noise, the steady stepsize range at each node is different for the SCMB algorithm.

For the CCMB adaptive algorithm, substitute $u_{c,i} = 1/g_i$ into (45) and (43), and then the sufficient condition for stability is that the stepsize is limited to the following range

$$0 \le \mu_c \le \min_k \frac{2}{\lambda_k^{\mathbf{R}_{ct}}}\ , \tag{46}$$

where $\lambda_k^{\mathbf{R}_{ct}}$ is the $k-$th eigenvalue of $\mathbf{R}_{ct}$ and

$$\mathbf{R}_{ct} = \mathcal{E}\{(\mathbf{l}_1 - |\bar{\mathbf{y}}_\mathbf{c}[n]|^2)\odot\mathbf{t}[n]\mathbf{t}^H[n]\}\ . \tag{47}$$

D. Steady-state excess MSE (EMSE) analysis

The cost function of the CMB algorithms is very complicated and it is very difficult to given a reasonable analysis of their steady-sate EMSE. Here we only focus on the RSB adaptive algorithm. For simplification, we will drop the subscript $\{\}_r$ in all of the parameters in the following analysis. At the nth snapshot, the MSE of the beamformer can be written as

$$\begin{aligned}&MSE[n] = \mathcal{E}|r[n] - \mathbf{w}^H[n]\mathbf{t}[n] - \eta[n]|^2 = \mathcal{E}|r[n] -\\&(\mathbf{w}_o - \mathbf{e}_w[n])^H\mathbf{t}[n] - \eta[n]|^2\\&= MSE_{min} + \mathcal{E}\{\mathbf{e}_w^H[n]\mathbf{t}[n]\varepsilon[n] + \mathbf{t}^H[n]\mathbf{e}_w[n]\varepsilon^*[n] +\\&\mathbf{e}_w^H[n]\mathbf{t}[n]\mathbf{t}^H[n]\mathbf{e}_w[n]\}\ ,\end{aligned} \tag{48}$$

where $MSE_{min} = \mathcal{E}|\varepsilon|^2 = \mathcal{E}|r[n] - \mathbf{w}_o^H\mathbf{t}[n] - \eta[n]|^2$ is the minimum MSE. When $n\to\infty$, the weight vector error $\mathbf{e}_w[n]\to\mathbf{0}$, and we can then define the steady-state EMSE as

$$\begin{aligned}\rho &= \lim_{n\to\infty} MSE[n] - MSE_{min}\\&= \lim_{n\to\infty}\mathcal{E}\{\mathbf{e}_w^H[n]\mathbf{t}[n]\mathbf{t}^H[n]\mathbf{e}_w[n]\}\ .\end{aligned} \tag{49}$$

For convenience, we define the $a\ priori$ estimation error as

$$e_a[n] = (\mathbf{w}_o - \mathbf{w}[n])^H\mathbf{t}[n] = \mathbf{e}_w^H[n]\mathbf{t}[n]\ . \tag{50}$$

Substitute it into (49), and we have

$$\rho = \lim_{n\to\infty}\mathcal{E}|e_a|^2\ . \tag{51}$$

The EMSE of a stochastic-gradient (SG) based algorithm has been analyzed using the energy conservation method [38], [39]. However, due to difference in signal model and the resultant element-wise multiplication, it is not straightforward to apply the standard analysis to our proposed adaptive algorithm. To analyze the EMSE, we first rewrite $\mathcal{E}|e_a|^2$ as

$$\rho = \lim_{n\to\infty}\mathcal{E}|e_{w,1}^*[n]t_1[n] + \cdots + e_{w,M}^*[n]t_M[n]|^2\ . \tag{52}$$

Now we establish the following assumption:

Assumption 1: when $n\to\infty$, $e_{w,i}$ and $e_{w,j}$ are uncorrelated for $i \neq j$.

Then (52) can be changed to

$$\rho = \lim_{n\to\infty}\mathcal{E}|e_a|^2 = \lim_{n\to\infty}\sum_{i=1}^{M}\mathcal{E}|e_{a,i}|^2 = \lim_{n\to\infty}\sum_{i=1}^{M}\mathcal{E}|e_{w,i}^*[n]t_i[n]|^2 \ , \quad (53)$$

with $e_{a,i} = e_{w,i}^*[n]t_i[n]$.

In the next we will focus on $\lim_{n\to\infty}|e_{a,i}|^2$ only. Note that (18) can be written as

$$e_{w,i}[n+1] = e_{w,i}[n] - \mu e_i^*[n]t_i[n] \ , \quad (54)$$

where $e_i[n] = (r^*[n] - \bar{y}_i^*[n])$. Taking the Hermitian transpose and multiplying $\mathbf{t}[n]$ with both sides of (54), we have

$$e_{p,i} = e_{a,i} - \mu e_i[n]|t_i[n]|^2 \ , \quad (55)$$

where $e_{p,i} = (w_{o,i} - w_i[n+1])^* t_i[n]$ is a posteriori estimation error. Solving for $e_i[n]$, we have

$$e_i[n] = \frac{e_{a,i} - e_{p,i}}{\mu|t_i[n]|^2} \ . \quad (56)$$

Substituting (56) into (54), we have

$$e_{w,i}[n+1] = e_{w,i}[n] - \frac{(e_{a,i} - e_{p,i})^*}{t_i^*[n]} \ . \quad (57)$$

We can rearrange and take the modulus of (57) as

$$|e_{w,i}[n+1] + \frac{e_{a,i}^*}{|t_i^*[n]}|^2 = |e_{w,i}[n] + \frac{e_{p,i}^*}{t_i^*[n]}|^2 \ . \quad (58)$$

Taking expectation and canceling the uncorrelated items, (58) can be simplified as

$$\mathcal{E}|e_{w,i}[n+1]|^2 + \mathcal{E}|\frac{e_{a,i}^*}{t_i^*[n]}|^2 = \mathcal{E}|e_{w,i}[n]|^2 + \mathcal{E}|\frac{e_{p,i}^*}{t_i^*[n]}|^2 \ , \quad (59)$$

which is the so-called energy conservation relation. However, different from the one derived from [38] that contains all of the reserved energy of the array, (59) is the reserved energy for one node only. When the filter operation is in steady state for $n \to \infty$, we can also write

$$\mathcal{E}|e_{w,i}[n+1]|^2 = \mathcal{E}|e_{w,i}[n]|^2 \ . \quad (60)$$

Then (59) is reduced to $\mathcal{E}\frac{|e_{a,i}^*|^2}{|t_i[n]|^2} = \mathcal{E}\frac{|e_{p,i}^*|^2}{|t_i[n]|^2}$. With (55),

$$\mathcal{E}\frac{|e_{a,i}^*|^2}{|t_i[n]|^2} = \mathcal{E}\frac{|(e_{a,i} - \mu e_i[n]|t_i[n]|^2)^*|^2}{|t_i[n]|^2} \ . \quad (61)$$

Expanding and simplifying (61), we can rearrange it as

$$\mathcal{E}\{\mu|t_i[n]|^2|e_i[n]|^2\} = 2\mathcal{E}\{\mathrm{Re}(e_{a,i}^* e_i[n])\} \ . \quad (62)$$

For $e_i[n]$, we have

$$\begin{aligned} &e_i[n] = r[n] - u_i\tilde{y}_i[n] \\ &= r[n] - u_i g_i \mathbf{w}^H[n]\mathbf{t}[n] - u_i g_i \eta[n] - u_i v_i[n] \\ &= r[n] - u_i g_i \mathbf{w}_o^H \mathbf{t}[n] - u_i g_i e_a[n] - u_i g_i \eta[n] - u_i v_i[n] \ . \end{aligned} \quad (63)$$

With the optimum solution, $\mathbf{w}_o^H\mathbf{t}[n] = r[n] + v[n]$, where $v[n]$ is the noise, respectively.

Assumption 2: In the steady-state, $|t_i[n]|^2$ and $|e_i[n]|^2$ are uncorrelated [38]. Then

$$\mathcal{E}(\mu|t_i[n]|^2|e_i[n]|^2) = \mu\mathcal{E}|t_i[n]|^2[a_i + b_i + |u_i g_i|^2|e_a|^2)] \quad (64)$$

and

$$\begin{aligned} 2\mathcal{E}\{\mathrm{Re}(e_{a,i}^* e_i[n])\} &= 2\mathcal{E}\{\mathrm{Re}(u_i g_i e_{a,i}^* \sum_{i=1}^{M} e_{a,i}[n])\} \\ &= 2\mathrm{Re}(u_i g_i)\mathcal{E}|e_{a,i}|^2 \ , \end{aligned} \quad (65)$$

where $a_i = |1 - u_i g_i|^2 P_s$ and $b_i = |u_i g_i|^2(\sigma_v^2 + \sigma_\eta^2) + |u_i|^2\sigma_v^2$, $P_s = \mathcal{E}|r[n]|^2$. Taking i from 1 to M and summing up (65) and (64), separately, let the two results be equal to each other, which leads to

$$\begin{aligned} &2\sum_{i=1}^{M}\{\mathrm{Re}(u_i g_i)|e_{a,i}|^2\} \\ &= \mu\mathcal{E}\sum_{i=1}^{M}|t_i[n]|^2[a_i + b_i + |u_i g_i|^2|e_a|^2)] \ . \end{aligned} \quad (66)$$

For the RSB adaptive algorithm, since the optimum coefficient $u_{r,i} = 1/g_i$, $a_i = 0$ and $b_i = \sigma_v^2 + \sigma_\eta^2 + \sigma_v^2/g_i^2$. Then (66) can be reduced to

$$2|e_a|^2 = \mu\sum_{i=1}^{M}|t_i|^2(\sigma_v^2 + \sigma_\eta^2 + \sigma_v^2/g_i^2 + |e_a|^2) \ . \quad (67)$$

We can obtain $|e_a|^2$ as

$$|e_a|^2 = \frac{\mu\sum_{i=1}^{M}|t_i|^2(\sigma_v^2 + \sigma_\eta^2 + \sigma_v^2/g_i^2)}{2 - \mu\sum_{i=1}^{M}|\bar{t}_i|^2} \ . \quad (68)$$

V. Simulation Results

In our MATLAB simulations, the signal power is normalized as $\sigma_s^2 = 1$ and the input SNR is adjusted by changing the noise power. We set $\sigma_\eta^2 = \sigma_v^2 = \sigma_n^2 = -20$ dB (i.e. link SNR = 20 dB) unless otherwise specified. However, for the traditional 2nd order statistics based algorithm, we assume that the links from the central node to all relay nodes have a higher SNR since this approach has a flexibility in selecting the central node among the source, relay nodes and destination. Therefore, we assume $\sigma_\varepsilon^2 = \sigma_\epsilon^2 = -30$ dB. $\mathbf{f}$ and $\mathbf{g}$ are generated following the Rayleigh distribution with unit variance. Since any node with low quality channel can not contribute to the performance very much, a relay node selection process is adopted and only the channel gain $|f_i|^2 > u_T$ and $|g_i|^2 > u_T$ will be selected for cooperation, where u_T is the threshold and here we use $u_T = 0.25$ for all simulations. The relay node number is $M = 30$ and the stepsizes are $\mu_r = 0.004$, $\mu_s = \mu_c = \mu_{it} = 0.008$ and $\mu_u = 0.001$ unless otherwise specified. The initial values for the weight vector is set as $\mathbf{w}[1] = 0.01(ones(M,1) + j \cdot ones(M,1))/\sqrt{2}$. The transmitted signals for both RSB and CMB algorithms are QPSK (Quadrature phase-shift keying) modulated and in the RSB case, the reference $r[n] = s[n]$ is assumed known at the destination.

1) Simulation 1: We first examine the MSE performance of the traditional 2nd order statistics based algorithm in (10), the direct RSB (D-RSB) algorithm in (7) (assuming $y[n]$ is known at each node), and the two proposed RSB adaptive algorithms: the unconstrained RSB (U-RSB) and the constrained RSB (C-RSB). The results are shown in Fig. 2. The normalized

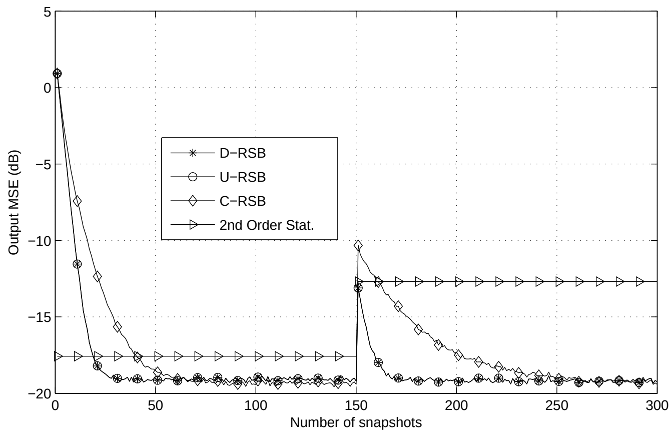


Fig. 2. Simulation 1: MSE for algorithms in (10) and (7), and the two proposed RSB adaptive algorithms.

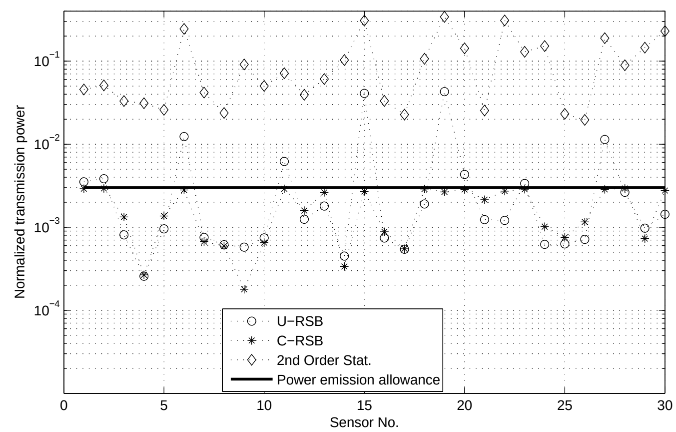


Fig. 3. Simulation 1: Individual node power emission for RSB and 2nd order statistics based algorithms ($M = 30$).

transmission power for each node is the same and $p_1 = \cdots = p_M = 0.003$. There are two stages in the simulation: at the first stage (snapshot number n from 1 to 150) we have $M = 30$ nodes; at the second stage (n from 151 onwards), 10 sensors exit the relay network. The purpose of this setup is to demonstrate the tracking capability of the proposed algorithms, which is based on the scenario that during the communication (adaptation), some relay nodes in the network have to exit the system due to reasons such as low battery power, or some nodes moving out of the effective channel, or more possibly, some nodes' own need for communication. From Fig. 2 we can see that all of the proposed algorithms have adapted to the new environment very quickly and reached the steady-state within only few snapshots. On the contrary, the traditional 2nd order statistics based algorithm shows significant performance loss after 10 sensors leave the network.

Moreover, the U-RSB result has almost completely overlapped with that of D-RSB, which means that the method by first feeding back $y[n]$ to all relay nodes and then estimating its value is quite effective. On the other hand, the unconstrained algorithms (U-RSB and D-RSB) had a faster convergence speed than C-RSB, but converged to a little higher steady-state MSE, which is because the convergence speed and steady-state MSE are affected by the values of stepsize. A larger setpsize leads to a faster convergence but comes with a higher steady-state MSE, while a smaller stepsize results in a slower convergence speed but also a lower steady-state MSE.

For the first stage (snapshot number n from 1 to 150, $M = 30$), in terms of MSE, compared to the 2nd order statistics based algorithm in (10), the proposed U-RSB has achieved a lower steady-state MSE. The reason for this is that the feedback for $\mathbf{f}$ and $\mathbf{w}_D$ introduced two estimation errors, while the proposed algorithm only has one estimation error (i.e. $y[n]$). In terms of overhead, the proposed U-RSB can converge to the same MSE as the 2nd order statistics based algorithm in (10) by using only 20 snapshots for training. According to the analysis in Section IV-A, the total overhead of the proposed algorithm is $3K + 1 = 61$ symbols, while for the 2nd order statistics based algorithm, the total overhead is $3M + 1 = 91$, an increase by a factor of 1.5. At the second stage with $M = 20$, the overhead of the 2nd order statistics based algorithm will be reduced to the same value (i.e. 61) as the proposed adaptive algorithms.

Fig. 3 shows the transmission power at each relay node at the first stage ($M = 30$). For the U-RSB, it varies significantly. 8 out of 30 relay nodes used more power than the normalized power allowance. However, for the C-RSB, all of the relay nodes used less power than the allowance. Considering that the constrained algorithm achieved a lower MSE with the given stepsize and data length, the power has been used much more efficiently in this case. Interestingly, to achieve the optimal solution by 2nd order statistics based algorithm, all of the individual relay node has consumed much more power than the proposed adaptive algorithms (i.e. U-RSB), which implies that the proposed U-RSB is more energy efficient.

2) Simulation 2: In this set of simulations, the performance of our proposed unconstrained SCMB (U-SCMB), unconstrained CCMB (U-CCMB), and unconstrained ICMB (U-ICMB) algorithms are compared with the constrained SCMB (C-SCMB), constrained CCMB (C-CCMB), and constrained ICMB (C-ICMB) algorithms. The results are shown in Fig. 4. We used individual power allowance $p_1 = \cdots = p_M = 0.003$. All the other settings are the same as in Simulation 1. From Fig. 4 we can see that the unconstrained algorithms (U-SCMB, U-CCMB, U-ICMB) have converged faster than the constrained algorithms, however, the steady-state MSE is a little higher as a cost. Again, the results show that the proposed adaptive algorithms has very prompt response to the node leaving the relay networks. Note that the three unconstrained algorithms (also the three constrained algorithms) have a very similar performance, which leads to almost overlapped learning curves.

As shown in Fig. 5, for the three unconstrained algorithms, most of the relay nodes used much more power than the set allowance and the distribution is extremely uneven. However, with the added constraints for all of the constrained CMB algorithms, the individual power consumption at each node has become less than the allowance.

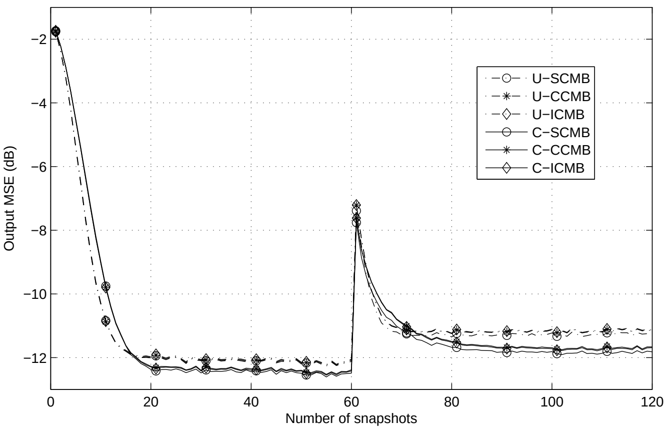


Fig. 4. Simulation 2: Output MSE versus the number of snapshots for the CMB algorithms, with $M = 30$ for the first 150 snapshots and $M = 20$ afterwards that.

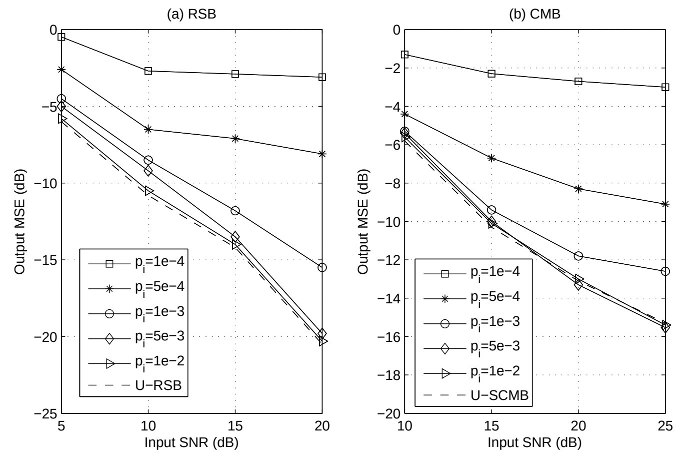


Fig. 6. Simulation 3: Output MSE versus the input SNR for different p_i. Solid lines are C-RSB and C-SCMB in sub-figure (a) and (b), respectively.

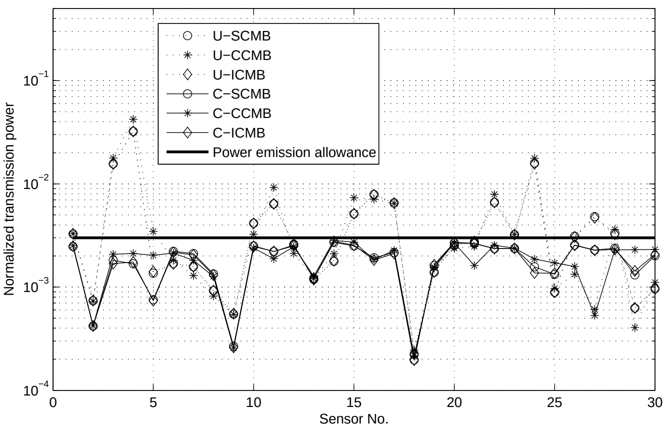


Fig. 5. Simulation 2: Individual node power emission for CMB algorithms ($M = 30$).

3) Simulation 3: The output MSE of the constrained algorithms versus power consumption (p_i) for different input SNR values is shown in Fig. 6, where the left one is for the RSB algorithms and all of the settings are the same as in Simulation 1 except for the variable p_i; the right side subplot of the figure is for the CMB algorithms with the same parameters as in Simulation 2. Since the CMB algorithms have a similar performance, in this set of simulations, we only compare the pair of SCMB (C-SCMB and U-SCMB) algorithms to try to find a proper individual power constraint level p_i. In order to show better steady-state MSE performance and reduce the large stepsize caused high error-floor, the stepsizes in this simulation are reduced to $\mu_r = 0.001$, $\mu_s = 0.003$, respectively. From the figure, with the input SNR increasing, all of the MSE values with different p_i have decreased and for the unconstrained algorithm, it dropped almost linearly.

For the RSB case, when normalized transmission power $p_i = 0.0001$ and 0.0005, the MSE performance of the constrained algorithm is much worse than the unconstrained one, especially for high input SNRs, which is due to that the power allowance for each node is so low that even with maximum allowed power, it still can not reduce the MSE effectively. Form the figure we can see that when $p_i = 0.005$ and larger, the constrained and unconstrained ones can reach nearly the same MSE for different SNRs. With an even larger p_i, the MSE of the constrained algorithm increased to the same level as the unconstrained one since in that case, the individual power constraint is so large that no normalization is really needed any more in the C-RSB, which is then reduced to the unconstrained one. A similar trend can be observed in CMB on the right side subplot of Fig. 6. In particular, for $p_i = 0.005$, the unconstrained algorithms outperformed the constrained ones at high SNRs (e.g. 20 dB and 25 dB), while at lower SNRs (e.g., 10 dB, 15 dB), it is the opposite.

4) Simulation 4: In this set of simulations, the excess MSE derived in (68) and the simulated results were compared, as shown in Fig. 7 for different SNR values from $5, 10, 15$ to 20 dB. With the number of samples increasing, the EMSE values dropped sharply at the beginning. For the small SNR (for example, 5dB) scenario, $|e_a|^2$ achieved the theoretic value at about the 2500-th snapshot. With the input SNR increasing, $|e_a|^2$ gets smaller according to (68), and it takes longer for the curves to converge to the theoretic values. For example, it takes 40000 snapshots for the 20dB case, much more than the 5 dB case. Overall, the results show that even with a high input SNR, our theoretical analysis matches the simulations very well.

5) Simulation 5: The uncoded bit error rate (BER) of the RSB and CMB algorithms at the steady state (after 60 snapshots in our simulations) in terms of the energy per bit (E_b) to noise power spectral density N_0 ratio (E_b/N_0) are shown in Fig. 8. For the RSB related results in Fig. 8 (a), we can see that the proposed algorithms (C-RSB and U-RSB) show around 2.5 dB gain compared to the 2nd-order statistics based algorithm (there are no enough samples to estimate the 2nd-order statistics), while the D-RSB and U-RSB curves almost coincide with each other completely, which is consistent with the BER performance in Simulation 1.

Note that the CMB scheme is a blind approach and due to less information available to the system, its overall perfor-

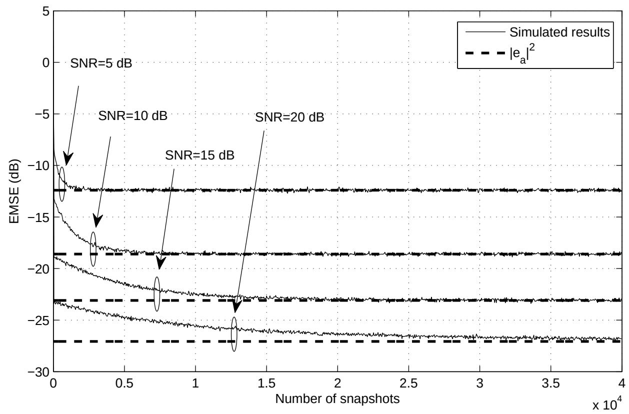


Fig. 7. Simulation 4: Excess MSE versus the number of snapshots for the adaptive unconstrained RSB algorithm.

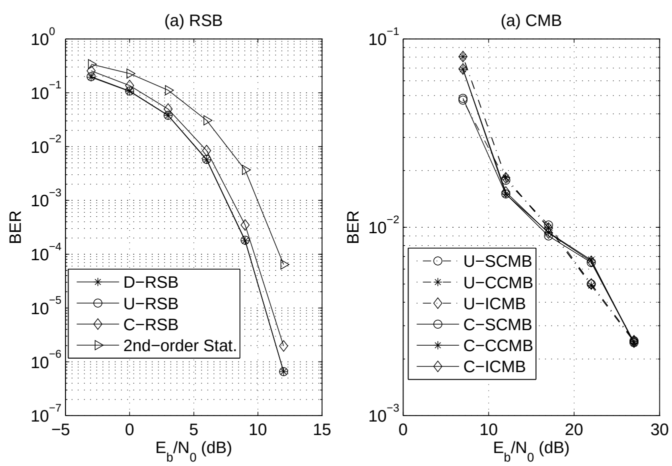


Fig. 8. Simulation 5: BER versus input E_b/N_0.

mance is not as good as the RSB scheme, especially at low SNR scenarios. For moderate and high SNR ranges, the BER curves of the CMB algorithms are shown in Fig. 8 (b), with a fixed $\mu_1 = 0.004$ for all CMB algorithms and all other parameters are kept the same as in Simulation 3. We see all of the algorithms have a similar performance, which matches the MSE analysis shown in Simulation 2.

VI. Conclusions

In this paper, by estimating the destination received data through a feedback link to the relay network, a class of adaptive distributed beamforming algorithms has been proposed which can achieve a globally optimum solution by employing only local channel state information. With the MMSE criterion, a reference signal based scheme is first derived, followed by a constant modulus based beamforming scheme when the reference signal is not available. The overhead and stepsize range for the derived algorithms were provided and the EMSE for the RSB case was studied based on the energy reservation method. The proposed algorithms have low computational complexity and overhead. As shown in simulations, they have outperformed the traditional 2nd order statistics based algorithm in terms of MSE and BER performance, overhead, tracking capability and energy efficiency.

References

[1] Z. Yi and I.-M. Kim, "Joint optimization of relay-precoders and decoders with partial channel side information in cooperative networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 447–458, 2007.

[2] R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 1754–1763, 2007.

[3] R. Mudumbai, D. R. Brown, U. Madhow, and H. V. Poor, "Distributed transmit beamforming: Challenges and recent progress," *IEEE Communications Magazine*, vol. 47, pp. 102–110, 2009.

[4] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed transmit beamforming using feedback control," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 411 –426, jan. 2010.

[5] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - part 1: System description," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[6] ——, "User cooperation diversity - part ii: Implementation aspects and performance analysis," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[7] S. Talwar, Y. Jing, and S. Shahbazpanahi, "Joint relay selection and power allocation for two-way relay networks," *IEEE Signal Processing Letters*, vol. 18, no. 2, pp. 91 –94, feb. 2011.

[8] M. Abdallah and H. Papadopoulos, "Beamforming algorithms for information relaying in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4772 –4784, oct. 2008.

[9] L. Zhang, W. Liu, and J. Li, "Low-complexity distributed beamforming for relay networks with real-valued implementation," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 5039–5048, Oct 2013.

[10] A. Adinoyi and H. Yanikomeroglu, "Cooperative relaying in multi-antenna fixed relay networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 533–544, Feb. 2007.

[11] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[12] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 362–371, Feb. 2004.

[13] Y. D. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Transactions on Information Theory*, vol. 55, pp. 2499–2517, 2009.

[14] L. Chen, K.-K. Wong, H. Chen, J. Liu, and G. Zheng, "Optimizing transmitter-receiver collaborative-relay beamforming with perfect CSI," *IEEE Communications Letters*, vol. 15, pp. 314–316, 2011.

[15] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z. Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Transactions on Signal Processing*, vol. 56, pp. 4306–4316, 2008.

[16] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1238–1250, Mar. 2010.

[17] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Transactions On Signal Processing*, vol. 57, no. 8, pp. 3053–3062, Aug. 2009.

[18] G. Zheng, K. K. Wong, A. Paulraj, and B. Ottersten, "Robust collaborative-relay beamforming," *IEEE Transactions on Signal Processing*, vol. 57, pp. 3130–3143, 2009.

[19] J. Kim, J. Choi, and J. Cioffi, "Cooperative distributed beamforming with outdated CSI and channel estimation errors," *IEEE Transactions on Communications*, vol. 62, no. 12, pp. 4269–4280, Dec 2014.

[20] A. El-Keyi and B. Champagne, "Collaborative uplink transmit beamforming with robustness against channel estimation errors," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 126 –139, jan. 2009.

[21] N. Khajehnouri and A. Sayed, "Distributed mmse relay strategies for wireless sensor networks," *IEEE Transaction on Signal Processing*, vol. 55, pp. 3336 – 3348, 2007.

[22] J. Choi, “MMSE-based distributed beamforming in cooperative relay networks,” *IEEE Transactions on Communications*, vol. 59, no. 5, pp. 1346–1356, May 2011.

[23] ——, “Distributed beamforming using a consensus algorithm for cooperative relay networks,” *IEEE Communications Letters*, vol. 15, no. 4, pp. 368–370, Apr. 2011.

[24] C.-C. Chen, C.-S. Tseng, J. Denis, and C. Lin, “Adaptive distributed beamforming for amplify-and-forward relay networks: Convergence analysis,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4167–4178, Aug 2014.

[25] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs: Prentice Hall, 1996.

[26] L. Zhang, W. Liu, and R. Langley, “A class of constrained adaptive beamforming algorithms based on uniform linear arrays,” *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3916–3922, July 2010.

[27] L. C. Godara, “Application of antenna arrays to mobile communications, Part II: beam-forming and direction-of-arrival estimation,” *Proceedings of the IEEE*, vol. 85, pp. 1195–1245, 1997.

[28] H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*. New York, U.S.A.: John Wiley & Sons, Inc., 2002.

[29] A. J. van der Veen and A. Paulraj, “A Analytical Constant Modulus Algorithm,” *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[30] C. R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, “Blind equalization using the constant modulus criterion: A review,” *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1927–1950, October 1998.

[31] J. Li and P. Stoica, *Robust adaptive beamforming*. New Jersey, U.S.A.: John Wiley & Sons, Inc., 2005.

[32] R. Gooch and J. Lundell, “CM array: an adaptive beamformer for constant modulus signals,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, New York, NY, 1986, pp. 2523–2526.

[33] L. Zhang, W. Liu, and R. J. Langley, “A class of constant modulus algorithms for uniform linear arrays with a conjugate symmetric constraint,” *Signal Processing*, vol. 90, pp. 2760–2765, 2010.

[34] J. Y. Mai and A. Sayed, “A feedback approach to the steady-state performance of fractionally spaced blind adaptive equalizers,” *IEEE Transactions on Signal Processing*, vol. 48, pp. 80–91, 2000.

[35] S. K. Nagaraj, S. Gollamudi and Y. F. Huang, “BEACON: An adaptive set-membership filtering technique with sparse updates,” *IEEE Transactions on Signal Processing*, vol. 47, pp. 2928–2941, 1999.

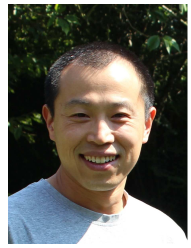
[36] P. Diniz, *Adaptive Filtering: Algorithms and Practical Implementations*, 3rd ed. Springer, Boston, MA, 2008.

[37] M. Honig, U. Madhow, and S. Verdu, “Blind adaptive multiuser detection,” *IEEE Transactions on Information Theory*, vol. 41, pp. 944–960, 1995.

[38] A. H. Sayed, *Adaptive Filters*. New Jersey, U.S.A.: John Wiley & Sons, Inc., 2008.

[39] Y. Cai and R. C. de Lamare, “Low-complexity variable step-size mechanism for code-constrained constant modulus stochastic gradient algorithms applied to CDMA interference suppression,” *IEEE Transactions on Signal Processing*, vol. 57, pp. 313–323, 2009.

Lei Zhang received his B.Eng. degree in Communication Engineering in 2005 and M.Sc. degree in Electromagnetic Fields and Microwave Technology in 2008, both from the Northwestern Polytechnic University, Xi’an, China, and Ph.D. in 2011 from the Communications Research Group, University of Sheffield, UK. He is currently a research fellow in Wireless Communications in the Institute of Communications (ICS), University of Surrey, UK. Before he joined ICS, he was a research engineer in Huawei technologies, China. His research interests include multi-antenna signal processing, air interface waveform design, cloud radio access networks, massive MIMO systems.



Wei Liu received his B.Sc. in 1996 and L.L.B. in 1997, both from Peking University, China, M.Phil. from University of Hong Kong, in 2001, and Ph.D. in 2003 from the School of Electronics and Computer Science, University of Southampton, UK. He later worked as a postdoc at Imperial College London. Since September 2005, he has been with the Department of Electronic and Electrical Engineering, University of Sheffield, UK, first as a lecturer, and then a senior lecturer. His research interests are in sensor array signal processing, blind signal processing, multirate signal processing and their various applications. He has now published about 170 journal and conference papers, three book chapters, and a research monograph about wideband beamforming (“Wideband Beamforming: Concepts and Techniques”, John Wiley & Sons, March 2010, webpage: www.zepler.net/$\sim$w.liu). He is a senior member of IEEE and currently an associate editor for IEEE Trans. Signal Processing.

Atta ul Quddus received the MSc degree in Satellite Communications and PhD degree in Mobile Cellular Communications, both from University of Surrey UK in 2000 and 2005, respectively. He is currently a Lecturer in Wireless Communications in the Institute of Communications (ICS), Department of Electronic Engineering, University of Surrey UK. During his research career, he has led several national and international research projects that also contributed towards 3GPP standardization. In 2004, he also won the Centre for Communications Systems (CCSR) Research Excellence Prize sponsored by Vodafone for his research on Adaptive Filtering algorithms. His current research interests include Machine Type Communication, Cloud Radio Access Networks, and Device to Device Communication.

Mehrdad Dianati received his BSc. and MSc. both in Electrical Engineering from Sharif University of Technology and K.N.T University of Technology, Iran. He completed his PhD in Electrical and Computer Engineering at the University of Waterloo in Canada. He is currently a Reader in Communications and Networking Systems. His research area mainly includes: 1) analysis and design of advanced cross layer optimised PHY/MAC layer techniques for 5G wireless access networks, and 2) communication network design for Intelligent Transport Systems. He has been involved in a number of national and international projects in the area of communications and networking as technical coordinator and work-package leader. In addition to his research experience, Mehrdad has 9 years of industrial as software/hardware developer and technical leader. He is an associate editor for IEEE Transactions on Vehicular Technology, IET Communications and Wiley’s Journal of Wireless Communications and Mobile. He has also been actively involved in organizing committees of numerous international conferences in recent years.

Rahim Tafazolli is a professor and the Director of the Institute for Communication Systems (ICS) and 5G Innovation Centre (5GIC), the University of Surrey in the UK. He has published more than 500 research papers in refereed journals, international conferences and as invited speaker. He is the editor of two books on “Technologies for Wireless Future” published by Wiley’s Vol.1 in 2004 and Vol.2 2006.