Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Joint adaptive power allocation and interference suppression algorithms based on the MSER criterion for wireless sensor networks*

Gui-jie WANG, Yun-long CAI, Min-jian ZHAO‡, Jie ZHONG
(*Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China*)
E-mail: wgj8711@zju.edu.cn; ylcai@zju.edu.cn; mjzhao@zju.edu.cn; zhongjie@zju.edu.cn
Received Feb. 6, 2014; Revision accepted May 5, 2014; Crosschecked Sept. 17, 2014

Abstract: In this study, a two-hop wireless sensor network with multiple relay nodes is considered where the amplify-and-forward (AF) scheme is employed. Two algorithms are presented to jointly consider interference suppression and power allocation (PA) based on the minimization of the symbol error rate (SER) criterion. A stochastic gradient (SG) algorithm is developed on the basis of the minimum-SER (MSER) criterion to jointly update the parameter vectors that allocate the power levels among the relay sensors subject to a total power constraint and the linear receiver. In addition, a conjugate gradient (CG) algorithm is developed on the basis of the SER criterion. A centralized algorithm is designed at the fusion center. Destination nodes transmit the quantized information of the PA vector to the relay nodes through a limited-feedback channel. The complexity and convergence analysis of the proposed algorithms are carried out. Simulation results show that the proposed two adaptive algorithms significantly outperform the other previously reported algorithms.

Key words: Cooperative communications, Adaptive filtering, Symbol error rate (SER), Interference suppression, Power allocation
doi:10.1631/jzus.C1400034 **Document code:** A **CLC number:** TP391.41

1 Introduction

Nowadays, because of the low cost, scalability, and robustness, the wireless sensor networks (WSNs) which allow a wide range of applications in the areas of defence, environment, health, and domestic use have received significant attention (Akyildiz *et al.*, 2002). WSNs usually deploy geographically distributed sensor nodes to transmit their data to the desired destination through multihop relays (Laneman *et al.*, 2004). The sensors typically have limited communication capability and energy resources in the network. Usually, each sensor in the WSNs produces a local analog or digital signal after making an observation of the desired signal, and then sends the information to the fusion center which combines the received sensor quantity to generate a final estimate of the observed information. Recently, many different relay transmission techniques have been developed, such as amplify-and-forward (AF), decode-and-forward (DF), and compress-and-forward (CF) (Laneman *et al.*, 2004; Kramer *et al.*, 2005; Rui, 2010). In fact, the use of spatial diversity with cooperation among the nodes can significantly enhance the performance and capacity of WSNs (Laneman *et al.*, 2004; Clarke and de Lamare, 2011; 2012).

Adaptive signal processing is of great importance in the modern information society, and it has been widely used in communication systems (de

‡ Corresponding author
* Project supported by the National Natural Science Foundation of China (Nos. 61101103 and 61201230) and the 5th Generation Mobile Communication Program in China (No. 2014AA01A707)

Lamare and Sampaio-Neto, 2003; Li and Hamouda, 2007; Chen *et al.*, 2011). Most of the existing adaptive detection methods are based on the minimum mean square error (MMSE) criterion for wireless communications (Verdu, 1998; Wang T *et al.*, 2012). However, MMSE is not the most appropriate metric in digital communication systems since it has been recognized that the method which minimizes the MSE criterion does not necessarily produce the minimum symbol error rate (MSER) or minimum bit error rate (MBER) performance. Several adaptive detection algorithms based on the minimization of SER have been proposed and the MSER algorithm (Chen *et al.*, 2004a; 2008a; 2008b) for adaptive interference suppression techniques is one of the most successful and suitable approaches.

The popular stochastic gradient (SG) descent algorithm is a least mean squares (LMS) adaptive method. The SG method is used to minimize a cost function with a format of a sum of differentiable functions. In addition, the conjugate gradient (CG) method is used mainly for the numerical solution of particular systems of linear equations. The matrix in these particular systems is positive-definite and symmetric. Due to the CG algorithm being an iterative method, it is widely used in the sparse systems which are difficult to solve by direct techniques such as the Cholesky decomposition. In addition, the CG algorithm can be employed to handle unconstrained optimization problems such as energy minimization (Straeter, 1971).

The majority of previous literature considers a scenario in which all nodes use the same power to transmit signals in the system. Without loss of generality, this 'constant power scheme' employed in real communication wastes more power than an adaptive power scheme. To the best of our knowledge, most of the existing algorithms employed in cooperative communication systems do not jointly consider the issues of interference mitigation and power allocation (PA). However, these two issues play an important role in sensor networks and ad-hoc wireless cooperative cellular systems (Souryal *et al.*, 2006; Fischione *et al.*, 2009).

In this study, we consider a two-hop WSN with multiple relay nodes where the AF scheme is used. Usually, a large number of densely deployed sensing devices in WSNs can transmit their signal to the desired user by multihop relays. The most important design characteristics of physical-layer methods employed for WSNs and communication protocols (Wang Y *et al.*, 2012) are of low complexity and high energy efficiency. Because of some limitations such as sensor node power, computational capacity, and memory, some power-constrained relay methods (Krishna *et al.*, 2008; Liu *et al.*, 2012) and PA strategies (Li *et al.*, 2007) have been proposed for WSNs to obtain the best possible quality of service (QoS) at the destinations. Hence, due to low complexity and low energy requirements, we propose a joint iterative PA and interference suppression algorithm on the basis of the minimization of an symbol error rate (SER) cost function for WSNs that employ the AF relaying strategy.

We propose an SG algorithm and a CG algorithm which are derived on the basis of the MSER expressions to jointly update the parameter vectors that allocate the power among the relays subject to a power constraint and the linear receiver. In addition, feedback PA coefficients are considered in the two proposed algorithms. Note that both relay nodes and destination nodes know the codebook designed off-line. Each destination node selects an index from the codebook of the quantized PA vector based on the fusion center estimation, and transmits it to the relay node through a limited-feedback channel. We refer to the proposed adaptive joint power allocation and interference suppression algorithm on the basis of MSER as JMSER. Simulation results show that the proposed adaptive algorithms significantly outperform the other state-of-the-art schemes. The main contributions of this study are summarized as follows:

1. On the basis of the minimization of the SER cost function for WSNs with quadrature phase shift keying (QPSK) modulation, the theory estimation algorithm which jointly considers PA and interference suppression is proposed.

2. Two sample-by-sample based jointly adaptive implementations of the theoretical JMSER solution are developed on the basis of the kernel density estimate of the probability density function (PDF).

3. The analyses of convergence and computational complexity of the proposed adaptive algorithms are carried out.

The superscripts 'T', 'H', and '*' denote the transpose, Hermitian transpose, and elementwise conjugate, respectively. A bold symbol denotes a matrix or a vector. The symbols $|\cdot|$, $\|\cdot\|$, $\Re(\cdot)$, $\Im(\cdot)$ $\mathrm{sgn}(\cdot)$, $\mathrm{diag}(\cdot)$, $\exp(\cdot)$, and $\mathrm{Prob}(\cdot)$ represent the 2-norm of a scalar, norm of a vector, real part of a complex number, imaginary part of a complex number, signum function, diagonal form of a vector, exponential function, and probability function, respectively.

2 System model and problem statement

In this study, we employ a system model similar to that introduced by Wang T *et al.* (2012). Their work is not an adaptive algorithm which is based on the MMSE criterion. However, our work focuses on the minimization of SER rather than MMSE, because we have known that the method which minimizes the MSE criterion does not necessarily produce the MSER or MBER performance. This point can also be proved by Section 6

As shown in Fig. 1, the WSN consists of N_s source nodes, N_r relay nodes, and N_d destination nodes. The received signal at the relay nodes can be expressed as

$$\boldsymbol{r} = \boldsymbol{H}_\mathrm{s}\boldsymbol{b} + \boldsymbol{n}_\mathrm{r}, \tag{1}$$

where $\boldsymbol{H}_\mathrm{s}$ denotes the $N_\mathrm{r} \times N_\mathrm{s}$ channel matrix between the source nodes and relay nodes, and it is given by $[\boldsymbol{h}_{\mathrm{s},1},\ \boldsymbol{h}_{\mathrm{s},2},\ \cdots,\ \boldsymbol{h}_{\mathrm{s},N_\mathrm{r}}]^\mathrm{T}$, where $\boldsymbol{h}_{\mathrm{s},r}$ ($r = 1,\ 2,\ \cdots,\ N_\mathrm{r}$) is an $N_\mathrm{s} \times 1$ vector. The quantity $\boldsymbol{b}$ denotes the $N_\mathrm{s} \times 1$ transmitted QPSK signal vector. The elements of $\boldsymbol{b}$ are given by $\pm 1 \pm \mathrm{j}$ ($\mathrm{j}^2 = -1$). The vector $\boldsymbol{n}_\mathrm{r}$ is a zero-mean circularly symmetric complex additive white Gaussian noise (AWGN) vector with covariance matrix $\sigma_n^2\boldsymbol{I}$, where $\boldsymbol{I}$ denotes an identity matrix of appropriate dimension. Next, we normalize the power of the received signal for each relay node as follows:

$$\boldsymbol{t} = \boldsymbol{G}\boldsymbol{r}, \tag{2}$$

where

$$\boldsymbol{G}=\mathrm{diag}\left(\sigma_\mathrm{b}^2||\boldsymbol{h}_{\mathrm{s},1}||^2 + \sigma_n^2, \cdots, \sigma_\mathrm{b}^2||\boldsymbol{h}_{\mathrm{s},N_\mathrm{r}}||^2 + \sigma_n^2\right)^{-\frac{1}{2}} \tag{3}$$

is a normalization matrix and σ_b^2 denotes the power of each transmitted signal.

The received signal at the destination can be expressed as

$$\boldsymbol{d} = \boldsymbol{H}_\mathrm{d}\boldsymbol{A}\boldsymbol{t} + \boldsymbol{n}_\mathrm{d}, \tag{4}$$

where $\boldsymbol{H}_\mathrm{d}$ denotes the $N_\mathrm{d} \times N_\mathrm{r}$ channel matrix between the relay nodes and destination nodes, given by $[\boldsymbol{h}_{\mathrm{d},1},\ \boldsymbol{h}_{\mathrm{d},2},\ \cdots,\ \boldsymbol{h}_{\mathrm{d},N_\mathrm{d}}]^\mathrm{T}$, $\boldsymbol{n}_\mathrm{d}$ denotes the $N_\mathrm{d} \times 1$ destination noise vector with zero mean and covariance matrix $\sigma_\mathrm{d}^2\boldsymbol{I}$, and $\boldsymbol{A} = \mathrm{diag}(a_1,\ a_2,\ \cdots,\ a_{N_\mathrm{r}})$ is a diagonal matrix whose element represents the amplification coefficient of each relay node. Here we use the property of matrix-vector multiplication $\boldsymbol{A}\boldsymbol{t} = \boldsymbol{T}\boldsymbol{a}$ in the system expression, where $\boldsymbol{T} = \mathrm{diag}(t_1,\ t_2,\ \cdots,\ t_{N_\mathrm{r}})$, $\boldsymbol{t} = [t_1,\ t_2,\ \cdots,\ t_{N_\mathrm{r}}]^\mathrm{T}$, and $\boldsymbol{a} = [a_1,\ a_2,\ \cdots,\ a_{N_\mathrm{r}}]^\mathrm{T}$. Hence, we rewrite Eq. (4) as follows:

$$\boldsymbol{d} = \boldsymbol{H}_\mathrm{d}\boldsymbol{T}\boldsymbol{a} + \boldsymbol{n}_\mathrm{d}. \tag{5}$$

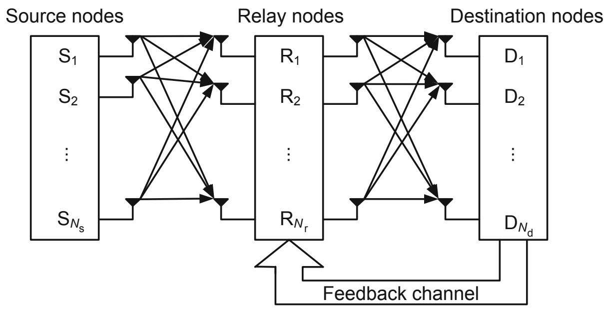

Fig. 1 A two-hop cooperative wireless sensor network

To simplify the derivation of the proposed algorithm in the next section, we employ the preceding related property of matrix-vector multiplication to rewrite Eq. (5) as follows:

$$\boldsymbol{d} = \boldsymbol{H}_\mathrm{d}\boldsymbol{G}\mathrm{diag}\left(\boldsymbol{H}_\mathrm{s}\boldsymbol{b}\right)\boldsymbol{a} + \boldsymbol{H}_\mathrm{d}\boldsymbol{G}\boldsymbol{N}_\mathrm{r}\boldsymbol{a} + \boldsymbol{n}_\mathrm{d}, \tag{6}$$

where $\boldsymbol{N}_\mathrm{r}$ is a diagonal matrix form of the relay noise vector $\boldsymbol{n}_\mathrm{r}$. The noise vector has zero mean and covariance matrix $\sigma_\mathrm{r}^2\boldsymbol{I}$. Notice that the second and third parts of Eq. (6) contain the additive noise. A fusion center which in practice contains the destination nodes can gather the channel state information and compute the optimal linear filters and the perfect PA vector coefficients. The fusion center can also send the information of the computed PA coefficients to relay nodes by a feedback channel.

Finally, the fusion center makes an estimation of the kth transmitted signal at the destination nodes as follows:

$$z_k = \boldsymbol{w}_k^\mathrm{H}\boldsymbol{d} = \hat{z}_k + n_k,\quad k = 1, 2, \cdots, N_\mathrm{s}, \tag{7}$$

where $\hat{z}_k = \boldsymbol{w}_k^\mathrm{H}\boldsymbol{H}_\mathrm{d}\boldsymbol{G}\mathrm{diag}\left(\boldsymbol{H}_\mathrm{s}\boldsymbol{b}\right)\boldsymbol{a}$, n_k is a Gaussian distributed variable with zero mean, $\boldsymbol{w}_k$ is an $N_\mathrm{d} \times 1$

complex-valued receiver weight vector, and z_k is a Gaussian distributed variable with a variance of $\sigma^2_{z_k}$. The variance of z_k is given by

$$\sigma^2_{z_k} = \sigma^2_{\mathrm{r}} \boldsymbol{a}^{\mathrm{H}} \overline{\boldsymbol{C}}^{\mathrm{H}}_k \overline{\boldsymbol{C}}_k \boldsymbol{a} + \sigma^2_{\mathrm{d}} \boldsymbol{w}^{\mathrm{H}}_k \boldsymbol{w}_k. \tag{8}$$

Let $\boldsymbol{w}^{\mathrm{H}}_k \boldsymbol{H}_{\mathrm{d}} \boldsymbol{G} = \overline{\boldsymbol{c}}_k$. Then $\overline{\boldsymbol{C}}_k = \mathrm{diag}(\overline{\boldsymbol{c}}_k)$ denotes the $N_{\mathrm{r}} \times N_{\mathrm{r}}$ diagonal matrix of vector $\overline{\boldsymbol{c}}_k$. Without loss of generality, in this work we assume that the first- and second-phase noises have the same power, that is $\sigma_{\mathrm{r}} = \sigma_{\mathrm{d}}$. Then, we obtain

$$\sigma^2_{z_k} = P_k \sigma^2_{\mathrm{d}}, \tag{9}$$

where $P_k = \boldsymbol{a}^{\mathrm{H}} \overline{\boldsymbol{C}}^{\mathrm{H}}_k \overline{\boldsymbol{C}}_k \boldsymbol{a} + \boldsymbol{w}^{\mathrm{H}}_k \boldsymbol{w}_k$.

The decision regarding the kth transmitted symbol b_k is made according to

$$\widehat{b}_k = \mathrm{sgn}(\Re(z_k)) + \mathrm{j}\,\mathrm{sgn}(\Im(z_k)). \tag{10}$$

3 Problem formulation for the joint PA-MSER scheme

In this section we jointly consider PA and interference suppression for the linear multisource transmission scheme, and then derive the theoretical MSER solution.

Since the WSN employs the QPSK modulation, we denote b^p_k for $1 \le p \le N$ as a possible transmitted symbol vector, where $N = 4^{N_{\mathrm{s}}}/4$. The MSER criterion is used in our proposed algorithm; hence, the error probability for the kth transmitted signal can be expressed as (Chen *et al.*, 2008b)

$$P_{\mathrm{e}}(\boldsymbol{w}_k, \boldsymbol{a}) = P_{E_{\mathrm{R}}} + P_{E_{\mathrm{I}}} - P_{E_{\mathrm{R}}} P_{E_{\mathrm{I}}}, \quad k = 1, 2, \cdots, N_{\mathrm{s}}, \tag{11}$$

where $P_{E_{\mathrm{R}}} = \mathrm{Prob}(y_{\mathrm{rs}} < 0)$, $P_{E_{\mathrm{I}}} = \mathrm{Prob}(y_{\mathrm{is}} < 0)$, $y_{\mathrm{rs}} = \mathrm{sgn}(\Re(b^p_k))\Re(\hat{z}_k)$ and $y_{\mathrm{is}} = \mathrm{sgn}(\Im(b^p_k))\Im(\hat{z}_k)$. The subscripts 'r' and 'i' represent the real and imaginary parts of a variable, respectively, and subscript 's' represents the source. We first consider the variance of the received signal estimation variable z_k. Given $b_k = +1 + \mathrm{j}$ ($k = 1, 2, \cdots, N_{\mathrm{s}}$), the theoretical PDF for y_{rs} and y_{is} can be expressed as follows:

$$\widetilde{p}(y_{\mathrm{rs}}| +1 + \mathrm{j}) = \frac{1}{N\sqrt{2\pi P_k}\sigma_{\mathrm{d}}} \sum_{p=1}^{N} \exp\left(- \frac{(y_{\mathrm{rs}} - \mathrm{sgn}[\Re(b^p_k)]\Re(\hat{z}_k))^2}{2\sigma^2_{\mathrm{d}} P_k} \right), \tag{12}$$

$$\widetilde{p}(y_{\mathrm{is}}| +1 + \mathrm{j}) = \frac{1}{N\sqrt{2\pi P_k}\sigma_{\mathrm{d}}} \sum_{p=1}^{N} \exp\left(- \frac{(y_{\mathrm{is}} - \mathrm{sgn}[\Im(b^p_k)]\Im(\hat{z}_k))^2}{2\sigma^2_{\mathrm{d}} P_k} \right). \tag{13}$$

The receiver and power allocation vectors $\boldsymbol{w}_k$ and $\boldsymbol{a}$ are defined as the receiving weight vector and PA vector, respectively. The upper bound of SER is given by

$$\widetilde{P}_{\mathrm{e}}(\boldsymbol{w}_k, \boldsymbol{a}) = P_{E_{\mathrm{R}}} + P_{E_{\mathrm{I}}} = \frac{1}{N} \left[\sum_{q=1}^{N} Q\left(\widetilde{C}_{\mathrm{R}}(\boldsymbol{w}_k, \boldsymbol{a}) \right) + \sum_{q=1}^{N} Q\left(\widetilde{C}_{\mathrm{I}}(\boldsymbol{w}_k, \boldsymbol{a}) \right) \right], \tag{14}$$

where

$$Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^{\infty} \exp(-\frac{x^2}{2}) \mathrm{d}x,$$

$$\widetilde{C}_{\mathrm{R}}(\boldsymbol{w}_k, \boldsymbol{a}) = \frac{\mathrm{sgn}(\Re(b^p_k))\Re(\hat{z}_k)}{\sigma_{\mathrm{d}}\sqrt{P_k}},$$

$$\widetilde{C}_{\mathrm{I}}(\boldsymbol{w}_k, \boldsymbol{a}) = \frac{\mathrm{sgn}(\Im(b^p_k))\Im(\hat{z}_k)}{\sigma_{\mathrm{d}}\sqrt{P_k}}.$$

The solution obtained by minimizing the upper bound (14) is practically equivalent to that of minimizing (11), since the bound $P_{\mathrm{e}}(\boldsymbol{w}_k, \boldsymbol{a}) < \widetilde{P}_{\mathrm{e}}(\boldsymbol{w}_k, \boldsymbol{a})$ is very tight. That is to say, $\widetilde{P}_{\mathrm{e}}(\boldsymbol{w}_k, \boldsymbol{a})$ is very close to the true SER $P_{\mathrm{e}}(\boldsymbol{w}_k, \boldsymbol{a})$ (Chen *et al.*, 2008b).

In the WSN, centralized detection is employed at the receiver end and the overall SER at the receiver is given by

$$\widetilde{P}_{\mathrm{E}}(\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_{N_{\mathrm{s}}}, \boldsymbol{a}) = \sum_{k=1}^{N_{\mathrm{s}}} \widetilde{P}_{\mathrm{e}}(\boldsymbol{w}_k, \boldsymbol{a}). \tag{15}$$

The JMSER method can be considered as the following optimization problem:

$$\left(\boldsymbol{w}^{\mathrm{JMSER}}_1, \boldsymbol{w}^{\mathrm{JMSER}}_2, \cdots, \boldsymbol{w}^{\mathrm{JMSER}}_{N_{\mathrm{s}}}, \boldsymbol{a}^{\mathrm{JMSER}}\right) = \arg \min_{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_{N_{\mathrm{s}}}, \boldsymbol{a}} \widetilde{P}_{\mathrm{E}}(\boldsymbol{w}_1, \cdots, \boldsymbol{w}_{N_{\mathrm{s}}}, \boldsymbol{a}). \tag{16}$$

Unlike the MMSE solution, it is very difficult to obtain a closed-form solution for the JMSER criterion. In this work we propose a jointly iterative optimization approach to obtain the optimal solution.

4 Adaptive JMSER algorithm

In this section, we derive the adaptive JMSER solution for the linear multisource detector and present two gradient search algorithms to find the JMSER solution. Here, the first algorithm is based on the SG method which is a gradient descent optimization method for minimizing a sum of differentiable objective functions. The second algorithm is based on the CG strategy which is used mainly for the numerical solution of particular systems of linear equations. The matrix in these particular systems is positive-definite and symmetric. In addition, we introduce the channel estimation algorithm employed in this work. In reality, the PDFs of $y_{\rm rs}$ and $y_{\rm is}$ are unknown. We adopt the temporal reference technique here to support the adaptive implementation of the JMSER technique.

4.1 SG-based adaptive JMSER algorithm

Effectively estimating the PDF is of great importance in adaptive implementation for the JMSER filtering design. We use the MSER criterion in our proposed algorithm. The error probability for the kth transmitted signal can be approximately expressed as follows:

$$\bar{P}_{\rm e}(\boldsymbol{w}_k,\boldsymbol{a}) = {\rm Prob}(y_{\rm r}<0)+{\rm Prob}(y_{\rm i}<0), \quad k=1,2,\cdots,N_{\rm s}, \tag{17}$$

where $y_{\rm r}={\rm sgn}(\Re(b_k))\Re(z_k)$, $y_{\rm i}={\rm sgn}(\Im(b_k))\Im(z_k)$.

This algorithm employs kernel density estimation (Silerman, 1996; Bowman and Azzalini, 1997; Chen *et al.*, 2004b), which can produce reliable PDF estimation with a short data record. In our application, it is obvious and natural to choose a Gaussian kernel function with a kernel width $\rho\sqrt{P_k}$ which is similar in form to the noise standard deviation $\sigma_{\rm d}\sqrt{P_k}$. Here ρ is the kernel width, working as a smoothing parameter and having a lower boundary $\rho=(\frac{4}{3N})^{\frac{1}{5}}\sigma_{\rm d}$ (Silerman, 1996).

A single-sample 'estimate' for PDFs $p(y_{\rm r}|+1+{\rm j})$ and $p(y_{\rm i}|+1+{\rm j})$ is adopted to derive a sample-by-sample adaptive algorithm, namely

$$p(y_{\rm r}|+1+{\rm j}) = \frac{1}{\sqrt{2\pi P_k}\rho}\exp\left(-\frac{(y_{\rm r}-{\rm sgn}(\Re(b_k))\Re(z_k))^2}{2\rho^2P_k}\right), \tag{18}$$

$$p(y_{\rm i}|+1+{\rm j}) = \frac{1}{\sqrt{2\pi P_k}\rho}\exp\left(-\frac{(y_{\rm i}-{\rm sgn}(\Im(b_k))\Im(z_k))^2}{2\rho^2P_k}\right). \tag{19}$$

Then the estimation of the error probability can be given by

$$\bar{P}_{\rm e}(\boldsymbol{w}_k,\boldsymbol{a}) = {\rm Prob}(y_{\rm r}<0)+{\rm Prob}(y_{\rm i}<0) = Q\left(C_{\rm R}(\boldsymbol{w}_k,\boldsymbol{a})\right)+Q\left(C_{\rm I}(\boldsymbol{w}_k,\boldsymbol{a})\right), \tag{20}$$

where

$$C_{\rm R}(\boldsymbol{w}_k,\boldsymbol{a})=\frac{{\rm sgn}(\Re(b_k))\Re(z_k)}{\rho\sqrt{P_k}},\quad C_{\rm I}(\boldsymbol{w}_k,\boldsymbol{a})=\frac{{\rm sgn}(\Im(b_k))\Im(z_k)}{\rho\sqrt{P_k}}.$$

The JMSER approach can be reorganized as the following optimization problem:

$$\left(\boldsymbol{w}_1^{\rm JMSER},\boldsymbol{w}_2^{\rm JMSER},\cdots,\boldsymbol{w}_{N_{\rm s}}^{\rm JMSER},\boldsymbol{a}^{\rm JMSER}\right) = \arg\min_{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{N_{\rm s}},\boldsymbol{a}} \bar{P}_{\rm E}(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{N_{\rm s}},\boldsymbol{a}), \tag{21}$$

where $\bar{P}_{\rm E}(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{N_{\rm s}},\boldsymbol{a})=\sum_{k=1}^{N_{\rm s}}\bar{P}_{\rm e}(\boldsymbol{w}_k,\boldsymbol{a})$ denotes the overall error probability that $N_{\rm s}$ source signals which need to be detected are contained.

The following SG algorithm (Haykin, 2002) can be used to find the MSER solution:

$$\boldsymbol{w}_k(n+1)=\boldsymbol{w}_k(n)-\mu_{\boldsymbol{w}_{\rm JS}}\frac{\partial\bar{P}_{\rm E}(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{N_{\rm s}},\boldsymbol{a})}{\partial\boldsymbol{w}_k^*}, \quad k=1,2,\cdots,N_{\rm s}, \tag{22}$$

$$\boldsymbol{a}(n+1)=\boldsymbol{a}(n)-\mu_{\boldsymbol{a}_{\rm JS}}\frac{\partial\bar{P}_{\rm E}(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{N_{\rm s}},\boldsymbol{a})}{\partial\boldsymbol{a}}, \tag{23}$$

where n indicates the snapshot (received symbol index), and $\mu_{\boldsymbol{w}_{\rm JS}}$, $\mu_{\boldsymbol{a}_{\rm JS}}$ are step sizes. By following the same approach, we have the expressions of the gradients as follows:

$$\nabla\bar{P}_{\rm E}(\boldsymbol{w}_k)=-\frac{1}{2\sqrt{2\pi}\rho}{\rm sgn}(\Re(b_k))\exp\left(-\frac{(\Re(z_k))^2}{2\rho^2P_k}\right)\overline{\boldsymbol{u}}_k -\frac{1}{2\sqrt{2\pi}\rho}{\rm sgn}(\Im(b_k))\exp\left(-\frac{(\Im(z_k))^2}{2\rho^2P_k}\right)\overline{\boldsymbol{v}}_k,$$

$$\nabla\bar{P}_{\rm E}(\boldsymbol{a})=-\sum_{k=1}^{N_{\rm s}}\frac{1}{2\sqrt{2\pi}\rho}{\rm sgn}(\Re(b_k))\exp\left(-\frac{[\Re(z_k)]^2}{2\rho^2P_k}\right)\widetilde{\boldsymbol{u}}_k -\sum_{k=1}^{N_{\rm s}}\frac{1}{2\sqrt{2\pi}\rho}{\rm sgn}(\Im(b_k))\exp\left(-\frac{[\Im(z_k)]^2}{2\rho^2P_k}\right)\widetilde{\boldsymbol{v}}_k,$$

where

$$\overline{\boldsymbol{u}}_k = \frac{\boldsymbol{d}}{\sqrt{P_k}} - \frac{\Re(z_k)(\boldsymbol{w}_k + \boldsymbol{H}_{\rm d}\boldsymbol{G}\boldsymbol{A}\boldsymbol{A}^{\rm H}\overline{\boldsymbol{c}}_k^{\rm H})}{(\sqrt{P_k})^3},$$
$$\overline{\boldsymbol{v}}_k = -\frac{{\rm j}\boldsymbol{d}}{\sqrt{P_k}} + \frac{\Im(z_k)(\boldsymbol{w}_k + \boldsymbol{H}_{\rm d}\boldsymbol{G}\boldsymbol{A}\boldsymbol{A}^{\rm H}\overline{\boldsymbol{c}}_k^{\rm H})}{(\sqrt{P_k})^3},$$

$$\widetilde{\boldsymbol{u}}_k = \frac{\boldsymbol{T}^{\rm H}\boldsymbol{H}_{\rm d}^{\rm H}\boldsymbol{w}_k}{\sqrt{P_k}} - \frac{\overline{\boldsymbol{C}}_k^{\rm H}\overline{\boldsymbol{C}}_k\boldsymbol{a}\Re(z_k)}{(\sqrt{P_k})^3},$$
$$\widetilde{\boldsymbol{v}}_k = -\frac{{\rm j}\boldsymbol{T}^{\rm H}\boldsymbol{H}_{\rm d}^{\rm H}\boldsymbol{w}_k}{\sqrt{P_k}} + \frac{\overline{\boldsymbol{C}}_k^{\rm H}\overline{\boldsymbol{C}}_k\boldsymbol{a}\Im(z_k)}{(\sqrt{P_k})^3}.$$

We have noticed that the decision boundary and SER really matter in the orientation of $\boldsymbol{w}_k$ instead of the size of $\boldsymbol{w}_k$. Hence, it is computationally advantageous to normalize $\boldsymbol{w}_k$ to a unit length after each iteration (Wang *et al.*, 2000):

$$\boldsymbol{w}_k(n+1) \leftarrow \frac{\boldsymbol{w}_k(n+1)}{\sqrt{\boldsymbol{w}_k^{\rm H}(n+1)\boldsymbol{w}_k(n+1)}}. \tag{24}$$

The vector $\boldsymbol{a}$ must be restricted by the total power constraint, which is described as $\boldsymbol{a}_k^{\rm H}(n+1)\boldsymbol{a}(n+1) \le P_{\rm r}$, where $P_{\rm r}$ is the total power for all relays. In this work, we use the following procedure to scale the power vector among the relay nodes at each iteration:

$$\boldsymbol{a}(n+1) \leftarrow \sqrt{P_{\rm r}}\frac{\boldsymbol{a}(n+1)}{\sqrt{\boldsymbol{a}^{\rm H}(n+1)\boldsymbol{a}(n+1)}}. \tag{25}$$

In this algorithm, the step sizes $\mu_{\boldsymbol{w}_{\rm JS}}$, $\mu_{\boldsymbol{a}_{\rm JS}}$, and kernel standard variance ρ should be chosen appropriately to achieve a desired convergence performance, in terms of both convergence speed and steady-state SER misadjustment. The algorithm is summarized in Algorithm 1.

Algorithm 1 JMSER-SG algorithm

```
Initialize w_1, w_2, ..., w_{N_s}, a. Set μ_{w_JS},
  μ_{a_JS}, and the number of cycles per symbol
For each received symbol do
    For each cycle do
        For each source transmitted symbol
          (k = 1, 2, ..., N_s)
            Update w_k using Eqs. (22) and (24)
        End
        Update a using Eqs. (23) and (25)
    End
End //We obtain w_1, w_2, ..., w_{N_s} and a for the
  // received symbol and move to the next symbols
```

4.2 CG-based adaptive JMSER algorithm

The CG method is implemented to minimize the cost function by searching for the optimal weight solution along a special direction (Boray and Srinath, 1992). We have known that the directions are determined sequentially at each iteration rather than specified beforehand according to some reported schemes. For the CG method, we obtain the direction by choosing the successive direction vectors as a conjugate version of the successive gradients obtained as the method progresses.

The starting point in the derivation is to consider how to minimize the cost function. In general, there are two steps from $\boldsymbol{w}_k(n-1)$ and $\boldsymbol{a}(n-1)$ to $\boldsymbol{w}_k(n)$ and $\boldsymbol{a}(n)$, respectively. The first step is to choose a search direction and the second step is to implement a line search along the chosen direction.

Here, we start to choose direction vectors $\boldsymbol{c}_1(n), \boldsymbol{c}_2(n), \cdots, \boldsymbol{c}_{N_{\rm s}}(n)$ and $\boldsymbol{g}(n)$ that indicate the directions in which $\boldsymbol{w}_k(n-1)$ and $\boldsymbol{a}(n-1)$ travel to $\boldsymbol{w}_k(n)$ and $\boldsymbol{a}(n)$, respectively. In general, for each step we choose a point

$$\boldsymbol{w}_k(n+1) = \boldsymbol{w}_k(n) + \mu_{\boldsymbol{w}_{\rm JC}}\boldsymbol{c}_k(n), \quad k = 1, 2, \cdots, N_{\rm s}, \tag{26}$$
$$\boldsymbol{a}(n+1) = \boldsymbol{a}(n) + \mu_{\boldsymbol{a}_{\rm JC}}\boldsymbol{g}(n). \tag{27}$$

The parameters $\mu_{\boldsymbol{w}_{\rm JC}}$ and $\mu_{\boldsymbol{a}_{\rm JC}}$ which control the rate of convergence, $\boldsymbol{c}_k(n)$ and $\boldsymbol{g}(n)$, are direction vectors, and n is the time index. The proposed algorithm generates a sequence of iterations $\boldsymbol{w}_k(1), \boldsymbol{w}_k(2), \cdots, \boldsymbol{w}_k(n)$, $\boldsymbol{a}(1), \boldsymbol{a}(2), \cdots, \boldsymbol{a}(n)$ such that $\bar{P}_{\rm E}(n) \le \bar{P}_{\rm E}(n-1)$ at each step; that is, the solution gets close to the minimum at each step. Without loss of generality, we can define this direction vector as

$$\boldsymbol{c}_k(n+1) = \phi_n\boldsymbol{c}_k(n) - \nabla\bar{P}_{\rm E}(\boldsymbol{w}_k(n+1)), \quad k = 1, 2, \cdots, N_{\rm s}, \tag{28}$$

$$\boldsymbol{g}(n+1) = \varphi_n\boldsymbol{g}(n) - \nabla\bar{P}_{\rm E}(\boldsymbol{a}(n+1)). \tag{29}$$

On the basis of the theory of CG algorithms, ϕ_n and φ_n are computed as follows to terminate the algorithm:

$$\phi_n = \frac{\|\nabla\bar{P}_{\rm E}(\boldsymbol{w}_k(n+1))\|}{\|\nabla\bar{P}_{\rm E}(\boldsymbol{w}_k(n))\|}, \tag{30}$$

$$\varphi_n = \frac{\|\nabla\bar{P}_{\rm E}(\boldsymbol{a}(n+1))\|}{\|\nabla\bar{P}_{\rm E}(\boldsymbol{a}(n))\|}. \tag{31}$$

Similar to the SG method, we have the expression of normalization as follows:

$$\boldsymbol{w}_k(n+1) \leftarrow \frac{\boldsymbol{w}_k(n+1)}{\sqrt{\boldsymbol{w}_k^{\rm H}(n+1)\boldsymbol{w}_k(n+1)}}. \tag{32}$$

Meanwhile, the PA vector $\boldsymbol{a}$ must be restricted by the total power constraint, the same as the preceding SG method, i.e., $\boldsymbol{a}_k^{\rm H}(n+1)\boldsymbol{a}(n+1) \le P_{\rm r}$. We also use the following procedure to scale the power vector among relay nodes at each iteration:

$$\boldsymbol{a}(n+1) \leftarrow \sqrt{P_{\rm r}}\frac{\boldsymbol{a}(n+1)}{\sqrt{\boldsymbol{a}^{\rm H}(n+1)\boldsymbol{a}(n+1)}}. \tag{33}$$

By running simulations, the updated algorithm is implemented by running multiple cycles over the recursions for $\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_{N_{\rm s}}, \boldsymbol{a}$. The algorithm is summarized in Algorithm 2.

Algorithm 2 JMSER-CG algorithm

Step 1: Initialize two step sizes $\mu_{\boldsymbol{w}_{\rm JC}} > 0, \mu_{\boldsymbol{a}_{\rm JC}} > 0$,
set the iteration index to $n = 1$,
given $\boldsymbol{a}(1)$, $\boldsymbol{w}_1(1), \cdots, \boldsymbol{w}_{N_{\rm s}}(1)$,
$\boldsymbol{c}_i(1) = -\nabla\overline{P}_{\rm E}(\boldsymbol{w}_i(1))$ $(i = 1, 2, \cdots, N_{\rm s})$,
$\boldsymbol{g}(1) = -\nabla\overline{P}_{\rm E}(\boldsymbol{a}(1))$;

Step 2: For $k = 1, 2, \cdots, N_{\rm s}$:
$\boldsymbol{w}_k(n+1) = \boldsymbol{w}_k(n) + \mu_{\boldsymbol{w}_{\rm JC}}\boldsymbol{c}_k(n)$
$\boldsymbol{w}_k(n+1) \leftarrow \frac{\boldsymbol{w}_k(n+1)}{\|\boldsymbol{w}_k(n+1)\|}$
$\phi_n = \frac{\|\nabla\overline{P}_{\rm E}(\boldsymbol{w}_k(n+1))\|}{\|\nabla\overline{P}_{\rm E}(\boldsymbol{w}_k(n))\|}$
$\boldsymbol{c}_k(n+1) = \phi_n\boldsymbol{c}_k(n) - \nabla\overline{P}_{\rm E}(\boldsymbol{w}_k(n+1))$
$k = k + 1$

Step 3: After $N_{\rm s}$ iterations:
$\boldsymbol{a}(n+1) = \boldsymbol{a}(n) + \mu_{\boldsymbol{a}_{\rm JC}}\boldsymbol{g}(n)$
$\beta = \sqrt{\frac{P_{\rm r}}{\|\boldsymbol{a}(n+1)\|^2}}$
$\boldsymbol{a}(n+1) \leftarrow \beta\boldsymbol{a}(n+1)$
$\varphi_n = \frac{\|\nabla\overline{P}_{\rm E}(\boldsymbol{a}(n+1))\|}{\|\nabla\overline{P}_{\rm E}(\boldsymbol{a}(n))\|}$
$\boldsymbol{g}(n+1) = \varphi_n\boldsymbol{g}(n) - \nabla\overline{P}_{\rm E}(\boldsymbol{a}(n+1))$
$n = n + 1$

4.3 Channel estimation

In this part, we introduce the channel estimation method of this work. A similar SG method is employed to estimate the information of the channel. In the following, we focus on the estimation of channel $\boldsymbol{H}_{\rm d}$ (Karami, 2007). Straightforwardly, the estimation of channel $\boldsymbol{H}_{\rm s}$ is obtained by the same approach.

Recall Eq. (6), and we have noticed that the second and third parts are the noise counter parts. Hence, to simplify the derivation, we can rewrite the $N_{\rm d} \times 1$ received vector $\boldsymbol{d}$ as follows:

$$\boldsymbol{d} = \boldsymbol{H}_{\rm d}\boldsymbol{G}\boldsymbol{A}\boldsymbol{H}_{\rm s}\boldsymbol{b} + \boldsymbol{n}, \tag{34}$$

where $\boldsymbol{n}$ denotes the noise embedded in the received signal.

The channel estimation problem is formulated as follows:

$$\widehat{H}_{\rm d} = \arg\min_{\boldsymbol{H}_{\rm d}} E[\|\boldsymbol{d} - \boldsymbol{H}_{\rm d}\boldsymbol{G}\boldsymbol{A}\boldsymbol{H}_{\rm s}\boldsymbol{b}\|^2]. \tag{35}$$

Using an SG recursion to solve the preceding problem (Haykin, 2002), the gradient of $\widehat{H}_{\rm d}$ is given by

$$\nabla\widehat{H}_{\rm d}(\boldsymbol{H}_{\rm d}) = (\boldsymbol{H}_{\rm d}\boldsymbol{G}\boldsymbol{A}\boldsymbol{H}_{\rm s}\boldsymbol{b} - \boldsymbol{d})\boldsymbol{b}^{\rm H}\boldsymbol{H}_{\rm s}^{\rm H}\boldsymbol{A}^{\rm H}\boldsymbol{G}^{\rm H}. \tag{36}$$

Hence, we have the estimation expression for channel $\widehat{H}_{\rm d}$:

$$\begin{aligned} &\widehat{H}_{\rm d}(n+1) \\ &= \widehat{H}_{\rm d}(n) + \mu_{\rm h}(\boldsymbol{d} - \boldsymbol{H}_{\rm d}(n)\boldsymbol{G}\boldsymbol{A}\boldsymbol{H}_{\rm s}\boldsymbol{b})\boldsymbol{b}^{\rm H}\boldsymbol{H}_{\rm s}^{\rm H}\boldsymbol{A}^{\rm H}\boldsymbol{G}^{\rm H}, \end{aligned} \tag{37}$$

where $\mu_{\rm h}$ is a step size. The channel algorithm is implemented using Eq. (37) with initial values. This algorithm jointly estimates the coefficients of the channels across all the links and for all relay nodes subject to a total power constraint.

5 Analysis of the proposed algorithms

In this section, we analyze the computational complexity and the convergence of the proposed algorithms.

5.1 Computational complexity analysis

We describe the computational complexity of the proposed JMSER algorithm in WSNs. Table 1 lists the computational complexity per iteration in terms of the numbers of multiplications and additions for JMSER-SG and JMSER-CG algorithms with joint linear receiver design and PA strategies, the LMS algorithm with fixed PA, and the MSER algorithm with fixed PA. The fixed PA schemes employ equal power for the relay nodes. Note that for the configuration with $N_{\rm s} = 2$ and $N_{\rm d} = 20$, the numbers of multiplications of the receiver for JMSER-SG, JMSER-CG, MSER, and LMS are given by 161, 241, 161, and 81, respectively. The numbers of additions for them are 160, 200, 160, and 80, respectively. The numbers of multiplications and additions of the PA vector for JMSER-SG and JMSER-CG are 161 and 201, respectively. The numbers of additions of the PA vector for JMSER-SG and JMSER-CG are 160 and 200, respectively. The number of multiplications of the weight vector for the JMSER-SG algorithm is equal to that for the MSER algorithm with fixed PA schemes. In general, compared to the LMS and MSER algorithms both with fixed PA, the total computational complexity of the JMSER-SG algorithm and JMSER-CG algorithm is slightly increased. In the simulations, we will show that the proposed algorithms significantly outperform the existing algorithms with slightly increased complexity.

Table 1 Computational complexity of the analyzed algorithms

Parameter	Algorithm	Multiplication	Addition
$\boldsymbol{w}_k$	JMSER-CG	$6N_\mathrm{s}N_\mathrm{d}+1$	$5N_\mathrm{s}N_\mathrm{d}$
	JMSER-SG	$4N_\mathrm{s}N_\mathrm{d}+1$	$4N_\mathrm{s}N_\mathrm{d}$
	MSER fixed	$4N_\mathrm{s}N_\mathrm{d}+1$	$4N_\mathrm{s}N_\mathrm{d}$
	LMS fixed	$2N_\mathrm{s}N_\mathrm{d}+1$	$2N_\mathrm{s}N_\mathrm{d}$
$\boldsymbol{a}$	JMSER-CG	$5N_\mathrm{s}N_\mathrm{d}+1$	$5N_\mathrm{s}N_\mathrm{d}$
	JMSER-SG	$4N_\mathrm{s}N_\mathrm{d}+1$	$4N_\mathrm{s}N_\mathrm{d}$

N_s: number of source nodes; N_d: number of destination nodes

5.2 Sufficient conditions for convergence

The proposed algorithms have two vectors, weight vector and PA vector. So, to develop the analysis and proofs, we need to define a metric space and the Hausdorff distance that will extensively be used. A metric space is an ordered pair $(\mathcal{S}, D)$, where $\mathcal{S}$ is a nonempty set, and D is a metric on $\mathcal{S}$, that is a function $D: \mathcal{S}\times\mathcal{S}\rightarrow\mathbb{R}$ such that for any $a, b, c\in\mathcal{S}$, the following conditions hold:

$D(a,b)\geq 0$, $D(a,b)=0$ iff $a=b$, $D(a,b)=D(b,a)$, and $D(a,c)\leq D(a,b)+D(b,c)$.

The Hausdorff distance measures how far two subsets of a metric space are from each other and is defined by

$$D_\mathrm{H}(A,B)=\max\left\{\sup_{a\in A}\inf_{b\in B}D(a,b), \sup_{b\in B}\inf_{a\in A}D(a,b)\right\}. \tag{38}$$

Here we first define $\boldsymbol{W}=[\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_{N_\mathrm{s}}]$ to simplify the analysis. It is noticed that $\overline{\boldsymbol{W}}$, $\overline{\boldsymbol{a}}\subset\mathcal{S}$. The proposed JMSER designs can be stated as an alternating minimization strategy on the basis of MSER, expressed as

$$\boldsymbol{W}_n\in\min_{\boldsymbol{W}\in\overline{\boldsymbol{W}}_n}\bar{P}_\mathrm{E}(\boldsymbol{W},\boldsymbol{a}_{n-1}), \tag{39}$$

$$\boldsymbol{a}_n\in\min_{\boldsymbol{a}\in\overline{\boldsymbol{a}}_n}\bar{P}_\mathrm{E}(\boldsymbol{W}_n,\boldsymbol{a}), \tag{40}$$

where the sequences of compact sets $\{\overline{\boldsymbol{W}}_n\}_{n\geq 0}$ and $\{\overline{\boldsymbol{a}}_n\}_{n\geq 0}$ converge to the sets $\overline{\boldsymbol{W}}$ and $\overline{\boldsymbol{a}}$, respectively. Although $\overline{\boldsymbol{W}}$ and $\overline{\boldsymbol{a}}$ are not given directly, we have the sequences of compact sets $\{\overline{\boldsymbol{W}}_n\}_{n\geq 0}$ and $\{\overline{\boldsymbol{a}}_n\}_{n\geq 0}$. The aim of our proposed JMSER designs is to find a sequence of $\boldsymbol{W}_n$ and $\boldsymbol{a}_n$ such that

$$\lim_{n\rightarrow\infty}\bar{P}_\mathrm{E}(\boldsymbol{W}_n,\boldsymbol{a}_n)=\bar{P}_\mathrm{E}(\boldsymbol{W}_\mathrm{opt},\boldsymbol{a}_\mathrm{opt}), \tag{41}$$

where $\boldsymbol{W}_\mathrm{opt}$ and $\boldsymbol{a}_\mathrm{opt}$ correspond to the optimal values of $\boldsymbol{W}_n$ and $\boldsymbol{a}_n$, respectively. To present a set of sufficient conditions under which the proposed algorithms converge, we need the so-called three- and four-point properties (Csiszar and Tusnády, 1984; Niesen *et al.*, 2009). Assume that there is a function $f: \mathcal{S}\times\mathcal{S}\rightarrow\mathbb{R}$ such that the following conditions are satisfied:

Three-point property $(\boldsymbol{W}, \widetilde{\boldsymbol{W}}, \boldsymbol{a})$

For all $n\geq 1$, $\boldsymbol{W}\in\overline{\boldsymbol{W}}_n$, $\boldsymbol{a}\in\overline{\boldsymbol{a}}_{n-1}$, and $\widetilde{\boldsymbol{W}}\in\arg\min_{\boldsymbol{W}\in\overline{\boldsymbol{W}}_n}\bar{P}_\mathrm{E}(\boldsymbol{W},\boldsymbol{a})$, we have

$$f(\boldsymbol{W},\widetilde{\boldsymbol{W}})+\bar{P}_\mathrm{E}(\widetilde{\boldsymbol{W}},\boldsymbol{a})\leq\bar{P}_\mathrm{E}(\boldsymbol{W},\boldsymbol{a}). \tag{42}$$

Four-point property $(\boldsymbol{W}, \boldsymbol{a}, \widetilde{\boldsymbol{W}}, \widetilde{\boldsymbol{a}})$

For all $n\geq 1$, $\boldsymbol{W}$, $\widetilde{\boldsymbol{W}}\in\overline{\boldsymbol{W}}_n$, $\boldsymbol{a}\in\overline{\boldsymbol{a}}_n$, and $\widetilde{\boldsymbol{a}}\in\arg\min_{\boldsymbol{a}\in\overline{\boldsymbol{a}}_n}\bar{P}_\mathrm{E}(\widetilde{\boldsymbol{W}},\boldsymbol{a})$, we have

$$\bar{P}_\mathrm{E}(\boldsymbol{W},\widetilde{\boldsymbol{a}})\leq\bar{P}_\mathrm{E}(\boldsymbol{W},\boldsymbol{a})+f(\boldsymbol{W},\widetilde{\boldsymbol{W}}). \tag{43}$$

These two properties are the mathematical expressions of the sufficient conditions for the convergence of the alternating minimization algorithms, which were stated in Csiszar and Tusnády (1984) and Niesen *et al.* (2009). They mean that if there exists a function $f(\boldsymbol{W},\widetilde{\boldsymbol{W}})$ with the parameter $\boldsymbol{W}$ during two iterations that satisfies the two inequalities (42) and (43) of JMSER, the convergence of our proposed JMSER designs that use the alternating minimization algorithm can be proved by the theorem below:

Theorem 1 Let $\{(\overline{\boldsymbol{W}}_n,\overline{\boldsymbol{a}}_n)\}_{n\geq 0}, \overline{\boldsymbol{W}}, \overline{\boldsymbol{a}}$ be compact subsects of the compact metric space $(\mathcal{S}, D)$ such that

$$\overline{\boldsymbol{W}}_n\xrightarrow{D_\mathrm{H}}\overline{\boldsymbol{W}}, \tag{44}$$

$$\overline{\boldsymbol{a}}_n\xrightarrow{D_\mathrm{H}}\overline{\boldsymbol{a}}, \tag{45}$$

and let $\bar{P}_{\mathrm{E}}$: $\mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$ be a continuous function. Then according to the proposed algorithms, we have

$$\lim_{n\to\infty} \bar{P}_{\mathrm{E}}(\boldsymbol{W}_n, \boldsymbol{a}_n) = \bar{P}_{\mathrm{E}}(\boldsymbol{W}_{\mathrm{opt}}, \boldsymbol{a}_{\mathrm{opt}}). \tag{46}$$

A general proof of this theorem as detailed in Csiszar and Tusnády (1984) and Niesen *et al.* (2009). The proposed JMSER designs can be stated as an alternating minimization strategy on the basis of the MSER criterion.

6 Simulation results

In this section, the SER performance of the proposed two algorithms is simulated and compared with those of two existing algorithms, namely the fixed PA methods based on the MSER and MMSE criteria with SG implementation. The fixed PA scheme is on the basis of equal PA. The two-hop modulated packet is with 1000 transmit symbols where we have 300 training symbols at the initial stage. We consider the AF cooperation scheme. The block fading channel is considered in our simulation whose elements are Rayleigh-distributed during the transmission of each packet. In our simulation the channel information $\boldsymbol{H}_{\mathrm{s}}$ and $\boldsymbol{H}_{\mathrm{d}}$ are obtained by preceding the proposed channel estimation algorithm in Section 4.3. The quasi-static fading channel (block-fading channel) whose elements are Rayleigh random variables (with zero mean and unit variance) is considered in our simulations and assumed to be invariant during the transmission of each packet. The noise at the relay and destination nodes is modeled as a circularly symmetric complex Gaussian random variable with zero mean. The simulations are averaged over 1000 runs. To study the performance with feedback channels, each complex amplification coefficient is quantized to a 4-bit binary value. The fusion center feeds the index that corresponds to the selected entry in the codebook back to the relay nodes. In addition, we select kernel width $\rho = \sqrt{2}\sigma_{\mathrm{d}}$ empirically to ensure a good performance in terms of the convergence rate and steady-state SER misadjustment.

Figs. 2–4 show the SER performances with the feedback PA vector. Figs. 5–7 show the performances in terms of different relays. The PA codebook design which employs the Lloyd algorithm is given in the Appendix.

Fig. 2 shows the SER performance versus the number of received symbols. In the simulation, the signal-to-noise ratio (SNR) is 15 dB, the number of source nodes is two, and we tune $\mu_{\boldsymbol{w}_{\mathrm{JC}}} = 0.01$, $\mu_{\boldsymbol{a}_{\mathrm{JC}}} = 0.04$ for the proposed JMSER-CG algorithm and $\mu_{\boldsymbol{w}_{\mathrm{JS}}} = 0.03$, $\mu_{\boldsymbol{a}_{\mathrm{JS}}} = 0.05$ for the proposed JMSER-SG algorithm. The parameters for the LMS receiver with fixed PA and the MSER receiver with fixed PA are $\mu_{\boldsymbol{w}_{\mathrm{M}}} = 0.03$, $\mu_{\boldsymbol{w}_{\mathrm{L}}} = 0.01$, and the fixed PA schemes employ equal power for the relay nodes. The above-mentioned parameters are detailed in Table 2. From the results we can see that the proposed CG algorithm with a perfect PA vector performs the best, followed by the proposed CG algorithm with feedback information, the JMSER-SG algorithm with a perfect PA vector, the JMSER-SG algorithm with feedback channel, the MSER receiver with fixed PA, and the LMS receiver with fixed PA.

Table 2 Step sizes of the analyzed algorithms

Algorithm	Weight vector	PA
JMSER-CG	$\mu_{\boldsymbol{w}_{\mathrm{JC}}} = 0.01$	$\mu_{\boldsymbol{a}_{\mathrm{JC}}} = 0.04$
JMSER-SG	$\mu_{\boldsymbol{w}_{\mathrm{JS}}} = 0.03$	$\mu_{\boldsymbol{a}_{\mathrm{JS}}} = 0.05$
MSER fixed	$\mu_{\boldsymbol{w}_{\mathrm{M}}} = 0.03$	–
LMS fixed	$\mu_{\boldsymbol{w}_{\mathrm{L}}} = 0.01$	–

Figs. 3 and 4 show the SER performances versus the number of source nodes and SNR, respectively. Both simulation results contain performance with the feedback PA. In addition, the initial PA vector and step sizes are tuned the same as in Fig. 2. From the results, we can see that the whole SER performance becomes worse and worse as the number of source nodes increases. The SER performance of JMSER-CG is always better than those of the other algorithms with SNR in the range of 1–15 dB. We see that the performance of the JMSER-CG algorithm with five relays using feedback quantized PA coefficients is also superior to those of the conventional algorithms. The performance of JMSER-CG with feedback 4-bit PA information approaches the performance of JMSER-SG with perfect PA information.

Fig. 5 shows the SER performance versus the number of received symbols. These simulation results show the SER performance with different relays in particular. The initial PA vector and step sizes are tuned the same as in Fig. 2. When all algorithms employ the same relays, the proposed JMSER-CG algorithm with a varying PA vector performs the

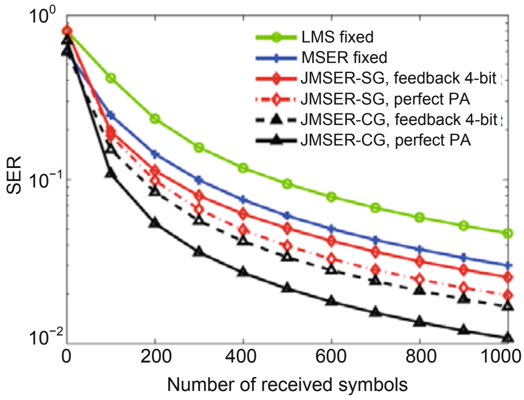

Fig. 2 Curves of SER performance versus the number of symbols with QPSK modulation (SNR=15 dB, $N_s = 2$, $N_r = 5$, $N_d = 20$, $N_{sym} = 1000$)

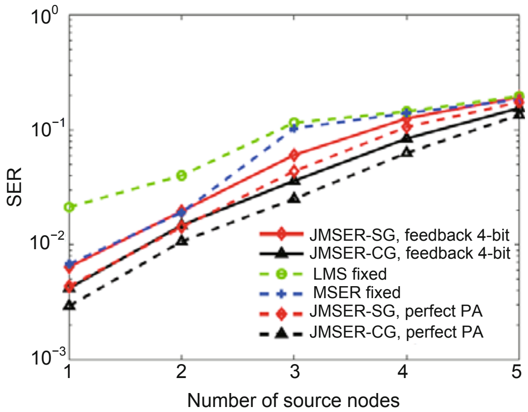

Fig. 3 Curves of SER performance versus the number of sources with QPSK modulation (SNR=15 dB, $N_r = 5$, $N_d = 20$, $N_{sym} = 1000$)

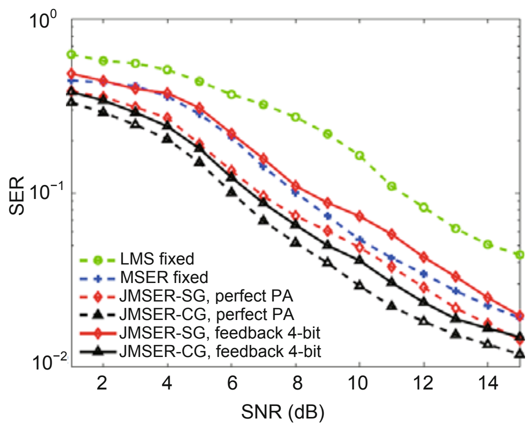

Fig. 4 Curves of SER performance versus SNR with QPSK modulation ($N_s = 2$, $N_r = 5$, $N_d = 20$, $N_{sym} = 1000$)

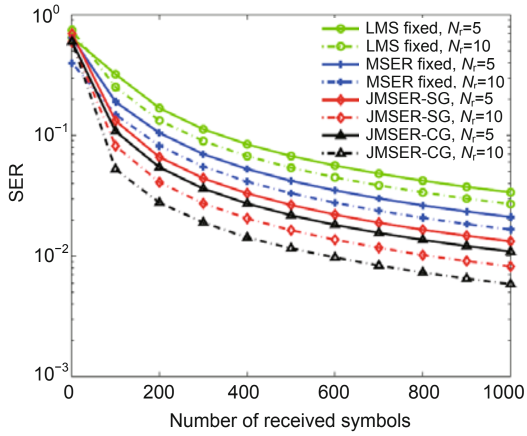

Fig. 5 Curves of SER performance versus the number of symbols with QPSK modulation for different relays (SNR=15 dB, $N_s = 2$, $N_d = 20$)

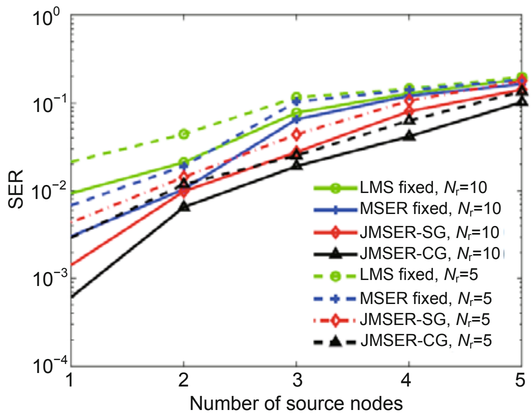

Fig. 6 Curves of SER performance versus the number of sources N_s with QPSK modulation for different relays (SNR=15 dB, $N_d = 20$, $N_{sym} = 1000$)

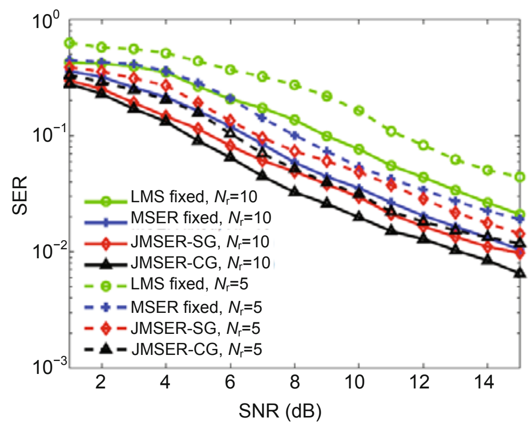

Fig. 7 Curves of SER performance versus SNR with QPSK modulation for different relays ($N_s = 2$, $N_d = 20$, $N_{sym} = 1000$)

best, followed by the proposed SG algorithm with a varying PA vector, the MSER receiver with fixed PA, and the LMS receiver with fixed PA. The algorithms with 10 relay nodes have better average SER performance as compared to those with five relays, reflecting the exploitation of spatial diversity.

Figs. 6 and 7 show the SER performances versus the number of source nodes and SNR, respectively. The PA vector and step sizes are tuned the same as in Fig. 2. We can see that the whole SER performance becomes worse as the number of source nodes increased. The SER performance of JMSER-CG is always better than those of the other algorithms with SNR in the range of 1–15 dB. In addition, all the simulation results show that the whole performance is enhanced with the increase in the number of relay nodes. In conclusion, although the JMSER-CG algorithm has a slightly increased complexity compared to the LMS and MSER algorithms, it significantly outperforms LMS and MSER in terms of the SER performance for different numbers of source nodes and different SNR.

7 Conclusions

In this study two joint iterative PA and interference suppression algorithms were proposed based on the minimization of the SER cost function for WSNs that employ the AF relaying strategy. The proposed two algorithms, including SG and CG, were derived on the basis of the MSER expression to jointly update the vector which allocates the power among the relays subject to a power constraint and the linear receiver. The SER performance of the proposed algorithms has been studied in terms of perfect relay PA and imperfect relay PA via the feedback channel. Simulation results showed that the proposed JMSER-CG algorithm achieves the best performance among the four algorithms, and that the two proposed adaptive algorithms significantly outperform the previously reported algorithms.

Acknowledgements

We wish to express our deepest gratitude to Prof. Rodrigo de LAMARE, who has given us the most valuable suggestions and made necessary corrections to the manuscript.

References

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., *et al.*, 2002. A survey on sensor networks. *IEEE Commun. Mag.*, **40**(8):102-114. [doi:10.1109/MCOM.2002.1024422]

Boray, G.K., Srinath, M.D., 1992. Conjugate gradient techniques for adaptive filtering. *IEEE Trans. Circ. Syst. I*, **39**(1):1-10. [doi:10.1109/81.109237]

Bowman, A.W., Azzalini, A., 1997. Applied Smoothing Techniques for Data Analysis. Oxford University Press, Oxford, UK.

Chen, J., Ueng, F., Lin, P., 2011. A low-complexity adaptive receiver for DS-CDMA systems in unknown code delay environment. *Int. J. Commun. Syst.*, **24**(2):225-238. [doi:10.1002/dac.1151]

Chen, S., Hanzo, L., Ahmad, N.N., *et al.*, 2004a. Adaptive minimum bit error rate beamforming assisted QPSK receiver. Proc. Int. Conf. on Communications, p.3389-3393. [doi:10.1109/ICC.2004.1313173]

Chen, S., Hanzo, L., Mulgrew, B., 2004b. Adaptive minimum symbol-error-rate decision feedback equalization for multilever pulse-amplitude modulation. *IEEE Trans. Signal Process.*, **52**(7):2092-2101. [doi:10.1109/TSP.2004.828944]

Chen, S., Livingstong, A., Du, H.Q., *et al.*, 2008a. Adaptive minimum symbol error rate beamforming assisted detection for quadrature amplitude modulation. *IEEE Trans. Wirel. Commun.*, **7**(4):1140-1145. [doi:10.1109/TWC.2007.060840]

Chen, S., Tan, S., Xu, L., *et al.*, 2008b. Adaptive minimum error-rate filtering design: a review. *Signal Process.*, **88**(7):1671-1697. [doi:10.1016/j.sigpro.2008.01.012]

Clarke, P., de Lamare, R.C., 2011. Joint transmit diversity optimization and relay selection for multi-relay cooperative MIMO systems using discrete stochastic algorithms. *IEEE Commun. Lett.*, **15**(10):1035-1037. [doi:10.1109/LCOMM.2011.082611.102262]

Clarke, P., de Lamare, R.C., 2012. Transmit diversity and relay selection algorithms for multirelay cooperative MIMO systems. *IEEE Trans. Veh. Technol.*, **61**(3):1084-1098. [doi:10.1109/TVT.2012.2186619]

Csiszar, I., Tusnády, G., 1984. Information geometry and alternating minimization procedures. *Statist. Dec.*, **1**:205-237.

de Lamare, R.C., Sampaio-Neto, R., 2003. Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels. *IEEE Commun. Lett.*, **7**(2):73-75. [doi:10.1109/LCOMM.2002.808373]

Fischione, C., Johansson, K.H., Sangiovanni-Vincentelli, A., *et al.*, 2009. Minimum energy coding in CDMA wireless sensor networks. *IEEE Trans. Wirel. Commun.*, **8**(2):985-994. [doi:10.1109/TWC.2009.080239]

Haykin, S.S., 2002. Adaptive Filter Theory. Prentice Hall, Englewood Cliffs, NJ, USA.

Karami, E., 2007. Tracking performance of least squares MIMO channel estimation algorithm. *IEEE Trans. Commun.*, **55**(11):2201-2209. [doi:10.1109/TCOMM.2007.908549]

Kramer, G., Gastpar, M., Gupta, P., 2005. Cooperative strategies and capacity theorems for relay networks. *IEEE Trans. Inform. Theory*, **51**(9):3037-3063. [doi:10.1109/TIT.2005.853304]

Krishna, R., Xiong, Z., Lambotharan, S., 2008. A cooperative MMSE relay strategy for wireless sensor networks.

IEEE Signal Process. Lett., **15**:549-552. [doi:10.1109/LSP.2008.925751]

Laneman, J.N., Tse, D.N.C., Wornell, G.W., 2004. Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Trans. Inform. Theory*, **50**(12):3062-3080. [doi:10.1109/TIT.2004.838089]

Li, M., Hamouda, W., 2007. An adptive multiuser detector for DS-CDMA systems in multipath fading channels. *Int. J. Commun. Syst.*, **20**(11):1299-1313. [doi:10.1002/dac.877]

Li, Y., Vucetic, B., Zhou, Z., *et al.*, 2007. Distributed adaptive power allocation for wireless relay networks. *IEEE Trans. Wirel. Commun.*, **6**(3):948-958. [doi:10.1109/TWC.2007.05256]

Liu, G., Xu, B., Chen, H., 2012. Decentralized estimation over noisy channels in cluster-based wireless sensor networks. *Int. J. Commun. Syst.*, **25**(10):1313-1329. [doi:10.1002/dac.1308]

Niesen, U., Shah, D., Wornell, G.W., 2009. Adaptive alternating minimization algorithms. *IEEE Trans. Inform. Theory*, **55**(3):1423-1429. [doi:10.1109/TIT.2008.2011442]

Rui, X., 2010. Decode-and-forward with partial relay selection. *Int. J. Commun. Syst.*, **23**(11):1443-1448. [doi:10.1002/dac.1128]

Silerman, B., 1996. Density Estimation. Chapman Hall, London, UK.

Souryal, M.R., Vojcic, B.R., Pickholtz, R.L., 2006. Adaptive modulation in ad hoc DS/CDMA packet radio networks. *IEEE Trans. Commun.*, **54**(4):714-725. [doi:10.1109/TCOMM.2006.873092]

Straeter, T.A., 1971. On the Extension of the Davidon-Broyden Class of Rank One, Quasi-Newton Minimization Methods to an Infinite Dimensional Hilbert Space with Applications to Optimal Control Problems. North Carolina State University at Raleigh.

Verdu, S., 1998. Multiuser Detection. Cambridge University Press, USA.

Wang, T., de Lamare, R.C., Schmeink, A., 2012. Joint linear receiver design and power allocation using alternating optimization algorithms for wireless sensor networks. *IEEE Trans. Veh. Technol.*, **61**(9):4129-4141. [doi:10.1109/TVT.2012.2212217]

Wang, X., Lu, W., Antomiou, A., 2000. Constrained minimum-BER multiuser detection. *IEEE Trans. Signal Process.*, **48**(10):2903-2909. [doi:10.1109/78.869045]

Wang, Y., Shi, P., Li, K., *et al.*, 2012. An energy efficient medium access control protocol for target tracking based on dynamic convey tree collaboration in wireless sensor networks. *Int. J. Commun. Syst.*, **25**(9):1139-1159. [doi:10.1002/dac.2355]

Appendix: Lloyd algorithm for the PA vector codebook design

Step 1: initialization phase.

Generate a training sequence that consists of source vectors $\boldsymbol{a}$ with coefficients that are independent and identically distributed with a complex Gaussian distribution with zero mean and unit variance.

Step 2: set $t = 1$.

Step 3: nearest neighbor rule.

All input vectors $\boldsymbol{a}$ that are closer to the codeword $\boldsymbol{a}_{i,t-1}$ than any other codeword should be assigned to the neighborhood of $\boldsymbol{a}_{i,t-1}$ of region Φ_i.

$\boldsymbol{a} \in \Phi_i$ if and only if $d(\boldsymbol{a}, \boldsymbol{a}_{i,t-1}) \leq d(\boldsymbol{a}, \boldsymbol{a}_{j,t-1})$, $\forall i, j = 1, 2, \cdots, 2^{\eta}$, where

$$d(\boldsymbol{a}, \boldsymbol{a}_{i,t-1}) = \sqrt{1 - |\boldsymbol{a}^{\mathrm{H}} \boldsymbol{a}_{i,t-1}|^2}$$

and η is the quantization bit number .

Step 4: centroid condition.

Take the ith region Φ_i as an example, whose local correlation matrix $\boldsymbol{\Upsilon}_i = E[\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} | \boldsymbol{a} \in \Phi_i]$. According to the centroid condition, the optimal vector $\boldsymbol{a}_{i,t}$ should maximize $\boldsymbol{u}_i^{\mathrm{H}} \boldsymbol{\Upsilon}_i \boldsymbol{u}_i$ subject to the unit norm constraint, that is,

$$\boldsymbol{a}_{i,t} = \arg \max_{\boldsymbol{u}_i^{\mathrm{H}} \boldsymbol{u}_i = 1} \boldsymbol{u}_i^{\mathrm{H}} \boldsymbol{\Upsilon}_i \boldsymbol{u}_i = \boldsymbol{e}_i,$$

where $\boldsymbol{e}_i$ is the eigenvector that corresponds to the largest eigenvalue of $\boldsymbol{\Upsilon}_i$.

Loop back to step 3 until convergence.