Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 88 (2008) 1413–1427
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

Performance analysis of an adaptive two-stage PIC CDMA receiver in AWGN channels ☆

Yu-Tao Hsieh[a,*], Wen-Rong Wu[b]

[a]SoC Technology Center, Industrial Technology Research Institute, Hsinchu 310, Taiwan, ROC
[b]Department of Communication Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, ROC

Received 18 April 2007; received in revised form 7 December 2007; accepted 7 December 2007
Available online 15 December 2007

Abstract

Parallel interference cancellation (PIC) is considered a simple yet effective multiuser detector for direct-sequence code-division multiple-access (DS-CDMA) systems. However, its performance may deteriorate due to unreliable interference cancellation in the early stages. Thus, a partial PIC detector, in which partial cancellation factors (PCFs) are introduced to control the interference cancellation level, has been developed as a remedy. Recently, an interesting adaptive multistage PIC algorithm was proposed. In this scheme, coefficients combining the channel responses and optimal PCFs are blindly trained with the least mean square (LMS) algorithm. The algorithm is simple to implement, inherently applicable to time-varying environments, and superior to the non-adaptive type of partial PICs. Despite its various advantages, its performance has not been theoretically analyzed yet. The contribution of this paper is to fill the gap by analyzing an adaptive two-stage PIC in AWGN channels. We explicitly derive the analytical results for optimal weights, weight-error means, and weight-error variances. Based on these results, we finally derive the output bit error rate (BER) for each user. Simulation results indicate that our analytical results highly agree with empirical ones.
© 2008 Elsevier B.V. All rights reserved.

Keywords: Parallel interference cancellation; Performance analysis; LMS algorithm

1. Introduction

Multiuser detection (MUD) is a technique for improving the performance of code-division multiple-access (CDMA) systems. Different from the conventional single-user approach, MUD does not treat interference from other users as noise. The development of MUD algorithms can be dated back to the seminal work of Verdu. He proposed a multiuser receiver utilizing the maximum-likelihood criterion [1] and showed a great performance enhancement. However, he also showed that the computational complexity grows exponentially with the user number. The high computational complexity adversely affects its real-world applications. Thus, a variety of low-complexity suboptimum receivers were then proposed [2–4].

The subtractive type interference cancellation is known to be a simple and effective MUD algorithm. This type of MUD involves only vector operations

☆Parts of this work were presented at IEEE Wireless and Networking Conference, WCNC 2007.

*Corresponding author. Tel.: +886 3 5914854; fax: +886 3 5829733.
E-mail addresses: ythsieh@itri.org.tw (Y.-T. Hsieh), wrwu@cc.nctu.edu.tw (W.-R. Wu).

0165-1684/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2007.12.002

making it a good candidate for real-world implementation. For a particular desired user, the subtractive-type canceller estimates interference from other users, regenerates it, and cancels it from the received signal. This canceller is usually implemented with a multistage structure, and the temporary data decision for a stage is obtained from its previous stage. The successive interference cancellation (SIC) cancels interference from other users one by one [5–7], while the parallel interference cancellation (PIC) cancels it all at one time [8–10]. To have the best performance, signal power ranking is necessary for SIC. Stronger signals usually have lower probability of decision errors and cancellation of these signals give more significant result. For these reasons, SIC has better performance where users have unbalanced powers. However, SIC requires additional complexity for power ranking and its processing delay is also larger. By contrast, PIC cancels the interference disregarding to the interference power distribution and is more suitable for power-balanced systems.

One problem associated with the PIC approach is that interference estimates may not be reliable in early stages. In other words, interference cancellation does not necessarily reduce interference. To alleviate this problem, partial PIC was then developed. Partial cancellation factors (PCFs) ranging from zero to unity were introduced to control the signal cancellation level. Optimal PCFs can be theoretically derived with given channel and noise statistics. One problem of the approach is that the computational complexity will become very high in time-varying environments. Also, when the required statistics are not properly estimated, the performance may be seriously affected. To remedy the problem, an adaptive approach using the least mean square (LMS) algorithm was then proposed for partial PICs [11]. Due to the special architecture proposed in [11], the adaptive algorithm does not require any training sequence. In other words, it is operated in a blind way. The basic idea behind the adaptive multistage PIC is fundamentally different from that of conventional partial PICs. The weight vector it derives for a user corresponds to a combination of the channel response and the PCF. There are many advantages using the adaptive PIC in [11]. It is simple to implement [12], and is inherently applicable in time-varying environments. Also, it does not have to conduct channel estimation, and its performance is better than non-adaptive PICs. Other related works can be found in [13–16]. In [14], the adaptive multistage PIC applied was applied to multi-rate systems. In [15,16], adapted weights are filtered before or after weight adaptation such that better cancellation performance can be obtained.

Although the adaptive multistage PIC has been studied by many researches, its performance has not been analyzed before. The difficulty arises from the nonlinear operation involved in the decision process, and its interaction with the LMS algorithm. There exist many theoretical results for the LMS algorithm; however, most of them consider the steady-state performance and are valid only for the small step size scenario. This cannot be applied in the problem considered here. This is because in an adaptive PIC, only the data in one bit interval are available. For better performance, a large step size must be used. Even with the large step size, the weights still cannot converge due to the short training sequence. As a result, the LMS algorithm is always in its transient state. One other obstacle is that the input to the LMS algorithm in a certain stage depends on the decision in the previous stage, and this complicates the problem furthermore. In this paper, we develop a method overcoming these problems outlined above. We explicitly derive the analytical results for optimal weights, weight-error means, and weight-error variances. Based on these results, we finally derive the output bit error rate (BER) for each user. Simulation results indicate that our analytical results highly agree with empirical ones. Although the analysis provides only the system behavior of the original structure in [11], it is straightforward to conduct the performance analysis for the improved structures such as [15] or [16] with the proposed algorithm.

The remainder of the paper is organized as follows. In Section 2, the framework of the adaptive multistage PIC is first reviewed. In Section 3, a complete derivation for the LMS convergence statistics in a single-user scenario is given, which includes optimal weights, weight-error means, and weight-error variances. In Section 4, the results obtained for the single-user scenario are extended to the two-user scenario. Finally, with some approximation techniques, analytical results for a general multiple-user scenario are derived in Section 5. The simulation results and discussions are presented in Section 6. Finally the conclusions are drawn in Section 7.

2. System model

Consider a synchronous K-user CDMA system in the additive white Gaussian noise (AWGN) channel. Let the spreading sequence of the kth user denoted by $x_k(n)$ with processing gain N and amplitude $\pm 1/\sqrt{N}$. Then the chip-sampled received signal in a certain bit interval can be represented as

$$r(n) = \sum_{k=1}^{K} a_k b_k x_k(n) + v(n),\quad n = 0, \ldots, N-1, \tag{1}$$

where a_k and b_k are the channel gain and data bit of the kth user, and $v(n)$ is the AWGN with variance σ^2. Without loss of generality, we let $a_k > 0$. We first describe the operation of a partial PIC receiver. The first stage operation is just the matched filtering. Let $y_k^{(i)}$ be the ith stage output of the kth user. Then,

$$\begin{aligned} y_k^{(1)} &= \sum_{n=0}^{N-1} x_k(n) r(n) \\ &= a_k b_k + \sum_{j \neq k} a_j b_j \rho_{jk} + \sum_{n=0}^{N-1} v_k(n), \end{aligned} \tag{2}$$

where $\rho_{jk} \sum_{n=0}^{N-1} x_j(n) x_k(n)$ defined as the time-averaged cross-correlation function between the ith and the jth user and $v_k(n) = x_k(n) v(n)$. We further denote the noise term in (2) as $\gamma_k = \sum_{n=0}^{N-1} v_k(n)$. It can be seen that in the received signal of (2), in addition to noise, the MAI affects the signal detection. Let $\hat{r}_k^{(i)}(n)$ denote an interference-subtracted signal for User k in the ith stage and $i > 1$. Then,

$$\begin{aligned} \hat{r}_k^{(i)}(n) &= r(n) - \sum_{j \neq k} g_j^{(i)} \hat{s}_j^{(i)}(n) \\ &= r(n) - \sum_{j \neq k} g_j^{(i)} a_j \hat{b}_j^{(i-1)} \cdot x_j(n), \end{aligned} \tag{3}$$

where $g_j^{(i)}$ denotes the PCF for the jth user in the ith stage and $\hat{s}_j^{(i)}(n)$ is the corresponding interference estimate being equal to $a_j \hat{b}_j^{(i-1)} x_j(n)$. Thus, the output signal in Stage i is then

$$y_k^{(i)} = \sum_{n=0}^{N-1} \hat{r}_k^{(i)}(n) x_k(n). \tag{4}$$

Finally, we can obtain the ith stage detected bit with $y_k^{(i)}$, i.e., $\hat{b}_j^{(i)} = \mathrm{sgn}[y_j^{(i)}]$, in which sgn[.] denotes the sign operation. As we can see, the knowledge of the channel gain is required in partial PIC. In general, optimal PCFs are derived through minimization of some cost functions.

The interference subtraction operation of the adaptive multistage PIC is similar to that of the conventional partial PIC. However, the PCF and the channel gain are merged into a single weight. Let $w_k^{(i)}$ be the weight of the kth user at the ith stage and $\mathbf{w}^{(i)} = [w_1^{(i)}, w_2^{(i)}, \ldots, w_K^{(i)}]^{\mathrm{T}}$. The optimal weight vector, denoted as $\mathbf{w}_{\mathrm{opt}}^{(i)}$, is obtained by minimizing a mean square error (MSE) function $J^{(i)}$, i.e., $\mathbf{w}_{\mathrm{opt}}^{(i)} = \min_{\mathbf{w}^{(i)}} J^{(i)}$. The MSE is given by

$$\begin{aligned} J^{(i)} &= E\{[r(n) - \tilde{r}^{(i)}(n)]^2\} \\ &= E\left\{ \left[r(n) - \sum_{k=1}^{K} w_k^{(i)} \hat{b}_k^{(i-1)} x_k(n) \right]^2 \right\}. \end{aligned} \tag{5}$$

Note that $\tilde{r}^{(i)}(n)$ denotes the reconstructed received signal (excluding noise). The optimum solution of (5) can be theoretically solved. However, it will require high computational complexity. A simple alternative is to use the adaptive filtering approach. The LMS algorithm is a well-known adaptive algorithm. Let $\mathbf{x}(n) = [x_1(n), x_2(n), \ldots, x_K(n)]^{\mathrm{T}}$. The LMS update equation for the ith stage processing (with $i-1$ stages of interference cancellation) can be expressed as

$$\begin{aligned} \tilde{r}^{(i)}(n) &= [\boldsymbol{\chi}^{(i)}(n)]^{\mathrm{T}} \mathbf{w}^{(i)}(n), \\ e^{(i)}(n) &= r(n) - \tilde{r}^{(i)}(n), \\ \mathbf{w}^{(i)}(n+1) &= \mathbf{w}^{(i)}(n) + \mu e^{(i)}(n) \boldsymbol{\chi}^{(i)}(n), \end{aligned} \tag{6}$$

where the input signals are described as $\boldsymbol{\chi}^{(i)}(n) = \hat{\mathbf{B}}^{(i-1)} \mathbf{x}(n)$ with $\hat{\mathbf{B}}^{(i)} \triangleq \mathrm{diag}\{\hat{b}_1^{(i)}, \hat{b}_2^{(i)}, \ldots, \hat{b}_K^{(i)}\}$. After the weights are trained for a bit interval, they are used to cancel the interference from other users such that the input to the kth user' slicer in the ith stage is

$$\hat{r}_k^{(i)}(n) = r(n) - \sum_{j \neq k} w_j^{(i)}(N) \hat{b}_j^{(i-1)} x_j(n). \tag{7}$$

Then the ith stage output for the kth user can be obtained as that in (4). The block diagrams for the adaptive multistage partial PIC and the LMS algorithm are shown in Figs. 1 and 2, respectively. In this paper, we will consider an adaptive two-stage PIC receiver. Since only the operation in the second stage is concerned, the superscripts for the first-stage outputs and for the second-stage inputs are omitted for notational simplicity. As a result, $y_k = y_k^{(1)}$, $\hat{b}_k = \hat{b}_k^{(1)}$, $\tilde{r}(n) = \tilde{r}^{(2)}(n)$, and $w_k(n) = w_k^{(2)}(n)$. It can be seen from (5)

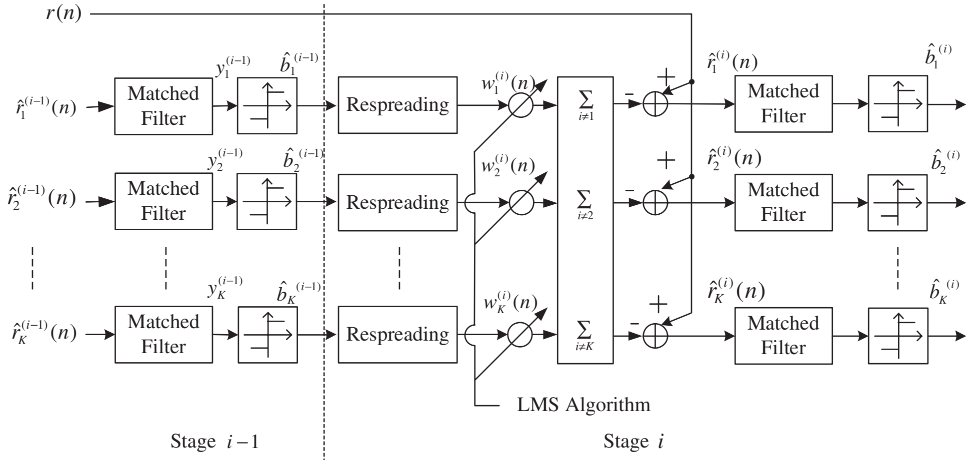

Fig. 1. Block diagram of an adaptive multistage PIC receiver.

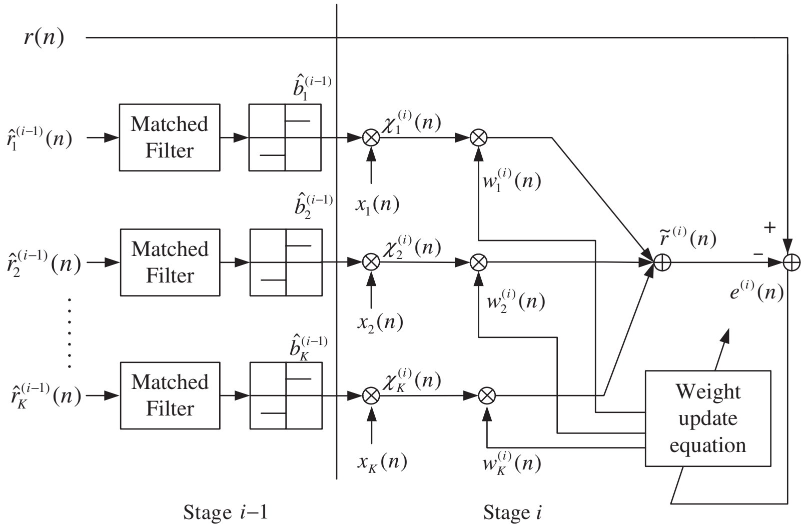

Fig. 2. LMS algorithm for an adaptive multistage PIC receiver.

that in the ideal condition, i.e., $\tilde{r}(n)$ is equal to the noise-free received signal, the ideal convergent weights will be

$$w_k(N) = \begin{cases} a_k, & \hat{b}_k = b_k, \\ -a_k, & \hat{b}_k \neq b_k. \end{cases} \tag{8}$$

Thus, the convergent weights depend on whether the bit decisions in the previous stage are correct or erroneous. This property makes the performance analysis difficult. The adaptive algorithm allows the weight of each user to attain the desired value bit-by-bit. This is the reason why the adaptive approach performs better than non-adaptive methods.

As mentioned, the adaptation period is constrained in one symbol period. This is because the optimal weight for User k may be $+a_k$ or $-a_k$ depending on the bit decision for each symbol. Although the LMS algorithm is simple, its convergence is slow and the weight may not converge to the desired value in such a short period. In addition, the resultant weight heavily depends on the parameters used in the LMS algorithm so is the cancellation performance. These parameters include the step size and initial weights. In the conventional approach, these parameters are determined heuristically. The weight initials are usually set as the channel gains, i.e., $w_k(0) = a_k$. This is reasonable since the bit error probability is usually low, most of the weights will start adaptation at their optimal values; only few weights are away from their optimal values by $2a_k$. A larger step size will accelerate the convergence speed for the weights with erroneous decision, but also inevitably introduces a larger variance for all weights. There is almost no research regarding the convergence analysis for the adaptive multistage PIC receiver and this is the motivation of our research. We will start the analysis with a single-user scenario. In this case, there is no MAI; however, the result can serve as a basis for the analysis of the two-user and general K-user scenarios.

3. Exact analysis for single-user scenario

3.1. Optimal weight analysis

Consider the CDMA system with only one active user, i.e., $K = 1$. Since only one user is present, we will omit the subscript k for notational simplicity. Thus, $x(n) = x_1(n)$, $a = a_1$, $b = b_1$, $y = y_1$, $v(n) = v_1(n)$, $\gamma = \gamma_1$, and $\chi(n) = \hat{b}_1 x_1(n)$. Note that by definition, $v(n) = x(n)v(n)$ and

$$\gamma = \sum_{n=0}^{N-1} v(n) = \sum_{n=0}^{N-1} x(n)v(n). \tag{9}$$

The matched filter output signal for a certain bit signal in (2) can be rewritten as

$$y = \sum_{n=0}^{N-1} x(n)r(n) = ab + \sum_{n=0}^{N-1} v(n) = ab + \gamma, \tag{10}$$

where the PN-code multiplied noise samples $v(n), n = 0, 1, \ldots, N-1$ are i.i.d. random variables with zero mean and variance $\sigma_v^2 = \sigma^2/N$. Note that in the following derivation, for simplicity we refer to the first stage decision, the first stage correct decision, and the first stage erroneous decision as the decision, the correct decision, and the erroneous decision, respectively. From (10), it is simple to see that the decision $\hat{b}$, which equals $\mathrm{sgn}[y]$, depends on the noise term γ. It is simple to derive the condition for correct or erroneous decision. Denote the set of γ for which the decision is correct as $\mathbb{V}^{\mathrm{c}}$, and that for which the decision is erroneous as $\mathbb{V}^{\mathrm{e}}$. Then,

$$\mathbb{V}^{\mathrm{c}} \triangleq \{\gamma | \hat{b} = b\} = \begin{cases} \gamma > -a & \text{for } b = 1, \\ \gamma < a & \text{for } b = -1 \end{cases} \tag{11}$$

and

$$\mathbb{V}^{\mathrm{e}} \triangleq \{\gamma | \hat{b} \neq b\} = \begin{cases} \gamma < -a & \text{for } b = 1, \\ \gamma > a & \text{for } b = -1. \end{cases} \tag{12}$$

We will first derive the optimal weight conditioned on γ and then take the expectation on the conditional optimal weight to obtain the final result. Since the input to the LMS filter depends on $\hat{b}$, the optimal weight will be different for $\hat{b} = b$ and $\hat{b} \neq b$. To facilitate the derivation, we first define some notations. Let a random variable z conditioned on γ be denoted as $\tilde{z}$, i.e., $\tilde{z} = \{z|\gamma\}$. Also let the conditional random variable with $\gamma \in \mathbb{V}^{\mathrm{c}}$ be denoted as $\tilde{z}^{\mathrm{c}}$, i.e., $\tilde{z}^{\mathrm{c}} = \{z|\gamma \in \mathbb{V}^{\mathrm{c}}\}$. Similarly, $\tilde{z}^{\mathrm{e}} = \{z|\gamma \in \mathbb{V}^{\mathrm{e}}\}$. Also let $\tilde{z}_M^{\mathrm{c}} = E_{v,x}\{\tilde{z}^{\mathrm{c}}\}$, $\tilde{z}_M^{\mathrm{e}} = E_{v,x}\{\tilde{z}^{\mathrm{e}}\}$, $z_M^{\mathrm{c}} = E_\gamma\{\tilde{z}_M^{\mathrm{c}}\}$, and $z_M^{\mathrm{e}} = E_\gamma\{\tilde{z}_M^{\mathrm{e}}\}$ where the subscript M denotes the correspondingvariable is a mean value, $E_{v,x}\{.\}$ denotes the expectation operated on $v(n)$ and $x(n)$, and $E_\gamma\{.\}$ denotes the expectation operated on γ. Note here that, γ is a function of $x(n)$ and $v(n)$ [from (9)]. When γ is set as a constant, $x(n)$ and $v(n)$ are still random but constrained. Let $E\{\cdot\}$ denote the expectation operated on all random variables. Then, we have $E\{z^{\mathrm{c}}\} = E_\gamma\{E_{v,x}\{\tilde{z}^{\mathrm{c}}\}\} = z_M^{\mathrm{c}}$, and $E\{z^{\mathrm{e}}\} = E_\gamma\{E_{v,x}\{\tilde{z}^{\mathrm{e}}\}\} = z_M^{\mathrm{e}}$. Using the similar rule, we define the optimal weight conditioned on $\gamma \in \mathbb{V}^{\mathrm{c}}$ as $\tilde{w}_{\mathrm{opt}}^{\mathrm{c}}$, and that on $\gamma \in \mathbb{V}^{\mathrm{e}}$ as $\tilde{w}_{\mathrm{opt}}^{\mathrm{e}}$. One step further, let the optimal weight for correct decision be $w_{\mathrm{opt}}^{\mathrm{c}}$ and that for erroneous decision be $w_{\mathrm{opt}}^{\mathrm{e}}$. We then have $w_{\mathrm{opt}}^{\mathrm{c}} = E_\gamma\{\tilde{w}_{\mathrm{opt}}^{\mathrm{c}}\}$ and $w_{\mathrm{opt}}^{\mathrm{e}} = E_\gamma\{\tilde{w}_{\mathrm{opt}}^{\mathrm{e}}\}$. In the sequel, we only consider the scenario of correct decision since the analysis for the erroneous decisions is similar. The conditional optimal weight is derived from the Wiener solution as $\tilde{w}_{\mathrm{opt}}^{\mathrm{c}} = (\tilde{\mathbf{Q}}^{\mathrm{c}})^{-1}\tilde{\mathbf{p}}^{\mathrm{c}}$ where $\tilde{\mathbf{Q}}^{\mathrm{c}} = E_{v,x}\{\tilde{\chi}^{\mathrm{c}}(n)^2\}$ and $\tilde{\mathbf{p}}^{\mathrm{c}} = E_{v,x}\{\tilde{\chi}^{\mathrm{c}}(n)\tilde{r}^{\mathrm{c}}(n)\}$. Note that $\tilde{\mathbf{Q}}^{\mathrm{c}} = 1/N$. We then have

$$\tilde{w}_{\mathrm{opt}}^{\mathrm{c}} = a + N\hat{b}\tilde{v}_M^{\mathrm{c}}(n). \tag{13}$$

The conditional mean for $\tilde{v}^{\mathrm{c}}(n)$ can be obtained by taking the conditional expectation on both sides of (9)

$$\tilde{\gamma}_M^{\mathrm{c}} = \sum_{n=0}^{N-1} \tilde{v}_M^{\mathrm{c}}(n) = N\tilde{v}_M^{\mathrm{c}}(n). \tag{14}$$

Note that $\tilde{\gamma}_M^{\mathrm{c}} = \tilde{\gamma}^{\mathrm{c}}$ since $\tilde{\gamma}^{\mathrm{c}}$ is assumed constant in the corresponding expectation operation. Thus, (13) can be rewritten as

$$\tilde{w}_{\mathrm{opt}}^{\mathrm{c}} = a + \hat{b}\tilde{\gamma}^{\mathrm{c}}. \tag{15}$$

Assuming that $b=1$, we can obtain the optimal weight for correct decision as (the result is identical for $b=-1$)

$$w^{\rm c}_{\rm opt} = E_\gamma\{\tilde{w}^{\rm c}_{\rm opt}\} = a + E_\gamma\{\tilde{\gamma}^{\rm c}\}. \tag{16}$$

Note that γ is a Gaussian random variable with zero mean and variance $\sigma^2_\gamma = N\sigma^2_v = \sigma^2$. Let $f(\cdot)$ denote a Gaussian probability density function. Thus, the second term in the right-hand side of (16) can be expressed as

$$E_\gamma\{\tilde{\gamma}^{\rm c}\} = \int \gamma f(\gamma|\mathbb{V}^{\rm c})\,{\rm d}\gamma = \frac{\int_{\mathbb{V}^{\rm c}} \gamma f(\gamma)\,{\rm d}\gamma}{\int_{\mathbb{V}^{\rm c}} f(\gamma)\,{\rm d}\gamma}. \tag{17}$$

3.2. Weight-error mean analysis

In this subsection, we will analyze the statistical properties of the LMS adaptive algorithm used in the two-stage PIC receiver. Define the weight error as

$$\tilde{\varepsilon}^{\rm c}(n) = \tilde{w}^{\rm c}(n) - w^{\rm c}_{\rm opt}. \tag{18}$$

Our objective is to find closed-form expressions for the mean of $\tilde{\varepsilon}^{\rm c}(n)$. Using the notation system defined above, we have $\varepsilon^{\rm c}_M(n) = E_\gamma\{\tilde{\varepsilon}^{\rm c}_M(n)\}$. Decomposing $\tilde{\varepsilon}^{\rm c}(n)$ into two parts as

$$\tilde{\varepsilon}^{\rm c}(n) = \tilde{\varepsilon}^{\rm c}(n) + \tilde{\delta}^{\rm c}, \tag{19}$$

where $\tilde{\varepsilon}^{\rm c}(n) = \tilde{w}^{\rm c}(n) - \tilde{w}^{\rm c}_{\rm opt}$ and $\tilde{\delta}^{\rm c} = \tilde{w}^{\rm c}_{\rm opt} - w^{\rm c}_{\rm opt}$. It is simple to see that $E_\gamma\{\tilde{\delta}^{\rm c}_M\} = 0$. Thus, $\varepsilon^{\rm c}_M(n) = E_\gamma\{\tilde{\varepsilon}^{\rm c}_M(n)\}$. Expanding $\tilde{\varepsilon}^{\rm c}(n)$ by (6), (13), and (15), we have

$$\begin{aligned}\tilde{\varepsilon}^{\rm c}(n) &= \tilde{w}^{\rm c}(n) - \tilde{w}^{\rm c}_{\rm opt} \\ &= \tilde{w}^{\rm c}(n-1) + \mu\tilde{\chi}^{\rm c}(n)\tilde{e}^{\rm c}(n-1) - \tilde{w}^{\rm c}_{\rm opt} \\ &= (1-\mu/N)\tilde{\varepsilon}^{\rm c}(n-1) \\ &\quad + \mu\hat{b}(\tilde{v}^{\rm c}(n-1) - \tilde{\gamma}^{\rm c}/N).\end{aligned} \tag{20}$$

Iterating (20), we can obtain

$$\begin{aligned}\tilde{\varepsilon}^{\rm c}(n) &= (1-\mu/N)^n\tilde{\varepsilon}^{\rm c}(0) \\ &\quad + \mu\hat{b}\left\{\sum_{i=0}^{n-1}\left(1-\frac{\mu}{N}\right)^{n-1-i}\tilde{v}^{\rm c}(i) - \frac{\tilde{\gamma}^{\rm c}}{N}\sum_{i=0}^{n-1}\left(1-\frac{\mu}{N}\right)^{i}\right\} \\ &= \alpha^n\tilde{\varepsilon}^{\rm c}(0) + \mu\hat{b}\left[\sum_{i=0}^{n-1}\alpha^{n-1-i}\tilde{v}^{\rm c}(i) - \frac{\tilde{\gamma}^{\rm c}}{N}\left(\frac{1-\alpha^n}{1-\alpha}\right)\right],\end{aligned} \tag{21}$$

where $\alpha = 1-\mu/N$ and $\tilde{\varepsilon}^{\rm c}(0) = \tilde{w}^{\rm c}(0) - \tilde{w}^{\rm c}_{\rm opt}$. Note that $w(0)$ is an deterministic initial value and $\tilde{w}^{\rm c}(0) = w(0)$. Taking expectation on both sides of (21) with respect to $v(n)$ and $x(n)$, we have

$$\begin{aligned}\tilde{\varepsilon}^{\rm c}_M(n) &= \alpha^n\tilde{\varepsilon}^{\rm c}(0) \\ &\quad + \mu\hat{b}\left[\sum_{i=0}^{n-1}\alpha^{n-1-i}\tilde{v}^{\rm c}_M(i) - \frac{\tilde{\gamma}^{\rm c}_M}{N}\left(\frac{1-\alpha^n}{1-\alpha}\right)\right].\end{aligned} \tag{22}$$

Using the result from (14), we have

$$\tilde{\varepsilon}^{\rm c}_M(n) = \alpha^n\tilde{\varepsilon}^{\rm c}(0). \tag{23}$$

From above, we know that $\varepsilon^{\rm c}_M(n) = E_\gamma\{\tilde{\varepsilon}^{\rm c}_M(n)\}$. Thus,

$$\varepsilon^{\rm c}_M(n) = \alpha^n E_\gamma\{\tilde{\varepsilon}^{\rm c}(0)\} = \alpha^n\varepsilon^{\rm c}(0), \tag{24}$$

where $\varepsilon^{\rm c}(0) = w(0) - w^{\rm c}_{\rm opt}$.

3.3. Weight-error power analysis

In this section, we will find closed-form expressions for $E\{[\tilde{\varepsilon}^{\rm c}(n)]^2\}$. Before proceeding further, we define additional notations. Given a random variable z, let $\tilde{z}^{\rm c}_V = E_{v,x}\{[\tilde{z}^{\rm c} - z^{\rm c}_M]^2\}$ and $z^{\rm c}_V = E\{[\tilde{z}^{\rm c} - z^{\rm c}_M]^2\}$. Thus, $z^{\rm c}_V = E_\gamma\{\tilde{z}^{\rm c}_V\}$. Using the definitions, we have $\varepsilon^{\rm c}_V(n) = E\{[\tilde{\varepsilon}^{\rm c}(n) - \varepsilon^{\rm c}_M(n)]^2\}$. We then have

$$E\{[\tilde{\varepsilon}^{\rm c}(n)]^2\} = \varepsilon^{\rm c}_V(n) + [\varepsilon^{\rm c}_M(n)]^2. \tag{25}$$

Also, $\tilde{\varepsilon}^{\rm c}_V(n) = E_{v,x}\{[\tilde{\varepsilon}^{\rm c}(n) - \varepsilon^{\rm c}_M(n)]^2\}$ and $\varepsilon^{\rm c}_V(n) = E_\gamma\{\tilde{\varepsilon}^{\rm c}_V(n)\}$. From (19), (21) and (23), we have

$$\begin{aligned}\tilde{\varepsilon}^{\rm c}_V(n) &= E_{v,x}\{[\tilde{\varepsilon}^{\rm c}(n) - \varepsilon^{\rm c}_M(n)]^2\} \\ &= E_{v,x}\{[\tilde{\varepsilon}^{\rm c}(n) + \tilde{\delta}^{\rm c} - \alpha^n\varepsilon^{\rm c}(0)]^2\} \\ &= E_{v,x}\left\{\left[\alpha^n(\tilde{\varepsilon}^{\rm c}(0) - \varepsilon^{\rm c}(0)) + \tilde{\delta}^{\rm c}\right.\right. \\ &\quad \left.\left. + \mu\hat{b}\left(\sum_{i=0}^{n}\alpha^{n-1-i}\tilde{v}^{\rm c}(i) - \frac{\tilde{\gamma}^{\rm c}}{N}\left(\frac{1-\alpha^n}{1-\alpha}\right)\right)\right]^2\right\} \\ &\quad + \mu^2 E_{v,x}\left\{\left[\left(\sum_{i=0}^{n-1}\alpha^{n-1-i}\tilde{v}^{\rm c}(i)\right.\right.\right. \\ &\quad \left.\left.\left. - \frac{\tilde{\gamma}^{\rm c}}{N}\left(\frac{1-\alpha^n}{1-\alpha}\right)\right)\right]^2\right\}.\end{aligned} \tag{26}$$

Note that the expectation term on the second term of the right-hand side of (26) is just $\tilde{\varepsilon}^{\rm c}_V(n)$. This can be seen from (21) and (23). We now evaluate this term

$$\tilde{\varepsilon}_V^{\rm c}(n) = \mu^2 E_{v,x}\left\{\left[\sum_{i=0}^{n-1}\alpha^{n-1-i}\tilde{v}^{\rm c}(i) - \frac{\tilde{\gamma}^{\rm c}}{N}\left(\frac{1-\alpha^n}{1-\alpha}\right)\right]^2\right\}$$
$$= \mu^2\left\{E_{v,x}\left\{\left[\sum_{i=0}^{n-1}\alpha^{n-1-i}\tilde{v}^{\rm c}(i)\right]^2\right\} - \frac{(\tilde{\gamma}_M^{\rm c})^2}{N^2}\left(\frac{1-\alpha^n}{1-\alpha}\right)^2\right\}. \quad (27)$$

From (27), we can see that we have to find the autocorrelation function of $\tilde{v}^{\rm c}(i)$, which is $E_{v,x}\{\tilde{v}^{\rm c}(i)\tilde{v}^{\rm c}(j)\}$. It can be shown that the function has the same value for $i \neq j$. Let

$$E_{v,x}\{\tilde{v}^{\rm c}(i)\tilde{v}^{\rm c}(j)\} = \begin{cases} p_v, & i=j, \\ q_v, & i \neq j. \end{cases} \quad (28)$$

To solve this problem, we first consider a simple two-chip case in which $N=2$. Then,

$$\tilde{v}^{\rm c}(0) + \tilde{v}^{\rm c}(1) = \gamma, \quad (29)$$

where the unconstrained variables $v^{\rm c}(0)$ and $v^{\rm c}(1)$ are two i.i.d. random variables with zero mean and variance σ_v^2. We can evaluate the conditional joint Gaussian probability function of $\{\tilde{v}^{\rm c}(0), \tilde{v}^{\rm c}(1)\}$ as

$$f(\tilde{v}^{\rm c}(0), \tilde{v}^{\rm c}(1))$$
$$= \frac{1}{2\pi\sigma_v^2}\exp\left\{-\frac{(\tilde{v}^{\rm c}(0))^2 + (\tilde{v}^{\rm c}(1))^2}{2\sigma_v^2}\right\}$$
$$= C\frac{1}{\sqrt{2\pi\cdot\sigma_v^2/2}}\exp\left\{-\frac{1}{2}\frac{[\tilde{v}^{\rm c}(0) - \frac{1}{2}\gamma]^2}{\frac{1}{2}\sigma_v^2}\right\}$$
$$= Cf(\tilde{v}^{\rm c}(0)), \quad (30)$$

where C is a normalization constant. From (30), we can obtain $\tilde{v}_M^{\rm c}(0) = \tilde{v}_M^{\rm c}(1) = \frac{1}{2}\gamma$, and $\tilde{v}_V^{\rm c}(0) = \frac{1}{2}\sigma_v^2$. Multiplying $\tilde{v}^{\rm c}(0)$ and taking expectation on the both sides of (29), we can obtain

$$[\tilde{v}_M^{\rm c}(0)]^2 + \tilde{v}_V^{\rm c}(0) + q_v = \gamma\tilde{v}_M^{\rm c}(0). \quad (31)$$

Substituting $\tilde{v}_M^{\rm c}(0)$ with $\gamma/2$ into (31), we can obtain q_v as $q_v = \gamma^2/4 - \sigma^2/(2N)$. Direct extension of the above derivation to $N>2$ is difficult since we have to evaluate multi-dimensional integrations. We now use a simple method to overcome this problem. First, we let N be even and rewrite the γ-constrained equation as

$$\tilde{\theta}^{\rm c}(0) + \tilde{\theta}^{\rm c}(1) = \gamma \quad \text{with} \begin{cases} \tilde{\theta}^{\rm c}(0) = \sum_{i=0}^{N/2-1}\tilde{v}^{\rm c}(i), \\ \tilde{\theta}^{\rm c}(1) = \sum_{i=N/2}^{N-1}\tilde{v}^{\rm c}(i). \end{cases} \quad (32)$$

The unconstrained variables $\theta^{\rm c}(0)$ and $\theta^{\rm c}(1)$ are i.i.d. random variables with the same distribution. We can then apply the result in (31) and obtain

$$(\tilde{\theta}_M^{\rm c}(0))^2 + \tilde{\theta}_V^{\rm c}(0) + q_\theta = \gamma\tilde{\theta}_M^{\rm c}(0), \quad (33)$$

where $q_\theta = E_{v,x}\{\tilde{\theta}^{\rm c}(0)\tilde{\theta}^{\rm c}(1)\}$. Note that $\tilde{\theta}_V^{\rm c}(0) = \frac{1}{2}\theta_V^{\rm c}(0)$ with $\theta_V^{\rm c}(0) = N\sigma_v^2/2$, and $\tilde{\theta}_M^{\rm c}(0) = \gamma/2$. Combing with (33), we can then obtain q_θ. Note that

$$q_\theta = E_{v,x}\left\{\sum_{i=0}^{N/2-1}\sum_{j=N/2}^{N-1}\tilde{v}^{\rm c}(i)\tilde{v}^{\rm c}(j)\right\} = \left(\frac{N}{2}\right)^2 q_v. \quad (34)$$

Thus, from (34) we can obtain the cross-correlation q_v for $N>2$ as

$$q_v = \frac{\gamma^2}{N^2} - \frac{\sigma_v^2}{N}. \quad (35)$$

Multiplying $\tilde{v}^{\rm c}(i)$ on both sides of (32) and taking expectation, we have $p_v + (N-1)q_v = \gamma^2/N$ where $\tilde{v}_M^{\rm c}(i) = \gamma/N$ for $N>2$ is used. Finally, we obtain

$$p_v = \frac{\gamma^2}{N^2} + \frac{N-1}{N}\sigma_v^2. \quad (36)$$

Simulation results show that the result (derived for an even N) is also very accurate for an odd N. We can then have an explicit expression of the first summation term on the right-hand side of (27) as

$$E_{v,x}\left\{\left[\sum_{i=0}^{n-1}\alpha^{n-1-i}\tilde{v}^{\rm c}(i)\right]^2\right\}$$
$$= q_v\left(\frac{1-\alpha^n}{1-\alpha}\right)^2 + (p_v - q_v)\left(\frac{1-\alpha^{2n}}{1-\alpha^2}\right). \quad (37)$$

Thus, combining (27), (35)–(37), we have rewrite (26) as

$$\tilde{\varepsilon}_V^{\rm c}(n) = \frac{\mu^2}{N^2}\left\{N\sigma^2\left(\frac{1-\alpha^{2n}}{1-\alpha^2}\right) - \sigma^2\left(\frac{1-\alpha^n}{1-\alpha}\right)^2\right\}, \quad (38)$$

where the relation $\sigma_v^2 = \sigma^2/N$ is used. By definition and (26), we have

$$\varepsilon_V^{\rm c}(n) = E_\gamma\{[\alpha^n(\tilde{\varepsilon}^{\rm c}(0) - \varepsilon^{\rm c}(0)) + \tilde{\delta}^{\rm c}]^2\} + \varepsilon_V^{\rm c}(n). \quad (39)$$

Note that the result in (38) is independent of γ; it is a function of noise variance and the step size only. Thus, $\varepsilon_V^{c}(n) = E_\gamma\{\tilde{\varepsilon}_V^{c}(n)\} = \tilde{\varepsilon}_V^{c}(n)$. Then the second term in the right-hand side of (39) can be evaluated using (38). Denote the first term in (39) as $\beta^{c}(n)$. Then,

$$\varepsilon_V^{c}(n) = \beta^{c}(n) + \varepsilon_V^{c}(n). \quad (40)$$

The term $\beta^{c}(n)$ can be further evaluated as

$$\begin{aligned}\beta^{c}(n) &= E_\gamma\{[\alpha^n(\tilde{\varepsilon}^{c}(0) - \varepsilon^{c}(0)) + \tilde{\delta}^{c}]^2\} \\ &= (1-\alpha^n)^2(E_\gamma\{[\tilde{w}^{c}_{opt}]^2\} - [w^{c}_{opt}]^2),\end{aligned} \quad (41)$$

where the second moment of $\tilde{w}^{c}_{opt}$ is given by $E_\gamma\{[\tilde{w}^{c}_{opt}]^2\} = E_\gamma\{(a+\gamma)^2 | \gamma \in \mathbb{V}^{c}\}$. Thus, we can obtain the weight-error power shown in (25) using (24), (38), (39) and (41) as $E\{[\tilde{\varepsilon}^{c}(n)]^2\} = \alpha^{2n}[\varepsilon^{c}(0)]^2 + \beta^{c}(n) + \tilde{\varepsilon}_V^{c}(n)$.

4. Analysis for two-user scenario

Extending the procedure developed in the previous section, we now proceed to analyze the two-user case. Only the exact analysis for the optimal weights and convergent weight-error means are derived, while the closed-form expression for the weight-error variance is difficult to obtain and would be represented by approximate analysis. As previous, we only present the results for correct decision (denoted with superscript 'c'). Derivations for erroneous decision are similar.

4.1. Optimal weight analysis

Define the two-by-two user correlation matrix as $\mathbf{R}$ and its entries to be $\mathbf{R}[i,j] = E\{x_i[n]x_j[n]\} = 1/N\sum_{n=0}^{N-1} x_i(n)x_j(n)$ under the assumption of ergodic property is assumed. We also let the time-averaged correlation between these two users' codes is given by $\rho = \sum_{n=0}^{N-1} x_1(n)x_2(n)$. Note that $N\rho$ is an integer. It is simple to show that

$$\mathbf{R} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \quad (42)$$

The matched filter output vector, denoted by $\mathbf{y} = [y_1, y_2]^{T}$, is then

$$\mathbf{y} = \sum_{n=0}^{N-1} \mathbf{x}(n) r(n) = \mathbf{RAb} + \sum_{n=0}^{N-1} \boldsymbol{v}(n), \quad (43)$$

where $\mathbf{b} = [b_1, b_2]^{T}$ is the data bit vector, $\mathbf{A} = \mathrm{diag}\{a_1, a_2\}$ is the channel amplitude matrix, and $\boldsymbol{v}(n) = [v_1(n), v_2(n)]^{T}$ is the noise vector after code multiplication. Let the second term in the right-hand side of (43) be denoted as $\boldsymbol{\gamma} = [\gamma_1, \gamma_2]^{T}$. Then $\boldsymbol{\gamma} = \sum_{n=0}^{N-1} \boldsymbol{v}(n)$, and $\mathbf{y} = \mathbf{RAb} + \boldsymbol{\gamma}$. As that in the single-user case, the decision in the first stage depends on the value of $\boldsymbol{\gamma}$. However, the problem here becomes more involved since the distribution of $\boldsymbol{\gamma}$ depends on ρ also. It can be shown that the joint probability density function for the random vector $\boldsymbol{\gamma}$ is Gaussian and

$$f(\boldsymbol{\gamma}) = \frac{1}{2\pi|\mathbf{C}_\gamma|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\gamma}^{T}\mathbf{C}_\gamma^{-1}\boldsymbol{\gamma}\right\}, \quad (44)$$

where the covariance matrix is given as

$$\mathbf{C}_\gamma \triangleq E\{\boldsymbol{\gamma}\boldsymbol{\gamma}^{T}\} = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}. \quad (45)$$

Now, the number of bits for decision is two. Let k be 1 or 2 and $j = 3-k$. Define the set for which User k's decision is correct as

$$\begin{aligned}\mathbb{V}_k^{c} &\triangleq \{\gamma_k | b_k = \hat{b}_k\} \\ &= \begin{cases} \gamma_k > -(a_k + a_j b_j \rho) & \text{for } b_k = 1, \\ \gamma_k < a_k - a_j b_j \rho & \text{for } b_k = -1. \end{cases}\end{aligned} \quad (46)$$

Similarly the noise subset for making erroneous decision is represented as

$$\begin{aligned}\mathbb{V}_k^{e} &\triangleq \{\gamma_k | b_k \neq \hat{b}_k\} \\ &= \begin{cases} \gamma_k < -(a_k + a_j b_j \rho) & \text{for } b_k = 1, \\ \gamma_k > a_k - a_j b_j \rho & \text{for } b_k = -1. \end{cases}\end{aligned} \quad (47)$$

We then extend our notations defined in the previous section. Let a random variable z conditioned on $\boldsymbol{\gamma}$ and then on ρ be denoted as $\tilde{z}$, i.e., $\tilde{z} = \{z|\boldsymbol{\gamma}|\rho\}$. Also let $\tilde{z}_M = E_{v,x}\{\tilde{z}\}$, $\check{z}_M = E_\gamma\{\tilde{z}_M\}$, and $z_M = E_\rho\{\check{z}_M\}$. We then have $z_M = E\{z\}$. Using the similar rule, we define the optimal weight conditioned on $\boldsymbol{\gamma}$ and then on ρ as $\tilde{\mathbf{w}}_{opt}$, the optimal weight conditioned on ρ as $\check{\mathbf{w}}_{opt}$, and the optimal weight as $\mathbf{w}_{opt}$. We then have $\check{\mathbf{w}}_{opt} = E_\gamma\{\tilde{\mathbf{w}}_{opt}\}$ and $\mathbf{w}_{opt} = E_\rho\{\check{\mathbf{w}}_{opt}\}$. Recall from (6) that $\boldsymbol{\chi}(n) = \hat{\mathbf{B}}\mathbf{x}(n)$ (with superscript $i = 2$ omitted) where $\hat{\mathbf{B}}$ is a diagonal matrix with $\hat{\mathbf{B}}[i,i] = \hat{b}_i$ for bit decision in the first stage. Then the optimal weight conditioned on $\boldsymbol{\gamma}$ and then on ρ can be represented as $\tilde{\mathbf{w}}_{opt} = \tilde{\mathbf{Q}}^{-1}\tilde{\mathbf{p}}$ where the correlation matrix of input signals is expressed by

$$\tilde{\mathbf{Q}} = E_{v,x}\{\tilde{\boldsymbol{\chi}}(n)\tilde{\boldsymbol{\chi}}(n)^{T}\} = \frac{1}{N}\hat{\mathbf{B}}\mathbf{R}\hat{\mathbf{B}}. \quad (48)$$

The cross-correlation vector is given by

$$\tilde{\mathbf{p}} \triangleq E_{v,x}\{\tilde{\boldsymbol{\chi}}(n)\tilde{r}(n)\} = \frac{\hat{\mathbf{B}}}{N}(\mathbf{RAb} + \tilde{\boldsymbol{\gamma}}). \tag{49}$$

Thus, the conditional optimal weight vector is

$$\tilde{\mathbf{w}}_{\text{opt}} = \mathbf{A}\hat{\mathbf{B}}\mathbf{b} + \hat{\mathbf{B}}\mathbf{R}^{-1}\tilde{\boldsymbol{\gamma}}. \tag{50}$$

As we can see from (50), the optimal weights depend on the decision patterns in $\hat{\mathbf{B}}$. There are four decision patterns, i.e., $\{\hat{b}_1 = b_1, \hat{b}_2 = b_2\}$, $\{\hat{b}_1 = b_1, \hat{b}_2 \neq b_2\}$, $\{\hat{b}_1 \neq b_1, \hat{b}_2 \neq b_2\}$, and $\{\hat{b}_1 \neq b_1, \hat{b}_2 = b_2\}$. Note that for each decision pattern, we have two bit patterns that $b_1 = b_2$ and $b_1 \neq b_2$. Let $\mathbb{U}^{ij}$ denote the set of $\boldsymbol{\gamma}$ yielding the ith decision for the jth bit pattern. For example,

$$\mathbb{U}^{11} = \{\boldsymbol{\gamma}|\gamma_1 \in \mathbb{V}_1^{\mathsf{c}}, \gamma_2 \in \mathbb{V}_2^{\mathsf{c}}, b_1 = b_2\}, \tag{51}$$

$$\mathbb{U}^{12} = \{\boldsymbol{\gamma}|\gamma_1 \in \mathbb{V}_1^{\mathsf{c}}, \gamma_2 \in \mathbb{V}_2^{\mathsf{c}}, b_1 \neq b_2\}. \tag{52}$$

Let $\tilde{z}^{ij} = \{z|\boldsymbol{\gamma} \in \mathbb{U}^{ij}|\rho\}$. Also let $\tilde{z}_M^{ij} = E_{v,x}\{\tilde{z}^{ij}\}$. We can have similar notations for optimal weights. Let the optimal weight conditioned on $\boldsymbol{\gamma} \in \mathbb{U}^{ij}$ and then on ρ as $\tilde{\mathbf{w}}_{\text{opt}}^{ij}$ and $\check{\mathbf{w}}_{\text{opt}}^{ij} = E_{\gamma}\{\tilde{\mathbf{w}}_{\text{opt}}^{ij}\}$. Then,

$$\check{\mathbf{w}}_{\text{opt}}^{ij} = \mathbf{A}\hat{\mathbf{B}}^i\mathbf{b}^j + \hat{\mathbf{B}}^i\mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{ij}\}, \tag{53}$$

where $\hat{\mathbf{B}}^i$ denotes $\hat{\mathbf{B}}$ associated with the ith decision pattern, and $\mathbf{b}^j$ denotes the jth bit pattern.

If we further assume that $b_1 = 1$, we have

$$\begin{aligned}\mathbb{U}^{11} &= \{\boldsymbol{\gamma}|\gamma_1 \in \mathbb{V}_1^{\mathsf{c}}, \gamma_2 \in \mathbb{V}_2^{\mathsf{c}}, b_1 = 1 = b_2\} \\ &= \{\boldsymbol{\gamma}|\gamma_1 > -(a_1 + a_2\rho), \gamma_2 > -(a_2 + a_1\rho)\}\end{aligned} \tag{54}$$

and

$$\begin{aligned}\mathbb{U}^{12} &= \{\boldsymbol{\gamma}|\gamma_1 \in \mathbb{V}_1^{\mathsf{c}}, \gamma_2 \in \mathbb{V}_2^{\mathsf{c}}, b_1 = 1 \neq b_2\} \\ &= \{\boldsymbol{\gamma}|\gamma_1 > -(a_1 - a_2\rho), \gamma_2 > -(a_2 - a_1\rho)\}.\end{aligned} \tag{55}$$

The conditional optimal weights for $\hat{\mathbf{B}}^1$ become

$$\begin{aligned}\check{\mathbf{w}}_{\text{opt}}^{11} &= \boldsymbol{a} + \mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{11}\}, \\ \check{\mathbf{w}}_{\text{opt}}^{12} &= \boldsymbol{a} + \mathbf{J}\mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{12}\},\end{aligned} \tag{56}$$

where $\boldsymbol{a} = [a_1, a_2]^{\mathrm{T}}$ and $\mathbf{J} \triangleq \text{diag}\{1, -1\}$. The result for $b = -1$ is identical to that in (56) since in (53) the product in $\hat{\mathbf{B}}^i\mathbf{b}^j$ or $\hat{\mathbf{B}}^i\mathbf{R}^{-1}E_{\gamma}^{ij}\{\gamma\}$ is independent of the value of b_1. The components in $E_{\gamma}\{\boldsymbol{\gamma}^{ij}\}$ are given by

$$E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{ij}\} = \frac{\int_{\mathbb{U}^{ij}} \boldsymbol{\gamma} f(\boldsymbol{\gamma})\,\mathrm{d}\boldsymbol{\gamma}}{\int_{\mathbb{U}^{ij}} f(\boldsymbol{\gamma})\,\mathrm{d}\boldsymbol{\gamma}}. \tag{57}$$

The complete set of $\boldsymbol{\gamma}$ for all decision and bit patterns is shown in Table 1, and the complete set of conditional optimal weights is given in Table 2.

Table 1
Sets of $\boldsymbol{\gamma}$ for all decision and bit patterns

$\mathbb{U}^{ij}$	$\hat{\mathbf{B}}^i\mathbf{b}^j$	Range for γ_1	Range for γ_I
$\mathbb{U}^{11}$	$\begin{bmatrix}1\\1\end{bmatrix}$	$\gamma_1 > -(a_1 + a_I\rho)$	$\gamma_I > -(a_1\rho + a_I)$
$\mathbb{U}^{12}$		$\gamma_1 > -(a_1 - a_I\rho)$	$\gamma_I < -(a_1\rho - a_I)$
$\mathbb{U}^{21}$	$\begin{bmatrix}1\\-1\end{bmatrix}$	$\gamma_1 > -(a_1 + a_I\rho)$	$\gamma_I < -(a_I + a_1\rho)$
$\mathbb{U}^{22}$		$\gamma_1 > -(a_1 - a_I\rho)$	$\gamma_I > -(a_1\rho - a_I)$
$\mathbb{U}^{31}$	$\begin{bmatrix}-1\\-1\end{bmatrix}$	$\gamma_1 < -(a_1 + a_I\rho)$	$\gamma_I < -(a_I + a_1\rho)$
$\mathbb{U}^{32}$		$\gamma_1 < -(a_1 - a_I\rho)$	$\gamma_I > -(a_1\rho - a_I)$
$\mathbb{U}^{41}$	$\begin{bmatrix}-1\\1\end{bmatrix}$	$\gamma_1 < -(a_1 + a_I\rho)$	$\gamma_I > -(a_I + a_1\rho)$
$\mathbb{U}^{42}$		$\gamma_1 < -(a_1 - a_I\rho)$	$\gamma_I < -(a_1\rho - a_I)$

Table 2
Complete list of conditional optimal weights

$\check{\mathbf{w}}_{\text{opt}}^{11} = \boldsymbol{a} + \mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{11}\}$	$\check{\mathbf{w}}_{\text{opt}}^{12} = \boldsymbol{a} + \mathbf{J}\mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{12}\}$
$\check{\mathbf{w}}_{\text{opt}}^{21} = -\boldsymbol{a} + \mathbf{J}\mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{21}\}$	$\check{\mathbf{w}}_{\text{opt}}^{22} = -\boldsymbol{a} + \mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{22}\}$
$\check{\mathbf{w}}_{\text{opt}}^{31} = \mathbf{J}\boldsymbol{a} - \mathbf{J}\mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{31}\}$	$\check{\mathbf{w}}_{\text{opt}}^{32} = \mathbf{J}\boldsymbol{a} - \mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{32}\}$
$\check{\mathbf{w}}_{\text{opt}}^{41} = -\mathbf{J}\boldsymbol{a} - \mathbf{J}\mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{41}\}$	$\check{\mathbf{w}}_{\text{opt}}^{42} = -\mathbf{J}\boldsymbol{a} - \mathbf{R}^{-1}E_{\gamma}\{\tilde{\boldsymbol{\gamma}}^{42}\}$

Our objective is to determine $\mathbf{w}_{\text{opt}}^{\mathsf{c}}$ by taking expectation on $\check{\mathbf{w}}_{\text{opt}}^{\mathsf{c}}$. As seen from Table 1, the region of γ_1 for correct decision is different from that of γ_2. Thus we have to determine the components of $\check{\mathbf{w}}_{\text{opt}}^{\mathsf{c}}$ user-by-user. The union of noise subsets for the first user to have correct decision is then

$$\mathbb{C}_1 = \mathbb{U}^{11} \cup \mathbb{U}^{12} \cup \mathbb{U}^{21} \cup \mathbb{U}^{22}, \tag{58}$$

where $\cup$ denotes the union operation. The occurrence probability for $\mathbb{U}^{ij}$ is obtained as $P_{ij} = \int_{\mathbb{U}^{ij}} f(\boldsymbol{\gamma})\,\mathrm{d}\boldsymbol{\gamma}$. Note from (44) that P_{ij} is dependent upon ρ. The optimal weight for the first user with a given ρ, denoted as $\check{w}_{\text{opt},1}^{\mathsf{c}}$, is

$$\check{w}_{\text{opt},1}^{\mathsf{c}} = \frac{1}{P_{\mathbb{C}_1}} \sum_{\mathbb{C}_1} \check{w}_{\text{opt},1}^{ij} P_{ij}, \tag{59}$$

where $P_{\mathbb{C}_1} = P_{11} + P_{12} + P_{21} + P_{22}$. The optimal weight for the second user with a given ρ, denoted as $\check{w}_{\text{opt},2}^{\mathsf{c}}$, is

$$\check{w}_{\text{opt},2}^{\mathsf{c}} = \frac{1}{P_{\mathbb{C}_2}} \sum_{\mathbb{C}_2} \check{w}_{\text{opt},2}^{ij} P_{ij}, \tag{60}$$

where $\mathbb{C}_2 = \mathbb{U}^{11} \cup \mathbb{U}^{12} \cup \mathbb{U}^{41} \cup \mathbb{U}^{42}$ and $P_{\mathbb{C}_2} = P_{11} + P_{12} + P_{41} + P_{42}$. The optimal weight is obtained through averaging $\check{\mathbf{w}}_{\text{opt}}^{\mathsf{c}}$ over all ρ values by

$$w^{\mathrm{c}}_{\mathrm{opt},i} = E_\rho\{\breve{w}^{\mathrm{c}}_{\mathrm{opt},i}\} = \frac{\sum_\rho \breve{w}^{\mathrm{c}}_{\mathrm{opt},i} P_\rho P_{\mathbb{C}_i}}{\sum_\rho P_\rho P_{\mathbb{C}_i}}, \tag{61}$$

where $i = 1, 2$ and the distribution for the correlation coefficient is given by

$$P_\rho = \frac{1}{2^N}\begin{pmatrix} N \\ N(1+\rho)/2 \end{pmatrix}. \tag{62}$$

4.2. Weight-error mean analysis

Define the weight error for correct decision as $\tilde{\varepsilon}^{\mathrm{c}}_i(n) = \tilde{w}^{\mathrm{c}}_i(n) - w^{\mathrm{c}}_{\mathrm{opt},i}, i = 1, 2$. From the optimal weights of the two-user case, we see that $\breve{\mathbf{w}}^{\mathrm{c}}_{\mathrm{opt}}$ is derived from $\breve{\mathbf{w}}^{ij}_{\mathrm{opt}}$'s. Thus, we also consider the conditional weight errors as $\tilde{\boldsymbol{\varepsilon}}^{ij}(n) = \tilde{\mathbf{w}}^{ij}(n) - \breve{\mathbf{w}}^{ij}_{\mathrm{opt}}$ where the conditional weights are defined as $\tilde{\mathbf{w}}^{ij}(n) = \{\mathbf{w}(n) | \gamma \in \mathbb{U}^{ij} | \rho\}$. As that in the single-user case, our goal is to determine closed-form expressions of $E\{\varepsilon^{\mathrm{c}}_i(n)\}$ being equal to $E_\rho\{\breve{\varepsilon}^{\mathrm{c}}_{M,i}(n)\}$. By definition $\breve{\mathbf{w}}^{ij}_{\mathrm{opt}} = E_\gamma\{\tilde{\mathbf{w}}^{ij}_{\mathrm{opt}}\}$. We then define

$$\tilde{\boldsymbol{\varepsilon}}^{ij}(n) = \tilde{\mathbf{w}}^{ij}(n) - \tilde{\mathbf{w}}^{ij}_{\mathrm{opt}},$$
$$\tilde{\boldsymbol{\delta}}^{ij} = \tilde{\mathbf{w}}^{ij}_{\mathrm{opt}} - \breve{\mathbf{w}}^{ij}_{\mathrm{opt}}. \tag{63}$$

We then have $\tilde{\boldsymbol{\varepsilon}}^{ij}(n) = \tilde{\boldsymbol{\varepsilon}}^{ij}(n) + \tilde{\boldsymbol{\delta}}^{ij}$. It is obvious that $\tilde{\boldsymbol{\delta}}^{ij}_M = [0, 0]^{\mathrm{T}}$. Thus we have $\tilde{\boldsymbol{\varepsilon}}^{ij}_M(n) = \tilde{\boldsymbol{\varepsilon}}^{ij}_M(n)$. Thus $\breve{\boldsymbol{\varepsilon}}^{ij}_M(n) = \breve{\boldsymbol{\varepsilon}}^{ij}_M(n) = E_\gamma\{\tilde{\boldsymbol{\varepsilon}}^{ij}_M(n)\}$. We then have

$$\begin{aligned}\tilde{\boldsymbol{\varepsilon}}^{ij}(n) &= \tilde{\mathbf{w}}^{ij}(n) - \tilde{\mathbf{w}}^{ij}_{\mathrm{opt}} \\ &= \tilde{\mathbf{w}}^{ij}(n-1) - \tilde{\mathbf{w}}^{ij}_{\mathrm{opt}} + \mu\tilde{\boldsymbol{\chi}}^{ij}(n-1)\tilde{e}^{ij}(n-1) \\ &= (\mathbf{I} - \mu\tilde{\boldsymbol{\Sigma}}^{ij}(n-1))\tilde{\boldsymbol{\varepsilon}}^{ij}(n-1) + \mu\tilde{\boldsymbol{\Psi}}^{ij}(n-1),\end{aligned} \tag{64}$$

where $\tilde{\boldsymbol{\Sigma}}^{ij}(n) \triangleq \tilde{\boldsymbol{\chi}}^{ij}(n)[\tilde{\boldsymbol{\chi}}^{ij}(n)]^{\mathrm{T}}$ and $\tilde{\boldsymbol{\Psi}}^{ij}(n) \triangleq \tilde{\boldsymbol{\chi}}^{ij}(n)\tilde{r}^{ij}(n) - \tilde{\boldsymbol{\Sigma}}^{ij}(n)\tilde{\mathbf{w}}^{ij}_{\mathrm{opt}}$. It can be easily shown by deduction that

$$\begin{aligned}\tilde{\boldsymbol{\varepsilon}}^{ij}(n) = &\prod_{m=0}^{n-1}(\mathbf{I} - \mu\tilde{\boldsymbol{\Sigma}}^{ij}(m))\tilde{\boldsymbol{\varepsilon}}^{ij}(0) \\ &+ \mu\sum_{m=0}^{n-1}\prod_{l=m+1}^{n-1}(\mathbf{I} - \mu\tilde{\boldsymbol{\Sigma}}^{ij}(l))\tilde{\boldsymbol{\Psi}}^{ij}(m) \\ = &\widetilde{\mathbf{W}}^{ij}(n,-1)\tilde{\boldsymbol{\varepsilon}}^{ij}(0) + \mu\sum_{m=0}^{n-1}\widetilde{\mathbf{W}}^{ij}(n,m)\tilde{\boldsymbol{\Psi}}^{ij},\end{aligned} \tag{65}$$

where the substitution $\widetilde{\mathbf{W}}^{ij}(n,m) \triangleq \prod_{l=m+1}^{n-1}(\mathbf{I} - \mu\tilde{\boldsymbol{\Sigma}}^{ij}(l))$ is used. It can be observed that the expectation of $\tilde{\boldsymbol{\Psi}}^{ij}(n)$ can be expressed through (48)–(50) as

$$\begin{aligned}\tilde{\boldsymbol{\Psi}}^{ij}_M(n) &= E_{v,x}\{\tilde{\boldsymbol{\chi}}^{ij}(n)\tilde{r}^{ij}(n) - \tilde{\boldsymbol{\Sigma}}^{ij}(n)\tilde{\mathbf{w}}^{ij}_{\mathrm{opt}}\} \\ &= E_{v,x}\{\tilde{\boldsymbol{\chi}}^{ij}(n)\tilde{r}^{ij}(n)\} - \tilde{\mathbf{Q}}^{ij}\tilde{\mathbf{w}}^{ij}_{\mathrm{opt}} \\ &= \frac{\hat{\mathbf{B}}^i}{N}(\mathbf{RAb}^j + \tilde{\gamma}) - \frac{\hat{\mathbf{B}}^i\mathbf{R}\hat{\mathbf{B}}^i}{N}[\mathbf{A}\hat{\mathbf{B}}^i\mathbf{b}^j + \hat{\mathbf{B}}^i\mathbf{R}^{-1}\tilde{\gamma}] \\ &= [0, 0]^{\mathrm{T}}.\end{aligned} \tag{66}$$

We also obtain

$$\begin{aligned}\widetilde{\mathbf{W}}^{ij}_M(n,m) &= E_{v,x}\left\{\prod_{m=0}^{n-1}(\mathbf{I} - \mu\tilde{\boldsymbol{\chi}}^{ij}(m)[\tilde{\boldsymbol{\chi}}^{ij}(m)]^{\mathrm{T}})\right\} \\ &= E_{v,x}\left\{\prod_{m=0}^{n-1}(\mathbf{I} - \mu\hat{\mathbf{B}}^i\mathbf{x}(m)[\mathbf{x}(m)]^{\mathrm{T}}\hat{\mathbf{B}}^i)\right\} \\ &= \prod_{m=0}^{n-1}[\mathbf{I} - \mu\hat{\mathbf{B}}^i\mathbf{R}\hat{\mathbf{B}}^i] \\ &= (\mathbf{I} - \mu\hat{\mathbf{B}}^i\mathbf{R}\hat{\mathbf{B}}^i)^n.\end{aligned} \tag{67}$$

We then have

$$\begin{aligned}\breve{\boldsymbol{\varepsilon}}^{ij}_M(n) = \breve{\boldsymbol{\varepsilon}}^{ij}_M(n) &= \left(\mathbf{I} - \frac{\mu}{N}\hat{\mathbf{B}}^i\mathbf{R}\hat{\mathbf{B}}^i\right)^n\breve{\boldsymbol{\varepsilon}}^{ij}_M(0) \\ &= \left(\mathbf{I} - \frac{\mu}{N}\hat{\mathbf{B}}^i\mathbf{R}\hat{\mathbf{B}}^i\right)^n\breve{\boldsymbol{\varepsilon}}^{ij}_M(0).\end{aligned} \tag{68}$$

Let $\breve{\boldsymbol{\varepsilon}}^{ij}_M(n) = [\breve{\varepsilon}^{ij}_{M,1}(n), \breve{\varepsilon}^{ij}_{M,2}(n)]^{\mathrm{T}}$. The weight-error mean for the first user conditioned on only the correct decision and ρ is represented by

$$\breve{\varepsilon}^{\mathrm{c}}_{M,1}(n) = \frac{1}{P_{\mathbb{C}_1}}\sum_{\mathbb{C}_1}\breve{\varepsilon}^{ij}_{M,1}(n)P_{ij}. \tag{69}$$

Similarly the weight-error mean for the second user is represented as

$$\breve{\varepsilon}^{\mathrm{c}}_{M,2}(n) = \frac{1}{P_{\mathbb{C}_2}}\sum_{\mathbb{C}_2}\breve{\varepsilon}^{ij}_{M,2}(n)P_{ij}. \tag{70}$$

Then, the averaged weight-error mean for correct decision can be obtain by

$$E\{\varepsilon^{\mathrm{c}}_i(n)\} = E_\rho\{\breve{\varepsilon}^{\mathrm{c}}_{M,i}(n)\} = \frac{\sum_\rho \breve{\varepsilon}^{\mathrm{c}}_{M,i}(n)P_\rho P_{\mathbb{C}_i}}{\sum_\rho P_\rho P_{\mathbb{C}_i}} \tag{71}$$

for $i = 1, 2$.

From the derivation presented above, it can be noted that the exact analysis for the first-order statistics in the two-user case is complicated. It is simple to image that the complexity of the second-order statistics will even more higher. In view of this, we then provide approximate solutions to the weight-error power behavior as discussed in the sequel.

4.3. Weight-error power analysis

We re-show (40) here as $\varepsilon^{\mathrm{c}}_V(n) = \beta^{\mathrm{c}}(n) + \varepsilon^{\mathrm{c}}_V(n)$ for reference convenience. Assuming that the first user is the desired user and considering (2) for the two-user case, we can obtain

$$y_1 = a_1 b_1 + a_2 b_2 \rho + \gamma_1. \tag{72}$$

From (72) with (10), we found that the desired user output in this two-user environment has one extra interference term which is $a_2 b_2 \rho$, compared to that in the single-user case. Our approach is to take the interference effect into account such that the result derived for the single-user case in (40) can be leveraged. There are two major terms in (40), i.e., $\varepsilon_V^{\mathrm{c}}(n)$ and $\beta^{\mathrm{c}}(n)$. From (38), we first can see that $\varepsilon_V^{\mathrm{c}}(n)$ is a function of step size and noise variance only. Thus, it is reasonable to treat $\varepsilon_V^{\mathrm{c}}(n)$ as the same as that in the single-user case. Secondly, from (41) we find that $\beta^{\mathrm{c}}(n)$ is directly related to optimal weights. Since one additional variable ρ is introduced, we modify (41) as $\beta^{\mathrm{c}}(n) = E_\rho\{\check{\beta}^{\mathrm{c}}(n)\}$ where

$$\check{\beta}^{\mathrm{c}}(n) \triangleq (1-\alpha^n)^2 E_{\gamma_1}\{(\tilde{w}^{\mathrm{c}}_{\mathrm{opt},1} - w^{\mathrm{c}}_{\mathrm{opt},1})^2\}. \tag{73}$$

The next problem is how to obtain the conditioned optimal weight of the desired user in the two-user case, i.e., $\tilde{w}^{\mathrm{c}}_{\mathrm{opt},1}$. As we can see, $\tilde{w}^{\mathrm{c}}_{\mathrm{opt},1}$ is influenced by ρ and γ_2, as shown in (50). It should be noted, however, that the analytical results in (50) cannot be applied directly since (41) or (73) only involve one-dimensional rather than two-dimensional integrations. To solve the problem, we change the desired signal in (72) from a_1 to $a_1 + a_2 b_2 \rho$. In other words, the interference is treated as a bias term with a constant ρ. Referring to the single-user case in (15), we can have the approximate solution (assuming $b_1 = 1$) as

$$\tilde{w}^{\mathrm{c}}_{\mathrm{opt},1} = a_1 + a_2 b_2 \rho + \tilde{\gamma}_1. \tag{74}$$

Note that in (74) the number of noise variables is reduced from two into one. As we can see, the first user decision is treated as independent of the second user decision. This means that we ignore some coupling relationship between two users such that the ρ-related integration region partition is reduced from two dimensions into one dimension. This can be obtained readily by properly combining the complementary partitions of $\mathbb{U}^{ij}$ in Table 1. In this case, the total number of decision patterns are degenerated from four into two, $b_1 = \hat{b}_1$ and $b_1 \neq \hat{b}_1$. We denote these decision patterns as the fifth and the sixth pattern. For each decision pattern, we have two bit patterns for different optimal weights, i.e., $b_1 = b_2$ and $b_1 \neq b_2$, as implied in (74). The noise space can be partitioned into two subsets accordingly. Thus, for $b_1 = \hat{b}_1$, we have two sets as ($b_1 = 1$)

$$\begin{aligned} \mathbb{U}^{51} &= \mathbb{U}^{11} \cup \mathbb{U}^{21}, \quad b_1 = b_2, \\ \mathbb{U}^{52} &= \mathbb{U}^{12} \cup \mathbb{U}^{22}, \quad b_1 \neq b_2. \end{aligned} \tag{75}$$

The conditional optimal weight on $\mathbb{U}^{5j}$, $j = 1, 2$ for correct decision is obtained to be

$$\tilde{w}^{5j}_{\mathrm{opt},1} = \begin{cases} a_1 + a_2\rho + \tilde{\gamma}_1, & j = 1, \\ a_1 - a_2\rho + \tilde{\gamma}_1, & j = 2. \end{cases} \tag{76}$$

We first denote the noise integration region and the corresponding occurrence probabilities for $\mathbb{U}^{51}$ and $\mathbb{U}^{52}$ as

$$\begin{aligned} \mathbb{B} &= \mathbb{U}^{51} \cup \mathbb{U}^{52}, \\ P_{\mathbb{B}} &= P_{51} + P_{52}. \end{aligned} \tag{77}$$

Then we rewrite (73) as

$$\check{\beta}^{\mathrm{c}}(n) = \frac{\check{\beta}^{51}(n) P_{51} + \check{\beta}^{52}(n) P_{52}}{P_{\mathbb{B}}}, \tag{78}$$

where the terms $\check{\beta}^{5j}(n)$, $j = 1, 2$ are obtained as

$$\check{\beta}^{5j}(n) = (1-\alpha^n)^2 E_{\gamma_1}\{(\tilde{w}^{5j}_{\mathrm{opt},1} - w^{\mathrm{c}}_{\mathrm{opt},1})^2\}. \tag{79}$$

The expectation term in (79) is analogous to (41) and can be derived as

$$E_{\gamma_1}\left\{(\tilde{w}^{5j}_{\mathrm{opt},1} - w^{\mathrm{c}}_{\mathrm{opt},1})^2\right\} = \begin{cases} E_{\gamma_1}\{(a_1 + a_2\rho + \gamma_1 - w^{\mathrm{c}}_{\mathrm{opt},1})^2 | \gamma_1 > -(a_1 + a_2\rho)\}, & j = 1, \\ E_{\gamma_1}\{(a_1 - a_2\rho + \gamma_1 - w^{\mathrm{c}}_{\mathrm{opt},1})^2 | \gamma_1 > -(a_1 - a_2\rho)\}, & j = 2. \end{cases} \tag{80}$$

Finally, $\beta^{\mathrm{c}}(n)$, being equal to $E_\rho\{\check{\beta}^{\mathrm{c}}(n)\}$, can be represented as

$$\beta^{\mathrm{c}}(n) = \frac{\sum_\rho \check{\beta}^{\mathrm{c}}(n) P_\rho P_{\mathbb{B}}}{\sum_\rho P_\rho P_{\mathbb{B}}}, \tag{81}$$

where P_ρ is defined in (62). The weight-error power for correct decision is expressed as

$$\begin{aligned} E\{[\tilde{\varepsilon}^{\mathrm{c}}_1(n)]^2\} &= E\{[\tilde{w}^{\mathrm{c}}_1(n) - w^{\mathrm{c}}_{\mathrm{opt},1}]^2\} \\ &= [\varepsilon^{\mathrm{c}}_{M,1}(n)]^2 + \varepsilon^{\mathrm{c}}_V(n) \\ &= \alpha^{2n}[\varepsilon^{\mathrm{c}}_{M,1}(0)]^2 + \beta^{\mathrm{c}}(n) + \varepsilon^{\mathrm{c}}_V(n), \end{aligned} \tag{82}$$

where $\varepsilon^{\mathrm{c}}_{M,1}(0)$ can be obtained from (71). Similarly the weight-error power for erroneous decision can be obtained.

5. Approximate analysis for multiple-user scenario

5.1. Analysis of weight behavior

In prior two sections, we have derived the exact analytical results for the optimal weight, the weight-error mean, and the weight-error variance for the single-user case, the optimal weights, and the weight-error means for the two-user case. We also provide the approximate analysis for weight-error power of the two-user case. In this section, we will extend the results to accommodate the general multiple-user case. Due to the difficulty of the problem, we will seek approximate rather exact solutions. In most cases, we will only give the result for correct decision (denoted with superscript 'c') and omit the derivation for erroneous decision.

First note that the received despread signal of each user composes of three parts, i.e., the desired signal, the MAI, and noise. According to previous analytical results, we have two alternative approaches to approximate the MAI term. The first is to regard the MAI as noise and to apply the single-user model. The other is to treat the MAI as the interference produced by a virtual user, and then to use the two-user model. To have better result, we use the second approach. Let

$$\sum_{j \neq k} a_j b_j \rho_{jk} \simeq a_I b_I \rho, \tag{83}$$

where $b_I \in \{\pm 1\}$ and

$$a_I = \left(\sum_{j \neq 1} a_j^2 \right)^{1/2}. \tag{84}$$

Here, a_I represent the equivalent amplitude and ρ the equivalent correlation of a virtual user. Using this model, we can fit the second-order statistics of MAI. Also note that b_I is virtual and we do not need its actual value in derivation. In the following analysis, we assume that the desired user is the first user. The matched filter output is given by

$$y_1 \simeq a_1 b_1 + a_I b_I \rho + \gamma_1. \tag{85}$$

Thus, we can keep the computational complexity comparable to that of the two-user case.

Using the two-user model in (85), the performance of the adaptive two-stage PIC in the general K-user scenario can be analyzed straightforwardly. All we have to do is to let the amplitude of the second user, a_2, be equal to a_I in (84), and apply corresponding parameters to Table 1. Optimal weights, weight-error means, and weight-error powers can be readily obtained with the procedure described in Section 4.

5.2. Output MSE and BER

Since we have derived the weight-error power for the adapted weights corresponding to both correct and erroneous decisions, we can then calculate the output MSE and then BER. As mentioned in (8), if the adapted weight of the kth user is a_k for $b_k = \hat{b}_k$, or $-a_k$ for $b_k \neq \hat{b}_k$, its interference to other users can be perfectly cancelled. Thus, the MSE introduced to other users when the weight obtained from correct decision at time n is used for cancellation is $E\{[\tilde{w}_k^{\mathsf{c}}(n)\hat{b}_k x_k(n) - a_k b_k x_k(n)]^2\} = E\{[\tilde{w}_k^{\mathsf{c}}(n) - a_k]^2\}/N$. Similarly, the MSE introduced to other users due to erroneous decision can be expressed as $E\{[\tilde{w}_k^{\mathsf{e}}(n) + a_k]^2\}/N$. As a result, the overall MSE, denoted as $\varpi_k(n)$, introduced to other users when the cancellation is performed is then

$$\varpi_k(n) = \frac{1}{N}\{P_{\mathrm{c},k} E\{[\tilde{w}_k^{\mathsf{c}}(n) - a_k]^2\} + P_{\mathrm{e},k} E\{[\tilde{w}_k^{\mathsf{e}}(n) + a_k]^2\}\}, \tag{86}$$

where $P_{\mathrm{c},k}$ and $P_{\mathrm{e},k}$ denote the probability of correct and that of erroneous decision in the first stage for the kth user, respectively. Note that these probabilities can be easily obtained using the method of Gaussian approximation. Substituting the weight-error power in (82) into (86), we can obtain $\varpi_k(n)$ as

$$\begin{aligned} \varpi_k(n) = \frac{1}{N}\Big\{ & P_{\mathrm{c},k}\{E\{[\tilde{\varepsilon}_k^{\mathsf{c}}(n)]^2\} \\ & + 2w_{M,k}^{\mathsf{c}}(n)(w_{\mathrm{opt},k}^{\mathsf{c}} - a_k) + a_k^2 - [w_{\mathrm{opt},k}^{\mathsf{c}}]^2\} \\ & + P_{\mathrm{e},k}\{E\{[\tilde{\varepsilon}_k^{\mathsf{e}}(n)]^2\} + 2w_{M,k}^{\mathsf{e}}(n)(w_{\mathrm{opt},k}^{\mathsf{e}} + a_k) \\ & + a_k^2 - [w_{\mathrm{opt},k}^{\mathsf{e}}]^2\}\Big\}. \end{aligned} \tag{87}$$

Note here that we extend our notations defined previously to the kth user. Assuming that the residual error resulting from imperfect interference cancellation is Gaussian distributed with zero mean and variance $\varpi_k(N)$, we can then obtain the BER for the kth user as

$$P(y_k^{(2)}) = \mathcal{Q}\left\{ \frac{a_k}{\sqrt{\sigma^2 + \sum_{j=1, j\neq k}^{K} \varpi_j(N)}} \right\}. \tag{88}$$

6. Simulation results

In this section we report some simulation results to evaluate the validity of our analytical results. We utilized the random codes as the spreading codes and the processing gain is set as $N = 31$. Only the AWGN channel was used throughout the simulations. For the first set of simulations, we compared theoretical optimal weights with empirical ones for various E_b/N_0 ($N_0/2 = \sigma^2$). Optimal weights for correct and erroneous decision were considered separately. Note that the channel gain was normalized to unity, i.e., $a_k = 1$ for each k. Thus all weights starting adaptation from $w_k(0) = 1$. Fig. 3 shows the result using the approximated analysis and that obtained empirically for five users. As depicted in the figure, optimal weights for correct decision are almost the same as channel gains, while weights for erroneous decision are not; their actual values depend on noise variance. The larger the E_b/N_0 ratio, the closer the optimal weight to $-a_k$. We also give optimal weights for 15 users (with various E_b/N_0) in Fig. 4. We can see that although the approximation is performed based on the two-user model, the results are very close to the true optimums. From these figures, we can observe that when the E_b/N_0 and the number of users vary, optimal weights for correct decision keep very close to channel gains which are unity, while those for erroneous decision vary.

We next consider the weight convergence of the LMS algorithm. Fig. 5 presents the analytical mean weights along with the empirical mean weights for a

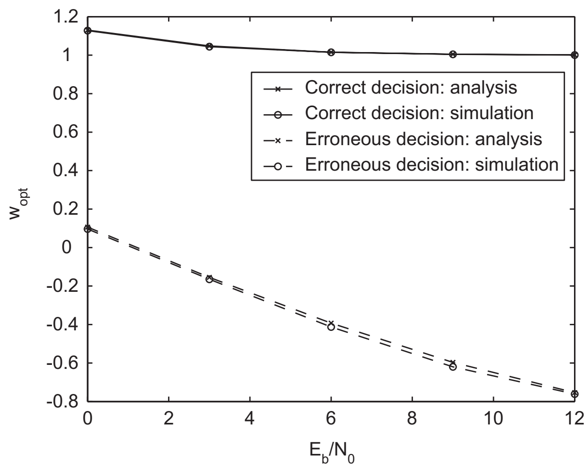

Fig. 3. Optimal weight comparison for five power-balanced users.

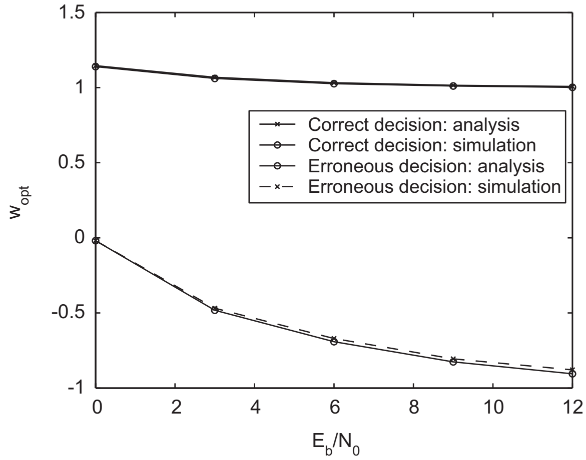

Fig. 4. Optimal weight comparison for 15 power-balanced users.

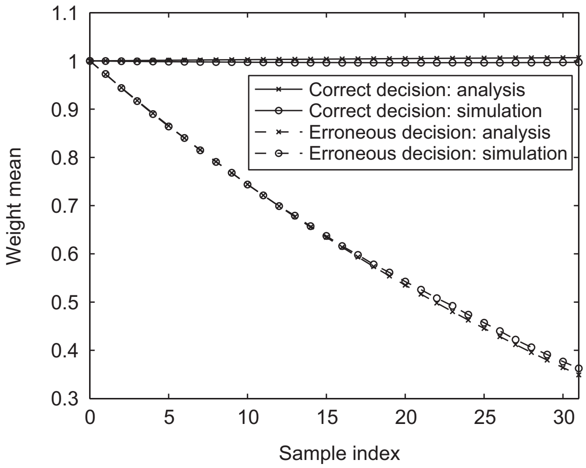

Fig. 5. Weight mean comparison for five power-balanced users ($\mu_0 = 0.02$ and $E_b/N_0 = 6$ dB).

five-user scenario. The powers of these users are equal and $E_b/N_0 = 6$ dB. The normalized step size is chosen as $\mu_0 = \mu/N = 0.02$. In the figure, we can observe that the analytical results match with the empirical mean weights quite well. Similar comparison for 15 power-balanced users with $E_b/N_0 = 6$ dB and $\mu_0 = 0.02$ is also shown in Fig. 6. There is some discrepancy between analytical and empirical results since the number of users is larger. The weight-error power comparison for the five-user case with $E_b/N_0 = 6$ dB and $\mu_0 = 0.02$ is given in Fig. 7. It is obvious that the analytic results perform close to simulated results. Also note that

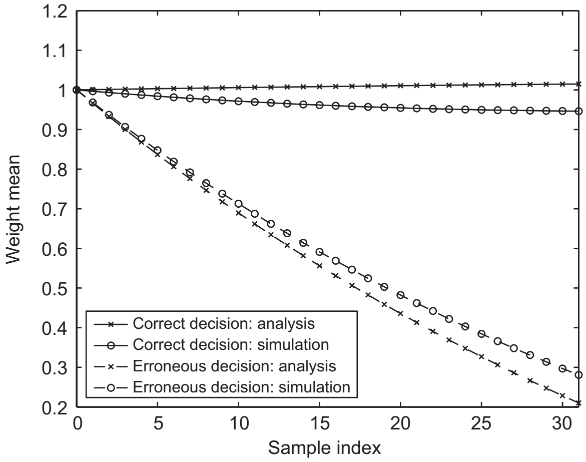

Fig. 6. Weight mean comparison for 15 power-balanced users ($\mu_0 = 0.02$ and $E_b/N_0 = 6\,\text{dB}$).

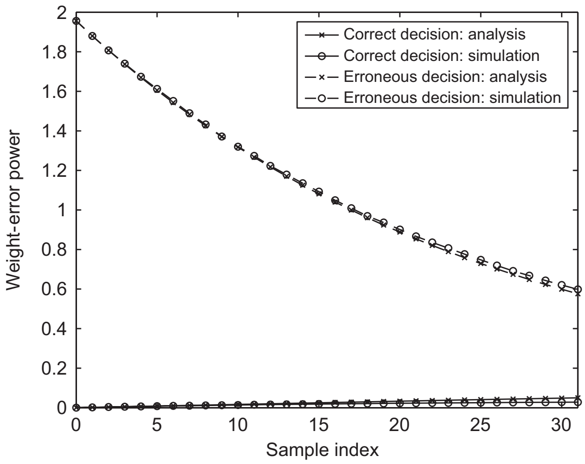

Fig. 7. Weight-error power comparison for five power-balanced users ($\mu_0 = 0.02$ and $E_b/N_0 = 6\,\text{dB}$).

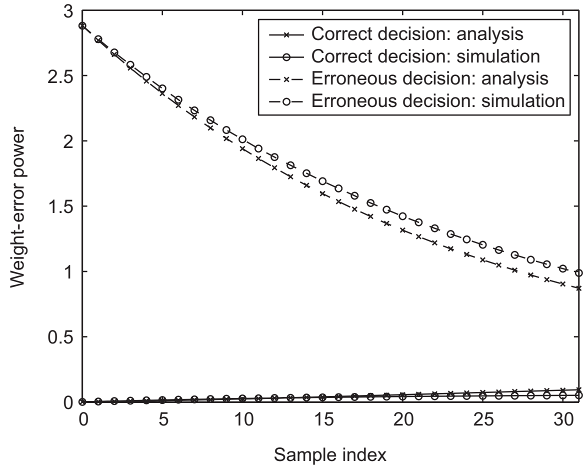

Fig. 8. Weight-error power comparison for 15 power-balanced users ($\mu_0 = 0.02$ and $E_b/N_0 = 6\,\text{dB}$).

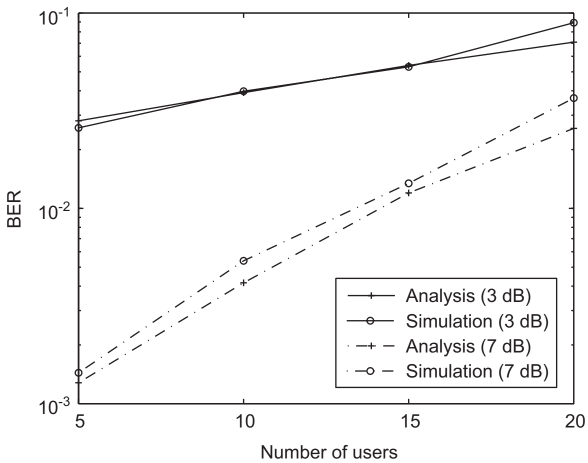

Fig. 9. Second-stage BER comparison for power-balanced cases ($\mu_0 = 0.02$).

the weight-error power incurred from correct decision is smaller than that from erroneous decision. This is because the weights for erroneous decision converge slower. The similar phenomenon can be observed when the user number is larger. In Fig. 8, the weight-error power for 15 users is examined ($E_b/N_0 = 6\,\text{dB}$ and $\mu_0 = 0.02$). As we can see, the analytic results are still accurate even for the erroneous decision of the 15-user case. We also give the BER comparison for the second stage output using (88) with different user numbers in Fig. 9. From the figure, we observe that the analytical and empirical results are quite close from low to moderate E_b/N_0 values.

7. Conclusions

In DS-CDMA communication systems, MAI is considered as the main factor limiting the system performance. Among many multiuser detection schemes, the PIC receiver is considered as a simple yet effective approach. It has been shown that the performance of the PIC can be further improved if interference is not fully cancelled, resulting in the partial PIC approach. The performance of a partial PIC depends heavily on the PCFs. Thus, how to determine PCFs is of great concern. The Wiener solution is known to be optimal. However, it requires high computational complexity. Also, it is not efficient when the receiver is operated in time-varying environments. The adaptive multistage PIC scheme was developed to solve the problems. It is computational complexity is low, it is inherently applicable to time-varying environment, and its performance is superior to that of non-adaptive ones. However, its performance has not been analyzed before. In this paper, we conduct performance analysis for an adaptive two-stage PIC receiver in the AWGN channel. We first derive optimal weights, weight-error means, and weight-error variances for a single-user scenario. Using the similar idea, we extend the results to the two-user scenario. Finally, with some approximations, we further extend the results for the two-user scenario to a general multiple-user scenario. Based on the results obtained, we then derive the output MSE and the BER. Simulation results show that the derived analytic results are accurate even when the number of users is large.

References

[1] S. Verdu, Minimum probability of error for asynchronous Gaussian multiple-access channels, IEEE Trans. Inform. Theory IT-32 (January 1986) 85–96.
[2] S. Moshavi, Multi-user detection for DS-CDMA communications, IEEE Comm. Mag. 34 (10) (1996) 124–136.
[3] A. Duel-Hallen, J. Holtzman, Z. Zvonar, Multiuser detection for CDMA systems, IEEE Personal Comm. 2 (2) (1995) 46–58.
[4] S. Verdu, Multiuser Detection, Cambridge University Press, Cambridge, UK, 1998.
[5] A.J. Viterbi, Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels, IEEE J. Select. Areas Comm. 8 (4) (1990) 641–649.
[6] J.M. Holtzman. Successive interference cancellation for direct sequence code division multiple access. Military Conf., MILCOM '94, vol. 3, 1994, pp. 997–1001.
[7] P. Patel, J.M. Holtzman, Analysis of simple successive interference cancellation scheme in DS/CDMA system, IEEE J. Select. Areas Comm. 12 (5) (1994) 796–807.
[8] M.K. Varanasi, B. Aazhang, Multistage detection in asynchronous code-division multiple-access communications, IEEE Trans. Comm. 38 (4) (1990) 509–519.
[9] R. Kohno, H. Imai, M. Hatori, S. Pasupathy, An adaptive canceller of cochannel interference for spread-spectrum multiple-access communication networks in a power line, IEEE J. Select. Areas. Comm. 8 (4) (1990) 691–699.
[10] Y.C. Yoon, R. Kohno, H. Imai, A spread-spectrum multi-access system with cochannel interference cancellation, IEEE J. Select. Areas Comm. 11 (7) (1993) 1067–1075.
[11] G. Xue, J. Weng, T. Le-Ngoc, S. Tahar, Adaptive multistage parallel interference cancellation for CDMA, IEEE J. Select. Areas Comm. 17 (10) (1999) 1815–1827.
[12] Y.B. Guo, J.R. Cavallaro, A low complexity and low power SoC design architecture for adaptive MAI suppression in CDMA systems, J. VLSI Signal Process. 44 (3) (2006) 195–217.
[13] S.R. Kim, I. Choi, S. Kang, J.G. Lee, Adaptive weighted parallel interference cancellation for CDMA systems, Electron. Lett. 34 (22) (1998) 2085–2086.
[14] S.H. Han, J.H. Lee, Adaptive multi-stage parallel interference cancellation receiver for multi-rate DS-CDMA system, IEICE Trans. Comm. E86-B (1) (2004) 2401–2405.
[15] Y.-F. Huang, J.-H. Wen, S.T. Wu, An adaptive decision-feedback multiuser detector using parallel interference cancellation for CDMA systems, in: Proceedings of the IEEE VTC 2001 Spring, vol. 3, May 2001, pp. 1770–1774.
[16] J. Bae, I. Song, D.H. Won, A selective and adaptive interference cancellation scheme for code decision multiple access systems, Signal Processing 83 (2) (2003) 259–273.