Low-Complexity Adaptive Step Size Constrained Constant Modulus SG Algorithms for Adaptive Beamforming

Lei Wang, Rodrigo C. de Lamare, and Yun Long Cai

Communications Research Group, Department of Electronics, The University of York, York, YO10 5DD, UK

Abstract

This paper proposes two low-complexity adaptive step size mechanisms for adaptive beamforming that employs stochastic gradient (SG) algorithms. The algorithms employ the constrained constant modulus (CCM) criterion as the design approach. A complexity comparison is given and a convergence analysis is developed for illustrating their properties. Theoretical expressions of the excess mean squared error (EMSE), in both the steady-state and tracking phases, are derived for the proposed algorithms by employing the energy conservation approach. Simulation experiments are presented for both the stationary and nonstationary scenarios, illustrating that the new algorithms achieve superior performance compared with existing methods, and verifying the accuracy of the convergence and tracking analyses.

Key words–Adaptive beamforming techniques, spatial-division multiple access (SDMA), constrained constant modulus (CCM), modified adaptive step size (MASS), time averaging adaptive step size (TAASS).

I. Introduction

Adaptive beamforming technology is a strategic and widely investigated technique for rejecting interference and improving the performance in current and future communications systems [1]-[4], such as spatial-division multiple access (SDMA) systems [5]-[6]. By employing little information of the desired signal, e.g., direction of arrival (DOA), to form the array direction response towards the desired user, it deals with interference cancellation, tracking of dynamic systems, robustness issues, and complexity reduction. There have been numerous works in the literature on the performance of beamforming algorithms [7]-[9].

This work is supported in part by the Department of Electronics, the University of York. This paper was accepted in part at the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'08), Las Vegas, USA, March 30 - April 4, 2008 [19].

Email addresses: lw517@york.ac.uk, rcdl500@ohm.york.ac.uk, yc521@york.ac.uk.

DOA of the desired signal in this work can be located by DOA estimation methods, such as ESPRIT [10] or ML [11].

The adaptive stochastic gradient (SG) method, which is commonly employed in the adaptive beamforming area, is a well-known technique for solving optimization problems with different criteria, e.g., minimum mean squared error (MMSE) [12], minimum variance (MV) [13], or constant modulus (CM) [14]-[17]. The results in [13] prove that the MV criterion leads to a computational efficient solution identical to that obtained from the minimization of the mean squared error (MSE). The CM algorithm exploits the low modulus fluctuation exhibited by communications signals using constant modulus constellations to extract them from the array input. It is well known that the performance of the CM method is superior to that of the MV [18] (hence we only consider the CM design here). However, the CM method may converge to the local minima, and typical designs do not usually obtain the optimal solutions. Xu and Liu [20] developed a SG algorithm on the basis of the CM criterion subject to a certain set of constraints to circumvent the local minima problem. Unfortunately, the performance of these approaches is sensitive to the step size. The small value of the step size will lead to slow convergence rate, whereas a large one will lead to high misadjustment or even instability. For accelerating the convergence, recursive least squares (RLS) algorithms were introduced in [21] using the CM approach for the performance improvement. Nevertheless, the RLS-based beamformer cannot avoid complicated computations caused by the required covariance matrix inversion.

A comparison of SG algorithms, which represent simple and low-complexity solutions but are subject to slow convergence, with RLS methods, which possess fast convergence but high computational load, suggests that it is preferable to adopt SG beamformers due to complexity and cost issues. However, its performance (fixed step size (FSS)) is strongly dependent on the choice of the step size [22]. Actually, the communication systems are nonstationary environments, which make it difficult to predetermine the step size. The adaptive step size (ASS) mechanism [23], [24] was used for this aim. The requirement of an additional update equation for the gradient of the weight vector with respect to the step size, which increases the extra computational load, limits their applications.

This paper has two contributions, the first of which is the derivation of two SG beamforming algorithms with the CCM approach, using two novel adaptive step size mechanisms. The origins of these mechanisms can be traced back to the works of [25] and [26] where low-complexity adaptive step size mechanisms were introduced for LMS algorithms. In contrast to the existing works, the mechanisms here are designed for CCM algorithms. The additional number of operations of the proposed algorithms does not depend on the number of sensor elements. In addition, the results are presented for the stationary and nonstationary environments, different levels of interference, and subject to steering vector mismatch, highlighting that the new mechanisms achieve better performance and convergence behavior than those of previous ones and exhibit an increased robustness.

The second contribution is in the study of characteristics of the proposed algorithms, including convergence, the steady-state and tracking analyses. The condition of the step size for the convergence guarantee of the weight vector is illustrated first. The mean and mean-square values of the step size, under the steady-state condition, are calculated for the computation of the EMSE in the steady-state and tracking scenarios. The EMSE here considers the effects of the additive white Gaussian noise (AWGN) and multiple access interference (MAI) when multiple users are introduced in the system. Notice that the classical approaches to the steady-state and tracking performance evaluation cannot be employed here since it is necessary to determine the covariance matrix of the weight error vector, which becomes a burden for CCM algorithms due to their inherent nonlinear updates [27]-[29]. Instead, the analyses in this paper exploit the energy-preserving relation that holds not only just for CM algorithms but for a general class of adaptive filters [30], [31]. This relation is carried out by a feedback structure that is composed of a lossless forward block and a feedback path. Computer simulations are performed to confirm the accuracy of the analyses.

The remaining of this paper is organized as follows: In the next section, we outline a system model for smart antennas. Based on this model, the adaptive CCM beamformer design using the SG method is presented. Section III derives the proposed adaptive step size mechanisms. Section IV is dedicated to the convergence, steady-state and tracking analyses of the new algorithms. Simulation results are provided and discussed in Section V, and conclusions are drawn in Section VI.

II. System Model and Problem Statement

A. Receiver Models

In order to describe the system structure, we assume that the sources are in the "far field" to approximate the spherically propagating wave with a plane wave and the propagation medium is lossless and nondispersive to make the propagation speed uniform. Now, let us consider the adaptive beamforming scheme in Fig. 1 and suppose that q narrowband signals impinge on the uniform linear array (ULA) of m ($q \leq m$) sensor elements. Note that the receiver model can be extended to arbitrary antenna arrays and the ULA here is adopted for simplicity. The ith snapshot's vector of sensor array outputs $\boldsymbol{x}(i) \in \mathcal{C}^{m\times 1}$ can be modeled as [32]

$$\boldsymbol{x}(i) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(i) + \boldsymbol{n}(i), \quad i = 1, \ldots, N \tag{1}$$

where $\boldsymbol{\theta} = [\theta_0, \ldots, \theta_{q-1}]^T \in \mathcal{C}^{q\times 1}$ is the signal DOAs, $\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_0), \ldots, \boldsymbol{a}(\theta_{q-1})] \in \mathcal{C}^{m\times q}$ is the matrix composed of the signal direction vectors $\boldsymbol{a}(\theta_k) = [1, e^{-2\pi j \frac{d}{\lambda_c} cos\theta_k}, \ldots, e^{-2\pi j(m-1)\frac{d}{\lambda_c} cos\theta_k}]^T \in$

$\mathcal{C}^{m\times 1}$, $\quad (k = 0, \ldots, q-1)$, where λ_c is the wavelength and $d = \lambda_c/2$ is the inter-element distance of the ULA, $\boldsymbol{s}(i) \in \mathcal{R}^{q\times 1}$ is the source data with its entries are uncorrelated with each other. $\boldsymbol{n}(i) \in \mathcal{C}^{m\times 1}$ is the white sensor noise, which is assumed to be a zero-mean spatially and white Gaussian process, N is the number of snapshots, and $(\cdot)^T$ stands for the transpose. The output of a narrowband beamformer is given by

$$y(i) = \boldsymbol{w}^H(i)\boldsymbol{x}(i) \tag{2}$$

where $\boldsymbol{w}(i) = [w_1(i), \ldots, w_m(i)]^T \in \mathcal{C}^{m\times 1}$ is the complex weight vector, and $(\cdot)^H$ stands for the Hermitian transpose.

B. Adaptive CCM Design Using the SG algorithm

The CCM criterion is a constrained derivation of the CM algorithm for minimizing the cost function

$$J_{CM} = E[(|y(i)|^p - R_p)^2], \quad i = 1, \ldots, N \tag{3}$$

where the constant R_p is suitably chosen in order to guarantee that the optimal weight solution is close to the global minimum (see, e.g., [14]) and $p \geq 1$ is an integer. The cost function, as it stands, is not amenable to an optimal solution with respect to the desired signal since it is a high order cost function if $p \geq 2$, i.e., multiple local minima occurs. A constrained condition is added to handle this issue, which is stated as

$$J_{CM} = E[(|y(i)|^2 - 1)^2], \quad i = 1, \ldots, N \qquad \text{subject to} \quad \boldsymbol{w}^H(i)\boldsymbol{a}(\theta_0) = 1 \tag{4}$$

where $\boldsymbol{a}(\theta_0)$ denotes the steering vector of the desired signal and is known at the receiver by employing the DOA algorithms. For mathematical convenience, $p = 2$ is selected [14] and we consider the cost function as the expected deviation of the squared modulus of the array output to a constant, say $\delta_p = 1$. The constrained optimization means that the technique minimizes the contribution of undesired interference while maintaining the gain along the look direction to be constant.

With respect to the SG algorithm, the beamformer optimizes the Lagrangian cost function described by

$$L_{CCM} = E[(|y(i)|^2 - 1)^2] + \lambda(\boldsymbol{w}^H(i)\boldsymbol{a}(\theta_0) - 1) \tag{5}$$

where λ is a scalar Lagrange multiplier. The solution can be obtained by setting the gradient terms of (5) with respect to $\boldsymbol{w}(i)$ equal to zero and using the constraint [13]. Thus, the weight estimate is given by

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) - \mu(i)(|y(i)|^2 - 1)y^*(i)[\boldsymbol{I} - \boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0)]\boldsymbol{x}(i) \tag{6}$$

where $\mu(i)$ is the step size, which is a constant for FSS and a variable value for ASS and $(\cdot)^*$ denotes complex conjugate. The coefficient in the second term on the right side is absorbed by $\mu(i)$ since it is a very small value. The symbol $\boldsymbol{I}$ denotes the identity matrix of appropriate dimensions. Note that the term $\boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0)$ is normalized. Regarding ASS [23], [24], another step size recursion, for the sake of following the change of the surrounding, need to be included by taking the gradient (6) in terms of $\mu(i)$ equal to zero and then using again the SG recursion, which requires an additional number of operations proportional to the number of sensor elements m. The difficulties encountered in both FSS and ASS algorithms motivate the development in this paper of new algorithms to exhibit favorable performance of the SG method.

III. Proposed Adaptive Step Size Mechanisms

In this section, two novel adaptive step size methods are described for adjusting the step size in order to track the change of the communication system. The additional computational complexity of these mechanisms is shown in this section as well.

A. *Modified Adaptive Step Size (MASS) Mechanism*

The first proposed algorithm based on the MASS mechanism employs the prediction error and uses the update rule

$$\mu(i+1) = \alpha\mu(i) + \gamma(|y(i)|^2 - 1)^2 \tag{7}$$

where $0 < \alpha < 1$, $\gamma > 0$ and $y(i)$ is the same as that in (2). The rationale for the MASS is that at early stage of adaptation, the error is large due to the second term on the right side of (7), causing the step size to increase, thus providing faster convergence rate. When the error decreases, the step size follows this trend, yielding small misadjustment around the optimum. The parameter γ is an independent variable for controlling the prediction error and scaling it at different levels. It is worth pointing out that the step size $\mu(i+1)$ should be restricted in a range as follows

$$\mu(i+1) = \begin{cases} \mu_{max} & \text{if } \mu(i+1) > \mu_{max} \\ \mu_{min} & \text{if } \mu(i+1) < \mu_{min} \\ \mu(i+1) & \text{otherwise} \end{cases} \tag{8}$$

where $0 < \mu_{min} < \mu_{max}$. The constant μ_{min} is chosen as a compromise between the satisfying level of steady-state misadjustment and the required minimum level of tracking ability while μ_{max} is normally selected close to the point of instability of the algorithm for providing the maximum convergence speed. The MASS is the result of several attempts to devise a simple and yet effective mechanism.

B. Time Averaging Adaptive Step Size (TAASS) Mechanism

The second mechanism, which is called TAASS, uses a time average estimate of the correlation of $(|y(i)|^2 - 1)$ and $(|y(i-1)|^2 - 1)$. The update rule is

$$\mu(i+1) = \alpha\mu(i) + \gamma v^2(i) \tag{9}$$

where $v(i) = \beta v(i-1) + (1-\beta)(|y(i)|^2 - 1)(|y(i-1)|^2 - 1)$ and $0 < \beta < 1$. The limits on $\mu(i+1)$, α and γ are similar to those of the MASS algorithm. The exponential weighting parameter β governs the averaging time constant, namely, the quality of the estimation. The quality $v(i)$ in (9) contains information that is useful to determine an accurate measure for the proximity of the adaptive beamformer coefficients to the optimal value. Here, β should be close to 1. For nonstationary environments, the time averaging window should be small for forgetting past data and leaving space for the current statistics adaptation, so, $\beta < 1$.

The difference between (7) and (9) is that in the presence of the instantaneous error energy, MASS does not perform as well as expected in the presence of measurement noise. A zero-mean independent disturbance associated with the error energy can not be deleted in the expression of the expectation of the step size (see [26]). As a result, the variation of the step size cannot reflect accurately the adaptation state before or after convergence. Besides, close to the optimum, this disturbance results in large fluctuations of the solution around this optimum. TAASS operates in a large $\mu(i)$ when the algorithm works far from the optimum with $\mu(i)$ reducing as it approaches the optimum even under this instantaneous error condition. The term $v^2(i)$ in (9) is introduced to achieve this task. There are two objectives for using $v(i)$ here. First, it rejects the effect of the uncorrelated noise sequence on the step-size update [26]. In the beginning, because of scarcity of transmitters' information, the error correlation estimate $v^2(i)$ is large and so $\mu(i)$ is large to increase the convergence rate and track the change of input data. As the system approaches the steady-state, the recovered signal $y(i) \to 1$ and $v^2(i)$ is very small , resulting in a small step size for ensuring low misadjustment near optimum. Second, the error correlation is generally a good measure of the proximity to the optimum. It is clear from (9) that the update of $\mu(i)$ is not affected by the independent disturbance noise but dependent on how far from the optimum. The experiments later demonstrate this superiority.

TABLE I

ADDITIONAL COMPUTATIONAL COMPLEXITY OF ADAPTIVE STEP SIZE ALGORITHMS

Algorithms	Number of operations per snapshot	
	Additions	Multiplications
ASS	$5m-1$	$4m+3$
MASS	1	3
TAASS	2	6

C. Computational Complexity

It is well known that addition and multiplicity complexities of the CCM-SG method [17] are $3m$ and $3m+4$, respectively, and of the CCM-RLS method [21] are $2m^2+m$ and $3m^2+5m+5$, respectively. Also, the computational complexity of the ASS algorithm is a linear monotonic increasing function of the number of sensor elements in AWGN model (see Table I). Therefore, the computational complexity becomes very large if the array size is big.

An important feature of the proposed algorithms is that they only require a few fixed number of operations for updating the step size. The additional computational complexity of the proposed adaptive step size mechanisms is listed in Table I for comparison with the existent ASS method. We remark that the number of arithmetic operations is estimated by taking into account the number of complex additions and multiplications required in the mechanisms ("additional" means that the complexity is considered with the exception of the weight update equation in each method).

IV. ANALYSIS OF THE PROPOSED ALGORITHMS

The characteristics of the proposed algorithms are investigated in this section. Firstly, a necessary condition for the stability of the developed algorithms is studied. Then, the derivations of the mean and the mean-square values of the step size in the steady-state scenario are given for the convergence analysis. On the basis of the optimum step size values, the steady-state EMSE is calculated by using the energy conservation relation originally developed in [33]. Finally, in a similar way, the tracking analysis of the new methods is represented to study their ability to track time variations under nonstationary environment. It is worth noting that for the analyses, several assumptions are taken to simplify them.

A. Convergence Analysis

1) Condition for Stability: In view of (6), the stability of the SG algorithm is determined by two factors, namely, the step size parameter $\mu(i)$ and the input vector $\boldsymbol{x}(i)$. For further analysis, we bring forth one assumption.

Assumption 1: we assume that $\mu(i)$ varies slowly around its mean value.

This assumption is approximately true if γ is small and α close to one, which will be shown in the simulations. Under this condition, we have

$$E\{\mu(i)[|y(i)|^2-1]y^*(i)[\boldsymbol{I}-\boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0)]\boldsymbol{x}(i)\} = E[\mu(i)]E\{[|y(i)|^2-1]y^*(i)[\boldsymbol{I}-\boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0)]\boldsymbol{x}(i)\} \tag{10}$$

and

$$E\{\mu(i)[|y(i)|^2-1]\boldsymbol{x}(i)\boldsymbol{x}^H(i)\}\boldsymbol{w}(i) = E[\mu(i)]\boldsymbol{R}_{CCM}\boldsymbol{w}(i) \tag{11}$$

where $\boldsymbol{R}_{CCM} = E\{[|y(i)|^2-1]\boldsymbol{x}(i)\boldsymbol{x}^H(i)\} \in \mathcal{C}^{m\times m}$.

Now, (6) can be written as

$$\boldsymbol{w}(i+1) = \{\boldsymbol{I}-\mu(i)[|y(i)|^2-1]\boldsymbol{u}(i)\boldsymbol{x}^H(i)\}\boldsymbol{w}(i) \tag{12}$$

where $\boldsymbol{u}(i) = [\boldsymbol{I}-\boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0)]\boldsymbol{x}(i) \in \mathcal{C}^{m\times 1}$.

By defining the weight error vector $\tilde{\boldsymbol{w}}(i)$ and substituting (12) into the expression, we get

$$\begin{aligned}\tilde{\boldsymbol{w}}(i+1) &= \boldsymbol{w}_{opt} - \boldsymbol{w}(i+1) \\ &= \{\boldsymbol{I}-\mu(i)[|y(i)|^2-1]\boldsymbol{u}(i)\boldsymbol{x}^H(i)\}\tilde{\boldsymbol{w}}(i) + \mu(i)[|y(i)|^2-1]\boldsymbol{u}(i)\boldsymbol{x}^H(i)\boldsymbol{w}_{opt}\end{aligned} \tag{13}$$

where $\boldsymbol{w}_{opt}$ denotes the weight vector optimum solution.

By employing *Assumption 1* and taking expectations on both sides of (13), we have

$$E[\tilde{\boldsymbol{w}}(i+1)] = \{\boldsymbol{I}-E[\mu(i)]\boldsymbol{R}_{ux}(i)\}E[\tilde{\boldsymbol{w}}(i)] \tag{14}$$

where $\boldsymbol{R}_{ux}(i) = E\{[|y(i)|^2-1]\boldsymbol{u}(i)\boldsymbol{x}^H(i)\} = [\boldsymbol{I}-\boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0)]\boldsymbol{R}_{CCM}$ and $\boldsymbol{R}_{ux}\boldsymbol{w}_{opt} = \boldsymbol{0}$ [34]. Therefore, $E[\boldsymbol{w}(i)] \rightarrow \boldsymbol{w}_{opt}$ or equivalently, $\lim_{i\rightarrow\infty} E[\tilde{\boldsymbol{w}}(i)] = \boldsymbol{0}$ represents the stable condition if and only if $\prod_{i=0}^{\infty}\{\boldsymbol{I}-E[\mu(i)]\boldsymbol{R}_{ux}\} \rightarrow 0$. Following the idea of the eigenstructure [22] with respect to the covariance matrix $\boldsymbol{R}_{ux}$, the sufficient condition for (14) to hold implies that

$$0 \leq E[\mu(\infty)] \leq \frac{2}{\lambda_{max}^{ux}} \tag{15}$$

where λ_{max}^{ux} is the maximum eigenvalue of $\boldsymbol{R}_{ux}$.

Obviously, the stability condition in (15) is applicable to both MASS and TAASS algorithms.

2) Steady-State Step Size Value for MASS: We consider $E[\mu(i)]$ and $E[\mu^2(i)]$ to be the first-order and second-order expectations of the step size of the proposed MASS mechanism. By taking expectations on both sides of (7), under *Assumption 1*,

$$E[\mu(i+1)] = \alpha E[\mu(i)] + \gamma E[(|y(i)|^2 - 1)^2] \tag{16}$$

Also, by making the square of (7) and taking expectations,

$$\begin{aligned} E[\mu^2(i+1)] &= \alpha^2 E[\mu^2(i)] + 2\alpha\gamma E[\mu(i)]E[(|y(i)|^2 - 1)^2] + \gamma^2 E[(|y(i)|^2 - 1)^4] \\ &\approx \alpha^2 E[\mu^2(i)] + 2\alpha\gamma E[\mu(i)]E[(|y(i)|^2 - 1)^2] \end{aligned} \tag{17}$$

where the third term on the right side of the first expression of (17) is negligible as compared to the other terms in the second approximation since γ^2 and $E[(|y(i)|^2 - 1)^4]$ are small in the steady-state condition.

The relations $\lim_{i\to\infty} E[\mu(i)] = \lim_{i\to\infty} E[\mu(i+1)] = E[\mu(\infty)]$ and $\lim_{i\to\infty} E[\mu^2(i)] = \lim_{i\to\infty} E[\mu^2(i+1)] = E[\mu^2(\infty)]$ hold if we consider the steady-state values. Before proceeding to the steady-state step size, one approximation should be given.

Approximation 1: $\lim_{i\to\infty} E[(|y(i)|^2 - 1)^2] = \xi_{min} + \xi_{ex}(\infty)$
where $\xi_{min} = |d(i)|^2 - \boldsymbol{p}^H \boldsymbol{R}^{-1} \boldsymbol{p}$ [22] is the minimum mean square error (MMSE), $\boldsymbol{p} = E[\boldsymbol{x}(i)d^*(i)]$ denotes the cross correlation between the input vector $\boldsymbol{x}(i)$ and the desired response $d(i)$, and $\xi_{ex}(\infty)$ is the excess mean square error (EMSE) associated with the CM cost function. Here, $\boldsymbol{R} = E[\boldsymbol{x}(i)\boldsymbol{x}^H(i)]$ represents the covariance matrix of the input vector.

This approximation is reasonable under the assumption that the CM algorithm converges to the optimum solution that is close to that of the actual MMSE design. The theorem stating that the optimal CM algorithm (CMA) minima roughly correspond to the MMSE minima was conjectured in [14] and then was thoroughly scrutinized by Zeng *et al.* in [35]. Since the MMSE criterion is achieved by minimizing the MSE of the estimation error, which is defined by the difference between $d(i)$ and $y(i)$, after rearrangement, the cost function of the MMSE design (see, e.g., [22]) becomes

$$J(i) = J_{min} + J_{ex}(i) \tag{18}$$

where J_{min} and $J_{ex}(i)$ are the MMSE and the EMSE correlated with the MMSE criterion. This methodology is similar to that of the CMA, which employs the constant modulus property to subtract it from the array output. Following this idea, the approximation is built up as $i \to \infty$.

Using *Approximation 1* to (16) and (17), we have

$$E[\mu(\infty)] = \frac{\gamma[\xi_{min} + \xi_{ex}(\infty)]}{1 - \alpha} \tag{19}$$

$$E[\mu^2(\infty)] = \frac{2\alpha\gamma^2[\xi_{min} + \xi_{ex}(\infty)]^2}{(1-\alpha)^2(1+\alpha)} \tag{20}$$

Further simplified formulations of (19) and (20) are more efficient if the inconvenient term $\xi_{ex}(\infty)$ can be ignored. In doing so, we consider another assumption.

Assumption 2: $(\xi_{min} + \xi_{ex}(\infty)) \approx \xi_{min}$ and $(\xi_{min} + \xi_{ex}(\infty))^2 \approx \xi_{min}^2$ if $\xi_{min} \gg \xi_{ex}(\infty)$.

We claim that this assumption is true if the SG algorithm operates in the steady-state, which is exactly the same condition we are working now. Therefore, it leads to

$$E[\mu(\infty)] \approx \frac{\gamma\xi_{min}}{1-\alpha} \tag{21}$$

$$E[\mu^2(\infty)] \approx \frac{2\alpha\gamma^2\xi_{min}^2}{(1-\alpha)^2(1+\alpha)} \tag{22}$$

Note that (21) and (22) can be used for computing the EMSE of the proposed algorithm. The following analysis will show their application.

3) Steady-State Step Size Value for TAASS: The same method above can be employed to derive the step size value for TAASS with a little more complicated procedure. Using *Assumption 1* for (9) and setting expectations, we get

$$E[\mu(i+1)] = \alpha E[\mu(i)] + \gamma E[v^2(i)] \tag{23}$$

Also, we take the square and then expectations on both sides of (9) to obtain

$$\begin{aligned} E[\mu^2(i+1)] &= \alpha^2 E[\mu^2(i)] + 2\alpha\gamma E[\mu(i)]E[v^2(i)] + \gamma^2 E[v^4(i)] \\ &\approx \alpha^2 E[\mu^2(i)] + 2\alpha\gamma E[\mu(i)]E[v^2(i)] \end{aligned} \tag{24}$$

where both γ^2 and $E[v^4(i)]$ of the third term are small in the steady-state compared with the second on the right side.

From the definition of $v(i)$ in (9), an alternative way of $v(i)$ is written as

$$v(i) = (1-\beta)\sum_{n=0}^{i-1}\beta^n[|\boldsymbol{w}^H(i-n)\boldsymbol{x}(i-n)|^2 - 1][|\boldsymbol{w}^H(i-n-1)\boldsymbol{x}(i-n-1)|^2 - 1] \tag{25}$$

and

$$\begin{aligned} v^2(i) = (1-\beta)^2 \sum_{n=0}^{i-1}\sum_{j=0}^{i-1}\beta^n\beta^j[|\boldsymbol{w}^H(i-n)\boldsymbol{x}(i-n)|^2 - 1][|\boldsymbol{w}^H(i-n-1)\boldsymbol{x}(i-n-1)|^2 - 1] \\ \cdot [|\boldsymbol{w}^H(i-j)\boldsymbol{x}(i-j)|^2 - 1][|\boldsymbol{w}^H(i-j-1)\boldsymbol{x}(i-j-1)|^2 - 1] \end{aligned} \tag{26}$$

We assume that the algorithm has converged in respect that the analysis is under the steady-state condition. At this point, the term $\{|\boldsymbol{w}^H(i)\boldsymbol{x}(i)|^2-1,\ i=1,\dots,N\}$ is treated as uncorrelated, i.e., $E[(|\boldsymbol{w}^H(i-n)\boldsymbol{x}(i-n)|^2-1)(|\boldsymbol{w}^H(i-j)\boldsymbol{x}(i-j)|^2-1)]=0\ \forall n\neq j$. Therefore, expectations of (26) can be simplified as

$$E[v^2(i)]=(1-\beta)^2\sum_{n=0}^{i-1}\beta^{2n}E[(|\boldsymbol{w}^H(i-n)\boldsymbol{x}(i-n)|^2-1)^2]E[(|\boldsymbol{w}^H(i-n-1)\boldsymbol{x}(i-n-1)|^2-1)^2] \quad (27)$$

Taking into account the relation $\lim_{i\rightarrow\infty}E[v^2(i)]=E[v^2(\infty)]$ and invoking *Approximation 1* and *Assumption 2* for (27) brings on

$$\begin{aligned} E[v^2(i)] &= \frac{(1-\beta)(\xi_{min}+\xi_{ex}(\infty))^2}{1+\beta} \\ &\approx \frac{(1-\beta)\xi_{min}^2}{1+\beta} \end{aligned} \quad (28)$$

where the elaborate derivation is given in Appendix.

In the steady-state environment, considering again the relations $\lim_{i\rightarrow\infty}E[\mu(i)]=\lim_{i\rightarrow\infty}E[\mu(i+1)]=E[\mu(\infty)]$ and $\lim_{i\rightarrow\infty}E[\mu^2(i)]=\lim_{i\rightarrow\infty}E[\mu^2(i+1)]=E[\mu^2(\infty)]$, substituting (28) into (23) and (24), respectively, we have the following

$$E[\mu(\infty)]\approx\frac{\gamma(1-\beta)(\xi_{min}+\xi_{ex}(\infty))^2}{(1-\alpha)(1+\beta)} \quad (29)$$

$$E[\mu^2(\infty)]\approx\frac{2\alpha\gamma^2(1-\beta)^2(\xi_{min}+\xi_{ex}(\infty))^4}{(1+\alpha)(1-\alpha)^2(1+\beta)^2} \quad (30)$$

Assumption 2 can be employed for (29) and extended to $(\xi_{min}+\xi_{ex}(\infty))^4\approx\xi_{min}^4$ in the steady-state for (30)

$$E[\mu(\infty)]\approx\frac{\gamma(1-\beta)\xi_{min}^2}{(1-\alpha)(1+\beta)} \quad (31)$$

$$E[\mu^2(\infty)]\approx\frac{2\alpha\gamma^2(1-\beta)^2\xi_{min}^4}{(1+\alpha)(1-\alpha)^2(1+\beta)^2} \quad (32)$$

It is observed that the first-order and second-order step size values associated with the TAASS approach are more complicated than those of the MASS method due to the presence of $v^2(i)$ in (9). Note that (31) and (32) can be used for computing the EMSE of the proposed TAASS algorithm.

B. Steady-State Analysis

The customary techniques are not suitable to figure out the proposed algorithms here due to their nonlinear update [28], [29], presenting difficulties for determining the covariance matrix of the weight error vector. Instead, we adopt an energy flow framework [27], [30], [33] to bypass the difficulties incurred in obtaining the steady-state result. The approach creates an energy-preserving connection, relying on a fundamental error variance relation, between adjacent iterations, to avoid the weight error variance. Furthermore, this approach allows for an unified treatment of a large class of algorithms [31].

Consider the noisy measurements

$$d(i) = \boldsymbol{w}_{opt}^H \boldsymbol{x}(i) + n(i) \tag{33}$$

where $\boldsymbol{w}_{opt}$, as mentioned before, is the unknown optimum weight vector and $n(i)$ accounts for measurement noise, which is stochastic in nature.

In the classic MSE measure,

$$MSE = \lim_{i\to\infty} E[|e(i)|^2] = \lim_{i\to\infty} E[|\tilde{\boldsymbol{w}}^H(i)\boldsymbol{x}(i) + n(i)|^2] \tag{34}$$

where $e(i) = d(i) - \boldsymbol{w}^H(i)\boldsymbol{x}(i)$ denotes the output estimation error and $\tilde{\boldsymbol{w}}(i) = \boldsymbol{w}_{opt} - \boldsymbol{w}(i)$ is the weight error vector.

Under a realistic assumption [36],

Assumption 3: The noise term $n(i)$ is independent and identically distributed (i.i.d.) and statically independent of the input vector $\boldsymbol{x}(i)$. The MSE is equivalent to

$$MSE = \sigma_n^2 + \lim_{i\to\infty} E[|\tilde{\boldsymbol{w}}^H(i)\boldsymbol{x}(i)|^2] \tag{35}$$

where σ_n^2 represents the variance of the measurement noise.

Now, we associate the following so-called a *priori* and a *posteriori* estimation errors

$$e_a(i) = \tilde{\boldsymbol{w}}^H(i)\boldsymbol{x}(i), e_p(i) = \tilde{\boldsymbol{w}}^H(i+1)\boldsymbol{x}(i). \tag{36}$$

Thus, MSE in (35) can be expressed by

$$MSE = \sigma_n^2 + \zeta_s \tag{37}$$

where $\zeta_s = \lim_{i\to\infty} E[|e_a(i)|^2]$ denotes the EMSE corresponding to the MSE measure in the steady-state. It is straightforward that calculating ζ_s is equivalent to finding the MSE.

Recall from (6) for compactness of notation

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) + \mu(i)\boldsymbol{u}(i)F_e(i) \tag{38}$$

where $z(i) = |\boldsymbol{w}^H(i)\boldsymbol{x}(i)|^2 - 1$ and $F_e(i) = -y^*(i)z(i)$.

By subtracting both sides of (38) from $\boldsymbol{w}_{opt}$, taking transpose, and multiplying by $\boldsymbol{x}(i)$ from the right, we find

$$e_p(i) = e_a(i) - \mu(i)F_e^*(i)\boldsymbol{u}^H(i)\boldsymbol{x}(i) \tag{39}$$

We notice that $\boldsymbol{u}^H(i)\boldsymbol{x}(i) = \boldsymbol{x}^H(i)(\boldsymbol{I} - \boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0))\boldsymbol{x}(i) = \boldsymbol{u}^H(i)\boldsymbol{u}(i)$. So, an alternative way of (39) is

$$e_p(i) = e_a(i) - \mu(i)F_e^*(i)\|\boldsymbol{u}(i)\|^2 \tag{40}$$

Substituting (40) into the weight vector error update equation from (38) and rearranging the outcome, we have

$$\tilde{\boldsymbol{w}}(i+1) = \tilde{\boldsymbol{w}}(i) - \frac{\boldsymbol{u}(i)}{\|\boldsymbol{u}(i)\|^2}[e_a^*(i) - e_p^*(i)] \tag{41}$$

Recombining (41) and squaring it, we obtain the energy evaluation

$$\|\tilde{\boldsymbol{w}}(i+1)\|^2 + \bar{\mu}(i)|e_a(i)|^2 = \|\tilde{\boldsymbol{w}}(i)\|^2 + \bar{\mu}(i)|e_p(i)|^2 \tag{42}$$

where $\bar{\mu}(i) = 1/\|\boldsymbol{u}(i)\|^2$.

This is an exact energy conservation relation that illustrates the energies of the weight error vectors between two successive time instants and which correspond to the energies of the a *priori* and a *posteriori* estimation errors. It is achieved by a closed loop configuration that is split into a feedforward path and a feedback path (see, [27]). Note that (42) is obtained without any assumptions.

In the steady-state condition (i.e., as $i \to \infty$), we can assume that $E[\|\tilde{\boldsymbol{w}}(i+1)\|^2] = E[\|\tilde{\boldsymbol{w}}(i)\|^2]$. Now, by taking expectations on both sides of (42), cancelling the effect of the weight error vector and using (40), we get

$$E[\bar{\mu}(i)|e_a(i)|^2] = E[\bar{\mu}(i)|e_a(i) - \frac{\mu(i)}{\bar{\mu}(i)}F_e^*(i)|^2] \tag{43}$$

Substituting $F_e(i) = [1 - |y(i)|^2]y^*(i)$ into (43), we have

$$E\{\mu(i)e_a^*(i)y(i)[1-|y(i)|^2]\}+E\{\mu(i)e_a(i)y^*(i)[1-|y(i)|^2]\}=E\{\frac{\mu^2(i)}{\bar{\mu}(i)}|y(i)|^2[1-|y(i)|^2]^2\} \quad (44)$$

The next step is to consider the MAI and AWGN in the system model, which will affect the array output. Therefore, the output estimation can be defined as

$$\begin{aligned} y(i) &= [\boldsymbol{w}_{opt}-\tilde{\boldsymbol{w}}(i)]^H\boldsymbol{x}(i) \\ &= \boldsymbol{w}_{opt}^H\boldsymbol{x}(i)-e_a(i) \\ &= s_0(i)+M(i)+n(i)-e_a(i) \end{aligned} \quad (45)$$

where $s_0(i)$ denotes the transmitted symbol of the desired source at time instant i, $M(i)=\sum_{k=1}^{q-1}\boldsymbol{w}_{opt}^H s_k(i)\boldsymbol{a}(\theta_k)$ is the output residual MAI caused as multiple users appear in the system, where $s_k(i)$ denotes the transmitted symbol of users with the exception of the desired one, and $n(i)=\boldsymbol{w}_{opt}^H\boldsymbol{n}(i)$ is the processed AWGN.

Before making use of (45), we provide the properties [37], approximation and assumptions for this definition. We write more compactly by dropping time index i and using s to present $s_0(i)$.

Properties:$\{s,M,n,e_a\}$ are zero-mean random variables, and $\{s,M,n\}$ are mutually independent. The random sources are independent and the processed noise n is Gaussian random variable.

Approximation 2: The residual MAI M is Gaussian and holds well [38].

Assumption 4: $\{s,M,n,e_a\}$ are mutually independent [37], which is realistic under *Properties* condition since the blind algorithms works independently of the transmitted signals.

Now, in a concise way, we make a *LP* and a *RP* terms to represent the left part and the right part of (44), respectively.

$$\begin{aligned} LP &= E[\mu e_a^* y(1-|y|^2)]+E[\mu e_a y^*(1-|y|^2)] \\ RP &= E[\frac{\mu^2}{\bar{\mu}}|y|^2(1-|y|^2)^2] \end{aligned} \quad (46)$$

Using (45) into (46) and employing *Approximation 2* and *Assumption 1, 4*, (46) becomes

$$\begin{aligned} LP &= 2E[\mu|e_a|^2(|s|^2+|M|^2+|n|^2+|e_a|^2-1)] \\ RP &= E[\frac{\mu^2}{\bar{\mu}}(J_1+J_2+J_3+J_4|e_a|^2)] \end{aligned} \quad (47)$$

where

$J_1=|s|^6+3|s|^4|M|^2+3|s|^4|n|^2-2|s|^4+|s|^2+3|s|^2|M|^4+3|s|^2|n|^4-4|s|^2|M|^2-4|s|^2|n|^2+6|s|^2|M|^2|n|^2$;

$J_2=|M|^6+|n|^6-2|M|^4-2|n|^4+|M|^2+|n|^2+3|M|^4|n|^2+3|M|^2|n|^4-4|M|^2|n|^2$;

$J_3 = |e_a|^6 - 2|e_a|^4 + 3|M|^2|e_a|^4 + 3|s|^2|e_a|^4 + 3|n|^2|e_a|^4$;

$J_4 = 6|s|^2|n|^2 + 6|M|^2|n|^2 + 6|s|^2|M|^2 - 4|M|^2 - 4|n|^2 + 3|s|^4 - 4|s|^2 + 3|n|^4 + 3|M|^4 + 1$.

Define $\sigma_0^2 = E[|s|^2]$, $\sigma_n^2 = E[|n|^2]$, and $\sigma_M^2 = E[|M|^2]$ as the desired signal power, the processed noise variance, and the residual MAI variance, respectively. In the steady-state scenario, we have $E[|s|^4] = (E[|s|^2])^2 = \sigma_0^4$ and $E[|s|^6] = (E[|s|^2])^3 = \sigma_0^6$. Similarly, $E[|n|^4] = \sigma_n^4$ and $E[|n|^6] = \sigma_n^6$; $E[|M|^4] = \sigma_M^4$ and $E[|M|^6] = \sigma_M^6$.

It is reasonable to assume that the higher order of $E[|e_a|^2]$ can be neglected since it is relatively small if $i \to \infty$. Under this condition, J_3 in (47) can be cancelled. Thus, we obtain

$$\begin{aligned} LP &\approx 2K_1 E[\mu] E[|e_a|^2] \\ RP &\approx \frac{E[\mu^2]}{E[\bar{\mu}]}(K_2 + K_3 + K_4 E[|e_a|^2]) \end{aligned} \tag{48}$$

where

$K_1 = \sigma_0^2 + \sigma_M^2 + \sigma_n^2 - 1$;

$K_2 = \sigma_0^6 + 3\sigma_0^4\sigma_M^2 + 3\sigma_0^4\sigma_n^2 - 2\sigma_0^4 + \sigma_0^2 + 3\sigma_0^2\sigma_M^4 + 3\sigma_0^2\sigma_n^4 - 4\sigma_0^2\sigma_M^2 - 4\sigma_0^2\sigma_n^2 + 6\sigma_0^2\sigma_M^2\sigma_n^2$;

$K_3 = \sigma_M^6 + \sigma_n^6 - 2\sigma_M^4 - 2\sigma_n^4 + \sigma_M^2 + \sigma_n^2 + 3\sigma_M^4\sigma_n^2 + 3\sigma_M^2\sigma_n^4 - 4\sigma_M^2\sigma_n^2$;

$K_4 = 6\sigma_0^2\sigma_n^2 + 6\sigma_M^2\sigma_n^2 + 6\sigma_0^2\sigma_M^2 - 4\sigma_M^2 - 4\sigma_n^2 - 4\sigma_0^2 + 3\sigma_0^4 + 3\sigma_n^4 + 3\sigma_M^4 + 1$.

Remind $LP = RP$ from (44) and rearrange (48), we get

$$\zeta_s \approx \frac{\frac{E[\mu^2]}{E[\bar{\mu}]}(K_2 + K_3)}{2E[\mu]K_1 - \frac{E[\mu^2]}{E[\bar{\mu}]}K_4} = \frac{E[\mu^2]E[\|\boldsymbol{u}\|^2](K_2 + K_3)}{2E[\mu]K_1 - E[\mu^2]E[\|\boldsymbol{u}\|^2]K_4} \tag{49}$$

where $E[\|\boldsymbol{u}\|^2] = \sum_{k=1}^{q-1} \sigma_k^2 \{\boldsymbol{a}^H(\theta_k)\boldsymbol{a}(\theta_k) - [\boldsymbol{a}^H(\theta_k)\boldsymbol{a}(\theta_0)\boldsymbol{a}^H(\theta_0)\boldsymbol{a}(\theta_k)/(\boldsymbol{a}^H(\theta_0)\boldsymbol{a}(\theta_0))]\} + \sigma^2(m-1)$ [37], where σ_k^2 denotes the kth user power with the exception of the desired one, and the noise term at the output is a Gaussian random variable of type $n \sim \mathcal{N}(0, \sigma_n^2)$, where $\sigma_n = \|\boldsymbol{w}_{opt}\|\sigma$. The term $\sigma_M = \sqrt{q-1}\boldsymbol{w}_{opt}^H \sigma_k \boldsymbol{a}(\theta_k)$ for $k \neq 0$. It is worth noting that all the terms in (49) are relevant to the steady-state $i \to \infty$. Substituting (21) and (22) into (49) corresponds to the MASS algorithm, whereas (31) and (32) corresponds to the TAASS algorithm.

The steady-state analysis of (49) can be further simplified if we impose that $\sigma_0^2 = 1$. In most cases, the residual MAI power σ_M^2 at the output is significantly lower than the output noise power σ_n^2 [37], i.e., $\sigma_M^2 \ll \sigma_n^2$, and thus, an even simpler expression can be given by

$$\zeta_s \approx \frac{E[\mu^2]E[\|\boldsymbol{u}\|^2]K_5}{2\sigma_n^2 E[\mu] - E[\mu^2]E[\|\boldsymbol{u}\|^2]K_6} \tag{50}$$

where $K_5 = \sigma_n^4 + \sigma_n^6$ and $K_6 = 2\sigma_n^2 + 3\sigma_n^4$.

Hence, we can substitute (21) and (22) into (50) to calculate ζ_s for the MASS, which after arrangement is given by

$$\zeta_s^{MASS} \approx \frac{\alpha\gamma\xi_{min}E[\|\boldsymbol{u}\|^2]K_5}{\sigma_n^2(1-\alpha^2) - \alpha\gamma\xi_{min}E[\|\boldsymbol{u}\|^2]K_6} \tag{51}$$

or substitute (31) and (32) into (50) to calculate ζ_{taass} for the TAASS, which is given by

$$\zeta_s^{TAASS} \approx \frac{\alpha\gamma(1-\beta)\xi_{min}^2E[\|\boldsymbol{u}\|^2]K_5}{\sigma_n^2(1-\alpha^2)(1+\beta) - \alpha\gamma(1-\beta)\xi_{min}^2E[\|\boldsymbol{u}\|^2]K_6} \tag{52}$$

The accuracy of the approximation is analyzed via simulations later.

C. Tracking Analysis

The energy-preserving relation has been verified to provide the tracking analysis in a nonstationary environment [31]. The derivation in this section is specific for the proposed algorithms and to the best of the author's knowledge, did not exist in the literature. It is processed on the foundation of the steady-state result and introduces further assumptions.

In the nonstationary surrounding, the optimum weight coefficients are not constant but assumed to vary following the model $\boldsymbol{w}_{opt}(i+1) = \boldsymbol{w}_{opt}(i) + \boldsymbol{q}(i)$, where $\boldsymbol{q}(i)$ denotes a random perturbation [22], [36]. This perturbation is introduced by the time variations of the system in the nonstationary condition. This update formulation is invoked to track the variation. Therefore, the weight vector error recursion can be expressed

$$\tilde{\boldsymbol{w}}(i+1) = \tilde{\boldsymbol{w}}(i) + \boldsymbol{q}(i) - \mu(i)\boldsymbol{u}(i)F_e(i) \tag{53}$$

By squaring (53), we get the energy evaluation

$$\|\tilde{\boldsymbol{w}}(i+1)\|^2 + \bar{\mu}(i)|e_a(i)|^2 = \|\tilde{\boldsymbol{w}}(i) + \boldsymbol{q}(i)\|^2 + \bar{\mu}(i)|e_p(i)|^2 \tag{54}$$

The further analysis relies on the assumption following

Assumption 5: The sequence $\{\boldsymbol{q}(i)\}$ is a stationary sequence of independent zero-mean vectors with positive definite covariance matrix $\boldsymbol{Q} = E[\boldsymbol{q}(i)\boldsymbol{q}^H(i)]$. It is independent of the sequence $\{\boldsymbol{u}(i)\}$ and $\{n(i)\}$.

Under this assumption, using $E[\tilde{\boldsymbol{w}}^H(i)\boldsymbol{q}(i)] = 0$ [39] and $E[\|\tilde{\boldsymbol{w}}(i+1)\|^2] = E[\|\tilde{\boldsymbol{w}}(i)\|^2]$ if the system operates in the steady-state condition and taking expectations, the energy equation (54) can be written as

$$E[\bar{\mu}(i)|e_a(i)|^2] = Tr(\boldsymbol{Q}) + E[\bar{\mu}(i)|e_a(i) - \frac{\mu(i)}{\bar{\mu}(i)}F_e^*(i)|^2] \tag{55}$$

In view of (55), it can be regarded as an extension of (43) with an addition of the system nonstationary contribution $Tr(\boldsymbol{Q})$. This is a useful advantage of the energy conservation approach over many classic methods since it allows us to reach tracking analysis by analyzing the stationary case results.

In a similar way as that in the study of the steady-state case, substituting $F_e(i) = -y^*(i)z(i)$ and (45) into (55), the extension can be obtained in a compact way

$$\begin{aligned} LP &\approx 2K_1 E[\mu]E[|e_a|^2] \\ RP &\approx Tr(\boldsymbol{Q}) + \frac{E[\mu^2]}{E[\bar{\mu}]}(K_2 + K_3 + K_4 E[|e_a|^2]) \end{aligned} \tag{56}$$

and

$$\zeta_t \approx \frac{Tr(\boldsymbol{Q}) + E[\mu^2]E[\|\boldsymbol{u}\|^2](K_2 + K_3)}{2E[\mu]K_1 - E[\mu^2]E[\|\boldsymbol{u}\|^2]K_4} \tag{57}$$

where ζ_t denotes the EMSE corresponding to the tracking condition, the terms K_1, K_2, K_3 and K_4 are the same as those in (48) and it is assumed that $\boldsymbol{q}(i)$ is independent of $\{s, M, n, e_a\}$, and higher orders of $\{|e_a|^2\}$ are ignored. Note that (57) is applicable to the analyses for both MASS and TAASS by inserting different values of first-order and second-order step sizes.

If $\sigma_0^2 = 1$ and $\sigma_M^2 \ll \sigma_n^2$ are imposed on (56), we can have a simpler expression

$$\zeta_t \approx \frac{Tr(\boldsymbol{Q}) + E[\mu^2]E[\|\boldsymbol{u}\|^2]K_5}{2\sigma_n^2 E[\mu] - E[\mu^2]E[\|\boldsymbol{u}\|^2]K_6} \tag{58}$$

where K_5 and K_6 are the same as those in (50). It supports again that the tracking analysis can be received by inspection from the stationary results, which indicates the superiority of the energy-preserving configuration.

Substituting (21)-(22) and (31)-(32) into (58), respectively, we can express the tracking analysis ζ_t^{MASS} for the MASS and ζ_t^{TAASS} for the TAASS as

$$\zeta_t^{MASS} \approx \frac{(1-\alpha)^2(1+\alpha)Tr(\boldsymbol{Q}) + 2\alpha\gamma^2\xi_{min}^2 E[\|\boldsymbol{u}\|^2]K_5}{2\sigma_n^2(1-\alpha^2)\gamma\xi_{min} - 2\alpha\gamma^2\xi_{min}E[\|\boldsymbol{u}\|^2]K_6} \tag{59}$$

and

$$\zeta_t^{TAASS} \approx \frac{(1-\alpha)^2(1+\alpha)(1+\beta)^2 Tr(\boldsymbol{Q}) + 2\alpha\gamma^2(1-\beta)^2\xi_{min}^4 E[\|\boldsymbol{u}\|^2]K_5}{2\sigma_n^2(1-\alpha^2)(1-\beta^2)\gamma\xi_{min}^2 - 2\alpha\gamma^2(1-\beta)^2\xi_{min}^4 E[\|\boldsymbol{u}\|^2]K_6} \tag{60}$$

V. SIMULATION RESULTS

In this section, we demonstrate the effectiveness of the proposed algorithms over existing methods through simulations and verify the accuracy of the analyses of the MASS and TAASS mechanisms. The experiments are carried out under stationary and nonstationary scenarios to assess the performance. All simulations are performed by an ULA containing $m = 16$ sensor elements with half-wavelength spacing. The noise is spatially and temporally white Gaussian noise. For each scenario, $K = 1000$ iterations are used to get each simulated curve. In all experiments, the desired signal power is $\sigma_0^2 = 1$. The BPSK modulation scheme is employed to modulate the signals.

A. BER Performance

Fig. 2 compares the proposed MASS and TAASS algorithms with the FSS, AAS, and RLS methods by showing the BER versus the input SNR. 1000 snapshots are considered. There are five interferers in the system, one interferer with 4dB above the desired user's power level, one with the same power level of the desired one and three with power 0.5dB lower than that of the desired user. Note that the actual spatial signature of the desired source is known exactly. We set the first element of the initial weight vector $\boldsymbol{w}(0)$ equals to the corresponding element of the steering vector $\boldsymbol{a}(\theta_0)$ of SOI. Other parameters are optimized with $\alpha = 0.98$, $\gamma = 10^{-3}$, $\mu_0 = 10^{-5}$, $\mu_{max} = 10^{-4}$ and $\mu_{min} = 10^{-6}$ for MASS and $\alpha = 0.98$, $\beta = 0.99$,$\gamma = 10^{-3}$, $\mu_0 = 10^{-4}$, $\mu_{max} = 3 \times 10^{-4}$ and $\mu_{min} = 10^{-6}$ for TAASS. The parameters for the existing methods are tuned in order to optimize the performance, allowing for a fair comparison with the proposed algorithms. It is observed that under low input SNR condition, the BER behaviors of all methods are worse, compared with those under high SNR condition. However, MASS and TAASS exhibit better performance than existent algorithms, especially for TAASS, which converges rapidly and approaches the RLS as SNR increases.

B. Steady-State Performance

Fig. 3 and Fig. 4 compare the proposed MASS and TAASS steady-state analysis with simulation results, separately. Fig. 3 corresponds to the MASS algorithm and includes two experiments. Fig. 3(a) works with 3 interferers with 0.5dB lower than the desired user's power and Fig. 3(b) runs with 1 interferer 2dB above the power level of the desired one, 1 interferer with the same power level of the desired user and 2 interferers with 0.5dB below the desired power. Both of them work with the input SNR= 20dB. It is evident that the proposed algorithms reach the steady-state quickly in the first situation and the simulation result is very close to the theoretical expression, which is calculated by using (51), where $\boldsymbol{p} = E[\boldsymbol{x}(i)d^*(i)] = \boldsymbol{a}(\theta_0)$, $\boldsymbol{w}_{opt} = \boldsymbol{R}^{-1}\boldsymbol{p} = \boldsymbol{R}^{-1}\boldsymbol{a}(\theta_0)$, and $\xi_{min} = 1 - \boldsymbol{a}^H(\theta_0)\boldsymbol{R}^{-1}\boldsymbol{a}(\theta_0)$. The fact that the number or/and the power level of the interferers increase, deteriorates the performance of both simulation and theory. However, the experiment results are still in good match with the analytical values. The fluctuation of the curve is due to the system noise and the users' status (e.g., power level and the number). The ignored terms in (51) due to the higher order effect do not affect the accuracy of the results.

The same conclusion can be drawn in the view of Fig. 4, which works in the same scenarios as that in Fig. 3 by using (52). Compared with Fig. 3, the steady-state performance in both experiments are better than those of MASS since it is not affected by the independent disturbance.

Fig. 5 demonstrates the accuracy of the expression for MASS analysis by employing (51), compared with TAASS analysis by using (52) instead. In this figure, the EMSE is expressed as a function of SNR for a 3 interferers system with lower power level of interferers (all 0.5dB lower than that of the desired power) and SNR$=$ 20dB. The step size is truncated between $\mu_{min} = 1 \times 10^{-5}$ and $\mu_{max} = 1 \times 10^{-3}$ for MASS and $\mu_{min} = 5 \times 10^{-5}$ and $\mu_{max} = 1 \times 10^{-3}$ for TAASS. The experiment confirms the steady-state analysis for the proposed algorithms that the EMSE decreases monotonically with SNR. A reasonable explanation is that as the SNR increases the system works in a favorable state to direct the main beam towards the desired user for eliminating the effects of interferers and making the processed noise approach zero. As this occurs, the cost function approaches zero at convergence and so, the steady-state EMSE approaches zero.

Clearly, the simulation accords greatly with the theory as shown in Fig. 5. Furthermore, if input SNR increases, the curves of the simulation are more and more close to the trends of the analyses, which corroborates the explanation above. We notice that the simulation result for TAASS at the initial stage may be inferior to that for MASS due to a little more complexity of the former, which impacts the convergence rate of the algorithm.

Fig. 6 shows the result in a more severe condition ,where there are 4 interferers with one 3dB above the desired user's power, one with the same power level of the desired one and two others with 0.5dB lower than that of the desired signal. Compared with Fig. 5, the EMSE performance is worsen by the increase of the interferers, both in the number and power level. Under this condition, it is immediately observed that the TAASS algorithm displays a better performance over the MASS method, especially in high SNR condition. In other words, the EMSE analysis for MASS underestimates the real theoretical result since the independent disturbance by using the instantaneous error energy as a measure. Nevertheless, it is better to select MASS if the cost is considered in priority due to its less complex load. The proposed TAASS algorithm adapts to some harsh environments in the light of the performance.

The mismatch (steering vector error) condition is considered in Fig. 7, which includes two experiments. Fig. 7(a) shows the BER of each method versus the number of snapshots. The system model is the same as that in the first experiment. The results illustrate that the proposed algorithms converge faster and have better performance than the existing SG-Type methods and also close to that of the RLS method. The steering vector mismatch scenario is shown in Fig. 7(b). The assumed DOA of the SOI is a constant value 1^o away from the actual direction. Compared with Fig. 7(a), Fig. 7(b) indicates that the mismatch problem leads to a worse performance for all the solutions. The convergence rate of all the methods reduces whereas the devised algorithms are more robust to this mismatch, especially for the TAASS approach, reaching the steady-state rapidly with much lower complexity than the RLS algorithm.

C. Nonstationary Scenario Performance

In this part, we will evaluate the performance of the proposed MASS and TAASS algorithms in nonstationary environment, namely, when the number of users changes.

In Fig. 8, The system starts with 4 interferers, two of which have the same power as that of the desired signal and the rest of them with the power 0.5 dB lower than the desired one. Two more users with one of them 2 dB above the desired user's power level and the other 0.5 dB lower than that of the desired user, enter the system at 1000 symbols. In this condition, the parameters are set to the same values as those in Fig. 2 except $\mu_{max} = 3 \times 10^{-3}$ for MASS and $\mu_{max} = 5 \times 10^{-3}$ for TAASS due to optimization. As can be seen from the figure, SINRs of all the methods reduce at the same time. It is clear that the performance degradation of the proposed ones is much less significant than those of the others (FSS and ASS). In addition, MASS and TAASS can quickly track the change and recover to a steady-state, just a little slower than the RLS method. This figure demonstrates that the proposed algorithms have fast convergence even though they are less complex. The experiment shows that the proposed techniques exhibit satisfied performance after an abrupt change, in a nonstationary environment where the number of users/interferers suddenly changes in the system.

Fig. 9 depicts the step size values of the proposed algorithms as a function of the received symbols in a nonstationary setup corresponding to Fig. 8. It verifies that both MASS and TAASS mechanisms behave faster convergence improvement that tracks the system change rapidly and then reaches to the steady-state performance.

VI. Concluding Remarks

This paper proposed two adaptive CCM SG algorithms by employing low-complexity variable step size mechanisms for the beamforming technique to improve the performance. We compared the computational complexity of the new algorithms with the existent methods and further investigated the characteristics of the new mechanisms via analyses of the convergence, steady-state and tracking performance. The theoretical expressions were derived, in terms of EMSE, using the energy-preserving approach for both stationary and nonstationary scenarios. Considering the effects of MAI and AWGN in the array output makes the analysis more accurate and suitable to the practical application. Simulation experiments were conducted to verify the analytical results. The signature mismatch and users' number change conditions were emulated for comparing the proposed algorithms with the classic methods in order to show their robustness. The computer experiments illustrate that the new algorithms are superior to the existent SG methods in terms of convergence behavior and output performance in both stationary and nonstationary scenarios even though they are less complex. The proposed algorithms have demonstrated excellent performance and can be used for beamforming design of SDMA systems and other smart antenna systems.

APPENDIX

DERIVATION OF (28)

Since we make the derivation in the steady-state condition, that means, $i \rightarrow \infty$, so by defining $B = \xi_{min} + \xi_{ex}(\infty)$ and employing *Assumption 1*, we get

$$E[v^2(i)] = (1-\beta)^2[B^2 + \beta^2 B^2 + \ldots + \beta^{2(i-1)}B^2] \tag{61}$$

In a compact way, we also define $C = [B^2 + \beta^2 B^2 + \ldots + \beta^{2(i-1)}B^2]$. By multiplying β^2 on both sides, we obtain

$$\beta^2 C = \beta^2 B^2 + \beta^4 B^2 + \ldots + \beta^{2(i-1)}B^2 + \beta^{2i}B^2 \tag{62}$$

With $0 < \beta < 1$, if $i \rightarrow \infty$, β^{2i} can be ignored in (62) and then we have

$$\beta^2 C = C - B^2 \tag{63}$$

Rearranging (63) to get C in the form of B and β,

$$C = \frac{B^2}{1-\beta^2} \tag{64}$$

Substituting (64) into (61) and using *Assumption 2*, we find

$$\begin{aligned} E[v^2(i)] &= \frac{(1-\beta)^2 B^2}{1-\beta^2} \\ &= \frac{(1-\beta)(\xi_{min} + \xi_{ex}(\infty))^2}{1+\beta} \approx \frac{(1-\beta)\xi_{min}^2}{1+\beta} \end{aligned} \tag{65}$$

REFERENCES

[1] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 27-34, Jan. 1982.

[2] Y. Kameda and J. Ohga, "Adaptive microphone-array system for noise reduction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1391-1400, Dec. 1986.

[3] A. B. Gershman, V. I. Turchin, and V. A. Zverev, "Experimental results of locationzation of moving underwater signal by adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 43, pp. 2249-2257, Oct. 1995.

[4] A. B. Gershman, E. Nemeth, and J. F. Bohme, "Experimental performance of adaptive beamforming in sonar environment with a towed array and moving interfering sources," *IEEE Trans. Signal Processing*, vol. 48, pp. 246-250, Jan. 2000.

[5] C. Farsakh and J. A. Nossek, "Spatial covariance based downlink beamforming in an SDMA mobile radio system," *IEEE Trans. Commun.*, vol. 46, pp. 1497-1506, Nov. 1998.

[6] J. Chung, C. S. Hwang, K. Kim, and Y. K. Kim, "A random beamforming technique in MIMO systems exploiting multiuser diversity," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 848-855, Jun. 2003.

[7] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. J. Electron. Commun.*, vol. 53, pp. 1365-1376, Dec. 1999.

[8] S. Anderson, M. Millnert, M. Viberg, and B. Wahlberg, "An adaptive array for mobile communication systems," *IEEE Trans. Vehicular Technology*, vol. 40, pp. 230-236, Feb. 1991.

[9] J. Li and P. Stoica, *Robust Adaptive Beamforming*, Hoboken, NJ: Wiley, 2006.

[10] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Signal Processing*, vol. 37, pp. 984-995, Jul. 1989.

[11] P. Stoica and A. B. Gershman, "Maximum-likelihood DOA estimation by data-supported grid search," *IEEE Letters Signal Processing*, vol. 6, pp. 273-275, Oct. 1999.

[12] U. Madhow and M. L. Honig, "MMSE interference suppresion for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178-3188, Dec. 1994.

[13] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, pp. 944-960, Jul. 1995.

[14] D. N. Godard, "Self-recovering equalization and carrier tracking in two-demensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867-1875, Nov. 1980.

[15] J. Treichler and B. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Signal Processing*, vol. 31, pp. 459-472, Apr. 1983.

[16] J. J. Shynk and R. P. Gooch, "The constant modulus array for cochannel signal copy and direction finding," *IEEE Trans. Signal Processing*, vol. 44, pp. 652-660, Mar. 1996.

[17] R. Jr. Johnson, P. Schniter, T. J. Endres, and J. D. Behm, "Blind equalization using the constant modulus criterion: a review," *IEEE Proceedings*, vol. 86, pp. 1927-1950, Oct. 1998.

[18] L. Wang and R. C. de Lamare, "Constrained constant modulus RLS-based blind adaptive beamforming algorithm for smart antennas," *IEEE Proceedings. International Conference on Wireless Communication System ISWCS*, pp. 657-661, Oct. 2007.

[19] L. Wang, Y. L. Cai, and R. C. de Lamare, "Low-complexity adaptive step size constrained constant modulus SG-based algorithms for blind adaptive beamforming," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2593-2596, Apr. 2008.

[20] Z. Y. Xu and P. Liu, "Code-constrained blind detection of CDMA signals in multipath channels," *IEEE Signal Processing Letters*, vol. 9, pp. 389-392, Dec. 2002.

[21] Y. Chen, T. Le-Ngoc, B. Champagne, and C. Xu, "Recursive least squares constant modulus algorithm for blind adaptive array," *IEEE Trans. Signal Processing*, vol. 52, pp. 1452-1456, May 2004.

[22] S. Haykin, *Adaptive Filter Theory, 4rd ed.*, Englewood Cliffs, NJ: Prentice-Hall, 1996.

[23] V. J. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive step size," *IEEE Trans. Signal Process.*, vol. 41, pp. 2075-2087, Jun. 1993.

[24] H. J. Kushner and J. Yang, "Analysis of adaptive step-size sa algorithms for parameter tracking," *IEEE Trans. Autom. Control*, vol. 40, pp. 1403-1410, Aug. 1995.

[25] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Trans. Signal Processing*, vol. 40, pp. 1633-1642, July 1992.

[26] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-Type algorithm: analysis and simulations," *IEEE Trans. Signal Processing*, vol. 45, pp. 631-639, Mar. 1997.

[27] J. Mai and A. H. Sayed, "A feedback approach to the steady-state performance of fractionally spaced blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 48, pp. 80-91, Jan. 2000.

[28] A. V. Keerthi, A. Mathur, and J. Shynk, "Misadjustment and tracking analysis of the constant modulus array," *IEEE Trans. Signal Processing*, vol. 46, pp. 51-58, Jan. 1998.

[29] O. W. Kwon, C. K. Un, and J. C. Lee, "Performance of constant modulus adaptive digital filters for interference cancellation," *IEEE Trans. Signal Processing*, vol. 26, pp. 185-196, Feb. 1992.

[30] M. Rupp and A. H. Sayed, "A time-domain feedback analysis of filtered-error adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1428-1439, June 1996.

[31] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Processing*, vol. 49, pp. 314-324, Feb. 2001.

[32] P. Stoica and A. Nehorai, "Music, maximum likelihood and cramer-rao bound," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 37, pp. 720-741, May. 1989.

[33] A. H. Sayed and M. Rupp, "A time-domain feedback analysis of adaptive algorithms via the small gain theorem," *Proc. SPIE*, vol. 2563, pp.458-469, July 1995.

[34] R. C. de Lamare and R. Sampaio-Neto, "Low-complexity variable step-size mechanisms for stochastic gradient algorithms in minimum variance CDMA receivers," *IEEE Trans. Signal Processing*, vol. 54, pp. 2302-2317, Jun. 2006.

[35] H. H. Zeng, L. Tong, and C. R. Johnson, "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Inf. Theory*, vol. 44, pp. 1523-1538, July 1998.

[36] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[37] J. B. Whitehead and F. Takawira, "Performance analysis of the linearly constrained constant modulus algorithm-based multiuser detector," *IEEE Trans. Signal Processing*, vol. 53, pp. 643-653, Feb. 2005.

[38] S. Verdu, *Multiuser Detection*. Cambridge, UK: Cambridge Univ. Press, 1998.

[39] A. H. Sayed, *Fundamentals of Adaptive Filtering*. John Wiley Sons, NY, 2003.