Standard and Quasi-Standard Stochastic Power Control Algorithms

Jie Luo, *Member, IEEE*, Sennur Ulukus, *Member, IEEE*, and Anthony Ephremides, *Fellow, IEEE*

Abstract—In an energy-efficient wireless communication system, transmit powers are minimized subject to predetermined signal-to-interference ratio (SIR) requirements. In this paper, a general framework for distributed stochastic power control (PC) algorithms is proposed, where the transmit powers are updated based on stochastic approximations. The proposed algorithms are distributed in the sense that no global information is needed in the power updates. Interference to each user is estimated locally via noisy observations. Two types of stochastic PC algorithms are studied: standard stochastic PC algorithms where the interference estimator is unbiased, and quasi-standard stochastic PC algorithms where the interference estimator is biased. The conditions under which the stochastic PC algorithms converge to the unique optimal solution are identified. Corresponding to two classes of iteration step-size sequences, two types of convergence, the probability one convergence and convergence in probability, are shown for both algorithms based on recent results in the stochastic approximation literature.

Based on the theoretical results, some well-known stochastic PC algorithms, such as stochastic PC with matched filter receivers, and joint stochastic PC with blind minimum mean-squared error (MMSE) interference suppression, are revisited; several new stochastic PC algorithms, such as stochastic PC with minimum-power base-station assignment, and stochastic PC with limited diversity, are proposed. It is shown that these algorithms fall into either the standard or the quasi-standard stochastic PC framework. Simulation results are given to illustrate the performance of the proposed algorithms in practical systems.

Index Terms—Distributed algorithms, stochastic approximation, stochastic power control.

I. INTRODUCTION

SINCE power is an important and limited resource in wireless communication systems, power control (PC) algorithms that minimize the transmission power while ensuring the quality of service (QoS) have been widely studied in the literature. Earlier work on PC viewed the problem as constrained optimization, and found the optimal transmit power for each user by solving the optimization problem directly [2]–[4]. Such algorithms are "centralized" since they assume that the optimization problem is solved at a centralized server using the knowledge of some global parameters such as the channel gains of all users to all base stations. When the size of the system increases, computational complexity and acquiring the knowledge of such global parameters become serious issues. Mainly due to this reason, many distributed PC algorithms have been developed [5], [6]. A general framework for the PC problem is developed in [7] where the PC problem is reduced to finding the smallest power vector $\boldsymbol{p}$ that satisfies $\boldsymbol{p} \geq \boldsymbol{I}(\boldsymbol{p})$, where $\boldsymbol{I}(\boldsymbol{p})$ is the interference vector, whose ith component I_i is the generalized interference user i experiences. Although $\boldsymbol{I}(\boldsymbol{p})$ may take different closed-form expressions in different PC problems, it was shown in [7] that, if the interference function $\boldsymbol{I}(\boldsymbol{p})$ is *standard* (see detailed definition in Section II), then the iteration $\boldsymbol{p}(n+1) = \boldsymbol{I}(\boldsymbol{p}(n))$ converges to the unique componentwise smallest feasible power vector. Such an iterative algorithm requires only local information and hence can be carried out in a distributed fashion. It also significantly reduces the complexity of the PC and allows the wireless systems to adapt the transmission powers to the possibly time-varying system parameters.

The algorithms studied in [7] are *deterministic* in the sense that they require the perfect knowledge of the received interference power. However, in a practical system, the interference power can only be estimated using noisy observations. Stochastic PC that uses noisy interference estimates was first considered in [8]. Under the assumption of conventional matched filter receivers, [8] showed that the stochastic PC, which updates the power vector iteratively using stochastic approximations, converges to the optimal power vector in the mean square error (MSE) sense as long as the step-size sequence in the iteration meets certain requirements. These results were later extended to the cases when a linear receiver or a decision feedback receiver is used [9], [10]. A fast PC algorithm with averaging was also proposed in [10]. The common feature of [8] and [10] is that, for both cases, the deterministic interference function is linear in transmit powers. Among the iterative PC algorithms, the joint PC and receiver optimization algorithm also attracted special attention. The convergence of the deterministic version was studied in [11] and the stochastic implementation was studied in [12], where a local convergence was shown for the stochastic PC algorithm given that the step size sequence satisfies certain requirements. Unlike the deterministic iterative PC, where the convergence was proved for a general class of algorithms [7], a common feature on the convergence results for stochastic PC in [8], [12],

[10] is that they assume specific closed-form expressions for the interference function.

The stochastic PC algorithms proposed in [8], [12], [10] are based on stochastic approximation methods, which were introduced in [13] and [14]. Since then, stochastic approximation-based algorithms have been studied extensively in many different contexts including communications, signal processing, and neural networks. These methods update the system state variable iteratively using noise-corrupted control variables. Under certain conditions on the system dynamics, when the iteration step size is decreased suitably, the effect of the noise can be made small, and the stochastic iterations converge asymptotically to the equilibrium point of the corresponding deterministic iterations. Usually, different types of convergence can be proved under different assumptions on the step-size sequence. Although various convergence proofs for general stochastic approximation algorithms are available in the literature, verifying the validity of the conditions under which these convergence proofs are carried out in specific practical problems, such as a stochastic PC problem, is not always easy, depending on the system specifications and the methods used in the proofs.

In this paper, we propose a general framework for *stochastic* PC algorithms where the transmitted powers are iteratively updated in a stochastic approximation fashion, using the noisy estimates of the interference. We consider the general PC problem where the deterministic interference function is *standard*, and it satisfies the Lipschitz condition (defined later). We define two types of stochastic interference functions. When the interference estimate is unbiased and a "growing condition" is satisfied by the estimation noise, we term the stochastic inference function the standard stochastic interference function. Similar to [7], we call the stochastic iterative PC algorithm which is obtained using a standard stochastic interference function, a standard stochastic PC algorithm. We show that the stochastic PC algorithms studied in [8] and [10] fall into this category. However, in the case when PC is combined with additional parameter optimization, such as joint PC and receiver filter optimization, and joint PC and base station assignment, the interference estimator may not be unbiased. For such situations, we define the stochastic interference estimate as the quasi-standard stochastic interference function, if the estimation bias satisfies an additional "bias condition." We call the corresponding iterative PC algorithm the quasi-standard stochastic power control algorithm.

Under certain conditions on the step-size sequence, we show that both algorithms (standard and quasi-standard) converge to the solution of the deterministic power control problem. Two types of convergence are proved for both standard and quasi-standard stochastic PC algorithms, corresponding to different conditions on the step-size sequence. When the step size is decreasing, and is decreasing faster than a certain rate, we show that both algorithms converge to the componentwise minimum power vector with probability one. In the case when the step size is decreasing slowly, we show that both algorithms converge to the minimum power vector *in probability*. The latter proof also indicates that, if the step size is fixed at a small constant α^*, the power vector will concentrate around the optimal power vector.

Although we present results for both standard and quasi-standard stochastic PC algorithms, we provide proofs only for the quasi-standard stochastic PC algorithms, since every standard stochastic PC algorithm is also a quasi-standard algorithm. The proofs provided in this paper are based on the recent literature on stochastic approximation methods [15], [16]. Specifically, in the convergence proofs, we focus on verifying the conditions of [15, Theorem 4.3] and [16, Theorem 2.3] while most of the details that already appeared in [15] and [16] are skipped. However, since the proof for the probability one convergence is relatively short, for the convenience of the reader, a brief but complete proof is given in the Appendix B.

Based on the theoretical analysis, we revisit the problem of stochastic PC with matched filter receivers [8]. We verify that it falls into the standard stochastic PC algorithm category. The joint stochastic PC and blind minimum MSE (MMSE) interference suppression [11], [9] is also revisited. Several new stochastic PC algorithms, including stochastic PC with minimum-power base station assignment, and stochastic PC with limited diversity, are proposed. By verifying that they are all quasi-standard stochastic PC algorithms, the convergence of these algorithms are established. For all these stochastic PC algorithms, modified versions, that allow parallel implementation of the stochastic PC and the parameter optimization, are presented. In such implementations, the PC algorithm does not have to wait for the convergence of the parameter optimization between consecutive power updates. It is also shown via these examples that failing to meet the conditions on the quasi-standard stochastic PC interference estimates may result in divergence of the algorithm, or the convergence of the algorithm to an undesired power vector.

The rest of the paper is organized as follows. In Section II, we present the definitions of standard deterministic PC algorithm [7], the standard stochastic PC algorithm, and the quasi-standard stochastic PC algorithm. The convergence analysis is presented in Section III. Examples of joint stochastic PC with blind MMSE interference suppression, the stochastic PC with minimum-power base station assignment, and the stochastic PC with limited diversity are proposed in Section IV. For each of these algorithms, we verify that they are either standard or quasi-standard. Computer simulation results are given in Section V to illustrate the performance of the proposed algorithms in practical systems. Conclusions are given in Section VI.

II. STANDARD AND QUASI-STANDARD PC ALGORITHMS

Suppose $\boldsymbol{p}$ is the power vector whose ith component, p_i, is the transmit power of user i. The signal-to-interference ratio (SIR) requirements of all users can be expressed as a single vector inequality as shown in [7]

$$\boldsymbol{p} \geq \boldsymbol{I}(\boldsymbol{p}) \tag{1}$$

where $\boldsymbol{I}(\boldsymbol{p}) = [I_1(\boldsymbol{p}), \dots, I_K(\boldsymbol{p})]^T$ is the interference function, and K is the number of users. Suppose $\boldsymbol{I}(\boldsymbol{p})$ is *standard* as defined by the following [7].

Definition 1: Interference function $\boldsymbol{I}(\boldsymbol{p})$ is standard if for all $\boldsymbol{p} \geq \boldsymbol{0}$, the following properties are satisfied.

- Positivity. $\boldsymbol{I}(\boldsymbol{p}) > \boldsymbol{0}$.
- Monotonicity. If $\boldsymbol{p} \geq \boldsymbol{p}'$, then $\bar{\boldsymbol{I}}(\boldsymbol{p}) \geq \bar{\boldsymbol{I}}(\boldsymbol{p}')$.
- Scalability. For all $\eta > 1$, $\eta \boldsymbol{I}(\boldsymbol{p}) > \boldsymbol{I}(\eta \boldsymbol{p})$.

It is shown in [7] that, if the PC is feasible, then the deterministic PC algorithm

$$\boldsymbol{p}(n+1) = \boldsymbol{I}(\boldsymbol{p}(n)) \qquad (2)$$

converges to the componentwise smallest feasible power vector $\boldsymbol{p}^*$. In fact, the inequality in (1) is satisfied with equality at $\boldsymbol{p}^*$, i.e., $\boldsymbol{p}^* = \bar{\boldsymbol{I}}(\boldsymbol{p}^*)$. Furthermore, for all α, such that $0 < \alpha \leq 1$, since $(1 - \alpha)\boldsymbol{p} + \alpha \boldsymbol{I}(\boldsymbol{p})$ is also a standard interference function, the PC algorithm

$$\boldsymbol{p}(n+1) = (1 - \alpha)\boldsymbol{p}(n) + \alpha \boldsymbol{I}(\boldsymbol{p}(n)) \qquad (3)$$

converges to $\boldsymbol{p}^*$, as well.

The above algorithm is deterministic in the sense that it requires perfect knowledge of $\boldsymbol{I}(\boldsymbol{p})$. In practical systems, we cannot know $\boldsymbol{I}(\boldsymbol{p})$ perfectly, but we may have a random estimate of $\boldsymbol{I}(\boldsymbol{p})$, denoted as $\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta})$, with $\boldsymbol{\theta}$ denoting the estimation noise. Consider now the following stochastic PC algorithm:

$$\boldsymbol{p}(n+1) = (1 - \alpha(n))\boldsymbol{p}(n) + \alpha(n)\tilde{\boldsymbol{I}}(\boldsymbol{p}(n), \boldsymbol{\theta}(n)) \qquad (4)$$

where $\tilde{\boldsymbol{I}}(\boldsymbol{p}(n), \boldsymbol{\theta}(n))$ is the noisy estimate of $\boldsymbol{I}(\boldsymbol{p}(n))$, and $\alpha(n)$ is the step size at the nth iteration. We define the standard stochastic interference function as follows.

Definition 2: Stochastic interference function $\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta})$ is standard if for all $\boldsymbol{p} \geq \boldsymbol{0}$, the following properties are satisfied.

- Mean condition. $E\left[\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta}) | \boldsymbol{p}\right] = \bar{\boldsymbol{I}}(\boldsymbol{p})$, and $\bar{\boldsymbol{I}}(\boldsymbol{p})$ is a standard deterministic interference function.
- Lipschitz condition. There exits a constant $K_1 > 0$, such that

$$\|\boldsymbol{I}(\boldsymbol{p}_1) - \boldsymbol{I}(\boldsymbol{p}_2)\|^2 \leq K_1 \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|^2. \qquad (5)$$

- Growing condition. There exits a constant $K_2 > 0$, such that

$$E\left[\|\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta}) - \boldsymbol{I}(\boldsymbol{p})\|^2 | \boldsymbol{p}\right] \leq K_2(1 + \|\boldsymbol{p}\|^2). \qquad (6)$$

When the stochastic interference function is standard, we call (4) a standard stochastic PC algorithm. It is easy to verify that the stochastic PC algorithms that have been studied in [8] and [10] are standard.

Although it seems that the mean condition in Definition 2 is quite natural, it may not be satisfied when there is an additional parameter optimization combined with the PC problem. Assume that the deterministic interference function can be written as $\boldsymbol{I}(\boldsymbol{p}) = \boldsymbol{I}(\boldsymbol{p}, \boldsymbol{X}^*(\boldsymbol{p}))$, where $\boldsymbol{X}^*(\boldsymbol{p})$ is an additional set of parameters obtained via minimizing a cost function $\boldsymbol{J}(\boldsymbol{p}, \boldsymbol{X})$, i.e.,

$$\boldsymbol{X}^*(\boldsymbol{p}) = \arg \min_{\boldsymbol{X}} \boldsymbol{J}(\boldsymbol{p}, \boldsymbol{X}). \qquad (7)$$

In the joint PC and MMSE interference suppression [11], [9], for example, $\boldsymbol{X}$ is the receiver filter coefficients, and $\boldsymbol{J}(\boldsymbol{p}, \boldsymbol{X})$ is the deterministic interference when $\boldsymbol{p}$ and $\boldsymbol{X}$ are given. We choose the filter coefficients to minimize all components of $\boldsymbol{J}(\boldsymbol{p}, \boldsymbol{X})$,

and $\boldsymbol{X}^*(\boldsymbol{p})$ turns out to be the MMSE filter coefficients for all users, given $\boldsymbol{p}$.

Due to the distributed optimization requirements and the stochastic nature of the problem, usually, we cannot guarantee that we choose $\boldsymbol{X}$ exactly at the optimal value. Hence,

$$E\left[\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta}) | \boldsymbol{p}\right] = \boldsymbol{I}(\boldsymbol{p})$$

may not be true in general. In a general joint stochastic PC and parameter optimization problem, we assume that

$$E\left[\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta}) | \boldsymbol{p}\right] = \boldsymbol{I}(\boldsymbol{p}) + \boldsymbol{g}(\boldsymbol{p}). \qquad (8)$$

We define the quasi-standard stochastic interference function as follows.

Definition 3: Stochastic interference function $\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta})$ is quasi-standard if for all $\boldsymbol{p} \geq \boldsymbol{0}$, the following properties are satisfied.

- Mean condition. $E\left[\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta}) | \boldsymbol{p}\right] = \boldsymbol{I}(\boldsymbol{p}) + \boldsymbol{g}(\boldsymbol{p})$, where $\boldsymbol{I}(\boldsymbol{p})$ is a standard deterministic interference function, and $\boldsymbol{g}(\boldsymbol{p})$ is a bias term.
- Bias condition. There exits a constant $K_3 > 0$ and a sequence $1 \geq \beta(n) \geq 0$, such that

$$\|\boldsymbol{g}(\boldsymbol{p}(n))\| \leq \beta(n)K_3(1 + \|\boldsymbol{p}(n)\|). \qquad (9)$$

- Lipschitz condition. There exits a constant $K_1 > 0$, such that

$$\|\boldsymbol{I}(\boldsymbol{p}_1) - \boldsymbol{I}(\boldsymbol{p}_2)\|^2 \leq K_1 \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|^2. \qquad (10)$$

- Growing condition. There exits a constant $K_2 > 0$, such that

$$E\left[\|\tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta}) - \boldsymbol{I}(\boldsymbol{p}) - \boldsymbol{g}(\boldsymbol{p})\|^2 | \boldsymbol{p}\right] \leq K_2(1 + \|\boldsymbol{p}\|^2). \qquad (11)$$

Note that a standard stochastic interference function is also quasi-standard, but the converse is not true. Whether an interference function is standard or quasi-standard depends on the design of the estimator. A biased quasi-standard interference estimator does not necessarily exclude the possibility that one may be able to design another estimator which is standard, for the same system.

III. CONVERGENCE OF STANDARD AND QUASI-STANDARD STOCHASTIC PC ALGORITHMS

In this section, we study the convergence of standard and quasi-standard stochastic PC algorithms. We show that the algorithms converge to the optimal power vector when the step-size sequence satisfies certain conditions. Two different types of convergence results are shown under different choices of the step-size sequence. The proofs presented in this paper are based on, and modified from, the recent papers on stochastic approximation methods, specifically, [15, Theorem 4.3] and [16, Theorem 2.3]. We only present proofs and necessary discussion that verify the conditions of [15, Theorem 4.3] and [16, Theorem 2.3], except that for the case of probability one convergence, a brief but complete proof of Theorem 1 is given in Appendix B.

Throughout this section, we assume that the PC problem is feasible, i.e., there exits a power vector $\boldsymbol{p}$ that satisfies the inequality in (1).

A. Convergence of the Mean Ordinary Differential Equation

Consider the deterministic PC algorithm in (3). Define a function $V(\boldsymbol{p})$ by

$$\nabla_{\boldsymbol{p}} V(\boldsymbol{p}) = \boldsymbol{p} - \boldsymbol{I}(\boldsymbol{p}), \qquad V(\boldsymbol{p}^*) = 0 \qquad (12)$$

where $\nabla_{\boldsymbol{p}}$ represents the gradient operator with respect to $\boldsymbol{p}$. Since $\boldsymbol{I}(\boldsymbol{p})$ is Lipschitz continuous, from (3), we obtain [17]

$$V(\boldsymbol{p}(n+1)) = V(\boldsymbol{p}(n)) - \alpha \| \boldsymbol{p}(n) - \boldsymbol{I}(\boldsymbol{p}(n)) \|^2 + O(\alpha^2). \quad (13)$$

Hence, if $\boldsymbol{p}(n) \neq \boldsymbol{p}^*$ and α is small enough, we have $V(\boldsymbol{p}(n+1)) < V(\boldsymbol{p}(n))$. Noting that the PC algorithm (3) converges to $\boldsymbol{p}^*$ from any initial point for all $0 < \alpha \leq 1$, if $\boldsymbol{p}(0) \neq \boldsymbol{p}^*$, we can always find a sequence $\boldsymbol{p}(0), \ldots, \boldsymbol{p}(n), \ldots$ such that $V(\boldsymbol{p}(0)) > \cdots > V(\boldsymbol{p}(n))$ and $\boldsymbol{p}(n) \to \boldsymbol{p}^*$ as $n \to \infty$. Therefore, we have

$$V(\boldsymbol{p}) > 0, \qquad \forall \boldsymbol{p} \neq \boldsymbol{p}^*. \qquad (14)$$

This shows that $V(\boldsymbol{p})$ is a Lyapunov function.

Define the points of time as $t(n) = \sum_{i=0}^{n-1} \alpha(i)$. In (3), we subtract $\boldsymbol{p}(n)$ from both sides and then divide both sides by $\alpha(n)$. Let $\alpha(n) \to 0$. Then, the deterministic PC algorithm can be approximated by the following ordinary differential equation (ODE):

$$\frac{d\boldsymbol{p}}{dt} = -[\boldsymbol{p}(t) - \boldsymbol{I}(\boldsymbol{p}(t))]. \qquad (15)$$

We term (15) as the characteristic ODE of the PC algorithm. Since the evolution of (15) is indeed a gradient descent search that minimizes $V(\boldsymbol{p})$, from the preceding analysis, we can see that $\boldsymbol{p}(t) \to \boldsymbol{p}^*$ as $t \to \infty$.

B. Probability One Convergence

In this subsection, we present the conditions under which the stochastic PC algorithms converge to $\boldsymbol{p}^*$ with probability one. We first show a bound on the power vector, which will be useful in the convergence proof of the main result.

Lemma 1: There exists a constant $K_4 > 0$, such that

$$\| \boldsymbol{p} \|^2 \leq K_4 (1 + \| \boldsymbol{p} - \boldsymbol{I}(\boldsymbol{p}) \|^2). \qquad (16)$$

The proof of Lemma 1 is given in Appendix A.

Next in Theorem 1, we characterize the conditions on the iteration step-size sequence such that the stochastic PC algorithm converges to $\boldsymbol{p}^*$ with probability one.

Theorem 1: Suppose the stochastic interference $\tilde{\boldsymbol{I}}(\boldsymbol{p}(n))$ is standard and the step-size sequence satisfies

$$\sum_{n=0}^{\infty} \alpha(n) = \infty, \qquad \sum_{n=0}^{\infty} \alpha(n)^2 < \infty. \qquad (17)$$

Or, if the stochastic interference $\tilde{\boldsymbol{I}}(\boldsymbol{p}(n))$ is quasi-standard and the step-size sequence satisfies

$$\sum_{n=0}^{\infty} \alpha(n) = \infty, \quad \sum_{n=0}^{\infty} \alpha(n)^2 < \infty, \quad \sum_{n=0}^{\infty} \alpha(n)\beta(n) < \infty \qquad (18)$$

with $\beta(n)$ being the sequence given in the bias condition (9), then the power vector in the corresponding standard or quasi-standard stochastic PC algorithm in (4) converges to $\boldsymbol{p}^*$ with probability one.

The proof of this theorem is based on the proof of [15, Theorem 4.3] with the help of Lemma 1, and the well-known martingale convergence theorem [18]. Since the proof is relatively short, for the sake of completeness, we give a complete proof of Theorem 1 in Appendix B.

C. Convergence in Probability

Due to the requirements on robustness and the capability to track time-varying environments, the iteration step-size sequences of many practical systems are not allowed to decrease to zero. In this subsection, we study the situation when either $\sum_{n=0}^{\infty} \alpha(n)^2 < \infty$, or $\sum_{0}^{\infty} \alpha(n)\beta(n) < \infty$, or both are violated. These include the situation when the step-size sequence decreases to zero slowly, and the situation when the step size is fixed at a small constant. We study the quasi-standard stochastic PC algorithm only; the standard stochastic PC case is covered by simply setting the bias term to zero. Since the proof of the main result given in Theorem 4 in this section can be derived from the proof of [16, Theorem 2.3] with minor modifications, we focus on verifying the conditions of [16, Theorem 2.3] and only sketch the main steps of the proof.

We first define the scaled ODE and the limit ODE. The convergence of the limit ODE is an important condition imposed in [16].

Given $r > 0$, define the scaled interference function $\boldsymbol{I}_r(\boldsymbol{p})$ as

$$\bar{\boldsymbol{I}}_r(\boldsymbol{p}) = \frac{\boldsymbol{I}(r\boldsymbol{p})}{r}. \qquad (19)$$

Correspondingly, we define the scaled ODE as

$$\frac{d\boldsymbol{p}}{dt} = -[\boldsymbol{p}(t) - \boldsymbol{I}_r(\boldsymbol{p}(t))] \qquad (20)$$

which can be written equivalently as

$$\frac{dr\boldsymbol{p}}{dt} = -[r\boldsymbol{p}(t) - \boldsymbol{I}(r\boldsymbol{p}(t))]. \qquad (21)$$

Since $\boldsymbol{p}^* - \boldsymbol{I}(\boldsymbol{p}^*) = \boldsymbol{0}$, $\boldsymbol{p}_r^* = \frac{\boldsymbol{p}^*}{r}$ is the only equilibrium, and is also a globally stable equilibrium of (20).

According to the scalability assumption in Definition 1, if $r > 1$, then $\boldsymbol{I}_r(\boldsymbol{p}) < \boldsymbol{I}(\boldsymbol{p})$. Therefore, noting that $\boldsymbol{I}_r(\boldsymbol{p}) > 0$, $\lim_{r \to \infty} \boldsymbol{I}_r(\boldsymbol{p})$ exists. Define the limit of the scaled interference function as

$$\boldsymbol{I}_{\infty}(\boldsymbol{p}) = \lim_{r \to \infty} \boldsymbol{I}_r(\boldsymbol{p}) = \lim_{r \to \infty} \frac{\boldsymbol{I}(r\boldsymbol{p})}{r}. \qquad (22)$$

Define the limit ODE as

$$\frac{d\boldsymbol{p}}{dt} = -[\boldsymbol{p}(t) - \boldsymbol{I}_{\infty}(\boldsymbol{p}(t))]. \qquad (23)$$

Then, we have the following theorem.

Theorem 2: $\boldsymbol{p}_\infty^* = \boldsymbol{0}$ is an asymptotically stable equilibrium, and is the only equilibrium for the limit ODE (23).

The proof of Theorem 2 is given in Appendix C.

In the quasi-standard stochastic PC algorithm, define

$$\boldsymbol{w}(\boldsymbol{p}, \boldsymbol{\theta}) = \tilde{\boldsymbol{I}}(\boldsymbol{p}, \boldsymbol{\theta}) - \boldsymbol{I}(\boldsymbol{p}) - \boldsymbol{g}(\boldsymbol{p}). \tag{24}$$

Similar to (19), we define the scaled noise and the scaled bias as

$$\boldsymbol{w}_r(\boldsymbol{p}, \boldsymbol{\theta}) = \frac{\boldsymbol{w}(r\boldsymbol{p}, \boldsymbol{\theta})}{r}, \qquad \boldsymbol{g}_r(\boldsymbol{p}) = \frac{\boldsymbol{g}(r\boldsymbol{p})}{r}. \tag{25}$$

When $r \geq 1$, from (9) and (11), we obtain

$$E\left[\|\boldsymbol{w}_r(\boldsymbol{p}, \boldsymbol{\theta})\|^2 | \boldsymbol{p}\right] \leq K_2(1 + \|\boldsymbol{p}\|^2)$$
$$\|\boldsymbol{g}_r(\boldsymbol{p}(n))\| \leq \beta(n)K_3(1 + \|\boldsymbol{p}(n)\|). \tag{26}$$

Both Theorem 2 and (26) are key results for the proof of the following theorem, which is one of the major steps toward the proof of the main result.

Theorem 3: Suppose there exist $\alpha^* > 0$ and $N > 0$, such that $\alpha(n) \leq \alpha^*$ and $\beta(n) \leq \sqrt{\alpha^*}, \forall n \geq N$, then in the quasi-standard stochastic PC algorithm, we can find a constant $K_7 < \infty$ that satisfies

$$\lim_{n\to\infty} \sup E\left[\|\boldsymbol{p}(n)\|^2\right] \leq K_7. \tag{27}$$

The proof can be carried out by following the proof of [16, Theorem 2.1], where the result of Theorem 2 is needed. Although the bias term, $\boldsymbol{g}(\boldsymbol{p})$, is not considered in [16], it can be taken into consideration by straightforward modifications to [16]; therefore, the proof is omitted in this paper.

Since Theorem 2 plays an important role in proving Theorem 3, which is indeed the major challenge in proving the main convergence result, it is worthwhile to explain briefly the purpose of showing Theorem 2. In standard and quasi-standard stochastic PC algorithms, the dynamics driven by the ODE is regarded as "force of convergence" since the deterministic PC iteration converges to the optimal solution. The dynamics driven by the noise is regarded as "force of divergence" because it can drive the power vector away from the optimal value. Since the limit ODE characterizes the system dynamics when $\|\boldsymbol{p}\|^2 \to \infty$, the convergence of the limit ODE implies that, the "force of convergence" is strong enough to dominate the system dynamics when $\|\boldsymbol{p}\|^2$ is large. This leads to the conclusion that $E[\|\boldsymbol{p}\|^2]$ is asymptotically bounded, as presented in Theorem 3.

Next, we present in the following theorem the main convergence result when the step size sequence is decreasing slowly, or is fixed at a small constant.

Theorem 4: Let $\alpha^* > 0$ as introduced in Theorem 3. For any $\epsilon > 0$, there exists $K_8 = K_8(\epsilon)$, which is *not* a function of α^*, such that in the quasi-standard stochastic PC algorithm

$$\lim_{n\to\infty} \sup P(\|\boldsymbol{p}(n) - \boldsymbol{p}^*\| \geq \epsilon) \leq K_8 \alpha^*. \tag{28}$$

Here $P(A)$ denotes the probability of event A.

Theorem 4 can be proved by following the proof of [16, Theorem 2.3], and making necessary modifications to include the bias term. Hence, it is omitted in this paper.

From Theorem 4, if $\alpha(n) \to 0$ and $\beta(n) \to 0$, for any $\epsilon > 0$, $P(\|\boldsymbol{p}(n) - \boldsymbol{p}^*\| < \epsilon) \to 1$. In other words, $\boldsymbol{p}(n)$ converges to $\boldsymbol{p}^*$ in probability. It is important to note that this result is obtained without requiring either $\sum_{n=0}^{\infty} \alpha(n)^2 < \infty$ or $\sum_0^{\infty} \alpha(n)\beta(n) < \infty$. In addition, when the step-size sequences $\alpha(n)$ and $\beta(n)$ are fixed, respectively, at small constants, (28) indicates that the power vector will be concentrated around the optimal value. However, usually the error between the power vector and the optimal value does not vanish for nonvanishing step-size sequences; this is the price we have to pay for the capability of adapting to time-varying environments.

IV. EXAMPLES OF STANDARD AND QUASI-STANDARD STOCHASTIC PC ALGORITHMS

In this section, we present several examples of stochastic PC algorithms; the deterministic versions of these PC algorithms have been studied in [7], [11], and [10]. In the stochastic implementations, we address the case when the step-size sequence decreases to zero asymptotically. However, as supported by Theorem 4, a practical system can certainly set the iteration step size at a fixed small constant.

A. System Model

Consider the uplink of a symbol synchronous wireless code-division multiple-access (CDMA) system with K users and M base stations. The chip matched filter output at the assigned base station of user i can be written as

$$\boldsymbol{z}_i = \sum_{j=1}^{K} \sqrt{p_j} \sqrt{h_{ij}} b_j \boldsymbol{s}_j + \boldsymbol{v}_i \tag{29}$$

where p_j is the transmit power of user j; h_{ij} is the channel gain of user j to the assigned base station of user i; b_j and $\boldsymbol{s}_j$ are the transmitted information symbol and the normalized signature sequence of user j, respectively; $\boldsymbol{v}_i$ is a white Gaussian noise vector with zero mean and $E[\boldsymbol{v}_i \boldsymbol{v}_i^T] = \sigma^2 \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix.

One should note that (29) also models the situation of receiver beamforming, as presented in [19], [20]. Suppose each base station is equipped with L antennas. Denote h_{ilj} as the channel gain from user j to the lth antenna ($l \in \{1, \ldots, L\}$) of the base station assigned to user i. The chip matched filter output of the lth antenna at the base station of user i can be written as

$$\boldsymbol{z}_{il} = \sum_{j=1}^{K} \sqrt{p_j} \sqrt{h_{ilj}} b_j \boldsymbol{s}_j + \boldsymbol{v}_{il} \tag{30}$$

where $\boldsymbol{v}_{il}$ is a white Gaussian noise vector with zero mean and $E[\boldsymbol{v}_{il}\boldsymbol{v}_{il}^T] = \sigma^2 \boldsymbol{I}$. Now, stack the matched filter outputs of the L antennas together. Define $\boldsymbol{z}_i = [\boldsymbol{z}_{i1}^T, \ldots, \boldsymbol{z}_{iL}^T]^T$; $\boldsymbol{v}_i = [\boldsymbol{v}_{i1}^T, \ldots, \boldsymbol{v}_{iL}^T]^T$. Define $h_{ij} = \sum_{l=1}^{L} h_{ilj}$, and consider

$$\tilde{\boldsymbol{s}}_j = \left[\sqrt{\frac{h_{i1j}}{h_{ij}}}\boldsymbol{s}_j^T, \ldots, \sqrt{\frac{h_{iLj}}{h_{ij}}}\boldsymbol{s}_j^T\right]^T$$

as the equivalent spreading sequence of user j (with a spreading factor of NL). We can write the stacked chip matched filter output as

$$z_i = \sum_{j=1}^{K} \sqrt{p_j} \sqrt{h_{ij}} b_j \tilde{s}_j + v_i. \qquad (31)$$

It is easily seen that (31) and (29) fall into the same category.[1] For simplicity, in this paper, we focus on the notation of (29).

Assume that the user symbols are detected by linear filters at the base stations. Let c_i denote the receiver filter of user i at its assigned base station. The receiver filter output of user i is

$$y_i = \sum_{j=1}^{K} \sqrt{p_j} \sqrt{h_{ij}} (c_i^T s_j) b_j + c_i^T v_i. \qquad (32)$$

B. Fixed Base Station Assignment With Matched Filter Receiver

The deterministic and stochastic versions of the PC problem when the base station uses conventional matched filters to detect user signals was studied in [7] and [8], respectively. In this case, $c_i = s_i$. Suppose that γ_i is the SIR target of user i, then the deterministic interference function can be written as [7]

$$I_i(p) = \frac{\gamma_i}{h_{ii}} \left(\sum_{j \neq i} p_j h_{ij} (s_i^T s_j)^2 + \sigma^2 \right). \qquad (33)$$

The stochastic estimate of $I_i(p)$ is [8]

$$\tilde{I}_i = \frac{\gamma_i}{h_{ii}} y_i^2 - \gamma_i p_i. \qquad (34)$$

From (32) and using the fact that b_i and b_j are independent of each other, and independent of v_k for all $i \neq j, k$, we can easily verify that [8]

$$E[\tilde{I}_i | p] = \frac{\gamma_i}{h_{ii}} \left(\sum_{j \neq i} p_j h_{ij} (s_i^T s_j)^2 + \sigma^2 \right) = I_i(p) \qquad (35)$$

which is the mean condition in Definition 2. Since $\bar{I}(p)$ is linear in p, the Lipschitz condition in Definition 2 follows immediately. Also, it is not difficult to see that $\tilde{I}$ satisfies the growing condition in Definition 2. Hence, $\tilde{I}$ is a standard stochastic interference function.

C. Fixed Base Station Assignment With Blind MMSE Receiver

The deterministic and stochastic implementations of the joint PC and MMSE interference suppression were studied in [11], [9], [12], and [1], respectively. In this subsection, we show that the stochastic version of the PC algorithm falls into the quasi-standard category.

From (32), the SIR of user i can be written as

$$\text{SIR}_i = \frac{p_i h_{ii} (c_i^T s_i)^2}{\sum_{j \neq i} p_j h_{ij} (c_i^T s_j)^2 + \sigma^2 (c_i^T c_i)}. \qquad (36)$$

Define

$$J_i(p, c_i) = \frac{\gamma_i}{h_{ii}} \frac{\sum_{j \neq i} p_j h_{ij} (c_i^T s_j)^2 + \sigma^2 (c_i^T c_i)}{(c_i^T s_i)^2} \qquad (37)$$

where γ_i is the SIR target of user i. Then the QoS requirements of the users can be written as

$$p_i \geq J_i(p, c_i). \qquad (38)$$

Now, consider the filter coefficients c_i as an optimization parameter. Since we want to minimize the powers, we should choose c_i such that $J_i(p, c_i)$ is minimized for each user. Define

$$I_i(p) = \min_{c_i} J_i(p, c_i). \qquad (39)$$

Note that c_i that minimizes $J_i(p, c_i)$ is the scaled version of the well-known MMSE filter [21]

$$c_i^* = k_i A_i^{-1} s_i \qquad (40)$$

where $A_i = \sum_{j \neq i} p_j h_{ij} s_j s_j^T + \sigma^2 I$ and $k_i > 0$ is an arbitrary constant. Substituting (40) into (39), we get

$$I_i(p) = \frac{\gamma_i}{h_{ii}} \frac{1}{s_i^T A_i^{-1} s_i}. \qquad (41)$$

It has been shown in [11] that $I(p)$ is a standard deterministic interference function, and the corresponding deterministic PC algorithm converges to the componentwise minimum power vector.

In the stochastic implementation of the combined PC and filter optimization, we assume that the information of other users is not available at the receiver, and therefore, A_i and c_i^* cannot be computed directly. Consequently, one has to construct the optimum filter coefficients c_i^* by using the local information only. When the powers are fixed, such a construction can be done by using the blind MMSE method introduced in [22]. The blind MMSE adaptation rule can be summarized as [22]

$$c_i(l+1) = c_i(l) - \mu(l) c_i(l)^T z_i(l) (z_i(l) - s_i^T z_i(l) s_i) \quad (42)$$

where l is the iteration index and $\mu(l)$ is the step size for the lth iteration in the blind MMSE algorithm. In this algorithm, c_i is initialized with an arbitrary vector satisfying $c_i(0)^T s_i = 1$. It is shown in [22] using the results of [23] that, for given fixed powers, $p(n)$, if $\mu(l)$ satisfies

$$\sum_{l=0}^{\infty} \mu(l) = \infty \quad \text{and} \quad \sum_{l=0}^{\infty} \mu(l)^2 < \infty$$

then (42) converges to $c_i(n)^*$ when $l \to \infty$ with probability one. Since (42) is also a stochastic approximation algorithm, given $p(n)$, we can write the characteristic ODE of (42) as

$$\frac{dc_i}{dt} = -2E[c_i(l)^T z_i(l)(z_i(l) - s_i^T z_i(l) s_i) | p]$$
$$= -2(A_i - s_i s_i^T A_i) c_i. \qquad (43)$$

Since $\frac{dc_i}{dt}$ is linear in c_i, it is easily seen that if c_i^* is a globally asymptotically stable equilibrium for (43), it must be a globally exponentially asymptotically stable equilibrium [24]. Therefore, according to [16, Theorem 2.3], (42) converges to $c_i(n)^*$ also in the MSE sense as long as the step-size sequence satisfies $\sum_{l=0}^{\infty} \mu(l) = \infty$ and $\mu(l) \to 0$.[2]

However, before performing the nth iteration on the stochastic PC (4), in order to ensure $c_i(n) = c_i(n)^*$, one has to perform an infinite number of iterations of (42). This is not acceptable in practical systems. When the number of blind

---

[1]Note that this conclusion only applies to the systems where a fixed single base station is assigned to each user.

[2]Similar results can also be found in [22] and [25].

MMSE iterations between the power updates is finite, we can write $\boldsymbol{c}_i(n)$ as

$$\boldsymbol{c}_i(n) = \boldsymbol{c}_i(n)^* + \boldsymbol{w}_i(n). \qquad (44)$$

Since (42) converges to $\boldsymbol{c}_i(n)^*$ in the MSE sense, we can always choose $\mu(l)$ and the number of iterations such that

$$E\left[\|\boldsymbol{w}_i(n)\|^2 | \boldsymbol{p}(n)\right] \leq \beta(n) K_9 \qquad (45)$$

is satisfied. Here $K_9 > 0$ is an arbitrary constant, and $\beta(n)$ is a sequence as in (9).

Define the stochastic estimate of $I_i(n)$ as

$$\tilde{I}_i(n) = \frac{\gamma_i}{h_{ii}} \left[(\boldsymbol{c}_i(n)^T \boldsymbol{z}_i(n))^2 - p_i(n)h_{ii}\right]. \qquad (46)$$

Then, unlike the previous case where the stochastic version of the PC algorithm was standard, in this case we show that the stochastic version of the PC algorithm is quasi-standard.

Theorem 5: $\check{\boldsymbol{I}}(n)$ in (46) is a quasi-standard stochastic interference function if (44) and (45) are satisfied.

The proof of Theorem 5 is presented in Appendix D.

Theorem 5 shows the convergence of the joint stochastic PC and filter optimization algorithm under the condition that (45) is satisfied by the filter updates between the power updates. Hence, one may still need to perform a large number of iterations of the blind MMSE (42) between two power updates (4) in order to ensure that (45) holds. Obviously, when $\beta(n) \to 0$, the number of steps required on the blind MMSE iteration grows to infinity. Noting that when $\alpha(n)$ is small, $\boldsymbol{c}^*(n+1)$ differs from $\boldsymbol{c}^*(n)$ only slightly, we can initialize the blind MMSE iteration of $\boldsymbol{c}(n+1)$ by $\boldsymbol{c}(n)$ which is an estimate of $\boldsymbol{c}^*(n)$. Based on this basic principle, we propose a modified version of the algorithm, which is similar to the one given in [9], as follows.

Combined stochastic PC and receiver optimization:

1. Initialize the iteration counter $n = 0$. Initialize $\boldsymbol{p}$, $\boldsymbol{c}_i(0) = \boldsymbol{s}_i$.
2. $\forall i$, compute $\tilde{I}_i(n)$ via

$$\tilde{I}_i(n) = \frac{\gamma_i}{h_{ii}} \left[(\boldsymbol{c}_i(n)^T \boldsymbol{z}_i(n))^2 - p_i(n)h_{ii}\right]. \qquad (47)$$

3. $\forall i$, compute $\tilde{\boldsymbol{c}}_i(n)$ via

$$\tilde{\boldsymbol{c}}_i(n) = \boldsymbol{c}_i(n)^T \boldsymbol{z}_i(n)(\boldsymbol{z}_i(n) - \boldsymbol{s}_i^T \boldsymbol{z}_i(n)\boldsymbol{s}_i). \qquad (48)$$

4. $\forall i$, update $p_i(n+1)$ and $\boldsymbol{c}_i(n+1)$ by

$$p_i(n+1) = p_i(n) - \alpha_p(n)(p_i(n) - \tilde{I}_i(n))$$
$$\boldsymbol{c}_i(n+1) = \boldsymbol{c}_i(n) - \alpha_{ci}(n)\tilde{\boldsymbol{c}}_i(n) \qquad (49)$$

where $\alpha_p(n)$, $\alpha_{ci}(n)$ are the step-size sequences for the PC and blind MMSE updates, respectively.

5. As in **[22]**, perform

$$\boldsymbol{c}_i(n+1) = \boldsymbol{c}_i(n+1) - (\boldsymbol{c}_i(n+1)^T \boldsymbol{s}_i - 1)\boldsymbol{s}_i \quad (50)$$

to ensure that $\boldsymbol{c}_i(n+1)^T \boldsymbol{s}_i = 1$ holds.

6. Stop when the power and filter coefficients converge. Otherwise, let $n = n + 1$, and go to step 2. $\qquad \square$

In the above algorithm, we update the power vector and the filter coefficients in parallel, i.e., the PC algorithm does not wait

for the convergence of the blind MMSE. Furthermore, we use different notations on the step-size sequences of the PC and the blind MMSE to indicate that they are not necessarily the same. However, we should clarify that, since parallel update introduces correlation between the noises in the PC and filter optimizations, the convergence proof in Theorems 1 and 4 do not ensure the convergence of this modified algorithm. Nevertheless, we will verify later via computer simulations that, when the step sizes are small enough, the above algorithm is convergent as well.

D. Minimum Power Base Station Assignment With Matched-Filter Receiver

Now assume that there are K_i base stations associated with user i (one base station may be associated to multiple users simultaneously) and that at each base station the receiver uses a matched filter. Denote these base stations as $i_1, i_2, \ldots, i_{K_i}$. The deterministic interference to user i at base station i_k is

$$I_i^{(i_k)}(\boldsymbol{p}) = \frac{\gamma_i}{h_{i_k i}} \left(\sum_{j \neq i_k} p_j h_{i_k j}(\boldsymbol{s}_i^T \boldsymbol{s}_j)^2 + \sigma^2\right). \qquad (51)$$

Define the minimum interference $\boldsymbol{I}(\boldsymbol{p})$ as

$$I_i(\boldsymbol{p}) = \min_{i_k} I_i^{(i_k)}. \qquad (52)$$

It is shown in [7] that $\boldsymbol{I}(\boldsymbol{p})$ is a standard deterministic interference function, and hence the corresponding deterministic PC algorithm converges to the optimal power vector. The base station assignment of (52) is termed the minimum power base station assignment [26].

In the stochastic implementation of the joint PC with minimum power base station assignment, at each base station, we still use (as in (34))

$$\tilde{I}_i^{(i_k)}(\boldsymbol{p}) = \frac{\gamma_i}{h_{i_k i}} y_{i_k}^2 - \gamma_i p_i \qquad (53)$$

as the stochastic estimate of $I_i^{(i_k)}(\boldsymbol{p})$. However, since

$$E[\tilde{I}_i^{(i_k)} | \boldsymbol{p}] = I_i^{(i_k)}$$

as long as the variance of the noise is not zero, we have

$$E[\min_{i_k} \tilde{I}_i^{(i_k)} | \boldsymbol{p}] < \min_{i_k} I_i^{(i_k)}. \qquad (54)$$

Hence, if we use $\min_{i_k} \tilde{I}_i^{(i_k)}$ as a stochastic estimate of I_i, $\beta(n)$ in the bias condition (9) in Definition 3 will not decrease to 0 as $n \to \infty$. To avoid this problem, we assume that, between the successive power updates, we obtain $L(n)$ samples of $\tilde{I}_i^{(i_k)}$, denoted as $\tilde{I}_i^{(i_k)}(1), \ldots, \tilde{I}_i^{(i_k)}(L(n))$. Define

$$\tilde{I}_i = \min_{i_k} \frac{\sum_{j=1}^{L(n)} \tilde{I}_i^{(i_k)}(j)}{L(n)}. \qquad (55)$$

We can see that $E[\tilde{\boldsymbol{I}} | \boldsymbol{p}] \to \boldsymbol{I}$ as $L(n) \to \infty$. It is also easy to verify that $\tilde{\boldsymbol{I}}$ satisfies the growing condition, and that $\boldsymbol{I}$ satisfies the Lipschitz condition. Hence, the stochastic PC algorithm (4) with $\tilde{\boldsymbol{I}}$ in (55) is quasi-standard as long as $L(n)$ is chosen properly.

Similar to the joint stochastic PC with filter optimization, if we require that the bias of the interference estimates go to zero, $L(n) \to \infty$ must be true. Clearly, getting a large number of samples between two power updates will significantly reduce the rate of the power updates. Hence, similar to the problem of joint stochastic PC with filter optimization, we propose a modified version of the algorithm as follows.

Combined stochastic PC and base station assignment:

1. Initialize the iteration counter $n = 0$. Initialize $\boldsymbol{p}(0)$. $\forall i$ and $\forall i_k$, initialize $\tilde{\chi}_i^{(i_k)}(0) = 0$. Initialize $L(0) = \lfloor \frac{K_{10}}{\alpha(0)} \rfloor$, where $K_{10} > 0$ is a carefully chosen constant and $\lfloor \frac{K_{10}}{\alpha(0)} \rfloor$ is the integer part of $\frac{K_{10}}{\alpha(0)}$.

2. $\forall i$ and $\forall i_k$, let $L(n) = \lfloor \frac{K_{10}}{\alpha(n)} \rfloor$. Compute $\tilde{I}_i^{(i_k)}(n)$ via

$$\tilde{I}_i^{(i_k)}(n) = \frac{\gamma_i}{h_{i_k i}} \left[(\boldsymbol{s}_i^T \boldsymbol{z}_i(n))^2 - p_i(n) h_{i_k i} \right]. \quad (56)$$

Compute $\tilde{\chi}_i^{(i_k)}(n)$ by

$$\tilde{\chi}_i^{(i_k)}(n) = \frac{L(n)-1}{L(n)} \tilde{\chi}_i^{(i_k)}(n) + \frac{1}{L(n)} \tilde{I}_i^{(i_k)}(n). \quad (57)$$

3. $\forall i$, find the minimum power base station assignment by

$$a_i = \arg \min_{i_k} \tilde{\chi}_i^{(i_k)}(n). \quad (58)$$

Compute $\tilde{\boldsymbol{I}}$ via

$$\tilde{I}_i(n) = \tilde{I}_i^{(a_i)}(n). \quad (59)$$

4. $\forall i$, update $p_i(n+1)$ by

$$p_i(n+1) = p_i(n) - \alpha(n)(p_i(n) - \tilde{I}_i(n)). \quad (60)$$

$\forall i, i_k$, update $\tilde{\chi}_i^{(i_k)}(n+1) = \tilde{\chi}_i^{(i_k)}(n)$

5. Stop when the power and base station assignments converge. Otherwise, let $n = n+1$, and go to step 2. $\qquad \square$

The convergence of the modified algorithm will be verified later via computer simulations.

E. Limited Diversity With Matched-Filter Receiver

Following the previous subsection, we assume again that there are K_i base stations associated with user i. Instead of choosing the base station with the minimum interference, we combine the signals at all K_i base stations to detect the signal of user i. Such a scenario was analyzed in [27] from the channel capacity point of view. Under the assumption that the interfering signals at different base stations appear to user i as uncorrelated noises, maximal ratio combining of the received signals for user i yields the SIR constraint of user i as [27]

$$p_i \sum_{i_k} \frac{h_{i_k i}}{\sum_{j \neq i} p_j h_{i_k j} (\boldsymbol{s}_i^T \boldsymbol{s}_j)^2 + \sigma^2} \geq \gamma_i. \quad (61)$$

This gives the deterministic interference $\boldsymbol{I}(\boldsymbol{p})$ as

$$I_i(\boldsymbol{p}) = \frac{\gamma_i}{\sum_{i_k} \frac{h_{i_k i}}{\sum_j p_j h_{i_k j} (\boldsymbol{s}_i^T \boldsymbol{s}_j)^2 + \sigma^2}} = \frac{1}{\sum_{i_k} \frac{1}{I_i^{(i_k)}(\boldsymbol{p})}} \quad (62)$$

where $I_i^{(i_k)}(\boldsymbol{p})$ is given by (51). It was shown in [7] that $\boldsymbol{I}(\boldsymbol{p})$ is a standard deterministic interference function, and therefore a deterministic PC based on this interference function converges to the optimal power vector.

Before studying the stochastic version of the PC algorithm, we first show that $\boldsymbol{I}(\boldsymbol{p})$ satisfies the Lipschitz condition as well. In fact

$$\left| \frac{\partial I_i}{\partial p_l} \right| = \gamma_i \frac{\sum_{i_k} \frac{h_{i_k i} h_{i_k l} (\boldsymbol{s}_i^T \boldsymbol{s}_l)^2}{\left(\sum_{j \neq i} p_j h_{i_k j} (\boldsymbol{s}_i^T \boldsymbol{s}_j)^2 + \sigma^2 \right)^2}}{\left(\sum_{i_k} \frac{h_{i_k i}}{\sum_{j \neq i} p_j h_{i_k j} (\boldsymbol{s}_i^T \boldsymbol{s}_j)^2 + \sigma^2} \right)^2}$$

$$\leq \gamma_i \max_{i_k} \left(\frac{h_{i_k l}}{h_{i_k i}} \right) \frac{\sum_{i_k} \left(\frac{h_{i_k i}}{\sum_{j \neq i} p_j h_{i_k j} (\boldsymbol{s}_i^T \boldsymbol{s}_j)^2 + \sigma^2} \right)^2}{\left(\sum_{i_k} \frac{h_{i_k i}}{\sum_{j \neq i} p_j h_{i_k j} (\boldsymbol{s}_i^T \boldsymbol{s}_j)^2 + \sigma^2} \right)^2}$$

$$\leq \gamma_i \max_{i_k} \left(\frac{h_{i_k l}}{h_{i_k i}} \right) \quad (63)$$

that is, all components of $\frac{\partial \boldsymbol{I}}{\partial \boldsymbol{p}}$ are bounded, and the Lipschitz continuity follows immediately.

As in previous examples, we use

$$\tilde{I}_i^{(i_k)} = \frac{\gamma_i}{h_{i_k i}} y_{i_k}^2 - \gamma_i p_i \quad (64)$$

as the stochastic estimate of $I_i^{(i_k)}$. However, we cannot directly replace $I_i^{(i_k)}$ by $\tilde{I}_i^{(i_k)}$ in (62) since in general

$$E \left[\frac{1}{\sum_{i_k} \frac{1}{\tilde{I}_i^{(i_k)}(\boldsymbol{p})}} \middle| \boldsymbol{p} \right] \neq I_i(\boldsymbol{p}). \quad (65)$$

In order to meet the bias condition requirement, we assume that, between the $(n-1)$th and the nth power updates, we fix the powers and receive $L(n)$ samples of $\tilde{I}_i^{(i_k)}$, denoted by $\tilde{I}_i^{(i_k)}(1), \ldots, \tilde{I}_i^{(i_k)}(L(n))$. Define

$$\tilde{\chi}_i^{(i_k)} = \frac{1}{L(n)} \sum_{l=1}^{L(n)} \tilde{I}_i^{(i_k)}(l) \quad (66)$$

and define the stochastic estimate of $\boldsymbol{I}(\boldsymbol{p})$ by

$$\tilde{I}_i(\boldsymbol{p}) = \frac{1}{\sum_{i_k} \frac{1}{\max \left(\tilde{\chi}_i^{(i_k)}(\boldsymbol{p}), \frac{\gamma_i \sigma^2}{h_{i_k i}} \right)}}. \quad (67)$$

It is easily seen that if $L(n) \to \infty$, $E \left[\tilde{I}_i(\boldsymbol{p}) | \boldsymbol{p} \right] \to I_i(\boldsymbol{p})$ holds. Hence, the bias condition is satisfied if $L(n)$ is chosen properly.

Furthermore, since $\max \left(\tilde{\chi}_i^{(i_k)}(\boldsymbol{p}), \frac{\gamma_i \sigma^2}{h_{i_k i}} \right) > 0$, we have

$$\tilde{I}_i(\boldsymbol{p}) = \frac{1}{\sum_{i_k} \frac{1}{\max \left(\tilde{\chi}_i^{(i_k)}(\boldsymbol{p}), \frac{\gamma_i \sigma^2}{h_{i_k i}} \right)}}$$

$$< \frac{1}{K_i} \sum_{i_k} \max \left(\tilde{\chi}_i^{(i_k)}(\boldsymbol{p}), \frac{\gamma_i \sigma^2}{h_{i_k i}} \right). \quad (68)$$

Hence, from the fact that there exists a constant K_{11}, such that

$$E \left[(\tilde{I}_i^{(i_k)})^2 | \boldsymbol{p} \right] \leq K_{11} (1 + \|\boldsymbol{p}\|^2) \quad (69)$$

and

$$\max \left(\tilde{\chi}_i^{(i_k)}, \frac{\gamma_i \sigma^2}{h_{i_k i}} \right)^2 < \left(\tilde{\chi}_i^{(i_k)} \right)^2 + \left(\frac{\gamma_i \sigma^2}{h_{i_k i}} \right)^2. \quad (70)$$

It can be verified that $\tilde{\boldsymbol{I}}(\boldsymbol{p})$ satisfies the growing condition. Therefore, the proposed algorithm is quasi-standard.

In this example, even though there is no other parameter optimization combined with the PC problem, we are only able to design a quasi-standard stochastic algorithm due to the difficulty in satisfying the mean condition in the standard stochastic PC algorithm. Finally, similar to the previous examples, to avoid the problem that $L(n) \to \infty$, we further modify the algorithm as follows.

Combined stochastic PC with limited diversity:

1. Initialize the iteration counter $n = 0$. Initialize $\boldsymbol{p}(0)$. $\forall i$ and $\forall i_k$, initialize $\tilde{\chi}_i^{(i_k)}(0) = 0$. Initialize $L(0) = \lfloor \frac{K_{12}}{\alpha(0)} \rfloor$.
2. $\forall i$ and $\forall i_k$, let $L(n) = \lfloor \frac{K_{12}}{\alpha(n)} \rfloor$. Compute $\tilde{I}_i^{(i_k)}(n)$ via

$$\tilde{I}_i^{(i_k)}(n) = \frac{\gamma_i}{h_{i_k i}} \left[(\boldsymbol{s}_i^T \boldsymbol{z}_i(n))^2 - p_i(n) h_{i_k i} \right]. \qquad (71)$$

Compute $\tilde{\chi}_i^{(i_k)}(n)$ by

$$\tilde{\chi}_i^{(i_k)}(n) = \frac{L(n) - 1}{L(n)} \tilde{\chi}_i^{(i_k)}(n-1) + \frac{1}{L(n)} \tilde{I}_i^{(i_k)}(n). \qquad (72)$$

3. Compute $\tilde{\boldsymbol{I}}(n)$ by

$$\tilde{I}_i(n) = \frac{1}{\sum_{i_k} \frac{1}{\max\left(\tilde{\chi}_i^{(i_k)}(n), \frac{\gamma_i \sigma^2}{h_{i_k i}}\right)}}. \qquad (73)$$

4. $\forall i$, update $p_i(n+1)$ by

$$p_i(n+1) = p_i(n) - \alpha(n)(p_i(n) - \tilde{I}_i(n)). \qquad (74)$$

$\forall i, i_k$, update $\tilde{\chi}_i^{(i_k)}(n+1) = \tilde{\chi}_i^{(i_k)}(n)$.
5. Stop when the powers converge. Otherwise, let $n = n+1$, and go to step 2. $\qquad \square$

V. COMPUTER SIMULATIONS

In this section, we present several computer simulations to verify the convergence of the modified algorithms proposed in Sections IV-C and IV-E. We denote the stochastic PC algorithm with matched-filter receiver as PCMF, the stochastic PC algorithm with blind MMSE receiver as PCMMSE, the stochastic PC algorithm with minimum power base station assignment as PCMP, and the stochastic PC algorithm with limited diversity as PCLD.

Example 1: In this example, we study the performance of PCMMSE proposed in Section IV-C. We choose $K = 4$ and the 5-length binary signature sequences are generated randomly as

$$[\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4] = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \qquad (75)$$

For simplicity, we set all channel gains to $h_{ij} = 1$, $\forall i, j$. The SIR targets of the users are chosen arbitrarily as

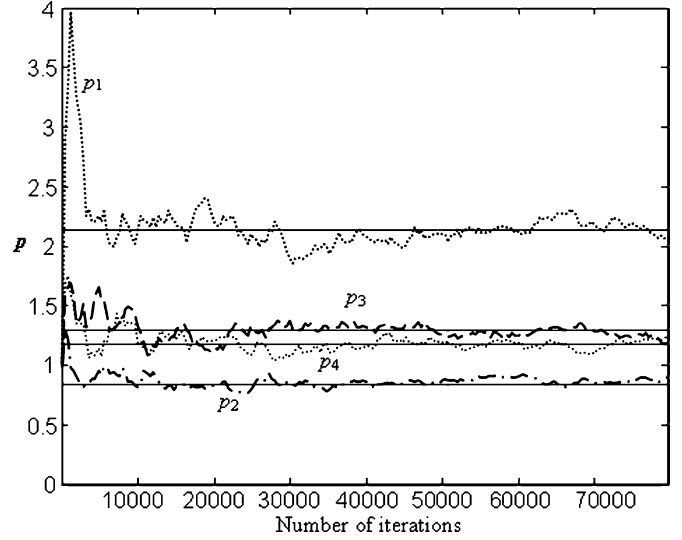$$[\gamma_1, \gamma_3, \gamma_3, \gamma_4] = [5.9, 4.9, 7.6, 6.9]. \qquad (76)$$
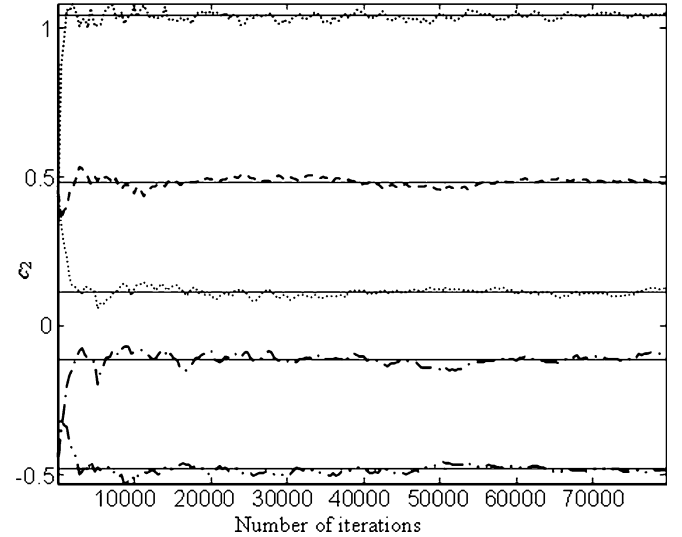

Fig. 1. Convergence of the powers of all users.

Fig. 2. Convergence of the filter coefficients of user 2.

And $\sigma^2 = 0.1$. We initialize all user powers at 1, and initialize the filter coefficients to $\boldsymbol{c}_i(0) = \boldsymbol{s}_i$. The step size sequence is chosen as $\alpha_p(n) = \alpha_{ci}(n) = \frac{10}{10000+n}$ so that the system has a reasonable initial convergence and the step size sequence does not decrease to zero too quickly. Fig. 1 shows the convergence of the transmitted power p_i of each user, and Fig. 2 shows the convergence of the filter coefficients $\boldsymbol{c}_2$ for user 2. The optimal values of the parameters which are obtained from the deterministic iterations are also provided as horizontal lines in these figures.

Example 2: In this example, we study the performances of PCMF, PCMMSE, PCMP, as well as PCLD. We consider a general multicell CDMA system on a rectangular grid. There are 25 base stations with coordinates $(1000i + 500, 1000j + 500)$ for $0 \le i, j \le 4$. We have 400 users, whose positions are randomly and independently generated. Both x and y coordinates of each user are uniformly distributed between $0 \sim 5000$ m. Fig. 3 shows the positions of the users and the base stations with symbols $\times$ and o, respectively.
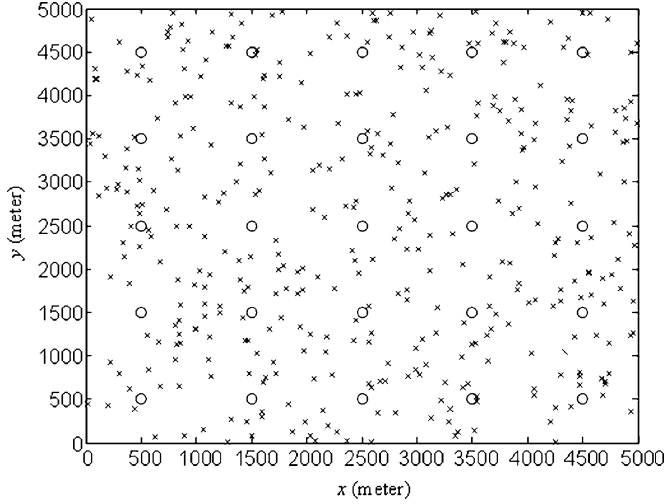
Fig. 3. Simulation environments for 400 users; o and $\times$ are base stations and users, respectively.
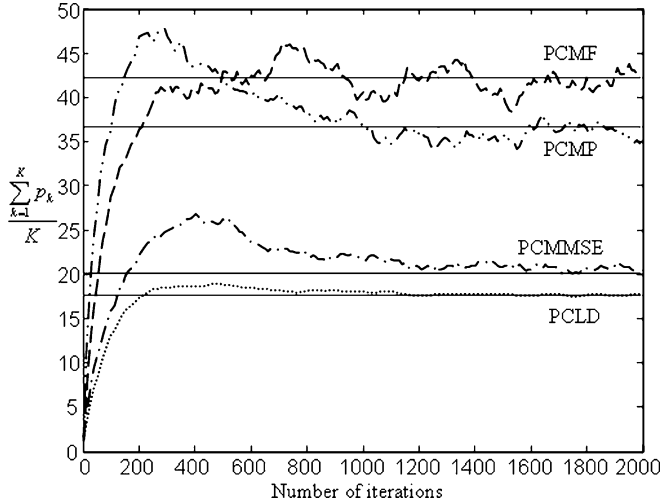

Fig. 4. Performance comparison in terms of average power.

The channel gain of user j to the assigned base station of user i is computed as $h_{ij} = (\frac{100}{d_{ij}})^4$, where d_{ij} is the distance between user j and the assigned base station of user i. The 150-length binary signature sequences are randomly generated. The target SIR is set at $\gamma_i = 4$ for all users. For PCMF and PCMMSE, each user is assigned to its closest base station, and for PCMP and PCLD, we associate the four closest base stations with each user. In this example, we fix the step sizes at $\alpha(n) = 0.01$, and all other settings are the same as in the previous example. In PCMF, we set $L(n)$ at $L(n) = 300$. In PCLD, we set $L(n)$ at $L(n) = 30$.

We first show the performances of different algorithms in terms of the average power in Fig. 4. The optimal solutions obtained via the corresponding deterministic PC algorithms are shown as horizontal lines, respectively. In addition to the convergence of the four algorithms, we can see that, to achieve the same SIR target, the average powers of PCMMSE and PCLD are much lower than those of the other two algorithms.

For both PCMP and PCLD, since $E[\tilde{I}_i^{(i_k)}|\boldsymbol{p}] = I_i^{(i_k)}$, one may think that plugging $\tilde{I}_i^{(i_k)}$ directly into the deterministic interfer-
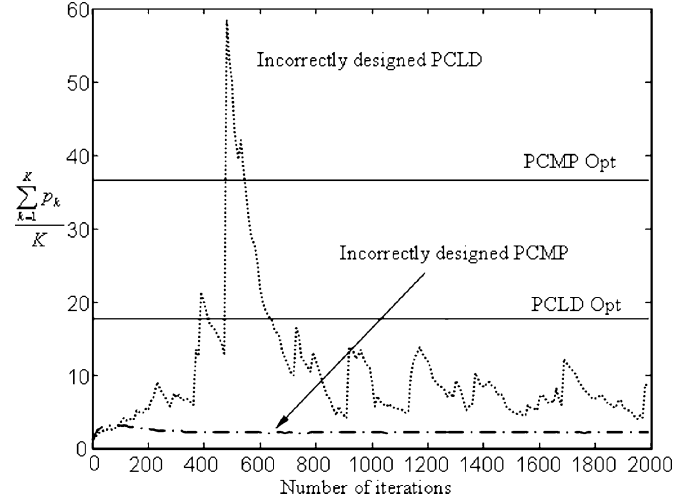

Fig. 5. Performance of incorrectly designed PC algorithms.

ence function should give a reasonably good stochastic estimate. This corresponds to using $\min_{i_k} \tilde{I}_i^{(i_k)}$ as the stochastic estimate of $I_i^{(i_k)}$ in PCMP, and using $\dfrac{1}{\sum_{i_k} \frac{1}{\tilde{I}_i^{(i_k)}}}$ as the stochastic estimate of I_i in PCLD. The trajectories of the average powers of these two algorithms are shown in Fig. 5. As we see in this figure, due to significant bias in the interference estimation, the incorrectly designed PCMP algorithm converges to an infeasible solution; powers converge to small values that violate the SIR requirements. On the other hand, due to the violation of the growing condition, the incorrectly designed PCLD algorithm does not even show a convergent behavior.

VI. CONCLUSION

This paper proposes a general framework for standard and quasi-standard stochastic PC algorithms. Based on the recent literature on the stochastic approximation algorithms, we show that, under certain mild conditions, both standard and quasi-standard algorithms converge to the optimal solution. Different types of convergence are shown under different assumptions on the iteration step-size sequence. Several existing stochastic PC algorithms are studied, and several new stochastic PC algorithms are proposed. We show that these algorithms are either standard or quasi-standard. In the examples of the quasi-standard stochastic PC algorithm, we further extend the algorithms so that the stochastic PC and the parameter optimization can be carried out in parallel. Convergence of the modified algorithms are verified by computer simulations.

APPENDIX A
PROOF OF LEMMA 1

Proof

Given an arbitrary constant $K_a > 1$, since $\boldsymbol{I}(K_a \boldsymbol{p}^*) < K_a \boldsymbol{p}^*$ with strict inequality, we can always find $0 < \epsilon < 1$, such that

$$\boldsymbol{I}(K_a \boldsymbol{p}^*) \leq K_a(1 - \epsilon)\boldsymbol{p}^*. \tag{77}$$

That is,

$$K_a \boldsymbol{p}^* - \boldsymbol{I}(K_a \boldsymbol{p}^*) \geq \epsilon K_a \boldsymbol{p}^*. \tag{78}$$

Furthermore, $\forall K_b \geq K_a$

$$
\begin{aligned}
K_b \boldsymbol{p}^* - \boldsymbol{I}(K_b \boldsymbol{p}^*) &= \frac{K_b}{K_a} K_a \boldsymbol{p}^* - \boldsymbol{I}\left(\frac{K_b}{K_a} K_a \boldsymbol{p}^*\right) \\
&\geq \frac{K_b}{K_a}[K_a \boldsymbol{p}^* - \boldsymbol{I}(K_a \boldsymbol{p}^*)] \\
&\geq \epsilon K_b \boldsymbol{p}^*.
\end{aligned}
\tag{79}
$$

This indicates that, for any standard interference function $\boldsymbol{I}(\boldsymbol{p})$, we can always find an $\epsilon > 0$ and $K_a > 1$ such that (79) is valid for all scaling constants $K_b \geq K_a$. Now, define

$$
\xi_i = \frac{p_i}{p_i^*}, \qquad k = \arg\max_i \xi_i.
\tag{80}
$$

Using (79), if $\xi_k \geq K_a$

$$
\|\boldsymbol{p} - \boldsymbol{I}(\boldsymbol{p})\| \geq p_k - I_k(\boldsymbol{p}) \geq \xi_k p_k^* - I_k(\xi_k \boldsymbol{p}^*) \geq \epsilon \xi_k p_k^* \tag{81}
$$

which gives

$$
\begin{aligned}
\|\boldsymbol{p}\|^2 &\leq \xi_k^2 \|\boldsymbol{p}^*\|^2 = \frac{\|\boldsymbol{p}^*\|^2}{\epsilon^2 \min_j \{p_j^*\}^2} \xi_k^2 \epsilon^2 \min_j \{p_j^*\}^2 \\
&\leq \frac{\|\boldsymbol{p}^*\|^2}{\epsilon^2 \min_j \{p_j^*\}^2} \|\boldsymbol{p} - \boldsymbol{I}(\boldsymbol{p})\|^2 \\
&= K_c \|\boldsymbol{p} - \boldsymbol{I}(\boldsymbol{p})\|^2.
\end{aligned}
\tag{82}
$$

On the other hand, if $\xi_k < K_a$, we have

$$
\|\boldsymbol{p}\|^2 < K_a^2 \|\boldsymbol{p}^*\|^2 = K_d.
\tag{83}
$$

Combing (82) and (83) gives us (16). $\qquad\qquad\square$

APPENDIX B
PROOF OF THEOREM 1

Proof

We give proof for the quasi-standard stochastic PC algorithm only since a standard stochastic PC algorithm is also quasi-standard.

To prove the theorem, we need to work with the Lyapunov function $V(\boldsymbol{p})$ defined in (12). To simplify the notation, we write $\boldsymbol{I}(\boldsymbol{p}(n))$, $\tilde{\boldsymbol{I}}(\boldsymbol{p}(n))$, and $\boldsymbol{g}(\boldsymbol{p}(n))$ as $\boldsymbol{I}(n)$, $\tilde{\boldsymbol{I}}(n)$, and $\boldsymbol{g}(n)$, respectively. Since $\boldsymbol{I}(\boldsymbol{p})$ is Lipschitz continuous, the eigenvalues of the Hessian matrix of $V(\boldsymbol{p})$ are bounded. Hence, denoting $V(\boldsymbol{p}(n))$ by $V(n)$, from (4), we can find a constant K_5, such that the truncated Taylor expansion on $V(n+1)$ satisfies

$$
\begin{aligned}
V(n+1) \leq V(n) &- \alpha(n)(\boldsymbol{p}(n) - \boldsymbol{I}(n))^T(\boldsymbol{p}(n) - \tilde{\boldsymbol{I}}(n)) \\
&+ \frac{\alpha(n)^2 K_5}{2} \|\boldsymbol{p}(n) - \tilde{\boldsymbol{I}}(n)\|^2.
\end{aligned}
\tag{84}
$$

Now, defining $E_n[\cdot]$ as the conditional expectation given $\boldsymbol{p}(n)$, we have

$$
\begin{aligned}
E_n[V(n+1)] \leq V(n) &- \alpha(n)\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2 \\
&+ \alpha(n)(\boldsymbol{p}(n) - \boldsymbol{I}(n))^T \boldsymbol{g}(n) \\
&+ \frac{\alpha(n)^2 K_5}{2} E_n[\|\boldsymbol{p}(n) - \tilde{\boldsymbol{I}}(n)\|^2] \\
\leq V(n) &- \alpha(n)\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2
\end{aligned}
$$

$$
\begin{aligned}
&+ \alpha(n)\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|\|\boldsymbol{g}(n)\| \\
&+ \alpha(n)^2 K_5(\|\boldsymbol{p}(n) - \bar{\boldsymbol{I}}(n)\|^2 + \|\boldsymbol{g}(n)\|^2) \\
&+ \frac{\alpha(n)^2 K_5}{2} E_n[\|\tilde{\boldsymbol{I}}(n) - \boldsymbol{I}(n) - \boldsymbol{g}(n)\|^2].
\end{aligned}
\tag{85}
$$

Using (9) and (11) in Definition 3, and (16) in Lemma 1, we can find a constant K_6, such that

$$
\begin{aligned}
E_n[V(n+1)] \leq V(n) &- \alpha(n)\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2 \\
&+ \alpha(n)^2 K_6(1 + \|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2) \\
&+ \alpha(n)\beta(n) K_6(1 + \|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2) \\
\leq V(n) &+ (\alpha(n)^2 + \alpha(n)\beta(n)) K_6 \\
&- \alpha(n)\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2 \\
&+ K_6 \alpha(n)(\alpha(n) + \beta(n))\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2.
\end{aligned}
\tag{86}
$$

Let us define

$$
\hat{V}(n) = V(n) + K_6 \sum_{i=n}^{\infty}(\alpha(i)^2 + \alpha(i)\beta(i)).
\tag{87}
$$

Then, from (86), we get

$$
\begin{aligned}
E_n[\hat{V}(n+1)] \leq \hat{V}(n) &- \alpha(n)\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2 \\
&+ K_6 \alpha(n)(\alpha(n) + \beta(n))\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2.
\end{aligned}
\tag{88}
$$

Since $\alpha(n), \beta(n) \to 0$ when $n \to \infty$, we can assume that there exists a constant N, such that, $\forall n \geq N$, $1 - K_6(\alpha(n) + \beta(n)) \geq 0$. Therefore,

$$
E_n[\hat{V}(n+1)] \leq \hat{V}(n), \qquad \forall n \geq N.
\tag{89}
$$

Since $\hat{V}(n) \geq 0$, (89) indicates that $\hat{V}(n)$ is a supermartingale sequence [18]. According to martingale convergence theorem [18], $\hat{V}(n)$ converges to a random variable $\hat{V}_\infty$ with probability one. This yields

$$
\begin{aligned}
E[V(N)] &+ K_6 \sum_{i=N}^{\infty}(\alpha(i)^2 + \alpha(i)\beta(i)) \\
&\geq E[V(N)] + K_6 \sum_{i=N}^{\infty}(\alpha(i)^2 + \alpha(i)\beta(i)) - E[\hat{V}_\infty] \\
&\geq \sum_{i=N}^{\infty} \alpha(i)(1 - K_6(\alpha(i) + \beta(i)))E[\|\boldsymbol{p}(i) - \boldsymbol{I}(i)\|^2]
\end{aligned}
\tag{90}
$$

where the last inequality follows from (88). Since $E[V(0)] < \infty$ and N is finite, we have $E[V(N)] < \infty$. Together with the assumption that $\sum_N^{\infty} \alpha(n)^2 < \infty$, $\sum_N^{\infty} \alpha(n)\beta(n) < \infty$, the left-hand side of (90) is finite. If $\hat{V}_\infty > 0$ has a positive probability, then $\|\boldsymbol{p}(n) - \boldsymbol{I}(n)\|^2 > 0$ has a positive probability. This and

$$
\sum_{n=0}^{\infty} \alpha(n) = \infty, \quad \sum_{n=0}^{\infty} \alpha(n)^2 < \infty, \quad \text{and} \quad \sum_0^{\infty} \alpha(n)\beta(n) < \infty
$$

lead to a contradiction. Therefore, $\hat{V}_\infty = 0$ and $\boldsymbol{p}(n) \to \boldsymbol{p}^*$ as $n \to \infty$, with probability one. $\qquad\square$

APPENDIX C
PROOF OF THEOREM 2

Proof

Since $p_r^* = \frac{p^*}{r}$ is an equilibrium of the scaled ODE (20), by taking $r \to \infty$, we can see that $p_\infty^* = 0$ is an equilibrium of (23).

Since $\forall \eta \neq 0$

$$I_\infty(\eta p) = \lim_{r \to \infty} \frac{I(\eta r p)}{r} = \eta \lim_{\eta r \to \infty} \frac{I(\eta r p)}{\eta r} = \eta I_\infty(p). \quad (91)$$

We observe that $I_\infty(p)$ is linear in p. Suppose $\bar{p} \neq 0$ is another equilibrium of (23), i.e., $\bar{p} = I_\infty(\bar{p})$. Define ξ_i and k as

$$\xi_i = \frac{\bar{p}_i}{p_i^*}, \qquad k = \arg \max_i \xi_i. \quad (92)$$

According to (91), since $\forall \eta \neq 0$, $\eta \bar{p}$ is also an equilibrium of (23), without losing the generality, we can assume that $\xi_k > 1$. Hence,

$$I_k(\bar{p}) \leq I_k(\xi_k p^*) < \xi_k p_k^* = \bar{p}_k. \quad (93)$$

This contradicts $\bar{p} = I_\infty(\bar{p})$ since

$$I_\infty(\bar{p}) = \lim_{r \to \infty} \frac{I(r\bar{p})}{r} < I(\bar{p}). \quad (94)$$

It is not difficult to verify that $p = 0$ is a globally stable equilibrium of (23). Therefore, we conclude that Theorem 2 holds. □

APPENDIX D
PROOF OF THEOREM 5

Proof

We first show that the deterministic interference $I(p)$ satisfies the Lipschitz condition. Taking the derivative of $I(p)$, we obtain

$$\frac{\partial I_i}{\partial p_j} = -\frac{\gamma_i}{h_{ii}} \frac{(s_i^T A_i^{-1} s_j)^2}{(s_i^T A_i^{-1} s_i)^2}. \quad (95)$$

From (95), we obtain

$$\frac{(s_i^T A_i^{-1} s_j)^2}{(s_i^T A_i^{-1} s_i)^2} \leq \frac{(s_i^T A_i^{-1} A_i^{-1} s_i)(s_j^T s_j)}{(s_i^T A_i^{-1} s_i)^2} = \frac{s_i^T A_i^{-1} A_i^{-1} s_i}{(s_i^T A_i^{-1} s_i)^2}. \quad (96)$$

Noting that A_i^{-1} is symmetric and positive definite, we define $A_i^{-1} = Q^H \Lambda Q$ as the eigenvalue decomposition of A_i^{-1}; and define λ_{max} as the maximum diagonal element of Λ. Since $A_i \geq \sigma^2 I$, we have $\lambda_{max} \leq \frac{1}{\sigma^2}$. This implies that

$$\frac{s_i^T A_i^{-1} A_i^{-1} s_i}{(s_i^T A_i^{-1} s_i)^2} = \frac{s_i^T Q^H \Lambda \Lambda Q s_i}{(s_i^T Q^H \Lambda Q s_i)^2}$$
$$= \frac{(\sqrt{\Lambda} Q s_i)^H \Lambda (\sqrt{\Lambda} Q s_i)}{\|\sqrt{\Lambda} Q s_i\|^2}$$
$$\leq \lambda_{max} \leq \frac{1}{\sigma^2}. \quad (97)$$

Combined (97) with (96), we can see that $I(p)$ is Lipschitz continuous.

Next, note that $c_i^T s_i = 1$ is satisfied at every step of the blind MMSE iteration (42), we have from (44) that $w_i^T s_i = 0$. Hence, from (40), we obtain $c_i^{*T} A_i w_i = k_i s_i^T w_i = 0$. Since

$$E[\tilde{I}_i | p, c_i] = \frac{\gamma_i}{h_{ii}} \frac{c_i^T A_i c_i}{(c_i^T s_i)^2} = \frac{\gamma_i}{h_{ii}} c_i^T A_i c_i. \quad (98)$$

Using (44), we obtain

$$c_i^T A_i c_i = c_i^{*T} A_i c_i^* + 2 c_i^{*T} A_i w_i + w_i^T A_i w_i$$
$$= \frac{1}{s_i^T A_i^{-1} s_i} + w_i^T A_i w_i$$
$$= \frac{1}{s_i^T A_i^{-1} s_i} + \sum_{j \neq i} (p_j h_{ij} \| s_j^T w_i \|^2 + \sigma^2 \| w_i \|^2)$$
$$\leq \frac{1}{s_i^T A_i^{-1} s_i} + \sum_{j \neq i} (p_j h_{ij} + \sigma^2) \| w_i \|^2. \quad (99)$$

Define

$$g(p(n)) = E[\tilde{I}(n) | p(n)] - I(n). \quad (100)$$

From (45) and (99), we see that $g(p(n))$ satisfies the bias condition in Definition 3.

Finally, using (29), (44), and (45), we can verify that $\tilde{I}(n)$ also satisfies the growing condition in Definition 3. This completes the proof that the stochastic interference function is quasi-standard. Then the convergence of the joint stochastic PC with blind MMSE interference suppression follows from Theorems 1 and 4. □

REFERENCES

[1] J. Luo, S. Ulukus, and A. Ephremides, "Probability one convergence in joint stochastic power control and blind MMSE interference suppression," in *Proc. Conf. Information Science and Systems (CISS 2003)*, Baltimore, MD, Mar. 2003.

[2] S. Grandhi, R. Vijayan, D. Goodman, and J. Zander, "Centralized power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 42, no. 6, pp. 466–468, Nov. 1993.

[3] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, no. 1, pp. 57–62, Feb. 1992.

[4] D. O'Neill, M. Chiang, D. Julian, and S. Boyd, "Optimal resource allocation with QoS constraints in wireless cellular and ad hoc networks," Elec. Eng. Dept., Stanford Univ., Stanford, CA, Working paper, 2002.

[5] G. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, Nov. 1993.

[6] S. Chen, N. Bambos, and G. Pottie, "On distributed power control for radio networks," in *Proc. IEEE Int. Conf. Communications*, New Orleans, LA, May 1994, pp. 1281–1285.

[7] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 5, pp. 1341–1347, Sep. 1995.

[8] S. Ulukus and R. Yates, "Stochastic power control for cellular radio systems," *IEEE Trans. Commun.*, vol. 46, no. 6, pp. 784–798, Jun. 1998.

[9] M. Varanasi, "Nonlinear multiuser receivers with distributed power control in cellular radio networks," in *Proc. 37th Annu. Allerton Conf. Communication, Control and Computers*, Allerton, IL, Sep. 1999, pp. 820–830.

[10] M. Varanasi and D. Das, "Fast stochastic power control algorithms for nonlinear multiuser receivers," *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1817–1827, Nov. 2002.

[11] S. Ulukus and R. Yates, "Adaptive power control and MMSE interference suppression," *ACM Wireless Netw.*, vol. 4, pp. 489–496, Nov. 1998.

[12] D. Das and M. Varanasi, "Blind algorithms for joint optimization of multiuser receiver and power control," in *Proc. 38th Annu. Allerton Conf. Communication, Control and Computers*, Monticello, IL, Oct. 2000.

[13] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, pp. 400–407, 1951.

[14] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Statist.*, vol. 23, pp. 462–466, 1952.

[15] H. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*. New York: Springer-Verlag, 1997.

[16] V. Borkar and S. Meyn, "The O.D.E. method for convergence of stochastic approximation and reinforcement learning," *SIAM J. Control Optim.*, vol. 38, pp. 447–469, Jan. 2000.

[17] D. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.

[18] S. Karlin and H. Taylor, *A First Course in Stochastic Processes*, 2nd ed. San Diego, CA: Academic, 1975.

[19] A. Yener, R. Yates, and S. Ulukus, "Interference management for CDMA systems through power control, multiuser detection and beamforming," *IEEE Trans. Commun.*, vol. 49, no. 7, pp. 1227–1239, Jul. 2001.

[20] ——, "Combined multiuser detection and beamforming for CDMA systems: Filter structures," *IEEE Trans. Veh, Technol.*, vol. 51, no. 5, pp. 1087–1095, Sep. 2002.

[21] S. Verdú, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.

[22] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 944–960, Jul. 1995.

[23] L. Gyorfi, "Adaptive linear procedures under general conditions," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 2, pp. 262–267, Mar. 1984.

[24] T. Caraballo, "Asymptotic exponential stability of stochastic partial differential equations with delay," *Stochastics and Stochastics Reps.*, vol. 33, pp. 27–47, 1990.

[25] D. Das and M. Varanasi, "Blind adaptive multiuser detection for cellular systems using stochastic approximation with averaging," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 2, pp. 310–319, Feb. 2002.

[26] R. Yates and C. Huang, "Integrated power control and base station assignment," *IEEE Trans. Veh. Technol.*, vol. 44, no. 3, pp. 638–644, Aug. 1995.

[27] S. Hanly, "Information capacity of radio networks," Ph.D. dissertation, Cambridge Univ., Cambridge, U.K., 1993.