# Stochastic approximation methods-Powerful tools for simulation and optimization: A survey of some recent work on multi-agent systems and cyber-physical systems

George Yin, Le Yi Wang, and Hongwei Zhang

Citation: AIP Conference Proceedings **1637**, 1263 (2014); doi: 10.1063/1.4907291
View online: http://dx.doi.org/10.1063/1.4907291
View Table of Contents: http://scitation.aip.org/content/aip/proceeding/aipcp/1637?ver=pdfcov
Published by the AIP Publishing

# Stochastic Approximation Methods–Powerful Tools for Simulation and Optimization: A Survey of Some Recent Work on Multi-Agent Systems and Cyber-Physical Systems

George Yin*, Le Yi Wang† and Hongwei Zhang**

*Department of Mathematics, Wayne State University, Detroit, MI 48202, USA
†Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA
**Department Computer Science, Wayne State University, Detroit, MI 48202, USA

**Abstract.** Stochastic approximation methods have found extensive and diversified applications. Recent emergence of networked systems and cyber-physical systems has generated renewed interest in advancing stochastic approximation into a general framework to support algorithm development for information processing and decisions in such systems. This paper presents a survey on some recent developments in stochastic approximation methods and their applications. Using connected vehicles in platoon formation and coordination as a platform, we highlight some traditional and new methodologies of stochastic approximation algorithms and explain how they can be used to capture essential features in networked systems. Distinct features of networked systems with randomly switching topologies, dynamically evolving parameters, and unknown delays are presented, and control strategies are provided.

**Keywords:** Stochastic approximation, stochastic optimization, simulation, cyber-physical system.
**PACS:** 02.50.Ey Stochastic processes, 02.50.Fz Stochastic analysis, 02.60.Cb Numerical simulation, 02.70.Uu Applications of Monte Carlo methods

## INTRODUCTION

We are entering a new era in which systems are increasingly high-dimensional, complex in structure, networked, time varying, and subject to different types of uncertainties. Such systems cannot be represented by stand-alone systems, and their control strategies and optimization cannot be obtained by deterministic root finding and extreme seeking algorithms for simple functions. There are pressing needs for designing fast, adaptive, and on-line procedures to develop strategies under stochastic environments. Stochastic approximation methodologies have emerged as a promising general framework in this pursuit.

It is well known that stochastic approximation methods are capable of treating a wide variety of estimation and optimization problems in which the precise forms of the functions are either not known or too complex to evaluate. Instead, one can use available information obtained from noisy observations to obtain approximate solutions. As a powerful tool, the methods of stochastic approximation (SA) was introduced in 1951; see [1]. In the original work, Robbins and Monro aimed for finding the root of a nonlinear real-valued function. Although it appeared to be similar to classical root finding problems for which many numerical algorithms exist, SA is different from the usual setup in that the precise form of the function is unknown but only noisy corrupted observations or measurements are available. A year later, Kiefer and Wolfowitz [2] proposed another algorithm in which in lieu of finding the zero of the function, the purpose was to find the minimizer of a real-valued function. Since the initiation of the methods of SA, there have been enormous literature developed with numerous applications. One of the most up-to-date treatment of SA is the work of Kushner and Yin [3]. Much relaxed conditions compared to the initial setup are used. Sophisticated mathematical tools based on stochastic analysis and dynamic systems have been developed. Complex constraints as well as set-valued systems (in terms of differential inclusion) are presented. Nowadays, stochastic approximation methods have enjoyed a wide range of applications in diverse fields, ranging from medical applications, production planning and flexible manufacturing systems, to financial engineering, learning and adaptive optimization, and networked systems.

The main purpose of the current paper is to review some recent developments of SA methodologies in consensus formation, networked systems, and cyber-physical systems. Using connected vehicles in platoon formation and coordination as a platform, we highlight some traditional and new methodologies of stochastic approximation algorithms and explain how they can be used to capture essential features in networked systems. Distinct features of networked systems with randomly switching topologies, dynamically evolving parameters, and unknown delays are presented,

*10th International Conference on Mathematical Problems in Engineering, Aerospace and Sciences*
AIP Conf. Proc. 1637, 1263-1272 (2014); doi: 10.1063/1.4907291

and control strategies are provided. Although some specific examples are considered in this paper, we hope that the methods and results will shed new lights on treating other systems. The rest of the paper is arranged as follows. First, we present the basic setup of stochastic approximation problems. In addition to the basic algorithms, some asymptotic results are also summarized. Then we present several classes of current interests. Due to page limitation, detailed technical developments are omitted. However, appropriate references are provided for additional reading.

## BASIC ALGORITHMS

This section presents the basic setup of stochastic approximation algorithms and some of their variants. We begin with the Robbins-Monro algorithm that aims at finding the zeros of a nonlinear function. Let $f : \mathbb{R}^r \mapsto \mathbb{R}^r$ be a continuous function. Suppose that we want to find $f(x) = 0$, but only noisy measurements

$$y_n = f(x_n) + \xi_n$$

are available, where $\{\xi_n\}$ is a sequence of random noise. Note that $n$ is a positive integer representing the number of measurements or observations up to this moment. For convenience, it is often thought as a "discrete time." The basic setup of the stochastic approximation algorithms proposed by Robbins and Monro takes the form

$$x_{n+1} = x_n + a_n y_n, \tag{1}$$

where $\{a_n\}$ is a sequence of nonnegative real numbers known as step sizes or gains such that $\sum_n a_n = \infty$ and $a_n \to 0$ as $n \to \infty$. The conditions on the step sizes indicate that they cannot be too big, otherwise the noise effect cannot be suppressed.

However, they cannot be too small either. If they are too small (i.e., $\sum_n a_n < \infty$), then the iterates produced may fail to converge to the desired value. To see this, take the noise-free case $\xi_n = 0$ and suppose that $f(\cdot)$ is a bounded function. Then[1]

$$\sum_{j=0}^{\infty} |x_{j+1} - x_j| \le \sum_{j=0}^{\infty} a_j |f(x_j)| \le K.$$

The above argument indicates that $\sum_j (x_{j+1} - x_j)$ converges absolutely. However,

$$\sum_{j=0}^{n} (x_{j+1} - x_j) = x_{n+1} - x_0$$

by telescoping. As a result, $x_n \not\to x^*$, where $x^*$ is the true parameter we are approximating, unless $x_0$ is sufficiently close to $x^*$.

In 1952, Kiefer and Wolfowitz proposed another class of stochastic approximation algorithms to locate the minima or maxima of a real-valued function. Suppose that we want to minimize a function $f(x)$, but only noisy observations $F(x,\zeta)$ are available. Suppose that $EF(x,\zeta) = f(x)$, but we know neither the form of $F(\cdot)$ nor that of $f(\cdot)$. To approximate/estimate the optimizer, we use the finite difference approximation to the gradient of $f(x)$. Denote the finite difference interval by $\{c_n\}$ (with $c_n \to 0$ as $n \to \infty$). Use $x_n$ to denote the $n$th estimate of the minimum. Suppose that for each $i$ and each $n$, we can observe

$$y_{n,i} = -\frac{F(x_n + c_n e_i, \zeta_n^+) - F(x_n - c_n, \zeta_n^-)}{2c_n},$$

where $\zeta_n^\pm$ are random noise. Denote $y_n = (y_{n,1}, \ldots, y_{n,r})$. Then the approximation algorithm is again given by $x_{n+1} = x_n + a_n y_n$, which is the same form as that of (1). By introducing

$$\xi_n = (\xi_{n,1}, \ldots, \xi_{n,r}),\ \beta_n = (\beta_{n,1}, \ldots, \beta_{n,r}),$$

[1] Here and throughout the paper, $K > 0$ is used as a generic constant; its value may change for different usage. Thus by our convention, $K + K = K$ and $KK = K$.

with

$$\xi_{n,i} = [f(x_n + c_n e_i) - F(x_n + c_n e_i, \zeta^+_{n,i})] - [f(x_n - c_n e_i) - F(x_n - c_n e_i, \zeta^-_{n,i})]$$
$$\beta_{n,i} = f_{x_i}(x_n) - \frac{f(x_n + c_n e_i) - f(x_n - c_n e_i)}{2c_n},$$

where $f_{x_i}(\cdot)$ and $f_x(\cdot)$ denote the partial derivative with respect to $x_i$ and gradient of $f(\cdot)$ w.r.t. $x$, respectively. Now the above algorithm can be rewritten as

$$x_{n+1} = x_n - a_n f_x(x_n) + a_n \frac{\xi_n}{2c_n} + a_n \beta_n. \tag{2}$$

In (2), $\xi_n$ represents the noise and $\beta_n$ denotes the bias. In lieu of the two-sided finite difference, one-sided finite difference can also be used. However, in practice, the two-sided finite difference is often more preferable since it has a smaller bias. This is easily seen by taking a Taylor expansion of the finite difference quotient in $\beta_n$.

Extending the algorithms mentioned above, we may consider algorithms of the form

$$x_{n+1} = x_n + a_n f_n(x_n, \xi_n). \tag{3}$$

As can be seen in (3), instead of a fixed function $f(\cdot)$, a time-varying function $f_n(\cdot)$ can be treated. The noise may appear in a non-additive way. Moreover, to be able to track slight parameter variation, one often uses an algorithm with constant step size of the form

$$x_{n+1} = x_n + \varepsilon f_n(x_n, \xi_n), \tag{4}$$

where $\varepsilon > 0$ is a small parameter. Either (3) or (4) is often used in conjunction with tracking analysis of time-varying parameters.

In another scenario, the values of $\{x_n\}$ are generated externally, but not by the experimenter. Suppose one still wants to find the zero of the function $f(\cdot)$, where $y_n = f(x_n, \xi_n)$. Then one can combine the stochastic approximation methods with nonparametric kernel estimation procedures and to approximate the root of equation $\overline{f}(x) = 0$ by another sequence $\{z_n\}$ according to

$$z_{n+1} = z_n + \frac{a_n}{h_n^r} \kappa\left(\frac{x_n - z_n}{h_n}\right) y_n, \tag{5}$$

where $\kappa(\cdot)$ is a kernel function, $a_n$ is the step size, and $h_n$ represents the window width. The kernel is crucial. If $x_n$ and $z_n$ are far away, $\kappa((x_n - z_n)/h_n)$ will be small and the measurement $y_n$ has little effect on the iteration. If $x_n$ and $z_n$ are close, a non-trivial amount will be added similar to the usual stochastic approximation.

During the past several decades, these stochastic approximation algorithms have been the main focus. Great attention has been devoted to proving convergence and rates of convergence and corresponding properties of the recursively defined stochastic algorithms. Now, we have rather comprehensive understanding and good techniques to treat such stochastic approximation problems.

## CYBER-PHYSICAL SYSTEMS AND CONSENSUS ISSUES

In this section, we first present a problem arising from cyber-physical systems (CPS). Then we study several algorithms that can be used to treat the problems in CPS. These algorithms involve new features beyond the traditional stochastic approximation framework.

A cyber-physical system (CPS) integrates control, communication, and computational systems with physical entities. Owing to its importance, significant research effort has been devoted to improving the link between information-processing and physical elements, and to increasing the adaptability, autonomy, efficiency, functionality, reliability, safety, and usability of cyber-physical systems. To illustrate, in this paper, we consider platoon formation of highway vehicles, which is a critical foundation to support autonomous or semi-autonomous vehicle control for enhanced safety, improved highway utility, increased fuel economy, and reduced emission toward intelligent transportation systems. Such problems pose great challenges from vehicle control, communications, coordinated control, and uncertainties. Starting from the model introduced in [4] for coordinated control of platoons using integrated network consensus decisions and vehicle control, we consider the problem without constraint for ease of presentation. To achieve suitable deployment of the team vehicles based on terrain and environmental conditions, we use the emerging technology of network consensus.

Let us begin with a longitudinal platoon control problem, see Fig. 1. There are $r+1$ vehicles driving in the same lane forming a platoon. The leading vehicle is regarded as a reference, whose position $p^0$ is used as the origin of the line coordinate (without loss of generality, assume $p^0 \equiv 0$), and its speed $v^0$ is the reference speed for the rest of the vehicles in the platoon to follow. Coordination of vehicle control aims to sustain a platoon formation, avoid collision, adjust the formation according to weather and road conditions, converge fast to a new formation after disturbances, reconfigure a formation after vehicle addition and departure. Therefore, inter-vehicle distances are the variables to be controlled. The following diagram serves as a demonstration.

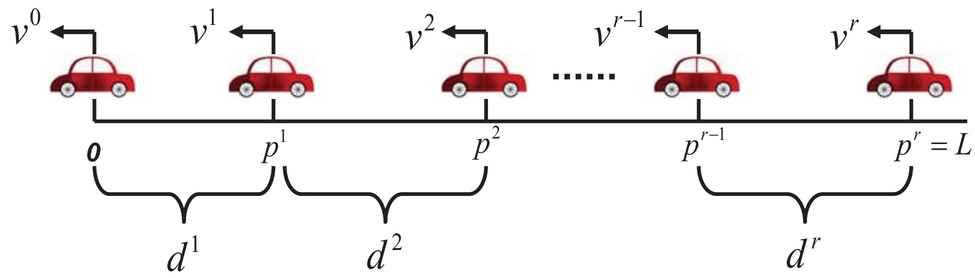


**FIGURE 1.** Platoon coordinates

We work with a discrete time $n$. In the platoon formation, the position of each vehicle is defined by the central point of its length and denoted by $p_n^j$, $j = 1,\ldots,r$, which is the distance of the $j$th vehicle to the leading vehicle. The vehicle speed will be denoted by $v_n^j \geq 0$, $j = 1,\ldots,r$. Let the inter-vehicle distances be defined as

$$d_n^j = p_n^j - p_n^{j-1},\ j = 1,\ldots,r.$$

The leading vehicle's speed $v_n^0$ at time $n$ is the speed target for all the other vehicles in the platoon to follow. Also, a desired distance $\beta$ between consecutive vehicles is a goal that balances efficiency and safety. In principle, $\beta$ is a function of weather, road condition, platoon traveling speed, terrain composition (uphill or downhill), and road curvatures, and consequently changes with time.

**Definition 1** A platoon is said to be in consensus in weakly (or in mean squares) if

$$v_n^j = v_n^0, \text{ and } d_n^j = \beta_n,\ j = 1,\ldots,r.$$

Denote $d_n = [d_n^1,\ldots,d_n^r]'$ and $v_n = [v_n^1,\ldots,v_n^r]'$, and consensus errors

$$e_n = \begin{bmatrix} e_n^1 \\ \vdots \\ e_n^r \end{bmatrix} = d_n - \beta \mathbb{1} \text{ and } \varepsilon_n = \begin{bmatrix} \varepsilon_n^1 \\ \vdots \\ \varepsilon_n^r \end{bmatrix} = v_n - v_n^0 \mathbb{1},$$

where $\mathbb{1} = [1,\ldots,1]'$. Starting at $t = 0$ with initial condition $e(0)$ and $\varepsilon(0)$, the goal of consensus control is to achieve convergence

$$e_n \to 0,\ \ \varepsilon_n \to 0 \text{ as } n \to \infty$$

in probability (or in mean squares, or with probability one). It is noted that to accommodate the time-varying environment, convergence speeds are of interest also.

The basic scheme of platoon formation employs a sensor-based network topology, in which a vehicle uses sensors to measure its own speed and relative distance to the vehicle ahead of it. As a result, $v^{j-1}$, $v^j$, $d^j$ are available to the $j$th vehicle in its control strategies. We demonstrate this sensor-based inter-vehicle information by a string topology shown in Figure 2. Nevertheless, inter-vehicle wireless communications allow enhanced information exchange among vehicles. Figure 3 indicates a more advanced inter-vehicle communication, in which the $j$th vehicle receives not only the parameters from the $(j-1)$th vehicle by sensors, but also the information from $(j-2)$th vehicle using wireless communications.

To mathematically state the problem, in this paper, we start with a simplest setting first. Then in the subsequent sections, we extend this simple formulation to include topology switching, possible non-trial delays, and asynchronous

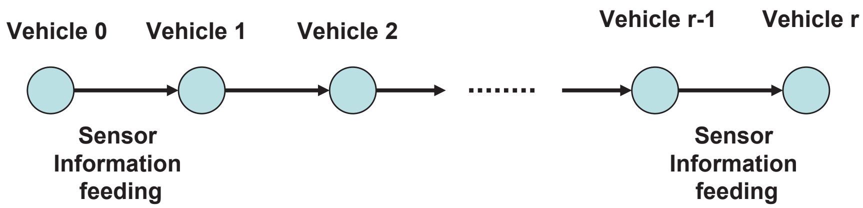


**FIGURE 2.** Sensor-based inter-vehicle communication networks

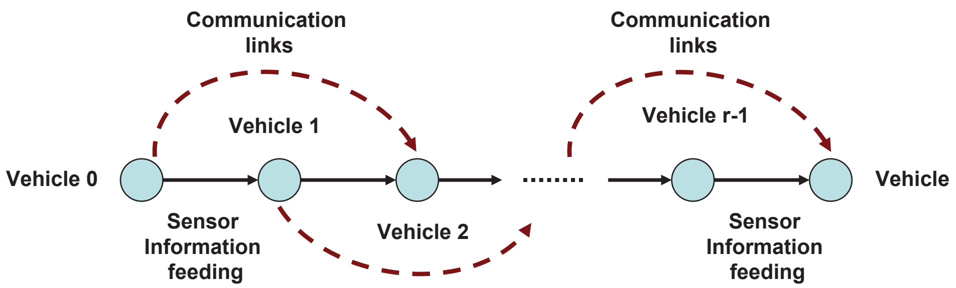


**FIGURE 3.** Information network topologies using inter-vehicle communications

computation and communication times. We begin with a team consisting of $r$ vehicles. Note that in our work in [4], there is an additional constraint on the states and also weights are allowed to represent variations among the vehicles. For clarity in conveying the key ideas, we disregard the constraint and weights throughout the paper.

The node control $u_n^i$ is determined by the link control $v_n^{ij}$. Since a positive transportation of quantity $v_n^{ij}$ on $(i,j)$ means a loss of $v_n^{ij}$ at node $i$ and a gain of $v_n^{ij}$ at node $j$, the node control at node $i$ is $u_n^i = -\sum_{(i,j)\in\mathscr{G}} v_n^{ij} + \sum_{(j,i)\in\mathscr{G}} v_n^{ji}$. The most relevant implication in this control scheme is that for all $n$, $\sum_{i=1}^r x_n^i = \sum_{i=1}^r x_0^i := \eta r$, for some $\eta \in \mathbb{R}$ that is the average of $x_0$. That is, $\eta = \sum_{i=1}^r x_0^i / r$. Consensus control seeks control algorithms that achieve $x_n \to \eta \mathbb{1}$, where $\mathbb{1}$ is the column vector of all 1s. A link $(i,j) \in \mathscr{G}$ entails an estimate, denoted by $\widehat{x}_n^{ij}$, of $x_n^j$ by node $i$ with estimation error $d_n^{ij}$, i.e.,

$$\widehat{x}_n^{ij} = x_n^j + d_n^{ij}. \tag{6}$$

The estimation error $d_n^{ij}$ is usually a function of the signal $x_n^j$ itself and depends on communication channel noises $\xi_n^{ij}$ in a nonadditive and nonlinear relation

$$d_n^{ij} = g(x_n^j, \xi_n^{ij}) \tag{7}$$

and can be spatially and temporally dependent. Most existing literature considers much simplified noise classes $d_n^{ij} = \xi_n^{ij}$ with i.i.d. assumptions.

A sampled and quantized signal $x$ in a networked system enters a communication transmitter as a source. To enhance channel efficiency and reduce noise effects, source symbols are encoded [5, 6]. Typical block or convolutional coding schemes such as Hamming, Reed-Solomon, or more recently the low-density parity-check (LDPC) code and Turbo code, often introduce a nonlinear mapping $v = f_1(x)$. The code word $v$ is then modulated into a waveform $s = f_2(v) = f_2(f_1(x))$ which is then transmitted. Even when the channel noise is additive, namely the received waveform is $w = s + d$ where $d$ is the channel noise, after the reverse process of demodulation and decoding, we have $y = g(w) = g(s+d) = g(f_2(f_1(x)) + d)$. As a result, the error term $g(f_2(f_1(x)) + d) - x$ in general is nonadditive and signal dependent. In addition, block and convolution coding schemes introduce temporally dependent noises. In our formulation, this aspect is reflected in dependent $\phi$-mixing noises on $\xi_n^{ij}$. These will be detailed later.

For simplification on system derivations, we use first $d_n^{ij} = \xi_n^{ij}$ in this section. Let $\widetilde{\eta}_n$ and $\xi_n$ be the $l_s$ dimensional vectors that contain all $\widehat{x}_n^{ij}$ and $\xi_n^{ij}$ in a selected order, respectively. Then, (6) can be written as $\widetilde{\eta}_n = H_1 x_n + \xi_n$, where $H_1$ is an $l_s \times r$ matrix whose rows are elementary vectors such that if the $\ell$th element of $\widetilde{\zeta}_n$ is $\widehat{x}^{ij}$ then the $\ell$th row in $H_1$ is the row vector of all zeros except for a "1" at the $j$th position. Each sensing link provides information $\delta_n^{ij} = x_n^i - \widehat{x}_n^{ij}$, an estimated difference between $x_n^i$ and $x_n^j$. This information may be represented, in the same arrangement as $\widetilde{\eta}_n$, by a vector $\delta_n$ of size $l_s$ containing all $\delta_n^{ij}$ in the same order as $\widetilde{\eta}_n$. $\delta_n$ can be written as $\delta_n = H_2 x_n - \widetilde{\eta}_n = H_2 x_n - H_1 x_n - \xi_n = H x_n - \xi_n$, where $H_2$ is an $l_s \times r$ matrix whose rows are elementary vectors such that if the $\ell$th element of $\widetilde{\zeta}(k)$ is $\widehat{x}^{ij}$ then the $\ell$th row in $H_2$ is the row vector of all zeros except for a "1" at the $i$th position, and $H = H_2 - H_1$. The reader is referred to [7] for basic matrix properties in graphs and to [8] for matrix iterative schemes. Due to network constraints, the information $\delta_n^{ij}$ can only be used by nodes $i$ and $j$. When the control is linear, time invariant, and memoryless, we have $v_n^{ij} = \mu g_{ij} \delta_n^{ij}$ where $g_{ij}$ is the link control gain on $(i,j)$ and $\mu$ is a global scaling factor that will be used in state updating algorithms as the recursive stepsize. Let $G$ be the $l_s \times l_s$ diagonal matrix that has $g_{ij}$ as its diagonal element. In this case, the node control becomes $u_n = -\mu H' G \delta_n$. For convergence analysis, we note that $\mu$ is a global control variable and we may represent $u_n$ equivalently as $u_n = -\mu(H'GHx_n - H'G\xi_n) = \mu(Mx_n + W\xi_n)$, with $M = -H'GH$ and $W = H'G$.

The following assumption is imposed on the network.

(1) All link gains are positive, $g_{ij} > 0$.

(2) $\mathscr{G}$ is strongly connected. Recall that a directed graph is called strongly connected if there is a path from each node in the graph to every other node.

We consider the state updating algorithm

$$x_{n+1} = x_n + \mu M x_n + \mu W d_n, \tag{8}$$

together with the constraint

$$\mathbb{1}' x_n = \eta r, \tag{9}$$

where $\mu > 0$ is a small stepsize, $M$ is a generator of a continuous-time Markov chain so $M\mathbb{1} = 0$, and $\{d_n\}$ is a noise sequence. Now. (8) together with (9) becomes an algorithm of consensus type; see the history and related references in [14, 15, 16, 17].

## Random Switching Topology

In lieu of the well-known consensus-type algorithms (8) and (9), we suppose the network topology evolves according to a random process. The rational is that the platoon formation is usual subject to certain random environment influence such as the weather, road conditions, traffic intensity changes with respect to time, etc. We illustrate the idea in [9] below. Suppose that $\alpha_n$ is a discrete-time Markov chain and the network topology $\mathscr{G}(\alpha_n)$ depends on $\alpha_n$. The Markov chain is used to model, for example, interrupts and rerouting of communication channels. At a given instance $n$, if $\alpha_n = i$, then $\mathscr{G}(\alpha_n) = \mathscr{G}(i)$, namely the topology switches according to the values of $\alpha_n$. To include topology switching and the extended noise class, the network states are updated according to

$$x_{n+1} = x_n + \mu M(\alpha_n) x_n + \mu \widetilde{W}(x_n, \alpha_n, \widetilde{\xi}_n), \tag{10}$$

where $\mu > 0$ is the step size of consensus control. For each $i \in \mathscr{M}$, $M(i)$ is a generator of a continuous-time Markov chain. The noise term $\widetilde{W}(\cdot,\cdot,\cdot) : \mathbb{R}^r \times \mathscr{M} \times \mathbb{R}^r \mapsto \mathbb{R}^r$ is allowed to have the following general structure

$$\widetilde{W}(x, i, \widetilde{\xi}) = W(i)\xi + \widehat{W}(x, i, \zeta), \quad \text{for each } x \in \mathbb{R}^r \text{ and } i \in \mathscr{M}. \tag{11}$$

When $W(i)$ is constant and $\widetilde{W} \equiv 0$, (11) is reduced to the standard additive noise without state dependence. Here, (11) includes state dependence, and nonadditive noises. We consider more general noise, which are necessary when we deal with networked systems. The nonadditive portion is a general nonlinear function of the analog state $x$, the Markov chain state $i \in \mathscr{M}$, as well as the noise source $\zeta$. To state more explicitly dependence on $\xi_n$ and $\zeta_n$, in lieu of using the notation $\widetilde{\xi}_n$, we rewrite the algorithm as

$$x_{n+1} = x_n + \mu M(\alpha_n) x_n + \mu W(\alpha_n)\xi_n + \mu \widehat{W}(x_n, \alpha_n, \zeta_n) \tag{12}$$

in what follows.

Depending on their relative values, three distinct scenarios emerge. Under suitable conditions, we show that when $0 < \varepsilon = O(\mu)$, a continuous-time interpolation of the iterates converges weakly to a system of randomly varying ordinary differential equations modulated by a continuous-time Markov chain. In this case a scaled sequence of tracking errors converges to a system of switching diffusion. When $0 < \varepsilon \ll \mu$, the network topology is almost non-switching during consensus control transient intervals, and hence the limit dynamic system is simply an autonomous differential equation. When $\mu \ll \varepsilon$, the Markov chain acts as a fast varying noise, and only its averaged network matrices are relevant, which results in a limit differential equation that is an average with respect to the stationary measure of the Markov chain. Simulation results are presented to demonstrate these findings.

To give some insight on the limit dynamics, we consider one case $\varepsilon = O(\mu)$. For simplicity, simply assume that $\varepsilon = \mu$. We also assume that the following conditions hold.

(H1) The observation noise $\{\xi_n\}$ is a sequence of stationary $\phi$-mixing sequence such that $E\xi_n = 0$, $E|\xi_n|^{2+\Delta} < \infty$ for some $\Delta > 0$, and that the mixing measure $\widetilde{\phi}_n$ satisfies

$$\sum_{k=0}^{\infty} \widetilde{\phi}_n^{\Delta/(1+\Delta)} < \infty, \tag{13}$$

where $\widetilde{\phi}_n = \sup_{A \in \mathscr{F}^{n+m}} E^{(1+\Delta)/(2+\Delta)} |P(A|\mathscr{F}_m) - P(A)|^{(2+\Delta)/(1+\Delta)}$, $\mathscr{F}_n = \sigma\{\xi_n; k < n\}$, $\mathscr{F}^n = \sigma\{\xi_n; k \geq n\}$.

(H2) Assume the following conditions.

(a) $\alpha_n$ is a discrete-time Markov chain with a finite state space $\mathscr{M} = \{1, \ldots, m\}$ representing the random environment and other random factors. The transition probability matrix of $\alpha_n$ is given by

$$P^{\varepsilon} = I + \varepsilon Q, \tag{14}$$

where $\varepsilon > 0$ is a small parameter, $I$ is an $m \times m$ identity matrix, and $Q = [q_{ij}] \in \mathbb{R}^{m \times m}$ is the generator of a continuous-time Markov chain, (i.e., $Q$ satisfies $q_{ij} \geq 0$ for $i \neq j$, $\sum_{j=1}^{m} q_{ij} = 0$ for each $i = 1, \ldots, m$).

(b) The noise sequence $\{\xi_n\}$ is given in (A1).

(c) The noise sequence $\{\zeta_n\}$ is a stationary sequence that is uniformly bounded such that for each $x \in \mathbb{R}^r$ and each $i \in \mathscr{M}$, $E\widehat{W}(x, i, \zeta_n) = 0$, and for any positive integer $m$,

$$\frac{1}{n} \sum_{j=m}^{m+n-1} E_m \widehat{W}(x, i, \zeta_j) \to 0 \text{ in probability}, \tag{15}$$

where $E_m$ denotes the conditioning on the $\sigma$-algebra $\mathscr{F}_m = \{x_j, \alpha_j, \xi_{j-1}, \zeta_{j-1} : j \leq m\}$.

(d) $\widehat{W}(\cdot, i, \zeta)$ is a continuous function for each $i \in \mathscr{M}$ and each $\zeta$ and $|\widehat{W}(x, i, \zeta)| \leq K(1 + |x|)$ for each $x \in \mathbb{R}^r$, $i \in \mathscr{M}$, and $\zeta$.

(e) $\{\alpha_n\}$, $\{\xi_n\}$, and $\{\zeta_n\}$ are mutually independent.

(H3) The generator $Q$ is irreducible.

We can proceed to obtain the following main convergence theorem. Note that the limit is not an ordinary differential equation but a differential equation with Markov switching.

**Theorem 2** *Assume* (H1) *and* (H2).

- *Then $(x^{\varepsilon}(\cdot), \alpha^{\varepsilon}(\cdot))$ is tight in $D([0,T] : \mathbb{R}^r \times \mathscr{M})$. Moreover, as $\varepsilon \to 0$, $(x^{\varepsilon}(\cdot), \alpha^{\varepsilon}(\cdot))$ converges weakly to $(x(\cdot), \alpha(\cdot))$ that is a solution of the martingale problem with operator $\mathscr{L}_1$. For any $f(\cdot,\cdot) : \mathbb{R}^r \times \mathscr{M} \mapsto \mathbb{R}$ satisfying for each $\alpha \in \mathscr{M}$, $f(\cdot, \alpha) \in C_0^1$ (space of continuously differentiable functions with compact support), $\mathscr{L}_1$ is defined as follows:*

$$\mathscr{L}_1 f(x,i) = (\nabla f(x,i))' M(i) x + Q f(x,\cdot)(i), \quad i \in \mathscr{M}, \tag{16}$$

*where*

$$Q f(x,\cdot)(i) = \sum_{j=1}^{m} q_{ij} f(x,j). \tag{17}$$

- *Assume that for each $\alpha \in \mathscr{M}$, $M(\alpha)$ is irreducible. Under the conditions of Theorem* 2*, the following assertions hold.*

(i) *The set $\mathsf{Z} = span\{\mathbb{1}\}$ is an invariant set.*

(ii) *The set* Z *is asymptotically stable in probability.*

- *Assume the conditions of Theorem* 2. *In the recursive algorithm, we also use the constraint* (9). *Then for any* $t_\varepsilon \to \infty$ *as* $\varepsilon \to 0$, $x^\varepsilon(\cdot + t_\varepsilon)$ *converges to the consensus solution* $\eta \mathbb{1}$ *in probability. That is for any* $\delta > 0$,

$$\lim_{\varepsilon \to 0} P(|x^\varepsilon(\cdot + t_\varepsilon) - \eta \mathbb{1}| \geq \delta) = 0.$$

To study the rate of convergence, define

$$z_n = \frac{x_n - y_n}{\sqrt{\mu}} = \frac{x_n - y_n}{\sqrt{\varepsilon}} \text{ since } \mu = \varepsilon. \tag{18}$$

Then it is readily verified that

$$z_{n+1} = z_n + \varepsilon M(\alpha_n) z_n + \sqrt{\varepsilon} W(\alpha_n) \xi_n + \varepsilon \mathrm{diag}(z_n) \Psi(\alpha_n, \zeta_n). \tag{19}$$

We pose another condition.

(H2') Condition (H2) holds with the following modifications. Either $\widehat{W}(x,\alpha,\zeta) = \mathrm{diag}(x)\Psi(\alpha,\zeta)$ or $\widehat{W}(x,\alpha,\zeta) = x\psi_1(\alpha,\zeta)$ where $\Psi(\alpha,\zeta) : \mathscr{M} \times \mathbb{R}^r \mapsto \mathbb{R}^r$ and $\psi_1(\alpha,\zeta) : \mathscr{M} \times \mathbb{R}^r \mapsto \mathbb{R}$ such that $\Psi(\cdot,\cdot)$ (resp. $\psi_1(\alpha,\zeta)$) is a bounded function, and that for each fixed $\alpha \in \mathscr{M}$ and each positive integer $m$, (15) is replaced by

$$\begin{aligned} &\frac{1}{n} \sum_{j=m}^{m+n-1} E_m \Psi(\alpha, \zeta_j) \to 0 \text{ in probability,} \\ &\sum_{j=n}^{\infty} |E_n \Psi(\alpha, \zeta_j)| < \infty, \end{aligned} \tag{20}$$

or

$$\begin{aligned} &\frac{1}{n} \sum_{j=m}^{m+n-1} E_m \psi_1(\alpha, \zeta_j) \to 0 \text{ in probability,} \\ &\sum_{j=n}^{\infty} |E_n \psi_1(\alpha, \zeta_j)| < \infty, \end{aligned} \tag{21}$$

where $\mathrm{diag}(x) = \mathrm{diag}(x', \ldots, x')$.

First, we can show that for sufficiently large $n$, $E|z_n|^2 = O(1)$. Define $z^\varepsilon(t) = z_n$ for $t \in [(n - N_\varepsilon)\varepsilon, (n - N_\varepsilon)\varepsilon + \varepsilon)$.

**Theorem 3** *Under conditions* (H1)–(H3), $(z^\varepsilon(\cdot), \alpha^\varepsilon(\cdot))$ *converges to* $(z(\cdot), \alpha(\cdot))$ *such that* $z(\cdot)$ *is a solution of the following Markov regime-switching stochastic differential equation*

$$dz = M(\alpha(t)) z dt + W(\alpha(t)) d\widehat{B}(t). \tag{22}$$

The study of the asymptotic properties are essentially in [9]. In the above, we only illustrated one of the cases regarding the relative size of $\varepsilon$ and $\mu$. Full considerations of multi-scale formulations can be done. The main reference for the two-time-scale system is [10].

## Algorithm with Delays

We still consider randomly regime-switching network topologies. In addition, we add another fold of complication by assuming that there are delays in the communication and computation. The delayed information makes the formulation more realistic. Suppose that the network topology depends on a discrete-time Markov chain. In our setup, the graph can take $m$ possible values. That is $\mathscr{M} = \{1, \ldots, m\}$. The Markov chain is used to model, for example, capacity of the network, random environment, and other random factors such as interrupts and rerouting of communication channels etc. Thus $\mathscr{G}(\alpha_n) = \sum_{l=1}^m \mathscr{G}(l) I_{\{\alpha_n = l\}}$. To illustrate, suppose that initially, the Markov chain is at $\alpha_0 = i$. Then the graph takes the value $\mathscr{G}(i)$. At a random instance $\tau_1$, the first jump of the Markov chain takes place so that $\alpha_{\tau_1} = j \neq i$, Then the graph switches to $\mathscr{G}(j)$ and holds that value for a random duration until the next jump of the Markov chain takes place.

Although stochastic approximation algorithms have been well studied, introduction of unbounded delays in the discrete-time algorithms makes the analysis non-classical. There are no established results on such systems at present. It is noted that in our formulation, the delays are of the order $O(1/\mu)$ in their relations to the adaptation stepsize $\mu$. Such a consideration is motivated by practical systems with non-negligible latency and time delays whose discretization always lead to unbounded discrete-time delays. The unbounded delays of order $O(1/\mu)$ makes it far more difficult to analyze algorithm convergence and rates of convergence. In addition, in our recursive algorithms, the iterate $x_{n+1}$ depends on $x_{n-\lfloor d/\mu \rfloor}$ but not on state $x_n$, which presents another difficulty. As a direct consequence of the above formulation, algorithm analysis becomes more challenging and non-standard, compared to traditional stochastic approximation algorithms. By taking appropriate interpolations, we obtain limit dynamic systems, which involve delays in equations rather than the usual differential equations. Because of the multiple scales in adaptation stepsize and dynamics switching frequency, the limit dynamic equation may become either an ordinary differential delay equation or a stochastic differential delay equation whose random switching is represented by a continuous-time Markov chain.

To include topology switching, to allow the use of correlated noise, and to permit delays in the measured states, suppose that $d > 0$ is a constant and consider the following stochastic approximation type algorithm

$$\begin{aligned} x_{n+1} &= x_n + \mu M(\alpha_n) x_{n-\lfloor d/\mu \rfloor} + \mu W(\alpha_n)[\xi_{n-\lfloor d/\mu \rfloor} + \widetilde{\xi}_n] \\ &\quad + \mu \widehat{W}(x_{n-\lfloor d/\mu \rfloor}, \alpha_n, \zeta_{n-\lfloor d/\mu \rfloor}) + \mu \widehat{W}(x_{n-\lfloor d/\mu \rfloor}, \alpha_n, \widetilde{\zeta}_n), \end{aligned} \tag{23}$$

with the initial segment $x_k$ for $k = -\lfloor d/\mu \rfloor, \ldots, 0$ being arbitrary, where $\mu > 0$ is the stepsize of consensus control algorithm, and $\lfloor d/\mu \rfloor$ denotes the integer part of $d/\mu$. For each $i \in \mathscr{M}$, $M(i)$ is the generator of a continuous-time Markov chain. The sequences $\{\xi_n\}$, $\{\widetilde{\xi}_n\}$, $\{\zeta_n\}$, and $\{\widetilde{\zeta}_n\}$ are random noise sequences with $\widehat{W}(\cdot,\cdot,\cdot) : \mathbb{R}^r \times \mathscr{M} \times \mathbb{R}^r \mapsto \mathbb{R}^r$.

When switching topologies are present and delays are allowed, care must be taken. In the traditional setup of stochastic approximation problems, the limit or averaged system is an ordinary differential equations (ODEs). Very often these limits are autonomous. Even if they are sometimes time inhomogeneous ordinary differential equations, these equations are non-random. As can be seen later, for the problem we are treating here, the limit is no longer an ODE, but an ODE with delays or randomly varying ODE with delays subject to switching modulated by the Markov switching process. In the literature of stochastic approximation, the rate of convergence study is normally associated with a limit stochastic differential equation. In the current problem, the limit ODE and SDE are however replaced by

$$\frac{dx(t)}{dt} = M(\alpha(t))x(t-d), \tag{24}$$

$$dz = M(\alpha(t))z(t-d)dt + W(\alpha(t))d\widehat{B}(t), \tag{25}$$

respectively. That is, they are delay equation with Markovian switching, and switching delay stochastic differential equation, respectively. The stability of (24) can be analyzed using the methods for switching diffusions [12] and treatment of delay equations with switching in [18].

## Algorithm with Asynchronous Computation and Communication

In this section, we enlarge the applicability of the platoon formation from another angle in that the computation and communications are done at random times and asynchronously. To carry out the recursive computational task, we consider a class of asynchronous and distributed algorithms in the following setup. Suppose that the state $x \in \mathbb{R}^r$ and there are $r$ processors participating in the computational task. For notational simplicity, we assume that each processor handles only one component. It is clear that this can be made substantially more general by allowing each processor to handle a vector of possibly different dimensions. However, the mathematical framework will be essentially the same albeit the complex notation. Suppose that for each $i \le r$, $\{Y_n^i\}$ is a sequence of positive integer-valued random variables (assuming the random sequence to be positive integer valued is for notational convenience) that are generally state and data dependent such that the $n$th iteration of processor $i$ takes $Y_{n-1}^i$ units of time. Define a sequence of "renewal-type" random computation times $\tau_n^i$ as

$$\tau_0^i = 0, \quad \tau_{n+1}^i = \tau_n^i + Y_n^i. \tag{26}$$

For each $i$, the sequence $\{Y_n^i\}$ is an inter-arrival time and $\{\tau_n^i\}$ is the corresponding "renewal" time. It is well known that $\widetilde{\alpha}_n$ is strongly Markov, so $\widetilde{\alpha}_{\tau_n^i}$ is a Markov chain.

Using constant stepsize $\mu > 0$, we consider the following asynchronous algorithm

$$x^i_{\tau^i_{n+1}} = x^i_{\tau^i_n} + \mu [M_{\tau^i_n}(\widetilde{\alpha}_{\tau^i_n}) x_{\tau^i_n}]^i + \mu [W_{\tau^i_n}(\widetilde{\alpha}_{\tau^i_n})\widetilde{\xi}^i_n]^i + \mu \widehat{W}^i_{\tau^i_n}(x_{\tau^i_n}, \widetilde{\alpha}_{\tau^i_n}, \widetilde{\zeta}^i_n), \quad i \le r, \tag{27}$$

where $\widetilde{\xi}^i_n \in \mathbb{R}^r$ and $\widetilde{\zeta}^i_n \in \mathbb{R}^r$ are the noise sequences incurred in the $(n+1)$st iteration. Note that the functions involved are time dependent. We use the same idea as in the setup of a fixed configuration, but allow more general structure. Note also that for each $n$ and $\alpha \in \mathscr{M}$, $M_n(\alpha)$ is not a generator of a Markov chain as the fixed $M$. We allow the non-additive noise be used. When $M_n(\iota) = M$ and $W_n(\iota) = W$ are constant matrices being generators of continuous Markov chains for all $n$ and all $\iota \in \mathscr{M}$, and $\widehat{W}_n \equiv 0$, the algorithm reduces to the existing standard consensus algorithm with additive noise. The nonadditive portion is a general nonlinear function of the analog state $x$, the Markov chain state $\iota \in \mathscr{M}$, the noise source $\zeta$, as well as $n$. The setup here is in line with [13]. Related asymptotic results can be obtained.

## CONCLUDING REMARKS

In this paper, we have surveyed some recent progress on applications of stochastic approximation methods to consensus control. Our primary motivations stem from platoon formation and maintenance in cyber-physical systems. Starting with the simplest basic problem, we have illustrated how random environments in terms of configuration changes, inclusion of delayed information, and asynchronous computation and communication can be incorporated. It is conceivable that such effort may lead to significant advances in cyber-physical systems and other related fields.

## ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under CNS-1136007.

## REFERENCES

1. H. Robbins and S. Monro, A stochastic approximation method, *Ann. Math. Statist.*, **22** (1951), 400–407.
2. J. Kiefer and J. Wolfowitz, Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, **23** (1952), 462–466.
3. H.J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed., Springer-Verlag, New York, NY, 2003.
4. L.Y. Wang, A. Syed, G. Yin, A. Pandya, H. Zhang, Control of vehicle platoons for highway safety and efficient utility: Consensus with communications and vehicle dynamics, *J. Syst. Sci. Complexity*, **27** (2014).
5. S. Haykin, *Digital Communications*, 4th ed., J. Wiley & Sons, 2001.
6. T.K. Moon, *Error Correction Coding, Mathematical Methods and Algorithms*, J. Wiley & Sons, 2005.
7. R.A. Brualdi, H.J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, 1991.
8. R.S. Varga, *Matrix Iterative Analysis*, Spring-Verlag, Berlin, 2000.
9. G. Yin, Y. Sun, and L.Y. Wang, Asymptotic properties of consensus-type algorithms for networked systems with regime-switching topologies, *Automatica*, **47** (2011) 1366–1378.
10. G. Yin and Q. Zhang, *Discrete-time Markov Chains: Two-time-scale Methods and Applications*, Springer, New York, NY, 2005.
11. G. Yin, L.Y. Wang, and Y. Sun, Stochastic recursive algorithms for networked systems with delay and random switching: Multiscale formulations and asymptotic properties, *SIAM J. Multiscale Modeling Simulation*, **9** (2011), 1087–1112.
12. G. Yin and C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Springer, New York, 2010.
13. G. Yin, Q. Yuan, L.Y. Wang, Asynchronous stochastic approximation algorithms for networked systems: Regime-switching topologies and multi-scale structure, *SIAM J. Multiscale Modeling Simulation*, **11** (2013), 813–839.
14. C.W. Reynolds, Flocks, herds, and schools: a distributed behavioral model, *Computer Graphics*, **21**(4): 25–34, July 1987.
15. T. Viseck, A. Czirook, E. Ben-Jacob, O. Cohen, and I. Shochet, Novel type of phase transition in a system of self-deriven particles, *Physical Review Letters*, **75** (6): 1226–1229, August, 1995.
16. J. Toner and Y. Tu, Flocks, herds, and schools: A quantitative theory of flocking, *Physical Review E*, **58** (4): 4828–4858, October 1998.
17. R. Olfati-Saber, J.A. Fax, and R.M. Murray. Consensus and cooperation in networked multi-agent systems, *IEEE Proc.*, vol. 95, no. 1, pp. 215-233, Jan. 2007.
18. F. Wu, G. Yin, and L.Y. Wang, Stability of a pure random delay system with two-time-scale Markovian switching, *J. Differential Eqs.*, **253** (2012), 878–905.