Published Online on 25 June 2015

Proc Indian Natn Sci Acad **81** No. 3 June 2015 pp. 569-582

DOI: 10.16943/ptinsa/2015/v81i3/48221

Medal Lecture

Playing Dice with the Universe: Algorithms for Random Environments

VIVEK S. BORKAR*

Department of Electrical Engineering, Indian Institute of Technology, Powai, Mumbai 400 076, India

(Delivered on 20 December 2014)

We describe stochastic recursive algorithms that find many applications in optimization under uncertainty, computational statistics, soft computing and machine learning, signal processing, communications, adaptive control, etc. The basic paradigm is that of the 'stochastic approximation' scheme introduced by Robbins and Monro in 1951. We describe the basic scheme and various theoretical and computational aspects related to its convergence and convergence rates. Several specific instances are also described, followed by more sophisticated versions such as distributed schemes and simulated annealing.

Key Words : Stochastic Recursive Algorithms; Stochastic Approximation; Convergence and Convergence Rates; Asynchronous and Multi-Scale Algorithms

Introduction

Stochastic approximation algorithms were introduced by (Robbins and Monro, 1951) as a scheme for finding roots of nonlinear equations given noisy measurements. That is, given a nonlinear function $h : \mathcal{R}^d \mapsto \mathcal{R}^d, d \geq 1$, and a black box that yields $h(x)$ + noise on input x, we wish to solve $h(x) = \theta$, the zero vector. Robbins and Monro came up with the iterative scheme

$$x(n+1) = x(n) + a(n)[h(x(n)) + M(n+1)],\ n \geq 0. \tag{1}$$

The term in square brackets is the aforementioned noisy measurement of $h(x(n))$ with $M(n+1)$ being the additive noise, assumed to be uncorrelated with past (i.e., its conditional expectation given all random

**Author for Correspondence : E-mail: borkar.vs@gmail.com; Tel.: 022 25769405 (Office); 9969066626 Mobile)*

variables realized till time n is zero). Contrary to its appearance, the scheme is quite general, because an iteration of the type

$$x(n+1) = x(n) + a(n)f(X(n), \xi(n+1))$$

with non-additive independent identically distributed (IID) noise $\xi(n), n \geq 0$, can be reduced to (1) by setting $h(x) := E[f(x, \xi(n))]$ and $M(n+1) := f(x(n), \xi(n+1)) - h(x(n))$. The ingenious idea of Robbins and Monro that clinched the issue was the choice of step-size sequence $a(n),\ n \geq 0$: these were chosen to be positive scalars satisfying

$$\sum_n a(n) = \infty,\ \sum_n a(n)^2 < \infty. \tag{2}$$

This ensures that $a(n) \to 0$, but sufficiently slowly. This fact proved to be enough to ensure probability one convergence of $x(n)$ to the set of zeros of h under mild technical conditions.

This stirred much interest in statistics community for a while and attracted an army of powerful researchers who further contributed in its wake. These include illustrious names such as Burkholder, Chung, Kesten and Kallianpur. But then the interest in statistics community waned somewhat. The domain in which this idea really took off is engineering, particularly electrical engineering, where it quickly became the standard paradigm for adaptive algorithms. So much so that along with Markov Chain Monte Carlo, it is a prime workhorse for statistical computations in engineering. The reasons for this success are not hard to guess. Stochastic approximation has some natural advantages: it is incremental, i.e., makes small changes per step, making it robust to noise, and has typically very low per iterate computation. These features conform to our intuitive requirements for adaptation and it is no wonder that the scheme fitted like a glove in engineering disciplines that require iterative adaptation. These include signal processing, adaptive control, communications, soft computing and machine learning, in the latter case particularly where high dimensional data is involved, so that sampling based schemes are almost obligatory.

What this survey attempts to do is to give a concise overview of the state-of-the-art of this vibrant field. The next section touches upon some theoretical issues. Section 3 will describe a few variants dictated by specific engineering situations. Section 4 lists a slew of special cases arising in applications. Section 5 describes some less known and more sophisticated variations. Section 6 concludes with some remarks.

This article is based on my Prashant Chandra Mahalanobis Medal Lecture delivered at INSA Annual Meeting in Goa on 20th December 2014. I am grateful to INSA for honouring me thus.

Theoretical Issues

The first major theoretical issue is convergence guarantees. While the early works used probabilistic tools such as 'almost supermartingales' for this purpose, much of recent work has been based upon the so called ODE (for '*Ordinary Differential Equation*') approach initiated by (Derevitskii and Fradkov, 1974) and further developed by (Ljung, 1977). The idea here is to compare (1) with the ODE

$$\dot{x}(t) = h(x(t)). \tag{3}$$

Recall the Euler scheme for numerical solution of (3) based on discretization:

$$x(n+1) = x(n) + ah(x(n)),\ n \geq 0,$$

where $a > 0$ is a small time step. Comparing with (1), one can view (1) as a noisy discretization of (3) with a slowly decreasing step-size. Under suitable technical conditions that imply in particular that (3) has a unique solution for all time for any initial condition, one can show that if the iterates $\{x(n)\}$ remain bounded with probability one (w. p. 1 for short), then (1) tracks the asymptotic behaviour of (3). This is so because the second condition in (2) ensures that the errors due to discretization and noise are asymptotically negligible, whereas the first condition ensures that the whole time axis is covered, which is essential in order to capture the asymptotic behaviour of (3). In particular, if the trajectories of (3) converge to a set of equilibria, i.e., the zero set of $h(\cdot)$, then so do the iterates w. p. 1. This can be refined further if the noise is 'rich enough' in all directions, to the iterate converge w. p. 1 to *stable* equilibria of (3). Often these are, or at least contain, the desired limit for (1). Of course, the ideal set-up is when (3) has a single globally asymptotically stable equilibrium x^*, when $x(n) \to x^*$ w. p. 1. The most general result in this context is due to (Benaim, 1996), which shows convergence to an '*internally chain transitive invariant set*' of (3) – see (Benaim, 1996) for a definition.

The condition that the iterates remain bounded is not free. There are several stability tests to ensure this, see, e.g., (Abounadi *et al*., 2002; Borkar and Meyn, 2000; Tsitsiklis, 1994). Another alternative is to project the iterates back to a prescribed bounded set, in which case they track a projected version of (3) (Kushner and Clark, 1978). A further variant is projection to a set that is slowly enlarged to fill up the whole space (Chen, 2002).

It is instructive to compare (1) with the classical strong law of large numbers (SLLN). Let $X_n, n \geq 1$, be IID random variables with mean μ and variance $\sigma^2 < \infty$. Let $S_0 = 0, S_n = \sum_{m=1}^{n} X_m, n \geq 0$. Then the SLLN states that $x(n) := \frac{S_n}{n} \to \mu$ w. p. 1. Some simple algebra shows that $x(n), n \geq 1$, defined thus satisfies (1) with $a(n) = \frac{1}{n+1}, h(x) := \mu - x$ and $M(n+1) = X_{n+1} - \mu$. The aforementioned convergence

analysis of (1) then implies the SLLN. This suggests that the convergence w. p. 1 of (1), when it occurs, is essentially a SLLN for a class of dependent random variables. For the classical SLLN, which captures 'typical' behaviour of sums of IID random variables, there are ancillary limit theorems to characterize the fluctuations around typical behaviour. One might then expect, and there indeed are, corresponding results for the iterates (1), e.g., central limit theorem and functional central limit theorem, 'almost sure' central limit theorem for empirical measures (think of histograms), law of iterated logarithms, etc. (Duflo, 1997). As in traditional statistics, these can be used to make statements about confidence intervals, rate of convergence, etc. A topic of greater interest to the engineering community is the sample complexity, i.e., for prescribed *accuracy parameter* ϵ and *confidence parameter* δ, the number of samples needed to be within $\epsilon > 0$ of the desired limit with probability $> 1 - \delta$, given as a function of ϵ, δ. See (Kamal, 2010) for a result along these directions.

Some useful observations are:

1. Smaller stepsizes reduce errors due to discretization and noise, but in view of their interpretation as time steps for the ODE, it is clear that they will imply slower convergence. Thus there is a trade-off involved. There is also some work on adaptive step-size selection.

2. If we add to the right hand side of (1) an additional error term $\eta(n+1)$ satisfying $\eta(n) \to 0$ w. p. 1, it does not affect the conclusion, since this is just another asymptotically negligible error. If, however, it does not go to zero but is sufficiently small, one can say, under reasonable hypotheses, that $x(n)$ will converge w. p. 1 to some small neighbourhood of the limit corresponding to no error, i.e., $\eta(n) \equiv 0$.

Variations

Some useful and commonplace variations are as follows.

1. *Constant stepsize:* This corresponds to setting $a(n) \equiv a > 0 \; \forall n$. Then the conditions (2) fail and w. p. 1 convergence cannot in general be expected. Nevertheless one can show a concentration of probability around the desired limit in most cases of interest. This scheme is preferred when one is operating in a slowly varying environment because a fixed step-size implies a constant timescale and the iterates can be expected to zero in and stay close to (in other words, track) the environment as long as this time scale is faster than the time scale on which the environment changes. On the other hand, the time scale of a decreasing step-size scheme as above would have eventually become slower than that of the environment and then the tracking ability would be lost.

2. *Differential inclusion limits:* There are situations where one gets a *differential inclusion* limit rather than an ODE limit, i.e.,

$$\dot{x}(t) \in h(x(t))$$

where $h(\cdot)$ is a set valued map, usually taking values in closed bounded convex sets and with a closed graph (i.e., $\{(x, y) : y \in h(x)\}$ is closed). This can happen, e.g., when $h(\cdot)$ is given as a maximizer or minimizer of something and happens to be non-uniquely specified, or when the original h of (3) is discontinuous and one is obliged to interpret the solution in a generalized sense (see, e.g., (Filippov, 1988)). Much, though not all, of the aforementioned theory has its counterparts for this case. See, e.g., (Benaim *et al*., 2005; Benaim *et al*., 2006), which also describe several interesting applications.

3. *Two time scale stochastic approximation:* Consider coupled iterations

$$\begin{aligned} x(n+1) &= x(n) + a(n)[h(x(n), y(n)) + M(n+1)], \\ y(n+1) &= y(n) + b(n)[g(x(n), y(n)) + M'(n+1)]. \end{aligned}$$

Here $\{a(n)\}, \{b(n)\}$ satisfy (2) and in addition, $\frac{b(n)}{a(n)} \to 0$. The latter condition implies that the second iteration moves on a slower algorithmic time scale than the first. Thus the first can be analyzed assuming the second 'quasi-static', i.e., $y(n) \approx$ a constant, whence $\{x(n)\}$ tracks the ODE

$$\dot{x}(t) = h(x(t), y)$$

where y enters parametrically. Suppose this ODE has a unique globally asymptotically stable equilibrium $\lambda(y)$. This implies $x(n) \approx \lambda(y(n))$. Then the slower iterates, which can be analyzed treating the faster ones as 'quasi-equilibrated', track the ODE

$$\dot{y}(t) = g(\lambda(y(t)), y(t)).$$

If this ODE has a unique globally asymptotically stable equilibrium y^*, then we get $(x(n), y(n)) \to (\lambda(y^*), y^*)$ w. p. 1. This idea can be used to implement nested algorithms: the fast iteration can be an inner loop (subroutine) whereas the slow one the outer loop. A common situation is averaging or estimation on the fast scale and optimization on the slow scale. Another situation is non-gradient primal-dual methods in approximate dynamic programming.

4. *Asynchronous implementations:* Consider a situation where different components are updated by different processors which communicate with each other across communication channels with delays.

The ith component, $1 \leq i \leq d$, is then updated by

$$x_i(n+1) = x(n) + a(n)I\{i \in Y(n)\}\Big(h_i(x_1(n-\tau_{1i}(n)), \cdots, x_d(n-\tau_{di}(n))) + M_i(n+1)\Big).$$

Here $Y(n) \in \{1, \cdots, d\}$ is the subset of components updated at time n, $I\{i \in Y(n)\} = 1$ if ith component is updated at time n and 0 if not, and $\tau_{ij}(n)$'s are random delays satisfying suitable conditional independence and conditional moment bounds. This, however, requires a global clock whose ticks correspond to the iterate index n. A more realistic model is given by

$$x_i(n+1) = x(n) + a(\nu(i,n))I\{i \in Y(n)\}\Big(h_i(x_1(n-\tau_{1i}(n)), \cdots, x_d(n-\tau_{di}(n))) + M_i(n+1)\Big).$$

Here $\nu(i,n) = \sum_{m=0}^{n} I\{i \in Y(m)\}$ is the 'local clock' of i, counting its own number of iterates. Then there is no need for a common global clock, in fact n can be any artificial tag that respects causality. Analysis shows that under reasonable conditions, delays contribute another asymptotically negligible error term and so do not affect the asymptotic dynamics, but asynchrony implicit in $Y(n) \neq \{1, \cdots, d\}$ does matter. The general limiting ODE turns out to be

$$\dot{x}(t) = \Lambda(t)h(x(t)),$$

where $\Lambda(t)$ is a diagonal matrix with non-negative entries. For special choice of step-sizes and under the condition that $\nu(i,n)/n$ remain uniformly bounded away from zero (i.e., all components are updated 'comparably often'), one can get $\Lambda(t)$ proportional to the identity matrix. If so, the limiting ODE is a time scaled version of (3), with exactly the same asymptotic behaviour.

One important instance of this situation is simulation based algorithms wherein the indices i correspond to the state of a simulated finite state Markov chain $X_n, n \geq 0$, and $Y(n) = \{X_n\}$, the singleton corresponding to its current position.

5. *Markov noise:* A more general stochastic approximation scheme is

$$x(n+1) = x(n) + a(n)[h(x(n), Y(n)) + M(n+1)]$$

where $Y(n), n \geq 0$, is the so called 'Markov noise', satisfying

$$P(Y(n+1) \in A | x(m), Y(m), M(m), m \leq n) = p_{x(n)}(A|Y(n)).$$

Here for each $x \in \mathcal{R}^d$, $p_x(\cdot|\cdot)$ is a transition probability kernel of an ergodic Markov chain with unique stationary distribution π_x. Under suitable technical hypotheses, the limiting ODE then is

$$\dot{x}(t) = \int \pi_{x(t)}(dy) h(x(t), y).$$

A more general 'controlled Markov' noise can be similarly handled (Borkar, 2006). See also (Anantharam and Borkar, 2012) for some non-standard noise models.

Examples

1. *Stochastic gradient scheme:* The most common situation in practice is stochastic gradient descent wherein $h(x) = -\nabla F(x)$, or ascent, where $h(x) = \nabla F(x)$, for a suitable performance measure F. A standard example of the former is the mean square error. For example in on-line nonlinear regression, we have IID pairs of observations $(X_n, Y_n), n \geq 1$, and we wish to fit a model

$$Y_n = f_\beta(X_n) + \epsilon_n.$$

Here f_β is to be chosen from a prescribed parametric family $\{f_\beta,\ \beta \in I\}$ for some index set $I \subset \mathcal{R}^d$, so as to minimize the mean square error

$$\frac{1}{2}E[\|\epsilon_n\|^2] = \frac{1}{2}E[\|Y_n - f_\beta(X_n)\|^2] := h(\beta).$$

The stochastic gradient scheme (for $I = \mathcal{R}^d$) then is

$$\beta(n+1) = \beta(n) - a(n)\Big(\nabla^\beta f_{\beta(n)}(X_n)\Big)^T [Y_n - f_{\beta(n)}(X_n)],$$

where $\nabla^\beta f(x)$ is the Jacobian matrix of $f_\beta(x)$ with respect to the β variable. Many parameter learning algorithms such as the Widrow-Hoff rule or the backpropagation algorithm for feedforward neural networks are of this form. (The latter uses the chain rule of calculus to split the computation of the gradient into simple local calculations). Standard choices for the parametrized function family are linear combinations of suitable bases such as trigonometric functions, wavelets, etc., splines, sigmoidal or radial basis function neural networks, etc. The limiting ODE

$$\dot{x}(t) = -\nabla F(x(t))$$

has local minima of F as its stable equilibria, ensuring w. p. 1 convergence to the set thereof.

Often the exact gradient is not available and an approximation has to be used. An old scheme for this is the Kiefer-Wolfowitz scheme (Kiefer and Wolfowitz, 1952) which uses finite difference approximation in each coordinate direction. The Simultaneous Perturbation Stochastic Approximation (SPSA) schemes of (Spall, 1992) replace this by finite difference along a single random direction, thereby saving a lot of computation. A more recent gradient estimation scheme for high dimensions uses Taylor expansion and regression (Mukherjee and Wu, 2006). As pointed out above, a good estimate ensures convergence to a neighbourhood of a local minimum.

2. *Liapunov systems:* These have limiting ODE for which there exists a Liapunov function that monotonically decreases along non-equilibrium trajectories, ensuring convergence to an equilibrium. Some examples are:

   (a) Given a unique saddle point (x^*, y^*) for a strictly convex-concave function $(x, y) \mapsto f(x, y)$ (i.e., $f(\cdot, y)$ is strictly convex and $f(x, \cdot)$ is strictly concave), a gradient descent-ascent (in x, resp. y variable) will converge to the saddle point. This follows from the fact that $\|x - x^*\|^2 + \|y - y^*\|^2$ serves as a Liapunov function. This situation arises in zero sum games and primal-dual schemes for constrained optimization.

   (b) Given a contraction map $f : \mathcal{R}^d \mapsto \mathcal{R}^d$ w.r.t. norm $\|x\|_p$ ($p \in [1, \infty]$) (i.e., $\|f(x) - f(y)\|_p \leq \alpha\|x - y\|_p$ for some $\alpha \in (0, 1)$), it is known that it has a unique fixed point x^*, i.e., $F(x^*) = x^*$. Then (1) with $h(x) = f(x) - x$ serves as a stochastic approximation scheme for computing x^*. Here $\|x - x^*\|_p$ serves as a Liapunov function. This can also be extended to some 'non-expansive' maps where $\alpha = 1$. This has important applications in algorithms for approximate dynamic programming arising in control engineering. Another situation where this works is 'anti-monotone' f, i.e., f satisfying

   $$\langle f(x) - f(y), x - y \rangle < 0 \ \ \forall\, x \neq y.$$

   (c) There are some instances of limiting ODEs of (1) taking values in the probability simplex

   $$S := \{p = [p_1, \cdots, p_d] : p_i \geq 0 \ \forall\, i, \ \sum_i p_i = 1\}.$$

   These necessarily require that $\sum_i h_i(\cdot) \equiv 0$ (so that the sum of components is preserved) and $p_i = 0 \implies h_i(p) \geq 0$ (so that the positive orthant is an invariant set). Of course, the actual stochastic algorithm may require projection to S at each step. Such algorithms are common when probabilities e.g., for routing, or randomized strategies in games, are being learnt. A

common Liapunov function that often works is the relative entropy w.r.t. the target equilibrium p^*, i.e.,

$$D(p||p^*) := \sum_i p_i^* \log\left(\frac{p_i^*}{p_i}\right).$$

One major instance of such ODEs is the replicator dynamics from evolutionary biology given by

$$\dot{p}_i(t) = p_i(t)[F_i(p(t)) - \sum_j p_j(t)F_j(p(t))],\ 1 \le i \le d.$$

In biology, $p_i(t)$ is the fraction of species i at time t and $F_i(p(t))$ is its payoff given the current population profile. The equation then captures the Darwinian evolution whereby the fraction increases or decreases according to whether this payoff is above or below the population average. This has been made the basis of many algorithms for network optimization. The equation need not converge to equilibria except in special cases, e.g., when either $F(\cdot) := [F_1(\cdot), \cdots, F_d(\cdot)]$ is anti-monotone or is the gradient of some 'potential' $\Psi : \mathcal{R}^d \mapsto \mathcal{R}$.

(d) Another special situation is the ODE

$$\dot{x}_i(t) = a_i(x(t))\left[b_i(x_i(t)) - \sum_j c_{ij} f_j\left(\sum_k c_{jk} g_k(x_k(t))\right)\right],$$

where $c_{ij} = c_{ji} \ge 0$ and $\ell := a_i, b_i, f_i, g_i$ or $\frac{d}{dx}g_i$ is > 0 and satisfies $\|\ell(x) - \ell(y)\| \le L\|x - y\|$ for a suitable $L > 0$. This has the Liapunov function

$$V(x) := \sum_i \left(\int_0^{\sum_j c_{ij} g_j(x_j)} f_i(y)dy - \int_0^{x_i} b_i(y)\frac{d}{dx}g_i(y)dy\right).$$

Such equations occur in a network optimization problem (Borkar, 2007), the Kelly-Maulloo-Tan dynamic pricing scheme (Kelly *et al*., 1998), and the Cohen-Grossberg model in neural networks (Cohen and Grossberg, 1983).

3. *Miscellaneous:* There are some special cases that do not fit the above, such as the following.

(a) *Eigenvector computations:* For computing the principal eigenvector of a non-negative irreducible matrix Q, $h(x) = \frac{Qx}{x_{i_0}} - x$ initialized at some $x(0) > 0$ and confined to positive orthant provides a stochastic approximation counterpart of the 'power method' (Borkar *et al*., 2014). Here i_0 is a fixed index. This also works for positive definite Q. Normalizing factors other than x_{i_0} are possible. A nonlinear version of this with $\min_u Q_u x$ in place of Qx for a parametrized

family Q_u (the minimum being pointwise) works as a scheme for approximate dynamic programming for risk-sensitive (i.e., multiplicative cost) control popular in finance (Borkar, 2002).

(b) *Global Newton:* Letting ∇f denote the Jacobian matrix of $f : \mathcal{R}^d \mapsto \mathcal{R}^d$, $h(x) = -\nabla f(x)^{-1} f(x)$ corresponds to the 'global Newton' scheme of Smale (Smale, 1976). Smale proved the convergence of this ODE to the zero set of f for 'almost all' initial conditions. The latter qualification can be ignored if the noise is rich enough. A case of special interest is $f := \nabla F$, $\nabla f := \nabla^2 F$ for some $F : \mathcal{R}^d \mapsto \mathcal{R}$, which corresponds to the Newton scheme for minimization of F. See (Bhatnagar *et al*., 2013) for stochastic approximation versions of this.

(c) *Cooperative systems:* If $\frac{\partial h_i(x)}{\partial x_j} \geq 0$ for $i \neq j$ and the Jacobian matrix is irreducible, then the ODE is called 'cooperaive' and a theorem of Hirsch ensures that if its trajectories are bounded, then they converge to the set of equilibria for 'almost all initial conditions' in a certain sense (Hirsch, 1985). Again, the qualification may be ignored in many practical cases. This has several applications, see, e.g., (Garg *et al*., 2005), for an application to dynamic pricing.

(d) *Principal Component Analysis (PCA):* A stochastic approximation version of Oja's scheme for PCA can be given (Borkar and Meyn, 2012), with the ODE limit

$$\dot{M}(t) = (1 + \text{trace}(M(t)M(t)^T)^{-1}[I - M(t)M(t)^T]WM(t).$$

Here W is the given $d \times d$ positive definite matrix and $M(t) \in \mathcal{R}^{d \times k}, k < d$. The algorithm converges to a basis for the k-dimensional subspace corresponding to k leading eigenvalues. See also (Diamantaras and Kung, 1996) for a scheme that gives the leading eigenvectors as well. This has applications to spectral clustering among other things.

Further Variants

Some related algorithms that stretch further the possibilities offered by stochastic approximation are as follows.

1. *Simulated annealing:* Consider the problem of minimizing a continuously differentiable function $V : \mathcal{R}^d \to \mathcal{R}$ with $\lim_{\|x\| \uparrow \infty} V(x) = \infty$ sufficiently fast so that $e^{-\alpha V(x)}$ is integrable for $\alpha > 0$. For this, replace (1) by

$$x(n+1) = x(n) + a(n)[-\nabla V(x(n)) + M(n+1)] + b(n)\xi(n+1),\ n \geq 0.$$

where $b(n) \downarrow 0$ at a rate faster than $a(n)$. This can be arranged so that $x(n)$ tracks the behaviour of the *stochastic differential equation*

$$dX(t) = -\nabla V(x(t))dt + \epsilon(t)dW(t)$$

for a standard d-dimensional Brownian motion $W(\cdot)$, where $\epsilon(t) \downarrow 0$ sufficiently slowly. For suitable choices of $a(n), b(n)$, one has $\epsilon^2(t) \geq \frac{C}{\log t}$ where $C \geq C_0 > 0$ for a certain constant C_0 related to the spectrum of the differential operator $\frac{1}{2}\Delta - \langle \nabla V, \cdot \rangle$. In this case, the distribution of $x(t)$ and hence that of $x(n)$, tracks the stationary distribution of the time-homogeneous stochastic differential equation

$$dX(t) = -\nabla V(x(t))dt + \epsilon_0 dW(t)$$

in the $\epsilon_0 \downarrow 0$ limit. The latter distribution can be shown to concentrate on the set M of global minima of V. Together, these considerations establish convergence in probability of $x(n)$ to M. This is the continuous state space analog of the celebrated simulated annealing algorithm for global optimization. See (Gelfand and Mitter, 1991) for details.

2. *Gossip-based algorithms:* Consider a distributed algorithm wherein processors sit on the nodes of a connected undirected graph $\mathcal{G}$ and the ith processor performs the d-dimensional iteration

$$x^i(n+1) = \sum_{j \in \mathcal{N}(i)} p(j|i)x^j(n) + a(n)[h^i(x^i(n)) + M^i(n+1)],\ 1 \leq i \leq k. \tag{4}$$

Here $\mathcal{N}(i) :=$ the set of neighbours of i in $\mathcal{G}$ and $p(\cdot|\cdot)$ an irreducible stochastic matrix compatible with $\mathcal{G}$. Let π denote the unique stationary distribution for $p(\cdot|\cdot)$. Then the limiting ODE is

$$\dot{x}^i(t) = \sum_j \pi(j)h^j(x^j(t)).$$

Compare this with the 'gossip' algorithm

$$x^i(n+1) = \sum_{j \in \mathcal{N}(i)} p(j|i)x^j(n)$$

for averaging, which leads to $x^i(n) \to \sum_j \pi(j)x^j(0)\ \forall i$. Thus the first term on the right of (4) is a gossip-like averaging that forces consensus among the processors and the second term is their individual stochastic approximation computations. First introduced in (Tsitsiklis *et al.*, 1986), this is a popular paradigm for coordination algorithms among multiple agents. More interesting variants involve random polling of neighbours (see, e.g., (Borkar *et al.*, 2014)) and mobile agents leading to time-varying $\mathcal{G}$ (see, e.g., (Chen *et al.*, 2012)).

Nonlinear versions of the 'gossip' part have been proposed in (Mathkar and Borkar, 2014), which either allow the averaging weights $p(\cdot|\cdot)$ to depend dynamically on $x(n)$, or replace them altogether by a nonlinear map. The latter is motivated by distributed projection schemes.

Conclusions

Despite its long history and considerable body of work, the field of stochastic iterative algorithms is still active, largely due to new issues being constantly thrown up by the technological imperatives, high dimensional data being a recent example. Some important current research themes are:

1. Algorithms that scale well with dimension in the high dimensional limit, both from point of view of computation and of memory and data retrieval.

2. Algorithms for multiagent scenarios such as autonomous agents (robot swarms / UAVs) and cyber-physical systems, that respect the communication constraints and dynamic network topologies.

3. Algorithms that operate as a component of a loop in closed loop control, particularly in networked control.

References

1. Abounadi J, Bertsekas D and Borkar V S (2002) Stochastic approximation for nonexpansive maps: applications to Q-learning algorithms *SIAM J Control and Optim* **41** 1-22
2. Anantharam V and Borkar V S (2012) Stochastic approximation with long range dependent and heavy tailed noise *Queueing Systems* **71** 221-242
3. Benaim M (1996) A dynamical system approach to stochastic approximation *SIAM J Control and Optim* **34** 437-472
4. Benaim M, Hofbauer J and Sorin S (2005) Stochastic approximations and differential inclusions *SIAM J Control and Optim* **44** 328-348
5. Benaim M, Hofbauer J and Sorin S (2006) Stochastic approximations and differential inclusions, Part II: Applications *Math Op Research* **31** 673-695
6. Bhatnagar S, Prasad H L and Prashanth L A (2013) *Stochastic Recursive Algorithms for Optimization: Simultaneous Perturbation Methods* Lecture Notes in Control and Information Sciences No. 434, Springer Verlag, London
7. Borkar V S (2002) Q-learning for risk-sensitive control, *Math Op Research* **27** 294-311
8. Borkar V S Some examples of stochastic approximation in communications, *Proc 2nd Euro-NF Workshop on Network Control and Optimization, Avignon, France, June 2007* (Eds: T Chahed and B Tuffin), Lecture Notes in Comp Sci No 4465, Springer Verlag, Berlin-Heidelberg, 150-157

9. Borkar V S (2006) Stochastic approximation with controlled Markov noise, *Systems and Control Letters* **55** 139-145
10. Borkar V S (2008) *Stochastic Approximation: A Dynamical Systems Viewpoint*, Hindustan Book Agency, New Delhi, and Cambridge Uni Press, Cambridge, UK
11. Borkar V S, Makhijani R and Sundaresan R (2014) Asynchronous gossip for averaging and spectral ranking *IEEE J on Selected Topics in Signal Processing* **8** 703-716
12. Borkar V S and Meyn S P (2000) The ODE method for convergence of stochastic approximation and reinforcement learning *SIAM J Control and Optim* **38** 447-469
13. Borkar V S and Meyn S P (2012) Oja's algorithm for graph clustering, Markov spectral decomposition, and risk sensitive control *Automatica* **48** 2512-2519
14. Chen H-F (2002) *Stochastic Approximation and its Applications* Kluwer Academic, Dordrecht
15. Chen G, Liu Z and Guo L (2012) The smallest possible interaction radius for synchronization of self-propelled particles *SIAM J Control and Optim* **50** 1950-1970
16. Cohen M A and Grossberg S (1983) Absolute stability of global pattern formation and parallel memory storage by competitive neural networks *IEEE Trans Systems, Man and Cybernetics* **SMC-13** 815-826
17. Derevitskii D P and Fradkov A L (1974) Two models for analyzing the dynamics of adaptation algorithms, *Automation and Remote Control* **35** 59-67
18. Duflo M (1997) *Random Iterative Models*, Springer Verlag, Berlin Heidelberg
19. Diamantaras K I and Kung S Y (1996) *Principal Components Neural Networks: Theory and Applications* John Wiley, New York
20. Filippov A F (1988) *Differential Equations with Discontinuous Righthand Sides*, Kluwer Academic, Dordrecht
21. Garg D, Borkar V S and Manjunath D (2005) Network pricing for QoS: a 'regulation' approach, in *Advances in Control, Communication Networks and Transportation Systems* (Ed: E. H. Abed), Birkhäuser, Boston, 137-157
22. Gelfand S B and Mitter S K (1991) Recursive stochastic algorithms for global optimization in $\mathcal{R}^d$, *SIAM J Control and Optim* **29** 999-1018
23. Hirsch M W (1985) Systems of differential equations that are competitive or cooperative II: convergence almost everywhere *SIAM J Math Analysis* **16** 423-439
24. Kamal S (2010) On the convergence, lock-in probability, and sample complexity of stochastic approximation *SIAM J Control and Optim* **48** 5178-5192
25. Kelly F P, Maulloo A and Tan D (1998) Rate control in communication networks: shadow prices, proportional fairness and stability *J Op Research Society* **49** 237-252
26. Kiefer J and Wolfowitz J (1952) Stochastic estimation of the maximum of a regression function *Annals of Math Stat* **23** 462-466
27. Kushner H J and Clark D (1978) *Stochastic Approximation Algorithms for Constrained and Unconstrained Systems* Springer Verlag, New York
28. Ljung L (1978) Analysis of recursive stochastic algorithms *IEEE Trans on Automatic Control* **22** 551-575
29. Mathkar A S and Borkar V S (2014) *Nonlinear gossip* (submitted)

30. Mukherjee S and Wu Q (2006) Estimation of gradients and coordinate covariation in classification, *J Machine Learning Research* **7** 2481-2514
31. Robbins H and Monro S (1951) A stochastic approximation method *Annals of Math Stat* **22** 400-407
32. Smale S (1976) A convergent process of price adjustment and global Newton methods *J Math Economics* **3** 107-120
33. Spall J C (1992) Multivariate stochastic approximation using a simultaneous perturbation gradient approximation *IEEE Trans Auto Control* **37** 332-341
34. Tsitsiklis J N (1994) Asynchronous stochastic approximation and Q-learning *Machine Learning* **16** 185-202
35. Tsitsiklis J N, Athans M and Bertsekas D P (1986) Distributed asynchronous deterministic and stochastic gradient optimization algorithms *IEEE Trans Auto Control* **AC-31** 803-812.