A Generalization of the Borkar-Meyn Theorem for Stochastic Recursive Inclusions

Arunselvan Ramaswamy ¹ and Shalabh Bhatnagar ²

¹arunselvan@csa.iisc.ernet.in
²shalabh@csa.iisc.ernet.in
¹,²Department of Computer Science and Automation, Indian Institute of Science, Bangalore - 560012, India.

March 3, 2015

Abstract

In this paper the stability theorem of Borkar and Meyn is extended to include the case when the mean field is a set-valued map. Two different sets of sufficient conditions are presented that guarantee the stability and convergence of stochastic recursive inclusions. Our work builds on the works of Benaim, Hofbauer and Sorin as well as Borkar and Meyn. As a corollary to one of the main theorems, a natural generalization of the Borkar and Meyn Theorem follows. In addition, the original theorem of Borkar and Meyn is shown to hold under slightly relaxed assumptions. Finally, as an application to one of the main theorems we discuss a solution to the ‘approximate drift problem’.

1 Introduction

Consider the following recursion in \( \mathbb{R}^d \) (\( d \geq 1 \)):

\[
x_{n+1} = x_n + a(n) \left[ h(x_n) + M_{n+1} \right], \quad \text{for } n \geq 0,
\]

where

(i) \( h : \mathbb{R}^d \to \mathbb{R}^d \) is a Lipschitz continuous function.

(ii) \( a(n) > 0, \) for all \( n, \) is the step-size sequence satisfying \( \sum_{n=0}^{\infty} a(n) = \infty \) and \( \sum_{n=0}^{\infty} a(n)^2 < \infty. \)

(iii) \( M_n, n \geq 0, \) is a sequence of martingale difference terms.

The stochastic recursion given by (1) is often referred to as a stochastic recursive equation (SRE). A powerful method to analyze the limiting behavior of (1) is the ODE (Ordinary Differential Equation) method. Here the limiting behavior of the algorithm is described in terms of the asymptotics of the solution to the ODE

\[
\dot{x}(t) = h(x(t)).
\]

This method was introduced by Ljung [10] in 1977. For a detailed exposition on the subject and a survey of results, the reader is referred to Kushner and Yin [9] as well as Borkar [7].
In 1996, Benaim [4] showed that the asymptotic behavior of a stochastic recursive equation can be studied by analyzing the asymptotic behavior of the associated o.d.e. However no assumptions were made on the dynamics of the o.d.e. Specifically, he developed sufficient conditions which guarantee that limit sets of the continuously interpolated stochastic iterates are compact, connected, internally chain transitive and invariant sets of the associated o.d.e. The results found in [4] are generalized in [3]. The assumptions made in [4] are sometimes referred to as the ‘classical assumptions’. One of the key assumptions used by Benaim to prove this convergence theorem is the almost sure boundedness of the iterates. In other words the iterates are stable. In 1999, Borkar and Meyn [8] developed sufficient conditions which guarantee both the stability and convergence of stochastic recursive equations. These assumptions were consistent with those developed in [4]. In this paper we refer to the main result of Borkar and Meyn colloquially as the Borkar-Meyn Theorem. In the same paper [8], several applications to problems from reinforcement learning have also been discussed. An extention to the Borkar-Meyn result for asynchronous stochastic iterates was given by Bhatnagar [6].

In 2005, Benaim, Hofbauer and Sorin [5] showed that the dynamical systems approach can be extended to the situation where the mean field is a set-valued map. In other words, they considered algorithms with iterates of the form:

\[ x_{n+1} = x_n + a(n) [y_n + M_{n+1}], \text{ for } n \geq 0, \text{ where} \]

(i) \( y_n \in h(x_n) \) and \( h : \mathbb{R}^d \to \{ \text{subsets of } \mathbb{R}^d \} \) is a Marchaud map. For a definition of Marchaud maps the reader is referred to section 2.1.

(ii) \( a(n) > 0, \text{ for all } n, \) is the step-size sequence satisfying \( \sum_{n=0}^{\infty} a(n) = \infty \) and \( \sum_{n=0}^{\infty} a(n)^2 < \infty. \)

(iii) \( M_n, n \geq 0, \) is a sequence of martingale difference terms.

Such iterates are referred to as stochastic recursive inclusions (SRI). An SRE given by (1) can be seen as a special case of an SRI given by (2). Also, a differential equation can be seen as a special case of a differential inclusion where the set \( \{ h(x) \} \) has cardinality one for all \( x \in \mathbb{R}^d. \)

The main aim of this paper is to extend the original Borkar-Meyn theorem to the case of stochastic recursive inclusions. We present two overlapping yet different sets of assumptions, in Sections 2.2 and 5.1 respectively, that guarantee the stability and convergence of a SRI given by (2). We present several interesting consequences to the main results (Theorems 2 and 4). Specifically, in Section 5.2 we show that the Borkar-Meyn theorem can be naturally extended to include the case where the infinity system is more generally allowed to be the set of accumulation points. We present a unified set of assumptions that takes care of both the original theorem and the aforementioned extension. While doing so we show how the assumptions of the Borkar-Meyn theorem could be relaxed.
The organization of this paper is as follows:

In section 2.1, we discuss the definitions and notations used in this paper. In section 2.2, we discuss the assumptions under which the iterates given by (2) are stable and converge to a closed, connected, internally chain transitive and invariant set. A preliminary result is also presented in this section.

In section 3.1, we present auxiliary results needed to prove the main result. In section 3.2, a stability theorem (Theorem 2) for stochastic recursive inclusions is proved under the assumptions outlined in section 2.2. Theorem 2 is one of two stability theorems presented in this paper.

In section 4 we discuss applications of Theorem 2. Specifically, in section 4.1 a solution to the problem of ‘approximate drift’ is discussed. For more details on the ‘approximate drift’ problem the reader is referred to Borkar [7]. In section 4.2, the Borkar-Meyn Theorem is proved under a relaxed set of assumptions.

In section 5.1, another set of assumptions that guarantee the stability and convergence to a closed, connected, internally chain transitive and invariant set of a stochastic recursive inclusion is discussed. In the same section we provide a brief outline of the proof of stability under these set of assumptions. This yields our second stability result for SRI: Theorem 4. In section 5.2, as an application of Theorem 4 the Borkar-Meyn Theorem [8] is generalized to when one of the key requirements does not hold. This result is summarized in Corollary 2. Finally, in section 5.3 we discuss how the assumptions described in sections 2.2, 5.1 could be relaxed.

2 Preliminaries and Assumptions

2.1 Definitions and Notations

The definitions and notations used in this paper are similar to those in Benaîm et. al. [5], Aubin et. al. [1], [2] and Borkar [7]. In this section, we present a few for easy reference.

A set-valued map \( h : \mathbb{R}^n \to \{\text{subsets of } \mathbb{R}^m \} \) is called a Marchaud map if it satisfies the following properties:

(i) For each \( x \in \mathbb{R}^n \), \( h(x) \) is convex and compact.

(ii) (point-wise boundedness) For each \( x \in \mathbb{R}^n \), \( \sup_{w \in h(x)} \|w\| < K (1 + \|x\|) \) for some \( K > 0 \).

(iii) \( h \) is an upper-semicontinuous map. We say that \( h \) is upper-semicontinuous, if given sequences \( \{x_n\}_{n \geq 1} \) (in \( \mathbb{R}^n \)) and \( \{y_n\}_{n \geq 1} \) (in \( \mathbb{R}^m \)) with \( x_n \to x \), \( y_n \to y \) and \( y_n \in h(x_n) \), \( n \geq 1 \), implies that \( y \in h(x) \). In other words the graph of \( h \), \( \{(x,y) : y \in h(x), x \in \mathbb{R}^n\} \), is closed in \( \mathbb{R}^n \times \mathbb{R}^m \).

Let \( H \) be a Marchaud map on \( \mathbb{R}^d \). The differential inclusion (DI) given by

\[
\dot{x} \in H(x) \tag{3}
\]

is guaranteed to have at least one solution that is absolutely continuous. The reader is referred to [1] for more details. We say that \( x \in \sum \) if \( x \) is an absolutely continuous map that satisfies (3). The set-valued semiflow \( \Phi \) associated with
is defined on $[0, +\infty) \times \mathbb{R}^d$ as: $\Phi_t(x) = \{x(t) \mid x \in \sum_x(0) = x\}$. Let $B \times M \subset [0, +\infty) \times \mathbb{R}^k$ and define

$$\Phi_B(M) = \bigcup_{t \in B, x \in M} \Phi_t(x).$$

Let $M \subset \mathbb{R}^d$, the $\omega$–limit set is defined by $\omega_M(M) = \bigcap_{t \geq 0} \Phi_{[t, +\infty)}(M)$. Similarly the limit set of a solution $x$ is given by $L(x) = \bigcap_{t \geq 0} x([t, +\infty])$.

$M \subset \mathbb{R}^d$ is invariant if for every $x \in M$ there exists a trajectory, $x$, entirely in $M$ with $x(0) = x$. In other words, $x \in \sum_x$ with $x(t) \in M$, for all $t \geq 0$.

Internally Chain Transitive Set: $M \subset \mathbb{R}^d$ is said to be internally chain transitive if $M$ is compact and for every $x, y \in M, \epsilon > 0$ and $T > 0$ we have the following: There exist $\Phi^1, \ldots, \Phi^n$ that are $n$ solutions to the differential inclusion $\dot{x}(t) \in h(x(t))$, a sequence $x_1(= x), \ldots, x_{n+1}(= y) \subset M$ and $n$ real numbers $t_1, t_2, \ldots, t_n$ greater than $T$ such that: $\Phi_{t_i}(x_i) \in N^\epsilon(x_{i+1})$ and $\Phi_{[0, t_i]}(x_i) \subset M$ for $1 \leq i \leq n$. The sequence $(x_1(= x), \ldots, x_{n+1}(= y))$ is called an $(\epsilon, T)$ chain in $M$ from $x$ to $y$.

Let $x \in \mathbb{R}^d$ and $A \subset \mathbb{R}^d$, then $d(x, A) := \inf\{\|a - y\| \mid y \in A\}$. We define the $\delta$-open neighborhood of $A$ by $N^\delta(A) := \{x \mid d(x, A) < \delta\}$. The $\delta$-closed neighborhood of $A$ is defined by $\overline{N}^\delta(A) := \{x \mid d(x, A) \leq \delta\}$.

$A \subset \mathbb{R}^d$ is an attracting set if it is compact and there exists a neighborhood $U$ such that for any $\epsilon > 0, \exists T(\epsilon) \geq 0$ such that $\Phi_{[t, +\infty)}(U) \subset N^\epsilon(A)$. Such a $U$ is called the fundamental neighborhood of $A$. In addition to being compact if the attracting set is also invariant then it is called an attractor. The basin of attraction of $A$ is given by $B(A) = \{x \mid \omega_M(x) \subset A\}$. It is called Lyapunov stable if for all $\delta > 0, \exists \epsilon > 0$ such that $\Phi_{[0, +\infty)}(N^\epsilon(A)) \subset N^\delta(A)$. We use $T(\epsilon)$ and $T_{\epsilon}$ interchangeably to denote the dependence of $T$ on $\epsilon$.

The open ball of radius $r$ around $0$ is represented by $B_r(0)$, while the closed ball is represented by $\overline{B}_r(0)$.

We define the lower and upper limits of sequences of sets. Let $\{K_n\}_{n \geq 1}$ be a sequence of sets in $\mathbb{R}^d$.

1. The lower limit of $\{K_n\}_{n \geq 1}$ is given by,

$$\liminf_{n \to +\infty} K_n := \{x \mid \lim_{n \to +\infty} d(x, K_n) = 0\}.$$

2. The upper-limit of $\{K_n\}_{n \geq 1}$ given by,

$$\limsup_{n \to +\infty} K_n := \{y \mid \lim_{n \to +\infty} d(y, K_n) = 0\}.$$ We may interpret that the upper-limit collects the accumulation points of $\{K_n\}_{n \geq 1}$.

### 2.2 The assumptions

Recall that we have the following recursion in $\mathbb{R}^d$:

$$x_{n+1} = x_n + a(n) [y_n + M_{n+1}],$$

where $y_n \in h(x_n)$. We state our assumptions below:

(A1) $h : \mathbb{R}^d \to \{\text{subsets of } \mathbb{R}^d\}$ is a Marchaud map.

(A2) $\{a(n)\}_{n \geq 0}$ is a scalar sequence such that: $a(n) \geq 0 \forall n, \sum_{n \geq 0} a(n) = \infty$ and $\sum_{n \geq 0} a(n)^2 < \infty$. Without loss of generality we let $\sup_{n} a(n) \leq 1$. 

4
\(\{M_n\}_{n \geq 1}\) is a martingale difference sequence with respect to the filtration \(\mathcal{F}_n := \sigma (x_0, M_1, \ldots, M_n), n \geq 0.\)

(i) \(\{M_n\}_{n \geq 1}\) is a square integrable sequence with \(E [M_{n+1} | \mathcal{F}_n] = 0\) a.s., for \(n \geq 0.\)

(ii) \(E [\|M_{n+1}\|^2 | \mathcal{F}_n] \leq K (1 + \|x_n\|^2),\) for \(n \geq 0\) and some constant \(K > 0.\) Without loss of generality assume that the same constant, \(K,\) works for both the point-wise boundedness condition of (A1) (see conditions (ii) of definition of Marchaud map in Section 2.1) and (A3).

For \(c \geq 1\) and \(x \in \mathbb{R}^d,\) define \(h_c(x) = \{y \mid cy \in h(cx)\}.\)

For each \(x \in \mathbb{R}^d,\) define \(h_\infty (x) := \varliminf_{c \to \infty} h_c(x)\) i.e. the closure of the lower limit of \(\{h_c(x)\}_{c \geq 1}.\)

(A4) \(h_\infty (x)\) is non-empty for all \(x \in \mathbb{R}^d.\) Further, the differential inclusion \(\dot{x} (t) \in h_\infty (x(t))\) has the origin as an attracting set and \(\mathcal{B}_1(0)\) is a subset of its fundamental neighborhood.

(A5) Let \(c_n \geq 1\) be an increasing sequence of integers such that \(c_n \uparrow \infty\) as \(n \to \infty.\) Further, let \(x_n \to x\) and \(y_n \to y\) as \(n \to \infty,\) such that \(y_n \in h_{c_n}(x_n), \forall n,\) then \(y \in h_\infty (x).\)

Assumptions (A1) – (A3) are the same as in Benaim [5]. However, the assumption on the stability of the iterates is replaced by (A4) and (A5). We show that (A4) and (A5) are sufficient conditions to ensure stability of iterates. We start by observing that \(h_c\) and \(h_\infty\) are Marchaud maps, where \(c \geq 1.\) Further, we show that the constant associated with the point-wise boundedness property is \(K\) of (A1) and (A3).

**Proposition 1.** \(h_\infty\) and \(h_c, c \geq 1,\) are Marchaud maps.

**Proof.** Fix \(c \geq 1\) and \(x \in \mathbb{R}^d.\) To prove that \(h_c(x)\) is compact, we show that it is closed and bounded. For \(n \geq 1,\) let \(y_n \in h_c(x)\) and let \(\lim y_n = y.\) It follows that \(cy_n \in h(cx)\) for each \(n \geq 1\) and \(\lim_{n \to \infty} cy_n = cy.\) Since \(h(cx)\) is closed, we have that \(cy \in h(cx)\) and \(y \in h_c(x).\) If we show that \(h_c\) is point-wise bounded then we can conclude that \(h_c(x)\) is compact. To prove the aforementioned, let \(y \in h_c(x),\) then \(cy \in h(cx).\) Since \(h\) satisfies (A1)(ii), we have that

\[
c \|y\| \leq K (1 + \|cx\|), \quad \text{hence}
\]

\[
\|y\| \leq K \left(\frac{1}{c} + \|x\|\right).
\]

Since \(c(\geq 1)\) and \(x\) is arbitrarily chosen, \(h_c\) is point-wise bounded and the compactness of \(h_c(x)\) follows. The set \(h_c(x) = \{z/c \mid z \in h(cx)\}\) is convex since \(h(cx)\) is convex and \(h_c(x)\) is obtained by scaling it by \(\frac{1}{c}.

Next, we show that \(h_c(x)\) is upper-semicontinuous. Let \(\lim_{n \to \infty} x_n = x, \lim_{n \to \infty} y_n = y\) and \(y_n \in h_c(x_n)\), \(\forall n \geq 1.\) We need to show that \(y \in h_c(x).\) We have that \(cy_n \in h(cx_n)\) for each \(n \geq 1.\) Since \(\lim_{n \to \infty} cx_n = cx\) and \(\lim_{n \to \infty} cy_n = cy,\) we conclude that \(cy \in h(cx)\) since \(h\) is assumed to be upper-semicontinuous.
It is left to show that \( h_\infty(x), x \in \mathbb{R}^d \) is a Marchaud map. To prove that \( \|z\| \leq K (1 + \|x\|) \) for all \( z \in h_\infty(x) \), it is enough to prove that \( \|y\| \leq K (1 + \|x\|) \) for all \( y \in \text{Liminf}_{t \to \infty} h_c(x) \). Fix some \( y \in \text{Liminf}_{t \to \infty} h_c(x) \) then there exist \( z_n \in h_n(x), n \geq 1 \), such that \( \lim_{n \to \infty} \|y - z_n\| = 0 \). We have that

\[
\|y\| \leq \|y - z_n\| + \|z_n\|.
\]

Since \( h_c, c \geq 1 \), is point-wise bounded (the constant associated is independent of \( c \) and equals \( K \)) the above inequality becomes

\[
\|y\| \leq \|y - z_n\| + K (1 + \|x\|).
\]

Letting \( n \to \infty \) in the above inequality, we obtain \( \|y\| \leq K (1 + \|x\|) \). Recall that \( h_\infty(x) = \text{Liminf}_{t \to \infty} h_c(x) \), hence it is compact.

Again, to show that \( h_\infty(x) \) is convex, for each \( x \in \mathbb{R}^d \), we start by proving that \( \text{Liminf}_{t \to \infty} h_c(x) \) is convex. Let \( u, v \in \text{Liminf}_{t \to \infty} h_c(x) \) and \( 0 \leq t \leq 1 \). We need to show that \( tu + (1 - t)v \in \text{Liminf}_{t \to \infty} h_c(x) \). Consider an arbitrary sequence \( \{c_n\}_{n \geq 1} \) such that \( c_n \to \infty \), then there exist \( u_n, v_n \in h_c(x) \) such that \( \|u_n - u\| \) and \( \|v_n - v\| \to 0 \) as \( c_n \to \infty \). Since \( h_c(x) \) is convex, it follows that \( tu_n + (1 - t)v_n \in h_c(x) \), further

\[
\lim_{c_n \to \infty} (tu_n + (1 - t)v_n) = tu + (1 - t)v.
\]

Since we started with an arbitrary sequence \( c_n \to \infty \), it follows that \( tu + (1 - t)v \in \text{Liminf}_{t \to \infty} h_c(x) \). Now we can prove that \( h_\infty(x) \) is convex. Let \( u, v \in h_\infty(x) \). Then \( \exists \{u_n\}_{n \geq 1} \) and \( \{v_n\}_{n \geq 1} \subseteq \text{Liminf}_{t \to \infty} h_c(x) \) such that \( u_n \to u \) and \( v_n \to v \) as \( n \to \infty \). We need to show that \( tu_n + (1 - t)v_n \in h_\infty(x) \), for \( 0 \leq t \leq 1 \). Since \( tu_n + (1 - t)v_n \in \text{Liminf}_{t \to \infty} h_c(x) \), the desired result is obtained by letting \( n \to \infty \) in \( tu_n + (1 - t)v_n \).

Finally, we show that \( h_\infty \) is upper-semicontinuous. Let \( \lim_{n \to \infty} x_n = x, \lim_{n \to \infty} y_n = y \) and \( y_n \in h_\infty(x_n), \forall n \geq 1 \). We need to show that \( y \in h_\infty(x) \). Since \( y_n \in h_\infty(x_n), \exists z_n \in \text{Liminf}_{t \to \infty} h_c(x_n) \) such that \( \|z_n - y_n\| < \frac{1}{n} \). Since \( z_n \in \text{Liminf}_{t \to \infty} h_c(x_n), n \geq 1 \), it follows that there exist \( c_n \) such that for all \( c \geq c_n, d(z_n, h_c(x_n)) < \frac{1}{n} \). In particular, \( \exists u_n \in h_c(x_n, n) \) such that \( \|z_n - u_n\| < \frac{1}{n} \). We choose the sequence \( \{c_n\}_{n \geq 1} \) such that \( c_{n+1} > c_n \) for each \( n \geq 1 \). We now have the following: \( \lim_{n \to \infty} u_n = y, u_n \in h_c(x_n) \forall n \) and \( \lim_{n \to \infty} x_n = x \). It follows directly from assumption (A5) that \( y \in h_\infty(x) \).

3 Stability and convergence of stochastic recursive inclusions

We begin by providing a brief outline of our approach to prove the stability of a SRI under assumptions (A1) – (A5). First we divide the time line, \([0, \infty)\), approximately into intervals of length \( T \). We shall explain later how we choose and fix \( T \). Then we construct a linearly interpolated trajectory from the given stochastic recursive inclusion; the construction is explained in the next paragraph. A sequence of ‘rescaled’ trajectories of length \( T \) is constructed as follows:
At the beginning of each $T$-length interval we observe the trajectory to see if it is outside the unit ball, if so we scale it back to the boundary of the unit ball. This scaling factor is then used to scale the ‘rest of the $T$-length trajectory’.

To show that the iterates are bounded almost surely we need to show that the linearly interpolated trajectory does not ‘run off’ to infinity. To do so we assume that this is not true and show that there exists a subsequence of the rescaled $T$-length trajectories that has a solution to $\dot{x}(t) \in h_\infty(x(t))$ as a limit point in $C([0, T], \mathbb{R}^d)$. We choose and fix $T$ such that any solution to $\dot{x}(t) \in h_\infty(x(t))$ with an initial value inside the unit ball is close to the origin at the end of time $T$. For example we could choose $T = T(1/8)$. We then argue that the linearly interpolated trajectory is forced to make arbitrarily large ‘jumps’ within time $T$. The Gronwall inequality is then used to show that this is not possible.

Once we prove stability of the recursion we invoke Theorem 3.6 & Lemma 3.8 from Benaim, Hofbauer and Sorin [5] to conclude that the limit set is a closed, connected, internally chain transitive and invariant set associated with $\hat{x}(t) \in h_\infty(x(t))$.

We construct the linearly interpolated trajectory $\overline{y}(t)$, for $t \in [0, \infty)$, from the sequence \{ \{x_n\} \} as follows: Define $t(0) := 0$, $t(n) := \sum_{i=0}^{n-1} a(i)$. Let $\overline{y}(t(n)) := x_n$ and for $t \in (t(n), t(n+1))$, let

$$\overline{y}(t) := \left( \frac{t(n+1) - t}{t(n+1) - t(n)} \right) \overline{y}(t(n)) + \left( \frac{t - t(n)}{t(n+1) - t(n)} \right) \overline{y}(t(n+1)).$$

We define a piecewise constant trajectory using the sequence \{ \{y_n\} \} as follows: $\overline{y}(t) := y_n$ for $t \in [t(n), t(n+1))$, $n \geq 0$.

We know that the DI given by $\hat{x}(t) \in h_\infty(x(t))$ has the origin as an attractor set. Let us fix $T := T(1/8)$, where $T(1/8)$ is as defined in section 2.1. Then, $|x(t)| < \frac{1}{2}$, for all $t \geq T_1/8$, where \{ \{x(t) : t \in [0, \infty)\} \} is a solution to $\dot{x}(t) \in h_\infty(x(t))$ with an initial value inside the unit ball around the origin.

Define $T_0 := 0$ and $T_n := \text{min}\{t(m) : t(m) \geq T_{n-1} + T\}$, $n \geq 1$. Observe that there exists a subsequence \{ \{m(n)\} \} of \{ \{n\} \} such that $T_n = t(m(n)) \forall n \geq 0$.

We construct the rescaled trajectory, $\hat{x}(t)$, $t \geq 0$, as follows: Let $t \in [T_n, T_{n+1})$ for some $n \geq 0$, then $\hat{x}(t) := \overline{y}_{t(n)}$, where $r(n) = ||\overline{y}(t(n))|| \vee 1$. Also, let $\hat{x}(T_{n+1}^-) := \lim_{t \uparrow T_{n+1}} \hat{x}(t)$, $t \in [T_n, T_{n+1})$. The corresponding ‘rescaled y iterates’ are given by $\hat{y}(t) := \overline{y}_{t(n)}^r$, and the rescaled martingale noise terms by $\hat{M}_{k+1} := \frac{M_{k+1}}{r(n)}$, $t(k) \in [T_n, T_{n+1})$, $n \geq 0$.

Consider the recursion at hand, i.e.,

$$\overline{y}(t(k+1)) = \overline{y}(t(k)) + a(k) \left( \overline{y}(t(k)) + M_{k+1} \right),$$

such that $t(k), t(k+1) \in [T_n, T_{n+1})$. Multiplying both sides by $1/r(n)$ we get the rescaled recursion:

$$\hat{x}(t(k+1)) = \hat{x}(t(k)) + a(k) \left( \hat{y}(t(k)) + \hat{M}_{k+1} \right).$$
noting that \( E \left[ \| \hat{M}_{k+1} \|^2 | \mathcal{F}_k \right] \leq K \left( 1 + \| \hat{x}(t(k)) \|^2 \right) \).

### 3.1 Characterizing limits, in \( C([0, T], \mathbb{R}^d) \), of the rescaled trajectories

The first two lemmas prove that the rescaled martingale difference noise converges almost surely. Although they can be found in Borakar [3], we present them here with proofs for the sake of completeness and easy reference.

**Lemma 1.** \( \sup_{t \in [0, T]} E \| \hat{x}(t) \|^2 < \infty \).

**Proof.** It is enough to show that
\[
\sup_{m(n) \leq k < m(n+1)} E \left[ \| \hat{x}(t(k)) \|^2 \right] \leq M,
\]
for some \( M > 0 \) that is independent of \( n \). Recall that \( T_n = t(m(n)) \) and \( T_{n+1} = t(m(n+1)) \). Let us fix \( n \) and \( k \) such that \( n \geq 0 \) and \( m(n) \leq k < m(n+1) \).

Now consider
\[
\hat{x}(t(k)) = \hat{x}(t(k-1)) + a(k-1) \left( \hat{y}(t(k-1)) + \hat{M}_k \right).
\]

Unfolding the above recursion we get,
\[
\hat{x}(t(k)) = \hat{x}(t(m(n))) + \sum_{l=m(n)}^{k-1} a(l) \left( \hat{y}(t(l)) + \hat{M}_{l+1} \right).
\]

Taking the expectation of the square of the norms on both sides of the above equation we get,
\[
E \| \hat{x}(t(k)) \|^2 = E \left[ \| \hat{x}(t(m(n))) + \sum_{l=m(n)}^{k-1} a(l) \left( \hat{y}(t(l)) + \hat{M}_{l+1} \right) \|^2 \right].
\]

It follows from the **Minkowski inequality** that,
\[
E^{1/2} \| \hat{x}(t(k)) \| \leq E^{1/2} \| \hat{x}(T_n) \|^2 + \sum_{l=m(n)}^{k-1} a(l) \left( E^{1/2} \| \hat{y}(t(l)) \|^2 + E^{1/2} \| \hat{M}_{l+1} \|^2 \right).
\]

From assumptions (A1) and (A2), it follows that for each \( m(n) \leq l < k - 1 \),
\[
\| \hat{y}(t(l)) \| \leq K(1 + \| \hat{x}(t(l)) \|) \] and \( E \left[ \| \hat{M}_{l+1} \|^2 | \mathcal{F}_l \right] \leq K \left( 1 + \| \hat{x}(t(l)) \|^2 \right), \]
respectively. We also observe that \( T_{n+1} - T_n \leq T + 1 \) (since \( \sup_n a(n) \leq 1 \)). Using these observations we obtain the following set of inequalities:
\[
E^{1/2} \| \hat{x}(t(k)) \|^2 \leq 1 + \sum_{l=m(n)}^{k-1} a(l) \left( KE^{1/2} \left( 1 + \| \hat{x}(t(l)) \|^2 \right) + \sqrt{K} E^{1/2} \left( 1 + \| \hat{x}(t(l)) \|^2 \right) \right),
\]
\[ E^{1/2} ||\dot{x}(t(k))||^2 \leq 1 + \sum_{l=m(n)}^{k-1} a(l) \left( K \left( 1 + E^{1/2} ||\dot{x}(t(l))||^2 \right) + \sqrt{K} \left( 1 + E^{1/2} ||\dot{x}(t(l))||^2 \right) \right), \]

\[ E^{1/2} ||\dot{x}(t(k))||^2 \leq \left[ 1 + (K + \sqrt{K})(T + 1) \right] + (K + \sqrt{K}) \sum_{l=m(n)}^{k-1} a(l)E^{1/2} ||\dot{x}(t(l))||^2. \]

Applying the discrete version of the Gronwall inequality we get,

\[ E^{1/2} ||\dot{x}(t(k))||^2 \leq \left[ 1 + (K + \sqrt{K})(T + 1) \right] + (K + \sqrt{K}) \left( T + 1 \right) + (K + \sqrt{K}) \left( T + 1 \right). \]

The claim follows by letting \( M = \left[ 1 + (K + \sqrt{K})(T + 1) \right] e^{(K + \sqrt{K})(T + 1)}. \]

**Lemma 2.** The rescaled sequence \( \{\hat{\zeta}_n\}_{n \geq 1} \), where \( \hat{\zeta}_n = \sum_{k=0}^{n-1} a(k) \hat{M}_{k+1} \), is convergent almost surely.

**Proof.** It is enough to prove that almost surely,

\[ \sum_{k=0}^{\infty} E \left[ a(k) \hat{M}_{k+1}^2 \mid F_k \right] < \infty. \]

Instead, we prove that

\[ E \left[ \sum_{k=0}^{\infty} a(k)^2 E \left[ \hat{M}_{k+1}^2 \mid F_k \right] \right] < \infty. \]

It follows as a consequence of assumption (A3) that,

\[ E \left[ \sum_{k=0}^{\infty} a(k)^2 E \left[ \hat{M}_{k+1}^2 \mid F_k \right] \right] \leq \sum_{k=0}^{\infty} a(k)^2 K \left( 1 + E ||\dot{x}(t(k))||^2 \right). \]

The claim now follows from assumption (A2) and Lemma [1]}

The rest of the lemmas are needed to prove the stability theorem, Theorem [2]. We begin by showing that the rescaled trajectories are bounded almost surely.

**Lemma 3.** \( \sup_{t \in [0,\infty)} ||\dot{x}(t)|| < \infty \) a.s.

**Proof.** Let \( A = \{ \omega \mid \{\hat{\zeta}_n(\omega)\}_{n \geq 1} \text{ converges} \} \), be the set on which \( \{\hat{\zeta}_n\}_{n \geq 1} \) converges. It is enough to prove that \( ||\dot{x}(t(m(n) + k))|| < K_\omega \), where \( T_n \leq t(m(n) + k) < T_{n+1} \) and \( K_\omega \) is a constant independent of \( n \). The constant may however be dependent on \( \omega \) (sample path), where \( \omega \in A \).
Let us consider the rescaled recursion:
\[ \hat{x}(t(m(n)+k)) = \hat{x}(t(m(n)+k-1)) + a(m(n)+k-1) \left( \hat{y}(t(m(n) + k - 1)) + \hat{M}_{m(n)+k} \right). \]

Unfolding the above recursion, we obtain
\[ \hat{x}(t(m(n) + k)) = \hat{x}(T_n) + \sum_{l=0}^{k-1} a(m(n) + l) \left( \hat{y}(t(m(n) + l)) + \hat{M}_{m(n)+l+1} \right). \quad (4) \]

Since \( \hat{\zeta}_n, \; n \geq 1 \), converges on \( A \), there exists \( M_w < \infty \), that may be sample path dependent, such that \( \| \sum_{l=0}^{k-1} a(m(n) + l)M_{m(n)+l+1} \| \leq M_w \), where \( M_w \) is independent of \( n \). Since \( \| \hat{x}(T_n) \| \leq 1 \) we have,
\[ \| \hat{x}(t(m(n) + k)) \| \leq 1 + \sum_{l=0}^{k-1} a(m(n) + l)\| \hat{y}(t(m(n) + l)) \| + M_w. \]

From (A1) it follows that,
\[ \| \hat{x}(t(m(n) + k)) \| \leq 1 + \sum_{l=0}^{k-1} a(m(n) + l) (1 + \| \hat{x}(t(m(n) + l)) \| ) + M_w. \]

Rearranging the terms in the above inequality and applying the discrete Gronwall inequality we get, \( \| \hat{x}(t(m(n) + k)) \| \leq (1 + M_w + (T + 1)K) e^{K(T+1)} \), a constant independent of both \( n \) and \( k \). The rest of the proof follows in a straightforward manner. \[ \square \]

Note that in the proof of the above lemma we get a bound on \( \| \hat{x}(t) \| \) that is dependent on \( T \). This is sufficient for our purposes, since we fix \( T \) to be \( T_{1/8} \).

Let \( x^n(t), \; t \in [0, T] \) be the solution (upto time \( T \)) to \( \dot{x}^n(t) = \hat{y}(T_n + t) \), with the initial condition \( x^n(0) = \hat{x}(T_n) \). Clearly, we have
\[ x^n(t) = \hat{x}(T_n) + \int_0^t \hat{y}(T_n + z) \; dz. \quad (5) \]

The following two lemmas are inspired by ideas from Benaim, Hofbauer and Sorin [8] as well as Borkar [7]. In the lemma that follows we show that the limit sets of \( \{ x^n(\cdot) \mid n \geq 0 \} \) and \( \{ \hat{x}(T_n + \cdot) \mid n \geq 0 \} \) coincide. We seek limits in \( C([0, T], \mathbb{R}^d) \).

**Lemma 4.** \( \lim_{n \to \infty} \sup_{t \in [T_n, T_n + T]} \| x^n(t) - \hat{x}(t) \| = 0 \; a.s. \)

**Proof.** Let \( t \in [(t(m(n) + k), t(m(n) + k + 1)) \) and \( t(m(n) + k + 1) \leq T_{n+1} \). We first assume that \( t(m(n) + k + 1) < T_{n+1} \). We have the following:
\[ \hat{x}(t) = \left( \frac{t(m(n) + k + 1) - t}{a(m(n) + k)} \right) \hat{x}(t(m(n)+k)) + \left( \frac{t - t(m(n) + k)}{a(m(n) + k)} \right) \hat{x}(t(m(n)+k+1)). \]
Substituting for $\dot{x}(t(m(n) + k + 1))$ in the above equation we get:

$$\dot{x}(t) = \left( \frac{t(m(n) + k + 1) - t}{a(m(n) + k)} \right) \dot{x}(t(m(n) + k)) + \left( \frac{t - t(m(n) + k)}{a(m(n) + k)} \right) \left( \dot{x}(t(m(n) + k)) + a(m(n) + k) \left( \dot{y}(t(m(n) + k)) + \dot{M}_{m(n)+k+1} \right) \right),$$

hence,

$$\dot{x}(t) = \dot{x}(t(m(n) + k)) + (t - t(m(n) + k)) \left( \dot{y}(t(m(n) + k)) + \dot{M}_{m(n)+k+1} \right).$$

Unfolding $\dot{x}(t(m(n) + k))$ over $k$ (see (4)) we get,

$$\dot{x}(t) = \dot{x}(T_n) + \sum_{l=0}^{k-1} a(m(n) + l) \left( \dot{y}(t(m(n) + l)) + \dot{M}_{m(n)+l+1} \right) + (t - t(m(n) + k)) \left( \dot{y}(t(m(n) + k)) + \dot{M}_{m(n)+k+1} \right). \quad (6)$$

Now, we consider $x^n(t)$, i.e.,

$$x^n(t) = \dot{x}(T_n) + \int_0^t \dot{y}(T_n + z) \ dz.$$

Splitting the above integral, we get

$$x^n(t) = \dot{x}(T_n) + \sum_{l=0}^{k-1} \int_{t(m(n)+l)}^{t(m(n)+l+1)} \dot{y}(z) \ dz + \int_{t(m(n)+k)}^t \dot{y}(z) \ dz.$$

Thus,

$$x^n(t) = \dot{x}(T_n) + \sum_{l=0}^{k-1} a(m(n) + l) \dot{y}(t(m(n) + l)) + (t - t(m(n) + k)) \dot{y}(t(m(n) + k)). \quad (7)$$

From (6) and (7), it follows that

$$\|x^n(t) - \dot{x}(t)\| \leq \left\| \sum_{l=0}^{k-1} a(m(n) + l) \dot{M}_{m(n)+l+1} \right\| + \left\| (t - t(m(n) + k)) \dot{M}_{m(n)+k+1} \right\|,$$

and hence,

$$\|x^n(t) - \dot{x}(t)\| \leq \left\| \dot{\zeta}_{m(n)+k} - \dot{\zeta}_{m(n)} \right\| + \left\| \dot{\zeta}_{m(n)+k+1} - \dot{\zeta}_{m(n)+k} \right\|.$$

If $t(m(n) + k + 1) = T_{n+1}$ then in the proof we may replace $\dot{x}(t(m(n) + k + 1))$ with $\dot{x}(T_{n+1})$. The arguments remain the same. Since $\dot{\zeta}_n$, $n \geq 1$, converges almost surely, the desired result follows.

The sets $\{x^n(t), t \in [0, T] \mid n \geq 0\}$ and $\{\dot{x}(T_n + t), t \in [0, T] \mid n \geq 0\}$ can be viewed as subsets of $C([0, T], \mathbb{R}^d)$. It can be shown that $\{x^n(t), t \in [0, T] \mid n \geq 0\}$ is equi-continuous and point-wise bounded. Thus from the Arzela-Ascoli theorem, $\{x^n(t), t \in [0, T] \mid n \geq 0\}$ is relatively compact. It follows from Lemma 4 that the set $\{\dot{x}(T_n + t), t \in [0, T] \mid n \geq 0\}$ is also relatively compact in $C([0, T], \mathbb{R}^d)$.
Lemma 5. Let \( r(n) \uparrow \infty \), then any limit point of \( \{ \hat{x}(T_n + t), t \in [0, T] : n \geq 0 \} \) is of the form \( x(t) = x(0) + \int_0^t y(s) \, ds \), where \( y : [0, T] \to \mathbb{R}^d \) is a measurable function and \( y(t) \in h_\infty(x(t)) \), \( t \in [0, T] \).

Proof. We define the notation \( [t] := \max \{ t(k) \mid t(k) \leq t \} \), \( t \geq 0 \). Let \( t \in [T_n, T_{n+1}) \), then \( \hat{y}(t) \in h_{r(n)}(\hat{x}([t])) \) and \( \| \hat{y}(t) \| \leq K(1 + \| \hat{x}([t]) \|) \) since \( h_{r(n)} \) is a Marchaud map (\( K \) is the constant associated with the point-wise boundedness property). It follows from Lemma 4 that \( \sup_{t \in [0, \infty)} \| \hat{y}(t) \| < \infty \) a.s. Using observations made earlier, we can deduce that there exists a sub-sequence of \( \mathbb{N} \), say \( \{ l \} \subseteq \{ n \} \), such that \( \hat{x}(T_l + t) \to x(t) \) in \( C([0, T], \mathbb{R}^d) \) and \( \hat{y}(m(t) + \cdot) \to y(\cdot) \) weakly in \( L_2([0, T], \mathbb{R}^d) \). From Lemma 3 it follows that \( x'(\cdot) \to x(\cdot) \) in \( C([0, T], \mathbb{R}^d) \).

Letting \( r(l) \uparrow \infty \) in

\[
x'(t) = x'(0) + \int_0^t \hat{y}(t(m(l) + z)) \, dz, \ t \in [0, T],
\]

we get \( x(t) = x(0) + \int_0^t y(z) \, dz \) for \( t \in [0, T] \). Since \( \| \hat{x}(T_n) \| \leq 1 \) we have \( \| x(0) \| \leq 1 \).

Since \( \hat{y}(T_l + \cdot) \to y(\cdot) \) weakly in \( L_2([0, T], \mathbb{R}^d) \), there exists \( \{ l(k) \} \subseteq \{ l \} \) such that

\[
\frac{1}{N} \sum_{k=1}^N \hat{y}(T_{l(k)} + \cdot) \to y(\cdot) \text{ strongly in } L_2([0, T], \mathbb{R}^d).
\]

Further, there exists \( \{ N(m) \} \subseteq \{ N \} \) such that

\[
\frac{1}{N(m)} \sum_{k=1}^{N(m)} \hat{y}(T_{l(k)} + \cdot) \to y(\cdot) \text{ a.e. on } [0, T].
\]

Let us fix \( t_0 \in \{ t \mid \frac{1}{N(m)} \sum_{k=1}^{N(m)} \hat{y}(T_{l(k)} + t) \to y(t), \ t \in [0, T] \} \), then

\[
\lim_{N(m) \to \infty} \frac{1}{N(m)} \sum_{k=1}^{N(m)} \hat{y}(T_{l(k)} + t_0) = y(t_0).
\]

Since \( h_\infty(x(t_0)) \) is convex and compact, to show that \( y(t_0) \in h_\infty(x(t_0)) \) it is enough to prove that \( \lim_{l(k) \to \infty} d(\hat{y}(T_{l(k)} + t_0), h_\infty(x(t_0))) = 0 \). If not, \( \exists \epsilon > 0 \) and \( \{ n(k) \} \subseteq \{ l(k) \} \) such that \( d(\hat{y}(T_{n(k)} + t_0), h_\infty(x(t_0))) > \epsilon \). Since \( \{ n(k) \} \) is norm bounded, it follows that there is a convergent sub-sequence. For the sake of convenience we assume that \( \lim_{k \to \infty} \hat{y}(T_{n(k)} + t_0) = y \), for some \( y \in \mathbb{R}^d \). Since \( \hat{y}(T_{n(k)} + t_0) \in h_{r(n(k))}(\hat{x}([T_{n(k)} + t_0])) \) and \( \lim_{k \to \infty} \hat{x}([T_{n(k)} + t_0]) = x(t_0) \), it follows from assumption (A5) that \( y \in h_\infty(x(t_0)) \). This leads to a contradiction.

Note that in the statement of Lemma 4 we can replace ‘\( r(n) \uparrow \infty \)’ by ‘\( r(l) \uparrow \infty \)’, where \( \{ r(l) \} \) is a subsequence of \( \{ r(n) \} \). Specifically we can conclude that any limit point of \( \{ \hat{x}(T_k + t), t \in [0, T] \}_{k \leq n} \) in \( C([0, T], \mathbb{R}^d) \), conditioned on \( r(k) \uparrow \infty \), is of the form \( x(t) = x(0) + \int_0^t y(z) \, dz \), where \( y(t) \in h_\infty(x(t)) \) for \( t \in [0, T] \). It should be noted that \( y(\cdot) \) may be sample path dependent.
3.2 The Stability Theorem

We are now ready to prove the stability of a SRI given by (2) under the assumptions (A1) – (A5). If \( \sup_{n} r(n) < \infty \), then the iterates are stable and there is nothing to prove. If on the other hand \( \sup_{n} r(n) = \infty \), there exists \( \{l\} \subseteq \{n\} \) such that \( r(l) \uparrow \infty \). It follows from Lemma 5 that any limit point of \( \{\hat{x}(T_{l} + t), t \in [0, T] : \{l\} \subseteq \{n\}\} \) is of the form \( x(t) = x(0) + \int_{0}^{t} y(s) \, ds \), where \( y(t) \in h_{\infty}(x(t)) \) for \( t \in [0, T] \). From assumption (A4), we have that \( \|x\| < 1/8 \) (\( T = T(1/8) \)). Since the time intervals are roughly \( T \) apart, for large values of \( r(n) \) we can conclude that \( \|\hat{x}(T_{n+1}^{-})\| < \frac{1}{4} \), where \( \hat{x}(T_{n+1}^{-}) = \lim_{t \uparrow t_{m(n+1)}} \hat{x}(t), \ t \in [T_{n}, T_{n+1}] \).

**Theorem 1** (Stability Theorem for DI). Under assumptions (A1) – (A5), \( \sup_{n} \|x(n)\| < \infty \) a.s.

**Proof.** As explained earlier it is sufficient to consider the case when \( \sup_{n} r(n) = \infty \). Let \( \{l\} \subseteq \{n\} \) such that \( r(l) \uparrow \infty \). Recall that \( T_{l} = t(m(l)) \) and that \( [T_{l} + T] = \max\{t(k) \mid t(k) \leq T_{l} + T\} \).

We have \( \|x(T)\| < \frac{1}{4} \) since \( x(t) \) is a solution, up to time \( T \), to the DI given by \( \hat{x}(t) \in h_{\infty}(x(t)) \) and we have fixed \( T = T(1/8) \). From Lemma 5, we conclude that there exists \( N \) such that all of the following happen:

(i) \( m(l) \geq N \implies \|\hat{x}([T_{l} + T])\| < \frac{1}{4} \).

(ii) \( n \geq N \implies a(n) < \frac{1}{8[K(1+K_{\omega})+M_{\omega}]} \).

(iii) \( n > m \geq N \implies \|\hat{\zeta}_{n} - \hat{\zeta}_{m}\| < M_{\omega} \).

(iv) \( m(l) \geq N \implies r(l) > 1 \).

In the above \( K_{\omega} \) and \( M_{\omega} \) are as explained in Lemma 3.

Let \( m(l) \geq N \) and \( t(m(l)+1) = t(m(l)+k+1) \) for some \( k \geq 0 \). Clearly from the manner in which the \( T_{n} \) sequence is defined, we have \( t(m(l)+k) = [T_{l} + T] \).

As defined earlier \( \hat{x}(T_{n+1}^{-}) = \lim_{t \uparrow t_{m(n+1)}} \hat{x}(t), \ t \in [T_{n}, T_{n+1}] \) and \( n \geq 0 \). Consider the equation

\[
\hat{x}(T_{n+1}^{-}) = \hat{x}(t(m(l)+k)) + a(m(l)+k) \left( \hat{y}(t(m(l)+k)) + \hat{M}_{m(l)+k+1} \right).
\]

Taking norms on both sides we get,

\[
\|\hat{x}(T_{n+1}^{-})\| \leq \|\hat{x}(t(m(l)+k))\| + a(m(l)+k)\|\hat{y}(t(m(l)+k))\| + a(m(l)+k)\|\hat{M}_{m(l)+k+1}\|.
\]

From the way we have chosen \( N \) we conclude that:

\[
\|\hat{y}(t(m(l)+k))\| \leq K (1 + \|\hat{x}(t(m(l)+k))\|) \leq K (1 + K_{\omega}) \text{ and that } \|\hat{M}_{m(l)+k+1}\| = \|\hat{\zeta}_{m(l)+k+1} - \hat{\zeta}_{m(l)+k}\| \leq M_{\omega}.
\]

13
Thus we have that,
\[ \|\hat{x}(T_{i+1})\| \leq \|\hat{x}(t(m(l) + k))\| + a(m(l) + k)(K(1 + K_\omega) + M_\omega). \]
Finally we have that \( \|\hat{x}(T_{i+1})\| < \frac{1}{2} \) and
\[ \frac{\|\hat{x}(T_{i+1})\|}{\|\hat{x}(T_i)\|} = \frac{\|\hat{x}(T_{i+1})\|}{\|\hat{x}(T_i)\|} < \frac{1}{2}. \]

Let \( t \leq T_l, t \in [T_n, T_{n+1}) \) and \( n+1 \leq l \), be the last time that \( x(t) \) jumps from inside the unit ball to the outside. From (8) we know that \( \|x(T_{n+1})\| < \|x(T_n)\| \) if \( \|x(T_n)\| \geq 1 \). This implies that the trajectory falls exponentially till it enters the unit ball. Since we have assumed that \( \hat{x}(T) \) is the last time before \( T_l \) when the trajectory jumps out of the unit ball, we have \( \hat{x}(T_{n+1}) \geq \|\hat{x}(T_l)\| \). Since \( r(l) \uparrow \infty \), \( \hat{x}(t) \) would be forced to make larger and larger jumps within an interval of \( T + 1 \). However, the maximum jump within any fixed time interval can be bounded by the Gronwall inequality. This leads to a contradiction.

We now state one of the main theorems of this paper.

**Theorem 2.** Under assumptions (A1) – (A5), almost surely, the sequence \( \{x_n\}_{n \geq 0} \) generated by the stochastic recursive inclusion, given by (3), is bounded and converges to a closed, connected, internally chain transitive and invariant set of \( \hat{x}(t) \in h(x(t)) \).

**Proof.** The stability of the iterates is shown in Theorem 1. The convergence can be proved under assumptions (A1) – (A3) and the stability of the iterates in exactly the same manner as in Theorem 3.6 & Lemma 3.8 of Benaim, Hofbauer and Sorin [5].

We have thus far shown that under assumptions (A1) – (A5) the SRI given by (2) is stable and converges to a closed, connected, internally chain transitive and invariant set.

### 4 An Application of Theorem 2

#### 4.1 The problem of approximate drifts in stochastic recursive equations

Let us recall the following SRE:
\[ x_{n+1} = x_n + a(n) (h(x_n) + M_{n+1}), \]
where \( h : \mathbb{R}^d \rightarrow \mathbb{R}^d \) is Lipschitz continuous, \( \{a(n)\}_{n \geq 0} \) is the step-size sequence and \( \{M_n\}_{n \geq 1} \) is the noise sequence.

The function \( h \) is colloquially referred to as the drift. In many applications the drift cannot be calculated accurately. This is referred to as the approximate drift problem. For more details the reader is referred to Chapter 5.3 of Borkar [7]. Suppose the room for error is at most \( \epsilon(> 0) \) then such an algorithm can be characterized by the following stochastic recursive inclusion:
\[ x_{n+1} = x_n + a(n) (y_n + M_{n+1}), \]
where \( y_n \in h(x_n) + \overline{B}_\epsilon(0) \) is an estimate of \( h(x_n) \) and \( \overline{B}_\epsilon(0) \) is the closed ball of radius \( \epsilon \) around the origin. We define a new set-valued map called the approximate drift by \( H(x) := h(x) + \overline{B}_\epsilon(0) \) for each \( x \in \mathbb{R}^d \). In the following discussion we assume that \( \epsilon \geq 0 \). When \( \epsilon = 0 \), the approximate drift algorithm described by (10) is really the SRE given by (9).

In this section we answer the following question: Suppose the recursion given by (9) is stable and convergent, what can we say about the approximate drift version described by (10)? We shall show that if the recursion (9) satisfies assumptions (A1)–(A5). We then invoke Theorem 2 to conclude that the iterates converge to a closed, connected, internally chain transitive and invariant set associated with \( \dot{x}(t) \in h(x(t)) + \overline{B}_\epsilon(0) = H(x(t)) \).

Before we proceed we recall the assumptions and the statement of the Borkar-Meyn Theorem [8].

(BM1) (i) The function \( h : \mathbb{R}^d \to \mathbb{R}^d \) is Lipschitz continuous, with Lipschitz constant \( L \). There exists a function \( h_{\infty} : \mathbb{R}^d \to \mathbb{R}^d \) such that \( \lim_{c \to \infty} \frac{h(cx)}{c} = h_{\infty}(x), \) for each \( x \in \mathbb{R}^d \).
(ii) \( h_c \to h_{\infty} \) uniformly on compacts, as \( c \to \infty \).
(iii) The o.d.e. \( \dot{x}(t) = h_{\infty}(x(t)) \) has the origin as the unique globally asymptotically stable equilibrium.

(BM2) \( \{a(n)\}_{n \geq 0} \) is a scalar sequence such that: \( a(n) \geq 0, \sum_{n \geq 0} a(n) = \infty \) and \( \sum_{n \geq 0} a(n)^2 < \infty \). Without loss of generality, we assume that \( \sup_n a(n) \leq 1 \).

(BM3) \( \{M_n\}_{n \geq 1} \) is a martingale difference sequence with respect to the filtration \( \mathcal{F}_n := \sigma(x_0, M_1, \ldots, M_n), n \geq 0 \). Thus, \( E[M_{n+1}|\mathcal{F}_n] = 0 \) a.s., \( \forall n \geq 0 \).
\( \{M_n\} \) is also square integrable with \( E[\|M_{n+1}\|^2|\mathcal{F}_n] \leq L(1 + \|x_n\|^2) \), for some constant \( L > 0 \). Without loss of generality, assume that the same constant, \( L \), works for both (BM1)(i) and (BM3).

Theorem 3 (Borkar-Meyn Theorem). Suppose (BM1)-(BM3) hold. Then \( \|x_n\| < \infty \) almost surely. Further, the sequence \( \{x_n\} \) converges almost surely to a (possibly sample path dependent) compact connected internally chain transitive invariant set of \( \dot{x}(t) = h(x(t)) \).

We assume that the recursion given by (9) satisfies assumptions (BM1)–(BM3). As explained earlier we have to prove that the approximate drift version given by (10) satisfies (A1)–(A5). We begin by proving that \( H(x) = h(x) + \overline{B}_\epsilon(0) \) is a Marchaud map. Since \( \overline{B}_\epsilon(0) \) is convex and compact for each \( x \in \mathbb{R}^d \), it follows that \( H(x) \) is convex and compact. Fix \( x \in \mathbb{R}^d \) and \( y \in H(x) \), then \( \|y\| \leq \|h(x)\| + \epsilon \) and \( \|y\| \leq \|h(0)\| + L\|x - 0\| + \epsilon \) since \( h \) is Lipschitz continuous with Lipschitz constant \( L \). If we set \( K := (\|h(0)\| + \epsilon) \vee L \), then we get \( \|y\| \leq K(1 + \|x\|) \). This shows that \( H \) is point-wise bounded. We now show that \( H \) is upper-semicontinuous. Assume \( \lim_{n \to \infty} x_n = x \). Since \( y_n \in h(x_n) + \overline{B}_\epsilon(z_n) \) for each \( n \geq 1 \). Since \( y_n \in H(x_n) \), we conclude that \( y_n = h(x_n) + z_n \).
for some $z_n \in \overline{B}_c(0)$, $n \geq 1$. Since $\{y_n\}_{n \geq 1}$ and $\{h(x_n)\}_{n \geq 1}$ are convergent sequences, it follows that $\{z_n\}_{n \geq 1}$ is also convergent. Let $z := \lim_{n \to \infty} z_n$. Since $\overline{B}_c(0)$ is compact we have that $z \in \overline{B}_c(0)$. Taking limits on both sides of $y_n = h(x_n) + z_n$, we get $y = h(x) + z$. Thus $y \in H(x)$.

It follows directly from $(BM2)$ and $(BM3)$ that assumptions $(A2)$ and $(A3)$ are satisfied by the approximate drift algorithm $(10)$.

Before showing that $(10)$ satisfies $(A4)$, we construct the following family of set-valued maps: $H_c(x) := \{h(x) + \frac{\epsilon}{c} y \mid y \in \overline{B}_c(0)\}$, i.e., $H_c(x) = h_c(x) + \overline{B}_c(0)$ for each $x \in \mathbb{R}^d$. We claim that the lower limit, $\liminf_{c \to \infty} H_c(x) = \{h_\infty(x)\}$. Since $\lim_{c \to \infty} d(h_c(x), H_\infty(x)) \leq \lim_{c \to \infty} ||h_\infty(x) - h_c(x)|| = 0$ it follows that $h_\infty(x) \in H_\infty(x)$. Now we show that $h_\infty(x)$ is the only element of $H_\infty(x)$. Let $y \in H_\infty(x)$, then $\exists z_n \in H_n(x), n \geq 1$, such that $\lim_{n \to \infty} z_n = y$. We have the following inequality:

$$||y - h_\infty(x)|| \leq ||y - z_n|| + ||z_n - h_n(x)|| + ||h_n(x) - h_\infty(x)||.$$ Letting $n \to \infty$ in the above inequality, we get $||y - h_\infty(x)|| = 0$, in other words $H_\infty(x) = \{h_\infty(x)\}$.

The differential inclusion $\dot{x}(t) \in H_\infty(x(t))$ is really the o.d.e. $\dot{x}(t) = h_\infty(x(t))$. Since the origin is assumed to be the unique globally asymptotically stable equilibrium point of $\dot{x}(t) = h_\infty(x(t))$, it follows that the origin is also Lyapunov stable. This can be used to show that the origin is an attracting set and that $\overline{B}_1(0)$ is its fundamental neighborhood. For a proof of this fact the reader is referred to Lemma 1 from Chapter 3 of Borkar [1].

Finally we show that $(A5)$ is satisfied. Let $e_n \uparrow \infty,$ $\lim_{n \to \infty} x_n = x$, $\lim_{n \to \infty} y_n = y$, and $y_n \in H_{e_n}(x_n)$ for each $n \geq 1$. We need to show that $y \in H_\infty(x)$ ($y = h_\infty(x)$). Consider the following set of inequalities:

$$\|y - h_\infty(x)\| \leq \|y - y_n\| + \|y_n - h_{e_n}(x_n)\| + \|h_{e_n}(x_n) - h_\infty(x)\|,$$

$$\|y - h_\infty(x)\| \leq \|y - y_n\| + \frac{\epsilon}{e_n} + \|h_{e_n}(x) - h_\infty(x)\| + \|h_{e_n}(x_n) - h_{e_n}(x)\|.$$ The second inequality follows from the first since $\|h_{e_n}(x_n) - h_\infty(x)\| \leq \|h_{e_n}(x_n) - h_{e_n}(x)\| + \|h_{e_n}(x) - h_\infty(x)\|$ and $\|y_n - h_\infty(x)\| \leq \frac{\epsilon}{e_n}$. Letting $n \to \infty$ in the second inequality we get $y = h_\infty(x)$. Thus $(A5)$ is satisfied.

Since assumptions $(A1)$–$(A5)$ are satisfied the following is a direct corollary to Theorem 2.

**Corollary 1.** If a SRE, given by $(9)$, satisfies $(BM1)$–$(BM3)$ then the corresponding approximate drift algorithm, given by $(10)$, is stable almost surely. In addition, it converges to a closed, connected, invariant and internally chain transitive set of $\dot{x}(t) \in H(x(t))$, where $H(x) = h(x) + \overline{B}_c(0)$.

Since the result holds true even when $\epsilon = 0$, it follows that the Borkar-Meyn Theorem emerges as a special case of Theorem 2.
4.2 Relaxation of assumption (BM1) of the Borkar-Meyn Theorem

The assumptions involved in the Borkar-Meyn Theorem are listed in Section 1.1. One of the assumptions, (BM1)(ii), requires that \( h_c \to h_\infty \) uniformly on compacts. In this section we discuss how to dispense with assumption (BM1)(ii).

As discussed in Section 1.1, the Borkar-Meyn Theorem emerges as a special case of Theorem [8]. This happens when \( \epsilon = 0 \) and \( H(x) = \{h(x)\} \) for all \( x \in \mathbb{R}^d \). Hence the SRI given by (10) is same as the SRE given by (9). In the foregoing discussion \( H \) and \( h \) are as defined in Section 1.1. In the previous section we have shown that if (9) satisfies assumptions (BM1) - (BM3) then it also satisfies assumptions (A1) - (A5). In proving this we did not use (BM1)(ii) i.e., \( h_c \to h_\infty \) uniformly on compacts.

Now, we discuss in brief how we work around using (BM1)(ii) in proving the Borkar-Meyn Theorem. The notations used in this paragraph are consistent with those found in Chapter 3 of Borkar [7]. We list a few below.

1. \( \phi_n(\cdot, x) \) denotes the solution to \( \dot{x}(t) \in h_{r(n)}(x(t)) \) with initial value \( x \).
2. \( \phi_\infty(\cdot, x) \) denotes the solution to \( \dot{x}(t) \in h_\infty(x(t)) \) with initial value \( x \).
3. \( x^n(t), t \in [0, T] \) denotes the solution to \( \dot{x}^n(t) = h_{r(n)}(\dot{x}(T_n + t)) \) with initial value \( x^n(0) = \hat{x}(T_n) \). Then \( x^n(t) = \phi_n(t, \hat{x}(T_n)) \), \( t \in [0, T] \).

For more details the reader is referred to Borkar and Meyn [8] or Chapter 3 of Borkar [7]. In proving the Borkar-Meyn Theorem as outlined in [8] (BM1)(ii) is used to show that for large values of \( r(n) \), \( \phi_n(t, \hat{x}(T_n)) \) is ‘close’ to \( \phi_\infty(t, \hat{x}(T_n)) \), \( t \in [0, T] \). Here we deviate from [8] in the definition of \( x^n(t) \), \( t \in [0, T] \). In this paper, \( x^n(\cdot) \) denotes the solution up to time \( T \) to \( \dot{x}^n(t) = \hat{y}(T_n + t) = h_{r(n)}(\hat{x}(T_n + t)) \) with \( x^n(0) = \hat{x}(T_n) \), where \( [\cdot] \) is defined in Lemma 5. In other words, we have the following:

\[
x^n(t) = \hat{x}(T_n) + \sum_{i=0}^{k-1} \int_{t(m(n)+l)}^{t(m(n)+l+1)} \hat{y}(z) \, dz + \int_{t(m(n)+k)}^{t} \hat{y}(z) \, dz.
\]

For \( t \in [t_n, t_{n+1}] \), \( \hat{y}(t) \) is a constant and equals \( \hat{y}(t_n) \). We get the following:

\[
x^n(t) = \hat{x}(T_n) + \sum_{i=0}^{k-1} a(m(n) + l)h_{r(n)}(\hat{x}(t[m(n)+l])) + (t - t(m(n) + k)) h_{r(n)}(\hat{x}(t[m(n) + k])).
\]

If the Borkar-Meyn Theorem is proven along of lines of Section 4.2 i.e., Lemmas 1-5 and Theorem 1 then we essentially show the following: If \( r(n) \uparrow \infty \) then the T-length trajectories given by \( \{x^n(\cdot), n \geq 0\} \) have \( \phi_\infty(x, t), t \in [0, T] \), as the limit point in \( C([0, T], \mathbb{R}^d) \), where \( x \in B_1(0) \). This is proven in Lemmas 4 and 5 the proofs of which do not require (BM1)(ii).
5 Another Stability Theorem for Stochastic Recursive Inclusions

In (A4) we assumed that Liminf$_c \rightarrow \infty h_c(x)$ is nonempty for all $x \in \mathbb{R}^d$. In this section we shall develop a stability criterion for the case when we can no longer make such an assumption. In other words, we work with a modified version of assumption (A4).

5.1 A Modification of Assumption (A4)

Recall the following SRI:

$$x_{n+1} = x_n + a(n) |y_n + M_{n+1}|, \text{ for } n \geq 0.$$ (11)

Since $h_c$ is point-wise bounded for each $c \geq 1$, we have $\sup_{y \in h_c(x)} \|y\| \leq K(1 + \|x\|)$, where $x \in \mathbb{R}^d$ (see Proposition 1). This implies that $\{y_c\}_{c \geq 1}$, where $y_c \in h_c(x)$, has at least one convergent subsequence. It follows from the definition of upper-limit of a sequence of sets (see Section 2.1) that Limsup$_{c \rightarrow \infty} h_c(x)$ is non-empty for every $x \in \mathbb{R}^d$. It is worth noting that Liminf$_{c \rightarrow \infty} h_c(x) \subseteq$ Limsup$_{c \rightarrow \infty} h_c(x)$ for every $x \in \mathbb{R}^d$. Another important point to consider is that the lower-limits of sequences of sets are harder to compute than their upper-limits, see Aubin $^2$ for more details.

Recall that $h_c(x) = \{y \mid cy \in h(cx)\}$, where $x \in \mathbb{R}^d$ and $c \geq 1$. Clearly the upper-limit, Limsup$_{c \rightarrow \infty} h_c(x) = \{y \mid \lim_{c \rightarrow \infty} d(y, h_c(x)) = 0\}$ is nonempty for every $x \in \mathbb{R}^d$. For $A \subseteq \mathbb{R}^d$, cl$(A)$ denotes the closure of the convex hull of $A$, i.e., the closure of the smallest convex set containing $A$.

Define $h_\infty(x) := \overline{\text{cl}}(\text{Limsup}_{c \rightarrow \infty} h_c(x))$.

Below we state the modification of assumption (A4) that we call (A6).

(A6) The differential inclusion $\dot{x}(t) \in h_\infty(x(t))$ has the origin as an attracting set and $\overline{B}_1(0)$ is a subset of some fundamental neighborhood of the origin.

Note that in (A4), $h_\infty(x) := \overline{\text{Liminf}_{c \rightarrow \infty} h_c(x)}$ while in (A6), $h_\infty(x) := \overline{\text{cl}}(\text{Limsup}_{c \rightarrow \infty} h_c(x))$. In this section we shall work with this new definition of $h_\infty$.

Proposition 2. $h_\infty$ is a Marchaud map.

Proof. From the definition of $h_\infty$ it follows that $h_\infty(x)$ is convex, compact for all $x \in \mathbb{R}^d$ and $h_\infty$ is point-wise bounded. It is left to prove that $h_\infty$ is an upper-semicontinuous map.

Let $x_n \rightarrow x$, $y_n \rightarrow y$ and $y_n \in h_\infty(x_n)$, for all $n \geq 1$. We need to show that $y \in h_\infty(x)$. We present a proof by contradiction. Since $h_\infty(x)$ is convex and compact, $y \notin h_\infty(x)$ implies that there exists a linear functional on $\mathbb{R}^d$, say $f$, such that $\sup_{z \in h_\infty(x)} f(z) \leq \alpha - \epsilon$ and $f(y) \geq \alpha + \epsilon$, for some $\alpha \in \mathbb{R}$.
The statements of the second stability theorem (Theorem 4) we address the following question: If \( \lim h_c(x) \) does not exist for all \( x \in \mathbb{R}^d \), then what are the sufficient conditions for the stability and convergence of the algorithm?
We take our cue from assumption (A6) in constructing the following replacement for (BM1) that we call (BM4).

(BM4) (i) The function \( h : \mathbb{R}^d \to \mathbb{R}^d \) is Lipschitz continuous, with Lipschitz constant \( L \). Define the set-valued map, \( h_\infty(x) := \overline{\text{lim}_{c \to \infty}}(\text{Limsup}_{c \to \infty} \{ h_c(x) \}) \), where \( x \in \mathbb{R}^d \).

Note that \( \text{Limsup}_{c \to \infty} \{ h_c(x) \} = \{ y | \lim_{c \to \infty} \| h_c(x) - y \| = 0 \} \).

(ii) The differential inclusion \( \dot{x}(t) \in \text{Limsup}_c(x(t)) \) has the origin as an attractor set and \( \overline{B}_1(0) \) is a subset of its fundamental neighborhood.

Recall that the function \( \text{Limsup} \) is defined for a sequence of sets (see Section 2.1) and \( h_c \) is a single valued map. For any \( x \in \mathbb{R}^d \) and \( c \geq 1 \), \( \{ h_c(x) \} \) is a set of cardinality one, hence \( \text{Limsup}_{c \to \infty} \{ h_c(x) \} \) is well-defined.

Let us recall the following SRE:

\[
x_{n+1} = x_n + a(n) \left( h(x_n) + M_{n+1} \right), \quad \text{for } n \geq 0.
\]

Observe that \( \text{Limsup}_{c \to \infty} \{ h_c(x) \} = \lim_{c \to \infty} h_c(x) \) when \( \lim_{c \to \infty} h_c(x) \) exists for each \( x \in \mathbb{R}^d \), where \( \text{Limsup} \) is the upper-limit of a sequence of sets (see section 2.1). It can be shown that if a recursion given by (12) satisfies assumptions (BM1)(i) and (BM1)(iii) then it also satisfies (BM4). Assumption (BM4) unifies the two possible cases: when the limit of \( h_c \), as \( c \to \infty \), exists and when it does not.

We claim that a recursion given by (12), satisfying assumptions (BM2), (BM3) and (BM4) will also satisfy (A1) – (A3), (A6) and (A5) (see section 5.1). From Theorem 3 it follows that the iterates are stable and converge to a closed, connected, internally chain transitive and invariant set of \( \dot{x}(t) = h(x(t)) \). The following generalization of the Borkar-Meyn Theorem is a direct consequence of Theorem 3.

Corollary 2 (Generalized Borkar-Meyn Theorem). Under assumptions (BM2), (BM3) and (BM4), almost surely the sequence \( \{ x_n \}_{n \geq 0} \) generated by the stochastic recursive equation (12), is bounded and converges to a closed, connected, internally chain transitive and invariant set of \( \dot{x}(t) = h(x(t)) \).

Proof. Assumptions (A1) – (A3) and (A6) follow directly from (BM2), (BM3) and (BM4). We show that (A5) is also satisfied. Let \( c_n \uparrow \infty, x_n \to x, y_n \to y \) and \( y_n \in h_{c_n}(x_n) \) (here \( y_n = h_{c_n}(x_n) \), \( \forall n \geq 1 \). It can be shown that \( \| h_{c_n}(x_n) - h_{c_n}(x) \| \leq L \| x_n - x \| \). Hence we get that \( h_{c_n}(x) \to y \). In other words, \( \lim_{c \to \infty} \| h_c(x) - y \| = 0 \). Hence we have \( y \in h_\infty(x) \). The claim now follows from Theorem 3.

In Section 4.1 we showed that the approximate drift version of an algorithm is stable and convergent if the original algorithm satisfies (BM1)(i), (iii), (BM2) and (BM3) (assumptions of the Borkar-Meyn Theorem). Similarly, we claim if a SRE given by (12) satisfies (BM2), (BM3) and (BM4), then the corresponding approximate drift version satisfies (A1) – (A3), (A6) and (A5). The arguments involved are similar to those discussed in Section 4.1. We omit the proof to avoid repetition. It follows from Theorem 3 that the iterates are stable and
converge to a closed, connected, internally chain transitive and invariant set of \( \dot{x}(t) \in h(x(t)) + \mathcal{F}(0) \).

5.3 Relaxation of Assumptions (A4) and (A6)

In assumptions (A4) and (A6), the infinity system \( (\dot{x}(t) \in h_\infty(x(t))) \) is assumed to have the origin as an attracting set. Here, we discuss how we can relax this assumption. In the proof of Theorem 1, we used \( \|x(T_{n+1})\| < \frac{1}{2} \) to assert that the trajectory, \( \mathcal{T}(\cdot) \), falls exponentially to within the unit ball. This helped us in proving the stability of the iterates. The same conclusion can be drawn even if we are able to prove that \( \|x(T_{n+1})\| < \delta \), for some \( 0 < \delta < 1 \). Once stability is established, ‘convergence’ of the algorithm follows, under appropriate assumptions, from Theorems 2 and 4. In this section we show that Theorems 2, 4 and Corollary 2 are valid even when the attracting set is ‘sufficiently’ close to the origin.

We modify assumptions (A4), (A6) and (BM4) as follows: In the aforementioned assumptions, we now assume that the associated differential inclusion, \( \dot{x}(t) \in h_\infty(x(t)) \), has an attracting set, \( \mathcal{A} \), such that \( \sup_{x \in \mathcal{A}} \|x\| < 1 \). Further, we assume that \( \overline{B}_1(0) \) is a subset of some fundamental neighborhood of \( \mathcal{A} \).

In what follows it should be noted that the definition of \( h_\infty \) changes with the assumptions being made. For example, when (MA4) is used we let \( h_\infty(x) := \liminf_{c \to \infty} h_c(x) \); when (MA6) is used we let \( h_\infty(x) := \limsup_{c \to \infty} h_c(x) \) and when (BM4) is used we let \( h_\infty(x) := \limsup_{c \to \infty} \{h_c(x)\} \).

**Proposition 3.**

1. Under assumptions (A1) – (A3), (MA4) and (A5), the statement of Theorem 2 is true.

2. Under assumptions (A1) – (A3), (MA6) and (A5), the statement of Theorem 4 is true.

3. Under assumptions (BM2), (BM3) and (MBM4), the statement of Corollary 2 is true.

**Proof.** We merely highlight the differences as the proofs essentially remain the same. Define \( \delta_1 := \sup_{x \in \mathcal{A}} \|x\| \) and choose \( \delta_2, \delta_3 \) and \( \delta_4 \) such that \( 0 \leq \delta_1 < \delta_2 < \delta_3 < \delta_4 < 1 \). Recall that \( T(\delta) \) is the time beyond which any solution
to the differential inclusion given by \( \dot{x}(t) \in h_\infty(x(t)) \), with starting point in the fundamental neighborhood of \( \mathcal{A} \), remains within \( N^δ(\mathcal{A}) \). Note that \( h_\infty \) is appropriately interpreted based on the theorem being proven. We define \( T := T(\delta_2 - \delta_1) \). The statements of Lemma 1 – Lemma 5 are true with the aforementioned definition of \( T \). In the proof of Theorem 1 we choose \( N \) such that the following hold: \( \forall \ m(l) \geq N, \|\dot{x}(T_l + T)\| < \delta_3 \); \( \forall \ n \geq N, \ a(n) < \frac{\delta_4 - \delta_3}{K(1 + K) + M_ω} \); \( \forall \ n > m \geq N, \|\zeta_n - \zeta_m\| < M_ω \) and \( \forall \ m(l) \geq N, r(l) > 1 \). Note that \( K_ω \) and \( M_ω \) are as explained in Lemma 3. Using similar arguments as before, we can conclude that \( \|\dot{x}(T_{l+1})\| < \delta_4 \), where \( l \) is such that \( m(l) \geq N \). Finally we can conclude that,

\[
\frac{\|\pi(T_{l+1})\|}{\|\dot{x}(T_l)\|} = \frac{\|\dot{x}(T_{l+1})\|}{\|\dot{x}(T_l)\|} < \delta_4.
\]

Since \( \delta_4 < 1 \), we can deduce that the trajectory falls exponentially till it hits the unit ball around the origin. The rest of the proof follows in a similar manner as before.

### 6 Conclusions

This paper presents an extension to the theorem of Borkar and Meyn that includes the case where the mean field is a set-valued map. Two different sets of assumptions are discussed which guarantee the ‘stability and convergence’ of a stochastic recursive inclusion. As an immediate application of Theorem 2 a solution to the ‘approximate drift problem’ is discussed. Further, the assumptions of the Borkar-Meyn Theorem are relaxed as a consequence of Theorem 2. As a corollary to Theorem 4 a generalization of the Borkar-Meyn Theorem is presented which includes the case when \( \lim_{c \to \infty} h_c(x) \) does not exist for all \( x \in \mathbb{R}^d \).

An important future direction would be to extend these results to the case when the set-valued drift is governed by a Markov process in addition to the iterate sequence. For the case of stochastic approximations, such a situation has been considered in [7], Chapter 6, where the Markov ‘noise’ is tackled using the ‘natural timescale averaging’ properties of stochastic approximation.

### References

[1] J. Aubin and A. Cellina. *Differential Inclusions: Set-Valued Maps and Viability Theory*. Springer, 1984.

[2] J. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, 1990.

[3] M. Benaïm. Dynamics of stochastic approximation algorithms. *Séminaire de probabilités XXXIII*, pages 1–68, 1999.

[4] M. Benaïm. A dynamical system approach to stochastic approximations. *SIAM J. Control Optim.*, 34(2):437–472, 1996.

[5] M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, pages 328–348, 2005.
[6] S. Bhatnagar. The Borkar-Meyn theorem for asynchronous stochastic approximations. *Systems & Control Letters*, 60(7):472–478, 2011.

[7] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, 2008.

[8] V. S. Borkar and S.P. Meyn. The O.D.E. method for convergence of stochastic approximation and reinforcement learning. *SIAM J. Control Optim*, 38:447–469, 1999.

[9] H. Kushner and G.G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.

[10] L. Ljung. Analysis of recursive stochastic algorithms. *Automatic Control, IEEE Transactions on*, 22(4):551–575, 1977.