ABSTRACT

Title of Dissertation: STOCHASTIC APPROXIMATION AND OPTIMIZATION FOR MARKOV CHAINS

John David Bartusek, Doctor of Philosophy, 2000

Dissertation directed by: Professor Armand M. Makowski
Department of Electrical Engineering

We study the convergence properties of the projected stochastic approximation (SA) algorithm used to find the root of an unknown steady state function of a parameterized family of Markov chains. The analysis is based on the ODE Method and we develop a set of application-oriented conditions which imply almost sure convergence and are *verifiable* in terms of typically available model data. Specific results are obtained for geometrically ergodic Markov chains satisfying a uniform Foster-Lyapunov drift inequality.

Stochastic optimization is a direct application of the above root finding problem if the SA is driven by a gradient estimate of steady state performance. We study the convergence properties of an SA driven by a gradient estimator which observes an increasing number of samples from the Markov chain at each step of the SA's recursion. To show almost sure convergence to the optimizer, a framework of verifiable conditions is introduced which builds on the general SA conditions proposed for the root finding problem.

We also consider a difficulty sometimes encountered in applications when selecting the set used in the projection operator of the algorithm. There often exists a well behaved positive recurrent region of the state process parameter space where the convergence conditions are satisfied; this being the ideal set to project on. Unfortunately, the boundaries of this projection set are usually not known *a priori* when implementing the SA. Therefore, we consider the convergence properties when the projection set is chosen to include regions outside the well behaved region. Specifically, we consider an SA applied to an M/M/1 which adjusts the service rate parameter when the projection set includes parameters which cause the queue to be transient.

Finally, we consider an alternative SA where the recursion is driven by a sample average of observations. We develop conditions implying convergence for this algorithm which are based on a uniform large deviation upper bound and we present specialized conditions implying this property for finite state Markov chains.

STOCHASTIC APPROXIMATION AND OPTIMIZATION FOR MARKOV CHAINS

by

John David Bartusek

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2000

Advisory Committee:

Professor Armand M. Makowski, Chairman/Advisor
Professor Michael C. Fu
Professor Steven I. Marcus
Professor Adrianos Papamarcou
Professor Mark A. Shayman

DEDICATION

To Lisa

ACKNOWLEDGEMENTS

I wish to sincerely thank my advisor Dr. Armand M. Makowski for his guidance, inspiration, and patience over the course of this project. His many ideas and suggestions were always very helpful and stimulating. Also, thanks go to Marie Duflo for some helpful email correspondence and the Electrical Engineering Department for providing me with several Teaching Assistantships. Special thanks go to Dr. Steven Marcus, Dr. Michael Fu, Dr. Mark Shayman and Dr. Adrian Papamarcou for serving on my dissertation committee. Thanks also to my friends and family who provided encouragement over the years. Especially, I wish to thank my patient and loving wife Lisa; without her steady encouragement and support, completion of this endeavor would certainly not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

In many contexts, it is necessary to find parameter values $\theta^\star$ which satisfy a nonlinear equation of the form

$$h(\theta) = 0, \quad \theta \in \Theta \tag{1.1}$$

for some mapping $h : \Theta \to \mathbb{R}^p$ where Θ is a subset of $\mathbb{R}^p$. A typical example arises when optimizing a performance measure $J : \Theta \to \mathbb{R}$; a task often equivalent to setting the gradient of J to zero. At other times it is desirable to maintain system performance at some given level $J^\star$, and this points to solving (1.1) with $h(\theta) = J(\theta) - J^\star$.

The overwhelming majority of methods for solving (1.1) are recursive in nature and produce a sequence of iterates $\{\theta_n,\ n = 0, 1, \ldots\}$ which eventually converge to the desired value(s) $\theta^\star$: Starting with an initial guess θ_0, the $(n+1)^{rst}$ iterate θ_{n+1} is computed on the basis of the previous iterate θ_n and past values of h, say $h(\theta_i)$, $i = 0, 1, \ldots, n$, (and sometimes derivatives of h at these points).

Unfortunately, it is often the case that h is not directly available, either because its functional form is unknown or because evaluation is computationally prohibitive. To remedy this difficulty, Robbins and Monro [88] proposed the class of algorithms known as *stochastic approximations* (SA). In their simplest form, such algorithms are unconstrained (i.e., $\Theta = \mathbb{R}^p$) and produce a sequence of iterates $\{\theta_n,\ n = 0, 1, \ldots\}$ through the recursion

$$\theta_0 \in \mathbb{R}^p, \qquad \theta_{n+1} = \theta_n + \gamma_{n+1} Y_{n+1}, \qquad n = 0, 1, \ldots \tag{1.2}$$

for some $\mathbb{R}^p$–valued "driving" process $\{Y_n,\ n = 0, 1, \ldots\}$, and some sequence of step-sizes $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$ which satisfy standard conditions, say $\gamma_n \downarrow 0$ and $\sum_{n=0}^{\infty} \gamma_{n+1} = \infty$.

It is customary to view Y_{n+1} as an *approximation* to $h(\theta_n)$. In their original paper, Robbins and Monro generated random variables (rvs) $\{Y_{n+1},\ n = 0, 1, \ldots\}$ according to

$$\mathbf{P}[Y_{n+1} \in B | Y_0, \ldots, Y_n] = \mu_{\theta_n}(B) \qquad n = 0, 1, \ldots \tag{1.3}$$

for some family of probability measures $\{\mu_\theta,\ \theta \in \mathbb{R}^p\}$ on $\mathbb{R}^p$ with the property that

$$h(\theta) = \int_{\mathbb{R}^p} y \mu_\theta(dy), \quad \theta \in \Theta.$$

The key issue in the study of algorithms such as (1.2) (and variations thereof) is the *convergence* of the iterate sequence $\{\theta_n,\ n = 0, 1, \ldots\}$ to the desired value(s) $\theta^\star$.

1.1 Extensions to the Robbins-Monro Algorithm

Over the years, increasingly more complex applications have lead to the use of *projected* versions of the stochastic approximation scheme (1.2) which take the form

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \Pi_\Theta \{\theta_n + \gamma_{n+1} Y_{n+1}\} \qquad n = 0, 1, \ldots \tag{1.4}$$

with Π_Θ denoting the nearest-point projection on Θ. It also became necessary to consider versions of (1.4) which are driven by processes $\{Y_n,\ n = 0, 1, \ldots\}$ with a more general statistical structure than (1.3). For instance, several authors [62, 70, 77, 97] have considered both (1.2) and (1.4) when

$$Y_{n+1} = H(\theta_n, X_{n+1}) \qquad n = 0, 1, \ldots \tag{1.5}$$

for some Borel mapping $H : \Theta \times \mathsf{X} \to \mathbb{R}^p$ and *state process* $\{X_n,\ n = 0, 1, \ldots\}$ evolving on $\mathsf{X} \subset \mathbb{R}^s$ which is Markov in the sense that

$$\mathbf{P}[X_{n+1} \in B | \theta_i, X_i, i = 0, 1, \ldots, n] = \int_B P_{\theta_n}(X_n, dy) \qquad n = 0, 1, \ldots$$

for some family of transition kernels $\{P_\theta(x, dy),\ \theta \in \Theta,\ x \in \mathsf{X}\}$ on X. Here, the state process sequence $\{X_{n+1}, n = 0, 1, 2, \ldots\}$ can either model noise in the estimation of certain steady state values or represent randomness in some underlying system being observed in real time or simulated on a computer. In any event, this more general algorithm is also proposed to find the zero $\theta^\star$ of an unknown function h given by the expectation of this driving/observation function H with respect to the P_θ-invariant steady state distribution π_θ:

$$h(\theta) = \int_{\mathsf{X}} H(\theta, x) \pi_\theta(dx). \tag{1.6}$$

In the literature, this function h is sometimes referred to as the *regression function.*

1.2 Typical Applications

These extended SA procedures have found applications in a wide variety of fields where it is desired to tune or optimize certain continuous-valued parameters of stochastic systems or find roots of an unknown steady state mean function $h : \Theta \to \mathbb{R}^p$. Due to the simplicity of the recursive step, SA's are typically implemented with very low overhead while the class of θ-dependent Markovian state processes is broad enough to include many stochastic systems which may have a parameter dependence and/or include noise and memory effects. We mention only a few of the many diverse applications of SA which have been proposed in the literature to date:

Network Management

- Online adjustment of protocol parameters via SA is carried out within a *performance management* tool for multiple access computer communication networks [67].

- Optimal source rates for Available Bit Rate traffic in Asynchronous Transfer Mode networks are found by optimizing the feedback control policy over the network via SA [7].

- Online minimization of call/connection setup time over a circuit-switched network is performed via SA-based load balancing in [99].

- A real-time distributed system modeling telephone switching systems is described in [60] where peripheral processors submit both high and low priority jobs to a single central processor under a distributed load control algorithm. Here, each control algorithm estimates the root of a unavailable nonlinear function via SA while taking observations from the central processor's load.

- In a cellular wireless network, SA estimates blocking probabilities for an efficient paging strategy [87].

Control/Optimization of Queueing Systems

- The performance of the GI/G/1 queue is optimized using an IPA gradient estimate where the parameter θ affects the service time distribution [20].

- Multiple queues compete for a single server in [72] where an SA is used to drive the long run average cost to a given value.

- A simple open loop low-overhead Call Admission Control (CAC) scheme is described in [66] which delays the customer's admission if the time since previous admission is less than a parameter. This parameter is recursively updated via an online SA-based gradient descent algorithm.

- A scheduling algorithm is load-balanced adaptively by finding a root to a nonlinear steady state equation via SA [9].

Communications

- Several examples of equalization in digital communications are considered in [5, 6, 76, 31].

- A digital phase-locked loop based on SA is proposed for carrier synchronization in burst-mode communication architectures [51].

- In ALOHA networks, a distributed SA-based algorithm computes the retransmission probabilities for each channel [52].

Neural Networks

- SA provides online estimation of neural network weights for controlling a nonlinear stochastic system [100, 101].
- An alternative SA algorithm for adjusting the neural networks is described in [96] which incorporates some SA averaging methods.

Manufacturing Systems and Inventory Control

- An optimal stationary inventory control policy is approximated by a linear switching curve and located via SA in [3].
- Performance of a multi-product multi-machine manufacturing system is optimized in [26] using a Perturbation Analysis gradient estimate where performance is measured by the cumulative system time.
- A partially observed binary replacement problem is formulated as an adaptive Markov decision control problem, and an SA is used to estimate unknown parameters needed for the long run average optimal control policy [1].

1.3 Research Objectives

Our main objective in this dissertation is to study the *convergence* of the iterates of the SA algorithm for applications with Markov chain noise. We seek an operational framework of conditions which are easily *verifiable* in terms of the available model data. Our focus tends toward projected algorithms because, as we will soon see, these algorithms are more likely in practice to yield a convergence result which is not conditional on any unverifiable events. Also, we wish to emphasize we are not necessarily trying to find the weakest possible conditions implying convergence; instead, we seek an operational convergence theory in terms of explicit conditions on the model data which cover most of the Markov chains typically encountered in applications. This does include applications where the state space is either very large, countably infinite, or even general. We also consider applications where the driving function H may possess a functional dependence on the parameter θ as well as applications where H may be unbounded in the state variable, such as queue occupancy based estimates.

Our approach to showing convergence relies on the ODE method [61] which, in most of its forms, proceeds in two separate steps. The first step relies on the Kushner–Clark Lemma [61] to identify a deterministic ODE given by

$$\dot{\theta}(t) = h(\theta(t)), \qquad t \geq 0,$$

the stability properties of which determine the limit points of the iterate sequence $\{\theta_n,\ n = 0, 1, \ldots\}$. The second step, which is probabilistic in nature and depends on the algorithm, involves showing that asymptotically the output sequence of the original SA behaves like the solution to the ODE. Although general conditions are given in [61] for successfully completing this second step, these conditions are often not immediately checkable in terms of the model data. Therefore, it is precisely our main goal to develop specific tools or indirect methods to facilitate verification of this second step.

A substantial body of research currently exists on adaptive algorithms and SA's, see for example [6, 17, 18, 28, 61, 64, 70, 72], as well as the many references cited within these. We wish to point out one effort [70] where the authors, by focusing on the class of *finite state* Markov chain state processes, show convergence criteria for the ODE method which is indeed verifiable in terms of the state process model data. Here, we take a similar approach to accomplish our main goal, and seek to extend the criteria in [70] to a broader class of state processes which may have a countable or a non-finite state space while also permitting an unbounded driving function H.

Benveniste, Métivier and Priouret (BMP) [6] have presented a *general* framework for Markov chain state processes dependent on a parameter θ which imply almost sure convergence of the SA's iterates. Unfortunately, these general conditions are also difficult to verify directly since they are based on smoothness properties of the Poisson equation solution. BMP then propose *specialized conditions* for certain geometrically ergodic Markov chains which imply their more general conditions. There still remain many problems where it is difficult to verify BMP's specialized conditions, and we seek to either extend BMP's framework or find new tools which may be applied to these problems.

We also note that the finite state space results of [70] have been extended to a particular countable state space Markov chain in [72] where the key convergence condition to be verified is Lipschitz continuity of the Poisson equation solution. While these results generalize to other similar problems, we seek extensions to this framework by working with a weaker form of Hölder continuity on the Poisson equation solution while considering a more general class of applications similar to those considered by BMP in [6].

We also have several secondary goals. For one, we pay particular attention to Markov chains which possess certain Foster-Lyapunov stability or drift properties, as studied by Meyn and Tweedie [79] and others [49], implying a geometric ergodicity. This complements the specialized results of BMP mentioned above. Second, since steady state optimization dominates the applications for SA, we explore convergence verification and particular problems encountered when the SA algorithm is driven by an estimate of the performance gradient. At present, there exists a substantial body of literature on various forms of gradient estimation [36, 41, 44, 46, 48, 55, 63, 85] as well as treatments of stochastic optimization [19, 20, 43, 47, 75, 83, 84, 106]. We focus entirely on one particular class of gradient estimation algorithms recently proposed by Cao and Wan [14] which we find well-suited for SA.

In addition, we also seek to highlight certain operational difficulties which may be encountered when attempting to verify convergence properties of SA's applied to even simple Markov chain problems. In particular, we find that the *selection* of an "appropriate" projection set Θ can be difficult for many problems such as those applications with some degree of uncertainty in the model data. We explore issues related to the selection of an "appropriate" set Θ to use in the projection operator and we consider convergence when a less than ideal projection set is selected.

1.4 The ODE Method

As we stated above, our approach is based on the Kushner-Clark ODE method [61] so let us briefly review this method. There are two versions of this result; one for *constrained* algorithms which use a projection operator and one for *unconstrained* algorithms which we now summarize.

The algorithm may be written in the following form which includes additional noise random variables $\{\beta_{n+1},\ n = 0, 1, \ldots\}$,

$$\begin{aligned} \theta_{n+1} &= \theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}\beta_{n+1}, \qquad n = 0, 1, \ldots. \\ &= \theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\varepsilon_{n+1} + \gamma_{n+1}\beta_{n+1}, \end{aligned} \tag{1.7}$$

where following 1.6 the *main noise* sequence is defined by

$$\varepsilon_{n+1} \doteq H(\theta_n, X_{n+1}) - h(\theta_n).$$

We also need to define the times

$$t_0 = 0 \qquad t_n = \sum_{k=0}^{n-1} \gamma_{k+1}, \qquad n = 1, 2, \ldots,$$

and the function

$$m(t) \doteq \begin{cases} \max\{n : t_n \leq t\}, & t \geq 0 \\ 0, & t < 0 \end{cases}. \tag{1.8}$$

Then, the *piecewise linear interpolated* function $\theta^0(t)$ is defined by

$$\begin{aligned} \theta^0(t_n) &\doteq \theta_n, \\ \theta^0(t) &\doteq \frac{(t_{n+1} - t)}{\gamma_{n+1}}\theta_n + \frac{(t - t_n)}{\gamma_{n+1}}\theta_{n+1}, \qquad t \in (t_n, t_{n+1}), \end{aligned}$$

for each $n = 0, 1, \ldots$; as well as the time shifts

$$\theta^n(t) = \begin{cases} \theta^0(t + t_n), & t \geq -t_n \\ \theta_0, & t \leq -t_n, \end{cases} \qquad n = 1, 2, \ldots.$$

Thus, we have a sequence of functions $\{\theta^n(\cdot) : n = 0, 1, \ldots\}$.

Lemma 1.1 (Kushner-Clark Lemma, unconstrained case) *Suppose:*

(KC1) $h(\cdot) : \mathbb{R}^p \to \mathbb{R}^p$ *is a continuous function,*

(KC2) $\{\gamma_n,\ n = 1, 2, \ldots\}$ *is a sequence of positive real numbers such that* $\gamma_{n+1} \to 0$, *and* $\sum_{n=1} \gamma_{n+1} = \infty$,

(KC3) *$\{\beta_n,\ n=1,2,\ldots\}$ is a bounded (w.p.1) sequence of $\mathbb{R}^p$-valued rvs such that $\beta_n \to 0$, w.p.1,*

(KC4) *$\{\varepsilon_n,\ n=1,2,\ldots\}$ is a sequence of $\mathbb{R}^p$ valued rvs such that for some $T>0$ and all $\epsilon > 0$*

$$\lim_{n\to\infty} \mathbf{P}\left\{ \sup_{j\geq n} \max_{t\leq T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} \gamma_{i+1}\varepsilon_{i+1} \right| \geq \epsilon \right\} = 0, \tag{1.9}$$

(B1) *the iterates $\{\theta_n,\ n=0,1,\ldots\}$ are bounded w.p. 1.*

Then, there is a null set Ω_0 such that $\omega \notin \Omega_0$ implies

1. *$\{\theta^n(\cdot,\omega),\ n=0,1,\ldots\}$ is equicontinuous, and also, the limit $\theta(\cdot)$ of any convergent subsequence of $\{\theta^n(\cdot,\omega),\ n=0,1,\ldots\}$ is bounded and satisfies the ODE*

$$\dot{\theta} = h(\theta), \qquad t \in (-\infty,\infty). \tag{1.10}$$

2. *Let $\theta^\star$ be a locally asymptotically stable in the sense of Lyapunov solution to (1.10), with domain of attraction $DA(\theta^\star)$. There is a compact set $Q \subset DA(\theta^\star)$ such that if $\theta_n(\omega) \in Q$ infinitely often, we have*

$$\theta_n(\omega) \to \theta^\star \qquad \text{as } n \to \infty.$$

Proof: See [61, Theorem 2.3.1]. ■

1.5 Projected Algorithms

The unconstrained Kushner-Clark Lemma may be considered a *conditional* convergence result since in applications verifying the *boundedness condition* (B1) and the *stability-recurrence condition*, $\mathbf{P}\{\theta_n \in Q \ i.o.\} = 1$ for some $Q \subset DA(\theta^\star)$ tend to be difficult. There are no generally applicable results which give verifiable conditions which readily imply (B1) and the stability recurrence, so one is typically faced with the often difficult task of verifying these conditions for each particular application (see, for example [6] [61, Section 4.7], and [31, 32]).

To remedy this situation, algorithms with a *projection* on a compact set Θ have been proposed [61, 68] to ensure boundedness as well as assist in showing stability recurrence, i.e,

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\varepsilon_{n+1} + \gamma_{n+1}\beta_{n+1} \right\} \qquad n=0,1,\ldots \tag{1.11}$$

Also, the projection operator can often be helpful in verifying (KC1), (KC3), and (KC4) since the parameter iterate is then known to fall within some compact set. For many applications, it is typically not a problem to apply an "appropriately selected" projection operator.

Of course, the *ideal case* is to choose the compact projection set $\Theta \subset \Theta$ so that $\theta^\star \in \Theta$, $\Theta \subset DA(\theta^\star)$, and (KC1)-(KC4) hold when the parameter is constrained to Θ. For many applications, if conditions (KC1)-(KC4) can be verified and assuming some very limited knowledge of the state process is available, it may in fact be possible to identify an ideal projection set. In this case, the projected version of the Kushner-Clark Lemma [61, p. 191] provides identical conditions for a.s. convergence for algorithm (1.11) without (B1). Also, the stability-recurrence condition may be satisfied if we have some knowledge of the stability regions of the ODE, as is often the case. Thus, we can choose $Q = \Theta \subset DA(\theta^\star)$, and in this setting, the constrained Kushner-Clark Lemma tells us that unconditionally the iterate will converge to θ^* with probability one. This *ideal case* is essentially the topic of the next two chapters as we mainly explore methods to verify (KC4) for a given projection set Θ.

It is not always possible to identify an ideal projection set. An interesting situation arises in the case that any of the conditions (KC1), (KC3), (KC4) only hold if the iterates are constrained to some well-behaved domain $D_s \subset \mathbb{R}^p$ which is unfortunately not known when implementing the algorithm. Due to this uncertainty, there may be no way to determine an ideal compact projection set satisfying both $\Theta \subset D_s$ and $\theta^\star \in \Theta$. If we take Θ too large, we are likely to cause $\Theta \not\subset D_s$ while if we take Θ too small, it is possible to have $\theta^\star \notin \Theta$. Since it is clear that such a "too small" projection will prevent the desired convergence, we want to study the convergence when the projection set is chosen "too large". In general this may be very difficult; but, we find that for a particular problem, by exploiting certain structural properties of the state process together with the dynamics of the SA algorithm, we may in fact be able to show convergence.

This last issue was motivated by a simple SA example which attempts to regulate the steady state mean queue size of a fixed arrival rate M/M/1 queue via recursive SA updating of the service rate parameter. If the arrival rate is unknown, an ideal projection set which constrains the SA's iterates to a positive recurrent region is not available *a priori*. We then study the convergence of an algorithm which uses a "comfortably large" but compact projection set which includes part of the transient region. The queue may at times operate in the transient region and the analysis becomes a bit more difficult for several reasons, not the least of which is the function $h(\theta)$ is not even defined in this transient region. Nevertheless, this example possesses a key property; if at any time the service rate is set to a parameter which causes the M/M/1 to be transient, then the queue size tends to increase toward infinity and the dynamics of the SA will tend to return the queue to the possibly recurrent region. While convergence in this setting seems intuitively reasonable given these structural properties, we establish an approach to rigorously proving convergence. This approach should generalize to other problems with a similar structure.

1.6 Summary of Results

The previous sections introduced the Kushner-Clark Lemma and some technical difficulties encountered when applying SA to actual problems. For the majority of this dissertation, we focus on algorithms projected on compact $\Theta \subset \mathbb{R}^p$, thus eliminating the difficulty related to the boundedness condition. In Chapter 2, we temporarily put aside our concerns regarding selection of an *ideal* projection set Θ and simply assume one is available. We then study the noise process $\{\varepsilon_{n+1}, n = 0, 1, \ldots\}$ and develop a general set of conditions which imply (KC4) holds for a broad class of Markov chain state processes dependent on the parameter $\theta \in \Theta$. Our analysis is adapted from the framework developed in BMP [6] and we offer several extensions and variations of their results which lead to a.s. convergence, specifically for the case of projected SA's.

The extensions come about for two reasons. First, we discover that one of BMP's general conditions, local Lipschitz continuity on the regression function $h(\cdot)$, is unnecessarily strong since only continuity is required in the ODE method. Second, and perhaps more troubling in applications, we find that one of BMP's specialized conditions for geometrically ergodic Markov chains can be difficult to verify for many problems. We are able to sufficiently weaken both of these conditions by slightly modifying their framework so that our new conditions are straightforward to verify for applications of interest to us.

The new *specialized* conditions are developed in Chapter 3 and assume a uniform Foster-Lyapunov drift inequality on the family of one-step Markov transition probabilities which are derived from drift equalities recently studied by Meyn and Tweedie in [79]. These results form a new framework of verifiable conditions given in terms of the Markov transition probabilities and we show they ultimately imply condition (KC4) of the Kushner-Clark Lemma holds. To demonstrate the application of these specialized conditions to a countable state Markov chain, we carry out the verification steps for an SA algorithm applied to a simple parameterized random walk with a single reflection at the origin.

We also consider *stochastic optimization* applications where performance is measured by a steady state mean or long run average

$$J(\theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\theta, X_n),$$

for some given performance function $f(\cdot,\cdot) : \Theta \times \mathsf{X} \to \mathrm{I\!R}$. The objective here is to find a point $\theta^\star$ such that the performance gradient $\nabla J(\theta)$ is equal to zero through an iterative procedure based on SA coupled with a gradient estimate $\widehat{G}$ of $\nabla J(\theta)$, i.e.

$$\theta_{n+1} = \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} \widehat{G}(\theta_n, \bar{X}_{n+1}) \right\}, \qquad n = 0, 1, 2, \ldots \tag{1.12}$$

In this case, the process $\{\bar{X}_n, \ n = 1, 2, \ldots\}$ is an augmented Markov chain related to $\{X_n, \ n = 1, 2, \ldots\}$. First in Chapter 4, we propose a particular gradient estimation algorithm which is an adaptation of an estimate studied in [14]. Then in Chapter 5, we present a series of results which show for SA algorithms using this gradient estimate driven by $\{\bar{X}_n, \ n = 1, 2, \ldots\}$, the convergence can be checked, for the most part, by verifying the conditions on $\{X_n, \ n = 1, 2, \ldots\}$ as in the root-finding problem. Thus, we have a checkable verification procedure using the specialized conditions proposed in our convergence framework of Chapters 2-3 which, if met, implies almost sure convergence of θ_n to the optimizer $\theta^\star$.

In Chapter 6, we return to the issue set aside earlier; namely, the difficulty in choosing an *ideal* projection set Θ. For a simple random walk model, we demonstrate a sample path technique which allows one to choose the projection set to be "comfortably large" in the sense that the conditions developed in Chapter 2 need only hold for iterates which lie in some well-behaved subset Q of the non-ideal projection set Θ. This leaves a complementary region where some of the conditions may not be satisfied. We find it is still possible to prove almost sure convergence, even though at times some sample paths cause the the state process to be unstable or transient. Here, the analysis is problem dependent although we develop several results which should hold for other problems with a similar structure.

Finally in Chapter 7, we consider a substantial change in direction, partly inspired by an approach in [29] dubbed "sampling controlled SA," where we study an alternative SA algorithm *driven by sample averages*. This algorithm differs from the traditional algorithm which simply observes a single sample and immediately makes a parameter update by instead waiting to collect several observations before computing a parameter update based on an average of these observations. For the traditional SA algorithms studied in Chapter 2, we use a martingale approach along with conditions on the solution to the Poisson equation to show convergence. For this algorithm driven by sample averages, we instead use a large deviations approach with an increasing observation window. We show convergence of this algorithm follows readily if a certain uniform large deviations upper bound holds for the state process and the observation window is lengthened towards infinity as the recursion advances. Although there can be benefits to avoiding the martingale approach for some problems with nonlinearities, this large deviations upper bound can be difficult to show for several of the countable state space problems of interest. As a result, we now feel the traditional fixed-step SA algorithm studied in the previous chapters usually is preferable. Nevertheless, we do show this uniform large deviations bound is in fact satisfied for both i.i.d. state processes as well as finite state Markov chains which are dependent on a parameter θ if only mild regularity conditions are in place. Furthermore, if the sample average is passed through a nonlinear function g driving the recursion, then this large deviations approach may offer advantages since the martingale methods can break down.

As a final comment, the various SA's studied here may be classified as to how the time scale of the parameter updates relate to the time scale of the state process evolution. For the traditional algorithm studied in Chapters 2, 3, and 6 a single parameter update occurs for each transition of the state process. For the stochastic optimization algorithm of Chapter 5 and the sample average algorithm of Chapter 7, the SA is driven by observations taken over a window on the state process of length ℓ_n which steadily increases towards infinity as the SA's iterates are updated, i.e. $\ell_n \to \infty$ as $n \to \infty$.

1.7 Some Definitions and Notation

1. The set of all real numbers is denoted by $\mathbb{R}$ and the set of all integers is denoted $\mathbb{Z}$.

2. For any set X endowed with a topology, measurability is always taken to mean Borel measurability and the corresponding Borel σ–field, i.e., the smallest σ–field on X generated by the open sets of the topology, is denoted by $\mathcal{B}(\mathsf{X})$.

3. We denote the n-step probability transition function $P^n(x, A)$ for $A \in \mathcal{B}(\mathsf{X})$ and $x \in \mathsf{X}$. For a probability measure π and a Borel function $f : \mathsf{X} \to \mathbb{R}$ we define

$$\begin{aligned} P^n(x, f) &\doteq \int_{\mathsf{X}} P^n(x, dy) f(y), \qquad n = 1, 2, \ldots, \text{ and} \\ \pi(f) &\doteq \int_{\mathsf{X}} \pi(dx) f(x) \end{aligned}$$

Also, we may find it convenient to occasionally use the slightly more compact operator notation

$$\begin{aligned} P^n f(x) &\doteq \int_{\mathsf{X}} P^n(x, dy) f(y), \qquad n = 1, 2, \ldots, \text{ and} \\ \pi f &\doteq \int_{\mathsf{X}} \pi(dx) f(x) \end{aligned}$$

4. An element v of $\mathbb{R}^p$ is denoted by its column vector and its transpose is denoted by v'. For elements v and w of some $\mathbb{R}^p$, we write $\langle v, w \rangle$ for their usual scalar product, so that $\|v\| \equiv \sqrt{\langle v, v \rangle} = \left(\sum_{i=1}^p v_i^2\right)^{1/2}$ denotes the Euclidean norm of v. Also, we will regularly use the Schwarz Inequality [50, p. 2]

$$|v \cdot w| \doteq |\langle v, w \rangle| \leq \|v\| \, \|w\|, \qquad v, w \text{ in } \mathbb{R}^p.$$

In a squared form, this inequality reads

$$\left(\sum_{i=1}^p v_i w_i\right)^2 \leq \left(\sum_{i=1}^p v_i^2\right)\left(\sum_{i=1}^p w_i^2\right), \qquad v, w \text{ in } \mathbb{R}^p.$$

5. For a Borel function $f : \mathsf{X} \to [1, \infty)$ we define, as in [79], the f-norm of two probability transition functions $P_1(x, A)$ and $P_2(x, A)$ as

$$\|P_1(x, \cdot) - P_2(x, \cdot)\|_f \doteq \sup_{g:|g| \leq f} |P_1(x, g) - P_2(x, g)| \tag{1.13}$$

Similarly, we will often apply the same f-norm to probability measures in place of kernels by simply defining a kernel from the probability measure, i.e. $\mu(x, A) = \mu(A)$ for all $x \in \mathsf{X}$.

6. Following the (perhaps non-standard) practice from [6], we will often write the function $H_\theta(x)$ as an equivalent expression for the function $H(x, \theta)$.

7. The infimum over an empty set is taken to be ∞ by convention.

8. We take e to be a column vector $e = (1, 1, \ldots, 1)'$ or $e = (1, 1, 1, \ldots)'$ as is appropriate.

Chapter 2

Convergence of Projected Stochastic Approximations

We establish a basic framework to study the convergence properties of the projected Stochastic Approximation (SA) algorithm via the Kushner-Clark Lemma. The first part of this chapter is a variation and extension of a series of results in Benveniste, Métivier, and Priouret (BMP) [6] which bound the noise terms for this algorithm. We make several alterations to BMP's framework and weaken their Lipschitz condition on the regression function yet we are still able to show almost sure convergence. The general conditions of this chapter, given in terms of the Poisson equation solution, are not necessarily easy to check but they serve as a foundation for the next chapter's results where specialized conditions implying these general conditions are developed.

2.1 The Algorithm

Consider the *projected* stochastic approximation algorithm defined by the recursion:

$$\begin{aligned} \theta_{n+1} &= \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}^2 \rho_{n+1}(\theta_n, X_{n+1}) \right\}, \qquad n = 1, 2, \ldots \\ \theta_0 &= \theta \end{aligned} \tag{2.1}$$

The iterates $\{\theta_n, n = 0, 1, \ldots\}$ evolve in some closed projection set $\Theta \subset \mathbb{R}^p$ and the Markovian state process samples $\{X_n, n = 0, 1, \ldots\}$ lie in X, some general state space. The algorithm is driven by the functions $H : \Theta \times \mathsf{X} \to \mathbb{R}^p$ and $\rho_{n+1} : \Theta \times \mathsf{X} \to \mathbb{R}^p$ for $n = 0, 1, \ldots$. The ρ_{n+1} terms essentially play the role of the β_{n+1} terms defined in the Kushner-Clark Lemma. We note that this is the form of the recursion appearing in [6] which differs slightly from (1.7) studied in [61].

The initial values of the algorithm are arbitrary, i.e. $\theta_0 = \theta$ in Θ and $X_0 = x$ in X. The deterministic step-size sequence $\{\gamma_n, n = 1, 2, \ldots\}$ is chosen to satisfy the following condition:

(S) $\gamma_n \downarrow 0, \qquad \sum_{n=0}^{\infty} \gamma_{n+1} = \infty, \qquad \sum_{n=1}^{\infty} \gamma_n^{1+\widehat{\ell}_1} < \infty, \qquad$ for some $0 < \widehat{\ell}_1 < 1$.

The *state process* $\mathcal{X} = \{X_n, n = 1, 2, \ldots\}$ is formally defined in the next section but is simply a θ-parameterized discrete time Markov chain. The one-step transition kernel $P_\theta(x, \cdot)$ may depend on the continuous variable θ and the probability distribution of the next state X_{n+1}

depends on both the current state X_n and current iterate θ_n. We also assume that a generic homogeneous Markov chain governed by the same one-step transition kernel $P_\theta(x,\cdot)$ where the parameter θ is held **fixed** at any $\theta \in \Theta$ is ergodic in the sense that there exists a P_θ-invariant measure π_θ and

$$\lim_{n\to\infty} \mathbf{E}_{\theta,x}\left[H(\theta, X_n)\right] = \pi_\theta(H_\theta) \doteq h(\theta), \qquad \theta \in \Theta,\ x \in \mathsf{X}. \tag{2.2}$$

This algorithm (2.1) attempts to find points $\theta^\star$ such that $h(\theta^\star) = 0$, and assuming certain conditions are met which we will soon introduce, the algorithm is able to find these points $\theta^\star$ despite the fact that h is unknown, the current estimate θ_n is regularly updating to a new value, and the Markov chain state process is not necessarily observed at any given parameter θ held fixed for a "long time".

The most common form of SA's take $\rho_n = 0$ for all $n = 0, 1, \ldots$, although there are several possible uses for these terms being nonzero. The process $\{\rho_n, n = 1, 2, \ldots\}$ can be used to model either additional noise, a time dependent perturbation, or even an auxiliary control input to the algorithm. For example, perturbations [6] to an SA algorithm with vector valued iterates may be introduced if the individual vector components are updated successively rather than simultaneously. Alternatively, if it is known that some initial (asymptotically decaying) bias exists in an estimate $H(\theta_n, X_n)$ of $h(\theta_n)$, then it may be desirable to supply an opposing bias through the ρ_n term. As such, the ρ_n term can be used to give the engineer designing an SA procedure some limited means of control over the transient phase of the SA run. This is possible because under the conditions we will soon propose, the ρ_n term has no effect on the *asymptotic* convergence of the iterates θ_n to $\theta^\star$.

2.2 The Basic Ingredients

Throughout the discussion, p and s are fixed positive integers denoting the dimensions of the parameter and state vector spaces, respectively. We assume given a closed subset Θ of $\mathbb{R}^p$, and a Borel subset X of $\mathbb{R}^s$. Let X^∞ be the infinite Cartesian product of X with itself, and denote by $\mathcal{B}(\mathsf{X}^\infty)$ the standard σ–field on X^∞. We write a generic element ξ of X^∞ as $\xi = (x, x_1, \ldots)$ where $x, x_1, \ldots$ are all elements of X. The coordinate process $\{\xi_n,\ n = 0, 1, \ldots\}$ is then simply defined by

$$\xi_0(\xi) \equiv x, \qquad \xi_n(\xi) \equiv x_n, \qquad \xi \in \mathsf{X}^\infty,\ n = 1, \ldots \tag{2.3}$$

We postulate the existence of a family $\{\mathbf{P}_{\theta,x},\ \theta \in \Theta,\ x \in \mathsf{X}\}$ of probability measures on $\mathcal{B}(\mathsf{X}^\infty)$ such that

$$\mathbf{P}_{\theta,x}[\xi_0 = x] = 1, \qquad \theta \in \Theta,\ x \in \mathsf{X}. \tag{2.4}$$

For technical reasons, we need to assume a measurable functional dependence in θ and x:

(P0) For every $L = 1, 2, \ldots$, the mapping $\Theta \times \mathsf{X} \to \mathbb{R} : (\theta, x) \to \mathbf{P}_{\theta,x}[\xi_n \in B_n,\ n = 1, \ldots, L]$ is Borel measurable for all possible choices of Borel subsets $B_1, \ldots, B_L$ in $\mathcal{B}(\mathsf{X})$.

In order to define the stochastic approximation procedures, we start with a sample space Ω equipped with a σ–field of events $\mathcal{F}$. The measurable space $(\Omega, \mathcal{F})$ is assumed large enough to carry a sequence of X–valued rvs $\{X_n,\ n = 0, 1, \ldots\}$. We define the Θ–valued rvs $\{\theta_n,\ n = 0, 1, \ldots\}$ through the deterministic recursion:

$$\begin{aligned} \theta_{n+1} &= \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}^2 \rho_{n+1}(\theta_n, X_{n+1}) \right\} \qquad n = 0, 1, \ldots \\ \theta_0 &= \theta \qquad \in \Theta. \end{aligned} \tag{2.5}$$

Here, $H : \Theta \times \mathsf{X} \to \mathbb{R}^p$ and $\rho_{n+1} : \Theta \times \mathsf{X} \to \mathbb{R}^p$ for $n = 0, 1, \ldots$ are Borel mappings for all θ in Θ and x in X. In (2.5), Π_Θ denotes the nearest-point projection operator on the set Θ. We assume the operator Π_Θ is well defined and if the nearest point is not unique, then some mechanism is in place to ensure that Π_θ is well defined. If the closed set Θ is convex, the Π_Θ is well defined without any such mechanism.

Next, we introduce the filtration $\{\mathcal{F}_n,\ n = 0, 1, \ldots\}$ on $(\Omega, \mathcal{F})$ by setting

$$\begin{aligned} \mathcal{F}_n &\equiv \sigma\{\theta_m, X_m,\ m = 0, 1, \ldots, n\} \\ &= \sigma\{\theta_0; X_m,\ m = 0, 1, \ldots, n\} \qquad n = 0, 1, \ldots \end{aligned}$$

since the rvs θ_m, $m = 1, 2, \ldots, n$, are fully determined by the rvs θ_0, X_0, and X_m, $m = 1, \ldots, n$.

Given a probability measure μ on $\mathcal{B}(\Theta \times \mathsf{X})$, we postulate the existence of a probability measure $\mathbf{P}$ on $(\Omega, \mathcal{F})$ such that

$$\mathbf{P}[\theta_0 \in A, X_0 \in B] = \mu(A \times B) \tag{2.6}$$

for all Borel subsets A and B of Θ and X, respectively, and satisfying

$$\mathbf{P}[X_{n+1} \in B | \mathcal{F}_n] = \mathbf{P}_{\theta_n, X_n}[\xi_1 \in B] \qquad n = 1, \ldots \tag{2.7}$$

for all Borel subsets B in $\mathcal{B}(\mathsf{X})$. The existence of such a set-up is readily justified by the Daniell–Kolmogorov consistency theorem [69, p. 94] on $\Theta \times \mathsf{X} \times \mathsf{X}^\infty$ in the usual manner.

Finally for each $\theta \in \Theta$, we also define the one-step transition function (kernel)

$$P_\theta(x, B) \doteq \mathbf{P}_{\theta, x}[\xi_1 = B], \qquad x \in \mathsf{X}, \theta \in \Theta, B \in \mathcal{B}(\mathsf{X}).$$

We make the following additional assumptions:

(π) For each $\theta \in \Theta$, there exists a unique P_θ-invariant probability measure π_θ on $(\Omega, \mathcal{F})$.

(H1) For all $\theta \in \Theta$, $H(\theta, \cdot) \doteq H_\theta(\cdot)$ is *integrable* under π_θ. Let us denote

$$h(\theta) \doteq \pi_\theta(H_\theta) \doteq \int_{\mathsf{X}} H_\theta(x) \pi_\theta(dx), \qquad \theta \in \Theta. \tag{2.8}$$

2.2.1 Fixed-θ Algorithm and Generic Markov Chains

At times in our analysis, we will make comparisons of the state process defined above in (2.5), (2.6) and (2.7) to a *fixed-θ* state process which is the result of a new *fixed-θ* algorithm whereby θ_n is not computed as in (2.5) but is instead held fixed so that

$$\theta_n = \theta \in \Theta \qquad \text{for all } n = 1, 2, \ldots$$

For this *fixed-θ* algorithm, the state process $\{X_n, n = 0, 1, \ldots\}$ is simply a *generic* homogeneous Markov chain governed by the one-step transition kernel $P_\theta(\cdot, \cdot)$ for the fixed $\theta \in \Theta$.

Additionally, in the same manner as we defined $\mathbf{P}$ in (2.6) and (2.7), the existence of some other probability $\widetilde{\mathbf{P}}$ in place of $\mathbf{P}$ on a common measurable space also follows if we replace the deterministic SA algorithm (2.5) with any other $\mathcal{F}_n$-measurable algorithm such that $\theta_n \in \Theta$ for all $n = 0, 1, \ldots$, such as this fixed-θ algorithm.

2.3 General Convergence Criteria

Here we list several *general conditions* directed at the Markovian state setup just described. Recall, Θ is a given closed subset of $\mathbb{R}^p$. In the next chapters as we look at specific applications, we may also require that Θ be compact although we do not assume this in general for this chapter.

2.3.1 Uniform Drift Conditions

This first condition (D0) will be a primary condition assumed for most results to follow:

(D0) There exists a function $V : \mathsf{X} \to [1, \infty)$ and a constant $1 \leq C_D < \infty$ such that

$$\mathbf{E}_{\theta,x}\left[V(X_n)\right] \leq C_D V(x), \qquad \text{for all } \theta \text{ in } \Theta,\ n = 0, 1, 2, \ldots, \text{ and } x \text{ in } \mathsf{X}.$$

These next conditions, assumed only for certain applications, are related to the stability of the Markov chain and we may refer to them as *uniform drift conditions* since they are *uniform* versions of a stability conditions studied extensively by Meyn and Tweedie [79]. Both (D1) and (D2) are closely related.

(D1) There exists a function $V : \mathsf{X} \to [1, \infty)$ and two constants $0 < \lambda < 1$ and $L < \infty$ such that

$$P_\theta V(x) \leq \lambda V(x) + L \qquad \text{for all } \theta \text{ in } \Theta \text{ and } x \text{ in } \mathsf{X}. \tag{2.9}$$

(D2) There exists an extended real valued function $V : \mathsf{X} \to [1, \infty]$, a measurable set C, and constants $\beta > 0$, $b < \infty$,

$$\sup_{\theta \in \Theta} \Delta_\theta V(x) \leq -\beta V(x) + b 1_C(x), \qquad x \in \mathsf{X}.$$

where we define $\Delta_\theta V(x) \doteq P_\theta V(x) - V(x)$.

The condition (D0) for us is fundamental and we will either assume it holds or provide conditions which imply (D0), such as (D1) or (D2) with V finite. In any event, under (D0) we implicitly define the function $V : \mathsf{X} \to [1, \infty)$.

2.3.2 Conditions Related to the Algorithm

For a function $V : \mathsf{X} \to [1, \infty)$ defined above, we assume the remaining conditions all hold for some positive constant r such that $0 < r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ whereby the real constant $\widehat{\ell}_1$ lies in the interval $(0, 1)$ and satisfies (S).

(H2) There exists constants $C_H < \infty$ and $C_\rho < \infty$ such that for all $x \in \mathsf{X}$:

$$\begin{aligned} \sup_{\theta\in\Theta} \|H(\theta, x)\| &\leq C_H V^r(x), \\ \sup_{\theta\in\Theta} \|\rho_n(\theta, x)\| &\leq C_\rho V^r(x), \qquad\qquad n = 1, 2, \ldots \end{aligned}$$

(P1) For all $(\theta, x) \in \Theta \times \mathsf{X}$, the following series converges:

$$\nu_\theta(x) \doteq \sum_{n=0}^{\infty} \left(\int_{\mathsf{X}} P_\theta^n(x, dy) H_\theta(y) - h(\theta) \right) < \infty,$$

and we identify $\nu_\theta(x)$ as the solution to the Poisson equation associated with $H(\theta, \cdot) = H_\theta(\cdot)$:

$$H_\theta(x) - h(\theta) = \nu_\theta(x) - \int_{\mathsf{X}} P_\theta(x, dy)\nu_\theta(y), \qquad x \in \mathsf{X},\ \theta \in \Theta.$$

(P2) There exists a constant $C_\nu < \infty$ such that

$$\begin{aligned} \|\nu_\theta(x)\| &\leq C_\nu V^r(x), \qquad \text{for all } \theta \in \Theta, \quad x \in \mathsf{X} \\ \|P_\theta \nu_\theta(x)\| &\leq C_\nu V^r(x), \qquad \text{for all } \theta \in \Theta, \quad x \in \mathsf{X} \end{aligned}$$

(P3) There exists a constant $C_\delta < \infty$ and such that

$$\|P_\theta \nu_\theta(x) - P_{\theta'} \nu_{\theta'}(x)\| \leq C_\delta V^r(x) \left\|\theta - \theta'\right\|^{\widehat{\ell}_1}, \qquad \text{for all } \theta, \theta' \in \Theta, \quad x \in \mathsf{X}.$$

where $\widehat{\ell}_1 \in (0, 1)$ is determined by (S).

We will loosely refer to algorithm (2.1) and this collection of conditions as our *general* framework or *general* conditions.

2.3.3 Remarks on our Conditions

The SA algorithm studied in [6] does not use a projection operator and, as a result, BMP's results lead to an almost sure convergence which is *conditional* on $\{\theta_n, n = 1, 2, \ldots\}$ being almost surely bounded (as in the unconstrained Kushner-Clark Lemma). BMP then augment their convergence theorem with an additional set of Lyapunov type stability conditions (see [6, p. 239]) under which they show the iterates $\{\theta_n, n = 1, 2, \ldots\}$ are in fact almost surely bounded. Since we specifically focus on the constrained problem, such an augmentation is not necessary. We believe that for many applications, while perhaps not necessary for convergence, it is usually not a problem to use an "appropriately sized" compact projection and in so doing it clearly simplifies the convergence analysis. In this sense, our results are different in that they address a somewhat less difficult problem since the boundedness property does not need to be shown if a compact projection is used. On the other hand, for those problems which do not permit use of a compact projection operator, BMP's global convergence results [6, Thm. 17 p. 239] remain a viable approach.

Our main condition (D0) can be compared to BMP's assumption (A.5) [1] from [6]:

(A.5) For any compact subset Q of (their parameter space) D and any $q > 0$, there exists $\mu_q(Q) < \infty$ such that for all n, $x \in \mathbb{R}^k$, $\theta \in \mathbb{R}^d$

$$\mathbf{E}_{x,\theta}\left[1_{\{\theta_k \in Q, k \leq n\}}(1 + \|X_{n+1}\|^q)\right] \leq \mu_q(Q)(1 + \|x\|^q).$$

We note that (D0) can be weaker than (A.5) for two reasons. First, our (D0) does not require the existence of the bound for **all** compact sets Q in the parameter space D but only for the projection set Θ. Second, (D0) permits an *arbitrary* function $V : \mathsf{X} \to [1, \infty)$ in the inequality while BMP's condition uses a function which must take the form $const(1 + \|x\|^q)$ for at least sufficiently large $q > 0$. Additionally, this more general function V is compatible with many of the results on V-uniformly ergodic Markov chains [79] which employ Foster-Lyapunov type drift inequalities similar to conditions (D1) and (D2) involving V.

The other significant departure from BMP's conditions in this chapter is the dropping of BMP's local Lipschitz condition on the regression function $h : \Theta \to \mathbb{R}^p$.

The remaining assumptions in our framework are essentially similar to the basic framework developed in [6, pp 213–220], although we will discuss some additional differences in Section 2.5.

2.3.4 Simple Consequences Related to the Drift Inequalities

1. Under the drift condition (D1), we are assuming that $V(x) \geq 1$ for all x in X hence we have the bound taken the supremum over arbitrary sample paths $\{\theta_i, i = 1, 2, \ldots\}$:

$$\mathbf{E}_{\theta,x}\left[V(X_n)\right] \leq \sup_{\{\theta_i \in \Theta, i=1,\ldots,n-1\}} P_\theta P_{\theta_1} P_{\theta_2} \cdots P_{\theta_{n-1}} V(x) \tag{2.10}$$

[1] Assumption (A.5) is given here simply for reference and the reader should refer to [6] to see this assumption in context.

$$
\begin{aligned}
&\leq \lambda^n V(x) + L\sum_{i=0}^{n-1}\lambda^i \\
&\leq V(x) + \frac{L}{1-\lambda} \\
&\leq \left(1+\frac{L}{1-\lambda}\right)V(x) \qquad \text{for all } n = 1,2,\ldots
\end{aligned}
\tag{2.11}
$$

Therefore, if we define the constant $C_D \doteq 1 + \frac{L}{1-\lambda}$ so that for all $n = 1, 2, \ldots$, then

$$
\sup_{\theta\in\Theta} \mathbf{E}_{\theta,x}\left[V(X_n)\right] \leq C_D V(x), \qquad x \in \mathsf{X}, \tag{2.12}
$$

and we see that (D1) implies (D0).

By the same argument under (D1), we also see that

$$
\sup_{\theta\in\Theta} P_\theta^n V(x) \leq C_D V(x), \qquad n = 1, 2, \ldots \tag{2.13}
$$

2. It follows readily from Jensen's Inequality that under (D0), which defines some function V, we may also show a (D0)-like condition involving V^r for any constant $0 \leq r \leq 1$, i.e.

$$
\mathbf{E}_{\theta,x}\left[V^r(X_n)\right] \leq \left(\mathbf{E}_{\theta,x}\left[V(X_n)\right]\right)^r \leq C_D^r V^r(x) \qquad \text{for all } \theta \text{ in } \Theta \tag{2.14}
$$

all $n = 0, 1, 2, \ldots$, and x in X.

3. The type of problems we consider will generally satisfy (D1), or (D2) with a finite V. Our general convergence results for this chapter will always assume (D0), while the conditions (D1) or (D2) serve as tools to show (D0) in applications.

4. If the family of Markov chains given by $\{P_\theta, \theta \in \Theta\}$ are irreducible and positive recurrent over Θ, then condition (D1), or condition (D2) with a finite V, implies via Theorem 14.3.7 in [79] that

$$
\sup_{\theta\in\Theta} \pi_\theta(V) < \infty. \tag{2.15}
$$

2.3.5 Relationship between (D1) and (D2)

There is a very strong relationship between (D1) and (D2) and we use the next lemma to go back and forth between the two. This lemma is a slight generalization of Lemma 15.2.8 in [79, p. 370] which takes into account the uniformity over Θ of (D1) and (D2):

Lemma 2.1 *The drift condition (D2) holds with a petite set [2] C if and only if V is unbounded off petite sets and (D1) holds.*

Proof: See the Appendix.

[2] See [79] for the definition.

2.4 The BMP Decomposition of the SA Algorithm

For the projected SA algorithm

$$\theta_{n+1} = \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}^2 \rho_{n+1}(\theta_n, X_{n+1}) \right\}, \quad n = 0, 1, \ldots$$

we define the sequence of *noise* terms $\{\varepsilon_{n+1}, n = 0, 1, \ldots\}$ so that

$$\varepsilon_{n+1} \doteq H(\theta_n, X_{n+1}) - h(\theta_n) + \gamma_{n+1} \rho_{n+1}(\theta_n, X_{n+1}). \tag{2.16}$$

We also define a *projection process* $\{z_{n+1}, \ n = 0, 1, \ldots\}$ as in [61, 64], so we can rewrite the algorithm as:

$$\begin{aligned} \theta_{n+1} &= \theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}^2 \rho_{n+1}(\theta_n, X_{n+1}) + \gamma_{n+1} z_{n+1} && (2.17) \\ &= \theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1} \varepsilon_{n+1} + \gamma_{n+1} z_{n+1}, \qquad n = 0, 1, 2, \ldots && (2.18) \end{aligned}$$

Our main goal now is to study the noise sequence in a manner that allows us to show condition (KC4) in the Kushner-Clark Lemma. We first perform a decomposition of the noise terms using a variation of a method from [6, p. 220].

Assuming condition (P1) so that the solution ν_θ to the Poisson equation exists, then

$$\begin{aligned} \varepsilon_{k+1} &= \{H(\theta_k, X_{k+1}) - h(\theta_k)\} + \gamma_{k+1} \rho_{k+1}(\theta_k, X_{k+1}) \\ &= \{\nu_{\theta_k}(X_{k+1}) - P_{\theta_k} \nu_{\theta_k}(X_{k+1})\} + \gamma_{k+1} \rho_{k+1}(\theta_k, X_{k+1}) \\ &= \{\nu_{\theta_k}(X_{k+1}) - P_{\theta_k} \nu_{\theta_k}(X_k)\} \\ &+ \{P_{\theta_k} \nu_{\theta_k}(X_k) - P_{\theta_k} \nu_{\theta_k}(X_{k+1})\} + \gamma_{k+1} \rho_{k+1}(\theta_k, X_{k+1}). \end{aligned}$$

If $m < n$, the step-size weighted sum of the noise terms is formed and then rearranged:

$$\begin{aligned} \sum_{k=m}^{n-1} \gamma_{k+1} \varepsilon_{k+1} &= \sum_{k=m}^{n-1} \gamma_{k+1} \{\nu_{\theta_k}(X_{k+1}) - P_{\theta_k} \nu_{\theta_k}(X_k)\} \\ &+ \sum_{k=m}^{n-1} \gamma_{k+1} \{P_{\theta_k} \nu_{\theta_k}(X_k) - P_{\theta_k} \nu_{\theta_k}(X_{k+1})\} + \sum_{k=m}^{n-1} \gamma_{k+1}^2 \rho_{k+1}(\theta_k, X_{k+1}) \\ &= \sum_{k=m}^{n-1} \gamma_{k+1} \{\nu_{\theta_k}(X_{k+1}) - P_{\theta_k} \nu_{\theta_k}(X_k)\} \\ &+ \gamma_{m+1} P_{\theta_m} \nu_{\theta_m}(X_m) + \sum_{k=m+1}^{n-1} \gamma_{k+1} \left\{ P_{\theta_k} \nu_{\theta_k}(X_k) - P_{\theta_{k-1}} \nu_{\theta_{k-1}}(X_k) \right\} \\ &- \gamma_n P_{\theta_{n-1}} \nu_{\theta_{n-1}}(X_n) + \sum_{k=m+1}^{n-1} (\gamma_{k+1} - \gamma_k) P_{\theta_{k-1}} \nu_{\theta_{k-1}}(X_k) \\ &+ \sum_{k=m}^{n-1} \gamma_{k+1}^2 \rho_{k+1}(\theta_k, X_{k+1}) \end{aligned}$$

Now define for $k = m, m+1, \ldots, n-1$:

$$\begin{aligned}
\varepsilon_{k+1}^{(1)} &\doteq \nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k) \\
\varepsilon_{k+1}^{(2)} &\doteq P_{\theta_k}\nu_{\theta_k}(X_k) - P_{\theta_{k-1}}\nu_{\theta_{k-1}}(X_k) \\
\varepsilon_{k+1}^{(3)} &\doteq \frac{\gamma_{k+1} - \gamma_k}{\gamma_{k+1}} P_{\theta_{k-1}}\nu_{\theta_{k-1}}(X_k) \\
\varepsilon_{k+1}^{(4)} &\doteq \gamma_{k+1}\rho_{k+1}(\theta_k, X_{k+1}) \\
\eta_{m;n} &\doteq \gamma_{m+1}P_{\theta_m}\nu_{\theta_m}(X_m) - \gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)
\end{aligned}$$

and if $m < n$ we have the decomposition which is a variant of Lemma 1 in [6, p. 222]:

$$\begin{aligned}
&\sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1} \qquad (2.19) \\
&= \sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(1)} + \sum_{k=m+1}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(2)} + \sum_{k=m+1}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(3)} + \sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(4)} + \eta_{m;n}
\end{aligned}$$

This approach taken in this decomposition is a version of the *state perturbation method*[3] described in [64].

2.5 Variations on the BMP Lemmas

In this section, we adapt to our framework each of the Lemma's 2 through 6 in [6, pp. 223-228] which provide a bound for each term in the decomposition (2.19). These adapted lemmas are then collected in Proposition 2.7 of the next section to show the overall sum of the step-size weighted noise is almost surely convergent to a finite rv. As we will later see, this is an approach to proving the Kushner-Clark (KC4)-type noise condition in the Kushner-Clark Lemma.

Lemma 2.2 (Variant of BMP Lemma 2) *Assume (D0), (P1), (P2) hold for any positive constant* $r \le \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies (S).*

1. *There exists a constant* $A_1 < \infty$ *such that for each* $m = 1, 2, \ldots$

$$\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}\right\|^2\right] \le A_1 V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.$$

   Moreover, $A_1 \le 4pC_\nu^2 C_D$.

2. $\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}$ *converges* $\mathbf{P}_{\theta,x}$*-a.s. to a finite rv.*

[3] Contrary to our nomenclature, in [64] the process $\{\theta_n,\ n = 1, 2, \ldots\}$ is referred to as the "state process".

Proof: Consider the sum

$$\bar{M}_n = \sum_{k=0}^{n-1} \gamma_{k+1} \left(\nu_{\theta_k}(X_{k+1}) - P_{\theta_k} \nu_{\theta_k}(X_k) \right), \qquad n = 1, 2, \ldots$$

which is a vector martingale since (by the Markov property)

$$\mathbf{E}\left[\nu_{\theta_k}(X_{k+1}) | \mathcal{F}_k \right] = P_{\theta_k} \nu_{\theta_k}(X_k).$$

The vector $\bar{M}_n$ is a p-dimensional vector, and although convergence results exist for vector martingales [76], we find it simpler to consider each of the p components separately by defining the i^{th} component vector as

$$M_n^{(i)} = \sum_{k=0}^{n-1} \gamma_{k+1} \left(\nu_{\theta_k}(X_{k+1}) - P_{\theta_k} \nu_{\theta_k}(X_k) \right)^{(i)}, \qquad n = 1, 2, \ldots,$$

For brevity, let us now drop the $^{(i)}$ in this definition and consider any of the p components of the vector martingale as:

$$M_n = \sum_{k=0}^{n-1} \gamma_{k+1} \left(\nu_{\theta_k}(X_{k+1}) - P_{\theta_k} \nu_{\theta_k}(X_k) \right), \qquad n = 1, 2, \ldots$$

Clearly, each component of M_n above also has the martingale property.

Incremental orthogonality and Pythagoras formula [108, p.110] yield

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[M_n^2 \right] &= \mathbf{E}_{\theta,x}\left[M_1^2 \right] + \sum_{k=2}^{n} \mathbf{E}_{\theta,x}\left[(M_k - M_{k-1})^2 \right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[\left(\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k) \right)^2 \right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[\mathbf{E}\left[\left(\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k) \right)^2 | \mathcal{F}_k \right] \right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[\mathbf{E}\left[\left(\nu_{\theta_k}(X_{k+1}) \right)^2 | \mathcal{F}_k \right] - \left(P_{\theta_k}\nu_{\theta_k}(X_k) \right)^2 \right] \\
&\leq \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[\left(\nu_{\theta_k}(X_{k+1}) \right)^2 \right] \\
&\leq C_\nu^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[V^{2r}(X_{k+1}) \right] \\
&\leq C_\nu^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[V^{1}(X_{k+1}) \right]
\end{aligned}$$

where we have used (P2) in the second to last line. The last line follows since $2r \leq \frac{1}{1+\hat{\ell}_1} \leq 1$. Applying (D0) to the last line we find

$$\mathbf{E}_{\theta,x}\left[M_n^2 \right] \leq C_\nu^2 C_D V(x) \sum_{k=0}^{n-1} \gamma_{k+1}^2$$

The bound in the first part of the lemma follows from Doob's inequality

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\bar{M}_n\right\|^2\right] &= \mathbf{E}_{\theta,x}\left[\sup_{n\le m}\sum_{i=1}^{p}\left(\bar{M}_n^{(i)}\right)^2\right] \\
&\le \sum_{i=1}^{p}\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left(M_n^{(i)}\right)^2\right] \\
&\le \sum_{i=1}^{p} 4\sup_{n\le m}\mathbf{E}_{\theta,x}\left[\left(M_n^{(i)}\right)^2\right] \\
&\le 4pC_\nu^2 C_D V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2, \qquad x\in \mathsf{X}, \quad \theta\in\Theta.
\end{aligned}$$

Under (S), $\sum_{k=0}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1}<\infty$ and it then follows that $\sum_{k=0}^{\infty}\gamma_{k+1}^2<\infty$ since $\gamma_k\downarrow 0$ and there exists a k' such that $\gamma_{k'}<1$, hence $\gamma_k^2\le\gamma_k^{1+\widehat{\ell}_1}$ for all $k\ge k'$.

For the convergence properties in the second part of the lemma, we note that $\sum_{k=0}^{\infty}\gamma_{k+1}^2<\infty$ which implies that each component martingale of the vector martingale converges a.s. to a finite random variable (as well as converging in L^2) since it is bounded in L^2 [108]. ■

Lemma 2.3 (Variant of BMP Lemma 3) *Assume (D0), (H2), (P1), (P3) for any positive* $r\le\frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0<\widehat{\ell}_1<1$ *is determined from (S). There exists a constant* $A_2<\infty$ *such that for all* $m=1,2,\ldots$,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(2)}\right\|\right)^2\right]\le A_2V(x)\left(\sum_{k=1}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)^2, \qquad x\in\mathsf{X},\quad\theta\in\Theta.$$

Moreover, $A_2\le 4C_\delta^2\left(C_H+\gamma_1C_\rho\right)^{2\widehat{\ell}_1}C_D$.

Proof: Under (P3),

$$\|P_\theta\nu_\theta(x)-P_{\theta'}\nu_{\theta'}(x)\|\le C_\delta V^r(x)\left\|\theta-\theta'\right\|^{\widehat{\ell}_1}, \qquad x\in\mathsf{X},\quad \theta,\theta'\in\Theta. \tag{2.20}$$

Also, the nearest point projection term is bounded by

$$\|z_k\|\le\|H(\theta_{k-1},X_k)+\gamma_k\rho_k(\theta_{k-1},X_k)\|$$

which follows since $\theta_k\in\Theta$ and, at the very least, the projection term can return the iterate to this point so $\theta_{k+1}=\theta_k\in\Theta$. Hence for $k=1,2,\ldots$ we have from (H2) and the definition of the SA that

$$\begin{aligned}
\|\theta_k-\theta_{k-1}\| &\le \gamma_k\|H(\theta_{k-1},X_k)+\gamma_k\rho_k(\theta_{k-1},X_k)+z_k\| &(2.21)\\
&\le 2\gamma_k\|H(\theta_{k-1},X_k)+\gamma_k\rho_k(\theta_{k-1},X_k)\| &(2.22)\\
&\le 2C_H\gamma_kV^r(X_k)+2C_\rho\gamma_k^2V^r(X_k) &\\
&\le 2\left(C_H+\gamma_1C_\rho\right)\gamma_kV^r(X_k) &(2.23)
\end{aligned}$$

Thus by (2.20) and (2.23),

$$
\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(2)}\right\|\right)^2\right] \\
&= \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|P_{\theta_k}\nu_{\theta_k}(X_k)-P_{\theta_{k-1}}\nu_{\theta_{k-1}}(X_k)\right\|\right)^2\right] \\
&\leq \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}C_\delta V^r(X_k)\left\|\theta_k-\theta_{k-1}\right\|^{\widehat{\ell}_1}\right)^2\right] \\
&\leq 4\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}C_\delta V^r(X_k)\left(C_H+\gamma_1 C_\rho\right)^{\widehat{\ell}_1}\gamma_k^{\widehat{\ell}_1}V^{r\widehat{\ell}_1}(X_k)\right)^2\right] \\
&= 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}V^{r(1+\widehat{\ell}_1)}(X_k)\right)^2\right] \\
&\leq 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}V^{1/2}(X_k)\right)^2\right].
\end{aligned}
$$

since $r(1+\widehat{\ell}_1) \leq 1/2$. By treating the sum as an inner product and applying the Schwarz inequality [50, p. 2] to the last line

$$
\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_k\left\|\varepsilon_{k+1}^{(2)}\right\|\right)^2\right] \\
&\leq 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}V(X_k)\right] \\
&\leq 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[V(X_k)\right] \\
&\leq 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)C_D V(x)\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}.
\end{aligned}
$$

■

Lemma 2.4 (Variant of BMP Lemma 4) *Assume (D0), (P1), (P2) for any positive constant* $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies (S). There exists a constant* $A_3 < \infty$ *such that for all* $m = 1, 2, \ldots,$,

$$
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(3)}\right\|\right)^2\right] \leq A_3 V(x)\gamma_1^2, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.
$$

Moreover, $A_3 \leq C_\nu^2 C_D$.

Proof: Applying (P2)

$$\begin{aligned}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(3)}\right\|\right)^2\right] &= \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\left\|P_{\theta_{k-1}}\nu_{\theta_{k-1}}(X_k))\right\|\right)^2\right] \\ &\le \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})C_\nu V^r(X_k)\right)^2\right]\end{aligned}$$

Next, the Schwarz inequality yields

$$\begin{aligned}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(3)}\right\|\right)^2\right] &\le C_\nu^2\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\right)\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})V^{2r}(X_k)\right] \\ &\le C_\nu^2\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\right)\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\,\mathbf{E}_{\theta,x}\left[V^1(X_k)\right] \\ &\le C_\nu^2\gamma_1\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})C_D V(x) \\ &\le C_\nu^2 C_D V(x)\gamma_1^2\end{aligned}$$

■

Lemma 2.5 (Variant of BMP Lemma 5) *Assume (D0), (P1), (H2) for any positive constant* $r \le \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies (S). There exists a constant* $A_4 < \infty$ *such that for all* $m = 1, 2, \ldots$,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(4)}\right\|\right)^2\right] \le A_4 V(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)^2, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.$$

Moreover, $A_4 \le C_D C_\rho^2$.

Proof: First we have from (H2):

$$\gamma_{k+1}\left\|\varepsilon_{k+1}^{(4)}\right\| = \gamma_{k+1}^2\left\|\rho_k(\theta_k, X_{k+1})\right\| \le \gamma_{k+1}^2 C_\rho V^r(X_{k+1})$$

Hence,

$$\begin{aligned}&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(4)}\right\|\right)^2\right] \\ \le\ & C_\rho^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2 V^r(X_{k+1})\right)^2\right] \\ \le\ & C_\rho^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)\sum_{k=0}^{m-1}\gamma_{k+1}^2 V^{2r}(X_{k+1})\right], \qquad m = 1, 2, \ldots\end{aligned}$$

where the last line follows from the Schwarz inequality. We have $r(1+\widehat{\ell}_1) \le 1/2$, so

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(4)}\right\|\right)^2\right] &\le C_\rho^2\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)\mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1}\gamma_{k+1}^2 V^1(X_{k+1})\right] \\
&\le C_\rho^2\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)C_D V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2, \qquad m=1,2,\ldots
\end{aligned}$$

■

Lemma 2.6 (Variant of BMP Lemma 6) *Assume (D0), (P1), (P2) for any positive constant* $r \le \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies (S).*

1. *There exists a constant* $A_5 < \infty$ *such that for each* $m = 1, 2, \ldots,$

$$\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m}\|\eta_{0;n}\|^2\right] \le A_5 V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.$$

   Moreover, $A_5 \le 4C_D C_\nu^2$.

2. *As* $n \to \infty$ *we have that* $\eta_{0;n}$ *converges a.s.*

Proof: Recall that $\eta_{0;n} \doteq \gamma_1 P_{\theta_0}\nu_{\theta_0}(X_0) - \gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)$ for $n = 1, 2, \ldots$.

First we have $X_0 = x$ a.s. and under (P2)

$$\begin{aligned}
\|\gamma_1 P_{\theta_0}\nu_{\theta_0}(x)\|^2 &\le \gamma_1^2 C_\nu^2 V^{2r}(x) \\
&\le \gamma_1^2 C_\nu^2 V^1(x)
\end{aligned}$$

Also, for each $m = 1, 2, \ldots$

$$\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m}\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\|^2\right] \le C_\nu^2\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m}\gamma_n^2 V^{2r}(X_n)\right]$$

Thus,

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m}\|\eta_{0;n}\|^2\right] &= \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m}\left\|\gamma_1 P_{\theta_0}\nu_{\theta_0}(X_0) - \gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\|^2\right] \\
&\le \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m}\left(2\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\|^2 + 2\|\gamma_1 P_{\theta_0}\nu_{\theta_0}(x)\|^2\right)\right] \\
&\le 2\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m}\gamma_k^2 V^{2r}(X_k)C_\nu^2\right] + 2\gamma_1^2 C_\nu^2 V^{2r}(x) \\
&\le 2C_\nu^2\sum_{k=1}^{m}\gamma_k^2\mathbf{E}_{\theta,x}\left[V^1(X_k)\right] + 2\gamma_1^2 C_\nu^2 V^1(x) \\
&\le 2C_D C_\nu^2 V^1(x)\sum_{k=1}^{m}\gamma_k^2 + 2\gamma_1^2 C_\nu^2 V^1(x) \\
&\le 4C_D C_\nu^2 V(x)\sum_{k=1}^{m}\gamma_{k+1}^2, \qquad m = 1, 2, \ldots
\end{aligned}$$

and recalling $C_D \geq 1$ for the last line.

To prove the lemma's second conclusion, we have for each $n = 1, 2, \ldots$

$$\begin{aligned}\mathbf{E}_{\theta,x}\left[\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\|^2\right] &\leq \gamma_n^2 C_\nu^2 \mathbf{E}_{\theta,x}\left[V^{2r}(X_n)\right] \\ &\leq \gamma_n^2 C_\nu^2 C_D V(x), \qquad x \in \mathsf{X}.\end{aligned}$$

Therefore,

$$\mathbf{E}_{\theta,x}\left[\sum_{n=1}^{\infty}\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\|^2\right] \leq C_\nu^2 C_D V(x)\sum_{n=1}^{\infty}\gamma_n^2 < \infty, \qquad x \in \mathsf{X},$$

This implies the sum $\sum_{n=0}^{\infty}\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\|^2$ converges to a finite rv $\mathbf{P}_{\theta,x}$-a.s. Hence, $\lim_{n\to\infty}\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-}}\right.$ 0 $\mathbf{P}_{\theta,x}$-a.s. and thus

$$\lim_{n\to\infty}\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\| = 0, \qquad \mathbf{P}_{\theta,x} - a.s.$$

Therefore,

$$\lim_{n\to\infty}\|\eta_{0;n} - \gamma_1 P_{\theta_0}\nu_{\theta_0}(x)\| = \lim_{n\to\infty}\left\|\gamma_n P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\| = 0, \qquad \mathbf{P}_{\theta,x} - a.s.$$

■

2.5.1 Remarks

The above lemmas and their proofs are similar to the development in BMP [6] despite the fact that we have made several significant changes. We acknowledge this similarity and consider these lemmas to be *variations* of BMP's originals. Let us summarize how these lemmas differ from BMP's:

1. As previously mentioned, we have changed BMP's growth in x factor, which is of the form $const(1 + \|x\|^q)$, to the more general function $V(x)$ compatible with the theory of V-uniformly ergodic Markov chains, and satisfying (D0). Although this change is not essential, we find it convenient and several applications of interest to us at this time satisfy the conditions for V-uniformly ergodicity, see Meyn and Tweedie [79] for many examples. As a result, BMP's assumption (A.5) is replaced by our (D0) (which can be weaker).

2. We have already mentioned that BMP make the assumption in their Lemma's 1 through 6 and Proposition 7 that the regression function $h : \Theta \to \mathbb{R}^p$ is locally Lipschitz. For our versions of these lemmas, we do not actually assume any continuity conditions whatsoever on h, although in later sections we shall consider either basic continuity or a Hölder continuity.

3. We identify a trade-off through the condition $r(1 + \widehat{\ell}_1) \leq 1/2$ which affects the space of allowable functions $H(\theta, x)$ permitted by condition (H2) versus the space of allowable step-size sequences $\{\gamma_{k+1},\ k = 0, 1, \ldots\}$ satisfying (S). This trade-off will be discussed in Section 3.8.

4. Since our focus in applications is not on the global unconstrained convergence problem but for projected SA algorithms on a *compact* projection set Θ, we turn our attention to verifying our conditions over the entire compact projection set Θ. We also treat as a special case the situation which arises when our conditions can only be verified for a particular subset Q of the projection set Θ and we develop at least one possible approach in Chapter 6 to show an unconditional convergence in this setting. While this particular approach is very much problem dependent, we do allow the precise boundaries of this set Q to be unknown *a priori*.

5. Although we are mainly interested in the application of projected algorithms yielding an *unconditional* almost sure convergence, our framework does extend to the unconstrained case, i.e. take the projection set Θ to be equal to $\mathbb{R}^p$, but the convergence results via the Kushner-Clark Lemma will then be conditional.

2.6 Main Properties of the Noise

In this section, we collect the results of the previous sections Lemmas for each term in the decomposition and show several bounds on the "step-size weighted sum of error" sequence. This next result is a variant of BMP's Proposition 7 [6].

Proposition 2.7 *Assume (D0), (P1)-(P3), (H1)-(H2) hold for some positive* $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ *and* $0 < \widehat{\ell}_1 < 1$ *satisfies (S). There exist finite constants* B_1, B_2, B_3 *such that for all* $m = 1, 2, \ldots$*:*

1. *We have*

$$\mathbf{E}_{\theta,x}\left[\sup_{n\leq m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}\right\|^2\right] \leq B_1 V(x)\left(1+\sum_{k=0}^{m-1}\gamma_{k+1}^{2\widehat{\ell}_1}\right)\sum_{k=0}^{m-1}\gamma_{k+1}^2. \tag{2.24}$$

2. *We have*

$$\mathbf{E}_{\theta,x}\left[\sup_{n\leq m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}\right\|^2\right] \leq B_2 V(x)\left(\gamma_1^{1-\widehat{\ell}_1}+\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1} \tag{2.25}$$

3. *And since* $\sum_{k=0}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1} < \infty$*:*

   (a)

$$\mathbf{E}_{\theta,x}\left[\sup_{n\geq 1}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}\right\|^2\right] \leq B_3 V(x)\sum_{k=0}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1}. \tag{2.26}$$

   (b) the series $\sum_k \gamma_{k+1}\varepsilon_{k+1}$ *converges* $\mathbf{P}_{\theta,x}$*-a.s.*

Proof: Using the BMP decomposition

$$
\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}\right\|^2\right] \\
&\le \mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left(\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}\right\| + \left\|\sum_{k=1}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(2)}\right\|\right.\right. \qquad (2.27)\\
&\qquad \left.\left. + \left\|\sum_{k=1}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(3)}\right\| + \left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(4)}\right\| + \|\eta_{0;n}\|\right)^2\right] \\
&\le 8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}\right\|^2\right] + 8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=1}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(2)}\right\|^2\right] \\
&\qquad +8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=1}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(3)}\right\|^2\right] + 8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(4)}\right\|^2\right] \\
&\qquad +8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\|\eta_{0;n}\|^2\right] \\
&\le 8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}\right\|^2\right] + 8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left(\sum_{k=1}^{n-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(2)}\right\|\right)^2\right] \\
&\qquad +8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left(\sum_{k=1}^{n-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(3)}\right\|\right)^2\right] + 8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left(\sum_{k=0}^{n-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(4)}\right\|\right)^2\right] \\
&\qquad +8\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\|\eta_{0;n}\|^2\right] \qquad (2.28)
\end{aligned}
$$

By Lemmas 2.2 through 2.6 we have shown the entire expression (2.28) is bounded by a sum of expressions which are one of the following types:

$$
const.V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2 \qquad \text{or} \qquad const.V(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)^2.
$$

It's clear under (S) that both of these converge.

By squaring the Schwarz inequality, i.e.

$$
\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)^2 \le \left(\sum_{k=0}^{m-1}\gamma_{k+1}^{2\widehat{\ell}_1}\right)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right),
$$

we have the first conclusion since

$$
\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}\right\|^2\right] \\
&\le const.V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2 + const.V(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)^2 \qquad (2.29)\\
&\le const.V(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)\left(1+\sum_{k=0}^{m-1}\gamma_{k+1}^{2\widehat{\ell}_1}\right). \qquad (2.30)
\end{aligned}
$$

The second part is a variation on the first part where we notice

$$\gamma_{k+1}^2 = \gamma_{k+1}^{1-\widehat{\ell}_1}\gamma_{k+1}^{1+\widehat{\ell}_1} \le \gamma_1^{1-\widehat{\ell}_1}\gamma_{k+1}^{1+\widehat{\ell}_1}, \qquad k = 1, 2, \ldots,$$

hence

$$\sum_{k=0}^{m-1} \gamma_{k+1}^2 \le \gamma_1^{1-\widehat{\ell}_1} \sum_{k=0}^{m-1} \gamma_{k+1}^{1+\widehat{\ell}_1}$$

Thus starting from (2.29),

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}\right\|^2\right] \\
&\le const.V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2 + const.V(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)^2 \\
&\le const.\gamma_1^{1-\widehat{\ell}_1}V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1} + const.V(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)^2 \\
&= V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\left(const.\gamma_1^{1-\widehat{\ell}_1} + const.\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)\right)
\end{aligned}$$

Finally, to show convergence of the series $\sum_k \gamma_{k+1}\varepsilon_{k+1}$, recall the decomposition:

$$\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1} = \sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)} + \sum_{k=1}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(2)} + \sum_{k=1}^{n-1}\gamma_1\varepsilon_{k+1}^{(3)} + \sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(4)} + \eta_{0;n} \qquad (2.31)$$

From Lemmas 2.2 and 2.6, the first term $\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}$ and the last term $\eta_{0;n}$ converge a.s. The remaining terms for $i = 2, 3, 4$ all satisfy

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{\infty}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(i)}\right\|\right)^2\right] < \infty, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.$$

Also, by Jensen's inequality, for $i = 2, 3, 4$,

$$\mathbf{E}_{\theta,x}\left[\sum_{k=0}^{\infty}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(i)}\right\|\right] < \infty, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.$$

which implies $\sum_{k=0}^{\infty}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(i)}\right\|$ each converge a.s. to a finite r.v. (since the series is positive term). Thus, for each $i = 2, 3, 4$, the series $\sum_k \gamma_{k+1}\varepsilon_{k+1}^{(i)}$ converges a.s. since each component vector converges absolutely. Therefore, the series $\sum_k \gamma_{k+1}\varepsilon_{k+1}$ converges almost surely to a finite rv. ■

2.6.1 A Modest Extension

If we look back at the developments up to now, it's clear a result nearly identical to Proposition 2.7 is also possible if the SA recursion is modified to the following form:

$$\theta_{n+1} = \Pi_\Theta\left\{\theta_n + \gamma_{n+1}H(\theta_n, X_{n+1}) + \gamma_{n+1}^{1+\widehat{\ell}_1}\rho_{n+1}(\theta_n, X_{n+1})\right\}, \qquad n = 1, 2, \ldots$$

The changes required to the BMP decomposition, the Lemmas and Proposition 2.7 are minimal and obvious. We do claim this to be an extension as the ρ_{n+1} terms decay more slowly with results nearly identical to Proposition 2.7 still achieved. In fact, this modest extension will become critical in Chapter 5 when we consider a stochastic optimization algorithm.

2.7 Finalizing the Convergence Analysis

So far, we have shown a bound involving the error term ε_n under a new set of conditions recast from BMP's framework. We have weakened the Lipschitz condition on the regression function $h(\cdot)$. We have also recast the framework leading to this bound using (D0) so we can easily apply Meyn and Tweedie's drift criteria results in the next chapter. Actually, up to this point, dropping the Lipschitz condition on h has had absolutely no effect on the outcome of our results because it has not entered into the lemma's so far. Now, as we desire to show convergence, we shall assume at minimum a simple continuity condition on h although in the next chapter we will strengthen this to a "Hölder" form. Lipschitz continuity, as was assumed by BMP, will not be required here. Let us define the conditions:

(H3) The regression function $h : \Theta \to \mathbb{R}^p$ is continuous.

(H4) For some $0 < \ell < 1$, there exists a $C_h < \infty$ and a $\delta > 0$ such that:

$$\|h(\theta) - h(\theta')\| \leq C_h \|\theta - \theta'\|^{\ell}, \qquad \theta, \theta' \in \Theta,\ \|\theta - \theta'\| \leq \delta.$$

Up to this point we have followed the BMP monograph [6] fairly closely and we have shown an analogous lemma for each lemma in BMP which bounds the corresponding terms of the BMP decomposition. To make the next step to show convergence we will not be able to continue using a version of BMP's approach under our modified framework; nor can we simply cite BMP's final convergence results [6, Theorems 13-15 , pp. 236–239] since these require the regression function h to be locally Lipschitz continuous. Attempts to relax BMP's Lipschitz condition on $h(\cdot)$ by adapting their development runs into immediate difficulties since BMP's Lemma 8 [6, p. 231] does not appear to extend to our framework without the Lipschitz condition on h, even under (H4). As a result, we are unable to even show that the iterates $\{\theta_n, n = 1, 2, \ldots\}$ tend to converge to the solution $\bar{\theta}(t)$ of the ODE

$$\begin{aligned} \dot{\bar{\theta}}(t) &= h(\bar{\theta}(t)), \qquad t \geq 0, \\ \bar{\theta}(0) &= \theta. \end{aligned}$$

via BMP's method under (H4).

Other versions of Métivier and Priouret's work [77, 76] can also be found in the literature. In particular, [77] is notable as it was instrumental in linking the convergence properties of the SA to a Lipschitz condition on the Poisson equation solution. Under our conditions, this in itself eliminates it as a candidate for showing convergence as we are are assuming the weaker

Hölder form (P3) on the Poisson equation solution. Interestingly though, [77] does not assume the regression function h is Lipschitz and merely assumes continuity, but unfortunately, it also assumes the following boundedness condition on the observation/measurement function $H(\theta, x)$:

(F) For every $R > 0$ there exists a constant M_R such that

$$\sup_{\|\theta\| \le R} \sup_x \|H(\theta, x)\| \le M_R$$

We feel (F) is undesirable for some of the queueing system applications we have in mind with driving functions $H(\cdot, \cdot)$ which are unbounded in the state variable $x \in \mathsf{X}$.

2.7.1 Kushner-Clark Lemma for the Projected Algorithm

The projected version of the Kushner-Clark Lemma [61, p. 191, Thm. 5.3.1] provides a means to show an unconditional convergence for the iterates produced by a projected algorithm. This is precisely the approach taken in [70] to show convergence of stochastic approximation iterates driven by *finite state* Markov chains. Ma, Makowski and Shwartz. [70] cited the work of Métivier and Priouret [77] coupled with the Kushner-Clark Lemma applied to a projected SA algorithm.

For some **compact** projection Θ satisfying (KC0) below, Kushner and Clark consider the following projected recursion

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \Pi_\Theta \left\{\theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\varepsilon_{n+1} + \gamma_{n+1}\beta_{n+1}\right\}, \qquad n = 0, 1, \ldots \tag{2.32}$$

We now summarize the Kushner Clark assumptions needed for this approach:

(KC0) $\Theta \doteq \{\theta : q_i(\theta) \le 0,\ i = 1, \ldots, s\}$ is the closure of its interior and is bounded. The $q_i : \mathbb{R}^p \to \mathbb{R}, i = 1, 2, \ldots, s$ are continuously differentiable functions defining Θ. At each boundary point $\theta \in \delta\Theta$, the gradients of the active constraints are linearly independent.

(KC1) $h(\cdot)$ is a continuous function.

(KC2) $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$ is a sequence of positive real numbers such that $\gamma_n > 0$, $\gamma_n \to 0$, and $\sum_{n=0}^{\infty} \gamma_{n+1} = \infty$.

(KC3) $\{\beta_{n+1},\ n = 0, 1, \ldots\}$ is a bounded (w.p.1) sequence tending to zero with probability one.

(KC4) There is a $T > 0$ such that for each $\epsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}\left[\sup_{j \ge n} \max_{t \le T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} \gamma_{i+1}\varepsilon_{i+1} \right| \ge \epsilon \right] = 0. \tag{2.33}$$

Recall the times

$$t_0 = 0, \qquad t_n \;=\; \sum_{k=0}^{n-1} \gamma_{k+1}, \qquad n = 1, 2, \ldots$$

and then construct the *piecewise linear interpolated* function $\theta^0(t)$ by defining

$$\begin{aligned} \theta^0(t_n) &= \theta_n, \\ \theta^0(t) &= \frac{(t_{n+1} - t)}{\gamma_n}\theta_n + \frac{(t - t_n)}{\gamma_n}\theta_{n+1}, \qquad t \in (t_n, t_{n+1}) \end{aligned}$$

for each $n = 0, 1, \ldots$. Also, define a sequence of functions $\{\theta^n(\cdot) : n = 0, 1, \ldots\}$ which are time shifts of the piecewise linear interpolated function:

$$\theta^n(t) = \begin{cases} \theta^0(t + t_n), & t \geq -t_n \\ \theta_0, & t \leq -t_n, \end{cases} \qquad n = 0, 1, \ldots$$

We shall refer to the following as the Kushner-Clark Lemma for the constrained case, see [61, p. 191].

Lemma 2.8 (Kushner-Clark) *Assume (KC0)-(KC4). There is a null set* Ω_0 *such that if* $\omega \notin \Omega_0$*, the following hold.*

1. $\theta^0(\cdot)$ *is bounded and uniformly continuous on* $[0, \infty)$.
2. *If* $\theta(\cdot)$ *is the limit of a convergent subsequence of* $\{\theta^n(\cdot)\}$*, then* $\theta(\cdot)$ *satisfies the ODE*

$$\dot{\theta} = \bar{\Pi}_\theta\{h(\theta)\}. \tag{2.34}$$

   where $\bar{\Pi}_\Theta\{h(\theta)\} = \lim_{0<\Delta\to 0} \frac{\Pi_\Theta\{\theta + \Delta h(\theta)\} - \theta}{\Delta}$.

3. *The set of stationary points of (2.34) is the set of Kuhn-Tucker points, denoted* KT*, where*

$$KT \doteq \left\{ \theta : \text{ there are } \lambda_i \geq 0 \text{ such that } \quad -h(\theta) + \sum_{i:q_i(\theta)=0} \lambda_i \nabla q_i(\theta) = 0 \right\}.$$

4. *Let* $\theta^\star$ *denote an asymptotically stable point (which must be in* KT*) of (2.34) with domain of attraction* $DA(\theta^\star)$*. If* $Q \subseteq DA(\theta^\star)$ *is compact and* $\theta_n \in Q$ *infinitely often, then* $\theta_n \to \theta^\star$ *as* $n \to \infty$.

2.7.2 Conditions (KC1)-(KC4)

Continuity of $h(\cdot)$ on Θ or (KC1) is certainly implied by our Hölder continuity condition (H4) in our framework. We are assuming the β_n terms are all zero as we have essentially modeled these terms as our ρ_n terms[4] which are included in (KC4). Thus, the significant remaining issue is to verify the noise condition (KC4).

Claim 2.9 *Condition (KC4) is implied by the almost sure convergence of*

$$\lim_{n\to\infty} \sum_{k=0}^{n} \gamma_{k+1}\varepsilon_{k+1}$$

to a finite rv.

Proof: Observe that we prove the claim if we show for all $\epsilon > 0$ that

$$\mathbf{P}\left[\limsup_{n\to\infty}\left[\sup_{j\geq n}\max_{t\leq T}\left\|\sum_{i-m(jT)}^{m(jT+t)-1}\gamma_{i+1}\varepsilon_{i+1}\right\|\geq\epsilon\right]\right]=0$$

Next, notice that for any $\epsilon > 0$ that

$$\limsup_{n\to\infty}\left[\sup_{j\geq n}\max_{t\leq T}\left\|\sum_{i-m(jT)}^{m(jT+t)-1}\gamma_{i+1}\varepsilon_{i+1}\right\|\geq\epsilon\right]\subset\limsup_{n\to\infty}\left[\sup_{m,p\geq n}\left\|\sum_{i=m}^{p}\gamma_{i+1}\varepsilon_{i+1}\right\|\geq\epsilon\right] \tag{2.35}$$

Since $\sum_{k=0}^{\infty}\gamma_{k+1}\varepsilon_{k+1}$ converges almost surely to a finite rv and thus forms a Cauchy sequence, the probability of the right hand side of (2.35) is zero for every epsilon $\epsilon > 0$. ■

2.8 Concluding Remarks

In this chapter we have developed a general framework and sufficient conditions to study the convergence for SA's taking observations from a Markov chain. In some ways these conditions are weaker than BMP's and they still imply the noise condition (KC4) so we can apply Kushner-Clark Lemma. Unfortunately, these general conditions are not much easier to verify, since they are in terms of Poisson's equation solution and the unknown regression function $h(\cdot)$. The next chapter remedies this situation by identifying verifiable conditions in terms of the transition probabilities which imply the general conditions of this chapter.

[4]There is a slight loss of generality in doing this since the β_{n+1} terms under (KC3) can almost surely converge to zero at a slower rate than the convergence rate of $\gamma_{n+1}^2\rho_{n+1}$ implied by (H2) and (S). We proceed nonetheless as this construction will be useful to bound the ε_n noise terms and the loss in generality appears to be minimal in applications. Note that we have offered a modified algorithm in the extension following Proposition 2.7 which brings the conditions slightly closer.

Chapter 3

Convergence for Geometrically Ergodic Markov Chains

For certain geometrically ergodic Markov chains, we introduce some new *specialized* conditions on the family of transition kernels $\{P_\theta, \theta \in \Theta\}$ and the driving function $H(\theta, x)$ for this large class of SA problems which ensure the more general conditions of the previous chapter are satisfied. We present a straightforward approach to check these specialized conditions which imply convergence via the ODE Method. As an example, we carry out the verification for an SA applied to a θ-dependent random walk with a reflection at the origin.

3.1 The Specialized Conditions

As in the previous chapter, we assume (D0) which identifies a function $V : \mathsf{X} \to [1, \infty)$ as well as (S) which identifies a constant $0 < \ell_1 < 1$ so that $\sum_{n=0}^{\infty} \gamma_{n+1}^{1+\widehat{\ell}_1} < \infty$. We also assume a fixed positive constant $r \le \frac{1}{2(1+\widehat{\ell}_1)}$.

Define now a new norm for vector valued functions $f : \mathsf{X} \to \mathbb{R}^p$ by a straightforward extension to the same norm for scalar valued functions defined in (1.13) and f-norm in [79]:

$$\|P_\theta^n(x, \cdot) - \pi_\theta(\cdot)\|_{V^r} \doteq \sup_{f : \|f\| \le V^r} \|P_\theta^n(x, f) - \pi_\theta(f)\| . \tag{3.1}$$

(Here, the symbol $\|\cdot\|$ with no subscript still represents the Euclidean norm.)

The following *specialized conditions* are defined in terms of the above $\widehat{\ell}_1$, r and V.

(E1) There exists constants $C_E < \infty$ and $0 < \rho < 1$ such that

$$\sup_{\theta \in \Theta} \|P_\theta^n(x, \cdot) - \pi_\theta(\cdot)\|_{V^r} \le C_E V^r(x) \rho^n, \qquad x \in \mathsf{X}, \quad n = 0, 1, 2, \ldots$$

(H5) There exists constants $C_5 < \infty$, $\delta_H > 0$, and $\widehat{\ell}_2 \in (\widehat{\ell}_1, 1)$ such that for all θ, $\theta' \in \Theta$ with $\|\theta - \theta'\| \le \delta_H$ and all $x \in \mathsf{X}$, we have

$$\|H(\theta, x) - H(\theta', x)\| \le C_5 V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_2}$$

(C) There exists constants $C_C < \infty$, $\delta_C > 0$, and $\widehat{\ell}_3 \in (\widehat{\ell}_2, 1]$ such that for each $n = 0, 1, \ldots$

$$\|P_\theta^n(x, \cdot) - P_{\theta'}^n(x, \cdot)\|_{V^r} \le n\, C_C V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_3} ,$$

for all θ, $\theta' \in \Theta$ with $\|\theta - \theta'\| \le \delta_C$, and all $x \in \mathsf{X}$.

3.1.1 Remarks

1. The linear growth with n of the bound in condition (C) will be very helpful in applications.
2. The assumptions (P1) and (E1) are both related to the drift criteria (D1), as explained in [79]. These relationships will be discussed later in Sections 3.4-3.5.
3. Condition (H5) can be somewhat restrictive; for example, it is not satisfied for $H(\theta, x) = 1_{\{x \geq \theta\}}$ which might be used to estimate the probability of buffer overflow in a queue. It turns out that for a certain subclass of problems, we may relax condition (H5) via an extension to our main results coming up. Thus, (H5) serves as a condition which should be checked first (since it's easily verified) and only if it should fail to be satisfied would we look into the extension.

3.2 Consequences of the Specialized Conditions

Here, we prove three theorems which imply the most difficult to verify of the general conditions proposed in the previous chapter leading to convergence for SA's; namely, conditions (H4), (P2), and (P3). These three theorems serve as an extension to BMP's Theorem 5 in [6, Chapter 2 (Part II)], which for reference is summarized here in the appendix.

First, a simple inequality is established which is used often in the theorems to follow.

Lemma 3.1 *Let ρ be a fixed real constant in the interval $(0,1)$. For every ℓ such that $0 < \ell < 1$, there exists a constant $C(\ell) < \infty$ such that*

$$\left| x \log_\rho x \right| \leq C(\ell) x^\ell, \qquad 0 < x \leq 1.$$

Proof: See the appendix.

This first main result identifies conditions to show (H4), i.e. that $h(\cdot)$ is Hölder continuous over all Θ.

Theorem 3.2 *Assume (S), (D1), (C), (H2), (H5) and (E1) with $\widehat{\ell}_2$ determined from (H5). Then, there exists a constant $C_h < \infty$ such that*

$$\|h(\theta) - h(\theta')\| \leq C_h \|\theta - \theta'\|^{\widehat{\ell}_2}, \qquad \theta, \theta' \in \Theta.$$

Proof: Fix a δ such that $\delta \leq \min\{\delta_C, \delta_H\}$ and $0 < \delta \leq 1$.

Case 1) θ and θ' are chosen in Θ so that $\|\theta - \theta'\| \geq \delta$:

$$\begin{aligned} &\|h(\theta) - h(\theta')\| \\ &= \|\pi_\theta H_\theta - \pi_{\theta'} H_{\theta'}\| \\ &\leq \|\pi_\theta H_\theta\| + \|\pi_{\theta'} H_{\theta'}\| \\ &\leq 2 \sup_{\theta \in \Theta} \|\pi_\theta H_\theta\| \end{aligned}$$

$$\begin{aligned}
&\leq 2\sup_{\theta\in\Theta}\|P_\theta^n(x,H_\theta)-\pi_\theta H_\theta\|+2\sup_{\theta\in\Theta}\|P_\theta^n(x,H_\theta)\|, \qquad x\in \mathsf{X}\\
&= 2\sup_{\theta\in\Theta}C_H\left\|P_\theta^n\left(x,C_H^{-1}H_\theta\right)-\pi_\theta\left(C_H^{-1}H_\theta\right)\right\|+2\sup_{\theta\in\Theta}C_H\left\|P_\theta^n\left(x,C_H^{-1}H_\theta\right)\right\|\\
&\leq 2C_H\sup_{\theta\in\Theta}\|P_\theta^n(x,\cdot)-\pi_\theta(\cdot)\|_{V^r}+2C_H\sup_{\theta\in\Theta}\|P_\theta^n(x,\cdot)\|_{V^r}
\end{aligned}$$

The first term is bounded using (E1) while the second term is bounded by (2.13) and Jensen's inequality under (D1) so that

$$\begin{aligned}
\|h(\theta)-h(\theta')\| &\leq 2C_HC_EV^r(x)\rho^n+2C_HC_D^rV^r(x), \qquad x\in\mathsf{X}\\
&\leq 2C_H\left(C_E+C_D^r\right)V^r(x), \qquad x\in\mathsf{X}.
\end{aligned}$$

Thus, by choosing some arbitrary x' in X and defining $K \doteq 2C_H(C_E+C_D^r)V^r(x')$ so

$$\begin{aligned}
\|h(\theta)-h(\theta')\| &\leq K\\
&\leq \frac{K}{\delta^{\widehat{\ell}_2}}\|\theta-\theta'\|^{\widehat{\ell}_2}.
\end{aligned}$$

Case 2) θ and θ' in Θ are chosen so that $\|\theta-\theta'\|<\delta\leq 1$. Under our assumptions, for any $n=1,2,\ldots$ and any $x\in\mathsf{X}$, we have

$$\begin{aligned}
&\|h(\theta)-h(\theta')\|\\
&= \|\pi_\theta H_\theta - P_\theta^nH_\theta(x)+P_\theta^nH_\theta(x)-P_{\theta'}^nH_\theta(x)\\
&\quad +P_{\theta'}^nH_\theta(x)-P_{\theta'}^nH_{\theta'}(x)+P_{\theta'}^nH_{\theta'}(x)-\pi_{\theta'}H_{\theta'}\|\\
&\leq \|\pi_\theta H_\theta-P_\theta^nH_\theta(x)\|+\|P_\theta^nH_\theta(x)-P_{\theta'}^nH_\theta(x)\|\\
&\quad +\|P_{\theta'}^nH_\theta(x)-P_{\theta'}^nH_{\theta'}(x)\|+\|P_{\theta'}^nH_{\theta'}(x)-\pi_{\theta'}H_{\theta'}\|\\
&\leq C_HC_EV^r(x)\rho^n+C_HC_CV^r(x)n\|\theta-\theta'\|^{\widehat{\ell}_3}\\
&\quad +C_5\|\theta-\theta'\|^{\widehat{\ell}_2}P_{\theta'}^nV^r(x)+C_HC_EV^r(x)\rho^n\\
&\leq V^r(x)\left\{2C_HC_E\ \rho^n+C_HC_Cn\|\theta-\theta'\|^{\widehat{\ell}_3}+C_D^rC_5\|\theta-\theta'\|^{\widehat{\ell}_2}\right\}, \qquad x\in\mathsf{X}.
\end{aligned}$$

In the second inequality above we have applied (C), (E1), and (H5) while the last inequality we have again applied (2.13) under (D1) with Jensen's inequality.

This last inequality is true for all $n=1,2,\ldots$, hence we may choose an integer $n \doteq \left\lceil \log_\rho\|\Delta\theta\|^{\widehat{\ell}_3}\right\rceil = \log_\rho\|\Delta\theta\|^{\widehat{\ell}_3}+u$ where the remainder u is such that $0\leq u<1$ and $\widehat{\ell}_3$ is form (C). If we let $\Delta\theta\doteq\theta-\theta'$, the bracketed term becomes

$$\begin{aligned}
&\left\{2C_HC_E\rho^n+C_HC_Cn\|\Delta\theta\|^{\widehat{\ell}_3}+C_D^rC_5\|\Delta\theta\|^{\widehat{\ell}_2}\right\}\\
&\leq 2C_HC_E\rho^{\log_\rho\|\Delta\theta\|^{\widehat{\ell}_3}}+C_HC_C(\log_\rho\|\Delta\theta\|^{\widehat{\ell}_3}+1)\|\Delta\theta\|^{\widehat{\ell}_3}+C_D^rC_5\|\Delta\theta\|^{\widehat{\ell}_2}\\
&\leq 2C_HC_E\|\Delta\theta\|^{\widehat{\ell}_3}+C_HC_C\|\Delta\theta\|^{\widehat{\ell}_3}\log_\rho\|\Delta\theta\|^{\widehat{\ell}_3}+C_HC_C\|\Delta\theta\|^{\widehat{\ell}_3}+C_D^rC_5\|\Delta\theta\|^{\widehat{\ell}_2}\\
&\leq (2C_HC_E+C_HC_C+C_D^rC_5)\|\Delta\theta\|^{\widehat{\ell}_2}+C_HC_C\|\Delta\theta\|^{\widehat{\ell}_3}\log_\rho\|\Delta\theta\|^{\widehat{\ell}_3}\\
&\leq \left(2C_HC_E+C_HC_C+C_D^rC_5+C_HC_CC(\widehat{\ell}_2/\widehat{\ell}_3)\right)\|\Delta\theta\|^{\widehat{\ell}_2}.
\end{aligned}$$

Here, we have used Lemma 3.1 in the last inequality with $0 < \widehat{\ell}_2/\widehat{\ell}_3 < 1$ and $C(\widehat{\ell}_2/\widehat{\ell}_3) < \infty$ a constant.

Finally, since we are free to choose any $x \in \mathsf{X}$, we choose a minimizing x in $V(x)$ for the tightest bound. Unifying the two cases, there exists a $C_h < \infty$ such that

$$\|h(\theta) - h(\theta')\| \leq C_h \|\theta - \theta'\|^{\widehat{\ell}_2}, \qquad \theta, \theta' \in \Theta.$$

■

The next theorem identifies sufficient conditions which imply (P2).

Theorem 3.3 *Assume (S), (P1), (E1), and (H2). Then for all $\theta, \theta' \in \Theta$, and $x \in \mathsf{X}$*

$$\|\nu_\theta(x)\| \leq C_\nu V^r(x), \tag{3.2}$$

$$\|P_\theta \nu_\theta(x)\| \leq C_\nu V^r(x), \tag{3.3}$$

where $C_\nu \leq C_H C_E (1-\rho)^{-1}$.

Proof: For $\nu_\theta(x) = \sum_{n=0}^{\infty} \left(\int P_\theta^n(x,dy) H_\theta(y) - h(\theta)\right)$ we have:

$$\begin{aligned}
\|\nu_\theta(x)\| &= \left\| \sum_{n=0}^{\infty} \left(\int P_\theta^n(x,dy) H_\theta(y) - h(\theta) \right) \right\| \\
&\leq \sum_{n=0}^{\infty} \left\| \int P_\theta^n(x,dy) H_\theta(y) - \pi_\theta(H_\theta) \right\| \\
&\leq C_H \sum_{n=0}^{\infty} \|P_\theta^n(x,\cdot) - \pi_\theta(\cdot)\|_{V^r} \qquad \text{by (H2) and definition of norm } \|\cdot\|_{V^r} \\
&\leq C_H C_E \sum_{n=0}^{\infty} V^r(x) \rho^n \qquad \text{from assumption (E1)} \\
&= \frac{C_H C_E}{1-\rho} V^r(x), \qquad \theta \in \Theta, \quad x \in \mathsf{X}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|P_\theta \nu_\theta(x)\| &= \left\| \sum_{n=1}^{\infty} \left(\int P_\theta^n(x,dy) H_\theta(y) - h(\theta) \right) \right\| \\
&\leq C_H C_E \sum_{n=1}^{\infty} V^r(x) \rho^n \\
&= \frac{C_H C_E \, \rho}{1-\rho} V^r(x) \\
&\leq \frac{C_H C_E}{1-\rho} V^r(x), \qquad \theta \in \Theta, \quad x \in \mathsf{X}.
\end{aligned}$$

■

The following theorem, which is probably the most significant of the three theorems here, shows that the Poisson equation solution ν_θ is also Hölder continuous under a set of assumptions which include our (C). The proof is lengthy because we have written out most of the steps in detail.

Theorem 3.4 *Assume (S), (H2), (H5), (P1), (E1), (C), and (D1) with the constants* $\widehat{\ell}_1$ *determined from (S),* $\widehat{\ell}_2$ *determined from (H5) and* $\widehat{\ell}_3$ *determined from (C). Then there exists a constant* $C_\delta < \infty$ *such that for all* $\theta, \theta' \in \Theta, \quad x \in \mathsf{X}$

$$\begin{aligned} \|\nu_\theta(x) - \nu_{\theta'}(x)\| &\leq C_\delta V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_1}, \qquad (3.4) \\ \|P_\theta \nu_\theta(x) - P_{\theta'} \nu_{\theta'}(x)\| &\leq C_\delta V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_1}. \qquad (3.5) \end{aligned}$$

Proof: Pick a δ such such $\delta \leq \min\{\delta_C, \delta_H\}$ and $0 < \delta \leq 1$. Again, we let $\Delta\theta = \theta - \theta'$ and consider the two cases of $\|\Delta\theta\| \leq \delta$ and $\|\Delta\theta\| > \delta$ separately.

We now show the Poisson equation solution satisfies (3.4). The case $\|\Delta\theta\| > \delta$ follows trivially from Theorem 3.3:

$$\begin{aligned} \|\nu_\theta(x) - \nu_{\theta'}(x)\| &\leq \|\nu_\theta(x)\| + \|\nu_{\theta'}(x)\| \\ &\leq 2\frac{C_H C_E}{1-\rho} V^r(x) \\ &\leq \frac{2}{\delta^{\widehat{\ell}_1}} \frac{C_H C_E}{1-\rho} V^r(x) \|\Delta\theta\|^{\widehat{\ell}_1}, \qquad \theta, \theta' \in \Theta, \quad \|\Delta\theta\| \geq \delta. \qquad (3.6) \end{aligned}$$

Now consider the case $\|\Delta\theta\| \leq \delta$ such that $\theta, \theta' \in \Theta$,

$$\begin{aligned} &\|\nu_\theta(x) - \nu_{\theta'}(x)\| \\ &= \left\| \sum_{n=0}^{\infty} \left(\int P_\theta^n(x, dy) H_\theta(y) - h(\theta) \right) - \sum_{n=0}^{\infty} \left(\int P_{\theta'}^n(x, dy) H_{\theta'}(y) - h(\theta') \right) \right\| \\ &\leq \sum_{n=0}^{\infty} \left\| \int P_\theta^n(x, dy) H_\theta(y) - h(\theta) - \int P_{\theta'}^n(x, dy) H_{\theta'}(y) + h(\theta') \right\| \qquad (3.7) \\ &\leq \sum_{n=0}^{N-1} \|P_\theta^n(x, H_\theta) - h(\theta) - P_{\theta'}^n(x, H_{\theta'}) + h(\theta') + P_\theta^n(x, H_{\theta'}) - P_\theta^n(x, H_{\theta'})\| \\ &+ \sum_{n=N}^{\infty} \|P_\theta^n(x, H_\theta) - \pi_\theta(H_\theta) - P_{\theta'}^n(x, H_{\theta'}) + \pi_{\theta'}(H_{\theta'})\| \qquad (3.8) \end{aligned}$$

where we have introduced some canceling terms in the last inequality. Continuing from (3.8)

$$\begin{aligned} &\|\nu_\theta(x) - \nu_{\theta'}(x)\| \\ &\leq \sum_{n=0}^{N-1} \{\|P_\theta^n(x, H_\theta) - P_\theta^n(x, H_{\theta'})\| + \|P_\theta^n(x, H_{\theta'}) - P_{\theta'}^n(x, H_{\theta'})\| + \|h(\theta) - h(\theta')\|\} \\ &+ \sum_{n=N}^{\infty} \|P_\theta^n(x, H_\theta) - \pi_\theta(H_\theta)\| + \sum_{n=N}^{\infty} \|P_{\theta'}^n(x, H_{\theta'}) - \pi_{\theta'}(H_{\theta'})\| \\ &\leq \sum_{n=0}^{N-1} \{\|P_\theta^n(x, H_\theta - H_{\theta'})\| + \|P_\theta^n(x, H_{\theta'}) - P_{\theta'}^n(x, H_{\theta'})\| + \|h(\theta) - h(\theta')\|\} \\ &+ 2 \sup_{\theta \in \Theta} \sum_{n=N}^{\infty} \|P_\theta^n(x, H_\theta) - \pi_\theta(H_\theta)\| \qquad (3.9) \end{aligned}$$

We look at several of the above terms. First, from (H5) followed by (2.13) under (D1) (and using Jensen's Inequality)

$$
\begin{aligned}
\|P_\theta^n(x, H_\theta - H_{\theta'})\| &\le C_5 \|\theta - \theta'\|^{\widehat{\ell}_2} P_\theta^n(x, V^r), \qquad n = 0, 1, \ldots, \quad x \in \mathsf{X} \\
&\le C_5 \|\theta - \theta'\|^{\widehat{\ell}_2} C_D^r V^r(x), \quad x \in \mathsf{X}.
\end{aligned}
$$

Second, from (H2) and (C),

$$
\begin{aligned}
\|P_\theta^n(x, H_{\theta'}) - P_{\theta'}^n(x, H_{\theta'})\| &\le C_H \|P_\theta^n(x, \cdot) - P_{\theta'}^n(x, \cdot)\|_{V^r} \\
&\le nC_C C_H V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_3}, \quad n = 0, 1, \ldots, \ x \in \mathsf{X}.
\end{aligned}
$$

Third, from Theorem 3.2 there exists a $C_h < \infty$ such that

$$
\|h(\theta) - h(\theta')\| \le C_h \|\theta - \theta'\|^{\widehat{\ell}_2}.
$$

Fourth, from (H2) and (E1)

$$
\|P_\theta^n(x, H_{\theta'}) - \pi_\theta(H_{\theta'})\| \le C_E C_H \rho^n V^r(x), \qquad n = 0, 1, \ldots, \quad x \in \mathsf{X}
$$

Substituting these bounds into (3.9) we find

$$
\begin{aligned}
&\|\nu_\theta(x) - \nu_{\theta'}(x)\| \\
&\le \sum_{n=0}^{N-1} \left\{ C_5 \|\Delta\theta\|^{\widehat{\ell}_2} C_D^r V^r(x) + nC_C C_H V^r(x) \|\Delta\theta\|^{\widehat{\ell}_3} + C_h \|\Delta\theta\|^{\widehat{\ell}_2} \right\} \\
&+ 2 \sup_{\theta \in \Theta} \sum_{n=N}^{\infty} C_E C_H \rho^n V^r(x) \\
&= NC_5 \|\Delta\theta\|^{\widehat{\ell}_2} C_D^r V^r(x) + \frac{N(N-1)}{2} C_C C_H \|\Delta\theta\|^{\widehat{\ell}_3} V^r(x) + NC_h \|\Delta\theta\|^{\widehat{\ell}_2} \\
&+ 2C_E C_H \frac{\rho^N}{1-\rho} V^r(x) \\
&\le V^r(x) \left\{ (C_D^r C_5 + C_h) N \|\Delta\theta\|^{\widehat{\ell}_2} + \frac{C_C C_H}{2} N(N-1) \|\Delta\theta\|^{\widehat{\ell}_3} + 2C_E C_H \frac{\rho^N}{1-\rho} \right\}
\end{aligned}
$$

since $V^r \ge 1$.

This last inequality is true for all integers $N \ge 1$ and we now set

$$
N = \left\lceil \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} \right\rceil = \left\lceil \frac{\ln \|\Delta\theta\|^{\widehat{\ell}_3}}{\ln \rho} \right\rceil = \frac{\ln \|\Delta\theta\|^{\widehat{\ell}_3}}{\ln \rho} + u \ge 1
$$

where the remainder is such that $0 \le u < 1$. Thus, the bracketed expression becomes

$$
\left\{ (C_D^r C_5 + C_h) N \|\Delta\theta\|^{\widehat{\ell}_2} + \frac{C_C C_H}{2} (N-1) N \|\Delta\theta\|^{\widehat{\ell}_3} + 2C_E C_H \frac{\rho^N}{1-\rho} \right\}
$$

$$
\begin{aligned}
&\leq \Big\{ (C_D^r C_5 + C_h)(\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + 1)\, \|\Delta\theta\|^{\widehat{\ell}_2} \\
&+ \frac{C_C C_H}{2} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} (\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + 1)\, \|\Delta\theta\|^{\widehat{\ell}_3} + 2C_E C_H \frac{\rho^{\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3}+1}}{1-\rho} \Big\} \\
&\leq (C_D^r C_5 + C_h)\, \|\Delta\theta\|^{\widehat{\ell}_2} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + (C_D^r C_5 + C_h)\, \|\Delta\theta\|^{\widehat{\ell}_2} \\
&+ \frac{C_C C_H}{2} \|\Delta\theta\|^{\widehat{\ell}_3} \left(\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3}\right) \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + \frac{C_C C_H}{2} \|\Delta\theta\|^{\widehat{\ell}_3} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} \\
&+ 2C_E C_H \frac{\|\Delta\theta\|^{\widehat{\ell}_3}}{1-\rho} \\
&\leq \left(C_D^r C_5 + C_h + \frac{C_C C_H}{2}\right) \|\Delta\theta\|^{\widehat{\ell}_2} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} \\
&+ \frac{C_C C_H}{2} \|\Delta\theta\|^{\widehat{\ell}_3} \left(\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3}\right) \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + \left(C_D^r C_5 + C_h + \frac{2C_E C_H}{1-\rho}\right) \|\Delta\theta\|^{\widehat{\ell}_2}
\end{aligned}
$$

We have $0 < \widehat{\ell}_1 < \widehat{\ell}_2 < \widehat{\ell}_3 \leq 1$ which are determined from (S) and (H5) and (C) so from Lemma 3.1,

$$
\begin{aligned}
\|\Delta\theta\|^{\widehat{\ell}_2} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} &= \frac{\widehat{\ell}_3}{\widehat{\ell}_2} \|\Delta\theta\|^{\widehat{\ell}_2} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_2} \\
&\leq \frac{\widehat{\ell}_3}{\widehat{\ell}_2} C(\widehat{\ell}_1/\widehat{\ell}_2)\, \|\Delta\theta\|^{\widehat{\ell}_1}
\end{aligned}
$$

Also from Lemma 3.1 and this last line,

$$
\begin{aligned}
\|\Delta\theta\|^{\widehat{\ell}_3} \left(\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3}\right) \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} &\leq C(\widehat{\ell}_2/\widehat{\ell}_3)\, \|\Delta\theta\|^{\widehat{\ell}_2} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} \\
&\leq \frac{\widehat{\ell}_3}{\widehat{\ell}_2} C(\widehat{\ell}_1/\widehat{\ell}_2) C(\widehat{\ell}_2/\widehat{\ell}_3)\, \|\Delta\theta\|^{\widehat{\ell}_1}
\end{aligned}
$$

Finally, since $0 < \widehat{\ell}_1 < \widehat{\ell}_2 < 1$, we clearly have

$$
\|\Delta\theta\|^{\widehat{\ell}_2} \leq \|\Delta\theta\|^{\widehat{\ell}_1}, \qquad \text{for } \|\Delta\theta\| \leq \delta \leq 1.
$$

Thus, for this case $\|\Delta\theta\| \leq \delta \leq 1$ such that $\theta, \theta' \in \Theta$, there exists a $C'_\nu < \infty$ such that

$$
\|\nu_\theta(x) - \nu_{\theta'}(x)\| = C'_\nu V^r(x)\, \|\Delta\theta\|^{\widehat{\ell}_1}.
$$

Unifying the bounds for the case $\|\Delta\theta\| \leq \delta$ with the case for $\|\Delta\theta\| > \delta$ we have

$$
\|\nu_\theta(x) - \nu_{\theta'}(x)\| \leq C_\nu V^r(x)\, \|\theta - \theta'\|^{\widehat{\ell}_1} \qquad \theta, \theta' \in \Theta, \quad x \in \mathsf{X}
$$

for a suitably large constant $C_\nu < \infty$.

Finally, (3.5) can be computed similarly or we can simply note that we have

$$\begin{aligned}
&\|P_\theta \nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\| \\
&= \left\| \sum_{n=1}^{\infty} \left(\int P_\theta^n(x,dy) H_\theta(y) - h(\theta) \right) - \sum_{n=1}^{\infty} \left(\int P_{\theta'}^n(x,dy) H_{\theta'}(y) - h(\theta') \right) \right\| \\
&\leq \sum_{n=1}^{\infty} \left\| \left(\int P_\theta^n(x,dy) H_\theta(y) - h(\theta) \right) - \left(\int P_{\theta'}^n(x,dy) H_{\theta'}(y) - h(\theta') \right) \right\| \\
&\leq \sum_{n=0}^{\infty} \left\| \left(\int P_\theta^n(x,dy) H_\theta(y) - h(\theta) \right) - \left(\int P_{\theta'}^n(x,dy) H_{\theta'}(y) - h(\theta') \right) \right\|
\end{aligned}$$

and the last line is bounded from (3.7). ■

Remark: At the end of their paper[70], Ma, Makowski, and Shwartz made a Hölder generalization to their main (Lipschitz) conditions for the finite state Markov chain case they consider. While similar in some ways, our approach here is substantially different from theirs and provides extensions to the non-finite state space case.

3.2.1 Possible Extensions

We mention two possible extensions:

Super-linear Condition (C)

From the proof of Theorems 3.2 and 3.4, it is evident that condition (C) can be further weakened to allow faster than linear growth in n, such as n^p for some fixed $p > 1$ by repeatedly applying Lemma 3.1. The revised condition is then

(C') There exist constants $p \geq 1$, $C_C < \infty$, $\delta_C > 0$ and $\widehat{\ell}_3 \in (\widehat{\ell}_2, 1]$ such that for each $n = 0, 1, \ldots$

$$\|P_\theta^n(x,\cdot) - P_{\theta'}^n(x,\cdot)\|_{V^r} \leq n^p \, C_C V^r(x) \, \|\theta - \theta'\|^{\widehat{\ell}_3},$$

for all θ, $\theta' \in \Theta$ with $\|\theta - \theta'\| \leq \delta_C$, and all $x \in \mathsf{X}$.

At present, we are not aware of any applications which might benefit from such an extension so we merely mention it as a corollary. The details of the proof are similar to Theorem 3.4.

Corollary 3.5 *Assume (S), (H2), (H5), (P1), (E1), (C'), and (D1) with the constants $\widehat{\ell}_1$ determined from (S), $\widehat{\ell}_2$ determined from (H5) and $\widehat{\ell}_3$ determined form (C'). Then, there exists a constant $C_\nu < \infty$ such that for all $\theta, \theta' \in \Theta$, $\quad x \in \mathsf{X}$*

$$\begin{aligned}
\|h(\theta) - h(\theta')\| &\leq C_h \|\theta - \theta'\|^{\widehat{\ell}_2} \\
\|\nu_\theta(x) - \nu_{\theta'}(x)\| &\leq C_\delta V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_1}, \\
\|P_\theta \nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\| &\leq C_\delta V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_1}.
\end{aligned}$$

Weakening of (H5)

We pointed out earlier that (H5) can be restrictive. For some applications, the result of Theorem 3.4 can be improved by noting that

$$
\begin{aligned}
&\|P_\theta \nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\| \\
&= \left\| \sum_{n=1}^{\infty} \left(\int P_\theta^n(x,dy) H_\theta(y) - h(\theta) \right) - \sum_{n=1}^{\infty} \left(\int P_{\theta'}^n(x,dy) H_{\theta'}(y) - h(\theta') \right) \right\| \\
&= \left\| \sum_{n=0}^{\infty} \left(\int P_\theta^n(x,dy) P_\theta(y, H_\theta) - h(\theta) \right) - \sum_{n=0}^{\infty} \left(\int P_{\theta'}^n(x,dy) P_{\theta'}(y, H_{\theta'}) - h(\theta') \right) \right\| .
\end{aligned}
$$

In order to show (3.5), we can replace condition (H5) in Theorem 3.4 with the following condition:

(H6) There exists constants $C_5 < \infty$ and $\delta_H > 0$ such that for all θ, $\theta' \in \Theta$ with $\|\theta - \theta'\| \le \delta_H$ and all $x \in \mathsf{X}$, we have

$$\|P_\theta(x, H_\theta) - P_{\theta'}(x, H_{\theta'})\| \le C_5 V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_2} .$$

With (H6) replacing (H5) we can show (3.5) by the same method as Theorem 3.4. The advantage of making this replacement is in some applications where (H5) does not hold, we may be able to show (H6) if the one-step transition kernel P_θ has a sufficient smoothing effect when integrated with H_θ. A similar observation was also noted in [6] for their framework.

3.3 Comparison to BMP's Results

In BMP's framework[6, p. 216], verifying the convergence properties of the SA involves verifying their condition (A.4) which says:

(A.4) There exists a function h on Θ, and for each $\theta \in \Theta$ a function $\nu_\theta(\cdot)$ on X such that

- (i) $h : \Theta \to \mathbb{R}^p$ is locally Lipschitz;
- (ii) $(I - P_\theta)\nu_\theta = H_\theta - h(\theta)$ for all $\theta \in \Theta$;
- (iii) for all compact subsets Q of Θ, there exist constants C_3, C_4, q_3, q_4, $\lambda \in [\frac{1}{2}, 1]$, such that for all $\theta, \theta' \in Q$

$$
\begin{aligned}
\|\nu_\theta(x)\| &\le C_3 \left(1 + \|x\|^{q_3}\right) && (3.10) \\
\|P_\theta \nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\| &\le C_4 \|\theta - \theta'\|^{\lambda} \left(1 + \|x\|^{q_4}\right) && (3.11)
\end{aligned}
$$

We now compare how one may verify (A.4) BMP's framework versus our conditions (H3), (P1)-(P3) in our framework. In BMP's framework, their Theorem 5 (summarized in the Appendix here) identifies a set of conditions which imply (A.4) holds. We make two main points comparing BMP's Theorem 5 to the last three theorems under our specialized conditions.

The first point is that BMP's local Lipschitz continuity condition on h, condition (A.4)-(i), is stronger than what is required for the Kushner-Clark Lemma. In Theorem 3.2 we weaken this by showing a collection of completely *verifiable* conditions which imply (H4) which in turn implies (H3). Condition (H4) is simply Hölder continuity

$$\|h(\theta) - h(\theta')\| \leq C_h \|\theta - \theta'\|^{\ell}, \qquad \text{for all } \theta, \theta' \in \Theta \text{ and } \|\theta - \theta'\| \leq \delta, \tag{3.12}$$

for some $0 < \ell < 1$ and $\delta > 0$. We found this seemingly minor change from BMP's local Lipschitz condition on $h(\theta)$ to the Hölder form (H4) actually paid substantial dividends as the conditions assumed in Theorem 3.2 are weaker than BMP's Theorem 5.

The second point compares the following assumption which is made by BMP in their Theorem 5 [6].

Assumption *For some compact subset $Q \subset \Theta$, there exists constants $K < \infty$ and $N_{p_1}(g) < \infty$ such that for all $n = 1, 2, \ldots$*

$$\|P_\theta^n g(x) - P_{\theta'}^n g(x)\| \leq K N_{p_1}(g) \|\theta - \theta'\| (1 + \|x\|^q), \qquad \theta, \theta' \in Q, \ x \in \mathsf{X}, \tag{3.13}$$

for any function g belonging to a particular class of functions denoted $Li(p_1)$. (See the appendix or [6] for definitions of $N_{p_1}(g)$ and $Li(p_1)$ although they are not particularly important to the discussion here.)

Theorems 3.2-3.4 show that, in our framework which include some reasonable assumptions also made by BMP, h is shown Hölder continuous while (3.13) is weakened to our condition (C), i.e. for all $n = 1, 2, \ldots$

$$\|P_\theta^n(x, \cdot) - P_{\theta'}^n(x, \cdot)\|_{V^r} \leq n \, C_C V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_3}, \qquad \text{for all } \theta, \theta' \in \Theta \text{ and } x \in \mathsf{X}.$$

If we may neglect any differences brought about by the substitution of $V^r(x)$ in our framework for $(1 + \|x\|^q)$ in BMP's, our condition (C) is in general substantially weaker than (3.13) since it allows *linear growth* with $n = 1, 2, \ldots$. We will find in practice (C) is much easier to verify than (3.13) and we will show later a simple birth-death Markov chain example where (C) is easily checked with very little work while it is unclear how one would verify (3.13).

To summarize, the main point we wish to make for this chapter is that (H5) and (C) together with (E1) form the basis of a new collection of *specialized* conditions for geometrically ergodic Markov chain state processes which offer significant advantages in terms of ease of verification over those presented in BMP's Theorem 5.

3.4 Sufficient Conditions for (E1)

In this section we summarize some ergodicity results from Chapters 15 and 16 of Meyn and Tweedie [79] which form a critical link to condition (E1). The next condition, taken verbatim from [79], states that a generic Markov chain $\mathcal{X} = \{X_n, n = 0, 1, \ldots\}$ described by a Markov transition function $P(x, A)$ for $A \in \mathcal{B}(\mathsf{X})$ undergoes a *geometric drift* towards a subset C of X. Condition (V4) is simply a fixed-θ version of (D2). Recall that $\Delta V(x) = PV(x) - V(x)$.

(V4) There exists an extended real valued function $V : \mathsf{X} \to [1, \infty]$, a measurable set C, and constants $\beta > 0$, $b < \infty$,

$$\Delta V(x) \leq -\beta V(x) + b 1_C(x), \qquad x \in \mathsf{X}.$$

We also need to define the (scalar) V-norm [79, Chapter 16] of two kernels

$$\nu_1(x, A),\ \nu_2(x, A) \qquad A \in \mathcal{B}(\mathsf{X}), x \in \mathsf{X}$$

as

$$|||\nu_1(x,\cdot) - \nu_2(x,\cdot)|||_V \doteq \sup_{x \in \mathsf{X}} \frac{\|\nu_2(x,\cdot) - \nu_2(x,\cdot)\|_V}{V(x)}. \tag{3.14}$$

Here, the norm $\|\cdot\|_V$ in the numerator of (3.14) is the version for scalar valued functions $f : \mathsf{X} \to \mathbb{R}$ defined in (1.13). As in (1.13), we can also apply (3.14) to measures such as the invariant π by simply defining the kernel $\pi(x, A) \doteq \pi(\cdot)$ for all $x \in \mathsf{X},\ A \in \mathcal{B}(\mathsf{X})$.

It follows immediately from Theorem 16.0.1 of [79, p. 383] that for any $\theta \in \Theta$ such that a generic fixed-θ Markov chain is ϕ-irreducible, aperiodic and satisfies a (V4) condition, there exists an $0 < \rho_\theta < 1$ and an $R_\theta < \infty$ such that

$$|||P_\theta^n - \pi_\theta|||_V \leq R_\theta \rho_\theta^n, \qquad n = 1, 2, \ldots. \tag{3.15}$$

In order to verify condition (E1), we seek an extension of (3.15) which holds for vector valued functions $f : \mathsf{X} \to \mathbb{R}^p$ dominated in Euclidean norm by V^r (instead of V), and which is uniform over all θ in Θ, i.e. an $R < \infty$ and a $0 < \rho < 1$ such that

$$\sup_{\theta \in \Theta} |||P_\theta^n - \pi_\theta|||_{V^r} \leq R \rho^n, \qquad n = 1, 2, \ldots.$$

The extension to (finite dimension) vector valued functions $f : \mathsf{X} \to \mathbb{R}^p$ is straightforward if we use for the numerator of (3.14) the norm defined in (3.1). Fix any $\theta \in \Theta$. For vector valued functions $f = (f^{(1)}, f^{(2)}, \ldots, f^{(p)})'$ such that $\|f\| \leq V^r$, each vector component is also dominated by V^r in absolute value, i.e. $\left|f^{(i)}\right| \leq V^r$ for $i = 1, \ldots, p$. Hence,

$$\begin{aligned} \|P_\theta^n - \pi_\theta\|_{V^r} &= \sup_{f : \|f\| \leq V^r} \|P_\theta^n(x, f) - \pi_\theta(f)\| \\ &= \sup_{f : \|f\| \leq V^r} \left(\sum_{i=1}^p \left| P_\theta^n(x, f^{(i)}) - \pi_\theta(f^{(i)}) \right|^2 \right)^{1/2} \\ &\leq \left(\sum_{i=1}^p \sup_{\left|f^{(i)}\right| \leq V^r} \left| P_\theta^n(x, f^{(i)}) - \pi_\theta(f^{(i)}) \right|^2 \right)^{1/2} \end{aligned} \tag{3.16}$$

Let us now assume that there exists an $R_\theta < \infty$ and a $0 < \rho_\theta < 1$ such that (3.15) holds for each vector component of $f : \mathsf{X} \to \mathbb{R}^p$ such that $\left|f^{(i)}\right| \leq V^r$., i.e.

$$\sup_{x \in \mathsf{X}} \frac{\sup_{\left|f^{(i)}\right| \leq V^r} \left| P_\theta^n(x, f^{(i)}) - \pi_\theta(f^{(i)}) \right|}{V^r(x)} \leq R_\theta \rho_\theta^n, \qquad i = 1, \ldots, p;\ n = 1, 2, \ldots. \tag{3.17}$$

Substituting this in (3.16),

$$\begin{aligned} \|P_\theta^n(x,\cdot) - \pi_\theta\|_{V^r} &\le \left(\sum_{i=1}^{p} R_\theta^2 \rho_\theta^{2n} V^{2r}(x)\right)^{1/2} \\ &= \sqrt{p} R_\theta \rho_\theta^n V^r(x) \end{aligned}$$

so for vector valued functions f such that $\|f\| \le V^r$ we have

$$\begin{aligned} \|\|P_\theta^n - \pi_\theta\|\|_{V^r} &= \sup_{x \in \mathsf{X}} \frac{\|P_\theta^n(x,\cdot) - \pi_\theta\|_{V^r}}{V^r(x)} \\ &\le \sqrt{p} R_\theta \rho_\theta^n. \end{aligned}$$

This shows that it is sufficient to check (E1) by checking each component function $f^{(i)}$ such that $\left|f^{(i)}\right| \le V^r$ for $i = 1, \ldots, p$.

Thus, there are two remaining issues to address and the uniformity over $\theta \in \Theta$ is addressed first, i.e. we need to find "uniform" upper bounds R, ρ for $R_\theta \le R$ and $\rho_\theta \le \rho < 1$ over all θ in Θ. Since these bounds do not need to be particularly tight, we seek a loose bound which holds under the broadest possible conditions. While bounds such as these are an active area of research [4, 80, 91, 104], we feel that [80] offers a promising approach because it allows computation of bounds on the convergence rate parameters based on the drift equation (D2) provided certain additional conditions are met. Hence, if these conditions are met, the uniformity of the convergence rate bound over Θ follows immediately from the uniformity of (D2).

We note that BMP avoid the uniformity issue by simply assuming in their Theorem 5, condition (i), that for all functions g in a class of functions $Li(p_1)$ (see appendix for definition) that

$$\|P_\theta^n g(x_1) - P_\theta^n g(x_2)\| \le K\rho^n(1 + \|x_1\|^{q_1} + \|x_2\|^{q_2}), \tag{3.18}$$

for all $\theta \in \Theta$, $x_1, x_2 \in \mathsf{X}$, and $n \ge 0$. It's clear that (3.18) is related to the geometric ergodicity of the chain, and in fact, a uniform "(E1)-type" condition follows readily from (3.18) via BMP's Lemma 1 [6, p 252]; but, the uniformity over $\theta \in \Theta$ in this "(E1)-type" condition is inherited from the assumed uniformity over $\theta \in \Theta$ in (3.18). Although we could take this approach, we feel it is less than completely satisfactory due to the possibly difficult task of finding or proving existence of a finite K in (3.18) for all $\theta \in \Theta$. We take a more direct approach which is enabled by some recent results on bounding and actually *computing* the geometric convergence rate parameters of certain Markov chains.

Next, we summarize the results from [80] to address the uniformity issue above while the the following section addresses the second remaining issue of the "smaller" dominating function V^r.

3.4.1 Computable Bounds for (E1)

This section summarizes results by Meyn and Tweedie [80] which yield a computable bound on the convergence rate parameters for certain fixed-θ Markov chains satisfying a (V4) condition.

If (V4) is strengthened to a uniform (D2) condition, then these results can be very useful for verifying (E1).

Theorem 3.6 (Meyn and Tweedie [80]) *Suppose that for some atom $\alpha \in \mathcal{B}(\mathsf{X})$ we have constants $\lambda < 1$, $b < \infty$ and a function $V \geq 1$ such that*

$$PV \leq \lambda V + b\mathbb{1}_\alpha. \tag{3.19}$$

Let $\vartheta = 1 - M_\alpha^{-1}$, where

$$M_\alpha = \frac{1}{(1-\lambda)^2}\left[1 - \lambda + b + b^2 + \zeta_\alpha(b(1-\lambda) + b^2)\right] \tag{3.20}$$

and

$$\zeta_\alpha = \sup_{|z|\leq 1}\left|\sum_{n=0}^{\infty}\left[P^n(\alpha,\alpha) - P^{n-1}(\alpha,\alpha)\right]z^n\right| \tag{3.21}$$

Then, $\mathcal{X}$ is V-uniformly ergodic, and for any $\rho > \vartheta$,

$$|||P^n - \pi|||_V \leq \frac{\rho}{\rho - \vartheta}\rho^n, \qquad n = 1, 2, \ldots \tag{3.22}$$

The value of this bound is that it is given in terms of the drift parameters and the Markov transition probabilities. As is apparent from this result, the key challenge lies in bounding the quantity ζ_α. In the special case that the chain is *strongly aperiodic*, Meyn and Tweedie also prove the following.

Theorem 3.7 (Meyn and Tweedie [80]) *Suppose that (3.19) holds for an atom $\alpha \in \mathcal{B}(\mathsf{X})$, and also that the atom is strongly aperiodic, i.e. for some $\delta > 0$,*

$$P(\alpha,\alpha) > \delta$$

Then

$$\zeta_\alpha \leq \frac{32 - 8\delta^2}{\delta^3}\left(\frac{b}{1-\lambda}\right)^2. \tag{3.23}$$

Meyn and Tweedie make very clear that this bound is not particularly tight. While other authors have shown tighter results, they usually take into account the specific structure of the chain to achieve it. A loose bound is adequate for our needs.

A similar result has also been extended to the general strongly aperiodic case where the drift inequality holds instead for a set $C \in \mathcal{B}(\mathsf{X})$.

Theorem 3.8 (Meyn and Tweedie [80]) *Suppose that $C \in \mathcal{B}(\mathsf{X})$ satisfies*

$$P(x, \cdot) \geq \delta\nu(\cdot), \qquad x \in C,$$

for some $\delta > 0$ and probability measure ν concentrated on C, and that there is a drift to C in the sense that for some $\lambda_C < 1$, some $b_C < \infty$ and a function $V \geq 1$,

$$PV \leq \lambda_C V + b_C \mathbb{1}_C, \tag{3.24}$$

where C, V also satisfy

$$V(x) \leq \upsilon_C < \infty, \qquad x \in C.$$

Then $\mathcal{X}$ is V-uniformly ergodic and

$$|||P^n - \pi|||_V \leq (1 + \gamma_C)\frac{\rho}{\rho - \vartheta}\rho^n, \qquad n = 1, 2, \ldots,$$

for any $\rho > \vartheta = 1 - M_C^{-1}$, for

$$M_C = \frac{1}{(1 - \check{\lambda})^2}\left[1 - \check{\lambda} + \check{b} + \check{b}^2 + \bar{\zeta}_C(\check{b}(1 - \check{\lambda}) + \check{b}^2)\right]$$

defined either in terms of the constants

$$\begin{aligned} \gamma_C &= \delta^{-2}[4b_C + 2\delta\lambda_C\upsilon_C], \\ \check{\lambda} &= \frac{\lambda_C + \gamma_C}{1 + \gamma_C} < 1, \\ \check{b} &= \upsilon_C + \gamma_C < \infty \end{aligned}$$

and the bound

$$\bar{\zeta}_C \leq \frac{4 - \delta^2}{\delta^5}\left(\frac{b_C}{1 - \lambda_C}\right)^2$$

or in the case where

$$\eta \doteq \inf_{x \in C} P(x, C) - \delta > 0$$

in terms of the constants

$$\begin{aligned} b_C^* &= \frac{b_C + \delta(\lambda_C\upsilon_C - \nu(V))}{1 - \delta} \\ \gamma_C &= \frac{(1 - \delta)b_C^*}{\delta\eta}, \\ b_\alpha^* &= \nu(V) - \lambda_C \\ \check{\lambda} &= \frac{\lambda_C + \gamma_C}{1 + \gamma_C} < 1, \\ \check{b} &= b_\alpha^* + \gamma_C < \infty \end{aligned}$$

and the bound

$$\bar{\zeta}_C = \frac{1 - \eta^2}{2\delta^4\eta}\left(\frac{b_C}{1 - \lambda_C}\right)^2.$$

Thus, we have immediately that for *uniformly strongly aperiodic* Markov chains which satisfy either a common drift inequality

$$P_\theta V \le \lambda V + b 1_\alpha, \qquad \theta \in \Theta$$

such that

$$\sup_{\theta\in\Theta} P_\theta(\alpha,\alpha) > \delta$$

or if

$$P_\theta V \le \lambda_C V + b_C 1_C, \qquad \theta \in \Theta$$

and

$$\sup_{\theta\in\Theta} P_\theta(x,\cdot) > \delta\nu(\cdot), \qquad x \in C$$

for some $0 < \delta < \infty$, then we have a uniform bound and (E1) follows.

While it remains an active area of research, efforts to extend these convergence rate bounds beyond the strongly aperiodic case have fallen short of a completely computable bound, although Meyn and Tweedie do present a somewhat less explicit result which contains one possibly unbounded parameter, the ζ_α term. Nevertheless, sometimes additional information on the chain can be exploited to bound this quantity ζ_α in certain circumstances. Examples of these techniques appear in [80]. The result is:

Theorem 3.9 (Meyn and Tweedie [80]) *Suppose again that (3.24) holds and that there exists an atom α such that for some $N \ge 1$ and $\delta_C > 0$,*

$$\sum_{j=1}^{N} P^j(x,\alpha) \ge \delta_C, \qquad x \in C.$$

Define the constants

$$\begin{aligned} \delta_N &= \delta_C/N^2, \\ b_k &= b_C(1+\delta_N^{-1})^k, \qquad k = 0,\cdots,N, \\ \lambda_k &= 1-(1-\lambda_C)/\prod_{i=0}^{k-1}(1+b_i/\delta_N), \qquad x = 0,\cdots,N. \end{aligned}$$

Then, there exists a function V_N with

$$V \le V_N \le V + b_N$$

such that

$$PV_N \le \lambda_N V_N + b_N 1_\alpha.$$

Thus, Theorem 3.6 holds using λ_N, b_N and with V_N in place of V, so that in terms of V we have

$$|||P^n - \pi|||_V \le [1+b_N]\frac{\rho}{\vartheta-\rho}\rho^n, \qquad n = 1,2,\ldots,$$

for $\rho > \vartheta$ where ϑ is defined in Theorem 3.6 using λ_N, b_N.

3.4.2 Condition (D2) and Exponents of V

Under (D2) with finite function V, the following result shows that we also have a similar (D2) drift condition for each of the "smaller" functions V^r for each *real* exponent $0 \le r \le 1$, i.e. we show if there exists two constants $0 < \lambda < 1$ and $b < \infty$ and a set C such that

$$P_\theta V(x) \le \lambda V(x) + b1_{\{C\}} \qquad \text{for all } \theta \text{ in } \Theta \text{ and } x \text{ in } \mathsf{X},$$

then there exists two constants $0 < \lambda_r < 1$ and $b_r < \infty$ such that

$$P_\theta V^r(x) \le \lambda_r V^r(x) + b_r 1_{\{C\}} \qquad \text{for all } \theta \text{ in } \Theta \text{ and } x \text{ in } \mathsf{X}. \tag{3.25}$$

It is important to note that C is the same set in both inequalities.

It should be clear that this allows us to verify (E1) by using (3.25) with the computable bounds of the previous section.

Theorem 3.10 *If (D2) holds for $V : \mathsf{X} \to [1, \infty)$ and some set C, then (D2) also holds for the function V^r under the same petite set C where r is any positive real in the interval $[0, 1]$.*

Proof: Suppose (D2) holds for the function $V : \mathsf{X} \to [1, \infty)$ and if we let $\lambda = 1 - \beta$, then

$$P_\theta V \le \lambda V + b1_{\{C\}}, \qquad \text{for all } \theta \in \Theta.$$

Consider any rational $q = n/d$ in the interval $[0, 1]$ so that $V^q = V^{n/d}$ for some integers $n \le d$. We have from Jensen's inequality

$$\begin{aligned}
P_\theta V^q &\le (P_\theta V)^{n/d} \\
&\le \left(\lambda V + b1_{\{C\}}\right)^{n/d} \\
&\le \lambda^{n/d} V^{n/d} + \frac{b1_{\{C\}}}{\lambda^{(d-n)/d}} \qquad \text{(claim proven below)} \\
&= \lambda^{n/d} V^q + \frac{b1_{\{C\}}}{\lambda^{(d-n)/d}}, \qquad \text{for all } \theta \in \Theta
\end{aligned} \tag{3.26}$$

where the last inequality step follows from a claim we now prove.

The inequality

$$\left(\lambda V + b1_{\{C\}}\right)^{n/d} \le \lambda^{n/d} V^{n/d} + \frac{b1_{\{C\}}}{\lambda^{(d-n)/d}}$$

is valid if and only if

$$\left(\lambda V + b1_{\{C\}}\right)^{n} \le \left(\lambda^{n/d} V^{n/d} + \frac{b1_{\{C\}}}{\lambda^{(d-n)/d}}\right)^{d}. \tag{3.27}$$

Using the binomial expansion, the left hand side of (3.27) can be rewritten

$$\left(\lambda V + b1_{\{C\}}\right)^{n} = \sum_{k=0}^{n} (\lambda V)^{n-k} \left(b1_{\{C\}}\right)^{k} \binom{n}{k} \tag{3.28}$$

and the right hand side of (3.27) can be rewritten

$$\begin{aligned}
\left(\lambda^{n/d}V^{n/d} + \frac{b1_{\{C\}}}{\lambda^{(d-n)/d}}\right)^d &= \sum_{k=0}^{d} \left(\lambda^{n/d}V^{n/d}\right)^{d-k} \left(\frac{b1_{\{C\}}}{\lambda^{(d-n)/d}}\right)^k \binom{d}{k}, \qquad n \leq d. \\
&= \sum_{k=0}^{d} \lambda^{\frac{n(d-k)}{d} - \frac{k(d-n)}{d}} V^{\frac{n(d-k)}{d}} b^k 1_{\{C\}} \binom{d}{k}, \qquad n \leq d \\
&= \sum_{k=0}^{d} \lambda^{n-k} V^{\frac{n(d-k)}{d}} b^k 1_{\{C\}} \binom{d}{k}, \qquad n \leq d. \qquad (3.29)
\end{aligned}$$

We now compare the summands on the right hand sides of (3.28) and (3.29) for each $k = 0, 1, 2, \ldots, n$. When $k = 0$, we trivially find that the summands are equal. Examining the case when $k = 1, \ldots, n$, we find:

1. The exponent satisfies $n - k \leq \frac{n(d-k)}{d}$ for all positive integers $n \leq d$ and $k = 1, \ldots, n$. Since $V \geq 1$ we have
$$V^{n-k} \leq V^{\frac{n(d-k)}{d}}.$$

2. For $n \leq d$ and all $k = 1, \ldots, n$,
$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \leq \frac{d!}{k!(d-k)!} = \binom{d}{k}.$$

These two inequalities imply the individual summands of (3.28) and (3.29) obey the inequality:

$$(\lambda V)^{n-k} b^k 1_{\{C\}} \binom{n}{k} \leq \lambda^{n-k} V^{\frac{n(d-k)}{d}} b^k 1_{\{C\}} \binom{d}{k}, \qquad k = 0, 1, \ldots, n.$$

For the case $k = n+1, \ldots, d$, since the summands on the right hand side of (3.29) are all positive, the claim is now proven and (3.26) holds, i.e. for any rational $q = n/d \in \mathcal{Q}$ we have

$$P_\theta V^q = \lambda^q V^q + \frac{b1_{\{C\}}}{\lambda^{(1-q)}}, \qquad \text{for all } \theta \in \Theta. \qquad (3.30)$$

Now let q_i be any sequence of rationals in the interval $(0, 1)$ which converge to the real number $r \in (0, 1)$, i.e. $r = \lim_{i\to\infty} q_i$. Then by the Dominated Convergence Theorem since $V^{q_i} \leq V$ for all $i = 1, 2, \ldots$ and $P_\theta V(x) < V(x) + b < \infty$ for any $x \in \mathsf{X}$ and $\theta \in \Theta$, we have the following for all $\theta \in \Theta$

$$\begin{aligned}
P_\theta(V^r) &= P_\theta(\lim_{i\to\infty} V^{q_i}) \\
&= \lim_{i\to\infty} P_\theta(V^{q_i}) \\
&\leq \lim_{i\to\infty} \left(\lambda^{q_i} V^{q_i} + \frac{b1_{\{C\}}}{\lambda^{1-q_i}}\right) \\
&= \lambda^r V^r + \frac{b}{\lambda^{1-r}} 1_{\{C\}},
\end{aligned}$$

where the inequality follows from (3.30). Finally, the case for $r = 0$ and $r = 1$ follow trivially. ■

Remark

This last result extends Theorem 15.2.9 of [79, p. 371] which covers the case $r = n/d = 1/2$ to arbitrary reals $r \in [0,1]$.

3.5 The Poisson Equation: Sufficient Conditions for (P1)

In this section, we briefly review a result which provides sufficient conditions for (P1), the *existence* of a family of solutions $\{\nu_\theta,\ \theta \in \Theta\}$ to the Poisson equation with "forcing" function $H_\theta : \mathsf{X} \to \mathbb{R}^p$:

$$\nu_\theta(x) - P_\theta \nu_\theta(x) = H_\theta(x) - \pi_\theta(H_\theta), \qquad x \in \mathsf{X},\ \theta \in \Theta. \tag{3.31}$$

Let us now consider separately each of the p component vectors in (3.31) while fixing an arbitrary $\theta \in \Theta$:

$$\nu_\theta^{(i)}(x) - P_\theta \nu_\theta^{(i)}(x) = H_\theta^{(i)}(x) - \pi_\theta(H_\theta^{(i)}), \qquad x \in \mathsf{X},\ i = 1, \ldots, p. \tag{3.32}$$

Then, following [79], let us define a new fixed-θ drift property:

(V3) For a function $f : \mathsf{X} \to [1,\infty)$, a set $C \in \mathcal{B}(\mathsf{X})$, a constant, $b < \infty$, and an extended-real valued function $V(x) : X \to [1,\infty]$

$$\Delta V(x) \le -f(x) + b1_C(x), \qquad x \in X.$$

We note that (V3) reduces to (V4) if we take the function f equal to βV.

Our condition (P1) is verified rather easily once the following result is at hand:

Theorem 3.11 (Meyn and Tweedie) *Suppose that $\{X_n, n = 0,1,\ldots\}$ is a ψ-irreducible, and that (V3) holds with V everywhere finite, $f \ge 1$, and C petite. If $\pi(V) < \infty$ then for some $R < \infty$ and any $|H| \le f$, the Poisson equation (3.32) admits a solution ν satisfying the bound $|\nu| \le R(V+1)$.*

Proof: See [79, p. 433] ■

We immediately see under (H2) and a (D2) condition involving V^r that integrability of V^r with respect to the invariant measure is the key additional condition to be verified to show (P1).

Consider the typical case involving a family of irreducible aperiodic Markov chains given by one step transitions kernels $\{P_\theta, \theta \in \Theta\}$. If condition (D2) holds with a finite V, we then have $\sup_{\theta\in\Theta} \pi_\theta(V) < \infty$ which follows from (2.15). We also have $\pi_\theta(V^r) < \infty$ for each $\theta \in \Theta$ by Jensen's inequality as well as a (D2) condition holding for the function V^r by Theorem 3.10. Thus, for forcing functions satisfying $\left|H_\theta^{(i)}\right| \le V^r$ for each $i = 1, \ldots, p$, or a quasi (H2) condition, we have immediately from Theorem 3.11 that for each $\theta \in \Theta$ the Poisson equation (3.32) admits a solution ν_θ which satisfies the bound

$$\left|\nu_\theta^{(i)}\right| \le R_\theta(V^r + 1), \qquad i = 1, \ldots, p \tag{3.33}$$

for some $R_\theta < \infty$. Clearly, the constant C_H in (H2) will proportionately scale this bound and (3.33) thus complements our Theorem 3.3. (Note that we do not have a uniform bound for $\sup_{\theta \in \theta} R_\theta < \infty$ from these results.)

Alternatively, the key condition $\sup_{\theta \in \Theta} \pi_\theta(V^r) < \infty$ also follows readily from (D0) and (E1) since for any n

$$\begin{aligned} \sup_{\theta \in \Theta} \pi_\theta(V^r) &\leq \sup_{\theta \in \Theta} \|P_\theta^n(x, \cdot) - \pi_\theta(\cdot)\|_V + \sup_{\theta \in \Theta} P_\theta^n(x, V), \qquad x \in \mathsf{X} \\ &\leq C_E \rho^n V(x) + C_D V(x), \qquad x \in \mathsf{X} \\ &< \infty. \end{aligned}$$

Uniqueness of the Poisson equation solution up to a constant for each $\theta \in \Theta$ also follows from Proposition 17.4.1 in [79] if we assume the Markov chain is positive Harris for all $\theta \in \Theta$. Meyn and Tweedie's Proposition 17.4.1 in [79] states that, for some constant c, and any two solutions $\widehat{\nu}_1$ and $\widehat{\nu}_2$ such that $\pi(|\widehat{\nu}_1| + |\widehat{\nu}_2|) < \infty$, then $\widehat{\nu}_1 = \widehat{\nu}_2 + c$ for a.e. $x \in \mathsf{X}$ with respect to π. Thus we see that under these circumstances, that (P1) holds.

For another approach, Makowski and Shwartz [73] also present verifiable conditions for existence and uniqueness (up to a constant) of the Poisson equation solution for *countable* state space Markov chains. Additionally, their continuity results over Θ for the solution ν_θ allow direct verification of the uniform bound condition (P2) for compact Θ.

3.6 Sufficient Conditions for (C)

Here, we consider the following key assumption related to the θ-dependence of the transition probabilities:

(M) There exists a $\delta_M > 0$, $C_P < \infty$, and an $\widehat{\ell}_3 \in (\widehat{\ell}_2, 1]$ such that,

$$|P_\theta(x, A) - P_{\theta'}(x, A)| \leq C_P P_\theta(x, A) \|\theta - \theta'\|^{\widehat{\ell}_3}, \qquad \text{for each } x \in \mathsf{X},\ A \in \mathcal{B}(\mathsf{X})$$

for all θ, θ' in Θ such that $\|\theta - \theta'\| \leq \delta_M$.

It's clear that condition (M) disallows the possibility the probability of any transition of the Markov chain which is positive can go to zero as θ ranges over the set Θ. We also note that an assumption similar to (M) was used in [73] to prove a Lipschitz condition on the θ-parameterized solution to Poisson's Equation.

The next result identifies sufficient conditions for (C).

Theorem 3.12 *Under (M),(D1), we have for all θ, θ' in Θ such that $\|\theta - \theta'\| \leq \delta_M$, x in X,*

$$\|P_\theta^n(x, \cdot) - P_{\theta'}^n(x, \cdot)\|_{V^r} \leq n 2 C_P C_D^{2r} V^r(x) \|\theta - \theta'\|^{\widehat{\ell}_3}, \qquad \textit{for all } n = 1, 2, \ldots.$$

Proof: For all θ, θ' in Θ with $\|\theta-\theta'\| \le \delta_M$, x in X and all $n = 1, 2, \ldots$,

$$\begin{aligned}
&\|P_\theta^n(x,\cdot) - P_{\theta'}^n(x,\cdot)\|_{V^r} \\
&\quad = \sup_{\|f\|\le V^r} \|P_\theta^n(x,f) - P_{\theta'}^n(x,f)\| \\
&\quad = \sup_{\|f\|\le V^r} \left\| P_\theta^n(x,f) - \sum_{i=1}^{n-1} P_\theta^{n-i} P_{\theta'}^i(x,f) + \sum_{i=1}^{n-1} P_\theta^{n-i} P_{\theta'}^i(x,f) - P_{\theta'}^n(x,f) \right\| \\
&\quad \le \sum_{i=1}^{n} \sup_{\|f\|\le V^r} \left\| P_\theta^{n-i+1} P_{\theta'}^{i-1}(x,f) - P_\theta^{n-i} P_{\theta'}^i(x,f) \right\| \\
&\quad \le \sum_{i=1}^{n} \sup_{\|f\|\le V^r} \left\| P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right) P_{\theta'}^{i-1}(x,f) \right\|
\end{aligned}$$

where $P_\theta - P_{\theta'}$ is a signed measure. Then, the summands can be bounded by first using a Hahn-Jordan decomposition followed by application of our conditions:

$$\begin{aligned}
&\sup_{\|f\|\le V^r} \left\| P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right) P_{\theta'}^{i-1}(x,f) \right\| \\
&\quad = \sup_{\|f\|\le V^r} \left\| P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right) (x, P_{\theta'}^{i-1} f) \right\| \\
&\quad \le \sup_{\|f\|\le V^r} \left\| P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right)^+ (x, P_{\theta'}^{i-1} f) - P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right)^- (x, P_{\theta'}^{i-1} f) \right\| \\
&\quad \le \sup_{\|f\|\le V^r} \left\| P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right)^+ (x, P_{\theta'}^{i-1} f) \right\| + \sup_{\|f\|\le V^r} \left\| P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right)^- (x, P_{\theta'}^{i-1} f) \right\| \\
&\quad \le \sup_{\|f\|\le V^r} \left\| C_P \|\theta-\theta'\|^{\widehat{\ell}_3} P_\theta^{n-i+1}(x, P_{\theta'}^{i-1} f) \right\| \\
&\qquad + \sup_{\|f\|\le V^r} \left\| C_P \|\theta-\theta'\|^{\widehat{\ell}_3} P_\theta^{n-i+1}(x, P_{\theta'}^{i-1} f) \right\| \\
&\quad \le 2C_P \|\theta-\theta'\|^{\widehat{\ell}_3} P_\theta^{n-i+1}(x, P_{\theta'}^{i-1} V^r) \\
&\quad \le 2C_P \|\theta-\theta'\|^{\widehat{\ell}_3} P_\theta^{n-i+1}(x, C_D^r V^r) \\
&\quad \le 2C_P \|\theta-\theta'\|^{\widehat{\ell}_3} C_D^{2r} V^r(x).
\end{aligned}$$

The last two lines each follow from (D1) and (2.13).

Hence, for all $n = 0, 1, 2, \ldots$,

$$\begin{aligned}
&\|P_\theta^n(x,\cdot) - P_{\theta'}^n(x,\cdot)\|_{V^r} \\
&\quad \le \sum_{i=1}^{n} \sup_{\|f\|\le V^r} \left\| P_\theta^{n-i} \left(P_\theta - P_{\theta'}\right) P_{\theta'}^{i-1}(x,f) \right\| \\
&\quad \le n 2 C_P C_D^{2r} \|\theta-\theta'\|^{\widehat{\ell}_3} V(x) \qquad x \in \mathsf{X}, \quad \theta, \theta' \in \Theta, \quad \|\theta-\theta'\| \le \delta_M,
\end{aligned}$$

which satisfies assumption (C). ∎

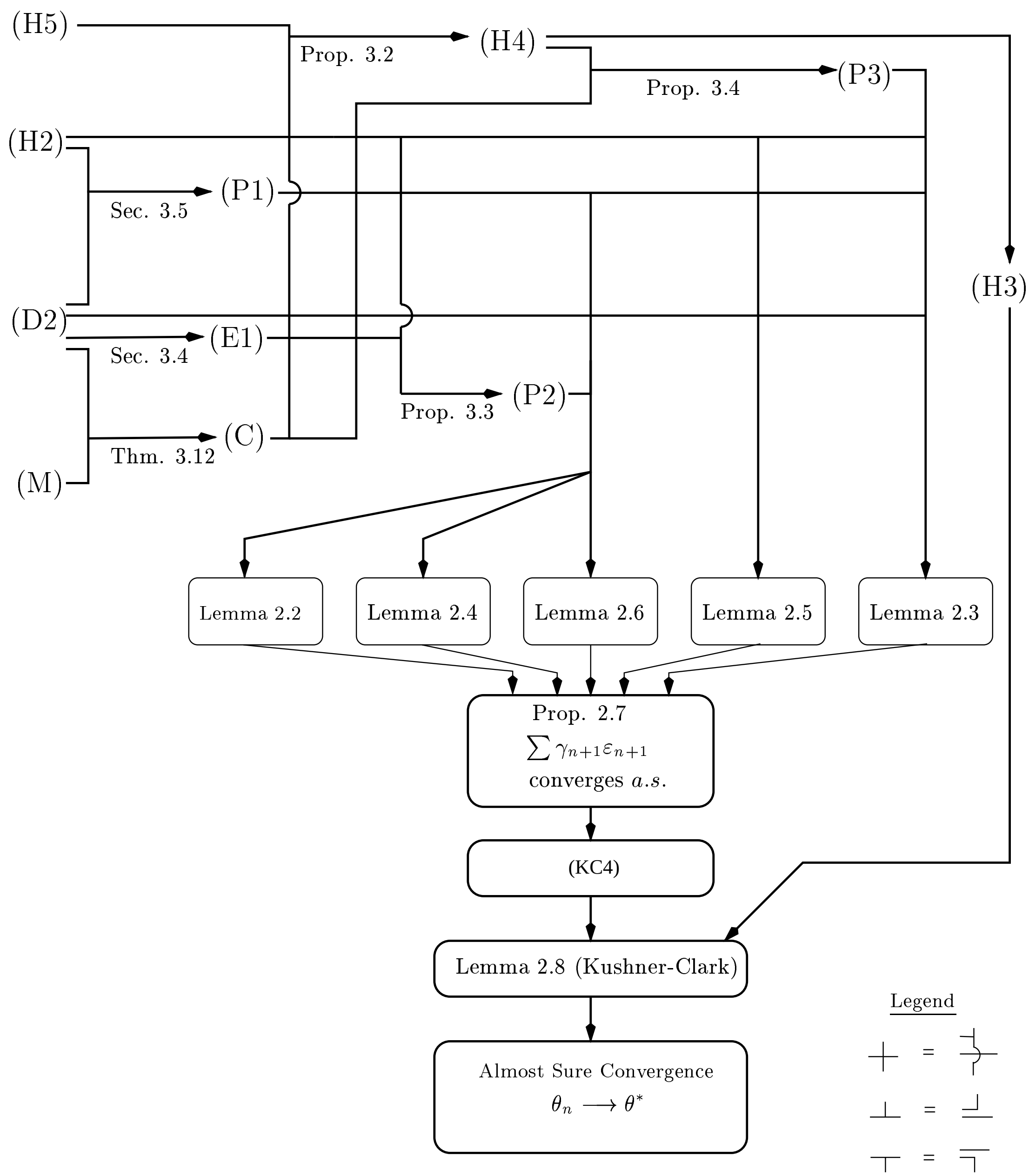

Figure 3.1: Relationship of various conditions and results.

3.7 Summary: Relationship among Conditions

The above Figure 3.1 indicates the relationship of many of the conditions we've presented within our framework which lead to the almost sure convergence of SA's. The arrows in this figure merely indicate that some relationship exists among the conditions and they do not necessarily represent an unconditional implication. For details on each relationship, the reader should refer to the particular result indicated under the arrow. Assumed throughout is either the constrained or unconstrained SA with condition (S) on the step-size sequence $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$.

It's clear from this diagram that the conditions (H2), (H5), (D2) and (M) to the upper left of the diagram are sufficient conditions leading to convergence of SA's under the setting of these last two chapters. Also, bear in mind the relationships between (D1) and (D2) which are not represented here.

3.8 Design Issue: Selection of the Exponent r

Suppose we assume (D1) with V unbounded off petite sets so (D2) holds as well. Given, any $r \leq 1$, we then showed that (D2) also holds for the smaller functions V^r, and subject to the constraint $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$, we find that the selection of r affects the analysis only in subtle ways.

First, under condition (E1) we have a test function V^r for which there exists a constant $C_E < \infty$ and $0 < \rho < 1$ such that

$$\sup_{\theta \in \Theta} \| P_\theta^n(x, \cdot) - \pi_\theta(\cdot) \|_{V^r} \leq C_E V^r(x) \rho^n, \qquad x \in \mathsf{X}, \quad n = 0, 1, 2, \ldots \tag{3.34}$$

The choice of r such that $0 < r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ clearly affects the bound in (E1) since V^r appears on both sides of (3.34). Nevertheless, choosing r smaller or larger in this range does not appear to have much effect on the overall convergence properties for the SA as seen in our analysis.

Second, we've shown in Section 3.5 that under (D1) and (E1),

$$\sup_{\theta \in \Theta} \pi_\theta V^r \leq C_L < \infty.$$

and this holds independently of $r \leq 1$.

Third, choosing a larger r within the constraint $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ permits use of a larger class of observation functions $H(\theta, x)$ as seen in (H2) and (H5). Thus, it's clear that for the greatest generality, r should be chosen so that $r = \frac{1}{2(1+\widehat{\ell}_1)}$. Thus, if we take $r = \frac{1}{2(1+\widehat{\ell}_1)}$ then there is a tradeoff in the choice of r versus the choice of $\widehat{\ell}_1$; with the r mainly affecting the space of permissible observable functions H satisfying (H2) and (H5) while $\widehat{\ell}_1$ is also affecting (H5) and the class of deterministic step-size sequences $\{\gamma_{n+1}, n = 0, 1, \ldots\}$ meeting the condition (S).

3.9 Bernoulli Random Walk with θ-Dependent Transitions

In the following example we demonstrate the use of our specialized conditions for an SA algorithm which attempts to regulate the mean number in an unbounded random walk with a single reflection at the origin. This discrete time Markov chain may be used to model the M/M/1 queue with adjustable service rate. We verify the conditions (D1)-(D2) and (M) on the Markov chain transition function and verify conditions (H2) and (H5) on the observation function. These conditions imply the remaining conditions in our general convergence framework.

3.9.1 The Model

The random walk $\{X_n, n = 0, 1, \ldots\}$ on the countable space $\mathbb{Z}^+$ is governed by a family of one step transition probabilities $\{P_\theta,\ \theta \in \Theta\}$ where we shall conveniently let $\Theta = [1/2 + \epsilon_1, 1 - \epsilon_2]$ for some small $\epsilon_1, \epsilon_2 > 0$ so that $1/2 < 1/2 + \epsilon_1 < 1 - \epsilon_2 < 1$. The transition probabilities are defined for each $\theta \in \Theta$ by

$$\begin{aligned} P_\theta(x, x+1) &= 1 - \theta, \qquad x \geq 0 \\ P_\theta(x, x-1) &= \theta, \qquad x > 0 \\ P_\theta(0, 0) &= \theta. \end{aligned}$$

and zero otherwise.

In this case, the steady state probabilities are analytically known and given by

$$\begin{aligned} \pi_\theta(n) &= (1 - \rho_\theta)\rho_\theta^n, \qquad \theta \in \Theta, \qquad n = 0, 1, \ldots \\ \rho_\theta &\doteq \frac{1 - \theta}{\theta}. \end{aligned}$$

The expected value of the random walk in steady state serves as our objective function:

$$\begin{aligned} J(\theta) &= \mathbf{E}_{\pi_\theta}[X] = \sum_{n=0}^{\infty} n\pi_\theta(n) \\ &= \frac{\rho_\theta}{1 - \rho_\theta}, \qquad \theta \in \Theta. \end{aligned} \tag{3.35}$$

3.9.2 The SA Algorithm

Our goal is for the SA algorithm to locate the value $\theta^\star$ such that the mean number in the system is some particular value, say L and we shall apply the projected SA with driving function

$$H(\theta, x) \doteq x - L$$

so that we find the zero of $h(\theta) = J(\theta) - L$.

Observations of the number in the system are made at each transition of the Markov chain and the next parameter iterate is computed from a projected SA which immediately updates the transition probabilities for the next state transition. This very simple recursion is given by

$$\begin{aligned} \theta_{n+1} &= \Pi_{[1/2+\epsilon_1, 1-\epsilon_2]}\left\{\theta_n + \gamma_{n+1}(X_{n+1} - L)\right\}, \qquad n = 0, 1, \ldots \\ \theta_0 &= \theta \qquad\qquad \text{(arbitrary in } \Theta) \end{aligned}$$

starting from any initial state value $X_0 = x \in \mathsf{X}$

Next, we verify in turn the specialized conditions which imply through this and last chapters results and the Kushner-Clark Lemma that $\theta_n \to \theta^\star$ $\mathbf{P}_{x,\theta}$- almost surely where $J(\theta^\star) = L$.

3.9.3 Verification of (D1) and (D2)

Let us try verifying (D1) for the test function $V(x) = Ks^x$ where s and K are some yet to be determined scalar parameters such that $s > 1$ and $1 \leq K < \infty$. We seek an s and K such that the uniform drift (D1) (or (D2)) holds over Θ. Recall:

(D1) $$P_\theta V(x) \leq \lambda V(x) + L, \qquad \text{for all } \theta \in \Theta,\ x \in \mathsf{X}.$$

For $x > 0$,

$$\begin{aligned} P_\theta V(x) &= \theta K s^{x-1} + (1-\theta) K s^{x+1} \\ &= \{\theta s^{-1} + (1-\theta)s\} K s^x \\ &= \{\theta s^{-1} + (1-\theta)s\} V(x), \qquad s > 1, \theta \in \Theta. \end{aligned} \tag{3.36}$$

Let us define the bracketed term as the function $\lambda(s,\theta) \doteq \theta s^{-1} + (1-\theta)s$ which is defined on $s \geq 1$ and $1/2 < \theta < 1$, and sketched in Figure 3.2. Setting the partial derivative given by

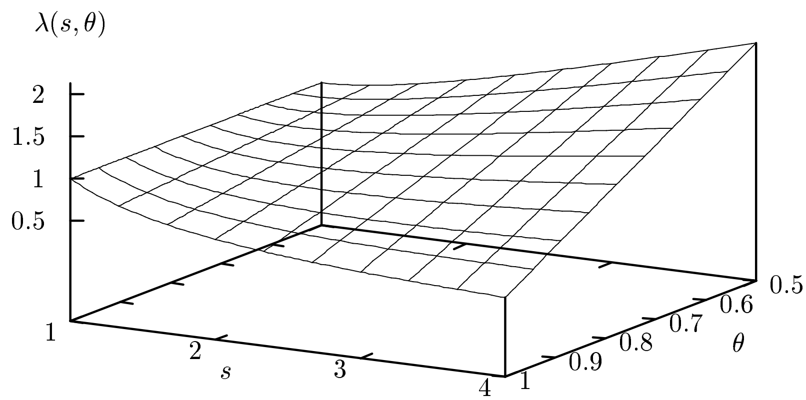

Figure 3.2: Surface plot of the function $\lambda(s,\theta)$.

$$\frac{\partial \lambda(s,\theta)}{\partial \theta} = -\theta s^{-2} + 1 - \theta, \qquad \theta \in \Theta, s \geq 1,$$

equal to zero, we find that for each $\theta \in \Theta$, the minimum of $\lambda(\cdot,\theta)$ is achieved when

$$s = \widehat{s}(\theta) \doteq \sqrt{\frac{\theta}{1-\theta}}, \qquad \theta \in \Theta.$$

Let us now review some facts about the function $\lambda(\cdot,\cdot)$:

1. Substituting in $s = \widehat{s}(\theta)$ from above:

$$\begin{aligned}\lambda(\widehat{s}(\theta), \theta) &= \theta\sqrt{\frac{1-\theta}{\theta}} + (1-\theta)\sqrt{\frac{\theta}{1-\theta}} \\ &= 2\sqrt{(1-\theta)\theta}, \qquad 1/2 < \theta < 1.\end{aligned}$$

2. The function $\lambda(\widehat{s}(\theta), \theta)$ is monotone decreasing on Θ and $\lim_{\theta \searrow 1/2} \lambda(\widehat{s}(\theta), \theta) = 1$.

3. We have $\lambda(\widehat{s}(1/2+\epsilon_1), 1/2+\epsilon_1) < 1$ for small $\epsilon_1 > 0$.

4. We have:

$$\begin{aligned}\lambda(\widehat{s}(1/2+\epsilon_1), \theta) &= \theta\sqrt{\frac{1/2-\epsilon_1}{1/2+\epsilon_1}} + (1-\theta)\sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}} \\ &\leq \lambda(\widehat{s}(1/2+\epsilon_1), 1/2+\epsilon_1), \qquad \text{for all } \theta \in [1/2+\epsilon_1, 1-\epsilon_2].\end{aligned}$$

Therefore, for the case $x > 0$ we have (3.36), and let us now fix $s = \widehat{s}(1/2+\epsilon_1) = \sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}} > 1$ which yield the easy bounds:

$$\begin{aligned}P_\theta V(x) &= \left\{\theta\sqrt{\frac{1/2-\epsilon_1}{1/2+\epsilon_1}} + (1-\theta)\sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}}\right\} V(x) \\ &\leq \left\{(1/2+\epsilon_1)\sqrt{\frac{1/2-\epsilon_1}{1/2+\epsilon_1}} + (1/2-\epsilon_1)\sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}}\right\} V(x) \\ &\leq 2\sqrt{(1/2-\epsilon_1)(1/2+\epsilon_1)}V(x), \qquad \text{for all } \theta \in [1/2+\epsilon_1, 1-\epsilon_2].\end{aligned}$$

So for the case $x > 0$, (D1) is satisfied with

$$\lambda = 2\sqrt{(1/2-\epsilon_1)(1/2+\epsilon_1)} < 1. \tag{3.37}$$

Now consider the case $x = 0$: We have

$$\begin{aligned}P_\theta V(0) &= \theta s^0 + (1-\theta)s \\ &= \theta V(0) + (1-\theta)s, \qquad s > 1.\end{aligned}$$

Again, we fix $s = \sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}} > 1$ so that

$$\begin{aligned}P_\theta V(0) &= \theta V(0) + (1-\theta)\sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}} \\ &\leq \lambda V(0) + \theta K + (1-\theta)\sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}} \\ &\leq \lambda V(0) + \max\left\{K, \sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}}\right\}\end{aligned}$$

Thus, letting $L = \max\left\{K, \sqrt{\frac{1/2+\epsilon_1}{1/2-\epsilon_1}}\right\}$, we have for verified (D1) with λ as given in (3.37), K arbitrary in $1 \le K < \infty$ and

$$V(x) = Ks^x = K\left(\frac{1/2+\epsilon_1}{1/2-\epsilon_1}\right)^{x/2}, \qquad x \in \mathsf{X}.$$

Furthermore, it's easy to check that we have also shown (D2) with petite set $C = \{x = 0\}$ and the parameters

$$\begin{aligned} \beta &= 1 - \lambda \\ b &= L. \end{aligned}$$

3.9.4 Verification of (M)

For this random walk example, (M) is very simple to check since there are only three classes of transitions to check; jump up, jump down, and the null transition $0 \to 0$. For the first class, verification reduces to checking for all $x \ge 0$:

$$\begin{aligned} |P_\theta(x, x+1) - P_{\theta'}(x, x+1)| &= |\theta - \theta'| \\ &\le \frac{P_\theta(x, x+1)}{\epsilon_2}|\theta - \theta'|, \qquad \text{for all } \theta,\ \theta' \in \Theta, \end{aligned}$$

where $\Theta = [1/2 + \epsilon_1, 1 - \epsilon_2]$. We are assuming $\epsilon_1, \epsilon_2 > 0$ and this demonstrates that this condition cannot be weakened.

Verification of the other classes of transitions,

$$|P_\theta(x, x-1) - P_{\theta'}(x, x-1)| \quad \text{and} \quad |P_\theta(0,0) - P_{\theta'}(0,0)|,$$

follows similarly. Thus, (M) is verified with $\widehat{\ell}_3 = 1$.

3.9.5 Verification of (E1)

Verification of (E1) involves finding computable bounds for the geometric convergence rate of the Markov chain. The structure of the chain suggests taking the recurrent atom α to be the 0 state. The easiest approach is simply to appeal to Theorem 3.7 since this chain is strongly aperiodic for all $\theta \in \Theta$ and satisfies (D2), i.e.

$$P_\theta V \le \lambda V + L 1_{\{x=0\}}, \qquad \text{for all } \theta \in \Theta,$$

with

$$\begin{aligned} V(x) &= Ks^x = K\left(\frac{1/2+\epsilon_1}{1/2-\epsilon_1}\right)^{x/2}, \qquad x \in \mathsf{X},\ \ 1 \le K < \infty \\ \lambda &= 2\sqrt{(1/2-\epsilon_1)(1/2+\epsilon_1)} \\ L &= K(1/2+\epsilon_1) + \sqrt{(1/2-\epsilon_1)(1/2+\epsilon_1)}. \end{aligned}$$

Then, Theorem 3.10 implies (D2) holds with V^r for the choice $r = \frac{1}{2(1+\widehat{\ell}_1)}$ so there exists a $\lambda_r < 1$ and a $L_r < \infty$ such that

$$P_\theta V^r \le \lambda_r V^r + L_r 1_{\{x=0\}}, \qquad \text{for all } \theta \in \Theta,$$

Thus, the bound (3.23) becomes

$$\zeta_\alpha \;\le\; \sup_{\theta\in\Theta} \frac{32-8\delta^2}{\delta^3}\left(\frac{L_r}{1-\lambda_r}\right)^2 < \infty$$

which holds for this example with

$$P_\theta(\alpha,\alpha) > \delta = 1/2 + \epsilon_1, \qquad\qquad \theta \in \Theta.$$

Hence, (E1) follows from Theorem 3.6 and Theorem 3.7.

An Alternative Method

While the above is adequate and perhaps the most straightforward method to show (E1), an alternative approach which may yield a tighter bound on ζ_α is to work directly with (3.21), i.e.

$$\zeta_\alpha \doteq \sup_{|z|\le 1}\left|\sum_{n=0}^{\infty}\left[P^n(\alpha,\alpha) - P^{n-1}(\alpha,\alpha)\right] z^n\right|. \tag{3.38}$$

Following [79], for an atom α and some probability P_α define the *renewal function* $u(n) \doteq P_\alpha(X_n = \alpha)$ and the *renewal variation*

$$Var(u) = \sum_{n=0}^{\infty} |u(n) - u(n-1)|\,. \tag{3.39}$$

Again, taking the atom α to be the state $x = 0$, clearly for any $\theta \in \Theta$,

$$\begin{aligned}
\zeta_0(\theta) &= \sup_{|z|\le 1}\left|\sum_{n=0}^{\infty}\left[P_\theta^n(0,0) - P_\theta^{n-1}(0,0)\right] z^n\right| \\
&\le \sum_{n=0}^{\infty}\left|P_\theta^n(0,0) - P_\theta^{n-1}(0,0)\right| \\
&\doteq Var_\theta(u),
\end{aligned}$$

where $\zeta_0(\theta)$ and $Var_\theta(u)$ are a θ-parameterized extensions of (3.38) and (3.39), respectively. From [80], for any $\theta \in \Theta$ we have the following bound on $Var_\theta(u)$:

$$\begin{aligned}
Var_\theta(u) &\le 1/2\left\{\frac{1}{(\theta-(1-\theta))^2(1-\theta)} - 1\right\} \\
&= 1/2\left\{\frac{1}{(2\theta-1)^2(1-\theta)} - 1\right\}, \qquad \theta \in \Theta.
\end{aligned}$$

Hence,

$$\zeta_0(\theta) \le \sup_{\theta\in[1/2+\epsilon_1,1-\epsilon_2]} Var_\theta(u) < \infty, \qquad \theta \in \Theta = [1/2+\epsilon_1, 1-\epsilon_2].$$

3.9.6 Verification of (H2) and (H5)

Above we have shown that

$$V(x) = Ks^x = K\left(\frac{1/2+\epsilon_1}{1/2-\epsilon_1}\right)^{x/2}, \qquad x \in \mathsf{X}.$$

for some arbitrary $1 \leq K < \infty$ is a solution to (D1) and (D2).

Verifying (H2) and (H5) is a simple matter of first selecting an exponent $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ where $0 < \widehat{\ell}_1 < 1$ satisfies (S) and there is no loss in generality in taking $r = \frac{1}{2(1+\widehat{\ell}_1)}$. Then any functions $H(\theta,x) : \Theta \times \mathsf{X} \to \mathbb{R}^p$ and $\rho(\theta,x) : \Theta \times \mathsf{X} \to \mathbb{R}^p$ satisfy (H2) if for some finite constants C_H, C_ρ they satisfy for all $x \in \mathsf{X}$

$$\sup_{\theta\in\Theta} \|H(\theta,x)\| \leq C_H V^r(x),$$
$$\sup_{\theta\in\Theta} \|\rho(\theta,x)\| \leq C_\rho V^r(x),$$

and satisfy (H5) if for some constants $C_5 < \infty$ and $\widehat{\ell}_2 \in (\widehat{\ell}_1, 1)$

$$\|H_\theta(x) - H_{\theta'}(x)\| \;\leq\; C_5 V^r(x)\,\|\theta-\theta'\|^{\widehat{\ell}_2},$$

for all $\theta, \theta' \in \Theta$ such that $\|\theta - \theta'\| < \delta_H$. It is immediately clear that the functions

$$H(\theta,x) = x - L, \qquad x \in \mathbb{Z}^+, \theta \in \Theta$$
$$\rho(\theta,x) = 0, \qquad x \in \mathbb{Z}^+, \theta \in \Theta$$

meet these conditions for any positive $r = \frac{1}{2(1+\widehat{\ell}_1)}$ with some suitably large constant $C_H < \infty$.

3.10 GSMP's and Continuous Time Markov Chains

The conditions we developed in this dissertation are for discrete time Markov chains but this framework can be extended to continuous time chains and Generalized Semi-Markov Processes (GSMP's) if discrete time conversion techniques [33, 56] are used. We note that a similar approach was also taken by [22].

Chapter 4

A Steady State Gradient Estimate for Markov Chains

This chapter studies gradient estimation for steady state performance of stochastic systems modeled as Markov chains which have a dependence on a parameter θ. A particular gradient estimate is considered for use with the stochastic approximation procedures.

4.1 Introduction

Suppose steady-state performance is given by

$$J(\theta) = \pi_\theta(f_\theta), \qquad \theta \in \Theta$$

where $f_\theta(x) : \Theta \times \mathsf{X} \to \mathbb{R}$ is a given performance function and π_θ is the invariant distribution at parameter θ for an irreducible positive recurrent Markov chain. Both f_θ and π_θ may depend on θ in Θ. The main goal in stochastic optimization is to minimize (or maximize) the objective function $J(\theta)$ over parameters $\theta \in \Theta$. We assume this optimizer $\theta^\star$ can be found by locating the fixed point $\theta^\star \in \Theta$ such that

$$\nabla J(\theta^\star) = 0.$$

Throughout this chapter and the next, we assume a compact parameter set Θ which is a subset of $\mathbb{R}^p$ and a Markov chain $\mathcal{X} = \{X_n \in \mathsf{X}, n = 0, 1, \ldots\}$ taking values on a now assumed *countable* state space X. Here $\mathcal{X}$ is governed by a family of one-step transition probabilities $\{P_\theta, \theta \in \Theta\}$ and is irreducible and positive recurrent for each parameter $\theta \in \Theta$. Let the one step transition probability matrix be given by $P_\theta = [p_{x,y}(\theta)]_{x,y}$. The performance function $f_\theta(x) = f(\theta, x) : \Theta \times \mathsf{X} \to \mathbb{R}$ is assumed differentiable with respect to θ for each $x \in \mathsf{X}$.

4.2 Gradient Estimation and Stochastic Approximations

Consider how one might estimate both $J(\theta)$ and $\nabla J(\theta)$ and, for simplicity, let us temporarily assume θ is scalar valued. If we are able to solve for π_θ directly from the stationary equation $\pi_\theta = \pi_\theta P_\theta$ and the normalizing equation $\pi_\theta e = 1$, then we can simply calculate the steady state performance as

$$J(\theta) = \sum_{x \in \mathsf{X}} \pi_\theta(x) f(\theta, x).$$

If, in addition, we are able to interchange the limit and expectation, the performance derivative may be given by

$$\frac{J(\theta)}{d\theta} = \sum_{x\in\mathsf{X}} \pi_\theta(x)\frac{df(\theta,x)}{d\theta} + \sum_{x\in\mathsf{X}} \frac{d\pi_\theta(x)}{d\theta} f(\theta,x).$$

Unfortunately, for many systems of interest, efficiently solving for $\pi(\theta)$, yet alone $\frac{d\pi_\theta(x)}{d\theta}$, is simply not possible, especially if the state space X is large or countably infinite.

Because of this difficulty, we are motivated to compute $J(\theta)$ via a long run sample average. By the Strong Law of Large Numbers (SLLN) for Markov chains [79, p. 411, Thm. 17.0.1] it follows that for each fixed $\theta \in \Theta$ and initial state $X_0 = x$

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\theta, X_k) &= \pi_\theta(f_\theta) \quad \mathbf{P}_{\theta,x}\text{-a.s.}, \\ &\doteq J(\theta), \end{aligned}$$

provided that $\{X_n, n = 0, 1, \ldots\}$ is a positive Harris recurrent chain and $\pi_\theta(|f_\theta|) < \infty$. Chains on a countable state space are always Harris recurrent if they are recurrent, [79, p. 201], thus this condition reduces to checking for or assuming positive recurrence. Both the positive recurrence and the integrability condition $\pi_\theta(|f_\theta|) < \infty$ seem completely reasonable and are satisfied uniformly over Θ in many interesting applications.

Suppose now the derivative (or gradient) can be computed similarly through a long run sample average so that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} g(\theta, X_n) = \frac{dJ(\theta)}{d\theta}, \qquad \mathbf{P}_{\theta,x} - a.s.$$

for some function g we have not defined yet. Then, we have a gradient estimation algorithm well suited for stochastic approximation since the estimate is taken from a single sample path and is related to the steady-state mean:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} g(\theta, X_n) = \pi_\theta(g(\theta,\cdot)), \qquad \mathbf{P}_{\theta,x}\text{-a.s.}$$

If such an estimate can be identified, we have an viable approach to stochastic optimization via SA with convergence following from the framework for the zero-finding problem of the previous chapters. Furthermore, we have an obvious function to drive the SA algorithm, i.e.

$$\theta_{n+1} = \Pi_\Theta\left\{\theta_n - \gamma_{n+1} g(\theta_n, X_{n+1})\right\}, \qquad n = 0, 1, \ldots \tag{4.1}$$

This algorithm is particularly useful if the function g should meet the convergence criteria for SA's which is truly verifiable in terms of the known model data. The key here is finding such functions $g(\theta, x)$ where

$$\frac{dJ(\theta)}{d\theta} = \pi_\theta\left(g(\theta,\cdot)\right), \qquad \theta \in \Theta.$$

4.2.1 Overview of Gradient Estimation for Markov Chains

Over the past several decades, many techniques have been proposed for various stochastic systems to estimate the gradient (or derivative) with respect to some parameter vector $\theta = (\theta_1, \ldots, \theta_p)$ of steady state performance,

$$\nabla J(\theta) = \left[\frac{\partial J(\theta)}{\partial \theta_1}, \frac{\partial J(\theta)}{\partial \theta_2}, \ldots, \frac{\partial J(\theta)}{\partial \theta_p}\right]', \qquad \theta \in \Theta \subset \mathbb{R}^p.$$

Examples include Infinitesimal Perturbation Analysis (IPA), Likelihood Ratio, Conditional Monte Carlo, Finite Difference, etc. and each method offers certain strengths for specific classes of problems. The last decade in particular has seen an explosion of research on *single sample path* gradient estimation which is well summarized in the recent book [36]. Let us briefly review a few selected highlights of gradient estimation research.

The longstanding alternative to the *single sample path* approach to gradient estimation is Finite Difference (FD) estimates [23, 45] of the form:

$$\frac{\widehat{J}(\theta + \delta\theta_i/2, N) - \widehat{J}(\theta - \delta\theta_i/2, N)}{\delta\theta_i}, \qquad \text{for each } i = 1, 2, \ldots, p, \tag{4.2}$$

for some small $\delta > 0$ and some convenient estimate of performance, such as

$$\widehat{J}(\theta, N) = \frac{1}{N}\sum_{n=0}^{N-1} f(\theta, X_n). \tag{4.3}$$

In a simulation environment, the use of *common random numbers* [23] to estimate the two terms $\widehat{J}(\theta \pm \delta\theta_i/2, N)$ involves running the two simulations with the same random number generator "seed" and same initial state $X_0 = x$. The use of common random numbers have been shown to reduce the variance of FD estimates but this is generally not possible when the system under study is observed in real time from a physical system and neither the "seed" nor the initial conditions can be selected arbitrarily. Using FD estimates in an SA algorithm with a shrinking step-size $\gamma_n \to 0$ and difference $\delta_n \to 0$ is known as the Kiefer-Wolfowitz procedure. One drawback to the original Kiefer-Wolfowitz procedure is the need to run the simulation twice *for each component* of the parameter vector in order to construct each gradient estimate. Recently, this requirement can be relaxed somewhat by using any one of the "random directions" methods [61, 102]. In any event, there is an obvious motivation to develop *single sample path* approaches which promise increased efficiency in general, as well as clear improvements for observation based gradient estimation.

The Likelihood Ratio (LR) method [44, 46, 48, 86, 94] is often proposed for chains which possess *structural parameters*, i.e. chains where the transition probabilities are dependent on the parameter. This parameter dependence is sometimes restricted to chains which do not cause any transition probabilities $p_{xy}(\theta)$ to increase from zero or decrease to zero (i.e. the chain "has no opening/closing arcs"). Unfortunately, the original LR method is less well suited for steady state estimation since it has been recognized [44, 86] that it suffers from unbounded variance which grows linearly with the length of the observation interval. More recently, several special techniques [109, 35] have been suggested to bound this variance.

IPA has been quite successful when the parameter varies the timing of events in a Discrete Event Dynamic System (DEDS) modeled as a Generalized Semi Markov Process (GSMP). Unfortunately, IPA is not generally capable of handling all types of parameter dependence [10, 54, 107] such as *structural* parameters where the tunable parameter continuously varies the transition probabilities (routing probabilities) of the GSMP.

On the other hand, Conditional Monte Carlo and SPA [35, 36, 37] gradient estimates often succeed where IPA fails by taking advantage of the smoothing properties of the conditional expectation. Also, we note that [13] has shown connections between conditional Monte Carlo and the LR method for steady state gradient estimation.

Glasserman [42] has proposed a technique to compute gradient estimates for continuous-time Markov chains which satisfy certain *structural conditions*. He has also proposed an extension to this method to compute gradients for discrete-time Markov chains [40] with respect to structural parameters. He considers the discrete-time chain as a skeleton of a continuous-time chain and uses his structural conditions to find the performance gradient with respect to the exponential holding times. In effect, he is converting the structural parameter to a timing parameter. Then using discrete-time conversion [33, 56], Glasserman's estimate is converted back to discrete-time where it is observed that the resulting estimator is actually a LR gradient estimate.

Recently, Dai and Ho [22, 24] proposed a class of derivative estimators they named Structural Infinitesimal Perturbation Analysis (SIPA) which handles structural parameters under certain conditions. SIPA estimates model perturbations in transition probabilities and utilize additional auxiliary simulated Markov chains which run in parallel to the nominal chain being simulated or observed to construct a derivative estimate. We also note that Fu and Hu have pointed out that SIPA is simply an implementation of Conditional Monte Carlo.

Dai and Ho's work appears to have inspired Cao et al. [12, 15] to develop some related sensitivity estimates based on some concepts they introduce called *realization factors* and *performance potentials*. Several implementations of specific gradient estimates derived from Cao and Chen's theoretical results are proposed and studied in [14]. We shall focus on one of these estimates in particular and identify some convenient alterations which adapt this estimate for use with SA.

4.2.2 Summary of Results

After reviewing Cao, Chen and Wan's approach to sensitivity analysis, we will propose a particular single sample path gradient estimation algorithm which observes the Markov chain for a fixed number of samples, say m samples, at the current iterate θ_n and computes the estimate. This estimate then may used in an SA to update the parameter to θ_{n+1} where the next gradient estimate is constructed over the next m samples, and so on. The algorithm is very simple, has low computational overhead, and can be used with SA in either a pure simulation or a real-time online observation setting. No auxiliary simulations are required for this gradient estimate, as is the case with SIPA estimates. Unfortunately, this estimate is in general biased for any finite m and, as such, will not in general yield convergence to the desired optimizer $\theta^{\star}$. We have a solution for this problem in the next chapter so this chapter focuses entirely on developing this gradient estimate.

We also reconsider a theoretical result by Cao and Chen [12] which provides three alternative expressions for the performance gradient under certain specific conditions. We propose two variations on their result under conditions which are more aligned with the framework of Chapters 2-3 for proving convergence of SA's. Specifically, we assume the family of parameterized Markov chains satisfies either form of the *uniform drift criteria* (D1) or (D2).

Also note, in previous chapters we used the subscript position on θ to index the sequence of parameter iterates generated by the SA algorithm, but for this chapter, the subscript position may also index the individual components of the parameter vector, i.e. for $\theta \in \Theta \subset \mathbb{R}^p$ we have $\theta = (\theta_1, \theta_2, \ldots, \theta_p)'$. For all θ in Θ, define a small change in the i^{th} component by $\Delta\theta_i$, and we only consider perturbed parameters $\theta' = \theta + \Delta\theta_i$ which fall in Θ. For an arbitrary perturbation in θ which is not restricted to the i^{th} component we shall omit the subscript i and simply write $\theta' = \theta + \Delta\theta$.

4.3 Cao-Chen-Wan Sensitivity Analysis

Cao and Chen in [12] address the general goal of sensitivity analysis for both continuous and discrete time Markov chains and they develop several useful tools for this problem. Their setting is somewhat different than we have proposed up to now so let us next summarize[1] their results. Later we shall propose some alterations motivated by the SA based constrained stochastic optimization problem using algorithm (4.1).

4.3.1 Continuous-Time Markov Chains

Suppose $\{X_t : t \geq 0\}$ is a continuous time Markov process on a countable state space X with infinitesimal generator $A = [a_{xy}]_{x,y}$. The process is assumed regular, positive recurrent, irreducible while the infinitesimal generator A obeys the conditions:

$$\begin{aligned} a_{xx} &< 0, \qquad \forall x \\ a_{xy} &\geq 0, \qquad x \neq y \\ \sup_x |a_{xx}| &< \infty. \end{aligned}$$

Without loss of generality, this infinitesimal generator may be normalized [27] so that $\sup_x |a_{xx}| \leq 1$.

Because they are seeking *sensitivity estimates* to a perturbation in the rates of the infinitesimal generator, the perturbed generator takes the form

$$A_\delta = A + \delta Q \tag{4.4}$$

[1] We have altered some parts of their notation slightly.

for some perturbation matrix Q and some small $\delta > 0$. The matrix Q is assumed to obey the equation $Qe = 0$. In this setting, they consider the sensitivity of the invariant distribution to a fixed perturbation defined by Q of the generator "in the direction of Q", i.e.

$$\frac{\partial \pi}{\partial Q} = \lim_{\delta \to 0} \frac{\pi_\delta - \pi}{\delta}.$$

Also, the derivative of the generator "in the direction of Q" follows readily from (4.4):

$$\begin{aligned} \frac{\partial A}{\partial Q} &= \lim_{\delta \to 0} \frac{A_\delta - A}{\delta} \\ &= Q. \end{aligned}$$

Suppose now that performance is defined by a given function $f : \mathsf{X} \to \mathbb{R}$. Then, under the assumption that $\frac{\partial J}{\partial Q} = \left(\frac{\partial \pi}{\partial Q}\right) f$ (where $\left(\frac{\partial \pi}{\partial Q}\right) f = \sum_{\mathsf{x} \in \mathsf{X}} \frac{\partial \pi(x)}{\partial Q} f(x)$), they calculate the performance gradient using f once a few more quantities are defined.

Let the Markov chain sample path with initial state $x \in \mathsf{X}$ be given by $\{X_t^{(x)}, \ \ t \geq 0\} \doteq \{X_t | X_0 = x, \ \ t \geq 0\}$. They define the *perturbation realization factor*:

$$d_{xy} \doteq \lim_{T \to \infty} \left\{ \mathbf{E}_y \left[\int_0^T f(X_t^{(y)}) dt \right] - \mathbf{E}_x \left[\int_0^T f(X_t^{(x)}) dt \right] \right\}, \qquad x, y \in \mathsf{X}$$

and form the matrix $D = [d_{xy}]_{xy}$. Also, the (nonunique) *performance potential* g_x is chosen for all $x \in \mathsf{X}$ such that

$$d_{xy} = g_y - g_x, \qquad x, y \in \mathsf{X}.$$

Cao and Chen point out that the performance potential is only unique up to an additive constant and if we take $J = \pi(f)$ they define it by

$$g_x \doteq \lim_{T \to \infty} \left\{ \mathbf{E}_x \left[\int_0^T f(X_t) dt \right] - TJ \right\}.$$

For some finite $T > 0$, an *estimate* of the performance potential can be given by

$$\widehat{g}_x(T) = \mathbf{E}_x \left[\int_0^T f(X_t) dt \right] - TJ;$$

but, since we intend to estimate the perturbation realization factor it is sufficient to use

$$\widehat{g}_x(T) = \mathbf{E}_x \left[\int_0^T f(X_t) dt \right]$$

owing to the nonuniqueness of this potential relative to additive constants. Finally, let

$$g = [g_0, g_1, g_2, \cdots]'.$$

This next lemma is taken verbatim from [12] and proved in Theorems 31 and 33 of [59]. Let us first define the matrix $M = [m_{xy}]_{x,y}$ with m_{xy} being the *mean first passage time* from state x to state y and note that a Markov process is designated *strong ergodic* if πM is finite [59].

Lemma 4.1 (Kemeny and Snell) *If the Markov process is strong ergodic, then the inverse*

$$(A - e\pi)^{-1} = -\sum_{k=0}^{\infty}(P - e\pi)^k \tag{4.5}$$

exists where $P = I + A$.

For an infinitesimal generator A, Cao and Chen [12] use a *group inverse* [78] defined as

$$A^{\#} \doteq (A - e\pi)^{-1} + e\pi \tag{4.6}$$

which is clearly related to the *fundamental matrix* $Z \doteq (A - e\pi)^{-1}$.

The next result appears in [12] and relates the quantities we have just defined to the performance gradient.

Theorem 4.2 (Cao and Chen) *Assume the Markov chain* $\mathcal{X} = \{X_t, t \geq 0\}$ *is strong ergodic and*

$$\pi(|f|) = \sum_{x \in \mathsf{X}} \pi(x)\,|f(x)| < \infty.$$

The derivative of the steady-state probability is

$$\frac{\partial \pi}{\partial Q} = -\pi Q A^{\#}.$$

Furthermore, if $(\frac{\partial \pi}{\partial Q})f = \left(\frac{\partial}{\partial Q}\right)(\pi f)$ *and the results of all operators are finite, then the performance derivative can be calculated by using the group inverse of* A*, denoted* $A^{\#}$*, or the realization matrix* D*, or the potential vector* g*:*

$$\begin{aligned} \frac{\partial J}{\partial Q} &= -\pi Q A^{\#} f \\ &= \pi Q D' \pi' \\ &= \pi Q g. \end{aligned}$$

Proof: See [12] ■

This is a nice result but it does appear difficult to verify the condition $(\frac{\partial \pi}{\partial Q})f = \left(\frac{\partial}{\partial Q}\right)(\pi f)$ without imposing assumptions such as *finite state space* or a *bounded* performance function. We want to avoid these conditions so we seek alternative means.

4.3.2 Discrete-Time Markov Chains

Similar results are presented for the case when the state process is a discrete-time Markov chain with one-step probability transition matrix P. Their proposed method [12] is to apply the above continuous time results in Theorem 4.2 by simply converting the discrete time chain to continuous time, i.e. by considering the discrete time chain as a uniformized embedded Markov chain in the Markov process with infinitesimal generator $A = P - I$. Consider a perturbation to the nominal transition matrix P given by $P_\delta = P + \delta Q$ for some small $\delta > 0$. Since $A = P - I$, the change in A is also δQ and results similar to Theorem 4.2 for discrete time Markov chains follow readily. See [12, 14] for details.

4.4 A Framework for Discrete Time Gradient Estimation

We have several goals in adapting Cao-Chen's theorem above to yield a similar expression for $\nabla J(\theta)$. First and foremost, the result should be based on conditions which are checkable within the SA framework we have developed in Chapters 2-3. Second, we must allow a θ-dependence in the performance function $f_\theta(x)$. Third, we must allow a more general (nonlinear) dependence on the transition probabilities.

Here, we carry out these alterations within the specialized framework of discrete time Markov chains which possess the uniform drift criteria (D2).

4.4.1 Conditions on the Transition Probabilities

We make these assumptions on the family of transition probabilities $\{P_\theta,\ \theta \in \Theta\}$:

(G1) For $P_\theta = [p_{x,y}(\theta)]_{x,y}$, let the gradient

$$\nabla p_{x,y}(\theta) = \left[\frac{\partial p_{x,y}(\theta)}{\partial \theta_1} \ \frac{\partial p_{x,y}(\theta)}{\partial \theta_2} \ \cdots \ \frac{\partial p_{x,y}(\theta)}{\partial \theta_p} \right]'$$

exist for each $x, y \in \mathsf{X}$ and $\theta \in \Theta$. i.e.

$$\lim_{\Delta\theta_i \to 0} \frac{p_{x,y}(\theta + \Delta\theta_i) - p_{x,y}(\theta)}{\Delta\theta_i} = \frac{\partial p_{x,y}(\theta)}{\partial \theta_i}.$$

(G2) For each $\theta \in \Theta$, there exists some $\delta > 0$ and constant $K_2 < \infty$ such that the following uniform bound holds for all $\theta + \Delta\theta \in \Theta$, such that $\|\Delta\theta\| < \delta$, and $x, y \in \mathsf{X}$:

$$|p_{x,y}(\theta + \Delta\theta) - p_{x,y}(\theta)| \le K_2 p_{x,y}(\theta) \|\Delta\theta\|.$$

Note that G(2) is a special case of Chapter 3's condition (M).

Consequences of (G1)-(G2)

1. Note that (G1) and (G2) clearly imply that:

$$\text{If } p_{x,y}(\theta) = 0 \text{ for any } x, y \in \mathsf{X} \text{ and } \theta \in \Theta, \text{ then } \nabla p_{x,y}(\theta) = 0. \tag{4.7}$$

2. For each $\theta \in \Theta$ and $i = 1, 2, \ldots, p$, the existence of a $K_2 < \infty$ such that:

$$\left| \frac{\partial p_{x,y}(\theta)}{\partial \theta_i} \right| \le K_2 p_{x,y}(\theta), \qquad x, y \in \mathsf{X}. \tag{4.8}$$

Also we readily see from (4.8) that

$$\begin{aligned} \|\nabla p_{x,y}(\theta)\| &= \left(\sum_{i=1}^{p} \left| \frac{\partial p_{x,y}(\theta)}{\partial \theta_i} \right|^2 \right)^{1/2}, \\ &= \sqrt{p} K_2 p_{x,y}(\theta) \\ &= \widetilde{K}_2 p_{x,y}(\theta), \end{aligned} \tag{4.9}$$

where we define $\widetilde{K}_2 \doteq \sqrt{p} K_2$.

3. Since the partial derivatives exist under (G1), let us define the following difference

$$r_{x,y}(\theta, \Delta\theta_i) \doteq p_{x,y}(\theta + \Delta\theta_i) - p_{x,y}(\theta) - \frac{\partial p_{x,y}(\theta)}{\partial \theta_i} \Delta\theta_i.$$

It then follows readily from (G2) and (4.8) that for each point $\theta \in \Theta$ and each component $i = 1, \ldots, p$ there exists some $K_2'' < \infty$ and $\delta > 0$ such that

$$|r_{x,y}(\theta, \Delta\theta_i)| \leq K_2'' p_{x,y}(\theta) |\Delta\theta_i|, \qquad \theta + \Delta\theta \in \Theta, \quad |\Delta\theta_i| < \delta, \quad x, y \in \mathsf{X}. \tag{4.10}$$

Finally, let us define the following notation for each $i = 1, 2, \ldots, p$

$$\begin{aligned} R_i(\theta, \Delta\theta_i) &\doteq [r_{x,y}(\theta, \Delta\theta_i)]_{x,y} \\ Q_{\theta,i} &\doteq \frac{\partial P_\theta}{\partial \theta_i} = \left[\frac{\partial p_{x,y}(\theta)}{\partial \theta_i}\right]_{x,y}. \end{aligned}$$

4.4.2 Conditions on the Performance Function

We have previously defined the performance function $f_\theta(x) = f(\theta, x) : \Theta \times \mathsf{X} \to \mathbb{R}$. Recall that r is a real number such that $0 < r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ where $\widehat{\ell}_1$ is defined in (S) and V satisfies either (D0), (D1) or (D2), one (or more) of which will always be assumed when using the following conditions.

(F1) There exists a constant $C_1 < \infty$ such that $|f_\theta(x)| \leq C_1 V^r(x)$ for all $x \in \mathsf{X}$ and $\theta \in \Theta$.

(F2) The function $f_\theta(x)$ is differentiable with respect to each component $\theta_i, i = 1, 2, \ldots, p$ for all $\theta \in \Theta, x \in \mathsf{X}$.

(F3) For each $\theta \in \Theta$, there exists a $\delta > 0$ and some $C_3' < \infty$ such that

$$|f_{\theta+\Delta\theta}(x) - f_\theta(x)| \leq C_3' \|\Delta\theta\| V^r(x), \qquad \|\Delta\theta\| \leq \delta, \; x \in \mathsf{X}.$$

A Consequence of (F2)-(F3)

Clearly, for each $\theta \in \Theta$ there exists a constant $C_3 < \infty$ such that $\left|\frac{\partial f_\theta(x)}{\partial \theta_i}\right| \leq C_3 V^r(x)$ for all $x \in \mathsf{X}$, and $i = 1, \ldots, p$.

4.4.3 Realization Factors and Performance Potentials

We next restate some of Cao and Chen's definitions defined earlier for continuous time Markov chains (realization factor, performance potential, etc.) to discrete time chains. These redefined quantities will now permit the performance function to depend on $\theta \in \Theta$, as well as the one step transition kernels $\{P_\theta, \theta \in \Theta\}$.

Let $\{X_n^{\{y\}}\} = \{X_n | X_0 = y,\ n \geq 0\}$ be a Markov chain sample path starting in state y and define the *first passage time* from state y to state x as

$$L^{\{y\}}(x) \doteq \inf\{n | X_n^{\{y\}} = x;\ n \geq 0\}.$$

With that, for all fixed $\theta \in \Theta$, Cao and Chen [12] have defined the *realization factor* and using a coupling argument have shown the following equalities under their conditions:

$$\begin{aligned} d_{xy}(\theta) &\doteq \lim_{n\to\infty} \left\{ \mathbf{E}_{\theta,y}\left[\sum_{k=0}^{n} f_\theta(X_k^{\{y\}})\right] - \mathbf{E}_{\theta,x}\left[\sum_{k=0}^{n} f_\theta(X_k^{\{x\}})\right]\right\} \\ &= \mathbf{E}_{\theta,y}\left[\sum_{k=0}^{L^{\{y\}}(x)-1} f_\theta(X_k^{\{y\}})\right] - \mathbf{E}_{\theta,y}\left[L^{\{y\}}(x)\right] J(\theta), \qquad x, y \in \mathsf{X} \\ &= \mathbf{E}_{\theta,y}\left[\sum_{k=0}^{L^{\{y\}}(x)-1} f_\theta(X_k^{\{y\}}) - J(\theta)\right] \end{aligned}$$

Additionally, they have defined what they call a *performance potential*

$$g_x(\theta) \doteq \lim_{n\to\infty} \left\{ \mathbf{E}_{\theta,x}\left[\sum_{k=0}^{n} f_\theta(X_k^{\{x\}})\right] - nJ(\theta)\right\}, \qquad x \in \mathsf{X}.$$

which obeys

$$d_{xy}(\theta) = g_y(\theta) - g_x(\theta) \tag{4.11}$$

As in Section 4.3, let us define for each $\theta \in \Theta$ the matrix $D_\theta \doteq [d_{xy}(\theta)]$ and column vector $g_\theta = [g_x(\theta)]_{x\in\mathsf{X}}$. Note we have generalized Cao and Chen's original definitions to allow for dependence on θ over Θ.

For an *estimate* of the performance potential taken from a finite number of steps N we can use

$$\widehat{g}_x(\theta, N) = \mathbf{E}_{\theta,x}\left[\sum_{k=0}^{N-1} f_\theta(X_k^{\{x\}}) - NJ(\theta)\right]$$

or, owing to the nonuniqueness relative to additive constants, it suffices to use

$$\widehat{g}_x(\theta, N) = \mathbf{E}_{\theta,x}\left[\sum_{k=0}^{N-1} f_\theta(X_k^{\{x\}})\right]$$

to estimate the realization factor. Let us also define $\widehat{g}_{\theta,N} = [\widehat{g}_x(\theta, N)]_{x\in\mathsf{X}}$.

Note: Observe that the following is one form of the Poisson equation solution which converges under appropriate conditions [79, Theorem 17.4.2]

$$\begin{aligned} g_x(\theta) &\doteq \lim_{n\to\infty} \mathbf{E}_{x,\theta}\left[\sum_{k=0}^{n}\left(f_\theta(X_k^{\{x\}}) - J(\theta)\right)\right] \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \mathbf{E}_{x,\theta}\left[f_\theta(X_k^{\{x\}}) - J(\theta)\right]. \end{aligned}$$

Additionally, the Poisson equation solution is unique (under certain conditions) up to an additive constant via Theorem 17.4.1 in [79]. Hence, the "performance potential" is simply a Poisson equation solution under these particular conditions.

4.4.4 Two Results on the Fundamental Matrix

Our first result depends on the existence of the fundamental matrix and the following lemma provides conditions for its existence. Before we begin, recall we have defined the matrix $M = [m_{x,y}]_{x,y}$ with $m_{x,y}$ being the *mean first passage time* from state x to state y and a Markov chain is labeled *strong ergodic* if πM is finite [59]. We have the following lemma for discrete time Markov chains governed by P which admits an invariant π.

Lemma 4.3 (Kemeny-Snell) *If the Markov chain is strong ergodic, then the fundamental matrix*

$$Z \doteq \sum_{k=0}^{\infty}(P - e\pi)^k \tag{4.12}$$

exists and Z is both a left and right inverse operator for $(I - P + e\pi)$. Thus

$$Z = \sum_{k=0}^{\infty}(P - e\pi)^k = (I - P + e\pi)^{-1}. \tag{4.13}$$

Proof: See [59, Thms. 31, 33]. ■

Unfortunately, *strong ergodicity* may be difficult to verify directly but there is one case where it is readily known, and that is the case of finite state, irreducible, positive recurrent chains.

We will also find useful the following result by Glynn and Meyn for general state space Markov chains, although we shall only apply these results for chains restricted to a countable state space. Recall that $L_V^\infty \doteq \{h : \sup_{x\in \mathsf{X}} \frac{|h(x)|}{V(x)} < \infty\}$.

Lemma 4.4 (Glynn and Meyn [49]) *Assume:*

1. *$\{P_\theta : \theta \in \Theta\}$ is a family of Markov transition functions where Θ denotes some open subset of Euclidean space.*

2. *Each of the corresponding Markov chains is ϕ_θ-irreducible.*

3. *For each $\theta_0 \in \Theta$, the following drift criterion holds for some $\delta > 0$ in some open ball $B^\delta(\theta_0)$ containing θ_0*

   $$P_\theta V \leq \lambda V + b1_C, \qquad \theta \in B^\delta(\theta_0) = \{\theta \in \mathbb{R}^p : \|\theta - \theta_0\| < \delta\}.$$

   for some common petite set C.

4. *$P_\theta \to P_{\theta_0}$ as $\theta \to \theta_0$ in the induced operator norm $|||\cdot|||_V$, i.e.*

   $$\lim_{\theta\to\theta_0} |||P_\theta - P_{\theta_0}|||_V = \lim_{\theta\to\theta_0} \sup_{\substack{h\in L_V^\infty \\ |h|_V=1}} |(P_\theta - P_{\theta_0})h|_V = 0 \tag{4.14}$$

Let $\{\pi_\theta : \theta \in \Theta\}$ denote the collection of invariant probabilities and, assuming the inverse is well defined, let $\{Z_\theta = (I - P_\theta + \Pi_\theta)^{-1} : \theta \in \Theta\}$ denote the collection of fundamental kernels where $\Pi_\theta(x, A) = \pi_\theta(A),\ x \in \mathsf{X},\ A \in \mathcal{B}(\mathsf{X})$. With these assumptions:

1. *Each of the kernels* $\{Z_\theta, P_\theta, \Pi_\theta : \theta \in \Theta\}$ *is a bounded linear transformation from* L_V^∞ *to* L_V^∞.

2. *The invariant probabilities converges in* V*-total variation norm, i.e.*

$$\lim_{\theta \to \theta_0} \|\pi_\theta - \pi_{\theta_0}\|_V = 0 \tag{4.15}$$

3. *We have*

$$\lim_{\theta \to \theta_0} |||Z_\theta - Z_{\theta_0}|||_V = 0 \tag{4.16}$$

hence if $f \in L_V^\infty$, *then Poisson's equation solution, given by* $\widehat{g}_\theta = Z_\theta f$, *converges in the* L_V^∞ *norm as* $\theta \to \theta_0$, *i.e.*

$$\lim_{\theta \to \theta_0} \sup_{x \in \mathsf{X}} \frac{|Z_\theta(x,f) - Z_{\theta_0}(x,f)|}{V(x)} = 0 \tag{4.17}$$

Proof: The first conclusion is reached via [49, Theorem 2.3] and the second and third follow via a generalization of Schweitzer's result [95, Theorem2]. For details, see [49, Section 4.2]. ■

4.5 Gradient Estimation for Discrete Time Markov Chains

In this section, we rework Cao and Chen's Theorem 4.2 specializing it to discrete time chains where transition probabilities are dependent on a parameter vector θ and satisfying a uniform drift criteria (D2). We apply some recent results on the smoothness of solutions to Poisson's equation for Markov chains which satisfy a uniform drift criteria (D2) and thus present an alternative version of Cao and Chen's Theorem 4.2. This new version allows the performance function to be unbounded and to have a functional dependence on θ, hence our theorem offers an extension to Cao and Chen's theorem under the special case of chains satisfying a uniform drift criteria (D2). Our version is directed at the SA framework of the previous two chapters.

We shall also present a second variant of this main result which, like Cao and Chen's version, provides various expressions for the steady-state performance gradient. While, the first theorem assumes the existence of the fundamental matrix (or strong ergodicity), the second theorem assumes the existence of a solution to the Poisson equation. This second version sometimes has advantages if strong ergodicity cannot be proven (as is often the case with countable state space chains) since it may be possible to prove existence of the solution to the Poisson equation by other means. Specifically, solutions to the Poisson equation can sometimes be established via probabilistic methods such as those presented in [73] or under drift conditions as in [79].

4.5.1 The (First) Main Result

For this section we assume the existence of the fundamental matrix Z_θ where Lemma 4.3 gives sufficient conditions for it to exist at any fixed $\theta \in \Theta$. Consider the following θ-parameterized *group inverse* we define as

$$\begin{aligned} P_\theta^{\#} &\doteq -e\pi_\theta + (I - P_\theta + e\pi_\theta)^{-1} \\ &= -e\pi_\theta + Z_\theta \\ &= -e\pi_\theta + \sum_{k=0}^{\infty}(P_\theta - e\pi_\theta)^k. \end{aligned}$$

This is simply the sum of a (negative) invariant matrix $e\pi_\theta$ and the fundamental matrix. Since we have

$$Z_\theta \doteq \sum_{k=0}^{\infty}(P_\theta - e\pi_\theta)^k = I + \sum_{k=1}^{\infty}(P_\theta^k - e\pi_\theta),$$

we immediately get

$$P_\theta^{\#} = \sum_{k=0}^{\infty}(P_\theta^k - e\pi_\theta).$$

Then, it's not too difficult to verify

$$(I - P_\theta)P_\theta^{\#} = P_\theta^{\#}(I - P_\theta) = I - e\pi_\theta. \tag{4.18}$$

Theorem 4.5 *Assume:*

1. *We have a family of discrete-time countable state positive recurrent irreducible Markov chains governed by one-step transition matrices* $\{P_\theta, \theta \in \Theta\}$ *where* $\{\pi_\theta,\ \theta \in \Theta\}$ *denotes the corresponding collection of invariant probabilities.*

2. *The fundamental matrix*

$$Z_\theta = \sum_{k=0}^{\infty}(P_\theta - e\pi_\theta)^k$$

exists and is the left and right inverse operator for $(I - P_\theta + e\pi_\theta)$ *for all* $\theta \in \Theta$.

3. *The matrix* $Q_{\theta,i} \doteq \left[\frac{\partial p_{x,y}}{\partial \theta_i}(\theta)\right]_{x,y}$ *is such that* $Q_{\theta,i}e = 0$ *for each* $i = 1, \ldots, p$.

4. *Conditions (D2), (G1)-(G2), (F1)-(F3) all hold.*

Then, for each $i = 1, 2, \ldots, p$, *the partial derivatives are given by*

$$\begin{aligned} \frac{\partial J(\theta)}{\partial \theta_i} &= \pi_\theta\left(\frac{\partial f_\theta}{\partial \theta_i}\right) + \pi_\theta Q_{\theta,i} P_\theta^{\#} f_\theta \\ &= \pi_\theta\left(\frac{\partial f_\theta}{\partial \theta_i}\right) + \pi_\theta Q_{\theta,i} g_\theta \\ &= \pi_\theta\left(\frac{\partial f_\theta}{\partial \theta_i}\right) + \pi_\theta Q_{\theta,i} \nu_\theta. \qquad \theta \in \Theta. \end{aligned}$$

where $\{\nu_\theta, \theta \in \Theta\}$ *is any of the Poisson equation solutions (which are unique only up to a constant).*

Proof: See the appendix.

Remark

In [12, 15], the authors also show the following equality (under their conditions) involving the realization matrix D:

$$\pi_\theta Q_{\theta,i} P_\theta^{\#} f_\theta = \pi_\theta Q_{\theta,i} D'_\theta \pi'_\theta = \pi_\theta Q_{\theta,i} g_\theta, \qquad \theta \in \Theta.$$

4.5.2 An Alternate Version using the Poisson Equation

We now rework the last result except here we do not explicitly assume strong ergodicity or the existence of the fundamental matrix Z_θ. Instead, we rely simply on the existence of a solution $\nu_\theta : \mathsf{X} \to \mathbb{R}$ to the Poisson equation for each $\theta \in \Theta$. We note that this was also assumed in our general SA framework of Chapter 2, i.e. condition (P1), so we are not adding any additional conditions by this. Chapter 3 provided sufficient conditions for (P1) hence those results may be applied here as well.

On a countable state space, the solution ν_θ as well as performance function f_θ may be represented as column vectors, so the Poisson equation can be stated as a matrix equation

$$\begin{aligned} f_\theta - e\pi_\theta(f_\theta) &= \nu_\theta - P_\theta \nu_\theta \\ &= (I - P_\theta)\nu_\theta, \qquad \theta \in \Theta. \end{aligned} \tag{4.19}$$

Observe that the last term on the left hand side is simply $J(\theta) = \pi_\theta(f_\theta)$ converted to a vector, the steady state performance.

Theorem 4.6 *Assume:*

1. *We have a family of discrete-time countable state positive recurrent irreducible Markov chains governed by one-step transition matrices $\{P_\theta, \theta \in \Theta\}$ where $\{\pi_\theta,\ \theta \in \Theta\}$ denotes the corresponding collection of invariant probabilities.*
2. *The matrix $Q_{\theta,i} \doteq \left[\frac{\partial p_{x,y}}{\partial \theta_i}(\theta)\right]_{x,y}$ is such that $Q_{\theta,i} e = 0$ for each $i = 1, \dots, p$.*
3. *For each $\theta \in \Theta$, the Poisson equation (4.19) admits a solution denoted ν_θ.*
4. *There exists a constant $C < \infty$ with*

$$\sup_{\theta \in \Theta} |\nu_\theta(x)| \le C V^r(x), \qquad \text{for all } x \in \mathsf{X}, \tag{4.20}$$

5. *For each $x \in \mathsf{X}$, the solution $\nu_\theta(x)$ is continuous on Θ.*
6. *The conditions (D2), (G1)-(G2), (F1)-(F3) all hold.*

Then, for each $i = 1, 2, \ldots, p$, the partial derivatives

$$\frac{\partial J(\theta)}{\partial \theta_i} = \pi_\theta \left(\frac{\partial f_\theta}{\partial \theta_i} \right) + \pi_\theta Q_{\theta,i} \nu_\theta \quad (4.21)$$

$$= \pi_\theta \left(\frac{\partial f_\theta}{\partial \theta_i} \right) + \pi_\theta Q_{\theta,i} g_\theta. \qquad \theta \in \Theta. \quad (4.22)$$

where $\{\nu_\theta, \theta \in \Theta\}$ is any of the Poisson equation solutions (which are unique only up to a constant).

Proof: See the appendix.

4.6 A Biased Gradient Estimate for Stochastic Approximation

We now consider a specific *biased* gradient estimation algorithm in the discrete time setting which is adapted from Cao and Wan's "3c estimator" in [14]. Cao and Wan have shown [14] that for some *fixed* positive integer m that

$$\lim_{M \to \infty} \frac{1}{M-m} \sum_{k=m}^{M-1} \left\{ \frac{\frac{\partial}{\partial \theta_i} p_{X_{k-m}, X_{k-m+1}}(\theta)}{p_{X_{k-m}, X_{k-m+1}}(\theta)} \right\} \sum_{j=0}^{m-1} f_\theta(X_{k-j}) = \pi_\theta Q_{\theta,i} \widehat{g}_{\theta,m} \quad \mathbf{P}_{\theta,x} - a.s. \quad (4.23)$$

for the case when the performance function $f_\theta(\cdot)$ does not depend on θ. Also, they suggest simply using the finite sample average on the left hand side in (4.23), i.e. a large fixed integer M such that $M >> m$, as a performance gradient estimate. Our setting is slightly more general in that it allows performance functions to depend on θ. As such, under the conditions of Theorem 4.5 or Theorem 4.6, we have shown a similar conclusion as Cao and Chen's result in Theorem 4.2. Let us define the *bias*

$$\beta_{i,m}(\theta) \doteq \pi_\theta Q_{\theta,i} \widehat{g}_{\theta,m} - \pi_\theta Q_{\theta,i} g_\theta$$

so that either Theorem 4.5 or Theorem 4.6, the limit (4.23), and a simple application of the SLLN for Markov chains yields

$$\frac{\partial J(\theta)}{\partial \theta_i} + \beta_{i,m}(\theta) = \pi_\theta\left(\frac{\partial f_\theta}{\partial \theta_i}\right) + \pi_\theta Q_{\theta,i} \widehat{g}_{\theta,m}$$

$$= \lim_{M \to \infty} \frac{1}{M-m} \sum_{k=m}^{M-1} \frac{\partial f_\theta}{\partial \theta_i}(X_k) \quad (4.24)$$

$$+ \lim_{M \to \infty} \frac{1}{M-m} \sum_{k=m}^{M-1} \left\{ \frac{\frac{\partial}{\partial \theta_i} p_{X_{k-m}, X_{k-m+1}}(\theta)}{p_{X_{k-m}, X_{k-m+1}}(\theta)} \right\} \sum_{j=0}^{m-1} f_\theta(X_{k-j}) \quad (4.25)$$

$$\mathbf{P}_{\theta,x} - a.s.$$

Note the similarity in the form of this estimate to the standard LR estimates [44, 46, 48]; with the main difference being that these estimates have a truncated inner sum window length of size m.

We now propose some alterations to (4.24)-(4.25) which are motivated by SA. The quantities inside the first limit (4.24) can be rewritten:

$$\begin{aligned} \lim_{M\to\infty} \frac{1}{M-m} \sum_{k=1}^{M-m} \frac{\partial f_\theta}{\partial \theta_i}(X_k) &= \lim_{N\to\infty} \frac{1}{Nm-m} \sum_{k=1}^{Nm-m} \frac{\partial f_\theta}{\partial \theta_i}(X_k) \\ &= \lim_{N\to\infty} \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{1}{m} \sum_{j=0}^{m-1} \frac{\partial f_\theta}{\partial \theta_i}(X_{nm-j}) \right). \end{aligned} \tag{4.26}$$

The left hand side of (4.26) converges a.s. via the SLLN for Markov chains, and the right hand side is simply a subsequence so it also converges to the same limit.

The quantities inside the second limit (4.25) can be reindexed as well:

$$\lim_{M\to\infty} \frac{1}{M-m} \sum_{k=m}^{M-1} \left\{ \frac{\frac{\partial}{\partial\theta_i} p_{X_{k-m},X_{k-m+1}}(\theta)}{p_{X_{k-m},X_{k-m+1}}(\theta)} \right\} \sum_{j=0}^{m-1} f_\theta(X_{k-j}) \tag{4.27}$$

$$= \lim_{N\to\infty} \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{1}{m} \sum_{l=0}^{m-1} \left\{ \frac{\frac{\partial}{\partial\theta_i} p_{X_{nm-m+l},X_{nm-m+l+1}}(\theta)}{p_{X_{nm-m+l},X_{nm-m+l+1}}(\theta)} \right\} \sum_{j=0}^{m-1} f_\theta(X_{nm+l-j}) \right)$$

$$= \pi_\theta Q_{\theta,i} \widehat{g}_{\theta,m}, \qquad \mathbf{P}_{\theta,x} - a.s. \tag{4.28}$$

$$= \mathbf{E}_{\pi_\theta} \left[\frac{\frac{\partial}{\partial\theta_i} p_{X_0,X_1}(\theta)}{p_{X_0,X_1}(\theta)} \sum_{j=0}^{m-1} f_\theta(X_{m-j}) \right]. \tag{4.29}$$

to a form similar to (4.26).

We now claim that we also have convergence to the same steady state expectation (4.29) if we replace the inner average over m terms by only the first term in the average, i.e.

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{\frac{\partial}{\partial\theta_i} p_{X_{nm-m},X_{nm-m+1}}(\theta)}{p_{X_{nm-m},X_{nm-m+1}}(\theta)} \sum_{j=0}^{m-1} f_\theta(X_{nm-j}) \right) \\ &= \mathbf{E}_{\pi_\theta} \left[\frac{\frac{\partial}{\partial\theta_i} p_{X_0,X_1}(\theta)}{p_{X_0,X_1}(\theta)} \sum_{j=0}^{m-1} f_\theta(X_{m-j}) \right], \quad \mathbf{P}_{\theta,x} - a.s. \end{aligned} \tag{4.30}$$

4.6.1 m-Window Process

Now for a fixed window size m, let us define $\{Y_n,\ n = m, m+1, \ldots\}$ as the vector formed by the $m+1$ most recent samples of $\{X_n,\ n = 0, 1, \ldots\}$, i.e.

$$Y_n \doteq (X_n, X_{n-1}, \ldots, X_{n-m}), \qquad n \geq m,$$

and let $\{Z_n,\ n = 1, \ldots\}$ be the m-skeleton of $\{Y_n,\ n = m, m+1, \ldots\}$:

$$\begin{aligned} Z_n &\doteq Y_{mn} \\ &= (X_{mn}, X_{mn-1}, \ldots, X_{mn-m}), \qquad n \geq 1, \end{aligned}$$

which we will refer to as the m-window process.

Clearly, both of the processes $\{Y_n;\ n = m, m+1, \ldots\}$ and $\{Z_n;\ n = 1, 2, \ldots\}$ are Markov if $\{X_k;\ k = 0, 1, \ldots\}$ is Markov. In fact, the m-window process defined by $Z_n = \{X_{mn}, X_{mn-1}, \ldots, X_{mn-m}\}$ is Markov with transition function $\bar{P}_\theta$ given by

$$\begin{aligned}
&\bar{P}_\theta[Z_{n+1} \in B_m \times B_{m-1} \times \ldots \times B_0 | Z_n = (x_0, x_{-1}, \ldots, x_{-m})] \\
&= P_\theta[X_{mn+m} \in B_m, \ldots, X_{mn+1} \in B_1, X_{mn} \in B_0 | X_{mn} = x_0, \ldots, X_{mn-m} = x_{-m}] \\
&= P_\theta[X_{mn+m} \in B_m, \ldots, X_{mn+1} \in B_1, X_{mn} \in B_0 | X_{mn} = x_0] \\
&= P_\theta[X_{mn+m} \in B_m, \ldots, X_{mn+1} \in B_1 | X_{mn} = x_0] 1_{B_0}(x_0) \\
&= \sum_{B_m} \sum_{B_{m-1}} \ldots \sum_{B_1} \prod_{l=1}^{m} P_\theta[X_{mn+l} = x_l | X_{mn+l-1} = x_{l-1}] 1_{B_0}(x_0)
\end{aligned}$$

for $B_l \in \mathcal{B}(\mathsf{X})$, for $l = 0, 1, \ldots, m$. Also, it should be obvious that the invariant distribution for $\{Z_n;\ n = 1, 2, \ldots\}$, denoted $\bar{\pi}_\theta$, is given by

$$\bar{\pi}_\theta(B_m \times B_{m-1} \times \ldots \times B_0) = \sum_{B_m} \sum_{B_{m-1}} \ldots \sum_{B_0} \prod_{l=1}^{m} P_\theta[X_l \in x_l | X_{l-1} = x_{l-1}] \pi_\theta(x_0)$$

4.7 A Modified Gradient Estimate

Now, in the same manner as Cao and Wan [14] used to show almost sure convergence of (4.23) we also have convergence via the Strong Law of Large Numbers for Markov Chains applied to $\{Z_n, n = 0, 1, \ldots\}$ so that

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{\frac{\partial}{\partial\theta_i} p_{X_{nm-m}, X_{nm-m+1}}(\theta)}{p_{X_{nm-m}, X_{nm-m+1}}(\theta)} \sum_{j=0}^{m-1} f_\theta(X_{nm-j}) \right) \\
&= \mathbf{E}_{\bar{\pi}_\theta} \left[\frac{\frac{\partial}{\partial\theta_i} p_{X_0, X_1}(\theta)}{p_{X_0, X_1}(\theta)} \sum_{j=0}^{m-1} f_\theta(X_{m-j}) \right], \quad \mathbf{P}_{\theta,x} - a.s. \qquad (4.31)
\end{aligned}$$

Define for each $\theta \in \Theta$ and each $i = 1, \ldots, p$ the function $\widehat{G}_i : \Theta \times \mathsf{X}^{m+1} \to \mathbb{R}$ as

$$\widehat{G}_i(\theta, Z_n) \doteq \frac{1}{m} \sum_{j=0}^{m-1} \frac{\partial f_\theta}{\partial\theta_i}(X_{nm-j}) + \left\{ \frac{\frac{\partial}{\partial\theta_i} p_{X_{nm-m}, X_{nm-m+1}}(\theta)}{p_{X_{nm-m}, X_{nm-m+1}}(\theta)} \right\} \sum_{j=0}^{m-1} f_\theta(X_{nm-j}), \quad n = 1, 2, \ldots,$$

and this will serve as our *gradient estimate* based on an observed window of $m+1$ samples $Z_n = (X_{mn}, X_{mn-1}, \ldots, X_{mn-m})$.

Next, we note for the estimate function $\widehat{G}_i(\theta, \cdot) : \mathsf{X}^{m+1} \to \mathbb{R}$, we have

$$\begin{aligned}
\mathbf{E}_\theta \left[\widehat{G}_i(\theta, Z_{n+1}) | Z_n \right] &= \mathbf{E}_\theta \left[\widehat{G}_i(\theta, X_{mn+m}, \ldots, X_{mn+1}, X_{mn}) | X_{mn}, \ldots, X_{mn-m} \right] \\
&= \mathbf{E}_\theta \left[\widehat{G}_i(\theta, X_{mn+m}, \ldots, X_{mn+1}, X_{mn}) | X_{mn} \right]
\end{aligned}$$

and for any $n = 1, 2, \ldots$

$$
\begin{aligned}
&\mathbf{E}_{\bar{\pi}_\theta}[\widehat{G}_i(\theta, Z_{n+1})] \\
&= \sum_{z_0 \in \mathsf{X}^{m+1}} \mathbf{E}_\theta \left[\widehat{G}_i(\theta, Z_{n+1}) | Z_0 = z_0\right] \bar{\pi}_\theta(z_0) \\
&= \sum_{\mathsf{X}^{m+1}} \mathbf{E}_\theta \left[\widehat{G}_i(\theta, X_{mn+m}, \ldots, X_{mn+1}, X_{mn}) | X_0 = x_0, \ldots, X_{-m} = x_{-m}\right] \bar{\pi}_\theta(z_0) \\
&= \sum_{\mathsf{X}^{m+1}} \mathbf{E}_\theta \left[\widehat{G}_i(\theta, X_{mn+m}, \ldots, X_{mn+1}, X_{mn}) | X_0 = x_0\right] \bar{\pi}_\theta(z_0 = (x_0, x_{-1}, \ldots, x_{-m})) \\
&= \sum_{x_0 \in \mathsf{X}} \mathbf{E}_\theta \left[\widehat{G}_i(\theta, X_{mn+m}, \ldots, X_{mn+1}, X_{mn}) | X_0 = x_0\right] \pi_\theta(x_0) \\
&= \mathbf{E}_{\pi_\theta}[\widehat{G}_i(\theta, X_{mn+m}, \ldots, X_{mn+1}, X_{mn})] \\
&= \mathbf{E}_{\pi_\theta}[\widehat{G}_i(\theta, Z_{n+1})]
\end{aligned}
$$

This with (4.31) shows the claim (4.30). Hence for each $\theta \in \Theta$ and for arbitrary $n = 1, 2, \ldots$:

$$
\begin{aligned}
&\pi_\theta \left(\frac{\partial f_\theta}{\partial \theta_i}\right) + \pi_\theta Q_{\theta,i} \widehat{g}_{\theta,m} \\
&= \mathbf{E}_{\pi_\theta} \left[\left(\frac{1}{m} \sum_{j=0}^{m-1} \frac{\partial f_\theta}{\partial \theta_i}(X_{nm-j}) \right) + \left\{ \frac{\frac{\partial}{\partial \theta_i} p_{X_{nm-m}, X_{nm-m+1}}(\theta)}{p_{X_{nm-m}, X_{nm-m+1}}(\theta)} \right\} \sum_{j=0}^{m-1} f_\theta(X_{nm-j}) \right],
\end{aligned}
$$

and although there is a bias, this does suggest a possible function to drive the stochastic approximation algorithm for optimization.

Let us write

$$
\begin{aligned}
\widehat{G}(\theta, Z_n) &\doteq \left[\widehat{G}_1(\theta, Z_n), \widehat{G}_2(\theta, Z_n), \cdots, \widehat{G}_p(\theta, Z_n)\right]' \\
\beta_m(\theta) &= [\beta_{1,m}(\theta), \beta_{2,m}(\theta), \ldots, \beta_{p,m}(\theta)]'
\end{aligned}
$$

and, as we have discussed, we have

$$
\begin{aligned}
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \widehat{G}(\theta, Z_n) &= \mathbf{E}_{\pi_\theta} \left[\widehat{G}(\theta, Z_1)\right], \qquad \mathbf{P}_{\theta,x} - a.s. \\
&= \nabla J(\theta) + \beta_m(\theta), \qquad \theta \in \Theta.
\end{aligned}
$$

Thus, if an algorithm of the form

$$
\theta_{n+1} = \Pi_\Theta \left\{\theta_n - \gamma_{n+1} \widehat{G}(\theta_n, Z_{n+1})\right\}, \qquad n = 0, 1, \ldots \tag{4.32}
$$

with a fixed m is used, we see that due to a possibly nonzero bias term, in general the iterates will not be convergent to the optimal $\theta^\star$ such that $\nabla J(\theta^\star) = 0$. The next chapter resolves this issue with a $\theta^\star$ convergent algorithm.

Chapter 5

Stochastic Optimization of Steady State Performance

We develop an SA algorithm which is appropriate for the gradient estimate of the last chapter. Since that estimate was biased for any fixed observation window, we use a sequence of increasing window lengths to achieve convergence to the optimal parameter.

5.1 Introduction

Consider the projected stochastic approximation algorithm defined by the recursion

$$\begin{aligned} \theta_{n+1} &= \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} H_{\ell_{n+1}}(\theta_n, \bar{X}_{n+1}) + \gamma_{n+1}^{1+\widehat{\ell}_1} \bar{\rho}_{n+1}(\theta_n, \bar{X}_{n+1}) \right\}, \qquad n = 0, 1, 2, \ldots \\ \theta_0 &= \theta \end{aligned} \tag{5.1}$$

which takes observations from the state process over a window

$$\bar{X}_{n+1} \doteq \left(X_{n+1,0}, X_{n+1,1}, \ldots, X_{n+1,\ell_{n+1}} \right), \qquad n = 0, 1, \ldots$$

Within this $(n+1)^{th}$ window, the state process $\{X_{n+1,k}, k = 0, 1, \ldots, \ell_{n+1}\}$ is simply a Markov chain taking values on X and governed by one-step transition kernel (or matrix) P_{θ_n} from the family $\{P_\theta, \theta \in \Theta\}$. Each observation window is initialized with the last sample of the previous window, i.e. $X_{n+1,0} = X_{n,\ell_n}$, and the first observation window is initialized at $X_{1,0} = x \in \mathsf{X}$. (See the next section for a complete description.)

The deterministic sequence $\{\ell_n, n = 0, 1, \ldots\}$ defines the length ℓ_n of the n^{th} *observation window* for each step of the algorithm. As before, a compact projection set $\Theta \subset \mathbb{R}^p$ is assumed and the algorithm is driven by the functions $H_{\ell_{n+1}} : \Theta \times \mathsf{X}^{1+\ell_{n+1}} \to \mathbb{R}^p$ and $\bar{\rho}_{n+1} : \Theta \times \mathsf{X}^{1+\ell_{n+1}} \to \mathbb{R}^p$ for $n = 0, 1, \ldots$.

In this chapter, we develop conditions for convergence of the iterates $\{\theta_n, n = 0, 1, \ldots\}$ when the ℓ_{n+1} increases slowly, on the order of $\ell_n \approx \log(n), \quad n = 1, 2, \ldots$. We continue to assume the previously defined (S) for the step-size sequence $\{\gamma_{n+1}, n = 0, 1, \ldots\}$ although we slightly strengthen it to the form ($\overline{\text{S}}$) below.

($\overline{\textbf{S}}$) For $0 < \widehat{\ell}_1 < 1$ from (5.1) the following holds:

a) $\quad \gamma_n > 0$ and $\gamma_n \geq \gamma_{n+1}$ for all $n = 1, 2, \ldots; \qquad \lim_{n\to\infty} \gamma_n = 0$

b) $\ell_n \in \mathbb{Z}^+$ and $\ell_n \leq \ell_{n+1}$ for all $n = 0, 1, \ldots$, $\lim_{n\to\infty} \ell_n = \infty$

c) $\sum_{n=0}^{\infty} \gamma_{n+1} = \infty$, $\sum_{n=1}^{\infty} \gamma_n^{1+\widehat{\ell}_1} \ell_{n+1}^4 < \infty$

d) $\sum_{k=1}^{\infty} (\gamma_k - \gamma_{k+1}) \ell_{k+1}^3 < \infty$

We note that the sequences defined by $\gamma_n = \frac{1}{n+1}$ and $\ell_n = \max(1, \lfloor \ln(n) \rfloor)$ satisfies $(\overline{\mathrm{S}})$ with any $0 < \widehat{\ell}_1 < 1$.

5.2 Basic Ingredients

Let X^∞ be the infinite Cartesian product of X with itself, and denote by $\mathcal{B}(\mathsf{X}^\infty)$ the standard σ–field on X^∞. We write a generic element ξ of X^∞ as $\xi = (x, x_1, \ldots)$ where $x, x_1, \ldots$ are all elements of X. The coordinate process $\{\xi_\ell,\ \ell = 0, 1, \ldots\}$ is then simply defined by

$$\xi_0(\xi) \doteq x, \qquad \xi_\ell(\xi) \doteq x_\ell, \qquad \xi \in \mathsf{X}^\infty,\ \ell = 1, \ldots$$

We postulate the existence of a family $\{\mathbf{P}_{\theta,x},\ \theta \in \Theta,\ x \in \mathsf{X}\}$ of probability measures on $\mathcal{B}(\mathsf{X}^\infty)$ such that

$$\mathbf{P}_{\theta,x}[\xi_0 = x] = 1, \qquad \theta \in \Theta,\ x \in \mathsf{X}.$$

For technical reasons, we again assume a measurable functional dependence in θ and x:

(P0) For every $L = 1, 2, \ldots$, the mapping $\Theta \times \mathsf{X} \to \mathbb{R} : (\theta, x) \to \mathbf{P}_{\theta,x}[\xi_\ell \in B_\ell,\ \ell = 1, \ldots, L]$ is Borel measurable for all possible choices of Borel subsets $B_1, \ldots, B_L$ in $\mathcal{B}(\mathsf{X})$.

In order to define the stochastic approximation procedures, we start with a sample space Ω equipped with a σ–field of events $\mathcal{F}$. The measurable space $(\Omega, \mathcal{F})$ is assumed large enough to carry a double array of X–valued rvs $\{X_{n,\ell},\ \ell = 1, \ldots, \ell_n;\ n = 0, 1, \ldots\}$ where we take the convention that $\ell_0 = 1$. We define the Θ–valued rvs $\{\theta_n,\ n = 0, 1, \ldots\}$ through the recursion (5.1), and for convenience later, we define $X_{n+1,0} = X_{n,\ell_n}$ for all $n = 0, 1, \ldots$.

Next, we introduce the filtration $\{\mathcal{F}_n,\ n = 0, 1, \ldots\}$ on $(\Omega, \mathcal{F})$ by setting

$$\begin{aligned} \mathcal{F}_n &\doteq \sigma\{\theta_m, X_{m,\ell},\ \ell = 1, \ldots, \ell_m,\ m = 0, 1, \ldots, n\} \\ &= \sigma\{\theta_0; X_{m,\ell},\ \ell = 1, \ldots, \ell_m,\ m = 0, 1, \ldots, n\} \qquad n = 0, 1, \ldots \end{aligned}$$

where the equality follows since the rvs θ_m, $m = 1, 2, \ldots, n$, are fully determined by the rvs θ_0, $X_{0,1}$, and $X_{m+1,\ell}$, $\ell = 1, \ldots, \ell_{m+1}$, $m = 0, 1, \ldots, n-1$.

Finally, given a probability measure ν on $\mathcal{B}(\Theta \times \mathsf{X})$, we postulate the existence of a probability measure $\mathbf{P}$ on $(\Omega, \mathcal{F})$ satisfying

$$\mathbf{P}[\theta \in B, X_{0,1} \in B_1] = \nu(B \times B_1), \qquad B \in \mathcal{B}(\Theta), B_1 \in \mathcal{B}(\mathsf{X})$$

and

$$\mathbf{P}[X_{n+1,\ell} \in B_\ell,\ \ell = 1, \ldots, \ell_{n+1} | \mathcal{F}_n] = \mathbf{P}_{\theta_n, X_{n,\ell_n}}[\xi_\ell \in B_\ell,\ \ell = 1, \ldots, \ell_{n+1}] \qquad n = 0, 1, \ldots \tag{5.2}$$

for Borel subsets $B_1, \ldots, B_{\ell_{n+1}}$ in $\mathcal{B}(\mathsf{X})$. The existence of such a set-up is readily justified by the Daniell–Kolmogorov consistency theorem [69, p. 94] on $\Theta \times \mathsf{X} \times \mathsf{X}^\infty$ in the usual manner. We shall also define the one-step transition probability $P_\theta(x, A) \doteq \mathbf{P}_{x,\theta}[\xi_1 = A]$ for all $\theta \in \Theta$ and $A \in \mathcal{B}(\mathsf{X})$.

5.3 The Windowed State Process

We shall now assume a countable state space X. For a given sequence $\{\ell_{n+1}, n = 0, 1, \ldots\}$, we have defined the windowed state process $\{\bar{X}_{n+1}, n = 0, 1, \ldots\}$ by setting

$$\bar{X}_{n+1} \doteq \left(X_{n+1,0}, X_{n+1,1}, \ldots, X_{n+1,\ell_{n+1}}\right), \qquad n = 0, 1, \ldots. \tag{5.3}$$

The state process $\{\bar{X}_{n+1}, n = 0, 1, \ldots\}$ is clearly inhomogeneous. The one-step transition probabilities for the windowed process (5.3), denoted $\bar{P}_{\theta_n}$, can be defined in terms of the transition probabilities P_{θ_n} which govern each transition within the n^{th} window. Clearly, the windowed process $\{\bar{X}_n;\ n = 0, 1, 2, \ldots\}$ is Markov under our construction. We have for any $\theta_n \in \Theta$

$$\begin{aligned}
&\bar{P}_{\theta_n}\left(\bar{x}_n, \bar{x}_{n+1}\right) \\
&\quad = \bar{P}_{\theta_n}\left((x_{n,0}, x_{n,1}, \ldots, x_{n,\ell_n}), (x_{n+1,0}, x_{n+1,1}, \ldots, x_{n+1,\ell_{n+1}})\right) \\
&\quad = \mathbf{P}_{\theta_n}[\bar{X}_{n+1} = (x_{n+1,0}, x_{n+1,1}, \ldots, x_{n+1,\ell_{n+1}}) | \bar{X}_n = (x_{n,0}, x_{n,1}, \ldots, x_{n,\ell_n})] \\
&\quad = \mathbf{P}_{\theta_n}[X_{n+1,k} = x_{n+1,k}, k = 0, \ldots, \ell_{n+1} | X_{n,0} = x_{n,0}, \ldots, X_{n,\ell_n} = x_{n,\ell_n}] \\
&\quad = \mathbf{P}_{\theta_n}[X_{n+1,k} = x_{n+1,k}, k = 0, \ldots, \ell_{n+1} | X_{n,\ell_n} = x_{n,\ell_n}] \\
&\quad = 1_{\{x_{n,\ell_n} = x_{n+1,0}\}} \prod_{k=0}^{\ell_{n+1}-1} P_{\theta_n}\left(x_{n+1,k}, x_{n+1,k+1}\right), \qquad n = 0, 1, \ldots
\end{aligned}$$

Note that as expected, $\bar{P}_{\theta_n}$ only depends on the *last* point x_{n,ℓ_n} of the window and not the entire window $\bar{x}_n$.

Let us now generalize the definition of $\bar{P}_{\theta_n}$ to allow transitions from an arbitrary size window of size ℓ' to one of size ℓ. Thus, for any parameter $\theta \in \Theta$, probability of any transition from any state $\bar{x} = (x_0, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}$ to any state $\bar{y} = (y_0, y_1, \ldots, y_\ell) \in \mathsf{X}^{\ell+1}$ is given by

$$\bar{P}_\theta\left(\bar{x}, \bar{y}\right) \doteq 1_{\{x_{\ell'} = y_0\}} \prod_{k=0}^{\ell-1} P_\theta\left(y_k, y_{k+1}\right).$$

It should be obvious that for any fixed $\theta \in \Theta$ and $\ell' \in \mathbb{Z}^+$, an invariant distribution for $\bar{P}_\theta\left(\bar{x}, \bar{y}\right)$ is given by

$$\bar{\pi}_\theta(\bar{x}) = \pi_\theta(x_0) \prod_{k=0}^{\ell'-1} P_\theta\left(x_k, x_{k+1}\right)$$

where $\pi_\theta(\cdot)$ is the invariant distribution for the one-step P_θ.

As a final point of notation, if in the above we have $\ell = \ell'$ then we shall write

$$\bar{P}_{\ell,\theta}\left(\bar{x}, \bar{y}\right) \doteq 1_{\{x_\ell = y_0\}} \prod_{k=0}^{\ell-1} P_\theta\left(y_k, y_{k+1}\right), \qquad \theta \in \Theta, \tag{5.4}$$

for $\bar{x}, \bar{y} \in \mathsf{X}^{\ell+1}$ where $\bar{x} = (x_0, x_1, \ldots x_\ell)$ and $\bar{y} = (y_0, y_1, \ldots, y_\ell)$.

Assumed Ergodicity

Given this framework, we assume a generalized ergodicity for any *generic* homogeneous Markov chain $\{\bar{X}_n, n = 0, 1, \ldots\}$ governed by (5.4) for any **fixed** $\ell \in \mathbb{Z}^+$ and **fixed** $\theta \in \Theta$ in the sense that

$$\lim_{n\to\infty} \mathbf{E}_{\theta,\bar{x}_0}\left[H_\ell(\theta, \bar{X}_n)\right] = \bar{\pi}_\theta(H_{\ell,\theta}) \doteq h_\ell(\theta), \qquad \theta \in \Theta,\ \bar{x}_0 \in \mathsf{X}^{\ell+1}. \tag{5.5}$$

(Note, for this generic chain here we assume θ is held fixed and not being updated by the SA algorithm.) Additionally, we assume that

$$\lim_{\ell\to\infty} h_\ell(\theta) = h(\theta), \qquad \theta \in \Theta.$$

5.4 The Increasing Window Size SA Algorithm

Here we develop a general form of SA algorithm. The recursion we apply takes the form:

$$\theta_{n+1} = \Pi_\Theta\left\{\theta_n + \gamma_{n+1} H_{\ell_{n+1}}(\theta_n, \bar{X}_{n+1}) + \gamma_{n+1}^{1+\widehat{\ell}_1} \bar{\rho}_{n+1}(\theta_n, \bar{X}_{n+1})\right\}, \quad n = 0, 1, \ldots$$

Inserting canceling terms, we find

$$\begin{aligned}\theta_{n+1} &= \Pi_\Theta\Big\{\theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\left(H_{\ell_{n+1}}(\theta_n, \bar{X}_{n+1}) - h_{\ell_{n+1}}(\theta_n)\right) \\ &\quad + \gamma_{n+1}\left(h_{\ell_{n+1}}(\theta_n) - h(\theta_n)\right) + \gamma_{n+1}^{1+\widehat{\ell}_1} \bar{\rho}_{n+1}(\theta_n, \bar{X}_{n+1})\Big\}.\end{aligned}$$

We now lump the *bias* term in with any applied perturbation $\bar{\rho}_n$ and define the total perturbation

$$\rho_{n+1}(\theta_n, \bar{X}_{n+1}) \doteq \gamma_{n+1}^{-\widehat{\ell}_1}\left(h_{\ell_{n+1}}(\theta_n) - h(\theta_n)\right) + \bar{\rho}_{n+1}(\theta_n, \bar{X}_{n+1}), \tag{5.6}$$

Then we define the overall noise

$$\varepsilon_{n+1} \doteq H_{\ell_{n+1}}(\theta_n, \bar{X}_{n+1}) - h_{\ell_{n+1}}(\theta_n) + \gamma_{n+1}^{\widehat{\ell}_1} \rho_{n+1}(\theta_n, \bar{X}_{n+1}). \tag{5.7}$$

and we have the *increasing window size* SA algorithm

$$\begin{aligned}\theta_{n+1} &= \Pi_\Theta\{\theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\varepsilon_{n+1}\} \\ &= \theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\varepsilon_{n+1} + \gamma_{n+1} z_{n+1}, \quad n = 0, 1, \ldots.\end{aligned} \tag{5.8}$$

It should be clear from this construction that if we want the the bias term to vanish, we will require the window sizes, defined by the sequence $\{\ell_{n+1},\ n = 1, 2, \ldots\}$, to be increasing towards infinity. The trick here is to increase the window size slowly enough so the increasing amount of noise can be controlled by the the SA algorithm.

5.4.1 A Stochastic Optimization Algorithm

The above algorithm in (5.8) is given in a general form similar to the basic SA algorithm of Chapter 2 and is designed to be used with the gradient algorithms of Chapter 4. Here we review the specific optimization algorithm proposed in Chapter 4.

First we note that $\bar{\rho}_n(\cdot,\cdot) = 0$ for all $n = 1, 2, \ldots$. Then, let us designate the p components of the vector $H_\ell(\theta, \bar{x})$ as $H_{\ell,i}(\theta, \bar{x})$ for $i = 1, \ldots, p$. Thus, for any $i = 1, \ldots, p$ we have

$$\begin{aligned} H_{\ell_{n+1},i}(\theta_n, \bar{X}_{n+1}) &= -\widehat{G}_i(\theta_n, \bar{X}_{n+1}) \\ &= -\left(\frac{1}{\ell_{n+1}} \sum_{j=1}^{\ell_{n+1}} \frac{\partial f_\theta(X_{n,j})}{\partial \theta_i} \right) - L_{\theta_n,i}(X_{n+1,0}; X_{n+1,1}) \sum_{k=1}^{\ell_{n+1}} f(X_{n+1,k}) \end{aligned}$$

where the *likelihood ratio* is

$$L_{\theta_n,i}(X_{n+1,0}; X_{n+1,1}) \doteq \frac{\partial P_{\theta_n}(X_{n+1,0}, X_{n+1,1})/\partial \theta_i}{P_{\theta_n}(X_{n+1,0}, X_{n+1,1})}.$$

If we assume all the conditions of Theorem 4.6 we may define

$$h(\theta) = -\nabla J(\theta) = -\pi_\theta(\nabla f_\theta) - \pi_\theta Q_\theta g_\theta, \qquad \theta \in \Theta,$$

and if we write $h_\ell(\theta) = [h_{\ell,1}(\theta) \ldots h_{\ell,p}(\theta)]'$, we have for each $i = 1, \ldots, p$ and $\ell = 1, 2, \ldots$

$$h_{\ell,i}(\theta) \doteq \mathbf{E}_{\bar{\pi}_\theta}\left[H_{\ell,i}(\theta, \bar{X}_n) \right] = \mathbf{E}_{\bar{\pi}_\theta}\left[-\frac{\partial f_\theta}{\partial \theta_i}(X_{n,1}) - L_{\theta,i}(X_{n,0}; X_{n,1}) \sum_{k=1}^{\ell} f(X_{n,k}) \right].$$

5.4.2 Additive Form of the Driving Function

If we continue to consider the gradient algorithm defined above, observe the function $H_{\ell_{n+1}}$ is additive in the sense that for any $\theta \in \Theta$,

$$H_{\ell_{n+1}}(\theta, \bar{X}_{n+1}) = \sum_{i=1}^{\ell_{n+1}} \left(\frac{\nabla_\theta f_\theta(X_{n,i})}{\ell_{n+1}} + L_\theta(X_{n,0}; X_{n,1}) f_\theta(X_{n,i}) \right). \tag{5.9}$$

Under the conditions we will make, the norm of the likelihood ratio $L_\theta(x, y)$ is bounded for all $\theta \in \Theta$ and all $x, y \in \mathsf{X}$. Furthermore, the one-step performance is assumed dominated by $|f(x)| \le C_1 V^r(x)$ where V is defined in the assumed (D1) or (D2) as in Chapter 2. Additionally, we will assume $\|\nabla_\theta f_\theta(x)\| \le C_2 V^r(x)$. Thus we have a bound for H_ℓ which grows with $\ell = 1, 2, \ldots$:

$$\begin{aligned} \|H_\ell(\theta, \bar{x})\| &\le \sum_{i=1}^{\ell} \left(\|\nabla_\theta f_\theta(x_i)\| + \|L_\theta(x_0; x_1)\| \, |f(x_i)| \right) \\ &\le \sum_{i=1}^{\ell} C V^r(x_i), \qquad \theta \in \Theta, \quad \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}, \end{aligned}$$

for some $C < \infty$.

Consider applying the framework of Chapter 2 to this SA with increasing window sizes. Since condition (H2) is clearly violated if we use an unbounded window length sequence $\{\ell_n,\ n = 0, 1, \ldots\}$, it should be apparent that we need to make some substantial changes to adapt the earlier SA framework for this optimization problem.

5.5 General Convergence Criteria

We start with verifying (or assuming) either condition (D1) or (D2) involving $\{P_\theta, \theta \in \Theta\}$ for some function $V : \mathsf{X} \to [1, \infty)$. Then, we define the functions

$$\begin{aligned} \bar{V}_0(x) &\doteq V(x), \qquad x \in \mathsf{X}, \\ \bar{V}_\ell(\bar{x}) &\doteq \sup\{V(x_i), i = 1, \ldots, \ell\}, \qquad \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}, \quad \ell = 1, 2, \ldots \end{aligned}$$

All the conditions here are analogs of those in Chapters 2 and 3 except we have added an "over-line" to distinguish that they apply to the windowed process.

($\overline{\mathbf{D0}}$) For the sequence of functions $\bar{V}_\ell : \mathsf{X}^{\ell+1} \to [1, \infty)$ for $\ell = 0, 1, 2, \ldots$ there exists a constant $1 \le C_D < \infty$ such that

$$\mathbf{E}_{\theta,x}\left[\bar{V}_{\ell_{n+1}}(\bar{X}_{n+1})\right] \le C_D \ell_{n+1} \bar{V}_0(x), \qquad n = 0, 1, 2, \ldots$$

for all θ in Θ and x in X.

For this sequence of functions $\bar{V}_\ell : \mathsf{X}^{\ell+1} \to [1, \infty)$, $\ell = 0, 1, \ldots$, we assume the remaining conditions all hold for some constant r such that $0 < r \le \frac{1}{2(1+\widehat{\ell}_1)}$ where the real constant $\widehat{\ell}_1$ lies in the interval $(0, 1)$ and satisfies ($\overline{\text{S}}$).

($\overline{\mathbf{H2}}$) There exists constants $C_H < \infty$ and $C_\rho < \infty$ such that for all $\ell = 1, 2, \ldots$

$$\sup_{\theta \in \Theta} \|H_\ell(\theta, \bar{x})\| \le C_H \sum_{i=1}^{\ell} V_\ell^r(x_i) \le C_H\ \ell\ \bar{V}_\ell^r(\bar{x}),$$

for all $\bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}$; and

$$\sup_{\theta \in \Theta} \|\rho_n(\theta, \bar{x})\| \le C_\rho\ \ell_n\ \bar{V}_{\ell_n}^r(\bar{x}), \qquad n = 1, 2, \ldots$$

for all $\bar{x} = (x_0, x_1, \ldots, x_{\ell_n}) \in \mathsf{X}^{\ell_n+1}$.

($\overline{\mathbf{P1}}$) For any $\ell = 1, 2, \ldots$, the following series converges

$$\nu_{\ell,\theta}(\bar{x}) \doteq \sum_{n=0}^{\infty} \left(\int_{\mathsf{X}^{\ell+1}} \bar{P}_{\ell,\theta}^n(\bar{x}, d\bar{y}) H_{\ell,\theta}(\bar{y}) - h_\ell(\theta) \right) < \infty, \qquad (\theta, \bar{x}) \in \Theta \times \mathsf{X}^{\ell+1}.$$

We identify $\nu_{\ell,\theta}(\bar{x})$ as the solution to the Poisson equation associated with $H_\ell(\theta, \cdot)$,

$$H_{\ell,\theta}(\bar{x}) - h_\ell(\theta) = \nu_{\ell,\theta}(\bar{x}) - \int_{\mathsf{X}^{\ell+1}} \bar{P}_{\ell,\theta}(\bar{x}, d\bar{y}) \nu_{\ell,\theta}(\bar{y}), \qquad \bar{x} \in \mathsf{X}^{\ell+1},\ \theta \in \Theta.$$

Recall our convention that $H_\ell(\theta, \cdot) = H_{\ell,\theta}(\cdot)$ and $h_\ell(\theta) = \bar{\pi}_\theta(H_{\ell,\theta}(\cdot))$.

($\overline{\textbf{P2}}$) There exists a constant $C_\nu < \infty$ such that for all $\ell = 1, 2, \ldots$:

$$\begin{aligned}
\|\nu_{\ell,\theta}(\bar{x})\| &\leq C_\nu \; \ell \; \bar{V}_\ell^r(\bar{x}), \qquad \text{for all } \theta \in \Theta, \quad \bar{x} \in \mathsf{X}^{\ell+1}, \\
\left\|\bar{P}_\theta \nu_{\ell,\theta}(\bar{x})\right\| &\leq C_\nu \; \ell \; \bar{V}_0^r(x_{\ell'}), \qquad \text{for all } \theta \in \Theta, \quad \bar{x} = (x_0, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}, \\
& \qquad\qquad \ell' = 1, 2, \ldots \qquad \text{arbitrary.}
\end{aligned}$$

($\overline{\textbf{P3}}$) There exists a constant $C_\delta < \infty$ such that for all $\ell = 1, 2, \ldots$

$$\begin{aligned}
\left\|\bar{P}_\theta \nu_{\ell,\theta}(\bar{x}) - \bar{P}_{\theta'} \nu_{\ell,\theta'}(\bar{x})\right\| &\leq C_\delta \; \ell^2 \; \bar{V}_0^r(x_{\ell'}) \, \|\theta - \theta'\|^{\widehat{\ell}_1}, \\
& \text{for all } \theta, \theta' \in \Theta, \quad \bar{x} = (x_0, x_1, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}, \\
& \qquad \ell' = 1, 2, \ldots \qquad \text{arbitrary,}
\end{aligned}$$

and $\widehat{\ell}_1 \in (0, 1)$ determined by ($\overline{\text{S}}$).

Note, the arbitrariness of ℓ' is due to the Markov property of the conditional expectation; i.e. all that it depends on is the most recent sample within the window no matter how long the window happens to be.

5.6 Decomposition of the Increasing Window SA Algorithm

Previously in Section 5.4 we showed the stochastic optimization algorithm takes the form of an SA. Our main goal now is to study the noise sequence in a manner that allows us to show condition (KC4) in the Kushner-Clark Lemma. We decompose the noise so we can break this large problem up into several manageable pieces. The decomposition is adapted from BMP's technique used in Chapter 2 except in this case here the increasing window length creates several difficulties we must attend to.

Earlier in (5.7) we derived the sequence of *noise* terms $\{\varepsilon_{k+1}, k = 0, 1, \ldots\}$ given by

$$\varepsilon_{k+1} \doteq H_{\ell_{k+1}}(\theta_k, \bar{X}_{k+1}) - h_{\ell_{k+1}}(\theta_k) + \gamma_{k+1}^{\widehat{\ell}_1} \rho_{k+1}(\theta_k, \bar{X}_{k+1}).$$

Assuming condition ($\overline{\text{P1}}$) so that the solution to the Poisson equation exists, then

$$\begin{aligned}
\varepsilon_{k+1} &= \left\{\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\ell_{k+1},\theta_k} \nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})\right\} + \gamma_{k+1} \rho_{k+1}(\theta_k, \bar{X}_{k+1}) \\
&= \left\{\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k} \nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right\} \\
&+ \left\{\bar{P}_{\theta_k} \nu_{\ell_{k+1},\theta_k}(\bar{X}_k) - \bar{P}_{\ell_{k+1},\theta_k} \nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})\right\} + \gamma_{k+1} \rho_{k+1}(\theta_k, \bar{X}_{k+1}) \\
&= \left\{\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k} \nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right\} \\
&+ \left\{\bar{P}_{\theta_k} \nu_{\ell_{k+1},\theta_k}(\bar{X}_k) - \bar{P}_{\theta_k} \nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1})\right\} \\
&+ \left\{\bar{P}_{\theta_k} \nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\ell_{k+1},\theta_k} \nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})\right\} + \gamma_{k+1} \rho_{k+1}(\theta_k, \bar{X}_{k+1})
\end{aligned}$$

If $m < n$, the step-size weighted sum of the noise terms is formed and then rearranged:

$$\begin{aligned}
\sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1} &= \sum_{k=m}^{n-1} \gamma_{k+1} \left\{ \nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k) \right\} \\
&+ \sum_{k=m}^{n-1} \gamma_{k+1} \left\{ \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k) - \bar{P}_{\theta_k}\nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1}) \right\} \\
&+ \sum_{k=m}^{n-1} \gamma_{k+1} \left\{ \bar{P}_{\theta_k}\nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\ell_{k+1},\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) \right\} \\
&+ \sum_{k=m}^{n-1} \gamma_{k+1}^2 \rho_{k+1}(\theta_k, \bar{X}_{k+1})
\end{aligned}$$

Rearranging the terms

$$\begin{aligned}
\sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1} &= \sum_{k=m}^{n-1} \gamma_{k+1} \left\{ \nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k) \right\} \\
&+ \sum_{k=m+1}^{n-1} \gamma_{k+1} \left\{ \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k) - \bar{P}_{\theta_{k-1}}\nu_{\ell_{k+1},\theta_{k-1}}(\bar{X}_k) \right\} \\
&+ \gamma_{m+1}\bar{P}_{\theta_m}\nu_{\ell_{m+1},\theta_m}(\bar{X}_m) - \gamma_n \bar{P}_{\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n) \\
&+ \sum_{k=m}^{n-1} \gamma_{k+1} \left\{ \bar{P}_{\theta_k}\nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\ell_{k+1},\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) \right\} \\
&+ \sum_{k=m+1}^{n-1} \left(\gamma_{k+1} - \gamma_k \right) \bar{P}_{\theta_{k-1}}\nu_{\ell_{k+1},\theta_{k-1}}(\bar{X}_k) \\
&+ \sum_{k=m}^{n-1} \gamma_{k+1}^2 \rho_{k+1}(\theta_k, \bar{X}_{k+1})
\end{aligned}$$

Now define for $k = m, m+1, \ldots, n-1$:

$$\begin{aligned}
\varepsilon_{k+1}^{(1)} &\doteq \nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k) \\
\varepsilon_{k+1}^{(2)} &\doteq \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k) - \bar{P}_{\theta_{k-1}}\nu_{\ell_{k+1},\theta_{k-1}}(\bar{X}_k) \\
\varepsilon_{k+1}^{(3)} &\doteq \frac{\gamma_{k+1} - \gamma_k}{\gamma_{k+1}} \bar{P}_{\theta_{k-1}}\nu_{\ell_{k+1},\theta_{k-1}}(\bar{X}_k) \\
\varepsilon_{k+1}^{(4)} &\doteq \gamma_{k+1}^{\widehat{\ell}_1} \rho_{k+1}(\theta_k, \bar{X}_{k+1}) \\
\varepsilon_{k+1}^{(5)} &\doteq \bar{P}_{\theta_k}\nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\ell_{k+1},\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) \\
\eta_{m;n} &\doteq \gamma_{m+1}\bar{P}_{\ell_{m+1},\theta_m}\nu_{\ell_{m+1},\theta_m}(\bar{X}_m) - \gamma_n \bar{P}_{\ell_{n+1},\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n).
\end{aligned}$$

Thus, for $m < n$ we have the decomposition

$$\begin{aligned}
\sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1} &= \sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(1)} + \sum_{k=m+1}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(2)} + \sum_{k=m+1}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(3)} \\
&+ \sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(4)} + \sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1}^{(5)} + \eta_{m;n} \qquad (5.10)
\end{aligned}$$

5.7 New Lemmas to Bound the Noise Terms

In this section, we prove a bound for all of the noise terms in the decomposition above.

Lemma 5.1 *Assume* ($\overline{\text{D0}}$)*,* ($\overline{\text{P1}}$)*,* ($\overline{\text{P2}}$) *hold for any positive constant* $r \le \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies* ($\overline{\text{S}}$)*.*

1. *There exists a constant* $A_1 < \infty$ *such that for each* $m = 1, 2, \ldots$

$$\mathbf{E}_{\theta,x}\left[\sup_{n\le m}\left\|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}\right\|^2\right] \le A_1\bar{V}_0(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2\ell_{k+1}^3, \qquad x\in\mathsf{X}, \quad \theta\in\Theta.$$

   Moreover, $A_1 \le 4pC_\nu^2 C_D$*.*

2. *The series* $\lim_{n\to\infty}\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}^{(1)}$ *converges* $\mathbf{P}_{\theta,x}$*-a.s. to a finite rv.*

Proof: Consider the sum

$$\bar{M}_n = \sum_{k=0}^{n-1}\gamma_{k+1}\left\{\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right\}, \qquad n = 1, 2, \ldots$$

which is a vector martingale since (by the Markov property)

$$\mathbf{E}\left[\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})|\mathcal{F}_k\right] = \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k).$$

The vector $\bar{M}_n$ is a p-dimensional vector, and although convergence results exist for vector martingales [76], we find it simpler to consider each of the p components separately by defining the i^{th} component vector as

$$M_n^{(i)} = \sum_{k=0}^{n-1}\gamma_{k+1}\left\{\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right\}^{(i)}, \qquad n = 1, 2, \ldots,$$

For brevity, let us now drop the $^{(i)}$ in this definition and consider any of the p components of the vector martingale as:

$$M_n = \sum_{k=0}^{n-1}\gamma_{k+1}\left\{\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right\}, \qquad n = 1, 2, \ldots$$

Clearly, each component of M_n above also has the martingale property.

Incremental orthogonality and Pythagoras formula [108, p.110] yield

$$\begin{aligned}\mathbf{E}_{\theta,x}\left[M_n^2\right] &= \mathbf{E}_{\theta,x}\left[M_1^2\right] + \sum_{k=2}^{n}\mathbf{E}_{\theta,x}\left[(M_k - M_{k-1})^2\right] \\ &= \sum_{k=0}^{n-1}\gamma_{k+1}^2\mathbf{E}_{\theta,x}\left[\left(\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right)^2\right]\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[\mathbf{E}\left[\left(\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right)^2 |\mathcal{F}_k\right]\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[\mathbf{E}\left[\left(\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})\right)^2 |\mathcal{F}_k\right] - \left(\bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k)\right)^2\right] \\
&\leq \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[\left(\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})\right)^2\right] \\
&\leq C_\nu^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \ell_{k+1}^2 \mathbf{E}_{\theta,x}\left[\bar{V}_{\ell_{k+1}}^{2r}(\bar{X}_{k+1})\right] \\
&\leq C_\nu^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \ell_{k+1}^2 \mathbf{E}_{\theta,x}\left[\bar{V}_{\ell_{k+1}}^{1}(\bar{X}_{k+1})\right]
\end{aligned}
$$

where we have used ($\overline{\text{P2}}$) in the second to last line. The last line follows since $2r \leq \frac{1}{1+\widehat{\ell}_1} \leq 1$. Applying ($\overline{\text{D0}}$) to the last line we find

$$
\mathbf{E}_{\theta,x}\left[M_n^2\right] \leq C_\nu^2 C_D \bar{V}_0(x) \sum_{k=0}^{n-1} \gamma_{k+1}^2 \ell_{k+1}^3
$$

The bound in the first part of the lemma follows from Doob's inequality

$$
\begin{aligned}
\mathbf{E}_{\theta,x}\left[\sup_{n\leq m} \left\|\bar{M}_n\right\|^2\right] &= \mathbf{E}_{\theta,x}\left[\sup_{n\leq m} \sum_{i=1}^{p} \left(\bar{M}_n^{(i)}\right)^2\right] \\
&\leq \sum_{i=1}^{p} \mathbf{E}_{\theta,x}\left[\sup_{n\leq m} \left(M_n^{(i)}\right)^2\right] \\
&\leq \sum_{i=1}^{p} 4 \sup_{n\leq m} \mathbf{E}_{\theta,x}\left[\left(M_n^{(i)}\right)^2\right] \\
&\leq 4pC_\nu^2 C_D \bar{V}_0(x) \sum_{k=0}^{m-1} \gamma_{k+1}^2 \ell_{k+1}^3, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.
\end{aligned}
$$

Under ($\overline{\text{S}}$), $\sum_{k=0}^{\infty} \gamma_{k+1}^{1+\widehat{\ell}_1} \ell_{k+1}^3 < \infty$ and it then follows that $\sum_{k=0}^{\infty} \gamma_{k+1}^2 \ell_{k+1}^3 < \infty$ since $\gamma_k \downarrow 0$ and there exists a k' such that $\gamma_{k'} < 1$, hence $\gamma_k^2 \leq \gamma_k^{1+\widehat{\ell}_1}$ for all $k \geq k'$.

For the convergence properties in the second part of the lemma, we note that $\sum_{k=0}^{\infty} \gamma_{k+1}^2 \ell_{k+1}^3 < \infty$ which implies that each component martingale of the vector martingale converges a.s. to a finite random variable (as well as converging in L^2) since it is bounded in L^2 [108]. ■

Lemma 5.2 *Assume* ($\overline{\text{D0}}$), ($\overline{\text{H2}}$), ($\overline{\text{P1}}$), ($\overline{\text{P3}}$) *for any positive* $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *is determined from* ($\overline{\text{S}}$). *There exists a constant* $A_2 < \infty$ *such that for all* $m = 1, 2, \ldots$,

$$
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_{k+1} \left\|\varepsilon_{k+1}^{(2)}\right\|\right)^2\right] \leq A_2 \bar{V}_0(x) \left(\sum_{k=1}^{m-1} \gamma_k^{1+\widehat{\ell}_1} \ell_{k+1}^{5/2+\widehat{\ell}_1}\right)^2, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.
$$

Moreover, $A_2 \leq 4C_\delta^2 \left(C_H + \gamma_1 C_\rho\right)^{2\widehat{\ell}_1} C_D$.

Proof: Under ($\overline{\mathrm{P3}}$),

$$\begin{aligned} &\left\|\bar{P}_{,\theta}\nu_{\ell_{n+1},\theta}(\bar{x}) - \bar{P}_{\theta'}\nu_{\ell_{n+1},\theta'}(\bar{x})\right\| \\ &\leq C_\delta \ell_{n+1}^2 \bar{V}_0^r(x_{\ell_n}) \left\|\theta - \theta'\right\|^{\widehat{\ell}_1} \\ &\leq C_\delta \ell_{n+1}^2 \bar{V}_{\ell_n}^r(\bar{x}) \left\|\theta - \theta'\right\|^{\widehat{\ell}_1}, \quad \bar{x} = (x_0, \ldots, x_{\ell_n}) \in \mathsf{X}^{\ell_n+1}, \quad \theta, \theta' \in \Theta. \end{aligned} \tag{5.11}$$

Also, the nearest point projection term is bounded by

$$\|z_k\| \leq \left\|H_{\ell_k}(\theta_{k-1}, \bar{X}_k) + \gamma_k \rho_k(\theta_{k-1}, \bar{X}_k)\right\|$$

which follows since $\theta_k \in \Theta$ and, at the very least, the projection term can return the iterate to this point so $\theta_{k+1} = \theta_k \in \Theta$. Hence for $k = 1, 2, \ldots$ we have from ($\overline{\mathrm{H2}}$) and the definition of the SA that

$$\begin{aligned} \|\theta_k - \theta_{k-1}\| &\leq \gamma_k \left\|H_{\ell_k}(\theta_{k-1}, \bar{X}_k) + \gamma_k^{\widehat{\ell}_1} \rho_k(\theta_{k-1}, \bar{X}_k) + z_k\right\| \\ &\leq 2\gamma_k \left\|H_{\ell_k}(\theta_{k-1}, \bar{X}_k) + \gamma_k^{\widehat{\ell}_1} \rho_k(\theta_{k-1}, \bar{X}_k)\right\| \\ &\leq 2C_H \gamma_k \ell_k \bar{V}_{\ell_k}^r(\bar{X}_k) + 2C_\rho \gamma_k^{1+\widehat{\ell}_1} \ell_k \bar{V}_{\ell_k}^r(\bar{X}_k) \\ &\leq 2\left(C_H + \gamma_1^{\widehat{\ell}_1} C_\rho\right) \gamma_k \ell_k \bar{V}_{\ell_k}^r(\bar{X}_k) \end{aligned} \tag{5.12}$$

Thus by (5.11) and (5.12),

$$\begin{aligned} &\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_{k+1} \left\|\varepsilon_{k+1}^{(2)}\right\|\right)^2\right] \\ &= \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_{k+1} \left\|\bar{P}_{\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_k) - \bar{P}_{\theta_{k-1}}\nu_{\ell_{k+1},\theta_{k-1}}(\bar{X}_k)\right\|\right)^2\right] \\ &\leq \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_{k+1} \ell_{k+1}^2 C_\delta \bar{V}_{\ell_k}^r(\bar{X}_k) \left\|\theta_k - \theta_{k-1}\right\|^{\widehat{\ell}_1}\right)^2\right] \\ &\leq 4\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_{k+1} \ell_{k+1}^2 C_\delta \bar{V}_{\ell_k}^r(\bar{X}_k) \left(C_H + \gamma_1 C_\rho\right)^{\widehat{\ell}_1} \gamma_k^{\widehat{\ell}_1} \ell_k^{\widehat{\ell}_1} \bar{V}_{\ell_k}^{r\widehat{\ell}_1}(\bar{X}_k)\right)^2\right] \\ &= 4C_\delta^2 \left(C_H + \gamma_1 C_\rho\right)^{2\widehat{\ell}_1} \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_k^{1+\widehat{\ell}_1} \ell_{k+1}^{2+\widehat{\ell}_1} V_{\ell_k}^{r(1+\widehat{\ell}_1)}(\bar{X}_k)\right)^2\right] \\ &\leq 4C_\delta^2 \left(C_H + \gamma_1 C_\rho\right)^{2\widehat{\ell}_1} \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_k^{1+\widehat{\ell}_1} \ell_{k+1}^{2+\widehat{\ell}_1} \bar{V}_{\ell_k}^{1/2}(\bar{X}_k)\right)^2\right]. \end{aligned}$$

since $r(1 + \widehat{\ell}_1) \leq 1/2$. By treating the sum as an inner product and applying the Schwarz inequality [50, p. 2] to the last line

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1} \gamma_k \left\|\varepsilon_{k+1}^{(2)}\right\|\right)^2\right]$$

$$
\begin{aligned}
&\leq 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\ell_{k+1}^{5/2+\widehat{\ell}_1}\right)\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\ell_{k+1}^{3/2+\widehat{\ell}_1}\bar{V}_{\ell_k}(\bar{X}_k)\right]\\
&\leq 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\left(\sum_{k=1}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}^{5/2+\widehat{\ell}_1}\right)\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\ell_{k+1}^{3/2+\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[\bar{V}_{\ell_k}(\bar{X}_k)\right]\\
&\leq 4C_\delta^2\left(C_H+\gamma_1 C_\rho\right)^{2\widehat{\ell}_1}\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\ell_{k+1}^{5/2+\widehat{\ell}_1}\right)C_D\bar{V}_0(x)\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\ell_{k+1}^{5/2+\widehat{\ell}_1}.
\end{aligned}
$$

■

Lemma 5.3 *Assume* ($\overline{\text{D0}}$), ($\overline{\text{P1}}$), ($\overline{\text{P2}}$) *for any positive constant* $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies* ($\overline{\text{S}}$). *There exists a constant* $A_3 < \infty$ *such that for all* $m = 1, 2, \ldots,$,

$$
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(3)}\right\|\right)^2\right] \leq A_3\gamma_1\bar{V}_0(x)\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\ell_{k+1}^3, \qquad x\in\mathsf{X},\quad \theta\in\Theta.
$$

Moreover, $A_3 \leq C_\nu^2 C_D$.

Proof: Applying ($\overline{\text{P2}}$)

$$
\begin{aligned}
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(3)}\right\|\right)^2\right] &= \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\left\|\bar{P}_{\theta_{k-1}}\nu_{\ell_{k+1},\theta_{k-1}}(\bar{X}_k))\right\|\right)^2\right]\\
&\leq \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})C_\nu\ell_{k+1}\bar{V}_{\ell_k}^r(\bar{X}_k)\right)^2\right]
\end{aligned}
$$

Next, the Schwarz inequality yields

$$
\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(3)}\right\|\right)^2\right]\\
&\leq C_\nu^2\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\right)\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\ell_{k+1}^2\bar{V}_{\ell_k}^{2r}(\bar{X}_k)\right]\\
&\leq C_\nu^2\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\right)\sum_{k=1}^{m-1}\left(\gamma_k-\gamma_{k+1}\right)\ell_{k+1}^2\mathbf{E}_{\theta,x}\left[\bar{V}_{\ell_k}^1(\bar{X}_k)\right]\\
&\leq C_\nu^2\gamma_1\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\ell_{k+1}^3 C_D\bar{V}_0(x)
\end{aligned}
$$

■

Lemma 5.4 *Assume* ($\overline{\text{D0}}$), ($\overline{\text{P1}}$), ($\overline{\text{H2}}$) *for any positive constant* $r \leq \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies* ($\overline{\text{S}}$). *There exists a constant* $A_4 < \infty$ *such that for all* $m = 1, 2, \ldots,$

$$
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon_{k+1}^{(4)}\right\|\right)^2\right] \leq A_4\bar{V}_0(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\ell_{k+1}^{3/2}\right)^2, \qquad x\in\mathsf{X},\quad \theta\in\Theta.
$$

Moreover, $A_4 \leq C_D C_\rho^2$.

Proof: First we have from ($\overline{\text{H2}}$):

$$\gamma_{k+1}\left\|\varepsilon^{(4)}_{k+1}\right\| = \gamma_{k+1}^{1+\widehat{\ell}_1}\left\|\rho_k(\theta_k,\bar{X}_{k+1})\right\| \le \gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}C_\rho\bar{V}^r_{\ell_{k+1}}(\bar{X}_{k+1})$$

Hence,

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon^{(4)}_{k+1}\right\|\right)^2\right]\\
&\le \; C_\rho^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\ell_{k+1}\gamma_{k+1}^{1+\widehat{\ell}_1}\bar{V}^r_{\ell_{k+1}}(\bar{X}_{k+1})\right)^2\right]\\
&\le \; C_\rho^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}^{3/2}\right)\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}^{1/2}\bar{V}^{2r}_{\ell_{k+1}}(\bar{X}_{k+1})\right], \qquad m=1,2,\ldots
\end{aligned}$$

where the last line follows from the Schwarz inequality. We have $r(1+\widehat{\ell}_1)\le 1/2$, so for $m=1,2,\ldots$

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon^{(4)}_{k+1}\right\|\right)^2\right] &\le C_\rho^2\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}^{3/2}\right)\mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}^{1/2}\bar{V}^1_{\ell_{k+1}}(\bar{X}_{k+1})\right]\\
&\le C_\rho^2\left(\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}^{3/2}\right)C_D\bar{V}_0(x)\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\ell_{k+1}^{3/2}.
\end{aligned}$$

■

Lemma 5.5 *Assume ($\overline{\text{D0}}$), ($\overline{\text{P1}}$), ($\overline{\text{P2}}$), ($\overline{\text{S}}$) for any positive constant $r\le 1$.*

Then, there exists a constant $A_5<\infty$ such that for all $m=1,2,\ldots$,

$$\mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon^{(5)}_{k+1}\right\|\right] \le A_5\bar{V}_0(x)\sum_{k=0}^{m-1}\gamma_{k+1}\ell_{k+2}^2 1_{\{\ell_{k+1}\neq\ell_{k+2}\}}, \qquad x\in\mathsf{X}, \quad \theta\in\Theta.$$

If we further assume the sequences $\{\gamma_n, n=1,2,\ldots\}$ and $\{\ell_n, n=1,2,\ldots\}$ are defined by

$$\begin{aligned}
\gamma_n &= \frac{1}{n+1}, \qquad n=1,2,\ldots\\
\ell_n &= \max\left(1,\lfloor\ln(n)\rfloor\right), \qquad n=1,2,3,\ldots,
\end{aligned}$$

then, there exists a constant $A_5'<\infty$ such that for all $m=1,2,\ldots$,

$$\mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1}\gamma_{k+1}\left\|\varepsilon^{(5)}_{k+1}\right\|\right] \le A_5'\bar{V}_0(x), \qquad x\in\mathsf{X}, \quad \theta\in\Theta.$$

Proof: Recall the term

$$\varepsilon^{(5)}_{k+1} \doteq \bar{P}_{\theta_k}\nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\ell_{k+1},\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1}).$$

This is simply the difference of the two Poisson equation solutions corresponding to forcing functions $H_{\ell_{k+1}}$ and $H_{\ell_{k+2}}$ as seen through the corresponding one-step expectation operator under $\bar{P}$. This difference, if nonzero, is cause by a difference in the length of the observation windows. For any k such that $\ell_{k+1} = \ell_{k+2}$ then $\varepsilon_{k+1}^{(5)} = 0$.

Therefore, we have

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1} \gamma_{k+1} \left\|\varepsilon_{k+1}^{(5)}\right\|\right] \\
&= \mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1} \gamma_{k+1} 1_{\{\ell_{k+1}\neq\ell_{k+2}\}} \left\|\bar{P}_{\theta_k}\nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1}) - \bar{P}_{\ell_{k+1},\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})\right\|\right] \\
&\leq \sum_{k=0}^{m-1} \gamma_{k+1} 1_{\{\ell_{k+1}\neq\ell_{k+2}\}} \mathbf{E}_{\theta,x}\left[\left\|\bar{P}_{\theta_k}\nu_{\ell_{k+2},\theta_k}(\bar{X}_{k+1})\right\| + \left\|\bar{P}_{\ell_{k+1},\theta_k}\nu_{\ell_{k+1},\theta_k}(\bar{X}_{k+1})\right\|\right] \\
&\leq \sum_{k=0}^{m-1} \gamma_{k+1} 1_{\{\ell_{k+1}\neq\ell_{k+2}\}} \mathbf{E}_{\theta,x}\left[C_\nu \ell_{k+2} \bar{V}_{\ell_{k+1}}^r(X_{k+1}) + C_\nu \ell_{k+1} \bar{V}_{\ell_{k+1}}^r(X_{k+1})\right] \\
&\leq 2C_\nu \sum_{k=0}^{m-1} \gamma_{k+1} \ell_{k+2} 1_{\{\ell_{k+1}\neq\ell_{k+2}\}} \mathbf{E}_{\theta,x}\left[\bar{V}_{\ell_{k+1}}^r(X_{k+1})\right] \\
&\leq 2C_\nu \bar{V}_0(x) \sum_{k=0}^{m-1} \gamma_{k+1} \ell_{k+2} 1_{\{\ell_{k+1}\neq\ell_{k+2}\}} \ell_{k+1} \\
&\leq 2C_\nu \bar{V}_0(x) \sum_{k=0}^{m-1} \gamma_{k+1} \ell_{k+2}^2 1_{\{\ell_{k+1}\neq\ell_{k+2}\}}, \qquad x \in \mathsf{X},\ \theta \in \Theta.
\end{aligned}$$

For the second part we look at the series

$$\begin{aligned}
\sum_{k=0}^{m-1} \gamma_{k+1} \ell_{k+2}^2 1_{\{\ell_{k+1}\neq\ell_{k+2}\}} &= \sum_{n=1}^{m} \gamma_n \ell_{n+1}^2 1_{\{\ell_n\neq\ell_{n+1}\}} \\
&= \sum_{n=1}^{m} \frac{(\max(1, \lfloor\ln(n+1)\rfloor))^2}{n+1} 1_{\{\ell_n\neq\ell_{n+1}\}}
\end{aligned}$$

The convergence of the series is determined by the tail so we may discard the first few terms of the series to determine convergence. If we start at $n = 10$, the summands (excluding the indicator function) are all decreasing and noting that $\ln(3) \approx 1.0986123$, we can also drop the max with 1 operator. Thus,

$$\begin{aligned}
\sum_{n=10}^{\infty} \frac{(\lfloor\ln(n+1)\rfloor)^2}{n+1} 1_{\{\ell_n\neq\ell_{n+1}\}} &\leq \sum_{n=10}^{\infty} \frac{\ln^2(n+1)}{n+1} 1_{\{\lfloor\ln(n)\rfloor\neq\lfloor\ln(n+1)\rfloor\}} \\
&\leq \sum_{\substack{n\in\mathbb{R}: \\ \ln(n+1)\in\{11,12,13,\ldots\}}} \frac{\ln^2(n+1)}{n+1} 1_{\{\lfloor\ln(n)\rfloor\neq\lfloor\ln(n+1)\rfloor\}} \\
&= \sum_{\substack{n\in\mathbb{R}: \\ \ln(n+1)\in\{11,12,13,\ldots\}}} \frac{\ln^2(n+1)}{n+1}.
\end{aligned}$$

Now substitute $m = \ln(n+1)$ so that $n+1 = e^m$,

$$\sum_{\substack{n\in\mathbb{R}:\\ \ln(n+1)\in\{11,12,13,\ldots\}}} \frac{\ln^2(n+1)}{n+1} = \sum_{m=11}^{\infty} m^2 e^{-m} < \infty.$$

■

Lemma 5.6 *Assume* ($\overline{\mathrm{D0}}$), ($\overline{\mathrm{P1}}$), ($\overline{\mathrm{P2}}$) *for any positive constant* $r \le \frac{1}{2(1+\widehat{\ell}_1)}$ *where the positive constant* $0 < \widehat{\ell}_1 < 1$ *satisfies* ($\overline{\mathrm{S}}$).

1. *There exists a constant* $A_6 < \infty$ *such that for each* $m = 1, 2, \ldots$,

$$\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} \|\eta_{0;n}\|^2\right] \le A_6 \bar{V}_0(x) \sum_{k=0}^{m-1} \gamma_{k+1}^2, \qquad x \in \mathsf{X}, \quad \theta \in \Theta.$$

   Moreover, $A_6 \le 4C_D C_\nu^2$.

2. *As* $n \to \infty$ *we have that* $\eta_{0;n}$ *converges* $\mathbf{P}_{\theta,x}$*-a.s.*

Proof: Recall that $\eta_{0;n} \doteq \gamma_1 \bar{P}_{\theta_0} \nu_{\ell_1,\theta_0}(\bar{X}_0) - \gamma_n \bar{P}_{\theta_{n-1}} \nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)$ for $n = 1, 2, \ldots$.
First we have $\bar{X}_0 = x$ a.s. and under ($\overline{\mathrm{P2}}$)

$$\left\|\gamma_1 \bar{P}_{\theta_0} \nu_{\ell_1,\theta_0}(x)\right\|^2 \le \gamma_1^2 \ell_1^2 C_\nu^2 \bar{V}_0^{2r}(x) \le \gamma_1^2 \ell_1^2 C_\nu^2 \bar{V}_0^1(x)$$

Also, for each $m = 1, 2, \ldots$

$$\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} \left\|\gamma_n \bar{P}_{\theta_{n-1}} \nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2\right] \le C_\nu^2 \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} \gamma_n^2 \ell_{n+1}^2 \bar{V}^{2r}(\bar{X}_n)\right]$$

Thus,

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} \|\eta_{0;n}\|^2\right] \\
&= \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} \left\|\gamma_1 \bar{P}_{\theta_0} \nu_{\ell_1,\theta_0}(\bar{X}_0) - \gamma_n \bar{P}_{\theta_{n-1}} \nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2\right] \\
&\le \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} \left(2\left\|\gamma_n \bar{P}_{\theta_{n-1}} \nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2 + 2\left\|\gamma_1 \bar{P}_{\theta_0} \nu_{\ell_1,\theta_0}(x)\right\|^2\right)\right] \\
&\le 2\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m} \gamma_k^2 \ell_{k+1}^2 \bar{V}_{\ell_k}^{2r}(\bar{X}_k) C_\nu^2\right] + 2\gamma_1^2 \ell_1^2 C_\nu^2 \bar{V}_0^{2r}(x) \\
&\le 2C_\nu^2 \sum_{k=1}^{m} \gamma_k^2 \ell_{k+1}^2 \mathbf{E}_{\theta,x}\left[\bar{V}_{\ell_k}^1(\bar{X}_k)\right] + 2\gamma_1^2 \ell_1^2 C_\nu^2 \bar{V}_0^1(x) \\
&\le 2C_D C_\nu^2 \bar{V}_0(x) \sum_{k=1}^{m} \gamma_k^2 \ell_{k+1}^3 + 2\gamma_1^2 \ell_1^2 C_\nu^2 \bar{V}_0^1(x) \\
&\le 4C_D C_\nu^2 \bar{V}_0(x) \sum_{k=1}^{m} \gamma_k^2 \ell_{k+1}^3, \qquad m = 1, 2, \ldots
\end{aligned}$$

and recalling $C_D \geq 1$ for the last line.

To prove the lemma's second conclusion, we have for each $n = 1, 2, \ldots$

$$\begin{aligned}\mathbf{E}_{\theta,x}\left[\left\|\gamma_n \bar{P}_{\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2\right] &\leq \gamma_n^2\ell_{n+1}^2 C_\nu^2 \mathbf{E}_{\theta,x}\left[\bar{V}^{2r}(\bar{X}_n)\right] \\ &\leq \gamma_n^2\ell_{n+1}^3 C_\nu^2 C_D \bar{V}_0(x), \qquad x \in \mathsf{X}.\end{aligned}$$

Therefore (and by the Monotone Convergence Theorem)

$$\begin{aligned}&\mathbf{E}_{\theta,x}\left[\sum_{n=1}^{\infty}\left\|\gamma_n \bar{P}_{\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2\right] \\ &= \lim_{m\to\infty}\mathbf{E}_{\theta,x}\left[\sum_{n=1}^{m}\left\|\gamma_n \bar{P}_{\ell_{n+1},\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2\right] \\ &\leq C_\nu^2 C_D \bar{V}_0(x)\sum_{n=1}^{\infty}\gamma_n^2\ell_{n+1}^3 < \infty, \qquad x \in \mathsf{X},\end{aligned}$$

This implies the sum $\sum_{n=0}^{\infty}\left\|\gamma_n \bar{P}_{\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2$ converges to a finite rv $\mathbf{P}_{\theta,x}$-a.s. Hence, $\lim_{n\to\infty}\left\|\gamma_n \bar{P}_{\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\|^2 = 0$ $\mathbf{P}_{\theta,x}$-a.s. and thus

$$\lim_{n\to\infty}\left\|\gamma_n \bar{P}_{\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\| = 0, \qquad \mathbf{P}_{\theta,x} - a.s.$$

Therefore,

$$\lim_{n\to\infty}\left\|\eta_{0;n} - \gamma_1\bar{P}_{\theta_0}\nu_{\ell_1,\theta_0}(x)\right\| = \lim_{n\to\infty}\left\|\gamma_n \bar{P}_{\theta_{n-1}}\nu_{\ell_{n+1},\theta_{n-1}}(\bar{X}_n)\right\| = 0, \qquad \mathbf{P}_{\theta,x} - a.s.$$

■

5.8 Almost Sure Convergence of the Increasing Window SA

In Chapter 2 we showed convergence via the Kushner-Clark Lemma based on the outcome of a similar set of lemmas which bounded the terms in the decomposition there. Based on the conditions we've made in the above lemmas, we have immediately that condition (KC4) of the Kushner Clark Lemma holds for the SA with increasing window-size.

Of course we will also require

($\overline{\mathbf{H3}}$) The function $h : \Theta \to \mathbb{R}^p$ is continuous.

Thus, if the remaining conditions in the Kushner-Clark Lemma hold then the almost sure convergence of the iterates follows immediately.

The next step is to develop some methods to verify ($\overline{\text{H3}}$) and the various conditions we've made in the above lemmas.

5.9 A Framework for Geometrically Ergodic Markov Chains

As in Chapter 3, we shall define some new *specialized conditions* in terms of the previously defined $\widehat{\ell}_1$, r and functions $\bar{V}_\ell : \mathsf{X}^{\ell+1} \to [1, \infty)$, $\ell = 0, 1, \ldots$.

5.9.1 A Norm for the Class of Gradient Estimates

First we define a norm which utilizes the additive boundedness property of the gradient estimate function. Recall the assumptions we have made for the optimization problem:

Definition 1 *For the n-step probability transition kernel (matrix) $P_{\ell,\theta}(\bar{x},\cdot)$ and its invariant measure $\bar{\pi}_\theta(\cdot)$, both on $\mathcal{B}(\mathsf{X}^{\ell+1})$, and some function $V : \mathsf{X} \to [1,\infty)$ let us define the norm over all functions $f_i(\bar{x}) = f_i(x_i)$ where $\bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}$ for each $i = 1, 2, \ldots$:*

$$\begin{aligned}
\left\|\bar{P}_{\ell,\theta}(\bar{x},\cdot) - \bar{\pi}_\theta(\cdot)\right\|_{\bar{V}_\ell} &\doteq \sup_{\substack{f_i : \mathsf{X} \to \mathbb{R}^p, i=1,\ldots,\ell \\ \|f_i(x_i)\| \le V(x_i)}} \left\|\bar{P}_{\ell,\theta}(\bar{x}, \sum_{i=1}^{\ell} f_i) - \bar{\pi}_\theta(\sum_{i=1}^{\ell} f_i)\right\| \\
&\doteq \sup_{\substack{f_i : \mathsf{X} \to \mathbb{R}^p, i=1,\ldots,\ell \\ \|f_i(x_i)\| \le V(x_i)}} \left\|\int_{\mathsf{X}^{\ell+1}} \bar{P}_{\ell,\theta}(\bar{x}, d\bar{y}) \sum_{i=1}^{\ell} f_i(y_i) - \int_{\mathsf{X}^{\ell+1}} \bar{\pi}_\theta(d\bar{x}) \sum_{i=1}^{\ell} f_i(x_i)\right\| \\
&= \sum_{i=1}^{\ell} \sup_{\substack{f_i : \mathsf{X} \to \mathbb{R}^p \\ \|f_i(x_i)\| \le V(x_i)}} \left\|\int_{\mathsf{X}^{\ell+1}} \bar{P}_{\ell,\theta}(\bar{x}, d\bar{y}) f_i(y_i) - \int_{\mathsf{X}^{\ell+1}} \bar{\pi}_\theta(d\bar{x}) f_i(x_i)\right\|, \\
&\qquad \ell = 1, 2, \ldots;\ \bar{x} \in \mathsf{X}^{\ell+1}. \qquad (5.13)
\end{aligned}$$

At first it may appear that this norm is not general enough to include the class of gradient estimates we have suggested in (5.9).

Claim 5.7 *Under the assumptions for some $C_0 < \infty$, $C_1 < \infty$, and $C_2 < \infty$:*

1. $\sup_{\theta\in\Theta} \|L_\theta(x_0, x_1)\| \le C_0, \qquad x_0, x_1 \in \mathsf{X}$
2. $\sup_{\theta\in\Theta} |f_\theta(x_i)| \le C_1 V^r(x_i), \qquad x_i \in \mathsf{X}, \qquad i = 1, \ldots, \ell,$
3. $\sup_{\theta\in\Theta} \|\nabla f_\theta(x_i)\| \le C_2 V^r(x_i), \qquad x_i \in \mathsf{X}, \qquad i = 1, \ldots, \ell,$

the norm (5.13) is equivalent to the norm which includes the class of gradient functions we have defined in (5.9).

Proof:

$$\begin{aligned}
&\sup_{\substack{\bar{g}_i : \mathsf{X} \to \mathbb{R}^p : \|g_i(x_i)\| \le C_2 V(x_i), \quad i=1,\ldots,\ell \\ f_i : \mathsf{X} \to \mathbb{R} : |f_i(x_i)| \le C_1 V(x_i), \quad i=1,\ldots,\ell \\ \bar{L} : \mathsf{X}^2 \to \mathbb{R}^p : \left\|\bar{L}(x_0,x_1)\right\| \le C_0}} \left\|\bar{P}_\theta^n\left(\bar{x}, \sum_{i=1}^{\ell} \frac{\bar{g}_i}{\ell} + \bar{L} f_i\right) - \bar{\pi}_\theta\left(\sum_{i=1}^{\ell} \frac{\bar{g}_i}{\ell} + \bar{L} f_i\right)\right\| \\
\le &\sum_{i=1}^{\ell} \sup_{\substack{\bar{g}_i : \mathsf{X} \to \mathbb{R}^p : \|g_i(x_i)\| \le C_2 V(x_i), \\ f_i : \mathsf{X} \to \mathbb{R} : |f_i(x_i)| \le C_1 V(x_i), \\ \bar{L} : \mathsf{X}^2 \to \mathbb{R}^p : \left\|\bar{L}(x_0,x_1)\right\| \le C_0}} \left\|\bar{P}_\theta^n\left(\bar{x}, \bar{g}_i + \bar{L} f_i\right) - \bar{\pi}_\theta\left(\bar{g}_i + \bar{L} f_i\right)\right\| \\
= &\sum_{i=1}^{\ell} \sup_{\bar{g}_i : \mathsf{X} \to \mathbb{R}^p : \|g_i(x_i)\| \le C_2 V(x_i),} \left\|\bar{P}_\theta^n\left(\bar{x}, \bar{g}_i\right) - \bar{\pi}_\theta\left(\bar{g}_i\right)\right\|
\end{aligned}$$

$$+ \sum_{i=1}^{\ell} \sup_{\substack{f_i : \mathsf{X} \to \mathbb{R} : |f_i(x_i)| \leq C_1 V(x_i), \\ \bar{L} : \mathsf{X}^2 \to \mathbb{R}^p : \left\| \bar{L}(x_0, x_1) \right\| \leq C_0}} \left\| \bar{P}_\theta^n \left(\bar{x}, \bar{L} f_i \right) - \bar{\pi}_\theta \left(\bar{L} f_i \right) \right\|.$$

Now, for any $i = 1, \ldots, \ell$ we check a summand of the last term and notice

$$\begin{aligned} & \sup_{\substack{f_i : \mathsf{X} \to \mathbb{R} : |f_i(x_i)| \leq C_1 V(x_i), \\ \bar{L} : \mathsf{X}^2 \to \mathbb{R}^p : \left\| \bar{L}(x_0, x_1) \right\| \leq C_0}} \left\| \bar{P}_\theta^n \left(\bar{x}, \bar{L} f_i \right) - \bar{\pi}_\theta \left(\bar{L} f_i \right) \right\| \\ = & \sup_{f_i : \mathsf{X} \to \mathbb{R} : |f_i(x_i)| \leq C_1 V(x_i)} \quad \sup_{\bar{L} : \mathsf{X}^2 \to \mathbb{R}^p : \left\| \bar{L}(x_0, x_1) \right\| \leq C_0} \left\| \bar{P}_\theta^n \left(\bar{x}, \bar{L} f_i \right) - \bar{\pi}_\theta \left(\bar{L} f_i \right) \right\| \\ = & \; 2C_0 \sup_{f_i : \mathsf{X} \to \mathbb{R} : |f_i(x_i)| \leq C_1 V(x_i)} \left| \bar{P}_\theta^n \left(\bar{x}, f_i \right) - \bar{\pi}_\theta \left(f_i \right) \right|. \end{aligned}$$

Thus,

$$\begin{aligned} & \sup_{\substack{\bar{g}_i : \mathsf{X} \to \mathbb{R}^p : \|g_i(x_i)\| \leq C_2 V(x_i), \quad i = 1, \ldots, \ell \\ f_i : \mathsf{X} \to \mathbb{R} : |f_i(x_i)| \leq C_1 V(x_i), \quad i = 1, \ldots, \ell \\ \bar{L} : \mathsf{X}^2 \to \mathbb{R}^p : \left\| \bar{L}(x_0, x_1) \right\| \leq C_0}} \left\| \bar{P}_\theta^n \left(\bar{x}, \sum_{i=1}^{\ell} \frac{\bar{g}_i}{\ell} + \bar{L} f_i \right) - \bar{\pi}_\theta \left(\sum_{i=1}^{\ell} \frac{\bar{g}_i}{\ell} + \bar{L} f_i \right) \right\| \\ \leq & \; (C_2 + 2C_0 C_1) \sum_{i=1}^{\ell} \sup_{f_i : \mathsf{X} \to \mathbb{R}^p : \|f_i(x_i)\| \leq V(x_i)} \left\| \bar{P}_\theta^n \left(\bar{x}, f_i \right) - \bar{\pi}_\theta \left(f_i \right) \right\|. \end{aligned}$$

■

5.9.2 Specialized Conditions

($\overline{\mathbf{E1}}$) There exists constants $C_E < \infty$ and $0 < \rho < 1$ such that for $\bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}$,

$$\sup_{\theta \in \Theta} \left\| \bar{P}_{\ell,\theta}^n(\bar{x}, \cdot) - \bar{\pi}_\theta(\cdot) \right\|_{\bar{V}_\ell^r} \leq \begin{cases} C_E \; \ell \; \bar{V}_\ell^r(\bar{x}) \rho^n, & n = 0; \; \ell = 1, 2, \ldots \\ C_E \; \ell \; V^r(x_\ell) \rho^n, & n = 1, 2, \ldots; \; \ell = 1, 2, \ldots \end{cases}$$

($\overline{\mathbf{H5}}$) There exists constants $C_5 < \infty$, $\delta_H > 0$, and $\widehat{\ell}_2 \in (\widehat{\ell}_1, 1)$ such that for all θ, $\theta' \in \Theta$ with $\|\theta - \theta'\| \leq \delta_H$ and all $\bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}$, we have

$$\begin{aligned} \|H_\ell(\theta, \bar{x}) - H_\ell(\theta', \bar{x})\| & \leq C_5 \|\theta - \theta'\|^{\widehat{\ell}_2} \sum_{i=1}^{\ell} V^r(x_i) \\ & \leq C_5 \bar{V}_\ell^r(\bar{x}) \ell \|\theta - \theta'\|^{\widehat{\ell}_2}, \ell = 1, 2, \ldots. \end{aligned}$$

($\overline{\mathbf{C}}$) There exists constants $C_C < \infty$, $\delta_C > 0$, and $\widehat{\ell}_3 \in (\widehat{\ell}_2, 1]$ such that for each $n = 0, 1, \ldots$ and $\ell = 1, 2, \ldots$

$$\left\| \bar{P}_\theta^n(\bar{x}, \cdot) - \bar{P}_{\theta'}^n(\bar{x}, \cdot) \right\|_{\bar{V}_\ell^r} \leq n \ell^2 C_C V^r(x_{\ell'}) \|\theta - \theta'\|^{\widehat{\ell}_3},$$

for all θ, $\theta' \in \Theta$ with $\|\theta - \theta'\| \leq \delta_C$, and all $\bar{x} = (x_0, x_1, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}$.

5.10 Consequences of the Specialized Conditions

Here, we prove three theorems which imply ($\overline{\text{H3}}$), ($\overline{\text{P2}}$), and ($\overline{\text{P3}}$).

Before we begin, let us state a simple bound which follows by iterating (D1) as in 2.10 which we will use later.

$$\begin{aligned}
\sup_{\theta\in\Theta} \bar{P}_\theta^n \bar{V}_\ell(\bar{x}) &\leq \sup_{\theta\in\Theta} \bar{P}_\theta^n \left(\sum_{i=1}^{\ell} V(\cdot_i)\right)(\bar{x}) \\
&\leq \ell C_D V(x_{\ell'}) \\
&\leq \ell C_D \bar{V}_\ell(\bar{x}), \qquad \bar{x} = (x_0, x_1, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}. \qquad (5.14)
\end{aligned}$$

Note in the above we take the convention that for a vector $\bar{x} = (x_0, \ldots, x_\ell) \in \mathbb{R}^\ell$ and some real valued function f on the reals, that $f(\cdot_i)$ refers to the function $f(x_i)$ over all $x_i \in \mathbb{R}$.

In order to prove ($\overline{\text{H3}}$) we will need the following Hölder continuity result on the functions h_ℓ. The method is proof is nearly identical to the proof of Theorem 3.2.

Lemma 5.8 *Assume* (D1), ($\overline{\text{C}}$), ($\overline{\text{H2}}$), ($\overline{\text{H5}}$) *and* ($\overline{\text{E1}}$) *with* $\widehat{\ell}_2$ *determined from* ($\overline{\text{H5}}$). *Then, there exists a constant* $C_h < \infty$ *such that*

$$\|h_\ell(\theta) - h_\ell(\theta')\| \leq \ell^2 C_h \|\theta - \theta'\|^{\widehat{\ell}_2}, \qquad \theta, \theta' \in \Theta,\ \ell = 1, 2, \ldots.$$

Proof: See the appendix. ■

Theorem 5.9 *Assume* (D1), ($\overline{\text{C}}$), ($\overline{\text{H2}}$), ($\overline{\text{H5}}$), ($\overline{\text{E1}}$), (H2), (E1), (P1), (G1), (G2) *plus irreducibility and positive recurrence over* Θ. *Then:*

1. *The function* $h : \Theta \to \mathbb{R}^p$ *given by*

$$h(\theta) = -\nabla_\theta J(\theta) = -\pi_\theta(\nabla_\theta f_\theta) - \pi_\theta Q_\theta g_\theta, \qquad \theta \in \Theta$$

   is continuous.

2. *There exists a* $0 < \rho < 1$ *and a* $C < \infty$ *such that*

$$\sup_{\theta\in\Theta} \|h_\ell(\theta) - h(\theta)\| \leq C\rho^\ell, \qquad \ell = 1, 2, \ldots$$

Proof:
Part 1: Note if $p > 1$ the likelihood ratio L_θ is a p-dimensional column vector.

$$\begin{aligned}
h_\ell(\theta) &= \mathbf{E}_{\bar{\pi}_\theta}\left[H_{\ell,\theta}(\bar{X}_n)\right] \\
&= \mathbf{E}_{\bar{\pi}_\theta}\left[\frac{1}{\ell}\sum_{i=1}^{\ell} \nabla_\theta f_\theta(X_{n,i}) + L_\theta(X_{n,0}, X_{n,1}) \sum_{k=1}^{\ell} f_\theta(X_{n,k})\right] \\
&= \mathbf{E}_{\pi_\theta}\left[\nabla_\theta f_\theta(X_{n,1})\right] + \mathbf{E}_{\pi_\theta}\left[L_\theta(X_{n,0}, X_{n,1}) \sum_{k=1}^{\ell} f_\theta(X_{n,k})\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{E}_{\pi_\theta}\left[\nabla_\theta f_\theta(X_{n,1})\right] + \sum_i \pi_\theta(i) \sum_j P_\theta(i,j) L_\theta(i,j) \mathbf{E}_\theta\left[\sum_{k=1}^{\ell} f_\theta(X_{n,k}) | X_{n,1} = j\right] \\
&= \mathbf{E}_{\pi_\theta}\left[\nabla_\theta f_\theta(X_{n,1})\right] \\
&\quad + \sum_i \pi_\theta(i) \sum_j P_\theta(i,j) L_\theta(i,j) \mathbf{E}_\theta\left[\sum_{k=1}^{\ell} f_\theta(X_{n,k}) - \pi_\theta(f_\theta) | X_{n,1} = j\right].
\end{aligned}
$$

The last step follows since for any finite constant c we have $\sum_j P_\theta(i,j)L_\theta(i,j)c = 0$ uniformly over all i and $\theta \in \Theta$.

Now we recognize the conditional expectation on the right,

$$
\mathbf{E}_\theta\left[\sum_{k=1}^{\ell} f_\theta(X_{n,k}) - \pi_\theta(f_\theta) | X_{n,1} = x\right], \qquad \ell = 1, 2, \ldots
$$

as a sequence which converges to the Poisson equation solution $\nu_\theta(x)$, i.e.

$$
f_\theta(x) - \pi_\theta(f_\theta) = \nu_\theta(x) - P_\theta \nu_\theta(x), \qquad x \in \mathbb{R}
$$

Let us define the sequence of functions

$$
\nu_{\ell,\theta}(x) = \mathbf{E}_\theta\left[\sum_{k=1}^{\ell} f_\theta(X_{n,k}) - \pi_\theta(f_\theta) | X_{n,1} = x\right], \qquad \ell = 1, 2, \ldots,
$$

so we have

$$
\begin{aligned}
\nu_\theta(x) = \lim_{\ell\to\infty} \nu_{\ell,\theta}(x) &= \lim_{\ell\to\infty} \mathbf{E}_\theta\left[\sum_{k=1}^{\ell} f_\theta(X_{n,k}) - \pi_\theta(f_\theta) | X_{n,1} = x\right] \\
&= \sum_{n=0}^{\infty} \mathbf{E}_\theta\left[f_\theta(X_{n,k}) - \pi_\theta(f_\theta) | X_{n,0} = x\right].
\end{aligned}
$$

Under (H2) and (E1) we have the bound

$$
|\nu_\theta(x)| \leq C_H \sum_{\ell=0}^{\infty} \left\| P_\theta^\ell(x,\cdot) - \pi_\theta(\cdot) \right\|_{V^r} \leq C_H C_E V^r(x) \sum_{\ell=0}^{\infty} \rho^\ell
$$

Now for all $m > \ell$ look at the sequence

$$
\begin{aligned}
\sup_{\theta\in\Theta} |\nu_{m,\theta}(x) - \nu_{\ell,\theta}(x)| &= \sup_{\theta\in\Theta} \left| \sum_{k=\ell+1}^{m} \mathbf{E}_\theta\left[f_\theta(X_{n,k}) - \pi_\theta(f_\theta) | X_{n,1} = x\right] \right| \\
&\leq C_H \sup_{\theta\in\Theta} \sum_{\ell=\ell+1}^{\infty} \left\| P_\theta^\ell(x,\cdot) - \pi_\theta(\cdot) \right\|_{V^r} \\
&\leq C_H C_E V^r(x) \frac{\rho^{\ell+1}}{1-\rho}
\end{aligned}
$$

and thus by the Cauchy criterion we have for each fixed $x \in \mathsf{X}$ a *uniform* convergence of $\lim_{\ell\to\infty} \nu_{\ell,\theta}(x) = \nu_\theta(x)$.

Again we consider any $m > \ell$ and look at

$$\sup_{\theta\in\Theta} \|h_m(\theta) - h_\ell(\theta)\| = \sup_{\theta\in\Theta} \left\| \sum_i \pi_\theta(i) \sum_j P_\theta(i,j) L_\theta(i,j) \left(\nu_{m,\theta}(j) - \nu_{\ell,\theta}(j)\right) \right\|$$

By condition (4.9) under (G1) and (G2) we have $\|L_\theta(i,j)\| \le \widetilde{K}_2 < \infty$ so that

$$\begin{aligned}
\sup_{\theta\in\Theta} \|h_m(\theta) - h_\ell(\theta)\| &= \sup_{\theta\in\Theta} \left\| \sum_i \pi_\theta(i) \sum_j P_\theta(i,j) L_\theta(i,j) \left(\nu_{m,\theta}(j) - \nu_{\ell,\theta}(j)\right) \right\| \\
&\le \widetilde{K}_2 \sup_{\theta\in\Theta} \sum_i \pi_\theta(i) \sum_j P_\theta(i,j) \left|\nu_{m,\theta}(j) - \nu_{\ell,\theta}(j)\right| \\
&\le \widetilde{K}_2 \sup_{\theta\in\Theta} \sum_j \pi_\theta(j) \left|\nu_{m,\theta}(j) - \nu_{\ell,\theta}(j)\right| \\
&\le \widetilde{K}_2 \sup_{\theta\in\Theta} \sum_j \pi_\theta(j) C_H C_E V^r(x) \sum_{k=\ell}^{m} \rho^k \\
&\le \widetilde{K}_2 C_H C_E \frac{\rho^\ell}{1-\rho} \sup_{\theta\in\Theta} \sum_j \pi_\theta(j) V^r(j)
\end{aligned}$$

Since $\sup_{\theta\in\Theta} \pi_\theta(V^r) < \infty$ by (2.15) we have shown that the sequence of functions $h_\ell(\theta)$ are also *uniformly* convergent, and by Lemma 5.8 they are each continuous, hence the limit $h(\theta)$ is also continuous.

Part 2: We can see from the above that the convergence rate of $h_\ell(\theta)$ to $h(\theta)$ is exponential.

■

The next theorem identifies sufficient conditions which imply ($\overline{\text{P2}}$).

Theorem 5.10 *Assume* (D1), ($\overline{\text{P1}}$), (E1), *and* ($\overline{\text{H2}}$). *Then there exists a $C_\nu < \infty$ such that for all $\theta, \theta' \in \Theta$, and $\ell, \ell' = 1, 2, \ldots$*

$$\|\nu_{\ell,\theta}(\bar{x})\| \le C_\nu \; \ell \; \bar{V}_\ell^r(\bar{x}), \qquad \bar{x} \in \mathsf{X}^{\ell+1} \tag{5.15}$$

$$\left\|\bar{P}_\theta \nu_{\ell,\theta}(\bar{x})\right\| \le C_\nu \; \ell \; V^r(x_{\ell'}), \qquad \bar{x} = (x_0, x_1, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}. \tag{5.16}$$

Proof: First note that for all $\bar{x} = (x_0, x_1, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}$ where ℓ' is an arbitrary positive integer that

$$\bar{P}_\theta \nu_{\ell,\theta}(\bar{x}) = \bar{P}_\theta \nu_{\ell,\theta}(x_{\ell'}) = \bar{P}_{\ell,\theta} \nu_{\ell,\theta}(x_{\ell'})$$

Therefore,

$$\begin{aligned}
\left\|\bar{P}_\theta \nu_{\ell,\theta}(\bar{x})\right\| &= \left\|\bar{P}_{\ell,\theta} \nu_{\ell,\theta}(x_{\ell'})\right\| \\
&= \left\| \sum_{n=1}^{\infty} \bar{P}_{\ell,\theta}^n(x_{\ell'}, H_{\ell,\theta}) - \pi_{\ell,\theta}(H_{\ell,\theta}) \right\| \\
&\le \sum_{n=1}^{\infty} \left\| \bar{P}_{\ell,\theta}^n(x_{\ell'}, H_{\ell,\theta}) - \bar{\pi}_\theta(H_{\ell,\theta}) \right\|
\end{aligned}$$

$$
\begin{aligned}
&\leq \quad C_H \sum_{n=1}^{\infty} \left\| \bar{P}^n_{\ell,\theta}(x_{\ell'},\cdot) - \bar{\pi}_\theta(\cdot) \right\|_{\bar{V}^r_\ell} \\
&\leq \quad C_H \sum_{i=1}^{\ell} \sum_{n=1}^{\infty} \left\| P^{\ell(n-1)+i}_\theta(x_{\ell'},\cdot) - \pi_\theta(\cdot) \right\|_{V^r} \\
&\leq \quad C_H \sum_{i=1}^{\ell} \sum_{n=1}^{\infty} C_E \rho^{\ell(n-1)+i} V^r(x_{\ell'}) \\
&\leq \quad C_H \ell \sum_{n=1}^{\infty} C_E \rho^{\ell(n-1)} V^r(x_{\ell'}) \\
&\leq \quad \ell C_H C_E V^r(x_{\ell'}) \sum_{k=0}^{\infty} \rho^{\ell k} \\
&\leq \quad \ell C_H C_E V^r(x_{\ell'}) \sum_{k=0}^{\infty} \rho^{k} \\
&= \quad \ell C_H C_E V^r(x_{\ell'}) \frac{1}{1-\rho}, \qquad \bar{x} = (x_0, x_1, \ldots, x_{\ell'}) \in \mathsf{X}^{\ell'+1}, \quad \theta \in \Theta,
\end{aligned} \tag{5.17}
$$

and for all $\ell = 1, 2, \ldots$.

For $\nu_{\ell,\theta}(\bar{x}) = \sum_{n=0}^{\infty} \left(\int \bar{P}^n_{\ell,\theta}(\bar{x}, dy) H_{\ell,\theta}(y) - h_\ell(\theta) \right)$ we have:

$$
\begin{aligned}
\|\nu_{\ell,\theta}(\bar{x})\| \quad &= \quad \left\| \sum_{n=0}^{\infty} \left(\int \bar{P}^n_{\ell,\theta}(\bar{x}, dy) H_{\ell,\theta}(y) - h_\ell(\theta) \right) \right\| \\
&\leq \quad \sum_{n=0}^{\infty} \left\| \int \bar{P}^n_{\ell,\theta}(\bar{x}, dy) H_\theta(y) - \pi_{\ell,\theta}(H_{\ell,\theta}) \right\| \\
&\leq \quad \|H_{\ell,\theta}(\bar{x}) - \pi_{\ell,\theta}(H_{\ell,\theta})\| + \sum_{n=1}^{\infty} \left\| \int \bar{P}^n_{\ell,\theta}(\bar{x}, dy) H_\theta(y) - \pi_{\ell,\theta}(H_{\ell,\theta}) \right\|
\end{aligned}
$$

the last term is bounded above in (5.17). Thus by $(\overline{\text{H2}})$ we have for all $\ell = 1, 2, \ldots$

$$
\begin{aligned}
\|\nu_{\ell,\theta}(\bar{x})\| \quad &\leq \quad C_H \ell \bar{V}^r_\ell(\bar{x}) + \sup_{\theta\in\Theta} \bar{\pi}_\theta \left(\sum_{i=1}^{\ell} V^r(\cdot_i) \right) + \ell C_H C_E V^r(x_\ell) \frac{1}{1-\rho} \\
&\leq \quad C_\nu \ell \bar{V}^r_\ell(\bar{x}), \qquad \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}
\end{aligned}
$$

for some $C_\nu < \infty$. The middle term in the first line is bounded as in (2.15) under the assumed irreducibility and positive recurrence. ∎

Theorem 5.11 *Assume* $(\overline{\text{S}})$, $(\overline{\text{H2}})$, $(\overline{\text{H5}})$, $(\overline{\text{P1}})$, $(\overline{\text{E1}})$, $(\overline{\text{C}})$, *and* (D1) *with the constants* $\widehat{\ell}_1$ *determined from* (S), $\widehat{\ell}_2$ *determined from* $(\overline{\text{H5}})$ *and* $\widehat{\ell}_3$ *determined from* (C). *Then there exists a constant* $C_\delta < \infty$ *such that for all* $\ell, \ell' = 1, 2, \ldots$, $\theta, \theta' \in \Theta$, $\quad \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell'+1}$

$$
\left\| \bar{P}_\theta \nu_{\ell,\theta}(\bar{x}) - \bar{P}_{\theta'} \nu_{\ell,\theta'}(\bar{x}) \right\| \leq C_\delta \ell^2 \bar{V}^r(x_{\ell'}) \left\| \theta - \theta' \right\|^{\widehat{\ell}_1}. \tag{5.18}
$$

Proof: Pick a δ such such $\delta \le \min\{\delta_C, \delta_H\}$ and $0 < \delta \le 1$. Again, we let $\Delta\theta = \theta - \theta'$ and consider the two cases of $\|\Delta\theta\| \le \delta$ and $\|\Delta\theta\| > \delta$ separately.

The case $\|\Delta\theta\| > \delta$ follows trivially from Theorem 5.10:

$$\begin{aligned} \left\|\bar{P}_\theta \nu_\theta(\bar{x}) - \bar{P}_{\theta'}\nu_{\ell,\theta'}(\bar{x})\right\| &\le \left\|\bar{P}_\theta \nu_\theta(\bar{x})\right\| + \left\|\bar{P}_{\theta'}\nu_{\ell,\theta'}(\bar{x})\right\| \\ &\le 2C_\nu \ell \bar{V}^r(x_{\ell'}) \\ &\le \frac{2}{\delta^{\widehat{\ell}_1}} C_\nu \ell V^r(x_{\ell'}) \left\|\Delta\theta\right\|^{\widehat{\ell}_1}, \qquad \theta, \theta' \in \Theta, \quad \|\Delta\theta\| \ge \delta. \end{aligned} \tag{5.19}$$

Now consider the case $\|\Delta\theta\| \le \delta$ such that $\theta, \theta' \in \Theta$,

$$\begin{aligned} &\left\|\bar{P}_\theta \nu_{\ell,\theta}(\bar{x}) - \bar{P}_{\theta'}\nu_{\ell,\theta'}(\bar{x})\right\| = \left\|\bar{P}_\theta \nu_{\ell,\theta}(x_{\ell'}) - \bar{P}_{\theta'}\nu_{\ell,\theta'}(x_{\ell'})\right\| \\ &\quad = \left\|\bar{P}_{\ell,\theta} \nu_{\ell,\theta}(x_{\ell'}) - \bar{P}_{\ell,\theta'}\nu_{\ell,\theta'}(x_{\ell'})\right\| \qquad (5.20) \\ &\quad = \left\|\sum_{n=1}^{\infty} \left(\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta}) - h_\ell(\theta)\right) - \sum_{n=1}^{\infty} \left(\bar{P}^n_{\ell,\theta'}(x_{\ell'}, H_{\ell,\theta'}) - h_\ell(\theta')\right)\right\| \\ &\quad \le \sum_{n=1}^{\infty} \left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta}) - h_\ell(\theta) - \bar{P}^n_{\ell,\theta'}(x_{\ell'}, H_{\ell,\theta'}) + h_\ell(\theta')\right\| \qquad (5.21) \\ &\quad \le \sum_{n=1}^{N-1} \left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta}) - \bar{P}^n_{\ell,\theta'}(x_{\ell'}, H_{\ell,\theta'}) + \bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta'}) - \bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta'})\right\| \\ &\quad + \sum_{n=1}^{N-1} \|h_\ell(\theta') - h_\ell(\theta)\| \\ &\quad + \sum_{n=N}^{\infty} \|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta}) - \bar{\pi}_\theta(H_\theta) - \bar{P}^n_{\ell,\theta'}(x_{\ell'}, H_{\ell,\theta'}) + \bar{\pi}_{\ell,\theta'}(H_{\ell,\theta'})\| \qquad (5.22) \end{aligned}$$

where we have introduced some canceling terms in the last inequality. Continuing from (5.22)

$$\begin{aligned} &\left\|\bar{P}_{\ell,\theta} \nu_{\ell,\theta}(x_{\ell'}) - \bar{P}_{\ell,\theta'}\nu_{\ell,\theta'}(x_{\ell'})\right\| \\ &\quad \le \sum_{n=1}^{N-1} \left\{\left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta}) - \bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta'})\right\| + \left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta'}) - \bar{P}^n_{\ell,\theta'}(x_{\ell'}, H_{\ell,\theta'})\right\|\right\} \\ &\quad + \sum_{n=1}^{N-1} \|h_\ell(\theta) - h_\ell(\theta')\| \\ &\quad + \sum_{n=N}^{\infty} \left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta}) - \bar{\pi}_{\ell,\theta}(H_{\ell,\theta})\right\| + \sum_{n=N}^{\infty} \left\|\bar{P}^n_{\ell,\theta'}(x_{\ell'}, H_{\ell,\theta'}) - \bar{\pi}_{\ell,\theta'}(H_{\ell,\theta'})\right\| \\ &\quad \le \sum_{n=0}^{N-1} \left\{\left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta} - H_{\ell,\theta'})\right\| + \left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta'}) - \bar{P}^n_{\ell,\theta'}(x_{\ell'}, H_{\ell,\theta'})\right\|\right\} \\ &\quad + \sum_{n=0}^{N-1} \|h_\ell(\theta) - h_\ell(\theta')\| + 2 \sup_{\theta \in \Theta} \sum_{n=N}^{\infty} \left\|\bar{P}^n_{\ell,\theta}(x_{\ell'}, H_{\ell,\theta}) - \bar{\pi}_{\ell,\theta}(H_{\ell,\theta})\right\| \end{aligned} \tag{5.23}$$

We look at several of the above terms. First, from ($\overline{\text{H5}}$) followed by (5.14) under (D1) (and using Jensen's Inequality)

$$\left\|\bar{P}^n_{\ell,\theta}(\bar{x}, H_{\ell,\theta} - H_{\ell,\theta'})\right\| \;\le\; \ell C_5 \left\|\theta - \theta'\right\|^{\widehat{\ell}_2} \bar{P}^n_{\ell,\theta}(\bar{x}, \bar{V}^r_\ell), \qquad n = 1, 2, \ldots, \quad \bar{x} \in \mathsf{X}^{\ell+1}$$

$$\leq \quad C_5 \ell^2 \|\theta - \theta'\|^{\widehat{\ell_2}} C_D^r V^r(x_\ell), \qquad \bar{x} \in \mathsf{X}^{\ell+1}.$$

Second, from ($\overline{\text{H2}}$) and ($\overline{\text{C}}$),

$$\begin{aligned}
\left\| \bar{P}^n_{\ell,\theta}(\bar{x}, H_{\ell,\theta'}) - \bar{P}^n_{\ell,\theta'}(\bar{x}, H_{\ell,\theta'}) \right\| &\leq C_H \left\| \bar{P}^n_{\ell,\theta}(\bar{x}, \cdot) - \bar{P}^n_{\ell,\theta'}(\bar{x}, \cdot) \right\|_{\bar{V}^r_\ell} \\
&\leq n\ell^2 C_C C_H V^r(x_\ell) \|\theta - \theta'\|^{\widehat{\ell_3}}, \\
&\quad n = 1, 2, \ldots, \ \bar{x} \in \mathsf{X}^{\ell+1}, \qquad \ell = 1, 2, \ldots.
\end{aligned}$$

Third, from Theorem 5.8 there exists a $C_h < \infty$ such that

$$\|h_\ell(\theta) - h_\ell(\theta')\| \leq \ell^2 C_h \|\theta - \theta'\|^{\widehat{\ell_2}}, \qquad \theta, \theta' \in \Theta, \quad , \ell = 1, 2, \ldots.$$

Fourth, from ($\overline{\text{H2}}$) and ($\overline{\text{E1}}$)

$$\begin{aligned}
\left\| \bar{P}^n_{\ell,\theta}(\bar{x}, H_{\ell,\theta'}) - \bar{\pi}_{\ell,\theta}(H_{\ell,\theta'}) \right\| \leq C_E C_H \ell \rho^n V^r(x_\ell), \qquad & n = 1, 2, \ldots, \quad \bar{x} \in \mathsf{X}^{\ell+1} \\
& \ell = 1, 2, \ldots.
\end{aligned}$$

Substituting these bounds into (5.23) we find

$$\begin{aligned}
&\left\| \bar{P}_{\ell,\theta}\nu_{\ell,\theta}(\bar{x}) - \bar{P}_{\ell,\theta}\nu_{\ell,\theta'}(\bar{x}) \right\| \\
&\leq \sum_{n=1}^{N-1} \left\{ \ell^2 C_5 \|\Delta\theta\|^{\widehat{\ell_2}} C_D^r V^r(x_\ell) + n\ell^2 C_C C_H V^r(x_\ell) \|\Delta\theta\|^{\widehat{\ell_3}} + C_h \ell^2 \|\Delta\theta\|^{\widehat{\ell_2}} \right\} \\
&+ 2\ell \sup_{\theta \in \Theta} \sum_{n=N}^{\infty} C_E C_H \rho^n V^r(x_\ell) \\
&\leq \ell^2 N C_5 \|\Delta\theta\|^{\widehat{\ell_2}} C_D^r V^r(x_\ell) + \ell^2 \frac{N(N-1)}{2} C_C C_H \|\Delta\theta\|^{\widehat{\ell_3}} V^r(x_\ell) + N C_h \|\Delta\theta\|^{\widehat{\ell_2}} \\
&+ \ell 2 C_E C_H \frac{\rho^N}{1-\rho} V^r(x) \\
&\leq \ell^2 V^r(x) \left\{ (C_D^r C_5 + C_h) N \|\Delta\theta\|^{\widehat{\ell_2}} + \frac{C_C C_H}{2} N(N-1) \|\Delta\theta\|^{\widehat{\ell_3}} + 2 C_E C_H \frac{\rho^N}{1-\rho} \right\}
\end{aligned}$$

since $V^r \geq 1$. Now the bracketed expression can be bounded by the same technique as in the proof of Theorem 3.4 to yield (5.18). ∎

5.11 Verification of the Specialized Conditions for the Windowed Process

In this section, we develop tools to verify the various conditions outlined in the previous section which imply convergence of the SA when driven by this gradient estimate. We present a series of results which show that under most conditions, the specialized conditions ($\overline{\text{D0}}$), ($\overline{\text{H2}}$), ($\overline{\text{H5}}$), ($\overline{\text{C}}$), ($\overline{\text{E1}}$), and ($\overline{\text{P1}}$) for the windowed process can be verified by checking these conditions for the original single step Markov chain. Furthermore, Theorems 5.8, 5.10, and 5.11 also imply ($\overline{\text{P2}}$), ($\overline{\text{P3}}$), and ($\overline{\text{H3}}$). Hence, all the conditions implying the (KC4) noise condition for increasing window SA algorithm are satisfied. The following sections identify what modest additional conditions are required for this approach.

5.11.1 Condition $(\overline{\text{D0}})$

The next theorem shows that $(\overline{\text{D0}})$ follows from (D1) for the standard single-step Markov chain governed by one step transition function P_θ.

Theorem 5.12 *Suppose there exists a function $V : \mathsf{X} \to [1, \infty)$ and a family of one-step transition functions $\{P_\theta, \theta \in \Theta\}$ such that* (D1) *holds. Define for $\ell = 1, 2, \ldots$ the functions*

$$\begin{aligned} \bar{V}_\ell(\bar{x}) &\doteq \bar{V}(x_0, x_1, \ldots, x_\ell) \doteq \sup\{V(x_j) : \; j = 1, \ldots, \ell_n\} \\ \bar{V}_0(x) &\doteq V(x) \end{aligned}$$

for $\bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}$ and $x \in \mathsf{X}$. Then, the windowed Markov chain $\{\bar{X}_n : \; n = 0, 1, \ldots\}$ has the $(\overline{\text{D0}})$ *property.*

Proof: We have

$$\begin{aligned} \mathbf{E}_\theta\left[\bar{V}_\ell(\bar{X}_{n+1}) | X_0 = x\right] &= \mathbf{E}_\theta\left[\sup_{i=1,\ldots,\ell}\{V(X_{n+1,i})\} | X_0 = x\right] \\ &\le \mathbf{E}_\theta\left[\sum_{i=1}^{\ell} V(X_{n,i}) | X_0 = x\right] \\ &= \sum_{i=1}^{\ell} \mathbf{E}_\theta\left[V(X_{n,i}) | X_0 = x\right] \end{aligned}$$

Then we can write any of the expectations $\mathbf{E}_\theta\left[V(X_{n,i}) | X_0 = x\right], \; i = 1, \ldots, \ell$ as an iterated conditional expectation. Each single-step conditional expectation is bounded by (D1), so if $X_{n,i}$ is m steps from X_0, we can write m iterations.

$$\begin{aligned} \mathbf{E}_\theta\left[V(X_{n,i}) | X_0 = x\right] &\le \sup_{\{\theta_i \in \Theta, i=1,\ldots,m-1\}} P_\theta P_{\theta_1} P_{\theta_2} \cdots P_{\theta_{m-1}} V(x) && (5.24) \\ &\le \lambda^m V(x) + L \sum_{i=0}^{m-1} \lambda^i \\ &\le \left(1 + \frac{L}{1-\lambda}\right) V(x) && (5.25) \end{aligned}$$

and this bound holds for any $m = 1, 2, \ldots$. If we define $C_D \doteq 1 + \frac{L}{1-\lambda}$ we conclude

$$\mathbf{E}_\theta\left[\bar{V}_\ell(\bar{X}_{n+1}) | X_0 = x\right] = \ell C_D V(x), \qquad n = 0, 1, 2, \ldots$$

which is $(\overline{\text{D0}})$ if we substitute in the sequence $\{\ell_{n+1}, n = 0, 1, 2, \ldots\}$ for ℓ. ■

5.11.2 Condition $(\overline{\text{H2}})$

Recall the gradient estimate we have proposed using in the SA algorithm. For some fixed integer ℓ, if we take

$$\bar{x}_n = (x_{n,0}, x_{n,1}, \ldots, x_{n,\ell}) \in \mathsf{X}^{\ell+1},$$

and if we make the convention that $\frac{0}{0} = 0$, then

$$\widehat{G}_i(\theta, \bar{x}) \doteq \left(\frac{1}{\ell} \sum_{j=1}^{\ell} \frac{\partial f_\theta}{\partial \theta_i}(x_j) \right) + \left\{ \frac{\frac{\partial}{\partial \theta_i} p_{x_0,x_1}(\theta)}{p_{x_0,x_1}(\theta)} \right\} \left(\sum_{j=1}^{\ell_n} f_\theta(x_j) \right),$$

$$\bar{x}_n \in \mathsf{X}^{\ell+1}, \quad n = 1, 2, \ldots, \quad \theta \in \Theta, \quad i = 1, \ldots, p.$$

Applying conditions (G2), (F1), (F3) to this gradient estimate, we have for each $i = 1, 2, \ldots, p$

$$\begin{aligned}
\left| \widehat{G}_i(\theta, \bar{x}) \right| &\leq \frac{1}{\ell} \sum_{j=1}^{\ell} \left| \frac{\partial f_\theta}{\partial \theta_i}(x_j) \right| + \left| \frac{\frac{\partial}{\partial \theta_i} p_{x_0,x_1}(\theta)}{p_{x_0,x_1}(\theta)} \right| \sum_{j=1}^{\ell} |f_\theta(x_j)| \\
&\leq \frac{1}{\ell_n} C_3 \sum_{j=1}^{\ell} V^r(x_j) + K_2 \sum_{j=1}^{\ell} V^r(x_j) \\
&= (C_3 + K_2 C_1) \sum_{j=1}^{\ell} V^r(x_j) \\
&\leq C'_H \sum_{j=1}^{\ell} V^r(x_j), \qquad \theta \in \Theta, \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1},
\end{aligned}$$

where we let $C'_H = C_3 + K_2 C_1$. Finally, in $\mathbb{R}^p$, since the norm $\|x\|_1 = \sum_{i=1}^{p} \left| x^{(i)} \right|$ is *equivalent* to the Euclidean norm, this implies there exists a constant $C_H < \infty$ such that

$$\left\| \widehat{G}(\theta, \bar{x}) \right\| \leq C_H \sum_{j=1}^{\ell} V^r(x_j), \qquad \theta \in \Theta, \bar{x} \in \mathsf{X}^{\ell+1}.$$

This shows that the function $\bar{V}^r$ provides a uniform upper bound on the windowed gradient estimate hence the first part of ($\overline{\text{H2}}$) is verified.

The second part of ($\overline{\text{H2}}$) involves the ρ_n term defined in (5.6).. From Theorem 5.9 there exists some $C < \infty$ and a $0 < \rho < 1$ such that

$$\gamma_n^{\widehat{\ell}_1} \|\rho_n(\theta, \cdot)\| \leq \sup_{\theta \in \Theta} \|h_{\ell_n}(\theta) - h(\theta)\| \leq C \rho^{\ell_n} = C e^{-\delta \ell_n}, \qquad n = 1, 2, \ldots$$

by taking $-\delta = \ln(\rho)$ so that

$$\|\rho_n(\theta, \cdot)\| \leq \frac{C e^{-\delta \ell_n}}{\gamma_n^{\widehat{\ell}_1}}, \qquad n = 1, 2, \ldots$$

Suppose now we use the sequences defined by

$$\gamma_n = \frac{1}{n+1}, \qquad \ell_n = \lfloor \alpha \ln(n+1) \rfloor, \qquad n = 1, 2, \ldots$$

for some chosen $\alpha > 0$. Then

$$\rho \rho_n(\theta, \cdot) \leq \frac{C e^{-\delta \alpha \ln(n+1)}}{(n+1)^{-\widehat{\ell}_1}} = \frac{C}{(n+1)^{\alpha\delta - \widehat{\ell}_1}}, \qquad n = 1, 2, \ldots. \tag{5.26}$$

and we clearly need $0 < \widehat{\ell}_1 \leq \alpha\delta$ to satisfy ($\overline{\text{H2}}$). For any $\widehat{\ell}_1$ satisfying ($\overline{\text{S}}$), this can be achieved through appropriate choice of α using a bound on ρ obtained from the results in Section 3.4.1 and computing $\delta = -\ln(\rho)$. Alternatively, we can simply note the step-size and observation window size sequences we are using satisfy ($\overline{\text{S}}$) for *any* $\widehat{\ell}_1$ in $0 < \widehat{\ell}_1 < 1$. Hence, we conclude that a sufficiently small $\widehat{\ell}_1$ exists so that (5.26) is bounded for all $n = 1, 2, \ldots$.

5.11.3 Condition ($\overline{\text{C}}$)

To show ($\overline{\text{C}}$), let us rework Theorem 3.12 for the window process and its transition kernels $\{\bar{P}_\theta, \theta \in \Theta\}$ with the goal of identifying conditions on the one-step transition kernel P.

For some fixed positive integer ℓ, we have defined

$$\bar{V}_\ell(\bar{x}) \doteq \sup_{i=1,2,\ldots,\ell} \{V(x_i)\}, \qquad \bar{x} \in \mathsf{X}^{\ell+1}$$

where $\bar{x} = (x_0, x_1, x_2, \ldots x_\ell) \in \mathsf{X}^{\ell+1}$. We also have for all exponents $0 < r \leq 1$

$$\bar{V}_\ell^r(\bar{x}) = \sup_{i=1,2,\ldots,\ell} \{V^r(x_i)\}, \qquad \bar{x} \in \mathsf{X}^{\ell+1}$$

Note that $\bar{V}(\bar{x})$ and $\bar{V}^r(\bar{x})$ do not depend on x_0.

Theorem 5.13 *Assume* (D1) *and* (G2). *Then, for all θ, θ' in Θ such that $\|\theta - \theta'\| \leq \delta_M$, $\bar{x}$ in $\mathsf{X}^{\ell+1}$,*

$$\left\| \bar{P}^n_{\ell,\theta}(\bar{x}, \cdot) - \bar{P}^n_{\ell,\theta'}(\bar{x}, \cdot) \right\|_{\bar{V}_\ell^r} \leq n 2 \ell^2 K_2 C_D^2 V^r(x_\ell) \left\|\theta - \theta'\right\|^{\widehat{\ell}_3}, \qquad \text{for all } n = 1, 2, \ldots.$$

Proof: Consider any $\theta, \theta' \in \Theta$ with $\|\theta - \theta'\| \leq \delta_m$, any $\bar{x} = (x_0, x_1, \cdots, x_\ell)$ in $\mathsf{X}^{\ell+1}$. Below let $f : \mathsf{X}^{\ell+1} \to \mathbb{R}^p$. We have for all $n = 1, 2, \ldots$

$$\begin{aligned}
&\left\| \bar{P}^n_{\ell,\theta}(\bar{x}, \cdot) - \bar{P}^n_{\ell,\theta'}(\bar{x}, \cdot) \right\|_{\bar{V}_\ell^r} \\
&= \sup_{\|f\| \leq \bar{V}_\ell^r} \left\| \bar{P}^n_{\ell,\theta}(\bar{x}, f) - \bar{P}^n_{\ell,\theta'}(\bar{x}, f) \right\| \\
&= \sup_{\|f\| \leq \bar{V}_\ell^r} \left\| \bar{P}^n_{\ell,\theta}(\bar{x}, f) - \sum_{i=1}^{n-1} \bar{P}^{n-i}_{\ell,\theta} \bar{P}^i_{\theta'}(\bar{x}, f) + \sum_{i=1}^{n-1} \bar{P}^{n-i}_{\ell,\theta} \bar{P}^i_{\theta'}(\bar{x}, f) - \bar{P}^n_{\ell,\theta'}(\bar{x}, f) \right\| \\
&\leq \sum_{i=1}^{n} \sup_{\|f\| \leq \bar{V}_\ell^r} \left\| \bar{P}^{n-i+1}_{\ell,\theta} \bar{P}^{i-1}_{\ell,\theta'}(\bar{x}, f) - \bar{P}^{n-i}_{\ell,\theta} \bar{P}^i_{\ell,\theta'}(\bar{x}, f) \right\| \\
&\leq \sum_{i=1}^{n} \sup_{\|f\| \leq \bar{V}_\ell^r} \left\| \bar{P}^{n-i}_{\ell,\theta} \left(\bar{P}_{\ell,\theta} - \bar{P}_{\ell,\theta'} \right) \bar{P}^{i-1}_{\ell,\theta'}(\bar{x}, f) \right\|
\end{aligned} \tag{5.27}$$

where $\bar{P}_{\ell,\theta} - \bar{P}_{\ell,\theta'}$ is a signed measure.

Continuing from (5.27) we find

$$\begin{aligned}
&\left\|\bar{P}^n_{\ell,\theta}(\bar{x},\cdot)-\bar{P}^n_{\ell,\theta'}(\bar{x},\cdot)\right\|_{\bar{V}^r} \\
&\leq \sum_{i=1}^{n}\sup_{\|f\|\leq\bar{V}^r_\ell}\left\|\bar{P}^{n-i}_{\ell,\theta}\left(\bar{P}_{\ell,\theta}-\bar{P}_{\ell,\theta'}\right)\bar{P}^{i-1}_{\ell,\theta'}(\bar{x},f)\right\| \\
&\leq \sum_{i=1}^{n}\left(\sup_{\|f\|\leq\bar{V}^r_\ell}\left\|\bar{P}^{n-i}_{\ell,\theta}\left(\bar{P}_{\ell,\theta}-\bar{P}_{\ell,\theta'}\right)^{+}\bar{P}^{i-1}_{\ell,\theta'}(\bar{x},f)\right\|\right. \qquad (5.28)\\
&\qquad\left.+\sup_{\|f\|\leq\bar{V}^r_\ell}\left\|\bar{P}^{n-i}_{\theta}\left(\bar{P}_{\ell,\theta}-\bar{P}_{\ell,\theta'}\right)^{-}\bar{P}^{i-1}_{\ell,\theta'}(\bar{x},f)\right\|\right)
\end{aligned}$$

Consider now

$$\left|\left(\bar{P}_{\ell,\theta}-\bar{P}_{\ell,\theta'}\right)^{\pm}(\bar{x},\bar{y})\right|\leq 1_{\{x_\ell=y_0\}}\left|\prod_{j=1}^{m}P_{\ell,\theta}(y_{j-1},y_j)-\prod_{j=1}^{m}P_{\ell,\theta'}(y_{j-1},y_j)\right|,\quad \bar{x},\bar{y}\in\mathsf{X}^{\ell+1}$$

and using the same expansion as above, we find

$$\begin{aligned}
&\left|\left(\bar{P}_{\ell,\theta}-\bar{P}_{\ell,\theta'}\right)^{\pm}(\bar{x},\bar{y}))\right| \\
&\leq 1_{\{x_\ell=y_0\}}\sum_{k=1}^{\ell}\left(\prod_{j=1}^{k-1}P_{\theta'}(y_{j-1},y_j)\right)\left|P_\theta(y_{k-1};y_k)-P_{\theta'}(y_{k-1};y_k)\right|\prod_{j=k+1}^{\ell}P_\theta(y_{j-1},y_j) \\
&\leq 1_{\{x_\ell=y_0\}}\sum_{k=1}^{\ell}\left(\prod_{j=1}^{k-1}P_{\theta'}(y_{j-1},y_j)\right)\|\Delta\theta\|^{\widehat{\ell_3}}K_2P_\theta(y_{k-1};y_k)\left(\prod_{j=k+1}^{\ell}P_\theta(y_{j-1},y_j)\right) \\
&\leq 1_{\{x_\ell=y_0\}}\|\Delta\theta\|^{\widehat{\ell_3}}K_2\sum_{k=1}^{\ell}\left(\prod_{j=1}^{k-1}P_{\theta'}(y_{j-1},y_j)\right)\left(\prod_{j=k}^{\ell}P_\theta(y_{j-1},y_j)\right)
\end{aligned}$$

Returning to (5.29) we thus get

$$\begin{aligned}
&\left\|\bar{P}^n_{\ell,\theta}(\bar{x},\cdot)-\bar{P}^n_{\ell,\theta'}(\bar{x},\cdot)\right\|_{\bar{V}^r_\ell} \\
&\leq \|\Delta\theta\|^{\widehat{\ell_3}}K_2\sum_{i=1}^{n}\sum_{k=1}^{\ell}\sup_{\|f\|\leq\bar{V}^r_\ell}\left\|\bar{P}^{n-i}_{\ell,\theta}P^{k-1}_{\theta'}P^{\ell-k+1}_{\theta}\bar{P}^{i-1}_{\ell,\theta'}(\bar{x},f)\right\| \\
&\leq \|\Delta\theta\|^{\widehat{\ell_3}}K_2\sum_{i=1}^{n}\sum_{k=1}^{\ell}\bar{P}^{n-i}_{\ell,\theta}P^{k-1}_{\theta'}P^{\ell-k+1}_{\theta}\bar{P}^{i-1}_{\ell,\theta'}(\bar{x},\bar{V}^r_\ell) \\
&\leq \|\Delta\theta\|^{\widehat{\ell_3}}K_2\sum_{i=1}^{n}\sum_{k=1}^{\ell}\bar{P}^{n-i}_{\ell,\theta}P^{k-1}_{\theta'}P^{\ell-k+1}_{\theta}\bar{P}^{i-1}_{\ell,\theta'}\left(\bar{x},\sum_{j=1}^{\ell}V^r(\cdot_j)\right) \\
&\leq \|\Delta\theta\|^{\widehat{\ell_3}}K_2C_Dn\ell^2V^r(x_\ell),\qquad \bar{x}=(x_0,x_1,\ldots,x_\ell)\in\mathsf{X}^{\ell+1}.
\end{aligned}$$

where the last line again follows from iterating (D1) as in (2.11) or (5.25). ■

5.11.4 Condition $(\overline{\text{E1}})$

Theorem 5.14 *Assume* (E1) *holds for the family* $\{P_\theta, \pi_\theta, \theta \in \Theta\}$ *with some function* $V : \mathsf{X} \to [1, \infty)$*. Then* $(\overline{\text{E1}})$ *is also satisfied for the windowed Markov chain* $\{\bar{X}_n : n = 1, 2, \ldots\}$ *where*

$$\bar{X}_n = (X_{n,0}, X_{n,1}, \ldots, X_{n,\ell}) \tag{5.29}$$

which is governed by $\{\bar{P}_\theta, \theta \in \Theta\}$*. Furthermore, the* ρ *in* $(\overline{\text{E1}})$ *is the same as that in* (E1)*.*

Proof: Under (E1) for the single-step chain governed by $\{P_\theta, \pi_\theta, \theta \in \Theta\}$ there exists some $C_E < \infty$ and $\rho < 1$ such that

$$\sup_{\theta \in \Theta} \|P_\theta^n(x, \cdot) - \pi_\theta(\cdot)\|_{V^r} \le C_E V^r(x) \rho^n, \qquad x \in \mathsf{X},\ n = 0, 1, \ldots.$$

Then for $\ell = 1, 2, \ldots$ and $n = 1, 2, \ldots$

$$\begin{aligned}
\sup_{\theta \in \Theta} \left\| \bar{P}_{\ell,\theta}^n(\bar{x}, \cdot) - \bar{\pi}_\theta(\cdot) \right\|_{\bar{V}_\ell^r} &= \sup_{\theta \in \Theta} \left\| \bar{P}_{\ell,\theta}^n(x_\ell, \cdot) - \bar{\pi}_\theta(\cdot) \right\|_{\bar{V}_\ell^r} \\
&\le \sum_{i=1}^{\ell} \sup_{\theta \in \Theta} \left\| P_\theta^{m(n-1)+i}(x_\ell, \cdot) - \pi_\theta(\cdot) \right\|_{V^r(\cdot)} \\
&\le C_E V^r(x_0) \sum_{i=1}^{\ell} \rho^{\ell(n-1)+i} \\
&\le C_E V^r(x_0) \rho^{\ell n} \sum_{i=1}^{\ell} \rho^{i-\ell} \\
&\le C_E V^r(x_0) \ell \rho^{\ell n} \\
&\le C_E V^r(x_0) \ell \rho^{n(\ell-1)} \rho^n \\
&\le C_E V^r(x_\ell) \ell \rho^n, \qquad \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1}.
\end{aligned}$$

For the case $n = 0$, we have

$$\sup_{\theta \in \Theta} \left\| \bar{P}_{\ell,\theta}^n(\bar{x}, \cdot) - \bar{\pi}_\theta(\cdot) \right\|_{\bar{V}_\ell^r} \le C_E \ell \bar{V}_\ell^r(\bar{x}).$$

■

5.11.5 Condition $(\overline{\text{P1}})$

Here, we check condition $(\overline{\text{P1}})$ for the full form of the gradient estimate. For each $i = 1, \ldots, p$ we must check the existence of the solution $\bar{\nu}_{\ell,\theta}^{(i)} : \Theta \times \mathsf{X}^{\ell+1} \to \mathbb{R}$ for the Poisson equation with forcing function $\widehat{G}_{i,\theta}(\bar{x})$:

$$\widehat{G}_{i,\theta}(\bar{x}) - \bar{\pi}_\theta(\widehat{G}_{i,\theta}) = \nu_{\ell,\theta}^{(i)}(\bar{x}) - \sum_{\bar{y} \in \mathsf{X}^{m+1}} \bar{P}_{\ell,\theta}(\bar{x}, \bar{y}) \nu_{\ell,\theta}^{(i)}(\bar{y}), \qquad \bar{x} \in \mathsf{X}^{\ell+1},\ \theta \in \Theta,$$

where $\widehat{G}_{i,\theta}(\bar{x}) = \widehat{G}_i(\theta, \bar{x})$ is the function used in the gradient estimate (4.32).

Condition ($\overline{\text{P1}}$) follows from($\overline{\text{E1}}$) and ($\overline{\text{H2}}$), both shown above, since for each $i = 1, \ldots, p$ they imply convergence of the sum for each $\ell = 1, 2, \ldots$

$$
\begin{aligned}
\bar{\nu}^{(i)}_{\ell,\theta}(\bar{x}) &\doteq \sum_{n=0}^{\infty} \left(\sum_{\bar{y} \in \mathsf{X}^{m+1}} \bar{P}^n_{\ell,\theta}(\bar{x}, d\bar{y}) \widehat{G}_{i,\theta}(\bar{y}) - \bar{\pi}(\widehat{G}_{i,\theta}) \right) \\
&\leq C_H \sum_{n=0}^{\infty} \left\| \bar{P}^n_{\ell,\theta}(\bar{x}, \cdot) - \bar{\pi}_\theta(\cdot) \right\|_{\bar{V}^r_\ell} \\
&< \infty, \qquad\qquad \bar{x} \in \mathsf{X}^{\ell+1},\ \theta \in \Theta.
\end{aligned}
$$

5.11.6 Condition ($\overline{\text{H5}}$)

To satisfy ($\overline{\text{H5}}$) , there must exist constants $\widehat{\ell}_2$ in $(\widehat{\ell}_1, 1)$, $\delta > 0$, and $C_5 < \infty$ so that for all θ, $\theta' \in \Theta$ such that $\|\theta - \theta'\| \leq \delta$, we have for $\ell = 1, 2, \ldots$

$$
\left\| \widehat{G}(\ell, \theta, \bar{x}) - \widehat{G}(\ell, \theta', \bar{x}) \right\| \leq C_5 \sum_{i=1}^{\ell} V^r(x_i) \left\| \theta - \theta' \right\|^{\widehat{\ell}_2}, \quad \text{for all } \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{m+1}.
$$

We shall propose two additional conditions which assume a common $\widehat{\ell}_2$ in $(\widehat{\ell}_1, 1)$ in both conditions:

(F4) There exists constants $C_4 < \infty$, $\delta_4 > 0$ and $\widehat{\ell}_2$ in $(\widehat{\ell}_1, 1)$ such that:

$$
\left| \frac{\partial f_{\theta+\Delta\theta}(x)}{\partial \theta} - \frac{\partial f_\theta(x)}{\partial \theta} \right| \leq C_4 \left\| \Delta\theta \right\|^{\widehat{\ell}_2} V^r(x)
$$

for all $x \in \mathsf{X}$ and $\theta, \theta' \in \Theta$ with $\|\Delta\theta\| \leq \delta_4$.

(G3) There exists constants $K_4 < \infty$, $\delta_G > 0$ and $\widehat{\ell}_2$ in $(\widehat{\ell}_1, 1)$ such that for all $i = 1, \ldots, p$ and all parameters $\theta, \theta' \in \Theta$ such that $\|\Delta\theta\| \leq \delta_G$ the transition probabilities partials satisfy

$$
\left| \frac{\partial}{\partial \theta_i} p_{x,y}(\theta) - \frac{\partial}{\partial \theta_i} p_{x,y}(\theta') \right| \leq K_4 \left\| \Delta\theta \right\|^{\widehat{\ell}_2} p_{x,y}(\theta), \qquad x, y \in \mathsf{X}. \tag{5.30}
$$

Theorem 5.15 *Assume* (F1)-(F4), (G1)-(G3). *Then, the gradient estimate* $\widehat{G}(\ell, \theta, \bar{x})$ *satisfies* ($\overline{\text{H5}}$) *in the form (5.30).*

Proof: First let δ_5 be defined as the minimum of 1, δ_4, δ_G, and δ which exists from (F3).

We shall treat the two main terms of the gradient estimate separately. Under (F4), the first term of $\widehat{G}$, given by

$$
\widehat{G}^{(1)}(\ell, \theta, \bar{x}) \doteq \frac{1}{\ell} \sum_{j=1}^{\ell} \frac{\partial f_\theta}{\partial \theta}(x_j), \qquad \bar{x} = (x_0, x_1, \ldots, x_\ell) \in \mathsf{X}^{\ell+1},\ \theta \in \Theta \tag{5.31}
$$

readily satisfies ($\overline{\text{H5}}$) for all $\theta, \theta' \in \Theta$ with $\|\theta - \theta'\| \leq \delta_5$.

Consider now the second term of the estimate:

$$\widehat{G}^{(2)}(\ell, \theta, \bar{x}) \doteq \frac{\nabla p_{x_0, X_1}(\theta)}{p_{X_0, X_1}(\theta)} \left(\sum_{j=1}^{\ell} f_\theta(X_j) \right), \qquad \bar{x} \in \mathsf{X}^{\ell+1}. \tag{5.32}$$

Recalling our convention that $\frac{0}{0} = 0$, we have for all parameters $\|\theta - \theta'\| \leq \delta_5$ and all $\bar{x} = (x_0, x_1, \dots x_\ell) \in \mathsf{X}^{\ell+1}$

$$\begin{aligned}
\left\| \widehat{G}^{(2)}(\ell, \theta, \bar{x}) - \widehat{G}^{(2)}(\ell, \theta', \bar{x}) \right\| &\leq \left\| \left\{ \frac{\nabla p_{x_0,x_1}(\theta)}{p_{x_0,x_1}(\theta)} - \frac{\nabla p_{x_0,x_1}(\theta')}{p_{x_0,x_1}(\theta')} \right\} \left(\sum_{j=1}^{\ell} f_\theta(x_j) \right) \right\| \\
&+ \left\| \frac{\nabla p_{x_0,x_1}(\theta')}{p_{x_0,x_1}(\theta')} \left\{ \left(\sum_{j=0}^{\ell} f_\theta(x_j) \right) - \left(\sum_{j=1}^{\ell} f_{\theta'}(x_j) \right) \right\} \right\| \\
&\leq C_1 \left\| \frac{\nabla p_{x_0,x_1}(\theta)}{p_{x_0,x_1}(\theta)} - \frac{\nabla p_{x_0,x_1}(\theta')}{p_{x_0,x_1}(\theta')} \right\| \sum_{j=1}^{\ell} V^r(x_j) \qquad (5.33) \\
&+ K_2 \left| \sum_{j=1}^{\ell} \left(f_\theta(x_j) - f_{\theta'}(x_j) \right) \right|. \qquad (5.34)
\end{aligned}$$

Next, we note that condition (G3) implies that if $p_{x,y}(\theta) = 0$ for some θ, then $\nabla p_{x,y}(\theta) = 0$ in a δ_G neighborhood of θ, hence we can safely rewrite the factor:

$$\begin{aligned}
&\left\| \frac{\nabla p_{x_0,x_1}(\theta)}{p_{x_0,x_1}(\theta)} - \frac{\nabla p_{x_0,x_1}(\theta')}{p_{x_0,x_1}(\theta')} \right\| \\
&\leq \left\| \frac{\nabla p_{x_0,x_1}(\theta) - \nabla p_{x_0,x_1}(\theta')}{p_{x_0,x_1}(\theta)} \right\| + \left\| \frac{\nabla p_{x_0,x_1}(\theta')}{p_{x_0,x_1}(\theta)} - \frac{\nabla p_{x_0,x_1}(\theta')}{p_{x_0,x_1}(\theta')} \right\| \\
&= \sqrt{p} K_4 \|\Delta\theta\|^{\widehat{\ell}_2} + \frac{\|\nabla p_{x_0,x_1}(\theta')\|}{p_{x_0,x_1}(\theta')} \frac{|p_{x_0,x_1}(\theta') - p_{x_0,x_1}(\theta)|}{p_{x_0,x_1}(\theta)} \\
&\leq \sqrt{p} K_4 \|\Delta\theta\|^{\widehat{\ell}_2} + \widetilde{K}_2 K_3 \|\theta - \theta'\| \\
&\leq K' \|\Delta\theta\|^{\widehat{\ell}_2}, \qquad \|\Delta\theta\| \leq \delta_5,
\end{aligned}$$

for some constant $K' < \infty$. Above, we have applied (G3) to bound the first term; and applied (4.9) and (G2) to bound the second term.

The term (5.34) is bounded easily from (F3) so there exists an $L < \infty$ such that

$$\begin{aligned}
\left| \sum_{j=1}^{\ell} \left(f_\theta(x_j) - f_{\theta'}(x_j) \right) \right| &\leq L \|\Delta\theta\| \sum_{j=1}^{\ell} V^r(x_j) \\
&\leq L \|\Delta\theta\|^{\widehat{\ell}_2} \sum_{j=1}^{\ell} V^r(x_j), \qquad \|\Delta\theta\| \leq \delta_5.
\end{aligned}$$

Therefore, for all θ in Θ with $\|\Delta\theta\| \leq \delta_5$ we combine the bounds for (5.33) and (5.34)

$$\begin{aligned}
\left\| \widehat{G}^{(2)}(\ell, \theta, \bar{x}) - \widehat{G}^{(2)}(\ell, \theta', \bar{x}) \right\| &\leq \|\Delta\theta\|^{\widehat{\ell}_2} (C_1 K' + L) \sum_{j=1}^{\ell} V^r(x_j), \\
&\qquad \bar{x} = (x_0, x_1, \dots, x_\ell) \in \mathsf{X}^{\ell+1}.
\end{aligned}$$

■

5.12 Example: Optimization of the Tandem M/M/1

We now demonstrate the SA-based stochastic optimization procedure we have just outlined in a very simple optimization problem. The performance measure, or objective function, we wish to minimize is the steady state total number of customers in the tandem queue consisting of two infinite buffer M/M/1 queues in series with adjustable service rates.

The first queue has a fixed arrival rate of $\lambda = 1$ and a parameterized service rate of

$$\mu_1(\theta) = \theta, \qquad \theta \in (1,3)$$

while the second queue, which takes customers exiting the first queue, has a service rate of

$$\mu_2(\theta) = 4 - \theta, \qquad \theta \in (1,3).$$

For this problem, since the arrival rate to the second queue also occurs at rate λ, the objective function has a simple closed form solution

$$J(\theta) = \frac{\lambda}{\mu_1(\theta) - \lambda} + \frac{\lambda}{\mu_2(\theta) - \lambda}, \qquad \theta \in (1,3),$$

and this is plotted in figure 5.1 as a function of θ. In this case, it's obvious the minimizing

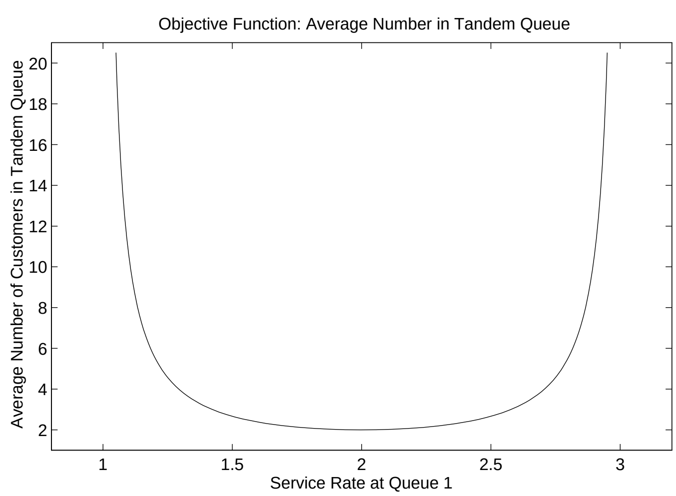

Figure 5.1: Steady state total number of customers in the tandem queue.

parameter is $\theta^\star = 2$, but nevertheless, let us perform a computer simulation of this tandem queue to see how the proposed derivative estimate performs when coupled with SA.

This simple optimization problem was also used by Meketon [75] who performed a nearly identical optimization problem (in an interesting survey paper on approaches to optimization) in which Meketon proposed SA with an IPA gradient estimate.

Looking closely at Meketon's simulation data in his graph [75], it is not in fact clear or convincing that those iterates are indeed converging to the optimal value in Meketon's simulated run, (with $\theta^\star = 3$ in this case). This is because for one, the simulation ends with the iterate at 3.24 with a slight upward drift, and second, the iterates remain on the high side of $\theta^\star = 3$ for all but the initial transient portion of the run. These observations could suggest a small bias present in this particular IPA estimate although it's certainly not clear from one simulation run. In any event, further study of this problem seems warranted.

Before we get to our simulation results let us remark that there are two aspects that could possibly make this a challenging system to to optimize via this SA method. First, the measurements we will be taking for our derivative estimate are unbounded as each queue has an infinite buffer and as such, it can lead to derivative estimates which are also unbounded. It's reasonable to conclude that unusually large derivative estimate could cause the next iterate to be forced a large distance from the previous one, and, if this were to occur when our parameter has essentially reached $\theta^\star$, then the "progress" the iterate had made towards the goal is undermined and the iterate essentially must start over. Second, notice from Figure 5.1 that the objective function is quite "flat" in the vicinity of $\theta^\star = 2$ so the performance derivative we will be estimating will not offer very much direction as the iterates approach this value $\theta^\star$. The trajectory of the mean ODE $\dot{\theta}(t) = -\frac{d}{d\theta}J(\theta(t))$ would thus approach $\theta^\star$ slowly as well, and we know from the ODE Method [61] to expect the iterates θ_n to approach the ODE trajectory. Thus, if we observe that the convergence is slow as the iterate nears $\theta^\star$, then this at least partially should be attributed to the "flatness" of the objective function we are seeking to optimize.

5.12.1 The Simulation Model and Optimization Algorithm

The sample path of the tandem M/M/1 queue is given by $\{(X_t^{(1)}, X_t^{(2)}), t \geq 0\}$ where $X_t^{(i)}$ represents the number of customers present in queue i at time t and is modeled as a continuous time Markov chain having infinitesimal generator $A_\theta = [a_{x,y}(\theta)]_{x,y}$. This system is *uniformizable* with the infinitesimal rates bounded uniformly over all $\theta \in \Theta$:

$$a \doteq \sup_{\theta\in\Theta} \sup_{x\in\mathsf{X}} |a_{x,x}(\theta)| < \infty, \qquad \theta \in \Theta.$$

Hence, we can uniformize [41, p. 118][92] and thus subordinate the continuous time Markov chain to a Poisson process of fixed rate a. Our approach will be to estimate the objective function by simulating the continuous time Markov chain as a discrete time Markov chain using the *equal holding time method* [56]. This has the advantage that it is not necessary to generate exponential holding times in the simulation. Thus, for each $\theta \in \Theta$, we get $a = \lambda + \mu_1(\theta) + \mu_2(\theta) = 5$ and the resulting discrete time Markov chain $\{(X_n^{(1)}, X_n^{(2)}),\ n = 0, 1, \ldots\}$, is governed by the one-step transition probabilities $P_\theta = I + A_\theta/a,\ \theta \in \Theta$.

Let us consider the probabilities corresponding to each transition event. Note the the simulation method has created some *null events*, i.e. transitions which result in revisiting the previous state with positive probability.

arrival at queue 1: For $(x^{(1)}, x^{(2)}) \in \{0, 1, \ldots\}^2$,

$$\mathbf{P}_\theta \left[(X_{n+1}^{(1)}, X_{n+1}^{(2)}) = (x^{(1)}+1, x^{(2)}) \middle| (X_n^{(1)}, X_n^{(2)}) = (x^{(1)}, x^{(2)}) \right] = \frac{\lambda}{5}$$

arrival at queue 2/departure from queue 1:
For $x^{(1)} = 1, 2, \ldots; x^{(2)} = 0, 1, \ldots,$

$$\begin{aligned}\mathbf{P}_\theta \left[(X_{n+1}^{(1)}, X_{n+1}^{(2)}) = (x^{(1)}-1, x^{(2)}+1) \middle| (X_n^{(1)}, X_n^{(2)}) = (x^{(1)}, x^{(2)}) \right] &= \frac{\mu_1(\theta)}{5} \\ &= \frac{\theta}{5}\end{aligned}$$

null departure from queue 1: For $x^{(2)} = 1, 2, \ldots$

$$\mathbf{P}_\theta \left[(X_{n+1}^{(1)}, X_{n+1}^{(2)}) = (0, x^{(2)}) \middle| (X_n^{(1)}, X_n^{(2)}) = (0, x^{(2)}) \right] = \frac{\mu_1(\theta)}{5} = \frac{\theta}{5},$$

null departure from empty state:

$$\begin{aligned}\mathbf{P}_\theta \left[(X_{n+1}^{(1)} = 0, X_{n+1}^{(2)}) = (0, 0) \middle| (X_n^{(1)}, X_n^{(2)}) = (0, 0) \right] &= \frac{\mu_1(\theta) + \mu_2(\theta)}{5} \\ &= \frac{4}{5},\end{aligned} \tag{5.35}$$

non-null departure from queue 2: For $x^{(1)} = 0, 1, \ldots, x^{(2)} = 1, 2, \ldots$

$$\begin{aligned}\mathbf{P}_\theta \left[(X_{n+1}^{(1)}, X_{n+1}^{(2)}) = (x^{(1)}, x^{(2)}-1) \middle| (X_n^{(1)}, X_n^{(2)}) = (x^{(1)}, x^{(2)}) \right] &= \frac{\mu_2(\theta)}{5} \\ &= \frac{4-\theta}{5}\end{aligned}$$

All remaining transitions have probability zero. The above events with nonzero probability can be represented by the following symbols:

$$\begin{aligned} A &= \text{Arrival at Queue 1} \\ D_1 &= \text{Departure from Queue 1 (and Arrival at Queue 2)} \\ N_1 &= \text{Null Departure from Queue 1} \\ D_2 &= \text{Departure form Queue 2} \\ N_2 &= \text{Null Departure from Queue 2} \end{aligned}$$

In the state $(X_n^{(1)}, X_n^{(2)}) = (0, 0)$, we do not distinguish between events N_1 and N_1 since the outcome is identical, and for greater efficiency, we simulate $N_1 \cup N_2$ at probability in (5.35).

For all states other than the zero state $(0,0)$, the likelihood ratios are computed:

$$\begin{aligned}
\frac{\frac{d}{d\theta}\mathbf{P}_\theta[A]}{\mathbf{P}_\theta[A]} &= \frac{0}{1/5} = 0 \\
\frac{\frac{d}{d\theta}\mathbf{P}_\theta[D_1]}{\mathbf{P}_\theta[D_1]} &= \frac{\frac{d}{d\theta}\mathbf{P}_\theta[N_1]}{\mathbf{P}_\theta[N_1]} = \frac{1/5}{\theta/5} = \frac{1}{\theta} \\
\frac{\frac{d}{d\theta}\mathbf{P}_\theta[D_2]}{\mathbf{P}_\theta[D_2]} &= \frac{\frac{d}{d\theta}\mathbf{P}_\theta[N_2]}{\mathbf{P}_\theta[N_2]} = \frac{-1/5}{(4-\theta)/5} = \frac{1}{\theta-4}
\end{aligned}$$

For the zero state $(0,0)$, the likelihood ratios are

$$\frac{\frac{d}{d\theta}\mathbf{P}_\theta[A]}{\mathbf{P}_\theta[A]} = \frac{0}{1/5} = 0, \qquad \frac{\frac{d}{d\theta}\mathbf{P}_\theta[N_1 \cup N_2]}{\mathbf{P}_\theta[N_1 \cup N_2]} = \frac{0}{4/5} = 0.$$

With that, we find the gradient estimate is

$$\widehat{G}(\theta, \bar{X}_{n+1}) \doteq \left\{ \frac{\frac{d}{d\theta}P_\theta(X_{n+1,0}, X_{n+1,1})}{P_\theta(X_{n+1,0}, X_{n+1,1})} \right\} \sum_{j=1}^{\ell_{n+1}} \left(X^{(1)}_{n+1,j} + X^{(2)}_{n+1,j} \right), \quad n = 1, 2, \ldots$$

where we take a window $\bar{X}_{n+1} = \left(X_{n+1,0}, \ldots, X_{n+1,\ell_{n+1}} \right)$.

A computer simulation was performed with this Markov chain using the projected algorithm:

$$\begin{aligned}
\theta_{n+1} &= \Pi_{[1.1,2.9]} \left\{ \theta_n - \gamma_{n+1} \widehat{G}(\theta_n, \bar{X}_{n+1}) \right\}, \quad n = 0, 1, \ldots \\
\theta_0 &= 1.25 \\
\bar{X}_0 &= (0,0)
\end{aligned}$$

We chose the rather arbitrary step-size sequence $\gamma_n = 1/(5+4n)$ for $n = 1, 2, \ldots$ and observation window sequence $\ell_n = 6\ln(n) + 5$ for $n = 1, 2, \ldots$ which satisfies ($\overline{\mathrm{S}}$). The results of the simulation are in Figures 5.2 and 5.3.

5.12.2 Simulation Results

Figure 5.2 shows the the early response of the stochastic optimization algorithm in two graphs; the lower graph shows the evolution of the service rate parameter θ_n at the first queue while the upper graph shows the computed theoretical objective function at each parameter value, i.e.

$$J(\theta_n) = \frac{\lambda}{\theta_n - \lambda} + \frac{1}{(4-\theta_n) - \lambda}, \qquad n = 1, 2, \ldots$$

During this early response, it's clear the noise is forcing the parameter over a wide range of values within the projection. The long term response of the algorithm in Figure 5.3 suggests asymptotic convergence of the iterates $\theta_n = \mu^{(1)}_n$ to the theoretically optimal value $\theta^\star = 2$.

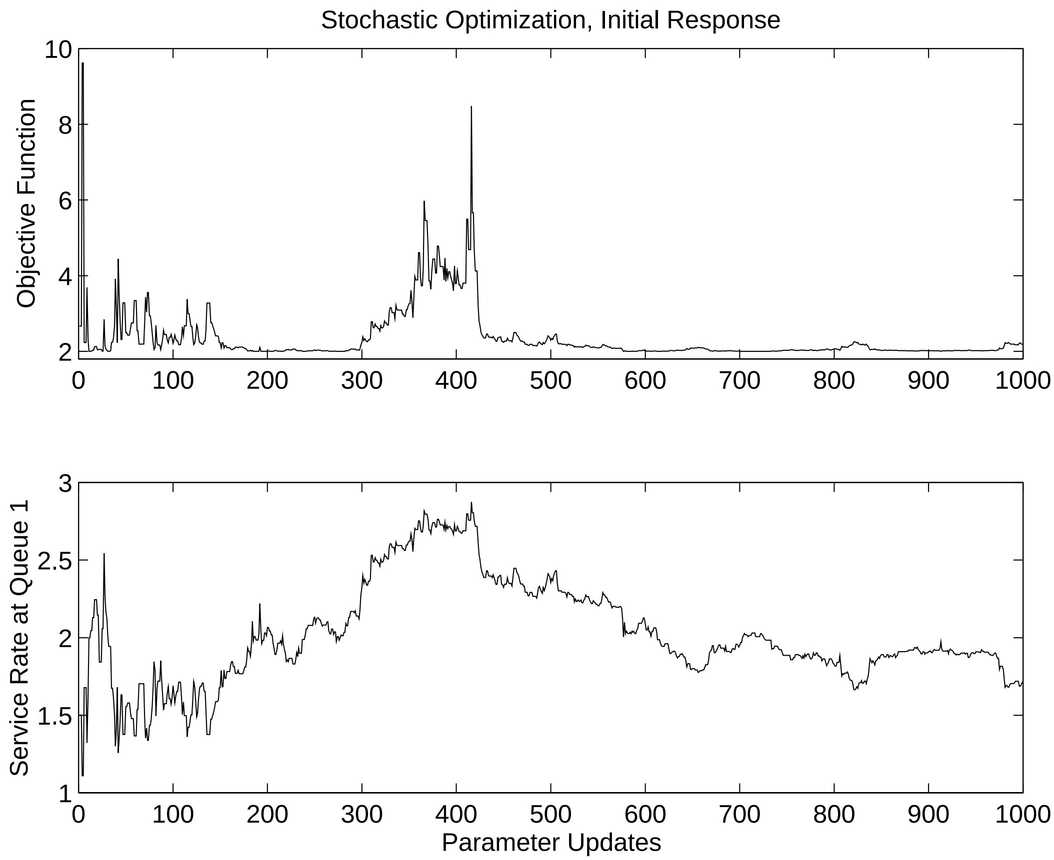

Figure 5.2: Early response of the simulated stochastic optimization algorithm.

5.12.3 Convergence Verification Procedure

In order to verify the conditions for convergence, we would start with showing (D1) or (D2) for the one-step chain for some function V which we need to find. Then we would check conditions (M), (E1), and (H2) as we did in Chapter 3 for the simple random walk example. Note that (G2) is equivalent to (M) here. For this example here, it can be verified that (D1) and (D2) follow (in the same manner as the random walk example) if we take

$$V(x^{(1)}, x^{(2)}) = Ke^{s(x^{(1)}+x^{(2)})}$$

for some constants $K < \infty$ and $s > 0$.

Then we check the conditions for the windowed chain used in the optimization algorithm. The results of this chapter then imply ($\overline{\text{D0}}$), ($\overline{\text{H2}}$), ($\overline{\text{C}}$), ($\overline{\text{E1}}$), ($\overline{\text{P1}}$) all hold for the windowed chain. The additional conditions (F1)-(F4) and (G1)-(G3) also follow readily and allow us to verify ($\overline{\text{H5}}$). The steps involved in checking these conditions are nearly identical to those carried out for the random walk example and the additional conditions (F1)-(F4) and (G1)-(G3) are straightforward to check.

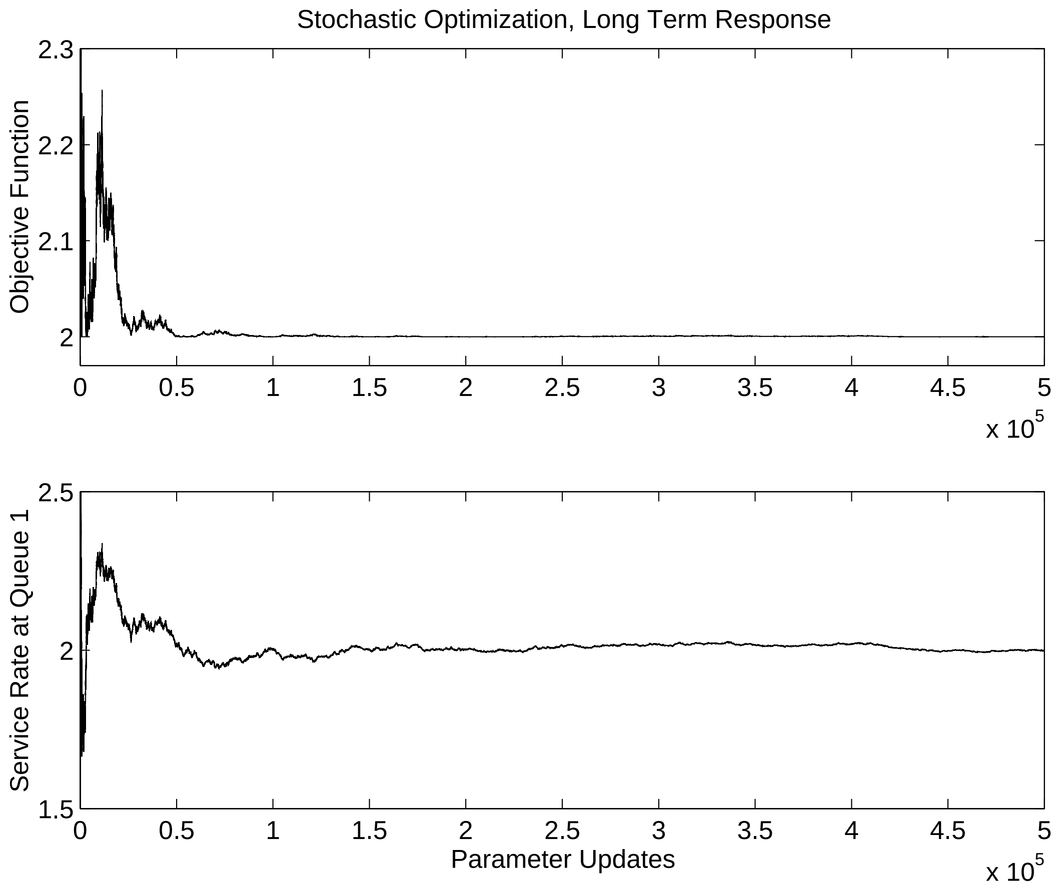

Figure 5.3: Long term response of the simulated stochastic optimization algorithm.

Chapter 6

An Algorithm for a Partially Transient Markov Chain

6.1 Introduction

The results of this chapter were first motivated by a difficulty we encountered when applying a projected SA to the M/M/1 queue. Consider an M/M/1 which has a fixed but *unknown* arrival rate $\lambda_{arrival}$ and we wish to iteratively approximate the location of the particular service rate parameter $\theta^\star = \mu^\star$ which achieves some given level of service defined by the steady-state mean queue size (or perhaps the steady-state mean waiting time). Since the arrival rate is unknown, we are unable to identify a compact projection set Θ to use with the SA algorithm which ensures the service rate iterates $\{\theta_n = \mu_n, n = 0, 1, \ldots\}$ remains within the positive recurrent region $(\lambda_{arrival}, \infty)$, i.e. we want to choose $\Theta \subset (\lambda_{arrival}, \infty)$. Unfortunately, if a non-ideal projection set is instead chosen which includes any part of the transient or null recurrent region $[0, \lambda_{arrival}]$, the framework developed in Chapter 2 breaks down for several reasons, one of which is simply the regression function $h(\cdot)$ is only defined on the interval $(\lambda_{arrival}, \infty)$ and we require existence as well as continuity of h on Θ.

Now consider the general SA problem and let us assume that all of the general conditions of Section 2.3 which lead to convergence of SA's (including positive recurrence) are satisfied over some subset D_s of $\mathbb{R}^p$. Except for unconstrained SA applications where the *recurrence condition* of the Kushner-Clark Lemma can be independently verified, the projected SA framework presented in Chapter 2 is most easily applied for cases where it's possible to identify an compact ideal projection set Θ_{ideal} which contains θ^* and lies within the ODE's domain of attraction $DA(\theta^\star)$. Since SA's are proposed for many "blind" equalization and regulation problems with unknown quantities, choosing an *ideal* projection set Θ_{ideal} for which $\theta^\star \in \Theta_{ideal}$, $\Theta_{ideal} \subset DA(\theta^\star)$, and $\Theta_{ideal} \subset D_s$ can be difficult or impossible. As a result, this difficulty often leads to a tradeoff in the choice of a suitable projection set Θ which approximates Θ_{ideal}. Choosing Θ to be "small" makes it more likely that $\Theta \subset D_s$ but can result in $\theta^\star \notin \Theta$ which prevents convergence to $\theta^\star$. Choosing Θ "large" makes it more likely that $\theta^\star$ falls in Θ but can result in Θ not being a subset of D_s. Our goal in this chapter is to develop an approach to proving convergence when Θ is chosen "large" which takes into account the possibility that Θ is not a subset of D_s. In this case, there may be a region within Θ where some of the general convergence conditions of Section 2.3 are not satisfied.

For the particular M/M/1 queue example above, we may easily choose a "large" compact set $\Theta \supset \Theta_{ideal}$ given by $\Theta = [0, M]$ where M is chosen so large that we are absolutely confident $\theta^\star << M$. We wish to study the behavior and convergence properties of this algorithm which generates iterates $\theta_n = \mu_n$ controlling the service rate which may fall in the transient/null-recurrent region $[0, \lambda_{arrival}]$ as well as the positive-recurrent region $(\lambda_{arrival}, M]$ of Θ.

In general, we study an extension to the framework in Chapter 2 where we wish to apply SA to a system where it is only assumed the convergence conditions are satisfied when the parameter θ is restricted to some compact and strict subset Q of $D_s \cap \Theta$. We find that the SA with "large projection set" can be applied to Markov chains which may contain a transient region within Θ and the desired a.s. convergence can still be shown provided one additional condition is satisfied; that, almost surely, the iterate θ_n always returns to this compact subset Q if it should ever leave Q. In this setting, the general conditions of Section 2.3 are only required to hold on a localized subset Q of $D_s \cap \Theta$ and we do not require that the regression function h be defined outside the subset D_s, i.e. we take $h : D_s \to \mathbb{R}^p$.

For the M/M/1 example above, we show that this key recurrence condition on the parameter iterates holds when the performance measure is the queue occupancy; and when combined with a *localized* ODE method we prove almost sure convergence $\theta_n \to \theta^\star$ with no unverified conditions. While this localization approach is not new (it is discussed in [64]), we do show in detail how the various steps may be carried out for a particular system and we feel this offers several insights as well as tools which should apply to other systems with properties similar to this M/M/1 example. The particular property that we exploit is that if any iterate should cause the queue to be set to operate in the transient region, the queue tends to grow unbounded as long as the iterate remains in this transient region and the response of the SA algorithm driven by the queue occupancy then tends to drive the iterate back to the positive recurrent subset Q almost surely. This recurrence property is related to the structure of the chain and therefore holds without *a priori* knowledge of the boundaries of Q or the value of $\lambda_{arrival}$.

6.2 Localized Conditions

Let us now define a new set of *localized conditions* which will be used later to show a.s. convergence of θ_n to $\theta^\star$ via a localized ODE method when accompanied with a parametric recurrence argument shown for the specific problem.

For this general setting, Θ is a compact subset of $\mathbb{R}^p$ used in the projection operator. The following conditions are nearly identical to those presented in Chapter 2 except here they hold only on a compact subset Q of Θ and unlike Chapter 2, here we simply require $0 \le r \le 1/4$. We still assume $\widehat{\ell}_1$ is in $(0, 1)$ and satisfies (S).

(D0') There exists a function $V : \mathsf{X} \to [1, \infty)$ and a constant $1 \le C_D < \infty$ such that for all $n = 0, 1, 2, \ldots$:

$$\mathbf{E}_{\theta,x}\left[1_{\{\theta_n \in Q\}} V(X_n)\right] \;\le\; C_D V(x), \qquad \text{and} \tag{6.1}$$

$$\mathbf{E}_{\theta,x}\left[1_{\{\theta_n\in Q\}}V(X_{n+1})\right] \;\le\; C_D V(x), \qquad \theta\in Q,\ x\in \mathsf{X}. \tag{6.2}$$

(D1') There exists a function $V : \mathsf{X} \to [1,\infty)$ and two constants $0 < \lambda < 1$ and $L < \infty$ such that

$$P_\theta V(x) \le \lambda V(x) + L, \qquad \text{for all } \theta \text{ in } Q \text{ and } x \text{ in } \mathsf{X}.$$

(π') For each $\theta \in Q$, there exists a unique P_θ-invariant probability π_θ on $(\Omega, \mathcal{F})$.

(H1') For all $\theta \in Q$, $H(\theta,\cdot) \doteq H_\theta(\cdot)$ is *integrable* under π_θ. Let us denote

$$h(\theta) \doteq \pi_\theta(H_\theta) \doteq \int_{\mathsf{X}} H_\theta(x)\pi_\theta(dx), \qquad \theta\in Q.$$

(H2') There exists constants $C_H < \infty$ and $C_\rho < \infty$ such that,

$$\begin{aligned}
\sup_{\theta\in Q}\|H(\theta,x)\| &\le C_H V^r(x), &&\text{for all } x\in\mathsf{X}\\
\sup_{\theta\in Q}\|\rho_n(\theta,x)\| &\le C_\rho V^r(x), &&\text{for all } x\in\mathsf{X}, \qquad n=1,2,\ldots
\end{aligned}$$

(H4') There exists a $C_h < \infty$ and a $\delta > 0$ such that for some $0 < \ell < 1$:

$$\|h(\theta) - h(\theta')\| \le C_h \|\theta-\theta'\|^{\ell}, \qquad \theta,\theta'\in Q,\ \|\theta-\theta'\|\le\delta.$$

(P1') For all $(\theta, x) \in Q \times \mathsf{X}$, the following series converges:

$$\nu_\theta(x) \doteq \sum_{n=0}^{\infty}\left(\int P_\theta^n(x,dy)H_\theta(y) - h(\theta)\right) < \infty,$$

and we identify $\nu_\theta(x)$ as the solution to the Poisson equation associated with $H(\theta,\cdot)$:

$$H_\theta(x) - h(\theta) = \nu_\theta(x) - \int P_\theta(x,dy)\nu_\theta(y), \qquad x\in\mathsf{X},\ \theta\in Q.$$

(P2') There exist a constant $C_\nu < \infty$ such that

$$\begin{aligned}
\|\nu_\theta(x)\| &\le C_\nu V^r(x), &&\text{for all } \theta\in Q, \quad x\in\mathsf{X}\\
\|P_\theta\nu_\theta(x)\| &\le C_\nu V^r(x), &&\text{for all } \theta\in Q, \quad x\in\mathsf{X}
\end{aligned}$$

(P3') There exists a constant $C_\delta < \infty$ such that

$$\|P_\theta\nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\| \le C_\delta V^r(x)\,\|\theta-\theta'\|^{\widehat{\ell}_1}, \qquad \text{for all } \theta,\theta'\in Q, \quad x\in\mathsf{X}.$$

where $\widehat{\ell}_1$ satisfies (S).

We may also have occasion to use localized versions of the *specialized conditions* of Chapter 3, the definitions of which are obvious.

6.2.1 An Immediate Consequence

Proposition 6.1 *Suppose there exists a function $V : \Theta \to [1, \infty)$ such that condition (6.1) of (D0') and (D1') both hold. Then both conditions of (D0') hold, i.e. that there exists a $C_D < \infty$ such that for all $n = 0, 1, 2, \ldots$:*

$$\begin{aligned}\mathbf{E}[V(X_{n+1})1_{\{\theta_n \in Q\}}|X_0 = x, \theta_0] &\leq C_D V(x), && \textit{and}\\ \mathbf{E}[V(X_n)1_{\{\theta_n \in Q\}}|X_0 = x, \theta_0] &\leq C_D V(x), && \textit{for all } \theta \textit{ in } Q \subset \Theta,\ x \textit{ in } \mathsf{X}.\end{aligned}$$

Proof: Suppose condition (6.1) is satisfied with $C'_D < \infty$ Conditioning and applying (D1'):

$$\begin{aligned}\mathbf{E}[V(X_{n+1})1_{\{\theta_n \in Q\}}|X_0, \theta_0] &= \mathbf{E}[\mathbf{E}[V(X_{n+1})|X_n, \theta_n]1_{\{\theta_n \in Q\}}|X_0, \theta_0]\\ &\leq \mathbf{E}[(\lambda V(X_n) + L)1_{\{\theta_n \in Q\}}|X_0, \theta_0]\\ &\leq \lambda \mathbf{E}[V(X_n)1_{\{\theta_n \in Q\}}|X_0, \theta_0] + L\\ &\leq \lambda C'_D V(x) + L\\ &\leq (C'_D + L)V(x)\end{aligned}$$

Now define $C_D = C'_D + L$. ■

6.3 A Noise Decomposition for Localized Domains

This section carries out a decomposition similar to what was used in Chapter 2 and is based on the decomposition and framework developed in [6]. We have slightly adapted BMP's decomposition to meet our needs for projected algorithms containing a localization Q within the projection set. The reader will notice the form and development of the decomposition is essentially identical to BMP's yet the algorithm and conditions for which the decomposition is valid are indeed different.

Consider the same Θ-*projected* stochastic approximation algorithm studied earlier in the previous chapters. We assume existence of a compact subset $Q \subset \Theta$ which we refer to as the *localization* on Θ. For convenience, the SA algorithm is initialized with an arbitrary parameter $\theta_0 = \theta$ in Q. The familiar recursion is defined for all $n = 0, 1, 2, \ldots$ by:

$$\begin{aligned}\theta_{n+1} &= \Pi_\Theta \left\{\theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}^2 \rho_{n+1}(\theta_n, X_{n+1})\right\} && (6.3)\\ &= \theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}^2 \rho_{n+1}(\theta_n, X_{n+1}) + \gamma_{n+1} z_{n+1}, && (6.4)\end{aligned}$$

and the deterministic step-size sequence $\{\gamma_n, n = 1, 2, \ldots\}$ is chosen to satisfy (S).

The *state process* $\mathcal{X} = \{X_n, n = 1, 2, \ldots\}$ behaves as a controlled Markov chain (see Section 2.2) in which the one step transition kernel $P_\theta(x, \cdot)$, which may depend on the continuous variable $\theta \in \Theta$, is controlled by the current iterate θ_n so X_{n+1} is governed by the one-step probability P_{θ_n, X_n}. Additionally, it is assumed that a generic time-homogeneous Markov chain governed by the same transition kernel $P_\theta(x, \cdot)$ with the parameter θ held fixed at any point in Q is ergodic in the sense that there exists a P_θ-invariant measure π_θ and

$$\lim_{n \to \infty} \mathbf{E}_{\theta,x}\left[H(\theta, X_n)\right] = \pi_\theta(H_\theta) \doteq h(\theta), \qquad \theta \in Q. \tag{6.5}$$

We assume the regression function $\|h(\theta)\| < \infty$ for all $\theta \in Q$, hence for any n such that $\theta_n \in Q$, we can write (6.4) as:

$$\begin{aligned}
\theta_{n+1} &= \theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1} z_{n+1} \\
&\quad + \gamma_{n+1} \left\{ H(\theta_n, X_{n+1}) - h(\theta_n) + \gamma_{n+1} \rho_{n+1}(\theta_n, X_{n+1}) \right\} \\
&= \theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1} z_{n+1} + \gamma_{n+1} \varepsilon_{n+1},
\end{aligned} \tag{6.6}$$

Rearranging we find

$$\gamma_{n+1}\varepsilon_{n+1} = \theta_{n+1} - \theta_n - \gamma_{n+1} h(\theta_n) - \gamma_{n+1} z_{n+1}; \tag{6.7}$$

this being valid for $\theta_n \in Q$ only, as we assume h is not defined on Q^c.

Following [6], we define the C^2 function $\phi : \Theta \to \mathbb{R}$ which has a bounded second derivative and consider the following generalization of (6.7):

$$\gamma_{n+1}\varepsilon_{n+1}(\phi) \doteq \phi(\theta_{n+1}) - \phi(\theta_n) - \gamma_{n+1} \nabla\phi(\theta_n) \cdot \{h(\theta_n) + z_{n+1}\} \tag{6.8}$$

For the compact set Q of Θ, BMP also define

$$\begin{aligned}
M_0 &\doteq \sup_{\theta \in Q} |\phi(\theta)| \\
M_1 &\doteq \sup_{\theta \in Q} \|\nabla\phi(\theta)\| \\
M_2 &\doteq \sup_{\theta \in Q} \left\|\nabla^2\phi(\theta)\right\| \\
\bar{M}_2 &\doteq \sup_{\theta \in \Theta} \left\|\nabla^2\phi(\theta)\right\|
\end{aligned}$$

Given two points $\theta, \theta' \in \Theta$, there exists a remainder $R(\phi, \theta, \theta')$ such that

$$\phi(\theta) - \phi(\theta') = (\theta - \theta') \cdot \nabla\phi(\theta) + R(\phi, \theta, \theta') \tag{6.9}$$

whereby $|R(\phi, \theta, \theta')| \le \bar{M}_2 \|\theta - \theta'\|^2$.

If we take $\theta_k \in Q$ (and we only know $\theta_{k+1} \in \Theta$), then from (6.7), (6.8) and (6.9),

$$\begin{aligned}
&\gamma_{k+1}\varepsilon_{k+1}(\phi) \\
&\quad = \phi(\theta_{k+1}) - \phi(\theta_k) - \gamma_{k+1}\nabla\phi(\theta_k) \cdot (h(\theta_k) + z_{k+1}) \\
&\quad = \nabla\phi(\theta_k) \cdot (\theta_{k+1} - \theta_k) - \gamma_{k+1}\nabla\phi(\theta_k) \cdot (h(\theta_k) + z_{k+1}) + R(\phi, \theta_k, \theta_{k+1}) \\
&\quad = \gamma_{k+1}\nabla\phi(\theta_k) \cdot \{H(\theta_k, X_{k+1}) - h(\theta_k)\} \\
&\qquad + \gamma_{k+1}^2 \nabla\phi(\theta_k) \cdot \rho_{k+1}(\theta_k, X_{k+1}) + R(\phi, \theta_k, \theta_{k+1})
\end{aligned} \tag{6.10}$$

where

$$\begin{aligned}
|R(\phi, \theta_k, \theta_{k+1})| &\le \bar{M}_2 \left\| \gamma_{k+1} H(\theta_k, X_{k+1}) + \gamma_{k+1}^2 \rho_{k+1}(\theta_k, X_{k+1}) + \gamma_{k+1} z_{k+1} \right\|^2 \\
&\le \bar{M}_2 \left(2 \left\| \gamma_{k+1} H(\theta_k, X_{k+1}) + \gamma_{k+1}^2 \rho_{k+1}(\theta_k, X_{k+1}) \right\| \right)^2 \\
&= 4\bar{M}_2 \gamma_{k+1}^2 \left\| H(\theta_k, X_{k+1}) + \gamma_{k+1} \rho_{k+1}(\theta_k, X_{k+1}) \right\|^2 .
\end{aligned} \tag{6.11}$$

For the fixed compact set Q and for each $i = 0, 1, 2, \ldots$ let us define the i^{th} *exit time* and the i^{th} *entrance time*, respectively

$$\begin{aligned}
\tau_i(Q) &\doteq \inf\{n \geq \upsilon_{i-1}(Q) : \theta_n \notin Q\} && (6.12)\\
\upsilon_i(Q) &\doteq \inf\{n \geq \tau_i(Q) : \theta_n \in Q\}, \qquad i = 1, 2, \ldots, && (6.13)
\end{aligned}$$

where by convention we define $\tau_0 = \upsilon_0 = 0$. Recall our assumption that $\theta_0 \in Q$.

For some $i = 1, 2, \ldots$, let us take $\upsilon_{i-1}(Q) \leq k < \tau_i(Q)$ so (6.10) is valid, and assuming condition (P1') so that the solution to the Poisson equation exists on Q, we reformulate (6.10) using some new terms $\widetilde{\varepsilon}^{(i)}_{k+1}$ (defined below in 6.15):

$$\begin{aligned}
\varepsilon_{k+1}(\phi) &= \nabla\phi(\theta_k) \cdot \{H(\theta_k, X_{k+1}) - h(\theta_k)\} + \widetilde{\varepsilon}^{(4)}_{k+1}\\
&= \nabla\phi(\theta_k) \cdot \{\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_{k+1})\} + \widetilde{\varepsilon}^{(4)}_{k+1}\\
&= \nabla\phi(\theta_k) \cdot \{\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)\}\\
&\quad + \nabla\phi(\theta_k) \cdot \{P_{\theta_k}\nu_{\theta_k}(X_k) - P_{\theta_k}\nu_{\theta_k}(X_{k+1})\} + \widetilde{\varepsilon}^{(4)}_{k+1}\\
&= \widetilde{\varepsilon}^{(1)}_{k+1} + \nabla\phi(\theta_k) \cdot \{P_{\theta_k}\nu_{\theta_k}(X_k) - P_{\theta_k}\nu_{\theta_k}(X_{k+1})\} + \widetilde{\varepsilon}^{(4)}_{k+1}
\end{aligned}$$

We retain BMP's notation and define the function $\psi_\theta(x) \doteq \nabla\phi(\theta) \cdot P_\theta\nu_\theta(x)$. For any $i = 1, 2, \ldots$, and for m, n such that $\upsilon_{i-1}(Q) \leq m < n \leq \tau_i(Q)$, the weighted sum of the noise terms is formed and then rearranged:

$$\begin{aligned}
&\sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1}(\phi)\\
&= \sum_{k=m}^{n-1} \gamma_{k+1}\left(\widetilde{\varepsilon}^{(1)}_{k+1} + \widetilde{\varepsilon}^{(4)}_{k+1}\right) + \sum_{k=m}^{n-1} \gamma_{k+1}\left\{\psi_{\theta_k}(X_k) - \psi_{\theta_k}(X_{k+1})\right\}\\
&= \sum_{k=m}^{n-1} \gamma_{k+1}\left(\widetilde{\varepsilon}^{(1)}_{k+1} + \widetilde{\varepsilon}^{(4)}_{k+1}\right) + \gamma_{m+1}\psi_{\theta_m}(X_m) + \sum_{k=m+1}^{n-1} \gamma_{k+1}\left\{\psi_{\theta_k}(X_k) - \psi_{\theta_{k-1}}(X_k)\right\}\\
&\quad + \sum_{k=m+1}^{n-1} (\gamma_{k+1} - \gamma_k)\psi_{\theta_{k-1}}(X_k) - \gamma_n\psi_{\theta_{n-1}}(X_n)
\end{aligned}$$

We thus have the following decomposition (which is quite similar in appearance to BMP's Lemma 1 [6, p. 222]):

$$\begin{aligned}
&\sum_{k=m}^{n-1} \gamma_{k+1}\varepsilon_{k+1}(\phi) && (6.14)\\
&= \sum_{k=m}^{n-1} \gamma_{k+1}\widetilde{\varepsilon}^{(1)}_{k+1} + \sum_{k=m+1}^{n-1} \gamma_{k+1}\widetilde{\varepsilon}^{(2)}_{k+1} + \sum_{k=m+1}^{n-1} \gamma_{k+1}\widetilde{\varepsilon}^{(3)}_{k+1} + \sum_{k=m}^{n-1} \gamma_{k+1}\widetilde{\varepsilon}^{(4)}_{k+1} + \widetilde{\eta}_{m;n}
\end{aligned}$$

where

$$\widetilde{\varepsilon}^{(1)}_{k+1} \doteq \nabla\phi(\theta_k) \cdot \{\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)\}$$

$$
\begin{aligned}
\widetilde{\varepsilon}^{(2)}_{k+1} &\doteq \psi_{\theta_k}(X_k) - \psi_{\theta_{k-1}}(X_k) \\
\widetilde{\varepsilon}^{(3)}_{k+1} &\doteq \frac{\gamma_{k+1} - \gamma_k}{\gamma_{k+1}} \psi_{\theta_{k-1}}(X_k) \\
\widetilde{\varepsilon}^{(4)}_{k+1} &\doteq \gamma_{k+1} \nabla\phi(\theta_k) \cdot \rho_{k+1}(\theta_k, X_{k+1}) + \frac{R(\phi, \theta_k, \theta_{k+1})}{\gamma_{k+1}} \\
\widetilde{\eta}_{m;n} &\doteq \gamma_{m+1}\psi_{\theta_m}(X_m) - \gamma_n \psi_{\theta_{n-1}}(X_n).
\end{aligned} \tag{6.15}
$$

6.3.1 A Partial Decomposition

In the coming sections, we shall also be interested in convergence properties of the following *partial decomposition sum* for **any** $n = 1, 2, \ldots$:

$$
\begin{aligned}
E_n \doteq\; & \sum_{k=0}^{n-1} \gamma_{k+1} \widetilde{\varepsilon}^{(1)}_{k+1} 1_{\{\theta_k \in Q\}} + \sum_{k=1}^{n-1} \gamma_{k+1} \widetilde{\varepsilon}^{(2)}_{k+1} 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}} \\
& + \sum_{k=1}^{n-1} \gamma_{k+1} \widetilde{\varepsilon}^{(3)}_{k+1} 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}} + \sum_{k=0}^{n-1} \gamma_{k+1} \widetilde{\varepsilon}^{(4)}_{k+1} 1_{\{\theta_k \in Q\}}.
\end{aligned} \tag{6.16}
$$

Note we take the convention that any summation $\sum_i^j$ is defined to be zero if $i > j$.

The decomposition is termed *partial* because the final $\widetilde{\eta}$ term in (6.14) is not present here and we will account for it separately when we work with this form of the decomposition. Observe that for any $i = 1, 2, \ldots$,

$$
\sum_{k=m}^{n-1} \gamma_{k+1} \varepsilon_{k+1}(\phi) = E_n - E_m + \widetilde{\eta}_{m;n}, \qquad \upsilon_{i-1}(Q) = m < n \le \tau_i(Q). \tag{6.17}
$$

6.4 Localized Versions of the BMP Lemmas

This section presents a series of five lemmas, analogous to those in Chapter 2 which bound each term in the decomposition of $\varepsilon(\phi)$ under localized conditions and with the greater generality of passing the noise through the function ϕ which was introduced in this chapter's decomposition. Since the proofs of these lemmas are similar to the versions in Chapter 2, we have placed them in the appendix.

Lemma 6.2 (Variant of BMP Lemma 2) *Assume (D0'), (P1'), (P2') for some positive $r \le 1/4$ and a set $Q \subset \Theta$ so that $\tau = \tau(Q)$.*

1. *There exists a constant $A_1 < \infty$ such that for each $m = 1, 2, \ldots$*

$$
\mathbf{E}_{\theta,x}\left[\sup_{n \le m} 1_{\{n \le \tau\}} \left| \sum_{k=0}^{n-1} \gamma_{k+1} \widetilde{\varepsilon}^{(1)}_{k+1} \right|^2 \right] \le A_1 V(x) \sum_{k=0}^{m-1} \gamma_{k+1}^2, \qquad x \in \mathsf{X}, \quad \theta \in Q.
$$

   Moreover, $A_1 \le 4 C_\nu^2 M_1^2 C_D$.

2. *On $\{\tau(Q) = \infty\}$, $\sum_{k=0}^{n-1} \gamma_{k+1} \widetilde{\varepsilon}^{(1)}_{k+1}$ converges $\mathbf{P}_{\theta,x}$-a.s. and in L^2 if $\sum_{k=0}^{\infty} \gamma_{k+1}^2 < \infty$.*

3. *The sum* $\sum_{k=0}^{n-1} \gamma_{k+1} 1_{\{\theta_k \in Q\}} \widehat{\varepsilon}_{k+1}^{(1)}$ *converges* $\mathbf{P}_{\theta,x}$*-a.s. and in* L^2 *if* $\sum_{k=0}^{\infty} \gamma_{k+1}^2 < \infty$.

Proof: See the appendix. ∎

Lemma 6.3 (Variant of BMP Lemma 3) *Assume (D0'), (H2'), (P1'), (P2'), (P3') for some positive* $r \leq \frac{1}{4}$ *and a set* $Q \subset \Theta$.

1. *There exists a constant* $A_2 < \infty$ *such that for all* $m = 1, 2, \ldots$,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1} \gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \leq A_2 V(x)\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)^2, \qquad x \in \mathsf{X}, \quad \theta \in Q,$$

where, $A_2 \leq C_D\left(8M_2^2 C_\nu^2 \left(C_H + \gamma_1 C_{\rho 3}\right)^2 \left(\frac{\gamma_1^2}{\gamma_1^{(1+\widehat{\ell}_1)}}\right)^2 + 8M_1^2 C_\delta^2 \left(C_H + \gamma_1 C_{\rho 3}\right)^{2\widehat{\ell}_1}\right)$.

2. *There exists a constant* $B_2 < \infty$ *such that for all* $m, n = 1, 2, \ldots$ *such that* $m > n$,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1} \gamma_{k+1} 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \;\leq\; B_2 V(x)\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}\right)^2, \qquad x \in \mathsf{X}, \quad \theta \in Q,$$

where, $B_2 \leq C_D\left(8M_2^2 C_\nu^2 \left(C_H + \gamma_1 C_{\rho 3}\right)^2 \left(\frac{\gamma_1^2}{\gamma_1^{(1+\widehat{\ell}_1)}}\right)^2 + 8M_1^2 C_\delta^2 \left(C_H + \gamma_1 C_{\rho 3}\right)^{2\widehat{\ell}_1}\right)$.

Proof: See the appendix. ∎

Lemma 6.4 (Variant of BMP Lemma 4) *Assume (D0'), (P1'), (P2') for some positive* $r \leq \frac{1}{4}$ *and* $Q \subset \Theta$.

1. *There exists a constant* $A_3 < \infty$ *such that for all* $m = 1, 2, \ldots,$,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1} \gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(3)}\right|\right)^2\right] \leq A_3 V(x)\gamma_1^2, \qquad x \in \mathsf{X}, \quad \theta \in Q.$$

   Moreover, $A_3 \leq M_1^2 C_\nu^2 C_D$.

2. *There exists a constant* $B_3 < \infty$ *such that for all* $m, n = 1, 2, \ldots,$ *with* $m < n$,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1} \gamma_{k+1} 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}}\left|\widehat{\varepsilon}_{k+1}^{(3)}\right|\right)^2\right] \leq B_3 V(x)\gamma_m^2, \qquad x \in \mathsf{X}, \quad \theta \in Q.$$

   Moreover, $B_3 \leq M_1^2 C_\nu^2 C_D$.

Proof: See the appendix. ∎

Lemma 6.5 (Variant BMP Lemma 5) *Assume (D0'), (P1'), (H2') for some positive* $r \le \frac{1}{4}$ *and* $Q \subset \Theta$.

1. *There exists a constant* $A_4 < \infty$ *such that for all* $m = 1, 2, \ldots$,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m\wedge\tau-1}\gamma_{k+1}\left|\widetilde{\varepsilon}^{(4)}_{k+1}\right|\right)^2\right] \le A_4 V(x)\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)^2, \qquad x \in \mathsf{X}, \quad \theta \in Q.$$

   Moreover, $A_4 \le C_D(C_{\rho 3} + 2M_2 C_H^2 + 2\gamma_1^2 C_{\rho 3}^2)^2$.

2. *There exists a constant* $B_4 < \infty$ *such that for all* $m = 0, 1, 2, \ldots$ *with integer* $n > m$

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}\left|\widetilde{\varepsilon}^{(4)}_{k+1}\right|\right)^2\right] \le B_4 V(x)\left(\sum_{k=m}^{n-1}\gamma_{k+1}^2\right)^2, \qquad x \in \mathsf{X}, \quad \theta \in Q.$$

   Moreover, $B_4 \le C_D(C_{\rho 3} + 2M_2 C_H^2 + 2\gamma_1^2 C_{\rho 3}^2)^2$.

Proof: See the appendix. ■

Lemma 6.6 (Variant of BMP Lemma 6) *Assume (D0'), (P1'), (P2') for some* $r \le \frac{1}{4}$ *and* $Q \subset \Theta$.

1. *There exists a constant* $A_5 < \infty$ *such that for each* $m = 1, 2, \ldots$,

$$\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m}\left|\widetilde{\eta}_{0;n}\right|^2\right] \le A_5 V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2, \qquad x \in \mathsf{X}, \quad \theta \in Q.$$

   Moreover, $A_5 \le 4C_D M_1^2 C_\nu^2$.

2. *On* $\{\tau(Q) = \infty\}$, $\widetilde{\eta}_{0;n}$ *converges a.s. and in* L^2 *if* $\sum \gamma_{k+1}^2 < \infty$.

Proof: See the appendix. ■

6.5 Main Properties of the Noise

In this section, we collect the results of each lemma for each term in the decomposition of $\varepsilon(\phi)$ and show several bounds related to the "step-size weighted sum of error" sequence (6.15). We let $\tau(Q) = \tau_1(Q)$.

Proposition 6.7 (BMP Prop. 7) *Assume (D0'), (P1')-(P3'), and (H1')-(H2') hold for some positive* $r \le \frac{1}{4}$ *and* $Q \subset \Theta$. *There exist constants* B_1, B_2, B_3, B_4 *such that for all* $m = 1, 2, \ldots$*:*

1. *We have*

$$\mathbf{E}_{\theta,x}\left[\sup_{n\le m}1_{\{n\le\tau(Q)\}}\left|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}(\phi)\right|^2\right] \le B_1 V(x)\left(1 + \sum_{k=0}^{m-1}\gamma_{k+1}^{2\widehat{\ell}_1}\right)\sum_{k=0}^{m-1}\gamma_{k+1}^2. \tag{6.18}$$

2. *We have*

$$\mathbf{E}_{\theta,x}\left[\sup_{n\le m} 1_{\{n\le\tau(Q)\}}\left|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}(\phi)\right|^2\right] \le B_2 V(x)\left(\gamma_1^{1-\widehat{\ell}_1}+\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1}\right)\sum_{k=0}^{m-1}\gamma_{k+1}^{1+\widehat{\ell}_1} \quad (6.19)$$

3. *On* $\{\tau(Q)=\infty\}$*, and if* $\sum_{k=0}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1}<\infty$*:*

   (a)

$$\mathbf{E}_{\theta,x}\left[\sup_{n\ge 1}\left|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}(\phi)\right|^2\right] \le B_3 V(x)\sum_{k=0}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1}. \quad (6.20)$$

   (b) *the series* $\sum_{k=0}^{\infty}\gamma_{k+1}\varepsilon_{k+1}(\phi)$ *converges* $\mathbf{P}_{\theta,x}$*-a.s. and in* L^2.

4. *If* $\sum_{k=0}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1}<\infty$*,*

   (a) *The partial decomposition sum* E_n*, defined in (6.16), converges* $\mathbf{P}_{\theta,x}$*-a.s and in* L^2.

   (b) *For any* $m<n$*, the remainder term* $\widetilde{\eta}_{m,n}1_{\{\theta_m\in Q\}}1_{\{\theta_{n-1}\in Q\}}$ *converges to* 0*,* $\mathbf{P}_{\theta,x}$*-a.s. as* $m,n\to\infty$*.*

Proof: Parts 1-3 follow similarly to the proof of Proposition 2.7 using the localized variants of the BMP Lemmas.
Part 4-a. We have for $n=1,2,\ldots$,

$$\begin{aligned} E_n &\doteq \sum_{k=0}^{n-1}\gamma_{k+1}\widetilde{\varepsilon}_{k+1}^{(1)}1_{\{\theta_k\in Q\}}+\sum_{k=1}^{n-1}\gamma_{k+1}\widetilde{\varepsilon}_{k+1}^{(2)}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}} \\ &\quad +\sum_{k=1}^{n-1}\gamma_{k+1}\widetilde{\varepsilon}_{k+1}^{(3)}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}+\sum_{k=0}^{n-1}\gamma_{k+1}\widetilde{\varepsilon}_{k+1}^{(4)}1_{\{\theta_k\in Q\}}. \end{aligned} \quad (6.21)$$

The first summation term in (6.21) converges almost surely and in L^2 by Lemma 6.2. It's also clear the remaining three summation terms all converge almost surely and in L^2 from the bounds in Lemmas 6.3, 6.4 and 6.5. (See the proof of Proposition 2.7 for details on how this is shown.)
Part 4-b. In the original derivation, $\widetilde{\eta}_{m;n}$ was only defined for $\upsilon_{i-1}\le m<n\le\tau_i$ for any $i=1,2,\ldots$ but we now extend the definition to arbitrary $m=0,1,2,\ldots$ and $n>m$ so that

$$\widetilde{\eta}_{m;n}\doteq\gamma_{m+1}\psi_{\theta_m}(X_m)-\gamma_n\psi_{\theta_{n-1}}(X_n).$$

From the Monotone Convergence Theorem we get

$$\begin{aligned} &\mathbf{E}_{\theta,x}\left[\sum_{m=1}^{\infty}\left|1_{\{\theta_m\in Q\}}\gamma_{m+1}\psi_{\theta_m}(X_m)\right|^2\right] \\ &\le \sum_{m=1}^{\infty}\gamma_{m+1}^2\mathbf{E}_{\theta,x}\left[1_{\{\theta_m\in Q\}}\left|\psi_{\theta_m}(X_m)\right|^2\right] \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{m=1}^{\infty} \gamma_{m+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_m\in Q\}} \left\|\nabla\phi(\theta_m)\right\|^2 \left\|P_{\theta_m}\nu_{\theta_m}(X_m)\right\|^2\right] \\
&\leq \sum_{m=1}^{\infty} \gamma_{m+1}^2 M_1^2 C_\nu^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_m\in Q\}} V^{2r}(X_m)\right] \\
&\leq \sum_{m=1}^{\infty} \gamma_{m+1}^2 M_1^2 C_\nu^2 C_D V(x) < \infty.
\end{aligned}$$

Therefore, the series $\sum_{m=1}^{\infty} \left|1_{\{\theta_m\in Q\}}\gamma_{m+1}\psi_{\theta_m}(X_m)\right|^2$ converges almost surely which implies

$$1_{\{\theta_m\in Q\}}\gamma_{m+1}\psi_{\theta_m}(X_m) \longrightarrow 0 \tag{6.22}$$

almost surely under $\mathbf{P}_{\theta,x}$.

Similarly

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sum_{n=1}^{\infty} \left|1_{\{\theta_{n-1}\in Q\}}\gamma_n\psi_{\theta_{n-1}}(X_n)\right|^2\right] \\
&\leq \sum_{n=1}^{\infty} \gamma_n^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_{n-1}\in Q\}} \left|\psi_{\theta_{n-1}}(X_n)\right|^2\right] \\
&\leq \sum_{n=1}^{\infty} \gamma_n^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_{n-1}\in Q\}} \left\|\nabla\phi(\theta_{n-1})\right\|^2 \left\|P_{\theta_{n-1}}\nu_{\theta_{n-1}}(X_n)\right\|^2\right] \\
&\leq \sum_{m=1}^{\infty} \gamma_n^2 M_1^2 C_\nu^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_{n-1}\in Q\}} V^{2r}(X_n)\right] \\
&\leq \sum_{m=1}^{\infty} \gamma_n^2 M_1^2 C_\nu^2 C_D V(x) < \infty,
\end{aligned}$$

The series $\sum_{n=1}^{\infty} \left|1_{\{\theta_{n-1}\in Q\}}\gamma_n\psi_{\theta_{n-1}}(X_n)\right|^2$ also converges almost surely which implies

$$\lim_{n\to\infty} 1_{\{\theta_{n-1}\in Q\}}\gamma_n\psi_{\theta_{n-1}}(X_n) = 0, \qquad \mathbf{P}_{\theta,x} - a.s.$$

Finally, the convergence to zero follows readily since

$$\begin{aligned}
&\left|\widetilde{\eta}_{m;n} 1_{\{\theta_m\in Q\}} 1_{\{\theta_{n-1}\in Q\}}\right| \\
&= \left|\gamma_{m+1} 1_{\{\theta_m\in Q\}} 1_{\{\theta_{n-1}\in Q\}} \psi_{\theta_m}(X_m) - \gamma_n 1_{\{\theta_m\in Q\}} 1_{\{\theta_{n-1}\in Q\}} \psi_{\theta_{n-1}}(X_n)\right| \\
&\leq \left|\gamma_{m+1} 1_{\{\theta_m\in Q\}} \psi_{\theta_m}(X_m)\right| + \left|\gamma_n 1_{\{\theta_{n-1}\in Q\}} \psi_{\theta_{n-1}}(X_n)\right|.
\end{aligned}$$

■

6.6 Application: Tuning the M/M/1 with Unknown Arrival Rate

As we stated in the introduction, we carry out the convergence analysis specifically for the M/M/1 queue with infinite buffer, unknown arrival rate $\lambda_{arrival} > 0$, and controlled service rate $\mu(\theta) = \theta$ which is constrained (projected) on the set $\Theta = [0, M]$ with $M < \infty$ arbitrarily large.

The performance measure of interest is the steady state mean queue size which, for this model, can be computed analytically if $\lambda_{arrival}$ and $\mu(\theta)$ are both known and $\frac{\lambda_{arrival}}{\mu(\theta)} < 1$:

$$\mathbf{E}_{\pi_\theta}[X] = \frac{\lambda_{arrival}}{\mu(\theta) - \lambda_{arrival}}. \tag{6.23}$$

Our goal is to find the particular service rate $\mu(\theta^\star)$ which achieves a steady state queue size of $N^\star$ customers in queue and in service.

We choose to model the M/M/1 queue size by *uniformizing* the continuous time Markov chain for the M/M/1 and simulating a corresponding discrete time chain [56] which is the following birth-death chain (a random-walk with a reflection at the origin). For each $x = 0, 1, \ldots$, the one-step transition probabilities are:

$$\begin{aligned} P_\theta[X_{n+1} = x + 1 | X_n = x] &= \frac{\lambda_{arrival}}{\lambda_{arrival} + \theta}, \\ P_\theta[X_{n+1} = (x-1)^+ | X_n = x] &= \frac{\theta}{\lambda_{arrival} + \theta}, \qquad \theta \in [0, M], \end{aligned} \tag{6.24}$$

with zero probability for all remaining transitions. This chain possesses a null transition from state 0 to state 0.

We apply a projected SA

$$\begin{aligned} \theta_{n+1} &= \Pi_\Theta \{\theta_n + \gamma_{n+1} H(\theta_n, X_{n+1})\}, \qquad n = 0, 1, \ldots \\ \theta_0 &= \theta \in \Theta \\ X_0 &= x \quad \in \mathsf{X} \end{aligned}$$

with projection set $\Theta = [0, M]$ and driving function $H(\theta, x) = x - N^\star$. Additionally, we shall assume an explicit step size sequence $\{\gamma_n,\ n = 1, 2, \ldots\}$ given by $\gamma_n = \frac{1}{n}$ for all $n = 1, 2, \ldots$. It's clear that at certain times the chain may in fact be operating at a parameter in the transient region so that $\frac{\lambda_{arrival}}{\mu(\theta)} > 1$.

Here, M serves to model the maximum mean service rate available in the queue and we choose $M < \infty$ simply to achieve the *boundedness condition* required for the Kushner-Clark Lemma. The use of a bounded service rate limit M is not seen as a significant limitation because M may be chosen arbitrarily large. Alternatively, this fixed upper bound on the projection set could likely be relaxed using a randomly increasing bound as in Chen's algorithm [17, 18], but we shall not explore this option as it is the effects of the transient region at the other extreme of Θ we wish to study here.

A similar birth-death chain was studied earlier in Chapter 3 in which the transition probabilities had a linear dependence on θ instead of the nonlinear dependence given by (6.24). In that earlier case, an ideal compact projection set was used with the SA which maintained positive recurrence for that Markov chain and the *specialized conditions* of Chapter 3 were explicitly verified which implied the more general convergence conditions of Section 2.3. In this case here, we do not assume accurate knowledge of the arrival rate $\lambda_{arrival}$ hence such an ideal projection set cannot be identified. Nevertheless, a compact subset $Q \subset (\lambda_{arrival}, M]$ does exist for which the *specialized conditions* hold if the iterates were constrained to Q and this is easily shown as in Chapter 3 with only a slight modification to account for the nonlinear dependence of the probabilities (6.24).

Unlike most applications of SA, the regression function for this example is analytically known, i.e.

$$h(\theta) = \frac{\lambda_{arrival}}{\theta - \lambda_{arrival}} - N^\star, \qquad \theta > \lambda_{arrival}, \tag{6.25}$$

and we define $\theta^\star$ as the point $h(\theta^\star) = 0$. We shall show that the algorithm above yields convergence of θ_n to $\theta^\star$ almost surely. Finally, we understand that it's very unlikely for SA to be proposed for the M/M/1 problem which we have stated here since since h is known and there are better methods to find $\theta^\star$. Nevertheless, this example serves as a test case to explore the effect of the transient region within the projection set on a tractable problem possessing an unbounded state space. It is our belief that many of the techniques presented below should extend to other problems where the regression function is not known.

6.7 Verification of (D0') for the Birth-Death Chain

We start with the verification of (D1') for this problem and recall we have shown in Section 3.9.3 that the test function $V(x) = Ks^x$ satisfied a (D1') condition. Actually, for the remainder of this chapter, we wish to change the form of V by defining a $\delta \doteq \delta(s) > 0$ such that

$$V(x) = Ks^x = Ke^{\delta x}.$$

Let us define $Q \subset \Theta$ to be the interval $Q = [q, M]$ for some satisfactory $q > \lambda_{arrival}$ such that $\lambda_{arrival} < q < \theta^*$. Recall from Section 3.9.3 that in proving the local (D1') condition we found what amounts to conditions on λ in (D1') and the boundaries of Q for each $\delta > 0$ (or $s > 1$), i.e.

$$P_\theta V \le \lambda V + L, \qquad \theta \in Q$$

Unfortunately, condition (D1') alone is not adequate for our main convergence theorems of this chapter and we next consider techniques involving excursions of the parameter from Q into the region $Q^c \cap \Theta$ to show the stronger (D0').

6.7.1 Last Exit and Return Decomposition

Let us fix an integer $n = 1, 2, \ldots$ and we assume $\theta_0 \in Q$. For any sample path $\{\theta_i, i = 0, 1, \ldots, n\} \in \mathcal{F}_n$ let us define a partition on the the event $\{\theta_n \in Q\}$. The first set in the partition is defined as those sample paths $\{\theta_i, i = 1, \ldots, n\}$ which remain in Q for **all** of the first n steps, denoted $A_n \doteq \{\theta_i \in Q, \ i = 0, 1, \ldots, n\}$. Next, for any sample path in the complementary event A_n^c we define the *last exit time* rv

$$\tau = \tau(Q) \doteq \sup\{1 \le k \le n-1 : \ \theta_k \notin Q, \theta_{k-1} \in Q\}.$$

Then we define the first return time after τ as

$$\upsilon = \upsilon(Q) \doteq \inf\{k > \tau : \ \theta_k \in Q\}$$

With these definitions, we define the remaining sets in the partition:

$$B^n_{\tau,\upsilon} \doteq \{\theta_{\tau-1} \in Q; \ \theta_j \notin Q, j = \tau, \cdots, \upsilon - 1; \ \theta_k \in Q, \ k = \upsilon, \ldots, n\}, \quad 1 \leq \tau < \upsilon \leq n.$$

It's clear that these events are disjoint, that $A_n \subset \{\theta_n \in Q\}$, $B^n_{\tau,\upsilon} \subset \{\theta_n \in Q\}$ for $1 \leq \tau < \upsilon \leq n$, and that

$$\{\theta_n \in Q\} = A_n \cup \bigcup_{1 \leq \tau < \upsilon \leq n} B^n_{\tau,\upsilon}, \qquad n = 1, 2, \ldots$$

First notice we have from (D1')

$$\begin{aligned} \mathbf{E}[V(X_{n+1})1_{\{\theta_n \in Q\}}|X_0, \theta_0] &= \mathbf{E}[\mathbf{E}[V(X_{n+1})|X_n, \theta_n]1_{\{\theta_n \in Q\}}|X_0, \theta_0] \\ &\leq \mathbf{E}[(\lambda V(X_n) + L)1_{\{\theta_n \in Q\}}|X_0, \theta_0] \\ &\leq \lambda \mathbf{E}[V(X_n)1_{\{\theta_n \in Q\}}|X_0, \theta_0] + L. \end{aligned} \tag{6.26}$$

Then, using the decomposition above

$$\mathbf{E}[V(X_n)1_{\{\theta_n \in Q\}}|X_0, \theta_0] = \mathbf{E}[V(X_n)1_{\{A_n\}}|X_0, \theta_0] + \sum_{1 \leq \tau < \upsilon \leq n} \mathbf{E}[V(X_n)1_{\{B^n_{\tau,\upsilon}\}}|X_0, \theta_0] \tag{6.27}$$

and we will now attempt to bound the quantities on the right hand side.

6.7.2 Sample Paths which remain in Q

The first term of (6.27) is evaluated as follows. For any $n = 1, 2, \ldots$

$$\begin{aligned} &\mathbf{E}\left[V(X_n)1_{\{A_n\}}|X_0 = x, \theta_0\right] \\ &= \mathbf{E}\left[V(X_n)1_{\{\theta_i \in Q, i=1,\ldots,n\}}|X_0 = x, \theta_0\right] \\ &\leq \mathbf{E}\left[V(X_n)1_{\{\theta_i \in Q, i=1,\ldots,n-1\}}|X_0 = x, \theta_0\right] \\ &= \mathbf{E}\left[\mathbf{E}[V(X_n)|X_{n-1}, \theta_{n-1}]1_{\{\theta_i \in Q, i=1,\ldots,n-1\}}|X_0 = x, \theta_0\right] \\ &\leq \mathbf{E}\left[(\lambda V(X_{n-1}) + L)1_{\{\theta_i \in Q, i=1,\ldots,n-1\}}|X_0 = x, \theta_0\right] \\ &= \lambda \mathbf{E}\left[V(X_{n-1})1_{\{\theta_i \in Q, i=1,\ldots,n-1\}}|X_0 = x, \theta_0\right] \qquad (6.28) \\ &\qquad + L\mathbf{P}[\{\theta_i \in Q, i = 1, \ldots, n-1\}|X_0 = x, \theta_0] \\ &\leq \lambda \mathbf{E}\left[V(X_{n-1})1_{\{\theta_i \in Q, i=1,\ldots,n-2\}}|X_0 = x, \theta_0\right] + L \\ &= \lambda \mathbf{E}\left[\mathbf{E}[V(X_{n-1})|X_{n-2}, \theta_{n-2}]1_{\{\theta_i \in Q, i=1,\ldots,n-2\}}|X_0 = x, \theta_0\right] + L \\ &\leq \lambda \mathbf{E}\left[(\lambda V(X_{n-2}) + L)1_{\{\theta_i \in Q, i=1,\ldots,n-2\}}|X_0 = x, \theta_0\right] + L \\ &= \lambda^2 \mathbf{E}\left[V(X_{n-2}))1_{\{\theta_i \in Q, i=1,\ldots,n-2\}}|X_0 = x, \theta_0\right] \qquad (6.29) \end{aligned}$$

$$
\begin{aligned}
&\quad + L + \lambda L\mathbf{P}[\{\theta_i \in Q, i = 1, \ldots, n-2\}|X_0, \theta_0] \\
&\leq \lambda^2 \mathbf{E}\left[V(X_{n-2}))1_{\{\theta_i \in Q, i=1,\ldots,n-3\}}|X_0 = x, \theta_0\right] + L(1+\lambda) \\
&\vdots \\
&\leq \lambda^n V(x) + L(1 + \lambda + \lambda^2 + \cdots + \lambda^{n-1}), \qquad x \in \mathsf{X}, \theta_0 \in Q. \qquad (6.30)
\end{aligned}
$$

6.7.3 Sample Paths which leave Q and return

For the second term of (6.27) we now consider sample paths that leave Q and return by time n, i.e. $1 \leq \tau < \upsilon \leq n$. First we find that

$$
\begin{aligned}
&\mathbf{E}[V(X_n)1_{\{B^n_{\tau,\upsilon}\}}|X_\upsilon, \theta_\upsilon] \\
&\quad = \mathbf{E}\left[V(X_n)1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n\}}|X_\upsilon, \theta_\upsilon\right] 1_{\{\theta_{\tau-1}\in Q; \theta_i \notin Q, i=\tau,\ldots,\upsilon-1; \theta_\upsilon \in Q\}} \\
&\quad \leq \mathbf{E}\left[V(X_n)1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-1\}}|X_\upsilon, \theta_\upsilon\right] 1_{\{\theta_{\tau-1}\in Q; \theta_i \notin Q, i=\tau,\ldots,\upsilon-1; \theta_\upsilon \in Q\}}.
\end{aligned}
$$

Then, in nearly the same manner as the last section we have for $n = \upsilon, \upsilon+1, \ldots$

$$
\begin{aligned}
&\mathbf{E}\left[V(X_n)1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-1\}}|X_\upsilon, \theta_\upsilon\right] \\
&\quad = \mathbf{E}\left[\mathbf{E}[V(X_n)|X_{n-1}, \theta_{n-1}]1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-1\}}|X_\upsilon, \theta_\upsilon\right] \\
&\quad \leq \mathbf{E}\left[(\lambda V(X_{n-1}) + L)1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-1\}}|X_\upsilon, \theta_\upsilon\right] \\
&\quad = \lambda\mathbf{E}\left[V(X_{n-1})1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-1\}}|X_\upsilon, \theta_\upsilon\right] \\
&\qquad\quad + L\mathbf{P}[\{\theta_i \in Q, i = \upsilon+1, \ldots, n-1\}|X_\upsilon, \theta_\upsilon] \\
&\quad \leq \lambda\mathbf{E}\left[V(X_{n-1})1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-2\}}|X_\upsilon, \theta_\upsilon\right] + L \\
&\quad = \lambda\mathbf{E}\left[\mathbf{E}[V(X_{n-1})|X_{n-2}, \theta_{n-2}]1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-2\}}|X_\upsilon, \theta_\upsilon\right] + L \\
&\quad \leq \lambda\mathbf{E}\left[(\lambda V(X_{n-2}) + L)1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-2\}}|X_\upsilon, \theta_\upsilon\right] + L \\
&\quad = \lambda^2\mathbf{E}\left[V(X_{n-2}))1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-2\}}|X_\upsilon, \theta_\upsilon\right] \\
&\qquad\quad + L + \lambda L\mathbf{P}[\{\theta_i \in Q, i = \upsilon+1, \ldots, n-2\}|X_\upsilon, \theta_\upsilon] \\
&\quad \leq \lambda^2\mathbf{E}\left[V(X_{n-2}))1_{\{\theta_i \in Q, i=\upsilon+1,\ldots,n-3\}}|X_\upsilon, \theta_\upsilon\right] + L(1+\lambda) \\
&\quad \vdots \\
&\quad \leq \lambda^{n-\upsilon}V(X_\upsilon) + L(1 + \lambda + \lambda^2 + \cdots + \lambda^{n-\upsilon-1}).
\end{aligned}
$$

Let us denote the set

$$
C_{\tau,\upsilon} \doteq \{\theta_{\tau-1} \in Q; \theta_i \notin Q, i = \tau, \ldots, \upsilon-1; \theta_\upsilon \in Q\}.
$$

Therefore,

$$
\begin{aligned}
&\mathbf{E}[V(X_n)1_{\{B^n_{\tau,\upsilon}\}}|X_0, \theta_0] \\
&\quad = \mathbf{E}[\mathbf{E}[V(X_n)1_{\{B^n_{\tau,\upsilon}\}}|X_\upsilon, \theta_\upsilon]|X_0, \theta_0]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \mathbf{E}\left[\left(\lambda^{n-v}V(X_v)+L(1+\lambda+\lambda^2+\cdots+\lambda^{n-v-1})\right)1_{\{C_{\tau,v}\}}|X_0,\theta_0\right] \\
&= \lambda^{n-v}\mathbf{E}[V(X_v)1_{\{C_{\tau,v}\}}|X_0,\theta_0] \qquad (6.31)\\
&\quad + L(1+\lambda+\lambda^2+\cdots+\lambda^{n-v-1})\mathbf{P}[\{C_{\tau,v}\}|X_0,\theta_0] \qquad (6.32)
\end{aligned}
$$

6.7.4 An Intermediate Expression

Using (6.30) and (6.32) we are now able to bound (6.27) as follows. For $n = 1, 2, \ldots$

$$
\begin{aligned}
&\mathbf{E}[V(X_n)1_{\{\theta_n\in Q\}}|X_0,\theta_0] \\
&\quad = \mathbf{E}[V(X_n)1_{\{A_n\}}|X_0,\theta_0] + \sum_{1\leq\tau<v\leq n}\mathbf{E}[V(X_n)1_{\{B^n_{\tau,v}\}}|X_0,\theta_0] \\
&\quad \leq V(x) + L(1+\lambda+\lambda^2+\cdots+\lambda^{n-1}) \\
&\quad + \sum_{1\leq\tau<v\leq n}\lambda^{n-v}\mathbf{E}[V(X_v)1_{\{C_{\tau,v}\}}|X_0,\theta_0] \\
&\quad + L(1+\lambda+\lambda^2+\cdots+\lambda^{n-v-1})\sum_{1\leq\tau<v\leq n}\mathbf{P}[\{C_{\tau,v}\}|X_0,\theta_0] \qquad (6.33)\\
&\quad \leq V(x) + 2L(1+\lambda+\lambda^2+\cdots+\lambda^{n-1}) \\
&\quad + \sum_{1\leq\tau<v\leq n}\lambda^{n-v}\mathbf{E}[V(X_v)1_{\{C_{\tau,v}\}}|X_0,\theta_0], \qquad X_0\in\mathsf{X},\theta_0\in Q. \qquad (6.34)
\end{aligned}
$$

The last line follows since the probabilities in the sum (6.33) are over mutually exclusive events, hence the sum is bounded by one.

Our approach to evaluating the final sum in (6.34) is to rewrite the above expression in the form below and develop a bound for the probability inside the integral:

$$
\begin{aligned}
&\mathbf{E}[V(X_n)1_{\{\theta_n\in Q\}}|X_0=x,\theta_0] \\
&\quad \leq V(x) + \frac{2L}{1-\lambda} \\
&\quad + \sum_{1\leq\tau<v\leq n}\lambda^{n-v}\int_0^\infty \mathbf{P}[V(X_v)1_{\{C_{\tau,v}\}}\geq t,|X_0=x,\theta_0]dt \qquad (6.35)
\end{aligned}
$$

We also note from (6.26) that

$$
\begin{aligned}
&\mathbf{E}[V(X_{n+1})1_{\{\theta_n\in Q\}}|X_0=x,\theta_0] \\
&\quad \leq \lambda\mathbf{E}[V(X_n)1_{\{\theta_n\in Q\}}|X_0=x,\theta_0]+L \\
&\quad = \lambda V(x) + 2L(\lambda+\lambda^2+\lambda^3+\cdots+\lambda^n) \\
&\quad + \lambda\sum_{1\leq\tau<v\leq n}\lambda^{n-v}\int_0^\infty \mathbf{P}_{\theta_0,x}[V(X_v)1_{\{C_{\tau,v}\}}\geq t,]dt + L \\
&\quad \leq V(x) + \frac{2L}{1-\lambda} \\
&\quad + \sum_{1\leq\tau<v\leq n}\lambda^{n-v}\int_0^\infty \mathbf{P}_{\theta_0,x}[V(X_v)1_{\{C_{\tau,v}\}}\geq t,]dt. \qquad (6.36)
\end{aligned}
$$

Hence, we have a common intermediate expression, i.e. (6.35) and (6.36), bounding both expectations in (D0').

6.7.5 A Key Sample Path Property

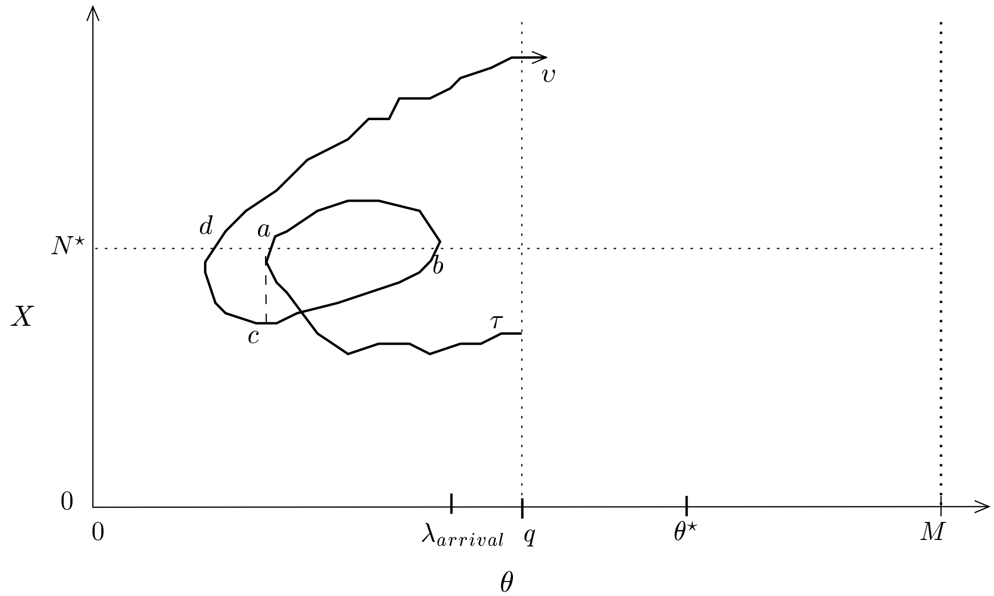

Figure 6.1: A sample path in $C_{\tau,v}$ with a "loop".

We now develop a key sample path property for events $C_{\tau,v}$ by considering the example shown in Figure 6.1 where we have one particular such sample path. This sample path has several points in time which are marked in Figure 6.1 and formally defined as follows:

$$\begin{aligned}
a &= \inf\{n \geq \tau+1 : X_n \geq N^\star\} \\
b &= \inf\{n \geq a+1 : X_n < N^\star\} \\
c &= \inf\{n \geq b+1 : \sum_{i=a}^{b-1} \gamma_i(X_i - N^\star) \leq -\sum_{i=b}^{n} \gamma_i(X_i - N^\star)\} \\
d &= \inf\{n \geq c+1 : X_n \geq N^\star\}
\end{aligned}$$

We note that for this example, the sample path segment does not come in contact with the projection boundary, hence we have $z_i = 0,\ i = \tau, \ldots, u-1$. (We shall address sample paths involving projections in the next example.)

Also, observe that *in general* we have $C_{\tau,v} \subset \left\{\sum_{i=\tau}^{v-1} \gamma_i(X_i - N^\star) < 0\right\}$ since

$$\begin{aligned}
\lambda_{arrival} > \theta_{v-1} &= \theta_{\tau-1} + \sum_{i=\tau}^{v-1} \gamma_i(X_i - N^\star) + \sum_{i=\tau}^{v-1} \gamma_i z_i \\
&\geq \lambda_{arrival} + \sum_{i=\tau}^{v-1} \gamma_i(X_i - N^\star).
\end{aligned}$$

The last line following because $z_i \geq 0$ on $i = \tau, \ldots, v-1$ for all sample paths in $C_{\tau,v}$.

The key to bounding the integrand in (6.36) is the following sample path property for paths in $C_{\tau,\upsilon}$ which allows us to show

$$\left\{\sum_{i=\tau}^{\upsilon-1}\gamma_i(X_i - N^\star) < 0\right\} \subset \left\{\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(X_i - N^\star) < q\right\}.$$

Claim 6.8 *For any sample path in $C_{\tau,\upsilon}$, we have*

$$\sum_{i=\tau}^{\upsilon-1}(X_i - N^\star) \leq (\upsilon-\tau)q. \tag{6.37}$$

Proof: We first consider the step-size weighted sum sample path segments $a \to b-1$ and $b \to c-1$. By the definitions of the point c, we have

$$-\sum_{i=b}^{c-1}\gamma_i(X_i - N^\star) \leq \sum_{i=a}^{b-1}\gamma_i(X_i - N^\star) \leq -\sum_{i=b}^{c}\gamma_i(X_i - N^\star)$$

so there exists a real number $\alpha \in [0,1]$ such that

$$\sum_{i=a}^{b-1}\gamma_i(X_i - N^\star) = -\sum_{i=b}^{c-1}\gamma_i(X_i - N^\star) - \alpha\gamma_c(X_c - N^\star) \tag{6.38}$$

Now, since $(X - N^\star) \geq 0$ for all $i = a, \ldots, b-1$, we have

$$\sum_{i=a}^{b-1}\gamma_i(X_i - N^\star) \geq \gamma_{b-1}\sum_{i=a}^{b-1}(X_i - N^\star) \geq \gamma_b\sum_{i=a}^{b-1}(X_i - N^\star). \tag{6.39}$$

Also, since $(X - N^\star) < 0$ for all $i = b, \ldots, c$, we have

$$-\sum_{i=b}^{c-1}\gamma_i(X_i - N^\star) - \alpha\gamma_c(X_c - N^\star) \leq -\gamma_b\sum_{i=b}^{c-1}(X_i - N^\star) - \alpha\gamma_b(X_c - N^\star). \tag{6.40}$$

Therefore, by applying the two inequalities (6.39) and (6.40) to (6.38) we have

$$\begin{aligned}\gamma_b\sum_{i=a}^{b-1}(X_i - N^\star) \leq \sum_{i=a}^{b-1}\gamma_i(X_i - N^\star) &= -\sum_{i=b}^{c-1}\gamma_i(X_i - N^\star) - \alpha\gamma_c(X_c - N^\star) \\ &\leq -\gamma_b\sum_{i=b}^{c-1}(X_i - N^\star) - \alpha\gamma_b(X_c - N^\star)\end{aligned}$$

or

$$\sum_{i=a}^{b-1}(X_i - N^\star) \leq -\sum_{i=b}^{c-1}(X_i - N^\star) - \alpha(X_c - N^\star). \tag{6.41}$$

If we compare this with (6.38), this says the result of removing the step-sizes from the weighted sum of these particular segments does not result in an increase in the sum and we have

$$\sum_{i=a}^{b-1}(X_i - N^\star) + \sum_{i=b}^{c-1}(X_i - N^\star) + \alpha(X_c - N^\star) \leq 0. \tag{6.42}$$

Next, we consider the remaining sample path segments $\tau \to a-1$, $c \to d-1$, and $d \to u-1$. We have

$$\begin{aligned}
&\sum_{i=\tau}^{a-1}(X_i - N^\star) + (1-\alpha)(X_c - N^\star) + \sum_{i=c+1}^{d-1}(X_i - N^\star) + \sum_{i=d}^{\upsilon-1}(X_i - N^\star) \\
&\leq \frac{1}{\gamma_\tau}\sum_{i=\tau}^{a-1}\gamma_i(X_i - N^\star) + (1-\alpha)\frac{\gamma_c}{\gamma_\tau}(X_c - N^\star) \\
&\quad + \frac{1}{\gamma_\tau}\sum_{i=c+1}^{d-1}\gamma_i(X_i - N^\star) + \frac{1}{\gamma_{\upsilon-1}}\sum_{i=d}^{\upsilon-1}\gamma_i(X_i - N^\star) \\
&\leq -\frac{1}{\gamma_\tau}\sum_{i=d}^{\upsilon-1}\gamma_i(X_i - N^\star) + \frac{1}{\gamma_\upsilon}\sum_{i=d}^{\upsilon-1}\gamma_i(X_i - N^\star) \\
&\leq (\frac{1}{\gamma_\upsilon} - \frac{1}{\gamma_\tau})q \\
&= (\upsilon - \tau)q.
\end{aligned}$$

The last line follows because we are assuming the explicit step size sequence $\gamma_i = \frac{1}{i}$ for $i = 1, 2, \ldots$.

Thus we have shown (6.37) for this particular sample path. Furthermore, it's not difficult to see that more complicated sample paths involving more "loops" can be considered via the same approach of canceling the upper portions of any loops (defined as the segment with $X_i > N^\star$ at all points in the segment) with some segment which *follows* (in time) the upper loop segment. After all such possible upper and lower loop segments have been canceled in this manner (and possibly using various linear combination factors, $\alpha_i \in (0, 1]$ in $i = 1, 2, \ldots$ to apportion certain single lower summands among the upper loops as we used in the first example), what remains are segments whose sum with the step sizes removed is bounded by $(\frac{1}{\gamma_\upsilon} - \frac{1}{\gamma_\tau})$ times the maximum distance to Q which is q for this problem.

Additionally, sample paths in $C_{\tau,\upsilon}$ which contact the leftmost projection boundary also satisfy (6.37). This can be seen by the simple loop-less sample-path of Figure 6.2. Let us redefine the point a as the first point in time where a positive projection occurs, i.e. $a = \inf\{n \geq \tau : z_n > 0\}$. Also, we redefine $b = \inf\{n > a : X_n > N^\star\}$.

Let us see what happens to the sum $\sum_{i=\tau}^{\upsilon-1}(X_i - N^\star)$. Observe that the projection terms z_i must be non-negative for all sample paths in $C_{\tau,\upsilon}$. Then, the segments $\tau \to a$ and $b \to \upsilon - 1$ are compared in nearly the same manner as the previous example so that

$$\sum_{i=\tau}^{a-1}(X_i - N^\star) + \sum_{i=b}^{\upsilon-1}(X_i - N^\star) \leq \left(\frac{1}{\gamma_\upsilon} - \frac{1}{\gamma_\tau}\right)\sum_{i=b}^{\upsilon-1}\gamma_i(X_i - N^\star) \leq (\upsilon - \tau)q$$

while the segment $\sum_{i=a}^{b-1}(X_i - N^\star) < 0$. Thus,

$$\sum_{i=\tau}^{\upsilon-1}(X_i - N^\star) \leq (\upsilon - \tau)\, q.$$

■

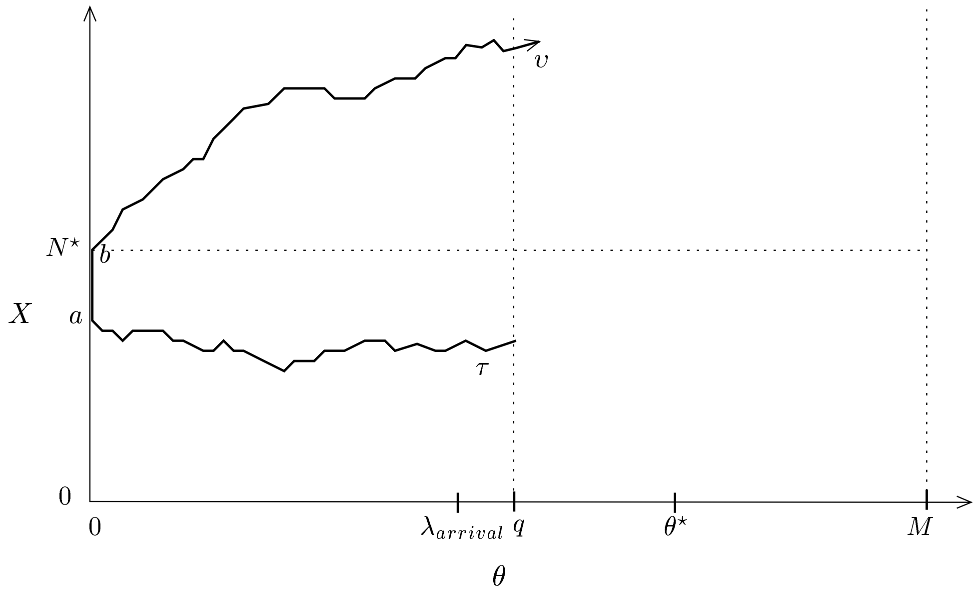

Figure 6.2: A sample path in $C_{\tau,\upsilon}$ with a positive projection term.

6.7.6 Stochastic Comparison Argument

We shall make two stochastic comparisons. First, we consider a generic (and time homogeneous) Markov chain $\{\widetilde{X}_i, i = \tau, \ldots, \upsilon\}$ which shares a common probability space with $\{X_i, i = \tau, \ldots, \upsilon\}$, which is driven by the same family of one step transition matrices $\{P_\theta, \theta \in \Theta\}$ but having a *fixed* θ set to the boundary of Q, i.e. $\theta = q$, and which is initiated with the same state as the nominal chain $\{X_i, i = \tau, \ldots, \upsilon\}$ at time τ, i.e. $\widetilde{X}_\tau = X_\tau$. Under this construction, it's clear that for all $y \in \mathbb{R}$,

$$\begin{aligned}
&\mathbf{P}\left[\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(X_i - N^\star) < y | X_\tau = x, \theta_\tau = \theta\right] \\
&\quad\leq \mathbf{P}\left[\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(\widetilde{X}_i - N^\star) < y | \widetilde{X}_\tau = X_\tau = x\right].
\end{aligned}$$

It is immediate the nominal chains sample path segment $\{X_i, i = \tau, \ldots, \upsilon - 1\}$ has parameters $\theta_i < q$ for all $i = \tau, \ldots, \upsilon - 1$.

Next, we employ a second stochastic comparison argument by defining a third Markov chain $\{\widetilde{\widetilde{X}}_i, i = \tau, \ldots, \upsilon\}$ which also shares a common probability space with $\{X_i, i = \tau, \ldots, \upsilon\}$ and $\{\widetilde{X}_i, i = \tau, \ldots, \upsilon\}$. This third Markov chain $\{\widetilde{\widetilde{X}}_i, i = \tau, \ldots, \upsilon\}$ is simply a version of $\{\widetilde{X}_i, i = \tau, \ldots, \upsilon\}$ with fixed parameter $\theta = q$ which evolves on a *bounded* state space $\{0, 1, 2, \ldots, B\}$. The transition probabilities are that of $\{\widetilde{X}_i, i = \tau, \ldots, \upsilon\}$ except we have inserted a reflection at the boundary B, i.e. the nonzero transition probabilities from state B are changed to:

$$\mathbf{P}[\widetilde{\widetilde{X}}_{i+1} = B | \widetilde{\widetilde{X}}_i = B] = \frac{\lambda_{arrival}}{\lambda_{arrival} + q}$$

$$\mathbf{P}[\widetilde{\widetilde{X}}_{i+1} = B-1 | \widetilde{\widetilde{X}}_i = B] \quad = \quad \frac{q}{\lambda_{arrival}+q}.$$

Next we observe that for any sample path in $C^n_{\tau,\upsilon}$ we must necessarily have $X_\tau \in \{0,1,\ldots,\lfloor N^\star \rfloor\}$ so if B is chosen to satisfy $B > \lfloor N^\star \rfloor$, we can make comparisons of these three chains given that they have the same starting value, i.e $\widetilde{\widetilde{X}}_\tau = \widetilde{X}_\tau = X_\tau$. With this construction, it's clear we have for all $y \in \mathbb{R}$

$$\begin{aligned}
&\mathbf{P}\left[\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(X_i - N^\star) < y | X_\tau = x, \theta_\tau = \theta\right] \\
&\leq \ \mathbf{P}\left[\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(\widetilde{X}_i - N^\star) < y | \widetilde{X}_\tau = x\right] \\
&\leq \ \mathbf{P}\left[\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(\widetilde{\widetilde{X}}_i - N^\star) < y | \widetilde{\widetilde{X}}_\tau = x\right], \\
&\qquad \text{for all } x \in \{0,1,\ldots,\lfloor N^\star \rfloor\}.
\end{aligned}$$

To see where we are going with this, we remark that the sample mean of $\{\widetilde{\widetilde{X}}_i, i = \tau,\ldots,\upsilon\}$ is readily known to possess a large deviations upper bound.

6.7.7 Large Deviations Bound

By the mean ergodic theorem for Markov chains, there exists some real $F = F(q,B)$ such that

$$\lim_{\upsilon\to\infty} \frac{1}{\upsilon-\tau}\sum_{I=\tau}^{\upsilon-1}\widetilde{\widetilde{X}}_i = F(q,B) \qquad \mathbf{P}-a.s. \tag{6.43}$$

The function $F(q,B)$ has known properties, i.e. for all fixed $q \in (\lambda_{arrival}, M]$, $F(q,B)$ is an increasing function of B; while for all fixed $1 \leq B < \infty$, $F(q,B)$ is monotone decreasing in q. Furthermore, $\lim_{B\to\infty}\lim_{q\searrow\lambda_{arrival}} F(q,B) = \infty$. We are free to choose a large $B < \infty$ and a small $q > \lambda_{arrival}$ arbitrarily close to $\lambda_{arrival}$. Thus, for any given $\epsilon > 0$, we are able to choose q, and B such that

$$F(q,B) - N^\star > q + \epsilon.$$

Therefore, the set

$$\left\{\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(\widetilde{\widetilde{X}}_i - N^\star) < q\right\} \subset \left\{\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(\widetilde{\widetilde{X}}_i - F(q,B)) < -\epsilon\right\}, \tag{6.44}$$

the latter being a standard form for large deviations bounds on the sample mean of a finite state Markov chain. Specifically, this Markov chain $\{\widetilde{\widetilde{X}}_i, i = \tau,\ldots,\upsilon-1\}$ is irreducible, aperiodic, positive recurrent and from the results of Theorem 7.10 in the next chapter there exists a $K(\epsilon) < \infty$ and a $c(\epsilon) > 0$ such that

$$\mathbf{P}\left[\frac{1}{\upsilon-\tau}\sum_{i=\tau}^{\upsilon-1}(\widetilde{\widetilde{X}}_i - F(q,B)) < -\epsilon \mid \widetilde{\widetilde{X}}_\tau = X_\tau = x\right] \leq K(\epsilon)e^{-c(\epsilon)(\upsilon-\tau)},$$

for all $\upsilon > \tau$ and $x \in \{0,1,\ldots,B\}$.

6.7.8 Bringing It All Together

For $t > 0$,

$$\begin{aligned} \mathbf{P}\left[V(X_\upsilon) \geq t, C_{\tau,\upsilon} | X_0 = x, \theta_0\right] &= \mathbf{P}\left[X_\upsilon \geq \frac{1}{\delta}\ln t, C_{\tau,\upsilon} | X_0 = x, \theta_0\right] \\ &\leq \mathbf{P}\left[C_{\tau,\upsilon} | X_0 = x, \theta_0\right] 1_{\{\tau + \lceil \frac{1}{\delta}\ln t \rceil - \lfloor N^\star \rfloor \leq \upsilon\}} \end{aligned} \tag{6.45}$$

since X_τ necessarily must satisfy $X_\tau \leq \lfloor N^\star \rfloor$ and the fact that the value of the Markov chain sample path $\{X_i, i = \tau, \ldots, \upsilon\}$ can only increase by $\upsilon - \tau$ customers in $\upsilon - \tau$ steps. Thus, $X_\upsilon \leq \lfloor N^\star \rfloor + \upsilon - \tau$ and if $\lceil \frac{1}{\delta}\ln t \rceil > \lfloor N^\star \rfloor + \upsilon - \tau$, the event $\{X_\upsilon \geq \lceil \frac{1}{\delta}\ln t \rceil\}$ necessarily must have probability zero. Now applying (6.44) to (6.45) followed by the large deviations bound, we have

$$\begin{aligned} &\mathbf{P}\left[V(X_\upsilon) \geq t, C_{\tau,\upsilon} | X_0 = x, \theta_0\right] \\ &\leq \mathbf{P}\left[\frac{1}{\upsilon - \tau}\sum_{i=\tau}^{\upsilon-1}(X_i - N^\star) < q | X_0 = x, \theta_0\right] 1_{\{\tau + \lceil \frac{1}{\delta}\ln t \rceil - \lfloor N^\star \rfloor \leq \upsilon\}} \\ &\leq \mathbf{P}\left[\frac{1}{\upsilon - \tau}\sum_{i=\tau}^{\upsilon-1}(\widetilde{\widetilde{X}}_i - F(q,B)) < -\epsilon | X_0 = x, \theta_0\right] 1_{\{t \leq e^{\delta(\upsilon - \tau + \lfloor N^\star \rfloor)}\}} \\ &\leq K(\epsilon) e^{-c(\epsilon)(\upsilon - \tau)} 1_{\{t \leq e^{\delta(\upsilon - \tau + \lfloor N^\star \rfloor)}\}} \end{aligned}$$

Finally, using this last expression we can bound the final term of (6.36),

$$\begin{aligned} &\sum_{1 \leq \tau < \upsilon \leq n} \lambda^{n-\upsilon} \int_0^\infty \mathbf{P}\left[V(X_\upsilon) \geq t, C_{\tau,\upsilon} | X_0, \theta_0\right] dt \\ &\leq K(\epsilon) \sum_{1 \leq \tau < \upsilon \leq n} \lambda^{n-\upsilon} \int_0^\infty e^{-c(\epsilon)(\upsilon - \tau)} 1_{\{t \leq e^{\delta(\upsilon - \tau + \lfloor N^\star \rfloor)}\}} dt \\ &\leq K(\epsilon) \sum_{1 \leq \tau < \upsilon \leq n} \lambda^{n-\upsilon} e^{-c(\epsilon)(\upsilon - \tau)} e^{\delta(\upsilon - \tau + \lfloor N^\star \rfloor)} \\ &\leq K(\epsilon) e^{\delta \lfloor N^\star \rfloor} \sum_{\upsilon=2}^{n} \lambda^{n-\upsilon} \sum_{\tau=1}^{\upsilon-1} e^{(\delta - c(\epsilon))(\upsilon - \tau)}. \end{aligned}$$

We immediately see that for this expression to be bounded for large n, we require δ to be such that $0 < \delta < c(\epsilon)$. Although, the value of $c(\epsilon)$ is not known, we are able to select the parameter δ in our test function $V(x) = e^{\delta x}$ to be an arbitrarily small value such that $\delta > 0$. Hence, we have the situation that such a δ exists but without specific knowledge of $c(\epsilon)$, V must be assumed to have an arbitrarily small parameter δ. This will clearly have effects on the remaining conditions in our framework. Nevertheless, we will find when we seek to verify condition (H2') that this does not pose any problem whatsoever for our chosen driving function $H(x,\theta) = x - N^\star$ since for each $\delta > 0$ there always exists some constant $C_H < \infty$ such that $\sup_{\theta \in Q} |H(x,\theta)| \leq C_H e^{r\delta x}$ for all $x \in \mathsf{X}$ and some $0 < r \leq 1/4$.

So with an arbitrarily small δ such that $0 < \delta < c(\epsilon)$, we have

$$\begin{aligned}
\sum_{1\leq\tau<v\leq v} \lambda^{n-v} \int_0^\infty \mathbf{P}_{\theta_0,x}\left[V(X_v) \geq t, C_{\tau,v}\right] dt &\leq K(\epsilon)e^{\delta\lfloor N^\star \rfloor} \sum_{v=2}^{n} \lambda^{n-v} \sum_{\tau=2}^{v-1} e^{(\delta-c(\epsilon))(v-\tau)} \\
&\leq K(\epsilon)e^{\delta\lfloor N^\star \rfloor} \sum_{v=1}^{n} \lambda^{n-v} \frac{e^{(\delta-c(\epsilon))}}{1-e^{(\delta-c(\epsilon))}} \\
&\leq K(\epsilon)e^{\delta\lfloor N^\star \rfloor} \left(\frac{1}{1-\lambda}\right)\left(\frac{e^{(\delta-c(\epsilon))}}{1-e^{(\delta-c(\epsilon))}}\right) \\
&\doteq C_B < \infty
\end{aligned}$$

We conclude that the common bound in (6.36) is finite and

$$\begin{aligned}
V(x) + \frac{2L}{1-\lambda} &+ \sum_{1\leq\tau<v\leq n} \lambda^{n-v} \int_0^\infty \mathbf{P}_{\theta_0,x}[V(X_v) \geq t, C_{\tau,v}]dt \\
&\leq \left(1 + \frac{2L}{1-\lambda} + C_B\right) V(x) \\
&\doteq C_D V(x), \qquad\qquad n = 1, 2, \ldots, \quad \theta_0 \in Q,
\end{aligned}$$

where $C_D < \infty$.

6.8 Asymptotic Analysis of an SA Applied to the M/M/1 with Unknown Arrival Rate

Here, we complete the convergence analysis for the SA algorithm applied to the simple Birth-Death chain which models the M/M/1 with unknown arrival rate $\lambda_{arrival}$ (see Section 6.6 for details).

Recall, we have assumed that $\theta^\star << M$. Also, for the unknown $\lambda_{arrival} > 0$, there exist two positive constants q_1, q_2 such that $0 < \lambda_{arrival} < q_2 < q_1 < \theta^\star$ which we use to define two nested compact sets $Q_1 = [q_1, M]$ and $Q_2 = [q_2, M]$. While h in (6.25) is only defined on $(\lambda_{arrival}, \infty)$, the SA algorithm uses a projection set $\Theta = [0, M]$. It's clear that Q_2 is a compact subset of the domain of attraction $DA(\theta^\star) = (\lambda_{arrival}, \infty)$ of the ODE and that $\theta^\star \in \Theta^o$, the interior of Θ.

Since the regression function $h(\theta)$ is only defined on the interval $(\lambda_{arrival}, \infty)$, it's clear the continuity condition of the Kushner-Clark Lemma is not satisfied for this problem in the region $[0, \lambda_{arrival}]$. Despite this difficulty, below we demonstrate a technique leading to almost sure convergence of the iterates to $\theta^\star$ which among other things, depends on showing that if θ_n ever leaves Q_2 it almost surely returns to the smaller set Q_1 in finite time. While this return property appears to be the key condition, we complete the convergence analysis by adapting several arguments from [6] to this problem and applying the Kushner-Clark Lemma upon return to Q_1. The conclusion is an almost sure convergence under completely verifiable conditions.

The entire development below is carried out under the specific circumstances of this particular M/M/1 SA problem, therefore all assumptions of this problem are assumed for all results of this section and not explicitly stated in each result.

6.8.1 Parametric Recurrence

We make the convention that $\tau_0(\cdot) = 0$ and we denote the i^{th} *return time* to Q_1 as

$$\upsilon_i(Q_1) \doteq \inf\{n \geq \tau_i(Q_2) : \theta_n \in Q_1\}, \qquad i = 0, 1, 2, \ldots. \tag{6.46}$$

Next, define the i^{th} *exit time* from Q_2

$$\tau_i(Q_2) \doteq \inf\{n \geq \upsilon_{i-1}(Q_1) : \theta_n \notin Q_2\}, \qquad i = 1, 2, \ldots \tag{6.47}$$

If the algorithm is initialized so $\theta_0 = \theta \in Q_1$, then $\upsilon_0(Q_1) = 0$.

Lemma 6.9 *For any $i = 1, 2, \ldots$, on the set $\{\omega : \tau_i(Q_2) < \infty\}$, we have $\upsilon_i(Q_1) < \infty$ $\mathbf{P}_{\theta,x}$-a.s.*

Proof:

Let $\tau = \tau_i(Q_2)$ and $\upsilon = \upsilon_i(Q_1)$ for any $i = 1, 2, \ldots$. Assume that $\{\tau < \infty\} \cap \{\upsilon = \infty\}$ and we now show a contradiction. The condition $\tau < \infty$ implies that on this set $\theta_\tau \in Q_2^c \cap \Theta = [0, q_2)$.

The process $\{(X_n, \theta_n), n = 0, 1, \ldots\}$ is a (non-homogeneous) Markov chain as defined in Section 2.2. We will make a comparison of $\{(X_n, \theta_n), n = 0, 1, \ldots\}$ to a second birth-death Markov chain $\{(\bar{X}_n, \theta_n), n = 1, 2, \ldots\}$ which is defined on a *common* probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The process $\{\bar{X}_n, n = 1, 2, \ldots\}$ has

$$\bar{X}_n = X_n, \qquad n = 0, 1, \ldots, \tau.$$

The chain $\{\bar{X}_n : n = \tau + 1, \tau + 2 \ldots\}$ has the θ-parameter of the one step transition probability $\mathbf{P}_{\theta,x}$ held **fixed** at $\theta = q_1$.

Since $\theta_k \in [0, q_1)$ for $\tau \leq k < \upsilon = \infty$, we have the following *strong stochastic ordering* property which is not hard to show [71] by induction via Lindley's recursion: For each $k = \tau, \tau + 1, \ldots$,

$$\begin{aligned} &\mathbf{P}\left[\sum_{\ell=\tau}^{k} \gamma_{\ell+1} X_{\ell+1} \geq y \,\middle|\, \theta_0 = \theta, X_0 = x;\ \tau < \infty, \upsilon = \infty\right] \\ &\geq \mathbf{P}\left[\sum_{\ell=\tau}^{k} \gamma_{\ell+1} \bar{X}_{\ell+1} \geq y \,\middle|\, \theta_0 = \theta, X_0 = x;\ \tau < \infty, \upsilon = \infty\right], \\ &\qquad \theta \in Q_2,\ x \in \mathsf{X},\ y \in \mathbb{R} \end{aligned} \tag{6.48}$$

Recall we have defined the following points along the regression function $h : (\lambda_{arrival}, M] \to \mathbb{R}$

$$\begin{aligned} h(\theta^\star) &\doteq \mathbf{E}_{\pi_{\theta^\star}}[X] - N^\star = 0 \\ h(q_1) &\doteq \mathbf{E}_{\pi_{q_1}}[X] - N^\star = N(q_1) - N^\star > 0. \end{aligned}$$

The above inequality follows from the fact that $\frac{d}{d\theta}h(\theta) < 0$ for all $\theta \in (\lambda_{arrival}, \infty)$. By adding a constant to both sides of (6.48) and setting $y = 0$, we also have the following property: For each $k = \tau, \tau+1, \ldots$,

$$\begin{aligned}
&\mathbf{P}\left[\sum_{\ell=\tau}^{k} \gamma_{\ell+1}\left(X_{\ell+1} - \frac{N(q_1)+N^\star}{2}\right) \geq 0 \middle| \theta_0 = \theta, X_0 = x;\ \tau < \infty, \upsilon = \infty\right] \\
&\geq \ \mathbf{P}\left[\sum_{\ell=\tau}^{k} \gamma_{\ell+1}\left(\bar{X}_{\ell+1} - \frac{N(q_1)+N^\star}{2}\right) \geq 0 \middle| \theta_0 = \theta, X_0 = x;\ \tau < \infty, \upsilon = \infty\right],
\end{aligned}$$

for all $\theta \in Q_2$ and $x \in \mathsf{X}$.

Now consider the *trivial* SA algorithm defined by setting $\Theta_{fixed} = \{q_1\}$, a single point, and

$$\begin{aligned}
\theta_{n+1} &= \Pi_{\Theta_{fixed}}\left\{\theta_n + \gamma_n(\bar{X}_n - N(q_1))\right\}, \qquad n = \tau, \tau+1, \ldots \\
\theta_\tau &= q_1 \\
\bar{X}_\tau &= X_\tau \\
X_0 &= x.
\end{aligned}$$

Let us also consider the (τ-delayed) step-size weighted sum of the noise process for this trivial algorithm:

$$S_{\tau,k} \doteq \sum_{\ell=\tau}^{k} \gamma_{\ell+1}\varepsilon_{\ell+1} = \sum_{\ell=\tau}^{k} \gamma_{\ell+1}(\bar{X}_\ell - N(q_1)), \qquad k \geq \tau$$

From Proposition 6.7, we have $S_{\tau,k}$ converges almost surely to a finite rv, hence

$$\lim_{k\to\infty} |S_{\tau,k}| < \infty, \quad \mathbf{P}_{\theta_0,x} - a.s.$$

and since the series $\sum_{\ell=\tau}^{\infty} \gamma_{\ell+1} = \infty$, we have

$$\lim_{k\to\infty} \frac{S_{\tau,k}}{\sum_{\ell=\tau}^{k} \gamma_{\ell+1}} = 0, \qquad \text{almost surely under } \mathbf{P}_{\theta_0,x}. \tag{6.49}$$

Hence for this fixed-θ Markov chain, there exists a null set $N \subset \Omega$ such that for all $\omega \in \Omega \setminus N$, and for all $\epsilon > 0$ there exists an $k'(\omega, \epsilon) < \infty$ such that

$$\left|\frac{S_{\tau,k}}{\sum_{\ell=\tau}^{k} \gamma_{\ell+1}}\right| < \epsilon \qquad \text{for all } k \geq k'(\omega, \epsilon).$$

In particular, this almost sure convergence holds for the case $\epsilon = \frac{N(q_1)-N^\star}{2} > 0$ and since we are only interested in the lower bound, for all $\omega \in \Omega \setminus N$ there exists an $k'(\omega)$ such that

$$-\frac{N(q_1)-N^\star}{2} < \frac{S_{\tau,k}}{\sum_{\ell=\tau}^{k} \gamma_{\ell+1}} \qquad \text{for all } k \geq k'(\omega). \tag{6.50}$$

Rearranging (6.50),

$$\sum_{\ell=\tau}^{k} \gamma_{\ell+1} \frac{N^\star - N(q_1)}{2} \leq \sum_{\ell=\tau}^{k} \gamma_{\ell+1}(\bar{X}_{\ell+1} - N(q_1)), \qquad \text{for all } k \geq k'(\omega)$$

or equivalently we have

$$\sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(\bar{X}_{\ell+1} - \frac{N(q_1)+N^\star}{2} \right) \geq 0, \qquad \text{for all } k \geq k'(\omega).$$

Now let

$$\begin{aligned} A_k &\doteq \left\{ \omega : \sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(X_{\ell+1} - \frac{N(q_1)+N^\star}{2} \right) \geq 0 \right\}, \qquad k = \tau, \tau+1, \ldots. \\ \bar{A}_k &\doteq \left\{ \omega : \sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(\bar{X}_{\ell+1} - \frac{N(q_1)+N^\star}{2} \right) \geq 0 \right\}, \qquad k = \tau, \tau+1, \ldots. \end{aligned}$$

Define $B_n \doteq \bigcap_{k \geq n} A_k$ for $n \geq \tau$, with the convention that $B_n = \emptyset$ if $n < \tau$. Similarly, define $\bar{B}_n \doteq \bigcap_{k \geq n} \bar{A}_k$ for $n \geq \tau$ and $\bar{B}_n = \emptyset$ if $n < \tau$. We have $B_n \nearrow B = \bigcup_{n=\tau}^{\infty} \bigcap_{k \geq n} A_k$ and a similar expression involving $\bar{B}_n$.

Thus, for the trivial algorithm,

$$\mathbf{P}\left[\bar{B}_n \middle| \; \theta_0, X_0 = x; \; \tau < \infty, \upsilon = \infty \right] \nearrow \mathbf{P}\left[\bar{B} \middle| \theta_0, X_0 = x; \; \tau < \infty, \upsilon = \infty \right] = 1. \tag{6.51}$$

But the strong stochastic ordering property (6.48) yields for each $n = 1, 2, \ldots,$

$$\mathbf{P}\left[B_n \middle| \; \theta_0, X_0 = x; \; \tau < \infty, \upsilon = \infty \right] \geq \mathbf{P}\left[\bar{B}_n \middle| \; \theta_0, X_0 = x; \; \tau < \infty, \upsilon = \infty \right],$$

and with (6.51), this implies

$$\mathbf{P}\left[B_n \middle| \theta_0, X_0 = x; \; \upsilon = \infty, \tau < \infty \right] \nearrow \mathbf{P}\left[B \middle| \theta_\tau, X_\tau = x; \; \upsilon = \infty, \tau < \infty \right] = 1$$

Thus for the original algorithm projected on Θ

$$\mathbf{P}\left[A_k \; a.a. \middle| \theta_0, X_0 = x; \; \upsilon = \infty, \tau < \infty \right] = \mathbf{P}_{\theta_0, x}\left[A_k \; a.a. \middle| \; \upsilon = \infty, \tau < \infty \right] = 1,$$

which implies there exist a null set $N \subset \{\upsilon = \infty\} \cap \{\tau < \infty\}$ such that for all $\omega \in \{\upsilon = \infty\} \cap \{\tau < \infty\} \setminus N$ there exists a $K(\omega) < \infty$ such that

$$\sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(X_{\ell+1} - \frac{N(q_1)+N^\star}{2} \right) \geq 0, \qquad \text{for all } k \geq K(\omega).$$

Now for all $\omega \in \{\upsilon = \infty\} \cap \{\tau < \infty\} \setminus N$, the original SA algorithm projected on $\Theta = [0, M]$ can be written

$$\begin{aligned} \theta_{k+1} &= \theta_\tau + \sum_{\ell=\tau}^{k} \gamma_{\ell+1}(X_{\ell+1} - N^\star + z_{\ell+1}), \qquad k \geq \tau \\ &= \theta_\tau + \sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(X_{\ell+1} - \frac{N(q_1)+N^\star}{2} \right) && (6.52) \\ &+ \sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(\frac{N(q_1)+N^\star}{2} - N^\star \right) && (6.53) \\ &+ \sum_{\ell=\tau}^{k} \gamma_{\ell+1} z_{\ell+1}^{-} 1_{\{\theta_\ell + X_{\ell+1} - N^\star > M\}} && (6.54) \\ &+ \sum_{\ell=\tau}^{k} \gamma_{\ell+1} z_{\ell+1}^{+} 1_{\{\theta_\ell + X_{\ell+1} - N^\star < 0\}}, \qquad k \geq \tau && (6.55) \end{aligned}$$

Clearly, each summand of the term (6.55) is positive, and above we have shown that except for a null set on $\{\upsilon = \infty\} \cap \{\tau < \infty\}$ there exists a $K(\omega) < \infty$ such that the sum in (6.52) is positive for all $k \geq K(w)$. Hence, on $\{\upsilon = \infty\} \cap \{\tau < \infty\}$ and for all $k \geq K(\omega)$, we have $\mathbf{P}_{\theta_0,x}$-almost surely that

$$\theta_{k+1} \geq \theta_\tau + \sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(\frac{N(q_1) + N^\star}{2} - N^\star \right) \quad (6.56)$$

$$+ \sum_{\ell=\tau}^{k} \gamma_{\ell+1} z_{\ell+1}^{-} 1_{\{\theta_\ell + X_{\ell+1} - N^\star > M\}} \quad (6.57)$$

The term (6.57) is zero unless of course the algorithm, before the projection operation, would attempt to place the next iterate at a point greater than M. In such a case, we find that the projection operator would then return the next iterate to the nearest point M so that $\upsilon < \infty$; leading to a contradiction (with $\upsilon = \infty$) for this case.

On the other hand, in the case that such a projection does not occur, then the term (6.57) is zero and as $k \to \infty$,

$$\sum_{\ell=\tau}^{k} \gamma_{\ell+1} \left(\frac{N(q_1) + N^\star}{2} - N^\star \right) \nearrow \infty,$$

and this increasing and unbounded lower bound (6.56) forces the iterate sequence to eventually return to $Q_1 = [q_1, M]$, hence $\upsilon < \infty$ again leads to a contradiction. ■

Thus we see $\mathbf{P}_{\theta,x}[\limsup_n \{\theta_n \in Q_2\}] = 1$ since we have shown the iterate almost surely returns to $Q_1 \subset Q_2$ if it leaves Q_2.

We now carry out a second lemma which complements Lemma 6.9 and is a slight variation of BMP's Lemma 12 in [6, p. 235].

Lemma 6.10 (Adaptation of BMP Lemma 12) *On* $\{\tau(Q_2) = \infty\}$,

$$\upsilon(Q_1) < \infty \quad \mathbf{P}_{\theta,x} - a.s., \qquad \text{for all } \theta \in Q_2, \text{ all } x \in \mathsf{X}.$$

Proof: Let us define $\phi : \Theta \to \mathbb{R}$

$$\phi(\theta) = \begin{cases} (\theta^\star - \theta)^3 + 1 & \theta \in [0, \theta^\star] \\ 1 & \theta \in [\theta^\star, M] \end{cases}$$

Let us assume $\{\upsilon(Q_1) = \tau(Q_2) = \infty\}$ which implies that θ_k remains in the interval $[q_2, q_1)$. Thus, no projection operation can occur and

$$\phi(\theta_{m(n,T)}) - \phi(\theta_n) = \sum_{k=n}^{m(n,T)-1} \gamma_{k+1} \frac{d\phi}{d\theta}(\theta_k) h(\theta_k) + \sum_{k=n}^{m(n,T)-1} \gamma_{k+1} \varepsilon_{k+1}(\phi)$$

Also for $\theta \in [q_2, q_1)$, there exists an $\alpha > 0$ such that $\frac{d\phi}{d\theta}(\theta) h(\theta) < -\alpha < 0$ and

$$- \sum_{k=n}^{m(n,T)-1} \gamma_{k+1} \frac{d\phi}{d\theta}(\theta_k) h(\theta_k) \geq \alpha \sum_{k=n}^{m(n,T)-1} \gamma_{k+1} \geq \alpha(T-1).$$

Furthermore, there exists a $\beta > 0$ such that

$$\phi(\theta_{m(n,T)}) - \phi(\theta_n) \geq -\left\{(\theta^\star - q_2)^3 - (\theta^\star - q_1)^3\right\} > \beta > 0.$$

Thus,

$$\sum_{k=n}^{m(n,T)-1} \gamma_{k+1}\varepsilon_{k+1}(\phi) \geq \alpha(T-1) - \beta \geq 1$$

with the second inequality holding for sufficiently large T and this contradicts item 3b) in Proposition 6.7. ■

Thus we see that if $\theta_0 \in Q_2 \setminus Q_1$ then it follows that $\upsilon_1(Q_1) < \infty$ $\mathbf{P}_{\theta,x} - a.s.$

Proposition 6.11 (Adaptation of BMP Prop. 10) *There exists a constant $B < \infty$ such that for all $\theta \in Q_1$ and $x \in \mathsf{X}$*

$$\mathbf{P}_{\theta,x}[\tau_1(Q_2) < \infty] \leq BV(x) \sum_{k=0}^{\infty} \gamma_{k+1}^{1+\widehat{\ell}_1}$$

Proof: Consider the function $\phi : \mathbb{R} \to [1,\infty)$ defined in Lemma 6.10. For all $n = 0, 1, \ldots$ we have

$$\phi(\theta_{n+1}) - \phi(\theta_n) = \gamma_{n+1}\varepsilon_{n+1}(\phi) + \gamma_{n+1}\frac{d\phi}{d\theta}(\theta_n)\left\{h(\theta_n) + z_{n+1}\right\}$$

Thus on $\{\tau(Q_2) < \infty\}$ we have

$$\begin{aligned}\phi(\theta_{\tau(Q_2)}) &= \phi(\theta_0) + \sum_{k=0}^{\tau(Q_2)-1} \gamma_{k+1}\frac{d\phi}{d\theta}(\theta_k)h(\theta_k) \\ &\quad + \sum_{k=0}^{\tau(Q_2)-1} \gamma_{k+1}\frac{d\phi}{d\theta}(\theta_k)z_{k+1} + \sum_{k=0}^{\tau(Q_2)-1} \gamma_{k+1}\varepsilon_{k+1}(\phi)\end{aligned}$$

A few observations: First, if $\theta_0 = \theta \in Q_1$, then $\phi(\theta_{\tau(Q_2)}) - \phi(\theta_0) \geq \phi(q_2) - \phi(q_1)$. Second, if $k < \tau(Q_2)$, we have $\frac{d\phi}{d\theta}(\theta_k)h(\theta_k) \leq 0$. Third, the (nearest point) projection term z_{k+1} can be nonzero only if θ_{k+1} equals 0 or M. Furthermore, z_{k+1} can be positive only if $\theta_{k+1} = 0$ and thus on $k+1 < \tau(Q_2)$, we have $z_{k+1} \leq 0$. Thus,

$$\begin{aligned}(\phi(q_2) - \phi(q_1))1_{\{\tau(Q_2)<\infty\}} &\leq 1_{\{\tau(Q_2)<\infty\}}\left|\sum_{k=0}^{\tau(Q_2)-1} \gamma_{k+1}\varepsilon_{k+1}(\phi)\right| \\ &\leq \sup_n 1_{\{n\leq\tau(Q_2)\}}\left|\sum_{k=0}^{n-1} \gamma_{k+1}\varepsilon_{k+1}(\phi)\right|.\end{aligned}$$

Squaring both sides, rearranging, and taking expectation we find

$$\begin{aligned}\mathbf{P}_{\theta,x}(\tau(Q_2) < \infty) &\leq \frac{\mathbf{E}_{\theta,x}\left[\sup_n 1_{\{n\leq\tau(Q_2)\}}\left|\sum_{k=0}^{n-1}\gamma_{k+1}\varepsilon_{k+1}(\phi)\right|^2\right]}{\left\{\phi(q_2) - \phi(q_1)\right\}^2} \\ &\leq BV(x)\sum_{k=0}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1},\end{aligned}$$

where the second inequality follows from Proposition 6.7 with some $B < \infty$. ■

6.8.2 Localization and Convergence

Before proving convergence, we prove some preliminary results used in the main convergence theorem.

Proposition 6.12 *On $\{\tau_1(Q_2) = \infty\}$, we have θ_n converges to $\theta^\star$, $\mathbf{P}_{\theta,x} - a.s.$ for all $\theta \in Q_2$ and $x \in \mathsf{X}$.*

Proof: This is a simple observation that on $\{\tau(Q_2) = \infty\}$ the original SA algorithm has the following equality:

$$\begin{aligned} \theta_{k+1} &= \Pi_\Theta \left\{\theta_k + \gamma_{k+1} H(\theta_k, X_{k+1}) + \gamma_{k+1}^2 \rho_{k+1}(\theta_k, X_{k+1})\right\}, \qquad k = 0, 1, \ldots \\ &= \Pi_{Q_2} \left\{\theta_k + \gamma_{k+1} H(\theta_k, X_{k+1}) + \gamma_{k+1}^2 \rho_{k+1}(\theta_k, X_{k+1})\right\}. \end{aligned}$$

The noise condition (KC4) follows from the example of Chapter 3 with only slight modification to account for the nonlinear transition probability dependence (6.24).

Thus, convergence follows immediately by the Kushner-Clark Lemma since $h(\theta)$ exists and is continuous on Q_2 while $Q_2 \subset DA(\theta^\star)$, $\theta^\star \in Q_2^o$, and θ_k is obviously bounded due to the compact projection. ■

Next, we define two distributions $\mathbf{P}_{n;\theta,x}$ and $\check{\mathbf{P}}_{n;\theta,x}$ as in [6, p. 233]. In general, let $\mathbf{P}_{n;\theta,x}$ denote the distribution of $\{(\theta_{n+k}, X_{n+k}), k = 0, 1, \ldots\}$ given $X_n = x$ and $\theta_n = \theta$ produced by the algorithm

$$\begin{aligned} \theta_{k+1} &= \Pi_\Theta \left\{\theta_k + \gamma_{k+1} H(\theta_k, X_{k+1}) + \gamma_{k+1}^2 \rho_{k+1}(\theta_k, X_{k+1})\right\}, \qquad k = n, n+1, \ldots \\ \theta_n &= \theta \\ X_n &= x \end{aligned}$$

Then $\check{\mathbf{P}}_{n;\theta,x}$ is defined as the distribution of $\{(\check{\theta}_{n+k}, \check{X}_{n+k}), k = 0, 1, \ldots\}$ produced by the same algorithm with a step-size sequence shifted forward by n, i.e.

$$\begin{aligned} \check{\theta}_{k+1} &= \Pi_\Theta \left\{\check{\theta}_k + \gamma_{n+k+1} H(\check{\theta}_k, \check{X}_{k+1}) + \gamma_{n+k+1}^2 \rho_{k+1}(\check{\theta}_k, \check{X}_{k+1})\right\}, \qquad k = 0, 1, \ldots \\ \check{\theta}_0 &= \theta \\ \check{X}_0 &= x \end{aligned}$$

It's clear that $\mathbf{P}_{n;\theta,x}$ is equivalent to $\check{\mathbf{P}}_{n;\theta,x}$.

For convenience, we follow the notation of [6] and define a more compact way of expressing almost sure convergence. For any $\epsilon > 0$ we denote the event

$$\{\theta_k \to \theta^\star\} \doteq \bigcup_{m \geq n} \bigcap_{k=m}^{\infty} \{\|\theta_k - \theta^\star\| \leq \epsilon\}.$$

Theorem 6.13 (Adaptation of BMP Thm. 13) *For all $\theta \in Q_1$ and $x \in \mathsf{X}$*

$$\mathbf{P}_{n;\theta,x}\left[\{\theta_k \to \theta^\star\}\right] \geq 1 - BV(x) \sum_{k=n}^{\infty} \gamma_{k+1}^{1+\widehat{\ell}_1}, \qquad n = 0, 1, \ldots.$$

Proof: By conditioning we have

$$\begin{aligned}\mathbf{P}_{n;\theta,x}\left[\{\theta_k \to \theta^\star\}\right] &= \mathbf{P}_{n;\theta,x}\left[\{\theta_k \to \theta^\star\}|\tau(Q_2)<\infty\right]\mathbf{P}_{n;\theta,x}\left[\tau(Q_2)<\infty\right] \\ &+ \mathbf{P}_{n;\theta,x}\left[\{\theta_k \to \theta^\star\}|\tau(Q_2)=\infty\right]\mathbf{P}_{n;\theta,x}\left[\tau(Q_2)=\infty\right]\end{aligned}$$

Then by Proposition 6.12, we have

$$\mathbf{P}_{n;\theta,x}\left[\{\theta_k \to \theta^\star\}|\tau(Q_2)=\infty\}\right] = 1,$$

hence

$$\begin{aligned}\mathbf{P}_{n,\theta,x}\left[\{\theta_k \to \theta^\star\}\right] &\geq \mathbf{P}_{n,\theta,x}[\tau(Q_2)=\infty] \\ &= 1-\mathbf{P}_{n,\theta,x}[\tau(Q_2)<\infty] \\ &= 1-BV(x)\sum_{k=n}^{\infty}\gamma_{k+1}^{1+\widehat{\ell}_1}.\end{aligned}$$

■

The following is an application of an argument in BMP's Theorem 15 applied to our problem.

Lemma 6.14 *On the event* $\limsup_n\{\theta_n \in Q_1\}$*, we have*

$$\liminf_n\{\theta_n \in Q_2\}, \quad \mathbf{P}_{\theta,x}-a.s.$$

Proof: Let us assume that $\{\theta_n \in Q_1\ i.o.\}$ and $\{\theta_n \in Q_2^c\ i.o.\}$ and we will show a contradiction. Under these assumptions θ_n successively visits Q_1 and Q_2^c so let us define as in (6.46)-(6.47) for each $n=0,1,\ldots$:

$$\begin{aligned}\tau_n &= \tau_n(Q_2) \\ \upsilon_n &= \upsilon_n(Q_1)\end{aligned}$$

Clearly, we have

$$\phi(\theta_{\tau_{n+1}})-\phi(\theta_{\upsilon_n}) \geq (\theta^\star-q_2)^3-(\theta^\star-q_1)^3>0 \qquad \text{for all } n=0,1,\ldots. \tag{6.58}$$

Then, by the same argument from the proof of Proposition 6.11 we have

$$\phi(\theta_{\tau_{n+1}})-\phi(\theta_{\upsilon_n}) \leq \sum_{k=\upsilon_n}^{\tau_{n+1}-1}\gamma_{k+1}\varepsilon_{k+1}(\phi), \qquad n=0,1,2,\ldots$$

Observe from (6.16) that

$$\sum_{k=\upsilon_n}^{\tau_{n+1}-1}\gamma_{k+1}\varepsilon_{k+1}(\phi) = E_{\tau_{n+1}}-E_{\upsilon_n}+\widetilde{\eta}_{\upsilon_n;\tau_{n+1}}, \qquad n=1,2,\ldots,$$

so that

$$0 < \left|(\theta^\star - q_2)^3 - (\theta^\star - q_1)^3\right| \leq \left|\sum_{k=\upsilon_n}^{\tau_{n+1}-1} \gamma_{k+1}\varepsilon_{k+1}(\phi)\right|$$
$$\leq \left|E_{\tau_{n+1}} - E_{\upsilon_n}\right| + \left|\widetilde{\eta}_{\upsilon_n;\tau_{n+1}}\right|, \qquad n = 1, 2, \ldots \quad (6.59)$$

In Proposition 6.7 we showed that E_n converges **P**-a.s. to a finite rv. Hence, by Cauchy's criterion, with the exception of a null subset of Ω, for each $\epsilon_1 > 0$, there exists a $N_1 = N_1(\epsilon_1, \omega) < \infty$ such that

$$|E_n - E_m| < \epsilon_1, \qquad \text{for all } m \geq N_1 \text{ and all } n > m. \quad (6.60)$$

This certainly implies that

$$\left|E_{\tau_{m+1}} - E_{\upsilon_m}\right| < \epsilon_1 \text{for all } m \geq N_1. \quad (6.61)$$

Also, from Proposition 6.7, we have $\eta_{m;n}1_{\{\theta_m \in Q_2\}}1_{\{\theta_{n-1} \in Q_2\}}$ converges to 0 almost surely under $\mathbf{P}_{\theta,x}$ as $m, n \to \infty$. Thus, except on another null set, for each $\epsilon_2 > 0$ there exists an $N_2 = N_2(\epsilon_2, \omega) < \infty$ such that

$$\left|\eta_{m;n}\ 1_{\{\theta_m \in Q_2\}}1_{\{\theta_n \in Q_2\}}\right| < \epsilon_2, \qquad \text{for all } m, n \geq N_2. \quad (6.62)$$

And this implies

$$\left|\eta_{\upsilon_m;\tau_{m+1}}\right| = \left|\eta_{\upsilon_m;\tau_{m+1}-1}\ 1_{\{\theta_{\upsilon_m} \in Q_2\}}1_{\{\theta_{\tau_{m+1}-1} \in Q_2\}}\right| < \epsilon_2, \qquad \text{for all } m \geq N_2. \quad (6.63)$$

Therefore, we find that except on the union of the two null sets that (6.61) and (6.63) contradicts (6.59) if ϵ_1 and ϵ_2 are chosen small enough. ■

Theorem 6.15 (Adaptation of BMP Thm. 15)

$$\mathbf{P}_{\theta,x}\left[\{\theta_k \to \theta^\star\}\right] = 1.$$

Proof: Together, Lemma's 6.9 and 6.10 imply that $\mathbf{P}_{\theta,x}\left[\limsup_n \{\theta_n \in Q_1\}\right] = 1$.

Thus,

$$\mathbf{P}_{\theta,x}\left[\{\theta_k \to \theta^\star\}^c\right] = \mathbf{P}_{\theta,x}\left[\{\limsup_n \{\theta_n \in Q_1\}\} \bigcap \{\theta_k \to \theta^\star\}^c\right]$$
$$\leq \mathbf{P}_{\theta,x}\left[\{\liminf_n \{\theta_n \in Q_2\}\} \bigcap \{\theta_k \to \theta^\star\}^c\right]$$

where the last inequality follows from Lemma 6.14. Continuing, we apply [8, Thm. 4.1]

$$\begin{aligned}
\mathbf{P}_{\theta,x}\left[\{\theta_k \to \theta^\star\}^c\right] &\leq \mathbf{P}_{\theta,x}\left[\{\liminf \{\theta_n \in Q_2\}\} \bigcap \{\theta_k \to \theta^\star\}^c\right] \\
&\leq \liminf_n \mathbf{P}_{\theta,x}\left[\{\theta_n \in Q_2\} \bigcap \{\theta_k \to \theta^\star\}^c\right] \\
&\leq \liminf_n \mathbf{E}_{\theta,x}\left[1_{\{\theta_n \in Q_2\}}\mathbf{P}_{n;\theta_n,X_n}\left[\{\theta_k \to \theta^\star\}^c\right]\right] \\
&\leq \liminf_n \mathbf{E}_{\theta,x}\left[B1_{\{\theta_n \in Q_2\}}V(X_n)\sum_{k \geq n}\gamma_{k+1}^{1+\widehat{\ell}_1}\right] \\
&\leq BC_D V(x) \liminf_n \sum_{k \geq n}\gamma_{k+1}^{1+\widehat{\ell}_1} \\
&= 0,
\end{aligned}$$

where we have applied Theorem 6.13 and (D0') in the last lines. ■

6.9 Concluding Remarks

This chapter demonstrates an approach to showing convergence described in [64] of combining a local Kushner-Clark ODE method with a parametric recurrence argument established to the specific problem. The method of proving recurrence combines a strong stochastic ordering property with some adapted versions of arguments in BMP. These new versions of BMP's results serve to demonstrate a possible approach in showing the parameter returns almost surely to a compact set. Although it is likely that this approach can be generalized to a certain extent, it is also clear that an important element of this stability argument depends on the strong stochastic ordering property which is specifically tailored to the problem at hand. This chapter outlines an approach which may prove useful in applying SA to other problems which may contain a region of instability or transience.

Chapter 7

Stochastic Approximations Driven by Sample Averages

7.1 Introduction

In this chapter, we study the convergence properties of an entirely different class of projected stochastic approximations which arise naturally in problems of on-line parametric optimization of discrete event dynamical systems, e.g., queueing systems and Petri net models [16, 2]. These algorithms are driven by *sample averages* defined on a well-structured state processes and operate at two different time scales, with state transitions occurring more frequently than parameter updates. For non-random integers $\{\ell_{n+1}, n = 0, 1, \ldots\}$, the stochastic approximations of interest are of the form

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} g(\theta_n, Y_{n+1}) \right\}, \qquad n = 0, 1, \ldots \tag{7.1}$$

with

$$Y_{n+1} = \frac{1}{\ell_{n+1}} \sum_{\ell=1}^{\ell_{n+1}} f(\theta_n, X_{n+1,\ell}) \qquad n = 0, 1, \ldots \tag{7.2}$$

for a state process $\{X_{n+1,\ell},\ \ell = 1, \ldots\}$ taking values in some state space X, and Borel mappings $f : \Theta \times \mathsf{X} \to \mathbb{R}^d$ and $g : \Theta \times \mathbb{R}^d \to \mathbb{R}^p$. In words, with iterate θ_n just returned by the algorithm, we observe or simulate the $(n+1)^{rst}$ state process for ℓ_{n+1} units of time with the understanding that the probability of the sequence $\{X_{n+1,\ell},\ \ell = 1, \ldots\}$ are fully determined by the parameter value θ_n and the final state reached in the previous evaluation interval, i.e., X_{n,ℓ_n}. At the end of the $(n+1)^{rst}$ evaluation interval, the sample average (7.2) is computed, and the algorithmic step is then completed by returning iterate θ_{n+1} according to (7.1).

Whenever such algorithms arise, we can invariably write $h(\theta) = g(\theta, F(\theta))$ for some *known* mapping g and some quantity $F(\theta)$ with is obtainable only through observation or simulation the state process at operating point θ. Fortunately, it is often the case that

$$F(\theta) = \lim_{L \uparrow \infty} \frac{1}{L} \sum_{\ell=1}^{L} f(\theta, \xi_\ell) \quad \mathbf{P}_{\theta,x} a.s. \tag{7.3}$$

where $\{\xi_\ell, \ell = 0, 1, \ldots\}$ is a generic X-valued random sequence modeling the time evolution of the system, and $\mathbf{P}_{\theta,x}$ denotes the probability measure on the set of system trajectories when starting in state x under parameter θ. This suggest that for ℓ_{n+1} large, under appropriate conditions on g, the rvs Y_{n+1} given by (7.2) and $g(\theta_n, Y_{n+1})$ can be viewed as good approximations to $F(\theta_n)$ and $h(\theta_n)$, respectively. Therefore, if the deterministic algorithm

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \Pi_\Theta \left\{ \theta_n + \gamma_{n+1} h(\theta_n) \right\}, \qquad n = 0, 1, \ldots$$

converges to some $\theta^\star$, then we should expect the stochastic version (7.1)–(7.2) to also converge, say almost surely, to the same point as the size of the sampling window grows unbounded.

Specifically, we develop a framework for investigating the a.s. convergence of the iterate sequence $\{\theta_n,\ n = 0, 1, \ldots\}$ generated by (7.1)–(7.2). We start essentially with no structural assumptions on the probability measures $\{\mathbf{P}_{\theta,x}, \theta \in \Theta, x \in \mathsf{X}\}$ governing the statistical behavior of the state process; it is only assumed that the law of large numbers such as (7.3) is in effect. Our focus is on charting a sequence of basic steps to help establish a.s. convergence; these steps point to a set of technical conditions that need to be verified for each specific application.

Our framework for this alternative algorithm also relies on the ODE method [61] which generally proceeds in two separate steps. The first step relies on the Kushner–Clark Lemma to identify a deterministic ODE, the stability properties of which determine the limit points of $\{\theta_n,\ n = 0, 1, \ldots\}$. The second step, which is probabilistic in nature and depends on the algorithm, involves showing that asymptotically (in the mode of convergence of interest) the output sequence of the original algorithm behaves like the solution to the ODE. Although general conditions are given in [61] for successfully completing this last step, these conditions are not usually checkable in terms of the model data. Nevertheless, in this chapter we show that this second step is determined by the *exponential* convergence of the rvs $\{g(\theta_n, Y_{n+1}) - h(\theta_n),\ n = 0, 1, \ldots\}$, i.e., for every $\epsilon > 0$, the convergence

$$\lim_{n \to \infty} \mathbf{P}\left[\| g(\theta_n, Y_{n+1}) - h(\theta_n) \| \geq \epsilon \right] = 0$$

takes place exponentially fast (with respect to the sequence of sample durations $\{\ell_{n+1},\quad n = 0, 1, \ldots\}$) This exponential convergence viewpoint was already implicit in the work of Dupuis and Simha [29] who consider schemes such as (7.1)–(7.2) but with *constant* step-sizes, i.e., $\gamma_{n+1} \doteq \gamma$, $n = 0, 1, \ldots$, and under the assumption that the rvs $\{X_{n+1,\ell},\ \ell = 1, \ldots\}$ are i.i.d. On the other hand, the work of Dupuis and Simha [29] does not make use of the ODE method but instead relies on the convergence properties of a deterministic discrete time algorithm associated with the original stochastic algorithm.

Going one step further, we give explicit conditions which ensure this exponential convergence. As in [29], we do so by invoking a *uniform* Large Deviations upper bound for the collection of probability measures $\{\mathbf{P}_{\theta,x},\ \theta \in \Theta,\ x \in \mathsf{X}\}$. Here, this upper bound is uniform in both the parameter θ *and* the initial condition x and with some functional $I : \mathbb{R}^d \to [0, \infty]$, and takes the form

$$\limsup_{L \to \infty} \frac{1}{L} \log \sup_{\theta \in \Theta,\ x \in \mathsf{X}} \mathbf{P}_{\theta,x} \left[\frac{1}{L} \sum_{\ell=1}^{L} f(\theta, \xi_\ell) - F(\theta) \in C \right] \leq -I(C) \qquad (7.4)$$

for every closed subset C of $\mathbb{R}^d$. We are able to find checkable conditions to ensure that (7.4) holds. The approach for doing this is in the spirit of the Ellis–Gartner Theorem [30, Thm. II.2. p.3]; in fact, we broaden the applicability of the ideas of Dupuis and Simha to more general classes of state processes.

To demonstrate the applicability of the results obtained herein, we specialize them to two specific classes of state processes. For the first class, the successive states form a sequence of i.i.d. rvs as in [29] so the results are only briefly outlined. In the second class, the state sequence is a finite state time–homogeneous Markov chain; an important class of processes often used in applications. In both cases we identify simple and checkable conditions that ensure the validity of a uniform Large Deviations upper bound.

The chapter is organized as follows: In Section 2 we introduce the basic building blocks that we use in Section 3 to formally define the class of stochastic approximations investigated here. The basic convergence result is stated as Theorem 1 in Section 4. Next, exponential convergence is shown in Section 5 to be the key condition for establishing a.s. convergence via the ODE method. In turn, this condition of exponential convergence is related in Section 6 to the existence of a uniform large deviations upper bounds. Conditions to ensure such uniform large deviations upper bounds are derived in Section 7. Several specific situations are treated in Sections 8 and 9, namely, the cases where the process driving the sample averages is i.i.d. and finite–state Markov; in all cases, we give concrete conditions for uniform large deviations upper bounds to exist.

7.2 The Basic Ingredients

Before defining the stochastic approximation procedures considered here, we devote this section to introducing the basic building blocks used in the formal definitions of Section 7.3. Throughout the discussion, p, s and d are fixed positive integers. We assume given a closed convex subset Θ of $\mathbb{R}^p$, and a Borel subset X of $\mathbb{R}^s$. Furthermore, let $f : \Theta \times \mathsf{X} \to \mathbb{R}^d$ and $g : \Theta \times \mathbb{R}^d \to \mathbb{R}^p$ denote fixed Borel mappings. Additional assumptions will imposed in due time.

We consider two sequences $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$ and $\{\ell_{n+1},\ n = 0, 1, \ldots\}$ which take values in $\mathbb{R}_+$ and $\mathbb{N}$, respectively. The following assumptions are enforced:

- **(S')** The $\mathbb{R}_+$–valued sequence $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$ is monotone decreasing with $\gamma_n \downarrow 0$ ($n \uparrow \infty$), under the usual divergence condition $\sum_{n=0}^{\infty} \gamma_{n+1} = \infty$.
- **(L)** The $\mathbb{N}$–valued sequence $\{\ell_{n+1},\ n = 0, 1, \ldots\}$ is monotone increasing and for all $\beta > 0$ satisfies the condition
$$\sum_{n=0}^{\infty} \exp(-\beta \ell_{n+1}) < \infty. \tag{7.5}$$

Condition **(L)** implies that $\ell_n \uparrow \infty$ as $n \to \infty$ but the reverse implication is not always true. Indeed, in the case $\ell_n = \lceil \log n \rceil$ for all $n = 1, 2, \ldots$ we see that (7.5) fails for $0 < \beta \leq 1$ since then $\sum n^{-\beta} = \infty$.

Let X^∞ be the infinite Cartesian product of X with itself, and denote by $\mathcal{B}(\mathsf{X}^\infty)$ the standard σ–field on X^∞. We write a generic element ξ of X^∞ as $\xi = (x, x_1, \ldots)$ where $x, x_1, \ldots$ are all elements of X. The coordinate process $\{\xi_\ell,\ \ell = 0, 1, \ldots\}$ is then simply defined by

$$\xi_0(\xi) \doteq x, \qquad \xi_\ell(\xi) \doteq x_\ell, \qquad \xi \in \mathsf{X}^\infty,\ \ell = 1, \ldots$$

We postulate the existence of a family $\{\mathbf{P}_{\theta,x},\ \theta \in \Theta,\ x \in \mathsf{X}\}$ of probability measures on $\mathcal{B}(\mathsf{X}^\infty)$ such that

$$\mathbf{P}_{\theta,x}[\xi_0 = x] = 1, \qquad \theta \in \Theta,\ x \in \mathsf{X}.$$

For technical reasons, we again assume a measurable functional dependence in θ and x:

(P0) For every $L = 1, 2, \ldots$, the mapping $\Theta \times \mathsf{X} \to \mathbb{R} : (\theta, x) \to \mathbf{P}_{\theta,x}[\xi_\ell \in B_\ell,\ \ell = 1, \ldots, L]$ is Borel measurable for all possible choices of Borel subsets $B_1, \ldots, B_L$ in $\mathcal{B}(\mathsf{X})$.

We also assume that a strong law of large numbers is in effect:

(P5) There exists a Borel mapping $F : \Theta \to \mathbb{R}^d$ such that for all θ in Θ and x in X, we have

$$\lim_{L \uparrow \infty} \frac{1}{L} \sum_{\ell=1}^{L} f(\theta, \xi_\ell) = F(\theta) \qquad \mathbf{P}_{\theta,x} - a.s.$$

7.3 Model and Assumptions

In order to define the stochastic approximation procedures, we start with a sample space Ω equipped with a σ–field of events $\mathcal{F}$. The measurable space $(\Omega, \mathcal{F})$ is assumed large enough to carry a double array of X–valued rvs $\{X_{n,\ell},\ \ell = 1, \ldots, \ell_n;\ n = 0, 1, \ldots\}$ where we use the convention $\ell_0 = 1$. We define the Θ–valued rvs $\{\theta_n,\ n = 0, 1, \ldots\}$ through the recursion

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \Pi_\Theta \left\{\theta_n + \gamma_{n+1} g(\theta_n, Y_{n+1})\right\} \qquad n = 0, 1, \ldots \tag{7.6}$$

where we use the notation

$$Y_{n+1} \doteq \frac{1}{l_{n+1}} \sum_{\ell=1}^{\ell_{n+1}} f(\theta_n, X_{n+1,\ell}). \qquad n = 0, 1, \ldots$$

In (7.6), Π_Θ denotes the nearest-point projection operator on the set Θ; it is well defined since Θ is assumed closed and convex.

Next, we introduce the filtration $\{\mathcal{F}_n,\ n = 0, 1, \ldots\}$ on $(\Omega, \mathcal{F})$ by setting

$$\begin{aligned} \mathcal{F}_n &\doteq \sigma\{\theta_m, X_{m,\ell},\ \ell = 1, \ldots, \ell_m,\ m = 0, 1, \ldots, n\} \\ &= \sigma\{\theta_0; X_{m,\ell},\ \ell = 1, \ldots, \ell_m,\ m = 0, 1, \ldots, n\} \qquad n = 0, 1, \ldots \end{aligned}$$

where the equality follows since the rvs θ_m, $m = 1, 2, \ldots, n$, are fully determined by the rvs θ_0, $X_{0,1}$, and $X_{m+1,\ell}$, $\ell = 0, 1, \ldots, \ell_{m+1}$, $m = 1, \ldots, n-1$.

Finally, given a probability measure ν on $\mathcal{B}(\Theta \times \mathsf{X})$, we postulate the existence of a probability measure $\mathbf{P}$ on $(\Omega, \mathcal{F})$ satisfying

$$\mathbf{P}[\theta \in B, X_{0,1} \in B_1] = \nu(B \times B_1), \qquad B \in \mathcal{B}(\Theta), B_1 \in \mathcal{B}(\mathsf{X})$$

and

$$\begin{aligned} &\mathbf{P}[X_{n+1,\ell} \in B_\ell,\ \ell = 1, \ldots, \ell_{n+1} | \mathcal{F}_n] \\ &= \mathbf{P}_{\theta_n, X_{n,\ell_n}}[\xi_\ell \in B_\ell,\ \ell = 1, \ldots, \ell_{n+1}] \qquad n = 1, \ldots \end{aligned}$$

for Borel subsets $B_1, \ldots, B_{\ell_{n+1}}$ in $\mathcal{B}(\mathsf{X})$. The existence of such a set-up is readily justified by the Daniell–Kolmogorov consistency theorem [69, p. 94] on $\Theta \times \mathsf{X} \times \mathsf{X}^\infty$ in the usual manner.

7.4 The Convergence Results

The presentation of the main convergence results is simplified by the following notation: Setting

$$h(\theta) \doteq g(\theta, F(\theta)), \qquad \theta \in \Theta$$

we define the $\mathbb{R}^p$–valued rvs $\{\varepsilon_{n+1},\ n = 0, 1, \ldots\}$ by

$$\varepsilon_{n+1} \doteq g(\theta_n, Y_{n+1}) - h(\theta_n) \quad n = 0, 1, \ldots \tag{7.7}$$

so that the recursion (7.6) now becomes

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \Pi_\Theta \left\{\theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\varepsilon_{n+1}\right\}. \qquad n = 0, 1, \ldots \tag{7.8}$$

The relevant assumptions concerning these quantities are the following:

(H3) The mapping $h : \Theta \to \mathbb{R}^p$ is continuous.

(E2) The $\mathbb{R}^p$–valued rvs $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$ converge exponentially to the zero vector, in the sense that for every $\epsilon > 0$, there exist a finite integer $n(\epsilon)$ and a positive constant $K(\epsilon)$ such that

$$\mathbf{P}[\|\varepsilon_{n+1}\| \geq \epsilon] \leq \exp\left(-\ell_{n+1} K(\epsilon)\right), \qquad n \geq n(\epsilon).$$

Sufficient conditions for **(E2)** are provided in Section 6 and follow from the availability of uniform large deviations upper bounds.

With the projection operator Π_Θ, we associate the transformation $\bar{\Pi}_\Theta : \Theta \times \mathbb{R}^p \to \mathbb{R}^p$ given by

$$\bar{\Pi}_\Theta(\theta, v) \doteq \lim_{\Delta \downarrow 0} \frac{\Pi_\Theta\{\theta + \Delta v\} - \theta}{\Delta}, \qquad \theta \in \Theta,\ v \in \mathbb{R}^p.$$

The limiting ODE corresponding to (7.8) is

$$\theta(0) \in \Theta, \qquad \frac{d\theta}{dt}(t) = \bar{\Pi}_\Theta \left\{\theta(t), h(\theta(t))\right\}, \qquad t \geq 0. \tag{7.9}$$

The unconstrained case corresponds to $\Theta \doteq \mathbb{R}^p$, in which case the recursion (7.6) reduces to

$$\theta_0 \in \Theta, \qquad \theta_{n+1} = \theta_n + \gamma_{n+1} h(\theta_n) + \gamma_{n+1}\varepsilon_{n+1} \qquad n = 0, 1, \dots \tag{7.10}$$

and the limiting ODE corresponding to (7.10) becomes

$$\theta(0) \in \mathbb{R}^p, \qquad \frac{d\theta}{dt}(t) = h(\theta(t)), \qquad t \geq 0. \tag{7.11}$$

The basic convergence result for this algorithm is contained in Theorem 7.1.

Theorem 7.1 *Consider the stochastic approximation scheme (7.8) under assumptions* **(S')**, **(L)**, **(P0)**, **(P5)**, **(H3)** *and* **(E2)**. *Let $\theta^\star$ be a point in the interior Θ^o which is a locally asymptotically stable solution to (7.9), and let $\mathcal{DA}(\theta^\star)$ denote its domain of attraction. Assume the following conditions hold:*

(i): The $\mathbb{R}^p$-valued random variables $\{\theta_n, n = 0, 1, \dots\}$ are bounded with probability one, i.e.

$$\mathbf{P}[\sup_n ||\theta_n|| < \infty] = 1. \tag{7.12}$$

(ii): There exists a compact set $Q \subseteq \mathcal{DA}(\theta^\star)$, such that

$$\mathbf{P}[\theta_n \in Q \quad i.o.]. \tag{7.13}$$

Then $\lim_{n\to\infty} \theta_n = \theta^\star \qquad \mathbf{P} - a.s.$

Theorem 7.1 is a simple consequence of the so–called ODE method as developed by Kushner and Clark [61] once we observe the following lemma. In some cases, it can be difficult to validate conditions (7.12) and (7.13) in Theorem 7.1. There is, however, one class of situations which naturally occur in practice where (7.12) is automatically satisfied, namely when Θ is a *compact* subset of $\mathbb{R}^p$. Furthermore, (7.13) is automatically satisfied when $\mathcal{DA}(\theta^\star) = \Theta$.

Given the gain sequence $\{\gamma_{n+1},\ n = 0, 1, \dots\}$, recall the sequence of times $\{t_n,\ n = 0, 1, \dots\}$ defined by

$$t_0 \doteq 0, \qquad t_{n+1} \doteq \sum_{i=0}^{n} \gamma_{i+1}, \qquad n = 0, 1, \dots$$

and set

$$m(t) \doteq \max\{n \in \mathbb{N} :\ t_n \leq t\}, \qquad t \geq 0.$$

Lemma 7.2 *Assume condition* **(S')**, **(L)** *and* **(E2)** *to be enforced. For every $T > 0$ and $\epsilon > 0$, we have*

$$\lim_{n\to\infty} \mathbf{P}\left[\sup_{j\geq n} \max_{0\leq t\leq T} \left\| \sum_{i=m(jT)}^{m(jT+t)-1} \gamma_{i+1}\varepsilon_{i+1} \right\| \geq \epsilon \right] = 0. \tag{7.14}$$

Proof: Fix $T > 0$ and $\epsilon > 0$. We readily observe that

$$
\begin{aligned}
& \mathbf{P} \left[\sup_{j \geq n} \max_{0 \leq t \leq T} \left\| \sum_{i=m(jT)}^{m(jT+t)-1} \gamma_{i+1} \varepsilon_{i+1} \right\| \geq \epsilon \right] \\
\leq & \ \mathbf{P} \left[\sup_{j \geq n} \max_{0 \leq t \leq T} \sum_{i=m(jT)}^{m(jT+t)-1} \gamma_{i+1} ||\varepsilon_{i+1}|| \geq \epsilon \right] \\
\leq & \ \mathbf{P} \left[\sup_{j \geq n} \sum_{i=m(jT)}^{m(jT+T)-1} \gamma_{i+1} \|\varepsilon_{i+1}\| \geq \epsilon \right] \\
\leq & \ \sum_{j=n}^{\infty} \mathbf{P} \left[\sum_{i=m(jT)}^{m(jT+T)-1} \gamma_{i+1} \|\varepsilon_{i+1}\| \geq \epsilon \right] \\
\leq & \ \sum_{j=n}^{\infty} \mathbf{P} \left[\max_{m(jT) \leq i < m(jT+T)} \|\varepsilon_{i+1}\| \cdot \sum_{i=m(jT)}^{m(jT+T)-1} \gamma_{i+1} \geq \epsilon \right]. \qquad (7.15)
\end{aligned}
$$

It is plain from (5.1)–(5.2) that $t_{m(jT+T)} \leq jT + T$ and $t_{m(jT)+1} = t_{m(jT)} + \gamma_{m(jT)+1} \geq jT$ for all $j = 0, 1, \ldots$. Therefore, we have

$$
\begin{aligned}
\sum_{i=m(jT)}^{m(jT+T)-1} \gamma_{i+1} &= t_{m(jT+T)} - t_{m(jT)} \\
&\leq (jT + T) - (jT - \gamma_{m(jT)+1}), \\
&\leq T + \gamma_1 \qquad j = 0, 1, \ldots \qquad (7.16)
\end{aligned}
$$

since the gain sequence $\{\gamma_{n+1}, \ n = 0, 1, \ldots\}$ is monotone decreasing. Combining (7.15) and (7.16), with $\epsilon' = \frac{\epsilon}{T+a_1}$, we get

$$
\begin{aligned}
& \sum_{j=n}^{\infty} \mathbf{P} \left[\max_{m(jT) \leq i < m(jT+T)} \|\varepsilon_{i+1}\| \cdot \sum_{i=m(jT)}^{m(jT+T)-1} \gamma_{i+1} \geq \epsilon \right] \\
\leq & \ \sum_{j=n}^{\infty} \mathbf{P} \left[\max_{m(jT) \leq i < m(jT+T)} \|\varepsilon_{i+1}\| \geq \epsilon' \right] \\
\leq & \ \sum_{j=n}^{\infty} \sum_{i=m(jT)}^{m(jT+T)-1} \mathbf{P} \left[\|\varepsilon_{i+1}\| \geq \epsilon' \right] \\
= & \ \sum_{i=m(nT)}^{\infty} \mathbf{P} \left[\|\varepsilon_{i+1}\| \geq \epsilon' \right]. \qquad (7.17)
\end{aligned}
$$

Next, upon invoking the exponential convergence condition **(E2)**, we can assert the existence of a finite integer $n(\epsilon')$ and of a positive constant $K(\epsilon')$ such that

$$
\mathbf{P}[\|\varepsilon_{i+1}\| \geq \epsilon'] \leq \exp\left(-\ell_{i+1} K(\epsilon')\right), \qquad i \geq n(\epsilon'). \qquad (7.18)
$$

Finally, we select a finite integer $n^\star$ such that $m(n^\star T) \geq n(\epsilon')$; such a selection is always possible since $\lim_{n \uparrow \infty} m(nT) = \infty$ by virtue of **(S')**. For all $n \geq n^\star$, we easily conclude from (7.17) and

(7.18) that

$$\mathbf{P}\left[\sup_{j\geq n}\max_{0\leq t\leq T}\left\|\sum_{i=m(jT)}^{m(jT+t)-1}\gamma_{i+1}\varepsilon_{i+1}\right\|\geq\epsilon\right]\leq\sum_{i=m(nT)}^{\infty}\exp\left(-\ell_{i+1}K(\epsilon')\right) \tag{7.19}$$

and the convergence (7.14) is now an immediate consequence of (7.19) and of the sumability condition **(L)** since $\lim_{n\uparrow\infty} m(nT)=\infty$. ■

7.4.1 Sufficient Conditions for (E2)

A sufficient condition for **(E2)** can be derived from uniform Large Deviations upper bounds as we now show. First a few definitions: With the coordinate process $\{\xi_\ell,\ \ell=0,1,\ldots\}$ defined on the measurable space $(\mathsf{X}^\infty,\mathcal{B}(\mathsf{X}^\infty))$, we write

$$\bar{S}_L(\theta)\doteq\frac{1}{L}\sum_{\ell=1}^{L}f(\theta,\xi_\ell),\qquad\theta\in\Theta.\qquad L=1,\ldots \tag{7.20}$$

Since condition **(P5)** can be rephrased as $\lim_{L\to\infty}\bar{S}_L(\theta)=F(\theta)\ \mathbf{P}_{\theta,x}$–a.s., the rate of convergence implied by **(E2)** thus suggests that the law of large numbers associated with the sample averages (7.20) be complemented by a Large Deviations upper bound. This is essentially the content of condition **(U1)**:

(U1) The collection of probability measures $\{\mathbf{P}_{\theta,x},\ \theta\in\Theta,\ x\in\mathsf{X}\}$ satisfies a *uniform* Large Deviations upper bound principle with respect to (the sample averages associated with) f if there exists a closed convex function $I:\mathbb{R}^d\to[0,\infty]$ such that

$$\limsup_{L\to\infty}\frac{1}{L}\log\sup_{\theta\in\Theta,x\in\mathsf{X}}\mathbf{P}_{\theta,x}[\bar{S}_L(\theta)-F(\theta)\in C]\leq-\inf_{z\in C}I(z).$$

for every closed subset C of $\mathbb{R}^d$.

We refer to I as the rate functional associated with this uniform Large Deviations upper bound principle. By itself condition **(U1)** is not sufficient for **(E2)**, so we supplement **(U1)** by imposing additional conditions **(U2)**–**(U3)** on the rate functional I:

(U2) The rate function I in **(U1)** is level compact, i.e., the set $\{z\in\mathbb{R}^d: I(z)\leq r\}$ is compact for all $r\geq 0$; and

(U3) The rate function I in **(U1)** has the property that $I(z)=0$ if and only if $z=0$.

In a brief but necessary interlude, we pause to establish the following consequence of **(U1)**–**(U3)**.

Lemma 7.3 *Assume* **(U2)–(U3)** *to hold for some closed convex rate function* $I : \mathbb{R}^d \to [0, +\infty]$. *Then, for every* $\delta > 0$, *we have*

$$K(\delta) \doteq \inf_{z \in C_\delta} I(z) > 0. \tag{7.21}$$

where $C_\delta \doteq \{z \in \mathbb{R}^d : \|z\| \geq \delta\}$.

Proof: We need only consider the case $0 \leq K(\delta) < \infty$, for otherwise (6.3) trivially holds. Therefore, there exists an C_δ-valued sequence $\{z_n, n = 1, 2, \ldots\}$ such that the values $\{I(z_n), n = 1, 2, \ldots\}$ are non–increasing with $\lim_{n\to\infty} I(z_n) = K(\delta)$. Hence, for every $\eta > 0$, there exists a positive integer $n(\eta)$ such that

$$K(\delta) \leq I(z_n) \leq K(\delta) + \eta, \qquad n \geq n(\delta). \tag{7.22}$$

Invoking the level–compactness condition **(U2)**, we conclude from (7.22) that a convergent C_δ-valued subsequence $\{z_{n_j}, j = 1, 2, \ldots\}$ can be extracted from $\{z_n, n \geq n(\delta)\}$. If $z^\star$ denotes the limit of this convergent subsequence, then $z^\star$ necessarily belongs to the closed set C_δ so that $z^\star \neq 0$. By the lower semicontinuity of I, we see that

$$K(\delta) = \lim_{j\uparrow\infty} I(z_{n_j}) \geq I(z^\star) > 0 \tag{7.23}$$

with the strict positivity follows from **(U3)** since $z^\star \neq 0$. ■

We also need some additional conditions on the mapping g.

(G) The mapping $\mathbb{R}^d \to \mathbb{R}^p : z \to g(\theta, z + F(\theta))$ is continuous at $z = 0$ uniformly in θ, i.e., for every $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ with the property that if $\|z\| < \delta(\epsilon)$, then

$$\sup_{\theta\in\Theta} \|g(\theta, z + F(\theta)) - g(\theta, F(\theta))\| < \epsilon. \tag{7.24}$$

In many situations of interest, the mapping g is independent of θ and takes the form

$$g(\theta, x) = \phi(x), \qquad \theta \in \Theta,\ x \in \mathbb{R}^d \tag{7.25}$$

for some Borel mapping $\phi : \mathbb{R}^d \to \mathbb{R}^p$. In such cases, condition **(G)** is guaranteed by requiring that ϕ be uniformly continuous on $\mathbb{R}^d$. This latter requirement is satisfied when ϕ is Lipschitz continuous, a condition obviously met for the frequent choice $\phi(x) \doteq x$.

Theorem 7.4 *Assume conditions* **(U1)–(U3)** *and* **(G)** *to hold. Then the rvs* $\{\varepsilon_{n+1},\ n = 0, 1, \ldots\}$ *satisfy condition* **(E2)**.

Proof: Fix $\epsilon > 0$, and for each $L = 1, 2, \ldots$, set

$$G_L(\theta, x) \doteq \mathbf{P}_{\theta,x}\left[\left\|g(\theta, \bar{S}_L(\theta)) - h(\theta)\right\| \geq \epsilon\right], \qquad \theta \in \Theta,\ x \in \mathsf{X}.$$

From the definition (7.7) we readily observe that

$$
\begin{aligned}
\mathbf{P}[\gamma_{n+1} \notin B_\epsilon] &= \mathbf{E}[\mathbf{P}[\|g(\theta_n, Y_{n+1}) - h(\theta_n)\| \geq \epsilon | \mathcal{F}_n]] \\
&= \mathbf{E}[G_{\ell_{n+1}}(\theta_n, X_{n,\ell_n})] \qquad n = 0, 1, \ldots
\end{aligned} \tag{7.26}
$$

where in the last equality we have made use of the requirement (7.7) on $\mathbf{P}$.

By virtue of the uniform continuity condition **(G)**, there exists $\delta(\epsilon) > 0$ such that (7.24) holds whenever $\|z\| < \delta(\epsilon)$. Hence, for each θ in Θ, the event

$$
\left[\left\|g(\theta, \bar{S}_L(\theta)) - h(\theta)\right\| \geq \epsilon\right] \subset \left[\left\|\bar{S}_L(\theta) - F(\theta)\right\| \geq \delta(\epsilon)\right].
$$

Therefore, with the notation of Lemma 7.3, we conclude that

$$
G_L(\theta, x) \leq \mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in C_{\delta(\epsilon)}], \qquad \theta \in \Theta,\ x \in \mathsf{X} \qquad L = 1, 2, \ldots \tag{7.27}
$$

where $C_{\delta(\varepsilon)} \doteq \{z \in \mathbb{R}^d : \|z\| \geq \delta(\epsilon)\}$.

Next, we pick η in the interval $(0, K(\delta(\epsilon)))$ which is non-empty due to Lemma 7.3. Under condition **(U1)** if $K(\delta(\epsilon))$ is finite, then there exists a finite integer $L(\eta)$ such that

$$
\sup_{\theta \in \Theta, x \in \mathsf{X}} \mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in C_{\delta(\epsilon)}] \leq e^{-L(K(\delta(\epsilon)) - \eta)}, \qquad L \geq L(\eta), \tag{7.28}
$$

while if $K(\delta(\epsilon)) = \infty$, then for every $R > 0$, there exists a finite integer $L(R)$ such that

$$
\sup_{\theta \in \Theta, x \in \mathsf{X}} \mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in C_{\delta(\epsilon)}] \leq e^{-LR}, \qquad L \geq L(R). \tag{7.29}
$$

In any event, either from (7.28) or (7.29), we can assert the existence of a finite integer $L^\star$ and of a strictly positive constant $K^\star$ such that

$$
\sup_{\theta \in \Theta, x \in \mathsf{X}} \mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in C_{\delta(\epsilon)}] \leq e^{-LK^\star}, \qquad L \geq L^\star. \tag{7.30}
$$

Using this information in (7.27), we readily conclude from (7.26) that **(E2)** indeed holds. ■

From the proof of Theorem 7.4 we see that the law of large numbers **(P5)** automatically holds under conditions **(U1)**–**(U3)**. This is a simple consequence of the bound (7.30) and of the Borel–Cantelli Lemma.

7.4.2 Sufficient Conditions for (U1)–(U3)

In this section, we develop a uniform large deviations upper bound for a parameterized sequence of dependent random variables. This result generalizes a similar result obtained by Dupuis and Simha [29] for i.i.d. rvs.

For $L = 1, 2, \cdots$, we define

$$
c_L(t, \theta, x) \doteq \frac{1}{L} \log \mathbf{E}_{\theta,x}[\exp\left(\left\langle t, L\bar{S}_L(\theta) - LF(\theta)\right\rangle\right)], \qquad t \in \mathbb{R}^d,\ \theta \in \Theta,\ x \in \mathsf{X} \tag{7.31}
$$

and

$$
c_L(t) \doteq \sup_{\theta \in \Theta, x \in \mathsf{X}} c_L(t, \theta, x), \qquad t \in \mathbb{R}^d. \tag{7.32}
$$

As in [30], we require that the following assumptions **(C1)**–**(C2)** hold:

(C1) For all t in $\mathbb{R}^d$, the limit $c(t) \doteq \lim_{L\to\infty} c_L(t)$ exists where we allow $+\infty$ both as a limit value and as an element in the sequence $\{c_L(t),\ L = 1, 2, \ldots\}$.

(C2) The mapping $c : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a closed convex function whose effective domain $\mathcal{D}(c) \doteq \{t \in \mathbb{R}^d : c(t) < \infty\}$ has a non–empty interior containing the point $t = 0$.

The Legendre–Fenchel transform of c is the closed convex mapping $I : \mathbb{R}^d \to [0, +\infty]$ defined by

$$I(z) \doteq \sup_{t\in\mathbb{R}^d} \{\langle t, z\rangle - c(t)\}, \qquad z \in \mathbb{R}^d, \tag{7.33}$$

and for notational convenience, we write

$$I(S') = \inf_{z\in A} I(z), \qquad A \subseteq \mathbb{R}^d.$$

The first result of this section shows that the conditions **(C1)**–**(C2)** are sufficient conditions for **(U1)**. The proof, which follows, is similar to that given by Dupuis and Simha for the i.i.d. case discussed in [29].

Theorem 7.5 *Assume* **(P0)**, **(P5)** *and* **(C1)**–**(C2)** *to hold. Then, for any closed subset C of $\mathbb{R}^d$, the inequality*

$$\limsup_{L\to\infty} \frac{1}{L} \log \sup_{\theta\in\Theta, x\in\mathsf{X}} \mathbf{P}_{\theta,x}\left[\bar{S}_L(\theta) - F(\theta) \in C\right] \leq -I(C) \tag{7.34}$$

holds.

Proof: Let C be a closed subset of $\mathbb{R}^d$. If $I(C) = 0$, then (7.34) automatically holds. Hence, we need only establish (7.34) when $0 < I(C)$, and thus two cases need to be considered, namely $0 < I(C) < \infty$ and $I(C) = \infty$.

Case 1: If $0 < I(C) < \infty$, then ϵ can be selected in the interval $(0, I(C))$. By Gärtner's covering lemma [30], there exist r distinct non–zero points $t_1, \ldots, t_r$ in $\mathcal{D}(c)$ such that

$$C \subset \bigcup_{i=1}^{r} H_+(t_i, I(C) - \epsilon) \tag{7.35}$$

where $H_+(t, \alpha) \doteq \{z \in \mathbb{R}^d : \langle t, z\rangle - c(t) \geq \alpha\}$.

The integer r and the points $t_1, \ldots, t_r$ depend on both ϵ and C, but *not* on θ and x. For each $i = 1, \ldots, r$, the point t_i belongs to $\mathcal{D}(c)$, so that $c(t_i)$ is finite and $c_L(t_i)$ is therefore also finite for L large enough, say $L \geq L'$ – it is plain that L' can be chosen the same for all $i = 1, \ldots, r$.

Fix θ in Θ, x in X and $L \geq L$. With these facts in mind, we readily see from (7.35) that

$$\begin{aligned}
&\mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in C] \\
&\quad\leq \sum_{i=1}^{r} \mathbf{P}_{\theta,x}\left[\left\langle t_i, \bar{S}_L(\theta) - F(\theta)\right\rangle - c(t_i) \geq I(C) - \epsilon\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{r} \mathbf{P}_{\theta,x}[\left\langle t_i, L\bar{S}_L(\theta) - LF(\theta)\right\rangle \geq L(c(t_i) + I(C) - \epsilon)] \\
&= \sum_{i=1}^{r} \mathbf{P}_{\theta,x}[\exp(\left\langle t_i, L\bar{S}_L(\theta) - LF(\theta)\right\rangle) \geq \exp\left(L(c(t_i) + I(C) - \epsilon)\right)] \\
&\leq \sum_{i=1}^{r} \mathbf{E}_{\theta,x}[\exp\left(\left\langle t_i, L\bar{S}_L(\theta) - LF(\theta)\right\rangle\right)] \exp\left(-L(c(t_i) + I(C) - \epsilon)\right) \\
&= \sum_{i=1}^{r} \exp\left(Lc_L(t_i, \theta, x)\right) \exp\left(-L(c(t_i) + I(C) - \epsilon)\right) \\
&= \sum_{i=1}^{r} \exp\left(L(c_L(t_i, \theta, x) - c(t_i))\right) \exp\left(-L(I(C) - \epsilon)\right) \\
&\leq \sum_{i=1}^{r} \exp(L(c_L(t_i) - c(t_i))) \exp\left(-L(I(C) - \epsilon)\right). \qquad (7.36)
\end{aligned}
$$

The last inequality follows from the fact that $c_L(t, \theta, x) \leq c_L(t)$ for all θ in Θ, x in X, and all $t \in \mathbb{R}^d$.

Since $\lim_{L\to\infty} c_L(t_i) = c(t_i)$, $i = 1, \ldots, r$, we can find for every $\delta > 0$, an integer $L^* = L^*(\delta)$ such that $L^* \geq L'$ and $|c_L(t_i) - c(t_i)| < \delta$, $i = 1, \ldots, r$, whenever $L \geq L^*$, and therefore

$$
\sup_{i=1,\ldots,r} \exp(L(c_L(t_i) - c(t_i))) \leq \exp(L\delta), \qquad L \geq L^*.
$$

Using this last fact, we conclude from (7.36) that

$$
\mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in C] \leq r \exp\left(-L(I(C) - \epsilon - \delta)\right), \qquad L \geq L^*
$$

whence

$$
\sup_{\theta\in\Theta, x\in\mathsf{X}} \mathbf{P}_{\theta,x}\left[\bar{S}_L(\theta) - F(\theta) \in C\right] \leq r \exp\left(-L(I(C) - \epsilon - \delta)\right), \qquad L \geq L^* \qquad (7.37)
$$

since the integer r and the points $t_1, t_2, \cdots, t_r$ depend on the set C and on the chosen ϵ, and the integer L^* depends on C, ϵ and the chosen $\delta > 0$. It then follows that

$$
\limsup_{L\to\infty} \frac{1}{L} \log \sup_{\theta\in\Theta, x\in\mathsf{X}} \mathbf{P}_{\theta,x}\left[\bar{S}_L(\theta) - F(\theta) \in C\right] \leq -(I(C) - \epsilon - \delta) \qquad (7.38)
$$

and (7.34) now follows since (7.38) holds for all ϵ in the interval $(0, I(C))$ and for all $\delta > 0$.

Case 2: If $I(C) = \infty$, then fix $R > 0$ and by Gärtner's covering lemma [30], there again exists r distinct non–zero points $t_1, \ldots, t_r$ in $\mathcal{D}(c)$ such that

$$
C \subset \bigcup_{i=1}^{r} H_+(t_i, R). \qquad (7.39)
$$

The integer r and the points $t_1, \ldots, t_r$ depend on both R and C, but *not* on θ and x. For each $i = 1, \ldots, r$, the point t_i belongs to $\mathcal{D}(c)$, so that $c(t_i)$ is finite and $c_L(t_i)$ is therefore also finite for L large enough, say $L \geq L''$ – it is again plain that L'' can be chosen the same for all $i = 1, \ldots, r$.

Fix θ in Θ and $L \geq L''$. By the same arguments as the one leading to (7.36), this time with the help of (7.39), we get

$$\begin{aligned}
&\mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in C] \\
&\leq \sum_{i=1}^{r} \mathbf{P}_{\theta,x}\left[\left\langle t_i, \bar{S}_L(\theta) - F(\theta)\right\rangle - c(t_i) \geq R\right] \\
&= \sum_{i=1}^{r} \mathbf{P}_{\theta,x}[\left\langle t_i, L\bar{S}_L(\theta) - LF(\theta)\right\rangle \geq L(c(t_i) + R)] \\
&= \sum_{i=1}^{r} \mathbf{P}_{\theta,x}[\exp(\left\langle t_i, L\bar{S}_L(\theta) - LF(\theta)\right\rangle) \geq \exp\left(L(c(t_i) + R)\right)] \\
&\leq \sum_{i=1}^{r} \mathbf{E}_{\theta,x}[\exp\left(\left\langle t_i, L\bar{S}_L(\theta) - LF(\theta)\right\rangle\right)] \exp\left(-L(c(t_i) + R)\right) \\
&= \sum_{i=1}^{r} \exp\left(Lc_L(t_i, \theta, x)\right) \exp\left(-L(c(t_i) + R)\right) \\
&= \sum_{i=1}^{r} \exp\left(L(c_L(t_i, \theta, x) - c(t_i))\right) \exp\left(-L(R)\right) \\
&\leq \sum_{i=1}^{r} \exp(L(c_L(t_i) - c(t_i))) \exp\left(-L(R)\right). \qquad (7.40)
\end{aligned}$$

The last inequality follows from the fact that $c_L(t, \theta, x) \leq c_L(t)$ for all θ in Θ and all $t \in \mathbb{R}^d$.

Since $\lim_{L\to\infty} c_L(t_i) = c(t_i)$, $i = 1, \ldots, r$, we can find for every $\delta > 0$, an integer $L^* = L^*(\delta)$ such that $L^* \geq L''$ and $|c_L(t_i) - c(t_i)| < \delta$, $i = 1, \ldots, r$, whenever $L \geq L^*$, and as in **Case 1**, we can conclude from 7.40 that

$$\sup_{\theta\in\Theta, x\in\mathsf{X}} \mathbf{P}_{\theta,x}\left[\bar{S}_L(\theta) - F(\theta) \in C\right] \leq r \exp\left(-L(R - \delta)\right), \qquad \text{for all } L \geq L^*. \qquad (7.41)$$

Therefore,

$$\limsup_{L\to\infty} \frac{1}{L} \log \sup_{\theta\in\Theta, x\in\mathsf{X}} \mathbf{P}_{\theta,x}\left[\bar{S}_L(\theta) - F(\theta) \in C\right] \leq -(R - \delta)$$

and since R and δ are arbitrary

$$\limsup_{L\to\infty} \frac{1}{L} \log \sup_{\theta\in\Theta, x\in\mathsf{X}} \mathbf{P}_{\theta,x}\left[\bar{S}_L(\theta) - F(\theta) \in C\right] \leq -\infty. \qquad (7.42)$$

Combining the two separate cases, we have for closed sets C such that $I(C) > 0$

$$\limsup_{L\to\infty} \frac{1}{L} \log \sup_{\theta\in\Theta, x\in\mathsf{X}} \mathbf{P}_{\theta,x}\left[\bar{S}_L(\theta) - F(\theta) \in C\right] \leq -I(C). \qquad (7.43)$$

■

We are now in a position to give a set of sufficient conditions for **(U1)**–**(U2)** to hold.

Lemma 7.6 *Assume conditions* **(P0)**, **(P5)** *and* **(C1)**–**(C2)** *to hold. Then the collection of probability measures* $\{P_{\theta,x},\ \theta \in \Theta,\ x \in \mathsf{X}\}$ *satisfies the uniform Large Deviations upper bound condition* **(U1)**. *The corresponding rate functional* I *given by (7.33) satisfies* **(U2)**.

Proof: That condition **(U1)** holds is immediate from Theorem 7.5 since the Legendre-Fenchel transform I given by (7.34) is a closed convex mapping.

Next, we show that I given by (7.33) is indeed level–compact. We do so by slightly modifying the arguments of Ellis' Theorem V.1, Part (f) in [30, pp. 6-7]: For $r \geq 0$, we consider the level set $K_r \doteq \{z \in \mathbb{R}^d : I(z) \leq r\}$ which is closed by the lower semicontinuity of c. From the definition of I, we see that

$$\langle t, z\rangle \leq I(z) + c(t) \leq r + c(t), \qquad t \in \mathbb{R}^d,\ z \in K_r \tag{7.44}$$

and therefore, for each $R > 0$, we get

$$\sup_{\|t\| \leq R} \langle t, z\rangle \leq r + \sup_{\|t\| \leq R} c(t) \qquad z \in K_r. \tag{7.45}$$

In view of **(C2)**, we can choose R such that the closed ball $B_R \doteq \{z \in \mathbb{R}^d : \|z\| \leq R\}$ is contained in the effective domain $\mathcal{D}(c)$, in which case c is continuous on B_R. Therefore, by standard results from real analysis, we can assert that

$$\sup_{\|t\| \leq R} |c(t)| \doteq A < \infty \qquad \text{and} \qquad \sup_{\|t\| \leq R} \langle t, z\rangle = R\,\|z\|\,. \tag{7.46}$$

Combining (7.45) and (7.46), we find $\|z\| \leq R^{-1}(r + A)$ for all z in K_r, and the level set K_r is thus compact since closed and bounded. ■

We address next the crucial condition **(U3)** on the rate functional I. We do so in two steps; the first step being contained in the next lemma and the second step appearing in Theorem 7.8.

Lemma 7.7 *Assume* **(P0)**, **(P5)** *and* **(C1)**–**(C2)** *to hold. If* $z = 0$, *then* $I(z) = 0$, *in which case* $I(\mathbb{R}^d) = 0$.

Proof: In order to show that $z = 0$ implies $I(z) = 0$, we proceed by contradiction, and assume $I(0) > 0$: We claim that $\epsilon > 0$ can always be selected small enough so that $I(B_\epsilon) > 0$, where again $B_\epsilon \doteq \{z \in \mathbb{R}^d : ||z|| \leq \epsilon\}$. Indeed, recall [90, Thm. 10.1, p. 82] that the convex function I is continuous on the interior of $\mathcal{D}(I)$ (which contains the origin $z = 0$). By choosing ϵ small enough, we can ensure that B_ϵ is contained in the interior of $\mathcal{D}(I)$, and that $I(z) > 0$ for all z in B_ϵ, this last fact by continuity under the assumption $I(0) > 0$. Continuity over the compact set B_ϵ yields $0 < I(B_\epsilon) < \infty$, and by Theorem 7.5, for $0 < \eta < I(B_\epsilon)$ there exists a finite integer $L^\star$ such that

$$\sup_{\theta \in \Theta, x \in \mathsf{X}} \mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in B_\epsilon] \leq e^{-L(I(B_\epsilon) - \eta)}, \qquad L \geq L^\star. \tag{7.47}$$

Now taking the limit in (7.47), we readily conclude that

$$\lim_{L\to\infty} \mathbf{P}_{\theta,x}[\bar{S}_L(\theta) - F(\theta) \in B_\epsilon] = 0, \qquad \theta \in \Theta,\ x \in \mathsf{X}, \tag{7.48}$$

or equivalently, that the sample averages (7.20) do not converge in probability, thus not a.s. This conclusion is in direct contradiction with **(P5)** and the assumption $I(0) > 0$ cannot hold. Thus $I(0) = 0$, and we readily get $I(\mathbb{R}^d) = 0$ from the fact that $I(z) \geq 0$ for all z in $\mathbb{R}^d$. ■

In order to show that $I(z) = 0$ implies $z = 0$, we need an additional condition on the function c defined in **(C1)**–**(C2)**.

(C3) The function c is (Fréchet-) differentiable at $t = 0$, i.e., its gradient $\nabla c(t)$ exists at $t = 0$, with $\nabla c(0) = 0$.

We are now ready to present the main result of this section:

Theorem 7.8 *Under* **(P0)**, **(P5)**, **(C1)**–**(C3)**, *the conditions* **(U1)**–**(U3)** *hold.*

Proof: Combining Theorem 7.5 with Lemmas 7.6 and 7.7, we see that all of **(U1)**–**(U3)** hold *except* for the property that I achieves its global minimum at the unique point $z = 0$, but this follows directly from [30, Thm. V.1 (g), pp. 6-7], under **(C3)**. ■

7.5 IID State Processes

We refer to the *i.i.d.* case as the situation characterized by some collection $\{\mu_\theta,\ \theta \in \Theta\}$ of probability measures on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$ such that for Borel subsets $B_1, \ldots, B_L$ in $\mathcal{B}(\mathsf{X})$,

$$\mathbf{P}_{\theta,x}[\xi_\ell \in B_\ell,\ \ell = 1, \ldots, L] = \prod_{\ell=1}^{L} \mu_\theta(B_\ell) \quad L = 1, \ldots \tag{7.49}$$

for all θ in Θ and x in X. Assumption **(P0)** is satisfied by requiring that the collection $\{\mu_\theta,\ \theta \in \Theta\}$ be measurable in the sense that for every Borel subset B in $\mathcal{B}(\mathsf{X})$, the mapping $\theta \to \mu_\theta(B)$ is Borel measurable. The validity of **(P5)** is guaranteed by the strong law of large numbers for i.i.d. sequences provided the moment condition

$$\int_\Theta |f(\theta, x)| d\mu_\theta(x) < \infty, \quad \theta \in \Theta$$

holds, in which case we have

$$F(\theta) = \int_\Theta f(\theta, x) d\mu_\theta(x), \quad \theta \in \Theta.$$

The Borel measurability of F then follows readily from the Borel measurability of f by standard arguments.

With (7.49), the definition (7.31) yields

$$c_L(t,\theta,x) = \log \int_{\mathsf{X}} e^{\langle t, f(\theta,x)-F(\theta)\rangle} d\mu_\theta(x), \quad t \in \mathbb{R}^d,\ \theta \in \Theta,\ x \in \mathsf{X} \tag{7.50}$$

for all $L = 1, 2, \ldots$, and **(C1)** holds in the form

$$\begin{aligned} c(t) &= \lim_{L\to\infty} \sup_{\theta\in\Theta, x\in\mathsf{X}} c_L(t,\theta,x) \\ &= \sup_{\theta\in\Theta} \log \int_{\mathsf{X}} e^{\langle t, f(\theta,x)-F(\theta)\rangle} d\mu_\theta(x), \quad t \in \mathbb{R}^d. \end{aligned} \tag{7.51}$$

For each θ in Θ and each x in X, the mapping $t \to c_L(t,\theta,x)$ given by (7.50) is convex by Hölder's Inequality [25, Lemma 2.2.31, p. 37], and is lower semicontinuous by Fatou's Lemma. Since for proper convex functions, closedness is equivalent to lower semicontinuity follows from of [90, Thm. 7.1, pp. 51-52], we conclude that the mapping $t \to c_L(t,\theta,x)$ is closed and convex. That c is closed and convex follows from the fact that both convexity [90, Thm. 10.8, p. 90], and closedness are preserved under the supremum operation. The other conditions **(C2)**–**(C3)** can be investigated in specific instances.

Example: As a simple example we consider the case when for each θ in Θ, the measure μ_θ is a Gaussian measure on $\mathbb{R}^d$ with mean $m(\theta)$ and covariance matrix $\Sigma(\theta)$. If $f(\theta,x) \doteq x$,

$$c(t) = \frac{1}{2} \sup_{\theta\in\Theta} \langle t, \Sigma(\theta)t\rangle, \quad t \in \mathbb{R}^d. \tag{7.52}$$

Consequently, $\nabla c(t)$ exists at $t = 0$ if there exists a symmetric positive semi–definite matrix Σ such that $\Sigma(\theta) \le \Sigma$ for all θ in Θ (where inequalities are with respect to the usual ordering on the cone of symmetric positive semi–definite matrices).

7.6 Markov Chains with Finite State Space

In the *Markovian* case, we assume the existence of a collection $\{K_\theta, \theta \in \Theta\}$ of measurable transition kernels $\mathsf{X} \times \mathcal{B}(\mathsf{X}) \to [0,1]$ such that

$$\mathbf{P}_{\theta,x}[\xi_{L+1} \in B | \xi_\ell,\ \ell = 0,1,\ldots,L] = K_\theta(\xi_L; B), \quad B \in \mathcal{B}(\mathsf{X}) \quad L = 0,1,\ldots \tag{7.53}$$

for all θ in Θ and x in X. Condition **(P0)** follows by requiring that for each x in X and each Borel subset B in $\mathcal{B}(\mathsf{X})$, the mapping $\theta \to K_\theta(x_L; B)$ is Borel measurable on Θ. Condition **(P5)** is guaranteed by imposing some ergodicity conditions on the Markov chains with transition kernels $\{K_\theta, \theta \in \Theta\}$.

Of particular interest for applications are the models involving finite state Markov chains. We develop this important case by finding explicit conditions on the one–step transition probabilities which ensure the various conditions discussed so far. The set–up is as follows: The state space X is a finite set, say with s elements, and following [25, 30], we identify X with the canonical basis $\{e_1, \ldots, e_s\}$ of $\mathbb{R}^s$, i.e., $\langle e_x, e_y\rangle = \delta_{xy}$, $x, y = 1, \ldots, s$; the notation x and

e_x, $x = 1, \ldots, s$, is used interchangeably. For each θ in Θ, with the transition kernel K_θ we associate the $s \times s$ stochastic matrix $P(\theta) \doteq (P_\theta(x,y))$ whose entries are defined by

$$P_\theta(x,y) \doteq K_\theta(x;\{y\}), \quad x, y \in \mathsf{X}. \tag{7.54}$$

In short, under each of the measures $\mathbf{P}_{\theta,x}$, the rvs $\{\xi_\ell,\ \ell = 0, 1, \ldots\}$ form a time–homogeneous Markov chain with one–step transition matrix P_θ.

Next, given the mapping $f : \Theta \times \mathsf{X} \to \mathbb{R}^d$, we seek to evaluate the corresponding quantities (7.31)–(7.32). Fixing t in $\mathbb{R}^d$, we define the $s \times s$ matrices $\{\Pi_{t,\theta}, \theta \in \Theta\}$ by

$$\Pi_{t,\theta}(x,y) \doteq P_\theta(x,y) e^{\langle t, f(\theta,y) - F(\theta)\rangle}, \quad \theta \in \Theta,\ x, y \in \mathsf{X}. \tag{7.55}$$

As in [25, pp. 58-61], we have

$$c_\ell(t,\theta,x) = \frac{1}{\ell} \log \left\langle e_x, \Pi_{t,\theta}^\ell e \right\rangle, \quad x \in \mathsf{X}, \theta \in \Theta \quad \ell = 1, 2, \ldots \tag{7.56}$$

where e is the element $(1, \ldots, 1)$ of $\mathbb{R}^d$. Armed with this notation, we can now turn to the main results of this section. We begin with an auxiliary result of a technical nature:

Lemma 7.9 *Consider the family of finite state space Markov chains with one–step transition matrices $\{P(\theta),\ \theta \in \Theta\}$. Suppose the following conditions are enforced:*

(i): For each θ in Θ, the one–step transition matrix is irreducible and aperiodic; and

(ii): For each x and y in X, the mappings $\theta \to P_\theta(x,y)$ and $\theta \to f(\theta,x)$ are continuous on Θ.

Then, for each t in $\mathbb{R}^d$, the following statements are true:

1. *For each θ in Θ, the non–negative matrix $\Pi_{t,\theta}$ is irreducible and primitive; its spectral radius $\rho(\Pi_{t,\theta})$ coincides with the largest positive eigenvalue of $\Pi_{t,\theta}$ which always has multiplicity one, and the eigenvector $u(\Pi_{t,\theta})$ corresponding to $\rho(\Pi_{t,\theta})$ can be selected such that*
$$m_{t,\theta} \doteq \min_i u_i(\Pi_{t,\theta}) > 0 \quad \textit{and} \quad \langle e, u(\Pi_{t,\theta})\rangle = 1; \tag{7.57}$$

2. *The mappings $\theta \to \rho(\Pi_{t,\theta})$ and $\theta \to u(\Pi_{t,\theta})$ are continuous on Θ.*

Proof: (Claim **1.**) Fix t in $\mathbb{R}^d$ and θ in Θ. Since the exponential factors entering the definition (7.56) are strictly positive, it is plain from (ii) that the non–negative matrix $\Pi_{t,\theta}$ is irreducible and primitive [38, Thm. 8, p. 80], and most of Claim **1** is now a simple rephrasing of the Perron–Frobenius theorem [57, Thm. 2.2, p. 545]. The existence of an eigenvector satisfying the normalization condition in (7.57) follows from the positivity condition in (7.57) and the scalability property of eigenvectors.

(Claim **2.**) Fix t in $\mathbb{R}^d$. For each θ in Θ, the stochastic matrix $P(\theta)$ is ergodic by virtue of (i), and therefore admits a unique invariant probability vector $\pi(\theta)$, i.e., $\pi(\theta)' = \pi(\theta)' P(\theta)$

and $\langle e, \pi(\theta)\rangle = 1$; we also have $F(\theta) = \sum_x \pi_x(\theta) f(\theta, x)$ by the Ergodic Theorem for Markov Chains [21, Thm. 2, p. 92]. With this in mind, we note that the continuity assumption (ii) on $\theta \to P(\theta)$ implies the continuity of $\theta \to \pi(\theta)$ since $\rho(\Pi_{t,\theta})$ has multiplicity one for all θ in Θ [58, p. 110]. Therefore, $\theta \to F(\theta)$ is also continuous by the continuity assumption (ii) on f. In short, from (7.56) and assumption (ii) we conclude that the matrix–valued mapping $\theta \to \Pi_{t,\theta}$ is (entrywise) continuous on Θ, whence the mapping $\theta \to \rho(\Pi_{t,\theta})$ is continuous since each eigenvalue is a continuous mapping on the space of square matrices [65, p. 225]. It is now a simple matter to see that the mapping $\theta \to u(\Pi_{t,\theta})$ is continuous: Indeed, for each θ in Θ, the conditions

$$[\Pi_{t,\theta} - \rho(\Pi_{t,\theta}) I_s]\, u = 0 \quad \text{and} \quad \langle e, u\rangle = 1 \tag{7.58}$$

uniquely determine the eigenvector $u(\Pi_{t,\theta})$ since $\rho(\Pi_{t,\theta})$ has multiplicity one. Using this characterization, we can now establish the desired continuity by adapting the arguments of [70, p. 39]. Another argument is available in [58, p. 110]. ■

The validity of the conditions **(C1)**–**(C3)** is now discussed:

Theorem 7.10 *Consider the family of finite state space Markov chains with one–step transition matrices $\{P(\theta),\ \theta \in \Theta\}$, under the assumptions (i)–(ii) of Lemma 7.9. If the parameter set Θ is a compact subset of $\mathbb{R}^p$, then conditions* **(C1)**–**(C3)** *hold with*

$$c(t) \doteq \lim_{\ell\to\infty} \sup_{\theta\in\Theta, x\in \mathsf{X}} c_\ell(t,\theta,x) = \sup_{\theta\in\Theta} \log \rho(\Pi_{t,\theta}), \quad t \in \mathbb{R}^d. \tag{7.59}$$

Proof: (Condition **(C1)**) Fix t in $\mathbb{R}^d$. As pointed out in the proof of Theorem 3.1.2 in [25, p. 60], the limit (7.60) exists and equals

$$c(t,\theta,x) \doteq \lim_{l\to\infty} c_\ell(t,\theta,x) = \log \rho(\Pi_{t,\theta}), \quad \theta \in \Theta,\ x \in \mathsf{X} \tag{7.60}$$

so that

$$\sup_{\theta\in\Theta} \log \rho(\Pi_{t,\theta}) \le \liminf_{\ell\to\infty} c_\ell(t) \tag{7.61}$$

by invoking the definition (7.31)–(7.32). The conclusion (7.59) (including the existence of the limit) will follow if we can establish that

$$\limsup_{\ell\to\infty} c_\ell(t) \le \sup_{\theta\in\Theta} \log \rho(\Pi_{t,\theta}). \tag{7.62}$$

To do so, we fix θ in Θ and x in X. In the notation of Lemma 7.9, $u(\Pi_{t,\theta})$ is the eigenvector of $\Pi_{t,\theta}$ associated with the eigenvalue $\rho(\Pi_{t,\theta})$ such that (7.58) holds. Using the representation (7.57), we readily get

$$\begin{aligned} c_\ell(t,\theta,x) &= \frac{1}{\ell} \log \left\langle e_x, \Pi_{t,\theta}^\ell e \right\rangle \\ &\le \frac{1}{\ell} \log \left\langle e_x, \Pi_{t,\theta}^\ell \frac{u(\Pi_{t,\theta})}{m_{t,\theta}} \right\rangle \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\ell}\log\left\langle e_x, \rho(\Pi_{t,\theta})^{\ell}\frac{u(\Pi_{t,\theta})}{m_{t,\theta}}\right\rangle \\
&\leq \log\rho(\Pi_{t,\theta}) + \frac{1}{\ell}\log\frac{\langle e_x, u(\Pi_{t,\theta})\rangle}{m_{t,\theta}} \\
&\leq \log\rho(\Pi_{t,\theta}) - \frac{1}{\ell}\log m_{t,\theta}. \qquad \ell = 1,2,\ldots \qquad (7.63)
\end{aligned}
$$

Next, upon taking the supremum in (7.63) , we see that

$$
c_\ell(t) = \sup_{\theta\in\Theta, x\in\mathsf{X}} c_\ell(t,\theta,x) \leq \sup_{\theta\in\Theta}\log\rho(\Pi_{t,\theta}) - \frac{1}{\ell}\log\left\{\inf_{\theta\in\Theta} m_{t,\theta}\right\} \quad \ell = 1,2,\ldots \tag{7.64}
$$

and the desired inequality (7.62) follows provided (7.59) can be strengthened to read $\inf_{\theta\in\Theta} m_{t,\theta} > 0$, or equivalently, $\min_i \inf_{\theta\in\Theta} u_i(\Pi_{t,\theta}) > 0$. This last condition is now an immediate consequence of the continuity result of Lemma 7.9 under the compactness condition on Θ.

(Condition **(C2)**) A careful inspection of the proof of (7.59) reveals that in fact we have shown

$$
c(t) = \lim_{\ell\to\infty}\sup_{\theta\in\Theta, x\in\mathsf{X}} c_\ell(t,\theta,x) = \sup_{\theta\in\Theta, x\in\mathsf{X}}\lim_{\ell\to\infty} c_\ell(t,\theta,x), \quad t\in\mathbb{R}^d \tag{7.65}
$$

With this in mind, fix θ in Θ and x in X: For each $\ell = 1,2,\ldots$, the mapping $t \to c_\ell(t,\theta,x)$ is convex (as can be seen by standard arguments [25, Lemma 2.3.9, p. 46] using Hölder's inequality). Therefore, the mapping $t \to c(t,\theta,x)$ is also convex since the pointwise limit of convex mappings is convex [90, Thm. 10.8, p. 90]. Hence, by (7.65), the mapping c is also convex since convexity is preserved under the supremum operation [90, Thm. 5.5, p. 35]. Next, it is plain from (7.60) and (7.65) that $\mathcal{D}(c) = \mathbb{R}^d$ since $0 < \sup_\theta \rho(\Pi_{t,\theta}) < \infty$ by the continuity result of Lemma 7.9 under the compactness condition on Θ. Therefore, c is continuous throughout $\mathbb{R}^d$, thus *a fortiori* closed.

(Condition **(C3)**) We need to establish that the mapping c is differentiable at $t = 0$ with $\nabla c(0) = 0$. We do so in three steps: **Step 1** – Fix θ in Θ and observe from Jensen's inequality that

$$
c_\ell(t,\theta,x) \geq \left\langle t, \mathbf{E}_{x,\theta}\left[\bar{S}_\ell(\theta)\right] - F(\theta)\right\rangle, \quad t\in\mathbb{R}^d,\ x\in\mathsf{X}, \quad \ell = 1,2,\ldots. \tag{7.66}
$$

It also follows from assumption (ii) of Lemma 7.9 that **(P5)** holds, whence $\lim_{\ell\to\infty}\mathbf{E}_{x,\theta}[\bar{S}_\ell(\theta)] = F(\theta)$ via the Bounded Convergence Theorem. Taking the limit in (7.66) and using this last limit result, we get $c(t,\theta,x) \geq 0$ for all t in $\mathbb{R}^d$ and x in X. Therefore, since $c(0,\theta,x) = 0$, we conclude that

$$
\inf_{t\in\mathbb{R}^d} c(t,\theta,x) = c(0,\theta,x) = 0, \quad t\in\mathbb{R}^d,\ x\in\mathsf{X}. \tag{7.67}
$$

Step 2 – Now, for any direction v in $\mathbb{R}^d$, the mapping $\lambda \to \Pi_{\lambda v,\theta}$ is entrywise analytic on $\mathbb{R}$. Hence, the mapping $\lambda \to \rho(\Pi_{\lambda v,\theta})$ is differentiable on $\mathbb{R}$, since in fact analytic on $\mathbb{R}$ [65, Thm. 7.7.1, p. 241] as the largest eigenvalue of $\Pi_{\lambda v,\theta}$ is guaranteed to be of multiplicity one by the Perron–Frobenius theory. Thus from differentiability and (7.67) we readily see by standard arguments that

$$\lim_{\lambda \to 0} D_v(\lambda; \theta) = 0 \tag{7.68}$$

where

$$D_v(\lambda; \theta) \doteq \frac{c(\lambda v, \theta, x)}{\lambda}, \quad \lambda \neq 0. \tag{7.69}$$

In particular, the convex mapping $t \to c(t; \theta, x)$ is Gâteaux-differentiable at $t = 0$ along any direction; its differentiability at $t = 0$ now follows from Theorem 25.2 of Rockafellar [90, p. 244].

Step 3 – It follows from convexity that $\lambda \to D_v(\lambda; \theta)$ is non–decreasing on $(0, \infty)$. Moreover, since $\theta \to \rho(\Pi_{\lambda v, \theta})$ is continuous on Θ for each $\lambda \neq 0$, we see that $\theta \to D_v(\lambda, \theta)$ is also continuous on Θ for each $\lambda > 0$. Therefore, starting with a decreasing sequence $\{\lambda_n, n = 0, 1, \ldots\}$ such that $\lambda_n \downarrow 0$ as $n \to \infty$, we see from (7.68) that $\lim_n D_v(\lambda_n, \theta) = 0$ monotonically for each θ in Θ. By Dini's Theorem [93, p. 195], this last convergence is taking place uniformly on the compact set Θ, i.e., for every $\epsilon > 0$, there exists a finite integer $N(\epsilon)$ such that

$$\sup_{\theta \in \Theta} |D_v(\lambda_n, \theta)| < \epsilon, \quad n \geq N(\epsilon). \tag{7.70}$$

Therefore, combining (7.69) and (7.70), we find that

$$\lim_{n \to \infty} \frac{c(\lambda_n v)}{\lambda_n} = \lim_{n \to \infty} \sup_{\theta \in \Theta} D_v(\lambda_n, \theta) = 0 \tag{7.71}$$

or equivalently, the mapping c is Gâteaux-differentiable at $t = 0$ along all directions, and c is indeed Frèchet-differentiable at $t = 0$ by virtue of Theorem 25.2 of [90, p. 244]. ■

7.7 Markov Chains with Countably Infinite State Space

In this section we reveal some limitations to the Large Deviations upper bounds we use and show that a very simple Markov chain on a countably infinite state space with unbounded function f can fail to meet the condition **(C1)**.

Example: We look at the Markov Chain $\{X_n, \ n = 0, 1, \ldots\}$ which is the random walk on the non-negative integers $\mathbb{N}$ with a reflection at the origin. We take for simplicity Θ to consist of the single point θ so uniformity over Θ is not the issue. An alternative representation for the M.C. can be derived if we define the i.i.d. process $\{U_n, n = 1, 2, \ldots\}$ where $P[U_n = 1] = p = 1 - P[U_n = -1]$ and

$$\begin{aligned}
X_{n+1} &= [X_n + U_{n+1}]^+, \quad n = 0, 1, \ldots \\
&= [[X_{n-1} + U_n]^+ + U_{n+1}]^+ \\
&= \max\{0, U_{n+1}, X_{n-1} + U_n + U_{n+1}\} \\
&= \max\{U_n + 1, U_n + U_{n+1}, X_{n-2} + U_{n-1} + U_n + U_{n+1}\} \\
&= \max\{0, U_n + 1, U_n + U_{n+1}, \ldots, U_2 + \ldots + U_{n+1}, X_0 + U_1 + \ldots + U_{n+1}\} \\
&\geq X_0 + U_1 + \ldots + U_{n+1}
\end{aligned} \tag{7.72}$$

In order to apply our large deviations results the M.C. must satisfy conditions (**C1**) – (**C3**) so let us look at the moment generating function. For $t \geq 0$

$$\begin{aligned}
\mathbf{E}\left[\exp\left(t\sum_{i=1}^{n} X_i\right)\right] &\geq \mathbf{E}\left[\exp\left(t\sum_{i=1}^{n}\left(X_0 + \sum_{j=1}^{i} U_j\right)\right)\right] \\
&= \mathbf{E}\left[\exp\left(ntX_0\right)\right]\mathbf{E}\left[\exp\left(t\sum_{i=1}^{n}\sum_{j=1}^{i} U_j\right)\right] \\
&= \mathbf{E}\left[e^{ntX_0}\right]\mathbf{E}\left[\exp\left(t\sum_{j=1}^{n} U_j(n-(j-1))\right)\right] \\
&= \mathbf{E}\left[e^{ntX_0}\right]\prod_{j=1}^{n}\mathbf{E}\left[e^{t(n-(j-1))U_j}\right] \\
&= \mathbf{E}\left[e^{ntX_0}\right]\prod_{\ell=1}^{n}\mathbf{E}\left[e^{t\ell U_j}\right] \\
&= \mathbf{E}[e^{ntX_0}]\prod_{\ell=1}^{n}\{qe^{-\ell t} + pe^{\ell t}\} \\
&= M_{X_0}(nt)\prod_{\ell=1}^{n} e^{\ell t}(p + qe^{-2\ell t})
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\frac{1}{n}\log\mathbf{E}\left[\exp\left(t\sum_{i=1}^{n} X_i\right)\right] \\
&\quad\geq \frac{1}{n}\log M_{X_0}(nt) + \frac{t}{n}\sum_{\ell}^{n}\ell + \frac{t}{n}\sum_{\ell=1}^{n}\log(p+qe^{-2\ell t}) \\
&\quad= \frac{1}{n}\log M_{X_0}(nt) + \frac{t(n+1)}{2} + \frac{1}{n}\sum_{\ell=1}^{n}\log(p+qe^{-2\ell t})
\end{aligned}$$

Taking $X_0 = x$, we get

$$\begin{aligned}
M_{X_0}(t) &= e^{tx}, \\
\frac{1}{n}\log M_{X_0}(nt) &= tx,
\end{aligned}$$

and

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\sum_{\ell=1}^{n}\log(p+qe^{-2\ell t}) &\geq \lim_{n\to\infty}\frac{1}{n}\sum_{\ell=1}^{n}\log(p) \\
&= \log(p).
\end{aligned}$$

Therefore,

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{E}\left[\exp\left(t\sum_{i=1}^{n} X_i\right)\right] = \infty, \quad t > 0$$

and the point $t = 0$ is not contained in $\mathcal{D}(c)$.

7.7.1 Uniform Markov Chains

It is known [105, 25] that general irreducible Markov chains satisfying a *uniform recurrence* or a *Doeblin condition* possess a large deviations principle. The condition (from [25]) is:

(U) There exists integers $0 < \ell \le N$ and a constant $M \ge 1$ such that for all $x, y \in \mathsf{X}$,

$$\mathbf{P}(x,\cdot) \le \frac{M}{N} \sum_{m=1}^{N} \mathbf{P}^m(y,\cdot)$$

where $\mathbf{P}^m(x,\cdot)$ is the m-step transition probability for initial state x

Since condition (U) is so restrictive as to preclude the simple example above, we must conclude that the large deviations approach we have taken in this chapter has some substantial limitations in regards to the class of state processes that we can accommodate. All is not lost however, and the next section makes the case for the large deviations approach, particularly when g is nonlinear.

7.8 Martingale Method for Convergence

As the approach of this chapter relies heavily on large deviations arguments, this requires the finiteness of certain exponential moments, thus leading naturally to the condition **(L)** on the window sizes $\{\ell_{n+1},\ n = 0, 1, \ldots\}$. Of course such a condition is dictated by the technique adopted here, and is certainly far from necessary as we now show through an example. We shall see that in some cases only finite second order moments suffice in order to yield a.s. convergence, and this in the absence of condition **(L)**, provided an additional condition is imposed on the gain sequence $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$, namely

$$\sum_{n=0}^{\infty} \gamma_{n+1}^2 < \infty. \tag{7.73}$$

To develop this point, we consider an unconstrained scheme (i.e. $\Theta = \mathbb{R}^p$) with $p = d = s$, and $g(\theta, x) = f(\theta, x) = x$ for all θ and x in $\mathbb{R}^p$, so that (1.5)–(1.7) takes the form

$$\theta_0 \in \mathbb{R}^p, \quad \theta_{n+1} = \theta_n + a_{n+1} \frac{1}{\ell_{n+1}} \sum_{\ell=1}^{\ell_{n+1}} X_{n+1,\ell}. \quad n = 0, 1, \ldots \tag{7.74}$$

We put ourselves in the *i.i.d.* case with the additional assumption that for each θ in $\mathbb{R}^p$, the probability measure μ_θ has finite mean $h(\theta)$ and covariance matrix $\Sigma(\theta)$. We assume that $h(\theta) \ne 0$ except for $\theta = \theta^\star$; we take $\theta^\star = 0$ for the sake of convenience. By following an argument of Gladyshev [39], we get the following result whose proof is in the appendix.

Proposition 7.11 *Under the foregoing assumptions on the probability measures* $\{\mu_\theta,\ \theta \in \mathbb{R}^p\}$*, we further assume the conditions*

$$\sup_{\delta^{-1} < ||\theta|| < \delta} \langle \theta, h(\theta) \rangle < 0, \quad \delta \in (0, 1) \tag{7.75}$$

and

$$||h(\theta)||^2 + Tr(\Sigma(\theta)) \leq K(1 + ||\theta||^2), \quad \theta \in \mathbb{R}^p \tag{7.76}$$

for positive constant K. If the gain sequence $\{\gamma_{n+1},\ n = 0, 1, \ldots\}$ satisfies both **(S')** *and (7.73), then $\lim_{n\to\infty} \theta_n = 0$ **P**–a.s. without any additional condition on the window size sequence $\{\ell_{n+1},\ n = 0, 1, \ldots\}$.*

It is plain under the i.i.d. assumption that there is no loss of generality in taking $f(\theta, x) \doteq x$ for all θ and x in $\mathbb{R}^p$. Moreover, projected versions of the algorithm can in principle be addressed by arguments similar to the ones given by Chong and Ramadge [19, Appendix A, p. 365]. Therefore, in the i.i.d. case with linear g, the above Proposition (and its variants) suggest conditions for a.s. convergence which are similar to those given for the standard Robbins–Monro scheme (without averaging), and probably weaker than the ones developed in this chapter so that the framework developed here then seems to provide little improvement, if any. However, the situation is quite different when g is nonlinear; the martingale arguments break down even in the i.i.d. case and the large deviations framework discussed in this chapter now leads to conditions for a.s. convergence.

Appendix A

Proofs and Auxiliary Results

A.1 A Proof of Lemma 2.1

Proof: If (D1) holds, then

$$P_\theta V \leq \lambda V + L, \qquad \text{for each } \theta \in \Theta.$$

Since V is unbounded off petite sets we can define the set

$$C \doteq \left\{ x \in X : V(x) \leq \frac{L}{1/2(1-\lambda)} \right\}$$

By Meyn and Tweedie's Lemma 15.2.8 [79, p. 370], if we let $\beta = \frac{1}{2}(1-\lambda)$ then we have

$$\Delta_\theta V \leq -\beta V - L 1_C, \qquad \text{for each } \theta \in \Theta.$$

which is (D2).

If (D2) holds, then

$$\begin{aligned} \sup_{\theta\in\Theta} \Delta_\theta V &= \sup_{\theta\in\Theta} P_\theta V - V \\ &= -\beta V + b 1_C \end{aligned}$$

which implies

$$\begin{aligned} P_\theta V &\leq V(1-\beta) + b1_C, \qquad \theta \in \Theta \\ &\leq V(1-\beta) + b, \qquad \theta \in \Theta \end{aligned}$$

It is enough to pick just one $\theta \in \Theta$ and together with Lemma 15.2.2 in [79], it follows that V is unbounded off petite sets. ■

A.2 A Proof of Lemma 3.1

Proof: Fix some arbitrary $\ell \in (0, 1)$. The inequality

$$\left| x \log_\rho x \right| \leq C(\ell) x^\ell, \qquad 0 < x \leq 1$$

holds if and only if

$$\left|\log_\rho x\right| \doteq \left|\frac{\ln x}{\ln \rho}\right| \leq C(\ell) x^{\ell-1}, \qquad 0 < x \leq 1$$

holds.

For all ℓ in $(0,1)$ and some yet to be defined constant $C(\ell) > 0$, define on $x \in (0,1]$:

$$\begin{aligned} f(x) &\doteq \left|\frac{\ln x}{\ln \rho}\right| \\ g(x) &\doteq C(\ell) x^{\ell-1}. \end{aligned}$$

The derivatives with respect to x on $(0,1]$ are calculated:

$$\begin{aligned} \frac{df}{dx}(x) &= \frac{1}{x \ln \rho} < 0, \\ \frac{dg}{dx}(x) &= C(\ell)(\ell-1)x^{\ell-2} \\ &= -C(\ell)(1-\ell)\frac{x^{\ell-1}}{x} \leq -C(\ell)(1-\ell)\frac{1}{x} < 0. \end{aligned}$$

We observe that for each ℓ such that $0 < \ell < 1$ there exists a constant $0 < C(\ell) < \infty$ such that

$$\frac{dg}{dx}(x) \leq \frac{df}{dx}(x) < 0, \qquad \text{for all } x \in (0,1].$$

We note that $f(1) = 0$ and $g(1) = C(\ell)$ and

$$\begin{aligned} f(1) - f(x) &= \int_x^1 \frac{df}{dx}(x)\ dx \\ &\geq \int_x^1 \frac{dg}{dx}(x)\ dx \\ &= g(1) - g(x), \qquad 0 < x \leq 1 \end{aligned}$$

Therefore,

$$\begin{aligned} f(x) &\leq g(x) - C(\ell) \\ &\leq g(x), \qquad 0 < x \leq 1 \end{aligned}$$

since we have chosen $C(\ell) > 0$. ■

A.3 Summary of BMP's Theorem 5

The results below are taken verbatim from [6].

BMP define two classes of functions: $Li(p)$ and $Li(Q, L_1, L_2, p_1, p_2)$ where Q is a compact subset of their parameter space D and the remaining arguments are constants.

Definition 2 (BMP) *Define for $p \geq 0$*

$$Li(p) \doteq \left\{ f : \sup_{x_1 \neq x_2} \frac{f(x_1) - f(x_2)}{\|x_1 - x_2\| \, (1 + \|x_1\|^p + \|x_2\|^p)} \right\}.$$

Definition 3 (BMP) *For a function g and an integer p, let*

$$N_p(g) \doteq \sup \left\{ \sup_x \frac{\|g(x)\|}{1 + \|x\|^{p+1}}, \ \sup_{x_1 \neq x_2} \frac{g(x_1) - g(x_2)}{\|x_1 - x_2\| \, (1 + \|x_1\|^p + \|x_2\|^p)} \right\}.$$

Definition 4 (BMP) *For $p_1, p_2, L_1, L_2 \geq 0$, define $Li(Q, L_1, L_2, p_1, p_2)$ to be by those functions $f(\theta, x)$ such that:*

(i) *for all $\theta \in Q$,*

$$N_{p_1}(f(\theta, \cdot)) \leq L_1,$$

(ii) *for all θ_1, $\theta_2 \in Q$, all $x \in \mathbb{R}^k$,*

$$\|f(\theta_1, x) - f(\theta_2, x)\| \leq L_2 \, \|\theta_1 - \theta_2\| \, (1 + \|x\|^{p_2}).$$

Theorem A.1 (BMP's Thm. 5) *Given $p_1 \geq 0$, $p_2 \geq 0$, we assume that there exist positive constants K_1, K_2, q_1, q_2, $\rho < 1$ such that:*

(i) *for all $g \in Li(p_1)$, $\theta \in Q$, $n \geq$, z_1, z_2:*

$$\|P_\theta^n g(x_1) - P_\theta^n g(x_2)\| \leq K_1 \rho^n N_{p_1}(g) \, (1 + \|x_1\|^{q_1} + \|x_2\|^{q_2})$$

(ii) *for all $\theta \in Q$, $n \geq 0$, z and all $m \leq q_1 \vee q_2$*

$$\int P_\theta^n(x; dx_1) \, (1 + \|x_1\|^m) \leq K_2 \, (1 + \|x\|^m)$$

(iii) *for all $g \in Li(p_1)$, θ, $\theta' \in Q$, $n \geq 0$, x,*

$$\|P_\theta^n g(x) - P_{\theta'}^n g(x)\| \leq K_3 N_{p_1}(g) \, \|\theta - \theta'\| \, (1 + \|x\|^{q_2})$$

Then, for any function $f(\theta, x)$ of class $Li(Q, L_1, L_2, p_1, p_2)$, there exist functions $h(\theta)$, ν_θ and constants C_1, C_2, $C(\ell)$, $0 < \ell < 1$ depending only on the L_j, p_j, such that:

(j) *for all θ, $\theta' \in Q$, $\|h(\theta) - h(\theta')\| \leq C_1 \, \|\theta - \theta'\|$,*

(jj) *for all $\theta \in Q$, $\|\nu_\theta(x)\| \leq C_2 \, (1 + \|x\|^{q_1})$*

(jjj) *for all θ, $\theta' \in Q$, all $\ell \in (0, 1)$ and for all $s = \max(p_2, q_1, q_2)$*

$$\begin{aligned} \|\nu_\theta(x) - \nu_{\theta'}(x)\| &\leq C(\ell) \, \|\theta - \theta'\|^\ell \, (1 + \|x\|^s) \\ \|P_\theta \nu_\theta(x) - P_{\theta'} \nu_{\theta'}(x)\| &\leq C(\ell) \, \|\theta - \theta'\|^\ell \, (1 + \|x\|^s) \end{aligned}$$

(jv) $\quad (I - \pi_\theta)\nu_\theta = f_\theta - h(\theta).$

Proof: See [6, page 260] ∎

A.4 A Version of Theorem 3.10 for (D1)

The next theorem is an alternate version of Theorem 3.10 for (D1).

Theorem A.2 *If (D1) holds for V, then (D1) also holds for the function $= V^r$ where r is a positive real in the interval* $[0, 1]$.

Proof:

Suppose (D1) holds for the function $V : \mathsf{X} \to [1, \infty)$ with some $\lambda < 1$ and $L < \infty$, i.e.

$$P_\theta V \leq \lambda V + L, \qquad \text{for all } \theta \in \Theta.$$

Consider any rational $q = n/d$ and let $V^q = V^{n/d}$ for some $n \leq d$. We have from Jensen's inequality

$$\begin{aligned} P_\theta V^q &\leq (P_\theta V)^{n/d} \\ &\leq (\lambda V + L)^{n/d} \\ &\leq \lambda^{n/d} V^{n/d} + \frac{L}{\lambda^{(d-n)/d}} \qquad \text{(claim proven below)} \\ &= \lambda^{n/d} V^q + \frac{L}{\lambda^{(d-n)/d}}, \qquad \text{for all } \theta \in \Theta \end{aligned} \tag{A.1}$$

where the last inequality step follows from a claim we now prove.

The inequality

$$(\lambda V + L)^{n/d} \leq \lambda^{n/d} V^{n/d} + \frac{L}{\lambda^{(d-n)/d}}$$

is valid if and only if

$$(\lambda V + L)^n \leq \left(\lambda^{n/d} V^{n/d} + \frac{L}{\lambda^{(d-n)/d}} \right)^d. \tag{A.2}$$

Using the binomial expansion, the left hand side of (A.2) can be rewritten

$$(\lambda V + L)^n = \sum_{k=0}^{n} (\lambda V)^{n-k} L^k \binom{n}{k} \tag{A.3}$$

and the right hand side of A.2 can be rewritten

$$\begin{aligned} \left(\lambda^{n/d} V^{n/d} + \frac{L}{\lambda^{(d-n)/d}} \right)^d &= \sum_{k=0}^{d} \left(\lambda^{n/d} V^{n/d} \right)^{d-k} \left(\frac{L}{\lambda^{(d-n)/d}} \right)^k \binom{d}{k}, \qquad n \leq d. \\ &= \sum_{k=0}^{d} \lambda^{\frac{n(d-k)}{d} - \frac{k(d-n)}{d}} V^{\frac{n(d-k)}{d}} L^k \binom{d}{k}, \qquad n \leq d \\ &= \sum_{k=0}^{d} \lambda^{n-k} V^{\frac{n(d-k)}{d}} L^k \binom{d}{k}, \qquad n \leq d. \end{aligned} \tag{A.4}$$

We now compare the summands on the right hand sides of A.3 and A.4 for each $k = 0, 1, 2, \ldots, n$. When $k = 0$, we trivially find that the summands are equal.

Examining the case when $k = 1, \ldots, n$, we find

1. The exponent $n-k \leq \frac{n(d-k)}{d}$ for all positive integers $n \leq d$ and $k = 1, \ldots, n$. Since $V \geq 1$ we have
$$V^{n-k} \leq V^{\frac{n(d-k)}{d}}.$$

2. For $n \leq d$ and $k = 1, \ldots, n$,
$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \leq \frac{d!}{k!(d-k)!} = \binom{d}{k}$$

These two inequalities imply the individual summands of (A.3) and (A.4) obey the inequality:
$$(\lambda V)^{n-k} L^k \binom{n}{k} \leq \lambda^{n-k} V^{\frac{n(d-k)}{d}} L^k \binom{d}{k}, \qquad k = 0, 1, \ldots, n.$$

For the case $k = n+1, \ldots, d$, since the summands on the right hand side of (A.4) are all positive, the claim is now proven and (A.1) holds. Thus for any rational $q = n/d \in \mathcal{Q}$ we have
$$P_\theta V^q = \lambda^q V^q + \frac{L}{\lambda^{(1-q)}}, \qquad \text{for all } \theta \in \Theta. \tag{A.5}$$
Now let q_i be any sequence of rationals in the interval $(0,1)$ which converge to the real number $r \in (0,1)$, i.e. $r = \lim_{i\to\infty} q_i$. Then by the Dominated Convergence Theorem since $V^{q_i} \leq V$ for all $i = 1, 2, \ldots$ and $P_\theta V(x) < V(x) + L < \infty$ for any $\theta \in \Theta$ and $x \in \mathsf{X}$ we have the following
$$\begin{aligned} P_\theta(V^r) &= P_\theta(\lim_{i\to\infty} V^{q_i}) \\ &= \lim_{i\to\infty} P_\theta(V^{q_i}) \\ &\leq \lim_{i\to\infty} \left(\lambda^{q_i} V^{q_i} + \frac{L}{\lambda^{1-q_i}}\right) \\ &= \lambda^r V^r + \frac{L}{\lambda^{1-r}}, \qquad \theta \in \Theta, \end{aligned}$$
where the inequality follows from (A.5). Finally, the case for $r = 0$ and $r = 1$ follow trivially.

■

Corollary A.3 *If for some $\lambda < 1$ and $L < \infty$ we have*
$$P_\theta V \leq \lambda V + L, \qquad \theta \in \Theta$$
then for any real r in the interval $[0,1]$, we have
$$P_\theta^m V^r \leq C_D(r) V^r \qquad m = 1, 2, \ldots, \qquad \theta \in \Theta$$
where
$$\begin{aligned} C_D(r) &= 1 + \frac{L}{\lambda^{(1-r)}(1-\lambda^r)}, \\ &= 1 + \frac{L}{\lambda^{(1-r)} - \lambda}. \end{aligned} \tag{A.6}$$

A.5 A Proof of Theorem 4.5

Proof: The proof is identical for each $i = 1, \ldots, p$ so let us now fix such an i. Let θ be a fixed point in Θ and consider a small perturbation in the i^{th} component vector, denoted $\Delta\theta_i$ such that $\theta' = \theta + \Delta\theta_i \in \Theta$.

Part 1 (Set up.): Expand the difference

$$\begin{aligned}
J(\theta') - J(\theta) &= \mathbf{E}_{\pi_{\theta'}}[f_{\theta'}(X_1)] - \mathbf{E}_{\pi_\theta}[f_\theta(X_1)] \\
&= \mathbf{E}_{\pi_{\theta'}}[f_\theta(X_1)] - \mathbf{E}_{\pi_\theta}[f_\theta(X_1)] + \mathbf{E}_{\pi_\theta}[f_{\theta'}(X_1) - f_\theta(X_1)] \\
&+ \Big\{\mathbf{E}_{\pi_{\theta'}}[f_{\theta'}(X_1) - f_\theta(X_1)] - \mathbf{E}_{\pi_\theta}[f_{\theta'}(X_1) - f_\theta(X_1)]\Big\}
\end{aligned}$$

so that

$$\begin{aligned}
\frac{\partial J(\theta)}{\partial\theta_i} &= \lim_{\Delta\theta_i \to 0} \frac{J(\theta') - J(\theta)}{\Delta\theta_i} \\
&= \lim_{\Delta\theta_i \to 0} \sum_{x \in \mathsf{X}} \frac{\pi_{\theta'}(x) - \pi_\theta(x)}{\Delta\theta_i} f_\theta(x) && \text{(A.7)} \\
&+ \lim_{\Delta\theta_i \to 0} \sum_{x \in \mathsf{X}} \pi_\theta(x) \frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i} && \text{(A.8)} \\
&+ \lim_{\Delta\theta_i \to 0} \sum_{x \in \mathsf{X}} (\pi_{\theta'}(x) - \pi_\theta(x)) \frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i} && \text{(A.9)}
\end{aligned}$$

We shall next consider the three limits (A.7) - (A.9) separately starting with the first.

Part 2-a (First limit, setup): We have the matrix equation

$$P_{\theta'} = P_\theta + \Delta\theta_i \; Q_{\theta,i} + R(\theta, \Delta\theta_i)$$

hence

$$\begin{aligned}
\pi_\theta P_{\theta'} &= \pi_\theta P_\theta + \Delta\theta_i \; \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i) \\
&= \pi_\theta + \Delta\theta_i \; \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i).
\end{aligned}$$

Inserting canceling terms on the left we have

$$-\pi_{\theta'} P_{\theta'} + \pi_\theta P_{\theta'} + \pi_{\theta'} = \pi_\theta + \Delta\theta_i \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i)$$

so that

$$(\pi_{\theta'} - \pi_\theta)(I - P_{\theta'}) = \Delta\theta_i \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i).$$

The *group inverse* $P_{\theta'}^{\#}$ exists by Lemma 4.3 under the assumption of strong ergodicity for all Θ, along with the series expansion

$$\begin{aligned}
P_{\theta'}^{\#} &\doteq -e\pi_{\theta'} + (I - P_{\theta'} + e\pi_{\theta'})^{-1} \\
&= -e\pi_{\theta'} + \sum_{k=0}^{\infty} (P_{\theta'} - e\pi_{\theta'})^k
\end{aligned}$$

Thus by (4.18),

$$(\pi_{\theta'} - \pi_\theta)(I - P_{\theta'})P_{\theta'}^{\#} = \Delta\theta_i \pi_\theta Q_{\theta,i} P_{\theta'}^{\#} + \pi_\theta R(\theta, \Delta\theta_i)P_{\theta'}^{\#}$$

which becomes

$$(\pi_{\theta'} - \pi_\theta)(I - e\pi_{\theta'}) = \Delta\theta_i \pi_\theta Q_{\theta,i} P_{\theta'}^{\#} + \pi_\theta R(\theta, \Delta\theta_i)P_{\theta'}^{\#}$$

or since $(\pi_{\theta'} - \pi_\theta)e = 0$

$$\left(\frac{1}{\Delta\theta_i}\right)(\pi_{\theta'} - \pi_\theta) = \pi_\theta Q_{\theta,i} P_{\theta'}^{\#} + \left(\frac{1}{\Delta\theta_i}\right)\pi_\theta R(\theta, \Delta\theta_i)P_{\theta'}^{\#}, \qquad \text{for small } |\Delta\theta_i| > 0.$$

Multiply on the right by the performance function column vector $f_\theta \doteq [f_\theta(x)\ ]_{x\in\mathsf{X}}$ for which we assume in (F1) that $f_\theta \in L^\infty_{V^r} = \{h : \sup_{x\in\mathsf{X}} \frac{|h(x)|}{V^r(x)} < \infty\}$:

$$\left(\frac{1}{\Delta\theta_i}\right)(\pi_{\theta'} - \pi_\theta)f_\theta = \pi_\theta Q_{\theta,i} P_{\theta'}^{\#} f_\theta + \left(\frac{1}{\Delta\theta_i}\right)\pi_\theta R(\theta, \Delta\theta_i)P_{\theta'}^{\#} f_\theta \tag{A.10}$$

We next consider separately the two terms on the right hand side of (A.10) as $\Delta\theta_i \to 0$ starting with the second term.

Part 2-b. (First Limit, second term.): Define $R'(\theta, \Delta\theta_i) \doteq \left(\frac{1}{\Delta\theta_i}\right) R(\theta, \Delta\theta_i)$ and the second term is

$$\pi_\theta R'(\theta, \Delta\theta_i)P_{\theta'}^{\#} f_\theta. \tag{A.11}$$

From Lemma 4.4 we find that under our conditions, the fundamental kernel is a mapping from $L^\infty_{V^r}$ to $L^\infty_{V^r}$ and by (2.15), $\pi_{\theta'}(V^r) < \infty$ so that $P_{\theta'}^{\#}$ is also a mapping from $L^\infty_{V^r}$ to $L^\infty_{V^r}$. We will show that this second term is zero in the limit as $\Delta\theta_i \to 0$, but, instead of (A.11) we can consider

$$\lim_{\Delta\theta_i \to 0} \pi_\theta R'(\theta, \Delta\theta_i) f = \lim_{\Delta\theta_i \to 0} \sum_x \pi_\theta(x) \sum_y \frac{r_{x,y}(\theta, \Delta\theta_i)}{\Delta\theta_i} f(y)$$

for **arbitrary** $f \in L^\infty_{V^r}$.

Note that if $p_{x,y}(\theta) = 0$ for any points x, y, θ, then by (4.10) $r_{x,y}(\theta, \Delta\theta_i)$ necessarily must be zero also in a small δ-neighborhood of $\Delta\theta_i = 0$. If we then take the convention that any fraction of the form $\frac{0}{0}$ is *defined* to be zero, then we can write the double sum as

$$\begin{aligned}
\pi_\theta R'(\theta, \Delta\theta_i) f &= \sum_x \pi_\theta(x) \sum_y \frac{r_{x,y}(\theta, \Delta\theta_i)}{\Delta\theta_i} f(y) \\
&= \sum_x \pi_\theta(x) \sum_y p_{x,y}(\theta) \frac{r_{x,y}(\theta, \Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i} f(y) \\
&= \sum_{x,y} \pi_\theta(x) p_{x,y}(\theta) \frac{r_{x,y}(\theta, \Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i} f(y).
\end{aligned}$$

To apply the Dominated Convergence Theorem, we find for all $|\Delta\theta_i| < \delta$

$$\sum_{x,y} \pi_\theta(x) p_{x,y}(\theta) \frac{r_{x,y}(\theta, \Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i} f(y) \le \sum_{x,y} \pi_\theta(x) p_{x,y}(\theta) K_3'' C V^r(y) < \infty, \tag{A.12}$$

where the first inequality follows from (4.10) and the fact that $f \in L^\infty_{V^r}$ (which defines some constant $C < \infty$) while the finiteness follows from (2.15). Thus,

$$\begin{aligned}
\lim_{\Delta\theta_i \to 0} \pi_\theta R'(\theta, \Delta\theta_i) f &= \lim_{\Delta\theta_i \to 0} \sum_{x,y} \pi_\theta(x) p_{x,y}(\theta) \frac{r_{x,y}(\theta, \Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i} f(y) \\
&= \sum_{x,y} \pi_\theta(x) p_{x,y}(\theta) \lim_{\Delta\theta_i \to 0} \frac{r_{x,y}(\theta, \Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i} f(y) \\
&= \sum_{x,y} \pi_\theta(x) p_{x,y}(\theta)\ 0 \\
&= 0
\end{aligned}$$

for **arbitrary** $f \in L^\infty_{V^r}$. Thus, the second term in (A.10) converges to zero.

Part 2-c. (First limit, first term) : In this part, we will be appealing to Lemma 4.4 so before we begin let us verify the fourth condition of the Lemma 4.4; that $P_\theta \to P_{\theta_0}$ as $\theta \to \theta_0$ in the induced operator norm $|||\cdot|||_{V^r}$, i.e.

$$\lim_{\theta \to \theta_0} |||P_\theta - P_{\theta_0}|||_{V^r} = \lim_{\theta \to \theta_0} \sup_{\substack{h \in L^\infty_{V^r} \\ |h|_{V^r}=1}} |(P_\theta - P_{\theta_0})h|_{V^r} = 0. \tag{A.13}$$

We have for some $h \in L^\infty_{V^r}$ such that $|h|_{V^r} \doteq \sup_{x \in \mathsf{X}} \frac{|h(x)|}{V^r(x)} = 1$

$$\begin{aligned}
|(P_\theta - P_{\theta_0})h| &= \left| \sum_y \left(p_{x,y}(\theta) - p_{x,y}(\theta_0)\right) h(y) \right| \\
&\leq \|\theta - \theta_0\| \sum_y \left(K_3 p_{x,y}(\theta_0)\right) V^r(y) \\
&\leq \|\theta - \theta_0\| \, C V^r(x)
\end{aligned}$$

for some constant $C < \infty$; the first inequality following from (G2) while the second inequality follows from (D0), and (D0) is implied by (D2). Thus,

$$|(P_\theta - P_{\theta_0})h|_V \leq \frac{\|\theta - \theta_0\| \, CV(x)}{V(x)} = C \, \|\theta - \theta_0\|$$

and clearly converges to zero in the limit as $\theta \to \theta_0$, hence we have shown (A.13).

We will soon consider the limit of the first term in (A.10):

$$\lim_{\theta' \to \theta} \pi_\theta Q_{\theta,i} P^\#_{\theta'} f_\theta$$

but before we begin, we note that the group inverse is given by

$$\begin{aligned}
P^\#_{\theta'} &\doteq \left(I - P_{\theta'} + e\pi_{\theta'}\right)^{-1} - e\pi_{\theta'} \\
&= \sum_{k=0}^{\infty} (P_{\theta'} - e\pi_{\theta'})^k - e\pi_{\theta'},
\end{aligned}$$

and is the difference of an *fundamental matrix* term and an *invariant matrix* term. From Lemma 4.4, if we let $f \in L^\infty_{V^r}$, under the conditions we have established, $P^\#_{\theta'} f$ converges to $P^\#_\theta f$ in the $L^\infty_{V^r}$ norm as $\theta' \to \theta$, i.e.

$$\lim_{\theta' \to \theta} \sup_{x \in \mathsf{X}} \frac{\left| P^\#_{\theta'}(x,f) - P^\#_\theta(x,f) \right|}{V^r(x)} = 0.$$

Hence, for every $\epsilon > 0$ there exists a $\delta > 0$ neighborhood sufficiently small, such that

$$0 \le \sup_{x \in \mathsf{X}} \frac{\left| P^\#_{\theta'}(x,f) - P^\#_\theta(x,f) \right|}{V^r(x)} \le \epsilon$$

for all θ' in this δ-neighborhood. Hence we have,

$$\left| P^\#_{\theta'}(x,f) - P^\#_\theta(x,f) \right| \le \epsilon V^r(x), \qquad \text{for all } x \in \mathsf{X}, \quad |\Delta\theta_i| \le \delta.$$

Thus, for all θ' in this neighborhood of θ,

$$\begin{aligned} \left| P^\#_{\theta'}(x,f) \right| &\le \left| P^\#_\theta(x,f) \right| + \epsilon V^r(x) \\ &\le CV^r(x), \qquad \text{for all } x \in \mathsf{X} \end{aligned} \tag{A.14}$$

for some constant $C < \infty$.

We now consider the limit of the first term in (A.10), namely:

$$\lim_{\substack{\Delta\theta_i \to 0 \\ \theta' = \theta + \Delta\theta_i}} \pi_\theta Q_{\theta,i} P^\#_{\theta'} f_\theta$$

Without loss of generality, instead of $P^\#_{\theta'} f_\theta$ we can consider an arbitrary family of functions $g_{\Delta\theta_i} : \mathsf{X} \to \mathbb{R}$ such that $|g_{\Delta\theta_i}| \le CV^r$ uniformly for all $\Delta\theta_i$ in the δ-neighborhood of zero. We have shown in (A.14) that this family of functions will include $P^\#_{\theta'} f_\theta$. Thus, we consider

$$\begin{aligned} \lim_{\substack{\Delta\theta_i \to 0 \\ \theta' = \theta + \Delta\theta_i}} \pi_\theta Q_{\theta,i} g_{\Delta\theta_i} &= \lim_{\substack{\Delta\theta_i \to 0 \\ \theta' = \theta + \Delta\theta_i}} \sum_{x \in \mathsf{X}} \pi_\theta(x) \sum_y \frac{\partial p_{x,y}(\theta)}{\partial \theta_i} g_{\Delta\theta_i}(y) \\ &= \lim_{\substack{\Delta\theta_i \to 0 \\ \theta' = \theta + \Delta\theta_i}} \sum_{x \in \mathsf{X}} \pi_\theta(x) \sum_y p_{x,y}(\theta) \frac{\frac{\partial p_{x,y}(\theta)}{\partial \theta_i}}{p_{x,y}(\theta)} g_{\Delta\theta_i}(y) \end{aligned}$$

where the last line holds because if $p_{x,y}(\theta) = 0$, then the partial derivative in the numerator is necessarily zero under (4.7) and this ratio is thus zero.

Again, to apply the Dominated Convergence theorem, we check for all $|\Delta\theta_i| < \delta$

$$\sum_{x,y \in \mathsf{X}} \pi_\theta(x) p_{x,y}(\theta) \frac{\frac{\partial p_{x,y}(\theta)}{\partial \theta_i}}{p_{x,y}(\theta)} g_{\Delta\theta_i}(y) \le \sum_{x,y \in \mathsf{X}} \pi_\theta(x) p_{x,y}(\theta) K_3 C V^r(y) < \infty \tag{A.15}$$

by (4.8) and (2.15). Thus,

$$\begin{aligned} &\lim_{\substack{\Delta\theta_i \to 0 \\ \theta' = \theta + \Delta\theta_i}} \sum_{x,y \in \mathsf{X}} \pi_\theta(x) p_{x,y}(\theta) \frac{\frac{\partial p_{x,y}(\theta)}{\partial \theta_i}}{p_{x,y}(\theta)} g_{\Delta\theta_i}(y) \\ &= \sum_{x,y \in \mathsf{X}} \pi_\theta(x) p_{x,y}(\theta) \frac{\frac{\partial p_{x,y}(\theta)}{\partial \theta_i}}{p_{x,y}(\theta)} \lim_{\substack{\Delta\theta_i \to 0 \\ \theta' = \theta + \Delta\theta_i}} g_{\Delta\theta_i}(y) \end{aligned}$$

for arbitrary $g_{\Delta\theta_i} \in L^{\infty}_{V^r}$.

Hence for our problem, we can take

$$g_{\Delta\theta_i}(y) = P^{\#}_{\theta+\Delta\theta_i}(y, f_\theta)$$

which is an element of $L^{\infty}_{V^r}$ for all $|\Delta\theta_i| \leq \delta$ and we thus have

$$\begin{aligned}
&\lim_{\substack{\Delta\theta_i \to 0 \\ \theta'=\theta+\Delta\theta_i}} \sum_x \pi_\theta(x) \sum_y \frac{\partial p_{x,y}(\theta)}{\partial \theta_i} P^{\#}_{\theta+\Delta\theta_i}(y, f_\theta) \\
&= \sum_x \pi_\theta(x) \sum_y \frac{\partial p_{x,y}(\theta)}{\partial \theta_i} \lim_{\substack{\Delta\theta_i \to 0 \\ \theta'=\theta+\Delta\theta_i}} P^{\#}_{\theta+\Delta\theta_i}(y, f_\theta)
\end{aligned}$$

We know from the Lemma 4.4 that for $f_\theta \in L^{\infty}_{V^r}$

$$\lim_{\theta' \to \theta} \sup_{x \in \mathsf{X}} \frac{\left|P^{\#}_{\theta'}(x, f_\theta) - P^{\#}_{\theta}(x, f_\theta)\right|}{V^r(x)} = 0$$

hence since $\mathsf{X} = \{x \in \mathsf{X} : V^r(x) < \infty\}$ we have

$$\lim_{\substack{\Delta\theta_i \to 0 \\ \theta'=\theta+\Delta\theta_i}} \pi_\theta Q_{\theta,i} P^{\#}_{\theta'} f_\theta = \pi_\theta Q_{\theta,i} P^{\#}_{\theta} f_\theta$$

Therefore, the limit in (A.10) yields:

$$\lim_{\Delta\theta_i \to 0} \left(\frac{1}{\Delta\theta_i}\right) (\pi_{\theta'} - \pi_\theta) f_\theta = \pi_\theta Q_{\theta,i} P^{\#}_{\theta} f_\theta$$

which concludes the first limit in (A.7).

Part 3 (Second limit): Condition (F3) implies that there exists a $\delta > 0$ such that for all $x \in \mathsf{X}$

$$\left|\frac{f_{\theta+\Delta\theta_i}(x) - f_\theta(x)}{\Delta\theta_i}\right| \leq C' V^r(x), \quad \text{for all } \Delta\theta_i \in (0, \delta), \tag{A.16}$$

for some $C' < \infty$. Since we have

$$\sum_{x \in \mathsf{X}} \pi_\theta(x) C' V^r(x) < \infty,$$

the second limit follows from the Dominated Convergence Theorem, i.e.

$$\begin{aligned}
\lim_{\Delta\theta_i \to 0} \sum_{x \in \mathsf{X}} \pi_\theta(x) \frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i} &= \sum_{x \in \mathsf{X}} \pi_\theta(x) \lim_{\Delta\theta_i \to 0} \frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i} \\
&= \sum_{x \in \mathsf{X}} \pi_\theta(x) \frac{\partial f_\theta(x)}{\partial \theta_i}.
\end{aligned}$$

Part 4 (Third limit): The convergence of the third limit,

$$\lim_{\Delta\theta_i \to 0} \sum_{x \in \mathsf{X}} \left(\pi_{\theta'}(x) - \pi_\theta(x)\right) \frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i} = 0 \tag{A.17}$$

follows via the same bound (A.16) coupled with (4.15) of Lemma 4.4.

Part 5 (Remaining equalities): The equality of

$$\pi_\theta Q_{\theta,i} P_\theta^{\#} f_\theta = \pi_\theta Q_{\theta,i} g_\theta, \qquad \theta \in \Theta,$$

follows since the *performance potential* $g_x(\theta)$ given by

$$g_x(\theta) \doteq \lim_{n\to\infty} \left\{ \mathbf{E}_{\theta,x} \left[\sum_{k=0}^{n} f_\theta(X_k^{\{x\}}) \right] - nJ(\theta) \right\}, \qquad x \in \mathsf{X}. \tag{A.18}$$

$$= \lim_{n\to\infty} \sum_{k=0}^{n} \mathbf{E}_{\theta,x} \left[f_\theta(X_k^{\{x\}}) - J(\theta) \right] \tag{A.19}$$

is nothing more than a solution to the Poisson equation with forcing function f_θ.

In prior chapters, we have denoted any Poisson equation solution as ν_θ. Here, we are assuming the fundamental matrix exists and is the inverse of $(I - P_\theta + e\pi_\theta)$. It is easily verified that a solution of the Poisson equation takes the form

$$\nu_\theta = Z_\theta f_\theta = (I - P_\theta + e\pi_\theta)^{-1} f_\theta$$

Additionally, the series form (A.19) is known to solve the Poisson equation and to exist (converge) by Theorem 17.4.2 in [79] under irreducibility, (D2), and (2.15). Furthermore, the solution in the form g_θ differs by at most a finite constant [79, Proposition 17.4.1] from the solution $P_\theta^{\#} f_\theta$, i.e.

$$\begin{aligned} P_\theta^{\#} f_\theta &= Z_\theta f_\theta - e\pi_\theta f_\theta \\ &= g_\theta + ec \end{aligned}$$

for some constant $|c| < \infty$. Therefore

$$\begin{aligned} \pi_\theta Q_{\theta,i} P_\theta^{\#} f_\theta &= \pi_\theta Q_{\theta,i} (g_\theta + ec) \\ &= \pi_\theta Q_{\theta,i} g_\theta \end{aligned}$$

since $Q_{\theta,i} e = 0$ by assumption. Similarly,

$$\begin{aligned} \pi_\theta Q_{\theta,i} P_\theta^{\#} f_\theta &= \pi_\theta Q_{\theta,i} (Z_\theta - e\pi_\theta) f_\theta \\ &= \pi_\theta Q_{\theta,i} (Z_\theta f_\theta - eJ(\theta)) \\ &= \pi_\theta Q_{\theta,i} \nu_\theta. \end{aligned}$$

or this holds for any of the family of solutions to the Poisson equation which differ by an additive constant. ■

A.6 A Proof of Theorem 4.6

This is an alternate version of last sections result which does not explicitly assume existence of an invertible fundamental matrix.

Proof: Let θ be a fixed point in Θ and consider a small perturbation in the i^{th} component vector, denoted $\Delta\theta_i$. The perturbation is assumed small enough so $\theta' = \theta + \Delta\theta_i \in \Theta$. The proof is identical for each $i = 1, \ldots, p$ so let us now fix such an i.

Part 1 (Set up.): Let us expand the difference

$$\begin{aligned} J(\theta') - J(\theta) &= \mathbf{E}_{\pi_{\theta'}}[f_{\theta'}(X_1)] - \mathbf{E}_{\pi_\theta}[f_\theta(X_1)] \\ &= \mathbf{E}_{\pi_{\theta'}}[f_{\theta'}(X_1)] - \mathbf{E}_{\pi_\theta}[f_{\theta'}(X_1)] + \mathbf{E}_{\pi_\theta}[f_{\theta'}(X_1) - f_\theta(X_1)] \end{aligned}$$

so that

$$\begin{aligned} \frac{\partial J(\theta)}{\partial \theta_i} &= \lim_{\Delta\theta_i \to 0} \frac{J(\theta') - J(\theta)}{\Delta\theta_i} \\ &= \lim_{\Delta\theta_i \to 0} \sum_{x \in \mathsf{X}} \frac{\pi_{\theta'}(x) - \pi_\theta(x)}{\Delta\theta_i} f_{\theta'}(x) \qquad \text{(A.20)} \\ &+ \lim_{\Delta\theta_i \to 0} \sum_{x \in \mathsf{X}} \pi_\theta(x) \frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i} \qquad \text{(A.21)} \end{aligned}$$

We shall next consider the two limits (A.20) - (A.21) separately starting with the first.

Part 2-a (First limit, setup): We have the matrix equation

$$P_{\theta'} = P_\theta + \Delta\theta_i \; Q_{\theta,i} + R(\theta, \Delta\theta_i)$$

hence

$$\begin{aligned} \pi_\theta P_{\theta'} &= \pi_\theta P_\theta + \Delta\theta_i \; \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i) \\ &= \pi_\theta + \Delta\theta_i \; \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i). \end{aligned}$$

Inserting canceling terms on the left we have

$$-\pi_{\theta'} P_{\theta'} + \pi_\theta P_{\theta'} + \pi_{\theta'} = \pi_\theta + \Delta\theta_i \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i)$$

so that

$$(\pi_{\theta'} - \pi_\theta)(I - P_{\theta'}) = \Delta\theta_i \pi_\theta Q_{\theta,i} + \pi_\theta R(\theta, \Delta\theta_i).$$

Multiplying on the right by the Poisson equation solution $\nu_{\theta'}$ we find,

$$(\pi_{\theta'} - \pi_\theta)(I - P_{\theta'})\nu_{\theta'} = \Delta\theta_i \pi_\theta Q_{\theta,i} \nu_{\theta'} + \pi_\theta R(\theta, \Delta\theta_i)\nu_{\theta'}$$

which becomes

$$(\pi_{\theta'} - \pi_\theta)(I - e\pi_{\theta'}) f_{\theta'} = \Delta\theta_i \pi_\theta Q_{\theta,i} \nu_{\theta'} + \pi_\theta R(\theta, \Delta\theta_i)\nu_{\theta'}$$

or since $(\pi_{\theta'} - \pi_\theta)e = 0$

$$\left(\frac{1}{\Delta\theta_i}\right)(\pi_{\theta'} - \pi_\theta)f_{\theta'} = \pi_\theta Q_{\theta,i}\nu_{\theta'} + \left(\frac{1}{\Delta\theta_i}\right)\pi_\theta R(\theta,\Delta\theta_i)\nu_{\theta'}, \qquad \text{for } |\Delta\theta_i| > 0. \tag{A.22}$$

We next consider separately the two terms on the right hand side of (A.22) as $\Delta\theta_i \to 0$ starting with the second term.

Part 2-b. (First Limit, second term.):

Define $R'(\theta,\Delta\theta_i) \doteq \left(\frac{1}{\Delta\theta_i}\right) R(\theta,\Delta\theta_i)$ and the second term is

$$\pi_\theta R'(\theta,\Delta\theta_i)\nu_{\theta'}. \tag{A.23}$$

Note again that if $p_{x,y}(\theta) = 0$, then $r_{x,y}(\theta,\Delta\theta_i)$ necessarily must be zero also in a small δ-neighborhood of $\Delta\theta_i = 0$. If we again take the convention that any fraction of the form $\frac{0}{0}$ is *defined* to be zero, then we can write the double sum as

$$\begin{aligned}\pi_\theta R'(\theta,\Delta\theta_i)\nu_{\theta'} &= \sum_x \pi_\theta(x) \sum_y \frac{r_{x,y}(\theta,\Delta\theta_i)}{\Delta\theta_i}\nu_{\theta'}(y) \\ &= \sum_x \pi_\theta(x) \sum_y p_{x,y}(\theta)\frac{r_{x,y}(\theta,\Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i}\nu_{\theta'}(y), \qquad |\Delta\theta| < \delta.\end{aligned}$$

To apply the Dominated Convergence Theorem, we find for all $|\Delta\theta_i| < \delta$

$$\sum_{x,y} \pi_\theta(x)p_{x,y}(\theta)\frac{r_{x,y}(\theta,\Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i}\nu_{\theta'}(y) \le \sum_{x,y} \pi_\theta(x)p_{x,y}(\theta)K_3''CV^r(y) < \infty, \qquad |\Delta\theta_i| < \delta \tag{A.24}$$

where the first inequality follows from (4.10) and the hypothesis while the second follows from (2.15). Thus

$$\begin{aligned}\lim_{\Delta\theta_i\to 0} \pi_\theta R'(\theta,\Delta\theta_i)\nu_{\theta'} &= \lim_{\Delta\theta_i\to 0} \sum_{x,y} \pi_\theta(x)p_{x,y}(\theta)\frac{r_{x,y}(\theta,\Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i}\nu_{\theta'}(y) \\ &= \sum_{x,y} \pi_\theta(x)p_{x,y}(\theta) \lim_{\Delta\theta_i\to 0} \frac{r_{x,y}(\theta,\Delta\theta_i)}{p_{x,y}(\theta)\Delta\theta_i}\nu_{\theta'}(y) \\ &= \sum_{x,y} \pi_\theta(x)p_{x,y}(\theta)\ 0 \\ &= 0.\end{aligned}$$

Thus, the second term in (A.22) converges to zero.

Part 2-c. (First limit, first term) :

We are assuming that for some constant $C < \infty$,

$$\sup_{\theta\in\Theta} |\nu_{\theta'}(x)| \le CV^r(x), \qquad \text{for all } x \in \mathsf{X}. \tag{A.25}$$

We now consider the limit of the first term in (A.22), namely:

$$\begin{aligned}\lim_{\substack{\Delta\theta_i\to 0\\ \theta'=\theta+\Delta\theta_i}} \pi_\theta Q_{\theta,i}\nu_{\theta'} &= \lim_{\substack{\Delta\theta_i\to 0\\ \theta'=\theta+\Delta\theta_i}} \sum_{x\in\mathsf{X}} \pi_\theta(x) \sum_y \frac{\partial p_{x,y}(\theta)}{\partial\theta_i}\nu_{\theta'}(y) \\ &= \lim_{\substack{\Delta\theta_i\to 0\\ \theta'=\theta+\Delta\theta_i}} \sum_{x\in\mathsf{X}} \pi_\theta(x) \sum_y p_{x,y}(\theta)\frac{\frac{\partial p_{x,y}(\theta)}{\partial\theta_i}}{p_{x,y}(\theta)}\nu_{\theta'}(y)\end{aligned}$$

where the last line follow because if $p_{x,y}(\theta) = 0$, then the partial derivative in the numerator is necessarily zero under (4.7) and this ratio is thus zero.

Again, to apply Dominated Convergence, we check

$$\sum_{x,y\in\mathsf{X}} \pi_\theta(x)p_{x,y}(\theta)\frac{\frac{\partial p_{x,y}(\theta)}{\partial\theta_i}}{p_{x,y}(\theta)}\nu_{\theta'}(y) \le \sum_{x,y\in\mathsf{X}} \pi_\theta(x)p_{x,y}(\theta)K_3CV^r(y) < \infty \tag{A.26}$$

for all $|\Delta\theta_i| < \delta$ by (4.8) and (2.15). Thus,

$$\begin{aligned}
\lim_{\substack{\Delta\theta_i\to 0\\ \theta'=\theta+\Delta\theta_i}} \sum_{x,y\in\mathsf{X}} \pi_\theta(x)p_{x,y}(\theta)\frac{\frac{\partial p_{x,y}(\theta)}{\partial\theta_i}}{p_{x,y}(\theta)}\nu_{\theta'}(y) &= \sum_{x,y\in\mathsf{X}} \pi_\theta(x)p_{x,y}(\theta)\frac{\frac{\partial p_{x,y}(\theta)}{\partial\theta_i}}{p_{x,y}(\theta)} \lim_{\substack{\Delta\theta_i\to 0\\ \theta'=\theta+\Delta\theta_i}} \nu_{\theta'}(y)\\
&= \sum_{x,y\in\mathsf{X}} \pi_\theta(y)\frac{\partial p_{x,y}(\theta)}{\partial\theta_i}\nu_\theta(y)
\end{aligned}$$

the last step following from the assumed continuity of ν_θ.

Therefore, the limit in (A.22) yields:

$$\lim_{\Delta\theta_i\to 0}\left(\frac{1}{\Delta\theta_i}\right)(\pi_{\theta'} - \pi_\theta)f_{\theta'} = \pi_\theta Q_{\theta,i}\nu_\theta$$

which concludes the first limit in (A.7).

Part 3 (Second limit): Condition (F3) implies that there exists a $\delta > 0$ and some $C' < \infty$ such that for all

$$\left|\frac{f_{\theta+\Delta\theta_i}(x) - f_\theta(x)}{\Delta\theta_i}\right| \le C'V^r(x), \qquad \Delta\theta_i \in (0,\delta),\ x\in\mathsf{X}.$$

Since we assume

$$\sum_{x\in\mathsf{X}} \pi_\theta(x)C'V^r(x) < \infty,$$

the result follows from the Dominated Convergence Theorem, i.e.

$$\begin{aligned}
\lim_{\Delta\theta_i\to 0}\sum_{x\in\mathsf{X}} \pi_\theta(x)\frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i} &= \sum_{x\in\mathsf{X}} \pi_\theta(x)\lim_{\Delta\theta_i\to 0}\frac{f_{\theta'}(x) - f_\theta(x)}{\Delta\theta_i}\\
&= \sum_{x\in\mathsf{X}} \pi_\theta(x)\frac{\partial f_\theta(x)}{\partial\theta_i}.
\end{aligned}$$

Part 4 (Last equality):

The equality of

$$\pi_\theta Q_{\theta,i}\nu_\theta = \pi_\theta Q_{\theta,i}g_\theta, \qquad \theta\in\Theta,$$

follows since the *performance potential* $g_x(\theta)$ given by

$$g_x(\theta) \doteq \lim_{n\to\infty}\left\{\mathbf{E}_{\theta,x}\left[\sum_{k=0}^n f_\theta(X_k^{\{x\}})\right] - nJ(\theta)\right\}, \qquad x\in\mathsf{X}. \tag{A.27}$$

$$= \lim_{n\to\infty}\sum_{k=0}^n \mathbf{E}_{\theta,x}\left[f_\theta(X_k^{\{x\}}) - J(\theta)\right] \tag{A.28}$$

is nothing more than a solution to the Poisson equation with forcing function f_θ.

Specifically, the series form (A.28) is known to solve the Poisson equation and to exist (converge) by Theorem 17.4.2 in [79] under irreducibility, (D2), and (2.15). Furthermore, the solution in the form g_θ differs by at most a finite constant [79, Proposition 17.4.1] from the solution ν_θ, i.e.

$$\nu_\theta \quad = \quad g_\theta + ec$$

for some constant $|c| < \infty$. Therefore

$$\begin{aligned} \pi_\theta Q_{\theta,i}\nu_\theta &= \pi_\theta Q_{\theta,i}\left(g_\theta + ec\right) \\ &= \pi_\theta Q_{\theta,i} g_\theta \end{aligned}$$

since $Q_{\theta,i}e = 0$ by assumption. ■

A.7 A Proof of Lemma 5.8

Proof: Fix a δ such that $\delta \leq \min\{\delta_C, \delta_H\}$ and $0 < \delta \leq 1$.

Case 1) θ and θ' are chosen in Θ so that $\|\theta - \theta'\| \geq \delta$:

$$\begin{aligned} &\|h_\ell(\theta) - h_\ell(\theta')\| \\ &\quad = \quad \|\bar{\pi}_\theta H_{\ell,\theta} - \bar{\pi}_{\theta'} H_{\ell,\theta'}\| \\ &\quad \leq \quad \|\bar{\pi}_\theta H_{\ell,\theta}\| + \|\bar{\pi}_{\ell,\theta'} H_{\ell,\theta'}\| \\ &\quad \leq \quad 2\sup_{\theta\in\Theta} \|\bar{\pi}_\theta H_{\ell,\theta}\| \\ &\quad \leq \quad 2\sup_{\theta\in\Theta} \left\| \bar{\pi}_\theta (\sum_{i=1}^{\ell} V^r(\cdot_i)) \right\| \\ &\quad \leq \quad \ell \pi(V^r) \end{aligned}$$

Since $\pi(V^r) < \infty$ by (2.15 there exists a $K < \infty$ such that

$$\begin{aligned} \|h_\ell(\theta) - h_\ell(\theta')\| &\leq \ell K \\ &\leq \ell \frac{K}{\delta^{\widehat{\ell_2}}} \|\theta - \theta'\|^{\widehat{\ell_2}} . \end{aligned}$$

Case 2) θ and θ' in Θ are chosen so that $\|\theta - \theta'\| < \delta \leq 1$. Under our assumptions, for any $n = 1, 2, \ldots$ and any $\bar{x} \in \mathsf{X}$, we have

$$\begin{aligned} &\|h_\ell(\theta) - h_\ell(\theta')\| \\ &\quad \leq \quad \left\|\bar{\pi}_\theta H_{\ell,\theta} - \bar{P}^n_{\ell,\theta} H_{\ell,\theta}(\bar{x})\right\| + \left\|\bar{P}^n_{\ell,\theta} H_{\ell,\theta}(\bar{x}) - \bar{P}^n_{\ell,\theta'} H_{\ell,\theta}(\bar{x})\right\| \\ &\quad\quad + \left\|\bar{P}^n_{\ell,\theta'} H_{\ell,\theta}(\bar{x}) - \bar{P}^n_{\ell,\theta'} H_{\ell,\theta'}(\bar{x})\right\| + \left\|\bar{P}^n_{\ell,\theta'} H_{\ell,\theta'}(\bar{x}) - \bar{\pi}_{\theta'} H_{\ell,\theta'}\right\| \\ &\quad \leq \quad \ell C_H C_E \bar{V}^r_\ell(\bar{x})\rho^n + C_H C_C \bar{V}^r_\ell(\bar{x}) n \ell^2 \|\theta - \theta'\|^{\widehat{\ell_3}} \\ &\quad\quad + \ell C_5 \|\theta - \theta'\|^{\widehat{\ell_2}} \bar{P}^n_{\ell,\theta'} \bar{V}^r_\ell(\bar{x}) + \ell C_H C_E \bar{V}^r_\ell(\bar{x})\rho^n \\ &\quad \leq \quad \ell^2 \bar{V}^r_\ell(\bar{x}) \left\{ 2C_H C_E\ \rho^n + C_H C_C n \|\theta - \theta'\|^{\widehat{\ell_3}} + C^r_D C_5 \|\theta - \theta'\|^{\widehat{\ell_2}} \right\}, \qquad \bar{x} \in \mathsf{X}^\ell. \end{aligned}$$

In the second inequality above we have applied ($\overline{\text{C}}$), ($\overline{\text{E1}}$), and ($\overline{\text{H5}}$) while the last inequality we have again applied (5.14) under (D1) with Jensen's inequality.

This last inequality is true for all $n = 1, 2, \ldots$, hence we may choose an integer $n \doteq \left\lceil \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} \right\rceil = \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + u$ where the remainder u is such that $0 \le u < 1$ and $\widehat{\ell}_3$ is from ($\overline{\text{C}}$). If we let $\Delta\theta \doteq \theta - \theta'$, the bracketed term becomes

$$\begin{aligned}
&\left\{2C_H C_E \rho^n + C_H C_C n \|\Delta\theta\|^{\widehat{\ell}_3} + C_D^r C_5 \|\Delta\theta\|^{\widehat{\ell}_2}\right\} \\
&\le \quad 2C_H C_E \rho^{\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3}} + C_H C_C (\log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + 1) \|\Delta\theta\|^{\widehat{\ell}_3} + C_D^r C_5 \|\Delta\theta\|^{\widehat{\ell}_2} \\
&\le \quad 2C_H C_E \|\Delta\theta\|^{\widehat{\ell}_3} + C_H C_C \|\Delta\theta\|^{\widehat{\ell}_3} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} + C_H C_C \|\Delta\theta\|^{\widehat{\ell}_3} + C_D^r C_5 \|\Delta\theta\|^{\widehat{\ell}_2} \\
&\le \quad (2C_H C_E + C_H C_C + C_D^r C_5) \|\Delta\theta\|^{\widehat{\ell}_2} + C_H C_C \|\Delta\theta\|^{\widehat{\ell}_3} \log_\rho \|\Delta\theta\|^{\widehat{\ell}_3} \\
&\le \quad \left(2C_H C_E + C_H C_C + C_D^r C_5 + C_H C_C C(\widehat{\ell}_2/\widehat{\ell}_3)\right) \|\Delta\theta\|^{\widehat{\ell}_2}.
\end{aligned}$$

Here, we have used Lemma 3.1 in the last inequality with $0 < \widehat{\ell}_2/\widehat{\ell}_3 < 1$ and $C(\widehat{\ell}_2/\widehat{\ell}_3) < \infty$ a constant.

Finally, since we are free to choose any $\bar{x} \in \mathsf{X}$, we choose a minimizing $\bar{x}$ in $\bar{V}_\ell^r(\bar{x})$ for the tightest bound. Unifying the two cases, there exists a $C_h < \infty$ such that

$$\|h_\ell(\theta) - h_\ell(\theta')\| \le \ell^2 C_h \|\theta - \theta'\|^{\widehat{\ell}_2}, \qquad \theta, \theta' \in \Theta,$$

for all $\ell = 1, 2, \ldots$. ∎

A.8 Localized Versions of the BMP Lemmas

Here, we adapt BMP Lemma's 2 through 6 [6, pp. 223-228], which provide a bound for each term of the decomposition, to our conditions and framework of this chapter. These adapted lemma's are then collected in Proposition 6.7 to show the the properties of the overall noise term.

A.8.1 A Proof of Lemma 6.2

Proof: Part 1. Consider the sum

$$S_n \doteq \sum_{k=0}^{n-1} \gamma_{k+1} 1_{\{k+1 \le \tau\}} \nabla\phi(\theta_k) \cdot (\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)), \qquad n = 1, 2, \ldots,$$

which is a martingale since

$$\mathbf{E}\left[1_{\{k+1 \le \tau\}} \nu_{\theta_k}(X_{k+1}) | \mathcal{F}_k\right] = 1_{\{k+1 \le \tau\}} P_{\theta_k} \nu_{\theta_k}(X_k).$$

Also, since the conditional expectation is a contraction in L^2 [108, p. 88], or via Jensen's Inequality, we have

$$\mathbf{E}_{\theta,x}\left[1_{\{k+1 \le \tau\}} |\nabla\phi(\theta_k) \cdot P_{\theta_k}\nu_{\theta_k}(X_k)|^2\right] \le \mathbf{E}_{\theta,x}\left[1_{\{k+1 \le \tau\}} |\nabla\phi(\theta_k) \cdot \nu_{\theta_k}(X_{k+1})|^2\right].$$

Incremental orthogonality and Pythagoras formula [108, p.110] along with the above results yield

$$
\begin{aligned}
\mathbf{E}_{\theta,x}\left[|S_n|^2\right] &= \mathbf{E}_{\theta,x}\left[|S_1|^2\right] + \sum_{k=2}^{n} \mathbf{E}_{\theta,x}\left[|S_k - S_{k-1}|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{k+1\le\tau\}} \left|\nabla\phi(\theta_k)\cdot(\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k))\right|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{k+1\le\tau\}} \|\nabla\phi(\theta_k)\|^2 \|\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)\|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 M_1^2 \mathbf{E}_{\theta,x}\left[1_{\{k+1\le\tau\}} \|\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)\|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 M_1^2 \Big\{\mathbf{E}_{\theta,x}\left[\mathbf{E}\left[1_{\{k+1\le\tau\}}\nu'_{\theta_k}(X_{k+1})\nu_{\theta_k}(X_{k+1})|\mathcal{F}_k\right]\right] \\
&\qquad\qquad - \mathbf{E}_{\theta,x}\left[1_{\{k+1\le\tau\}}P_{\theta_k}\nu'_{\theta_k}(X_k)P_{\theta_k}\nu_{\theta_k}(X_k)\right]\Big\} \\
&\le \sum_{k=0}^{n-1} \gamma_{k+1}^2 M_1^2 \mathbf{E}_{\theta,x}\left[\mathbf{E}\left[1_{\{k+1\le\tau\}}\nu'_{\theta_k}(X_{k+1})\nu_{\theta_k}(X_{k+1})|\mathcal{F}_k\right]\right] \\
&\le C_\nu^2 M_1^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{k+1\le\tau\}} V^{2r}(X_{k+1})\right] \\
&\le C_\nu^2 M_1^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{k+1\le\tau\}} V^{1}(X_{k+1})\right]
\end{aligned}
$$

where we have used (P2') in the second to last line. The last line follows since $r \le 1/4$. Applying (D0') to the last line we find

$$
\mathbf{E}_{\theta,x}\left[S_n^2\right] \le M_1^2 C_\nu^2 C_D V(x) \sum_{k=0}^{n-1} \gamma_{k+1}^2
$$

The bound in the first part of the lemma follows then from Doob's inequality

$$
\begin{aligned}
\mathbf{E}_{\theta,x}\left[\sup_{n\le m} 1_{\{n\le\tau\}} \left|\sum_{k=0}^{n-1} \gamma_{k+1}\widehat{\varepsilon}_{k+1}^{(1)}\right|^2\right] &\le \mathbf{E}_{\theta,x}\left[\sup_{n\le m} |S_n|^2\right] \le 4 \sup_{n\le m} \mathbf{E}_{\theta,x} |S_n|^2 \\
&\le 4 C_\nu^2 M_1^2 C_D V(x) \sum_{k=0}^{m-1} \gamma_{k+1}^2, \qquad x \in \mathsf{X}, \quad \theta \in Q.
\end{aligned}
$$

Part 2. For the convergence properties in the second part of the lemma, we note that on $\{\tau(Q) = \infty\}$ we have the bound

$$
\sum_{k=0}^{\infty} \gamma_{k+1}^2 < \infty
$$

which implies the martingale converges a.s. and also in L^2 since it is bounded in L^2 [108].

Part 3. Now consider the sum

$$Z_n \doteq \sum_{k=0}^{n-1} \gamma_{k+1} 1_{\{\theta_k \in Q\}} \nabla\phi(\theta_k) \cdot (\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)), \qquad n = 1, 2, \ldots,$$

which is a martingale since

$$\mathbf{E}\left[1_{\{\theta_k \in Q\}} \nu_{\theta_k}(X_{k+1}) | \mathcal{F}_k\right] = 1_{\{\theta_k \in Q\}} P_{\theta_k}\nu_{\theta_k}(X_k).$$

Also, since the conditional expectation is a contraction in L^2 [108, p. 88], or via Jensen's Inequality, we have

$$\mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} \left|\nabla\phi(\theta_k) \cdot P_{\theta_k}\nu_{\theta_k}(X_k)\right|^2\right] \leq \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} \left|\nabla\phi(\theta_k) \cdot \nu_{\theta_k}(X_{k+1})\right|^2\right].$$

Incremental orthogonality and Pythagoras formula [108, p.110] along with the above results yield

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[|Z_n|^2\right] &= \mathbf{E}_{\theta,x}\left[|Z_1|^2\right] + \sum_{k=2}^{n} \mathbf{E}_{\theta,x}\left[|Z_k - Z_{k-1}|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} \left|\nabla\phi(\theta_k) \cdot (\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k))\right|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} \|\nabla\phi(\theta_k)\|^2 \|\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)\|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 M_1^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} \|\nu_{\theta_k}(X_{k+1}) - P_{\theta_k}\nu_{\theta_k}(X_k)\|^2\right] \\
&= \sum_{k=0}^{n-1} \gamma_{k+1}^2 M_1^2 \Big\{\mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} \mathbf{E}\left[\nu'_{\theta_k}(X_{k+1})\nu_{\theta_k}(X_{k+1}) | \mathcal{F}_k\right]\right] \\
&\qquad\qquad - \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} P_{\theta_k}\nu'_{\theta_k}(X_k) P_{\theta_k}\nu_{\theta_k}(X_k)\right]\Big\} \\
&\leq \sum_{k=0}^{n-1} \gamma_{k+1}^2 M_1^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} \mathbf{E}\left[\nu'_{\theta_k}(X_{k+1})\nu_{\theta_k}(X_{k+1}) | \mathcal{F}_k\right]\right] \\
&\leq C_\nu^2 M_1^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} V^{2r}(X_{k+1})\right] \\
&\leq C_\nu^2 M_1^2 \sum_{k=0}^{n-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{\theta_k \in Q\}} V^{1}(X_{k+1})\right]
\end{aligned}$$

where we have used (P2') in the second to last line. The last line follows since $r \leq 1/4$. Applying (D0') to the last line we find

$$\mathbf{E}_{\theta,x}\left[Z_n^2\right] \;\leq\; M_1^2 C_\nu^2 C_D V(x) \sum_{k=0}^{n-1} \gamma_{k+1}^2$$

and since $\sum_{k=0}^{\infty} \gamma_{k+1}^2 < \infty$, the martingale Z_n converges $\mathbf{P}_{\theta,x}$-a.s. and in L^2. ■

A.8.2 A Proof of Lemma 6.3

Proof:

Under (P3'),

$$\|P_\theta \nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\| \leq C_\delta V^r(x)\,\|\theta-\theta'\|^{\widehat{\ell}_1}, \qquad x \in \mathsf{X}, \quad \theta,\theta' \in Q. \tag{A.29}$$

For $k = 1, 2, \ldots$ we have from (H2') and the definition of the SA that

$$\begin{aligned}
\|\theta_k - \theta_{k-1}\| &\leq \gamma_k \|H(\theta_{k-1}, X_k) + \gamma_k \rho_k(\theta_{k-1}, X_k) + z_k\| && \text{(A.30)}\\
&\leq 2\gamma_k \|H(\theta_{k-1}, X_k) + \gamma_k \rho_k(\theta_{k-1}, X_k)\| && \text{(A.31)}\\
&\leq 2C_H \gamma_k V^r(X_k) + 2C_{\rho 3}\gamma_k^2 V^r(X_k) \\
&\leq 2\left(C_H + \gamma_1 C_{\rho 3}\right)\gamma_k V^r(X_k) && \text{(A.32)}
\end{aligned}$$

Above, the projection term is bounded by

$$\|z_k\| \leq \|H(\theta_{k-1}, X_k) + \gamma_k \rho_k(\theta_{k-1}, X_k)\|$$

which follows since $\theta_k \in \Theta$ and at the very least, the projection term can return the iterate to this point so $\theta_{k+1} \in \Theta$.

We next observe that for any $\theta, \theta' \in Q$

$$\begin{aligned}
|\psi_\theta(x) - \psi_{\theta'}(x)| &= |\nabla\phi(\theta)\cdot P_\theta\nu_\theta(x) - \nabla\phi(\theta')\cdot P_{\theta'}\nu_{\theta'}(x)| \\
&\leq |\nabla\phi(\theta)\cdot P_\theta\nu_\theta(x) - \nabla\phi(\theta')\cdot P_\theta\nu_\theta(x)| + |\nabla\phi(\theta')\cdot P_\theta\nu_\theta(x) - \nabla\phi(\theta')\cdot P_{\theta'}\nu_{\theta'}(x)| \\
&\leq \|\nabla\phi(\theta) - \nabla\phi(\theta')\|\,\|P_\theta\nu_\theta(x)\| + \|\nabla\phi(\theta')\|\,\|P_\theta\nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\| \\
&\leq M_2\|\theta-\theta'\|\,\|P_\theta\nu_\theta(x)\| + M_1\|P_\theta\nu_\theta(x) - P_{\theta'}\nu_{\theta'}(x)\|.
\end{aligned}$$

Hence, by (P2') and (P3')

$$|\psi_\theta(x) - \psi_{\theta'}(x)| \leq M_2\|\theta-\theta'\|\,C_\nu V^r(x) + M_1 C_\delta \|\theta-\theta'\|^{\widehat{\ell}_1} V^r(x) \tag{A.33}$$

Part 1. By (A.29) and (A.32),

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1}\gamma_{k+1}\left|\widetilde{\varepsilon}^{(2)}_{k+1}\right|\right)^2\right] \\
&= \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left|\psi_{\theta_k}(X_k) - \psi_{\theta_{k-1}}(X_k)\right| 1_{\{k+1\leq\tau\}}\right)^2\right] \\
&\leq \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\left(M_2C_\nu\|\theta_k-\theta_{k-1}\| + M_1C_\delta\|\theta_k-\theta_{k-1}\|^{\widehat{\ell}_1}\right)V^r(X_k)1_{\{k+1\leq\tau\}}\right)^2\right] \\
&\leq 2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}M_2C_\nu\|\theta_k-\theta_{k-1}\|\,V^r(X_k)1_{\{k+1\leq\tau\}}\right)^2\right] \\
&\quad +2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}M_1C_\delta\|\theta_k-\theta_{k-1}\|^{\widehat{\ell}_1}V^r(X_k)1_{\{k+1\leq\tau\}}\right)^2\right]
\end{aligned}$$

Applying (A.32) to this last line

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1}\gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \\
&\leq 8M_2^2C_\nu^2\left(C_H+\gamma_1C_{\rho3}\right)^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\gamma_kV^{2r}(X_k)1_{\{k+1\leq\tau\}}\right)^2\right] \\
&\quad+8M_1^2C_\delta^2\left(C_H+\gamma_1C_{\rho3}\right)^{2\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_{k+1}\gamma_k^{\widehat{\ell}_1}V^{r(1+\widehat{\ell}_1)}(X_k)1_{\{k+1\leq\tau\}}\right)^2\right] \\
&\leq 8M_2^2C_\nu^2\left(C_H+\gamma_1C_{\rho3}\right)^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_k^2V^{2r}(X_k)1_{\{k+1\leq\tau\}}\right)^2\right] \\
&\quad+8M_1^2C_\delta^2\left(C_H+\gamma_1C_{\rho3}\right)^{2\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}V^{r(1+\widehat{\ell}_1)}(X_k)1_{\{k+1\leq\tau\}}\right)^2\right]
\end{aligned}$$

Since for all $k=1,2,\ldots,$ we have $\gamma_k^2=\gamma_1^2\left(\frac{\gamma_k^2}{\gamma_1^2}\right)\leq\gamma_1^2\left(\frac{\gamma_k^2}{\gamma_1^2}\right)^{\frac{1+\widehat{\ell}_1}{2}}=\left(\frac{\gamma_1^2}{\gamma_1^{(1+\widehat{\ell}_1)}}\right)\gamma_k^{(1+\widehat{\ell}_1)}$ and $V^{r(1+\widehat{\ell}_1)}(x)\leq V^{2r}(x)$

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1}\gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \\
&\leq\left(8M_2^2C_\nu^2\left(C_H+\gamma_1C_{\rho3}\right)^2\left(\frac{\gamma_1^2}{\gamma_1^{(1+\widehat{\ell}_1)}}\right)^2+8M_1^2C_\delta^2\left(C_H+\gamma_1C_{\rho3}\right)^{2\widehat{\ell}_1}\right) \\
&\quad\times\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}V^{2r}(X_k)1_{\{k+1\leq\tau\}}\right)^2\right]
\end{aligned}$$

Let A_2' be the large constant term in parentheses and if we apply the Schwarz inequality [50, p. 2]

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1}\gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] &\leq A_2'\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}V^{2r}(X_k)1_{\{k+1\leq\tau\}}\right)^2\right] \\
&\leq A_2'\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}V^{4r}(X_k)1_{\{k+1\leq\tau\}}\right] \\
&\leq A_2'\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[V(X_k)1_{\{k+1\leq\tau\}}\right] \\
&\leq A_2'\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}C_DV(x) \\
&= A_2V(x)\left(\sum_{k=1}^{m-1}\gamma_k^{1+\widehat{\ell}_1}\right)^2
\end{aligned}$$

with $A_2 = C_D A_2'$. Above, we have used the assumption that $r \leq 1/4$.

Part 2. By (A.29) and a simple bound on the square of a sum,

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \\
&= \mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left|\psi_{\theta_k}(X_k)-\psi_{\theta_{k-1}}(X_k)\right|\right)^2\right] \\
&\leq 2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}M_2C_\nu 1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left\|\theta_k-\theta_{k-1}\right\|V^r(X_k)\right)^2\right] \\
&\quad +2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}M_1C_\delta 1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left\|\theta_k-\theta_{k-1}\right\|^{\widehat{\ell}_1}V^r(X_k)\right)^2\right]
\end{aligned}$$

If we apply (A.32) to this we get

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \\
&\leq 8M_2^2C_\nu^2\left(C_H+\gamma_1C_{\rho3}\right)^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}\gamma_k1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}V^{2r}(X_k)\right)^2\right] \\
&\quad + 8M_1^2C_\delta^2\left(C_H+\gamma_1C_{\rho3}\right)^{2\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{m-1}\gamma_{k+1}\gamma_k^{\widehat{\ell}_1}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}V^{r(1+\widehat{\ell}_1)}(X_k)\right)^2\right] \\
&\leq 8M_2^2C_\nu^2\left(C_H+\gamma_1C_{\rho3}\right)^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_k^2 1_{\{\theta_{k-1}\in Q\}}V^{2r}(X_k)\right)^2\right] \\
&\quad + 8M_1^2C_\delta^2\left(C_H+\gamma_1C_{\rho3}\right)^{2\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}1_{\{\theta_{k-1}\in Q\}}V^{r(1+\widehat{\ell}_1)}(X_k)\right)^2.\right]
\end{aligned}$$

Since for all $k = m, m+1, \ldots,$ we have $\gamma_k^2 = \gamma_1^2\left(\frac{\gamma_k^2}{\gamma_1^2}\right) \leq \gamma_1^2\left(\frac{\gamma_k^2}{\gamma_1^2}\right)^{\frac{1+\widehat{\ell}_1}{2}} = \left(\frac{\gamma_1^2}{\gamma_1^{(1+\widehat{\ell}_1)}}\right)\gamma_k^{(1+\widehat{\ell}_1)}$ and $V^{r(1+\widehat{\ell}_1)}(x) \leq V^{2r}(x)$

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left|\widehat{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \\
&\leq \left(8M_2^2C_\nu^2\left(C_H+\gamma_1C_{\rho3}\right)^2\left(\frac{\gamma_1^2}{\gamma_1^{(1+\widehat{\ell}_1)}}\right)^2 + 8M_1^2C_\delta^2\left(C_H+\gamma_1C_{\rho3}\right)^{2\widehat{\ell}_1}\right) \\
&\quad \times\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}1_{\{\theta_{k-1}\in Q\}}V^{2r}(X_k)\right)^2\right]
\end{aligned}$$

Let B_2' be the large constant term in parentheses and if we apply the Schwarz inequality

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left|\widetilde{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right]\\
&\leq B_2'\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}1_{\{\theta_{k-1}\in Q\}}V^{2r}(X_k)\right)^2\right]\\
&\leq B_2'\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}\right)\mathbf{E}_{\theta,x}\left[\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}1_{\{\theta_{k-1}\in Q\}}V^{4r}(X_k)\right]\\
&\leq B_2'\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}\right)\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}\mathbf{E}_{\theta,x}\left[1_{\{\theta_{k-1}\in Q\}}V(X_k)\right]\\
&\leq B_2'\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}\right)\sum_{k=m}^{m-1}\gamma_k^{1+\widehat{\ell}_1}C_DV(x)\\
&= B_2V(x)\left(\sum_{k=m}^{n-1}\gamma_k^{1+\widehat{\ell}_1}\right)^2
\end{aligned}$$

with $B_2 = C_D B_2'$. Above, we have used the assumption that $r \leq 1/4$.

Also,

$$\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{\infty}\gamma_{k+1}1_{\{\theta_k\in Q\}}1_{\{\theta_{k-1}\in Q\}}\left|\widetilde{\varepsilon}_{k+1}^{(2)}\right|\right)^2\right] \leq B_2V(x)\left(\sum_{k=m}^{\infty}\gamma_k^{1+\widehat{\ell}_1}\right)^2 \tag{A.34}$$

■

A.8.3 A Proof of Lemma 6.4

Proof: First we note that

$$\sup_{\theta\in Q}|\psi_\theta(x)| = \left|\frac{d\phi}{d\theta}(\theta)\cdot P_\theta\nu_\theta(x)\right| \leq \left\|\frac{d\phi}{d\theta}(x)\right\| \|P_\theta\nu_\theta(x)\| \leq M_1C_\nu V^r(x).$$

Part 1. We have

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1}\gamma_{k+1}\left|\widetilde{\varepsilon}_{k+1}^{(3)}\right|\right)^2\right] &= \mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\left|\psi_{\theta_{k-1}}(X_k)\right|1_{\{k+1\leq\tau\}}\right)^2\right]\\
&\leq M_1^2C_\nu^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})V^r(X_k)1_{\{k+1\leq\tau\}}\right)^2\right].
\end{aligned}$$

Next, the Schwarz inequality yields

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=1}^{m\wedge\tau-1}\gamma_{k+1}\left|\widetilde{\varepsilon}_{k+1}^{(3)}\right|\right)^2\right]\\
&\leq M_1^2C_\nu^2\left(\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})\right)\mathbf{E}_{\theta,x}\left[\sum_{k=1}^{m-1}(\gamma_k-\gamma_{k+1})V^{2r}(X_k)1_{\{k+1\leq\tau\}}\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq M_1^2 C_\nu^2 \left(\sum_{k=1}^{m-1}(\gamma_k - \gamma_{k+1})\right) \sum_{k=1}^{m-1} (\gamma_k - \gamma_{k+1}) \mathbf{E}_{\theta,x} \left[V^1(X_k) 1_{\{k+1 \leq \tau\}}\right] \\
&\leq M_1^2 C_\nu^2 \gamma_1 \sum_{k=1}^{m-1} (\gamma_k - \gamma_{k+1}) C_D V(x) \\
&\leq M_1^2 C_\nu^2 C_D V(x) \gamma_1^2.
\end{aligned}
$$

Part 2. Here we have

$$
\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1} \gamma_{k+1} 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}} \left|\widehat{\varepsilon}_{k+1}^{(3)}\right|\right)^2\right] \\
&\quad = \mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1} 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}} (\gamma_k - \gamma_{k+1}) \left|\psi_{\theta_{k-1}}(X_k)\right|\right)^2\right] \\
&\quad \leq M_1^2 C_\nu^2 \mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1} (\gamma_k - \gamma_{k+1}) 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}} V^r(X_k)\right)^2\right].
\end{aligned}
$$

Next, the Schwarz inequality yields

$$
\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1} \gamma_{k+1} 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}} \left|\widehat{\varepsilon}_{k+1}^{(3)}\right|\right)^2\right] \\
&\leq M_1^2 C_\nu^2 \left(\sum_{k=m}^{n-1}(\gamma_k - \gamma_{k+1})\right) \mathbf{E}_{\theta,x}\left[\sum_{k=m}^{n-1}(\gamma_k - \gamma_{k+1}) 1_{\{\theta_k \in Q\}} 1_{\{\theta_{k-1} \in Q\}} V^{2r}(X_k)\right] \\
&\leq M_1^2 C_\nu^2 \left(\sum_{k=m}^{n-1}(\gamma_k - \gamma_{k+1})\right) \sum_{k=m}^{n-1} (\gamma_k - \gamma_{k+1}) \mathbf{E}_{\theta,x}\left[1_{\{\theta_{k-1} \in Q\}} V^1(X_k)\right] \\
&\leq M_1^2 C_\nu^2 \gamma_m \sum_{k=m}^{n-1} (\gamma_k - \gamma_{k+1}) C_D V(x) \\
&\leq M_1^2 C_\nu^2 C_D V(x) \gamma_m^2.
\end{aligned}
$$

■

A.8.4 A Proof of Lemma 6.5

Proof: First we have from (H2') and (6.11):

$$
\begin{aligned}
\gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(4)}\right| &= \left|\gamma_{k+1}^2 \nabla\phi(\theta_k) \cdot \rho_{k+1}(\theta_k, X_{k+1}) + R(\phi, \theta_k, \theta_{k+1})\right| \\
&\leq \gamma_{k+1}^2 \left\|\nabla\phi(\theta_k)\right\| \cdot \left\|\rho_{k+1}(\theta_k, X_{k+1})\right\| + \left|R(\phi, \theta_k, \theta_{k+1})\right| \\
&\leq \gamma_{k+1}^2 M_1 C_{\rho 3} V^r(X_{k+1}) + \gamma_{k+1}^2 M_2 \left\|H(\theta_k, X_{k+1}) + \gamma_{k+1}\rho_{k+1}(\theta_k, X_{k+1})\right\|^2 \\
&\leq \gamma_{k+1}^2 M_1 C_{\rho 3} V^r(X_{k+1}) + \gamma_{k+1}^2 2M_2 (C_H^2 + \gamma_{k+1}^2 C_{\rho 3}^2) V^{2r}(X_{k+1}) \\
&\leq \gamma_{k+1}^2 (M_1 C_{\rho 3} + 2M_2 C_H^2 + 2\gamma_1^2 C_{\rho 3}^2) V^{2r}(X_{k+1}) \\
&= \gamma_{k+1}^2 B V^{2r}(X_{k+1})
\end{aligned}
$$

where $B \doteq M_1 C_{\rho 3} + 2M_2 C_H^2 + 2\gamma_1^2 C_{\rho 3}^2$.

Part 1. With these results,

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m\wedge\tau-1}\gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(4)}\right|\right)^2\right] &\le B^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2 V^{2r}(X_{k+1})1_{\{k+1\le\tau\}}\right)^2\right] \\
&\le B^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)\sum_{k=0}^{m-1}\gamma_{k+1}^2 V^{4r}(X_{k+1})1_{\{k+1\le\tau\}}\right], \\
&\qquad m = 1,2,\ldots
\end{aligned}$$

where the last line follows from the Schwarz inequality. We have $4r \le 1$, so

$$\begin{aligned}
\mathbf{E}_{\theta,x}\left[\left(\sum_{k=0}^{m\wedge\tau-1}\gamma_{k+1}\left|\widehat{\varepsilon}_{k+1}^{(4)}\right|\right)^2\right] &\le B^2\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)\mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1}\gamma_{k+1}^2 V^{1}(X_{k+1})1_{\{k+1\le\tau\}}\right] \\
&\le B^2\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)C_D V(x)\sum_{k=0}^{m-1}\gamma_{k+1}^2 \\
&\le A_4\left(\sum_{k=0}^{m-1}\gamma_{k+1}^2\right)^2 V(x), \qquad m = 1,2,\ldots
\end{aligned}$$

where $A_4 = B^2 C_D$.

Part 2. Similarly,

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}\left|\widehat{\varepsilon}_{k+1}^{(4)}\right|\right)^2\right] \\
&\le B^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}^2 1_{\{\theta_k\in Q\}}V^{2r}(X_{k+1})\right)^2\right] \\
&\le B^2\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}^2\right)\sum_{k=n}^{m-1}\gamma_{k+1}^2 1_{\{\theta_k\in Q\}}V^{4r}(X_{k+1})\right], \qquad m = 0,1,2,\ldots,\; n > m,
\end{aligned}$$

where the last line follows from the Schwarz inequality. We assume $4r \le 1$, so

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\left(\sum_{k=m}^{n-1}\gamma_{k+1}1_{\{\theta_k\in Q\}}\left|\widehat{\varepsilon}_{k+1}^{(4)}\right|\right)^2\right] \\
&\le B^2\left(\sum_{k=m}^{n-1}\gamma_{k+1}^2\right)\mathbf{E}_{\theta,x}\left[\sum_{k=m}^{n-1}\gamma_{k+1}^2 1_{\{\theta_k\in Q\}}V^{1}(X_{k+1})\right] \\
&\le B^2\left(\sum_{k=m}^{n-1}\gamma_{k+1}^2\right)C_D V(x)\sum_{k=m}^{n-1}\gamma_{k+1}^2 \\
&\le B_4\left(\sum_{k=m}^{n-1}\gamma_{k+1}^2\right)^2 V(x), \qquad m = 0,1,2,\ldots;\; n > m,
\end{aligned}$$

where $B_4 = B^2 C_D$. ∎

A.8.5 A Proof of Lemma 6.6

Proof: Part 1. Recall $\widetilde{\eta}_{0;n} \doteq \gamma_1 \nabla\phi(\theta) \cdot P_\theta \nu_\theta(x) - \gamma_n \nabla\phi(\theta_{n-1}) \cdot P_{\theta_{n-1}} \nu_{\theta_{n-1}}(X_n)$ for $n = 1, 2, \ldots$.

First we have

$$|\gamma_1 \nabla\phi(\theta) \cdot P_\theta \nu_\theta(x)|^2 \le \gamma_1^2 M_1^2 C_\nu^2 V^{2r}(x) \le \gamma_1^2 M_1^2 C_\nu^2 V^1(x)$$

and for each $m = 1, 2, \ldots$

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} 1_{\{n\le\tau\}} \left|\gamma_n \nabla\phi(\theta_{n-1}) \cdot P_{\theta_{n-1}} \nu_{\theta_{n-1}}(X_n)\right|^2\right] \\
&\le \; M_1^2 C_\nu^2 \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} \gamma_n^2 V^{2r}(X_n) 1_{\{n\le\tau\}}\right]
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} 1_{\{n\le\tau\}} \left|\widetilde{\eta}_{0;n}\right|^2\right] \\
&= \; \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} 1_{\{n\le\tau\}} \left|\gamma_1 \nabla\phi(\theta) \cdot P_\theta \nu_\theta(x) - \gamma_n \nabla\phi(\theta_{n-1}) \cdot P_{\theta_{n-1}} \nu_{\theta_{n-1}}(X_n)\right|^2\right] \\
&\le \; \mathbf{E}_{\theta,x}\left[\sup_{1\le n\le m} 1_{\{n\le\tau\}} \left(2\left|\gamma_n \nabla\phi(\theta_{n-1}) \cdot P_{\theta_{n-1}} \nu_{\theta_{n-1}}(X_n)\right|^2 + 2\left|\gamma_1 \nabla\phi(\theta) \cdot P_\theta \nu_\theta(x)\right|^2\right)\right] \\
&\le \; 2M_1^2 \mathbf{E}_{\theta,x}\left[\sum_{k=0}^{m-1} \gamma_{k+1}^2 1_{\{n\le\tau\}} V^{2r}(X_{k+1}) C_\nu^2\right] + 2M_1^2 \gamma_1^2 C_\nu^2 V^{2r}(x) \\
&\le \; 2M_1^2 C_\nu^2 \sum_{k=0}^{m-1} \gamma_{k+1}^2 \mathbf{E}_{\theta,x}\left[1_{\{n\le\tau\}} V^1(X_{k+1})\right] + 2\gamma_1^2 M_1^2 C_\nu^2 V^1(x) \\
&\le \; 2M_1^2 C_D C_\nu^2 V(x) \sum_{k=0}^{m-1} \gamma_{k+1}^2 + 2\gamma_1^2 M_1^2 C_\nu^2 V(x) \\
&\le \; 4C_D M_1^2 C_\nu^2 V(x) \sum_{k=0}^{m-1} \gamma_{k+1}^2, \qquad m = 1, 2, \ldots
\end{aligned}$$

and recalling $C_D \ge 1$ for the last line.

Part 2. To prove the second statement, we use a similar argument to yield

$$\mathbf{E}_{\theta,x}\left[\sum_{n=0}^{\infty} \left|\widetilde{\eta}_{0;n}\right|^2\right] \le 4C_D M_1^2 C_\nu^2 V(x) \sum_{k=0}^{\infty} \gamma_k^2 < \infty, \qquad x \in \mathsf{X},$$

and this implies that the sum $\sum_{n=0}^{\infty} \left|\widetilde{\eta}_{0;n}\right|^2$ converges $\mathbf{P}_{\theta,x} - a.s.$ and hence $\lim_{n\to\infty} \widetilde{\eta}_{n;0} = 0$, $\mathbf{P}_{\theta,x} - a.s.$

■

A.9 A Proof of Proposition 7.11

Proof: There is no loss of generality in assuming θ_0 to be non–random, as we do from now on. We begin by writing (7.74) in the form

$$\theta_0 \in \mathbb{R}^p, \quad \theta_{n+1} = \theta_n + \gamma_{n+1}\{h(\theta_n) + \varepsilon_{n+1}\}, \quad n = 0, 1, \ldots$$

and by noting that under the i.i.d. assumption, we have

$$\mathbf{E}[\varepsilon_{n+1}|\mathcal{F}_n] = 0 \quad \text{and} \quad \mathbf{E}[||\varepsilon_{n+1}||^2|\mathcal{F}_n] = \frac{1}{\ell_{n+1}} Tr[\Sigma(\theta_n)]. \quad n = 0, 1, \ldots \tag{A.35}$$

With the notation $R(\theta) \doteq \langle \theta, h(\theta) \rangle$ for all θ in $\mathbb{R}^p$, we readily get from (A.35) that

$$\begin{aligned} \mathbf{E}[||\theta_{n+1}||^2|\mathcal{F}_n] &= ||\theta_n||^2 + 2\gamma_{n+1}R(\theta_n) + \gamma_{n+1}^2 \left[||h(\theta_n)||^2 + \frac{1}{\ell_{n+1}} Tr[\Sigma(\theta_n)] \right] \\ &\le ||\theta_n||^2 + \gamma_{n+1}^2 K(1 + ||\theta_n||^2) + \frac{\gamma_{n+1}^2}{\ell_{n+1}} K(1 + ||\theta_n||^2) \qquad \text{(A.36)} \\ &\le (1 + 2K\gamma_{n+1}^2)||\theta_n||^2 + 2K\gamma_{n+1}^2 \quad n = 0, 1, \ldots \qquad \text{(A.37)} \end{aligned}$$

where in (A.36) we used (7.75)–(7.76). Next we introduce the integrable rvs $\{M_n,\ n = 0, 1, \ldots\}$ by setting

$$M_0 \doteq ||\theta_0||^2, \quad M_{n+1} \doteq \alpha_{n+1}||\theta_{n+1}||^2 - \beta_{n+1} \quad n = 0, 1, \ldots \tag{A.38}$$

with

$$\alpha_{n+1} \doteq \prod_{i=0}^{n} \left(1 + 2K\gamma_{i+1}^2\right)^{-1} \quad \text{and} \quad \beta_{n+1} \doteq \sum_{i=0}^{n} 2K\gamma_{i+1}^2 \alpha_{i+1}. \quad n = 0, 1, \ldots$$

We observe that (A.37) is equivalent to the supermartingale property

$$\mathbf{E}[M_{n+1}|\mathcal{F}_n] \le M_n \quad \mathbf{P} - a.s. \qquad n = 0, 1, \ldots \tag{A.39}$$

so that

$$\sup_n \mathbf{E}[M_n] \le ||\theta_0||^2. \tag{A.40}$$

We also note the easy bounds

$$A \le \alpha_{n+1} \le 1 \quad \text{and} \quad 0 \le \beta_{n+1} \le B \qquad n = 0, 1, \ldots \tag{A.41}$$

where

$$A \doteq \exp[-2K \sum_{i=0}^{\infty} \gamma_{i+1}^2] \quad \text{and} \quad B \doteq \lim_{n\to\infty} \beta_n; \tag{A.42}$$

from (7.73) we see that $0 < A \le 1$ and $B < \infty$. From (A.38), with $\alpha_0 = 1$ and $\beta_0 = 0$, we readily obtain the inequalities

$$M_n + \beta_n \ge A||\theta_n||^2 \quad \text{and} \quad |M_n| \le \alpha_n||\theta_n||^2 + \beta_n. \quad n = 0, 1, \ldots \tag{A.43}$$

Using (A.41)–(A.42) we conclude from (A.40) and the first inequality in (A.43) that

$$\sup_n \mathbf{E}[||\theta_n||^2] < \infty,$$

whence $\sup_n \mathbf{E}[|M_n|] < \infty$ by the second part of (A.43). Therefore, by the basic martingale convergence theorem [57, Thm. 5.1., p. 278], the supermartingale $\{M_n,\ n = 0, 1, \ldots\}$ converges $\mathbf{P}$–a.s. to a finite rv, and so does also the sequence $\{||\theta_n||^2,\ n = 0, 1, \ldots\}$.

It remains to show that $\lim_{n\to\infty} ||\theta_n||^2 = 0$ $\mathbf{P}$–a.s. To do this, we take expectations on both sides of (B.3) and get

$$\begin{aligned} \mathbf{E}[||\theta_{n+1}||^2] &= \mathbf{E}[||\theta_n||^2] + 2\gamma_{n+1}\mathbf{E}[R(\theta_n)] + \gamma_{n+1}^2 \mathbf{E}\left[||h(\theta_n)||^2 + \frac{1}{\ell_{n+1}} Tr[\Sigma(\theta_n)]\right] \\ & \qquad n = 0, 1, \ldots \end{aligned} \tag{A.44}$$

After adding these relations for $k = 0, 1, \ldots, n$ and canceling appropriate terms, we are then left with the relation for $n = 0, 1, \ldots$

$$\mathbf{E}[||\theta_{n+1}||^2] = ||\theta_0||^2 + \sum_{k=0}^{n} \gamma_{k+1}\mathbf{E}[R(\theta_k)] + \sum_{k=0}^{n} \gamma_{k+1}^2 \mathbf{E}\left[\mathbf{E}[||h(\theta_k)||^2] + \frac{1}{\ell_{k+1}} Tr[\Sigma(\theta_k)\right]. \tag{A.45}$$

Upon using the inequality (7.76) and the bound $\sup_n \mathbf{E}[||\theta_n||^2] < \infty$ obtained earlier, we easily conclude from (A.45) that

$$0 \le -\sum_{k=0}^{\infty} \gamma_{k+1}\mathbf{E}[R(\theta_k)] < \infty. \tag{A.46}$$

Therefore, $\lim_{k\to\infty} \gamma_{k+1}\mathbf{E}[R(\theta_k)] = 0$ and under **(S')** a simple argument by contradiction shows that we must necessarily have $\liminf_{k\to\infty} \mathbf{E}[R(\theta_k)] = 0$. In other words, along a subsequence, say $\{n_j,\ j = 1, 2, \ldots\}$, we have $\lim_{j\to\infty} \mathbf{E}[R(\theta_{n_j})] = 0$, whence $\lim_{j\to\infty} R(\theta_{n_j}) = 0$ in probability (under $\mathbf{P}$). Consequently, along a further subsequence, still denoted $\{n_j,\ j = 1, 2, \ldots\}$, we have $\lim_{j\to\infty} R(\theta_{n_j}) = 0$ $\mathbf{P}$–a.s. Using this last fact in conjunction with (7.75) readily yields $\lim_{j\to\infty} \theta_{n_j} = 0$ $\mathbf{P}$–a.s. and the desired conclusion now follows. ■

BIBLIOGRAPHY

[1] Aristotle Arapostathis, Emmanuel Fernández-Gaucherand, and Steven I. Marcus. Analysis of an adaptive control scheme for a partially observed markov chain. In *Proceedings of the 29th Conference on Decision and Control*, pages 1438–1444, 1990.

[2] F. Archetti, A. Gaivoronski, and A. Sciomachen. Sensitivity analysis and optimization of stochastic petri nets. *Discrete Event Dynamic Systems: Theory and Applicaitons*, 3:5–37, 1993.

[3] Sridhar Bashyam, Michael C. Fu, and Bharat K. Kaku. Application of perturbation analysis to multiproduct capacitated production-inventory control. In *Proceedings of the American Control Conference*, pages 2170–2174, 1995.

[4] P.H. Baxendale. Uniform estimates for geometric ergodicity of recurrent markov processes. Technical report, University of Southern California, 1993.

[5] Albert Benveniste and M. Goursat. Blind equalizers. *IEEE Transactions on Communications*, 32:871–883, August 1984.

[6] Albert Benveniste, Michel Métivier, and Pierre Priouret. *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, Berlin, 1990.

[7] Shalabh Bhatnagar. *Multiscale Stochastic Approximation Schemes with Applications to ABR Service in ATM Networks*. PhD thesis, Indian Institute of Science, Bangalore - 560 012, India, July 1997.

[8] Patrick Billingsley. *Probability and Measure*. John Wiley and Sons, New York, 1986.

[9] Flavio Bonomi and Anurag Kumar. Adaptive optimal load balancing in a nonhomogeneous multiserver system with a central job scheduler. *IEEE Transactions on Computers*, 39(10):1232–1250, October 1990.

[10] Xi-Ren Cao. First-order perturbation analysis of a single multi-class finite source queue. *Performance Eval.*, 7:31–41, 1987.

[11] Xi-Ren Cao. The relations among potentials, perturbation analysis, and markov decision processes. *Discrete Event Dynamic Systems: Theory and Applications*, 8(1):71–87, March 1998.

[12] Xi-Ren Cao and Han-Fu Chen. Perturbation realization, potentials, and sensitivity analysis of markov processes. *IEEE Transactions on Automatic Control*, 42(10):1382–1393, October 1997.

[13] Xi-Ren Cao, Michael C. Fu, and Jian-Qiang Hu. On performance potentials and conditional monte carlo for gradient estimation for markov chains. *Annals of Operations Research*, 87:263–272, 1999.

[14] Xi-Ren Cao and Yat wah Wan. Algorithms for sensitivity analysis of markov systems through potentials and perturbation realization. *IEEE Transactions on Control Systems Technology*, 6(4):482–494, July 1998.

[15] Xi-Ren Cao, Xue-Ming Yuan, and Li Qiu. A single sample path-based performance sensitivity formula for markov chains. *IEEE Transactions on Automatic Control*, 41(12):1814–1817, December 1996.

[16] Christos Cassandras. *Discrete Event Systems: Modeling and Performance Analysis*. Irwin, 1993.

[17] Hanfu Chen. Stochastic approximation procedures with randomly varying truncations. *Scientia Sinica*, 29(9), September 1986.

[18] Hanfu Chen. Stochastic approximations and its new applications. In *Proceedings of the 1994 Hong Kong International Workshop on New Directions in Control and Manufacturing*, pages 2–12, 1994.

[19] Edwin K. Chong and Peter J. Ramadge. Convergence of recursive optimization algorithms using infinitesimal perturbation analysis. *Journal of Discrete Event Dynamical Systems*, 1:339–372, 1992.

[20] Edwin K. Chong and Peter J. Ramadge. Optimization of queues using an infinitesimal perturbation analysis-based stochastic algorithm with general update times. *SIAM Journal of Control and Optimization*, 31(1):1–35, 1993.

[21] Kai Lai Chung. *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, Berlin, 1967.

[22] Liyi Dai. Sensitivity analysis of stationary performance measures for markov chains. *Mathematical Comput. Modelling*, 23(11/12):143–160, 1996.

[23] Liyi Dai. Rate of convergence for derivative estimation of discrete-time markov chains via finite difference approximation with common random numbers. *SIAM Journal of Applied Mathematics*, 57(3):731–751, June 1997.

[24] Liyi Dai and Yu-Chi Ho. Structural infinitesimal perturbation analysis (sipa) for derivative estimation of discrete event dynamic systems. *IEEE Transactions on Automatic Control*, 40(7):1154–1166, July 1995.

[25] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications.* Jones and Bartlett Publishers, Boston, MA, 1993.

[26] Jastej S. Dhingra and Gilmer L. Blankenship. Online parameter optimization for a muli-product, multi-machine manufacturing system. In *Proceedings of the 31st Conference on Decision and Control*, pages 1676–1677, 1992.

[27] N. M. Van Dijk. On a simple proof of uniformization for continuous and discrete-state continuous-time markov chains. *Advances in Applied Probability*, 22:749–750, 1990.

[28] Marie Duflo. *Random Iterative Models.* Springer-Verlag, 1997.

[29] Paul Dupuis and Rahul Simha. On sampling controlled stochastic approximation. *IEEE Transactions on Automatic Control*, 36(8):915–924, August 1991.

[30] Richard C. Ellis. Large deviations for a general class of random vectors. *Annals of Probability*, 12(1):1–12, January 1984.

[31] E. Eweda and O. Macchi. Convergence analysis of self-adaptive equalizers. *IEEE Transactions on Information Theory*, pages 161–176, 1984.

[32] E. Eweda and O. Macchi. Convergence of an adaptive linear estimation algorithm. *IEEE Transactions on Automatic Control*, pages 119–127, 1984.

[33] Bennett L. Fox. Discrete-time conversion for simulating semi-markov processes. *Operations Research Letters*, 5(4):191–196, October 1986.

[34] Bennett L. Fox and Paul Glasserman. Estimating derivatives via poisson's equation. *Probability in the Engineering and Informational Sciences*, 5:415–428, 1991.

[35] M.C. Fu, J.Q. Hu, and L.Y. Shi. Likelihood ratio methods via conditional monte carlo and splitting. In *Proceedings of the Allerton Conference on Communication, Control, and Computing*, pages 958–967, 1993.

[36] Michael Fu and Jian-Qiang Hu. *Conditional Monte Carlo*. Kluwer Academic Publishers, 1997.

[37] Michael C. Fu and Jian-Qiang Hu. Smoothed perturbation analysis derivative estimation for markov chains. *Operations Research Letters*, 15:241–251, 1994.

[38] F. R. Gantmacher. *The Theory of Matrices*, volume 2. Chelsea Publishing Company, New York, NY, 1960.

[39] E. G. Gladyshev. On stochastic approximation. *Theory of Probability and its Applications*, 10(2):275–278, 1965.

[40] Paul Glasserman. Discrete-time 'inversion" and derivative estimation for markov chains. *Operations Research Letters*, 9:305–313, September 1990.

[41] Paul Glasserman. *Gradient Estimation Via Perturbation Analysis*. Kluwer Academic Press, Boston, 1991.

[42] Paul Glasserman. Derivative estimates from simulation of continuous time markov chains. *Operations Research*, 40(2):292–308, March-April 1992.

[43] Peter W. Glynn. Optimization of stochastic systems. In *Proceedings of the 1986 Winter Simulation Conference*, pages 52–59, 1986.

[44] Peter W. Glynn. Likelihood ratio gradient estimation: An overview. In *Proceedings of the 1987 Winter Simulation Conference*, pages 366–375, 1987.

[45] Peter W. Glynn. A gsmp formalism for discrete event systems. *Proceedings of the IEEE*, 77(1):11–20, January 1989.

[46] Peter W. Glynn. Likelihood ratio derivative estimators for stochastic systems. In P. Heidelberger E.A. MacNair, K.J. Musselman, editor, *Proceedings of the 1989 Winter Simulation Conference*, pages 374–380, 1989.

[47] Peter W. Glynn. Optimization of stochastic systems via simulation. In *Proceedings of the 1989 Winter Simulation Conference*, pages 90–105, 1989.

[48] Peter W. Glynn. Likelihood ratio derivative estimators for stochastic systems. *Communications of the ACM*, 33(10):75–84, October 1990.

[49] Peter W. Glynn and Sean P. Meyn. A lyapunov bound for solutions to poisson's equation. *Annals of Probability*, 24, 1996.

[50] Israel Gohberg and Seymour Goldberg. *Basic Operator Theory*. Birkhauser, 1981.

[51] Nandakumar Gopalakrishnan and Michael P. Fitz. Burst mode carrier synchronization with digital phase-locked loops. In *IEEE MILCOM*, pages 324–328, 1992.

[52] Bruce Hajek. Stochastic approximation methods for decentralized control of multiaccess communications. *IEEE Transactions on Information Theory*, 31(2):176–184, March 1985.

[53] P. Hall and C.C. Heyde. *Martingale Limit Theory and its Applications*. Academic Press, 1980.

[54] Philip Heidelberger, Xi-Ren Cao, Michael A. Zazanis, and Rajan Suri. Convergence properties of infinitesimal perturbation analysis estimates. *Management Science*, 34(11):1281–1302, November 1988.

[55] Yu-Chi Ho and Xi-Ren Cao. *Perturbation Analysis of Discrete Event Dynamical Systems*. Kluwer Academic Press, Boston, 1991.

[56] Arie Hordijk, Donald Iglehart, and Rolf Schassberger. Discrete time methods for simulating continuous time markov chains. *Advances in Applied Probability*, 8:772–788, 1976.

[57] Samuel Karlin and Howard M. Taylor. *A First Course in Stochastic Processes*. Academic Press, San Diego, 1975.

[58] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, first edition, 1966.

[59] J. G. Kemeny and J. L. Snell. Potentials for denumerable markov chains. *J. Math. Anal. Appl.*, 3:196–260, 1961.

[60] Anurag Kumar. Adaptive load control of the central processor in a distributed system with a star topology. *IEEE Transactions on Computers*, 38(11):1502–1512, November 1989.

[61] Harold J. Kushner and Dean S. Clark. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Springer-Verlag, New York, 1978.

[62] Harold J. Kushner and Adam Shwartz. An invariant measure approach to the convergence of stochastic approximations with state-dependent noise. *SIAM Journal of Control and Optimization*, 21:1–35, 1983.

[63] Harold J. Kushner and Felisa Vazquez-Abad. Estimation of the derivative of a stationary measure with respect to a control parameter. *Journal of Applied Probability*, 1992.

[64] Harold J. Kushner and G. George Yin. *Stochastic Approximation Algorithms and Applications*. Springer, 1997.

[65] Peter Lancaster. *Theory of Matrices*. Academic Press, New York, 1969.

[66] Dan Lee. *On Open-Loop Admission Control into a Queueing System*. PhD thesis, Massachusetts Institute of Technology, 1992.

[67] Suk Lee and Asok Ray. Performance management of multiple access communication networks. *IEEE Journal on Selected Areas in Communications*, 11(9):1426–1437, December 1993.

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

[69] M. Loeve. *Probability Theory II.* Springer-Verlag, 4th edition, 1977.

[70] D.-J. Ma, A. M. Makowski, and A. Shwartz. Stochastic approximations for finite-state markov chains. *Stochastic Processes and their Applications*, 35:27–45, 1990.

[71] Armand M. Makowski. Seminar on stochastic ordering (class notes), Spring 1993. University of Maryland.

[72] Armand M. Makowski and Adam Shwartz. Stochastic approximations and adaptive contorl of a discrete-time single-server network with random routing. *SIAM Journal on Control and Optimization*, 30(6):1476–1506, November 1992.

[73] Armand M. Makowski and Adam Shwartz. On the poisson equation for countable markov chains: Existence of solutions and parameter dependence by probabilistic methods. Technical Report 94-87, Institute of Systems Research, University of Maryland, College Park, Maryland, 1994. (September 1997 version).

[74] V.A. Malyshev and F.M. Spieksma. Intrinsic convergence rate of countable markov chains. *Markov Processes and Related Fields*, 1995.

[75] Marc S. Meketon. Optimization in simulation: A survey of recent results. In *Proceedings of the 1987 Winter Simulation Conference*, pages 58–67, 1987.

[76] Michel Métivier. Mathematical analysis of stochastic algorithms. Systems Research Center Lecture Series, University of Maryland, College Park, MD, 1988.

[77] Michel Métivier and Pierre Priouret. Applications of a kushner and clark lemma to general classes of stochastic algorithms. *IEEE Transactions on Information Theory*, 30(2):140–151, March 1984.

[78] C. D. Meyer, Jr. The role of the group generalized inverse in the theory of finite markov chains. *SIAM Review*, 17:443–464, 1975.

[79] Sean P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, 1993.

[80] Sean P. Meyn and R.L. Tweedie. Computable bounds for geometric convergence rates of markov chains. *The Annals of Applied Probability*, 4(4):981–1011, 1994.

[81] J. Neveu. *Discrete-Parameter Martingales.* North-Holland, Amsterdam, 1975. Translated by T. P. Speed.

[82] S. Orey. Large deviations in ergodic theory. In K.L Chung E. Cinlar and R.K. Getoor, editors, *Seminar on Stochastic Processes, 1984.* Birkhauser, 1986.

[83] G. Ch. Pflug. On-line optimization of simulated markovian processes. *Mathematics of Operations Research*, 15(3):"381–395", August 1990.

[84] G. Ch. Pflug. On-line optimization of simulated markovian processes. *Mathematics of Operations Research*, 15(3):381–395, August 1990.

[85] G. Ch. Pflug. Gradient estimates for the performance of markov chains and discrete event processes. *Annals of Operations Research*, 39:172–193, 1992.

[86] Martin I. Reiman and Alan Weiss. Sensitivity analysis for simulations via likelihood ratios. *Operations Research*, 37(5):830–844, September-October 1989.

[87] Ramin Rezaiifar. *Optimal Resource Allocation in Cellular Networks*. PhD thesis, University of Maryland, 1996.

[88] H. Robbins and S. Monro. A stochastic approximation method. *Annals of Mathematical Statistics*, 22:400–407, 1951.

[89] H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications. In Jagdish S. Rustagi, editor, *Optimization Methods in Statistics*, pages 233–257. Academic Press, New York, 1971.

[90] R. Tyrell Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[91] J.S Rosenthal. Minorization conditions and convergence rates for markov chain monte carlo. Technical report, University of Toronto, 1993.

[92] Sheldon Ross. *Stochastic Processes*. Wiley, New York, 1983.

[93] H.L. Royden. *Real Analysis*. MacMillan Publishing Company, Third Edition, New York, NY, 1988.

[94] Reuven Y. Rubinstein and Alexander Shapiro. *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*. John Wiley and Sons, 1993.

[95] Paul J. Schweitzer. Perturbation theory and finite markov chains. *Journal of Applied Probability*, 5:401–413, 1968.

[96] P.S. Shcherbakov, S. N. Tikhonov, J. D. Mason, and K. Warwick. Learning in neural networks and stochastic approximation methods with averaging. *IEE Colloquium on Advances in Neural Networks for Control and Systems*, 1994.

[97] Adam Shwartz. *Convergence of Stochastic Approximations: The Invariant Measure Approach*. PhD thesis, Brown University, 1982.

[98] Rahul Simha and James F. Kurose. Stochastic approximation schemes for a load balancing problem. Technical report 89-10, COINS, 1989.

[99] Rahul Simha and James F. Kurose. On-line minimization of call setup time via load balancing: A stochastic approximation approach. *IEEE Transactions on Communications*, 42(2/3/4):1141–1152, February-March-April 1994.

[100] James C. Spall and John A. Cristion. Direct adaptive contro of nonlinear systems using neural netowrks and stochastic approximation. In *Proceedings of the 31st Conference on Decision and Control*, 1992.

[101] James C. Spall and John A. Cristion. A neural network controller for systems with unmodeled dynamics with applications to wastewater treatment. In *1994 IEEE International Symposium on Intelligent Control*, pages 273–278, 1994.

[102] J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, March 1992.

[103] F.M. Spieksma. *Geometrically Ergodic Markov Chains and The Optimal Control of Queues*. PhD thesis, University of Leiden, 1991.

[104] F.M. Spieksma. Spectral conditions and bounds for the rate of convergence of countable markov chains. Technical report, University of Leiden, 1993.

[105] D.W. Stroock. *An Introduction to the Theory of Large Deviations*. Springer-Verlag, 1984.

[106] Qian-Yu Tang and Han-Fu Chen. Convergence of perturbation analysis based optimization algorithm with fixed-number of customers period. *Discrete Event Dynamic Systems: Theory and Applications*, 4(4), 1994.

[107] Felisa J. Vázquez-Abad and Pierre L'Ecuyer. Comparing alternative methods for derivative estimation when ipa does not apply directly. In *Proceedings of the 1991 Winter Simulation Conference*, pages 1004–1011, 1991.

[108] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.

[109] B. Zhang and Y.-C. Ho. Performance gradient estimation for very large finite markov chains. *IEEE Transactions on Automatic Control*, 36(11):1218–1227, November 1991.