Abstract

Title of Dissertation: Estimation of hidden Markov models for partially-observed, risk-sensitive, control problems.

Bernard Adrian Frankpitt, Doctor of Philosophy, 2001

Dissertation directed by: John Baras
Department of Electrical Engineering

This dissertation formulates and solves a combined estimation and optimal control problem for a finite-state, discrete-time, partially observed, controlled, hidden Markov model with unknown state transition and output transition matrices. The cardinality of the state and the cost structure are assumed known. The control implemented at each step is a randomized approximation to an optimal risk-sensitive control, and is calculated with the current value of the plant estimates. The degree of randomization is determined by the value of a positive parameter which is allowed to decay to zero at a constant rate. As the parameter decays to zero the control converges to an optimal control for a moving horizon risk sensitive criterion.

The main contribution of the dissertation is the presentation of a stochastic approximation proof for the asymptotic convergence of the algorithm for combined estimation and control. The proof requires the development of a potential theory for the Markov

chain that captures the combined dynamics of the hidden Markov model, the estimator and the control algorithm. The potential kernel associated with this chain is shown to be regular with respect to variation in the plant estimates which influence the kernel both directly though the estimation algorithm, and indirectly through the control algorithm.

Estimation of hidden Markov models for partially-observed, risk-sensitive, control problems.

by

Bernard Adrian Frankpitt

Dissertation submitted to the Faculty of the Graduate School
of The University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2001

Advisory Committee:

John Baras, Chairman/Advisor
Professor Steven Marcus
Professor Carlos Berenstein
Professor Prakash Narayan

Table of Contents

List of Figures

Chapter 1

Introduction

This dissertation addresses the problem of designing estimators for partially-observed feedback-control systems that optimize a risk sensitive cost functional. The difficult part of this problem is the provision of assurances both that the estimator converges to the true estimate of the model, and that the controller converges to the optimal risk-sensitive control.

Theoretical considerations were the major influence for the choice of problem. The simple structure of a finite state hidden Markov model permits extensive analysis, and reveals clearly the interaction between the part of the system that implements optimal control, and the part that implements model estimation. For example, the notion of persistent excitation has a precise formulation in the system studied here as a requirement on the control policy that ensures a primitivity condition on a Markov chain, and the formulation of the estimator convergence problem as a stochastic approximation problem provides insight into why the control algorithm needs to be regular with respect to parameter variations in the model.

The work in this dissertation potentially has direct application to control problems for systems that are intrinsically difficult to model. Under these circumstances the combination of a risk-sensitive control strategy with a coarse finite state system model developed from input-output data may give better performance than control based on a poor *a priori* model. The estimation algorithms and the analysis developed in this dissertation would apply directly to such a circumstance.

The dissertation draws on ideas from control theory, stochastic approximation, and ergodic theory for discrete-time Markov processes. This introduction selectively references this broad literature focusing on the sources of the main ideas that directly motivate the current work, the works that contribute to the formulation and solution of the central problem in the dissertation, and the works that have strongly influenced the dissertation through discussions of broader issues.

The problem of combined estimation and control has long been studied in the fields of adaptive control, stochastic control, and intelligent control. One of the earliest formulations of the problem was made by Feldbaum [13] who recognized that when designing an optimal control for a plant with unknown model, there is a trade-off between the cost of acquiring additional information, and the cost of using a suboptimal control designed from an imperfect model. Feldbaum formulated the identification problem in a stochastic framework, and combined maximum likelihood estimates of the model from the input-output data with control algorithms that minimized an accrued cost. Feldbaum (or his translator) coined the name "Dual Control Problem" to describe the interaction between estimation and optimal control, and provided an apt nautical analogy by using the term "sounding" to describe the way that variation in the choice of plant inputs provides data for the model estimation process.

Following Feldbaum's work, a number of authors investigated the problem of combined control and estimation in a variety of frameworks including Sworder [42] who investigated the interaction between optimal control and estimation for a linear system with a quadratic cost functional, and Witten [53] who looked at the conflict between estimation and control for a two-armed bandit problem. Witten points out that the "Dual Control Problem" is in fact a general systems problem that occurs in any system where "the need to balance the costs of further exploratory activity against the costs of poor system knowledge creates a basic conflict that can not be avoided".

The early work on combined estimation and control spawned the field of adaptive control which followed two variations. The model reference adaptive control schemes are deterministic schemes which use Lyapunov theory to prove convergence of either the model estimates or the system trajectory. In a pair of papers [30, 31] Narendra and Kudva summarize earlier work and present a sound methodology for the linear, finite dimensional, time invariant case. More recent work of authors such as Kokotovich et al. cast the deterministic adaptive control problem in a nonlinear setting which gives new insight into the problems of convergence, and allows the techniques developed for linear systems to be extended to certain classes of nonlinear systems — particularly systems that are feedback linearizable. An important aspect of the work of Kokotovich *et al.* is the use of perturbation analysis to separate the fast dynamics of the plant from the slower dynamics of the estimator. A similar classification of the parts of the system by their characteristic time-scales applies to the system considered in this thesis.

A second approach to adaptive control is the self tuning regulator of Astrom and Wittenmark [3]. These authors take a point of view closer to that of Fieldbaum (*op. cit.*) and cast the combined estimation and control problem in a stochastic setting by introducing a noise disturbance into the linear system. The authors estimate unknown system parameters with a recursive least squares algorithm, and regulate the variance of the output process with a minimum variance controller. The self tuning regulator has additional importance as one of the stochastic approximation problems that Ljung[26, 25] chose as examples to illustrate an analysis technique that was latter dubbed "The

ODE Method". A more recent variant of the same technique forms the basis of the analysis in this dissertation.

At the same time that the field of adaptive control was developing, workers in a parallel field that has come to be known as intelligent control[1] were applying the results of work on machine learning to the "Dual Control Problem". The grand aim of intelligent control is to address the following statement of C.E. Shannon and J. McCarthy which is taken from the preface to 1956 monograph "Automata Studies" [39].

> "Among the most challenging questions of our time are the corresponding analytic and synthetic problems: How does the brain function? Can we design a machine that will simulate the brain?"

An important difference between the intelligent control approach and the adaptive control approach is that while in adative control the primary emphasis is on showing convergence of a parameterized estimated model to a "true" model, the primary emphasis in intelligent control is on learning a control strategy that improves an observed perfomance metric, such as an incremental cost, often without explicit reference to an underlying system model. A good example is the work of Widrow *et al.* on a machine that learns an optimal strategy for the game of blackjack [51]. The automata that the authors describe has two parts: The first part implements a decision rule by passing a weighted sum of input stimuli through an output step-nonlinearity (a structure sim-

[1]It scarcely seems credible that McCulloch and Pitts[29] published their famous mechanistic model of neural activity in 1943, but the date of this event is a good indication of the extent to which engineers in the second half of this century have looked to anthropomorphic biological analogy as a source of inspiration and justification for their work. Another influential early expression of this program is the book "Cybernetics" by Norbet Wiener [52]

ilar to that of the McCulloch-Pitts neuron). This decision element is augmented with an adaptive critic, an architecture that adaptively adjusts the input weights of the decision element based on a comparison of recent average performance with long-term average performance. The authors analyze the convergence properties of this system with a probabilistic argument. The ideas that underly the adaptive critic have been developed by a number of authors. The work on temporal difference learning by Sutton [41] and on Q-learning by Watkins and Dayan [45] falls into this category, and from a control point of view[2] there is the work of Werbos [46, 47] which explicitly links adaptive critics with dynamic programming. Werbos points out that the adaptive critic part of the architecture computes an approximation to a value function (or, in alternative architectures, an approximation to the gradient of a value function) and that the actor part of the architecture provides a feedback control policy implemented as a mapping from an observed state to a choice of control. The critic influences the actor by adjusting parameters in the actor architecture so that the updated control policy has an estimated value function that is smaller at each value of the state than the value function for the existing control policy. In the language of reinforcement learning, the updates in the parameters of the critic architecture are value iterations, and the updates in the parameters for the actor architecture are policy iterations.

Work on reinforcement learning and intelligent control has lead to some spectacular successes in recent years both in applications and in theoretical understanding of the algorithms [12, 8]. Stochastic approximation techniques have proved to be an important tool for obtaining convergence proofs for the architectures [44]

The simplicity of finite state models makes them an attractive choice for system

[2]Barto *et al.* [5] present a review that links the work of researchers in AI to parallel work of control theorists

models. Bertsekas and Tsitsiklis [8] and many workers in AI use controllers that are either explicitly or implicitly formulated as finite state models. The advantage of finite state models for these applications is that the dynamic programming equation satisfied by an optimal state feedback policy has computable, exact solutions. In the case of optimal output feedback policies for partially observed systems the dynamic programming equation is perfomed on a finite dimensional information state rather than a finite state set, but even here structural propeties of the solution to the dynamic programming equation provide a way to economically calculate exact solutions for finite horizon problems [14].

Motivation to look at partially observed problems comes from a traditional control theory view of systems structure. Feedback control is based on the premise that histories of observed system signals contain information that enables the construction of optimal future trajectories for controllable system parameters. A control algorithm is a mapping from the histories of output signals to input signals which represent planned future control parameter trajectories. Since signal histories are difficult objects to manipulate, the control mappings from output signals to input signals are constructed in two stages. The first stage is an information compression stage which maps an output signal history onto a state which is a simpler object, often a point on a finite dimensional manifold. The state encodes the useful information about the history output signals. The second stage is a mapping from the state to the input parameters. The advantage of this factored approach is that when the state-space is well chosen, the selection of an optimal map between the state-space and the system's control parameters is a simpler problem then the original problem of choosing an optimal mapping between an output signal space and a space of input signals.

The useful information in the signal spaces of stochastic systems is contained in the probability distributions which determine the signal statistics. A mapping from an output signal process of a stochastic system to a state signal process should preserve relevant information from the output signal process in the marginal distribution of the state trajectory. In an optimal control problem, the information of interest that the output process provides is the conditional expectation of the cost, conditioned on the past values of the output process. A choice of a state process with the property that the marginal distribution of the state process provides a set of sufficient statistics for this conditional expectation is called an information state[49, 50, 19, 4]. The structure of an information state depends on the forms of both the cost functional and the stochastic system. In the case of the linear quadratic Gaussian regulator problem, a linear system with a additive i.i.d. Gaussian state and output noise processes and a cost functional that is a quadratic function of the state and input processes, the minimum mean square error estimate of the state process provides an information state which is a vector space that is isomorphic to the underlying state space. This fact underlies the certainty equivalence principle for the linear quadratic Gaussian regulator. In general, the structure of an information state is not so simple. In the case of the hidden Markov model with a risk sensitive cost functional, the information space is isomorphic to a space of functions defined on the state space of the underlying Markov model. This situation is typical of quite general control problems, and provides a strong theoretical motivation for the study of risk-sensitive control for finite state, hidden Markov models as a simple setting for control problems that exhibit important features of a much broader class of stochastic control problems.

Combined estimation and control of a hidden Markov model is interesting because it retains the aspects of the dual control problem that are common to diverse areas in control theory, yet discards the sophisticated structural features that differentiate problems in one area from problems in another. The objective of the research in this dissertation is to gain a thorough understanding of the analysis of the dual control problem for this reduced system in the hope that the methods developed to attack the problem at the simple level will translate to equivalent methods for complicated systems with more sophisticated structure. In light of this objective the dissertation attempts to present a solution to the analysis problem in terms of general abstract frameworks that, with luck, will transcend the simple setting.

The design of estimators for the parameters of a hidden Markov model is a long-standing problem in estimation theory. The first treatment of the properties of the maximum likelihood estimator is that of Baum and Petrie [6]. The authors show that the maximum likelihood estimator minimizes the entropy of the observed process. They then define a relative entropy function and devise an algorithm that minimizes the relative entropy function. This algorithm is the EM method. The classic EM algorithm is an iterative algorithm that works by making sequential passes though the data set, at each pass the estimate of the model parameters is used to compute the relative entropy of the output process as a function of the model parameters, and minimizer of this relative entropy function is chosen as the the updated parameter estimate. Krishnmurthy and Moore[21] use work of Titterington[43] to develop a recursive version of the EM algorithm. A simple adaptation of this recusive algorithm is used in Chapter 3 as the basis for a recursive estimator.

The use of a recursive estimator gives the combined control estimation problem the character of a stochastic approximation algorithm. This is the point of view from which the dissertation atacks the analysis of the convergence properties of the system. Stochastic approximation algorithms go back at least as far as the work of Robbins and Munro [35] on the algorithm that bears their names, but the style of analysis pursued here originates in the work of Ljung [25] who introduced the ODE technique for analysis of convergence. The book of Kushner and Clark [22], which became a standard reference for the subject, extends the method introduced by Ljung to more general noise models, treats weak convergence of the algorithm as well as a.s. convergence of the paths, and considers constrained systems. The essential idea in the ODE method is to view the sequence of iterates from the stochastic approximation algorithm as a noisy, discrete approximation to a continuous ordinary differential equation. A martingale argument bounds the culmulative error of the approximation, and a Lyapunov argument establishes convergence of the associated ODE. An advantage of selecting an estimation algorithm based on minimization of a relative entropy is that the entropy function becomes a good candidate for the Lyapunov function in the ODE method.

Since the work of Kushner and Clark on stochastic approximation algorithms the field has grown very large. In the more recent book of Benveniste et al. [7], the authors present an analysis of a general stochastic approximation algorithm that uses a potential theory of the underlying Markov chain[3] to bound the cumulative error. The authors are able to relax many of the requirements that earlier analyses placed on the stochastic approximation algorithms, in particular their treatment allows the evolution of the stochastic approximation to be generated by discontinuous functions. This dissertation directly uses the convergence results from Benveniste et al. [7] to establish the main convergence theorems for the combined estimation and control problem.

Adaptive control problems are a natural source of examples for stochastic approximation methods. Ljung [25] applies his ODE method in one of his first applications to the analysis of the asymptotic properties of Aström's self regulating tuner. More recently a number of authors have analyzed systems that are quite similar to the sys-

[3] A good account of the potential theory for Markov chains is the book by Revuz[34]

tem considered in this dissertation. Arapostathis and Marcus analyze an adaptive estimation algorithm for a partially observed Markov chain in [2]. LeGland and Mevel [24, 23] use similar methods to treat more general systems. In a series of papers [15, 16, 1, 17] Fernandéz-Gaucherand *et al.* pose the problem of jointly estimating the parameters for the partially observed Markov chain while executing a control policy that is optimal for an average cost functional in the limit as the parameter estimates converge to the true model. The analysis uses Stochastic Approximation methods to show convergence of the estimates, and an argument similar to that of Shwartz and Makowski [40] to prove convegence of the control to the optimal control for the exact model. The example used is a very simple two state model of a production problem taken from the operations research literature [37] and [48]. A feature common to all the work cited is the extent of the technical difficulties that the authors encountered in their analysis of a seemingly simple problem. In particular, the analysis of underlying Markov chains for the adaptive control problems is complicated, and, at least in the case of Fernandéz-Gaucherand *et al.*, the discontinuous nature of the optimal controls for the discounted problem prevents a straight forward application of the theory developed by Shwartz and Makowski.

The convergence analysis of the combined estimation and control problem in this dissertation adds two important innovations to the approaches of Arapostathis and Marcus [2] and Le Gland and Mevel [23]. Here the analysis combines the control and observation process with the state process to form an enlarged Markov chain. In addition, the requirement for optimality in the control is slightly relaxed, and a close-to-optimal randomized control strategy replaces the optimal strategy. If θ is used to denote an estimate of the hidden Markov model parameters, then the randomized control is a conditional probability distribution $\upsilon_\theta(du|y_{k,k+\Delta}, u_{k,k+\Delta})$. This conditional distribution is computed as a Gibbs distribution using the value function as a Hamiltonian. A temperture parameter in the Gibbs distribution determines the difference in cost between the randomized contol and an optimal control. By using a randomized strategy, the analysis avoids the problems that Fernandéz-Gaucherand *et al.* encountered with discontinuities in the control. The manner in which the control strategy is chosen ensures that the augmented Markov chain consisting of state, output and input processes is irreducible, and that the control strategy depends continuously on θ, the estimate of the hidden Markov model parameters.

The dissertation also introduces an innovative structure for the Markov process that provides the random perturbation in the stochastic approximation formulation. A problem that arises when applying the stochastic approximation framework to the combined estimation and control problem is how to divide the structure of the estimator between the random process that provides the perturbations, and the parameter values that constitute the iterates in the stochastic approximation. The output, state, and input processes provide a minimal set of sufficient statistics for the controlled hidden Markov model, and it is clear that these processes, which combine to form a Markov chain, should be grouped with the random component of the stochastic approximation. Likewise, the iterates of the estimates of the state transition and output matrices should be grouped with the iterates in the stochastic approximation. The problem comes with the recursively defined estimator quantities that comprise the state information in the estimator. These quantities are the recursive estimates of state occupation probabilities and state transition probabilities in the controlled system's underlying Markov chain. The approach that the dissertation takes is to include these quantities with the random component rather than with the iterates of the stochastic approximation algorithm. This decision makes the problem of finding a Lyapunov function for the stochastic approximation problem easier at the expense of making the problem of analysing the cumulative error from the random perturbations harder.

A large portion of the new material in the dissertation deals with the analysis of the Markov process that provides the random perturbations in the stochastic approximation problem. The Markov process factors into a subchain that provides sufficient statistics for the whole process, and factors that evolve on products of probability simplexes by random transformations that are functions of the sub-chain. The structure is identical to that of a random walk, except that the statistics of the random transformations are governed by a Markov chain rather than an iid process. This dissertation coins the phrase "Markov modulated random walk" for this new stochastic structure. The analysis of the culmulative error in the stochastic approximation problem requires a potential theory for the Markov process, and an important contribution of the dissertation is the development of the necessary ergodic theory for the Markov modulated random walks.

The dissertation is structured in seven chapters, the first of which is this introduction. Chapter two introduces the controlled hidden Markov model that is the central object of the combined control and estimation problem, and introduces an algorithm for computing finite-horizon, risk-sensitive controllers for such systems. Chapter three introduces the estimator component of the combined control and estimation algorithm. Chapters four, five and six present the analysis of the convergence properties of the combined control and estimation algorithm. Chapter four reformulates the problem as a stochastic approximation problem, introduces the Markov modulated random walks that provide the structure of the perturbation component of the stochastic approximation formulation, describes the application of the ODE method, and formulates the assumptions that are used by the the ODE method's convergence theorem. Chapter five develops the potential theory that is needed to verify the assumptions that apply to the stochastic approximation problem's Markov chain. Chapter six exihibits a Lyapunov function for the stochastic aproximation problem, and formulates the main convergence results for the combined control and estimation problem. The seventhth chapter provides conclusions and comments about the work.

Chapter 2

Formulation and solution of the risk sensitive control problem

In this dissertation system models use a state chosen from a finite state-space X to characterize a system at a moment in time. The cardinality of the finite state-space is N, and the states are represented by unit vectors $\{e_1, ..., e_N\}$ of $\mathbb{R}^N$. A discrete time stochastic process with values in the state space X models the evolution of the system in time. The underlying probabilistic structure is an abstract probability space $(\Omega, \mathcal{F}, P)$, which is chosen with the understanding that the sigma algebra $\mathcal{F}$ contains all sigma algebras of interest. For example,the state process, $x_l,\ l = 0, 1, 2, ...$, induces a filtration $\mathcal{X}_l$ on $\mathcal{F}$ as follows. For a fixed l, the $\mathcal{X}_l$ is the smallest sigma algebra contained in $\mathcal{F}$ on which the finite products of random variables $(x_{i_1}, \ldots, x_{i_p}, i_p < l)$ are all measurable functions on Ω, and $\mathcal{X}_\infty$ is the smallest sigma algebra generated by $\cup_{l=0}^{\infty} \mathcal{X}_l$. The process x_l is a Markov chain if it satisfies

$$P(x_{l+1} \mid \mathcal{X}_l) = P(x_{l+1} \mid x_l).$$

The state transition matrix for a Markov chain is the matrix A_l, a function of time l, with entries defined by

$$A_{l;ij} = P(x_{l+1} = e_j \mid x_l = e_i). \tag{2.1}$$

At each time l the rows of A_l are probability densities on X. A specification of both a distribution for the random variable x_0, the initial state in the model, and the values for the entries in the transition matrices A_l completely determines a probability measure P on the sigma algebra $\mathcal{X}_\infty$.

A controlled hidden Markov model consists of a finite set of controls U with cardinality P, a finite set of outputs Y with cardinality M, a finite state set X with cardinality N, and the following rules for generating a state process and an output process from an input sequence u_k. A Markov transition matrix A_u is associated with each control u, so that a sequence of controls, u_l, and an initial state, x_0, will generate a Markov chain, x_l, with transition probabilities governed by the transition matrices

$$A_{l;ij} = A_{u_l;ij} \qquad l = 0, 1, \ldots\, . \tag{2.2}$$

The finite set of outputs Y is represented by $\{e_1, \ldots, e_M\}$, the set of unit vectors in $\mathbb{R}^M$. The output sequence y_l is a random process that has values in Y. The distribution for the process y_l is determined by the distribution of the state process, and by the conditional probabilities

$$P(y_l = e_m \mid x_l = e_i) = B_{im}. \tag{2.3}$$

where B_{ij} is a $N \times M$ matrix with rows that are probability densities over Y.

A fundamental assumption that holds throughout the dissertation is that there exists a constant $\rho > 0$ such that

$$\begin{aligned} A_{u;ij} &> \rho \qquad \forall i, j, \text{ such that } 1 \leq i, j \leq N, \text{ and } \forall u \in U \\ B_{im} &> \rho \qquad \forall i, m \text{ such that } 1 \leq i \leq N \text{ and } 1 \leq m \leq M \end{aligned} \tag{2.4}$$

This strong assumption simplifies many of the ergodicity results and convergence proofs that are presented in the dissertation. In particular, the lower bound on the values of A_u ensures that the Markov chain generated by A_{u_l} is both recurrent and irreducible. In many instances weaker assumptions will lead to the same results at the cost of more intricate arguments. Seneta [38] provides a detailed account of recurrence structures and ergodic theorems for finite Markov chains.

Let $\mathcal{M}$ denote the space of probability distributions on U endowed with the weak topology[1], and $\mathcal{M}_\eta$ denote the compact subset of distributions that satisfy $\mu\{u\} \geq \eta$ for all $u \in U$. A randomized control policy of length K is a specification of a sequence of probability distributions $\mu_0, \mu_1, \ldots, \mu_{K-1}$, where μ_l is the distribution of values taken by the random variable u_l.

The output process y_l and the input process u_l generate filtrations $\mathcal{Y}_l$ and $\mathcal{U}_l$ on $\mathcal{F}$; the filtration generated by the combined state and output processes is denoted $\mathcal{G}_l$, and the filtration $\mathcal{O}_l$ is the filtration generated by the sequence of pairs (u_{l-1}, y_l). $\mathcal{O}_l$ can be interpreted as a time-indexed specification of the information that past records of the input and output processes provide about the system. Given some filtration $\mathcal{F}_l$, a deterministic control policy corresponds to the degenerate case when the conditional distributions $P(u_l \mid \mathcal{F}_l)$ have point support, an open loop policy has μ_l a measurable function of the initial condition $\mathcal{X}_0$ for all l, a state feedback policy has μ_l adapted to the filtration $\mathcal{X}_l$, and an output feedback policy has μ_l adapted to the observation filtration $\mathcal{O}_l$.

A control policy $\mu = \mu_0, \mu_1, ..., \mu_{K-1}$ induces a probability distribution on $\mathcal{G}_K$ with

[1]Throughout the dissertation weak topology on probability measures should be understood in the sense of probability theory. From the point of view of analysis this would be a weak* topology. Since U is finite this distinction is not important in this instance, as in either case the topology is the topology induced by the Euclidean norm in $\mathbb{R}^P$

density

$$P^{\mu}(u_{0,K-1}, x_{0,K}, y_{0,K}) = \langle x_K, By_K \rangle \langle x_0, \pi_0 \rangle \prod_{l=0}^{K-1} \langle x_l, A_{u_l} x_{l+1} \rangle \langle x_l, By_l \rangle \langle u_l, \mu_l \rangle. \tag{2.5}$$

The products in the angle brackets are the normal Euclidean inner products, and π_0 is the probability distribution for the random variable x_0.

A risk sensitive control problem is defined on a hidden Markov model by specifying a cost functional with a particular form. Given a running cost, $\phi(x, u)$, which is a function of both the state and the input, and a final cost $\phi_f(x)$, which is a function of the state only, the finite horizon, risk sensitive cost associated with ϕ, ϕ_f, risk γ and horizon K is the functional

$$\mathcal{J}^{\gamma}(\mu) = \mathbf{E}^{\mu} \left[\exp \frac{1}{\gamma} \left(\phi_f(x_K) + \sum_{l=0}^{K-1} \phi(x_l, u_l) \right) \right]. \tag{2.6}$$

The expectation is taken with respect to the distribution P^{μ} in (2.5). If $\mathfrak{M}$ is a class of control policies, then a solution to the risk sensitive control problem is a control policy $\mu_k^* \in \mathfrak{M}$ that minimizes $\mathcal{J}^{\gamma}(\mu)$. The important class of control policies in this dissertation is the class of output feedback policies. A control policy μ_l is a randomized output feedback policy when for any l, and for any $f : U \to \mathbb{R}$, the random variable $\langle \mu_l, f \rangle$ is measurable with respect to the sigma-algebra $\mathcal{O}_l$ in the observation filtration.

A search for an optimal feedback policy encounters an immediate difficulty. Because the incremental cost is expressed as a function of the state process, an application of dynamic programming produces a control policy that is adapted to the state filtration $\mathcal{X}_l$, but not to the observation filtration $\mathcal{O}_l$. This difficulty is resolved by a reformulation of the dynamics of the plant and the associated cost function. The new plant dynamics generate a state process that is adapted to the observation filtration, while maintaining invariance between the two formulations of the cost of an output feedback policy.

The state in the new state process is called an information state, and the corresponding dynamics are called the information state dynamics.

The information state dynamics are generated by taking conditional expectations of the accrued cost up to the present time that are conditioned with respect to the σ-algebras in the output filtration $\mathcal{O}_k$. Manipulation of conditional expectations produces a recursive formula that governs the evolution of the information state dynamics. Manipulation of the conditional expectations is greatly simplified when the σ-algebras generated by the marginal distributions in the observation process $\{y_k\}$ are mutually independent, and independent of the marginal distributions $\{x_k\}$ under the underlying probability measure. Consequently, a new probability measure $P^\dagger$ that meets this requirement is defined on $(\Omega, \mathcal{F})$ by the densities

$$P^\dagger(u_{0,K-1}, x_{0,K}, y_{0,K}) = \frac{1}{M}\langle x_0, \pi_0\rangle \prod_{l=0}^{K-1} \frac{1}{M}\langle x_l, A_{u_l} x_{l+1}\rangle\langle u_l, \mu_l\rangle.$$

The importance of this definition lies in the properties that the process y_k possesses with respect to the conditional probability distribution that is formed by conditioning $P^\dagger$ on the input σ-algebra $\mathcal{U}_K$. With respect to this conditional distribution, the output process y_k is i.i.d. For each k the distribution of y_k is given by $P^\dagger(y_k = e_m) = 1/M$ for $1 \leq m \leq M$. Furthermore, under the measure $P^\dagger$ the distributions for the processes y_k and x_k are mutually independent. The measure P^μ is absolutely continuous with respect to $P^\dagger$ on each sigma algebra $\mathcal{G}_k$, and has Radon Nikodym derivative

$$\left.\frac{dP^\mu}{dP^\dagger}\right|_{\mathcal{G}_k} = \Lambda_K = \prod_{l=0}^{K} M\langle x_l, B y_l\rangle.$$

The risk sensitive cost function with risk γ and horizon K is written in terms of $P^\dagger$ as

$$\mathcal{J}^\gamma(\mu) = \mathbf{E}^\dagger\left[\Lambda_K \exp\frac{1}{\gamma}\left(\phi_f(x_K) + \sum_{l=0}^{K-1}\phi(x_l, u_l)\right)\right].$$

The appropriate notion of state for the risk-sensitive output-feedback control problem is the information state which, at each time k, is a function on the state-space X defined by

$$\sigma_k^\gamma(x) = \mathbf{E}^\dagger \left[I_{\{x_k=x\}} \Lambda_k \exp\left(\frac{1}{\gamma} \sum_{l=0}^{k-1} \phi(x_l, u_l) \right) \mid \mathcal{O}_k \right].$$

Comparison of this expression with the expression for the conditional expectation under a change of measure given in the conditional Bayes theorem of Elliot et al. [11, Theorem 3.2] reveals the information state σ_k to be an un-normalized conditional expectation of the component of the cost that is incurred before time k. The expectation is taken with respect to the P^μ measure, and is conditioned on both the σ-algebra $\mathcal{O}_k$ and the event $x_k = x$.

The finite horizon cost functional is expressed in terms of the information state by the following formula which is derived in Appendix B.

$$\mathcal{J}^\gamma(\mu) = \mathbf{E}^\dagger \left[\langle \sigma_K^\gamma(\cdot), \exp(\phi_f(\cdot)/\gamma) \rangle \right]. \tag{2.7}$$

The initial value of the information state is

$$\sigma_0^\gamma(x) = \langle \pi_0, x \rangle M \langle x, By_0 \rangle, \tag{2.8}$$

and a recursion that describes the evolution of the information state is calculated as follows:

$$\begin{aligned}\sigma_k^\gamma(x'') &= \mathbf{E}^\dagger\left[I_{\{x_k=x''\}}\Lambda_k \exp\left(\frac{1}{\gamma}\sum_{l=0}^{k-1}\phi(x_l,u_l)\right) \mid \mathcal{O}_k\right] \\ &= \mathbf{E}^\dagger\Bigg[I_{\{x_k=x''\}}\exp(\phi(x_{k-1},u_{k-1})/\gamma)M\langle x_k, By_k\rangle \\ &\qquad\qquad \Lambda_{k-1}\exp\left(\frac{1}{\gamma}\sum_{l=0}^{k-2}\phi(x_l,u_l)\right) \mid \mathcal{O}_k\Bigg] \qquad (2.9) \\ &= \sum_x \Bigg(\langle x, A_{u_{k-1}}x''\rangle \exp(\phi(x,u_{k-1})/\gamma)M\langle x'', By_k\rangle \\ &\qquad\qquad \mathbf{E}^\dagger\left[I_{\{x_{k-1}=x\}}\Lambda_{k-1}\exp\left(\frac{1}{\gamma}\sum_{l=0}^{k-2}\phi(x_l,u_l)\right) \mid \mathcal{O}_{k-1}\right]\Bigg)\end{aligned}$$

The last step in the derivation relies on the mutual independence of the processes x_k and y_k under the conditional probability[2]. The need for independence here is the reason for the particular choice of $P^\dagger$. If the information state is represented as a vector on $\mathbb{R}^N$, then the recursion can be written in the form

$$\sigma_k = \Sigma(u_{k-1}, y_k)\sigma_{k-1}, \qquad (2.10)$$

in which the matrix $\Sigma(u,y)$ is given by the formula

$$\Sigma(u,y) = M\,\mathrm{diag}(\langle\cdot, By\rangle)A_u^\top\,\mathrm{diag}(\exp(1/\gamma\,\phi(\cdot,u))) \qquad (2.11)$$

with the understanding that vectors in $\mathbb{R}^N$ are formed by applying the functions $\langle\cdot, By\rangle$ and $\exp(1/\gamma\,\phi(\cdot,u))$ to the N basis vectors $\{e_1,\ldots,e_n\}$.

In summary, Equation (2.10) describes a linear dynamic system with a state process σ_k^γ that is vector valued and $\mathcal{O}_k$ adapted. The cost functional $\mathcal{J}^\gamma(\mu)$ on the input policy μ is expressed as a functional on the information state at the final time σ_K by (2.7). A

[2] Loève's textbook [28] provides a good account of the methods for manipulating conditional expectations

feedback controller on the information state process determines a control policy that is $\mathcal{O}_k$-adapted, and consequently is an output-feedback control for the hidden Markov model formulation of the problem.

If A'_u and B' are estimates for the matrices A and B, and σ_0 is an approximation to the information state at time 0, then

$$\sigma'_k = \Sigma'(u_{k-1}, y_k) \ldots \Sigma'(u_1, y_2)\Sigma'(u_0, y_1)\sigma_0$$

is an estimate of σ_k, the information state at time k. This estimate is only useful when it comes with assurances that the estimate converges as $k \to \infty$, and that the estimators σ_k are regular with respect to variation in the estimates A'_u and B'. The following results establish convergence and regularity of the estimates.

The operators $\Sigma(u, y)$ are examples of positive operators which map the cone $\mathcal{K} = \{\sigma \in \mathbb{R}^N : \sigma_i > 0,\ 1 \leq i \leq N\}$ into itself. The projective pseudo-metric $d(\sigma, \sigma') = \sup_{i,j} \log(\sigma_i \sigma'_j / \sigma_j \sigma'_i)$ is useful when studying ergodicity properties of positive operators on $\mathcal{K}$. The sets $P_\sigma = \{\sigma' \in \mathcal{K} : d(\sigma, \sigma') = 0\}$ form a partition of $\mathcal{K}$, and the quotient space is a complete metric space that is homeomorphic to the $\mathbb{R}^{N-1}$ plane. In general, the quotient space will be identified with the probability simplex $\Theta = \{\sigma \in \mathcal{K} : \sum_i \sigma_i = 1\}$. The following result from Nussbaum [32] relates the projective metric restricted to the probability simplex to the metric induced by the L_1 norm $\|\cdot\|$

Lemma 1. *If $x, y \in \mathcal{K}$ satisfy $\|x\| = \|y\| = 1$, then*

$$\|x - y\| \leq 3(\exp(d(x, y)) - 1)$$

The next two lemmas prove ergodicity and regularity results for the operator Σ. The ergodicity property proved in the first lemma is the weak ergodicity of Seneta [38], the averaging trick used to get the contraction goes back to Markov.

Lemma 2. *For any choice of u and y. The operator $\Sigma(u,y)$ is a strict contraction on the space $\{\sigma \in \mathbb{R}^N : \sigma_i > 0,\ 1 \leq i \leq N\}$ with respect to the projective pseudo-metric $d(\sigma, \sigma') = \sup_{i,j} \log(\sigma_i \sigma'_j / \sigma_j \sigma'_i)$. The operator induced by $\Sigma(u,y)$ on the real projective space Ω is well defined, denote this operator by $\hat{\Sigma}(u,y)$. Finally, the ranges of $\hat{\Sigma}(u,y)$ lie in a compact set.*

Proof. For any u, y, the operator $\Sigma(u,y)$ commutes with homotheties[3].

The operators defined component-wise by

$$\sigma_i \to \langle e_i, By \rangle \sigma_i$$
$$\sigma_i \to \frac{1}{\gamma} \phi(e_i, u) \sigma_i$$

are both isometries with respect to the projective pseudo-metric. That $\sigma \to A_u^\top \sigma$ is a contraction follows from the following averaging argument. For all i,

$$\frac{(A_u \sigma)_i}{(A_u \tau)_i} = \frac{\sum_j A_{u;ji} \sigma_j}{\sum_k A_{u;ki} \tau_k} = \sum_j \frac{A_{u;ji} \tau_j}{\sum_k A_{u;ki} \tau_k} \frac{\sigma_j}{\tau_j}$$

Let $P_{i,j} = \frac{A_{u;ji}\tau_j}{\sum_k A_{u;ki}\tau_k}$, then the entries of $P_{i,j}$ are strictly positive, the rows sum to 1, and

$$\sum_j \frac{A_{u;ji} \tau_j}{\sum_k A_{u;ki} \tau_k} \frac{\sigma_j}{\tau_j} = \sum_j P_{i,j} \frac{\sigma_j}{\tau_j} < \sup_j \frac{\sigma_j}{\tau_j}.$$

Consequently,

$$\sup_i \frac{(A_u \sigma)_i}{(A_u \tau)_i} \leq \rho \sup_j \frac{\sigma_j}{\tau_j}.$$

For some constant ρ that depends on the entries in A_u, but satisfies $0 \leq \rho < 1$. □

Flemming and Hernández-Hernández, [18] write the information state σ as a product $\sigma = r\theta$. The first factor $r = |\sigma|$, is a positive real number. The second factor,

[3] Homotheties are contractions and dilations.

$\theta = \sigma/|\sigma|$, lies on the probability simplex which is given the topology induced by the projective pseudo-metric d defined in Lemma 2. The simplex is identified with the projective plane Θ. This decomposition gives a useful geometric view of the space of information states as a linear fibration with the projective plane Θ as base space and the positive reals $\mathbb{R}$ as fiber. It is a consequence of Lemma 2 that the information state recursion defined by the operator Σ has a decomposition as the product $\hat{\Sigma} \otimes |\Sigma|$. The first factor, $\hat{\Sigma}$, which acts on the base space, is the contraction operator defined in Lemma 2, and $|\Sigma| : \Theta \times \mathbb{R} \to \mathbb{R}$ is the multiplicative operator on the fibre that is defined by $(\theta, r) \to r|\Sigma(\sigma(\theta))|$.

The following lemma establishes regularity of the operator Σ with respect to perturbations in A_u and B.[4]

Lemma 3. *Suppose that the operators* $A_u^\top$, $\mathrm{diag}(\exp(1/\gamma\phi(\cdot, u)))$ *and* B *are all uniformly bounded by the constant* C_1*. If* A' *and* B' *are perturbations of* A *and* B *that satisfy* $|A^\top - A'^\top| < \epsilon$*, and* $|B - B'| < \epsilon$ *then, for any information state* $\sigma \in \mathbb{R}^{+N}$

$$|\Sigma(u, y)\sigma - \Sigma'(u, y)\sigma| < 2\epsilon C_1^2 |\sigma|$$

[4]In the next lemma, and throughout the dissertation multiplicative constants are denoted by the symbol C. Subscripts will differentiate between different constants that appear in a single context, such as the following lemma, but the subscripts will not remain consistent between a result that is presented in one context and used in another.

Proof. Let $\max\{|B - B'|, |A_u - A'_u| : u \in U\} < \epsilon$, then for $\sigma \in \mathbb{R}^N$,

$$\begin{aligned}
|\Sigma(u,y)\sigma - &\Sigma'(u,y)\sigma| \\
&= |\operatorname{diag}(\langle \cdot, By\rangle) A_u^\top \operatorname{diag}(\exp(1/\gamma\phi(\cdot,u)))\sigma \\
&\qquad - \operatorname{diag}(\langle \cdot, B'y\rangle) A_u'^\top \operatorname{diag}(\exp(1/\gamma\phi(\cdot,u)))\sigma| \\
&= |\operatorname{diag}(\langle \cdot, (B - B')y\rangle) A_u^\top \operatorname{diag}(\exp(1/\gamma\phi(\cdot,u)))\sigma \\
&\qquad + \operatorname{diag}(\langle \cdot, B'y\rangle)(A'_u - A_u)^\top \operatorname{diag}(\exp(1/\gamma\phi(\cdot,u)))\sigma| \\
&\le 2\epsilon C_1^2 |\sigma|.
\end{aligned}$$

□

The interpretation of the information state as an un-normalized conditional expectation of the cost accrued up to the present time, conditioned on the values for the present state, would seem to indicate that an estimate for the information state must reflect the entire history of the system. While this is true of the information state as a whole, the fibration $\sigma = r\theta$ separates the distant past from the recent history. The historical values of incurred costs are accrued in the fibre r, and the recent history of the system dynamics is stored in the base point θ. It will turn out that the optimal feedback control will depend only on θ, the projection of the information state into the base-space. So, equipped with this clairvoyance, a buffer containing the last Δ observations of the input and output processes u_l and y_l before the present time k, and estimates A'_u and B' of the system kernels A_u and B, define an estimate of the factor θ_k of information state at time k, by the finite recursion

$$\theta'_{k-\Delta} = (1/N)\,\mathbf{1}$$

$$\theta'_{l+1} = \Sigma'(u_l, y_l)\theta'_l, \qquad k - \Delta \le l \le k - 1.$$

Lemma 2 and Lemma 3 combine to produce the following bound for the error in the estimate.

Proposition 4. *If the buffer length Δ can be chosen arbitrarily large, then the error in the estimate for the information state is bounded by*

$$|\theta_k - \theta'_k| < \epsilon C_2$$

Where C_2 is a constant, and ϵ is a uniform bound on the induced L^1 norms $|A_u^\top - A_u'^\top|$, and $|B - B'|$.

Proof. Define the operators $\hat{\Sigma}_{l_1..l_2}$ and $\hat{\Sigma}'_{l_1..l_2}$ by the products

$$\hat{\Sigma}_{l_1..l_2} = \hat{\Sigma}(u_{l_2}, y_{l_2}) \ldots \hat{\Sigma}(u_{l_1}, y_{l_1})$$
$$\hat{\Sigma}'_{l_1..l_2} = \hat{\Sigma}'(u_{l_2}, y_{l_2}) \ldots \hat{\Sigma}'(u_{l_1}, y_{l_1}),$$

then the following inequalities yield a bound on the error in the estimate.

$$\begin{aligned}
|\theta_k - \theta'_k| &= \hat{\Sigma}_{k-\Delta..k-1}\theta_{k-\Delta} - \hat{\Sigma}'_{k-\Delta..k-1}\theta'_{k-\Delta} \\
&= \sum_{l=k-\Delta}^{k-2} \hat{\Sigma}_{l..k-1}\hat{\Sigma}'_{k-\Delta..l}\theta_{k-\Delta} - \hat{\Sigma}_{l+1..k-1}\hat{\Sigma}'_{k-\Delta..l+1}\theta'_{k-\Delta} \\
&= \sum_{l=k-\Delta}^{k-2} \Big(\hat{\Sigma}_{l..k-1}\hat{\Sigma}'_{k-\Delta..l}\theta_{k-\Delta} - \hat{\Sigma}_{l..k-1}\hat{\Sigma}'_{k-\Delta..l}\theta'_{k-\Delta} \\
&\qquad + \hat{\Sigma}_{l..k-1}\hat{\Sigma}'_{k-\Delta..l}\theta'_{k-\Delta} - \hat{\Sigma}_{l+1..k-1}\hat{\Sigma}'_{k-\Delta..l+1}\theta'_{k-\Delta} \Big) \\
&\leq \Delta\rho^\Delta D + \sum_{l=k-\Delta}^{k-2} \rho^{k-l}\epsilon C_1 D \\
&\leq \Delta\rho^\Delta D + \frac{\rho^2}{1-\rho}\epsilon C_1 D
\end{aligned}$$

D and ρ are fixed constants, consequently, provided C_2 is chosen to satisfy $C_2 > \rho^2 D/(1-\rho)$, the estimate in the statement of the proposition will hold for all sufficiently large Δ. □

The reason for defining an information state is to convert the output feedback problem to an equivalent, fully observed, state feedback problem which is easier to solve.

Baras and James [4] use the following dynamic programming argument to compute optimal deterministic control. They define a time-dependent value function $S^\gamma(\sigma, l)$ by

$$S^\gamma(\sigma, l) = \min_{u_l \ldots u_{K-1} \in U} \mathbf{E}^\dagger \left[\langle \sigma_K^\gamma(\cdot), \phi_f(\cdot) \rangle \mid \sigma_l^\gamma = \sigma \right], \qquad 0 \leq l < K, \tag{2.12}$$

and associate with this value function the time-dependent dynamic programming equation

$$\begin{cases} S^\gamma(\sigma, l) = \min_{u \in U} \mathbf{E}^\dagger \left[S^\gamma(\Sigma^\gamma(u, y_{l+1})\sigma, l+1) \mid \sigma \right] \\ S^\gamma(\sigma, K) = \langle \sigma(\cdot), \phi_f(\cdot) \rangle. \end{cases} \tag{2.13}$$

The following theorem from [4] establishes that the deterministic control policy that at each step chooses the control that minimizes the expectation in (2.12) is an optimal feedback control policy on the information state recursion.

Theorem 5. *The value function* S^γ *defined by (2.12) is the unique solution to the dynamic programming equation (2.13). Conversely, assume that* S^γ *is the solution of the dynamic programming equation (2.13). Suppose that* μ^* *is a policy such that for each* $l = 0, \ldots, k-1$, $u_l^* = \bar{u}_l^*(\sigma_l^\gamma) \in \mathcal{M}$, *where* $\bar{u}_l^*(\sigma)$ *achieves the minimum in (2.13). Then* μ^* *is an optimal output feedback controller for the risk-sensitive stochastic control problem with cost functional (2.6).*

Given a hidden Markov model, and a risk-sensitive cost function, Theorem 5 together with the information state recursion (2.10) provide an algorithm for calculating optimal, deterministic output feedback controllers. A standard objection to dynamic programming methods is that the dynamic programming equations are expensive to solve, especially in the time dependent case. But Fernández-Gaucherand and Marcus [14] show that in the finite horizon, finite state case, the case considered here, the value function is a concave piece-wise linear function of the finite-dimensional information state, and the direct use of dynamic programming to solve the optimal risk-sensitive control problem is feasible for modest sized state-spaces.

An obvious, but erroneous, approach to the problem posed in this dissertation, the problem of combined estimation and control, is to simply compute an optimal, deterministic control with the current estimate of the system model, and update the control policy as the model is updated with successive iterations of the estimation algorithm. This approach faces two problems. The first is the problem of persistent excitation. The parameter estimation algorithm relies on the statistical information that it receives from observations of input and output values to reconstruct the state and output transition matrices. If an optimal control is used, the control algorithm avoids states with high costs, and consequently transitions to and from these states are poorly represented in the statistical information available to the parameter estimation algorithm. The second problem is that the mapping between a system model and the optimal control for that model is often ill-conditioned, a small change in the model can produce a large change in the optimal control. In the case of finite valued deterministic controls it is hard to imagine any sort of non-trivial topology on the control space that would make the optimal control problem well-conditioned. Both these problems are avoided by the use of randomized approximations to the optimal risk-sensitive control, rather than the deterministic optimal control itself. These randomized policies are created by solving a randomized regularization of the dynamic programming equation (2.13), the equation that is satisfied by the optimal, deterministic information-state feedback controller.

For a finite horizon K, recursively define a value function $V_K(\sigma, l)$ depending on the information state σ and the time $0 \leq l \leq K$ as follows:

1. At the final time K the value function is defined by

$$V_K(\sigma, K) = \langle \sigma, \phi_f(\cdot) \rangle$$

2. given the value function at time $l+1$, define an energy functional on the information state σ and the control u by

$$H(\sigma, u, l) = \mathbf{E}_u^{\dagger}[\log V_K(\Sigma(u, y)\sigma, l+1)]$$

3. Require that the randomized feedback control at time l be distributed according to the conditional Gibbs distribution with density given by

$$\upsilon_l(\sigma; du) = Z_l(\sigma)^{-1} \exp\left(\frac{-H(\sigma, u, l)}{\eta}\right) du \tag{2.14}$$

In which the partition function $Z_l(\sigma)$ is defined as

$$Z_l(\sigma) = \int_U \exp\left(\frac{-H(\sigma, u, l)}{\eta}\right) du.$$

4. Define the value function recursion by

$$V_K(\sigma, l) = \int_U \mathbf{E}_u^{\dagger}[V_K(\Sigma(u, y)\sigma, l+1)]\upsilon_l(\sigma; du) \tag{2.15}$$

The following theorem lists basic properties of the control policy that are needed later for the proof of combined estimation and control.

Theorem 6. *The policy υ satisfies the following properties*

(i) The policy is a strictly positive measure on U (needed for ergodicity results later)

(ii) The policy is continuous with respect to changes in the HMM. (Needed later to prove regularity results for potential kernels)

(iii) *In the limit as $\eta \to 0$ the randomized policy converges (in the weak topology) to the finite horizon, deterministic policy of Theorem 5.*

Proof. The proof follows from the definition of the control policy.

(i) This is a consequence of boundedness of the value function.

(ii) For a fixed value of η, the measures $\upsilon_l(\sigma; du)$ are continuous[5] functions of the energy functional H, which in turn is a continuous function of the information state σ. (The partition function $Z_l(\sigma)$ is bounded away from zero.) A backward recursive argument proves that the energy functional is continuous with respect to variations in the hidden Markov model parameters, A_u and B, and Proposition 4 establishes continuity of the the information state with respect to variations in the hidden Markov model.

(iii) This is an application of a simple form of Laplace's approximation theorem.

□

While Theorem 6 is adequate for the purposes of this dissertation, it is not a very satisfactory result from a control theory point of view. An intuitive justification for considering a finite horizon control is that as the length of the horizon K is extended, the control converges to an invariant value. But while it is easy to prove the theorem for a fixed value of K, it is not so easy to produce a proof that provides estimates that are uniform for arbitrarily large values values of K. In effect what is needed is a form of backwards ergodicity result that says that as the length of the horizon is increased, the influence of the choice of final cost ϕ_f on the control policy diminishes. Flemming

[5]Since the space of measures over a finite set of discrete points is a finite dimensional vector space, a precise definition of the metric on the space of measures is not required.

and Hernández-Hernández have proved just such a result for the deterministic case. This result allowed them to define an analogue to the optimal average cost control that occurs as a large horizon limit of the finite horizon policy for quadratic cost functionals. Unfortunately the method that Flemming and Hernández-Hernández use, which relies on the transformation of the optimal control problem to an equivalent dynamic game problem, does not have an obvious analogue for the stochastic regularization presented above. The remainder of this section demonstrates how an adaption of the method that Fleming and Hernández-Hernández use leads to a nonlinear eigenvalue problem formulation for the average cost limit to the regularized problem.

The argument starts by postulating that the average cost problem is well-posed in the sense that in the limit as K becomes large, the solutions $V_K(\sigma, 0)$ are well approximated by

$$V_K(\sigma, 0) \approx r \exp(\rho K + W(\theta)). \tag{2.16}$$

This assumption implies the existence of a fixed point equation that must be satisfied by $W(\theta)$ and ρ if this asymptotic limit is to hold.

When $|\sigma| = 1$, σ can be identified with a canonical injection of a point $\theta \in \Theta$, and in the sequel, when θ is used to denote a point in the information-state space, this injection is understood. Assume that equation (2.16) holds exactly, then for $\theta \in \Theta$,

$$\begin{aligned} V_K(\sigma, 0) &= \exp(\rho K + W(\theta)) \\ V_{K+1}(\sigma, 0) &= \exp(\rho(K+1) + W(\theta)) \end{aligned} \tag{2.17}$$

Let $w(\theta) = \exp(W(\theta))$, $g(u, y, \sigma) = |\Sigma(u, y)\sigma|$ and $G(u, y, \sigma) = \log g(u, y, \sigma)$, then the recursive definition of the value function implies that

$$\begin{aligned} V_{K+1}(\Sigma(u,y)\sigma,1) &= V_K(\Sigma(u,y)\sigma,0) \\ &= V_K(|\Sigma(u,y)\sigma|\hat{\Sigma}(u,y)\theta,0) \\ &= e^{\rho K} g(u,y,\sigma) w(\hat{\Sigma}(u,y)\theta) \end{aligned} \tag{2.18}$$

Equation (2.17) yields a simplified expression for the measure $\upsilon_0(\sigma; du)$. When the horizon is $K+1$,

$$\begin{aligned} \upsilon_0(\sigma; du) &= \frac{\left(\mathbf{E}_u^{\dagger}(V_K(\Sigma(u,y)\sigma, l+1))\right)^{-1/\eta} \; du}{\int_U \left(\mathbf{E}_u^{\dagger}(V_K(\Sigma(u,y)\sigma, l+1))\right)^{-1/\eta} \; du} \\ &= \frac{\left(\mathbf{E}_u^{\dagger}[g(u,y,\sigma) w(\hat{\Sigma}(u,y)\theta)]\right)^{-1/\eta} \; du}{\int_U \left(\mathbf{E}_u^{\dagger}[g(u,y,\sigma) w(\hat{\Sigma}(u,y)\theta)]\right)^{-1/\eta} \; du} \end{aligned} \tag{2.19}$$

Which is independent of K.

Equations (2.17) and (2.18) yield the following implicit equation for the function W and the quantity ρ.

$$e^{\rho} w(\theta) = \int \mathbf{E}_u^{\dagger}[g(u,y,\theta) w(\hat{\Sigma}(u,y)\theta)] \upsilon_0(\theta; du) \tag{2.20}$$

Define the nonlinear, homogeneous of degree 1 operator Λ by

$$[\Lambda w](\theta) = \int \mathbf{E}_u^{\dagger}[g(u,y,\theta) w(\hat{\Sigma}(u,y)\theta)] \upsilon_0(\theta; du). \tag{2.21}$$

Then the fixed point equation (2.20) can be written as the non-linear eigenvalue equation

$$\Lambda w = e^{-\rho} w \tag{2.22}$$

The operator Λ is nonlinear because the argument w appears both explicitly in the integrand on the right hand side of equation (2.21), and implicitly in the measure $\upsilon_0(\theta; du)$.

The search for a well posed solution to the average cost control problem is equivalent to establishing the existence of stable solutions to this non-linear eigenvalue problem. General non-linear eigenvalue problems are notoriously difficult to solve, but the positive, and homogeneous properties of the operator in this case provide hope that this problem is tractable. Ergodicity results for non-linear operators are an area of active research, and authors such as Nussbaum [32] have provided results for families of operators that are similar to the operator in equation (2.22).

Chapter 3

Estimating the parameters for the hidden Markov model

This chapter introduces an algorithm for on-line estimation of the hidden Markov model parameters. An architecture for the combined controller and estimator is illustrated in Figure 3.1. The estimator monitors the system inputs and outputs, estimates, on-line, a hidden Markov model for the plant, and feeds this estimate to the controller algorithm. This controller algorithm uses the estimate to compute the value of the operators (2.11), which it in turn uses in the value function recursion (2.15) that is the basis of the computation of the feedback control policy.

The estimation algorithm and its derivation are taken from Krishnamurthy and Moore [21] with small adaptations that account for the differences between the models treated by Krishnamurthy and Moore and controlled hidden Markov model considered here. The derivation takes the form of a formal stochastic approximation to the maximum likelihood estimator and produces an algorithm that is similar to the standard expectation maximization algorithm. Where the expectation maximization algorithm repeatedly processes the entire data-set, the algorithm derived here uses a shifting window of the data to recursively update filtered estimates of the model parameters.

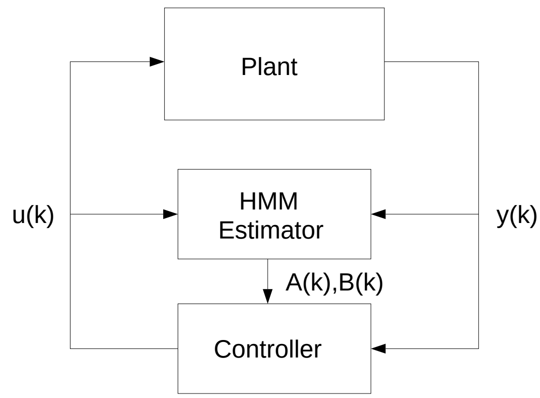

Figure 3.1: Controller Architecture

Let the symbol λ denote an unconstrained choice of parameters for the hidden Markov model. The transition matrices A_u and B are specified completely by the parameters λ, and through them λ determines a probability measure on $(\Omega, \mathcal{F})$. Given an input sequence $u_0 \ldots u_{K-1}$, and a model λ, let $f(y_{0,k} \mid \lambda)$ be the probability distribution function for the sequence of outputs $y_0, \ldots, y_K$. If λ^* is a particular choice of model, then a convexity arguement based on Jensen's inequality proves that the Kullback Leibler measure

$$J(\lambda) = \mathbf{E}[\log f(y_{0,k} \mid \lambda) \mid \lambda^*]$$

has a global maximum at $\lambda = \lambda^*$ provided that equality almost everywhere between $f(y_{0,k} \mid \lambda)$ and $f(y_{0,k} \mid \lambda^*)$ implies that $\lambda = \lambda^*$. Consequently, an algorithm that produces a sequence of estimates λ_k with the property that $J(\lambda_k)$ converges to the global maximum $J(\lambda^*)$ will converge to the maximum likelihood estimate of the model parameters. Unfortunately, given an observation process $y_{0,k}$, it is not easy to obtain recursive estimates for the log-likelihoods $\log f(y_{0,k} \mid \lambda)$, instead, log-likelihoods of the combined output and state processes are used. Specifically, Krishnamurthy and

Moore show that if λ and λ' are two parameter values, and

$$\mathcal{Q}_k(\lambda', \lambda) = \mathbf{E}[\log f(x_{0,k}, y_{0,k} \mid \lambda) \mid y_{0,k}, \lambda']$$
$$\bar{\mathcal{Q}}_k(\lambda', \lambda) = \mathbf{E}[\mathcal{Q}_k(\lambda', \lambda) \mid \lambda^*]$$

then $\bar{\mathcal{Q}}_k(\lambda', \lambda) > \bar{\mathcal{Q}}_k(\lambda', \lambda')$ implies that $J(\lambda) > J(\lambda')$. This observation leads to the following off-line expectation maximization algorithm: Estimate λ_{l+1} by

$$\lambda_{l+1} = \max_{\lambda} \mathcal{Q}(\lambda_l, \lambda).$$

The λ_l are the models at consecutive passes through the data.

A sequential algorithm that gives a sequence of stochastic approximations to the model parameters λ^* using an observation sequence of length K is: Estimate λ_{k+1} by

$$\lambda_{k+1} = \max_{\lambda} \mathcal{Q}_{k+1}(\Lambda_k, \lambda).$$

with

$$\mathcal{Q}_{k+1}(\Lambda_k, \lambda) = \mathbf{E}[\log f(x_{0,k+1}, y_{0,k+1} \mid \lambda) \mid y_{0,K}, \Lambda_k] \tag{3.1}$$

and $\Lambda_k = (\lambda_1, \lambda_2, \ldots, \lambda_k)$. This time λ_k is the estimate of the model based on log-likelihoods of the combined state and output sequence $x_{0,k}, y_{0,k}$, and the probability measure used in evaluating the conditional expectation in (3.1) is the empirical measure generated by the conditional distributions

$$f(x_l = e_j \mid x_{l-1} = e_i) = A_{u_{l-1};i,j}(l \wedge k) \qquad A_u(0) = \hat{A}_u$$
$$f(y_l = e_m \mid x_l = e_i) = B_{i,m}(l \wedge k) \qquad B(0) = \hat{B}$$
$$f(x_0) = \hat{\pi}$$

in which $\hat{A}$, $\hat{B}$ and $\hat{\pi}$ are initial estimates of the state transition matrix, the output matrix and the density of the initial state, and $A(l \wedge k)$ and $B(l \wedge k)$ are the estimates of the state transition matrix and output transition matrix computed either from the parameter estimate λ_l when $l < k$, or from the parameter estimate λ_k when $l \geq k$.

The first two terms of the Taylor's expansion for $\mathcal{Q}_{k+1}(\Lambda_k, \cdot)$ about λ_k are

$$\begin{aligned}\mathcal{Q}_{k+1}(\Lambda_k, \lambda) \approx \mathcal{Q}_{k+1}(\Lambda_k, \lambda_k) + (\lambda - \lambda_k)^\top \left.\frac{\partial \mathcal{Q}_{k+1}(\Lambda_k, \lambda)}{\partial \lambda}\right|_{\lambda=\lambda_k} \\ + \frac{1}{2}(\lambda - \lambda_k)^\top \left.\frac{\partial^2 \mathcal{Q}_{k+1}(\Lambda_k, \lambda)}{\partial \lambda^2}\right|_{\lambda=\lambda_k} (\lambda - \lambda_k).\end{aligned} \tag{3.2}$$

Let λ_{k+1} be the value of λ that maximizes the right hand side of (3.2), then

$$\lambda_{k+1} = \lambda_k + I_{k+1}^{-1}(\lambda_k) S(\lambda_k, y_{k+1}) \tag{3.3}$$

in which I_{k+1}, the Fisher information matrix for the combined output and state processes, is given by $I_{k+1}(\lambda_k) = -\partial^2 \mathcal{Q}_{k+1}/\partial\lambda^2|_{\lambda=\lambda_k}$, and $S(\lambda_k, y_{k+1})$, the score at time k is given by $S(\lambda_k, y_{k+1}) = \partial \mathcal{Q}_{k+1}/\partial\lambda|_{\lambda=\lambda_k}$. If $\lambda_{k+1} - \lambda_k$ is small, then λ_{k+1} is a close approximation to the value of λ that maximizes $\mathcal{Q}_{k+1}(\Lambda_k, \lambda)$.

To proceed requires an explicit expression for $\mathcal{Q}$ and its derivatives in terms of the parameter estimates λ_l. If A_u and B are the matrices associated with a particular choice of λ, and π is the estimate of the initial state distribution, then the probability density for the state and output process is

$$\begin{aligned} f(y_{0,k+1}, x_{0,k+1} \mid \lambda) &= f(y_{k+1}, x_{k+1} \mid y_{0,k}, x_{0,k}, \lambda) f(y_{0,k}, x_{0,k} \mid \lambda) \\ &= \langle x_{k+1}, B y_{k+1} \rangle \langle x_k, A_{u_k} x_{k+1} \rangle f(y_{0,k}, x_{0,k} \mid \lambda) \\ &= \langle x_0, B y_0 \rangle \langle \pi, x_0 \rangle \prod_{l=0}^{k} \langle x_{l+1}, B y_{l+1} \rangle \langle x_l, A_{u_l} x_{l+1} \rangle \end{aligned}$$

Taking logarithms gives[1]

$$\log f(y_{0,k+1}, x_{0,k+1} \mid \lambda) = \sum_i \sum_j \sum_u n^u_{i,j}(k+1) \log A_{u;i,j} + \sum_i \sum_m m_{i,m}(k+1) \log B_{i,m} + \sum_i \delta_{e_i}(x_0) \log \pi_i,$$

in which $n^u_{i,j}(k)$ is the number of transitions from $x_{l-1} = e_i$, $u_{l-1} = u$ to $x_l = e_j$ in the state sequence $x_{0,k}$, and $m_{i,m}(k)$ is the number of times $x_l = e_i$ and $y_l = e_m$ in the combined sequences $x_{0,k}, y_{0,k}$. Taking conditional expectations with respect to the probability measure induced by the sequence of model estimates Λ_l gives

$$\begin{aligned} &\mathbf{E}[\log f(y_{0,k+1}, x_{0,k+1} \mid \lambda) \mid y_{0,K}, \Lambda_k] \\ &\quad = \sum_i \mathbf{E}[\delta(x_0 = e_i) \mid y_{0,K}, \Lambda_k] \log \pi_i \\ &\qquad + \sum_i \sum_j \sum_u \log A_{u;i,j} \mathbf{E}[n^u_{i,j}(k+1) \mid y_{0,K}, \Lambda_k] \\ &\qquad + \sum_i \sum_m \log B_{i,m} \mathbf{E}[m_{i,m}(k+1) \mid y_{0,K}, \Lambda_k]. \end{aligned} \tag{3.4}$$

Define the conditional densities $\zeta_{l|K,\Lambda_k}$ and $\gamma_{l|K,\Lambda_k}$ by

$$\zeta_{l|K,\Lambda_k}(i,j) = f(x_l = e_j, x_{l-1} = e_i \mid y_{0,K}, \Lambda_k)$$
$$\gamma_{l|K,\Lambda_k}(i) = f(x_l = e_i \mid y_{0,K}, \Lambda_k),$$

then in terms of ζ and γ the conditional expectations on the right hand side of equation

[1] $\delta_a(x)$ is used throughout the dissertation to denote the Dirac delta, i.e. the distribution on the space in which x takes values, with unit mass supported at the point a.

(3.4) are:

$$\begin{aligned}\mathbf{E}[n^u_{i,j}(k) \mid y_{0,K}, \Lambda_k] &= \sum_{l=1}^{k} f(x_l = e_j, x_{l-1} = e_i \mid y_{0,K}, \Lambda_k)\delta_u(u_{l-1}) \\ &= \sum_{l=0}^{k} \zeta_{l|K,\Lambda_k}(i,j)\delta_u(u_{l-1}) \\ \mathbf{E}[m_{i,m}(k) \mid y_{0,K}, \Lambda_k] &= \sum_{l=0}^{k} f(x_l = e_i \mid y_{0,K}, \Lambda_k)\delta_{e_m}(y_l) \\ &= \sum_{l=0}^{k} \gamma_{l|K,\Lambda_k}(i)\delta_{e_m}(y_l) \\ \mathbf{E}[\delta_{e_i}(x_0) \mid y_{0,K}, \Lambda_k] &= f(x_0 = e_i \mid y_{0,K}, \Lambda_k) = \gamma_{0|K,\Lambda_k}(i)\end{aligned}$$

Substituting the above expressions for the conditional expectations into the right hand side of equation (3.4) gives

$$\begin{aligned}\mathcal{Q}_{k+1}(\Lambda_k, \lambda) = \sum_i \gamma_{0|K,\Lambda_k}(i) \log \pi_i + \sum_i \sum_j \sum_{l=1}^{k+1} \zeta_{l|K,\Lambda_k}(i,j) \log\langle e_i, A_{u_{l-1}} e_j\rangle \\ + \sum_i \sum_{l=0}^{k+1} \gamma_{l|K,\Lambda_k}(i) \log\langle e_i, B y_l\rangle. \qquad (3.5)\end{aligned}$$

Equation (3.5) expresses $\mathcal{Q}$ in a suitable form for the calculation of the gradient term and the Fisher information matrix that appear in the update equation — equation (3.3). Working from (3.5), $\mathcal{Q}_{k+1}(\Lambda_k, \lambda)$ is written in terms of $\mathcal{Q}_k(\Lambda_k, \lambda)$ as

$$\begin{aligned}\mathcal{Q}_{k+1}(\Lambda_k, \lambda) &= \mathcal{Q}_k(\Lambda_k, \lambda) \\ &+ \sum_i \sum_j \zeta_{k+1|K,\Lambda_k}(i,j) \log\langle e_i, A_{u_1} e_j\rangle + \sum_i \gamma_{k+1|K,\Lambda_k}(i) \log\langle e_i, B y_{k+1}\rangle.\end{aligned}$$

and differentiation with respect to λ gives the score vector:

$$\begin{aligned} S(\lambda_k, y_{k+1}) &= \left.\frac{\partial \mathcal{Q}_{k+1}(\Lambda_k, \lambda)}{\partial \lambda}\right|_{\lambda=\lambda_k} \\ &= \left.\frac{\partial \mathcal{Q}_k(\Lambda_k, \lambda)}{\partial \lambda}\right|_{\lambda=\lambda_k} + \sum_i \sum_j \zeta_{k+1|K,\Lambda_k}(i,j) \left.\frac{\partial \log\langle e_i, A_{u_k} e_j\rangle}{\partial \lambda}\right|_{\lambda=\lambda_k} \\ &\quad + \sum_i \gamma_{k+1|K,\Lambda_k}(i) \left.\frac{\partial \log\langle e_i, B y_{k+1}\rangle}{\partial \lambda}\right|_{\lambda=\lambda_k} . \end{aligned}$$

Under the assumptions that $\mathcal{Q}_k(\Lambda_k, \lambda) \approx \mathcal{Q}_k(\Lambda_{k-1}, \lambda)$, and that λ_k is close to the minimizer of $\mathcal{Q}_k(\lambda_{k-1}, \lambda)$ the first term is close to zero, and the score is closely approximated by

$$\begin{aligned} S(\lambda_k, y_{k+1}) \approx \sum_i \sum_j \zeta_{k+1|K,\Lambda_k}(i,j) \left.\frac{\partial \log\langle e_i, A_{u_k} e_j\rangle}{\partial \lambda}\right|_{\lambda=\lambda_k} \\ + \sum_i \gamma_{k+1|K,\Lambda_k}(i) \left.\frac{\partial \log\langle e_i, B y_{k+1}\rangle}{\partial \lambda}\right|_{\lambda=\lambda_k} . \end{aligned} \quad (3.6)$$

The choice of entries for the matrices A_u and B is constrained by the Markov conditions that $\sum_j A_{u;ij} = 1$ and $\sum_m B_{i,m} = 1$. Consequently, the model (A_u, B) can be expressed in terms of a vector λ with dimension $P \times N \times (N-1) + N \times (M-1)$. The components of λ are indexed as $\lambda = (\lambda^{A_u}_{i,j}, \lambda^B_{i,m})$, and the form of $\mathcal{Q}_{k+1}(\Lambda_k, \lambda)$ ensures that the gradient and Hessian have corresponding structure. For a given choice of u and i the index j in $\lambda^{A_u}_{i,j}$ skips one value q_i in the range $1 \leq q_i \leq N$, similarly, the index m in $\lambda^B_{i,m}$ skips one value p_i in the range $1 \leq p_i \leq M$. The matrix entries $A_{u;ij}$ and B_{im} are expressed in terms of the parameters λ^{A_u} and λ^B by the formulae

$$\begin{aligned} A_{u;ij} &= \lambda^{A_u}_{i,j} \qquad j \neq q_i & B_{im} &= \lambda^B_{i,m} \qquad m \neq p_i \\ A_{u;iq_i} &= 1 - \sum_{\substack{j=1 \\ j \neq q_i}} \lambda^{A_u}_{i,j} & B_{ip_i} &= 1 - \sum_{\substack{m=1 \\ m \neq p_i}} \lambda^B_{i,m}, \end{aligned} \quad (3.7)$$

First derivatives with respect to λ are

$$\frac{\partial}{\partial \lambda^{A_u}_{i,j}} = \frac{\partial}{\partial A_{u;i,j}} - \frac{\partial}{\partial A_{u;i,q_i}} \qquad \frac{\partial}{\partial \lambda^B_{i,m}} = \frac{\partial}{\partial B_{i,m}} - \frac{\partial}{\partial B_{i,p_i}}$$

and the second derivatives are

$$\frac{\partial^2}{\partial {\lambda_{i,j}^{A_u}}^2} = \frac{\partial^2}{\partial {A_{u;i,j}}^2} + \frac{\partial^2}{\partial {A_{u;i,q_i}}^2} \qquad \frac{\partial^2}{\partial \lambda_{i,j}^{A_u} \partial \lambda_{i,j'}^{A_u}} = \frac{\partial^2}{\partial {A_{u;i,q_i}}^2} \qquad j \neq j'$$

$$\frac{\partial^2}{\partial {\lambda_{i,m}^{B}}^2} = \frac{\partial^2}{\partial {B_{i,m}}^2} + \frac{\partial^2}{\partial {B_{i,p_i}}^2} \qquad \frac{\partial^2}{\partial \lambda_{i,m}^{B} \partial \lambda_{i,m'}^{B}} = \frac{\partial^2}{\partial {B_{i,p_i}}^2} \qquad m \neq m'$$

The parameters $A_{u;i,j}$ and $B_{i,m}$ occur in separate terms of the expression for $\mathcal{Q}_{k+1}(\Lambda_k, \lambda)$. As a result, the form of λ induces corresponding block structures on the score vector $S(\lambda_k, y_{k+1})$ and the Fisher information matrix $I_{k+1}(\lambda_k)$. Evaluating the partial derivatives in the score vector gives:

$$S(\lambda, y_{k+1}) = [S^{A_u}(\lambda, y_{k+1}), S^B(\lambda, y_{k+1})]^\top \tag{3.8}$$

with one S^{A_u} block for each input $u \in U$. Each block S^{A_u} is given by

$$S^{A_u}(\lambda, y_{k+1}) = [S^{A_u}(1), S^{A_u}(2), \ldots, S^{A_u}(N)]^\top \tag{3.9}$$

$$S^{A_u}(i) = [s^{A_u}(i,1), \ldots, s^{A_u}(i, q_i - 1), s^{A_u}(i, q_i + 1), \ldots, s^{A_u}(i, N)]^\top \tag{3.10}$$

with

$$s^{A_u}(i,j) = g_{i,j} - g_{i,q_i} \qquad j \neq q_i \tag{3.11}$$

$$g_{i,j} = \frac{\zeta_{k+1|K,\Lambda_k}(i,j)}{A_{u;i,j}} \delta_u(u_k). \tag{3.12}$$

Likewise, the block S^B is given by

$$S^B(\lambda, y_{k+1}) = [S^B(1), S^B(2), \ldots, S^B(N)]^\top \tag{3.13}$$

$$S^B(i) = [s^B(i,1), \ldots, s^B(i, p_i - 1), s^B(i, p_i + 1), \ldots, s^B(i, M)]^\top \tag{3.14}$$

with

$$s^B(i,m) = h_{i,m} - h_{i,p_i} \qquad m \neq p_i \tag{3.15}$$

$$h_{i,m} = \frac{\gamma_{k+1|K,\Lambda_k}(i)}{B_{i,m}} \delta(y_{k+1} = f_m). \tag{3.16}$$

The delta functions that occur in the score vector come from the inner products in summands of equation (3.6). At each time k the output y_k selects one column of the matrix B to be updated, and the input u_k selects one transition matrix A_{u_k}. If $y_k = e_m$ the first term in equation (3.15) is non-zero, if $y_k = e_{p_i}$ the second term is selected, and if y_k is something else, then $s^B(i, m) = 0$ at that particular value of k.

The Fisher information matrix I_{k+1} has a block diagonal structure

$$I_{k+1}(\lambda) = \text{diag}(I_{k+1}^{A_u}(\lambda), I_{k+1}^{B}(\lambda)) \tag{3.17}$$

again, there is one I^{A_u} block for each $u \in U$. Each block I^{A_u} has the form

$$I_{k+1}^{A_u}(\lambda) = \text{diag}(P_1, \ldots, P_N),$$

with

$$\begin{aligned} P_i &= M_i + C_i, \\ M_i &= \text{diag}(\mu_{i,j}), \qquad 1 < j < N; \quad j \neq q_i, \end{aligned} \tag{3.18}$$

and

$$\mu_{i,j} = \left(\frac{\sum_{l=1}^{k+1} \zeta_{l|K,\Lambda_k}(i,j)\delta_u(u_{l-1})}{{A_{u;i,j}}^2} \right). \tag{3.19}$$

For each i, C_i is an $(N-1) \times (N-1)$ matrix with identical elements all equal to

$$c_i = \left(\frac{\sum_{l=1}^{k+1} \zeta_{l|K,\Lambda_k}(i,q_i)\delta_u(u_{l-1})}{{A_{u;i,q_i}}^2} \right)^{1/2}. \tag{3.20}$$

A standard matrix inversion Lemma [20, p. 655] produces expressions for the inverses of the blocks P_i,

$$P_i^{-1} = M_i^{-1} - \left(\frac{1}{{c_i}^2} + \sum_{j \neq q_i}^{N} \frac{1}{\mu_{i,j}} \right)^{-1} F_i F_i^{\top} \tag{3.21}$$

in which F_i is an $N-1$ dimensional column vector with entries ${\mu_{i,j}}^{-1}$, $j \neq q_i$. These inverses are used to construct the blocks ${I_{k+1}^{A_u}}^{-1}$ in the inverse of the Fisher information matrix.

An identical derivation for an expression for the block ${I_{k+1}^{B}}^{-1}$ of the inverse of the Fisher information matrix in terms of blocks Q_i^{-1} results in

$$
{I_{k+1}^{B}}^{-1} = \text{diag}(Q_i^{-1}) \tag{3.22}
$$

$$
Q_i^{-1} = N_i^{-1} - \left(\frac{1}{{d_i}^2} + \sum_{m \neq p_i}^{M} \frac{1}{\nu_{i,m}} \right)^{-1} E_i E_i^{\top} \tag{3.23}
$$

$$
\nu_{i,m} = \left(\frac{\sum_{l=1}^{k+1} \gamma_{l|K,\Lambda_k}(i) \delta_{f_m}(y_l)}{{B_{i,m}}^2} \right) \tag{3.24}
$$

$$
d_i = \left(\frac{\sum_{l=1}^{k+1} \gamma_{l|K,\Lambda_k}(i) \delta_{f_{p_i}}(y_l)}{{B_{i,p_i}}^2} \right)^{1/2} \tag{3.25}
$$

with E_i an $M-1$ dimensional column vector with entries ${\nu_{i,m}}^{-1}$, $m \neq p_i$.

Substituting the explicit expressions for the score vectors (3.8 – 3.16), and the Fisher information matrices (3.17 – 3.25) into the parameter update equation (3.3) produces a formula for λ_{k+1} in terms of λ_k and the conditional densities $\zeta_{k+1|K,\Lambda_k}$, and $\gamma_{k+1|K,\Lambda_k}$. Substituting this estimate in (3.7) gives the following formulae for the updates of the transition matrix estimate.

$$
\begin{aligned}
A_{u;ij}(k+1) &= A_{u;ij}(k) + \frac{1}{\mu_{i,j}} \left(g_{i,j} - \frac{\sum_{r=1}^{N} g_{i,r}/\mu_{i,r}}{\sum_{r=1}^{N} {\mu_{i,r}}^{-1}} \right), \\
&\qquad\qquad\qquad\qquad i,j = 1 \ldots N; \\
B_{im}(k+1) &= B_{im}(k) + \frac{1}{\nu_{i,m}} \left(h_{i,m} - \frac{\sum_{r=1}^{M} h_{i,r}/\nu_{i,r}}{\sum_{r=1}^{M} {\nu_{i,r}}^{-1}} \right), \\
&\qquad\qquad i = 1 \ldots N, \quad m = 1 \ldots M;
\end{aligned} \tag{3.26}
$$

Equations (3.26) along with (3.12, 3.15, 3.19, 3.24) give estimates for the hidden Markov model parameters in terms of the conditional densities $\zeta_{k+1|K,\Lambda_k}$ and $\gamma_{k+1|K,\Lambda_k}$

and reduce the problem of deriving recursive estimators for the hidden Markov model parameters to the problem of finding recursive formulae for these densities. Moore and Krishnamurthy tackle this problem by applying methods similar to those used in the backwards-forwards recursions of the standard EM algorithm described by Rabiner [33].

Two auxiliary distributions

$$\alpha_{l|\Lambda_k}(i) = f(x_l = e_i \mid y_{0,l}, \Lambda_k)$$
$$\beta_{l|K,\Lambda_k}(j) = f(y_{l+1,K} \mid x_{l+1} = e_j, \Lambda_k)$$

are introduced. $\alpha_{l|\Lambda_k}(j)$, which is the conditional density of the state at time l conditioned on the prior observations, is computed by the recursive formula,

$$\alpha_{l|\Lambda_k}(j) = \frac{\sum_i \langle e_j, B(l \wedge k) y_l \rangle A_{u_{l-1};ij}(l \wedge k) \alpha_{l-1|\Lambda_k}(i)}{\sum_j \sum_i \langle e_j, B(l \wedge k) y_l \rangle A_{u_{l-1};ij}(l \wedge k) \alpha_{l-1|\Lambda_k}(i)}. \tag{3.27}$$

The recursion is initialized with

$$\alpha_{0|\Lambda_k}(j) = \frac{\langle e_j, B(0) y_0 \rangle \pi_j}{\sum_j \langle e_j, B(0) y_0 \rangle \pi_j}.$$

$\beta_{l|K,\Lambda_k}(i)$ is computed with the backwards recursion

$$\beta_{l|K,\Lambda_k}(i) = \sum_j \beta_{l+1|K,\Lambda_k}(j) A_{u_{l+1};ij}((l+2) \wedge k) \langle e_i, B((l+1) \wedge k) y_{l+1} \rangle \tag{3.28}$$

which is initialized with $\beta_{K|K,\Lambda_k}(i) = 1$. The densities $\zeta_{l|K,\Lambda_k}$ and $\gamma_{l|K,\Lambda_k}$ are expressed in terms of α and β by the formulae

$$\begin{aligned} \zeta_{l|K,\Lambda_k}(i,j) &= f(x_l = e_j, x_{l-1} = e_i \mid y_{0,K}, \Lambda_k) \\ &= \frac{\alpha_{l-1,\Lambda_k}(i) A_{u_{l-1};ij}(l \wedge k) \beta_{l-1|K,\Lambda_k}(j)}{\sum_{i,j} \alpha_{l-1,\Lambda_{l-1}}(i) A_{u_{l-1};ij}(l \wedge k) \beta_{l-1|K,\Lambda_k}(j)} \end{aligned} \tag{3.29}$$

$$\gamma_{l|K,\Lambda_k}(i) = \frac{\sum_j \beta_{l,K|\Lambda_k}(j) A_{u_l;ij}((l+1) \wedge k) \alpha_{l,|\Lambda_k}(i)}{\sum_i \sum_j \beta_{l,K|\Lambda_k}(j) A_{u_l;ij}((l+1) \wedge k) \alpha_{l,|\Lambda_k}(i)}. \tag{3.30}$$

Details of the derivations of these formulae are given in Appendix A.

The derivation given in this chapter closely follows the development of the basic recursive estimation algorithm described by Krishnamurthy and Moore [21]. In the same paper the authors make two changes to the algorithm that improve its implementation. The first change is to use a fixed length buffer of "future" observations to calculate the a-posteriori probabilities β. Instead of computing $\beta_{l|K,\Lambda_k}$, the a-posteriori probability for the entire observation sequence form l to K, a buffer of Δ observations is used, and $\beta_{l|K,\Lambda_k}$ is replaced by the a-posteriori probabilities $\beta_{l|l+\Delta,\Lambda_k}$ which are calculated from the next Δ observations. The effect of this change is to produce a fixed-lag recursive estimator. The use of the interval Δ for both the fixed-lag in the estimator and the buffer length for the information state estimate in the controller is an intentional simplification in notation. Although the collusion of the intervals does not affect the asymptotic properties of the algorithm, in practice the lengths need not, and probably should not, be identical.

The second change that Krishnamurthy and Moore propose is the incorporation of a "forgetting factor" in the computation of the Fisher information matrices. Specifically, the expressions defining μ, c, ν, and d in equations (3.19, 3.20, 3.24, 3.25) and are replaced by

$$\mu_{i,j} = \left(\frac{\sum_{l=1}^{k+1} \rho^{k-l+1} \zeta_{l|K,\Lambda_k}(i,j) \delta(u_{l-1} = u)}{{A_{u;i,j}}^2} \right)$$

$$c_i = \left(\frac{\sum_{l=1}^{k+1} \rho^{k-l+1} \zeta_{l|K,\Lambda_k}(i,q_i) \delta(u_{l-1} = u)}{{A_{u;i,q_i}}^2} \right)^{1/2}$$

$$\nu_{i,m} = \left(\frac{\sum_{l=1}^{k+1} \rho^{k-l+1} \gamma_{l|K,\Lambda_k}(i) \delta(y_l = f_m)}{{B_{i,m}}^2} \right)$$

$$d_i = \left(\frac{\sum_{l=1}^{k+1} \rho^{k-l+1} \gamma_{l|K,\Lambda_k}(i) \delta(y_l = f_{p_i})}{{B_{i,p_i}}^2} \right)^{1/2}$$

Another way to implement "exponential forgetting" is the following. Let κ be a fixed positive integer, and write the update equation (3.26) for the transition matrix estimate $A_{u;ij}(k)$ as

$$\begin{aligned} A_{u;ij}(k+1) &= A_{u;ij}(k) + \frac{1}{\mu_{i,j}} \left(g_{i,j} - \frac{\sum_{r=1}^{N} g_{i,r}/\mu_{i,r}}{\sum_{r=1}^{N} {\mu_{i,r}}^{-1}} \right) \\ &= A_{u;ij}(k) + \frac{1}{k+\kappa} \frac{1}{\mu_{i,j}/(k+\kappa)} \left(g_{i,j} - \frac{\sum_{r=1}^{N} g_{i,r}/(\mu_{i,r}/(k+\kappa))}{\sum_{r=1}^{N} \left(\mu_{i,r}/(k+\kappa)\right)^{-1}} \right). \end{aligned} \tag{3.31}$$

When $k >> \kappa$, the quantity $\mu_{i,r}/(k+\kappa)$ is an empirical estimate of the probability of occurrence of the transition $u_k = u, x_k = e_i \to x_{k+1} = e_r$. An alternative way to estimate this probability is with the recursive estimator

$$Z_{k+1}^{u} = Z_{k}^{u} + (1-\rho)\delta_u(u_k)(\zeta_{k|k+\Delta,\Lambda_k} - Z_k^u), \tag{3.32}$$

the matrix element $Z_{k+1}^{u}(i,r)$ gives the frequency estimate that corresponds to $\mu_{i,r}/(k+\kappa)$. Substitute $Z_{k+1}^{u}(i,r)$ for $\mu_{i,r}/(k+\kappa)$ in the second line of (3.31) to give a new recursive update for the transition matrix estimates

$$A_{u;ij}(k+1) = A_{u;ij}(k) + \frac{1}{k+\kappa} \frac{1}{Z_k^u(i,j)} \left(g_{i,j} - \frac{\sum_{r=1}^{N} g_{i,r}/(Z_k^u(i,r))}{\sum_{r=1}^{N} \left(Z_k^u(i,r)\right)^{-1}} \right). \tag{3.33}$$

The corresponding update equation for the estimates $B_{im}(k)$ is

$$B_{im}(k+1) = B_{im}(k) + \frac{1}{k+\kappa} \frac{1}{\Gamma_k(i,m)} \left(h_{i,m} - \frac{\sum_{r=1}^{M} h_{i,r}/\Gamma_k(i,r)}{\sum_{r=1}^{M} \left(\Gamma_k(i,r)\right)^{-1}} \right), \tag{3.34}$$

with Γ_k being the recursively defined quantity

$$\Gamma_{k+1} = \Gamma_k + (1-\rho)(\gamma_{k|k+\Delta,\Lambda_k} y_k^{\top} - \Gamma_k) \tag{3.35}$$

Chapter 4

Analysis of the control and estimation algorithm.

Part 1: Stochastic approximation formulation.

The next three chapters present an analysis of the combined control and estimation algorithm as a stochastic approximation algorithm. The first section of this chapter reformulates the combined control and estimation problem in new compact notation that emphasis the division of the system structure between a Markov process and the parameter trajectory of a stochastic difference equation. The section introduces a new stochastic structure called a Markov modulated random walk, and provides a decomposition of the Markov process in terms of these structures and an underlying finite Markov chain. The second section in the chapter provides an overview of the ODE method, and formulates the assumptions from the premise of the convergence theorems in terms of the specific structure of the combined estimation and control problem. Verification of three of the seven assumtions presented in this chapter requires nontrivial analysis. The next two chapters present this analysis.

4.1 Decomposition of the Markov chain

The analysis of the combined control and estimation algorithm proceeds by the identification and exploitation of structure in the complex Markov process that is associated with the algorithm. The state-space of this Markov process has two levels of factorization. At each level the Markov process is expressed as a product of an embedded Markov process that provides a sufficient statistic for the large process, and a dependent structure, a Markov modulated random walk, that is defined as follows.

Consider a discrete-time Markov process $\{T_k\}$ that takes values in a state space $\mathcal{T} = \mathcal{S} \times \mathfrak{S}$ that is a product of a set $\mathcal{S}$ and a semigroup $\mathfrak{S}$. Assume that the set $\mathcal{S}$ and the semigroup $\mathfrak{S}$ each have measurable structures that are compatible with the sigma algebra on $\mathcal{T}$. Let T_k have components $T_k = (S_k, s_k)$, where $\{S_k\}$ is a Markov process in $\mathcal{S}$ with transition kernel $\Pi(S_k; dS_{k+1})$. If there exists a measurable function $g : \mathcal{S} \to \mathfrak{S}$ such that for all k, $s_k = g(S_k)s_{k-1}$, then the random process $\{s_k\}$ is the Markov modulated random walk generated by the Markov process $\{S_k\}$, and the map g.

The transition kernel for the process $\{T_k\}$ is $\Pi(T_a; dT_b) = \delta_{g(S_b)s_a}(s_b)\Pi(S_a; dS_b)$, where $\Pi(S_a; dS_b)$ is the transition kernel for the Markov process S_k. The meaning of this notation becomes clearer when the action of the kernel on a function is considered. If $f : \mathcal{T} \to \mathbb{R}$ is a measurable function, then

$$\mathbf{E}[f(T_b) \mid T_a] = \int f(T_b)\Pi(T_a; dT_b) = \int f(S_b, g(S_b)s_a)\Pi(S_a; dS_b)$$

If the map induced by the function g on the space of measures (the push-forward of g) is written as $d_\star g$, then $d\mu_b = [d_\star g](\Pi(S_a; dS_b))$ is a sequence of time dependent probability kernels on $\mathfrak{S}$, and the marginal distribution for s_{k+1} conditioned on $T_k =$

(S_k, s_k) is given by the kernel[1] $\Pi(S_a, s_a; ds_b) = \delta_{s_a} * d\mu_b(s_b)$. In the special case in which the kernel $\Pi(S_a; dS_b) = \mathbf{1}(S_a)d\nu(S_b)$ is independent of S_a, the push-forward $d\mu_b = [d_\star g](d\nu) = d\mu$ is a constant measure, and the Markov random walk s_k is a conventional random walk (with independent increments) on the semigroup $\mathfrak{S}$.

Turning now to the chain that underlies the combined control estimation problem, it is time for some new notation. Let $k \in \mathbb{N}$ be the instant in time after the controller has read the value of the k'th output y_k, but before the k'th input u_k is computed. Define the following random variables:

$X_k^x = x_k$ is the state of the controlled hidden Markov model defined in Section 2. X_k^x takes values in a finite set of size N which is represented by the N canonical basis vectors in $\mathbb{R}^N$.

$X_k^u = u_{k-\Delta,k-1}$ is a buffer containing the last Δ values for the control u. The control takes values in a finite set of size P, and the values of X_k^u are represented by elements of the set formed by taking Cartesian products of length Δ of copies of the set of canonical basis vectors in $\mathbb{R}^P$.

$X_k^y = y_{k-\Delta+1,k}$ is a buffer containing the last Δ values for the output (including the k'th value). The output takes values in a finite set of size M, and the values of X_k^y are represented by elements of the set formed by taking Cartesian products of length Δ of copies of the set of canonical basis vectors in $\mathbb{R}^M$.

$X_k^\alpha = \alpha_{k-\Delta,k-\Delta+1}$ is buffer of length two containing the values for the time-lagged empirical density for the state calculated at times k and $k-1$. Denote the probability simplex over the state by Ω^α, then X^α takes values in the Cartesian product

[1]The convolution of two measures μ and ν defined on a semigroup $\mathfrak{S}$ is the measure defined by the formula $\mu * \nu(f) = \int_{\mathfrak{S}\times\mathfrak{S}} f(gh)\, d\mu(g)d\nu(h)$

$\Omega^\alpha \times \Omega^\alpha$.

X_k^ζ is the empirical estimate of the density for the conditional joint distribution of successive states in the Markov chain, conditioned on the value of the input at the transition. Let Ω^ζ denote the probability simplex in $\mathbb{R}^{N^2}$, then the densities of joint distributions of successive states take values in Ω^ζ, and X_k^ζ takes values in the Cartesian product of P copies of Ω^ζ.

X_k^γ is the empirical estimate of the density for the distribution of the state of the Markov chain conditioned on the associated output. X_k^γ takes values in the Cartesian product of M copies of Ω^α.

Throughout the remainder of the dissertation n-dimensional probability simplices Ω are identified with subsets of real projective space $\mathbb{P}^n$. This identification simplifies formulas such as those for the filter equations by removing the need to keep track of normalizing constants. Conversely, if $X \in \mathbb{P}^n$, then an expression such as $\min_i X^i$ should be interpreted as applying to a representation of X as a point in the probability simplex Ω.

Let $\tilde{X}_k = (X_k^x, X_k^u, X_k^y)$ be the product of the discrete random variables, and let $X_k = (X_k^x, X_k^u, X_k^y, X_k^\alpha, X_k^\zeta, X_k^\gamma)$ be the product of all the random variables. The evolution of the discrete time random process X_k captures the combined dynamics of the controlled hidden Markov model that forms the plant, the control algorithm, and the estimator. The random variables X^u and X^y which represent buffers of length Δ are used both in both the estimator and the control algorithms. There is no reason, other than convenience, why the lag in the estimator algorithm, which is used to smooth the a-posteriori estimates of the state occupation and state transition, should be the same length as the buffer that is used by the moving horizon controller to compute the current value for the information state.

The range spaces for the random variables will be denoted by the corresponding script symbols, so, for example, $\tilde{\mathcal{X}}$ denotes the finite set of values taken by the process $\tilde{X}_k$, $\mathcal{X}^\alpha = \Omega^\alpha \times \Omega^\alpha$ is the continuous range space for the random variable X^α, and $\mathcal{X}$, the range space for the complete process X_k, is a complicated space formed from a finite number of disconnected continuous components.

The following proposition summarizes the information available about the evolution, and the proof provides a detailed decomposition of the structure in the transition kernel.

Proposition 7. *The random process $\{X_k\}$ is Markov, the chain $\{\tilde{X}_k\}$ is a Markov sub-chain, and for all l, the random variable X_l, together with the chain $\{\tilde{X}_k,\ k \geq l\}$ form a set of sufficient statistics for the process $\{X_k,\ k \geq l\}$.*

Proof. The proof proceeds by using the formulae from previous chapters to write explicit expressions for the Markov transition kernels $\Pi_\theta(\tilde{X}; d\tilde{X})$ and $\Pi_\theta(X; dX)$.

The dynamics of the discrete chain is determined from the description of the controlled hidden Markov model and the control algorithm in Chapter 2. Recalling the definitions of X^x, X^u, and X^y, $\tilde{X}_a \in \tilde{\mathcal{X}}$ has a representation as a tensor product of canonical basis vectors

$$\tilde{X}_a = x_a \otimes y_a^0 \otimes u_a^{-1} \otimes y_a^{-1} \otimes u_a^{-2} \otimes \ldots \otimes y_a^{-\Delta+1} \otimes u_a^{-\Delta}, \tag{4.1}$$

so if $\tilde{X}_k = \tilde{X}_a$ then $x_a = x_k$, $y_a^{-i} = y_{k-i}$ and $u_a^{-i} = u_{k-i}$. The factor $\tilde{X}_a$ can only take a finite number of values. The distribution of the sequence of discrete random variables $\tilde{X}_k$ is determined by a transition kernel $\Pi_\theta(X_a; d\tilde{X}_b)$ that depends only on the discrete part of the random variable X_a. This transition kernel can be written as a large, sparse, stochastic matrix M of rank $N(MP)^\Delta$. If the rows and columns of

M are indexed by the finite set of values that form the range of $\tilde{X}$, then a formula for the entry $M_{\tilde{X}_b}^{\tilde{X}_a}$ can be determined from the definitions of the controlled Markov chain transition matrices given in equations (2.1), (2.2), (2.3), and from the randomized, moving-horizon, output-feedback control strategy υ_θ.

$$M_{\tilde{X}_a}^{\tilde{X}_b} = \langle x_a, A_{u_b^{-1}} x_b \rangle \langle x_a, B y_b^0 \rangle \langle \upsilon_\theta(y_a^0, u_a^{-1}, \ldots, y_a^{-\Delta+1}, u_a^{-\Delta}), u_b^{-1} \rangle \\ \times \delta_{y_a^0}(y_b^{-1}) \delta_{u_a^{-1}}(u_b^{-2}) \ldots \delta_{y_a^{-\Delta+2}}(y_b^{-\Delta+1}) \delta_{u_a^{-\Delta+1}}(u_b^{-\Delta}) \quad (4.2)$$

That M is stochastic follows from A_u and B being stochastic, and from the fact that υ takes vector values that are probability densities over the finite input set U. The matrix M depends on the parameter θ solely through the feedback strategy υ_θ.

The evolution of the continuous factors of X is determined by the estimation algorithm presented in Chapter 3. Write $X^\alpha = (X^{\alpha,1}, X^{\alpha,2})$, $X^\zeta = (X^{\zeta,1}, \ldots, X^{\zeta,P})$ and $X^\gamma = (X^{\gamma,1}, \ldots, X^{\gamma,M})$ so $X^{\alpha,i}$ and $X^{\gamma,m}$ take values in the probability simplex Ω^α and $X^{\zeta,p}$ takes values in the probability simplex Ω^ζ. Let $\mathfrak{X}^\alpha = \Omega^\alpha \times \Omega^\alpha$, $\mathfrak{X}^\zeta = \Omega^\zeta \times \overset{P}{\cdots} \times \Omega^\zeta$ and $\mathfrak{X}^\gamma = \Omega^\alpha \times \overset{M}{\cdots} \times \Omega^\alpha$. For $S \in \{\mathfrak{X}^\alpha, \mathfrak{X}^\zeta, \mathfrak{X}^\gamma\}$, let $\mathfrak{S}(S)$ denote the semigroup of (not necessarily invertible) affine transformations of S into itself. The evolution of the random processes X^α, X^ζ and X^γ are described directly in terms of the action of elements of Markov modulated random walks on the affine spaces $\mathfrak{X}^\alpha$, $\mathfrak{X}^\zeta$, and, $\mathfrak{X}^\gamma$. Define the maps $g^\alpha : \tilde{\mathfrak{X}} \to \mathfrak{S}(\mathfrak{X}^\alpha)$, $g^\zeta : \tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \to \mathfrak{S}(\mathfrak{X}^\zeta)$ and $g^\gamma : \tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \to \mathfrak{S}(\mathfrak{X}^\gamma)$ by

$$(g^\alpha(\tilde{X}) X^\alpha)^1 = \text{diag}(B(\theta) y^{-\Delta+1}) A_{u^{-\Delta}}^\top(\theta) X^{\alpha,1}, \qquad (g^\alpha(\tilde{X}) X^\alpha)^2 = X^{\alpha,1} \quad (4.3)$$

$$(g^\zeta(\tilde{X}, X^\alpha) X^\zeta)^p = (1 - q\delta_{e_p}(u^{-\Delta})) X^{\zeta,p} + q\delta_{e_p}(u^{-\Delta})\zeta \quad (4.4)$$

$$(g^\gamma(\tilde{X}, X^\alpha) X^\gamma)^m = (1 - q\delta_{e_m}(y^{-\Delta})) X^{\gamma,m} + q\delta_{e_m}(y^{-\Delta})\gamma \quad (4.5)$$

with ζ and γ defined by

$$\zeta(i,j) = \left. \frac{X_i^{\alpha,2} A_{u;ij} \beta_j^{-\Delta}}{\sum_{i,j} X_i^{\alpha,2} A_{u;ij} \beta_j^{-\Delta}} \right|_{u=u^{-\Delta}} \tag{4.6}$$

$$\gamma(i) = \left. \frac{\sum_j \beta^{-\Delta+1}(j) A_{u;ij} X_i^{\alpha,1}}{\sum_i \sum_j \beta^{-\Delta+1}(j) A_{u;ij} X_i^{\alpha,1}} \right|_{u=u^{-\Delta+1}}, \tag{4.7}$$

and β^{-l} calculated from the buffered input and output values X^u and X^y by the backward recursion

$$\beta^{-(l+1),i} = \sum_j \beta^{-l,j} A_{u;ij} B_{i,m} \Big|_{u=u^{-l}, e_m = y^{-l}} \qquad \beta^0 = \mathbf{1} \tag{4.8}$$

The map g^α comes directly from the recursion (3.27), the equations for β^{-l}, ζ and γ come from (3.28 – 3.30), and the mappings g^ζ and g^γ come from equations (3.32) and (3.35), the recursive formulae for empirical estimates of the conditional densities X^ζ and X^γ. The ranges of these maps all lie in semigroups of affine transformations on the appropriate probability simplices.

The mappings defined in (4.3–4.5) give rise to Markov modulated random walks $(\tilde{X}_k, s_k^\alpha)$, $(\tilde{X}_k, X_k^\alpha, s_k^\zeta)$, and $(\tilde{X}_k, X_k^\alpha, s_k^\gamma)$ through the definitions,

$$s_0^\alpha = \mathbf{I}_{\mathcal{X}^\alpha} \qquad s_k^\alpha = g^\alpha(\tilde{X}_k) s_{k-1}^\alpha$$

$$s_0^\zeta = \mathbf{I}_{\mathcal{X}^\zeta} \qquad s_k^\zeta = g^\zeta(\tilde{X}_k, X_k^\alpha) s_{k-1}^\zeta$$

$$s_0^\gamma = \mathbf{I}_{\mathcal{X}^\gamma} \qquad s_k^\gamma = g^\gamma(\tilde{X}_k, X_k^\alpha) s_{k-1}^\gamma$$

and the random processes X^α, X^ζ and X^γ obey evolution equations that are determined by the actions of transformations which are randomly drawn from the corresponding semi-groups s^α, s^ζ and s^γ.

$$X_{l+1}^\alpha = s_{l+1}^\alpha X_0^\alpha \qquad X_{l+1}^\zeta = s_{l+1}^\zeta X_0^\zeta \qquad X_{l+1}^\gamma = s_{l+1}^\zeta X_0^\gamma.$$

Equations (4.2) and (4.3–4.5) combined with the formula for the transition kernel of a Markov modulated random walk yield the following explicit expression for the kernel

$\Pi_\theta(X_a; dX_b)$.

$$\Pi_\theta(X_a; dX_b) = M_{\tilde{X}_b}^{\tilde{X}_a} \delta_{X_a^{\alpha,1}}(X_b^{\alpha,2}) \delta_{g^\alpha(\tilde{X}_b) X_a^{\alpha,1}}(X_b^{\alpha,1}) \times \prod_p \delta_{g^{\zeta,p}(\tilde{X}_b, X_b^\alpha) X_a^{\zeta,p}}(X_b^{\zeta,p}) \prod_m \delta_{g^{\gamma,m}(\tilde{X}_b, X_b^\alpha) X_a^{\gamma,m}}(X_b^{\gamma,m}) \quad (4.9)$$

The claims that the chain $\{X_k\}$ is Markov and that the discrete sub-chain with the initial state forms a set of sufficient statistics both follow from the form of the kernel (4.9). □

The following definitions for the measures in the convolution kernels complete the notational definitions for the Markov modulated random walks that were introduced in Proposition 7.

$$\mu_k^\alpha = [d_* g^\alpha](\Pi_\theta(\tilde{X}_{k-1}; d\tilde{X}_k))$$
$$\mu_k^\zeta = [d_* g^\zeta](\Pi_\theta(\tilde{X}_{k-1}, X_{k-1}^\alpha; d(\tilde{X}_k, X_k^\alpha)))$$
$$\mu_k^\gamma = [d_* g^\gamma](\Pi_\theta(\tilde{X}_{k-1}, X_{k-1}^\alpha; d(\tilde{X}_k, X_k^\alpha)))$$

Finally, let $\mathcal{S}$ is an affine space, and $\mathfrak{S}(\mathcal{S})$ is the semigroup of affine transformations of $\mathcal{S}$ into itself. The following definitions map random walks defined on the semigroup $\mathfrak{S}(\mathcal{S})$, to random walks in the affine space $\mathcal{S}$. The orbit of a point $S_0 \in \mathcal{S}$ under the action of the semigroup $\mathfrak{S}$ is described by the continuous map

$$o_{S_0} : \mathfrak{S}(\mathcal{S}) \to \mathcal{S}$$
$$o_{S_0}(g) = gS_0$$

where the product gS_0 is the image of the point S_0 under the transformation g. The orbit mapping o_{S_0} induces a continuous mapping $d_\star o_{S_0}$ from the space of measures on the semi-group to the space of measures on the affine space, and the sequence of conditional distributions $\Pi^k(g_0; dg_k)$ map to a sequence of conditional distributions $\Pi^k(S_0; dS_k)$ through the equations

$$\Pi(S_0; dS_k) = d_\star o_{S_0}\, \Pi^k(g_0; dg_k).$$

For a more concrete picture of the situation, recall that objects in the spaces $\mathfrak{X}^\alpha$, $\mathfrak{X}^\zeta$, and $\mathfrak{X}^\gamma$ are empirical estimates of distributions associated with the discrete chain $\tilde{X}_k$. The random walks on the semigroups $\mathfrak{S}(\mathfrak{X}^\alpha)$, $\mathfrak{S}(\mathfrak{X}^\zeta)$, and $\mathfrak{S}(\mathfrak{X}^\gamma)$ induce stochastic processes on the corresponding spaces of empirical estimates. These stochastic processes are the sequences of empirical estimates that are generated by the recursive estimation algortihm.

4.2 A stochastic approximation formulation of the estimation algorithm

The notation that was introduced in the preceding section for the description of the underlying Markov chain permits a compact representation of the estimation algorithm in the form of a general stochastic approximation algorithm. Let κ be a positive integer, and let $\gamma_k = k + \kappa$. Using the notation of Benveniste *et al.*, the parameter update equation becomes

$$\theta_{k+1} = \theta_k + \gamma_{k+1} H(\theta_k, X_k). \tag{4.10}$$

Given initial values $X_0 = X$ and $\theta_0 = \theta$ for the state of the Markov chain, and the value of the parameter estimates, the recursion (4.10) and the transition kernel (4.9) define a distribution $P_{X,\theta}$ for the chain (θ_k, X_k). The particular form of $H(\theta_k, X_k)$ ensures that the values θ_k, which are a sequence of estimates for the entries in the matrices A_u and B, satisfy the constraints required of probability kernels.

Equation (4.10) summarizes the two recursive equations (3.33, 3.34). When written in the new notation of this chapter, the part of $H(\theta_k, X_k)$ that updates the estimates of the state transition matrices A_u is:

$$\delta_u(u^{-\Delta}) \frac{\frac{A^2_{u;ij}}{X^\zeta_k(i,j)} \left(\sum_{r=1}^N \frac{A^2_{u;ir}}{X^\zeta_k(i,r)} \left(\frac{\zeta(i,j)}{A_{u;ij}} - \frac{\zeta(i,r)}{A_{u;ir}} \right) \right)}{\sum_{r=1}^N \frac{A^2_{u;ir}}{X^\zeta_k(i,r)}} \tag{4.11}$$

while the part that updates of the estimates for the output transition matrix B is

$$\delta_{f_r}(y^{-\Delta+1}) \frac{\frac{B^2_{im}}{X^\gamma_k(i,m)} \left(\sum_{r=1}^M \frac{B^2_{ir}}{X^\gamma_k(i,r)} \left(\frac{\gamma(i)}{B_{im}} - \frac{\gamma(i)}{B_{ir}} \right) \right)}{\sum_{r=1}^M \frac{B^2_{ir}}{X^\gamma_k(i,r)}} \tag{4.12}$$

in which $u^{-\Delta}$ and $y^{-\Delta+1}$ are the buffered values of input and output, and ζ and γ are given by equations (4.6) and (4.7).

Examination of equations (4.11) and (4.12) indicates that, for fixed $X \in \mathcal{X}$, the function $\theta \rightsquigarrow H(\theta, X)$ is uniformly bounded on $\theta \in \Theta$. If, on the other hand, θ is a fixed point in Θ, then the function $X \rightsquigarrow H(\theta, X)$ blows up as $X \to \partial\mathcal{X}$. In particular, as $X^\zeta(i,j) \to 0$, the terms in (4.11) that update the estimate $A_{u;ij}$ have growth $O(1/X^\zeta(i,j))$. Similarly, as $X^\gamma(i,m) \to 0$, the terms in (4.12) that update the estimate B_{im} have growth $O(1/X^\zeta(i,j))$.

Equation (4.10) is in the form of the general stochastic approximation algorithm considered by Benveniste *et al.* [7, Part2, equation 1.1.1]

$$\theta_{k+1} = \theta_k + \gamma_{k+1} H(\theta_k, X_k) + \gamma^2_{k+1} \rho_k(\theta_k, X_k). \tag{4.13}$$

In the particular case of equation (4.10), the function ρ is identically zero. Part 2 of the book by Benveniste *et al.* [7] presents a general analysis of the convergence properties of this algorithm, and eventually this dissertation will use theorems from that work to establish the convergence of the combined estimation and control algorithm. The premises of the major theorems in Part 2 of Benveniste *et al.* rely on a set of seven non-trivial assumptions about the random process X_k, the generator function $H(\theta, X)$, and the sequence of step-sizes γ_k. A major objective of the work in this dissertation is to establish a theoretical framework for the combined estimation and control problem that supports the assumptions on which the stochastic approximation results of Benveniste *et al.* rely. The three most problematic assumptions that Benveniste *et al.* require form the conclusions of the theorems presented in the next couple of chapters. The remainder of this chapter outlines the major ideas in the stochastic approximation theory presented by Benveniste *et al.* and introduces the theory's seven basic assumptions in the contexts in which Benveniste *et al.* use them to advance the theory.

The fit between the general stochastic approximation theory of Benveniste *et al.* and the convergence problem presented in this dissertation is not perfect. Whereas Benveniste *et al.* allow the random processes X_l and θ_l to evolve in Euclidean spaces, here $x_l \in \mathcal{X}$ and $\theta_l \in \Theta$, where $\mathcal{X}$ and Θ are both Cartesian products of compact subsets of Euclidean spaces. The difference in the range of θ is not very important. Uniform boundedness of the function $H(\theta, X)$ with respect to θ means that the natural embedding of Θ into an appropriately dimensioned Euclidean space provides a suitable metric structure for the stochastic approximation theory. The case for $\mathcal{X}$ is different, $H(\theta, X)$ blows up as $X \to \partial\mathcal{X}$ at an asymptotic rate that is inversely proportional to the Euclidean distance between X and the boundary $\partial\mathcal{X}$. Under the natural embedding of $\mathcal{X}$ into a Euclidean space, the function $H(\theta, X)$ is neither locally Lipschitz with respect to X, nor is it bounded by a function of polynomial growth, and both deficiencies create problems for the stochastic approximation theory. The solution is to choose a metric for the space $\mathcal{X}$ that better suits the requirements of the stochastic approximation analysis. An appropriate metric for $\mathcal{X}$ is a hyperbolic metric that effectively puts the boundary $\partial\mathcal{X}$ at an infinite distance from points in the interior of $\mathcal{X}$.

Specifically, the following definitions define a suitable metric $\mathfrak{d}_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ for the space $\mathcal{X}$.

$$\mathfrak{d}_{\mathcal{X}}(X_a, X_b) = \begin{cases} \infty & \tilde{X}_a \neq \tilde{X}_b \\ \max\{\mathfrak{d}_\alpha(X_a^\alpha, X_b^\alpha), \mathfrak{d}_\zeta(X_a^\zeta, X_b^\zeta), \mathfrak{d}_\gamma(X_a^\gamma, X_b^\gamma)\} & \tilde{X}_a = \tilde{X}_b \end{cases}$$

where

$$\mathfrak{d}_\alpha(X_a^\alpha, X_b^\alpha) = \max\left\{\log\sup_{i,i'}\left(\frac{X_a^{\alpha,1}(i)}{X_b^{\alpha,1}(i)}\frac{X_b^{\alpha,1}(i')}{X_a^{\alpha,1}(i')}\right), \log\sup_{i,i'}\left(\frac{X_a^{\alpha,2}(i)}{X_b^{\alpha,2}(i)}\frac{X_b^{\alpha,2}(i')}{X_a^{\alpha,2}(i')}\right)\right\}$$

$$\mathfrak{d}_\zeta(X_a^\zeta, X_b^\zeta) = \max_{1\le p\le P}\left\{\log\sup_{i,j,i',j'}\left(\frac{X_a^{\zeta,p}(ij)}{X_b^{\zeta,p}(ij)}\frac{X_b^{\zeta,p}(i'j')}{X_a^{\zeta,p}(i'j')}\right)\right\}$$

$$\mathfrak{d}_\gamma(X_a^\gamma, X_b^\gamma) = \max_{1\le m\le M}\left\{\log\sup_{i,i'}\left(\frac{X_a^{\gamma,m}(i)}{X_b^{\gamma,m}(i)}\frac{X_b^{\gamma,m}(i')}{X_a^{\gamma,m}(i')}\right)\right\}$$

The selection of a hyperbolic metric on $\mathcal{X}$ affects the forms of the statements in the theory that use growth bounds or moment conditions. Where, in [7], Benveniste *et al.* use a growth bound of the form $|f(X)| < C(1 + |X|^s)$ for some non-negative integer s, the appropriate bound here will be an expression of the form $|f(X)| < M\beta_s(X)$ with $\beta_s(X)$ defined as follows. For $X^\alpha \in \mathcal{X}^\alpha$, $X^\zeta \in \mathcal{X}^\zeta$ and $X^\gamma \in \mathcal{X}^\gamma$ define

$$\beta_s(X^\alpha) = \max_{s_1+s_2=s}\left\{\sup_i |X^{\alpha,1}(i)|^{-s_1} \times \sup_i |X^{\alpha,2}(i)|^{-s_2}\right\}$$

$$\beta_s(X^\zeta) = \max_{\sum_p s_p=s}\left\{\prod_p \sup_{i,j} |X^{\zeta,p}(i,j)|^{-s_p}\right\}$$

$$\beta_s(X^\gamma) = \max_{\sum_m s_m=s}\left\{\prod_m \sup_i |X^{\gamma,m}(i)|^{-s_m}\right\},$$

and if $X = (\tilde{X}, X^\alpha, X^\zeta, X^\gamma) \in \mathcal{X}$, let

$$\beta_s(X) = \max_{s_\alpha+s_\zeta+s_\gamma=s}\left\{\beta_{s_\alpha}(X^\alpha)\beta_{s_\zeta}(X^\zeta)\beta_{s_\gamma}(X^\gamma)\right\}. \tag{4.14}$$

Accordingly, where Benveniste *et al.* use a moment condition of the form

$$\int \Pi^k(X_0; dX_k)|X_k|^q < C,$$

the appropriate moment condition here will be

$$\int \Pi^k(X_0; dX_k)\beta_q(X_k) < C.$$

Much of the ergodic theory presented in the next chapter relies on notions of weak convergence of measures, and it is important to develop the theory in the context of an appropriate space of test functions. A key requirement for the stochastic approximation theory is the existence of a potential theory for the Markov transition kernel on a suitable function space. A suitable function space, in this context is one that includes the family of generators $H(\theta, \cdot)$ which are considered as functions of the state-space $\mathcal{X}$. The potential theory rests on the weak ergodicity theory for the chain X_l, and the choice of a space of test functions in the ergodicity theory determines the function space in which the potential theory is applicable.

The setting that Benveniste *et al.* use is the following. Markov chain states are represented by variables $z = (x, e) \in \mathbb{R}^k \times E$ with E a finite set. Given g, a function on $\mathbb{R}^k \times E$, and a constant $p \geq 0$, they define

$$\begin{aligned}
\|g\|_{\infty,p} &= \sup_{x,e} \frac{|g(x,e)|}{1+|x|^p} \\
[g]_p &= \sup_{x_1 \neq x_2, e \in E} \frac{|g(x_1,e) - g(x_2,e)|}{|x_1 - x_2|(1+|x_1|^p + |x_2|^p)} \\
\mathit{Li}(p) &= \{g : [g]_p < +\infty\} \\
N_p(g) &= \sup\{\|g\|_{\infty,p+1}, [g]_p\}
\end{aligned}$$

The space of test functions that is used to develop the ergodic theory is the space of Borel functions on $\mathbb{R}^k \times E$ that are bounded, and Lipschitz, and the space of functions for which the potential theory is developed is the space $Li(Q, L_1, L_2, p_1, p_2)$ which is defined as follows.

Let Q be a compact subset of $\mathbb{R}^d$, and, given a function $f(\theta, x, e)$ on $\mathbb{R}^d \times \mathbb{R}^k \times E$, let f_θ denote the function $(x, e) \rightarrow f(\theta, x, e)$. A function $f(\theta, x, e)$ is of the class $Li(Q, L_1, L_2, p_1, p_2)$ if

(i) for all $\theta \in Q$, $N_{p_1}(f_\theta) \leq L_1$

(ii) for all $\theta_1, \theta_2 \in Q$, all $(x, e) \in \mathbb{R}^k \times E$,

$$|f(\theta_1, x, e) - f(\theta_2, x, e)| \leq L_2|\theta_1 - \theta_2|(1 + |x|^{p_2}),$$

and $Li(Q)$ denotes the set of functions f which belong to $Li(Q, L_1, L_2, p_1, p_2)$ for some values of L_1, L_2, p_1 and p_2.

The important features of the spaces $Li(Q, L_1, L_2, p_1, p_2)$ are that weighted integrals of the functions with respect to finite measures are always bounded, and that functions in the space display a uniform regularity with respect to a family of weighted Lipschitz semi-norms. Appropriate definitions for the context of a hyperbolic geometry on the space $\mathcal{X}$ are:

$$\begin{aligned}
\|g\|_{\infty,p} &= \sup_X \frac{|g(X)|}{\beta_p(X)} \\
[g]_p &= \sup_{\substack{X_1 \neq X_2 \\ \tilde{X}_1 = \tilde{X}_2}} \frac{|g(X_1) - g(X_2)|}{\mathfrak{d}(X_1, X_2)(\beta_p(X_1) + \beta_p(X_2))} \\
Li(p) &= \{g : [g]_p < +\infty\} \\
N_p(g) &= \sup\{\|g\|_{\infty,p+1}, [g]_p\}
\end{aligned}$$

The definition of the space $Li(Q, L_1, L_2, p_1, p_2)$ is almost the same as the definition that Benveniste *et al.* use with the Euclidean metrics, only condition (ii) becomes

(ii) for all $\theta_1, \theta_2 \in Q$, all $X \in \mathfrak{X}$,

$$|f(\theta_1, X) - f(\theta_2, X)| \leq L_2 |\theta_1 - \theta_2| \beta_{p_2}(X),$$

The development of the stochastic approximation theory in Benveniste *et al.* [7] relies only on the metric properties of the Euclidean metric, not on the specific form. As a consequence, the extension from the Euclidean theory to the more general metric theory requires nothing more than the substitutions of appropriate metrics, moments, and function seminorms.

Benveniste *et al.* employ the ODE method first used by Ljung [25] to analyze the asymptotic properties of the algorithm. This method associates the random sequence of successive parameter estimates, θ_k, with a piece-wise constant path $\theta(t)$, $t \geq 0$. The mapping between the index k and the time parameter t is determined by the step sizes γ_k, and relies on the assumption:

Assumption 1. *[7, p. 213, A.1] $(\gamma_l)_{l \in \mathbb{N}}$ is a decreasing sequence (in the broad sense) of positive real numbers such that $\sum_l \gamma_l = +\infty$.*

This is certainly true of the sequence $\gamma_l = 1/\kappa + l$ which is decreasing, and has divergent partial sums. The path $\theta(t)$ is parameterized as follows. An increasing sequence of times t_k is defined by $t_0 = 0$, and the sequence of partial sums $t_k = \sum_{l=1}^{k} \gamma_l$. A partial inverse mapping $m : \mathbb{R} \to \mathbb{N}$ is defined on the interval $[0, \infty)$ by $m(t) = \sup\{k : k \leq t\}$, and the path $\theta(t)$ is defined by $\theta(t) = \theta_{m(t)}$.

The analysis is broken down into two logical tasks. The first task is to show that for any length of time T, the piecewise constant trajectories $\theta(t)$ approximate solutions of an associated ordinary differential equation

$$d/dt\, \bar{\theta}(t) = h(\bar{\theta}(t)), \quad t \geq t_0, \quad \bar{\theta}(t_0) = a_0 \tag{4.15}$$

on the interval $[t_0, t_0 + T]$, and that the error in the approximation can be made arbitrarily small provided t_0 is chosen large enough. The second task is to show that equation (4.15) has a stable equilibrium solution at $\bar{\theta}(t) = \theta_*$ where θ_* is the true value of the parameter. The successful completion of these two tasks permits the inference that the recursion (4.10) converges to the true parameter at least when initialized in a neighborhood of the true value.

In practice, the division between the tasks is not so precise. Benveniste *et al.* use a Lyapunov technique to demonstrate stability of the ODE. So, rather than proving convergence of the iterates of the stochastic formulae to trajectories, they show that an evaluation of the Lyapunov function on the iterates of the approximation algorithm θ_k, produces a sequence in $\mathbb{R}$ that converges to a sequence formed by evaluating the Lyapunov function at points $\bar{\theta}(t_k)$ sampled from the trajectory of the ODE. In other words, the authors show that for any length of time T, and for t_0 sufficiently large, the values of the Lyapunov function and its first two derivatives evaluated on the trajectory $\theta(t)$ stay arbitrarily close to values evaluated on corresponding points of $\bar{\theta}(t)$ on the interval $[t_0, t_0 + T]$. With an assumption of uniformly bounded derivatives for the Lyapunov function, an application of the mean value theorem provides the means of dealing with this added complexity.

Let $\bar{\theta}(t; , t_0, a_0)$ denote the solution of equation (4.15) with initial condition $\bar{\theta}(t_0) = a_0$, then the set of equations:

$$\begin{gathered} \bar{\theta}(t_0) = \bar{\theta}_0 = a_0 \qquad \bar{\theta}(t) = \bar{\theta}_{m(t)} \\ \bar{\theta}_k = \bar{\theta}_{k-1} + \gamma_{k+1} h(\bar{\theta}_k) \end{gathered} \tag{4.16}$$

provides a variable step-size, Euler approximation to the trajectories $\bar{\theta}(t; , t_0, a_0)$. The accumulated error between the Euler scheme and the random sequence after K steps is

$$e_K = \sum_{k=1}^{K-1} \epsilon_k \tag{4.17}$$

with

$$\begin{aligned}\epsilon_k &= \theta_{k+1} - \theta_k - \gamma_{k+1} h(\theta_k) \\ &= \gamma_{k+1}(H(\theta_l, X_k) - h(\theta_k)).\end{aligned} \tag{4.18}$$

In light of equations (4.17) and (4.18), it is possible to rephrase the application of the ODE method as the problem of first selecting a generator h, for a stable ODE, and then bounding the mean of the square of the accumulated conditional error $\mathbf{E}[e_K^2|X_0, \theta_0]$ on a sufficiently large set of trajectories θ_k. The first step in the solution is to assume that on D, an open subset of Θ, the random quantity $H(\theta, X)$ has the form of a stable generator $h(\theta)$ with added noise that can be eliminated by averaging $H(\theta, X)$ with respect to the random variable X. The problem of bounding the mean square cumulative error is then a problem of analyzing time averages of the error process ϵ_k. This analysis requires careful restriction of the possible form of the probability distribution of the error process e_K. The next two assumptions from Benveniste *et al.* restrict the form of the probability law for e_K indirectly through restrictions on the stochastic structure of the process X_k and the growth in X of the function $H(\theta, X)$.

Assumption 2. *[7, p. 213, A.2] There exists a family* $\{\Pi_\theta : \theta \in \mathbb{R}^d\}$ *of transition probabilities* $\Pi_\theta(x, A)$ *on* $\mathbb{R}^k$ *such that, for any Borel subset* A *of* $\mathbb{R}^k$*, we have*

$$P[X_{n+1} \in A \mid F_n] = \Pi_{\theta_n}(X_n, A).$$

Assumption 3. *[7, p. 216, A.3] For any Compact subset* Q *of* D*, there exist constants*

C_1, C_2, q_1, *and* q_2 *(depending on* Q*), such that for all* $\theta \in Q$*, and all* n *we have*

$$|H(\theta, x)| \leq C_1(1 + |x|^{q_1})$$
$$|\rho_n(\theta, x)| \leq C_2(1 + |x|^{q_2})$$

Proposition 7, and equations (3.33) and (3.34) establish that the process X_l and the function $H(\theta, X)$ in the parameter update equation (4.10) satisfy the following variations of Assumptions 2 and 3.

Assumption 2-*bis*. *There exists a family* $\{\Pi_\theta : \theta \in \Theta\}$ *of transition probabilities* $\Pi_\theta(X, A)$ *on* $\mathcal{X}$ *such that, for any Borel subset* A *of* $\mathcal{X}$*, we have*

$$P[X_{l+1} \in A \mid F_l] = \Pi_{\theta_l}(X_l, A)$$

Assumption 3-*bis*. *For any Compact subset* Q *of* D*, there exist constants* C_1*, and* q_1*,(depending on* Q*), such that for all* $\theta \in Q$*, and all* n *we have*

$$|H(\theta, X)| \leq C_1 \beta_{q_1}(X)$$

For a specific trajectory θ_k, the error e_K depends on the sequence of parameter estimates $\{\theta_k,\ 0 < k < K\}$, the iterates of the Markov chain $\{X_k,\ 0 < k < K\}$, the parameter κ in the formula for γ_k, and K, the number of steps taken. Unboundedness in the factors $H(\theta_k, X_{k+1}) - h(\theta_k)$ as $X_k \to \partial\mathcal{X}$, and divergence of the series $\sum_k \gamma_k$ mean that the summations (4.17) are not absolutely convergent for all trajectories. The problem of unboundedness in the summands is circumvented by modifying the argument of the expectation $\mathbf{E}[e_K^2|X_0, \theta_0]$ to include only those parts of each trajectory where $H(\theta_k, X_{k+1}) - h(\theta_k)$ remains small, and the second problem, which stems from the divergence of the series $\sum_k \gamma_k$, is circumvented (after some work) with a martingale convergence theorem that takes advantage of the centralizing tendency in the distribution of the sum of random variables. The use of these two devices produces two results. The first is an approximation result for the expected error that is a function of the initial value X_0, and is conditioned on the behavior of the trajectories θ_k. The second is an estimate of the probability that a trajectory starting with initial value θ_0 remains well behaved.

The factor $H(\theta_k, X_{k+1}) - h(\theta_k)$ that appears in each summand of the cumulative error becomes large when X_{k+1} lies close to the boundary of $\mathcal{X}$, in particular, when the empirical densities $X^{\zeta}_{k+1}(i,j)$ and X^{γ}_{k+1} have minima close to 0. The effect of excursions of X_{k+1} towards $\partial\mathcal{X}$ on the expected cumulated error are mitigated by the use of a stopping rule on the incremental error process ϵ_k. The stopping rule rejects from the computation of the expectation terms in the error summation that come from parts of the trajectory θ_k that are likely to be associated with problematic points X_k in the accompanying chain. The stopping rule is based on two criteria. The first criterion relies on the selection of a compact set $\Theta^c \subset \Theta$ and stopping time $\tau_1(\Theta^c) = \inf\{k : \theta_k \notin \Theta^c\}$. The second criterion relies on the selection of a positive constant ϵ, and uses the stopping time $\tau_2(\epsilon) = \inf\{k \geq 1, |\theta_k - \theta_{k-1}| > \epsilon\}$. Each sample path of the process (θ_k, X_k) is stopped at the smaller of the two stopping times, which is itself a stopping time $\tau(\epsilon, Q) = \min\{\tau_1(\Theta^c), \tau_2(\epsilon)\}$. The original expression for the expected cumulative error, $\mathbf{E}[e_K^2|X_0, \theta_0]$ is replaced by a new expression

$$\mathbf{E}\,[\sup_{k<K} I(k \leq \tau(\epsilon, \Theta^c))\, e_k^2 \,|\, X_0, \theta_0]. \tag{4.19}$$

In the modified expression for the expected cumulative error, the value for the cumulative error that each trajectory contributes to the expectation is the maximum value that the partial cumulative errors can take given that the summation stops when the trajectory either hits the stopping condition, or the index reaches the upper limit K.

Notice, for instance, that the expression (4.19) bounds the expression

$$\mathbf{E}\left[I(k \le \tau(\epsilon, \Theta^c))\, e_k^2 \,|\, X_0, \theta_0\right]$$

which is the conditional expectation of the cumulative error accumulated over the interval $0 < l < k$, and conditioned on the event that $k < \tau(\epsilon, \Theta^c)$.

The proof that the stochastic approximation iterates converge in probability to the approximate trajectories of the underlying ODE relies on bounding both the error expression (4.19) and the probability of the event $k < \tau(\epsilon, \Theta^c)$. The following assumption on the moments of the Kernels $\Pi_\theta^l(X_a; dX_b)$, and on the regularity of the kernels with respect to both X_a and the parameter θ restrict the distribution of trajectories sufficiently that the bound on the expected error (4.19) and the probability that a trajectory satisfies the stopping rule $P(k < \tau(\epsilon, \Theta^c))$ both remain small.

In the Euclidean case the form of the assumption is:

Assumption 4. *[7, p. 290, A$'$.5] For all $q \ge 1$, and for any compact subset Q of D, there exist $r \in \mathbb{N}$, and constants $\bar{\alpha} < 1$, C_1, C_2, K_1, and K_2, such that*

(i) $$\sup_{\theta \in Q} \int \Pi_\theta^r(X_a; dX_b)|X_b|^q \le \bar{\alpha}|X_a|^q + C_1$$

(ii) $$\sup_{\theta \in Q} \int \Pi_\theta(X_a; dX_b)|X_b|^q \le C_2|X_a|^q + C_1$$

For any Borel function g on $\mathcal{X}$ such that $[g]_q \le \infty$

(iii) $$\sup_{\theta \in Q} |\Pi_\theta g(X_1) - \Pi_\theta g(X_2)| \le K_1[g]_q|X_1 - X_2|(1 + |X_1|^q + |X_2|^q)$$

For all $\theta, \theta' \in Q$, and for any Borel function g with $[g]_q \le \infty$

(iv) $$|\Pi_\theta g(X) - \Pi_{\theta'} g(X)| \le K_2[g]_q|\theta - \theta'|(1 + |X|^{q+1})$$

In the hyperbolic geometry the assumption becomes:

Assumption 4-*bis*. *For all* $q \geq 1$*, and for any compact subset* Q *of* D*, there exist* $r \in \mathbb{N}$*, and constants* $\bar{\alpha} < 1$*,* C_1*,* C_2*,* K_1*, and* K_2*, such that*

(i) $$\sup_{\theta \in Q} \int \Pi_\theta^r(X_a; dX_b)\beta_q(X_b) \leq \bar{\alpha}\beta_q(X_a) + C_1$$

(ii) $$\sup_{\theta \in Q} \int \Pi_\theta(X_a; dX_b)\beta_q(X_b) \leq C_2\beta_q(X_a) + C_1$$

For any Borel function g on $\mathcal{X}$ *such that* $[g]_q \leq \infty$

(iii) $$\sup_{\theta \in Q} |\Pi_\theta g(X_1) - \Pi_\theta g(X_2)| \leq K_1[g]_q \mathfrak{d}(X_1, X_2)(\beta_q(X_1) + \beta_q(X_2))$$

For all $\theta, \theta' \in Q$*, and for any Borel function* g *with* $[g]_q \leq \infty$

(iv) $$|\Pi_\theta g(X) - \Pi_{\theta'} g(X)| \leq K_2[g]_q |\theta - \theta'| \beta_{q+1}(X))$$

With control on both the conditional error (4.19) and the probability of the event $k < \tau(\epsilon, \Theta^c)$, a two step procedure establishes the approximation result. The first step establishes that the expected error after K iterations of the algorithm converges as $K \to \infty$ to a finite limit that depends on the choices of Θ^c and κ. The second step establishes two things. First, that the probability that the trajectory θ_k remains inside the set Θ^c converges to 1 as $\kappa \to \infty$, so long as θ_0 lies well inside Θ^c. Second, that the limit on the error estimates from the first step converge to 0 as $\kappa \to \infty$.

The sum is divided into two parts, one with partial sums that form a Martingale sequence with bounded second moments, and another which converges absolutely. The method that Benveniste *et al.* use to accomplish this decomposition rests on the observation that if the difference $H(\theta_k, X_{k+1}) - h(\theta_k)$ that appears in equation (4.18) is written as the right hand side of a Poisson equation

$$(I - \Pi_\theta(X))v_\theta = H(\theta, X) - h(\theta), \tag{4.20}$$

then the error terms in equation (4.18) can be rewritten in terms of the solutions $v_\theta(X_k)$ and conditional expectations (filtered observations) of the solutions $\Pi_\theta(X_k)v_\theta$ evaluated at points on the chain. If solutions v_{θ_k} to equation (4.20) exist on a region of state-space that includes each of the points X_k in the Markov chain, then a substitution of equation (4.20) into the expression for the error gives

$$\begin{aligned} e_K &= \sum_{k=0}^{K-1} \gamma_k (v_{\theta_k}(X_k) - [\Pi_{\theta_k} v_\theta](X_k)) \\ &= \gamma_1 v_{\theta_0}(X_0) - \gamma_K v_{\theta_{K-1}}(X_{K-1}) \\ &\quad + \sum_{k=0}^{K-1} \gamma_{k+2} v_{\theta_{k+1}}(X_{k+1}) - \gamma_{k+1}[\Pi_{\theta_k} v_{\theta_k}](X_k) \\ &= \gamma_1 v_{\theta_0}(X_0) - \gamma_K v_{\theta_{K-1}}(X_{K-1}) \\ &\quad + \sum_{k=0}^{K-1} \gamma_k (v_{\theta_k}(X_{k+1}) - [\Pi_{\theta_k} v_{\theta_k}](X_k)) \\ &\quad + \sum_{k=0}^{K-1} \frac{1}{k+2}(v_{\theta_{k+1}}(X_{k+1}) - v_{\theta_k}(X_{k+1})) \\ &\quad + \sum_{k=0}^{K-1} \Big(\gamma_{k+2} - \gamma_{k+1}\Big) v_{\theta_k}(X_{k+1}) \end{aligned}$$

The sequence of partial sums associated with the summation in the third term form a martingale which is bounded in mean square, and an application of Doob's $\mathcal{L}^p$ inequality [36, p. 152] establishes convergence of the series to a finite limit. An assumption of Lipschitz regularity of the solutions to the Poisson equation with respect to variation in θ produces $O(k^{-2})$ bounds on the terms in the second summation on the right hand side, and a uniform boundedness assumption for the solutions on a region in state-space that includes the chain X_k yields $O(k^{-2})$ bounds on the terms of the third summations.

Of course, an important prerequisite to the success of this approach is the existence of solutions to the Poisson equation (4.20) which are locally Lipschitz with respect to the parameter θ, and Benveniste *et al.* make the following assumption.

Assumption 5. *[7, p. 216, A.4] There exists a function h on D, and for each $\theta \in D$ a function $v_\theta(\cdot)$ on $\mathbb{R}^k$ such that*

(i) *h is locally Lipschitz on D;*

(ii) *$(I - \Pi_\theta)v_\theta = H_\theta - h(\theta)$ for all $\theta \in D$;*

(iii) *for all compact subsets Q of D, there exist constants C_3, C_4, q_3, q_4, $\lambda \in [1/2, 1]$, such that for all $\theta, \theta' \in Q$*

$$|v_\theta(X)| \leq C_3(1 + |X|^{q_3})$$
$$|\Pi_\theta v_\theta(X) - \Pi_{\theta'} v_{\theta'}(X)| \leq C_4|\theta - \theta'|^\lambda(1 + |X|^{q_4})$$

Once again, the statement of the bounds in the assumption need to be adjusted for a hyperbolic geometry. The modified assumption is:

Assumption 5-*bis*. *Let D be an open subset of Θ. There exists a function h on the interior of D, and, for each $\theta \in D$, a function $v_\theta(\cdot)$ on $\mathcal{X}$ such that*

(i) *h is locally Lipschitz on D;*

(ii) *$(I - \Pi_\theta)v_\theta = H_\theta - h(\theta)$ for all $\theta \in D$;*

(iii) *for all compact subsets Q of D, there exist constants C_2, C_3, q_2, q_3, $\lambda \in [1/2, 1]$, such that for all $\theta, \theta' \in Q$*

$$|v_\theta(X)| \leq C_2\beta_{q_2}(X)$$
$$|\Pi_\theta v_\theta(X) - \Pi_{\theta'} v_{\theta'}(X)| \leq C_3|\theta - \theta'|^\lambda\beta_{q_3}(X)$$

Much of the next chapter is devoted to establishing this prerequisite in the context of the combined estimation and control problem. The problem of finding solutions to the Poisson equation is at the root of the branch of probability called potential theory, and examination of a case where the potential theory is well established indicates the direction in which to proceed. When X_k is a positive Harris recurrent, and aperiodic Markov chain with invariant distribution m_θ, then the range of the Potential operator $(I - \Pi_\theta)$ is contained in the space of continuous functions $\{f : m_\theta(f) = 0\}$. An application of the measure μ_θ as a functional to both sides of Equation (4.20) fixes the generator $h(\theta)$ of the ODE to be $\mu_\theta(H(\theta, \cdot))$. Furthermore, the same conditions on the chain X_k guarantee that if f is a function on the state space which satisfies $\mu_\theta(f) = 0$, then the series $\sum_1^k [\Pi^l f](X)$ converges pointwise, and the function $v_\theta(X) = \lim_k \sum_1^k \Pi^l(H(\theta, \cdot) - h(\theta))(X)$ exists, and is a solution to the Poisson equation. Unfortunately the chain that was described in Section 4.1 does not have the Harris recurrence property, and the Potential theory for Harris chains [34] is not applicable, however there are other means of demonstrating geometric ergodicity for sequence of kernels Π_θ^l, and these prove to be good enough to establish existence and regularity of solutions to the Poisson equation.

Benveniste *et al.* develop the ODE theory for a general step size γ_k. In order to ensure that a summation equivalent to (4.17) is finite, they require that, in addition to satisfying Assumption 1, the step size should satisfy the following.

Assumption 6. *[7, p. 301] There exists $\alpha > 1$ such that $\sum \gamma_k^\alpha < +\infty$.*

When $\gamma_k = 1/(k + \kappa)$ and κ is any positive integer, any choice of α with $\alpha > 1$ satisfies the assumption.

The second task in the ODE method is the proof that the approximating ODE has asymptotically stable solutions in a neighborhood of a stable equilibrium that corresponds to the value of the true parameter in the stochastic estimation problem. This proof is a classical problem in non-linear ODEs and is solved with an appropriate Lyapunov function. Benveniste *et al.* formally assume the existence of a Lyapanov function as follows.

Assumption 7. *[7, p. 233] There exists a positive function U of class C^2 on D such that $U(\theta) \to C \leq +\infty$ if $\theta \to \partial D$ or $|\theta| \to +\infty$ and $U(\theta) < C$ for $\theta \in D$ satisfying*

$$U'(\theta) \cdot h(\theta) \leq 0 \text{ for all } \theta \in D \tag{4.21}$$

The natural choice of Lyapunov function for the case that this dissertation presents is the Kullback-Leibler entropy function for the estimation problem. Section 6.1 shows that a Lyapunov function that is based on the Kullback-Leibler entropy function satisfies Assumption 7.

Chapter 5

Analysis of the control and estimation algorithm.

Part 2: Potential theory for the Markov process.

The analysis of this chapter verifies that Assumptions (A.4) and (A.5) of the previous chapter are satisfied by the combined estimation and control problem. Both of these assumptions are concerned with ergodic properties of the Markov process in the stochastic approximation scheme, and this chapter develops the ergodic theory that is required for the analysis. The first section provides an ergodic theory for the discrete subchain. Apart from the requirement that the invariant vector should be regular with respect to perturbations in the state transition matrix, the theroy in this section is standard. The second section develops an ergodic theory for the Markov modulated random processes. The material in this section is new. Finally, the third section uses the results of the first two sections to verify that the assumptions hold.

5.1 Ergodic theory for the discrete subchain

A potential theory for the chain X_k is established by considering first the discrete subchain $\tilde{X}$ using methods developed by Arapostathis and Marcus [2] and Le Gland and

Mevel [23]. The principal difference between the problems treated in these two papers and the problem treated here is that here the transition kernels Π_θ depend on the parameter θ through the control algorithm. Consequently, regularity properties of the kernels with respect to the parameter θ are needed along with geometric ergodic properties in order to apply the potential theory in the stochastic approximation analysis.

The first task is to establish that the Markov chain $\{\tilde{X}_k\}$ is irreducible and acyclic. This is done by establishing conditions on the transition matrices A_u, the output matrix B, and the control policy $v(du; u_{1,\Delta}, y_{1,\Delta})$ that ensure that the matrix M is primitive.

The state space $\tilde{\mathfrak{X}}$ is a finite set which can be mapped onto a finite rectangle in a $2\Delta+3$ dimensional lattice. The lattice structure corresponds to the product structure on the state space which is described in Section 4.1, equation (4.1). The transition kernel, which has the matrix representation defined in equation (4.2), defines a directed graph on the lattice rectangle via the mapping

$$\tilde{X}_a \rightarrow \tilde{X}_b \quad \text{if} \quad M_{\tilde{X}_a}^{\tilde{X}_b} > 0$$

The problem of determining whether the kernel $\Pi(\tilde{X}_a; d\tilde{X}_b) = M$ is primitive is equivalent to the problem of determining connectedness of the associated graph.

Lemma 8. *The kernel is primitive with index of primitivity r if and only if any point on the directed graph is connected to any other by a path that traverses r, or fewer, edges.*

Proof. Let a and b be any two indices of the matrix M, and Suppose that the corresponding vertices on the lattice v_a and v_b are connected by a path

$$v_a \rightarrow v_1 \rightarrow v_2 \rightarrow \ldots \rightarrow v_{r-1} \rightarrow v_b$$

with r edges, then, giving the matrix indices the same labels as the vertices, the product

$M_{v_a}^{v_1} M_{v_1}^{v_2} \ldots M_{v_{r-1}}^{v_b}$ is positive, and the forward product M^r satisfies

$$
\begin{aligned}
(M^r)_a^b &= \sum_{i_1,\ldots,i_{r-1} \in \tilde{\mathfrak{X}}} M_a^{i_1} M_{i_1}^{i_2} \ldots M_{i_{r-1}}^{i_b} \\
&\geq M_{v_a}^{v_1} M_{v_1}^{v_2} \ldots M_{v_{r-1}}^{v_b} \\
&> 0
\end{aligned}
$$

Conversely, if $(M^r)_a^b > 0$ then for at least one choice of index set $\{i_1, i_2, \ldots, i_{r-1}\}$ the product $M_a^{i_1} M_{i_1}^{i_2} \ldots M_{i_{r-1}}^{i_b}$ is positive. Consequently the path $v_a \to v_1 \to v_2 \to \ldots \to v_{r-1} \to v_b$ with vertices $v_1, \ldots v_{r-1}$ corresponding to the indices $i_1, \ldots, i_{r-1}$ connects the vertex v_a to v_b by a path that traverses r edges. □

$$
\begin{gathered}
x_b \otimes y_b^0 \otimes u_b^{-1} \otimes y_b^{-1} \otimes u_b^{-1} \otimes y_b^{-2} \otimes u_b^{-2} \otimes \ldots \otimes y_b^{-\Delta+1} \otimes u_b^{-\Delta} \\
\downarrow \\
x_a \otimes y_a^{-\Delta} \otimes u_a^{-\Delta} \otimes y_b^0 \otimes u_b^0 \otimes y_b^{-1} \otimes u_b^{-1} \otimes \ldots \otimes y_b^{-(\Delta-1)} \otimes u_b^{-(\Delta-1)} \\
\downarrow \\
x_a \otimes y_a^{-(\Delta-1)} \otimes u_a^{-(\Delta-1)} \otimes y_a^{-\Delta} \otimes u_a^{-\Delta} \otimes y_b^0 \otimes u_b^0 \otimes \ldots \otimes y_b^{-(\Delta-2)} \otimes u_b^{-(\Delta-2)} \\
\downarrow \\
\ldots \\
\downarrow \\
x_a \otimes y_a^0 \otimes u_a^0 \otimes y_a^{-1} \otimes u_a^{-1} \otimes y_a^{-2} \otimes u_a^{-2} \otimes \ldots \otimes y_a^{-\Delta} \otimes u_a^{-\Delta}
\end{gathered}
$$

Figure 5.1: Path construction for Proposition 9

Proposition 9. *Let A_u and B be the state transition and output matrices for a controlled hidden Markov model with entries that satisfy the inequalities*

$$A_{u;ij} > \delta \quad B_{im} > \delta$$

for some constant $\delta > 0$. Fix $\eta > 0$ and let $\{\tilde{X}_n\}$ be the discrete chain defined in Section 4.1 and determined by the finite-horizon, risk-sensitive control algorithm described in Chapter 2. Then, the transition kernel $\Pi(\tilde{X}_a; d\tilde{X}_b)$ associated with the chain $\{\tilde{X}_n\}$ is primitive with index of primitivity $\Delta + 1$, and the chain itself is irreducible and acyclic.

Proof. Consider the directed graph induced by the discrete transition kernel $\Pi(\tilde{X}_a; d\tilde{X}_b)$ on the state space $\tilde{\mathfrak{X}}$. The result will follow from Lemma 8 provided that the nodes in this graph representing any two points $\tilde{X}_a$ and $\tilde{X}_b$ in $\tilde{\mathfrak{X}}$ are connected by a path containing $\Delta + 1$ or fewer links.

Any two nodes corresponding to points $\tilde{X}_a$ and $\tilde{X}_b$ are connected by a link provided that $\Pi(\tilde{X}_a; \{\tilde{X}_b\}) = M_{\tilde{X}_a}^{\tilde{X}_b} > 0$. Recall the explicit representation from equation (4.2)

$$M_{\tilde{X}_a}^{\tilde{X}_b} = \langle x_a, A_{u_b^{-1}} x_b \rangle \langle x_a, B y_b^0 \rangle \langle \upsilon_\theta(y_a^0, u_a^{-1}, \ldots, y_a^{-\Delta+1}, u_a^{-\Delta}), u_b^{-1} \rangle$$
$$\times\, \delta_{y_a^0}(y_b^{-1}) \delta_{u_a^{-1}}(u_b^{-2}) \ldots \delta_{y_a^{-\Delta+2}}(y_b^{-\Delta+1}) \delta_{u_a^{-\Delta+1}}(u_b^{-\Delta})$$

The assumptions in the premise of the proposition directly ensure that the first two factors are non-zero, and indirectly ensure that the third factor is non-zero through an application of Theorem 6. The remaining factors all have the form $\delta_{z_a^l}(z_b^{l-1})$ with z substituted by u or y as appropriate, and these factors are non zero only if $z_a^l = z_b^{l-1}$. Consequently, a link connects the node representing $\tilde{X}_a$ to the one representing $\tilde{X}_b$ if and only if $u_b^{-l} = u_a^{-(l+1)}$ and $y_b^{-l} = y_b^{-(l+1)}$ for all l with $1 \leq l \leq \Delta$

Now let $\tilde{X}_a$ and $\tilde{X}_b$ be two arbitrary points in $\tilde{\mathfrak{X}}$, then the path illustrated in Figure 5.1 connects the node representing $\tilde{X}_a$ to the node representing $\tilde{X}_b$ in less than $\Delta + 1$ links. □

Primitivity of the discrete kernel implies that the kernel is recurrent with a single recurrence class, and is a sufficient condition for the Perron-Frobenius theorem to apply.

Specifically, the following proposition lists well known facts about primitive transition kernels for finite state Markov chains. Proofs are given in Seneta [38], Revuz [34] and Le Gland and Metivier [23]

Proposition 10. *Let $\Pi(\tilde{X}; d\tilde{X})$ be a primitive kernel on a finite discrete space $\tilde{\mathfrak{X}}$. The following are true:*

(i) The eigenvalue 1 has strictly positive left and right eigenvectors.

(ii) The eigenvectors associated with the eigenvalue $1.$ are unique up to multiplication by a scalar, The left eigenvector ν is the invariant measure for the kernel.

(iii) 1 is a simple root of the characteristic equation for Π.

(iv) $1 > r > |\lambda|$, for any eigenvalue $\lambda \neq 1$.

(v) The kernel is both recurrent and acyclic

(vi) The kernel is ergodic. There exists $c \in (0,1)$ such that for any probability measure ν_a on $\tilde{\mathfrak{X}}$, $|\nu_a \Pi^n - \nu| < c^n$

(vii) The kernel possesses a well defined potential theory. In particular if $f : \tilde{X} \to \mathbb{R}$ is a bounded function on $\tilde{\mathfrak{X}}$ (in fact, a bounded vector in $\mathbb{R}^n$) then the Poisson equation

$$(I - \Pi)v(\tilde{X}) = f(\tilde{X}) - \nu(f)$$

has a solution

$$v(\tilde{X}) = \sum_{0}^{\infty} \Pi^r (f - \nu(f)\mathbf{1})(\tilde{X})$$

The following lemma provides a basis for regularity results for the discrete kernel

Lemma 11. *The eigenvector corresponding to the Perron-Frobenius eigenvalue of a primitive stochastic matrix M is a continuous function of the matrix parameters.*

Proof. Assume that $|M - M'| = \epsilon$, and let $\Delta = 1/\epsilon(M - M')$, then the rows of Δ sum to 0, and $|\Delta| = 1$. From the Perron-Frobenius theorem

$$|\nu - \nu'| = \lim_{k\to\infty} |\nu_a M^k - \nu_a(M + \epsilon\Delta)^k|,$$

and,

$$\begin{aligned}|\nu_a M^k - &\nu_a(M + \epsilon\Delta)^k| \\ &= \left|\nu_a \left(M^k - M^k + \sum_{i=1}^{k} \epsilon^i \sum_{\substack{l_1+\dots+l_i \\ =k-i}} M^{l_1}\Delta M^{l_2}\dots\Delta M^{l_i}\right)\right| \\ &\leq \epsilon \left|\sum_{l=0}^{k-1} \nu_a M^l \Delta M^{k-l-1}\right| + \left|\sum_{i=2}^{k} \epsilon^i \sum_{\substack{l_1+\dots+l_i \\ =k-i}} \nu_a M^{l_1}\Delta M^{l_2}\dots\Delta M^{l_i}\right|\end{aligned}$$

The bound $|M^l - \mathbf{1}\nu| < K_0 c^l$ where K_0 is a constant independent of M, and $0 < c < 1$, produces a bound $|\Delta M^l| < K_1 c^l$. Applying this bound to each term in the first summation on the right hand side of the inequality gives a bound for the summation as a whole of ϵK_2 with K_2 a constant depending on M but not on k. Assume that $\epsilon K_1 < 1$, then repeated use of the bound $|\Delta M^l| < K_1 c^l$ in the inner summation of the second term on the right hand side of the inequality gives the bound

$$\begin{aligned}|\sum_{i=2}^{k} \epsilon^i \sum_{\substack{l_1+\dots+l_i \\ =k-i}} \nu_a M^{l_1}\Delta M^{l_2}\dots\Delta M^{l_i}| &\leq \sum_{i=2}^{k} \epsilon^i \sum_{\substack{l_1+\dots+l_i \\ =k-i}} |M^{l_1}||\Delta M^{l_2}|\dots|\Delta M^{l_i}| \\ &\leq \sum_{i=2}^{k} \epsilon^i \sum_{l=0}^{k-i} K_1^i c^{k-i-l} \\ &\leq \epsilon^2 K_1^2 \frac{1}{(1-\epsilon K_1)(1-\rho)}\end{aligned}$$

a bound that is again independent of k. So for any primitive stochastic matrix M, it is possible to choose $K > \max\{K_1, K_2\}$, and $\epsilon_0 > 0$, both dependent on M such that if M' is a primitive stochastic matrix satisfying $|M - M'| < \epsilon$ for some $\epsilon \leq \epsilon_0$ then

$$|\nu - \nu'| < K\epsilon$$

□

Recall the situation that the dissertation addresses: the control of a plant that is described by a hidden Markov model with fixed, but unknown, state transition matrices A_u, and output matrix B. The control is computed using the stochastic approximation to the finite-horizon dynamic programming algorithm that was introduced in Chapter 2 along with an estimate for the plant transition matrices that is denoted by the variable θ. If θ is fixed, then the evolution of the controlled plant is described by a time-invariant Markov chain with a finite state set that is represented in (4.1) as a tensor product of canonical basis vectors. The evolution of the process is governed by a kernel $\Pi_\theta(\tilde{X}_l; d\tilde{X}_{l+1})$ which has a matrix representation $M_{\tilde{X}_a}^{\tilde{X}_b}$ with entries that depend on the choice of θ only through the factor $\langle \upsilon_\theta(y_a^0, u_a^{-1}, \ldots, y_a^{-\Delta+1}, u_a^{-\Delta}), u_b^{-1} \rangle$ in equation (4.2). Proposition 10, Lemma 11, and the fact that the control policy υ_θ is Lipschitz continuous with respect to the estimate θ combine to establish the conclusions of the following Proposition.

Proposition 12. *Let A_u and B be state and output transition matrices for a controlled hidden Markov model that satisfy the conditions in the premise of Proposition 9.*

(i) If θ is an estimate for the hidden Markov model parameters, $\tilde{\mathcal{X}}$ is the state space for the Markov chain that describes the evolution of the controlled hidden Markov model, then the transition kernel for the Markov chain $\Pi_\theta(\tilde{X}_l, d\tilde{X}_{l+1})$

has an invariant measure $\tilde{\nu}_\theta$, and for any positive constant D, there exists a positive constant K_1 and a constant ρ in the interval $0 < \rho < 1$ such that for all $g : \tilde{\mathfrak{X}} \to \mathbb{R}$ bounded by D, for all $\tilde{X}_a, \in \tilde{\mathfrak{X}}$, and for all $n > 0$

$$|(\Pi_\theta^n g)(\tilde{X}_a) - \tilde{\nu}_\theta(g)| \leq DK_1\rho^n$$

(ii) There exists a positive constants K_2 and ϵ_0, such that that if θ and θ' are two estimates for the hidden Markov model parameters with $|\theta - \theta'| < \epsilon \leq \epsilon_0$, then the invariant measures $\tilde{\nu}_\theta$ and $\tilde{\nu}_{\theta'}$ associated with the transition kernels Π_θ and $\Pi_{\theta'}$ satisfy

$$|\tilde{\nu}_\theta - \tilde{\nu}_{\theta'}| < K_2\epsilon$$

5.2 Ergodic theory for random walks on semi-groups

Proposition 7 characterizes the transition kernel for the continuous part of the chain X_k in terms of Markov modulated random walks. In the case where the underlying Markov chain is ergodic, the asymptotic properties of the random walk associated with the invariant distribution of the underlying chain determines the asymptotic properties for the Markov modulated random walk. This section develops a potential theory for random walks on semigroups and then shows how this theory can be extended to Markov modulated random walks.

Let $\mathfrak{G}$ be a topological semigroup with identity $\mathfrak{e}$, let ρ be a metric on $\mathfrak{G}$ that is compatible with the topology. The metric is right invariant if, for any $h_1, h_2, g \in \mathfrak{G}$, $\rho(h_1 g, h_2 g) = \rho(h_1, h_2)$. A member $g \in \mathfrak{G}$ is contractive if, for any $g_1, g_2 \in \mathfrak{G}$,

$$\rho(gg_1, gg_2) \leq \rho(g_1, g_2),$$

and is strictly contractive if the inequality is strict. An element $g_0 \in \mathfrak{S}$ has zero rank if for all $g \in \mathfrak{S}$, $g_0 g = g_0$. The elements of zero rank form a semigroup (without identity) since if $g_0 \in \mathfrak{S}$ is of zero rank then for all $g \in \mathfrak{S}$, gg_0 has zero rank.

In all the propositions $\mathfrak{S}$ is a finite dimensional topological semigroup with identity $\mathfrak{e}$, and metric ρ. The sub-semigroup of elements of zero rank is $\mathfrak{S}_0$ and there exists a continuous function $c : \mathfrak{S} \to [0, 1]$ with $c(\mathfrak{e}) = 1$, and $c(g) = 0$ for all $g \in \mathfrak{S}_0$.

A random walk on $\mathfrak{S}$ is a sequence of random variables $g_0, g_1, \ldots, g_k$ taking values in $\mathfrak{S}$. The distribution of g_0 is determined by the marginal probability measure μ_0, and the process distribution on the whole sequence is determined by the sequence of Markov convolution kernels $\Pi(g_l; dg_{l+1}) = d(\delta_{g_l} * \mu_{l+1})(g_{l+1})$. When the generators $\mu_1, \mu_2, \ldots$ are identical distributions, the random walk is homogeneous.

A random walk on $\mathfrak{S}$ is contractive if under each of the measures μ_l, $l \geq 1$, the elements of $\mathfrak{S}$ are contractive with probability 1. The random walk is strongly contractive if it is contractive, and there exist constants $0 \leq c_0 < 1$, and $1 \geq c_1 > 0$ such that for each $l \geq 1$ there exists a set $\mathcal{H}_l \subset \operatorname{supp} \mu_l$ with $\mu_l(\mathcal{H}_l) > c_1$, and for all $h \in \mathcal{H}_l$, $g_1, g_2 \in \mathfrak{S}$,

$$\rho(hg_1, hg_2) \leq c_0 \rho(g_1, g_2). \tag{5.1}$$

A third, slightly weaker, definition for a contractive random walk is a consequence of the following observation.

Lemma 13. *If a sequence of measures μ_l, $l = 1, 2 \ldots$, each with support on a semigroup $\mathfrak{S}$, generates a contractive random walk then the sequence of measures formed by taking r convolution products, $\mu_k^r = \mu_{kr} * \overset{r}{\ldots} * \mu_{(k-1)r+1}$ also generates a contractive random walk on $\mathfrak{S}$.*

A random walk is r-strongly contractive if the random walk generated by the r-convolution product μ_k^r is strongly contractive.

Lemma 14. *Let g_k be an r-strongly contractive random walk, and let c_0 and c_1 be the constants defined in (5.1). Then, for all c, in the interval $c_0^{c_1} < c < 1$, there exists (with probability 1) a constant $K > 0$ such that for all $k > K$, and for all $g, g' \in \mathfrak{S}$, $\rho(g_k g, g_k g') < c^k \rho(g, g')$.*

Proof. Let $g_k = h_k h_{k-1} \ldots h_2 h_1$, then the definition of an r-strongly contractive random walk implies the following (random) bound

$$\begin{aligned} \log \rho(g_k g, g_k g') &\leq \sum_{i=1}^{\lceil k/r \rceil} \log(c_0) \mathbf{1}_{\mathcal{H}_i}(h_i) + \log \rho(g, g') \\ &= \left\lceil \frac{k}{r} \right\rceil \log(c_0) \left(\left\lceil \frac{k}{r} \right\rceil^{-1} \sum_{i=1}^{\lceil k/r \rceil} \mathbf{1}_{\mathcal{H}_i}(h_i) \right) + \log \rho(g, g') \end{aligned}$$

where $\mathbf{1}_{\mathcal{H}_i}$ is the indicator function for the set $\mathcal{H}_i$ with the property that for all $g_1, g_2 \in \mathfrak{S}$ and $h \in \mathcal{H}_i$ the inequality in (5.1) holds. Since $\mu_i(\mathcal{H}_i) > c_1$, The process $\mathbf{1}_{\mathcal{H}_i}(h_i)$ is bounded below by a Bernoulli process which at each i takes the value 1 with probability c_1. From the strong law of large numbers

$$\lim_{k\to\infty} \frac{r}{k} \sum_{i=1}^{\lceil k/r \rceil} \mathbf{1}_{\mathcal{H}}(h_i) \geq c_1 \qquad \text{almost surely,}$$

and it follows that for any $\epsilon > 0$ there exists K such that for any $k > K$,

$$\lim_{k\to\infty} \frac{r}{k} \sum_{i=1}^{\lceil k/r \rceil} \mathbf{1}_{\mathcal{H}}(h_i) < \mu(\mathcal{H}) + \epsilon \tag{5.2}$$

Given c in the interval $c_0^{c_1} < c < 1$, choose $\epsilon = c_1 - (\log c / \log c_0)$, then $0 < \epsilon < c_1$, and

$$\rho(g_k g, g_k g') < c^{\lceil k/r \rceil} \rho(g, g') \tag{5.3}$$

for all $k > K$. □

One approach to the analysis of the asymptotic behavior of a random walk on the semigroup is to embed the open subset of non-singular elements of the semigroup inside a suitable group, and then to use the established theory for random walks on groups [34] to provide results for the embedded semigroup. The following example demonstrates that this approach is not productive, and provides motivation for the theory that follows.

Let $\mathfrak{A}$ be the group of invertible affine transformations of the N dimensional hyperplane in $\mathbb{R}^N$ that contains the probability simplex $\Omega(N)$, and let $\mathfrak{T}$ be the semigroup of affine transformations in $\mathfrak{A}$ that map $\Omega(N)$ into itself. Define a metric on $\mathfrak{A}$ by $\rho(g_1, g_2) = \sup_{m\in\Omega(N)} |g_1 m - g_2 m|$ so that $\mathfrak{A}$ is a topological group with respect to the topology induced by ρ. Recall[1] that a random walk $\{g_k\}$ on the topological group $\mathfrak{A}$ is topologically recurrent if for every open set $O \subset \mathfrak{A}$ that contains the group's identity element,

$$P[\limsup_{k\to\infty}\{g_k \in O\}] = 1. \tag{5.4}$$

Otherwise it is transient.

Proposition 15. *Let g_k be an r-strongly contractive random walk on $\mathfrak{T}$ with transition kernel $\Pi(g_a, dg_b) = d(\delta_{g_a} * \mu)(g_b)$. Then, when considered as a random walk on the group $\mathfrak{A}$, g_k is topologically transient. In fact, if $\mathfrak{T}_0$ denotes the sub semigroup of zero-rank elements, than any set compactly contained in $\mathfrak{T}\setminus\mathfrak{T}_0$ is transient with respect to g_k.*

Proof. Define a function $\gamma : \mathfrak{T} \to \mathbb{R}$ by

$$\gamma(g) = \sup_{g_1, g_2 \in \mathfrak{T}} \rho(gg_1, gg_2)/\rho(g_1, g_2), \tag{5.5}$$

then γ is continuous with respect to the metric ρ, γ takes values on the interval $[0, 1]$, γ maps the identity to the value 1, and if $\gamma(g) = 0$, then g has rank zero. Since the

[1] See, for example, Revuz. [34]

random walk g_k is r-strongly contractive, and it follows from Lemma 14 that for all $1 > c > c_0$, with probability 1 there exists a positive K such that for all $k > K$ and for all $g, g' \in \mathfrak{T}$, $\rho(g_k g, g_k g') < c^k \rho(g, g')$. Consequently, for all $k > K$

$$\gamma(g_k) = \sup_{g,g'} \frac{\rho(g_k g, g_k g')}{\rho(g, g')} \leq c^k \to 0 \qquad \text{as } k \to \infty. \tag{5.6}$$

if $\mathcal{H}$ be a neighborhood of the semigroup identity that is compactly contained in $\mathfrak{T} \setminus \mathfrak{T}_0$, then γ is bounded away from 0 in $\mathcal{H}$, and the exponential bound on $\gamma(g_k)$ in (5.6) ensures (with probability 1) that g_k hits $\mathcal{H}$ only finitely many times. □

Lemma 14 and Proposition 15 are similar in spirit to the exponential forgetting results of, for example, Le Gland and Mevel [23].

An obvious corollary of Proposition 15 and the compactness of the sub-semigroup of zero-rank elements is that if the contractive random walk has a recurrent set, it must be contained in the sub-semigroup of rank zero elements. This observation indicates a key shortcoming in using the group embedding approach as the basis for a potential theory. While the potential kernel associated with the random walk is proper[2] on the sub-semigroup of invertible elements, it is no longer proper when the elements of deficient rank are included. The group embedding approach of Proposition 15 pushes the singular set out to a boundary, and therefore avoids the central issue, which is the characterization of the asymptotic behavior of the random walk. The failure of the group embedding approach highlights the problems caused by an absence of continuous inverses, an inherent feature of semigroup structures.

[2]If $(E, \mathcal{E})$ is a measurable space, then a kernel $K : E \times \mathcal{E} \to \mathbb{R}$ is proper if there exists an increasing sequence of sets E_n with $E = \cup E_n$ such that for all n, $K(\cdot, E_n)$ is bounded. If P is a positive kernel on $(E, \mathcal{E})$, then the potential kernel associated with P is defined by $G = I + P + P^2 \ldots$. When the sequence converges (in an appropriate operator topology), the potential kernel is a left inverse to the Poisson operator $f \to (I - P)f$. See Revuz. [34]

Another approach is needed, one which directly attacks the problem of determining an appropriate setting in which random walks on semigroups exhibit recurrence properties. If $P(x; dx)$ is a probability kernel with respect to the measurable space $(E, \mathcal{E})$, and h is a function on E that takes values in the interval $[0, 1]$, then the generalized resolvent associated with P and the function h is the kernel $U_h(x; dx)$ given by the series

$$U_h = \sum_{n \geq 0} (P I_{1-h})^n P$$

where the operator I_{1-h} is defined by $[I_{1-h} f](x) = (1 - h(x)) f(x)$ (see Revuz [34]). The series is guaranteed to converge uniformly with respect to x and in the variational topology with respect to the space of measures on $\mathcal{E}$ when the function h takes values in the open interval $(0, 1)$. The potential kernel G associated with a general probability kernel P is formally defined by $G = I + U_0$.

The notion of ν-irreducibility is central to the standard development of a potential theory for Markov chains.

Definition 1. *(Revuz [34, Definition 2.1]) A chain X on a measure space $(E, \mathcal{E})$ is said to be* ν-irreducible *if there exists a probability measure ν on $\mathcal{E}$ which is absolutely continuous with respect to $U_c(x, \cdot)$ for all $x \in E$ and all constant functions c taking a value in* $(0, 1)$.

The importance of this notion is that if a Markov chain is ν-irreducible for some measure ν, then either the potential kernel for the chain is strictly proper, or the chain is recurrent in the following sense[34].

Definition 2. *A chain X defined on a measure space $(E, \mathcal{E})$ is said to be* recurrent in the sense of Harris *if there exists a positive, σ-finite, invariant measure m such that*

$m(A) > 0$ *implies that*

$$P_{X_0}\left[\sum_{n=1}^{\infty} \mathbf{1}_A(X_n) = \infty\right] = 1$$

for all x *in* E.

A chain that is Harris recurrent is clearly m-irreducible when m is the invariant measure postulated in the definition.

The key to establishing a recurrence theory for random walks on semigroups lies in identifying the short coming in these definitions. Let the transition kernel for the random walk be $\delta_x * \mu(x)$ for some measure μ supported on the invertible elements of $\mathfrak{S}$, and let ν be any measure supported on the set $\mathfrak{S}_0$ of elements with zero rank, then ν is singular with respect to $U_c(x; dx)$ whenever x is an invertible element of $\mathfrak{S}$. Since an invariant measure for the random walk must be supported on $\mathfrak{S}_0$, the random walk cannot be Harris recurrent, since this would imply irreducibility as well. The problem lies in the implicit choice of total variation norm as the topology of convergence for sequences of measures, and the solution is to look for a topology on the space of measures that supports convergence of a sequence of measures to a measure that is mutually singular to every measure in the sequence.

An appropriate family of topologies for a suitable recurrence theory are topologies generated by various notions of weak convergence of probability measures. If E is a metric space, $C(E)$, the space of bounded, continuous linear functions that map E to $\mathbb{R}$, then a sequence of bounded linear functionals $\lambda_k : C(E) \to \mathbb{R}$ converges to a bounded linear functional λ in the *weak*-$*$ topology if, for any $f \in C(E)$, $\lim_{k\to\infty} |\lambda_k(f) - \lambda(f)| \to 0$. When the bounded linear functionals are probability measures the convergence is generally called weak convergence. Under the weak topology on probability measures, the notion analogous to the notion of ν-irreducibility given in Definition 1 is the following.

Definition 3. *A chain X on a measure space $(E, \mathcal{E})$ is said to be* weakly ν-irreducible *if there exists a probability measure ν on $\mathcal{E}$ such that for all $x \in E$ and all constant functions c taking a value in $(0, 1)$, the implication $U_c(x, \cdot)h = 0 \Rightarrow \nu h = 0$ holds for all non-negative continuous functions $h : E \to \mathbb{R}$.*

The properties of the weak topology on the space of probability distributions is standard probability theory and can be found, for example, in Loève [27]. Aspects of the theory, and small embellishments that are pertinent in the present context are introduced in the following paragraphs.

The choice of the function space $C(E)$, the continuous bounded functions, determines which sequences of measures converge in the weak topology. If a larger function space is chosen, then the corresponding weak topology is larger, and fewer measures converge. In particular, let $\beta : E \to \mathbb{R}$ be a positive, continuous function bounded away from zero, let $C_\beta(E)$ be the space of functions $f : E \to \mathbb{R}$ that are bounded by $f(X) < K\beta(X)$ for some constant K, then $C^*_\beta(E)$ will denote the space of bounded linear functionals on $C_\beta(E)$ with the topology of weak-$*$ convergence. The following lemma has the fortunate consequence that topological properties of the space $C^*(E)$ also also properties of $C^*_\beta(E)$

Lemma 16. *The two spaces $C^*(E)$ and $C^*_\beta(E)$ are homeomorphic*

Proof. The mapping $\mu(\cdot) \to \mu(\beta^{-1}\cdot)$ is an open continuous bijection between the two spaces. ☐

Let E^n denote the Cartesian product of n copies of E. If $X \in E^n$, then $X = (X_1, \ldots, X_n)$ has n components, and each component lies in a copy of the space E. If $\beta : E \to \mathbb{R}$ is a weighting function on E, then define $\beta : (E^n) \to \mathbb{R}$ by

$$\beta(X) = \sup\{\beta(X_1), \ldots, \beta(X_n)\}.$$

The function space $C_\beta(E^n)$ and the dual space of measures $C^*_\beta(E^n)$ are natural extensions of the spaces $C_\beta(E)$ and $C^*_\beta(E)$.

Although the view of probability measures as linear functionals on a topological vector space is not a drastic departure from the standard presentation of probability theory, some care needs to be exercised in the definitions of distribution density functions. Conditional densities and probability kernels are of particular importance in the context of this work. Given a probability space $(\Omega, \mathcal{F}, P)$, a second sigma algebra $\mathcal{F}_1 \subset \mathcal{F}$, and a metric space E with weight function β, let $S : \Omega \to E$ be an $\mathcal{F}$ measurable random variable with a probability distribution $\mu \in C^*_\beta(E)$ that is absolutely continuous with respect to P. The conditional density for the conditional distribution of S, conditioned with respect to $\mathcal{F}_1$, maps any function $f \in C_\beta(E)$ to an $\mathcal{F}_1$ measurable function on Ω, and any point $\omega \in \Omega$ to a probability measure in $C^*(E)$. The conditional density, which is written as $\mu(S|\omega)$, is defined through the identity

$$\int_A \int f(S)\, d\mu(S|\omega)\, dP(\omega) = \int_A f(S(\omega))\, dP(\omega). \qquad \forall A \in \mathcal{F}_1$$

The existence of a conditional density is ultimately a consequence of the Radon Nikodym theorem, which ensures the existence of the conditional expectation $E[f(S)|\mathcal{F}_1]$. This conditional expectation has a representation as an $\mathcal{F}_1$ measurable function $\omega \to E[f(S)|x]$, and, for fixed x, this function is a positive, bounded, linear functional on $C_\beta(E)$ with the property that $E[\mathbf{1}|x] = 1$. Consequently, for each x, the linear functional has a representation as a probability density $\mu(S|x) \in C^*_\beta(E)$.

The definition of the conditional density $\mu(S|x)$ corresponds to Doob's definition of a conditional density in the weak sense [10]. The most important examples of a conditional density in the present work are probability kernels[3] $\Pi(S_a; dS_b)$ where the σ-algebra $\mathcal{F}_1$ is generated by a second random variable $S_a : \Omega \to E$.

[3]Note that the order of the arguments of a probability kernel is the reverse of the order of the ar-

When the underlying space (in this case the semigroup $\mathfrak{S}$) is Polish[4], the restriction of the weak topology to the probability simplex is metrizable, and, in fact, is the topology induced by the Lévy metric

$$d(\mu, \mu') = \min_{\delta} \Big\{ \mu(F) < \mu'(F^{\delta}) + \delta \text{ and } \mu'(F) < \mu(F^{\delta}) + \delta \quad \forall F \text{ closed subset of } \Omega \Big\}$$

where F^{δ} denotes a δ neighborhood about the closed set F defined in a metric that is compatible with the topology on $\mathfrak{S}$. (A proof of this result is given by Deuschel and Stroock [9, p. 65] who attribute the result to Lévy and Prohorov.) The definition of the Lévy metric can be reworded as follows:

Proposition 17. *Let μ and μ' be two measures on $\mathfrak{S}$, then $d(\mu, \mu') < \delta$ if and only if for any closed set $F \subset \Omega$ one of the following inequalities is true*

$$\mu(F) < \mu'(F) < \mu(F^{\delta}) + \delta$$
$$\mu'(F) < \mu(F) < \mu'(F^{\delta}) + \delta$$

It is a direct consequence of Lemma 16 that the topologies $C^*_{\beta}(E)$ are also metrizable. The inequalities in Proposition 17, and the definition of the Lévy metric become

$$\mu(\mathbf{1}_F \beta) < \mu'(\mathbf{1}_F \beta) < \mu(\mathbf{1}_{F^{\delta}} \beta) + \delta$$
$$\mu'(\mathbf{1}_F \beta) < \mu(\mathbf{1}_F \beta) < \mu'(\mathbf{1}_{F^{\delta}} \beta) + \delta$$

The next lemma and proposition are used later to establish regularity results. The lemma presents a characterization of the Lévy metric in terms of Lipschitz functions.

guments in a conditional density. Whereas $\mu(S|\omega)$ is a conditional density on the random variable S conditioned on the elements of the sigma-algebra of the random variable ω, $\Pi(S_a; dS_b)$ is a density on the random variable S_b conditioned on elements of the sigma-algebra of the random variable S_a.

[4] A complete separable metric space

Again, to cope with spaces of functions with bounded growth, the definition of a Lipschitz function is altered so that the Lipschitz constant is weighted by the growth bound. Let E be a metric space with metric ρ, and let $\beta : E \to \mathbb{R}$ be a positive continuous weight function that is bounded away from 0. A function $f : E \to \mathbb{R}$ is Lipschitz with respect to β if there exists a constant L_f such that $\forall X_1, X_2 \in E$

$$|f(X_1) - f(X_2)| \leq L_f \rho(X_1, X_2)(\beta(X_1) + \beta(X_2))$$

Lemma 18. *let μ and μ' be two measures on $\mathfrak{S}$, and suppose that for any Lipschitz function f with Lipschitz constant L_f,*

$$|\mu(f) - \mu'(f)| < \epsilon L_f$$

then $d(\mu, \mu') < \sqrt{\epsilon}$

Proof. Let ρ be a metric on $\mathfrak{S}$ that is compatible with the topology, let F be a closed subset of $\mathfrak{S}$, choose $\delta > \sqrt{\epsilon}$, and let $h : \mathfrak{S} \to [0, 1]$ be the function

$$h(g) = \begin{cases} \beta(g) & g \in F \\ \frac{1}{\delta}\rho(g, E \setminus F^\delta)\beta(g) & g \in (E \setminus F) \cap F^\delta \\ 0 & g \in E \setminus F \end{cases}$$

Then h is Lipschitz with $L_h = 1/\delta$, and

$$\mu(\mathbf{1}_F \beta) < \mu(h) \leq \mu'(h) + \epsilon L_h < \mu'(\mathbf{1}_{F^\delta}\beta) + \epsilon L_h < \mu'(\mathbf{1}_{F^\delta}\beta) + \delta.$$

The other inequality in the definition, and hence the proof of the proposition is demonstrated by the same argument with μ and μ' swapped. ☐

Proposition 19. *Let ν be a measure on a set $\mathbb{S}$, let E be a metric space with metric ρ, and a weight function β, and let $g : \mathbb{S} \to E$ and $g' : \mathbb{S} \to E$ be two mappings from*

$\mathcal{S}$ *to* E *that satisfy* $\nu(\beta \circ g) < M$, *and* $\nu(\beta \circ g) < M$ *for some constant* $M > 0$. *Let* $\mu = [d_\star g]\nu$ *and* $\mu' = [d_\star g']\nu$ *be the two measures on* E *induced by* ν *and the maps* g *and* g'. *Then,*

$$d(\mu, \mu') < \sqrt{\sup_{S \in \operatorname{supp} \nu} \rho(g(S), g'(S))}$$

Proof. Let f be Lipschitz with respect to β on E, and have Lipschitz constant L_f, then

$$\begin{aligned} |\mu f - \mu' f| &= |\nu(f \circ g) - \nu(f \circ g')| \\ &= |\nu(f \circ g - f \circ g')| \\ &\leq L_f \sup_{S \in \operatorname{supp} \nu} \rho(g(S), g'(S)) \end{aligned}$$

and the result follows from an application of Lemma 18. ☐

The Lévy Metric is convex in the following sense

Proposition 20. *Let* $\{\mu_\alpha : \alpha \in A\}$ *and* $\{\mu'_\alpha : \alpha \in A\}$ *be two families of probability measures in* $C^*_\beta(E)$ *that are indexed by the same countable set* A, *and that satisfy the inequalities* $d_\beta(\mu_\alpha, \mu'_\alpha) < \delta$ *for fixed* δ. *If* $\mu = \sum_\alpha c_\alpha \mu_\alpha$, *and* $\mu' = \sum_\alpha c_\alpha \mu'_\alpha$ *where* c_α *is any sequence of non-negative numbers that satisfy* $\sum c_\alpha = 1$, *then* $d_\beta(\mu, \mu') < \delta$.

Proof. Let F be a closed measurable subset of E, and $B \subset A$ contain the indices in A that satisfy $\mu_\alpha(\mathbf{1}_F \beta) \leq \mu'_\alpha(\mathbf{1}_F \beta) \leq \mu_\alpha(\mathbf{1}_{F^\delta} \beta) + \delta$. The indices in $A \setminus B$ satisfy $\mu'_\alpha(\mathbf{1}_F \beta) < \mu_\alpha(\mathbf{1}_F \beta) \leq \mu'_\alpha(\mathbf{1}_{F^\delta} \beta) + \delta$. Define

$$\mu_b = \sum_{\alpha \in B} c_\alpha \mu_\alpha, \quad \mu'_b = \sum_{\alpha \in B} c_\alpha \mu'_\alpha, \quad \mu_a = \sum_{\alpha \in A \setminus B} c_\alpha \mu_\alpha, \text{ and } \mu'_a = \sum_{\alpha \in A \setminus B} c_\alpha \mu'_\alpha.$$

If $c_a = \sum_{\alpha \in A \setminus B} c_\alpha$ and $c_b = \sum_{\alpha \in B} c_\alpha$ then $c_a + c_b = 1$,

$$\mu_b(\mathbf{1}_F \beta) \leq \mu'_b(\mathbf{1}_F \beta) \leq \mu_b(\mathbf{1}_{F^\delta} \beta) + c_b \delta, \tag{5.7}$$

$$\text{and} \quad \mu'_a(\mathbf{1}_F \beta) < \mu_\alpha(\mathbf{1}_F \beta) \leq \mu'_a(\mathbf{1}_{F^\delta} \beta) + c_a \delta. \tag{5.8}$$

Suppose that $\mu(\mathbf{1}_F\beta) = \mu_b(\mathbf{1}_F\beta) + \mu_a(\mathbf{1}_F\beta) < \mu_b'(\mathbf{1}_F\beta) + \mu_a'(\mathbf{1}_F\beta) = \mu'(\mathbf{1}_F\beta)$, then it follows from the second of the inequalities in (5.7) that $\mu_b'(\mathbf{1}_F\beta) \leq \mu_b(\mathbf{1}_{F^\delta}\beta) + \delta$, and it follows from the first of the inequalities in (5.8) that $\mu_a'(\mathbf{1}_F\beta) \leq \mu_a(\mathbf{1}_F\beta) \leq \mu_a(\mathbf{1}_{F^\delta}\beta) + \delta$. Consequently

$$\mu_b(\mathbf{1}_F\beta) + \mu_a(\mathbf{1}_F\beta) < \mu_b'(\mathbf{1}_F\beta) + \mu_a'(\mathbf{1}_F\beta) \leq \mu_b(\mathbf{1}_{F^\delta}\beta) + \mu_a(\mathbf{1}_{F^\delta}\beta) + \delta. \quad (5.9)$$

If, on the other hand, $\mu(\mathbf{1}_F\beta) = \mu_b(\mathbf{1}_F\beta) + \mu_a(\mathbf{1}_F\beta) \geq \mu_b'(\mathbf{1}_F\beta) + \mu_a'(\mathbf{1}_F\beta) = \mu'(\mathbf{1}_F\beta)$, then a symmetric argument implies that

$$\mu_b'(\mathbf{1}_F\beta) + \mu_a'(\mathbf{1}_F\beta) \leq \mu_b(\mathbf{1}_F\beta) + \mu_a(\mathbf{1}_F\beta) \leq \mu_b'(\mathbf{1}_{F^\delta}\beta) + \mu_a'(\mathbf{1}_{F^\delta}\beta) + \delta. \quad (5.10)$$

One or the other of (5.9) or (5.10) is true for any choice of a measurable set F, and therefore $d_\beta(\mu, \mu') < \delta$. □

The following proposition uses Lemma 18 to establish a result about independence that will be useful later.

Proposition 21. *Let x, y, and z be three random variables taking values in a metric space E with an associated function space $C_\beta(E)$. The probability measures that x, y, and z generate on E, and the measures generated by the joint distributions are all measures in $C_\beta^*(E)$, and are written as $\mu(x)$, $\mu(y)$, $\mu(x,z)$, etc. Conditional distributions are written $\mu(x|z)$, etc.*

If the conditional distributions $\mu(z|y,x)$ and $\mu(z|y)$ satisfy $\mu(z|y,x) = \mu(z|y)$, and if the map $\mu(z|y) : C_\beta(E) \to C_\beta(E)$ is bounded with respect to Lipschitz seminorms, with bound M, then

$$d(\mu(x,z), \mu(x)\mu(z)) < \sqrt{M d(\mu(x,y), \mu(x)\mu(y))}.$$

Proof. Let $f \in C_\beta(E)$ with Lipschitz constant L_f.

$$\begin{aligned}\int f(x,z)\,d\mu(x,z) &= \int f(x,z)\,d\mu(x,y,z)\\ &= \iint f(x,z)\,d\mu(x|yz)d\mu(y,z)\\ &= \iiint f(x,z)\,d\mu(z|y)d\mu(x|y)d\mu(y)\\ &= \iiint f(x,z)\,d\mu(x|y)d\mu(z|y)d\mu(y)\\ &= \iint f(x,z)\,d\mu(z|y)d\mu(x,y).\end{aligned}$$

Also,

$$\begin{aligned}\iint f(x,z)\,d\mu(x)d\mu(z) &= \iiint f(x,z)\,d\mu(x)d\mu(z|y)d\mu(y)\\ &= \iiint f(x,z)\,d\mu(z|y)d\mu(x)d\mu(y).\end{aligned}$$

Let $g(x,y) = \int f(x,z)\,d\mu(z|y)$, then

$$\begin{aligned}&\left|\int f(x,z)\,d\mu(x,z) - \iint f(x,z)\,d\mu(x)d\mu(z)\right|\\ &\qquad\le \left|\int g(x,y)\,d\mu(x,y) - \iint g(x,y)\,d\mu(x)d\mu(y)\right|\\ &\qquad\le L_g d(\mu(x,y),\mu(x)\mu(y)) \le M L_f d(\mu(x,y),\mu(x)\mu(y)).\end{aligned}$$

It follows from Lemma 18 that

$$d(\mu(x,z),\mu(x)\mu(z)) \le \sqrt{M d(\mu(x,y),\mu(x)\mu(y))}$$

□

Returning to the problem of establishing ν-irreducibility and recurrency for random walks on semigroups, given a semigroup $\mathfrak{S}$, let $K_0 \subset \mathfrak{S}$ be a compact subset of $\mathfrak{S}$, and let $\mathcal{G}$ be a set of measures on $\mathfrak{S}$. If $g_0 \in K_0$, and $\mu_1, \ldots, \mu_k$ are measures in $\mathcal{G}$, define

$$m_{g_0,k} = \mu_k * \overset{k}{\cdots} * \mu_1 * \delta_{g_0}. \tag{5.11}$$

The measures $m_{g_0,k}$ are the marginal distributions of a random walk g_k with starting value g_0, and the generalized resolvent U_c in Definition 3 is given by the formula

$$U_c(g_0; dg) = \sum_{k=1}^{\infty} c^k m_{g_0,k}.$$

If g_0 is fixed, the sum converges to a Radon measure when $0 \leq c < 1$.

One approach to establishing ν-irreducibility of the random walks is to examine the limit set of the sequence of measures $m_{g_0,k}$. ν-irreducibility is easily established when this limit set is independent of the initial point g_0. Relative compactness of the sequence of measures in the weak topology ensures the existence of a limit set, but rather than demonstrating relative compactness directly, it is easier to deal with the notion of tightness.

The standard definition of tightness for a set of probability measures, and the associated compactness theorem are quoted from Loève:

Definition 4 ([27, p194]). *Let $\mathcal{X}$ be a metric space with Borel field $\mathcal{S}$. A family $\mathcal{P}$ of probability measures on $\mathcal{S}$ is said to be* tight *if for every $\epsilon > 0$ there is a compact K_ϵ such that $P(\mathcal{X} \setminus K_\epsilon) < \epsilon$ for all $P \in \mathcal{P}$.*

The following proposition relates tightness of a family of measures to relative compactness in the weak topology.

Proposition 22 ([27, p195]). *Let $\mathcal{X}$ be a separable complete metric space. Then a family $\mathcal{P}$ of probability measures on its Borel field $\mathcal{S}$ is relatively compact in the $C^*(\mathcal{X})$ topology if and only if $\mathcal{P}$ is tight. In fact the "if" part holds for general metric spaces $\mathcal{X}$.*

To ensure tightness of the set of measures $m_{g_0,l}$ in the random walk, a suitable primitivity condition is imposed on the sequences of generates $\mu_1, \mu_2 \ldots$. The primitivity condition restricts the the permissible sequences of generators to a set $\mathbb{S} \subset \mathbb{N}^{\mathcal{G}}$. The difficult part of establishing a recurrence theory for a family of random walks is to impose primitivity conditions on the generator sequences in the family that are strong enough to ensure tightness in the set of measures $m_{g_0,k}$, yet weak enough to be satisfied by an interesting set of random walks.

The following definition and lemma provide sufficient conditions on the generators of a random walk g_k for the sequence of probability measures $m_{g_0,k}$ to be tight under the standard definition.

Definition 5. *A sequence $\mu_1, \mu_2 \ldots$ of generators for a random walk on a semigroup $\mathfrak{S}$ is p-*strongly primitive *if there exists a compact subset $K \subset \mathfrak{S}$ that is absorbing from the left*[5]*, and a constant $0 \leq \eta < 1$ such that $\mu_k^p(\mathfrak{S} \setminus K) < \eta$, when $\mu_k^p = \mu_k * \cdots * \mu_{k+p}$ is the convolution product of any p consecutive measures from the sequence $\mu_1, \mu_2, \ldots$.*

In the case of a homogeneous walk with initial point g_0, the measures in the sequence of generators μ_1, μ_2 are identical. The generator measure μ is p-primitive if the sequence $\mu_1 = \mu, \mu_2 = \mu, \ldots$ is primitive. The following lemma is typical of results

[5] K is absorbing from the left if for all $g \in \mathfrak{S}$, and all $g_a \in K$, $g_a g \in K$. Recall that the semigroups have representations as affine transformation groups acting on projective space. This thesis adopts the convention of writing these representations with the semigroup element acting on the object from the projective space from the left. So, for example, if $g_1, g_2, \ldots, g_k$ are successive affine transformations on a vector x in a projective space, then the orbit of x is $\{x,\ g_1 x,\ g_2 g_1 x,\ \ldots,\ g_k \ldots g_2 g_1 x\}$. Some confusion inevitably arises when the projective space is a finite dimensional probability space, and the affine map is a probability kernel that is represented by a matrix of finite rank. In this case the usual convention is to write the vector as a row vector, and to have the transition kernel act on the row vector by right matrix multiplication.

that relate primitivity of the generators to tightness of the set of iterated kernels.

Lemma 23. *If g_k is a random walk with a p-strongly primitive generator μ, and initial point g_0, then the family of measures $\{m_{g_0,k} : k \geq 0\}$ is tight.*

Proof. A generalization of the lemma is proved below. ☐

The definition of tightness extends, with some modification, to the weighted topologies introduced earlier. Let $C^*_\beta(\mathcal{X})$ be the dual space of bounded linear functionals on $C_\beta(\mathcal{X})$, again with the weak-$*$ topology generated by the topology on $C_\beta(\mathcal{X})$. A definition of tightness that is appropriate for this weaker topology is the following.

Definition 6. *Let $\beta : \mathcal{X} \to \mathbb{R}$ be a continuous positive function. A family of measures $\mathcal{P}$ is tight with respect to the $C^*_\beta(\mathcal{X})$ topology if for all ϵ there exists a compact set $K_\epsilon \subset \mathcal{X}$ such that $P(\mathbf{1}_{\mathcal{X}\setminus K_\epsilon}\beta) < \epsilon$ for all $P \in \mathcal{P}$.*

In the case where β is a constant function equal to 1, tightness with respect to the C^*_β topology is equivalent to the standard definition of tightness. The corresponding compactness proposition is the following.

Proposition 24. *Let $\mathcal{X}$ be a separable complete metric space with Borel σ-algebra $\mathcal{S}$, and let β be a continuous positive function on $\mathcal{X}$ that is bounded away from 0. A family $\mathcal{P}$ of probability measures on the Borel field $\mathcal{S}$ is relatively compact in the weak topology of $C^*_\beta(\mathcal{X})$ if and only if $\mathcal{P}$ is tight with respect to $C^*_\beta(\mathcal{X})$.*

Proof. The mapping $\mu(\cdot) \mapsto \mu(\beta\cdot)$ is a homeomorphism between the measure spaces $C^*_\beta(\mathcal{X})$ and $C^*(\mathcal{X})$. The same mapping maps families of measures in $C^*_\beta(\mathcal{X})$ that are tight with respect to β, to families in $C^*(\mathcal{X})$ that are tight under the standard definition. It follows that the proposition is a direct consequence of Proposition 22. ☐

A requirement for p-strong primitivity of the sequence of generators $\mu_1, \mu_2, \dots$ is no longer sufficient to ensure tightness of the sequence of measures $m_{g_0,1}, m_{g_0,2} \dots$ in the larger weighted topologies. Extra conditions such as the one included in the premise of the following lemma are needed.

Lemma 25. *Let $\mathcal{G} = \mu_1, \mu_2, \dots$ be a p-strongly primitive sequence of generators with uniformly bounded support, let η and H be the constant and compact set that are associated with the definition of p-strong primitivity, and let $a = \sup\{\beta(g) : g \in \cup_{\mu \in \mathcal{G}} \operatorname{supp} \mu\}$. If $a\eta^{1/p} < 1$ then the measures*

$$m_{g_0,l} = \delta_{g_0} * \mu_1 * \cdots * \mu_l$$

are tight with respect to β. In fact, let $K_l = \{g \in \mathfrak{S} : \beta(g) < a^l c\}$ where $c = \max\{\beta(g_0), \max_{h \in H} \beta(h)\}$. Then K_l is an increasing sequence of compact sets such that $\cup K_l = \mathfrak{S}$, and

$$m_{k,g_0}(\mathbf{1}_{\mathfrak{S} \setminus K_l} \beta) \leq C(a\eta^{1/p})^l. \tag{5.12}$$

Proof. From the definition of the measures m_{l,g_0},

$$\begin{aligned} m_{k,g_0}(K_l \setminus K_{l-1}) &\leq m_{k,g_0}(\mathfrak{S} \setminus K_{l-1}) \\ &\leq 1 - P(g_{r+p} \dots g_{r+2} g_{r+1} \in H \text{ for some} \\ &\qquad\qquad r \in \{k-l-p+1, \dots, k-p-1, k-p\}\Big) \\ &= 1 - \Big(1 - P\Big(g_{r+p} \dots g_{r+2} g_{r+1} \notin H, \text{ for all} \\ &\qquad\qquad r \in \{k-l-p+1, \dots, k-p-1, k-p\}\Big)\Big) \\ &\leq P\Big(g_{r+p} \dots g_{r+2} g_{r+1} \notin H \text{ for all} \\ &\qquad\qquad r \in \{k - \lfloor l/p - 1 \rfloor p, \dots, k-2p, k-p\}\Big) \\ &= \eta^{\lfloor l/p - 1/p \rfloor} \leq \eta^{(l-2p)/p} \end{aligned}$$

Consider now the quantity $m_{g,k}(\mathbf{1}_{\mathfrak{S}\setminus K_l})$. It follows from the definition of K_k that $g_k \in K_k$ for every sample path, so if $k < l$ then $m_{g,k}(\mathbf{1}_{\mathfrak{S}\setminus K_l}\beta) = 0$. When $k \geq l$, the above estimates of $m_{k,g_0}(K_l \setminus K_{l-1})$ give:

$$\begin{aligned} m_{k,g_0}(\mathbf{1}_{\mathfrak{S}\setminus K_l}\beta) &\leq \sum_{r=l}^{k} m_{k,g_0}(\mathbf{1}_{K_{r+1}\setminus K_r}\beta) \\ &\leq \sum_{r=l}^{k} m_{k,g_0}(K_{r+1} \setminus K_r) a^{r+1} c \\ &\leq c \sum_{r=l}^{k} \eta^{(r-2p)/p} a^{r+1} \\ &= c\frac{a}{\eta^2}(a\eta^{1/p})^l \sum_{r=0}^{k-l} (a\eta^{1/p})^r \\ &\leq c\frac{a}{(1 - a\eta^{1/p})\eta^2}(a\eta^{1/p})^l \end{aligned}$$

Since, by assumption, $a\eta^{1/p} < 1$, the right hand side is smaller than any fixed $\epsilon > 0$ provided l is sufficiently large. The choice of l, and by implication K_l depends on c, ϵ, η, and a, but not on k. ☐

The topology $C^*_\beta(\mathfrak{S})$ is the appropriate topology for a useful recurrence theory. A geometric forgetting result equivalent to Lemma (15) is

Proposition 26. *Let g_k be an r-strongly contractive Markov random walk on the semi-group $\mathfrak{S}$ with generator μ, let β be a strictly positive continuous function on $\mathfrak{S}$, and let $f : \mathfrak{S} \to \mathbb{R}$ be a function with magnitude bounded by $|f| < M\beta$ and which is Lipschitz with respect to β with Lipschitz constant L_f. Then, there exist constants $1 > c_0 \geq 0$, and $C > 0$, both independent of f such that for any $g_0, g_0' \in \mathfrak{S}$*

$$|(m_{g_0,l} - m_{g_0',l})f| < CL_f c_0^l \rho(g_0, g_0') \tag{5.13}$$

Proof. Because g_k is r-strongly contractive, there exists a set $H \subset \mathfrak{S}$ with $\mu * \overset{r}{\cdots} * \mu(H) = \eta > 0$, and a constant $0 \leq c < 1$ such that $\rho(hg, hg') < c\rho(g, g')$ for all

$g, g' \in \mathfrak{S}$. For all $l > 0$,

$$|(m_{g_0,l} - m_{g_0',l})f| \leq \int_{\mathfrak{S}} |f(h_l h_{l-1} \dots h_1 g_0) - f(h_l h_{l-1} \dots h_1 g)|$$

$$d\mu(h_l) d\mu(h_{l-1}) \dots d\mu(h_1).$$

$$\leq L_f \rho(g_0, g_0') \sum_{i=0}^{\lfloor l/r \rfloor} \binom{\lfloor l/r \rfloor}{i} c^i \eta^i (1-\eta)^{\lfloor l/r \rfloor - i}$$

$$= L_f \rho(g_0, g_0')(1 - (1-c)\eta)^{\lfloor l/r \rfloor}$$

The result follows when $c_0 = (1 - (1-c)\eta)^{1/r}$ and $C = (1 - (1-c)\eta)^{-1}$. □

The following proposition forms the basis for a recurrence theory for r-strongly contractive, and p-strongly primitive random walks on semigroups.

Proposition 27. *Let $\mathfrak{S}$ be a semigroup with a metric ρ that is compatible with the semigroup operator, and a positive weight function β that has compact sub-level sets, and is bounded away from zero. Let g_k be a homogeneous random walk on $\mathfrak{S}$ with Markov transition kernel $\Pi(g_a; dg_b) = d(\delta_{g_a} * \mu)(g_b)$. If the walk is r-strongly contractive, the generator μ is p-strongly primitive, r and p are finite, positive integers r and p, and the walk satisfies the conditions in the premise of Lemma 25, then there exists a probability measure m with support contained in $\mathfrak{S}_0$, the elements of zero rank, such that m is invariant with respect to the transition kernel. Furthermore, if f is any positive function that is bounded by $f(g) < M\beta(g)$, that is Lipschitz with respect to β with Lipschitz constant L_f, and that satisfies $m(f) > 0$, then:*

$$P_{g_0}\left[\sum_{k=1}^{\infty} f(g_k) = \infty\right] = 1 \tag{5.14}$$

independent of the choice of g_0.

Proof. Since g_k satisfies the premise of Lemma 25, the sequence of measures $m_{g_0,k}$ is tight with respect to the $C^*_\beta(\mathfrak{S})$ topology for any g_0 in a compact subset of $\mathfrak{S}$.

It follows from Proposition 24 that the sequence of measures is relatively compact in $C^*_\beta(\mathfrak{S})$, that is, each infinite subsequence of $m_{g_0,k}$ has a sub-subsequence $m_{g_0,q(k)}$ that converges to a measure m_q. Consider two convergent subsequences $m_{g_0,q(k)}$, and $m_{g_0,q'(k)}$ that have limiting measures m_q and $m_{q'}$. If $m_q \neq m_{q'}$, then there exists a Lipschitz continuous function f uniformly bounded on $\mathfrak{S}$ by $f(g) < M\beta(g)$, and a constant ϵ such that $|m'_q(f) - m_q(f)| = \epsilon > 0$. Choose k_0 sufficiently large that $|m_{g_0,q'(k)}(f) - m_{g_0,q(k)}(f)| > \epsilon/2$ for all $k > k_0$, and $\mu_{g_0,l}(\mathbf{1}_{\mathfrak{S}\setminus K_{k_0}} 2M\beta) < \epsilon/4$. Without loss of generality, assume that $q'(k) > q(k)$ for all k, so for all k there exists $l_k \geq 0$ such that $q'(k) = q(k) + l_k$, and $m_{g_0,q'(k)} = \mu * \overset{q(k)}{\cdots} * \mu * m_{g_0,l_k}$. Then,

$$
\begin{aligned}
|(m_{g_0,q'(k)} - m_{g_0,q(k)})f| &= |(\mu * \overset{q(k)}{\cdots} * \mu * m_{g_0,l_k} - \mu * \overset{q(k)}{\cdots} * \mu * \delta_{g_0})f| \\
&\leq \int_{g\in K_{k_0}} \left(\int |f(h_{q(k)}h_{q(k)-1}\dots h_1 g) - f(h_{q(k)}h_{q(k)-1}\dots h_1 g_0)| \right. \\
&\qquad\qquad \left. d\mu(h_k)d\mu(h_{k-1})\dots d\mu(h_1) \right) dm_{g_0,l_k}(g) + \epsilon/4.
\end{aligned}
$$

It follows from Proposition 26 that the inner integral converges to 0 with increasing k uniformly with respect to all $g \in K_{k_0}$. This contradicts the hypothesis that it is possible to choose a Lipschitz function f form $C_\beta(\mathfrak{S})$ that separates the measures m_q and $m_{q'}$.

So, it is a consequence of Proposition 26, and the tightness of the sequence of measures, that $m_{g_0,k}$ converges in $C^*_\beta(\mathfrak{S})$ to a unique measure m, that m is independent of the initial point g_0, and that the convergence is uniform when g_0 is chosen from a compact subset of $\mathfrak{S}$. In addition, it follows from the method of construction that m is invariant with respect to the random walk kernel, and from Proposition 15 that the support of m is contained in the set of zero rank elements, $\mathfrak{S}_0$.

Again, let $f : \mathfrak{S} \to \mathbb{R}$ be a positive function that is Lipschitz with respect to β, with Lipschitz constant L_f, and bounded by $|f| < M\beta$. Let $f_r = f \wedge M\beta_r$, so that f_r

is a strictly increasing sequence of Lipschitz functions that are integrable with respect to to the measure m. For each r,

$$\begin{aligned}\infty > m(f_r) &= \lim_{k\to\infty} m_{g_0,k}(f_r) \qquad \forall g_0 \\ &= \lim_{k\to\infty} \delta_{g_0}\Pi^k(f_r) \geq 0\end{aligned}$$

Since, by the monotone convergence theorem, $m(f_r) \to m(f)$, there exists r_0, $\epsilon > 0$ and positive integer k_0 such that for all $r > r_0$ and $k > k_0$, $\delta_{g_0}\Pi^k(f_r) \geq \epsilon$. For any $g_0 \in \mathfrak{G}$, if $w < \epsilon$ then $P_{g_0}[f_r(g_k) < w] < c < 1$ when $c = (M\beta_r - \epsilon)/(M\beta_r - w)$, and for any positive integers n_0 and p

$$\begin{aligned}P_{g_0}\left[\sum_{n=1}^{\infty} f_r(g_n) < pw\right] &< P_{g_0}\left[\sum_{n=1}^{k_0(n_0+p)} f_r(g_n) < pw\right] \\ &< P_{g_0}\left[\sum_{n=1}^{n_0+p} f_r(g_{nk_0}) < pw\right].\end{aligned}$$

The sum $\sum_{n=1}^{n_0+p} f_r(g_{nk_0}) < pw$ only if $f_r(g_{ik_0}) < \omega$ for at least n_0 values i, $1 \leq i \leq n_0 + p$. If $n < n_0$, let $\mathcal{K}_n$ denote the set of ordered mappings $k(i)$ from the integers $1 \leq i \leq n$ into the integers $k_0, 2k_0, \ldots (n_0+p)k_0$, then

$$P_{g_0}\left[\sum_{n=1}^{\infty} f_r(g_n) < pw\right] < \sum_{n=n_0}^{n_0+p} \sum_{k(i)\in\mathcal{K}_n} P_{g_0}\left[f_r(g_{k(i)}) < w \quad 1 \leq i \leq n\right] \tag{5.15}$$

The probability in the summation is estimated by

$$
\begin{aligned}
P_{g_0}[f_r(g_{k(i)}) < w, \quad 1 \le i \le n] \\
&= \int_{\{f(g_{k(i)})<w,\ 1\le i\le n\}} dP_{g_0}(g_{k(n)}, g_{k(n-1)}, \dots g_{k(1)}) \\
&= \int_{\{f(g_{k(i)})<w,\ 1\le i\le n-1\}} P_{g_{k(n-1)}}\left[f(g_{k(n)}) < w\right] \\
&\qquad\qquad dP_{g_0}(g_{k(n-1)}, g_{k(n-2)} \dots, g_{k(1)}) \\
&\le c\int_{\{f(g_{k(i)})<w,\ 1\le i\le n-1\}} dP_{g_0}(g_{k(n-1)}, g_{k(n-2)} \dots, g_{k(1)}) \\
&\le c^n
\end{aligned}
$$

substituting this estimate back in (5.15) gives

$$
\begin{aligned}
P_{g_0}\left[\sum_{n=1}^{\infty} f_r(g_n) < pw\right] &< \sum_{n=n_0}^{n_0+p} \sum_{k(i)\in\mathcal{K}_n} c^n \\
&< \sum_{n=n_0}^{n_0+p} \binom{n_0+p}{n} c^n \\
&< p\frac{(n_0+p)^p}{p!} c^{n_0}
\end{aligned}
$$

For any fixed p, the bound on the right can be made arbitrarily small by letting $n_0 \to \infty$. So, there exists $w > 0$ such that for every $r > r_0$ and for all $p > 0$,

$$
P_{g_0}\left[\sum_{k=1}^{\infty} f_r(g_k) < pw\right] = 0
$$

And a second application of the monotone convergence theorem, this time to the probability measure on the random walk, yields the result. □

Proposition 27 demonstrates a condition that is similar to Harris recurrence and potentially provides the basis for a potential theory for random walks on semigroups that mirrors the analogous theory for Harris recurrent random walks on groups [34]. This level of generality is not required in the present application because the range of the Poisson operator is restricted to a Lipschitz space. The existence of the invariant measure m, which is established by Proposition 27, and the geometric ergodicity result in Proposition 26 prove to be enough when extended to Markov modulated random walks. Eventually an application of the results from Part II, Chapter 2 of Benveniste *et al.* [7] will give a satisfactory potential theory for the purpose of establishing convergence of the stochastic approximation.

In addition to the existence of an invariant measure, and the geometric ergodicity result of Proposition 26, the stochastic approximation theory that was introduced in Chapter 4 requires that the solution to the Poisson equation, equation (4.20), with a Lipschitz function on the right hand side should be regular with respect to variation in the parameter θ. The next lemma and proposition are the key results in establishing this regularity.

Lemma 28. *Let S be a compact subset of a manifold that has a regular embedding in an N-dimensional space. Let ρ be a metric on the manifold that is compatible with the manifold topology, let d be the associated Lévy metric, and let $K = 2^N - 1$. If μ and μ' are two measures supported on S and separated by $d(\mu, \mu') < \delta$, then there exists a decomposition*

$$\mu = \sum_{\alpha} \mu_\alpha + \mu_0, \qquad \mu' = \sum_{\alpha} \mu'_\alpha + \mu'_0,$$

such that $\mu_0(S), \mu'_0(S) < K\delta$, $\mu_\alpha(S) = \mu'_\alpha(S)$ for all α, and if $g_1 \in \operatorname{supp} \mu_\alpha$ and $g_2 \in \operatorname{supp} \mu'_\alpha$ then $\rho(g_1, g_2) < K\delta$.

Proof. Tile the set S with a tessellation with the following properties:

(a) The tessellation can be partitioned into M sets, where M is a finite number independent of δ, in such a way that every pair of tiles in the same set of the partition are separated by a distance greater than $2M\delta$.

(b) Each tile in the tessellation has minimum diameter of at least $2M\delta$.

(c) Each tile in the tessellation has maximum diameter that is bounded by a fixed multiple of δ.

One way to achieve this tessellation is first, to regularly embed the manifold in a Euclidean space $\mathbb{R}^N$, and then generate the tessellation with a cubic subdivision of the N-dimensional embedding space, each cube having sides of length $2^{N+1}\delta$. The result is a tessellation with a partition of size $M = 2^N$, each cell in the partition is a cube with minimum diameter $2M\delta$, and maximum diameter $\sqrt{N}2M\delta$.

Label the M sets in the partition P_i with i taking odd values $i = 2m - 1$, $m = 1, \ldots, M$, and label the tiles in the tessellation S_α. Consider the first set in the partition, P_1, and divide the tiles in this set into two subsets. A tile S_α remains in P_1 when $\mu(S_\alpha) \geq \mu'(S_\alpha)$, and a tile S_α that satisfy the opposite inequality $\mu(S_\alpha) < \mu'(S_\alpha)$ is transferred to a new set P_2.

Consider, now the set P_1, and split this set into two subsets, Q_1 and R_1 . Place in Q_1 the tiles $S_\alpha \in P_1$ for which $\mu'(S_\alpha) \leq \mu(S_\alpha) \leq \mu'(S_\alpha^\delta)$, and place in R_1, the tiles from the P_1 that satisfy $\mu'(S_\alpha^\delta) < \mu(S_\alpha) \leq \mu'(S_\alpha^\delta) + \delta$. It follows from the characterization of the Lévy metric in Proposition 17 that every tile in P_1 ends up in either Q_1 or R_1. For every α such that $S_\alpha \in Q_1$, let $\mu_\alpha = \mu \mathbf{1}_{S_\alpha}$ and let $\mu'_\alpha = \mu' \mathbf{1}_{S_\alpha} + c_\alpha \mathbf{1}_{S_\alpha^\delta \setminus S_\alpha}$ where

$$c_\alpha = \frac{\mu_\alpha(S_\alpha) - \mu'(S_\alpha)}{\mu'(S_\alpha^\delta) - \mu'(S_\alpha)}.$$

For every α such that $S_\alpha \in R_1$, let $\mu'_\alpha = \mu' \mathbf{1}_{S_\alpha^\delta}$, let $\mu_\alpha = d_\alpha \mu_\alpha \mathbf{1}_{S_\alpha}$ where $d_\alpha = \mu'_\alpha(S_\alpha^\delta)/\mu(S_\alpha)$. Finally, let $\mu_1 = \mu - \sum_{\alpha : S_\alpha \in R_1} \mu_\alpha$, then another application of Proposition 17, the fact that any two tiles in R_1 are separated by a minimum distance greater than 2δ, and the fact that $\mu(S_\alpha) - \mu'(S_\alpha^\delta) > 0$ together establish that

$$\begin{aligned}\mu_1(\cup\{S_\alpha : S_\alpha \in R_1\}) &= \sum\{\mu(S_\alpha) - \mu'(S_\alpha^\delta) : S_\alpha \in R_1\} \\ &= \mu(\cup\{S_\alpha : S_\alpha \in R_1\}) - \mu'((\cup\{S_\alpha : S_\alpha \in R_1\})^\delta) \\ &< \delta.\end{aligned}$$

At this point in the construction there exists a positive measure μ_1, and two sets of positive measures $\{\mu_\alpha\}$, and $\{\mu'_\alpha\}$ indexed by the set $A_1 = \{\alpha : S_\alpha \in P_1\}$, with the properties that: for all $\alpha \in A_1$, $\mu_\alpha < \mu$, and $\mu'_\alpha < \mu'$, the support of each μ'_α is contained in S_α^δ (the δ neighborhood of S_α), the support of each μ_α is contained in S_α, $\mu'_\alpha(S) = \mu_\alpha(S)$, the support of μ_1 is contained in the set $\bigcup_{S_\alpha \in P_1} S_\alpha$, $\mu_1(S) \leq \delta$, and for any measurable set $F \subset \bigcup S_\alpha$, $\mu(F) = \mu_1(F) + \sum \mu_\alpha(F)$ and $\mu'(F) = \sum \mu'_\alpha(F)$.

Move now, to the tiles S_α in the set P_2. A symmetric construction to that used for the tiles in P_1 produces a positive measure μ_2, and two sets of positive measures μ_α and μ'_α indexed by the set $A_2 = \{\alpha : S_\alpha \in P_2\}$, with the properties that: for all $\alpha \in A_2$, $\mu_\alpha < \mu$, $\mu'_\alpha < \mu'$, the support of each μ_α is contained in S_α^δ, the support of each μ'_α is contained in S_α, $\mu'_\alpha(S) = \mu_\alpha(S)$, the support of μ_2 is contained in the set $\bigcup_{S_\alpha \in P_2} S_\alpha$, $\mu_2(S) \leq \delta$, and for any measurable set $F \subset \bigcup S_\alpha$, $\mu(F) = \sum \mu_\alpha(F)$ and $\mu'(F) = \mu_2(F) + \sum \mu'_\alpha(F)$.

Redefine μ as $\mu - \sum_{\alpha \in A_1} \mu_\alpha - \mu_1$ and μ' as $\mu' - \sum_{\alpha \in A_2} \mu'_\alpha - \mu_2$. Now, $d(\mu, \mu') < 2\delta$ and the same procedure as before results in measures μ_3, μ_4, μ_α and μ'_α with support in the sets associated with partition P_3. Continue in this way through all of the remaining partitions P_5 through P_{2M-1}. Let $\mu_0 = \mu_1 + \mu_3 \cdots + \mu_{2M-1}$ and $\mu'_0 = \mu_2 + \mu_4 \cdots + \mu_{2M}$, then μ_0, μ'_0, μ_α, and μ'_α provide the required partitions. If $K = 2^N - 1$ then the measures $\mu_0(S)$ and $\mu'_0(S)$ are both bounded by $K\delta$, and for any α, if $g_1 \in \operatorname{supp} \mu_\alpha$, and $g_2 \in \operatorname{supp} \mu'_\alpha$, then $\rho(g_1, g_2) < K\delta$. $\square$

Proposition 29. *Let $\mathfrak{S}$ be a semigroup that is a finite dimensional manifold. Let ρ be a metric that is compatible with the manifold topology, and suppose that left multiplication in the semigroup is r-strongly contractive with respect to ρ. Let β be a positive weighting function on $\mathfrak{S}$ that has compact sub-level sets, and is bounded away from zero. Let μ_l be a sequence of measures on $\mathfrak{S}$ with support contained in a compact set S of diameter D, and let μ_l satisfy the following two properties:*

(i) There exists $\delta > 0$ such that if μ_l' is a second sequence, and $\forall l$, $\operatorname{supp} \mu_l' \subset S$, and $d(\mu_l, \mu_l') < \delta$, then μ_l' is p-strongly primitive with respect to β.

(ii) The inhomogeneous random walk g_k with initial point $g_0 \in \mathfrak{S}$, and transition kernel $\Pi(g_l; g_{l+1}) = d(\delta_{g_l} * \mu_l)(g_{l+1})$ is strongly contractive.

For $g_0 \in K$, define $m_{g_0,k} = \delta_{g_0} * \mu_1 * \cdots * \mu_k$.

Given $g_0 \in \mathfrak{S}$, and $\{\mu_l',\ l = 1, 2, \ldots\}$, a second sequence of measures that satisfies condition (i), let g_l' be the non-homogeneous random walk with initial point g_0 and transition kernels $\Pi(g_l; g_{l+1}) = d(\delta_{g_l} * \mu_l')(g_{l+1})$, and define a second sequence of measures $m_{g_0,k}' = \delta_{g_0} * \mu_1' * \overset{k}{\cdots} * \mu_k'$.

There exists a constant C such that for $g_0 \in \mathfrak{S}$, and any function f that is Lipschitz with respect to β with Lipschitz constant L_f,

$$|m_{g_0,k}(f) - m_{g_0,k}'(f)| \leq \delta C L_f \beta(g_0). \tag{5.16}$$

Proof. The proof begins with an expansion of the left hand side of the inequality (5.16) as an integral.

$$
\begin{aligned}
&|m_{g_0,k}(f) - m'_{g_0,k}(f)| \\
&\qquad = \left| \int f(h_k \ldots h_1 g_0) d\mu_k(h_k) \ldots d\mu_1(h_1) - d\mu'_k(h_k) \ldots d\mu'_1(h_1) \right| \\
&\qquad = \left| \sum_{l=1}^{k} \int f(h_k \ldots h_1 g_0) d\mu_k(h_k) \ldots d\mu_l(h_l) d\mu'_{l-1}(h_{l-1}) \ldots d\mu'_1(h_1) \right. \\
&\qquad\qquad\qquad \left. - d\mu_k(h_k) \ldots d\mu_{l+1}(h_{l+1}) d\mu'_l(h_l) \ldots d\mu'_1(h_1) \right|
\end{aligned}
\tag{5.17}
$$

Let $g_{l-1} = h_{l-1} \ldots h_1 g_0$, then an application of Fubini's theorem to the l^{th} term in the summation yields an iterated integral

$$
\int \left[\int f(h_k \ldots h_l g_{l-1}) d\mu_k(h_k) \ldots d\mu_{l+1}(h_{l+1}) (d\mu_l(h_l) - d\mu'_l(h_l)) \right] dm'_{g_0,l-1}(g_{l-1}).
$$

Write the inner integral as

$$
\begin{aligned}
I_l(g_{l-1}) = &\int f(h_k, \ldots, h_l g_{l-1}) d\mu_k(h_k) \ldots d\mu_{l+1}(h_{l+1}) d\mu_l(h_l) \\
&- \int f(h_k, \ldots, h_l g_{l-1}) d\mu_k(h_k) \ldots d\mu_{l+1}(h_{l+1}) d\mu'_l(h_l)
\end{aligned}
\tag{5.18}
$$

By hypothesis, for all l, the measures μ_l and μ'_l are separated in the Lévy metric by $d(\mu_l, \mu'_l) < \delta$, and an application of Lemma 28 produces a decomposition for the measures μ_l and μ'_l.

$$
\mu_l = \sum_{\alpha} \mu_\alpha + \mu_0, \qquad \mu'_l = \sum_{\alpha} \mu'_\alpha + \mu'_0
$$

where $\mu_0(\mathfrak{S}), \mu'_0(\mathfrak{S}) < C_1\delta$, $\mu_\alpha(\mathfrak{S}) = \mu'_\alpha(\mathfrak{S})$ for all α, and if $h_l \in \operatorname{supp} \mu_\alpha$ and $h'_l \in \operatorname{supp} \mu'_\alpha$ then $\rho(h_l, h'_l) < C_2\delta$. Note that the particular decomposition varies with l, but this dependency on l is dropped in the notation.

Substituting the decomposition into (5.18) gives

$$|I_l(g_{l-1})| = \Bigg| \sum_\alpha \Bigg(\int f(h_k \dots h_l g_{l-1}) d\mu(h_k) \dots d\mu(h_{l+1}) d\mu_\alpha(h_l)$$
$$- \int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu'_\alpha(h_l)) \Bigg)$$
$$+ \int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_0(h_l)$$
$$- \int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu'_0(h_l) \Bigg| \tag{5.19}$$

Using the equality $\mu_\alpha(\mathfrak{S}) = \mu'_\alpha(\mathfrak{S})$, the summand in the summation can be rewritten as

$$\int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_\alpha(h_l)$$
$$- \int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu'_\alpha(h_l))$$
$$= \frac{1}{\mu_\alpha(\mathfrak{S})} \int \left(\int f(h_k \dots h_l g_{l-1}) - f(h_k \dots h'_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) \right) d\mu_\alpha(h_l) d\mu'_\alpha(h'_l) \tag{5.20}$$

The Lipschitz property of f allows the inner integral to be bounded by a function of g_{l-1} as follows;

$$\left| \int (f(h_k \dots h_l g_{l-1}) - f(h_k \dots h'_l g_{l-1})) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) \right|$$
$$\leq L_f \int \rho(h_k \dots h_l g_{l-1}, h_k \dots h'_l g_{l-1}) \max_{h \in \{h_l, h'_l\}} 2\beta(h_k \dots h g_{l-1})$$
$$d\mu(h_k) \dots d\mu(h_{l+1})$$

Since left multiplication in the semigroup is contractive, it follows from Lemma 14 and the conditions on the supports of μ_α and μ'_α that

$$\rho(h_k \dots h_l g_{l-1}, h_k \dots h'_l g_{l-1}) \leq c^{k-l+1} \delta$$

and

$$\left|\int (f(h_k \dots h_l g_{l-1}) - f(h_k \dots h_l' g_{l-1})) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1})\right| \\ \leq C c^{k-l-1} \delta L_f B_l(g_{l-1}) \quad (5.21)$$

with

$$B_l(g_{l-1}) = 2 \int \max_{h \in S} \beta(h_k \dots h g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}).$$

The bound in (5.21) is independent of h_l, so summing equation (5.20) over α gives a bound on the summation on the right hand side of (5.19) of

$$\left|\sum_\alpha \left(\int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_\alpha(h_l) \right.\right. \\ \left.\left. - \int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_\alpha'(h_l))\right)\right| \\ \leq C c^{k-l-1} \delta L_f B_l(g_{l-1}) \quad (5.22)$$

A similar argument produces a bound on the last two terms on the right hand side of equation (5.19).

$$\left|\int f(h_k \dots h_l g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_0(h_l) \right. \\ \left. - \int f(h_k \dots h_l' g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_0'(h_l')\right| \\ \leq \frac{1}{\delta} L_f \int \rho(h_k \dots h_l g_{l-1}, h_k \dots h_l' g_{l-1})(\beta(h_k \dots h_l g_{l-1}) + \beta(h_k \dots h_l' g_{l-1})) \\ \times d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_0(h_l) d\mu_0'(h_l')$$

This time $\rho(h_l, h_l') < D$, and

$$\rho(h_k \dots h_l g_{l-1}, h_k \dots h_l' g_{l-1}) \leq c^{k-l-1} D$$

so, using the bounds $\mu_0(\mathfrak{S}) < \delta$ and $\mu_0'(\mathfrak{S}) < \delta$,

$$\begin{aligned}\Bigg|\int f(h_k \dots h_l g_{l-1})\, d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_0(h_l) \\ - \int f(h_k \dots h_l' g_{l-1}) d\mu_k(h_k) \dots d\mu_{l+1}(h_{l+1}) d\mu_0'(h_l')\Bigg| \\ \leq \frac{1}{\delta} c^{k-l-1} D B_l(g_{l-1}) L_f \mu_0(\mathfrak{S}) \mu_0'(\mathfrak{S}) \\ \leq C c^{k-l-1} \delta L_f B_l(g_{l-1}), \end{aligned} \tag{5.23}$$

for some constant C that is independent of l.

Combining (5.22) and (5.23) produces a bound in (5.19) produces

$$I_l(g_{l-1}) \leq C c^{k-l-1} \delta L_f B_l(g_{l-1}),$$

and a corresponding bound on the l'th term in the summation on the right side of (5.17) of

$$\begin{aligned}\int I_l(g_{l-1})\, dm_{g_0,l-1}'(g_{l-1}) &\leq C c^{k-l-1} \delta L_f \int B_l(g_{l-1})\, dm_{g_0,l-1}'(g_{l-1}) \\ &\leq C c^{k-l-1} \delta L_f \big(1 - a\sqrt[p]{\eta}\big)^{-1} \beta(g_0) \end{aligned} \tag{5.24}$$

With each term in the summation on the right hand side of (5.17) bound by (5.24), summation over l gives a bound for the left hand side of

$$|m_{g_0,k}(f) - m_{g_0,k}'(f)| \leq C \delta L_f \big(1 - a\sqrt[p]{\eta}\big)^{-1} \beta(g_0) \frac{c}{1-c}, \tag{5.25}$$

and a suitable redefinition of the constant C completes the proof. □

The following corollary to Proposition 29 is a consequence of Proposition.

Corollary 30. *With the same conditions on $\mathfrak{S}$ as imposed in Proposition 29, let μ' and μ' be two measures on $\mathfrak{S}$ that both satisfy the conditions imposed on μ in the premise of Proposition 29. Let m and m' be invariant measures with respect to the Markov* *transition kernels* $\Pi(g_a; dg_b) = d(\delta_{g_a} * \mu)(g_b)$ *and* $\Pi'(g_a; dg_b) = d(\delta_{g_a} * \mu')(g_b)$*, the existence of which are predicated in the conclusion of Proposition 27. If* μ *and* μ' *are separated in the Lévy metric by* $d(\mu, \mu') < \delta$*, then there exists a constant* C*, independent of* δ *such that if* f *is a Lipschitz function with Lipschitz constant* L_f*,*

$$|m(f) - m'(f)| \leq CL_f\delta$$

Proof. A direct application of the Proposition gives the bound

$$|m_{g_0,k}(f) - m'_{g_0,k}(f)| \leq CL_f\delta\beta(g_0).$$

However, it follows from Proposition 27 that both $m_{g_0,k}$ and $m'_{g_0,k}$ have weak limits that are independent of the initial point g_0, and as a consequence, taking the limit as $k \to \infty$, and redefining the constant C gives

$$|m(f) - m'(f)| \leq CL_f\delta$$

□

The proposition also provides a method to deal with perturbations of random walks that introduce a small amount of dependence between successive increments.

Corollary 31. *Let* μ *be the generator for a Markov random walk on a semigroup* $\mathfrak{S}$*, and let* μ *and* $\mathfrak{S}$ *satisfy the conditions of Proposition 29. Let* $h \to \mu(h)$ *be a mapping from* $\mathfrak{S}$ *to* $C^\star_\beta(\mathfrak{S})$ *such that for all* $h \in \mathfrak{S}$*,* $d(\mu, \mu(h)) < \delta$*, and define a second (non-Markov) random walk with transition kernel* $\Pi(g_l; dg_{l+1}) = \delta_{g_l} * \mu(h_{l-1})$*, where* h_{l-1} *is the increment between* g_{l-1} *and* g_l*. Then, if the measure* $m_{g_0,k}$ *is the probability distribution of the* k*'th point in the walk with generator* μ *conditioned on the initial point* g_0*, and* $m'_{g_0,k}$ *the corresponding distribution for the walk with transition kernels constructed from the measures* $\mu(h_l)$*, then*

$$|m_{g_0,k}(f) - m'_{g_0,k}(f)| \leq CL_f\delta\beta(g_0)$$

Proof. The proof of the corollary is already contained in the proof of the theorem. ☐

A second consequence of Lemma 28 is the following Proposition

Proposition 32. *Let* $\Pi(g_a; dg_b) = (\delta_{g_a} * \mu)(dg_b)$ *be a transition kernel of a contractive random walk on an* M*-dimensional semigroup* $\mathfrak{S}$*, and let* μ *have compact support. Define a mapping* $\Pi : C^*_\beta(\mathfrak{S}) \to C^*_\beta(\mathfrak{S})$ *by*

$$m_a(dg_a) \to \int_{\mathfrak{S}} m_a(dg_a)\Pi(g_a; dg_b) = m_b(dg_b)$$

Π *is a uniformly continuous mapping with respect to the Lévy metric of the subspace of compactly supported measures onto itself. Further, if* $\mathcal{P} = \{\Pi_\mu\}$ *is the family of such contractive kernels indexed by the compactly supported generator* μ*, then* $\mathcal{P}$ *is an equicontinuous family.*

Proof. Let m_a and m'_a be two compactly supported measures on $\mathfrak{S}$ that are separated by $d(m_a, m'_a) = \delta$. Let $K = 2^M - 1$, and let

$$m'_a = \sum_\alpha m'_{a\alpha} + m'_{a0} \qquad m_a = \sum_\alpha m_{a\alpha} + m_{a0} \tag{5.26}$$

Be two decompositions that having the properties of the decompositions in the conclusion of Lemma 28. Since the map induced by the kernel Π is linear, m_b and m'_b have decompositions

$$m'_b = \sum_\alpha m'_{b\alpha} + m'_{b0} \qquad m_b = \sum_\alpha m_{b\alpha} + m_{b0}$$

where $m_{b\alpha} = m_{a\alpha} * \mu$ and $m'_{b\alpha} = m'_{a\alpha} * \mu$. The following conclusions are immediate

$$m_{b\alpha}(\mathfrak{S}) = m'_{b\alpha}(\mathfrak{S}) \quad \text{and} \quad m_{b0}(\mathfrak{S}) = m'_{b0}(\mathfrak{S}) < K\delta. \tag{5.27}$$

In addition, let E be a measurable subset of $\mathfrak{S}$, and let $g_b = hg_a \in E$ for some $h \in \operatorname{supp} \mu$, and $g_a \in m_{a\alpha}$. Since left multiplication by an element of the support of μ is a contraction in $\mathfrak{S}$, it follows that $g'_b = hg'_a \in E^{K\delta}$ for all g'_a that satisfy $\rho(g'_a, g_a) < K\delta$. But it is a property of the decompositions in (5.26) that all $g'_a \in \operatorname{supp} m'_{a\alpha}$ lie within $K\delta$ of g_a. As a consequence of this argument, the analogous argument with the roles of primed and non-primed variables reversed, and the first of the conclusions in (5.27)

$$
\begin{aligned}
& m_{b\alpha}(E) \leq m'_{b\alpha}(E) \leq m_{b\alpha}(E^{K\delta}) \\
\text{or} \qquad & m'_{b\alpha}(E) \leq m_{b\alpha}(E) \leq m'_{b\alpha}(E^{K\delta})
\end{aligned}
$$

And summing over the index α gives

$$
\begin{aligned}
& m_b(E) \leq m'_b(E) \leq m_b(E^{K\delta}) \\
\text{or} \qquad & m'_b(E) \leq m_b(E) \leq m'_b(E^{K\delta})
\end{aligned}
$$

This statement together with the second of the conclusions in (5.27) leads to the conclusion that $d(m_b, m'_b) < K\delta$ and therefore the first conclusion of the proposition: that the mapping Π is continuous. The second conclusion of the Proposition, that $\mathcal{P}$ is an equicontinuous family, is a consequence of the independence of the constant K with respect to the kernel Π. $\square$

In the theorems that follow $T_k = (S_k, g_k)$ is a Markov modulated random walk. The random process S_k, is a discrete time Markov process that takes values in a metric space $(\mathcal{S}, \rho_{\mathcal{S}})$, and has transition kernel $\Pi(S_a; dS_b)$. The process g_k takes values in the semigroup $\mathfrak{S}$ with metric $\rho_{\mathfrak{S}}$, and has dynamics determined by the equation $g_k = g(S_k)g_{k-1}$, where $g : \mathcal{S} \to \mathfrak{S}$ is a bounded Lipschitz continuous function with respect to ρ_S and $\rho_{\mathfrak{S}}$. The function g generates continuous maps $d_\star g : C^*_{\beta_S}(S) \to C^*_{\beta_{\mathfrak{S}}}(\mathfrak{S})$ and $d^\star g : C_{\beta_{\mathfrak{S}}}(\mathfrak{S}) \to C_{\beta_S}(S)$, where $\beta_{\mathcal{S}} : \mathcal{S} \to \mathbb{R}$ and $\beta_{\mathfrak{S}} : \mathfrak{S} \to \mathbb{R}$ are weight functions on the metric space $\mathcal{S}$ and the semigroup $\mathfrak{S}$.

Theorem 33. *Let $\Pi(S_a; dS_b)$ be the transition kernel for a Markov process S_k on $\mathcal{S}$ with an invariant measure ν. Let ν_0 be the marginal distribution for the initial value S_0, and let S_k satisfy the geometric ergodicity condition*

$$d_{\mathcal{S}}(\Pi^k(S_0; dS_k), \nu) < C_1 c^k \beta_{\mathcal{S}}(S_0)$$

and the regularity condition $d_{\mathcal{S}}(\nu_1 \Pi, \nu_2 \Pi) < C d_{\mathcal{S}}(\nu_1, \nu_2)\ \forall \nu_1, \nu_2 \in C^*_\beta(S)$.

Suppose that the map $g : \mathcal{S} \to \mathfrak{G}$ and the chain S_k generate a Markov random walk with transition kernel

$$\Pi(S_k, g_k; dg_{k+1}) = \delta_{g_k} * [d_\star g](\Pi(S_k; dS_{k+1})),$$

and suppose also that the sequence of measures $[d_\star g](\Pi(S_k; dS_{k+1}))$ is p-strongly primitive and r-strongly contractive with probability 1.

Then, there exists a measure m, and positive constants C and c with $0 \leq c < 1$ such that for any Lipschitz function $f : \mathfrak{G} \to \mathbb{R}$ with Lipschitz constant L_f

$$|\Pi^k(S_0, g_0; dg_k)(f) - m(f)| \leq C L_f c^k \beta_{\mathfrak{G}}(g_0) \beta_S(S_0)$$

Proof. The proof uses an appropriate grouping of increments to show that the Markov modulated random walk is well approximated by a second random walk with independent, identically distributed increments.

Let f be Lipschitz with respect to β. The quantity $\Pi^k(S_0, g_0; dg_k)(f)$ can be explicitly written as the iterated integral

$$\int f(g(S_k)g(S_{k-1}) \ldots g(S_1)g_0)\Pi(S_0; dS_1) \ldots \Pi(S_{k-2}; dS_{k-1})\Pi(S_{k-1}; dS_k) \quad (5.28)$$

Write $k = jl + r$. The interpretation of this decomposition is that the random walk is allowed to run for r increments, and after that the increments are compounded in groups of l. Use this decomposition to rewrite the integral in (5.28) as

$$\int f(h^l_j h^l_{j-1} \dots h^l_1 g_r)\Pi^l(S_{l(j-1)+r}; d(S_{lj+r}, h^l_j))\Pi^l(S_{l(j-2)+r}; d(S_{l(j-1)+r}, h^l_{j-1}))$$
$$\dots \Pi^l(S_r; d(S_{l+r}, h^l_1))\Pi^r(S_0, g_0; d(S_r, g_r))$$

where

$$h^l_i = g(S_{il+r})g(S_{il+r-1}) \dots g(S_{i(l-1)+r+1}) \quad \text{for } 1 \le i \le j. \tag{5.29}$$

For any $S_0 \in \mathcal{S}$, the kernel $\Pi^l(S_0; d(S_l, h^l))$ acts on a Lipschitz function $f : \mathfrak{S} \times \mathcal{S} \to \mathbb{R}$ by the formula

$$\int f(S_l, h^l)\Pi^l(S_0; d(S_l, h^l)) = \int f(S_l, g(S_l) \dots g(S_1))\Pi(S_{l-1}; dS_l) \dots \Pi(S_0; dS_1).$$

The essential steps in the proof compute appropriate values for the remainder r and the divisor l.

The value of r determines the the length of the initial transient. If S_a, the initial value of the Markov process, is restricted to a compact subset of $\mathcal{S}$, then during the transient the sequences of measures $\Pi^j(S_a; dS_b)$, $0 \le j \le r$, converge to a ball of radius δ about ν, the invariant measure of the transition kernel $\Pi(S_a; dS_b)$. The convergence is uniform in the initial value S_a. Meanwhile, the position of the Random walk $h_j h_{j-1} \dots g_0$, $0 \le j \le r$ is controlled by the a-priori bounds of Lemma 9. This bound is a result of p-strong primitivity of the Markov random walk, and a bound on the range of the map $g : \mathcal{S} \to \mathfrak{S}$.

Lemma 34. *For all $\delta > 0$ there exists r, sufficiently large that for all $S_0 \in \mathcal{S}$, and for all $k \ge r$, $d_{\mathcal{S}}(\Pi(S_0; dS_k), \nu) < \delta$. Also, almost surely, $\beta_{\mathfrak{S}}(g_r) < C/\delta \ \beta_{\mathcal{S}}(S_0)\beta_{\mathfrak{S}}(g_0)$*

Proof. A premise of the theorem states that there exists a real constant $C > 0$ and a positive integer c such that for all $S_0 \in \mathbb{S}$,

$$d_{\mathbb{S}}(\Pi^k(S_0; dS_k), \nu) \leq Cc^k \beta_{\mathbb{S}}(S_0)$$

Choose r so that

$$Cc^r \beta_{\mathbb{S}}(S_0) \leq \delta \leq Cc^{r-1} \beta_{\mathbb{S}}(S_0),$$

then the condition on the range of $g : \mathbb{S} \to \mathfrak{G}$ implies that

$$\begin{aligned} \beta_{\mathfrak{G}}(g_r) &\leq a^r \beta_{\mathfrak{G}}(g_0) \\ &\leq \frac{1}{cc^{r-1}} \beta_{\mathfrak{G}}(g_0) \\ &\leq \frac{C}{\delta} \beta_{\mathbb{S}}(S_0) \beta_{\mathfrak{G}}(g_0) \end{aligned}$$

with an appropriate re-definition of the constant C. □

Turning now to the divisor l, define the l-cumulants of the random walk by

$$h_i^l = h_{il+r} h_{il+r-1} \ldots h_{(i-1)l+r+2} h_{(i-1)l+r+1}, \quad i \geq 1.$$

The value of the divisor l should be large enough that, under the process distribution for the Markov modulated random walk, successive l-cumulants are close to being independently and identically distributed.

First an auxiliary lemma.

Lemma 35. *Under the conditions of the premise of the theorem the probability measures $\nu_{0,k}$ that describe the joint distributions of S_0 and S_k satisfy*

$$d_\beta(\nu_{0,k}, \nu_0 \otimes \nu) \leq C_2 c^k$$

Proof. Let $f \in C_\beta(\mathcal{S}^2)$ with $|f| \le \beta$, then

$$\left| \int f(S_0, S_k)\, d\nu_{0,k}(S_0, S_k) - \int f(S_0, S_k)\, d\nu_0(S_0) d\nu(S_k) \right|$$
$$= \left| \int \big(f(S_0, S_k)\, \Pi^k(S_0; dS_k) - f(S_0, S_k)\, d\nu(S_k) \big) d\nu_0(S_0) \right|$$
$$\le \int C_1 c^k \beta(S_0)\, d\nu_0(S_0) = C_2 c^k,$$

when $C_2 = C_1 \nu_0(\beta)$. Since the only condition on the choice of f is the uniform bound, the result follows. ☐

Lemma 36. *For all k, let $\nu_k = \Pi(S_0; dS_k)$ denote the distribution of the random variable S_k. For l sufficiently large, there exists a measure μ^l on $\mathfrak{G}$, and constants $\epsilon, \delta > 0$ such that if $d_{\mathcal{S}}(\nu_{il+r}, \nu) < \delta$, then*

$$d_{\mathfrak{G}}(\Pi^l(S_{il+r}, g_{il+r}; dg_{(i+1)l+r}), d[\delta_{g_{il+r}} * \mu^l](g_{(i+1)l+r})) < \epsilon$$

almost surely with respect to the distribution of $g_{il+r} \in \mathfrak{G}$ and $S_{il+r} \in \mathcal{S}$.

Proof. Renumber the indices so that the index $il + r$ becomes 0, and the index $l(i + 1) + r$ becomes l, and partition the interval l as $l = l_1 + l_2$. The proof rests on two claims.

Claim 1: For all $\delta > 0$, there exists l_1 such that the joint distribution of S_0, and $h_{l_1} = g(S_{l_1})$ satisfies

$$d(\nu_0(S_0) d_\star g[\Pi^{l_1}(S_0; \cdot)](h_l), d\nu_0(S_0) \otimes d_\star g[\nu](h_{l_1})) \le \delta, \tag{5.30}$$

where ν is the invariant distribution for the kernel $\Pi(S_a; dS_b)$.

Claim 2: Let $\mathcal{M} \subset C^*_\beta(\mathfrak{G})$ be a tight family of measures. For all η there exists l_2, a measure μ_{l_2}, and a constant C that depends on l_2, such that for all conditional measures $m(\cdot|\cdot)$ taking values $m(\cdot|S) \in \mathcal{M}$, and for all δ, if $d(\nu', \nu) < \delta$ then

$$d\big((\nu' m(\cdot|\cdot))\Pi^{l_2}(S_0, g^0; dh^{l_2}), \mu_{l_2}\big) < C\delta + \eta. \tag{5.31}$$

The following argument uses the two claims to prove the lemma. The assumptions of p-primitivity for the generators $d_\star g[\nu_k]$, and boundedness for the map g, together with Lemma 25 ensure that the measures m_{l_1} lie in a tight family, irrespective of the value of l_1. So, given $\delta > 0$, choose $\eta = \delta/2$ and use Claim 2 to compute l_2, μ_{l_2} and C such that the inequality (5.31) holds when $m = m_{l_1}$, and $\nu' = \nu_{l_1}$. Now choose δ such that $C\delta = \epsilon/4$, and use Claim 1, and the assumption that $d_{\mathcal{S}}(\Pi^k(S_0; dS_{l_1}), \nu) < Cc^k \beta_{\mathcal{S}}(S_0)$ to compute a value l_1 such that both the inequality (5.30) holds, and $d(\nu_{l_1}, \nu) < \delta$. Since the map $\Pi_2^l(S_{l_1}, \mathfrak{e}; d(h^{l_2}) : C_\beta(\mathfrak{S}) \to C_\beta(\mathfrak{S})$ is a contraction with respect to Lipschitz seminorms, Proposition 21 implies that the conclusion of the lemma is true with $l = l_1 + l_2$ and $\mu_l = \mu_{l_2}$.

Proof of claim 1: To prove the first claim, choose $f : \mathcal{S} \times \mathfrak{S} \to \mathbb{R}$ to be a function that is Lipschitz with respect to $\beta_{\mathcal{S}}$ in the first variable, and Lipschitz with respect to $\beta_{\mathfrak{S}}$ in the second variable. Define a second function $e : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ by $e(S_a, S_b) = f(S_a, g(S_b))$. Since $g : \mathcal{S} \to \mathfrak{S}$ is Lipschitz, the function e is also Lipschitz. Allowing the two measures that form the argument to the Lévy metric on the left hand side of (5.30) to act on the function f gives

$$\int f(S_0, g_{l_1}) d\nu_0(S_0) d_\star g[\Pi^{l_1}(S_0; \cdot)](g_{l_1}) = \int e(S_0, S_l) d\nu_0(S_0) \Pi^l(S_0; dS_{l_1})$$

and

$$\int f(S_0, g_{l_1}) d\nu_0(S_0) d_\star g[\nu(\cdot)](g_{l_1}) = \int e(S_0, S_l) d\nu_0(S_0) d\nu(S_{l_1}).$$

If $\nu_{0,k}$ and $\nu_0 \otimes \nu$ are written out explicitly as $d\nu_0(S_0)\Pi^l(S_0; dS_{l_1})$ and $d\nu_0(S_0)d\nu(S_{l_1})$ then Lemma 35 states that

$$d(d\nu_0(S_0)\Pi^l(S_0; dS_{l_1}), d\nu_0(S_0)d\nu(S_{l_1})) < Cc^{l_1}$$

for some constants $C \geq 0$ and $0 \leq c < 1$. Therefore, it follows from Lemma 18 that for every value $\delta > 0$, there exists a value for l_1 sufficiently large that inequality (5.30) holds.

Proof of claim 2: Let m be a measure in $C^*_\beta(\mathfrak{S})$, and ν and ν' be two measures in $C^*_\beta(\mathcal{S})$ and let $\Pi((S_a, g_a); d(S_b, g_b))$ be a transition kernel for a Markov modulated random walk, then

$$\begin{aligned}
&\int f(h^l)\,(\nu'\otimes m)\Pi^{l-k}((S_k, h^k); d(S_l, h^l)) - \int f(h^l)\,(\nu\otimes m)\Pi^{l-k}((S_k, h^k); d(S_l, h^l)) \\
&= \sum_\alpha \Bigg(\int f(h^l)\,\delta_{S'_k}(dS_k)\Pi(S_k; dS_{k+1})\,\nu'_\alpha(dS'_k)(m * d_\star g[\delta_{S'_k}])(dh^{k+1})\,\Pi^{l-k-1} \\
&- \frac{1}{\nu_\alpha(\mathcal{S})} \int f(h^l)\,\nu_\alpha(dS_k)\Pi(S_k; dS_{k+1})\,\nu'_\alpha(dS'_k)(m * d_\star g[\delta_{S'_k}])(dh^{k+1})\,\Pi^{l-k-1} \\
&+ \frac{1}{\nu_\alpha(\mathcal{S})} \int f(h^l)\,\nu_\alpha(dS_k)\Pi(S_k; dS_{k+1})\,\nu'_\alpha(dS'_k)(m * d_\star g[\delta_{S'_k}])(dh^{k+1})\,\Pi^{l-k-1} \\
&\quad - \int f(h^l)\,\nu_\alpha(dS_k)\Pi(S_k; dS_{k+1})\,\delta_{S_k}(dS'_k)(m * d_\star g[\delta_{S'_k}])(dh^{k+1})\,\Pi^{l-k-1} \Bigg)
\end{aligned}$$

Performing the integration with respect to S_k in the first pair of terms, and with respect to S'_k in the second pair of terms gives, with condensed notation,

$$\begin{aligned}
&\int f(h^l)\,(\nu'\otimes m)\Pi^{l-k}((S_k, h^k); d(S_l, h^l)) - \int f(h^l)\,(\nu\otimes m)\Pi^{l-k}((S_k, h^k); d(S_l, h^l)) \\
&\quad = \sum_\alpha \int \left(\int f(h^l)\,\hat{\nu}'_\alpha \otimes \tilde{m}'\Pi^l_{k+1} - \int f(h^l)\hat{\nu}_\alpha \otimes \tilde{m}'\Pi^l_{k+1} \right) \nu_\alpha(dS'_k) \\
&\quad + \sum_\alpha \int \left(\int f(h^l)\,\tilde{\nu}_\alpha \otimes \hat{m}'\Pi^l_{k+1} - \int f(h^l)\tilde{\nu}_\alpha \otimes \hat{m}\Pi^l_{k+1} \right) \nu'_\alpha(dS_k) \qquad (5.32)
\end{aligned}$$

The measures that are factors in the tensor products on the right hand side are defined as follows. In the first term, the measures are functions of S'_k,

$$\hat{\nu}'_\alpha = \Pi(S'_k; dS_{k+1}), \qquad \hat{\nu}_\alpha = \frac{1}{\nu_\alpha(\mathcal{S})} \int \Pi(S_k; dS_{k+1})\,d\nu_\alpha(S_k),$$

$$\text{and} \qquad \tilde{m}' = m * \delta_{g(S'_k)}.$$

While in the second term, the measures are functions of S_k,

$$\tilde{\nu}'_\alpha = \Pi(S_k; dS_{k+1}), \qquad \hat{m}' = \frac{m * d_\star g[\nu'_\alpha]}{\nu_\alpha(\mathcal{S})}, \quad \text{and} \quad \hat{m} = m * \delta_{g(S_k)}.$$

Write the left hand side of equation (5.32) as $[T^{l-k}(\delta R)](f)$, where $l-k$ is the number of terms in Π^{l-k}, the iterated kernel, δ is the bound on the distance $d(\nu, \nu')$, and R is a bound on the support of m. Examine the expressions for the measures $\hat{\nu}$ and $\hat{\nu}'$. It is true that for any α, and for any $S'_k \in \operatorname{supp} \nu'_\alpha$, $d(\delta_{S'_k}, \nu_\alpha/\nu_\alpha(\mathcal{S})) < Kd(\nu', \nu)$, and therefore $d(\hat{\nu}', \hat{\nu}) < C\delta$ for some constant C that is independent of ν and ν'. The premise of the proposition contains the assumption that the kernel $\Pi(S_a; dS_b)$ is contractive with respect to the Lévy metric. It is a consequence of this assumption that the magnitude of the first term on the right hand side of (5.32) is bounded by $[T^{l-k-1}(C_1\delta R)](f)$ for some constant C_1.

Examine the expressions for the measures $\hat{m}$ and $\hat{m}'$. The bound $d(\delta_{S'_k}, \nu_\alpha/\nu_\alpha(\mathcal{S})) < Kd(\nu', \nu)$, and Proposition 32 implies that $d(\hat{m}, \hat{m}') < C\delta$. Since $\Pi(S_a, g_a; d(S_b, g_b))$ is a kernel for a contractive random walk, it follows again from Proposition 32 that the second term in on the right hand side of (5.32) is bounded by $C_2\delta$.

An application of these observations to the expression in equation (5.32) produces the recursive bound

$$[T^{l-k}(\delta R)](f) \leq [T^{l-k-1}(C_1\delta R)](f) + C_2\delta$$

Computing the recursion for k iterations gives the bound

$$[T^{l-k}(\delta R)](f) \leq C_1^k \delta R + \sum_{i=1}^{k} C_2 C_1^{i-1}\delta < C_1^k C_3 R\delta.$$

Now let $m_0(\cdot|\cdot) : \mathcal{S} \to C^*_\beta(\mathfrak{S})$ be a conditional measure that takes values in the tight family of measures $\mathcal{M}$. Let ρ denote the metric on $\mathfrak{S}$. The tightness condition means that there exists a compact set $K \subset \mathfrak{S}$ such that $m(\cdot|S)[\beta \mathbf{1}_{\mathfrak{S}\setminus K}] < \epsilon/2$ outside a set of measure 0 on $\mathcal{S}$. Let $\mu_{l_2} = \nu\Pi^{l_2}(S_0, \mathfrak{e}; dh^{l_2})$. The left hand side of (5.31), written as an integral, is bounded by

$$\begin{aligned}
&\left| \int f(h^{l_2})\nu'(dS_0)m(dh_0|S_0))\Pi^{l_2}(S_0, h^0; dh^{l_2}) - \int f(h^{l_2})\mu_{l_2}(dh^{l_2}) \right| \\
&\leq \left| \int f(h^{l_2})\nu'(dS_0)m(dh_0|S_0))\Pi^{l_2}(S_0, h^0; dh^{l_2}) \right. \\
&\qquad\qquad \left. - \int f(h^{l_2})\nu'(dS_0)\Pi^{l_2}(S_0, \mathfrak{e}; dh^{l_2}) \right| \\
&\quad + \left| \int f(h^{l_2})\nu'(dS_0)\Pi^{l_2}(S_0, \mathfrak{e}; dh^{l_2}) - f(h^{l_2})\nu(dS_0)\Pi^{l_2}(S_0, \mathfrak{e}; dh^{l_2}) \right|
\end{aligned}$$

The first difference on the right hand side has the form of the left hand side of equation (5.32). Since K is compact $D = \sup\{d(\mathfrak{e}, g) : g \in K\}$ is finite, and the r-contractive property of the Markov modulated random walk provides a bound for the second difference. □

The theorem now follows from an application of Corollary 31. All that remains is to demonstrate how the requirements in the premises of the corollary are fulfilled. Consider again the original Markov modulated random walk g_k with modulating process S_k. Lemmas 34 and 36, state that for any $\delta > 0$, provided k is sufficiently large, it is possible to choose a factorization $k = lj + r$ with divisor l and a remainder r such that $\beta_{\mathfrak{G}}(g_r) < (C/\delta)\beta_{\mathfrak{G}}(g_0)\beta_{\mathcal{S}}(S_0)$, and if g^l_j denotes the walk g_k left shifted by r terms, and rewritten with increments grouped in groups of l, then the transition kernel for g^l, which can be written as

$$\Pi(S_{r+jl}, g^l_j; dg^l_{j+1}) = \delta_{g^l_j} * \Pi(S_{r+jl}, h^l_j; \cdot),$$

satisfies $d(\Pi(S_{r+jl}, g^l_j; dg^l_{j+1}), \delta_{g^l_j} * \mu^l) < \delta$, where $d\mu_l(h) = \nu\Pi^{l_2}(\cdot, \mathfrak{e}; dh)$ is the measure defined in the proof of Lemma 36. So provided μ^l has the contractive and primitivity properties required by Corollary 31, the theorem is proved. □

Theorem 33 gives an ergodicity result for the Markov modulated random walk g_k, but the estimator problem requires an analogous result for the Markov process (S_k, g_k) that combines the modulating process S_k with the random walk that it generates. The following result describes the situation when the modulating process S_k is i.i.d.

Proposition 37. *Let S_k be an i.i.d. process with distribution ν, and let g_k be the random walk generated by the associated measure $\mu = [d_\star g]\nu$ with invariant distribution m. Then the combined process (S_k, g_k) is ergodic, and has an invariant distribution on (S_b, g_b) given by $(m * \delta_{g(S_b)}(g_b))\nu(S_b)$*

Proof. First, demonstrate that $(m * \delta_{g(S_b)}(g_b))\nu(S_b)$ is an invariant distribution.

$$\begin{aligned}
&(m * \delta_{g(S_a)}(g_a))\nu(S_a)\Pi[S_a, g_a; d(S_b, g_b)]f(S_b, g_b) \\
&\qquad = (m * \delta_{g(S_a)}(g_a))\nu(S_a) \int f(S_b, g_b) d(\delta_{g_a} * \delta_{g(S_b)})(g_b) d\nu(S_b) \\
&\qquad = \int f(S_b, g_b) d\left(\int (\delta_{g_a} * \delta_{g(S_b)}) d(m * \delta_{g(S_a)})(g_a) d\nu(S_a)\right)(g_b) d\nu(S_b)) \\
&\qquad = \int f(S_b, g_b) d(m * \mu * \delta_{g(S_b)})(g_b) d\nu(S_b) \\
&\qquad = \int f(S_b, g_b) d(m * \delta_{g(S_b)})(g_b) d\nu(S_b)
\end{aligned}$$

Geometric convergence is demonstrated as follows. Let $f : \mathcal{S} \times \mathfrak{G} \to \mathbb{R}$ be Lipschitz (in both its variables), and let $h(S_a, g_a) = \int f(S_b, g_b) d(\delta_{g_a} * \delta_g(S_b))(g_b) d\nu(S_b)$, then h is Lipschitz with Lipschitz constant $L_h = KL_f$, and

$$\begin{aligned}
&|\Pi^k(S_0, g_0; d(S_b, g_b))f - \int f(S_b, g_b)(m * \delta_{g(S_b)}(g_b)) d\nu(S_b)| \\
&\qquad = \left| \int f(S_b, g_b) d(\delta_{g_0} * \mu * \overset{k-1}{\cdots} * \mu * \delta_{g(S_b)} - m * \delta_{g(S_b)})(g_b) d\nu(S_b) \right| \\
&\qquad = |m_{g_0,k} h - mh| \le KL_f c^k
\end{aligned}$$

□

The case when (S_k, g_k) is a Markov modulated walk introduces additional complexity. While the invariant distribution of g_k coincides with the invariant distribution of the random walk generated by ν, the invariant distribution of S_k, the same correspondence does not hold for the invariant distribution of the combined process. As a consequence, the proof of existence and regularity results for the combined process requires a new approach.

Theorem 38. *Let g_k be a Markov modulated random walk generated by a Markov process S_k. Suppose that process S_k and the map g satisfy the conditions in the premise of Theorem 33, and that ν, the invariant measure for the random walk on $\mathcal{S}$, is compactly supported. Then the Markov process (S_k, g_k) has an invariant distribution m, and there exist positive constants $c \in (0, 1)$ and K such that for any Lipschitz function $f : \mathcal{S} \times \mathfrak{G} \to \mathbb{R}$, with Lipschitz constant L_f,*

$$|\Pi^k(S_a, g_a; d(S_b, g_b))f - mf| \leq KL_f c^k \beta_{\mathfrak{G}}(g_a)\beta_{\mathcal{S}}(S_a) \tag{5.33}$$

Proof. Let $f : \mathcal{S} \times \mathfrak{G} \to \mathbb{R}$ be Lipschitz continuous with Lipschitz constant L_f. Let $\mathcal{S}_\alpha$, $\alpha \in \mathcal{A}$ be a finite partition of a subset of $\mathcal{S}$ that satisfies the following three conditions.

(i)	$\overline{\mathrm{supp}\,\nu} \subset \cup_\alpha \mathcal{S}_\alpha$	
(ii)	$\nu(\mathcal{S}_\alpha) \geq 0$	$\forall \alpha$
(iii)	$\mathrm{diag}(\mathcal{S}_\alpha) < \delta$	$\forall \alpha$

The existence of such a partition is a consequence of the compactness of the support of ν. For each $\alpha \in \mathcal{A}$, let f_α be the approximation to the restriction of f to the set $\mathcal{S}_\alpha \times \mathfrak{G}$ that is formed by averaging f over the set $\mathcal{S}_\alpha$. So, for $S \in \mathcal{S}$ and $g \in \mathfrak{G}$

$$f_\alpha(S, g) = \mathbf{1}_{\mathcal{S}_\alpha}(S) \otimes \frac{\nu(\mathbf{1}_{\mathcal{S}_\alpha} f(\cdot, g))}{\nu(\mathcal{S}_\alpha)}$$

Since f is Lipschitz, the error $|f - \sum_\alpha f_\alpha \mathbf{1}_{\mathcal{S}_\alpha}|$ is uniformly bounded on $\bigcup \mathcal{S}_\alpha \times \mathfrak{S}$.

Define a sequence of random times $k_\alpha(l)$ by the condition that $S_{k_\alpha(l)}$ is the l'th element of the sequence S_k that lies in $\mathcal{S}_\alpha$. Ergodicity of the chain S_k ensures that the values in the sequence $k_\alpha(l)$ are finite with probability 1. For each α define a family of distributions η_k on the positive real numbers $x \in \mathbb{R}^+$ by $\eta_k(x) = P[\{k = k_\alpha(kx)\}]$ then $\eta_k(x)$ is supported on the discrete set $\{x : x = l/k,\ l <= k\}$, $\lim_{k\to\infty} \mathbf{E}[\eta_k(x)] = \nu(\mathcal{S}_\alpha)$ and a large deviation result gives a bound that is exponential in k for the probability $P_{\eta_k}[|l/k - \nu_\alpha| > \epsilon]$.

Let $R_{\alpha,l} \in \Pi\mathcal{S}$ be the Markov chain that is constructed by forming the random length Cartesian products

$$R_{\alpha,l} = (S_{k_\alpha(l)+1}, \overset{k_\alpha(l+1)-k_\alpha(l)}{\ldots}, S_{k_\alpha(l+1)}).$$

Geometric ergodicity of the Markov process R_l follows from the geometric ergodicity of the underlying process S_k. Define a map $g : \Pi\mathcal{S} \to \mathfrak{S}$ by

$$g(R_l) = g(S_{k_\alpha(l)+1}) \ldots g(S_{k_\alpha(l+1)}),$$

a sequence of measures: μ_l on $\mathfrak{S}$ by

$$\mu_l = [d_* g]\Pi^l(R_0; \cdot)$$

and a sequence of transition kernels $\Pi^l_\alpha(g_a; dg_b)$ by the convolution product

$$\Pi^l_\alpha(g_a; dg_b) = \delta_{g_a} * \mu_1 * \cdots * \mu_l$$

The transition kernels Π^l_α define a Markov modulated random walk on $\mathfrak{S}$ with R_l as the modulating Markov process. Let $m_\alpha = \lim_{l\to\infty} \Pi^l_\alpha(g_a; dg_b)$, Theorem 33 ensures both the existence of the following geometric bound

$$|\Pi^l_\alpha(S_a, g_a; dg_k)(f) - m_\alpha(f)| \leq C_\alpha L_f c^l_\alpha \beta_{\mathfrak{S}}(g_a)\beta_{\mathcal{S}}(S_a).$$

For sufficiently large k, the kernel $\Pi^k((S_a, g_a); d(S_b, g_b))$ can be approximated to within a fixed multiple of δ by

$$\begin{aligned}
\Pi^k((S_a, g_a); d(S_b, g_b))f &\approx \sum_\alpha \sum_l P[\{k = k_\alpha(l)\}]\Pi^l_\alpha(g_a; dg_b) f_\alpha \\
&\approx \sum_\alpha \nu(\mathbb{S}_\alpha) m_\alpha f_\alpha
\end{aligned}$$

Where the large deviation result for the random time k_α, and the geometric ergodicity of the random walk have been used in the second approximation. ☐

Corollary 39. *Let $\Pi_\theta(S_a; dS_b)$ and $\Pi'_{\theta'}(S_a; dS_b)$ both satisfy the conditions of Lemma 33 for some measures ν and ν' separated in the Lévy metric by $d(\nu, \nu') < K_0|\theta - \theta'|$. Then there exists a constant K such that for any Lipschitz function $f : \mathbb{S} \times \mathfrak{G} \to \mathbb{R}$ with Lipschitz constant L_f,*

$$|\Pi^k_\theta(S_0, g_0; d(S_k, g_k))(f) - \Pi^k_{\theta'}(S_0, g_0; d(S_k, g_k))(f)| \le K L_f |\theta - \theta'| \beta_{\mathfrak{G}}(g_a) \beta_{\mathbb{S}}(S_a) \tag{5.34}$$

Proof. Let $m_{g_0,k}$ and $m'_{g_0,k}$ be the measures defined in the proof of Proposition 27.

$$\begin{aligned}
&|\Pi^k_\theta(S_0, g_0; d(S_k, g_k))(f) - \Pi^k_{\theta'}(S_0, g_0; d(S_k, g_k))(f)| \\
&\qquad \le |\Pi^k_\theta(S_0, g_0; d(S_k, g_k))(f) - m_{g_0,k}(f)| + |m_{g_0,k}(f) - m'_{g_0,k}(f)| \\
&\qquad\qquad + |m'_{g_0,k}(f) - \Pi^k_{\theta'}(S_0, g_0; d(S_k, g_k))(f)|
\end{aligned}$$

The result is a consequence of the preceding theorem and Corollary 30. ☐

5.3 Potential Theory for the Underlying Chain

This section combines the results from the first two sections in this chapter with the description of estimator Markov chain from the first section of Chapter 4 to prove that assumptions 4-*bis* and 5-*bis* from section 2 of Chapter 4 hold for the combined estimation and control stochastic approximation problem. Throughout this section Q is assumed to be a compact subset of the interior of the parameter space Θ, and the parameter $\theta \in Q$ is assumed to be a fixed point in Q. In particular, this means that the entries in the matrices $A_u(\theta)$ and $B(\theta)$, the estimates of the underlying state transition matrices A_u, and the ouptut matrix B are bounded away from zero by

$$\inf_{i,j,u} A_{u;ij}(\theta) > \delta(Q) \qquad \text{and} \qquad \inf_{i,m} B_{im}(\theta) > \delta(Q)$$

where $\delta(Q) > 0$ is a constant that depends on the choice of the compact set Q. Also, the standing assumptions from Chapter 2 about the underlying Markov model hold. Namely, there exists $\delta > 0$ such that

$$\inf_{i,j,u} A_{u;ij} > \delta \qquad \text{and} \qquad \inf_{i,m} B_{im} > \delta,$$

and η, the randomization parameter in the control algorithm (2.14), is strictly positive. The first task is to establish that the characterization of the Markov chain X_k that was given in Section 1 of Chapter 4 fits the requirements of the theory in Section 2 of this chapter.

Recall that the state space for the Markov chain has a projective structure

$$\mathfrak{X} \xrightarrow{\pi} \tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \xrightarrow{\pi} \tilde{\mathfrak{X}}, \tag{5.35}$$

and that the Markov transition kernel $\Pi(X_a, dX_b)$ is constructed from the discrete kernel $\Pi(\tilde{X}_a, d\tilde{X}_b)$ by using the actions of Markov modulated random walks on the sets $\mathfrak{X}^\alpha$, $\mathfrak{X}^\zeta$ and $\mathfrak{X}^\gamma$. The Markov modulated walks are generated by the maps

$$\tilde{\mathfrak{X}} \xrightarrow{g^\alpha} \mathfrak{S}(\Omega^\alpha) \tag{5.36}$$

$$\tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \xrightarrow{g^\zeta \otimes g^\gamma} \mathfrak{S}(\mathfrak{X}^\zeta) \otimes \mathfrak{S}(\mathfrak{X}^\gamma). \tag{5.37}$$

Where $g^\zeta \otimes g^\gamma$, the outer group product of the two maps g^ζ and g^γ, factors into two Markov modulated random walks

$$\tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \xrightarrow{g^\zeta} \mathfrak{S}(\mathfrak{X}^\zeta) \tag{5.38}$$

$$\tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \xrightarrow{g^\gamma} \mathfrak{S}(\mathfrak{X}^\gamma). \tag{5.39}$$

The product semigroup $\mathfrak{S}(\mathfrak{X}^\zeta) \otimes \mathfrak{S}(\mathfrak{X}^\gamma)$ is a 'block-diagonal' sub-semigroup of the semigroup $\mathfrak{S}(\mathfrak{X}^\zeta \times \mathfrak{X}^\gamma)$.

The Markov modulated random walk theory applies to both levels of the projective structure in (5.35), and the application of the theory to each level is considered in turn. The random walk generated by (5.36) is considered first. The underlying chain here is the discrete chain $\tilde{X}_l$ that was analyzed in the first section of this chapter. Proposition 12 establishes that the transition kernel for the discrete chain $\Pi(\tilde{X}_a; d\tilde{X}_b)$ has an invariant distribution $\tilde{\nu}_\theta$, and that the sequence of kernels $\Pi^l(\tilde{X}_a; d\tilde{X}_b)$ converges uniformly in X_a, and geometrically in l to the invariant measure. The mapping (5.36) with $g^\alpha : \tilde{X} \to \mathfrak{S}(\Omega^\alpha)$ defined by (4.3) induces a mapping $d_\star g^\alpha$ from measures on $\tilde{\mathfrak{X}}$ to measures on $\mathfrak{S}(\Omega^\alpha)$. The map $d_\star g^\alpha$ along with the Markov chain $\tilde{X}_k$ with transition kernel $\Pi(\tilde{X}_a; d\tilde{X}_b)$ generates a sequence of generators for a Markov modulated random walk on $\mathfrak{S}$ through the formula

$$\mu_k = d_\star g^\alpha \Pi(\tilde{X}_{k-1}; d\tilde{X}_k). \tag{5.40}$$

For $X \in \mathfrak{X}^\alpha = (X^1, X^2)$, let $\beta_s(X) = \max\{\sup_i |X_i^1|^{-s}, \sup_i |X_i^2|^{-s}\}$, and define a metric $\rho_\alpha : \mathfrak{S}(\mathfrak{X}^\alpha) \times \mathfrak{S}(\mathfrak{X}^\alpha) \to \mathbb{R}$ by

$$\rho_\alpha(g_1, g_2) = \sup_{X \in \mathfrak{X}^\alpha} \frac{\mathfrak{d}_\alpha(g_a X, g_2 X)}{\beta_1(X)}. \tag{5.41}$$

Lemma 40. *The sequence of generators μ_k defined in equation (5.40) are r-strongly contractive and p-strongly primitive. With $r = p = 2$.*

Proof. If $g \in \operatorname{supp}\mu_k$, then $g = g^\alpha(\tilde{X})$ for some $\tilde{X} \in \tilde{\mathfrak{X}}$. Consequently, if $X \in \mathfrak{X}^\alpha$ is written as an ordered pair $X = (X^1, X^2)$, then

$$\begin{aligned} (gX)^1 = g^\alpha(\tilde{X})X^1 = \operatorname{diag}(B(\theta)y^{-\Delta+1})A^\top_{u^{-\Delta}}(\theta)X^1 \\ (gX)^2 = X^1. \end{aligned} \tag{5.42}$$

Recall that the factors X^1 and X^2 are points in the probability simplex Ω^α which carries a hyperbolic metric, and is embedded in the projective plane. When $\theta \in Q$, left multiplication by the diagonal matrix $\operatorname{diag}(B(\theta)y^{-\Delta+1})$, considered as an operator on the projective plane, is an isometry on Ω^α with the hyperbolic metric, and left multiplication by the matrix $A^\top_{u^{-\Delta}}(\theta)$, is a strict contraction. Consequently, there exists a constant c, that depends on the choice of the compact set Q through the bound $\delta(Q)$, that satisfies the bounds $0 \le c < 1$, and that supports the following statement. For all $X \in \mathfrak{X}^\alpha$, for all $g_1, g_2 \in \mathfrak{S}(\mathfrak{X}^\alpha)$, and for all $g \in \operatorname{supp}\mu_k$

$$\begin{aligned} \mathfrak{d}_\alpha(gg_1X, gg_2X) \le c\,\mathfrak{d}_\alpha(g_1X, g_2X) \\ \Rightarrow \qquad \frac{\mathfrak{d}_\alpha(gg_1X, gg_2X)}{\beta(X)} \le c \sup_{X\in\mathfrak{X}^\alpha} \frac{\mathfrak{d}_\alpha(g_1X, g_2X)}{\beta(X)} = c\rho(g_1, g_2). \end{aligned}$$

The right hand side of the inequality is independent of X, so, taking the supremum of the left hand side over X gives the bound

$$\rho(gg_1, gg_2) \le c\rho(g_1, g_2).$$

Because this inequality holds for all g in the support of μ, it follows that the random walk is a strict contraction, and therefore a strong contraction.

Define $\beta_s(g) = \sup_{X\in\mathfrak{X}^\alpha}(\beta(gX)/\beta(X))$. Let K be a subset of $\mathfrak{S}$ defined by

$$K = \{g \in \mathfrak{S} : \beta_s(g) \le \frac{1}{\delta}\}$$

for some $\delta > 0$. If $g_1, g_2 \in K$, then

$$\begin{aligned}
\rho(g_1, g_2) &= \sup_{X \in \mathfrak{X}^\alpha} \sup_{i,j,l} \left[\log \left(\frac{(g_1X)_i^l}{(g_2X)_i^l} \frac{(g_2X)_j^l}{(g_1X)_j^l} \right) \beta(X)^{-1} \right] \\
&\leq \sup_{X \in \mathfrak{X}^\alpha} \sup_{i,j,l} \left[\log \left(\frac{1-(g_1X)_j^l}{(g_1X)_j^l} \right) \beta(X)^{-1} + \log \left(\frac{1-(g_2X)_i^l}{(g_2X)_i^l} \right) \beta(X)^{-1} \right] \\
&\leq \sup_{X \in \mathfrak{X}^\alpha} \sup_{i,j,l} \left[\frac{(1-2(g_1X)_j^l)}{(g_1X)_j^l} \beta(X)^{-1} + \frac{(1-2(g_2X)_i^l)}{(g_2X)_i^l} \beta(X)^{-1} \right] \leq \frac{2}{\delta}
\end{aligned}$$

K is a bounded closed subset of $\mathfrak{S}$ with respect to the metric $\mathfrak{d}$, and is therefore compact.

Returning to the structure of the map $g^\alpha; \tilde{X} \to \mathfrak{S}(\mathfrak{X}^\alpha)$ that is described in equations (5.42), let $g_1, g_2 \in \text{Range}(g^\alpha)$. Given an arbitrary point $X_0 \in \mathfrak{X}^\alpha$, let $X_1 = g_1X_0$, and $X_2 = g_2g_1X_0$. If X_1 and X_2 are written again as ordered pairs $X_1 = (X_1^1, X_1^2)$ and $X_2 = (X_2^1, X_2^2)$ then, it follows from the requirement that $A_{u;ij}(\theta) > \delta$ for all u, i and j that $\sup_i |X_{1,i}^1|^{-s} < 1/\delta$, and therefore that $\beta_s(X_2) < 1/\delta$. Not only is K absorbing from the right, but the random walk g_k is strictly p-primitive with $p = 2$, and therefore strongly p-primitive. ☐

Fix $X \in \mathfrak{X}_\alpha$, and let $o_X : \mathfrak{S}(\mathfrak{X}_\alpha) \to \mathfrak{X}_\alpha$ denote the mapping $o_X g \mapsto gX$. Since the members of the semigroup $\mathfrak{S}(\mathfrak{X}_\alpha)$ are all continuous transformations on $\mathfrak{X}_a$, it follows that o_X is a continuous map for all $X \in \mathfrak{X}_\alpha$. The map o_X, and the derived maps $d^\star o_X : C(\mathfrak{X}_\alpha) \to C(\mathfrak{S}(\mathfrak{X}_\alpha))$, the pullback on the space of bounded continuous functions, and $d_\star o_X : C^*(\mathfrak{S}(\mathfrak{X}_\alpha)) \to C^*(\mathfrak{X}_\alpha)$, the push-forward on the dual space of Radon measures provide mechanisms to map the properties of random walks on the semigroup $\mathfrak{S}(\mathfrak{X}_\alpha)$ to the properties of random sequences on the base space $\mathfrak{X}_\alpha$.

The metrics $\mathfrak{d}_\alpha$ on $\mathfrak{X}_\alpha$ and ρ_α on $\mathfrak{S}(\mathfrak{X})$, the weight functions β_s, and the Lipschitz-style seminorms for functions in $C(\mathfrak{X}_\alpha)$ and $\mathfrak{S}(\mathfrak{X}_\alpha)$ are defined in a way that allow properties of a random walk g_k on the group to be mapped to properties of a corresponding random sequence X_k on the space $\mathfrak{X}_a$.

Lemma 41. *Given $\beta : \mathfrak{X}_\alpha \to \mathbb{R}$ be a weighting function on $\mathfrak{X}_\alpha$, let*

$$\beta(g) = \sup_{X \in \mathfrak{X}_\alpha} ((\beta(gX))/\beta(X))$$

be the corresponding weight on the semigroup $\mathfrak{S}(\mathfrak{X}_\alpha)$. Define a seminorm $[\cdot]_{\beta_s}$ on the space of continuous functions $C(\mathfrak{X}_\alpha)$ by

$$[f]_\beta = \sup_{X_1 \neq X_2} \frac{|f(X_1) - f(X_2)|}{\mathfrak{d}(X_1, X_2)(\beta(X_1) + \beta(X_2))}$$

and an analogous seminorm on $C(\mathfrak{S}(\mathfrak{X}))$ by

$$[f]_\beta = \sup_{g_1 \neq g_2} \frac{|f(g_1) - f(g_2)|}{\rho(g_1, g_2)(\beta(g_1) + \beta(g_2))}.$$

Then, $[d^\star o_X(f)]_\beta \leq [f]_\beta$ for all functions $f : \mathfrak{X} \to \mathbb{R}$ with finite seminorm $[f]_\beta$.

Proof. Let X be any element of $\mathfrak{X}_\alpha$, and $g_a, g_b \in \mathfrak{S}(\mathfrak{X}_\alpha)$.

$$\begin{aligned} |[d^\star o_X(f)](g_a) - [d^\star o_X(f)](g_b)| &= |f(g_a X) - f(g_b X)| \\ &\leq [f]_\beta \mathfrak{d}_\alpha(g_a X, g_b X) \left(\frac{\beta(g_a X) + \beta(g_b X)}{\beta(X)} \right) \beta(X) \\ &\leq [f]_\beta \rho_\alpha(g_a, g_b)(\beta(g_a) + \beta(g_b)) \end{aligned}$$

□

Proposition 42. *Let $\theta \in Q$.*

(i) The Markov process $(\tilde{X}_l, X_l^\alpha)$ with transition kernel

$$\Pi_\theta(\tilde{X}_a, X_a^\alpha; d(\tilde{X}_b, X_b^\alpha)) = M_{\tilde{X}_b}^{\tilde{X}^a} \delta_{X_a^{\alpha,1}}(X_b^{\alpha,2}) \delta_{g^\alpha(\tilde{X}_b) X_a^{\alpha,1}}(X_b^{\alpha,1})$$

has an invariant measure ν^α.

(ii) *There exist positive constants* K, *and* c *with* $0 \leq c < 1$ *such that for any Lipschitz function* $f : \tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \to \mathbb{R}$ *with Lipschitz constant* L_f,

$$|\Pi_\theta^k(\tilde{X}_a, X_a^\alpha; d(\tilde{X}_b, X_b^\alpha))f - \nu^\alpha f| \leq K L_f c^k$$

(iii) *Let* θ' *be a second point in* Q, *then there exists a positive constant* K *such that for any Lipschitz function* $f : \tilde{\mathfrak{X}} \times \mathfrak{X}^\alpha \to \mathbb{R}$ *with Lipschitz constant* L_f, *and for all* $k > 0$,

$$|\Pi_\theta^k(\tilde{X}_a, X_a^\alpha; d(\tilde{X}_b, X_b^\alpha))f - \Pi_{\theta'}^k(\tilde{X}_a, X_a^\alpha; d(\tilde{X}_b, X_b^\alpha))f| \leq K L_f |\theta - \theta'|$$

Proof. As indicated in Proposition 7, the Markov chain $(\tilde{X}_l, X_l^\alpha)$ has a decomposition as a product of a discrete time Markov chain $\tilde{X}_l$ and the action of a Markov modulated random walk s_l on the space $\mathfrak{X}^\alpha = \Omega^\alpha \times \Omega^\alpha$. The Markov modulated random walk is defined on the semigroup $\mathfrak{S}(\Omega^\alpha)$, and has the transition kernel

$$\Pi_\theta(\tilde{X}_a, s_a; ds_b) = \delta_{s_a} * [d_\star g^\alpha](\Pi_\theta(\tilde{X}_a; \cdot))(ds_b). \tag{5.43}$$

The proof of the proposition relies on an application of Theorem 38, and the first order of business is to verify that the premises of that theorem are satisfied. Consider the following properties of the Markov modulated process with the transition kernel of equation (5.43):

(i) Since the space of the modulating process, $\mathcal{S} = \tilde{\mathfrak{X}}$, is finite and discrete, it trivially satisfies the requirements that there exist a metric $\mathfrak{d} : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ and a weight function $\beta_{\mathcal{S}}$. The space of measures $C_\beta^*(\tilde{\mathfrak{X}})$ is a finite dimensional vector space, the probability distributions form a finite dimensional probability simplex, and the Levy metric d is just the metric induced by the finite dimensional supremum norm.

(ii) From Proposition 12, $\Pi(S_a; dS_b) = M_{\tilde{X}^b}^{\tilde{X}^a}$ has an invariant distribution $\nu_{\tilde{\mathfrak{X}}}$, and there exist constants C and c such that for all $\tilde{X}_0$,

$$d(\Pi^k(\tilde{X}_0; d\tilde{X}_k), \nu_{\tilde{\mathfrak{X}}}) < Cc^k \beta_{\tilde{X}}(\tilde{X}_0)$$

(iii) $\mathfrak{S} = \mathfrak{S}(\Omega^\alpha)$, and $\rho_{\mathfrak{S}}$ is the metric defined in equation (5.41). Since $\tilde{\mathfrak{X}}$ is finite, the map $g_\alpha : \tilde{\mathfrak{X}} \to \mathfrak{S}(\Omega^\alpha)$ is trivially Lipschitz continuous with respect to the metric $\rho_{\mathfrak{S}}$ and any metric on $\tilde{\mathfrak{X}}$ that is compatible with the discrete topology.

(iv) The Markov chain $\tilde{X}_k$ generates a Markov modulated random walk on $\mathfrak{S}(\Omega^\alpha)$ through the map $g^\alpha : \tilde{\mathfrak{X}} \to \mathfrak{S}(\Omega^\alpha)$. Lemma 40 demonstrates that the Markov random walk is both p-primitive and r-contractive with $p = r = 2$.

Items (i) – (iv) establish the premises of Theorem 38, and an application of the theorem proves that the transition kernel in (5.43) has an invariant measure m^α, that the projection of the support of m^α onto the semigroup $\mathfrak{S}(\mathfrak{X}^\alpha)$ is contained in $\mathfrak{S}_0(\mathfrak{X}^\alpha)$, the sub-semigroup of zero-rank elements, and that the iterated kernels of the random walk converge geometrically to the invariant measure in the weak topology of probability measures on the semigroup.

Consider now the map

$$o_{X_a^\alpha} : \mathfrak{S}(\mathfrak{X}^\alpha) \to \mathfrak{X}^\alpha$$

$$o_{X_a^\alpha}(s) = sX_a^\alpha$$

The derived map $d_\star o_{X_a^\alpha} : C^*(\mathfrak{S})(\mathfrak{X}^\alpha) \to C^*(\mathfrak{X}^\alpha)$ maps measures on the semigroup to measures on the affine space $\mathfrak{X}^\alpha$. For example, the convolution kernel $\delta_{s_a} * m^\alpha$ evaluated at $s_a = \mathbf{I}$ on $\mathfrak{S}(\mathfrak{X}^\alpha)$ maps to the kernel $sX_a^\alpha\, dm(s)$ on $\mathfrak{X}^\alpha$.

If the Markov modulated random walk s_l is initialized at $s_0 = \mathbf{I}$, then the Markov process $(\tilde{X}_l, X_l^\alpha)$ can be written as $(\tilde{X}_l, o_{X_a^\alpha}(s_{l-1}), o_{X_a^\alpha}(s_l))$, and the sequence of derived maps

$$C^*(\tilde{\mathfrak{X}}) \xrightarrow{d_\star g^\alpha} C^*(\mathfrak{S}(\mathfrak{X}^\alpha)) \xrightarrow{d_\star o_{X^\alpha}} C^*(\mathfrak{X}^\alpha)$$

provides an alternative formula for the transition kernel (42):

$$\begin{aligned}
&\Pi_\theta(\tilde{X}_a, X_a^\alpha; d(\tilde{X}_b, X_b^\alpha)) \\
&\quad = M_{\tilde{X}_b}^{\tilde{X}^a} \delta_{X_a^{\alpha,1}}(X_b^{\alpha,2}) \delta_{g^\alpha(\tilde{X}_b) X_a^{\alpha,1}}(X_b^{\alpha,1})\, d\tilde{X}_b\, dX_b^\alpha \\
&\quad = M_{\tilde{X}_b}^{\tilde{X}^a} \delta_{X_a^{\alpha,1}}(X_b^{\alpha,2}) [d_\star o_{X_a^\alpha}](\delta_{\mathbf{1}}(\cdot) * [d_\star g^\alpha](\Pi_\theta(\tilde{X}_a; \cdot)))(X_b^\alpha)\, d\tilde{X}_b\, dX_b^\alpha
\end{aligned}$$

The advantage of the second, factored form of the kernel is that asymptotic properties of the iterated kernel can be derived directly from the properties of the iterated kernel for the Markov modulated random walk on the semi-group. In particular, the push-forward, $\nu^\alpha = d_\star o(m^\alpha)$ is the invariant measure proclaimed in statement (i) of the proposition, and the convergence estimate (5.33) in Theorem 38 combined with Lemma 41 establishes the estimate in statement (ii). The second conclusion of Proposition 12 combined with Corollary 39 and Lemma 41 establishes statement (iii) of the proposition. ☐

Corollary 43. *The projection of the support of the invariant measure ν^α onto the space $\mathfrak{X}^\alpha$ is compactly contained in the interior of $\mathfrak{X}^\alpha$*

Proof. Since $\theta \in Q$, a compact set in the interior of Θ, the entries in the matrices $A_u(\theta)$ and $B(\theta)$ are all bounded away from 0. Consequently, the mapping on $\mathfrak{X}^\alpha$ that is generated by any $g \in \operatorname{supp} \mu^\alpha$ (this mapping is written out explicitly in equation (5.42)) has a range that is a subset of $\mathfrak{X}_0^\alpha$, a compact subset of $\mathfrak{X}^\alpha$ defined by $X^{\alpha;i} > c$ for some constant $c > 0$ and for all $X^\alpha \in \mathfrak{X}^\alpha$, and for $i = 1, 2$. If $g \in \mathfrak{S}_0(\mathfrak{X}^\alpha)$, $h \in \operatorname{supp} \mu^\alpha$, and $X_a^\alpha \in \mathfrak{X}^\alpha$, then $o_{X_a^\alpha}(hg) \in \mathfrak{X}_0^\alpha$, and the claim follows. ☐

Consider now the map in Equation (5.39). The following paragraphs apply the Markov modulated random walk theory to this map. The families of affine transformations $g^{\zeta}(\tilde{X}, X^{\alpha}) : \mathcal{X}^{\zeta} \to \mathcal{X}^{\zeta}$ and $g^{\gamma}(\tilde{X}, X^{\alpha}) : \mathcal{X}^{\gamma} \to \mathcal{X}^{\gamma}$ are defined in equations (4.4) and (4.5) by the formulae

$$g^{\zeta,p}(\tilde{X}, X^{\alpha})X^{\zeta,p} = (1 - q\delta_{e_p}(u^{-\Delta}))X^{\zeta,p} + q\delta_{e_p}(u^{-\Delta})\zeta \tag{5.44}$$

$$g^{\gamma,m}(\tilde{X}, X^{\alpha})X^{\gamma,m} = (1 - q\delta_{e_m}(y^{-\Delta}))X^{\gamma,m} + q\delta_{e_m}(y^{-\Delta})\gamma. \tag{5.45}$$

The transformations depend on X^{α} through the smoothed distribution estimates $\zeta \in \Omega^{\zeta}$ and $\gamma \in \Omega^{\gamma}$ which are defined by equations (4.6) and (4.7)

$$\zeta^{ij} = \left. \frac{X_i^{\alpha,2} A_{u;ij}(\theta)\beta_j^{-\Delta}}{\sum_{i,j} X_i^{\alpha,2} A_{u;ij}(\theta)\beta_j^{-\Delta}} \right|_{u=u^{-\Delta}}$$
$$\gamma^{i} = \left. \frac{\sum_j \beta^{-\Delta+1}(j) A_{u;ij}(\theta) X_i^{\alpha,1}}{\sum_i \sum_j \beta^{-\Delta+1}(j) A_{u;ij}(\theta) X_i^{\alpha,1}} \right|_{u=u^{-\Delta+1}},$$

The smoothing quantities $\beta^{-\Delta}$ and $\beta^{-(\Delta+1)}$ are defined recursively by the formulae in equation (4.8)

$$\beta^{-(l+1),i} = \sum_j \beta^{-l,j} A_{u;ij}(\theta) B_{i,m}(\theta) \Bigg|_{u=u^{-l}, e_m=y^{-l}} \qquad \beta^0 = \mathbf{1}$$

Since the parameter θ lies in the compact set Q, all the components of the matrices $A_u(\theta)$ and $B(\theta)$ are bounded away from zero. Also, from Corollary 43, the components of $X^{\alpha,2}$ are bounded away from zero when $(\tilde{X}, X^{\alpha}) \in \operatorname{supp} \nu^{\alpha}$. These facts lead to the following corollary to Proposition 42.

Corollary 44. *Let Q be a compact subset of Θ, let $q \in Q$, and let $(\tilde{X}, X^{\alpha}) \in \operatorname{supp} \nu^{\alpha}$. then the smoothed distribution estimates ζ and γ, which are functions of $\tilde{X}$, X^{α} and θ, are uniformly bounded away from zero by a bound that is a function of Q and $\operatorname{supp} \nu^{\alpha}$.*

Define weight functions $\beta_\zeta : \mathcal{X}^\zeta \to \mathbb{R}$ and $\beta_\gamma : \mathcal{X}^\zeta \to \mathbb{R}$ by

$$\beta_\zeta(X^\zeta) = \min_p \min_{ij} |X^{\zeta,p;ij}|^{-1}$$
$$\beta_\gamma(X^\gamma) = \min_m \min_i |X^{\gamma,m;i}|^{-1}$$

The weight functions β_ζ and β_γ together with the metrics $\mathfrak{d}_\zeta$ and $\mathfrak{d}_\gamma$ on the spaces $\mathcal{X}^\zeta$ and $\mathcal{X}^\gamma$ generate metrics on the semigroups $\mathfrak{S}(\mathcal{X}^\zeta)$ and $\mathfrak{S}(\mathcal{X}^\gamma)$ through the definitions

$$\rho_\zeta(g_a, g_b) = \sup_{X \in \mathcal{X}^\zeta} \frac{\max_p \mathfrak{d}_\zeta((g_a X)^p, (g_b X)^p)}{\beta_\zeta(X)}$$
$$\rho_\gamma(g_a, g_b) = \sup_{X \in \mathcal{X}^\gamma} \frac{\max_p \mathfrak{d}_\gamma((g_a X)^p, (g_b X)^m)}{\beta_\gamma(X)}.$$

The following proposition uses these metrics to characterize the contractive nature of the random walks generated by g^ζ and g^γ.

Proposition 45. *The Markov modulated random walks on $\mathfrak{S}(\mathcal{X}^\zeta)$ and $\mathfrak{S}(\mathcal{X}^\gamma)$ generated by the maps $g^\zeta : \tilde{\mathcal{X}} \times \mathcal{X}^\alpha \to \mathfrak{S}(\mathcal{X}^\zeta)$, $g^\gamma : \tilde{\mathcal{X}} \times \mathcal{X}^\alpha \to \mathfrak{S}(\mathcal{X}^\gamma)$, and the Markov chain $(\tilde{X}_k, X_k^\alpha)$ from Proposition 42 are p-strong contractions with respect to the metrics ρ_ζ and ρ_γ when $p > P$.*

Proof. Consider first the random walk on $\mathfrak{S}(\mathcal{X}^\zeta)$ that is generated by $\mu^\zeta = d_\star g^\zeta(\nu^\alpha)$. The semigroup $\mathfrak{S}(\mathcal{X}^\zeta)$ is the outer product of P copies of the group $\mathfrak{S}(\Omega^\zeta)$ of transformations on Ω^ζ, the probability simplex on $\mathbb{R}^{N\times N}$. If g is in the support of the measure μ^ζ, then $g = g^\zeta(X^\alpha)$ for some $X^\alpha \in \operatorname{supp} \nu^\alpha$, and, repeating the definition of g^ζ in equation (5.44), for any $X^\zeta \in \mathcal{X}$,

$$g^{\zeta,p}(\tilde{X}, X^\alpha) X^{\zeta,p} = (1 - q\delta_{e_p}(u^{-\Delta})) X^{\zeta,p} + q\delta_{e_p}(u^{-\Delta})\zeta.$$

Equation (4.4) has P components, each one a transformation of one of the P components of the space $\mathcal{X}^\alpha$. For each value of $(\tilde{X}, X^\alpha)$, u can take only one of the P

values e_p, so $P-1$ of the components in equation (4.4) represent the identity transformation, and one component is the transformation

$$g^{\zeta,p}(\tilde{X}, X^{\alpha})X^{\zeta,p} = (1-q)X^{\zeta,p} + q\zeta \tag{5.46}$$

Let g_a and g_b be two elements of $\mathfrak{S}(\mathcal{X}^{\zeta})$, let X^{ζ} be an arbitrary point in $\mathcal{X}^{\zeta}$, and let $X_a^{\zeta} = g_a X^{\zeta}$ and $X_b^{\zeta} = g_b X^{\zeta}$.

$$\begin{aligned} &\mathfrak{d}_{\zeta}(g^{\zeta,p}(\tilde{X}, X^{\alpha})X_a^{\zeta,p}, g^{\zeta,p}(\tilde{X}, X^{\alpha})X_b^{\zeta,p}) \\ &\quad = \sup_{ij,i'j'} \log\left[\left(\frac{(1-q)X_a^{\zeta,p;ij} + q\zeta^{ij}}{(1-q)X_b^{\zeta,p;ij} + q\zeta^{ij}}\right)\left(\frac{(1-q)X_b^{\zeta,p;i'j'} + q\zeta^{i'j'}}{(1-q)X_a^{\zeta,p;i'j'} + q\zeta^{i'j'}}\right)\right] \end{aligned} \tag{5.47}$$

Now, if $(X_a^{\zeta,p;ij}/X_b^{\zeta,p;ij}) \geq 1$, then

$$\begin{aligned} \frac{(1-q)X_a^{\zeta,p;ij} + q\zeta^{ij}}{(1-q)X_b^{\zeta,p;ij} + q\zeta^{ij}} &= 1 + \frac{(1-q)X_b^{\zeta,p;ij}}{(1-q)X_b^{\zeta,p;ij} + q\zeta^{ij}}\left(\frac{X_a^{\zeta,p;ij}}{X_b^{\zeta,p;ij}} - 1\right) \\ &\leq \left(1 + \left(\frac{X_a^{\zeta,p;ij}}{X_b^{\zeta,p;ij}} - 1\right)\right)^c, \end{aligned}$$

and, since both the factors that form the argument of the logarithm on the right hand side of (5.47) are greater than 1,

$$\mathfrak{d}_{\zeta}(g^{\zeta,p}(\tilde{X}, X^{\alpha})X_a^{\zeta,p}, g^{\zeta,p}(\tilde{X}, X^{\alpha})X_b^{\zeta,p}) \leq c\,\mathfrak{d}_{\zeta}(X_a^{\zeta,p}, X_b^{\zeta,p}), \tag{5.48}$$

where

$$c = \sup_{X^{\zeta}\in\mathcal{X}^{\zeta}} \sup_{p,i,j}\left(\frac{(1-q)X^{\zeta,p;ij}}{(1-q)X^{\zeta,p;ij} + q\zeta^{ij}}\right).$$

By Corollary 44, ζ^{ij} is bounded away from 0, and so there exists a bound c_0 which depends only on q and the compact sets $Q \subset \Theta$ and $\operatorname{supp}\nu^{\alpha}$, and which satisfies $0 \leq c \leq c_0 < 1$ for all $X_b^{\zeta} \in \Omega^{\zeta}$. As a consequence it is permissible to reinterpret the multiplier c in inequality (5.48) as a constant less than 1.

Consider again the mapping in equation (5.44). All P of the component transformations are contractions, and one is a strict contraction. It follows from the result in Theorem 6 that proves strict positivity of the control measures $dv(u)$, that if $(\tilde{X}_k, X_k^\alpha), \ldots, (\tilde{X}_{k+P}, X_{k+P}^\alpha)$ are P consecutive points in the random process $(\tilde{X}_k, X_k^\alpha)$, then for any $X_a^\zeta, X_b^\zeta \in \mathfrak{X}^\zeta$ the element

$$g = g^\zeta(\tilde{X}_{k+P}, X_{k+P}^\alpha) \ldots g^\zeta(\tilde{X}_k, X_k^\alpha) \in \mathfrak{S}(\mathfrak{X}^\zeta)$$

satisfies the inequality $\max_p \mathfrak{d}_\zeta((gX_a^\zeta)^p, (gX_b^\zeta)^p) < c \max_p \mathfrak{d}_\zeta(X_a^{\zeta,p}, X_b^{\zeta,p})$ with positive probability. In particular, when $X_a^\zeta = g_a X^\zeta$, and $X_b^\zeta = g_b X^\zeta$ for arbitrary $g_a, g_b \in \mathfrak{S}(\mathfrak{X}^\zeta)$, and for arbitrary $X \in \mathfrak{X}^\zeta$,

$$\frac{\max_p \mathfrak{d}_\zeta((gg_a X^\zeta)^p, (gg_b X^\zeta)^p)}{\beta_\zeta(X^\zeta)} \leq c \frac{\max_p \mathfrak{d}_\zeta((g_a X^\zeta)^p, (g_b X^\zeta)^p)}{\beta_\zeta(X^\zeta)}$$

with positive probability. Taking a supremum over X^ζ first on the right hand side, and then on the left hand side yields the result that the random walk generated by g^ζ is p-strongly contractive for all $p \geq P$.

The proof that the random walk on $\mathfrak{S}(\mathfrak{X}^\gamma)$ generated by $\mu^\gamma = d_\star g^\gamma(\nu^\alpha)$ is m-strongly contractive for all $m \geq M$ follows an analogous argument. ☐

Define a weight function $\beta_{\zeta\gamma} : \mathfrak{X}^\zeta \times \mathfrak{X}^\gamma \to \mathbb{R}$ by

$$\beta_{\zeta\gamma}(X) = \max\{|\sup_{p,i,j} X^{\zeta,p;i,j}|^{-1}, |\sup_{m,i} X^{\gamma,m;i}|^{-1}\},$$

and a metric on the product space $\mathfrak{X}^\zeta \times \mathfrak{X}^\gamma$ by

$$\mathfrak{d}_{\zeta\gamma}((X_a^\zeta, X_a^\gamma), (X_b^\zeta, X_b^\gamma)) = \max\{\mathfrak{d}_\zeta(X_a^\zeta, X_a^\zeta), \mathfrak{d}_\gamma(X_a^\gamma, X_a^\gamma)\},$$

The corresponding metric on the semigroup $\mathfrak{S}(\mathfrak{X}^\zeta \times \mathfrak{X}^\gamma)$ is,

$$\rho_{\zeta\gamma}(g_a, g_b) = \sup_{X \in \mathfrak{X}^\gamma \times \mathfrak{X}^\gamma} \frac{\mathfrak{d}_{\zeta\gamma}(g_a X, g_b X)}{\beta_{\zeta\gamma}(X)},$$

and the following corollary to Proposition 45 is a consequence of the block structure of the direct product map $g^\zeta \otimes g^\gamma$.

Corollary 46. *The Markov modulated random walk on $\mathfrak{S}(\mathcal{X}^\zeta \times \mathcal{X}^\gamma)$ generated by the map $g^\zeta \otimes g^\gamma : \tilde{\mathcal{X}} \times \mathcal{X}^\alpha \to \mathfrak{S}(\mathcal{X}^\zeta) \times \mathfrak{S}(\mathcal{X}^\gamma)$, and the Markov chain $(\tilde{X}_k, X_k^\alpha)$ from Proposition 42 is a p-strong contraction with respect to the metrics $\rho_{\zeta\gamma}$ when $p > P$.*

Strong contractivity of the random walks is not enough to ensure geometric ergodicity. In addition, the random walks have to satisfy the primitivity condition from the premise of Lemma 25.

Define a weighting function $\beta : \mathfrak{S}(X^\zeta) \otimes \mathfrak{S}(X^\gamma) \to \mathbb{R}$ on the semigroup by $\beta(g) = \sup_{X \in \mathcal{X}^\zeta \otimes \mathcal{X}^\gamma} \beta_{\zeta\gamma}(gX)/\beta_{\zeta\gamma}(X)$. Let K_α be a compact neighborhood of $\operatorname{supp}(\nu_\alpha)$, where ν_α is the invariant distribution on $\mathcal{X}^\alpha$ postulated in Proposition 42.

Define a random sequence of generators μ_k by

$$\mu_k = d_\star[g^\zeta \otimes g^\gamma](\Pi^k((\tilde{X}_0, X_0^\alpha); d(\tilde{X}_k, X_k^\alpha))) \tag{5.49}$$

Lemma 47. *The sequence of generators μ_k satisfies the premises of Lemma 25 with probability 1. In particular, with probability 1:*

1. *The sequence $\mu_1, \mu_2, \ldots$ is R-strongly primitive, where $R = \max\{P, M\}$, P is the cardinality of the finite control set, and M is the cardinality of the finite observation space. I.e. there exist a left absorbing compact set K, and a constant $0 \le \eta < 1$ such that for all k,*

$$\mu_k * \mu_{k+1} * \cdots * \mu_{k+R}(\mathfrak{S} \setminus K) \le \eta.$$

2. *The sequence of measures $\mu_1, \mu_2, \ldots$ has uniformly bounded support, and if*

$$a = \sup\{\beta(g) : g \in \cup_k \operatorname{supp} \mu_k\},$$

then $a\eta^{1/R} < 1$.

Proof. Since the map $g^{\zeta,\gamma} = g^{\zeta} \otimes g^{\gamma}$ is a tensor product into the outer product of semigroups $\mathfrak{S}(\mathfrak{X}^{\zeta}) \otimes \mathfrak{S}(\mathfrak{X}^{\gamma})$, it is permissible to prove the lemma one map at a time. The methods of proof for each of the two maps g^{ζ} and g^{γ} are essentially the same, and only the proof for the map g^{ζ} is given.

Part 1: P-strong primitivity.

Consider the map $g^{\zeta,p}(\tilde{X}, X^{\alpha}) : \mathfrak{X}^{\zeta} \to \mathfrak{X}^{\zeta}$

$$g^{\zeta,p}(\tilde{X}, X^{\alpha})X^{\zeta,p} = (1 - q\delta_{e_p}(u^{-\Delta}))X^{\zeta,p} + q\zeta. \tag{5.50}$$

With ζ, which is a function of $\tilde{X}$ and X^{α}, defined by

$$\zeta^{ij} = \left.\frac{X_i^{\alpha,2} A_{u;ij} \beta_j^{-\Delta}}{\sum_{ij} X_i^{\alpha,2} A_{u;ij} \beta_j^{-\Delta}}\right|_{u=u^{-\Delta}}, \tag{5.51}$$

and $\beta^{-\Delta}$ defined by the backwards recursion

$$\beta_i^{-(l+1)} = \left.\sum_j \beta_j^{-l} A_{u;ij}(\theta) B_{i,m}(\theta)\right|_{u=u^{-l}, e_m=y^{-l}}.$$

with $\beta^0 = \mathbf{1}$. Recall that the parameter θ is restricted to a domain Q with the property that $A_{u;ij}(\theta) \geq \delta(Q)$ and $B_{im}(\theta) \geq \delta(Q)$. The constant $\delta(Q)$ lies within the bounds $0 < \delta(Q) < 1/\min\{N, M\}$ and depends only on the choice of domain Q. Consequently,

$$\begin{aligned}
\min_i \beta_i^{-(l+1)} &= \min_i \sum_j \beta_j^{-l} A_{u;ij}(\theta) B_{i,m}(\theta)\Big|_{u=u^{-l}, e_m=y^{-l}} \\
&\geq N \min_j \beta_j^{-l} \min_{ij} A_{u,ij}(\theta) \min_{i,m} B_{i,m}(\theta) \\
&\geq N\delta(Q)^2 \min_j \beta_j^{-l} \\
&\geq \delta(Q) \min_j \beta_j^{-l}.
\end{aligned}$$

Induction on the index l gives $\beta^{-\Delta} \geq \delta(Q)^{\Delta}$, and a similar argument gives an upper bound of $\beta^{-\Delta} \leq N^{\Delta}$.

The denominator in (5.51) is a convex combination of the components of $\beta^{-\Delta}$ and is therefore bounded above by N^{Δ}. Also, if $(\tilde{X}, X^{\alpha})$ lie in a compact subset of $\tilde{\mathcal{X}} \times \mathcal{X}^{\alpha}$, then $X_i^{\alpha,2} \geq c$ for come positive constant c, and the numerator in (5.51) is bounded below by $c\delta(Q)^{\Delta+1}$. It follows that the components of ζ are uniformly bounded below by a bound that depends on the choice of the parameter set Q, and the support of the random variable X^{α}.

The facts that the invariant distribution ν_{α} has compact support, that the initial point X_0^{α} is restricted to a compact subset of $\mathcal{X}^{\alpha}$, and that the kernel for the Markov modulated random walk on $\mathfrak{S}(\mathcal{X}^{\alpha})$ is strongly contractive together imply that the random process X_k^{α} is restricted to a compact subset of $\mathcal{X}^{\alpha}$. Consequently, if $(\tilde{X}_k, X_k^{\alpha})$ is the underlying Markov process for the Markov modulated process on $\mathfrak{S}(\mathcal{X}^{\zeta}) \otimes \mathfrak{S}(\mathcal{X}^{\gamma})$ and ζ_k is the random sequence of empirical distributions generated from $(\tilde{X}_k, X_k^{\alpha})$ by equation (5.51), then the components of all of the ζ_k are uniformly bounded below with probability 1. Call this lower bound c.

Now consider a sequence of generators $\mu_k * \mu_{k+1} * \mu_{k+P-1}$, where P is the cardinality of the set from which the controls u_k are drawn, and each generator μ_i is the projection of the generator defined in (5.49) onto $C_{\beta}^{\star}(\mathfrak{S}(\mathcal{X}^{\zeta}))$. Let $\mathcal{R} \subset \mathcal{X}^{\zeta}$ be the set of points $X^{\zeta} \in \mathcal{X}^{\zeta}$ such that $\min_{p,i,j} X^{p;ij} \geq c$. $\mathcal{R}$ is a compact subset of $\mathcal{X}^{\zeta}$. Let K be the the subset of the semigroup $\mathfrak{S}(\mathcal{X}^{\zeta})$ consisting of elements with representations as maps on $g : \mathcal{X}^{\zeta} \to \mathcal{R}$. K is a compact subset of $\mathfrak{S}(\mathcal{X}^{\zeta})$ that is absorbing from the left. Furthermore, as a result of the uniform bound on the random sequence ζ_k, the restriction that the control distributions υ_k are strictly positive, and the form of the mapping (5.50)

$$\mu_k * \mu_{k+1} * \mu_{k+P-1}(\mathfrak{S}(\mathcal{X}^{\zeta}) \setminus K) \leq \eta < 1$$

for some constant η, and the Markov random walk generated by g_{ζ} is strongly primitive.

Part 2: The growth condition.

Fix k, and consider more closely the measure

$$\mu_k = d_\star g_\zeta(\Pi^k((\tilde{X}_0, X_0^\alpha); d(\tilde{X}_k, X_k^\alpha)))$$

Each semigroup element $g \in \operatorname{supp} \mu_k$ has is a direct product of P factors. Each factor has a representation as a transformation from one of the P identical probability simplices $\mathfrak{X}^{\zeta;p}$ into itself. If $u_k \neq e_p$, the map on the p'th simplex is the identity, if $u_k = e_p$ the map on the p'th simplex is $X^p \to (1-q)X^p + \zeta$ where ζ is a function of a point $(\tilde{X}_k, X_k^\alpha)$ in the support of the kernel $\Pi^k((\tilde{X}_0, X_0^\alpha); d(\tilde{X}_k, X_k^\alpha))$. The first part of the proof demonstrated that there exists a constant $c > 0$ such that with probability 1, and for any k, $\min_{i,j} \zeta^{i,j} > c$. Consider the quantity $\beta(g)$, when $g \in \operatorname{supp} \mu_k$:

$$\beta(g) = \sup_{X \in \mathfrak{X}} \frac{\beta(gX)}{\beta(X)} = \sup_{X \in \mathfrak{X}} \max_p \frac{\beta((gX)^p)}{\beta(X^p)}$$

The $P-1$ factors of g that are the identity satisfy $(gX)^p = X^p$, and $\beta((gX)^p)/\beta(X^p) = 1$ for all X. For the other factor $(gX)^p = (1-q)X^p + q\zeta$, and

$$\begin{aligned}\beta((gX)^p) &= \max_{i,j}((gX)^{p;ij})^{-1} \\ &= \max_{i,j}((1-q)(X^{p;ij}) + q\zeta^{ij})^{-1} \leq (1-q)\max_{i,j}(X^{p;ij})^{-1} + q/c.\end{aligned}$$

Also, since $\min_{i,j} X^{p;ij} < 1/N^2$,

$$\beta((gX)^p)/\beta(X^p) \leq 1 - q + qN^2/c$$

and $\beta(g) \leq a = \max\{1, 1-q+qN^2/c\}$.

Since $\beta : \mathfrak{S}(\mathfrak{X}^\zeta) \to \mathbb{R}$ is continuous with respect to the metric ρ on $\mathfrak{S}(\mathfrak{X}^\zeta)$, the existence of the bound a implies that the set $\cup_k \operatorname{supp} \mu_k$ is bounded. Furthermore from the form of the bound it is clear that a sufficiently small choice of q will cause the bound to lie arbitrarily close to 1. In particular, when q is sufficiently small, $a\eta^{1/P} < 1$.

$\square$

Proposition 48. *Let $\theta \in Q$.*

(i) The Markov process $X_l = (\tilde{X}_l, X_l^\alpha, X_l^\zeta, X_l^\gamma)$ with transition kernel

$$\Pi_\theta(X_a; dX_b) = M_{\tilde{X}_b}^{\tilde{X}^a} \delta_{X_a^{\alpha,1}}(X_b^{\alpha,2}) \delta_{g^\alpha(\tilde{X}_b) X_a^{\alpha,1}}(X_b^{\alpha,1}) \times \prod_p \delta_{g^{\zeta,p}(\tilde{X}_b, X_b^\alpha) X_a^{\zeta,p}}(X_b^{\zeta,p}) \prod_m \delta_{g^{\gamma,m}(\tilde{X}_b, X_b^\alpha) X_a^{\gamma,m}}(X_b^{\gamma,m})$$

has an invariant measure ν.

(ii) There exist positive constants K, and c with $0 \le c < 1$ such that for any β_1-Lipschitz function $f : \mathcal{X} \to \mathbb{R}$ with Lipschitz constant L_f,

$$|\Pi_\theta^k(X_a; dX_b) f - \nu f| \le K L_f c^k \beta_2(X_a)$$

(iii) Let θ' be a second point in Q, then there exists a positive constant K such that for any β_1-Lipschitz function $f : \mathcal{X} \to \mathbb{R}$ with Lipschitz constant L_f, and for all $k > 0$,

$$|\Pi_\theta^k(X_a; dX_b) f - \Pi_{\theta'}^k(X_a; dX_b) f| \le K L_f |\theta - \theta'| \beta_2(X_a)$$

Proof. The proof follows the same argument as the proof of Proposition 42, only here the sequence of derived maps that factor the kernel are

$$C^*(\tilde{\mathcal{X}} \times \mathcal{X}^\alpha) \xrightarrow{d_\star(g^\zeta \otimes g^\gamma)} C^* \mathfrak{S}(\mathcal{X}^\zeta \times \mathcal{X}^\gamma) \xrightarrow{d_\star o_{X^\zeta \times X^\gamma}} C^*(\mathcal{X}^\zeta \times \mathcal{X}^\gamma),$$

and the Markov process $(\tilde{X}, X^\alpha)$ generates a Markov modulated random walk on $\mathfrak{S}(\mathcal{X}^\zeta \times \mathcal{X}^\gamma)$ with transition kernel

$$\Pi(s_a; ds_b) = [\delta_{s_a} * d_\star(g^\zeta \otimes g^\gamma)](ds_b).$$

Once again the argument in the proof depends on applications of Theorem 38 and Corollary 39. The proof begins by establishing that the conditions that form the premises of these results hold.

(i) *The space* $\mathcal{S} = \tilde{\mathcal{X}} \times \mathcal{X}^\alpha$ *supports a metric* $\mathfrak{d} : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ *and a weight function* $\beta_{\mathcal{S}}$.

Define the metric by

$$\mathfrak{d}((\tilde{X}_a, X_a^\alpha), (\tilde{X}_b, X_b^\alpha)) = \begin{cases} \tilde{X}_a = \tilde{X}_b & \mathfrak{d}_\alpha(X_a^\alpha, X_b^\alpha) \\ \tilde{X}_a \neq \tilde{X}_b & \infty, \end{cases}$$

and the weight function by

$$\beta((\tilde{X}, X^\alpha)) = \inf_{i,j} |X^{\alpha;j}(i)|^{-1}$$

(ii) *Let* d *be the Lévy metric that is induced on the probability measures in* $C_\beta^*(\mathcal{S})$ *by the definition of* $\mathfrak{d}_\alpha$ *in . The kernel* $\Pi(S_a; dS_b)$ *has a compactly supported, invariant distribution* ν*, and there exists constants* C *and* c *with* $0 \leq c < 1$ *such that for all* X_0*,*

$$d(\Pi(S_0; dS_k), \nu) < Cc^k \beta(S_0).$$

Proposition 42 proves the existence of the invariant measure ν, and in combination with Lemma 18 establishes the geometric bound on the convergence of the iterated kernels in the Lévy metric. Corollary 43 establishes that the support of the measure is compact.

(iii) *The mappings* g^ζ *and* g^γ *defined in equations (5.44) and (5.45) are Lipschitz continuous with respect to the metric* $\mathfrak{d}$ *on* $\tilde{\mathcal{X}} \times \mathcal{X}^\alpha$*, and the metrics* ρ_ζ *and* ρ_γ *on* $\mathfrak{S}(\mathcal{X}^\zeta)$ *and* $\mathfrak{S}(\mathcal{X}^\gamma)$ *respectively.*

The map g^ζ is defined in equation 5.44 by the formula

$$g^{\zeta,p}(\tilde{X}, X^\alpha)X^{\zeta,p} = (1 - q\delta_{e_p}(u^{-\Delta}))X^{\zeta,p} + q\delta_{e_p}(u^{-\Delta})\zeta$$

with

$$\zeta^{ij} = \left. \frac{X_i^{\alpha,2} A_{u;ij}(\theta)\beta_j^{-\Delta}}{\sum_{i,j} X_i^{\alpha,2} A_{u;ij}(\theta)\beta_j^{-\Delta}} \right|_{u=u^{-\Delta}} .$$

If $(\tilde{X}_a, X_a^\alpha)$ and $(\tilde{X}_b, X_b^\alpha)$ are two points in the domain of g^ζ, then

$$\rho(g^\zeta(\tilde{X}_a, X_a^\alpha), g^\zeta(\tilde{X}_b, X_b^\alpha)) \le C\,\mathfrak{d}((\tilde{X}_a, X_a^\alpha), (\tilde{X}_b, X_b^\alpha)) = \infty$$

for any $C > 0$ unless $\tilde{X}_a = \tilde{X}_b$. Assume that $\tilde{X}_a = \tilde{X}_b = \tilde{X}$, and consider the action of $g^\zeta(\tilde{X}_a, X_a^\alpha)$ and $g^\zeta(\tilde{X}_b, X_b^\alpha)$ on the component X^p of some $X \in \mathfrak{X}^\zeta$. If $u^{-\Delta} \neq e_p$, then $[g^\zeta(\tilde{X}_a, X_a^\alpha)X]^p = [g^\zeta(\tilde{X}_b, X_b^\alpha)X]^p = X^p$, and the only component of $g^\zeta(\tilde{X}_b, X_b^\alpha)X$ that makes a contribution to the distance $\rho(g^\zeta(\tilde{X}_a, X_a^\alpha), g^\zeta(\tilde{X}_b, X_b^\alpha))$ is the component corresponding to $u^{-\Delta}$. For this component, $[g^\zeta(\tilde{X}_a, X_a^\alpha)X]^p = (1-q)X^p + q\zeta_a$, and $[g^\zeta(\tilde{X}_b, X_b^\alpha)X]^p = (1-q)X^p + q\zeta_b$. It is always the case that

$$\max_{iji'j'} \frac{((1-q)X_{ij}^p + q\zeta_a^{ij})((1-q)X_{i'j'}^p + q\zeta_b^{i'j'})}{((1-q)X_{ij}^p + q\zeta_b^{ij})((1-q)X_{i'j'}^p + q\zeta_a^{i'j'})} \le \max_{iji'j'} \frac{\zeta_a^{ij}\zeta_b^{i'j'}}{\zeta_b^{ij}\zeta_a^{i'j'}},$$

and since

$$\frac{\zeta_a^{ij}\zeta_b^{i'j'}}{\zeta_b^{ij}\zeta_a^{i'j'}} = \frac{X_{a,i}^{\alpha,2} X_{b,i'}^{\alpha,2}}{X_{b,i}^{\alpha,2} X_{a,i'}^{\alpha,2}}$$

it follows that

$$\rho(g^\zeta(\tilde{X}, X_a^\alpha), g^\zeta(\tilde{X}, X_b^\alpha)) \le \mathfrak{d}((\tilde{X}, X_a^\alpha), (\tilde{X}, X_b^\alpha))$$

Which establishes Lipschitz continuity with a Lipschitz constant less than 1.

An analogous argument proves case for the map g^γ.

(iv) *The Markov modulated random walks generated by Markov process* $(\tilde{X}_k, X^\alpha)$ *and the maps* g^ζ *and* g^γ *are both* p*-primitive and* r*-contractive. In the case of the Markov modulated random walk on* $\mathfrak{S}(\mathcal{X}^\zeta)$, $p = r = P$*, and in the case of the Markov modulated random walk on* $\mathfrak{S}(\mathcal{X}^\gamma)$, $p = r = M$

Proposition 45 and Lemma 47 prove this statement.

Statements (i) – (iv) establish that the premises of Theorem 38 hold, and the remainder of the proof is entirely analogous to the corresponding proof of Proposition 42 ☐

The geometric ergodicity and regularity results in Propositions 42 and 48 provide the basis for establishing Assumption 4-*bis* on the boundedness of moments, and Assumption 5 about the existence of weak solutions to the Poisson equation. The moment condition is dealt with first in the following lemmas and proposition.

Lemma 49. *Let* $\theta \in Q$*, a compact subset of* Θ*, and let* ν *be the invariant measure that was postulated in Proposition 48, then the projections of the support of* ν *onto the spaces* $\mathcal{X}^\alpha$, $\mathcal{X}^\zeta$ *and* $\mathcal{X}^\gamma$ *are bounded with respect to the hyperbolic metrics* $\mathfrak{d}_\alpha$, $\mathfrak{d}_\zeta$ *and* $\mathfrak{d}_\gamma$.

Proof. The case for the projection onto $\mathcal{X}^\alpha$ is a direct consequence of Corollary 43, the case for the projection onto $\mathcal{X}^\zeta$ is proved here, and the case for the projection onto $\mathcal{X}^\gamma$ can be proved in a completely analogous fashion.

Let $\mu = d_\star g^\zeta \otimes g^\gamma(\nu^\alpha)$. If $s \in \operatorname{supp} \mu$, then the component of s that operates on $\mathcal{X}^\zeta$ is $g^\zeta(\tilde{X}, X^\alpha)$ for some $(\tilde{X}, X^\alpha) \in \operatorname{supp} \nu^\alpha$. If the mapping g^ζ is written explicitly, then the action of s on the p'th simplex in $\mathcal{X}^\zeta$ is described by the equation

$$sX^{\zeta,p} = g^{\zeta,p}(\tilde{X}, X^\alpha)X^{\zeta,p} \tag{5.52}$$

$$= (1 - q\delta_{e_p}(u^{-\Delta}))X^{\zeta,p} + q\delta_{e_p}(u^{-\Delta})\zeta \tag{5.53}$$

The proof of Proposition 45 established that the components of the quantity ζ appearing in the right hand side of equation (5.52) are all bounded away from zero by a bound ϵ that depends only on Q. Consequently, if X_l is a trajectory governed by a random walk generated by the measure μ, and if the a-priori lower bound on the components of the projection of the initial point in the walk, X_0, onto the p'th component of $\mathcal{X}^\zeta$ is 0, then each time during the walk that $u_l^{-\Delta} = e_p$, the lower bound on $X_k^{\zeta,p}$, for $k > l$, moves closer to ϵ by a factor $(1 - q)$. When the $u_l^{-\Delta} = e_p$, the mapping $s_l = g^\zeta(\tilde{X}_l, X_l^\alpha)$ is the identity and $X_l^{\zeta,p} = sX_{l-1}^{\zeta,p} = X_{l-1}^{\zeta,p}$.

The randomized control for the underlying controlled hidden Markov model is constrained so that the distribution of input values $v_{l-\Delta}(e_p) > \eta \geq 0$ for any p and any time $l > \Delta$. Consequently, under the invariant measure ν^α,

$$\nu^\alpha(\{(\tilde{X}, X^\alpha) : u^{-\Delta} = e_p\}) > \eta, \tag{5.54}$$

and under the measure μ, $\mu\{s : sX^{\zeta,p} \neq X^{\zeta,p}\} > \eta$.

Finally, each point in the support of ν is the range of a zero-rank mapping in the support of the invariant distribution m, and for any integer k,

$$m = \mu * \mu * \overset{k \text{ times}}{\cdots} * \mu * m.$$

It follows from a zero-one law argument that for any p, if $X^{\zeta,p}$ is the projection of a point in the support of ν onto the p'th simplex in the space $\mathcal{X}^\zeta$, then for any i, j $X^{\zeta,p;ij} > \epsilon$. Since the points of $\operatorname{supp} \nu$ are bounded away from the boundary of $\mathcal{X}^\zeta$ in the Euclidean metric, it follows that $\operatorname{supp} \nu$ is bounded with respect to the hyperbolic metric $\mathfrak{d}_\zeta$. □

Lemma 50. *Let Λ_δ be a family of probability measures that are defined on $\mathcal{X}$ and that have support in a δ neighborhood of $\operatorname{supp} \nu$, with respect to the metric $\mathfrak{d}_{\mathcal{X}}$. For each* *positive integer q, the set of moments*

$$\left\{ \int \beta_q(X)\ d\lambda(X)\ : \lambda \in \Lambda_\delta \right\}$$

is bounded by a bound that depends only on q, $\operatorname{supp}\nu$ and δ.

Proof. Since $\operatorname{supp}\nu$ is a compact set with respect to the metric $\mathfrak{d}_{\mathcal{X}}$, a δ neighborhood of $\operatorname{supp}\nu$ is also compact. The weight functions $\beta_q(X)$ are all continuous with respect to the metric $\mathfrak{d}_{\mathcal{X}}$, and consequently for every q, β_q is bounded on the δ neighborhood of $\operatorname{supp}\nu$. It follows that for each q the q-th moments with respect to measures in Λ_δ are uniformly bounded by the the same bound. ☐

Lemma 51. *Fix $\theta_0 \in Q$, choose $X_0 \in \mathcal{X}$, and consider the probability space of Markov chain trajectories that start at X_0, and have statistics consistent with the kernel $\Pi_\theta(X; dX)$. Let ν be the invariant distribution for $\Pi_\theta(X; dX)$, then:*

(i) the distance

$$\sup_{X_b \in \operatorname{supp}\nu} \mathfrak{d}(X_l, X_b)$$

is uniformly bounded for all trajectories, and

(ii) for all $\epsilon > 0$, there exists constants $1 > c \geq 0$ and $C > 0$ such that

$$\boldsymbol{P}[\mathfrak{d}(X_l, \operatorname{supp}\nu) > \epsilon] \leq Cc^l$$

Proof.

(i) Let $\sup_{X_b \in \operatorname{supp}\nu} \mathfrak{d}(X_l, X_b) = d$. Written out in terms of its components, $X_l = (\tilde{X}_l, X_l^\alpha, X_l^\zeta, X_l^\gamma)$. Since each of the continuous components is determined by the action of an r-strongly contractive semigroup for a large enough integer r, and the measure ν is invariant under the action of the kernel Π_θ, it follows that for all $l \geq 0$

$$\sup_{X_b \in \operatorname{supp}\nu} \mathfrak{d}(X_l, X_b) < d$$

on every trajectory X_l.

(ii) For each trajectory X_l starting at X_0, Choose $X_b \in \operatorname{supp}\nu$, with $\tilde{X}'_b = \tilde{X}_0$, and let X'_l be a trajectory generated by the same Markov modulated random walks as generate X_l, but with initial point $X'_0 = X_b$. The trajectory X'_l remains within $\operatorname{supp}\nu$ (w.p.1), and since the random walks are r-strongly contractive, and hence contractive, $\mathfrak{d}(X_l, \operatorname{supp}\nu) < \mathfrak{d}(X_l, X'_l)$ (w.p.1). Also, from the definition of r-strongly contractive walk, there exist constants $0 \leq c_0 < 1$ and $0 < c_0 \leq 1$ such that for all trajectories, and for all l

$$\boldsymbol{P}\left[\mathfrak{d}(X_{l+r}, X'_{l+r}) < c_0 \mathfrak{d}(X_l, X'_l) \;|X_l, X'_l\right] > c_1$$

Choose $K > (\ln\epsilon - \ln d)/\ln c_0$, then the binomial theorem gives,

$$\boldsymbol{P}[\mathfrak{d}(X_{lr}, \operatorname{supp}\nu) \geq \epsilon] \leq \sum_{k=0}^{K} \binom{l}{k} c_1^k (1-c_1)^{l-k}.$$

Choosing $c = \sqrt[r]{(1-c_1)}$, and C sufficiently large yields the required geometric bound. □

The following proposition establishes that Assumption 4-*bis* holds.

Proposition 52. *For all $q \geq 1$, and for any compact subset Q of D, there exist $r \in \mathbb{N}$, and constants $\bar{\alpha} < 1$, C_1, C_2, K_1, and K_2, such that*

(i) $$\sup_{\theta\in Q} \int \Pi^r_\theta(X_a; dX_b)\beta_q(X_b) \leq \bar{\alpha}\beta_q(X_a) + C_1$$

(ii) $$\sup_{\theta\in Q} \int \Pi_\theta(X_a; dX_b)\beta_q(X_b) \leq C_2\beta_q(X_a) + C_1$$

For any Borel function g on $\mathcal{X}$ such that $[g]_q \leq \infty$

(iii) $$\sup_{\theta\in Q} |\Pi_\theta g(X_1) - \Pi_\theta g(X_2)| \leq K_1[g]_q \mathfrak{d}(X_1, X_2)(\beta_q(X_1) + \beta_q(X_2))$$

For all $\theta, \theta' \in Q$, and for any Borel function g with $[g]_q \leq \infty$

(iv) $$|\Pi_\theta g(X) - \Pi_{\theta'} g(X)| \leq K_2[g]_q|\theta - \theta'|\beta_{q+1}(X))$$

Proof. For every $\theta \in Q$, the kernel $\Pi(X_a; dX_b)$ has an invariant measure ν_{theta}, and as a consequence of Lemma 49, the set $S = \cup_{\theta \in Q} \bar{\text{supp}}\, \nu_\theta$ is compact. So, for any $d > 0$, the set

$$S_d = \{X \in \mathcal{X} : \mathfrak{d}(X, S) \leq d\}$$

Is also compact, and for any values of q and d the weighting function β_q is bounded on S_d.

For each $X_a \in \mathcal{X}$, define $d_a = \sup_{X \in S} \mathfrak{d}(X_a, X)$.. Because, for all $\theta \in Q$, the Markov transition kernel Π_θ is a contractive random walk on $\mathcal{X}$, it follows that all trajectories that start at X_a must remain inside the set S_a. The bound

$$B_a = \sup_{X \in S_a} \beta_q(X)$$

is a function of B_a, and as $\beta_q(X_a) \to \infty$, the ratio $\beta_q(X_a)/B_a \to 1$. Fix d, so that the bound $B = \sup_{X \in S_d} \beta_q(X)$ is sufficiently large that if $X_a \in \mathcal{X}$, and $B_a > B$, then $\beta_q(X_a)/B_a < 2$.

Fix X_a. As a result of Lemma 51, it is possible to choose r sufficiently large that for all $\theta \in Q$, $\boldsymbol{P}[X_r \notin S_d] < 1/4$. The weight β_q is bounded on the set $\mathcal{X} \setminus S_d$ by $2\beta_q(X_a)$, and on the set S_d by B. Substituting these estimates into the integral in the left hand side of inequality (i) produces the bound on the right hand side when $\bar{\alpha} = 1/2$, and $C_1 = B$. In addition, because $B_a < 2\beta_q(X_a)$, inequality (ii) holds when the bound on the right hand side has $C_2 = 2$. ☐

The final task in this chapter is to prove that Assumption 5-*bis* holds for the the specific case of the estimator Markov chain. This is done with an appropriate adaptation of a theorem from Benveniste *et al.* [7]. The theorem uses an assumption of geometric ergodicity for the Markov chain to establish an appropriate Potential Theory for the Markov transition kernel. The central object in the Potential Theory is the integral equation

$$(I - \Pi_\theta)\nu_\theta = f_\theta - h(\theta),$$

which is known as the Poisson equation. The quoted theorem establishes conditions under which a regular family of regular solutions exist for the parameterized integral equation. The parameterized transition kernel is $(\Pi_\theta, \theta \in Q)$, and it acts on the state space $\mathbb{R}^k \times E$. The definitions for the norms and function spaces used in the statement of the theorem were given in Section 4.2.

Theorem 53. *[7, Part II, Chapter 2, Theorem 5] Given $p_1 \geq 0$, $p_2 \geq 0$, assume that there exist positive constants K_1, K_2, K_3, q_1, q_2, $\rho < 1$ such that:*

(i) for all $g \in$ Li$(p_1), \theta \in Q, n \geq 0$, z_1, and z_2:

$$|\Pi_\theta^n g(z_1) - \Pi_\theta^n g(z_2)| \leq K_1 \rho^n N_{p_1}(g)(1 + |x_1|^{q_1} + |x_2|^{q_1})$$

(ii) for all $\theta \in Q$, $n \geq 0$, z, and all $m \leq q_1 \vee q_2$,

$$\sup_e \int \Pi_\theta^n(x, e; dx_1 de_1)(1 + |x_1|^m) \leq K_2(1 + |x|^m)$$

(iii) for all $g \in$ Li$(p_1), \theta, \theta' \in Q$, $n \geq 0$, z,

$$|\Pi_\theta^n g(z) - \Pi_{\theta'}^n g(z)| \leq K_3 N_{p_1}(g)|\theta - \theta'|(1 + |x|^{q_2})$$

Then for any function $f(\theta, z)$ of class Li(Q, L_1, L_2, p_1, p_2), there exist functions $h(\theta)$, $\nu_\theta(\cdot)$ and constants C_1, C_2, $C(\lambda)$, $0 < \lambda < 1$ depending only on the L_j, p_j, such that:

(j) for all $\theta, \theta' \in Q$, $|h(\theta) - h(\theta')| \leq C_1|\theta - \theta'|$

(jj) for all $\theta \in Q$, $|\nu_\theta(x, e)| \leq C_2(1 + |x|^{q_1})$

(jjj) *for all* $\theta, \theta' \in Q$ *all* $\lambda \in (0,1)$ *and for* $s = \max(p_2, q_1, q_2)$

$$|\nu_\theta(x,e) - \nu_{\theta'}(x,e)| \le C(\lambda)|\theta - \theta'|^\lambda(1 + |x|^s)$$
$$|\Pi_\theta \nu_\theta(x,e) - \Pi_{\theta'} \nu_{\theta'}(x,e)| \le C(\lambda)|\theta - \theta'|^\lambda(1 + |x|^s)$$

(jv) $(I - \Pi_\theta)\nu_\theta = f_\theta - h(\theta)$

A statement of the theorem that is appropriate to the current situation and terminology is the following.

Theorem 53-*bis*. *Given* $p_1 \ge 0$, $p_2 \ge 0$*, assume that there exist positive constants* K_1, K_2, K_3, q_1, q_2, $\rho < 1$ *such that:*

(i) *for all* $g \in$ *Li*(p_1), $\theta \in Q, n \ge 0$, X_1*, and* X_2*:*

$$|\Pi_\theta^n g(X_1) - \Pi_\theta^n g(X_2)| \le K_1 \rho^n N_{p_1}(g)(\beta_{q_1}(X_1) + \beta_{q_1}(X_2))$$

(ii) *for all* $\theta \in Q$, $n \ge 0$, X*, and all* $m \le q_1 \vee q_2$,

$$\sup_{\tilde{X}} \int \Pi_\theta^n(X_a; dX_b)\beta_m(X_b) \le K_2 \beta_m(X_a)$$

(iii) *For all* $g \in$ *Li*(p_1), $\theta, \theta' \in Q$, $n \ge 0$, $X \in \mathfrak{X}$,

$$|\Pi_\theta^n g(X) - \Pi_{\theta'}^n g(X)| \le K_3 N_{p_1}(g)|\theta - \theta'|\beta_{q_2}(X)$$

Then for any function $f(\theta, X)$ *of class Li*(Q, L_1, L_2, p_1, p_2)*, there exist functions* $h(\theta)$, $v_\theta(\cdot)$ *and constants* C_1, C_2, $C(\lambda)$, $0 < \lambda < 1$ *depending only on the* L_j, p_j*, such that:*

(j) *for all* $\theta, \theta' \in Q$, $|h(\theta) - h(\theta')| \le C_1|\theta - \theta'|$

(jj) *for all* $\theta \in Q$, $|v_\theta(X)| \le C_2 \beta_{q_1}(X)$

(jjj) for all $\theta, \theta' \in Q$ *all* $\lambda \in (0,1)$ *and for* $s = \max(p_2, q_1, q_2)$,

$$|v_\theta(X) - v_{\theta'}(X)| \leq C(\lambda)|\theta - \theta'|^\lambda \beta_s(X)$$
$$|\Pi_\theta v_\theta(X) - \Pi_{\theta'} v_{\theta'}(X)| \leq C(\lambda)|\theta - \theta'|^\lambda \beta_s(X)$$

(jv) $(I - \Pi_\theta)v_\theta = f_\theta - h(\theta)$

If D is an open subset of the parameter space Θ, then on examining Equations (3.33) and (3.34) and the definition of the class $L(Q, L_1, L_2, p_1, p_2)$ that for any point $\theta \in D$ the function $H(\theta, X)$ in Equation (4.10) is of class $L(Q, L_1, L_2, p_1, p_2)$ for some compact set $Q \subset D$ that contains θ, for $p_1 = p_2 = 1$, and for choices of L_1 and L_2 that depend on Q. Consequently, provided that the premise of Theorem 53-*bis* is valid, the conclusions imply that each of the conditions in Assumption 5-*bis* hold. It remains only to show that the conditions in the premise of Theorem 53-*bis* are indeed satisfied for the estimator Markov chain.

Proposition 54. *Given that* $p_1 = p_2 = 1$*, there exist constants* K_1, K_2, K_3, q_1, q_2 *and* ρ *such that the Markov process* X_k *with transition function given by equation (4.9) satisfies conditions (i) (ii) and (iii) in the premise of Theorem 53*-bis

Proof. The proof follows from earlier results: From Proposition 48 the Markov process X_k with transition function given by equation (4.9) has an invariant measure ν, and if f is a β_1-Lipschitz function, then

$$|\Pi_\theta^k(X_a; dX_b)f - \nu f| \leq KL_f c^k \beta_2(X_a)$$

It follows from the triangle inequality that if X_1 and X_2 are two points in $\mathcal{X}$, then

$$\begin{aligned}|\Pi_\theta^k f(X_1) - \Pi_\theta^k f(X_2)| &\leq |\Pi_\theta^k(X_1; dX_b)f - \Pi_\theta^k(X_1; dX_b)f| \\ &\leq KL_f c^k(\beta_2(X_1) + \beta_2(X_2)) \\ &\leq Kc^n N_{p_1(f)}(\beta_2(X_1) + \beta_2(X_2)).\end{aligned}$$

when $p_1 = 1$. The step to the last inequality uses the bound $L_f < N_1(f)$. The first inequality in the premise of Theorem 53-*bis* holds with $q_1 = 2$.

The second inequality in the premise of Theorem 53-*bis* is implied by inequality (ii) in Proposition 52, and conclusion (iii) of Proposition 48 establishes the third inequality in the premise of Theorem 53-*bis*. □

Chapter 6

Analysis of the control and estimation algorithm.

Part 3: Convergence of the stochastic approximation.

This chapter completes the analysis of the convergence properties of the combined control and estimation algorithm that started in Chapter 4. The first section in this chapter verifies Assumption (A.7) from Chapter 4 for the combined estimation and control algorithm by exhibiting a suitable Lyapunov function. The second section of this chapter quotes an appropriate stochastic approximation theorem from Benveniste *et al.* [7], and adapts the theory to the specific requirements of the estimation and control problem. The section then formally states the main convergence results for the estimation and control problem presented in the dissertation. These results are a direct consequence of the stochastic approximation convergence theorem.

6.1 The Lyapunov Function

This section deals with the second of the two tasks mentioned in the outline of the ODE method in Section 4.2, the task of proving asymptotic stability of the associated ODE. A Lyapunov function is given for the ODE. The choice of Lyapunov function is guided by the derivation of the recursive estimation algorithm as a stochastic approximation to a second order minimization algorithm for the function $Q(\lambda) = \mathbf{E}[\log f(x_{0,k}, y_{0,k})]$.

Recall that the logarithm of the probability density for the process distribution of $\{x_k, y_k\}$ with model $\theta(\lambda)$ is given by the formula

$$\log f(y_{0,k+1}, x_{0,k+1} \mid \lambda) = \sum_i \sum_j \sum_u n_{u;i,j}(k+1) \log A_{u;i,j} + \sum_i \sum_m m_{i,m}(k+1) \log B_{i,m} + \sum_i \delta_{e_i}(x_0) \log \pi_i,$$

The components $X_k^{\zeta,u}(i,j)$ and $X_k^{\eta,m}(i)$ of the state vector X provide empirical receding horizon estimates of the values of the transition and occupation frequencies $n_{u;i,j}(k)/k$ and $m_{i,m}(k)/k$, and if $\theta^* = \theta(\lambda^*)$ denotes the unknown values of the constrained parameters for the controlled hidden Markov model, then the Leibler Kullback measure for the distribution of k observations of the full information process is approximated (for large k) by

$$-\mathbf{E}_{\theta^*}[\log\ f(y_{0,k+1}, x_{0,k+1} \mid \lambda)] \approx -k\, \mathbf{E}_{\theta^*}\left[\sum_{u,i,j} X_{k+1}^{\zeta,u}(i,j) \log A_{u;i,j} + \sum_{i,m} X_{k+1}^{\gamma,m}(i) \log B_{i,m}\right].$$

Define $U(\theta)$ to be the approximation to the relative entropy rate

$$U(\theta) = -m_{\theta^*}\left(\sum_{u,i,j} X_\theta^{\zeta,u}(i,j) \log A_{u;i,j} + \sum_{i,m} X_\theta^{\gamma,m}(i) \log B_{i,m}\right) \tag{6.1}$$

The next two lemmas show that $U(\theta)$ has the properties needed for a Lyapunov function.

Lemma 55. *The Lyapunov function has continuous first and second derivatives.*

The proof of this lemma is follows the proof that Baum and Petrie use in [6] to establish differentiability of the entropy rate function for a hidden Markov model. The idea of the proof is to use the strong ergodicity properties of the Markov chain to approximate the expectation with respect to the invariant measure in equation (6.1) with the converging sequence of conditional expectations that are generated by the Markov transition kernel. A combination of this approximation argument and a standard convergence theorem from Lebesgue integration theory justifies the transposition of differentiation operators with the expectation operator in the expression for the Lyapunov function (equation (6.1)).

Lemma 56. *Let Ω denote the (linear) manifold of values taken by the model A_u, B, and let $\nabla_\Omega U$ be the gradient of the function U on Ω, then there exists an open neighborhood O of θ^* in U such that for all $\theta \in O$,*

$$\nabla_\Omega U|_\theta \cdot m_\theta(H(X, \theta)) \leq 0 \tag{6.2}$$

with equality holding only when $\theta = \theta^*$

Proof. Choose for local coordinates on Ω the same unconstrained parameterization that was used in Chapter 3, equation (3.7), and for each row of the matrix A_u let q_i be the index of the element in the row that does not appear in the parameter $\lambda(\theta)$. The structure of the manifold Ω ensures that the scalar product in (6.2) can be expressed as a sum with one term in the sum for each of the matrices A_u and B. The gradient of U is given by

$$-\nabla_\Omega U(\theta) = \nabla_\Omega \; m_{\theta^*} \left(\sum_{u,i,j} X_\theta^{\zeta,u}(i,j) \log A_{u;i,j} + \sum_{i,m} X_\theta^{\gamma,m}(i) \log B_{i,m} \right)$$

and when the gradient is represented in the basis induced by the local co-ordinates, the $(u;ij)$'th component, $j \neq q_i$, which is the component corresponding to the matrix entry $A_{u;ij}$ is given by the expression.

$$\frac{m_{\theta^*}[X_\theta^{\zeta,u}(i,j)]}{A_{u;ij}} - \frac{m_{\theta^*}[X_\theta^{\zeta,u}(i,q_i)]}{A_{u;iq_i}}$$

The stochastic flow $H(X, \theta)$ in equation (6.2) is defined in Section 4.2 as the update of the parameter estimates given in equations (3.33, 3.34). Using the same block structures that were introduced for the score vector and Fisher information matrix in Chapter 3, if $j \neq q_i$, the $(u; ij)$'th component of the flow $H(X, \theta)$ is

$$\begin{aligned} P_i^{-1} &\left(\frac{\zeta(i,j)}{A_{u;ij}} \delta_{u-\triangle}(u) - \frac{\zeta(i,q_i)}{A_{u;iq_i}} \delta_{u-\triangle}(u) \right) \\ &= \frac{A_{u;ij}^2}{X^{\zeta,u}(i,j)} \left(\frac{\zeta(i,j)\delta_{u-\triangle}(u)}{A_{u;ij}} - \frac{\zeta(i,q_i)\delta_{u-\triangle}(u)}{A_{u;iq_i}} \right) \\ &\qquad - \frac{A_{u;ij}^2}{X^{\zeta,u}(i,j)} \left(\sum_{l=1}^{N} \frac{A_{u;il}^2}{X^{\zeta,u}(i,l)} \right)^{-1} \\ &\qquad\qquad \times \left(\sum_{l \neq q_i} \frac{A_{u;il}^2}{X^{\zeta,u}(i,l)} \left(\frac{\zeta(i,l)\delta_{u-\triangle}(u)}{A_{u;il}} - \frac{\zeta(i,q_i)\delta_{u-\triangle}(u)}{A_{u;iq_i}} \right) \right) \end{aligned}$$

The quantities $A_{u;ij}$ in this formula are components of the parameter $\theta \in \Omega$, and ζ is the function of the random variable X given by equation (4.6). On expansion of the second term, the right hand side of the equation becomes

$$\frac{A_{u;ij}\zeta(i,q_i)\delta_{u-\triangle}(u)}{X^{\zeta,u}(i,j)} - \frac{A_{u;ij}^2}{X^{\zeta,u}(i,j)} \left(\sum_{l=1}^{N} \frac{A_{u;il}^2}{X^{\zeta,u}(i,l)} \right)^{-1} \left(\sum_{l=1}^{N} \frac{\zeta(i,l)\delta_{u-\triangle}(u)}{X^{\zeta,u}(i,l)} A_{u;il} \right).$$

The expression for the flow and the expression for the gradient, ∇U, are substituted into the inner product on the right hand side of (6.2). The sum of the terms with indices $(u; ij)$, $j \neq q_i$ is

$$\begin{aligned} m_\theta \Bigg[\sum_{u,i} \sum_{j \neq q_i} &\left(\frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))\zeta(i,j)\delta_{u-\triangle}(u)}{X_\theta^{\zeta,u}(i,j)} - \frac{A_{u;ij}\zeta(i,j)\delta_{u-\triangle}(u) m_{\theta^*}(X^{\zeta,u}(i,q_i))}{A_{u;iq_i}X^{\zeta,u}(i,j)} \right) \\ &- \left(\sum_{l=1}^{N} \frac{A_{u;il}^2}{X^{\zeta,u}(i,l)} \right)^{-1} \left(\sum_{l=1}^{N} \frac{\zeta(i,l)\delta_{u-\triangle}(u)}{X^{\zeta,u}(i,l)} A_{u;il} \right) \\ &\qquad \times \left(\sum_{j \neq q_i} \frac{A_{u;ij}^2}{X^{\zeta,u}(i,j)} \left(\frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))}{A_{u;ij}} - \frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,q_i))}{A_{u;iq_i}} \right) \right) \Bigg] \end{aligned}$$

Fix u and i, and consider a single term of the first summation. An expansion of the second term in the argument of the j summation leads to a simplified expression

$$\sum_{j=1}^{N} \frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))\zeta(i,j)\delta_{u-\Delta}(u)}{X^{\zeta,u}(i,j)} - \left(\sum_{j=1}^{N} \frac{A_{u;ij} m_{\theta^*}(X_\theta^{\zeta,u}(i,j))}{X^{\zeta,u}(i,j)}\right) \times \left(\sum_{l=1}^{N} \frac{A_{u;il}^2}{X^{\zeta,u}(i,l)}\right)^{-1} \left(\sum_{l=1}^{N} \frac{\zeta(i,l)\delta_{u-\Delta}(u)}{X^{\zeta,u}(i,l)} A_{u;il}\right) \quad (6.3)$$

First consider the case when $\theta = \theta^*$. It follows from the consistency of the empirical estimator X^ζ that $m_{\theta*}(X^{\zeta,u}(i,j)) = \pi(i)A_{u;ij}$ where $\pi(i)$ is the invariant measure of the probability kernel A_u. Making this substitution in (6.3) gives the value 0 for all u and i.

For the case when $\theta \neq \theta^*$, define a discrete probability measure on N points by assigning to the point j the mass

$$P_j = \frac{A_{u;ij}^2}{X^{\zeta,u}(i,j)} \left(\sum_{l=1}^{N} \frac{A_{u;il}^2}{X^{\zeta,u}(i,l)}\right)^{-1},$$

then for fixed i and u, the expression (6.3) becomes

$$\delta_{u-\Delta}(u) \sum_{l=1}^{N} \frac{A_{u;il}^2}{X^{\zeta,u}(i,l)} \left(\sum_{j} \frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))\zeta(i,j)}{A_{u;ij}^2} P_j - \sum_{j} \frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))}{A_{u;ij}} P_j \sum_{j} \frac{\zeta(i,j)}{A_{u;ij}} P_j\right). \quad (6.4)$$

Define a family of random variables $\epsilon(u;i,l)$ by $X^{\zeta,u}(i,l) = A_{u;il}\pi(i)(1+\epsilon(u;i,l))$. Under the probability distribution m_θ, $\epsilon(u;i,l)$ is a zero-mean random variable with higher moments controlled by the estimator parameter q in equation (4.4). When u and i are fixed, the random variable P_j has an expansion

$$P_j = A_{u;ij}\Big(1 - \epsilon(u;i,j) + \epsilon^2(u;i,j) - \sum_{l} A_{u;il}\epsilon^2(u;i,l) + o(\epsilon^2(u;i,j))\Big), \quad (6.5)$$

and the expected value of the difference in (6.4) with respect to the measure m_θ has an approximation

$$m_\theta\Big(\sum_j \frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))\zeta(i,j)}{A_{u;ij}^2}P_j - \sum_j \frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))}{A_{u;ij}}P_j \sum_j \frac{\zeta(i,j)}{A_{u;ij}}P_j\Big)$$

$$= \sum_j m_\theta(\zeta(i,j)\epsilon(u;i,j))\hat{\pi}_i + \sum_j \frac{m_{\theta^*}(X_\theta^{\zeta,u}(i,j))}{A_{u;ij}}$$

$$\times\, m_\theta\Big(\sum_l \zeta(i,l)A_{u;ij}\epsilon(u;i,j) - \zeta(i,j)\epsilon(u;i,j)\Big), \qquad (6.6)$$

where $\hat{\pi}_i = \sum_l m_{\theta^*}(X_\theta^{\zeta,u}(i,l))$. Recall that the random variable $\zeta(i,j)$ is an *a-posteriori* probability calculated from the prior $X^{\zeta;u}(i,j)$, so bias in $\zeta(i,j)$ is proportional to bias in $X^{\zeta;u}(i,j)$, and the quantity $m_\theta(\zeta(i,j)\epsilon(u;i,j))$ is strictly positive. In addition, $m_\theta\Big(\sum_l \zeta(i,l)A_{u;ij}\Big) = m_\theta(\zeta(i,j))$, and since $A_{u;ij}$ and $\zeta(i,l)$ are both bounded away from zero, the first term on the right hand side of (6.6) dominates the second provided that $A^*_{u;ij}/A_{u;ij}$ is not too large. Since $A^*_{u;ij}$ is bounded away from zero, provided that q is sufficiently small that the linear terms in the approximation (6.5) dominate the higher order terms, there exists O an open neighborhood of A^* such that the inequality (6.2) holds for all $A \in O$ □

Proposition 57. *Let θ^* be the value in the parameter space Θ that maps onto the state transition and output matrices defined in equations (2.2) and (2.3), and let m_θ be the invariant distribution for the Markov kernel $\Pi_\theta(X_a; dX_b)$ defined in equation (4.9). The stochastic approximation problem with update equation (4.10) satisfies Assumption 7 of Chapter 4 when $h(\theta) = m_\theta(H(X,\theta))$ and $U(\theta)$ is defined by equation (6.1).*

Proof. Lemma 55 established regularity of the function U, and Lemma 56 establishes that the directional derivative of U in the direction of $h(\theta)$ is non positive in an open neighborhood of θ^*. The assumption is satisfied for any open set D that is compactly contained in the open neighborhood of θ^* on which the directional derivative of U is non-negative. $\square$

6.2 The stochastic approximation result.

Convergence of the combined estimation and control argument is established by the direct application of a stochastic approximation result from Benveniste *et al.*. Most of the work in the dissertation is directed towards establishing that the assumptions in the premise of the theorem hold in the particular case of the stochastic approximation problem posed in Section 4.2.

The stochastic approximation scheme that Benveniste *et al.* treat is described by the equation

$$\theta_{n+1} = \theta_n + \gamma_{n+1} H(\theta_n, X_{n+1}) + \gamma_{n+1}^2 \rho_{n+1}(\theta_n, X_{n+1}). \tag{6.7}$$

Let $P_{x,a}$ denote the distribution of (X_k, θ_k) with $X_0 = x$, $\theta_0 = a$, and let F be a compact subset of D that satisfies:

$$F = \{\theta; U(\theta) \leq c_0\} \supset \{\theta; U'(\theta) \cdot h(\theta) = 0\} \tag{6.8}$$

Let α be the exponent that assumption (A.6) postulates, and define $q_0(\alpha) = \sup\{2, 2(\alpha - 1)\}$. Theorem 17 from part II, Chapter 3 of Benveniste *et al.* states the following:[1]

Theorem 58. *[7, p. 304, Theorem 17] We assume (A.1), (A.2), (A.3), (A.4), (A.5), (A.6) and (A.7), and suppose that F is a compact set satisfying (6.8). Then, for any compact subset Q of D, and $q \geq q_0(\alpha)$, there exist constants B, s, such that for all*

[1] The statement of the theorem is altered to keep references in the statement of the theorem consistent with labeling scheme in the Dissertation

$a \in Q$ and all $x \in \mathbb{R}^k$:

$$P_{x,a}(\theta_k \text{ converges to } F) \geq 1 - B(1 + |x|^s) \sum_{k \geq 1} \gamma_k^{1+q/2}.$$

The assumptions (A.1) – (A.7) in the premise of the theorem are the seven assumptions listed in Section 4.2. The multiplicative constant B and the exponent s are functions of Q, but independent of X_0. In particular, s is a function of the exponents in assumptions (A.4) and (A.5).

Both the assumptions in Theorem 58 and the theorem itself are stated in terms of a discrete-time Markov process that evolves on the Euclidean space $\mathbb{R}^k$. The following restatement of the theorem extends its applicability to discrete time Markov processes that evolve in the more general metric spaces that this dissertation uses.

Theorem 58-*bis*. *Assume that Assumptions 1, 2-*bis*, 3-*bis*, 4-*bis*, 5-*bis*, 6 and 7 from Section 4.2 all hold for some open set D that is compactly contained in Θ. Suppose, also, that F is a compact set that satisfies (6.8). Then for any compact subset Q of D, and $q \geq 1$, there exist constants B and s, such that for all $a \in Q$, and all $X \in \mathcal{X}$:*

$$P_{X,a}(\theta_k \text{ converges to } F) \geq 1 - B\beta_s(X) \sum_{k \geq 1} \gamma_k^{1+q}.$$

Theorem 58-*bis* is the key ingredient in the proof of the main results of the dissertation. The results are posed in the notation of Chapters 2 and 3. Let Θ denote the space of possible values for state transition matrices A_u and output matrices B of the controlled, finite state, finite output, hidden Markov model from Chapter 2. Θ is a subset of a linear sub-manifold in a finite dimensional Euclidean space, and inherits a topology from the Euclidean space. Elements of Θ are denoted by (A_u, B). Given a control policy that determines $u(k)$ as a function of previous outputs $y(k)$ and an initial state x_0, the hidden Markov model together with the control policy and an initial value x_0 for the hidden Markov model state determine the statistics of the input process $u(k)$, and the output process $y(k)$. The input and output processes $u(k)$ and $y(k)$ together with initial values for the recursively defined quantities α, Z and Γ, an initial estimate $\theta(0) = (A_u(0), B(0))$, and the recursive estimation equations (3.33) and (3.34) determine a sequence of estimates $\theta(k) = (A_u(k), B(k))$ for θ^*. This sequence of estimates depends on the parameter κ through the recursive estimation equations.

Theorem 59. *Let A_u^* and B^* satisfy the inequalities in equation (2.4), choose the randomized, finite-horizon, risk-sensitive output-feedback control policy defined in equation (2.14), and let the elements of the initial values for α, Z and Γ be bounded away from 0. There exists $D \subset \Theta$, an open neighborhood of $\theta^* = (A_u^*, B^*)$ such that if Q is a compact subset of D, and O is an open neighborhood of θ^*, and $0 < \epsilon \leq 1$ then if $\theta_0 = (A_u(0), B(0)) \in Q$, and κ is sufficiently small, the sequence of estimates $\theta(k)$ is eventually contained in the set O with probability bounded below by $1 - \epsilon$.*

Proof. Provided the assumptions in the premise of Theorem 58-*bis* hold, the conclusion of Theorem 59 is a consequence of the conclusion of Theorem 58-*bis* For, if θ^* is the point in parameter space corresponding to (A_u^*, B^*), then θ^* is a local minimum of the Lyapunov function U, which is twice differentiable in a neighborhood of θ^*. Consequently, if O is an open neighborhood of θ^*, then for δ sufficiently small, there exists a compact set set $F = \{\theta : U(\theta) \leq U(\theta^*) + \delta\}$ such that $F \subset O$. From the conclusion of Theorem 58-*bis*,

$$P_{X,a}(\theta_k \text{ converges to } F) \geq 1 - B\beta_s(X) \sum_{k \geq 1} \gamma_k^{1+q},$$

and for any $\epsilon > 0$ the inequality $B\beta_s(X) \sum_{k\geq 1} \gamma_k^{1+q} < \epsilon$ holds provided that κ is sufficiently small.

The only remaining task is the checking of the assumptions in the premise of Theorem 58-*bis*. Assumptions 1 is the assumption that $\sum \gamma_k$ is divergent. This is true since $\gamma_k = 1/(\kappa + k)$. Assumption 2-*bis* is the assumption that the random perturbations in the stochastic approximation algorithm have statistics that are governed by a discrete time Markov Process. Proposition 7 demonstrates this in the current context. Assumption 3-*bis* bounds the size of the generator $H(\theta, X)$ in equation (4.10) as the the random perturbation X becomes 'large'. It is clear from the form of equations (4.11) and (4.12), and the definition of weighting function $\beta_s(X)$ in equation (4.14) that there exists a constant C independent of X such that $|H(\theta, X)| < C\beta_1(X)$, and that Assumption 3-*bis* is satisfied. Assumption 4-*bis* places bounds on the moments of the iterated kernels of the Markov chains, and Assumption 5-*bis* asserts the existence of regular solutions to the Poisson equation $(I - \Pi_\theta)v_\theta = H_\theta - h(\theta)$. Propositions 52 and 54 establish that these assumptions hold. Assumption 6 requires that the series $\sum \gamma_k^\alpha$ is summable for some $\alpha > 1$. Since $\gamma_k = 1/(\kappa + k)$, this is true for any $\alpha > 1$. Finally, Proposition 57 proves the existence of the Lyapunov function that Assumption 7 requires. $\square$

Theorem 59 and theorem 6 together prove the existence of an assymptotically-optimal, risk-sensitive, output feedback controller for a finite-state, partially-observed hidden Markov model with unknown state transition and output matrices.

Chapter 7

Conclusion

The work contained in this dissertation is directed towards a single result: a proof that, with a careful choice of initial parameters, an adaptive control policy for a restricted, simple class of systems has good asymptotic behavior. The result is not a strong one, and falls well short of the long term objective: the development of a unifying theory for the dual control problem that is applicable to the variety of applications discussed in the introduction. This state of affairs leads to the question "Is the result worth the effort?". The answer to the question is "Yes!". The dissertation approaches the dual control problem in a setting that is sophisticated enough to exhibit some of the problems that make the analysis difficult, yet simple enough that the analysis is tractable. The benefit of this approach lies not in the convergence result itself, but in the insights that the analysis gives about dual control problems in general, and in the potential application of methods developed for the simple problem to broader classes of problems. This final chapter reviews the analysis in the dissertation, commenting on the problems that the analysis avoids, the insights that the analysis provides, and how new techniques might be developed to treat more general problems.

The Formulation of the Control problem:

There is an inconsistency in the definition of the control problem in Chapter 2. The problem is posed as an output feedback problem, without knowledge of the state transition and output matrices, yet the incremental cost is given as a function of the state and input. A more natural exposition of the problem would specify the incremental cost as a separate output process of the system, providing additional information that is not available for feedback control.

A home heating control provides an example of such a problem. The processes that the controller on the wall observes is the temperature at the controller's thermostat. The incremental cost process is a combination of a measure of discomfort for the occupants of the house, and the gas bill that comes every couple of months. The variables that comprise the incremental cost are not directly available to the controller, and the observation process that is available to the controller does not provide sufficient statistics for the incremental cost process.

An output feedback problem posed in this way provides no guide to the state model. Given a cost criterion such as quadratic mean, risk sensitive, or minimax, the 'natural' way to approach the control and estimation problem is to choose the space for the information state, and estimate the information state recursion operator in equation (2.10) directly, and use the dynamic programming equation (2.13) to compute the optimal control. The finite state model with an incremental cost function provides a tractable alternative to the difficult problem of directly estimating the information state recursion. The formation in Chapter 2 simplifies the problem one stage further by assuming that the cost functional is known *a-priori*. This simplification permits the use of an existing estimation algorithm in Chapter 3.

Randomized Policies:

The use of a Gibbs distribution for the randomized policy in equation (2.14) ensures both continuity of the policy with respect to perturbations in the estimated model, and "persistent excitation" in a way that naturally avoids costly choices of control.

The notion of persistent excitation in adaptive and intelligent control captures the idea that if an adaptive controller is to converge to an optimal policy, then the controller needs to 'explore' the state space to determine an empirical estimates for the cost function and before deciding which action minimizes the incremental cost. The analysis in this dissertation shows that the concept of persistent excitation is closely tied to the ergodicity properties of the chain $\tilde{X}$ which includes the state process of the underlying model. Propositions 9 and 10 use the good support properties of the family of randomized policies $\upsilon_\theta(u)$ to prove primitivity of the kernel $\Pi(\tilde{X}_a; d\tilde{X}_b)$, the existence of an invariant measure $\tilde{\nu}$ for Π, and the weak convergence of the sequence of densities $\nu_0\Pi^k$ to the invariant density $\tilde{\nu}$. The ergodicity result for the kernel of the chain $\tilde{X}$ implies that the chain is recurrent on the entire state space and provides a concrete meaning to the concept of persistent excitation. It is quite easy to trace in the analysis of Section 5.3 the contribution of persistent excitation to the the convergence result for the combined control and estimation algorithm. Proposition 42 uses the ergodicity result for the kernel of the chain $\tilde{X}$ to derive a similar ergodicity result for the chain $(\tilde{X}, X^\alpha)$, and Proposition 48 uses the result for $(\tilde{X}, X^\alpha)$ in turn to prove geometric ergodicity for the full chain $(\tilde{X}, X^\alpha, X^\gamma, X^\zeta)$. Propositions 52 and 54 use the ergodicity result for the full chain to establish Assumptions 4-*bis* and 5-*bis* of the stochastic approximation theorem, Theorem 58-*bis*

In addition to providing a policy that satisfies the requirement of persistent excitation, the Gibbs policy ensures continuity of the control with respect to variations in the model estimates. This condition on the control, which is intuitively a practical requirement for an adaptive control system, is also a requirement for the stochastic approximation convergence proof. Continuity in this instance can be precisely interpreted by imposing a Levy metric on the space of distributions for the randomized controls, and any convenient metric on the model parameters. The continuity result for the control policy propagates through the analysis in a fashion similar to the propogation of the persistent excitation condition, and ultimately forms a component of the result that establishes Assumption 5-*bis* of the stochastic approximation theorem.

Stochastic Approximation Formulation:

The essence of the ODE approach to stochastic approximation results is to divide the evolution into a deterministic component and a stochastic component. The deterministic component is interpreted as a discrete approximation to a stable ODE with an associated Lyapunov function, and the random component is interpreted as a random perturbation acting on the ODE trajectories. The two difficult problems are the determination of a Lyapunov function for the auxiliary ODE, and the proof that the error accumulated by the random perturbations does not affect the eventual convergence of the trajectories. In practical applications of the ODE method such as the application to the problem in this dissertation provide a degree of choice over where to draw the line between the deterministic and stochastic components of the decomposition. Associated with this choice is a tradeoff between the difficulty of determining a Lyapunov function, and the difficulty of bounding the cumulative random error. This dissertation chooses to make the Lyapunov problem easy at the expense of the problem of bounding the error. A suitable Lyapunov function falls out of the derivation of the estimation algorithm in Chapter 3. The estimation is a second order gradient search for the minimum of the Leibler Kullback measure for the full information process, and Chapter

6.2 uses an approximation to the relative entropy rate for a Lyapunov function.

The cost of this approach of that the is that the occupation frequency estimators which are denoted by Z^u_{k+1} and Γ_{k+1} in Chapter 3 become part of the 'random perturbation' even though these objects are deterministic functions of the much simpler Markov chain $\tilde{X}$. In the problem that this dissertation addresses the Markov chain $\tilde{X}$ is a finite state Markov chain with a primitive kernel, and this simple chain provides sufficient statistics for the full Markov chain X which evolves over a complex topological sum of projective spaces. The geometric decomposition of the chain X in Chapter 4.1, and the new results that establish a potential theory for Markov modulated random walks in Chapter 5 provide the tools that manage the complexity of X.

Potential Theory for the Estimator Markov Chain.

The potential theory for the estimator chain, which is developed in Chapters 4.1 and 5, provides much of the original material in the dissertation as well as much of the detailed argument. The two aspects of the work that are unusual are the need for regularity results with respect to model variations, and the unusual structure of the random walks.

Ergodic theory for Markov modulated random walks on semigroups is new. Previous authors such as Marcus, and Le Gland and Mevel have used theory of inhomogeneous products of matrices (Seneta) to get ergodic results. The approach taken here is more precise, and fits well with the established theory of random walks on groups. A major difference between the group theory and the semigroup theory is the choice of metric on the measure space. The singularity of limit distributions in the semi-group case forces the use of a weak measure topology for the ergodic theory. In addition, the hyperbolic metric on the distribution supports means that the weak topology has to be weighted appropriately.

The key to a successful ergodic theory is a definition of primitivity in the generator that matches the notion of tightness in the measure topology. In this respect Definition 5 and the growth condition in Lemma 25 are the key requirements for the existence of invariant measures. In cases where the underlying models have more structure and less forgiving primitivity assumptions, more care will be needed here in establishing that suitable primitivity conditions hold for the corresponding estimator state processes.

Lemma 28 is the important result both for moving from random walks to Markov modulated random walks, and for establishing regularity of the invariant measures with respect to parametric variations in the parameter. The compactness argument can be relaxed in the presence of tightness, and this provides the key to extending the argument to more general spaces.

Asymptotic Convergence and Domains of Attraction: The ODE method is ultimately an asymptotic method. If the evolving system is close to a stable attractor, then it will converge to the attractor. While such results are important, they are also profoundly unsatisfactory. In practical applications an estimate of the size and location of the domains of attraction is equally important for system design. Unfortunately good estimates are generally hard to come by, but in this case the approximation (6.6) in the proof of Lemma 56 provides guidance. The auxiliary ODE converges provided that q is small enough that the linear term dominates higher order terms in the expansion (6.5). This condition confirms the intuitive knowledge that small step sizes help convergence.

Appendix A

Recursive formulae for the empirical densities ζ and γ

This appendix provides detailed derivations of the recursive formulae given in Section 3

$$\begin{aligned}\zeta_{l|K,\Lambda_k}(i,j) &= f(x_l = e_j, x_{l-1} = e_i \mid y_{0,K}, \Lambda_k)\\ &= \frac{f(x_l = e_j, x_{l-1} = e_i, y_{0,K} \mid \Lambda_k)}{\sum_{i,j} f(x_i = e_j, x_{l-1} = e_i, y_{0,K} \mid \Lambda_k)}.\end{aligned}$$

The density in the numerator is evaluated by

$$\begin{aligned}&f(x_l = e_j, x_{l-1} = e_i, y_{0,K} \mid \Lambda_k)\\ &\qquad = f(x_l = e_j, y_{l,K} \mid x_{l-1} = e_i, y_{0,l-1}, \Lambda_k) f(x_{l-1} = e_i, y_{0,l-1} \mid \Lambda_k) \quad \text{(A.1)}\end{aligned}$$

with

$$\begin{aligned}&f(x_l = e_j, y_{l,K} \mid x_{l-1} = e_i, y_{0,l-1}, \Lambda_k)\\ &\qquad = f(y_{l,K} \mid x_l = e_j, \Lambda_k) f(x_l = e_j \mid x_{l-1} = e_i, y_{0,l-1}, \Lambda_k)\\ &\qquad = f(y_{l,K} \mid x_l = e_j, \Lambda_k) A_{u_{l-1};ij}(l \wedge k). \quad \text{(A.2)}\end{aligned}$$

Substituting (A.2) in (A.1) yields:

$$\begin{aligned}&f(x_l = e_j, x_{l-1} = e_i, y_{0,K} \mid \Lambda_k)\\ &\qquad = A_{u_{l-1};ij}(l \wedge k) f(y_{l,K} \mid x_l = e_j, \Lambda_k) f(x_{l-1} = e_i, y_{0,l-1} \mid \Lambda_k) \quad \text{(A.3)}\end{aligned}$$

Let

$$\tilde{\alpha}_{l|\Lambda_k}(i) = f(x_l = e_i, y_{0,l} \mid \Lambda_k)$$
$$\alpha_{l|\Lambda_k}(i) = f(x_l = e_i \mid y_{0,l}, \Lambda_k) = \frac{\tilde{\alpha}_{l|\Lambda_k}(i)}{\sum_i \tilde{\alpha}_{l|\Lambda_k}(i)}$$
$$\beta_{l|K,\Lambda_k}(j) = f(y_{l+1,K} \mid x_{l+1} = e_j, \Lambda_k).$$

Note that, as a consequence of the definition of the empirical densities, if $k \geq l$ then

$$\alpha_{l|\Lambda_k}(i) = \alpha_{l|\Lambda_l}(i).$$

Using the expressions for α and β in (A.3) gives

$$f(x_l = e_j, x_{l-1} = e_i, y_{0,K} \mid \Lambda_k) = \tilde{\alpha}_{l-1,\Lambda_k}(i) A_{u_{l-1};ij}(l \wedge k) \beta_{l-1|K,\Lambda_k}(j)$$

and

$$\begin{aligned} \zeta_{l|K,\Lambda_k}(i,j) &= f(x_l = e_j, x_{l-1} = e_i \mid y_{0,K}, \Lambda_k) \\ &= \frac{\alpha_{l-1,\Lambda_k}(i) A_{u_{l-1};ij}(l \wedge k) \beta_{l-1|K,\Lambda_k}(j)}{\sum_{i,j} \alpha_{l-1,\Lambda_{l-1}}(i) A_{u_{l-1};ij}(l \wedge k) \beta_{l-1|K,\Lambda_k}(j)}. \end{aligned} \tag{A.4}$$

Equation (A.4) gives an expression for ζ in terms of the forward and backward estimates α and β. An analogous expression is obtained for the other empirical density γ as follows:

$$\begin{aligned} \gamma_{l|K,\Lambda_k}(i) &= f(x_l = e_i \mid y_{0,K}, \Lambda_k) \\ &= \frac{f(x_l = e_i, y_{0,K} \mid \Lambda_k)}{\sum_i f(x_l = e_i, y_{0,K} \mid \Lambda_k)} \end{aligned}$$

in which

$$
\begin{aligned}
f(x_l = e_i, y_{0,K} \mid \Lambda_k) &= \sum_j f(x_{l+1} = e_j, x_l = e_i, y_{0,K} \mid \Lambda_k) \\
&= \sum_j f(x_{l+1} = e_j, y_{l+1,K} \mid x_l = e_i, \Lambda_k) f(x_l = e_i, y_{0,l} \mid \Lambda_k) \\
&= \sum_j f(y_{l+1,K} \mid x_{l+1} = e_j, \Lambda_k) A_{u_l;ij}((l+1) \wedge k) \\
&\qquad f(x_l = e_i, y_{0,l} \mid \Lambda_k) \\
&= \sum_j \beta_{l,K|\Lambda_k}(j) A_{u_l;ij}((l+1) \wedge k) \tilde{\alpha}_{l|\Lambda_k}(i).
\end{aligned}
$$

Putting this together gives

$$
\gamma_{l|K,\Lambda_k}(i) = \frac{\sum_j \beta_{l,K|\Lambda_k}(j) A_{u_l;ij}((l+1) \wedge k) \alpha_{l,|\Lambda_k}(i)}{\sum_i \sum_j \beta_{l,K|\Lambda_k}(j) A_{u_l;ij}((l+1) \wedge k) \alpha_{l,|\Lambda_k}(i)} \tag{A.5}
$$

All that remains is the derivation of the recursive formulae for α and β.

$$
\begin{aligned}
\alpha_{l|\Lambda_k}(j) &= f(x_l = e_j \mid y_{0,l}, \Lambda_k) \\
&= \sum_i f(x_l = e_j, x_{l-1} = e_i, \mid y_{0,l}, \Lambda_k) \\
&= \frac{\sum_i f(y_l, x_l = e_j \mid x_{l-1} = e_i, \Lambda_k) f(x_{l-1} = e_i \mid y_{0,l-1}, \Lambda_k)}{\sum_j \sum_i f(y_l, x_l = e_j \mid x_{l-1} = e_i, \Lambda_k) f(x_{l-1} = e_i \mid y_{0,l-1}, \Lambda_k)} \\
&= \frac{\sum_i \langle e_j, B(l \wedge k) y_l \rangle A_{u_{l-1};ij}(l \wedge k) \alpha_{l-1|\Lambda_k}(i)}{\sum_j \sum_i \langle e_j, B(l \wedge k) y_l \rangle A_{u_{l-1};ij}(l \wedge k) \alpha_{l-1|\Lambda_k}(i)}
\end{aligned}
$$

with

$$
\begin{aligned}
\alpha_{0|\Lambda_k}(j) &= f(x_0 = e_j \mid y_0, \Lambda_k) \\
&= \frac{\langle e_j, B(0) y_0 \rangle \pi_j}{\sum_j \langle e_j, B(0) y_0 \rangle \pi_j}.
\end{aligned}
$$

Finally, β is calculated using the backward recursion

$$\begin{aligned}
\beta_{l|K,\Lambda_k}(i) &= f(y_{l+1,K} \mid x_{l+1} = e_i, \Lambda_k) \\
&= \sum_j f(y_{l+2,K}, y_{l+1}, x_{l+2} = e_j \mid x_{l+1} = e_i, \Lambda_k) \\
&= \sum_j \; f(y_{l+2,K}, x_{l+2} = e_j \mid x_{l+1} = e_i, \Lambda_k) f(y_{l+1}, \mid x_{l+1} = e_i, \Lambda_k) \\
&= \sum_j \; f(y_{l+2,K} \mid x_{l+2} = e_j, x_{l+1} = e_i, \Lambda_k) \\
&\qquad\qquad \times f(x_{l+2} = e_j \mid x_{l+1} = e_i, \Lambda_k) f(y_{l+1}, \mid x_{l+1} = e_i, \Lambda_k) \\
&= \sum_j \beta_{l+1|K,\Lambda_k}(j) A_{u_{l+1};ij}((l+2) \wedge k) \langle e_i, B((l+1) \wedge k) y_{l+1} \rangle
\end{aligned}$$

with

$$\beta_{K|K,\Lambda_k}(i) = 1.$$

Appendix B

Derivation of cost functional for information state dynamics

This appendix contains a derivation of the expression for the cost functional given in Equation 2.7. The result depends on a fundamental lemma which gives a formula for the transformation induced on a conditional expectation by a transformation in the underlying measure. Elliot et al. call this lemma the conditional Bayes' theorem, and give a proof for it in [11].

Let $(\Omega, \mathcal{X}, P)$ be a probability space, and let $\mathcal{Y}$ be a sub-sigma algebra of $\mathcal{X}$. Let $f(x)$ be a random variable on $(\Omega, \mathcal{X})$, and recall that $\mathbf{E}\left[f(x) \mid \mathcal{Y}\right]$, the conditional expectation of $f(x)$ with respect to $\mathcal{Y}$, is defined as the unique $\mathcal{Y}$ measurable function that satisfies the equation

$$\int \mathbf{1}_Y \mathbf{E}\left[f(x) \mid \mathcal{Y}\right]\, dP_{\mathcal{Y}} = \int f(x)\, dP \qquad \forall\, Y \in \mathcal{Y},$$

where $dP_{\mathcal{Y}}$ in the integral on the left-hand side is the restriction of the measure dP to the sigma algebra $\mathcal{Y}$. Existence and uniqueness of conditional expectations is guaranteed by the Radon Nikodym theorem.

Lemma 60. *Let $P^{\dagger}$ be a second probability measure on $(\Omega, \mathcal{X})$. Suppose that P is* *absolutely continuous with respect to $P^\dagger$, and that $dP = \Lambda\, dP^\dagger$, then:*

$$\mathbf{E}\left[f(x) \mid \mathcal{Y}\right] = \frac{\mathbf{E}^\dagger\left[\Lambda f(x) \mid \mathcal{Y}\right]}{\mathbf{E}^\dagger\left[\Lambda \mid \mathcal{Y}\right]}$$

Proof. Consider first the restricted measures $dP_\mathcal{Y}$ and $dP^\dagger_\mathcal{Y}$. For any $Y \in \mathcal{Y}$,

$$\begin{aligned}
\int \mathbf{1}_Y\, dP_\mathcal{Y} &= \int \mathbf{1}_Y\, dP \\
&= \int \mathbf{1}_Y \Lambda\, dP^\dagger \\
&= \int \mathbf{1}_Y \mathbf{E}^\dagger\left[\Lambda \mid \mathcal{Y}\right]\, dP^\dagger_\mathcal{Y}
\end{aligned}$$

and it follows that $dP_\mathcal{Y} = \mathbf{E}^\dagger\left[\Lambda \mid \mathcal{Y}\right]\, dP^\dagger_\mathcal{Y}$. Again, let Y be any element of $\mathcal{Y}$, and $f(x)$ be a random variable on $(\Omega, \mathcal{X})$.

$$\begin{aligned}
\int \mathbf{1}_Y \mathbf{E}\left[f(x) \mid \mathcal{Y}\right] \mathbf{E}^\dagger\left[\Lambda \mid \mathcal{Y}\right]\, dP^\dagger_\mathcal{Y} &= \int \mathbf{1}_Y \mathbf{E}\left[f(x) \mid \mathcal{Y}\right]\, dP_\mathcal{Y} \\
&= \int \mathbf{1}_Y f(x)\, dP \\
&= \int \mathbf{1}_Y \Lambda f(x)\, dP^\dagger \\
&= \int \mathbf{1}_Y \mathbf{E}^\dagger\left[\Lambda f(x) \mid \mathcal{Y}\right]\, dP^\dagger_\mathcal{Y}
\end{aligned}$$

□

Returning to the notation of Section 2, recall that the information state is defined by the formula

$$\sigma_k^\gamma(x) = \mathbf{E}^\dagger\left[I_{\{x_k = x\}} \Lambda_k \exp \frac{1}{\gamma} \left(\sum_{l=0}^{k-1} \phi(x_l, u_l) \right) \mid \mathcal{Y}_k \right].$$

With the use of the lemma and the formula for the information state, the cost function is rewritten in terms of the information state as follows

$$
\begin{aligned}
\mathcal{J}^{\gamma}(u) &= \mathbf{E}\left[\textstyle\sum_j \mathbf{E}\left[I_{\{x_K=e_j\}} \exp 1/\gamma \left(\phi_f(x_K) + \sum_{l=0}^{K-1}\phi(u_l, x_l)\right) \mid \mathcal{Y}_K\right]\right] \\
&= \mathbf{E}\left[\textstyle\sum_j \exp(\phi_f(e_j)/\gamma)\right. \\
&\qquad \left.\mathbf{E}^{\dagger}\left[I_{\{x_K=e_j\}}\Lambda_K \exp\left(1/\gamma \sum_{l=0}^{K-1}\phi(u_l, x_l)\right) \mid \mathcal{Y}_K\right] / \mathbf{E}^{\dagger}\left[\Lambda_K \mid \mathcal{Y}_K\right]\right] \\
&= \mathbf{E}\left[\textstyle\sum_j \exp(\phi_f(e_j)/\gamma)\sigma_K^{\gamma}(e_j)/\mathbf{E}^{\dagger}\left[\Lambda_K \mid \mathcal{Y}_K\right]\right] \\
&= \mathbf{E}^{\dagger}\left[\Lambda_K \langle \sigma_K^{\gamma}, \exp(\phi_f(e_j)/\gamma\rangle / \mathbf{E}^{\dagger}\left[\Lambda_K \mid \mathcal{Y}_K\right]\right] \\
&= \mathbf{E}^{\dagger}\left[\mathbf{E}^{\dagger}\left[\Lambda_K \langle \sigma_K^{\gamma}(\cdot), \exp(\phi_f((\cdot))/\gamma\rangle / \mathbf{E}^{\dagger}\left[\Lambda_K \mid \mathcal{Y}_K\right] \mid \mathcal{Y}_K\right]\right] \\
&= \mathbf{E}^{\dagger}\left[\langle \sigma_K^{\gamma}(\cdot), \exp(\phi_f(\cdot)/\gamma\rangle \mathbf{E}^{\dagger}\left[\Lambda_K \mid \mathcal{Y}_K\right] / \mathbf{E}^{\dagger}\left[\Lambda_K \mid \mathcal{Y}_K\right]\right] \\
&= \mathbf{E}^{\dagger}\left[\langle \sigma_K^{\gamma}(\cdot), \exp(\phi_f(\cdot)/\gamma\rangle\right].
\end{aligned}
$$

Bibliography

[1] Aristotle Arapostathis, Fernández-Gaucherand, and Steven I. Marcus. Analysis of an adaptive control scheme for a partially observed controlled Markov chain. In *Proceedings of the 29'th conference on decision and control*, pages 1438–1444, 1990.

[2] Aristotle Arapostathis and Steven I. Marcus. Analysis of an identification algorithm arising in the adaptive estimation of Markov chains. *Mathematics of Control, Signals and Systems*, 3:1–29, 1990.

[3] K. J. Astrom and B. Wittenmark. On self-tuning regulators. *Automatica*, 9:185–199, 1973.

[4] J.S. Baras and M.R. James. Robust and risk-sensitive output feedback control for finite state machines and hidden Markov models. *To appear in Mathematics of Controls, Signals and Systems.*, 1993.

[5] Andew C. Barto, Steven J. Bradtke, and Satinder P. Singh. Learning to act using real-time dynamic programming. *Artificial Intelligence*, 72:81–138, 1995.

[6] Leonard E. Baum and Ted Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *Annals of Mathematical Statistics*, 37:1554–1563, 1996.

[7] Albert Benveniste, Michel Métivier, and Pierre Priouret. *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, 1990. Translation of “Algorithmes adaptatifs et approximations stochastiques”, Masson, Paris, 1987.

[8] Dimitri P. Bertsekas and John N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, Mass, U.S.A., 1996.

[9] Jean-Dominique Deuschel and Daniel W. Stroock. *Large Deviations*, volume 137 of *Pure and applied mathematics*. Academic Press, 1989.

[10] J. L. Doob. *Stochastic Processes*. John Wiley and Sons, 1953. Wiley Classics Library edition, 1990.

[11] R.J. Elliot, L. Aggoun, and J.B. Moore. *Hidden Markov Models*. Springer-Verlag, 1995.

[12] White *et al.* *A Handbook of Intelligent Control*. Van Norstrand Reinhold, 1992.

[13] A. A. Feldbaum. Dual control theory I. *Automation and Remote Control*, 21:874–880, 1960. Translated from Automatika i Telemekhanika, Vol 21 # 11.

[14] E. Fernández-Gaucherand and S.I. Marcus. risk-sensitive optimal control of hidden Markov models: structural results. Technical Report TR 96-79, Institute for Systems Research, University of Maryland, College Park, Maryland 20742 U.S.A., 1996. obtainable from http://www.isr.umd.edu.

[15] Emannuel Fernández-Gaucherand, Aristotle Arapostathis, and Steven I. Marcus. On the adaptive control of a partially observable Markov decision process. In *Proceedings of the 27’th conference on decision and control*, pages 1204–1210, 1988.

[16] Emmanuel Fernández-Gaucherand, Aristotle Arapostathis, and Steven I. Marcus. On partially observable Markov decision processes with an average cost criterion. In *Procedings of the 28'th conference on decision and control*, pages 1267–1272, 1989.

[17] Emmanuel Fernández-Gaucherand, Aristotle Arapostathis, and Steven I. Marcus. Analysis of an adaptive control scheme for a partially observed controlled Markov chain. *IEEE Transactions on Automatic Control*, 38(6):987–993, 1993.

[18] W. H. Fleming and D. Hernández-Hernández. risk sensitive control of finite state machines on an infinite horizon ii. preprint, Division of Applied Mathematics, Brown University, 1997. available from http://www.dam.brown.edu.

[19] Mathew R. James, John S. Baras, and Robert J. Elliot. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE transactions on automatic control*, 39(4):780–791, 1994.

[20] Thomas Kailath. *Linear systems.* Prentice Hall, first edition, 1980.

[21] V Krishnamurthy and J. B. Moore. On-line estimation of hidden markov model parameters based on the. *IEEE Transactions on Signal Processing*, 41(8):2557 – 2573, August 1993.

[22] Harold J. Kushner and Dean S. Clark. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Applied mathematical statistics 26. Springer-Verlag, 1978.

[23] Francois Le Gland and Laurent Mevel. Geometric ergodicity in hidden Markov models. Technical Report No. 1028, IRISA/INRIA, Campus de Beaulieu, 35042 Rennes Cédex, France., 1996.

[24] François LeGland and Mével Laurent. Recursive identification of HMM's with observations in a finite set. In *Proceedings of the 34th conference on Decision and Control*. IEEE, 1995.

[25] L. Ljung. Analysis of recursive stochastic algorithms. *IEEE Transactions on Automatic Control*, AC-22:551 – 575, 1977.

[26] Lennart Ljung. On positive real transfer functions and the convergence fo some recursive schemes. *IEEE Transactions on Automatic Control*, AC-22(4):539–551, 1977.

[27] M. Loève. *Probability Theory I*. Springer-Verlag, 1977.

[28] M. Loève. *Probability Theory II*. Springer-Verlag, 1978.

[29] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biophysics*, 5:115–133, 1943.

[30] Kumpati S Narendra and Prabhakar Kudva. Stable adaptive schemes for system identification and control — part 1. *IEEE Transactions in Systems, Man and Cybernetics*, SMC-4(6):542–551, 1974.

[31] Kumpati S Narendra and Prabhakar Kudva. Stable adaptive schemes for system identification and control — part 1. *IEEE Transactions in Systems, Man and Cybernetics*, (6):552–560, 1974.

[32] Roger D. Nussbaum. *Hilbert's projective metric and iterated nonlinear maps.*, volume 75 of *Memoirs of the American Mathematical Society*. American Mathematical Society, September 1988.

[33] Lawrence R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257 – 285, 1989.

[34] D. Revuz. *Markov Chains*. North-Holland, 1975.

[35] H. Robbins and S. Munro. A stochastic approximation method. *Annals of Mathematical Statistics*, 22:400–407, 1951.

[36] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales*, volume 1. Wiley, second edition, 1994.

[37] Sheldon M. Ross. Quality control under Markovian deterioration. *Management Science*, 17(9):587–596, 1971.

[38] E. Seneta. *Non-negative matrices and Markov chains*. Springer-Verlag, New York, second edition, 1981.

[39] C. E. Shannon and McCarthy J. *Automata Studies*, chapter Preface, page v. Princeton, 1956.

[40] Adam Shwartz and Armand Makowski. Comparing policies in markov decision processes: Mandel's lemma revisited. *Mathematics of Operations Research*, 15(1):155 – 174, 1990.

[41] R. S. Sutton. Learning to predict by methods of temporal differences. *Machine Learning*, 3:9–44, 1988.

[42] Sworder. A study of the relationship between identification and optimization in adaptive control problems. *Journal of The Franklin Institute*, 281:198–213, 1968.

[43] D. M. Titterington. Recursive parameter estimation using incomplete data. *J. Roy. Stat. Soc. B*, 46(2):257–267, 1984.

[44] John N. Tsitsiklis and Benjamin Van Roy. An analysis of temporal-difference learning with function approximation. Technical report, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, 1996. Report # LIDS-P-2322.

[45] C.J.C.H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8:279–292, 1992.

[46] Paul J. Werbos. *Handbook of intelligent control, neural, fuzzy and adaptive approaches*, chapter Chapter 3. Neurocontrol and supervised learning: an overview and evaluation, pages 65–91. Van Nostrand Reinhold, 1992.

[47] Paul J. Werbos. *Handbook of intelligent control, neural, fuzzy and adaptive approaches*, chapter Chapter 4. Approximate dynamic programming for real time control and neural modelling, pages 493–526. Van Nostrand Reinhold, 1992.

[48] Chelsea C. White. A Markov quality control process subject to partial observation. *Management Science*, 23(8):843–852, 1977.

[49] P. Whittle. Risk-sensitive linear/quadratic/gaussian control. *Adv. Appl. Prob.*, 13:764–777, 1981.

[50] Peter Whittle. A risk-sensitive maximum principle: the case of imperfect state observation. *IEEE Transactions On Automatic Control*, 36(7):793 – 801, 1991.

[51] Bernard Widrow, Narendra K Gupta, and Sidhartha Maitra. Punish/reward: Learning with a critic in adaptive threshold systems. *IEEE Transactions in Systems, Man and Cybernetics*, TSMC-3:455–465, 1973.

[52] Norbert Wiener. *Cybernetics*. Wiley, 1948.

[53] Ian H. Witten. The apparent conflict between estimation and control — a survey of the two-armed bandit problem. *Journal of The Franklin Institute*, 301(1&2):161 – 188, 1976.