Proceedings of the 36th
Conference on Decision & Control
San Diego, California USA • December 1997

Combined Estimation and Control of HMMs

Bernard Frankpitt[1] and John S. Baras[2]
Dept. of Electrical Engineering and Institute for Systems Research
University of Maryland, College Park, MD 20742

Abstract

The principal contribution of this paper is the presentation of the potential theoretical results that are needed for an application of stochastic approximation theory to the problem of demonstrating asymptotic stability for combined estimation and control of a plant described by a hidden Markov model. We motivate the results by briefly describing a combined estimation and control problem. We show how the problem translates to the stochastic approximation framework. We also show how the Markov chain that underlies the stochastic approximation problem can be decomposed into factors with discrete and continuous range. Finally, we use this decompostition to develop the results that are needed for an application of the ODE method to the stochastic control problem.

1 Introduction

Problems of combined estimation and control have a long history, and the LQG case is standard material for stochastic control texts. Treatment of controlled hidden Markov models is more recent, the work of Fernández-Gaucherand *et al.* [6, 7] treats a situation similar to that treated here with different methods. The methods that we use are based on existing work in stochastic approximation. In particular we use a recursive estimation scheme based on Krishnamurthy and Moore [8], and an approach from Benveniste *et al.* [3] to prove convergence of the estimation scheme. The difference between this work and the preceding work is that by considering randomized strategies we can show convergence of the model estimate and the control without recourse to special reset conditions that are required in [7].

2 The Plant, Controller and Estimator

The plant that we consider is a controlled hidden Markov model with N states, P input values, and M outputs. It has an unknown state transition matrix

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

and an unknown output matrix

$$B_{ij} = P(y_k = e_j \mid x_k = e_i). \tag{2}$$

Throughout the paper values of the state, input and output will be represented by canonical basis vectors in $\mathbb{R}^N$, $\mathbb{R}^P$ and $\mathbb{R}^M$, probability densities on finite spaces will be represented as vectors in the appropriate probability simplex, functions of finite sets will be represented as vectors in the appropriate Euclidean space, expectations will be denoted by inner products, and probability kernels on finite spaces will be represented as matrices.

We consider an output feedback control problem in which the feedback controller is recomputed from updated estimates of the plant. As the estimates of the plant converge, we wish the controller to converge to a moving horizon controller that minimizes a risk-sensitive cost functional.

$$\mathcal{J}^\gamma(\mu) = \mathbf{E}^\mu\left[\exp\frac{1}{\gamma}\left(\phi_f(x_K) + \sum_{l=0}^{K-1}\phi(x_l, u_l)\right)\right]. \tag{3}$$

in which the incremental cost functions ϕ and the final cost ϕ_f are assumed given. Baras and James [2] show that the optimal output policy is equivalent to a feedback policy, computed on am information-state process

$$\sigma_k = \Sigma(u_{k-1}, y_k)\sigma_{k-1}, \tag{4}$$

with linear recursion operator given by

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

The value function $V(\sigma, k)$, which is a function of time k, is computed by backwards dyanmic programming each time the parameter estimates are updated. For small state-spaces and reasonable horizons this is a feasible computation especially when advantage is taken of structural properties of the value function [5].

The random choice of control is governed by the "Gibbs Distribution"

$$\mu_{\sigma_k}(u) = \frac{e^{-H(\sigma_k,u)/\eta}}{Z(\sigma_k)}$$

[1] e-mail: frankpit@isr.umd.edu
[2] Lockheed Martin Chair in Systems Engineering.
e-mail: baras@isr.umd.edu

in which the energy

$$H(\sigma_k, u) = \mathbf{E}[V(\Sigma(u,y)\sigma_k, k+1) \mid \sigma_k]$$

is calculated using current estimates of the matrices A and B, and $Z(\sigma)$ is a "partition function" that normalizes the distribution. A randomized policy provides the regularity needed later for proof of convergence of the combined estimation and control problem. Continuing the statistical mechanics analogy, the parameter η is a temperature that controls how close the randomized policy is to the optimal (deterministic) policy. The current value of the information state σ_k is computed from the recursion (4) and a buffer of the last Δ inputs and outputs.

The estimator architecture is based on the work of Krishnamurthy and Moore [8]. A maximum likelihood estimator for a hidden Markov model with parameterization λ^0 minimizes the Kullback Leibler measure

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

Here $f(y_{0,k} \mid \lambda)$ is used to denote the distribution function induced by the parameter λ on the sequence of random variables $y_{0,k}$.

The central part of the estimator is a finite buffer containing the last Δ values of the input and output processes (the length is chosen to be the same as the length of the controller buffer in order to simplify the presentation). This buffer is used to update smoothed recursive estimates of the various densities required for the estimator. These densities are: $\alpha_k = f(x_{k-\Delta} \mid y_{0,k-\Delta})$, which is calculated with the recursion

$$\alpha_k(j) = \frac{\sum_i \langle e_j, \hat{B}y_{l-\Delta}\rangle \hat{A}_{u_{l-\Delta-1};ij}\alpha_{k-1}(i)}{\sum_j \sum_i \langle e_j, \hat{B}y_{l-\Delta}\rangle \hat{A}_{u_{l-\Delta-1};ij}\alpha_{k-1}(i)}, \quad (5)$$

$\beta_k = f(y_{k-\Delta+1,k} \mid x_{k-\Delta})$ which is computed with the backwards recursion

$$\beta_l(i) = \sum_j \beta_{l+1}(j)\hat{A}_{u_{l+1};ij}\langle e_i, \hat{B}y_{l-\Delta+1}\rangle, \quad (6)$$

$\zeta_k = f(x_{k-\Delta}, x_{k-\Delta-1} | y_{0,k})$ and $\gamma_k = f(x_{k-\Delta} \mid y_{0,k})$ which are given in terms of α and β by

$$\zeta_l(i,j) = \frac{\alpha_{l-\Delta-1}(i)A_{u_{l-\Delta-1};ij}\beta_{l-\Delta-1}(j)}{\sum_{i,j}\alpha_{l-\Delta-1}(i)A_{u_{l-\Delta-1};ij}\beta_{l-\Delta-1}(j)}$$

$$\gamma_l(i) = \frac{\sum_j \beta_{l-\Delta}(j)A_{u_{l-\Delta};ij}\alpha_{l-\Delta}(i)}{\sum_i\sum_j \beta_{l-\Delta}(j)A_{u_{l-\Delta-1};ij}\alpha_{l-\Delta}(i)}.$$

The empirical estimates of state frequency conditioned on output, and pair frequency conditioned on input are given by the recursions

$$Z_{l+1}(p,i,j) = (1 - q\delta_{u_{l-1}}(e_p))Z_l(u,i,j) + q\delta_{u_{l-1}}(e_p)\zeta_l(i,j) \quad (7)$$

and

$$\Gamma_{l+1} = (1 - q\delta_{y_l}(e_m))\Gamma_l + q\delta_{y_l}(e_m)\gamma_l(i) \quad (8)$$

Where e_p and e_m are canonical basis vectors, and q is a discount factor $0 < q < 1$.

3 The stochastic approximation problem

The result is a combined control and estimation algorithm that can be written in the form of a standard stochastic approximation problem [3]

$$\theta_{k+1} = \theta_k + \frac{1}{k}H(X_k, \theta_k) \quad (9)$$

where θ represents the estimates of the model parameters A and B, $X_k = \{x_k, u_{k-\Delta,k}, y_{k-\Delta,k}, \alpha_{k,k-1}, Z_k, \Gamma_k\}$ is a Markov chain with a transition kernel $\Pi_\theta(X; dX)$ which is a function of the parameter θ, and the parts of H that correspond to the updates of A_u and B are given by

$$\frac{\frac{\hat{A}^2_{u;ij}}{Z_l(u,i,j)}\left(\sum_{r=1}^N \frac{\hat{A}^2_{u;ir}}{Z_l(u,i,r))}\left(\frac{\zeta_{k|K,\Lambda_k}(i,j)}{\hat{A}_{u;ij}} - \frac{\zeta_{k|K,\Lambda_k}(i,r)}{\hat{A}_{u;ir}}\right)\right)}{\sum_{r=1}^N \frac{\hat{A}^2_{u;ir}}{Z_l(u,i,r)}} \quad (10)$$

and

$$\left(\frac{\hat{B}^2_{im}}{\Gamma_l(i,m)}/\sum_{r=1}^N \frac{\hat{B}^2_{ir}}{\Gamma_l(i,r)}\right)\left(\sum_{r=1}^N \frac{\hat{B}^2_{ir}}{\Gamma_l(i,r)} \times \left(\frac{\gamma_{k|K,\Lambda_k}(i)\delta_{y_k}(f_m)}{\hat{B}_{im}} - \frac{\gamma_{k|K,\Lambda_k}(i)\delta_{y_k}(f_r)}{\hat{B}_{ir}}\right)\right) \quad (11)$$

respectively.

Demonstrating asymptotic convergence for the recursion (9) in a neighborhood of the true parameter θ is equivalent to establishing asymptotic stability of the combined control and estimation architecture. Following Benveniste *et al.* [3], asymptotic convergence of (9) is shown by the ODE method [10] which views the iterates of equation (9) as approximations to points on stable trajectories of an appropriately chosen ODE. The generator for the ODE is the mean $h(\theta) = \int H(X,\theta)\, d\mu_\theta(X)$, where μ_θ is the invariant measure for the chain X_k. Benveniste *et al.* use regular solutions to the Poisson equation $(I - \Pi_\theta)\nu_\theta = H(X,\theta) - h(\theta)$ to bound the deviation of θ_k from the trajectories of $\dot{\hat{\theta}} = h(\hat{\theta})$. The remainder of this paper deals with the problem of establishing the existence of the invariant mesure μ, and the existence of regular solutions to the Poisson equation.

4 Characterizing the Markov chain

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. The Markov chain X that underlies the stochastic approximation problem has the following structure. $X_k = (X_k^x, X_k^u, X_k^y, X_k^\alpha, X_k^\varsigma, X_k^\gamma)$ where:

$X_k^x = x_k$ is the state of the controlled hidden Markov model defined by (1) and (2).

$X_k^u = u_{k-\Delta,k-1}$ is a buffer containing the last Δ values for the control u.

$X_k^y = y_{k-\Delta+1,k}$ is a buffer containing the last Δ values for the output (including the k'th value).

$X_k^\alpha = \alpha_{k-\Delta-1,k-\Delta}$ is buffer of length two containing the values for the 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 $\Omega^\alpha \times \Omega^\alpha$.

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

X_k^γ is the empirical density for the distibution 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.

In addition, let $\tilde{X}_k = (X_k^x, X_k^u, X_k^y)$ denote the product of the finite valued factors in X. 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_k^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.

Proposition 1 *The random process $\{X_k\}$ is Markov, the chain $\{\tilde{X}_k\}$ is a Markov sub-chain, and for all l, the random varaible 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 Section 2 to write explicit expressions for the Markov transistion kernels $\Pi(\tilde{X}; d\tilde{X})$ and $\Pi(X; dX)$.

From the formulae for the plant, and controller it is clear that when the value of the estimate θ is fixed, the distribution of the finite valued random variable $\tilde{X}_k$ depends only on the finite valued random variable $\tilde{X}_{k-1}$. Consequently the transition kernel $\Pi_\theta(\tilde{X}_a; d\tilde{X}_b)$ can be represented by a stochastic matrix

$$\begin{aligned} M_{\tilde{X}_a}^{\tilde{X}_b} &= \langle x_a, A_{u_b^{-1}} x_b \rangle \langle x_a, B y_b^0 \rangle \\ &\quad \times \langle \mu(y_a^0, u_a^{-1}, \ldots, y_a^{-\Delta+1}, u_a^{-\Delta}), u_b^{-1} \rangle \\ &\times \delta_{y_a^0}(y_b^{-1}) \delta_{u_a^{-1}}(u_b^{-2}) \ldots \delta_{y_a^{-\Delta+2}}(y_b^{-\Delta+1}) \delta_{u_a^{-\Delta+1}}(u_b^{-\Delta}) \end{aligned} \tag{12}$$

The evolution of the processes X_k^α, X_k^ς and X_k^γ captures the dyamics of the estimation algorithm presented in section 2. This dynamics is described by random sequences of transformations on the probability simplices that compose the ranges of X_k^α, X_k^ς and X_k^γ in a way that is analogous to the way that random walks are constructed. Write $X^\alpha = (X^{\alpha,1}, X^{\alpha,2})$, $X^\varsigma = (X^{\varsigma,1}, \ldots, X^{\varsigma,P})$ 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 $X^{\varsigma,p}$ takes values in the probability simplex Ω^ς. For $\Omega \in \{\Omega^\alpha, \Omega^\varsigma\}$ let $\mathfrak{S}(\Omega)$ denote the semigroup of (not necessarily invertible) affine transformations of Ω into itself. Equations (5), (7) and (8) determine maps maps $g^\alpha : \tilde{\mathcal{X}} \to \mathfrak{S}(\Omega^\alpha)$, $g^\varsigma : \tilde{\mathcal{X}} \times \mathcal{X}^\alpha \to \mathfrak{S}(\Omega^\varsigma)$ and $g^\gamma : \tilde{\mathcal{X}} \times \mathcal{X}^\alpha \to \mathfrak{S}(\Omega^\gamma)$ from the state $\tilde{X}$ to apropriate transformation semigroups. The evolution of the random processes X^α, X^ς and X^γ are described by the equations

$$\begin{aligned} X_{l+1}^{\alpha,1} &= g^\alpha(\tilde{X}_l) X_l^{\alpha,1} \qquad X_{l+1}^{\alpha,2} = X_l^{\alpha,1} \\ X_{l+1}^{\varsigma,p} &= g^\varsigma(\tilde{X}_l, X_{l+1}^\alpha) X_l^{\varsigma,p} \\ X_{l+1}^{\gamma,m} &= g^\gamma(\tilde{X}_l, X_{l+1}^\alpha) X_l^{\gamma,m}. \end{aligned}$$

and the proof of the claim in the proposition about sufficient statistics follows from the form of these equations.

The mappings g^α, g^ς and g^γ push forward measures defined on their domains to measures on the semigroups that comprise their ranges. Writing $d_\star g(\Pi(\tilde{X}_a; d\tilde{X}_b))$ for the push-forward of the transition kernel through g, the transition kernels for the continuous factors have the explicit formulae as convolutions.

$$\begin{aligned} &\Pi_\theta(X_a; dX_b^{\alpha,1}) = \delta_{X_a^{\alpha,1}} * d_\star g^\alpha(\Pi(\tilde{X}_a; d\tilde{X}_b)) \\ &\Pi_\theta(X_a; dX_b^{\alpha,2}) = \delta_{X_a^{\alpha,1}} \\ &\Pi_\theta(X_a; dX_b^{\varsigma,p}) = \delta_{X_a^{\varsigma,p}} \\ &\qquad * d_\star g^\varsigma(\Pi(\tilde{X}_a; d\tilde{X}_b) \otimes \Pi(X_a; dX_b^{\alpha,2})) \quad (13) \\ &\Pi_\theta(X_a; dX_b^{\gamma,m}) = \delta_{X_a^{\gamma,m}} \\ &\qquad * d_\star g^\varsigma(\Pi(\tilde{X}_a; d\tilde{X}_b) \otimes \Pi(X_a; dX_b^{\alpha,1})) \end{aligned}$$

■

5 Potential theory for the underlying chain

An ergodic theory of the chain X_k is established by considering first the discrete sub-chain $\tilde{X}$ using methods developed by Arapostathis and Marcus [1] and Le Gland and Mevel [9]. The 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.

Lemma 2 *Suppose that for some $\delta > 0$ the entries of the state transition matrix A, the output transition matrix B and the control policies μ_k evaluated on the singleton subsets of U satisfy*

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

then the transition matrix M associated with the chain $\{\tilde{X}_n\}$ is primitive with index of primitivity $\Delta + 1$, and the chain itself is irreducible and acyclic.

Proof: Define a directed graph with nodes labeled by the states $\tilde{X}$ and directed links defined by the positive elements of the transition matrix M. Primitivity of the kernel can then be established by showing that any two points on the graph are connected, and that the connection is composed of a bounded number of 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 kernels. [13, 12, 9]

Proposition 3 *Let $\Pi(\tilde{X}; d\tilde{X})$ be a primitive kernel on a finite discrete space $\tilde{X}$. The following are true:*

(i) The unique left eigenvector μ is the invariant measure for the kernel.

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

(iii) If $f : \tilde{X} \to \mathbb{R}$ is a bounded function on $\tilde{X}$ (in fact, a bounded vector in $\mathbb{R}^n$) then the Poisson equation

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

has a solution

$$\nu(\tilde{X}) = \sum_0^\infty \Pi^r(f - \mu(f)\mathbf{1})(\tilde{X})$$

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

Lemma 4 *The eigenvector corresponding to the Perron-Frobenius eigenvalue of a primitive stochastic matrix A is a continuous function of the matrix parameters.*

Proof: Let A be a primitive stochastic matrix, and v the corresponding left eigenvector. Let Δ be a matrix zero row sums that satisfies the condition that $A - \Delta$ is positive, and, for $\epsilon \in (0,1]$ let v_ϵ be the invariant measure corresponding to the matrix $A + \epsilon\Delta$. Let $u_\epsilon = (v_\epsilon - \langle v_\epsilon, v\rangle v)$, then $u_\epsilon \perp v$ and the Perron-Frobenius theorem gives

$$|u_\epsilon - u_\epsilon A|_2 > (1-\rho)|u_\epsilon|_2 \tag{14}$$

also, an expansion of the eigenvector equation for v_ϵ gives

$$u_\epsilon - u_\epsilon A = -\epsilon v_\epsilon \Delta$$

which, upon substitution into the right hand side of (14) gives

$$\begin{aligned}\epsilon^2(v_\epsilon\Delta/(1-\rho))^2 &> |u_\epsilon|_2^2 \\ &= (1 - \langle v_\epsilon, v\rangle)(1 + 2\langle v_\epsilon, v\rangle).\end{aligned}$$

Now, since v_ϵ and v both have positive entries, $0 < \langle v_\epsilon, v\rangle < 1$, and consequently

$$|v_\epsilon - v|^2 < 2\epsilon^2(|v_\epsilon\Delta|/(1-\rho))^2$$

■

Proposition 1 characterizes the transition kernel for the continuous part of the chain X_k. The explicit expressions in (13) have forms that are similar to the form of the kernel of a random walk on a semigroup $\mathfrak{S}$ which can be expressed as a convolution $\Pi(g; dg) = \delta_g * m$ where $x \in \mathfrak{S}$ and m is a measure on the semigroup $\mathfrak{S}$. In this case the measure m rather than being constant is a function of a Markov process $\tilde{X}$. 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. So the existence of a regular potential kernel for the random walk $\delta_x * \mu$ where the measure μ is the push-forward of the invariant measure of the chain $\tilde{X}$ is sufficient to ensure a potential theory for the processes in Proposition 1.

A metric ρ on the probability simplex Ω induces a metric on the semigroup of affine transformations $\mathfrak{S}(\Omega)$ as follows. For any two elements $g_1, g_2 \in \mathfrak{S}$ define the distance between g_1 and g_2 by

$$\rho(g_1, g_2) = \sup_{x\in\Omega} |g_1 x - g_2 x| \tag{15}$$

Define a continuous function $c : \mathfrak{S} \to \mathbb{R}$ by

$$c(g) = \sup_{x_1, x_2 \in \Omega} \frac{\rho(gx_1, gx_2)}{\rho(x_1, x_2)}. \tag{16}$$

The semigroup $\mathfrak{S}$ is contractive with respect to the metric ρ if the function c takes values in the interval $[0, 1]$. A random walk with measure m is strictly contractive when there exists a constant $c_0 < 1$ such that $c(g) \leq c_0$ for all g in the support of m. An element $X \in \mathfrak{S}$ has zero rank if the mapping of left multiplication by X on $\mathfrak{S}$ is a constant map.

The key to the problem of defining a suitable notion of recurrence lies in the selection of the right topology for the convergence of measures $\Pi^n(X; dX)$. An appropriate topology is the topology of weak convergence (of probability measures). When the underlying space (in this case the semigroup $\mathfrak{S}$) is Polish[1], the restriction of the weak topology to the probability simplex is metrizable, and, in fact, is the topology induced by the Lévy metric

$$d(\mu, \mu') = \min_{\delta} \Big\{ \mu(F) < \mu'(F^\delta) + \delta$$

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

Where F^δ denotes a δ neighborhood about the closed set F defined in the metric on $\mathfrak{S}$ [4, p. 65] The properties of the weak topology on the space of probability distributions is standard Probability Theory and can be found in Loève [11]

Proposition 5 *Let X_k be a contractive random walk on a semigroup $\mathfrak{S}$ with Markov transition kernel $\Pi(X; dY) = \delta_X * d\mu(Y)$ for some measure μ with finite support contained in $\mathfrak{S}$. 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 for X_k, and for any positive continuous function f with $m(f) > 0$,*

$$P_{X_0}\left[\sum_{k=1}^{\infty} f(X_k) = \infty\right] = 1$$

independent of the choice of X_0.

Proof: Given a point $x_0 \in \mathfrak{S}_0$ define a sequence of measures $m_{x_0,k} = \delta_{x_0}\Pi^k$. From the contractive property of the walk it follows that if x_0' and x_0 are distinct initial points, then the measure $m_{x_0,k}$ converges weakly to $m_{x_0'}$, and $d(m_{x_0,k}, m_{x_0,k+1}) \to 0$ as $k \to \infty$. Compactness of the underyling semigroup $\mathfrak{S}$ implies that the set of measures $m_{x_0,k}$ are tight, and therefore relatively compact in the space of probability measures with the topology of weak convergence and it follows that $m_{x_0,k}$ converges weakly to a unique invariant measure m with supported on the elements of $\mathfrak{S}$ with zero rank.

Suppose that $m(f) > 0$, then then there exists $\epsilon > 0$ and positive integer k_0 such that for all $k > k_0$, $\delta_{x_0}\Pi^k(f) \geq \epsilon$. Let M be an upper bound on f, then it follows that $P[f(x_k) < \epsilon/2] < \epsilon/2M$, and that for any positive integers $n_0 > p$

$$P_{x_0}[\sum_{n=1}^{\infty} f(x_n) < p\epsilon/2] < P_{x_0}[\sum_{n=1}^{k_0(n_0+p)} f(x_n) < n_0\epsilon/2]$$
$$< p\frac{(n_0+p)^p}{n_0!}(\epsilon/2M)^{n_0}$$

The bound on the right can be made arbitrarily small by fixing p and letting $n_0 \to \infty$. Choosing $x_0 = X_0$ the result

$$P_{X_0}\left[\sum_{k=1}^{\infty} f(X_k) = \infty\right] = 1$$

follows. ∎

Proposition 5 demonstrates a condition that is similar to Harris recurrence and provides the basis for a potential theory for random walks on semigroups that mirrors the analogous theory for Harris recurrent random walks on groups.

The following corollary is a summary of the consequences of Proposition 6.

Corollary 6 *X_k has the following properties:*

(i) X_k is m-irreducible.

(ii) X_k is aperiodic and finite

(iii) Let $h(\theta) = mH(X, \theta)$, then the Poisson equation has a solution

$$\nu_\theta = \lim_n \sum_0^n (\Pi^m H(X, \theta) - h(\theta))$$

For proofs of these facts see Revuz [12]. (iii) is a consequence of the random walk X_k satisfying the conditions for a normal chain.

The final two results establish regularity of the invariant measures with respect to variations in the parameter θ. The first is a technical lemma.

Lemma 7 *Let μ and μ' be two measures supported on S, a compact subset of a finite dimensional manifold,*

[1] A complete separable metric space

and separated in the Levy metric 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 $\mu_0(S), \mu'_0(S) < \delta$, $\mu_\alpha(S) = \mu'_\alpha(S)$ for all α, and if $g_1 \in \operatorname{supp} \mu_\alpha$ and $g_2 \in \operatorname{supp} \mu'_\alpha$ then $\rho(g_1, g_2) < 6\delta$. If $g_l \in \operatorname{supp} \mu_\alpha$ and $g'_l \in \operatorname{supp} \mu'_\alpha$.

Proof: The proof requires a decomposition of the set S into a number of groups of well-spaced, uniformly small subsets of S. Careful use of the definition of the Lévy metric permits the decomposition of μ and μ' into measures supported on the small subsets sets in the partition, and a 'surplus' measure with large support that is small in size. ■

Let μ_θ and $\mu_{\theta'}$ be two measures on $\mathfrak{S}$, and let Π_θ and $\Pi_{\theta'}$ be the transition kernels for the random walks generated by μ_θ and $\mu_{\theta'}$.

Proposition 8 *Let μ_θ and $\mu_{\theta'}$ be two measures on $\mathfrak{S}$ with associated transition kernels Π_θ and $\Pi_{\theta'}$ with invariant measures m_θ and $m_{\theta'}$ and let D be a bound on the diameter of the set Ω, then for any Lipschitz function $f : \Omega \to \mathbb{R}$ with Lipshitz constant L_f,*

$$|m_\theta(f) - m_{\theta'}(f)| \leq \delta(D + 6) L_f \frac{1}{1 - c}. \qquad (17)$$

Proof: Let $\mu = \mu_\theta$ and $\mu' = \mu_{\theta'}$, and let $m_{x_0,k}$ be the sequences of measures defined in the proof of Proposistion 5, and $m'_{x_0,k}$ be a second sequence of measures defined in the same way except using μ' in place of μ.

$$\begin{aligned} &|m_{\theta,k}(f) - m_{\theta',k}(f)| \\ &= \left| \sum_{l=1}^{k} \int f(g_k \dots g_1 x_0) \right. \\ &\qquad \Big(d\mu(g_k) \dots d\mu(g_l) d\mu'(g_{l-1}) \dots d\mu'(g_1) \\ &\qquad\qquad \left. - d\mu(g_k) \dots d\mu(g_{l+1}) d\mu'(g_l) \dots d\mu'(g_1) \Big) \right| \end{aligned}$$

The proof proceeds by using the lemma, and the contractive propery of the semigroup to give exponential bounds for the terms in the summation. Summing, and letting $k \to \infty$ then gives the result. ■

References

[1] 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.

[2] 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.

[3] 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.

[4] Jean-Dominique Deuschel and Daniel W. Stroock. *Large Deviations*, volume 137 of *Pure and applied mathematics.* Academic Press, 1989.

[5] E. Fernandéz-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.

[6] Emamanuel Fernández-Gaucherand, Aristotle Arapostathis, and Steven I. Marcus. A methodology for the adaptive control of Markov chains under partial state information. In *Procedings of the 31'st conference on decision and control*, pages 2750–2751, 1992.

[7] Emmanuel Fernandéz-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.

[8] 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.

[9] 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.

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

[11] M. Loève. *Probability Theory I.* Springer-Verlag, 1977.

[12] D. Revuz. *Markov Chains.* North-Holland, 1975.

[13] E. Seneta. *Non-negative matrices and Markov chains.* Springer-Verlag, New York, second edition, 1981.