RULES OF THUMB VERSUS DYNAMIC PROGRAMMING

Technical Appendix

Martin Lettau

Humboldt University,

Tilburg University, and CEPR and

Harald Uhlig

Tilburg University, and CEPR

Appendix B Technical Appendix. Additional calculations, propositions and proofs.

Appendix B.1 Rewriting the accounting scheme (13).

The following additional lemma is important for proving the main results.

LEMMA 1 *1. Let $k : \mathcal{S} \rightarrow \{1, \ldots, K\}$ be any mapping from states into rule indices. Find the implied decision function $h(s) = r_{k(s)}(s)$, the utilities $u(s, r_k(s))$ and the probabilities $\Pi_{h,i,j}$ for transiting from state s_i to state s_j. Find the unique, invariant distribution μ_h over states for the probabilities $\Pi_{h,i,j}$. Make a list of all vectors $Y = [s, k, s', k']$, such that the rule indices are consistent with the states, $k = k(s)$ and $k' = k(s')$. For this list, find the probabilities $\hat{\Pi}_{Y,\tilde{Y}}$ for transiting from some "old" vector Y (think: Y_t) to some "new" vector $\tilde{Y} = [\tilde{s}, \tilde{k}, \tilde{s}', \tilde{k}']$ (think: Y_{t+1}) as follows. The transition probability is zero, if the "old" state $\tilde{s}$ listed in the "new" vector $\tilde{Y}$ does not coincide with the "new" state s' in the "old" vector Y. Otherwise, $\hat{\Pi}$ coincides with the transition probability for going from the "new" state s' in the old vector Y to the "new" state $\tilde{s}'$ in the "new" vector $\tilde{Y}$. Formally,*

$$\hat{\Pi}_{Y,Y'} = \begin{cases} 0, & \text{if } \tilde{s} \neq s' \\ \Pi_{h,i,j}, & \text{if } \tilde{s} = s' = s_i \text{ and } \tilde{s}' = s_j \end{cases}$$

Then, there is a unique invariant distribution , over vectors Y for the transition probabilities $\hat{\Pi}$, given by

$$, ([s_i, k, s_j, k']) = \mu_h(s_i)\Pi_{h,i,j}$$

2. Define a function $\phi(\theta)$ via

$$\phi(\theta) = E_\Gamma\left[f(\theta, Y)\right] = \sum_Y , (Y) f(\theta, Y)$$

Let $\nu(k) = \sum_{s|k(s)=k} \mu_h(s)$ be the probability of choosing rule k. Let $\tilde{K} = \{k : \nu(k) \neq 0\}$. Let $u \in \mathbf{R}^{\tilde{K}}$ be a vector with entries

$$u_k = \frac{1}{\nu(k)} \sum_{\{s|k(s)=k\}} \mu_h(s) u(s, r_k(s))$$

and let $B \in \mathbf{R}^{\tilde{K},\tilde{K}}$ be a matrix with entries

$$B_{k,l} = \frac{\sum_{\{i|k=k(s_i)\}} \sum_{\{j|l=k(s_j)\}} \mu_h(s_i)\Pi_{h,i,j}}{\nu(k)}$$

For any vector θ, let $\tilde{\theta}$ be the vector of only the entries $\theta_k, k \in \tilde{K}$. Then, any vector θ with $\tilde{\theta} = u + \beta B\tilde{\theta}$ satisfies $\phi(\theta) = 0$ and conversely. Furthermore, $I - \beta B$ is invertible.

Proof

1. Since all $\Pi_{h,i,j}$ are strictly positive, , must be unique since any Y' can be reached from any other Y in two "steps" with positive probability. To show the invariance of the given distribution, calculate

$$\begin{aligned} , ([s_j, k', s_m, k'']) &= \sum_{Y|s'=s_i} , {}_Y \hat{\Pi}_{Y,\tilde{Y}} \\ &= \sum_{Y=[s_i,k,s_j,k']} \mu_h(s_i)\Pi_{h,i,j}\Pi_{h,j,m} \\ &= \mu_h(s_j)\Pi_{h,j,m}, \end{aligned}$$

which was to be shown.

2. Rewrite $\phi(\theta) = 0$ for entry $\bar{k}$ as

$$\nu(\bar{k})\theta_{\bar{k}} = \sum_{Y:k=\bar{k}} , {}_Y \left(u(s, r_{\bar{k}}(s) + \beta\theta_{k'}\right)$$

For $\nu(k) \neq 0$, continue

$$\begin{aligned} \nu(\bar{k})\theta_{\bar{k}} &= \sum_{i|k(s_i)=\bar{k}} \mu_h(s_i) \left(u(s_i, r_{\bar{k}}(s_i) + \beta \sum_j \Pi_{h,i,j}\theta_{k(s_j)} \right) \\ &= \nu(\bar{k})u_{\bar{k}} + \beta \sum_{k'} \nu(\bar{k}) B_{\bar{k},k'}\theta_{k'} \end{aligned}$$

The invertibility of $I - \beta B$ follows, since B is a Markov matrix.

•

With the aid of this lemma, we see that the algorithm stated in the main text for computing candidate asymptotic attractors is equivalent to calculating solutions θ_∞ to $\phi(\theta_\infty) = 0$ as defined in the lemma above. The proofs will use this equivalent reformulation of the algorithm. We first prove a special version of our general theorem, since that is rather straightforward to do.

Appendix B.2 A simple special case of Theorem 1.

PROPOSITION 4 *Let θ_∞ be an asymptotic attractor. In the special situation where convergence to θ_∞ is almost sure for any value of Y_{t_0}, and where the ordering of the entries in θ_t coincides with the ordering of the entries in θ_∞ almost surely for all $t \geq t_0$, θ_∞ is also a candidate asymptotic attractor.*

Proof of Proposition 4. *Take expectations with respect to the invariant distribution , over the initial state Y_{t_0} in equation (26) and sum from $t = t_0$ to some T. Since , is the invariant distribution, this amounts to taking expectations with respect to the invariant distribution over each future Y_t as well. Exploiting almost sure convergence yields*

$$\theta_\infty = \theta_{t_0} - \sum_{t=t_0}^{\infty} \gamma_{t+1}\phi(\theta_t). \tag{29}$$

Assume now that (17) does not hold and that instead, say, $\phi(\theta_\infty) > \epsilon > 0$. Since $\phi(\theta_t) \to \phi(\theta_\infty)$ and hence $\phi(\theta_t) > \epsilon/2$ for $t \geq T$, some T, a contradiction follows from (29), (12) and the finiteness of θ_∞. •

Appendix B.3 More Proofs

Proof of Proposition 2. *With equation (20), we get*

$$v_h(s_i) - x(s_i) = \beta \sum_j \Pi_{h,i,j} \left(v_h(s_j) - \theta_{k(s_j)} \right)$$

Taking expectations with respect to μ_h *and exploiting, that* $\mu_h' \Pi_h = \mu_h'$, *one finds*

$$E_{\mu_h}[v_h(s) - \theta_{k(s)}] = \beta \mu_h' \Pi_h \left(v_h(\cdot) - \theta_{k(\cdot)} \right) = \beta E_{\mu_h}[v_h(s) - \theta_{k(s)}],$$

where the middle term should be interpreted as matrix notation. Since $\beta \neq 1$, *the first equality obtains. For the second, calculate similarly*

$$E_{\mu_h}[v_h(s)] = \mu_h' u(\cdot, h(\cdot)) + \beta \mu_h' \Pi_h v_h(\cdot) = E_{\mu_h}[u(s, h(s))] + \beta E_{\mu_h}[v_h(s)]$$

where the middle term should be interpreted as matrix notation. •

Proof of Proposition 3. *Let* $\Delta_i \equiv v^*(s_i) - z_i$ *and note that* $\Delta_1 > 0$. *Comparing the equations for* $v^*(s_1)$ *and* z_1, *it follows that*

$$\Delta_1 = \frac{\beta \left(1 - \Pi_{h^*,1,1}\right)}{1 - \beta \Pi_{h^*,1,1}} \Delta_2. \tag{30}$$

Hence $\Delta_2 > \Delta_1 > 0$. *Since* $z_2 \geq z_1$, *we find* $v(s_2) > v(s_1)$. •

While we were not able to prove the conjecture 1, we were able to prove the following proposition, which needs an additional assumption: $\max_{s \in \mathcal{S}^{(2)}} v_h(s) \geq z_2$. On the other hand, the proposition is true for arbitrary policy functions h rather than just $\mathcal{H}^*$, the optimal one. If one uses this optimal policy function h^*, we know of examples where the assumption $\max_{s \in \mathcal{S}^{(2)}} v_{h^*}(s) \geq z_2$ is violated, but the conclusion nonetheless holds true for v^*: thus, conjecture 1.

PROPOSITION 5 *Suppose there are two rules. Let the first rule* r_1 *be active in all states and coincide with some decision function* h. *Let the second rule* r_2 *be active in only a* *strict subset $\mathcal{S}^{(2)}$ of all states. Suppose each rule is active, i.e. suppose that $z_2 > z_1$, where z_k is the asymptotic strength of classifier k. Suppose that $\max_{s \in \mathcal{S}^{(2)}} v_h(s) \geq z_2$. Then,*

$$\min_{i \in \mathcal{S}/\mathcal{S}^{(2)}} v_h(s_i) < \max_{i \in \mathcal{S}^{(2)}} v_h(s_i).$$

Proof of Proposition 5. *Let $\mathcal{S}^{(1)} = \mathcal{S}/\mathcal{S}^{(2)}$ be the set of states in which the first rule is active. Let $\tilde{h}$ be the decision function implied by the two rules: $\tilde{h}$ coincides with h on $\mathcal{S}^{(1)}$. Let $\mu_{\tilde{h}}$ be the stationary distribution across states, resulting from $\tilde{h}$. Subtract (20) from equation (18) for $i \in \mathcal{S}^{(1)}$ and sum the resulting equations, using $\mu_{\tilde{h}}(s_i)$ as weights. Slightly rewritten, we get*

$$\begin{aligned} & \sum_{i \in \mathcal{S}^{(1)}} \mu_{\tilde{h}}(s_i) \left(v(s_i) - x_i\right) - \beta \sum_{i \in \mathcal{S}^{(1)}} \sum_{j \in \mathcal{S}^{(1)}} \mu_{\tilde{h}}(s_i) \Pi_{\tilde{h},i,j} \left(v(s_j) - z_1\right) \\ & = \beta \sum_{i \in \mathcal{S}^{(1)}} \sum_{j \in \mathcal{S}^{(2)}} \mu_{\tilde{h}}(s_i) \Pi_{\tilde{h},i,j} \left(v(s_j) - z_2\right). \end{aligned} \tag{31}$$

Let

$$\psi = \sum_{j \in \mathcal{S}^{(1)}} \left(\mu_{\tilde{h}}(s_j) - \beta \sum_{i \in \mathcal{S}^{(1)}} \mu_{\tilde{h}}(s_i) \Pi_{\tilde{h},i,j} \right)$$

and note that $\psi > 0$, since $\mu_{\tilde{h}}(s_j) = \sum_i \mu_{\tilde{h}}(s_i) \Pi_{\tilde{h},i,j}$. For the left hand side (lhs) of equation (31), we get

$$\begin{aligned} lhs &= \sum_{j \in \mathcal{S}^{(1)}} \left(\mu_{\tilde{h}}(s_j) - \beta \sum_{i \in \mathcal{S}^{(1)}} \mu_{\tilde{h}}(s_i) \Pi_{\tilde{h},i,j} \right) \left(v(s_j) - z_1\right) \\ &> \psi \left(\min_{s \in \mathcal{S}^{(1)}} v(s) - z_1 \right). \end{aligned}$$

For the right hand side (rhs) of equation (31), we get with (21)

$$\begin{aligned} rhs &\leq \beta \left(\sum_{i \in \mathcal{S}^{(1)}} \sum_{j \in \mathcal{S}^{(2)}} \mu_{\tilde{h}}(s_i) \Pi_{\tilde{h},i,j} \right) \left(\max_{s \in \mathcal{S}^{(2)}} v(s) - z_2 \right) \\ &= \left(\sum_{j \in \mathcal{S}^{(1)}} \beta \mu_{\tilde{h}}(s_j) - \beta \sum_{i \in \mathcal{S}^{(1)}} \sum_{j \in \mathcal{S}^{(1)}} \mu_{\tilde{h}}(s_i) \Pi_{\tilde{h},i,j} \right) \left(\max_{s \in \mathcal{S}^{(2)}} v(s) - z_2 \right) \end{aligned}$$

$$\leq \quad \psi\left(\max_{s\in\mathcal{S}^{(2)}} v(s) - z_2\right),$$

where we relabelled the summation index from i to j in the first sum of the second line. Comparing, we get

$$\min_{s\in\mathcal{S}^{(1)}} v(s) \leq \max_{s\in\mathcal{S}^{(2)}} v(s) + (z_1 - z_2) < \max_{s\in\mathcal{S}^{(2)}} v(s)$$

as claimed. •

We suspect that one of the key difficulties in proving conjecture 1 lies in comparing strengths and values. The following conjecture isolates that difficulty and is implied by conjecture 1. It has remained resilient to a proof so far, although extensive numerical experiments did not result in a counterexample.

CONJECTURE 2 *Let h be some decision function and let $(\mathcal{S}^{(1)}, \mathcal{S}^{(2)})$ be a partition of the state space $\mathcal{S}$. Let there be two rules with the first rule r_1 active only on $\mathcal{S}^1$ and the second rule active only on $\mathcal{S}^2$, and where both rules coincide with the decision function h, when they are active. Suppose that the asymptotic strength of the second rule is higher than the asymptotic rule of the first rule, $z_2 > z_1$. Then,*

$$\min_{s\in\mathcal{S}^{(1)}} v(s) < \max_{s\in\mathcal{S}^{(2)}} v(s).$$

Appendix B.4 Calculations and Propositions for section 7

We now proceed to formal statements and proofs of the claims in the illustration section: most of it is very simple.

PROPOSITION 6 *Suppose there is only one rule $r \in \mathcal{R}$. Then there is a unique candidate asymptotic attractor $\theta_\infty \in \mathbf{R}$ and it satisfies $\theta_\infty = E_{\mu_r} v_r$. In particular, if $r = h^*$, then $\theta_\infty = E_{\mu_{h^*}} v^*$.*

Proof of Proposition 6. *In this case (17) reduces to*

$$\theta_\infty = (I - \beta)^{-1} E_{\mu_h}[u_h].$$

We therefore need to show that

$$E_{\mu_h}[v_h] = (1 - \beta)^{-1} E_{\mu_h}[u_h]$$

or

$$(1 - \beta)\mu^T v_h = \mu_h u_h.$$

But this follows immediately from (10) and from the fact that

$$\mu^T = \mu^T \Pi_h,$$

which completes the proof. •

PROPOSITION 7 *Let h^* be a decision function with $v^* = v_{h*}$ and suppose that h^* is unique. Suppose, furthermore, that all K rules are applicable in at most one state and that for each $s \in \mathcal{S}$, there is exactly one rule with $r(s) = h^*(s)$; denote its index with $k^*(s)$. Define θ_∞ by assigning for each state s strength $v^*(s)$ to the classifier with index $k^*(s)$. For all other classifiers c applicable in some state s, assign some strength strictly below $v^*(s)$. Then θ_∞ is a candidate asymptotic attractor which implements the dynamic programming solution.*

Proof of Proposition 7. *Compare equation (17) to equation (10) and note that* $B = \Pi_h$ and $\vec{u} = u_h$. •

PROPOSITION 8 *If all rules are applicable in all states (i.e. all rules are total), then for each rule, there is an asymptotic attractor θ_∞, where that rule is the only asymptotically active rule.*

Proof of Proposition 8. *To find θ_∞, let h the decision function from that chosen rule, assign asymptotic strengths $E_h[v_h]$ to that rule and assign strictly lower strengths to all other rules. (17).* •

Here is the promised example, where the union of two rules can no longer be active. Let $\beta = 0.999$, let there be $n = 4$ states, $K = 4$ rules, and $m = 2$ actions. The utilities are

$$\begin{array}{lllll} & s = s_1 & s = s_2 & s = s_3 & s = s_4 \\ a = a_1: & u(s_1, a_1) = 10, & u(s_2, a_1) = 3.8, & u(s_3, a_1) = 5.9, & u(s_4, a_1) = 0.12 \\ a = a_2: & u(s_1, a_2) = 3.08, & u(s_2, a_2) = 3.05, & u(s_3, a_2) = -1, & u(s_4, a_1) = -1. \end{array}$$

Given the first action, the transition probabilities $\pi_{s,a_1}(s')$ are :

$$\begin{array}{lllll} & s' = s_1 & s' = s_2 & s' = s_3 & s' = s_4 \\ s = s_1: & \pi_{s_1,a_1}(s_1) = 0.001, & \pi_{s_1,a_1}(s_2) = 0.997, & \pi_{s_1,a_1}(s_3) = 0.001, & \pi_{s_1,a_1}(s_4) = 0.001, \\ s = s_2: & \pi_{s_2,a_1}(s_1) = 0.001, & \pi_{s_2,a_1}(s_2) = 0.001, & \pi_{s_2,a_1}(s_3) = 0.997, & \pi_{s_2,a_1}(s_4) = 0.001, \\ s = s_3: & \pi_{s_3,a_1}(s_1) = 0.001, & \pi_{s_3,a_1}(s_2) = 0.997, & \pi_{s_3,a_1}(s_3) = 0.001, & \pi_{s_3,a_1}(s_4) = 0.001, \\ s = s_4: & \pi_{s_4,a_1}(s_1) = 0.997, & \pi_{s_4,a_1}(s_2) = 0.001, & \pi_{s_4,a_1}(s_3) = 0.001, & \pi_{s_4,a_1}(s_4) = 0.001. \end{array}$$

Given the second action, the transition probabilities $\pi_{s,a_2}(s')$ are :

$$\begin{array}{lllll} & s' = s_1 & s' = s_2 & s' = s_3 & s' = s_4 \\ s = s_1: & \pi_{s_1,a_2}(s_1) = 0.001, & \pi_{s_1,a_2}(s_2) = 0.001, & \pi_{s_1,a_2}(s_3) = 0.997, & \pi_{s_1,a_2}(s_4) = 0.001, \\ s = s_2: & \pi_{s_2,a_2}(s_1) = 0.001, & \pi_{s_2,a_2}(s_2) = 0.001, & \pi_{s_2,a_2}(s_3) = 0.001, & \pi_{s_2,a_2}(s_4) = 0.997, \\ s = s_3: & \pi_{s_3,a_2}(s_1) = 0.997, & \pi_{s_3,a_2}(s_2) = 0.001, & \pi_{s_3,a_2}(s_3) = 0.001, & \pi_{s_3,a_2}(s_4) = 0.001, \\ s = s_4: & \pi_{s_4,a_2}(s_1) = 0.001, & \pi_{s_4,a_2}(s_2) = 0.997, & \pi_{s_4,a_2}(s_3) = 0.001, & \pi_{s_4,a_2}(s_4) = 0.001. \end{array}$$

The rules are initially

$$\begin{array}{lllll} & s = s_1 & s = s_2 & s = s_3 & s = s_4 \\ r_1: & r_1(s_1) = 1, & r_1(s_2) = 0, & r_1(s_3) = 1, & r_1(s_4) = 0 \\ r_2: & r_2(s_1) = 0, & r_2(s_2) = 1, & r_2(s_3) = 0, & r_2(s_4) = 1 \\ r_3: & r_3(s_1) = 2, & r_3(s_2) = 0, & r_3(s_3) = 0, & r_3(s_4) = 0 \\ r_4: & r_4(s_1) = 0, & r_4(s_2) = 2, & r_4(s_3) = 0, & r_4(s_4) = 0. \end{array}$$

One can calculate, that there is an asymptotic attractor with the ordering $(3,1,4,2)$ for the rules. The strength vector is given by $(3038.9; 3036.0; 3039.0; 3036.1)$. Imagine now, merging rules r_3 and r_4 into a single rule $\tilde{r}_3$, given by

$$\begin{array}{lllll} & s=s_1 & s=s_2 & s=s_3 & s=s_4 \\ \tilde{r}_3: & \tilde{r}_3(s_1)=2, & \tilde{r}_3(s_2)=2, & \tilde{r}_3(s_3)=0, & \tilde{r}_3(s_4)=0. \end{array}$$

In other words, now there are only the three rules r_1, r_2 and $\tilde{r}_3$. Then, there is no asymptotic attractor in which rule $\tilde{r}_3$ is asymptotically active:

1. The ordering $(3,1,2)$ results in strengths $(3040.4; 3034.6; 3037.5)$,
2. The ordering $(1,3,2)$ results in strengths $(4392.5; 4388.2; 4386.9)$,
3. The ordering $(2,3,1)$ results in strengths $(4843.3; 4842.3; 4841.5)$.

In all three cases, the consistency condition is violated.

For the example in subsection 7.2, here are the calculations for cases I and II:

Case I: $z_1 > z_2$**:** In this case, classifier 1 is activated in states $s=1$ and $s=3$ and classifier 2 is activated in state $s=2$. Thus, action $a=1$ is taken in all three states: $h(s)\equiv 1$. It follows that $\mu_h(1)=\mu_h(2)=\mu_h(3)=1/3$. To find the strengths solve the equations

$$\begin{aligned} z_1 &= u(1,1)+\frac{\beta}{3}(2z_1+z_2) \\ z_2 &= \frac{\beta}{3}(2z_1+z_2). \end{aligned}$$

This case can thus be obtained if and only if

$$u(1,1)>0. \tag{32}$$

Case II: $z_2 > z_1$**:** The decision function in this case is $h(1)=2$, $h(2)=1$, $h(3)=1$. Consequently $\mu_h(1)=\mu_h(3)=1/4$ and $\mu_h(2)=1/2$. The strengths are calculated by solving

$$\begin{aligned} z_1 &= u(1,1)+\frac{\beta}{3}(z_1+2z_2) \\ z_2 &= \frac{1}{3}(u(1,2)+\beta z_2)+\frac{2\beta}{9}(z_1+2z_2). \end{aligned}$$

It is easily checked that this case can be obtained if and only if

$$u(1,2) > 3u(1,1). \tag{33}$$

Here are the calculations to arrive at the inequality (25) in section 7.3. Similar to case I above, one needs to solve the equations

$$\begin{aligned} \kappa_1 z_1 &= u(1,1) + \frac{\beta}{3}(2\kappa_1 z_1 + z_2) \\ z_2 &= \frac{\beta}{3}(2\kappa_1 z_1 + z_2). \end{aligned}$$

One gets $\kappa z_1 = u(1,1) + z_2$, which can be used in the second equation to calculate z_2,

$$z_2 = \frac{2\beta}{3(1-\beta)} u(1,1)$$

and thus

$$z_1 - z_2 = \frac{\frac{3-\beta}{\kappa_1} - 2\beta}{3(1-\beta)} u(1,1)$$

Since $\beta < 1$ and $u(1,1) < 0$, we get $z_1 - z_2 > 0$ if and only if $\frac{3-\beta}{\kappa_1} - 2\beta > 0$ and thus equation (25).

Appendix B.5 A Theorem about Markov Stochastic Approximation Algorithms

In this section, we use the notation of Métivier and Priouret (1984). For a general overview and introduction to stochastic approximation algorithms, see Sargent (1992) and Ljung, Pflug and Walk (1992).

For each $\theta \in \mathbf{R}^d$ consider a transition probability $\hat{\Pi}_\theta(y; dx)$ on $\mathbf{R}^k$. This transition probability defines a controlled Markov chain on $\mathbf{R}^d$.

Define a stochastic algorithm by the following equations:[13]

$$\theta_{n+1} = \theta_n - \gamma_{n+1} f(\theta_n, Y_{n+1}) \tag{34}$$

[13] The algorithm here is subscripted with n rather than t.

where $f(\theta, y)$ is a function $f : \mathbf{R}^d \times \mathbf{R}^k \to \mathbf{R}^d$,

$$P\left[Y_{n+1} \in B \mid \mathcal{F}_n\right] = \hat{\Pi}_{\theta_n}(Y_n; B) \tag{35}$$

where $P\left[Y_{n+1} \in B \mid \mathcal{F}_n\right]$ is the conditional probability of the event $\{Y_{n+1} \in B\}$ given $\theta_0, \ldots, \theta_n$, $Y_0, \ldots, Y_n$.

We call $\psi \to \hat{\Pi}_\theta \psi$ the operator $\hat{\Pi}_\theta \psi(x) \equiv \int \psi(y) \hat{\Pi}_\theta(x; dy)$. Assume the following:

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

$$\sup_{|\theta| \le R} \sup_x \mid f(\theta, x) \mid \le M_R.$$

(M1) For every θ, the Markov chain $\hat{\Pi}_\theta$ has a unique invariant probability $\,_\theta$.

(M2) There exist $p \geq 2$ and positive constants $\alpha_R < 1$, K_R for which $\sup_{|\theta| \le R} \int \mid y \mid^p \hat{\Pi}_\theta(x; dy) \le \alpha_R \mid x \mid^p + K_R$.

(M3) For every function[14] v with the property $\mid v(x) \mid \le K(1 + \mid x \mid)$ and every θ, θ', $\mid \theta \mid \le R$, $\mid \theta' \mid \le R$,

$$\sup_x \mid \hat{\Pi}_\theta v(x) - \hat{\Pi}_{\theta'} v(x) \mid \le \tilde{K}_R \mid \theta - \theta' \mid \sup_{x \ne x'} \frac{\mid v(x) - v(x') \mid}{\mid x - x' \mid}.$$

(M4) For every θ the Poisson equation

$$(1 - \hat{\Pi}_\theta) v_\theta = f(\theta, \cdot) - \int f(\theta, y) \,_\theta(dy) \tag{36}$$

has a solution v_θ with the following properties of (M5).

(M5) For all R there exist constants M_R and C_R so that

a) $\sup_{|\theta| \le R} \mid v_\theta(x) - v_\theta(x') \mid \le M_R \mid x - x' \mid$,

b) $\sup_{|\theta| \le R} \mid v_\theta(x) \mid \le C_R(1 + \mid x \mid)$,

c) $\mid v_\theta(x) - v_{\theta'}(x) \mid \le C_R \mid \theta - \theta' \mid (1 + \mid x \mid)$ for $\mid \theta \mid \le R$, $\mid \theta' \mid \le R$.

[14] The functions v here and in the next two assumptions have no (or at least no apparent) connection with the value functions in the main body of the paper.

Let

$$\phi(\theta) \equiv \int f(\theta, y) , _\theta(dy) = E_{\Gamma_\theta}[f(\theta, y)].$$

Métivier and Priouret (1984) have shown the following theorem.

THEOREM 2 *Consider the algorithm defined by (34) and (35) and assume that (F) and (M1) through (M5) are satisfied. Suppose that* (γ_n) *is decreasing with* $\sum_n \gamma_n = +\infty$ *and* $\sum_n \gamma_n^{1+(p/2)} < \infty$*, where* $p \geq 2$ *is the constant entering (M2). Let* $\Omega_1 \equiv \{\sup_n \mid \theta_n \mid < \infty\}$*. Then there is a set* $\tilde{\Omega}_1 \subset \Omega_1$ *such that* $P(\Omega_1 \backslash \tilde{\Omega}_1) = 0$ *and with the following property: for every* θ^* *that is a locally asymptotically stable point of the equation*

$$\frac{d\theta(t)}{dt} = -\phi(\theta(t))$$

with domain of attraction $D(\theta^*)$ *and for every* $\omega \in \tilde{\Omega}_1$ *such that for some compact* $A \subset D(\theta^*)$*,* $\theta_n(\omega) \in A$ *for infinitely many n, the following holds:*

$$\lim_n \theta_n(\omega) = \theta^*$$

Remarks:

1. Suppose Y is always a member of some finite set $\{y_1, \ldots, y_q\}$ and assume that (M1) is satisfied. The operator $\hat{\Pi}_\theta$ can then simply be understood as a matrix operating on $\mathbf{R}^q$ via $\hat{\Pi}_\theta v_i = \sum_j (\hat{\Pi}_\theta)_{ij} v_j$, where $v_i \equiv v(y_i)$ for any given function $v : \mathbf{R}^k \rightarrow \mathbf{R}$. In particular, the q-dimensional vector corresponding to the function v_θ in (M4) can always be found by inverting the matrix $(I - \hat{\Pi}_\theta)$ on its range and applying it to the q-dimensional vector corresponding to the right-hand side of equation (36), noting that the right-hand side of that equation is indeed in the range of $I - \hat{\Pi}_\theta$, since it is orthogonal to the q-dimensional vector representing the unique invariant probability $, _\theta$.

2. Suppose θ is always in some compact subset of $\mathbf{R}^d$ and Y is always a member of some discrete, finite subset of $\mathbf{R}^k$. Then assumptions (M2), (M4) and (M5a) and (M5b) are trivially satisfied and p in (M2) can be chosen to be $p = \infty$.

3. Suppose, $\hat{\Pi}_\theta$ is independent of θ. Then assumption (M3) is trivially satisfied.