Multimode Value Iterations

Vishal Raman, Sunay Joshi

July 3, 2026

(Question marks are left at places where the argument doesn't seem rigorous enough. Also, some of the remarks are actually questions.)

In this write-up, we consider Multimode Value Iterations (MMVI) for PDMPs. Briefly, in each iteration of MMVI, we relax modes in the order $1, 2, \ldots, m$. As opposed to vanilla Value Iterations (VI), the MMVI algorithm is asynchronous. We claim that MMVI yields geometric convergence to the true value functions. Further, we derive stronger convergence rates for almost-causal PDMPs.

Remark 0.1. *Using Proposition 5.5, can we claim that MMVI is faster than vanilla VI? (?)*

1 Notation

Consider a PDMP in the domain $\Omega \subset \mathbb{R}^2$ with modes $1, \ldots, m$. The control at each point x is the direction of motion $a \in \mathbb{S}^1$, the speed of motion at x in mode i is $f_i(x)$, the running cost in mode i is $K_i(x)$, and the process terminates when the agent reaches point in the exit set Q and incurs a terminal penalty of $q(x) \geqslant 0$. We assume that q is independent of mode (?). Let $\lambda_{ij} : \Omega \to \mathbb{R}$ denote the transition rate from mode i to mode j for all $1 \leqslant i \neq j \leqslant m$.

We begin by working in full generality, making no assumptions on the rates λ_{ij} or the specific structure of the PDMP.

Given functions $v_i : \Omega \to \mathbb{R}$ for $1 \leqslant i \leqslant m$, we can define the m-tuple $v = (v_1, \ldots, v_m)$. If each v_i is an estimate of the value function in mode i, then we can think of v as the entire estimated value function across all modes. We sometimes refer to v as a "cost vector" below. Given $v = (v_1, \ldots, v_m)$ and $w = (w_1, \ldots, w_m)$, we say that $v \leqslant w$ if $v_i(x) \leqslant w_i(x)$ for all i, x. We denote the true value function for the PDMP by $u = (u_1, \ldots, u_m)$.

Given $v_i : \Omega \to \mathbb{R}$, we use the sup-norm $\|v_i\|_\infty = \sup_{x \in \Omega} |v_i(x)|$. Given $v = (v_1, \ldots, v_m)$, we define $\|v\|_\infty = \max_{i \in [m]} \|v_i\|_\infty$.

Similarly, given functions $a_i : \Omega \to \mathbb{S}^1$ for $1 \leqslant i \leqslant m$, we write $a = (a_1, \ldots, a_m)$. If each a_i is a policy in mode i, then we can think of a as prescribing a policy for the entire PDMP across all modes.

Let $\mathcal{A}$ denote the set of all policies a. Note that by Tychonoff's Theorem $\mathcal{A}$ is compact, since $\mathcal{A} = \prod_{x\in\Omega} \mathbb{S}^1$ is a product of compact sets.

Given a policy a_i in mode i, define μ_i as $\mu_i(s) = \int_s^{t_i^*} \sum_{j\neq i} \lambda_{ij}(y_i(r))dr$, where the path y_i corresponds to the policy a_i via $\dot{y}_i(r) = f_i(y_i(r))a_i(y_i(r))$, and $t_i^* = \inf\{t|y_i(t) \in Q\}$.

Let t_i^M be an upper bound on the travel time to the exit in mode i, assuming that there are no switches. Let $\lambda_i^M = \max_{x\in\Omega} \sum_{j\neq i} \lambda_{ij}(x)$ be an upper bound on the rate of switching out of mode i.

Remark 1.1. *Do we need to restrict $\mathcal{A}$ in any way? I think we need to consider $a = (a_1, \ldots, a_m)$ with smooth functions $a_i : \Omega \to \mathbb{S}^1$ in order to utilize ODE results below. Also, we need assumptions on $\mathcal{A}$ to ensure the argmins are unique (?).*

2 Operators

To facilitate our proofs, we introduce operator notation. Given an operator H and a cost vector $v = (v_1, \ldots, v_m)$, let $Hv = ((Hv)_1, \ldots, (Hv)_m)$. We define three operators T_a, T, F that are somewhat analogous to various versions of the discrete Bellman operator.

Definition 2.1. *Given a policy $a = (a_1, \ldots, a_m)$, the operator T_a is defined as*

$$(T_a v)_i(x) = e^{-\mu_i(t)} \int_t^{t_i^*} e^{\mu_i(s)} \left[K_i + \sum_{j\neq i} \lambda_{ij} v_j \right] ds + e^{-\mu_i(t)} q(y_i(t_i^*)),$$

where the path y_i corresponds to the policy a_i via $\dot{y}_i = f_i a_i$ and $y_i(t) = x$. We let $T_a v$ denote the entire updated cost vector.

Remark 2.2. *One can show that $w_i(x) = (T_a v)_i(x)$ is the solution to the PDE*

$$f_i(x)\nabla w_i(x) \cdot a_i(x) + K_i(x) + \sum_{j\neq i} \lambda_{ij}(x)(v_j(x) - w_i(x)) = 0$$

In other words, $(T_a v)_i$ is the mode i cost that one obtains when evaluating the policy a using estimated value functions v_j for all $j \neq i$.

Definition 2.3. *For all $x \in \Omega$ and for all i, let $(Tv)_i(x) = \min_{a\in\mathcal{A}} (T_a v)_i(x)$. We let Tv denote the entire updated cost vector.*

Proposition 2.4. *For all $a \in \mathcal{A}$ and v, we have $Tv \leqslant T_a v$.*

Proof. This is clear from the definition of Tv. □

Proposition 2.5. *There exists $a^* \in \mathcal{A}$ such that $(Tv)_i(x) = (T_{a^*} v)_i(x)$ for all i, x.*

Proof. Given i, x, define the policy $a^{i,x} = \operatorname{argmin}_{a\in\mathcal{A}} (T_a v)_i(x)$. Given i, x, let $(a^*)_i(x) = (a^{i,x})_i(x)$.

Let y^1 denote the trajectory in mode i defined by $a^{i,x}$, with $y^1(t) = x$. Let y^2 denote the trajectory in mode i defined by a^*, with $y^2(t) = x$.

Note that for all $s \geqslant t$, we have $a^{i,x}(y^1(s)) = a^*(y^1(s))$. Since $y^1(t) = y^2(t) = x$, the uniqueness of trajectories in a vector field implies that $y^1(s) = y^2(s)$ for all $s \geqslant t$. Hence $(T_{a^*}v)_i(x) = (T_{a^{i,x}}v)_i(x) = (Tv)_i(x)$, as desired (?). □

Remark 2.6. *a^* can be thought of as a greedy one-step optimal policy.*

Remark 2.7. *One can show that $w_i(x) = (Tv)_i(x)$ is the solution to the PDE*

$$-f_i(x)|\nabla w_i(x)| + K_i(x) + \sum_{j \neq i} \lambda_{ij}(x)(v_j(x) - w_i(x)) = 0$$

In other words, $(Tv)_i$ is the mode i cost that one obtains when choosing the optimal policy a^* using estimated value functions v_j for all $j \neq i$ (?).

We now define the MMVI operator F.

Definition 2.8. *The operator F is defined by the following procedure: begin by setting $(Fv)_1 = (Tv)_1$. Then, process $i \in \{2, \ldots, m\}$ sequentially, setting $(Fv)_i = (T\widehat{v})_i$, where $\widehat{v} = ((Fv)_1, \ldots, (Fv)_{i-1}, v_i, \ldots, v_m)$. The result is the updated estimate $Fv = ((Fv)_1, \ldots, (Fv)_m)$.*

Remark 2.9. *F can be thought of as applying Gauss–Seidel relaxations on the modes.*

Remark 2.10. *The operators T_a, T, F are not really Bellman operators. From the remarks above, we see that they are really "PDE operators" that correspond to applying Gauss-Jacobi or Gauss-Seidel updates to a system of m PDEs.*

3 Algorithms

Below, we summarize the vanilla VI algorithm and the MMVI algorithm.

Vanilla VI consists of the following steps:

1. Initialization: Fix u^0 and $\varepsilon > 0$.
2. Value Iterations:

   If $\|Tu^n - u^n\| \geqslant \varepsilon$:

   - Set $u^{n+1} = Tu^n$ and return to Step 2.

   Else:

   - Return u^n.

The MMVI Algorithm is identical to vanilla VI, except for the fact that it repeatedly applies the operator F, not the operator T.

4 Monotonicity

We show that T_a is monotone. By Proposition 2.5, this implies that T is monotone.

Definition 4.1. *Given an operator H, we say that H is monotone if $v \leqslant w$ implies $Hv \leqslant Hw$.*

Proposition 4.2. *T_a is monotone.*

Proof. Note that

$$(T_a v)_i(x) = e^{-\mu_i(t)} \int_t^{t_i^*} e^{\mu_i(s)} \left[K_i + \sum_{j \neq i} \lambda_{ij} v_j \right] ds + e^{-\mu_i(t)} q(y_i(t_i^*)) \tag{1}$$

$$\leqslant e^{-\mu_i(t)} \int_t^{t_i^*} e^{\mu_i(s)} \left[K_i + \sum_{j \neq i} \lambda_{ij} w_j \right] ds + e^{-\mu_i(t)} q(y_i(t_i^*)) = (T_a w)_i(x), \tag{2}$$

as desired. □

Corollary 4.3. *T is monotone.*

5 Contractivity

We now prove contractivity of all operators in the general case. For all i, let $C_i = 1 - \exp(-t_i^M \lambda_i^M) < 1$.

Proposition 5.1. *T_a is a contraction with Lipschitz constant $\max_i C_i$.*

Proof. Fix i and x. Assume WLOG that $(T_a v)_i(x) - (T_a \tilde{v})_i(x) > 0$. Then

$$(T_a v)_i(x) - (T_a \tilde{v})_i(x) = e^{-\mu_i(t)} \int_t^{t_i^*} e^{\mu_i(s)} \cdot \sum_{j \neq i} \lambda_{ij} (v_j - \tilde{v}_j) ds \tag{3}$$

$$\leqslant e^{-\mu_i(t)} \int_t^{t_i^*} e^{\mu_i(s)} \cdot \sum_{j \neq i} \lambda_{ij} |v_j - \tilde{v}_j| ds \tag{4}$$

$$\leqslant \max_{j \neq i} \|v_j - \tilde{v}_j\|_\infty \cdot e^{-\mu_i(t)} \int_t^{t_i^*} e^{\mu_i(s)} \cdot \sum_{j \neq i} \lambda_{ij} ds \tag{5}$$

$$= \max_{j \neq i} \|v_j - \tilde{v}_j\|_\infty \cdot (1 - \exp(-\mu_i(t))) \tag{6}$$

$$\leqslant C_i \max_{j \neq i} \|v_j - \tilde{v}_j\|_\infty \tag{7}$$

$$\leqslant C_i \|v - \tilde{v}\|_\infty , \tag{8}$$

where $\|v - \tilde{v}\|_\infty$ is simply notation for $\max_j \|v_j - \tilde{v}_j\|_\infty$. Taking the max over x yields

$$\|(T_a v)_i - (T_a \tilde{v})_i\|_\infty \leqslant C_i \max_{j \neq i} \|v_j - \tilde{v}_j\|_\infty ,$$

where we think of the LHS as a distance between mode i value function estimates. This implies that

$$\|(T_a v)_i - (T_a \tilde{v})_i\|_\infty \leqslant C_i \|v - \tilde{v}\|_\infty$$

(We will use both of these intermediate bounds below.)

Finally, taking the max over i yields

$$\|T_a v - T_a \tilde{v}\|_\infty \leqslant \max_i C_i \cdot \|v - \tilde{v}\|_\infty ,$$

as desired. □

Proposition 5.2. *T is a contraction with Lipschitz constant $\max_i C_i$.*

Proof. Fix i and x. Assume WLOG that $(Tv)_i(x) - (T\tilde{v})_i(x) > 0$. Let $a^*, \tilde{a}^*$ satisfy $Tv = T_{a^*} v$ and $T\tilde{v} = T_{\tilde{a}^*} \tilde{v}$. Then by optimality,

$$\begin{aligned}
(Tv)_i(x) - (T\tilde{v})_i(x) &= (T_{a^*} v)_i(x) - (T_{\tilde{a}^*} \tilde{v})_i(x) && (9)\\
&\leqslant (T_{\tilde{a}^*} v)_i(x) - (T_{\tilde{a}^*} \tilde{v})_i(x) && (10)\\
&\leqslant \|(T_{\tilde{a}^*} v)_i - (T_{\tilde{a}^*} \tilde{v})_i\|_\infty && (11)\\
&\leqslant C_i \|v - \tilde{v}\|_\infty , && (12)
\end{aligned}$$

where in the last step we used Proposition 5.1. In fact, from the proof of Proposition 5.1 we have the slightly stronger bound

$$(Tv)_i(x) - (T\tilde{v})_i(x) \leqslant C_i \max_{j \neq i} \|v_j - \tilde{v}_j\|_\infty$$

Taking the max over i yields

$$\|Tv - T\tilde{v}\|_\infty \leqslant \max_i C_i \cdot \|v - \tilde{v}\|_\infty ,$$

as desired. □

Proposition 5.3. *F is a contraction with*

$$\|(Fv)_i - (F\tilde{v})_i\|_\infty \leqslant C_i \|v - \tilde{v}\|_\infty$$

and thus

$$\|Fv - F\tilde{v}\|_\infty \leqslant \max_i C_i \cdot \|v - \tilde{v}\|_\infty$$

Proof. We proceed by induction on $i \geqslant 1$. For the base case, note that by the definition of $(Fv)_1$ and Proposition 5.2, we have $\|(Fv)_1 - (F\tilde{v})_1\|_\infty \leqslant C_1 \|v - \tilde{v}\|_\infty$. Next, assume the result holds up to $i-1$ and consider $\|(Fv)_i - (F\tilde{v})_i\|_\infty$ for $i \geqslant 2$. By definition, $(Fv)_i = (T\widehat{v})_i$ and $(F\tilde{v})_i = (T\widehat{\tilde{v}})_i$. Thus

$$\begin{aligned}
&\|(Fv)_i - (F\tilde{v})_i\|_\infty = \left\|(T\widehat{v})_i - (T\widehat{\tilde{v}})_i\right\|_\infty && (13)\\
&\leqslant C_i \left\|\widehat{v} - \widehat{\tilde{v}}\right\|_\infty && (14)\\
&= C_i \max\{\|(Fv)_1 - (F\tilde{v})_1\|_\infty , \ldots , \|(Fv)_{i-1} - (F\tilde{v})_{i-1}\|_\infty , \|v_i - \tilde{v}_i\|_\infty , \ldots , \|v_m - \tilde{v}_m\|_\infty\} && (15)
\end{aligned}$$

By the inductive hypothesis, each of the first $i-1$ entries of the maximum are less than $\|v - \tilde{v}\|_\infty$. Thus the entire expression is bounded by by $C_i \|v - \tilde{v}\|_\infty$, completing the induction. □

Remark 5.4. *A smaller Lipschitz constant for F will be derived in the almost-causal case in Proposition 6.3.*

Finally, we show that F-iterations (i.e. MMVI) converge faster than T-iterations (i.e. vanilla VI) starting with a strictly-improving underestimate of the value function. This result is the analogue of Proposition 3.3 in Chapter 1 of Bertsekas Volume 2.

Proposition 5.5. *Suppose that v satisfies $v \leqslant Tv \leqslant u$. Then $T^k v \leqslant F^k v \leqslant u$ for all $k \geqslant 1$.*

Proof. We must show that $(T^k v)_i \leqslant (F^k v)_i \leqslant v_i(x)$ for all k, i. We proceed by double induction on $(k, i) \in \mathbb{N} \times [m]$.

The base case of $(k, i) = (1, 1)$ is clear, as $(Fv)_1 = (Tv)_1$. Assume that the result holds up to and including $k \geqslant 1$, and consider the claim for $k + 1$. Note that $(F^{k+1}v)_i = (T(\widehat{F^k v}))_i$, where $\widehat{F^k v} = ((F^{k+1}v)_1, \ldots, (F^{k+1}v)_{i-1}, (F^k v)_i, \ldots, (F^k v)_m)$. By the inductive hypothesis, $(F^{k+1}v)_j \geqslant (T^{k+1}v)_j \geqslant (T^k v)_j$ for $1 \leqslant j \leqslant i - 1$ and $(F^k v)_j \geqslant (T^k v)_j$ for $i \leqslant j \leqslant m$. Therefore $\widehat{F^k v} \geqslant T^k v$ and $(F^{k+1}v)_i = (T\widehat{F^k v})_i \geqslant (T^{k+1}v)_i$, as claimed. The upper bound is similar (?). □

6 Almost-Causal PDMPs

We now consider MMVI applied to an "almost-causal" PDMP.

Definition 6.1. *A PMDP is almost-causal if $\lambda_{ij} = 0$ for all i, j with $i \geqslant 2$ and $j > i$.*

In other words, the only allowed transitions are between mode i and mode j with $j \leqslant i$. For mode 1, we allow transitions to any other mode.

Remark 6.2. *Really, the above definition should say "there exists an ordering", but we do the above so that F corresponds to the best order of relaxations (?).*

Note that the Lipschitz constant of vanilla VI in Proposition 5.2 is at most $\max_i C_i$. We show that in the almost-causal case, F has a stronger guarantee. Specifically, its Lipschitz constant is at most $C_1 \leqslant \max_i C_i$.

Proposition 6.3. *For an almost-causal PDMP, F satisfies*

$$\|(Fv)_1 - (F\tilde{v})_1\|_\infty \leqslant C_1 \cdot \max_{j \geqslant 2} \|v_j - \tilde{v}_j\|_\infty$$

and

$$\|(Fv)_i - (F\tilde{v})_i\|_\infty \leqslant C_i \, \|(Fv)_1 - (F\tilde{v})_1\|_\infty$$

for $i \geqslant 2$. Thus F is a contraction with Lipschitz constant C_1.

Proof. We proceed by induction on $i \geqslant 1$. The result for $i = 1$ is clear using the stronger versions of the contraction bound for T in the proof of Proposition 5.2. Next, assume the result holds up to $i-1$ and consider $\|(Fv)_i - (F\tilde{v})_i\|_\infty$ for $i \geqslant 2$. Then

$$\|(Fv)_i - (F\tilde{v})_i\|_\infty = \left\|(T\widehat{v})_i - (T\widehat{\tilde{v}})_i\right\|_\infty \tag{16}$$

$$\leqslant C_i \max_{j<i} \left\|\widehat{v}_j - \widehat{\tilde{v}}_j\right\|_\infty \tag{17}$$

$$= C_i \max\{\|(Fv)_1 - (F\tilde{v})_1\|_\infty, \ldots, \|(Fv)_{i-1} - (F\tilde{v})_{i-1}\|_\infty\} \tag{18}$$

By the inductive hypothesis, the maximum evaluates to $\|(Fv)_1 - (F\tilde{v})_1\|_\infty$ and we find an upper bound of $C_i \|(Fv)_1 - (F\tilde{v})_1\|_\infty$, completing the induction. □

We now consider the rate of convergence of MMVI in the almost-causal case. Let $\tilde{C} = \max_{i \geqslant 2} C_i$.

Proposition 6.4. *For an almost-causal PDMP, the MMVI algorithm yields geometric convergence according to* $\|u^n - u\|_\infty \leqslant C_1 \tilde{C} \|u^{n-1} - u\|_\infty$ *for* $n \geqslant 2$.

Proof. Let u_i^n denote the mode i value function estimate at stage n of the MMVI algorithm. For $n \geqslant 2$, we claim that

$$\|u_1^n - u_1\|_\infty \leqslant C_1 \tilde{C} \left\|u_1^{n-1} - u_1\right\|_\infty$$

and for $n \geqslant 1$, we claim that

$$\|u_i^n - u_i\|_\infty \leqslant C_i \|u_1^n - u_1\|_\infty$$

for $i \geqslant 2$.

The second inequality follows directly from the second inequality in Proposition 6.3. To see why the first inequality holds, we apply the first inequality in Proposition 6.3 with $v = u^{n-1}$ and $\tilde{v} = u$. This yields

$$\|u_1^n - u_1\|_\infty \leqslant C_1 \cdot \max_{j \geqslant 2} \left\|u_j^{n-1} - u_j\right\|_\infty \tag{19}$$

$$\leqslant C_1 \tilde{C} \cdot \left\|u_1^{n-1} - u_1\right\|_\infty, \tag{20}$$

as desired. Thus the MMVI algorithm yields geometric convergence according to $\|u^n - u\|_\infty \leqslant C_1 \tilde{C} \|u^{n-1} - u\|_\infty$, as desired. □

Remark 6.5. *The result in Proposition 6.4 does not hold for the first iteration, $n = 1$. For $n = 1$, we only have a Lipschitz constant of C_1.*

Remark 6.6. *Note that if $\lambda_{1j} = 0$ for all $j \geqslant 2$ and the PDMP is causal, then $C_1 = 1 - \exp(0) = 0$ and Proposition 6.4 implies that $\|u^1 - u\|_\infty = 0$, as expected.*

Remark 6.7. *We are assuming that $Tu = u$; do we need more conditions for u to exist? (?)*

7 Finer Analysis

Below, we show an improvement of the bound given in Proposition 5.1. The bound above is weakened by step 5 in the inequality chain.

Proposition 7.1. *For all $a \in \mathcal{A}$, v, and $\tilde{v}$, we have*

$$\|(T_a v)_i - (T_a \tilde{v})_i\|_\infty \leqslant C_i \cdot \max_{x \in \Omega} \frac{\sum_{j \neq i} \lambda_{ij} \|v_j - \tilde{v}_j\|_\infty}{\sum_{j \neq i} \lambda_{ij}}$$

for all $i \in [m]$.

Proof. Note that

$$\begin{aligned}
(T_a v)_i(x) - (T_a \tilde{v})_i(x) &\leqslant e^{-\mu_i(t)} \int_t^{t_i^*} e^{\mu_i(s)} \cdot \sum_{j \neq i} \lambda_{ij} |v_j - \tilde{v}_j| ds && (21)\\
&= e^{-\mu_i(t)} \int_t^{t_i^*} \left[\sum_{j \neq i} \lambda_{ij} \right] e^{\mu_i(s)} \cdot \frac{\sum_{j \neq i} \lambda_{ij} |v_j - \tilde{v}_j|}{\sum_{j \neq i} \lambda_{ij}} ds && (22)\\
&\leqslant e^{-\mu_i(t)} \cdot \max_{x \in \Omega} \frac{\sum_{j \neq i} \lambda_{ij} |v_j - \tilde{v}_j|}{\sum_{j \neq i} \lambda_{ij}} \cdot \int_t^{t_i^*} \left[\sum_{j \neq i} \lambda_{ij} \right] e^{\mu_i(s)} ds && (23)\\
&\leqslant C_i \cdot \max_{x \in \Omega} \frac{\sum_{j \neq i} \lambda_{ij} |v_j - \tilde{v}_j|}{\sum_{j \neq i} \lambda_{ij}} && (24)\\
&\leqslant C_i \cdot \max_{x \in \Omega} \frac{\sum_{j \neq i} \lambda_{ij} \|v_j - \tilde{v}_j\|_\infty}{\sum_{j \neq i} \lambda_{ij}} && (25)
\end{aligned}$$

where the second-to-last step used the same exact-integrand trick as in Proposition 5.1. □

Corollary 7.2. *For all v and $\tilde{v}$, we have*

$$\|(Tv)_i - (T\tilde{v})_i\|_\infty \leqslant C_i \cdot \max_{x \in \Omega} \frac{\sum_{j \neq i} \lambda_{ij} \|v_j - \tilde{v}_j\|_\infty}{\sum_{j \neq i} \lambda_{ij}}$$

for all $i \in [m]$.

Remark 7.3. *Note that in Corollary 7.2 if we bound $\|v_j - \tilde{v}_j\|_\infty$ by $\|v - \tilde{v}\|_\infty$ for all j, we recover Proposition 5.2.*

Remark 7.4. *Unlike the bound $C_i \|v - \tilde{v}\|_\infty$ given in Proposition 5.1, these bounds involve a weighted average of $\|v_j - \tilde{v}_j\|_\infty$ values based on the switching rates λ_{ij}. Therefore, these are helpful for finer analysis of iterations. We give an example in the next section.*

8 Examples of Finer Analysis

We apply the stronger bound in Corollary 7.2 to better understand the following examples.

8.1 Example 1: Outlier Mode

Consider an almost-causal PDMP. Suppose we have an initialization u^0 such that u_i^0 is within ε of u_i for all $i \neq k$, while u_k^0 is a distance $\delta \gg \varepsilon$ away from u_k. In other words, mode k is an outlier. According to the bound in the remark following Proposition 6.4, $\|u^1 - u\|_\infty \leqslant C_1 \|u^0 - u\|_\infty = C_1 \delta$. However, in the event that λ_{1k} is small, we can derive a better bound using Corollary 7.2.

Proposition 8.1. *Consider u^0 such that $\|u_i^0 - u_i\|_\infty < \varepsilon$ for $i \neq k$ and $\|u_k^0 - u_k\|_\infty = \delta$, where $\delta \gg \varepsilon$. Suppose that $\max_{x \in \Omega} \frac{\lambda_{1k}}{\sum_{j \geqslant 2} \lambda_{1j}} < \eta$. Then*

$$\|u^1 - u\|_\infty \leqslant C_1((1-\eta)\varepsilon + \eta\delta),$$

i.e.

$$\|u^1 - u\|_\infty \leqslant C_1 \bigg((1-\eta)\frac{\varepsilon}{\delta} + \eta \bigg) \|u^0 - u\|_\infty$$

Proof. It suffices to show that $\|u_1^1 - u_1\|_\infty \leqslant C_1 \bigg((1-\eta)\frac{\varepsilon}{\delta} + \eta \bigg) \|u^0 - u\|_\infty$; the second inequality in Proposition 6.3 implies the result.

To see this, we apply Corollary 7.2:

$$\|(Tu^0)_1 - u_1\|_\infty \leqslant C_1 \cdot \max_{x \in \Omega} \frac{\sum_{j \geqslant 2} \lambda_{1j} \|u_j^0 - u_j\|_\infty}{\sum_{j \geqslant 2} \lambda_{1j}} \tag{26}$$

$$\leqslant C_1 \cdot \max_{x \in \Omega} \left[\frac{\sum_{j \geqslant 2, j \neq k} \lambda_{1j}}{\sum_{j \geqslant 2} \lambda_{1j}} \varepsilon + \frac{\lambda_{1k}}{\sum_{j \geqslant 2} \lambda_{1j}} \delta \right] \tag{27}$$

$$\leqslant C_1 \cdot \bigg((1-\eta)\varepsilon + \eta\delta \bigg), \tag{28}$$

as desired. □

Remark 8.2. *This improves on Proposition 6.4 when $n = 1$, as Proposition 6.4 only yields a contraction coefficient of C_1 when $n = 1$. (See the remark following Proposition 6.4.)*

8.2 Example 2: The Lollipop

We consider the "lollipop" PDMP, which consists of a ring adjoined to a line. One can see that the lollipop is almost-causal by relabeling the modes as in Figure 1.

The following proposition shows that these two components converge at different rates under MMVI. Further, the rate of convergence of the ring utilizes Corollary 7.2.

Proposition 8.3. *Consider the lollipop PDMP with constant rate of switching λ_1 into the ring and constant rate of switching λ_2 into the line. Then the MMVI algorithm yields convergence in one iteration for modes $2, \ldots, \ell$. Further, the rest of the PDMP converges to the value function geometrically according to $\|u^n - u\|_\infty \leqslant C' \|u^{n-1} - u\|_\infty$ for $n \geqslant 2$, where $C' = \frac{\lambda_1}{\lambda_1 + \lambda_2} C_1 \cdot C_m \cdots C_{\ell+1}$.*

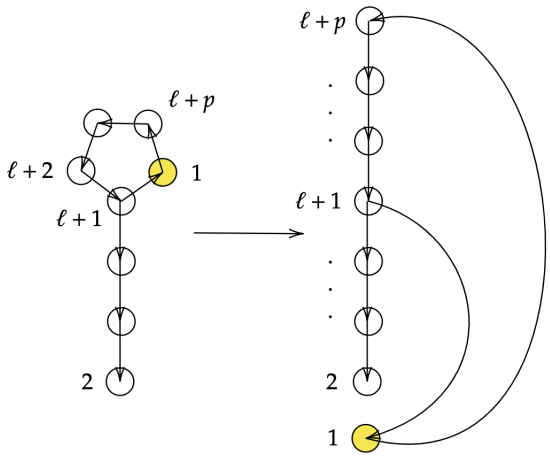


Figure 1: The almost-causal lollipop PDMP.

Proof. It is clear that modes $2, \dots, \ell$ converge in one iteration, as the sub-PDMP consisting of modes $2, \dots, \ell$ is causal.

We now consider the modes in the ring. First note that by Corollary 7.2 (or even the intermediate bounds in Proposition 5.2), we have

$$\|u_1^n - u_1\|_\infty \leqslant C_1 \|u_m^{n-1} - u_m\|_\infty$$

for $n \geqslant 1$. Now by Corollary 7.2 on mode $\ell + 1$, we find

$$\|u_{\ell+1}^n - u_{\ell+1}\|_\infty \leqslant \frac{\lambda_1}{\lambda_1 + \lambda_2} C_{\ell+1} \|u_1^n - u_1\|_\infty$$

for $n \geqslant 1$. Finally, by Corollary 7.2 (or even the intermediate bounds in Proposition 5.2), we have

$$\|u_m^n - u_m\|_\infty \leqslant C_m \cdots C_{\ell+2} \cdot \frac{\lambda_1}{\lambda_1 + \lambda_2} C_{\ell+1} \|u_1^n - u_1\|_\infty$$

for $n \geqslant 1$. Therefore for $n \geqslant 2$, we have

$$\|u_1^n - u_1\|_\infty \leqslant C_1 \|u_m^{n-1} - u_m\|_\infty \leqslant C' \|u_1^{n-1} - u_1\|_\infty \leqslant C' \|u^{n-1} - u\|_\infty$$

As $\|u_k^n - u_k\|_\infty \leqslant \|u_1^n - u_1\|_\infty$ for all k for all $n \geqslant 1$, we find that $\|u^n - u\|_\infty \leqslant C' \|u^{n-1} - u\|_\infty$, as desired. □

Remark 8.4. *The result in Proposition 8.3 does not hold for the first iteration, $n = 1$. For $n = 1$, we only have a Lipschitz constant of C_1.*

9 Worst Ordering

(This section is in progress.)

We can also consider what happens when we use the worst possible order of Gauss-Seidel relaxations in the almost-causal case. Specifically, we consider the operator that relaxes modes m through 1, against the direction of causality.

Definition 9.1. *The "bad" MMVI operator G is given by the following procedure: begin by setting $(Gv)_i = (Tv)_i$ for $i \in \{2, \dots, m\}$. Then set $(Gv)_1 = (T\widehat{v})_1$, where $\widehat{v} = (v_1, (Gv)_2, \dots, \dots, (Gv)_m)$. The result is the updated estimate Gv.*

Using similar techniques to the above, we can show the following guarantees for G-iterations:

$$\|u_i^n - u_i\|_\infty \leqslant C_i \max_{j<i} \|u_j^{n-1} - u_j\|_\infty \leqslant C_i \|u^{n-1} - u\|_\infty$$

for $i \geqslant 2$ and $n \geqslant 1$, and

$$\|u_1^n - u_1\|_\infty \leqslant C_1 \max_{j \geqslant 2} \|u_j^n - u_j\|_\infty \leqslant C_1 \tilde{C} \|u^{n-1} - u\|_\infty$$

for $n \geqslant 2$. As a result, our Lipschitz constant is $\tilde{C} = \max_{i \geqslant 2} C_i$ for $n \geqslant 2$, which is greater than the guarantee we obtained for "good" MMVI. Thus the "bottleneck" appears to be the synchronous component $\{2, \dots, m\}$.

Remark 9.2. *Note that unlike F, G updates modes $\{2, \dots, m\}$ synchronously. Also, note that unlike F, G requires a good guess for all of modes $\{2, \dots, m\}$ to converge quickly.*

10 Policy Iterations

(This section is in progress.)

Given a policy $a = (a_1, \dots, a_m)$, let $v_a = ((v_a)_1, \dots, (v_a)_m)$ denote the cost of evaluating the policy. Note that v_a is obtained by solving the following system of m linear PDEs for $w = v_a$:

$$f_i(x) \nabla w_i(x) \cdot a_i(x) + K_i(x) + \sum_{j \neq i} \lambda_{ij}(x)(w_j(x) - w_i(x)) = 0 \quad \forall i \in [m]$$

Vanilla Policy Iterations (PI) consists of the following steps (?):

1. Initialization: Fix a policy a^0.
2. Policy Evaluation: Compute v_{a^n}.
3. Policy Improvement: Let a^{n+1} satisfy $Tv_{a^n} = T_a v_{a^n}$. Return to Step 2.

Remark 10.1. *A multimode version of PI could modify the Policy Improvement step as follows: first compute a such that $Tv_{a^n} = T_a v_{a^n}$, and set $(a^{n+1})_1 = a_1$. Then process $i \in \{2, \dots, m\}$ sequentially such that at step i, we compute a such that $Tv_{\widehat{a}^n} = T_a v_{\widehat{a}^n}$, where $\widehat{a}^n = ((a^{n+1})_1, \dots, (a^{n+1})_{i-1}, a_i^n, \dots, a_m^n)$, and we set $(a^{n+1})_i = a_i$. In other words, we do a Gauss-Seidel update on the policy as it is being extracted. Perhaps one could show a stronger rate of superlinear convergence for "MMPI"?*

11 Quasi-Causal PDMPs

(This section is in progress.)

In this section, we consider an alternative definition of almost-causal. (We might switch the definitions of almost- and quasi-causal later...)

Definition 11.1. *A PDMP is η-quasi-causal if $\eta_i = \max_{x\in\Omega} \frac{\sum_{j>i}\lambda_{ij}}{\sum_{j\neq i}\lambda_{ij}} \leqslant \eta$ for all $m-1 \leqslant j \leqslant 1$.*

Consider "good" MMVI applied to an η-quasi-causal PMDP. For convenience, let $e_i^n = \|u_i^n - u_i\|_\infty$ and $e^n = \|u^n - u\|_\infty$.

For a given mode i, we consider all backwards transitions to mode j with $j > i$ as transitioning to "mode 0", and we let $f_{i0} = \frac{\sum_{j>i}\lambda_{ij}}{\sum_{j\neq i}\lambda_{ij}}$ at a given point x. Grouping all modes j with $j > i$ into "mode 0" and applying Corollary 7.2, we have

$$e_i^{n+1} \leqslant C_i[f_{i0}e^n + \sum_{j<i} f_{ij}e_j^{n+1}]$$

for all $m \leqslant i \leqslant 2$ and for $i = 1$ we have

$$e_1^{n+1} \leqslant C_1 f_{10} e^n$$

We now "unravel" the system of inequalities to find a bound of the form $e_i^{n+1} \leqslant Ae^n + (1-A)e_1^n$ for all $2 \leqslant i \leqslant m$. Let M denote the *zero-indexed* $(m+1)$-by-$(m+1)$ lower triangular matrix whose entries are $M_{ij} = C_i f_{ij}$ for $1 \leqslant i \leqslant m$ and $j \leqslant i$, $M_{00} = 1$, and zero elsewhere. Then

$$e_i^{n+1} \leqslant [M^{i-1}]_{i0} \cdot e^n + [M^{i-1}]_{i1} \cdot e_1^{n+1}$$

is a bound on e_i^{n+1} at a given point x. Note that the above bound is almost a weighted average of e^n and e_1^{n+1}, modulo the factors C_j. This matrix reflects the unraveling of the bounds, because it roughly corresponds to a Markov chain with absorbing states being modes 0 and 1.

If one were to bound the entries of the matrix M by η at this point, the bound on e_i^{n+1} would no longer be a weighted average, and it would yield nothing. It seems like we have to find a way to directly bound $[M^{i-1}]_{i0}$.

Note that the underlying Markov chain is almost symmetric between modes 0 and 1; at any mode i, there is a probability f_{i0} of transitioning to mode 0 and a probability f_{i1} of transitioning to mode 1. Can we assume a relation between f_{i0} and f_{i1} that allows us to simplify our bound? For example, if we assume that $\frac{f_{i0}}{f_{i1}} \leqslant \zeta$ for all i, then we know that $[M^{i-1}]_{i0} \leqslant \zeta[M^{i-1}]_{i1}$; each path that leads to mode 0 corresponds to a path leading to mode 1. Let $s = [M^{i-1}\mathbf{1}]_i$. Then since $e_1^{n+1} \leqslant C_1 f_{10} e^n \leqslant e^n$, our bound is less than the convex combination

$$e_i^{n+1} \leqslant \frac{\zeta}{1+\zeta} s \cdot e^n + \frac{1}{1+\zeta} s \cdot e_1^{n+1}$$

Is it reasonable to assume that $\frac{f_{i0}}{f_{i1}} \leqslant \zeta$, when the PDMP is η-quasi-causal? Also, can we bound s itself in terms of the C_j's? Consider a simple case to start, where C_j's are equal or λ_{ij}'s are independent of x? (Note that the above bound has to take ζ as $\max_{x\in\Omega}$...)

12 Next Steps

Here are some ideas for next steps:

- Add in conditions where they seem to be missing. In particular, justify the existence of the value function u and the identity $Tu = u$?
- Compute lower bounds on the Lipschitz constants of each of the operators in the general and almost-causal cases? This would allow us to definitively say which iterative algorithm is faster.
- Building on the previous section, consider multimode analogues of Policy Iterations and Value-Policy Iterations? The Improvement Lemma should follow from the fact that all the operators are contractions.
- Consider MMVI for in scenarios with incomplete information, like POMDPs and OOPDMPs?