Check for updates

Adv. Appl. Probab. **58,** 993–1024 (2026)
*doi:*10.1017/apr.2025.10044

NUMERICAL METHOD FOR SOLVING IMPULSE CONTROL PROBLEMS IN PARTIALLY OBSERVED PIECEWISE DETERMINISTIC MARKOV PROCESSES

ALICE CLEYNEN,* *CNRS, The Australian National University*
BENOÎTE DE SAPORTA ,** *University of Montpellier*

Abstract

Designing efficient and rigorous numerical methods for sequential decision-making under uncertainty is a difficult problem that arises in many applications frameworks. In this paper we focus on the numerical solution of a subclass of impulse control problems for the piecewise deterministic Markov process (PDMP) when the jump times are hidden. We first state the problem as a partially observed Markov decision process (POMDP) on a continuous state space and with controlled transition kernels corresponding to some specific skeleton chains of the PDMP. We then proceed to build a numerically tractable approximation of the POMDP by tailor-made discretizations of the state spaces. The main difficulty in evaluating the discretization error comes from the possible random jumps of the PDMP between consecutive epochs of the POMDP and requires special care. Finally, we discuss the practical construction of discretization grids and illustrate our method on simulations.

Keywords: Continuous time Markov process; dynamic programming; hidden process; numerical approximation; partially observed Markov decision process

2020 Mathematics Subject Classification: Primary 60J25; 93E20
Secondary 60J05; 93E11

1. Introduction

A large number of problems in science, including resource management, financial portfolio management, monitoring complex systems, and medical treatment design (to name just a few), can be characterized as sequential decision-making problems under uncertainty. In such problems, an agent interacts with a dynamic, stochastic, and incompletely known process, with the goal of finding an action-selection strategy that optimizes some performance measure over several time steps. In optimal stopping problems, for instance when the agent has to decide when to replace some component of a production chain before full deterioration, the policy typically does not influence the underlying process until replacement, when the process starts again with the same dynamics. However, in many decision-making problems an important aspect is the

Received 4 November 2025; accepted 4 November 2025.
* Postal address: IMAG, University of Montpellier, CNRS, Montpellier, France and John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia. Email: alice.cleynen@umontpellier.fr
** Postal address: IMAG, University of Montpellier, CNRS, Montpellier, France. Email: benoite.de-saporta@umontpellier.fr

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

effect of the agent's policy on the data collection, different policies naturally yielding different behaviors of the process at hand.

In this paper we focus on the numerical solution of a subclass of impulse control problems for the piecewise deterministic Markov process (PDMP) when the process is partially observed, in the very challenging case when jump times are hidden. PDMPs are continuous-time processes with a hybrid state space: they can have both discrete and Euclidean variables. The dynamics of a PDMP is determined by three local characteristics: the flow, the jump intensity, and the Markov jump kernel (see [14, Section 24]). Between jumps, trajectories follow a deterministic flow. The frequency of jumps is determined by the intensity, and the post-jump location is selected by the Markov kernel.

General impulse control for PDMPs allows the controller to act on the process dynamics by choosing intervention dates and new states to restart from at each intervention. This family of problems was first studied by Costa and Davis in [13]. It has received a lot of attention since (see, for instance, [1, 12, 18, 20]), and was further extended, for instance, in [23, 31]. In these papers, the authors define a rigorous mathematical framework to state such control problems and establish some optimality equations such as dynamic programming equations for the value function. Numerical methods to compute an approximation of the value function and an ϵ-optimal policy are also briefly or extensively described, for instance, in [13, 16, 17]. They rely on discretization of the state space, either with direct cartesian grids or with dynamic grids obtained from simulations of the inter-jump-time-post-jump-location discrete-time Markov chain embedded in the PDMP.

In all the papers cited above, the process is supposed to be perfectly observed at all times. However, in many real-life applications, such as dynamic reliability and biological systems, continuous measurements may not be available, and measures may be corrupted by noise. The field of dynamic reliability is particularly suitable for controlled PDMP models (see, for instance, [9, 18, 21] or [27, p. 96]). Consider a complex system with interacting components subject to degradation or failure. The controller aims to design a maintenance policy to optimize some suitable target, such as maximizing production while minimizing downtime or maintenance costs. The discrete variables correspond to the states of the different components, and the Euclidean variables are those that influence the process, for instance the age of the components, or some physical characteristics such as pressure or temperature. It is natural in this context to suppose that the partial or total failure of some components may be hidden. The physical variables are measured with captors or monitoring systems, and observations may thus be noisy (see, for instance, [4, 5, 34]). Biological systems can also be successfully modeled by PDMPs: one can cite examples in the neurosciences [32, 33], population dynamics and genomics [10, 22, 26], and medical treatment [8, 30] (to name just a few). Continuous-time observation of biological systems is often impossible, especially *in vivo*, and again measurements may be indirect and corrupted by noise.

Designing efficient and mathematically sound approximation methods to solve continuous-time and continuous-state-space impulse control problems under partial observation is very challenging, especially for processes with jumps such as PDMPs. Relevant literature is scarce. One can mention [2] or [6] (in the special easier case of optimal stopping) where the position of the process is observed through noise, but the jump times are still perfectly observed, so that the properties of the inter-jump-time-post-jump-location chain can still be fully exploited. Another related recent work is [7] where the author studies an optimal control problem for pure jump processes (corresponding to PDMPs with constant flow) under partial observations. However, they consider continuous control instead of impulse control and the observations are not corrupted by noise.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

A first step towards solving the impulse control problem for PDMPs under partial and noisy observations with hidden jump times was made by the authors in [8] in the special easier case of optimal stopping. They have shown that when trajectories of the process can be simulated, a double discretization allows the value function to be approximated with general error bounds, and provides a candidate policy with excellent performance. In this paper, we address a more general class of impulse control problems under partial observations where decisions do influence both the data collection and the dynamics of the process. More specifically, we focus on a subclass of impulse control problems with four specific properties. First, the lapse between interventions can only take a finite number of values. Second, observations are collected at intervention times and corrupted by noise. For instance, in the dynamic reliability framework these two properties correspond to collecting data only at inspection dates, and the delay until the next inspection is scheduled in a given range (one month, one year, . . .). Third, interventions only act on the discrete variables. For instance, in the dynamic reliability framework, we do not allow maintenance that would reset the Euclidean variables to some nominal value. At an intervention, the settings of the system or the environment can be punctually modified, leading to a different set of local characteristics without changing the value of the Euclidean variables. Fourth and finally, the state space of the controlled PDMP has no active boundary, so that the process is punctuated only by random jumps and not by boundary jumps.

This setting is more general than the setting of continuous-time Markov decision processes (CTMDPs). Indeed, in CTMDPs, the state space is supposed to be finite or countable, whereas we have a hybrid state space here. In CTMDPs, the process evolves in continuous time, and actions are taken at each jump time. The jump times are determined by exponential clocks from the controlled transition rates matrix and are not directly controller-chosen (see, for instance, [25, Section 2.2]). In our setting of controlled PDMPs, natural jumps may occur between interventions, an intervention does not correspond to a natural jump of the process, and intervention dates are directly controller-chosen.

In our particular PDMP setting, the impulse control problem can be stated as a partially observed Markov decision process (POMDP), and this is the first main contribution of this paper. This POMDP is in discrete time, with epochs corresponding to intervention dates. However, the state space is not discrete, and becomes infinite-dimensional when turned into the corresponding fully observed Markov decision process (MDP) for the filter on the belief space (see, for instance, [3, Section 5.3]). While such MDPs have theoretical exact optimal solutions, they are numerically intractable. In addition, the filter process cannot be simulated, even for fixed policies, preventing the use of direct simulation-based discretization procedures. Our second main contribution is to propose and prove the convergence of an algorithm to approximate the value function and explicitly build a policy close to optimality. Our approach is based on tailor-made discretizations of the state spaces of the Euclidean variable of the PDMP and then of the corresponding discrete filter. It takes into account the major difficulty that the combinatorics associated to the possible decisions is too large to explore the whole belief space through simulations. Our third main contribution is to discuss the practical construction of the discretization grids, which is no easy problem. Our results are illustrated by simulations on an example of medical treatment optimization for cancer patients.

The paper is organized as follows. In Section 2 we state our generic optimization problem and turn it into a POMDP. In Section 3 we give our main assumptions and our resolution strategy, detailing construction of the approximate value function and candidate policy. In Section 4, we provide a real-life example of medical treatment optimization fitting our framework and briefly present other general classes of applications fitting our framework.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

We provide experimental results and discussion of the practical construction of discretization grids in Section 5. Finally, in Section 6 we provide a short conclusion. We postpone the main proofs to the Appendix, and give details and specifications of the model used for the simulation study together with further discussion on the construction of grids in the supplementary materials.

2. Problem statement

We start by specifying the special class of impulse control problem for PDMPs we are focusing on, then we show how our control problem can be expressed as a POMDP.

2.1. Impulse control for hidden PDMPs

Let us first define the class of controlled PDMPs we consider. Let $\mathcal{M} = L \times M$ be a two-dimensional finite set. We will call regimes or modes its elements. We use a product set to distinguish between modes in L that can be controller-chosen and modes in M that cannot. For all regimes $(\ell, m) \in \mathcal{M}$, let E_m^ℓ be a bounded Borel subset of $\mathbb{R}^d$ endowed with a norm $\|\cdot\|$. Set $E = \{(\ell, m, \mathrm{x}), \ell \in L, m \in M, \mathrm{x} \in E_m^\ell\}$, and $E^\ell = \{(m, \mathrm{x}), m \in M, \mathrm{x} \in E_m^\ell\}$ for all $\ell \in L$. A PDMP on the state space E is determined by three local characteristics:

- The flow $\mathbf{\Phi}(x, t) = (\ell, \mathbf{\Phi}^\ell(m, \mathrm{x}, t)) = (\ell, m, \Phi_m^\ell(\mathrm{x}, t))$ for all $x = (\ell, m, \mathrm{x})$ in E and $t \geq 0$, where $\Phi_m^\ell : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^d$ is continuous and satisfies a semi-group property $\Phi_m^\ell(\cdot, t+s) = \Phi_m^\ell(\Phi_m^\ell(\cdot, t), s)$, for all $t, s \in \mathbb{R}_+$. This describes the deterministic trajectory between jumps.
- The jump intensity $\boldsymbol{\lambda}(x) = \boldsymbol{\lambda}^\ell(m, \mathrm{x}) = \lambda_m^\ell(\mathrm{x})$ for all $x = (\ell, m, \mathrm{x})$ in E, where $\lambda_m^\ell : E_m^\ell \to \mathbb{R}^+$ is a measurable function such that for any $x = (\ell, m, \mathrm{x})$ in E, there exists $\epsilon > 0$ such that

$$\int_0^\epsilon \lambda_m^\ell(\Phi_m^\ell(\mathrm{x}, t))\, \mathrm{d}t < +\infty.$$

  For all $x = (\ell, m, \mathrm{x})$ in E and $t \geq 0$, set

$$\mathbf{\Lambda}(x, t) = \mathbf{\Lambda}^\ell(m, \mathrm{x}, t) = \Lambda_m^\ell(\mathrm{x}, t) = \int_0^t \lambda_m^\ell(\Phi_m^\ell(\mathrm{x}, s))\, \mathrm{d}s.$$

- The Markov kernel $\mathbf{Q}$ on $(\mathcal{B}(E), E)$ represents the transition measure of the process and allows us to select the new location after each jump. It satisfies $\mathbf{Q}(\{x\}|x) = 0$ for all $x \in E$. We also write $Q_m^\ell(\cdot\ |\mathrm{x}) = \mathbf{Q}(\cdot|x) = \mathbf{Q}^\ell(\cdot|m, \mathrm{x})$ for all $x = (\ell, m, \mathrm{x}) \in E$ and add the additional constraint that Q cannot change the value of ℓ, as ℓ is intended to be controller-chosen, that is, Q_m^ℓ sends E^ℓ onto itself.

The formal probabilistic apparatus necessary to precisely define controlled trajectories and to formally state the impulse control problem is rather cumbersome, and will not be used in the sequel. Therefore, for the sake of simplicity, we only present an informal description of the optimal control problem. The interested reader is referred to [13], [15, section 54], or [23] for a formal setting. Our optimization problem will be rigorously stated as a POMDP in Section 2.2.

We consider a finite-horizon problem. Let $H > 0$ be the optimization horizon, $\delta > 0$ a fixed minimal lapse such that $N = H/\delta$ is an integer. Note that δ is not supposed to be small. A general impulse strategy $\mathcal{S} = (\ell_n, r_n)_{0 \leq n \leq N-1}$ is a sequence of non-anticipative E-valued random

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Algorithm 1 Simulation of a trajectory of a controlled PDMP between interventions n and $n+1$ from state $x_n = (\ell_n, m_n, \mathrm{x}_n)$

1: $s \leftarrow 0$
2: $\ell \leftarrow \ell_n$, $m \leftarrow m_n$, $\mathrm{x} \leftarrow \mathrm{x}_n$, $x \leftarrow (\ell, m, \mathrm{x})$
3: **while** $s < r_n$ **do**
4: $S \sim \lambda_m^\ell(\mathrm{x})$
5: $X_{\tau_n+s+t} \leftarrow \mathbf{\Phi}(x, t)$ for $0 \le t \le \min\{S, r_n - s\}$
6: $\mathrm{x} \leftarrow \Phi_m^\ell(\mathrm{x}, \min\{S, r_n - s\})$
7: **if** $S < r_n - s$ **then**
8: $s \leftarrow s + S$
9: $(\ell, m', \mathrm{x}') \sim Q_m^\ell(\cdot\,|\mathrm{x})$
10: $m \leftarrow m'$, $\mathrm{x} \leftarrow \mathrm{x}'$, $x \leftarrow (\ell, m, \mathrm{x})$
11: **else**
12: $s \leftarrow r_n$
13: **end if**
14: **end while**

variables on a measurable space $(\Omega, \mathcal{F})$ and of non-anticipative intervention lapses. In this work, we only consider a subclass of strategies where ℓ_n takes values in L and r_n is a multiple of δ: $r_n \in \mathbb{T}$, where $\mathbb{T}$ is a subset of $\delta^{1:N} = \{\delta, 2\delta, \dots, N\delta\}$ that contains δ. This means that on the one hand the controller can only act on the process by changing its regime, that is, by selecting the local characteristics to be applied until the next intervention, and on the other hand the lapse between consecutive interventions belongs to the finite set $\mathbb{T}$. The trajectories of the PDMP controlled by strategy $\mathcal{S}$ are constructed recursively between intervention dates τ_n (defined recursively by $\tau_0 = 0$ and $\tau_{n+1} = \tau_n + r_n$) as described in Algorithm 1. In line 5 of Algorithm 1, $S \sim \lambda_m^\ell(\mathrm{x})$ means that S has the survival function

$$\mathrm{P}_x(S > t) = \mathrm{e}^{-\int_0^t \lambda_m^\ell(\Phi_m^\ell(\mathrm{x},s))\,\mathrm{d}s} = \mathrm{e}^{-\mathbf{\Lambda}(x,t)}.$$

Note that boundary jumps are not allowed. The strategy $\mathcal{S}$ induces a family of probability measures $\mathrm{P}_x^{\mathcal{S}}$, $x \in E$, on a suitable probability space $(\Omega, \mathcal{F})$. Associated to strategy $\mathcal{S}$, we define the following expected total cost for a process starting at $x \in E$:

$$\mathcal{J}(\mathcal{S}, x) = \mathrm{E}_x^{\mathcal{S}}\Big[\int_0^H c_r(X_s)\,\mathrm{d}s + \sum_{n=0}^{N-1} c_i(X_{\tau_n}, \ell_n, r_n) + c_t(X_H)\Big], \tag{2.1}$$

where $\mathrm{E}_x^{\mathcal{S}}$ is the expectation with respect to $\mathrm{P}_x^{\mathcal{S}}$, c_r is the running cost, c_i the intervention cost, and c_t the terminal cost.

The last ingredient needed to state the optimization problem is to define admissible strategies. Again, the rigorous definition will be given in Section 2.2 in the framework of POMDPs. Informally, decisions can only be taken in view of some discrete-time noisy observations of the process, instead of the exact value of the process at all times. More specifically, we make the following assumptions.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

- Observations are only available at decision times τ_n.
- The controller-chosen regimes $\ell \in L$ are observed, while the uncontrolled regimes $m \in M$ are hidden.
- The Euclidean variable is observed through noise: at time τ_n, if the state of the process is $x = (\ell, m, \mathrm{x})$, the controller receives observation $y_n = F(x) + \epsilon_n = F(\mathrm{x}) + \epsilon_n$, where (ϵ_n) are real-valued independent and identically distributed random variables with density f independent of the controlled PDMP. We further assume that the random variables y_n take values in a compact interval I of the real line.

Denote by $\mathbb{S}$ the set of all admissible strategies. Our aim is to compute an approximation of the value function

$$\mathcal{V}(x) = \inf_{\mathcal{S} \in \mathbb{S}} \mathcal{J}(\mathcal{S}, x),$$

and explicitly construct a strategy close to optimality. We also assume that the process starts from a fixed point $x_0 = (\ell_0, m_0, \mathrm{x}_0) \in E$.

2.2. Partially observed Markov decision process

Finding a suitably rigorous way to state an impulse control problem for PDMPs with hidden jumps and under noisy observation is by no means straightforward, especially as regards defining admissible strategies (see, for instance, [1] or [11, Section 1.1]). This is our first main contribution in this paper. As decision dates are discrete, we use the framework of POMDPs to rigorously state our control problem. In the sequel, we drop the regime $\ell \in L$ from the state x and include it in the action instead to better fit the standard POMDP notation. Denote $E_m = \bigcup_{\ell \in L} E_m^\ell$ and $E_M = \{(m, \mathrm{x}), m \in M, \mathrm{x} \in E_m\}$. Define on E_M the following hybrid distance, for $x = (m, \mathrm{x})$ and $x' = (m', \mathrm{x}') \in E_M$:

$$\|x - x'\| = (\|\mathrm{x} - \mathrm{x}'\|)\mathbb{1}_{m=m'} + \infty \mathbb{1}_{m \neq m'}.$$

Let $(\mathbb{X}, \mathbb{A}, \mathbb{K}, R, c, C)$ be the POMDP with the following characteristics.

• The state space is $\mathbb{X} = \{\xi = (x, y, w) \in (E_M \times \mathbb{O})\} \cup \{\Delta\}$, with $\mathbb{O} = \{\gamma \in I \times \delta^{1:N}\}$ the observation space. It gathers the values X_n of the hidden PDMP at dates τ_n as well as the observation process Y_n, and one additional observed variable W_n that is the time elapsed since the beginning. State Δ is a cemetery state where the process is sent after the horizon H is reached.

• The action space is $\mathbb{A} = (L \times \mathbb{T}) \cup \{\breve{d}\}$, where $\breve{d}$ is an empty decision that is taken when the horizon H is reached. This purely technical decision sends the process to the cemetery state Δ.

• The constraints set $\mathbb{K} \subset \mathbb{X} \times \mathbb{A}$ is such that its sections $\mathbb{K}(\xi) = \{d \in \mathbb{A}; (\xi, d) \in \mathbb{K}\}$ satisfy $\mathbb{K}(\Delta) = \{\breve{d}\}$ and $\mathbb{K}(x, y, w) = \mathbb{K}(w)$ for $(x, y, w) \in \mathbb{X} - \{\Delta\}$ as constraints concern the remaining control time. More specifically, one has:

- $\mathbb{K}(w) \subset L \times (\mathbb{T} \cap \{\delta, 2\delta, \ldots, N\delta - w\})$ if $w \leq \delta(N-1)$, to force the last intervention to occur exactly at the horizon time H;
- $\mathbb{K}(N\delta) = \breve{d}$, denoting that no intervention is possible after the horizon has been reached.

Note that as $\delta \in \mathbb{T}$, the set $\mathbb{T} \cap \{\delta, 2\delta, \ldots, N\delta - w\}$ is never empty.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

• The controlled transition kernels R are defined as follows: for any bounded measurable function g on $\mathbb{X}$, any $\xi \in \mathbb{X}$ and $d \in \mathbb{K}(\xi)$, one has

$$Rg(\xi, \check{d}) = g(\Delta),$$
$$Rg(\xi, d) = \int_I \int_E g(x', y', w + r) f(y' - F(x')) P(\mathrm{d}x' | x, d)\, \mathrm{d}y',$$

if $\xi = (x, y, w)$, $d = (\ell, r) \neq \check{d}$, and $Rg(\xi, d) = 0$ otherwise, where $P(\cdot|x, d)$ is the distribution of X_r conditionally on $X_0 = x$ under regime ℓ, if $d = (\ell, r)$. Note that since, conditionally on X_n, Y_n is independent of any past X_k and Y_k, kernel R does not depend on y.

• The terminal cost function $C : \mathbb{X} \to \mathbb{R}_+$ satisfies $C(\Delta) = 0$ and $C(x, y, w) = c_t(x)$ for $(x, y, w) \in \mathbb{X} - \{\Delta\}$.

• The cost-per-stage function $c : \mathbb{K} \times \mathbb{X} \to \mathbb{R}_+$ satisfies $c(\Delta, \check{d}, \cdot) = 0$, $c(\xi, \check{d}, \cdot) = C(\xi)$ if $\xi \neq \Delta$. Ideally, to match equation (2.1), for $d = (\ell, r)$ one should choose c as

$$c(x, d) = \mathrm{E}_x^\ell \left[\int_0^r c_r(X_s)\, \mathrm{d}s \right] + c_i(x, d).$$

However, the integral part has no simple analytical expression and we choose some simpler proxies instead (see Section 5.1).

• The optimization horizon is finite and equals N.

Classically, the sets of observable histories are defined recursively by $H_0 = \mathbb{O} \cup \{\Delta\}$ and $H_n = H_{n-1} \times \mathbb{A} \times (\mathbb{O} \cup \{\Delta\})$. A decision rule at time n is a measurable mapping $g_n : H_n \to \mathbb{A}$ such that $g_n(h_n) \in \mathbb{K}(\gamma_n)$ for all histories $h_n = (\gamma_0, d_0, \gamma_1, d_1, \ldots, \gamma_n)$. A sequence $\pi = (g_n)_{0:N-1} = (g_0, \ldots, g_n, \ldots, g_{N-1})$ where g_k is a decision rule at time k is called an admissible policy. Let Π_N denote the set of all admissible policies. The controlled trajectory of the POMDP following policy $\pi = (g_n)_{0:N-1} \in \Pi_N$ is defined recursively by $\mathbb{X}i_0 \in \mathbb{X}$ and, for $0 \leq n \leq N - 1$:

- $A_n = g_n(\mathbb{X}i_n)$;
- $\mathbb{X}i_{n+1} \sim R(\cdot\,|\mathbb{X}i_n, A_n)$.

Note that the cemetery state Δ ensures that all trajectories have the same length N, even if they do not have the same number of actual decisions ($d \neq \check{d}$). Then one can define the total expected cost of policy $\pi \in \Pi_N$ starting at $\xi_0 \in \mathbb{X}$ as

$$J_\pi(\xi_0) = \mathrm{E}_{\xi_0}^\pi \left[\sum_{n=0}^{N-1} c(\mathbb{X}i_n, A_n, \mathbb{X}i_{n+1}) + C(\mathbb{X}i_N) \right],$$

and our control problem corresponds to the optimization problem

$$V(\xi_0) = \inf_{\pi \in \Pi_N} J_\pi(\xi_0).$$

Now the problem is rigorously stated, our next aim is to compute an approximation of the value function V and explicitly construct a strategy close to optimality.

3. Resolution strategy and assumptions

Our second main contribution is to propose a numerical approach to (approximately) solve our POMDP. The first difficulty in solving our POMDP comes from the fact that it is partially

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

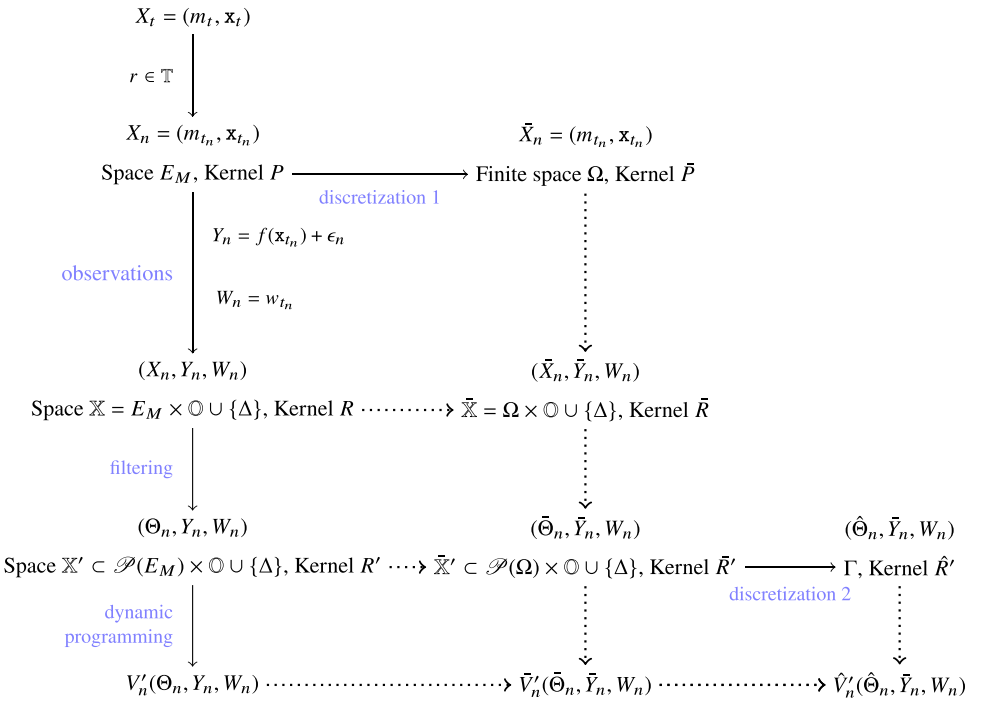


FIGURE 1. Graphical sketch of the approximation procedure. The first column corresponds to the original PDMP/POMDP/fully observed MDP. The second column corresponds to the first discretization. The third column corresponds to the second discretization

observed. Thus, our first step is to convert it into an equivalent fully observed MDP on a suitable belief space $\mathbb{X}'$ by introducing a filter process. This is done in Section 3.2. While dynamic programming equations hold true for the fully observed MDP and in theory provide the exact optimal strategy for the optimization problem, there are two main difficulties to its practical resolution. On the one hand, the state space $\mathbb{X}'$ is continuous and infinite-dimensional. On the other hand, the filter process cannot be simulated, even for fixed policies, preventing the use of direct simulation-based discretization procedures.

Our approach to (approximately) solve the belief MDP is to construct a numerically tractable approximation of its value function based on discretizations of the dynamic programming equations. To obtain a finite-dimensional process and an approximation of the filter process that can be simulated, we first discretize the Euclidean part of the state space E of the PDMP (X_t). Then we discretize the state space of the approximated filtered process in order to obtain a numerically tractable approximation. Finally, we obtain an explicit candidate policy by solving the resulting discretized approximation of the dynamic programming equations. This procedure is detailed in Section 3.3 and illustrated in Figure 1.

In order to keep track of the discretization errors throughout the different steps described above, we need our transition kernels to be sufficiently regular. Hence, we start this section with stating the assumptions required on the parameters of the PDMP.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

3.1. Regularity assumptions

The first set of assumptions will ensure the regularity of kernel P. First, we require the local characteristics of the PDMP to be sufficiently regular. This is a classical assumption, which usually ensures the regularity of the kernel of the inter-jump-time-post-jump-location Markov chain embedded in the PDMP (see, for instance, [18, Section 1.4]). However, here the kernel P is more complex as any number of jumps may occur between two interventions. Therefore, we need an additional restrictive assumption (Assumption 3.5) to limit the number of jumps between consecutive interventions. Denote by $\overline{\delta} = \max\{r \in \mathbb{T}\}$ the maximum time between consecutive interventions.

Assumption 3.1. *The flow* $\mathbf{\Phi}$ *is Lipschitz continuous: there exists a positive constant* $[\Phi]$ *such that for all* $\ell \in L$, $x, x' \in E$, $0 \le t, t' \le \overline{\delta}$, *one has*

$$|\mathbf{\Phi}^{\ell}(x,t) - \mathbf{\Phi}^{\ell}(x',t')| \le [\Phi](\|x - x'\| + |t - t'|).$$

Assumption 3.2. *The jump intensity* $\boldsymbol{\lambda}$ *is Lipschitz continuous along the flow: there exists a positive constant* $[\lambda]$ *such that for all* $\ell \in L$, x *and* $x' \in E_M$, $0 \le t \le \overline{\delta}$, *one has*

$$|\boldsymbol{\lambda}^{\ell}(\Phi^{\ell}(x,t)) - \boldsymbol{\lambda}^{\ell}(\Phi^{\ell}(x',t))| \le [\lambda]\|x - x'\|.$$

Assumption 3.3. *The jump intensity* $\boldsymbol{\lambda}$ *is bounded: there exists a positive constant* $\|\lambda\|$ *such that for all* $x \in E_M$, *and* $\ell \in L$, *one has*

$$|\boldsymbol{\lambda}^{\ell}(x)| \le \|\lambda\|.$$

In the sequel, for all bounded measurable functions h we note that

$$\mathbf{Q}^{\ell}h(x) = \int_{E_M} h(x')\mathbf{Q}^{\ell}(\mathrm{d}x', x).$$

Assumption 3.4. *The Markov kernel* $\mathbf{Q}$ *is Lipschitz continuous along the flow in the following sense: there exists a positive constant [Q] such that for all functions* $h \in BL_1(E)$, *for all* $\ell \in L$, *for all* $x, x' \in E_M$, *one has*

$$\left|\mathbf{Q}^{\ell}h(x) - \mathbf{Q}^{\ell}h(x')\right| \le [Q]\|x - x'\|.$$

Assumption 3.5. *For any control* $\ell \in L$, *for any* $x \in E_M$, *the maximum number of jumps of the PDMP starting at time 0 from x and following control* ℓ *up to time* $\overline{\delta}$ *is bounded by* $N_j < +\infty$.

This last assumption is further discussed in Appendix B.2. We also require additional regularity assumptions on the observation process and the cost functions.

Assumption 3.6. *There exist non-negative real constants* L_Y, $\underline{f}$, *and* $\overline{f}$ *such that for all* $(x, x') \in E_M$ *and* $y \in I$, *one has*

$$\begin{aligned} \left|f(y - F(x)) - f(y - F(x'))\right| &\le L_Y\|x - x'\|, \\ 0 < \underline{f} \le f(y - F(x)) &\le \overline{f} < +\infty. \end{aligned}$$

We also set $L_f = L_Y|I|$ *and* $B_f = \overline{f}|I|$, *where* $|I| = \int_I dy$ *is the length of interval I.*

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Assumption 3.7. *There exist non-negative real constants L_c, L_C, B_c, and B_C such that for all x, x',x'' in E_M and $d \in L \times \mathbb{T}$, one has*

$$\begin{aligned} |C(x)| &\le B_C, \\ |c(x, d, x')| &\le B_c, \\ |C(x) - C(x')| &\le L_C \|x - x'\|, \\ |c(x, d, x') - c(x'', d, x')| &\le L_c \|x - x''\|, \\ |c(x, d, x') - c(x, d, x'')| &\le L_c \|x' - x''\|. \end{aligned}$$

3.2. Equivalent MDP on the belief state

To convert a POMDP into an equivalent fully observed MDP is classical, therefore details are omitted. The interested reader may consult, for instance, [3, Section 5.3] or [6, 8] for similar derivations. For $n \le N$, set $\mathcal{F}_n^O = \sigma(Y_k, W_k, 0 \le k \le n)$, the σ-field generated by the observations up to n. Let

$$\Theta_n(A) = \mathrm{P}(X_n \in A | \mathcal{F}_n^O) = \mathrm{P}(\mathbb{X}i_n \in A \times \{(Y_n, W_n)\} | \mathcal{F}_n^O)$$

denote the filter or belief process for the unobserved part of the process, for any Borel subset A of E_M. The standard prediction-correction approach yields a recursive construction for the filter.

Proposition 3.1. *For any $n \ge 0$, conditionally on $(Y_{n+1}, W_{n+1}) = (y', w')$, $d = (\ell, r) \in L \times \mathbb{T}$, and $\Theta_n = \theta$, one has $\Theta_{n+1} = \Psi(\theta, y', w', d)$ with*

$$\Psi(\theta, y', w', d)(A) = \frac{\int_{E_M} \int_{E_M} f(y' - F(x')) \mathbb{1}_A(x') P(\mathrm{d}x' | x, d) \theta(\mathrm{d}x)}{\int_{E_M} \int_{E_M} f(y' - F(x')) P(\mathrm{d}x' | x, d) \theta(\mathrm{d}x)},$$

for any Borel subset A of E_M.

Let $\mathcal{P}(A)$ denote the set of probability measures on set A. The equivalent fully observed MDP $(\mathbb{X}', \mathbb{A}, \mathbb{K}', R', c', C')$ is defined as follows.

- The state space is a subset $\mathbb{X}'$ of $(\mathcal{P}(E_M) \times \mathbb{O}) \cup \{\Delta\}$ satisfying the following constraints: all $\xi = (\theta, y, w) \in \mathbb{X}'$ satisfy $\theta(E_{m_0}) \ge (\underline{f}\bar{f}^{-1})^{\frac{w}{\delta} \vee 1} \mathrm{e}^{-w\|\lambda\|}$, where m_0 is the mode of the fixed starting point. It is necessary to restrict the state space to ensure the regularity of our operators (see Appendix B.3). This constraint is technical, and is guaranteed as soon as the process starts in mode m_0 (see Appendix A.1).
- The action space is still $\mathbb{A} = (L \times \mathbb{T}) \cup \{\check{d}\}$.
- The constraints set $\mathbb{K}' \subset \mathbb{X}' \times \mathbb{A}$ is such that its sections $\mathbb{K}'(\xi) = \{d \in \mathbb{A};\ (\xi, d) \in \mathbb{K}'\}$ satisfy $\mathbb{K}'(\Delta) = \check{d}$ and $\mathbb{K}'(\theta, y, w) = \mathbb{K}'(w) = \mathbb{K}(w)$ for $(\theta, y, w) \in \mathcal{P}(E_M) \times \mathbb{O}$.
- The controlled transition kernels R' are defined as follows: for any bounded measurable function g on $\mathbb{X}'$, any $\xi \in \mathbb{X}'$, and $d \in \mathbb{K}'(\xi)$, one has $R'g(\xi, d) = g(\Delta)$ if $d = \check{d}$, and

$$R'g(\xi, d) = \int_{E_M} \int_I \int_{E_M} g\big(\Psi(\theta, y', w + r, d), y', w + r\big) f(y' - F(x')) P(\mathrm{d}x' | x, d)\, \mathrm{d}y' \theta(\mathrm{d}x)$$

if $\xi = (\theta, y, w)$, $d = (\ell, r) \ne \check{d}$. See Appendix A.1 for the proof that R' maps $\mathbb{X}'$ onto itself.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

• The non-negative cost-per-stage function $c' : \mathbb{K}' \to \mathbb{R}_+$ and the terminal cost function $C' : \mathbb{X}' \to \mathbb{R}_+$ are defined by $C'(\Delta) = c'(\Delta, \check{d}, \cdot) = 0$, and for $\xi = (\theta, \gamma) \in \mathcal{P}(E_M) \times \mathbb{O}$ and $d \in \mathbb{K}'(\xi)$,

$$c'(\xi, d) = c'(\theta, \gamma, d) = \int_{E_M^2} c(x, d, x') P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x),$$
$$C'(\xi) = C'(\theta, \gamma) = \int_{E_M} C(x)\theta(\mathrm{d}x).$$

• The optimization horizon is still N.

Denote by $(\mathbb{X}i'_n)$ a trajectory of the fully observed MDP. The cost of strategy $\pi \in \Pi_N$ is

$$J'(\pi, \xi'_0) = \mathrm{E}^{\pi}_{\xi'_0}\left[\sum_{n=0}^{N-1} c'(\mathbb{X}i'_n, A_n) + C'(\mathbb{X}i'_N)\right],$$

and the value function of the fully observed problem is

$$V'(\xi'_0) = \inf_{\pi \in \Pi_N} J'(\pi, \xi'_0).$$

If $\xi'_0 = (\delta_{x_0}, y_0, w_0)$, one has $V'(\xi'_0) = V(x_0, y_0, w_0)$, so that solving the partially observed MDP is equivalent to solving the fully observed one. In addition, the value function V' satisfies the well-known dynamic programming equations (see, for instance, [2, Section 5.3]).

Theorem 3.1. *For $\xi \in \mathbb{X}'$, set $v'_N(\xi) = C'(\xi)$, and for $0 \le n \le N-1$, define by backwards induction*

$$v'_n(\xi) = \min_{d \in \mathbb{K}'(\xi)} \left\{c'(\xi, d) + R' v'_{n+1}(\xi, d)\right\}.$$

Let $\xi'_0 = (\delta_{(m_0, \mathrm{x}_0)}, y_0, 0) \in \mathbb{X}'$. Then we have $v'_0(\xi'_0) = V'(\xi'_0) = V((m_0, \mathrm{x}_0), y_0, 0)$.

3.3. Approximations of the value function

Our approximation of the value function is based on discretizations of the underlying state spaces.

3.3.1. *First discretization.* Let $\Omega = \{\omega^1, \ldots, \omega^{n_\Omega}\}$ be a finite grid on E_M containing at least one point in each mode $m \in M$. Let p_Ω denote the nearest-neighbour projection from E_M onto Ω for the distance defined in Section 2.1. In particular, p_Ω preserves the mode. Set $\Omega_m = E_m \cap \Omega$ for all $n \in M$. Let $(C_i)_{1 \le i \le K}$ be a Voronoi tessellation of E_M associated to Ω. Namely, $(C_i)_{1 \le i \le n_\Omega}$ is a partition of E_M such that for all $1 \le i \le n_\Omega$, one has

$$C_i \subset \{x \in E_M\, ; \|x - \omega^i\| \le \|x - \omega^j\|\ \forall j \in 1 : n_\Omega\}.$$

We will denote by $\mathcal{D}_i$ the diameter of cell C_i: $\mathcal{D}_i = \sup\{\|x - x'\|; x, x' \in C_i\}$. As stated in Section 2.1, the state spaces are supposed to be bounded. Denote $\|E_M\| = \sup\{\|x\|, x \in E_M\}$. Hence, all $\mathcal{D}_i$ are finite and bounded by $\|E_M\|$.

We define the controlled kernels $\bar{P}$ from $E_M \times (L \times \mathbb{T})$ onto Ω as

$$\bar{P}(\omega^j|x, d) = P(C_j|x, d),$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

for all $x \in E$, $d \in L \times \mathbb{T}$, and $1 \le j \le n_\Omega$. In particular, the restriction of $\bar{P}$ to Ω is a controlled Markov kernel on Ω. We now replace kernel P by kernel $\bar{P}$ in the dynamic programming equations of Theorem 3.1 in order to obtain our first approximation.

Let $\bar{\mathbb{X}}' \subset \{\xi = (\theta, \gamma) \in \mathbb{X}' ; \theta(\Omega) = 1, \} \cup \{\Delta\}$ satisfying the following constraints: all $\bar{\xi} = (\bar{\theta}, y, w) \in \bar{\mathbb{X}}'$ satisfy $\bar{\theta}(\Omega_{m_0}) \ge (\underline{f}\bar{f}^{-1})^{\frac{w}{\delta} \vee 1} \mathrm{e}^{-w\|\lambda\|}$, where m_0 is the mode of the fixed starting point. Let $\bar{\Psi}$ be the approximate filter operator defined by replacing the integrals with respect to P in the filter operator Ψ defined in Proposition 3.1 by integrals with respect to $\bar{P}$. The approximated controlled transition kernels $\bar{R}'$ are defined as follows: for any bounded measurable function g on $\bar{\mathbb{X}}'$, any $\xi \in \bar{\mathbb{X}}'$ and $d \in \mathbb{K}'(\xi)$, one has $\bar{R}'g(\xi, d) = g(\Delta)$ if $d = \check{d}$, and

$$\bar{R}'g(\xi, d) = \int_I \sum_{\omega^i \in \Omega} \sum_{\omega^j \in \Omega} g\big(\bar{\Psi}(\bar{\theta}, y', w + r, d), y', w + r\big) f(y' - F(\omega^j)) \bar{P}(\omega^j | \omega^i, d) \bar{\theta}(\omega^i) \, \mathrm{d}y'$$

if $\xi = (\bar{\theta}, y, w)$, $d = (\ell, r)$. Note that this kernel does not depend on y, and see Appendix A.2 for the proof that $\bar{R}$ sends $\bar{\mathbb{X}}'$ onto itself. The approximated cost-per-stage function is defined by $\bar{c}'(\Delta, \check{d}) = 0$ and

$$\bar{c}'(\bar{\theta}, \gamma, d) = \sum_{\omega^i \in \Omega} \sum_{\omega^j \in \Omega} c(\omega^i, d, \omega^j) \bar{P}(\omega^j | \omega^i, d) \bar{\theta}(\omega^i),$$

for all $\xi = (\bar{\theta}, \gamma) \in \bar{\mathbb{X}}'$ and $d \in \mathbb{K}'(\xi)$. Finally, for all $\xi \in \bar{\mathbb{X}}'$, set $\bar{v}'_N(\xi) = C'(\xi)$, and for $0 \le n \le N - 1$, define by backwards induction

$$\bar{v}'_n(\xi) = \min_{d \in \mathbb{K}(\xi)} \left\{ \bar{c}'(\xi, d) + \bar{R}' \bar{v}'_{n+1}(\xi, d) \right\}.$$

If the grid Ω is precise enough, $\bar{P}$ should be a good approximation of kernel P and thus one can expect that functions $\bar{v}'_n$ are good approximations of our value functions v'_n. Indeed, we have the following result.

Theorem 3.2. *Under Assumptions 3.1–3.7, for all $\xi' \in \bar{\mathbb{X}}'$, one has*

$$|v'_N(\xi') - \bar{v}'_N(\xi')| = 0,$$
$$|v'_n(\xi') - \bar{v}'_n(\xi')| \le C_{v'_n} \sup_{j \in \{1, \dots, n_\Omega\}} \mathcal{D}_j, \quad 0 \le n < N,$$

where $C_{v'_n}$ depends only on n, N, δ, and the regularity constants of the parameters.

The proof of this result is given in Appendix B.7 and is based on the form of the kernels, their regularity (especially that of kernel P), and the dynamic programming equation.

The main gain with this first approximation is that the filter operator $\bar{\Psi}$ now involves only finite weighted sums and therefore the corresponding approximate filter process can now be simulated. However, functions $\bar{v}'_n$ still cannot be computed because of the continuous integration in y' and because the state space $\bar{\mathbb{X}}'$ is still continuous. We now proceed to a second discretization in order to obtain numerically tractable approximations of our value functions.

3.3.2. *Second discretization.* Let $\Gamma = \{\rho^1, \dots, \rho^{n_\Gamma}\}$ be a finite grid on $\bar{\mathbb{X}}'$ containing at least Δ, one element $\xi = (\theta, \gamma)$ such that $\theta(E_M) = 1$. Let p_Γ denote the nearest-neighbour projection

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

from $\bar{\mathbb{X}}'$ onto Γ. Here, the distance on $\bar{\mathbb{X}}'$ is such that $d(\xi, \xi') = d_\Omega(\bar\theta, \bar\theta') + |y - y'|$ if $\xi = (\bar\theta, y, w)$, $\xi' = (\bar\theta', y', w')$, with $w = w'$, $d(\xi, \xi') = +\infty$ otherwise, where d_Ω is defined as

$$d_\Omega(\bar\theta, \bar\theta') = \left[\sum_{k=1}^{n_\Omega} |\bar\theta(\omega^k) - \bar\theta'(\omega^k)|^2\right]^{1/2},$$

meaning that $\mathcal{P}(\Omega)$ is identified with the n_Ω-dimensional simplex endowed with the L^2-norm. Let $(\bar C_i)_{1\le i\le n_\Gamma}$ be a Voronoi tessellation of $\bar{\mathbb{X}}'$ associated to the grid Γ. We will denote by $\bar{\mathcal{D}}_i$ the diameter of cell $\bar C_i$. Note that the space $\bar{\mathbb{X}}'$ is compact, so that all $\bar{\mathcal{D}}_i$ are finite and bounded by $1 + |I|$. The controlled kernels $\hat R'$ are defined on $\bar{\mathbb{X}}'$ as follows: for any $\rho^j \in \Gamma$, any $\xi \in \bar{\mathbb{X}}'$ and $d \in \mathbb{K}'(\xi)$, one has

$$\hat R'(\rho^j|\xi, d) = \bar R'(\bar C_j|\xi, d).$$

In particular, the restriction of $\hat R'$ to Γ is a family of controlled Markov kernels on Γ. For $1 \le j \le n_\Gamma$, set $\hat v'_N(\rho^j) = C'(\rho^j)$, and for $0 \le n \le N-1$ and $1 \le j \le n_\Gamma$, define by backwards induction

$$\hat v'_n(\rho^j) = \min_{d\in\mathbb{K}'(\rho^j)} \left\{\bar c'(\rho^j, d) + \hat R' \hat v'_{n+1}(\rho^j, d)\right\}.$$

Again, if the grid Γ is precise enough, $\hat R'$ should be a good approximation of kernel $\bar R'$ and thus one can expect that functions $\hat v'_n$ are good approximations of functions $\bar v'_n$ and thus of our value functions v'_n. More precisely, we have the following result.

Theorem 3.3. *Under Assumptions 3.6 and 3.7, for all points $\rho^j \in \Gamma$, one has*

$$|\hat v'_N(\rho^j) - \bar v'_N(\rho^j)| = 0,$$
$$|\hat v'_n(\rho^j) - \bar v'_n(\rho^j)| \le C_{\bar v'_n} \sup_{j\in\{1,\dots,n_\Gamma\}} \bar{\mathcal{D}}_j,$$

where $C_{\bar v'_n}$ depends only on n, N, δ, and the regularity constants of the parameters.

The proof of this result is given in Appendix C.3. The main gain with this second approximation is that integration against kernel $\hat R'$ boils down to computing a finite weighted sum, and functions $\hat v'_n$ are defined on a finite state space which makes the dynamic programming equations fully tractable numerically.

3.4. Candidate strategy

We can now construct a computable strategy using the fully discretized value function. The idea is to first build an approximate filter using the operator $\bar\Psi$ and project the resulting filter together with the current observation onto grid Γ. Then one selects the next decision using the dynamic programming equation on the grid Γ. More precisely, suppose that the process starts from point $\xi_0 = (m_0, \mathrm{x}_0, y_0, 0)$, such that the initial observation is $\gamma_0 = (y_0, 0)$. One can recursively compute an approximate filter $(\bar\theta_n)$ and the corresponding decisions (d_n) as follows. First, set

$$\bar\theta_0 = \delta_{(m_0,\mathrm{x}_0)}, \quad d_0(\bar\theta_0, \gamma_0) = \mathrm{argmin}_{d\in\mathbb{K}'(\bar\theta_0,\gamma_0)}\left\{\bar c'(p_\Gamma(\bar\theta_0, \gamma_0), d) + \hat R' \hat v'_1(p_\Gamma(\bar\theta_0, \gamma_0), d)\right\}.$$

Suppose one has constructed the sequence $(\bar\theta_n, d_n)$ up to stage $k-1$. Then after receiving the kth observation γ_k, the next approximate filter and decisions are

$$\bar\theta_k = \bar\Psi(\bar\theta_{k-1}, \gamma_k, d_{k-1}), \ d_k(\bar\theta_k, \gamma_k) = \mathrm{argmin}_{d\in\mathbb{K}'(\bar\theta_0,\gamma_0)}\left\{\bar c'(p_\Gamma(\bar\theta_k, \gamma_k), d) + \hat R' \hat v'_1(p_\Gamma(\bar\theta_k, \gamma_k), d)\right\}.$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

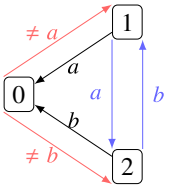


FIGURE 2. State graph for the changes of mode in the medical example. Letters indicate under which treatments the jumps are possible. Black arrows indicate changes to remission phase, red arrows changes to *standard* relapse, and blue arrows changes to *therapeutic escape*

with the convention that the last (not required) decision is $d_N = \check{d}$. Note that it should be better to use operator $\bar{\Psi}$ on $\bar{\theta}_{k-1}$ then project the result onto grid Γ than to use operator $\hat{\Psi}$ (obtained by replacing integration with respect to θ by integration with respect to $p_\Gamma(\theta)$ in the definition of $\bar{\Psi}$) as it should generate a smaller error. Although it is reasonable to think that this strategy should be close to optimality, it is an open problem to actually prove it as the sequence $(\bar{\theta}_n, \gamma_n)$ is not generated by the kernel $\bar{R}'$. Indeed, we use here the observations generated by the original sequence X_n with kernel P and not that generated by kernel $\bar{P}$, as in practice only the original observations are available. Its performance is assessed in Section 5 on a specific example.

4. Examples of impulse control problems

In this section, we give the specific medical example on which we ran our approximation procedure, and provide details of other control problems fitting into our framework.

4.1. Optimization of patient follow-up

We consider an example of patient follow-up that we will use in the simulation study in Section 5. The mode m corresponds to the overall state of the patient ($m = 0$, sound; $m = 1$, disease 1; $m = 2$, disease 2; $m = 3$: death of the patient). The Euclidean variable is two-dimensional $\mathrm{x} = (\zeta, u)$. Variable $\zeta \in [K_i, K_s]$ corresponds to some marker of the disease that can be measured, and $u \geq 0$ the time since the last jump in order to encompass semi-Markov dynamics. The control ℓ corresponds to the medical treatment ($\ell = \emptyset$, no treatment; $\ell = a$, efficient for disease 1 and slows the progression of disease 2; $\ell = b$, efficient for disease 2 and slows the progression of disease 1). Decision dates correspond to visits to the medical centre when the marker is measured and a new treatment is selected and applied until the next visit. Horizon H is 2400 days with possible visits every 15, 30, or 60 days ($\delta = 15$, $\overline{\delta} = 60$).

If treatment $\emptyset$ is applied, the patient may randomly jump from $m = 0$ to any of the two disease states $m \in \{1, 2\}$. In the sound state $m = 0$, the marker level is constant. In the disease states ($m = 1$ or $m = 2$), the marker level follows an increasing flow and no other jump is possible.

If treatment a is applied, the patient may only randomly jump from $m = 0$ to the other disease state $m = 2$. In the disease state $m = 1$, the marker follows a decreasing flow and may randomly jump to the other disease state $m = 2$. In the disease state $m = 2$, the marker follows an increasing flow and no other change of state is possible. The effects of treatment b are similar: decrease of the marker in disease $m = 2$, increase in disease $m = 1$. The specific values of the local characteristics of the PDMP and proof of their verifying the assumptions of Section 3.1 can be found in the supplementary material, and the state graph for mode transitions is given in Figure 2.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

4.2. Examples in biology

One may consider other kinds of therapeutic treatments, such as in [30] where the treatments correspond to interleukin injections in HIV-infected patients under antiretroviral treatment that must be suitably dosed and timed in order to maintain lymphocytes levels over a prescribed limit. In this example, the PDMP has no free modes and a controller-chosen mode corresponding to non-proliferating and proliferating lymphocytes, which dynamics can be influenced by punctual interleukin injections. Natural jumps correspond to the efficiency duration of an injection. Hence, there is at most one jump between two injections. It is natural to suppose that the lymphocytes are observed through noisy proxies at discrete dates.

Other examples in biology may cover, for instance, gene expression or chemostats. The simplest PDMP models for gene expression include a (possibly controller-triggered) promoter that can be on or off, and continuous levels of RNA and proteins (see, for instance, [26]). If only the controller can turn the promoter on or off, there are no natural jumps. One may also consider a network of interacting genes with partial control over promoters and feedback loops. Chemostats are bioreactors used to cultivate some micro-organisms. They can be modelled by PDMPs (see, for instance, [24]), and are controlled by tuning the flow of fresh medium. Again, it is natural to consider hidden information in these frameworks.

4.3. Dynamic reliability

The field of dynamic reliability is particularly suitable for controlled PDMP models, as explained in the introduction. The number of jumps in a given time interval is usually bounded for non-reparable systems, for instance as soon as there is a finite number of components, each with a single nominal state and a finite number of successive possible degradation states up to total failure. It is also common not to observe the state of each component at all times.

The simplest examples in this field are pure jump Markov processes. These fit our assumptions as the flow and intensity are constant in every mode, and the state space is discrete, hence all kernels are regular. See [28] for such an example in the simpler case of optimal stopping. The next important subclass is when the mode corresponds to the states of the components, and the Euclidean variables to their age or use-time (see, for instance, [4, 19] or [18, Section 1.8.4] for a repair-workshop model with similar variables). In this case, the control ℓ may correspond to a change of settings.

A typical example of a more complex controlled PDMP in this field is the heated hold-up tank, a well-known test case for dynamic reliability (see, for instance, [21, 29, 35]. The system consists of a tank containing a fluid whose level is controlled by three components: two inlet pumps and one outlet valve. A thermal power source heats up the fluid. The failure rate of the components depends on the temperature, the position of the three components monitors the liquid level in the tank, and in turn, the liquid level determines the temperature. The pumps and valve can be either functioning or stuck, which corresponds to the free modes m, and turned on or off (if functioning), corresponding to the controller-chosen mode ℓ. In any setting ℓ, there are at most three jumps (corresponding to the failure of the pumps and valve). The flow and jump parameters have analytical expressions and satisfy our regularity assumptions.

5. Simulation study

In this section we consider the medical example of Section 4.1. We briefly discuss the choice of the cost function and the grid construction, then state competing approaches and compare their performance to our approach, based on Monte Carlo simulations.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

5.1. Choice of the cost functions

The candidate policy depends on the cost functions, therefore the latter must be carefully chosen. We denote by $c_i(x, d) = C_V$ a fixed cost per visit that takes into account an emotional burden for the patient and health-care expenses for the check-up. For the counterpart of the integral of the running cost, we choose en expression of the form

$$\tilde{c}(x, d, x') = \kappa |\zeta' - K_i| r + \beta r \mathbb{1}_{\{\ell \neq \emptyset\}}.$$

Parameter $\beta > 0$ represents the cost of the treatment, $\kappa > 0$ is a scale parameter and $|\zeta' - K_i| r$ is a (crude) proxy of the integral of the process. Our cost function is then $c(x, d, x') = C_V + \tilde{c}(x, d, x')$. Calibrating those parameters is a difficult task that is not discussed here.

5.2. Construction of the grids

Because decisions influence the dynamics of the process, grids cannot be constructed with techniques based on simulations, such as quantization. For instance, on a horizon of 2400 days with possible visits every 15, 30, or 60 days, this leads to approximately 10^{152} possible strategies. Grids therefore have to be chosen by expert knowledge, and transition probabilities computed accordingly. In the supplementary materials we discuss the construction of parsimonious grids that provide satisfying performance.

5.3. Strategies in competition and performance criteria

To evaluate the performance of our approach (OS in the comparison tables) we compare our work with several other strategies.

The (unachievable) gold standard strategy is the *see all* (SA) where decisions are taken while observing the process X_{t_n} by choosing the optimal treatment for the current mode. In this SA strategy, we do not allow choice of the next visit date.

The *filter* strategy corresponds to choosing the optimal treatment for the estimated mode: at each new observation y_n the approximate filter $\bar{\theta}_n$ is computed based on the first discretization. Then the most probable mode $\hat{m}_n$ is obtained as

$$\hat{m}_n = \arg \max_{m=0:3} \sum_{j;\omega^j=(m,\cdot)} \bar{\theta}_n^j.$$

Note that to set up this strategy, only the first discretization is needed. As this discretization is the least computationally extensive, it might be worth designing larger grids Ω to obtain better mode estimates. However, this strategy does not take into account knowledge of the dynamics of the process, hence we expect it to have higher costs. It is not able to select the next decision date either.

The *standard* strategy is classically used in the hospitals and is based on a threshold s_{rel} for relapse, and on a fixed duration for treatment *tlength*. While the observations remain below s_{rel}, the patient does not receive treatment and visits are scheduled every 2 months. When s_{rel} is reached, the practitioner gives treatment b (corresponding to the most frequent relapse type 2) and the next visit is scheduled in 15 days. If at the next visit the observed marker level has decreased, treatment b is maintained with visits every 15 days for a total duration of *tlength* days. Otherwise the practitioner changes treatment for treatment a (with visits every 15 days, and same total duration). Once treatment is stopped, visits are scheduled every 2 months and the strategy is repeated.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

TABLE 1. Strategy comparisons. Performance of each strategy measured in terms of Monte Carlo costs.

	Visits	Real cost (sd)
	Choice	**203 (3.03)**
OS	15 days	291.36 (2.66)
	60 days	215.98 (3.2)
	15 days	321.28 (2.2)
Filter	60 days	223.8 (4.03)
	15 days	203.15 (0.89)
See all	60 days	**113.72** (1.79)
Standard		301.09 (5.69)

The *fixed dates* optimal strategies are the strategies based on our discretization and dynamic programming approach where the choice of the next visit date is not allowed. We will investigate the fixed dates every 15 and 60 days (FD-15 and FD-60 in the comparison tables).

The performance criterion is the real cost of each strategy evaluated on the real process X (averaged over 1000 simulations).

5.4. Results

First, note that all code is available online at https://github.com/acleynen/PDMP-control. All results presented here are based on 1000 simulations. The exact specifications can be found in the supplementary materials.

A first interesting result relates to the allocation of the computational burden. We compared, for a fixed number of elements of the second grid (n_Γ), the performance of our approach with different numbers of points in the first grid (n_Ω), the smallest ones being included in the largest. The performance of the approach drastically improves with a higher number of points in Ω, leading to a very crude discretization of $\mathcal{P}(\Omega)$, compared to a less refined discretization of $\mathbb{X}$ but leading to a more refined discretization of $\mathcal{P}(\Omega)$. In some cases, adding elements to Γ in fact worsened the results, as novel elements allowed consideration of filters with less certainty in the values of the process, most often leading to worse decisions. This is illustrated in Table 5 of the supplementary materials.

We therefore compared our approach with competing strategies on the largest Ω grid, containing 133 elements in Ω and 144 elements in Γ. Figure 3 illustrates a controlled trajectory with treatments indicated by shape and disease indicated by colour, for our approach. One can note that in most cases, the first time the projected filter has a most likely element in a disease mode, treatment is not administered but the time lapse to the next visit is reduced. Only on the next visit will treatment be allocated. This explains the differences of performance between the different strategies given in Table 1 and shows the added value of taking into account the knowledge of the dynamics of the process. It performs significantly better.

6. Conclusion

We have proposed a numerically feasible scheme to approximate the value function of an impulse control problem for a class of hidden PDMPs where control actions influence the dynamic of the process. Approximations rely on discretization of the observed and unobserved state spaces, for which we propose strategies to explore the large-dimensional belief spaces that depend on the process characteristics while maintaining error bounds for this approximation explicitly depending on the parameters of the problem. Code is freely available on GitHub.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

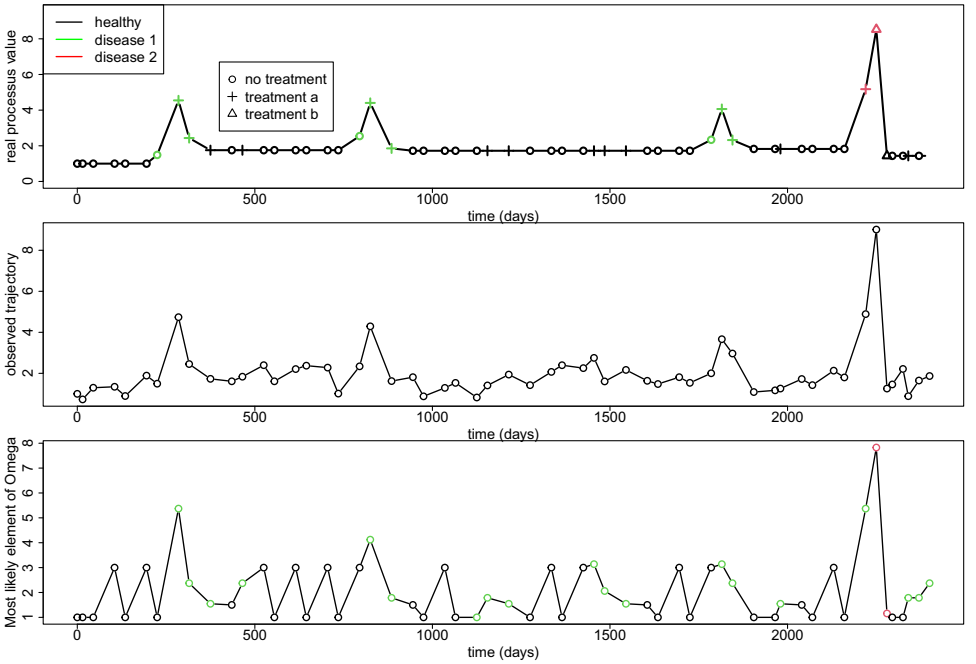


FIGURE 3. Example of a controlled (OS) simulated strategy. The top panel illustrates the (hidden) values of the true process, with modes indicated by colour, and allocated treatment indicated by point shape. The middle panel illustrates the observed trajectory (noisy version of the Euclidean variable from the panel above). The bottom panel illustrates the most likely element of Ω estimated by the filter of the process

This provides a mathematical setting to study realistic processes for instance in a disease-control framework, which we illustrate in simulations. This is a promising start for the study of even more realistic disease-control frameworks.

The important open question concerns the optimality of our candidate strategy. It cannot be directly linked to our various operators, but we are hopeful that further work will enable us to prove theoretically that it is close to optimality.

Appendix A. Stability of kernels R' and $\bar{R}'$

A.1. Kernel R' maps $\mathbb{X}'$ onto itself

Set $\xi \in \mathbb{X}'$ and $d \in \mathbb{K}'(\xi)$, and denote by $\mathbb{X}i$ a random variable with distribution $R'(\cdot|xi, d)$. If $d = \bar{d}$ then $\mathbb{X}i = \Delta \in \mathbb{X}'$. Otherwise, by definition one has $\xi = (\theta, y, w)$ such that $\theta(E_{m_0}) \geq (\underline{f}/\bar{f})^{\frac{w}{\delta} \vee 1} \mathrm{e}^{-w\|\lambda\|}$, and $d = (\ell, r) \in \mathbb{K}'(\xi)$. Then one has $\mathbb{X}i = (\Theta, Y, W)$ and, conditionally on Y, $\bar{\Theta} = \Psi(\theta, Y, w + r, d)$. Now, the definition of the filter operator Ψ and Assumption 3.6 yield

$$\begin{aligned}\Theta(E_{m_0}) &= \frac{\int_{E_M}\int_{E_M} f(Y - F(x'))\mathbb{1}_{E_{m_0}}(x')P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x)}{\int_{E_M}\int_{E_M} f(Y - F(x'))P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x)} \\ &\geq \left(\underline{f}/\bar{f}\right)\int_{E_{m_0}} P(E_{m_0}|x, d)\theta(\mathrm{d}x).\end{aligned}$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

By definition of kernel P, if $x \in E_{m_0}$ and $d = (\ell, r)$, the probability of being in E_{m_0} after a time r and starting from E_0 at time 0 is greater than the probability of not jumping at all between times 0 and r. Thus, one has

$$P(E_{m_0}|x, d) \geq \mathrm{e}^{-\mathbf{\Lambda}^{\ell}(x,r)} \geq \mathrm{e}^{-r\|\lambda\|},$$

thanks to Assumption 3.3. Then one has

$$\Theta(E_{m_0}) \geq \frac{\underline{f}\mathrm{e}^{-r\|\lambda\|}\theta(E_{m_0})}{\bar{f}} \geq \left(\underline{f}/\bar{f}\right)^{\frac{w+r}{\delta}\vee 1} \mathrm{e}^{-(w+r)\|\lambda\|},$$

as $\underline{f}/\bar{f} \leq 1$ and $r \geq \delta$.

A.2. Kernel $\bar{R}'$ maps $\bar{\mathbb{X}}'$ onto itself

Set $\bar{\xi} \in \bar{\mathbb{X}}'$ and $d \in \mathbb{K}'(\xi)$, and denote by $\mathbb{X}i$ a random variable with distribution $\bar{R}'(\cdot|\bar{\xi}, d)$. If $d = \check{d}$ then $\mathbb{X}i = \Delta \in \bar{\mathbb{X}}'$. Otherwise, by definition one has $\bar{\xi} = (\bar{\theta}, y, w)$ such that $\bar{\theta}(\Omega_{m_0}) \geq (\underline{f}/\bar{f})^{\frac{w}{\delta}\vee 1}\mathrm{e}^{-w\|\lambda\|}$, and $d = (\ell, r) \in \mathbb{K}'(\xi)$. In this case, one has $\mathbb{X}i = (\Theta, Y, W)$ and, conditionally on Y, $\Theta = \bar{\Psi}(\bar{\theta}, Y, w + r, d)$, and

$$\begin{aligned}\Theta(\Omega_{m_0}) &= \frac{\sum_{\omega^i\in\Omega}\sum_{\omega^j\in\Omega_{m_0}} f(Y - F(\omega^j))\bar{P}(\omega^j|\omega^i, d)\bar{\theta}(\omega^i)}{\sum_{\omega^i\in\Omega}\sum_{\omega^k\in\Omega} f(Y - F(\omega^k))\bar{P}(\omega^k|\omega^i, d)\bar{\theta}(\omega^i)} \\ &\geq \frac{\underline{f}}{\bar{f}} \sum_{\omega^i\in\Omega_{m_0}} \bar{P}(\Omega_{m_0}|\omega^i, d)\bar{\theta}(\omega^i),\end{aligned}$$

by Assumption 3.6. again. A similar argument as in the previous section directly yields $\bar{P}(\Omega_{m_0}|\omega^i, d) = P(E_0|x, d) \geq \mathrm{e}^{-r\|\lambda\|}$ as the Voronoi cells form a partition of E_M that preserves the mode, and

$$\Theta(\Omega_{m_0}) \geq \frac{\underline{f}\mathrm{e}^{-r\|\lambda\|}\bar{\theta}(\Omega_{m_0})}{\bar{f}} \geq \left(\underline{f}/\bar{f}\right)^{\frac{w+r}{\delta}\vee 1} \mathrm{e}^{-(w+r)\|\lambda\|},$$

hence the result.

Appendix B. Error bounds for the first discretization

The proof of Theorem 3.2 is split into a number of intermediate propositions and lemmas stated in the sequel. Most of the proofs rely on adequate splitting of expressions into simpler terms, and exploitation of Lipschitz regularity. As they are sometimes long and partly straightforward, some details are omitted. We highlight the main ideas and point out where we use our specific assumptions.

B.1. Function spaces

We start by defining some function spaces in order to evaluate the regularity of our operators. Denote by $BL(\mathbb{X})$ the set of real-valued, bounded measurable functions h on $\mathbb{X}$ that are Lipschitz continuous and set

$$\|h\| = \sup_{\xi\in\mathbb{X}} \|h(\xi)\|, \quad [h] = \sup_{m\in M}\sup_{\gamma\in\mathbb{O}}\sup_{\mathrm{x}\neq\mathrm{x}'\in E_m} \frac{|h(m, \mathrm{x}, \gamma) - h(m, \mathrm{x}', \gamma)|}{\|\mathrm{x} - \mathrm{x}'\|}.$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Note also $BL_1(\mathbb{X}) = \{h \in BL(\mathbb{X});\ \|h\| + [h] \le 1\}$, and $BL(E)$ the restriction of $BL(\mathbb{X})$ to functions defined on E_M. Denote by $BL(\mathbb{X}, \bar{\delta})$ the set of real-valued, bounded measurable functions w on $\mathbb{X} \times [0, \bar{\delta}]$ that are Lipschitz continuous and set

$$\|w\| = \sup_{(\xi,t)\in\mathbb{X}\times\mathbb{R}} \|w(\xi, t)\|, \quad [w] = \sup_{m\in M} \sup_{\gamma\in\mathbb{O}} \sup_{\mathtt{x}\neq\mathtt{x}'\in E_m\times\mathbb{R}} \frac{|w(m, \mathtt{x}, \gamma, t) - w(m, \mathtt{x}', \gamma, t')|}{\|\mathtt{x} - \mathtt{x}'\| + |t - t'|}.$$

Note also $BL(E, \bar{\delta})$ the restriction of $BL(\mathbb{X})$ to functions defined on $E_M \times \mathbb{R}$. For θ and $\bar{\theta}$ two probability measures in $\mathcal{P}(E_M)$, define the distance $d_E(\theta, \bar{\theta})$ by

$$d_E(\theta, \bar{\theta}) = \sup_{h\in BL_1(\mathbb{X})} \sup_{\gamma\in\mathbb{O}} \left| \int h(x, \gamma)\theta(\mathrm{d}x) - \int h(x, \gamma)\bar{\theta}(\mathrm{d}x) \right|.$$

Let $BLP(\mathbb{X}')$ be the set of real-valued, bounded measurable functions φ on $\mathbb{X}'$ that are Lipschitz continuous and set

$$\|\varphi\| = \sup_{\xi\in\mathbb{X}'} \|\varphi(\xi)\|, \quad [\varphi] = \sup_{(\theta,\gamma)\neq(\theta',\gamma)\in\mathbb{X}'-\{\Delta\}} \frac{|\varphi(\theta, \gamma) - \varphi(\bar{\theta}, \gamma)|}{d_E(\theta, \bar{\theta})}.$$

B.2. Regularity of operator P

The generic form of the transition kernels P of the skeleton chains with time span r, in the special case where at most two jumps can occur before time r, is

$$Ph(x, d) = I^\ell[h](x, r) + J^\ell\left[I^\ell[h]\right](x, r) + J^\ell\left[J^\ell\left[H^\ell[h]\right]\right](x, r),$$

where operator H corresponds to no jump and no more possible jumps, operator I to no jump in the case where more jumps can still occur, and operator J to one random jump. For any bounded measurable functions h on E_M and w on $E_M \times [0, H]$, all $\ell \in L$, $x \in E_M$, and $0 \le t \le \bar{\delta}$, these operators are defined as

$$H^\ell[h](x, t) = h(\mathbf{\Phi}^\ell(x, t)),$$

$$I^\ell[h](x, t) = h(\mathbf{\Phi}^\ell(x, t))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x,t)},$$

$$J^\ell[w](x, t) = \int_0^t \int_{E_M} w(x', t - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x, s))\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x,s)}\,\mathrm{d}s.$$

If more jumps are allowed, the iterates are longer and the inner term is $I^\ell[h]$ for all iterates except the last one, and $H^\ell[h]$ for the last one: for at most N_j jumps, one has

$$Ph(x, d) = \sum_{k=0}^{N_j-1} (J^\ell)^k\left[I^\ell[h]\right](x, r) + (J^\ell)^{N_j}\left[H^\ell[h]\right](x, r).$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

The sequel is detailed in the case of $N_j = 2$ but can be readily extended to any other finite value. To ensure the regularity of P, one thus needs that the three operators above are sufficiently regular, and that this property is preserved through iterations.

Lemma B.1 . *Under Assumption 3.1, for all $\ell \in L$, operator H^ℓ sends BL(E) onto BL(E, $\overline{\delta}$): for all functions $h \in BL(E)$, for all $x^1, x^2 \in E_M$, $t^1, t^2 \in [0, \overline{\delta}]$, one has*

$$|H^\ell[h](x^1, t^1)| \leq \|h\|,$$
$$|H^\ell[h](x^1, t^1) - H^\ell[h](x^2, t^2)| \leq [h][\Phi](\|x^1 - x^2\| + |t^1 - t^2|).$$

Set $C_H = [\Phi]$.

Proof. The first statement is obvious. Let us turn to the Lipschitz regularity. For all functions $h \in BL(E)$, one has

$$\begin{aligned}|H^\ell[h](x^1, t^1) - H^\ell[h](x^2, t^2)| &= |h(\mathbf{\Phi}^\ell(x^1, t^1)) - h(\mathbf{\Phi}^\ell(x^2, t^2))| \\ &\leq [h][\Phi](\|x^1 - x^2\| + |t^1 - t^2|),\end{aligned}$$

by Assumption 3.1. □

Lemma B.2. *Under Assumptions 3.1–3.3, for all $\ell \in L$, operator I^ℓ sends BL(E) onto BL(E, $\overline{\delta}$): for all functions $h \in BL(E)$, for all $x^1, x^2 \in E_M$, $t^1, t^2 \in [0, \overline{\delta}]$, one has*

$$|I^\ell[h](x^1, t^1)| \leq \|h\|,$$
$$|I^\ell[h](x^1, t^1) - I^\ell[h](x^2, t^2)| \leq ([h][\Phi] + \|h\|\overline{\delta}[\lambda])\|x^1 - x^2\| + ([h][\Phi] + \|h\|\|\lambda\|)|t^1 - t^2|.$$

Set $C_I = [\Phi] + \overline{\delta}[\lambda] + [\Phi] + \|\lambda\|$.

Proof. The first statement is obvious. Let us turn to the Lipschitz regularity. For all functions $h \in BL(E)$, one has

$$\begin{aligned}&|I^\ell[h](x^1, t^1) - I^\ell[h](x^2, t^2)| \\ &\quad = |h(\mathbf{\Phi}^\ell(x^1, t^1))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,t^1)} - h(\mathbf{\Phi}^\ell(x^2, t^2))e^{-\mathbf{\Lambda}^\ell(x^2,t^2)}| \\ &\quad \leq |H^\ell[h](x^1, t^1) - H^\ell[h](x^2, t^2)|\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,t^1)} + |H^\ell[h](x^2, t^2)|\left|\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,t^1)} - \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^2,t^2)}\right| \\ &\quad \leq [h][\Phi](\|x^1 - x^2\| + |t^1 - t^2|) + \|h\|(\overline{\delta}[\lambda]\|x^1 - x^2\| + \|\lambda\||t^1 - t^2|)\end{aligned}$$

by Lemma B.1 and Assumptions 3.1–3.3. □

Lemma B.3. *Under Assumptions 3.1–3.4, for all $\ell \in L$, operator J^ℓ sends BL(E, $\overline{\delta}$) onto itself: for all functions $w \in BL(E, \overline{\delta})$, for all $x^1, x^2 \in E_M$, $t^1, t^2 \in [0, \overline{\delta}]$, one has*

$$|J^\ell[w](x^1, t^1)| \leq \|w\|,$$
$$|J^\ell[w](x^1, t^1) - J^\ell[w](x^2, t^2)| \leq ([Q][w][\Phi] + \|w\|\overline{\delta}([\lambda] + \overline{\delta}\|\lambda\|))\|x^1 - x^2\| + 2[w]|t^1 - t^2|.$$

Set $C_J = [Q][\Phi] + \overline{\delta}([\lambda] + \overline{\delta}\|\lambda\|) + 2$.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Proof. The first statement is obvious. Let us turn to the Lipschitz regularity. Suppose without loss of generality that $t^1 \le t^2$. For all functions $w \in BL(E, \overline{\delta})$, one has

$$\begin{aligned}
&|J^\ell[w](x^1, t^1) - J^\ell[w](x^2, t^2)| \\
&\quad = \Big| \int_0^{t^1} \int_{E_M} w(x', t^1 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^1, s))\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)}\,\mathrm{d}s \\
&\qquad - \int_0^{t^2} \int_{E_M} w(x', t^2 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^2, s))\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^2, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^2,s)}\,\mathrm{d}s\Big| \\
&\quad \le \Big| \int_0^{t^1} \Big| \int_{E_M} w(x', t^1 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^1, s))\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)} \\
&\qquad - \int_{E_M} w(x', t^2 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^2, s))\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^2, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^2,s)}\Big|\,\mathrm{d}s \\
&\qquad + \int_{t^1}^{t^2} \int_{E_M} |w(x', t^2 - s)|\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^2, s))\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^2, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^2,s)}\,\mathrm{d}s \\
&\quad \le A + B.
\end{aligned}$$

The second term B is readily bounded by $|t^1 - t^2|\|w\|$. Let us turn to the first term. One has

$$\begin{aligned}
A \le &\int_0^{t^1} \Big| \int_{E_M} w(x', t^1 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^1, s)) - \int_{E_M} w(x', t^2 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^2, s))\Big| \\
&\times \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)}\,\mathrm{d}s \\
&+ \int_0^{t^1} \int_{E_M} |w(x', t^2 - s)|\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^2, s))\Big|\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)} - \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^2, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^2,s)}\Big|\,\mathrm{d}s \\
\le &A_1 + A_2.
\end{aligned}$$

For term A_1, one has

$$\begin{aligned}
A_1 \le &\int_0^{t^1} \int_{E_M} |w(x', t^1 - s) - w(x', t^2 - s)|\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^1, s))\boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)}\,\mathrm{d}s \\
&+ \int_0^{t^1} \Big| \int_{E_M} w(x', t^2 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^1, s)) - \int_{E_M} w(x', t^2 - s)\mathbf{Q}^\ell(\mathrm{d}x'|\mathbf{\Phi}^\ell(x^2, s))\Big| \\
&\times \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)}\,rmds \\
\le &A_{11} + A_{12}.
\end{aligned}$$

On the one hand, one has

$$A_{11} \le [w]|t^1 - t^2| \int_0^{t^1} \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s))\mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)}\,\mathrm{d}s \le [w]|t^1 - t^2|.$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

On the other hand, for fixed t^2, s, the mapping $h_{t^2,s} : x \mapsto w(x, t^2 - s)$ is clearly in $BL(E)$, with $\|h_{t^2,s}\| \le \|w\|$ and $[h_{t^2,s}] \le [w]$. Hence, Assumption 3.4 yields

$$\begin{aligned} A_{12} &= \int_0^{t^1} \left| \mathbf{Q}^\ell h_{t^2,s}(\mathbf{\Phi}^\ell(x^1, s)) - \mathbf{Q}^\ell h_{t^2,s}(\mathbf{\Phi}^\ell(x^2, s)) \right| \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s)) \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)} \, \mathrm{d}s \\ &\le [Q][w] \int_0^{t^1} \left| \mathbf{\Phi}^\ell(x^1, s) - \mathbf{\Phi}^\ell(x^2, s) \right| \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s)) \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)} \, \mathrm{d}s \\ &\le [Q][w][\Phi] \|x^1 - x^2\| \int_0^{t^1} \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s)) \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)} \, \mathrm{d}s \le [Q][w][\Phi] \|x^1 - x^2\|, \end{aligned}$$

by Assumption 3.1. Let us now consider term A_2. One has

$$\begin{aligned} A_2 &= \int_0^{t^1} \int_{E_M} |w(x', t^2 - s)| \mathbf{Q}^\ell(\mathrm{d}x' | \mathbf{\Phi}^\ell(x^2, s)) \left| \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s)) \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)} - \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^2, s)) \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^2,s)} \right| \mathrm{d}s \\ &\le \|w\| \int_0^{t^1} \left| \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^1, s)) \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^1,s)} - \boldsymbol{\lambda}^\ell(\mathbf{\Phi}^\ell(x^2, s)) \mathrm{e}^{-\mathbf{\Lambda}^\ell(x^2,s)} \right| \mathrm{d}s \\ &\le \|w\| \bar{\delta}[\lambda] \|x^1 - x^2\| + \|w\| \bar{\delta}^2 \|\lambda\| \|x^1 - x^2\|, \end{aligned}$$

by Assumptions 3.1, 3.2 and 3.3. The result follows. □

Theorem 1. *Under Assumptions 3.1–3.5, operator P sends BL(E) onto itself. More specifically, for all $d = (\ell, r) \in L \times \mathbb{T}$, for all functions $h \in BL_1(E)$, for all $x^1, x^2 \in E_M$, one has*

$$\begin{aligned} |Ph(x^1, d)| &\le \|h\|, \\ |Ph(x^1, d)) - Ph(x^2, d)| &\le C_P \|x^1 - x^2\|, \end{aligned}$$

where C_P is a constant depending only on the parameters of the problem.

Proof. The first statement is obvious as P is a Markov kernel. For all x^1, x^2 in E_M, one has

$$\begin{aligned} |Ph(x^1, d)) - Ph(x^2, d)| &\le |I^\ell[h](x^1, r) - I^\ell[h](x^2, r)| \\ &\quad + |J^\ell \left[I^\ell[h] \right](x^1, r) - J^\ell \left[I^\ell[h] \right](x^2, r)| \\ &\quad + |J^\ell \left[J^\ell \left[H^\ell[h] \right] \right](x^1, r) - J^\ell \left[J^\ell \left[H^\ell[h] \right] \right](x^1, r)(x^2, r)| \\ &\le A + B + C. \end{aligned}$$

Lemma B.2 applied to $h \in BL(E)$ yields

$$A = |I^\ell[h](x^1, r) - I^\ell[h](x^2, r)| \le ([h][\Phi] + \|h\| \bar{\delta}[\lambda]) \|x^1 - x^2\|.$$

The same lemma yields that the mapping $w : (x, t) \mapsto I^\ell[h](x, t)$ is in $BL(E, \bar{\delta})$, with $\|w\| \le \|h\|$, and $[w] \le C_I(\|h\| + [h])$. Then in turn, Lemma B.3 applied to w yields

$$\begin{aligned} B &= |J^\ell \left[I^\ell[h] \right](x^1, r) - J^\ell \left[I^\ell[h] \right](x^2, r)| \ = |J^\ell \left[w \right](x^1, r) - J^\ell \left[w \right](x^2, r)| \\ &\le ([Q][w][\Phi] + \|w\| \bar{\delta}([\lambda] + \bar{\delta} \|\lambda\|)) \|x^1 - x^2\|. \end{aligned}$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

The combination of Lemmas B.1 and B.3 also yields that the mapping $w' : (x, t) \mapsto J^\ell \left[H^\ell[h]\right](x, t)$ is still in $BL(E, \bar{\delta})$, with $\|w'\| \leq \|H^\ell[h]\| \leq \|h\|$ and

$$[w'] \leq C_J(\|H^\ell[h]\| + [H^\ell[h]]) \leq C_J(\|h\| + C_H(\|h\| + [h]) = (C_J + C_H)\|h\| + C_H[h].$$

Lemma B.3 for w' now yields

$$\begin{aligned} C &= |J^\ell \left[J^\ell \left[H^\ell[h]\right]\right](x^1, r) - J^\ell \left[J^\ell \left[H^\ell[h]\right]\right](x^1, r)(x^2, r)| \\ &= |J^\ell \left[w'\right](x^1, r) - J^\ell \left[w'\right](x^1, r)(x^2, r)| \\ &\leq ([Q][w'][\Phi] + \|w'\|\bar{\delta}([\lambda] + \bar{\delta}\|\lambda\|))\|x^1 - x^2\|, \end{aligned}$$

hence the result with $C_P \leq \|h\|([Q][\Phi](C_H + C_I + C_J) + 3\bar{\delta}[\lambda] + 2\bar{\delta}^2\|\lambda\|)) + [h][\Phi](1 + [Q](C_I + C_H))$. □

Note that further iterations of Lemma B.3 also yield that operator P sends $BL(E)$ onto itself for any bounded number of allowed jumps in the interval $[0, r]$. Note also that Theorem 1 is still valid on $BL(\mathbb{X})$.

B.3. Regularity of operator R'

We next need to investigate the regularity of the filter operator. The proof is based on the explicit form of the filter operator. The lower bounds on $\theta(E_{m_0})$ are required to bound the terms in the denominators.

Lemma B.4. *Under Assumptions 3.1–3.5, there exists some positive constant C_Ψ such that for all $w' \in [0, H]$, $d = (\ell, r) \in L \times \mathbb{T}$, and $\theta, \bar{\theta} \in \mathcal{P}(E_M)$ satisfying $\theta(E_{m_0}) > 0$, $\bar{\theta}(E_{m_0}) > 0$, we have*

$$\int_I d_E(\Psi(\theta, y', w', d), \Psi(\bar{\theta}, y', w', d))dy' \leq \frac{C_\Psi}{\theta(E_{m_0})\bar{\theta}(E_{m_0})} d_E(\theta, \bar{\theta}).$$

Proof. Let $g \in BL_1(\mathbb{X})$, let $y' \in I$, $w' \in [0, H]$, and let $\gamma \in \mathbb{O}$. We have

$$\begin{aligned} &d_E(\Psi(\theta, y', w', d), \Psi(\bar{\theta}, y', w', d)) \\ &\quad = \int_{E_M} g(x, \gamma)\Psi(\theta, y', w', d)(\mathrm{d}x) - \int_{E_M} g(x, \gamma)\Psi(\bar{\theta}, y', w', d)(\mathrm{d}x) \\ &\quad = \left(\int_{E_M^2} f(y' - F(x'))g(x', \gamma)P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x) \right. \\ &\qquad \left. - \int_{E_M^2} f(y' - F(x'))g(x', \gamma)P(\mathrm{d}x'|x, d)\bar{\theta}(\mathrm{d}x) \right) \frac{1}{\int_{E_M^2} f(y' - F(x'))P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x)} \\ &\qquad + \int_{E_M^2} f(y' - F(x'))g(x', \gamma)P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x) \\ &\qquad \times \left(\frac{1}{\int_{E_M^2} f(y' - F(x'))P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x)} - \frac{1}{\int_{E_M^2} f(y' - F(x'))P(\mathrm{d}x'|x, d)\bar{\theta}(\mathrm{d}x)} \right) \\ &\quad = \frac{A}{B} + C. \end{aligned}$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

First, as g is in $BL_1(\mathbb{X})$, then $\varphi : (x, \gamma) \mapsto \int_{E_M} f(y' - F(x'))g(x', \gamma)P(\mathrm{d}x'|x, d)$ is in $BL(\mathbb{X})$ with $\|\varphi\| \leq \bar{f}$ and $[\varphi] \leq (\bar{f} + L_Y)C_P$ thanks to Theorem 1, so that one has

$$|A| \leq \big(\bar{f}(1 + C_P) + L_Y C_P\big)d_E(\theta, \bar{\theta}).$$

On the other hand, one has

$$|B| \geq \underline{f} \int_{E_M} P(E_M|x, d)\theta(\mathrm{d}x) \geq \underline{f} \int_{E_{m_0}} P(E_{m_0}|x, d)\theta(\mathrm{d}x).$$

As seen in Section A.1, if $x \in E_0$ and $d = (\ell, r)$, one has

$$P(E_0|x, d) = \mathrm{e}^{-\mathbf{\Lambda}^{\ell}(x,r)} \geq \mathrm{e}^{-r\|\lambda\|} \geq \mathrm{e}^{-\bar{\delta}\|\lambda\|},$$

hence one has

$$|B| \geq \underline{f}\mathrm{e}^{-\bar{\delta}\|\lambda\|}\theta(E_{m_0}).$$

Similarly, the last term can be bounded by

$$|C| \leq \bar{f}\frac{\bar{f}(1 + C_P) + L_Y C_P}{\underline{f}^2\mathrm{e}^{-2\bar{\delta}\|\lambda\|}\theta(E_{m_0})\bar{\theta}(E_{m_0})}d_E(\theta, \bar{\theta}).$$

As I is a bounded interval, the result follows. □

We can now turn to the regularity of operator R'.

Proposition B.1. *Under Assumptions 3.1–3.6, there exists a positive constant $C_{R'}$ such that for all $g \in BLP_1(\mathbb{X}')$, $\gamma = (y, w) \in \mathbb{O}$, $d \in \mathbb{K}'(\gamma)$, and for all $\theta, \bar{\theta}$ in $\mathcal{P}(E_M)$ such that $(\theta, \gamma) \in \mathbb{X}'$, $(\bar{\theta}, \gamma) \in \mathbb{X}'$, one has*

$$\big|R'g(\theta, y, d) - R'g(\bar{\theta}, y, d)\big| \leq C_{R'}d_E(\theta, \bar{\theta}).$$

In particular, operator R' maps $BLP(\mathbb{X}')$ onto itself, and for $g \in BLP(\mathbb{X}')$ one has $\|R'g\| \leq \|g\|$ and $[R'g] \leq (\|g\| + [g])C_{R'}$.

Proof. If $d = \check{d}$, then $R'g(\theta, \gamma, \check{d}) = R'g(\bar{\theta}, \gamma, \check{d}) = g(\Delta)$, hence $|R'g(\theta, \gamma, \check{d}) - R'g(\bar{\theta}, \gamma, \check{d})| = 0$. If $d = (\ell, r) \neq \check{d}$, then by assumption one has $\theta(E_{m_0}) > 0$ and $\bar{\theta}(E_{m_0}) > 0$. In addition, one has $w < N\delta$, such that

$$R'g(\theta, \gamma, d) = \int_I \int_{E_M^2} g\big(\Psi(\theta, y', w + r, d), y', w + r\big)f(y' - F(x'))P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x)\,\mathrm{d}y'.$$

Set

$$h_\theta(x, \gamma) = \int_I g\big(\Psi(\theta, y', w + r, d), y', w + r\big)f(y' - F(x))dy',$$

if $x \in E_M$ and $\gamma = (y, w + r)$ and 0 otherwise, and $g_{\theta,d}(x, \gamma) = Ph_\theta(x, \gamma, d)$. From Assumption 3.6, h_θ is in $BL(\mathbb{X})$ with $\|h_\theta\| \leq B_f$ and $[h_\theta] \leq L_f$. Theorem 1 thus yields that $g_{\theta,d}$ in still in $BL(\mathbb{X})$ with $\|g_{\theta,d}\| \leq B_f$ and $[g_{\theta,d}] \leq (B_f + L_f)C_P$. Therefore, one has

$$\begin{aligned}
&|R'g(\theta, \gamma, d) - R'g(\bar{\theta}, \gamma, d)| \\
&\leq \int_{E_M} \big|g_{\theta,d}(x, \gamma) - g_{\bar{\theta},d}(x, \gamma)\big|\,\bar{\theta}(\mathrm{d}x) + \left|\int_{E_M} g_{\theta,d}(x, \gamma)\theta(\mathrm{d}x) - \int_{E_M} g_{\theta,d}(x, y)\bar{\theta}(\mathrm{d}x)\right| \\
&\leq \int_{E_M} \int_I d_E(\Psi(\theta, y', w + r, d), \Psi(\bar{\theta}, y', w + r, d))\,\mathrm{d}y'\bar{\theta}(\mathrm{d}x) + \big([g_{\theta,d}] + \|g_{\theta,d}\|\big)\,d_E(\theta, \bar{\theta}),
\end{aligned}$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

as $g \in BLP_1(\mathbb{X}')$. Note in particular that the assumptions on θ and $\bar{\theta}$ guarantee that the assumptions of Lemma B.4 are satisfied. Hence, we conclude using Lemma B.4 and the minoration of $\theta(E_{m_0})$ and $\bar{\theta}(E_{m_0})$ from the definition of $\mathbb{X}'$. □

B.4. Projection error for operators P and $\bar{P}$

For $\omega = (m, \mathrm{x}) \in \Omega$, we denote $\omega_1 = m$, $\omega_2 = \mathrm{x}$. For any function $h \in BL(\mathbb{X})$, $\xi = (x, \gamma) \in \Omega \times \mathbb{O}$, and $d \in \mathbb{K}(\xi)$, denote

$$\bar{P}h(\xi, d) = \int_{E_M} h(x', \gamma)\bar{P}(\mathrm{d}x'|x, d).$$

Proposition B.2. *For all $i \in \{1, \dots, n_\Omega\}$, $\gamma \in \mathbb{O}$, $d \in L \times \mathbb{T}$, and all $g \in BL_1(\mathbb{X})$, we have*

$$\left|Pg(\omega^i, \gamma, d) - \bar{P}g(\omega^i, \gamma, d)\right| \leq \sup_{j \in \{1,\dots,n_\Omega\}} \mathcal{D}_j.$$

Proof. For $x \in E$, denote by C_x the Voronoi cell of x. For $g \in BL_1(\mathbb{X})$, $i \in \{1, \dots, n_\Omega\}$, $\gamma = (y, w) \in \mathbb{O}$, $d \in L \times \mathbb{T}$, one has

$$\begin{aligned}
\left|Pg(\omega^i, \gamma, d) - \bar{P}g(\omega^i, \gamma, d)\right| &\leq \left| \int_{E_M} g(x', \gamma')P(\mathrm{d}x'|\omega^i, d) - \sum_{j=1}^{n_\Omega} g(\omega^j, \gamma')P(C_j|\omega^i, d)\right| \\
&\leq \sum_{j=1}^{n_\Omega} \int_{C_j} \left|g(x', \gamma') - g(\omega^j, \gamma')\right| P(\mathrm{d}x'|\omega^i, d) \\
&\leq \sum_{j=1}^{n_\Omega} \int_{C_j} |x' - \omega^j| P(\mathrm{d}x'|\omega^i, d) \leq \sup_{j \in \{1,\dots,n_\Omega\}} \mathcal{D}_j,
\end{aligned}$$

as g is Lipschitz-continuous in its first variable. □

B.5. Projection error for operators R' and $\bar{R}'$

We first need to evaluate the error between Ψ and $\bar{\Psi}$.

Lemma B.5. *Under Assumptions 3.1–3.6, there exists some positive constant $\bar{C}_\Psi$ such that for all $\gamma' = (y', w') \in \mathbb{O}$, $d = (\ell, r) \in L \times \mathbb{T}$, and all $\bar{\theta} \in \mathcal{P}(\Omega)$ satisfying $\bar{\theta}(\Omega_{m_0}) > 0$, we have*

$$\int_I d_E(\Psi(\bar{\theta}, \gamma', d), \bar{\Psi}(\bar{\theta}, \gamma', d))\, \mathrm{d}y' \leq \frac{\bar{C}_\Psi}{\bar{\theta}(\Omega_{m_0})} \sup_{j \in \{1,\dots,\ell\}} \mathcal{D}_j.$$

Proof. Set $\gamma \in \mathbb{O}$, $\gamma' = (y', w') \in \mathbb{O}$, $d = (\ell, r) \in L \times \mathbb{T}$, $\bar{\theta} \in \mathcal{P}(\Omega)$ satisfying $\bar{\theta}(\Omega_{m_0}) > 0$, and $g \in BL_1(\mathbb{X})$. We have

$$\begin{aligned}
&\int_{E_M} g(x, \gamma)\Psi(\bar{\theta}, y', w', d)(\mathrm{d}x) - \int_{E_M} g(x, \gamma)\bar{\Psi}(\bar{\theta}, y', w', d)(\mathrm{d}x) \\
&= \frac{\sum_{\omega^i \in \Omega} \int_{E_M} f(y' - F(x'))g(x', \gamma)P(\mathrm{d}x'|\omega^i, d)\bar{\theta}(\omega^i)}{\sum_{\omega^i \in \Omega} \int_{E_M} f(y' - F(x'))P(\mathrm{d}x'|\omega^i, d)\bar{\theta}(\omega^i)} \\
&\quad - \frac{\sum_{\omega^i \in \Omega} \sum_{\omega^j \in \Omega} f(y' - F(\omega^j))g(\omega^j, \gamma)\bar{P}(\omega^j|\omega^i, d)\bar{\theta}(\omega^i)}{\sum_{\omega^i \in \Omega} \sum_{\omega^k \in \Omega} f(y' - F(\omega^k))\bar{P}(\omega^k|\omega^i, d)\bar{\theta}(\omega^i)}.
\end{aligned}$$

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Using a similar splitting and similar arguments as in the proof of Lemma B.4 together with Proposition B.2 yields the expected result, as $\varphi : (x, \gamma) \mapsto \int_{E_M} f(y' - F(x'))g(x', \gamma)P(\mathrm{d}x'|x, d)$ is still in $BL(\mathbb{X})$ with $\|\varphi\| \leq \bar{f}$ and $[\varphi] \leq \left(\bar{f} + L_Y\right)C_P$ thanks to Theorem 1. □

Proposition B.3. *Under Assumptions 3.1–3.6, there exists a positive constant $\bar{C}_{R'}$ such that for all $g \in BLP_1(E)$, $(\theta, \gamma) \in \bar{\mathbb{X}}'$, and $d \in \mathbb{K}'(\gamma)$, one has*

$$\left|R'g(\bar{\theta}, \gamma, d) - \bar{R}'g(\bar{\theta}, \gamma, d)\right| \leq \bar{C}_{R'} \sup_{j\in\{1,\dots,\ell\}} \mathcal{D}_j.$$

Proof. This is a direct consequence of Propositions B.2 and B.5 with $\bar{C}_{R'} = L_f + 2B_f + L_f\bar{C}_\Psi$, as $\bar{\theta}(\Omega_{m_0})$ is uniformly bounded from below. □

B.6. Regularity and approximation error for the cost functions

The last preliminary result we need in order to prove Theorem 3.2 is to ensure that the cost functions c', C'' belong to the appropriate function spaces and the error between c' and $\bar{c}'$ is controlled.

Lemma B.6. *Under Assumption 3.7, the cost function C' is in $BLP(\mathbb{X}')$.*

Proof. Under Assumption 3.7, C is clearly in *BL(E)* with $\|C\| \leq B_C$ and $[C] \leq L_C$. Thus, C′ is still bounded by B_C, and for all (θ, γ) and $(\bar{\theta}, \gamma)$ in $\mathbb{X}'$ one has

$$|C'(\theta, \gamma) - C'(\bar{\theta}, \gamma)| = \left| \int_{E_M} C(x)\theta(\mathrm{d}x) - \int_{E_M} C(x)\bar{\theta}(\mathrm{d}x)\right| \leq (B_C + L_C)d_E(\theta, \bar{\theta}).$$

Thus, C′ is in $BLP(\mathbb{X})$ with $\|C'\| \leq B_C$ and $[C'] \leq B_C + L_C$. □

Lemma B.7. *Under Assumptions 3.1–3.5 and 3.7, for all $d \in \mathbb{A}$ the function $c'_d : \xi \mapsto c'(\xi, d)$ is in $BLP(\mathbb{X}')$.*

Proof. If $d = \check{d}$ then $c' = C'$ and the result is true. Otherwise, under Assumption 3.7, c'_d is bounded by B_c, and for fixed x, application $c_{x,d} : x' \mapsto c(x, d, x')$ is in *BL(E)* with $\|c_{x,d}\| \leq B_c$ and $[c_{x,d}] \leq L_c$. Let us study the application $Pc : x \mapsto \int_{E_M} c(x, d, x')P(\mathrm{d}x'|x, d)$ for fixed $d \in L \times \mathbb{T}$. It is still bounded by B_c, and for x^1, x^2 in E, one has

$$\begin{aligned}|Pc(x^1) - Pc(x^2)| &\leq \int_{E_M} |c(x^1, d, x') - c(x^2, d, x')|P(\mathrm{d}x'|x^1, d) \\ &\quad + \left| \int_{E_M} c(x^2, d, x')P(\mathrm{d}x'|x^1, d) - \int_{E_M} c(x^2, d, x')P(\mathrm{d}x'|x^2, d)\right| \\ &\leq L_c\|x^1 - x^2\| + C_P(B_c + L_c)\|x^1 - x^2\|\end{aligned}$$

by applying Theorem 1 to $c_{x^2,d}$. Hence, application Pc is still in BL(E) hence in $BL(\mathbb{X})$. Then for all (θ, γ) and $(\bar{\theta}, \gamma)$ in $\mathbb{X}'$ one obtains

$$\begin{aligned}|c'_d(\theta, \gamma) - c'_d(\bar{\theta}, \gamma)| &= \left| \int_{E_M^2} c(x, d, x')P(\mathrm{d}x'|x, d)\theta(\mathrm{d}x) - \int_{E_M^2} c(x, d, x')P(\mathrm{d}x'|x, d)\bar{\theta}(\mathrm{d}x)\right| \\ &\leq (B_c + L_c)(C_P + 1)d_E(\theta, \bar{\theta}).\end{aligned}$$

Hence the result. □

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Lemma B.8. *Under Assumption 3.7, for all $\xi \in \bar{\mathbb{X}}'$, $d \in \mathbb{K}'(\xi)$, one has*

$$|c'(\xi, d) - \bar{c}'(\xi, d)| \leq (B_c + L_c) \sup_{j\in\{1,\dots,\ell\}} \mathcal{D}_j.$$

Proof. The result the follows directly form Proposition B.2 and the facts that $x \mapsto c(\omega^i, d, x)$ is in BL(E) and $c'(\cdot, \check{d}) = \bar{c}'(\cdot, \check{d})$. □

B.7. Proof of Theorem 3.2

We establish the result by (backward) induction on n, with the additional statement that v'_n is in $BLP(\mathbb{X}')$ for all n.

For $n = N$, $v'_N = C'$ is in $BLP(\mathbb{X}')$ by Lemma B.6. In addition, set ξ in $\bar{\mathbb{X}}'$. Then by definition one has

$$|v'_N(\xi) - \bar{v}'_N(\xi)| = |C'(\xi) - C'(\xi)| = 0.$$

Suppose the result holds true for some $n + 1 \leq N$. By the induction hypothesis, v'_{n+1} is in $BLP(\mathbb{X}')$, thus $R'v'_{n+1}$ is also in $BL(\mathbb{X}')$ by Proposition B.1. For all $d \in \mathbb{A}$, c'_d is in $BLP(\mathbb{X}')$ by Lemma B.6. As $\mathbb{K}(\xi = (\theta, \gamma)) = \mathbb{K}(\gamma)$ depends only on γ and $BLP(\mathbb{X}')$ is clearly stable by finite maximum, it follows that v'_n is also in $BLP(\mathbb{X}')$.

Set ξ in $\bar{\mathbb{X}}'$. If $\xi = \Delta$, recall that $\mathbb{K}'(\Delta) = \{\check{d}\}$, so that on the one hand one has

$$v'_n(\xi) = c'(\xi, \check{d}) + R'v'_{n+1}(\xi, \check{d}) = 0 + v'_{n+1}(\Delta),$$

and on the other hand one has

$$\bar{v}'_n(\xi) = \bar{c}'(\xi, \check{d}) + \bar{R}'\bar{v}'_{n+1}(\xi, \check{d}) = 0 + \bar{v}'_{n+1}(\Delta).$$

Hence, one has $|v'_n(\Delta) - \bar{v}'_n(\Delta)| = |v'_{n+1}(\Delta) - \bar{v}'_{n+1}(\Delta)| = 0$ by induction.

If $\xi = (\bar{\theta}, \gamma) \in \bar{\mathbb{X}}' - \{\Delta\}$, with $\gamma = (y, w)$, then by definition one has $\theta(\Omega_{m_0}) \geq (\underline{f}/\bar{f})^{\frac{w}{\delta}\vee 1} \mathrm{e}^{-w\|\lambda\|}$. The dynamic programming equations yield

$$\begin{aligned}
|v'_n(\bar{\theta}, \gamma) - \bar{v}'_n(\bar{\theta}, \gamma)| &\leq \max_{d\in\mathbb{K}'(\xi)} |c'(\xi, d) - \bar{c}'(\xi, d)| + \max_{d\in\mathbb{K}'(\xi)} |R'v'_{n+1}(\xi, d) - \bar{R}'v'_{n+1}(\xi, d)| \\
&\quad + \max_{d\in\mathbb{K}'(\xi)} |\bar{R}'v'_{n+1}(\xi, d) - \bar{R}'\bar{v}'_{n+1}(\xi, d)| \\
&= A_1 + A_2 + A_3.
\end{aligned}$$

The fist term A_1 is bounded thanks to Lemma B.8 by $(B_c + L_c) \sup_{j\in\{1,\dots,\ell\}} \mathcal{D}_j$.

The second term A_2 is bounded by Proposition B.3 as v'_n is in $BLP(\mathbb{X}')$:

$$A_2 \leq (\|v'_{n+1}\| + [v'_{n+1}])\bar{C}_{R'}\big(\underline{f}^{-1}\bar{f}\big)^{N\vee 1} \mathrm{e}^{N\delta\|\lambda\|} \sup_{j\in\{1,\dots,\ell\}} \mathcal{D}_j.$$

The last term A_3 is bounded by the induction hypothesis and using the fact that $\bar{R}'$ is a Markov kernel:

$$A_3 \leq \sup_{\xi'\in\bar{\mathbb{X}}'} |v'_{n+1}(\xi') - \bar{v}'_{n+1}(\xi')| \leq \|v'_{n+1}\| \sup_{j\in\{1,\dots,\ell\}} \mathcal{D}_j.$$

Hence the result.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Appendix C. Error bounds for the second discretization

The proof of Theorem 3.3 follows similar lines to that of Theorem 3.2 Again, it is based on Lipschitz regularity properties of the operators involved, thus proofs are omitted. We introduce additional function spaces. Let $BLP(\bar{\mathbb{X}}')$ be the set of real-valued, bounded measurable functions φ on $\bar{\mathbb{X}}'$ that are Lipschitz continuous and set

$$\|\varphi\| = \sup_{\xi\in\bar{\mathbb{X}}'} \|\varphi(\xi)\|, \quad [\varphi] = \sup_{(\theta,\gamma)\neq(\theta',\gamma)\in\bar{\mathbb{X}}'-\{\Delta\}} \frac{|\varphi(\theta,\gamma)-\varphi(\bar{\theta},\gamma)|}{d_E(\theta,\bar{\theta})}.$$

Denote also $BLP_1(\bar{\mathbb{X}}') = \{\varphi \in BLP(\bar{\mathbb{X}}'): \|\varphi\|_{E,\mathcal{P}} + [\varphi]_{E,\mathcal{P}} \leq 1\}$.

C.1. Regularity of operator $\bar{R}'$

We first need to investigate the regularity of the approximate filter operator.

Lemma C.1. *Under Assumption 3.6, there exists some positive constant $C_{\bar{\Psi}}$ such that for all $\gamma' = (y', w') \in \mathbb{O}$, $d = (\ell, r) \in L \times \mathbb{T}$, and all $\theta, \bar{\theta} \in \mathcal{P}(\Omega)$ satisfying $\theta(\Omega_{m_0}) > 0$, $\bar{\theta}(\Omega_{m_0}) > 0$, we have*

$$\int_I d_\Omega(\bar{\Psi}(\theta, y', z', d), \bar{\Psi}(\bar{\theta}, y', z', d))\, \mathrm{d}y' \leq \frac{C_{\bar{\Psi}}}{\theta(\Omega_{m_0})\bar{\theta}(\Omega_{m_0})} d_\Omega(\theta, \bar{\theta}).$$

We can now turn to the regularity of operator $\bar{R}'$.

Proposition C.1. *Under Assumption 3.6, there exists a positive constant $C_{\bar{R}'}$ such that for all $g \in BLP_1(\bar{\mathbb{X}}')$, for all $\theta, \bar{\theta}$ in $\mathcal{P}(\Omega)$, $(y, \bar{y}, w) \in I^2 \times \{0, 1\} \times [0, H]$, and $d = (\ell, r) \in \mathbb{A} - \{\check{d}\}$, such that $(\theta, y, w) \in \bar{X}'$ and $(\bar{\theta}, \bar{y}, w) \in \bar{X}'$, one has*

$$\left|\bar{R}'g(\theta, y, z, d) - \bar{R}'g(\bar{\theta}, \bar{y}, z, d)\right| \leq C_{\bar{R}'} d_\Omega(\theta, \bar{\theta}).$$

In particular, operator $\bar{R}'$ maps $BLP(\bar{\mathbb{X}}')$ onto itself, and for $g \in BLP(\bar{\mathbb{X}}')$ one has $\|\bar{R}'g\| \leq \|g\|$ and $[\bar{R}'g] \leq (\|g\| + [g])C_{\bar{R}'}$.

Proposition C.2. *For all $g \in BLP_1(\bar{\mathbb{X}}')$ and $1 \leq j \leq n_\Gamma$, one has*

$$\left|\bar{R}'g(\rho^j, d) - \hat{R}'g(\rho^j, d)\right| \leq \sup_{\{1\leq j\leq n_\Gamma\}} \bar{\mathcal{D}}_j.$$

Proof. One has

$$\begin{aligned}\left|\bar{R}'g(\rho^j, d) - \hat{R}'g(\rho^j, d)\right| &= \left|\int_{\bar{\mathbb{X}}'} g(\xi)\bar{R}'(\mathrm{d}\xi|\rho^j, d) - \sum_{k=1}^{n_\Gamma} g(\rho^k)\bar{R}'(\bar{C}_k|\rho^j, d)\right| \\ &\leq \sum_{k=1}^{n_\Gamma} \int_{\bar{C}_k} \left|g(\xi) - g(\rho^k)\right| \bar{R}'(\mathrm{d}\xi|\rho^j, d) \\ &\leq \sum_{k=1}^{n_\Gamma} \int_{\bar{C}_k} d(\xi, \rho^k)\bar{R}'(\mathrm{d}\xi|\rho^j, d) \leq \sup_{\{1\leq k\leq n_\Gamma\}} \bar{\mathcal{D}}_k,\end{aligned}$$

hence the result. □

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

C.2. Regularity of the cost functions

We finally need to check that the cost functions c', C'' also belong to $BLP(\bar{\mathbb{X}}')$.

Lemma C.2 *Let g be a function from $\mathbb{X}'$ onto $\mathbb{R}$ belonging to $BLP(\mathbb{X})$. Then the restriction of g to $\bar{X}'$ is in $BLP(\bar{\mathbb{X}}')$.*

In particular, the restrictions of C' and $\bar{c}'$ to $\bar{X}'$ belong to $BLP(\bar{\mathbb{X}}')$.

C.3. Proof of Theorem 3.3

We establish the result by (backward) induction on k, with the additional statement that $\bar{v}'_k$ is in $BLP(\bar{\mathbb{X}}')$ for all k.

For $k = N$ and $1 \leq j \leq n_\Gamma$, one has

$$\left|\bar{v}'_N(\rho^j) - \hat{v}'(\rho^j)\right| = \left|C'(\rho^j) - C'(\rho^j)\right| = 0,$$

by definition. In addition, $\bar{v}' = C'$ is in $BLP(\bar{\mathbb{X}}')$.

Suppose the result holds true for some $k+1 \leq N$. By the induction hypothesis, $\bar{v}'_{n+1}$ is in $BLP(\bar{\mathbb{X}}')$, thus $\bar{R}'v'_{n+1}$ is also in $BLP(\bar{\mathbb{X}}')$ by Proposition C.1. For all $d \in \mathbb{A}$, $\bar{c}'_d$ is in $BLP(\bar{\mathbb{X}}')$. As $BLP(\bar{\mathbb{X}}')$ is clearly stable by finite maximum, it follows that $\bar{v}'_n$ is also in $BLP(\bar{\mathbb{X}}')$. Now, the dynamic programming equations yield

$$\begin{aligned}\left|\bar{v}'_k(\rho^j) - \hat{v}'_k(\rho^j)\right| \leq & \max_{d\in\mathbb{K}'(\rho^j)} \left|\bar{R}'\bar{v}'_{k+1}(\rho^j, d) - \hat{R}'\bar{v}'_{k+1}(\rho^j, d)\right| \\ & + \max_{d\in\mathbb{K}'(\rho^j)} \left|\hat{R}'\bar{v}'_{k+1}(\rho^j, d) - \hat{R}'\hat{v}'_{k+1}(\rho^j, d)\right|.\end{aligned}$$

The last term on the right-hand side is smaller than $\max_{1\leq j\leq n_\Gamma} |\bar{v}'_{k+1}(\rho^j) - \hat{v}'_{k+1}(\rho^j)|$ as $\hat{R}'$ is a Markov kernel. The first term on the right-hand side is bounded by Proposition C.2 as $\bar{v}'_n$ is in $BLP(\bar{\mathbb{X}}')$,

$$\max_{d\in\mathbb{K}'(\rho^j)} |\bar{R}'\bar{v}'_{k+1}(\rho^j, d) - \hat{R}'\bar{v}'_{k+1}(\rho^j, d)| \leq C_{\bar{R}'}(\|\bar{v}'_{k+1}\| + [\bar{v}'_{k+1}]) \sup_{j\in\{1,\ldots,n_\Gamma\}} \bar{\mathcal{D}}_j,$$

which concludes the proof.

Funding information

This work was partially supported by ANR Project HSSM-INCA (ANR-21-CE40-0005), by the European Union Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement no. 890462) and has been realized with the support of MESO@LR-Platform at the University of Montpellier.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article

Data

All code related to the simulations found in Section 5 can be found at https://github.com/acleynen/PDMP-control.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/apr.2025.10044.

References

[1] ALMUDEVAR, A. (2001). A dynamic programming algorithm for the optimal control of piecewise deterministic markov processes. *SIAM J. Control Optim.* **40**, 525–539.

[2] BÄUERLE, N. AND LANGE, D. (2018). Optimal control of partially observable piecewise deterministic Markov processes. *SIAM J. Control Optim.* **56**, 1441–1462.

[3] BÄUERLE, N. AND RIEDER, U. (2011). *Markov Decision Processes with Applications to Finance*. *Universitext*. Springer, Berlin.

[4] BAYSSE, C., BIHANNIC, D., GÉGOUT-PETIT, A., PRENAT, M., DE SAPORTA, B. AND SARACCO, J. (2013). Maintenance optimisation of optronic equipment. *Chem. Eng. Trans.* **33**, 709–714.

[5] BAYSSE, C., BIHANNIC, D., GÉGOUT-PETIT, A., PRENAT, M. AND SARACCO, J. (2014). Hidden Markov model for the detection of a degraded operating mode of optronic equipment. *J. Soc. Fr. Stat.* **155**, 48–61.

[6] BRANDEJSKY, A., DE SAPORTA, B. AND DUFOUR, F. (2013). Optimal stopping for partially observed piecewise-deterministic Markov processes. *Stochastic Process. Appl.* **123**, 3201–3238.

[7] CALVIA, A. (2020). Stochastic filtering and optimal control of pure jump markov processes with noise-free partial observation. *ESAIM: Control Optim. Calc. Var.* **26**, 25.

[8] CLEYNEN, A. AND DE SAPORTA, B. (2018). Change-point detection for piecewise deterministic Markov processes. *Automatica* **97**, 234–247.

[9] COCOZZA-THIVENT, C. *et al.*et al. (2021). *Markov Renewal and Piecewise Deterministic Processes*. Springer, Switzerland.

[10] COSTA, M. (2016). A piecewise deterministic model for a prey-predator community. *Ann. Appl. Probab.* **26**, 3491–3530.

[11] COSTA, O. AND DUFOUR, F. (2013). *Continuous Average Control of Piecewise Deterministic Markov Processes*. *Springer Briefs in Mathematics*. Springer, New York.

[12] COSTA, O. AND RAYMUNDO, C. (2000). Impulse and continuous control of piecewise deterministic Markov processes. *Stochastics Stochastic Rep.* **70**, 75–107.

[13] COSTA, O. L. AND DAVIS, M. (1989). Impulse control of piecewise-deterministic processes. *Math. Control Signals Syst.* **2**, 187–206.

[14] DAVIS, M. (1984). Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B* **46**, 353–388. With discussion.

[15] DAVIS, M. (1993). *Markov Models and Optimization*. *Monographs on Statistics and Applied Probability*, Vol. 49. Chapman & Hall, London.

[16] DE SAPORTA, B. AND DUFOUR, F. (2012). Numerical method for impulse control of piecewise deterministic Markov processes. *Automatica* **48**, 779–793.

[17] DE SAPORTA, B., DUFOUR, F. AND GEERAERT, A. (2017). Optimal strategies for impulse control of piecewise deterministic Markov processes. *Automatica* **77**, 219–229.

[18] DE SAPORTA, B., DUFOUR, F. AND ZHANG, H. (2016). *Numerical Methods for Simulation and Optimization of Piecewise Deterministic Markov Processes*. *Mathematics and Statistics Series*. ISTE, London; John Wiley & Sons, Hoboken, NJ. Application to reliability.

[19] DE SAPORTA, B., DUFOUR, F. AND ZHANG, H. (2019). Dynamic optimization of maintenance policies for multistate system. In *ESREL 2019: 29th European Safety and Reliability Conference*.

[20] DEMPSTER, M. AND YE, J. (1995). Impulse control of piecewise deterministic Markov processes. *Ann. Appl. Probab.* **5**, 399–423.

[21] DEVOOGHT, J. (1997). Dynamic reliability. In *Advances in Nuclear Science and Technology*, pp. 215–278. Springer, Boston.

[22] DOUMIC, M., HOFFMANN, M.and KRELL, N. AND Robert, L. (2015). Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli* **21**, 1760–1799.

[23] DUFOUR, F., HORIGUCHI, M. AND PIUNOVSKIY, A. (2016). Optimal impulsive control of piecewise deterministic Markov processes. *Stochastics* **88**, 1073–1098.

[24] FRITSCH, C., HARMAND, J. AND CAMPILLO, F. (2015). A modeling approach of the chemostat. *Ecol. Model.* **299**, 1–13.

[25] GUO, X., HERNÁNDEZ-LERMA, O., GUO, X. AND HERNÁNDEZ-LERMA, O. (2009). *Continuous-Time Markov Decision Processes*. Springer, Berlin.

[26] HERBACH, U., BONNAFFOUX, A., ESPINASSE, T. AND GANDRILLON, O. (2017). Inferring gene regulatory networks from single-cell data: a mechanistic approach. *BMC Syst. Biol.* **11**, 105.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press

[27] LANGE, D. K. (2017). Cost optimal control of piecewise deterministic Markov processes under partial observation. *PhD thesis*, Karlsruher Institut für Technologie (KIT), Karlsruhe.
[28] MAKIS, V. AND JIANG, X. (2003). Optimal replacement under partial observations. *Math. Oper. Res.* **28**, 382–394.
[29] MARSEGUERRA, M. AND ZIO, E. (1996). Monte Carlo approach to PSA for dynamic process systems. *Reliab. Eng. Syst. Safety* **52**, 227–241.
[30] PASIN, C., DUFOUR, F., VILLAIN, L., ZHANG, H. AND THIÉBAUT, R. (2018). Controlling IL-7 injections in HIV-infected patients. *Bull. Math. Biol.* **80**, 2349–2377.
[31] PIUNOVSKIY, A., PLAKHOV, A., TORRES, D. F. AND ZHANG, Y. (2019). Optimal impulse control of dynamical systems. *SIAM J. Control Optim.* **57**, 2720–2752.
[32] RIEDLER, M. AND THIEULLEN, M. (2015). Spatio-temporal hybrid (PDMP) models: central limit theorem and Langevin approximation for global fluctuations. Application to electrophysiology. *Bernoulli* **21**, 647–696.
[33] RIEDLER, M., THIEULLEN, M. AND WAINRIB, G. (2012). Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrane models. *Electron. J. Probab.* **17**(55), 48.
[34] SOLEIMANI, M., CAMPEAN, F. AND NEAGU, D. (2021). Integration of hidden Markov modelling and Bayesian network for fault detection and prediction of complex engineered systems. *Reliab. Eng. Syst. Safety* **215**, 107808.
[35] ZHANG, H., DUFOUR, F., DUTUIT, Y. AND GONZALEZ, K. (2008). Piecewise deterministic Markov processes and dynamic reliability. *Proc. Inst. Mech. Eng. Part O: J. Risk Reliab*. 222, 545–551.

https://doi.org/10.1017/apr.2025.10044 Published online by Cambridge University Press