J. Appl. Prob. **57,** 497–512 (2020)
*doi:*10.1017/jpr.2020.18
© *Applied Probability Trust* 2020

OPTIMAL STOPPING FOR MEASURE-VALUED PIECEWISE DETERMINISTIC MARKOV PROCESSES

BERTRAND CLOEZ,* *INRA*
BENOÎTE DE SAPORTA,** AND
MAUD JOUBAUD,** *University of Montpellier*

Abstract

This paper investigates the random horizon optimal stopping problem for measure-valued piecewise deterministic Markov processes (PDMPs). This is motivated by population dynamics applications, when one wants to monitor some characteristics of the individuals in a small population. The population and its individual characteristics can be represented by a point measure. We first define a PDMP on a space of locally finite measures. Then we define a sequence of random horizon optimal stopping problems for such processes. We prove that the value function of the problems can be obtained by iterating some dynamic programming operator. Finally we prove via a simple counter-example that controlling the whole population is not equivalent to controlling a random lineage.

Keywords: Markov processes; measure space; optimal stopping; dynamic programming; population dynamics

2010 Mathematics Subject Classification: Primary 60J25
Secondary 60G40; 92D25; 93E20

1. Introduction

Piecewise deterministic Markov processes (PDMPs) form a general class of non-diffusion processes that was introduced by Mark Davis in the 1980s [9, 10]. Such processes have deterministic trajectories punctuated by random jumps. They belong to the family of hybrid processes with a discrete component called the mode or regime interacting with a Euclidean component. PDMPs can model a wide area of phenomena from insurance and queuing problems [10], finance [1], and reliability [11] to neuroscience [14, 19], population dynamics [2, 7], and many other fields. In this paper we are especially interested in population dynamics applications. In this area, special cases of PDMPs include for instance growth-fragmentation processes for one or several interacting species [5, 6, 8, 12]. In that case, commonly, the deterministic part is the growth process that may depend, for example, on the age of the individual, on its size, on the quantity of available nutrient, and the jumps correspond to fragmentation or division (for cells), birth or death events, abrupt changes in the environment, etc.

In Davis's original construction, PDMPs are defined on subsets of $\mathbb{R}^d$, for some dimension d that may change when the process jumps. In this paper we are interested in extending the definition of PDMPs to measure-valued state spaces. Infinite-dimensional PDMPs have already been introduced in [4] (see also [14] and 19]). In those papers, PDMPs take values

Received 13 September 2018; revision received 23 July 2019.
* Postal address: MISTEA, INRA, Montpellier SupAgro, University of Montpellier, Montpellier, France.
** Postal address: IMAG, Univ. Montpellier, CNRS, Montpellier, France. Email: benoite.de-saporta@umontpellier.fr

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

in a separable Hilbert space and model spatio-temporal phenomena occurring on neuronal membranes. Our approach differs as we are interested in measure-valued PDMPs to deal with population dynamics models. Instead of modeling the dynamics of only a single individual by a finite-dimensional PDMP, we aim to take into account simultaneously the dynamics of all the individuals in the branching population when the population remains small and the stochastic approach is relevant and large-scale approximations do not hold. Such a population can be represented by a point measure, hence the need to define measure-valued PDMPs. The measure-valued process representation in population dynamics is used, for example in [3] and [13], with fragmentation-type processes. It is a particular case of measure-valued PDMPs, with no deterministic dynamics between jumps and exponential distributions for the jump times.

After constructing measure-valued PDMPs, we define a sequence of random horizon optimal stopping problems for measure-valued PDMPs and prove that the value functions can be obtained by iterating some dynamic programming operator. We also exhibit a sequence of ε-optimal stopping times. Our approach extends the results in [15] to the infinite-dimensional setting. In addition, we solve the optimal stopping problem for non-continuous reward functions.

When dealing with a branching population, some important characteristics of the global population, e.g. laws of large numbers for functionals of the individuals, can be obtained by simply studying a suitably weighted random tagged lineage, by means of many-to-one formulas; see e.g. [3], [12], and [16]. Here we prove that this property does not hold true for the optimal stopping problem. We provide a simple counter-example of cell division where stopping a suitably chosen tagged cell and stopping the whole population yield different value functions.

The paper is organized as follows. In Section 2 we construct measure-valued PDMPs. In Section 3 we state and solve the optimal stopping problem for such processes. Finally, in Section 4, we compare the value functions of the optimal stopping problems for the whole population and a tagged lineage.

2. Construction of measure-valued PDMPs

This section is dedicated to the construction of piecewise deterministic Markov processes taking values in some measure space. Our construction of measure-valued PDMPs follows the same lines as in [9]: we first define the hybrid state space in which the process evolves, then we define the local characteristics giving the dynamics of trajectories between jumps, the jump times and the post-jump locations, and prove that a strong Markov process with such characteristics can be constructed. Finally we provide a toy example of such processes. We start by setting some notation that will be used throughout the paper.

2.1. Notation

Let d be a positive integer. We let $\mathcal{B}$ denote the σ-field of Borel sets on $\mathbb{R}^d$ and let $\mathcal{B}_b$ denote its subset of bounded Borel sets. More generally, for any topological space E, we let $\mathcal{B}(E)$ denote its Borel σ-field, $B(E)$ its set of measurable bounded real-valued functions, $\overline{E}$ its closure, and ∂E its boundary.

Let $\mathfrak{M}$ be the set of locally finite measures on $(\mathbb{R}^d, \mathcal{B})$ and let $\mathfrak{N} = \{\mu \in \mathfrak{M}; \text{ for all } B \in \mathcal{B}_b, \mu(B) \in \mathbb{N}\}$ be the set of locally finite point measures. Note that any μ in $\mathfrak{N}$ can be expressed as a (possibly infinite) sum of Dirac distributions; the Dirac distribution with point mass at x, for x in $\mathbb{R}^d$, will be denoted by δ_x.

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

Let $C_c(\mathbb{R}^d)$ be the set of continuous real-valued functions with compact support on $\mathbb{R}^d$. For any measure $\mu \in \mathfrak{M}$ and function $f \in C_c(\mathbb{R}^d)$, set

$$\mu f := \int_{\mathbb{R}^d} f(x)\mu(\mathrm{d}x).$$

We endow $\mathfrak{M}$ with the vague topology. Recall that the vague convergence for a sequence of measures $(\mu_n)_n \subset \mathfrak{M}$ to a measure $\mu \in \mathfrak{M}$ is defined by

$$\mu_n f \xrightarrow[n\to\infty]{} \mu f \quad \text{for all } f \in C_c(\mathbb{R}^d).$$

We denote it by

$$\mu_n \xrightarrow[n\to\infty]{v} \mu.$$

It is easily seen that that $\mathfrak{N}$ is a closed subset of $\mathfrak{M}$ for the vague topology.

For any real numbers a and b, $a \vee b$ and $a \wedge b$ denote the maximum and minimum respectively between a and b.

2.2. State space of measure-valued PDMPs

Let K be a finite set called the modes or regimes space. For any mode $v \in K$, let E_v be an open subset of $\mathfrak{M}$, representing the state space in mode v. The global state space is then

$$E = \{(v, \zeta) \in K \times \mathfrak{M} \mid v \in K, \zeta \in E_v\}.$$

We endow this set E with the σ-field $\mathcal{E}$ generated by sets of the form $A = \{v\} \times A_v$ for all Borel sets $A_v \in \mathcal{B}(E_v)$. For instance, if one can consider the temporal evolution of a cell population characterized by state variables such as age, size, growth, maturity, or protein content, then the quantity $\zeta \in \mathfrak{N}$ represents such state variables and the mode can be the experiment conditions. In the following example, we consider a simple case with only one mode and where the variable state corresponds to the cell size.

Example 2.1. Consider a population of three individuals at a given time. Their sizes are denoted by $(x_i)_{1\leqslant i\leqslant 3}$. This population is identified with the measure

$$\zeta = \sum_{i=1}^{3} \delta_{x_i} \in \mathfrak{N}.$$

It gives a complete view of the population: all the information is contained in ζ. As time goes by, the number of individuals may increase or decrease, leading to more or fewer terms in this Dirac sum representation, but it will remain a measure in $\mathfrak{N}$. This representation is then easier to manipulate than a changing-dimension vector.

On the measure space $\mathfrak{M}$, we introduce a particular metric ρ for the vague topology in order to have a Polish space (i.e. a separable completely metrizable topological space). This property will be used in Section 2.3 for the explicit construction of the stochastic process. As shown in [17, Appendix], a suitable choice for ρ is constructed as follows. Let $\mathcal{C}$ be a countable basis of open bounded subsets in $\mathbb{R}^d$ closed under finite unions. For all C in $\mathcal{C}$, there exists a sequence $(C_n)_n$ in $\mathcal{B}_b$ and an increasing sequence $(f_{C,n})_n$ in $C_c(\mathbb{R}^d)$ such that

$$f_{C,n} \xrightarrow[n\to\infty]{} \mathbb{1}_C \quad \text{and} \quad \mathbb{1}_{C_n} \leqslant f_{C,n} \leqslant \mathbb{1}_C.$$

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

Since $\mathcal{C}$ is countable, the set $\{f_{C,n} | C \in \mathcal{C}, n \in \mathbb{N}\}$ is also countable, and we then number those functions $f_1, f_2, \ldots$. Any measure μ is completely determined by the set $\{\mu f_k, k \in \mathbb{N}\}$. Now, for all μ and μ' in $\mathfrak{M}$, we define the distance ρ by

$$\rho(\mu, \mu') := \sum_{k \geqslant 1} \frac{1}{2^k} (1 - \exp(-|\mu f_k - \mu' f_k|)).$$

From this metric on $\mathfrak{M}$, we define a metric on E related to its hybrid structure: any two points in E with different modes must be arbitrarily far away from each other. For all $x = (v, \zeta)$ and $x' = (v', \zeta')$ in E, set

$$\rho_0(x, x') := \begin{cases} \dfrac{2}{\pi} \arctan(\rho(\zeta, \zeta')) & \text{if } v = v', \\ 1 & \text{otherwise.} \end{cases}$$

Thus $\rho_0(x, x')$ equals 1 if and only if x and x' have different modes. With this metric, a sequence $(x_n)_n = (v_n, \zeta_n)_n$ converges to $x = (v, \zeta)$ in (E, ρ_0) if and only if there exists some m in $\mathbb{N}$ such that

$$\begin{cases} v_n = v \text{ for } n \geqslant m, \\ \zeta_{m+k} \xrightarrow[k \to \infty]{} \zeta \text{ in } (E_v, \rho). \end{cases}$$

We have

$$\overline{E} = \{(v, \zeta) \in K \times \mathfrak{M} \mid v \in K, \zeta \in \overline{E}_v\},$$

the closure of E for the distance ρ_0. The following statement is then straightforward.

Lemma 2.1. *The metric space $(\overline{E}, \rho_0)$ is a Polish space.*

2.3. Construction of measure-valued PDMPs

We now introduce the three local characteristics of the PDMP specifying the dynamics of trajectories between jumps, the jump times, and the post-jump locations.

- The flow Φ is defined by $(x, t) \mapsto \Phi(x, t) = (v, \Phi_v(\zeta, t))$ for all $x = (v, \zeta)$ in E and non-negative t, where the functions $\Phi_v : \mathfrak{M} \times \mathbb{R} \to \mathfrak{M}$ are continuous and have a semi-group property: for all $s, t \geqslant 0$, we have $\Phi_v(\cdot, t + s) = \Phi_v(\Phi_v(\cdot, s), t)$. The flow describes the deterministic trajectory of the process between jumps. Let

  $$t^*(x) = \inf\{t > 0, \Phi_v(t, \zeta) \in \partial E_v\}$$

  be the deterministic time the flow takes to reach the boundary of the domain starting from $x = (v, \zeta) \in E$, with the usual convention $\inf \emptyset = +\infty$. An infinite exit time t^* means that the process cannot reach the boundary in finite time.

- The jump intensity $\lambda : E \to \mathbb{R}_+$ is a measurable function, with a local integrability property: for all $x = (v, \zeta) \in E$ there exists some $\varepsilon > 0$ such that

  $$\int_0^\varepsilon \lambda(v, \Phi_v(\zeta, s)) \, \mathrm{d}s < \infty.$$

  It determines the frequency of the jumps.

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

- The Markov kernel $\mathcal{Q}\colon \overline{E} \times \mathcal{B}(E) \to [0, 1]$ selects the post-jump locations. It has the following property:
$$\mathcal{Q}(x, \{x\}) = 0 \quad \text{for all } x \in E,$$
meaning that the process cannot have a no-move jump.

From these local characteristics, one can construct a stochastic process similarly to [10, Section 24] as follows. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the canonical space for a sequence $(U_n)_{n\geqslant 1}$ of independent random variables with uniform distribution on $[0, 1]$. The sample path of an E-valued PDMP $(X_t(\omega))_{t\geqslant 0}$ starting from a fixed initial point $x = (v, \zeta) \in E$ and for some $\omega \in \Omega$ is defined iteratively. Let

$$F(x, t) = \mathbb{1}_{\{t<t^*(x)\}} \exp\left(-\int_0^t \lambda(v, \Phi_v(\zeta, s))\, \mathrm{d}s\right) \tag{2.1}$$

for $(x, t) = (v, \zeta, t) \in E \times \mathbb{R}_+$, let Ψ_1 be the function from $E \times [0, 1]$ onto $\mathbb{R}_+$ defined by

$$\Psi_1(x, u) = \inf\{t \geqslant 0; F(x, t) \leqslant u\},$$

and let $S_1(\omega) = T_1(\omega) = \Psi_1(x, U_1(\omega))$ be the first jump time of the process. Thus $F(x, \cdot)$ is the survivor function of T_1.

As the spaces $\mathfrak{M}$ and $\overline{E}$ are Polish, one can use [18, p. 6] to obtain that $\{\mathcal{Q}(x, \cdot)\}_{x\in\overline{E}}$ is a collection of probability measures on $\overline{E}$, with a measurable dependence on the parameter x in $\overline{E}$. Then there is a measurable function $\Psi_2\colon E \times [0, 1] \to E$ such that the distribution of $\Psi_2(y, U)$ is $\mathcal{Q}(y, \cdot)$ for any random variable U with uniform distribution on $[0, 1]$. Hence one can then define

$$\begin{aligned} X_t(\omega) &= \Phi_v(\zeta, t), \qquad \text{for } 0 \leqslant t < T_1, \\ X_{T_1}(\omega) &= \Psi_2((v, \Phi_v(\zeta, T_1(\omega))), U_2(\omega)). \end{aligned}$$

Hence the trajectory X_t follows the deterministic flow starting from $X_0 = x$ until the first jump time T_1. At T_1 a new location $Z_1 = X_{T_1}$ is drawn according to the Markov kernel $\mathcal{Q}$. Namely, the law $\mathcal{L}(Z_1 \mid T_1)$ of Z_1 conditionally on T_1 is

$$\mathcal{L}(Z_1 \mid T_1) = \mathcal{Q}(\Phi(x, S_1), \cdot).$$

The process now restarts from X_{T_1} following the same steps. Define

$$S_2(\omega) = \Psi_1(X_{T_1}(\omega), U_3(\omega)), \qquad T_2(\omega) = S_1(\omega) + S_2(\omega),$$

and set

$$\begin{aligned} X_t(\omega) &= \Phi(X_{T_1}(\omega), t - T_1(\omega)), \qquad \text{for } T_1(\omega) \leqslant t < T_2(\omega), \\ X_{T_2}(\omega) &= \Psi_2((v, \Phi(X_{T_1}(\omega), S_2(\omega))), U_4(\omega)), \end{aligned}$$

and so on. In order to avoid explosion issues, the following assumption will hold throughout the paper.

Assumption 2.1. *For all (x,t) in* $E \times \mathbb{R}_+$,

$$\mathbb{E}_x\left[\sum_{n=1}^{\infty} \mathbb{1}_{(T_n<t)}\right] < \infty.$$

Hence the trajectories of (X_t) are well-defined for all $t \geqslant 0$.

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

The positive random variables $S_1, S_2, \dots$ are the times between two consecutive jumps or inter-jump times. For notational convenience, we set $T_0 = S_0 = 0$ and $T_n := \sum_{i=1}^{n} S_n$ the nth jump time of the process. Note that we have $T_1 = S_1$ and $S_{n+1} = T_{n+1} - T_n$. The sequence $(Z_n)_n$ with $Z_n = X_{T_n}$ describes the post jump locations of the process. By construction, all the randomness of the continuous-time process $(X_t)_{t \geqslant 0}$ is contained in the discrete-time process $(Z_n, S_n)_{n \geqslant 0}$.

2.4. Structure of stopping times and Markov property

The aim of this section is to prove that the special structure of stopping times for finite-dimensional PDMPs given in [10, Theorem A2.3] still holds in the measure space context. This yields the Markov property and will be important for the study of the optimal stopping problem in Section 3.

Let $D_E[0, +\infty)$ be the set of functions on $\mathbb{R}_+$ with values in E that are right-continuous with left limits. Let $\tilde{X}_t$ denote the coordinate function $\tilde{X}_t(f) = f(t)$ for $f \in D_E[0, +\infty)$. Let $(\mathcal{F}_t^0)_{t \geqslant 0}$ denote the natural filtration of $(\tilde{x}_t)$ and $\mathcal{F}^0 = \vee_{t \geqslant 0} \mathcal{F}_t^0$. Under Assumption 2.1, for each starting point $x \in E$, the construction in the previous section defines a measurable mapping Ψ_x from Ω onto $D_E[0, +\infty)$ such that $\tilde{X}_t(\Psi_x(\omega)) = X_t(\omega)$. Let $\mathbb{P}_x$ denote the image measure of $\mathbb{P}$ by Ψ_x. This defines a family of measures $(\mathbb{P}_x)_{x \in E}$ on $D_E[0, +\infty)$. We shall identify $\tilde{X}_t$ and X_t. For any probability measure ν on E, define the measure $\mathbb{P}_\nu$ on $(D_E[0, +\infty), \mathcal{F}^0)$ by $\mathbb{P}_\nu(\cdots) = \int_E \mathbb{P}_x(\cdots)\nu(\mathrm{d}x)$. Now let $\mathcal{F}_t^\nu$ be the completion of $\mathcal{F}_t^0$ with all $\mathbb{P}_\nu$-null sets of $\mathcal{F}^0$ and define $\mathcal{F}_t = \cap_{\nu \in \mathcal{P}(E)} \mathcal{F}_t^\nu$, where $\mathcal{P}(E)$ is the set of probability measures on E.

We can now state a crucial result on the structure of stopping times for our process.

Theorem 2.1. *A non-negative random variable τ is a $(\mathcal{F}_t)_{t \geqslant 0}$ stopping time if and only if there exists a sequence $(R_n)_{n \in \mathbb{N}}$ of non-negative $(\mathcal{F}_{T_n})_{n \in \mathbb{N}}$-adapted random variables such that*

$$\tau = \sum_{n=1}^{\infty} R_{n-1} \wedge S_n.$$

Proof. The proof follows the same lines as in [11, Section 1.7]. It is based on Theorem A2.3 in [10], which is valid for any right-continuous piecewise constant process taking values in a Borel set and the one-to-one correspondence between our process (X_t) and the right-continuous piecewise constant $E \times \mathbb{N}$-valued process (η_t) defined by

$$\eta_t = (X_0, 0), \ t < T_1, \quad \eta_t = (X_{T_n}, n), \ T_n \leqslant t < T_{n+1}$$

(see [10, (25.1)]). □

Theorem 2.2. *The process $(X_t)_{t \geqslant 0}$ on $(D_E[0, +\infty), \mathcal{F}, (\mathcal{F}_t), (\mathbb{P}_x)_{x \in E})$ is a strong Markov process.*

Proof. The strong Markov property is proved in the same way as in [10, Section 25], using Theorem 2.1 and the one-to-one correspondence with the piecewise constant process (η_t) defined above. □

2.5. Toy example

We now develop Example 2.1 into a more generic model for some cells population. Typically, cells grow and divide into two daughter cells that start growing and then divide and so on.

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

2.5.1. *Single-cell model.* We first define a single-cell model that follows standard finite-dimensional PDMP dynamics, where one randomly selects a single daughter cell at each division. This model was further studied in [12]. We consider a model with a single mode, hence the state space is simply $\mathbb{R}_+$. For ξ in $\mathbb{R}_+$ and $t > 0$, the size at time t for a cell with initial size ξ is given by $\phi(\xi, t) = \xi \exp(rt)$, where $r > 0$ is the common growth rate for all the cells. The jump intensity giving the division dates is given by $l(\xi) = \xi^\alpha$ for some positive number α. This is a simplified model consistent with the statistical evidence that division is triggered by the cell size rather than by its age; see [20]. The jump kernel is simply a division by two, i.e. $Q(x, A) = \delta_{x/2}(A)$ for any Borel subset A of $\mathbb{R}_+$.

2.5.2 *Population model.* We now consider the previous cell growth-division dynamics but for the whole population instead of a single randomly selected cell. Again we consider a single mode so that the space E is simply $\mathfrak{N}$.

Let $\zeta = \sum_{i=1}^n \delta_{x_i}$ be in $\mathfrak{N}$ be an initial state of n cells with respective sizes $(x_i)_{1 \leqslant i \leqslant n}$. Each cell grows following the previous dynamics, so that globally the flow is

$$\begin{aligned} \Phi : \mathfrak{N} \times \mathbb{R}_+ &\to \mathfrak{N} \\ (\zeta, t) &\mapsto \textstyle\sum_{i=1}^n \delta_{\phi(x_i,t)}. \end{aligned}$$

The first jump time corresponds to the first splitting, i.e. the minimum between n exponentially distributed random variables. Thus the jump intensity is

$$\begin{aligned} \lambda : \mathfrak{N} &\to \mathbb{R} \\ \zeta &\mapsto \textstyle\sum_{i=1}^n l(x_i). \end{aligned}$$

Given that the jth cell is the one that splits, the post-jump location is

$$Z_1 = \sum_{\substack{i=1 \\ i \neq j}}^n \delta_{x_i} + 2\delta_{x_j/2},$$

where cell number j was removed and two new cells are added with half the size of cell j. For notational convenience, we set

$$\zeta_{(j)} := \zeta - \delta_{x_j} + 2\delta_{x_j/2}.$$

Thus, for A in $\mathcal{E}$, the Markov jump kernel of the PDMP is given by

$$\mathcal{Q}(\zeta, A) = \sum_{j=1}^n \frac{l(x_j)}{\sum_{i=1}^n l(x_i)} \mathbb{1}_A(\zeta_{(j)}).$$

2.5.3. *Time-augmented population model.* Sometimes it is convenient to add time in the state variable of a PDMP, for instance if we want to use time-dependent jump intensity, or trigger some jump when a certain lapse of time has passed, or to study control problems with time-dependent reward functions. In [10, Section 31] Davis proves that for finite-dimensional PDMPs, the time-augmented process is still a PDMP. The same property holds in the framework of our toy example.

We consider the time-augmented process $(\widetilde{X}_t)_t$ defined as the PDMP on $\mathfrak{N}$ starting from the initial state $\tilde{\zeta} = \sum_{i=1}^n \delta_{(x_i,t_i)}$. The flow is now given by

$$\begin{aligned} \widetilde{\Phi} : \mathfrak{N} \times \mathbb{R}_+ &\to \mathfrak{N} \\ (\tilde{\zeta}, t) &\mapsto \textstyle\sum_{i=1}^n \delta_{(\phi(x_i,t),t_i+t)}, \end{aligned}$$

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

for any $\tilde{\zeta}$ in $\mathfrak{N}$ of the form $\tilde{\zeta} = \sum_{i=1}^{n} \delta_{(x_i,t_i)}$. The jump intensity is

$$\begin{array}{rccc} \widetilde{\lambda}: & \mathfrak{N} & \to & \mathbb{R} \\ & \tilde{\zeta} = \sum_{i=1}^{n} \delta_{(x_i,t_i)} & \mapsto & \sum_{i=1}^{n} l(x_i), \end{array}$$

and the jump kernel is simply

$$\widetilde{\mathcal{Q}}\left(\sum_{i=1}^{n} \delta_{(x_i,t_i)}, A\right) = \sum_{j=1}^{n} \frac{l(x_j)}{\sum_{i=1}^{n} l(x_i)} \mathbb{1}_A \left(\sum_{i=1}^{n} \delta_{(x_i,t_i)} - \delta_{(x_j,t_i)} + 2\delta_{(x_j/2,t_i)}\right).$$

Remark 2.1. Following the definition of the time-augmented process by Davis, we could have defined the measure-valued time-augmented process on $\mathfrak{N} \times \mathbb{R}_+$ with the flow

$$\left(\sum_{i=1}^{n} \delta_{x_i}, u\right) \to \left(\sum_{i=1}^{n} \delta_{\phi_{u,t}(x_i)}, u + t\right).$$

But formally PDMPs on product spaces such as $\mathfrak{N} \times \mathbb{R}_+$ are not included in our previous construction. However, in our main examples, we only consider $\tilde{\zeta}$ in the form $\sum_{i=1}^{n} \delta_{(x_i,u)}$ which roughly means that all cells have the same current time.

We will further study this toy example in Section 4 to prove that controlling the whole population is not equivalent to controlling a suitably chosen random lineage.

3. Optimal stopping problem

We now turn to the main aim of this paper: defining and solving the optimal stopping problem for measure-valued PDMPs. Roughly speaking, one wants to stop the process at the best time in order to maximize some reward depending on the state of the process when stopped. More precisely, let $(X_t)_{t\geqslant 0}$ be an E-valued PDMP on $(D_E[0, +\infty), \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, (\mathbb{P}_x)_{x\in E})$ and let $g \in B(\overline{E})$ be some non-negative reward function. Let $\mathcal{M}$ denote the set of stopping times with respect to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ and, for all positive integer N, let $\mathcal{M}_N$ be the set of stopping times bounded by the Nth time jump T_N of the PDMP, that is,

$$\mathcal{M}_N = \{\tau \in \mathcal{M};\ \tau \leqslant T_N\}.$$

For all $x \in E$, set

$$\mathbb{V}(x) := \sup_{\tau\in\mathcal{M}_N} \mathbb{E}[g(X_\tau) \mid X_0 = x]. \tag{3.1}$$

Thus $\mathbb{V}(x)$ is the best possible (average) performance when stopping a PDMP starting from $X_0 = x$ before its Nth jump. Function $\mathbb{V}$ is called the *value function* of the optimal stopping problem. Solving an optimal stopping problem consists of characterizing the value function as the solution of some recursive equations called *dynamic programming* equations and exhibiting a family of *ε-optimal stopping times* $\tau_\varepsilon \in \mathcal{M}_N$ such that

$$\mathbb{V}(x) - \varepsilon \leqslant \mathbb{E}[g(X_{\tau_\varepsilon}) \mid X_0 = x] \leqslant \mathbb{V}(x).$$

In this section we first define some suitable dynamic programming operators and a family of stopping times. Then we prove that the value functions can be constructed by iteration

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

of the dynamic programming operators and that the stopping times are ε-optimal. This section is inspired by the study of the optimal stopping problem for finite (fixed) dimension PDMPs derived in [15]. The structure of the proof closely follows that of [15]. Difficulties are essentially technical: constructing the suitable state space, its topology, etc., to keep similar properties for the construction, as we did in the previous section. We simplify some arguments regarding the infinite-horizon problem in Theorem 3.2 and we investigate in detail the structure of stopping times (see Lemma 3.1). In addition, in [15], the reward function g is assumed to be continuous along the flow. We do not enforce this assumption here.

3.1. Dynamic programming operators

We start with some additional notation and assumption. For w in $B(E)$, x in E, and l a measurable real-valued function on E, we denote for short

$$l\mathcal{Q}w(x) := l(x) \times \mathcal{Q}w(x) = l(x) \int_E w(y)\mathcal{Q}(x, \mathrm{d}y).$$

The following assumption is made for simplicity reasons. It is satisfied in most real-life examples, for instance when monitoring a population until some finite horizon time.

We now define some operators on $B(E)$. Let $\mathbf{H}$ be the operator from $B(\overline{E})$ onto $B(\overline{E} \times \mathbb{R}_+)$ defined for all h in $B(\overline{E})$, x in $\overline{E}$ and $t \in \mathbb{R}_+$, by

$$\mathbf{H}h(x, t) = h(\Phi(x, t \wedge t^*(x)))\, \mathrm{e}^{-\Lambda(x, t\wedge t^*(x))},$$

and let $\mathbf{I}$ be the operator from $B(E)$ onto $B(E \times \mathbb{R}_+)$ defined for all w in $B(E)$, x in E, and $t \in \mathbb{R}_+$, by

$$\mathbf{I}w(x, t) = \int_0^{t\wedge t^*(x)} \lambda \mathcal{Q}w(\Phi(x, s))\, \mathrm{e}^{-\Lambda(x,s)}\, \mathrm{d}s,$$

where

$$\Lambda(x, t) = \int_0^t \lambda(\Phi(x, s))\, \mathrm{d}s.$$

We also introduce operator $\mathbf{K}$ from $B(E)$ onto $B(E)$, defined for all w in $B(E)$ and x in E as

$$\mathbf{K}w(x) = \int_0^{t^*(x)} \lambda \mathcal{Q}w(\Phi(x, s))\, \mathrm{e}^{-\Lambda(x,s)}\mathrm{d}s + \mathcal{Q}w(\Phi(x, t^*(x)))\, \mathrm{e}^{-\Lambda(x,t^*(x))}.$$

It is straightforward to see that these operators can be expressed as expectations involving the embedded Markov chain (Z_n, S_n) defined in Section 2.3.

Proposition 3.1. *For all h in $B(\overline{E})$ and w in $B(E)$, x in E and $t \geqslant 0$, we have*

$$\begin{aligned}
\mathbf{H}h(x, t) &= \mathbb{E}_x[h(X_{t\wedge t^*(x)})\, \mathbb{1}_{S_1 > t\wedge t^*(x)}] = h(\Phi(x, t \wedge t^*(x)))\mathbb{P}_x(S_1 > t \wedge t^*(x)),\\
\mathbf{I}w(x, t) &= \mathbb{E}_x[w(Z_1)\mathbb{1}_{S_1 \leqslant t\wedge t^*(x)}],\\
\mathbf{K}w(x) &= \mathbb{E}_x[w(Z_1)].
\end{aligned}$$

Finally we let $\mathbf{J}$ and $\mathbf{L}$ denote the operators from $B(E) \times B(\overline{E})$ onto $B(E \times \mathbb{R}_+)$ and $B(E)$ respectively defined for all w in $B(E)$, h in $B(\overline{E})$, x in E, and $t \in \mathbb{R}_+$, by

$$\begin{aligned}
\mathbf{J}(w, h)(x, t) &= \mathbf{H}h(x, t) + \mathbf{I}w(x, t),\\
\mathbf{L}(w, h)(x) &= \sup_{t\geqslant 0}\{\mathbf{J}(w, h)(x, t)\} \vee \mathbf{K}w(x).
\end{aligned}$$

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

We can now propose an iterative construction of value functions by iterating operator $\mathbf{L}$. For all $x \in E$, set

$$\mathbb{V}_0(x) = g(x),$$
$$\mathbb{V}_n(x) = \mathbf{L}(\mathbb{V}_{n-1}, g)(x) \quad \text{for all } n \geqslant 1,$$

where g is the reward function of the optimal stopping problem. Clearly the functions $\mathbb{V}_n$ are in $B(E)$. Roughly speaking, operator $\mathbf{L}$ represents the best compromise between stopping at the best location along the deterministic trajectory following the flow ($\sup_{t \geqslant 0} \mathbf{J}$ part) or waiting until the next jump ($\mathbf{K}$ part).

3.2. Family of stopping times

Now we introduce a family of random variables and prove they are stopping times. They will be candidate ε-optimal stopping times for our optimal stopping problem. For all $x \in E$, $n \in \mathbb{N}$, and $\varepsilon > 0$, set

$$r_{n,\varepsilon}(x) = \begin{cases} t^*(x) & \text{if } \mathbf{K}\mathbb{V}_n(x) > \sup_{t>0} \mathbf{J}(\mathbb{V}_n, g)(x, t), \\ \inf\{s \geqslant 0; \mathbf{J}(\mathbb{V}_n, g)(x, s) \geqslant \sup_{t>0} \mathbf{J}(\mathbb{V}_n, g)(x, t) - \varepsilon\} & \text{otherwise.} \end{cases}$$

For all $\varepsilon > 0$ and $n \geqslant 2$, also set

$$R_{1,\varepsilon} = r_{0,\varepsilon}(Z_0),$$
$$R^{\varepsilon}_{n,0} = r_{n-1,\varepsilon/2}(Z_0),$$
$$R^{\varepsilon}_{n,k} = r_{n-k-1,\varepsilon/2^k}(Z_k)\mathbb{1}_{(R^{\varepsilon}_{n,k-1} \geqslant S_k)}, \quad 1 \leqslant k \leqslant n-1.$$

Then $R^{\varepsilon}_{n,k}$ is clearly $\mathcal{F}_{T_k}$-measurable for $0 \leqslant k \leqslant n-1$. Finally, define $S_{1,\varepsilon} := r_{0,\varepsilon}(Z_0) \wedge T_1 = R^{\varepsilon}_{1,0} \wedge S_1$, and by iteration,

$$S_{n,\varepsilon} := \begin{cases} R^{\varepsilon}_{n,0} & \text{if } T_1 > R^{\varepsilon}_{n,0}, \\ T_1 + \theta(T_1) S_{n-1,\varepsilon/2} & \text{else } T_1 \leqslant R^{\varepsilon}_{n,0}, \end{cases}$$

where $\theta(t)$ is the shift operator with lag t on $D_E[0, +\infty)$, namely for $f \in D_E[0, +\infty)$, $\theta(t)f(\,\cdot\,) = f(t + \cdot)$.

In order to prove that the $S_{n,\varepsilon}$ are stopping times in $\mathcal{M}_n$, we first study the effect of the shift operator on $R^{\varepsilon}_{n,k}$.

Lemma 3.1. *For all $\varepsilon > 0$ and $n \geqslant 2$ and $1 \leqslant k \leqslant n-1$, on the set $(T_1 \leqslant S_{n,2\varepsilon})$, we have*

$$R^{2\varepsilon}_{n,k} = \theta(T_1)(R^{\varepsilon}_{n-1,k-1}).$$

Proof. For $n = 2$, by definition, we have $R^{\varepsilon}_{1,0} = r_{0,\varepsilon}(Z_0)$ and hence $\theta(T_1)(R^{\varepsilon}_{1,0}) = r_{0,\varepsilon}(Z_1)$ and $R^{2\varepsilon}_{2,1} = r_{0,\varepsilon}(Z_1)$ on $(T_1 \leqslant S_{2,2\varepsilon}) = (T_1 \leqslant R^{\varepsilon}_{2,0})$. Hence the result holds.

For $n \geqslant 3$, we prove the result by induction on k using similar arguments and the fact that $(T_1 \leqslant S_{n,\varepsilon}) = (R^{2\varepsilon}_{n,0} \geqslant T_1)$. □

We now prove that $S_{n,\varepsilon}$ is a stopping time using the characterization of Theorem 2.1.

Lemma 3.2. *For all $\varepsilon > 0$ and $n \in \mathbb{N}^*$, we have*

$$S_{n,\varepsilon} = \sum_{k=1}^{n} R^{\varepsilon}_{n,k-1} \wedge S_k.$$

In particular, $S_{n,\varepsilon}$ is a stopping time and $S_{n,\varepsilon} \leqslant T_n$.

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

Proof. We proceed by induction on n. For $n=1$, by definition, $S_{1,\varepsilon} = r_{0,\varepsilon}(Z_0) \wedge T_1 = R^{\varepsilon}_{1,k-1} \wedge S_1$ and the result is true.

Suppose the result holds for $n-1$. From the definition, on the event

$$(r_{n-1,\varepsilon/2}(Z_0) = R^{\varepsilon}_{n,0} < S_1),$$

we have $S_{n,\varepsilon} = R^{\varepsilon}_{n,0}$ and $R^{\varepsilon}_{n,1} = 0 < S_2$, $R^{\varepsilon}_{n,2} = 0 < S_3$, and so on, so that $\sum_{k=1}^{n} R^{\varepsilon}_{n,k-1} \wedge S_k = R^{\varepsilon}_{n,0}$ and the result is valid.

On the event $(R^{\varepsilon}_{n,0} \geqslant S_1)$, by definition $S_{n,\varepsilon} = T_1 + \theta(T_1) S_{n-1,\varepsilon/2}$. Now, the induction hypothesis and the previous lemma yield

$$T_1 + \theta(T_1) S_{n-1,\varepsilon/2} = S_1 + \sum_{k=1}^{n-1} R^{\varepsilon}_{n,k} \wedge S_{k+1} = \sum_{k=1}^{n} R^{\varepsilon}_{n,k-1} \wedge S_k,$$

on $(R^{\varepsilon}_{n,0} \geqslant S_1)$. Hence the result. □

3.3. Characterization of the value function

We can now state and prove our main result, namely that the value function $\mathbb{V}$ of the optimal stopping problem (3.1) equals the Nth iterate $\mathbb{V}_N$ and that $S_{N,\varepsilon}$ is an ε-optimal stopping time.

Theorem 3.1. *Let x be in E and $n \in \mathbb{N}$. Then, for all $\varepsilon > 0$, $S_{n,\varepsilon}$ is a stopping time in $\mathcal{M}_n$ and*

$$\mathbb{V}_n(x) = \sup_{S \in \mathcal{M}_n} \mathbb{E}_x[g(X_S)], \tag{3.2}$$

$$\mathbb{E}_x[g(X_{S_{n,\varepsilon}})] \geqslant \mathbb{V}_n(x) - \varepsilon. \tag{3.3}$$

Proof. By an induction argument we will prove equation (3.3) and the following inequality:

$$\mathbb{E}_x[g(X_{S \wedge T_n})] \leqslant \mathbb{V}_n(x) \quad \text{for all } S \in \mathcal{M},\ n \in \mathbb{N}. \tag{3.4}$$

These two equations imply (3.2). Indeed, for all S in $\mathcal{M}_\infty$, we have $S \wedge T_n \in \mathcal{M}_n$; then, for all S in $\mathcal{M}_\infty$, n in $\mathbb{N}$ and x in E, (3.3) and (3.4) yield

$$\mathbb{V}_n(x) - \varepsilon \leqslant \mathbb{E}_x[g(X_{S_{n,\varepsilon}})] \leqslant \sup_{S \in \mathcal{M}_n} \mathbb{E}_x[g(X_S)] = \sup_{S \in \mathcal{M}_n} \mathbb{E}_x[g(X_{S \wedge T_n})] \leqslant \mathbb{V}_n(x),$$

which is valid for all positive ε, hence the result.

It remains now to prove (3.4) and (3.3) by induction.

The case $n=1$ is based on Theorem 2.1 Indeed, we obtain

$$\mathbb{E}_x[g(X_{S \wedge T_1})] = \mathbb{E}_x[g(X_{R_0 \wedge T_1})],$$

for some $\mathcal{F}_0$-measurable random variable R_0. From this, we deduce

$$\begin{aligned}
\mathbb{E}_x[g(X_{S \wedge T_1})] &= \mathbb{E}_x[g(X_{R_0}) \mathbb{1}_{T_1 > R_0}] + \mathbb{E}_x[g(Z_1) \mathbb{1}_{T_1 \leqslant R_0}] \\
&= \mathbf{H} g(x, R_0) + \mathbf{I} g(x, R_0) \\
&= \mathbf{J}(g, g)(x, R_0) \\
&\leqslant \sup_{t \geqslant 0} \{\mathbf{J}(g, g)(x, t)\} \\
&\leqslant \mathbb{V}_1(x),
\end{aligned}$$

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

which proves (3.4) for $n = 1$. To prove (3.3), we distinguish two cases.

- If $\mathbf{K}g(x) > \sup_{t\geqslant 0} \mathbf{J}(g, g)(x, t)$, then $\mathbb{V}_1(x) = \mathbf{K}g(x) = \mathbb{E}_x[g(Z_1)]$ and also

$$S_{1,\varepsilon} = t^*(x) \wedge T_1 = T_1.$$

  Thus

$$\mathbb{E}_x[g(X_{S_{1,\varepsilon}})] = \mathbb{E}_x[g(Z_1)] = \mathbb{V}_1(x) \geqslant \mathbb{V}_1(x) - \varepsilon.$$

- Otherwise, $\mathbb{V}_1(x) = \sup_{t\geqslant 0} \mathbf{J}(g, g)(x, t)$ and, by definition of $\mathbf{J}$,

$$\begin{aligned}\mathbb{E}_x[g(X_{S_{1,\varepsilon}})] &= \mathbb{E}_x[g(X_{r_{0,\varepsilon}(x)\wedge T_1})] \\ &= \mathbb{E}_x[g(X_{r_{0,\varepsilon}(x)})\mathbb{1}_{T_1\geqslant r_{0,\varepsilon}(x)}] + \mathbb{E}_x[g(Z_1)\mathbb{1}_{T_1\leqslant r_{0,\varepsilon}(x)}] \\ &= \mathbf{J}(g, g)(x, r_{0,\varepsilon}(x)).\end{aligned}$$

We then deduce, by definition of $r_{0,\varepsilon}$,

$$\mathbb{E}_x[g(X_{S_{1,\varepsilon}})] \geqslant \sup_{t\geqslant 0}\{\mathbf{J}(g, g)(x, t)\} - \varepsilon = \mathbb{V}_1(\mu) - \varepsilon.$$

This completes the proof of (3.3) for $n = 1$.

Now, let $N \geqslant 1$, suppose that (3.4) and (3.3) hold true for all $n \leqslant N$, and prove that they hold for $n = N + 1$.

Begin by proving (3.4). Again, this is based on Theorem 2.1 As S is a stopping time, it can be decomposed as

$$S = \sum_{n=1}^{\infty} R_{n-1} \wedge S_n,$$

where $(R_n)_{n\in\mathbb{N}}$ are non-negative $(\mathcal{F}_{T_n})_{n\in\mathbb{N}}$-adapted random variables. In particular, on $(S \geqslant T_1)$ we have

$$S = S_1 + \sum_{n=2}^{\infty} R_{n-1} \wedge S_n = T_1 + \theta(T_1)(S'),$$

for some stopping time S'. Thus we have

$$\begin{aligned}\mathbb{E}_x[g(X_{S\wedge T_{n+1}})] &= \mathbb{E}_x[g(X_{S\wedge T_1})\mathbb{1}_{S<T_1}] + \mathbb{E}_x[g(X_{S\wedge T_{n+1}})\mathbb{1}_{S\geqslant T_1}] \\ &= \mathbb{E}_x[g(X_{R_0\wedge T_1})\mathbb{1}_{R_0<T_1}] + \mathbb{E}_x[g(X_{S\wedge T_{n+1}})\mathbb{1}_{R_0\geqslant T_1}] \\ &= \mathbb{E}_x[g(X_{R_0})\mathbb{1}_{R_0<T_1}] + \mathbb{E}_x[\mathbb{E}_x[g(X_{(T_1+\theta(T_1)(S'))\wedge T_{n+1}})|\mathcal{F}_1]\mathbb{1}_{R_0\geqslant T_1}] \\ &= \mathbb{E}_x[g(X_{R_0})\mathbb{1}_{R_0<T_1}] + \mathbb{E}_x[\mathbb{E}_{Z_1}[g(X_{S'\wedge T_n})]\mathbb{1}_{R_0\geqslant T_1}] \qquad (3.5) \\ &\leqslant \mathbb{E}_x[g(X_{R_0})\mathbb{1}_{R_0<T_1}] + \mathbb{E}_x[\mathbb{V}_n(Z_1)\mathbb{1}_{R_0\geqslant T_1}] \qquad (3.6) \\ &\leqslant \sup_{r\geqslant 0} \mathbb{E}_x[g(X_r)\mathbb{1}_{r<T_1}] + \mathbb{E}_x[\mathbb{V}_n(Z_1)\mathbb{1}_{r\geqslant T_1}] \\ &\leqslant \sup_{r\geqslant 0} \mathbf{J}(\mathbb{V}_n, g)(x, r) \vee \mathbf{K}\mathbb{V}_n(x) \\ &= \mathbb{V}_{n+1}(x).\end{aligned}$$

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

Line (3.5) is obtained thanks to the strong Markov property. The induction hypothesis is applied in line (3.6). This achieves the induction for (3.4).

Let us now prove (3.3). We have

$$\begin{aligned}
\mathbb{E}_x[g(X_{S_{n+1,2\varepsilon}})] &= \mathbb{E}_x[g(X_{S_{n+1,2\varepsilon}})\mathbb{1}_{R^\varepsilon_{n,0}<T_1}] + \mathbb{E}_x[g(X_{S_{n+1,2\varepsilon}})\mathbb{1}_{R^\varepsilon_{n,0}\geqslant T_1}] \\
&= \mathbb{E}_x[g(X_{R^\varepsilon_{n,0}})\mathbb{1}_{R^\varepsilon_{n,0}<T_1}] + \mathbb{E}_x[g(X_{T_1+\theta(T_1)S_{n,\varepsilon}})\mathbb{1}_{R^\varepsilon_{n,0}\geqslant T_1}] \\
&= \mathbb{E}_x[g(X_{R^\varepsilon_{n,0}})\mathbb{1}_{R^\varepsilon_{n,0}<T_1}] + \mathbb{E}_x[\mathbb{E}_{Z_1}[g(X_{S_{n,\varepsilon}})]\mathbb{1}_{R^\varepsilon_{n,0}\geqslant T_1}] \\
&\geqslant \mathbb{E}_x[g(X_{R^\varepsilon_{n,0}})\mathbb{1}_{R^\varepsilon_{n,0}<T_1}] + \mathbb{E}_x[\mathbb{V}_n(Z_1)\mathbb{1}_{R^\varepsilon_{n,0}\geqslant T_1}] - \varepsilon \times \mathbb{P}_x(R^\varepsilon_{n,0}\geqslant T_1) \\
&\geqslant \mathbb{E}_x[g(X_{R^\varepsilon_{n,0}})\mathbb{1}_{R^\varepsilon_{n,0}<T_1}] + \mathbb{E}_x[\mathbb{V}_n(Z_1)\mathbb{1}_{R^\varepsilon_{n,0}\geqslant T_1}] - \varepsilon.
\end{aligned}$$

From the definition of $R_{n,\varepsilon}$, we readily obtain the following.

- If $\mathbf{K}\mathbb{V}_n(x) > \sup_{t\geqslant 0}\mathbf{J}(\mathbb{V}_n, g)(x,t)$, then $\mathbb{V}_{n+1}(x) = \mathbf{K}\mathbb{V}_n(x) = \mathbb{E}_x[\mathbb{V}_n(Z_1)]$ and $R^\varepsilon_{n,0} = t^*(x)$. So the previous inequality becomes
$$\begin{aligned}
\mathbb{E}_x[g(X_{S_{n+1,2\varepsilon}})] &\geqslant \mathbb{E}_x[g(X_{t^*(x)})\mathbb{1}_{t^*(x)<T_1}] + \mathbb{E}_x[\mathbb{V}_n(Z_1)\mathbb{1}_{T_1\leqslant t^*(x)}] - \varepsilon \\
&= \mathbf{K}\mathbb{V}_n(x) - \varepsilon \\
&= \mathbb{V}_{n+1}(x) - \varepsilon.
\end{aligned}$$
- Otherwise, $\mathbb{V}_{n+1}(x) = \sup_{t\geqslant 0}\mathbf{J}(\mathbb{V}_n, g)(x,t)$, and the previous inequality becomes
$$\begin{aligned}
\mathbb{E}_x[g(X_{S_{n+1,2\varepsilon}})] &\geqslant \mathbf{J}(\mathbb{V}_n, g)(x, r_{0,\varepsilon}(x)) - \varepsilon \\
&\geqslant \sup_{t\geqslant 0}\mathbf{J}(\mathbb{V}_n, g)(x,t) - \varepsilon - \varepsilon \\
&= \mathbb{V}_{n+1}(x) - 2\varepsilon.
\end{aligned}$$

Thanks to these two cases, we prove (3.3) and we complete the proof. □

3.4. Unbounded horizon

We now consider the unbounded horizon optimal stopping problem

$$\mathbb{V}_\infty(x) = \sup_{\tau\in\mathcal{M}_\infty} \mathbb{E}_x[g(X_\tau)],$$

where $\mathcal{M}_\infty$ denotes the family of all stopping times which are $\mathbb{P}_x$-a.s. finite for all $x \in E$. The value function $\mathbb{V}_\infty$ can be readily obtained by an infinite number of iterations of the dynamic programming operator $\mathbf{L}$.

Theorem 3.2. *The sequence $(\mathbb{V}_n)_n$ converges pointwise to the function $\mathbb{V}_\infty$.*

Proof. Let us fix $x \in E$. First, it is obvious that

$$\mathbb{V}_n(x) \leqslant \mathbb{V}_\infty(x) \quad \text{for all } n \in \mathbb{N},$$

using Theorem 3.1 and that $\mathcal{M}_n \subset \mathcal{M}_\infty$. Thus $\limsup_{n\to\infty}\mathbb{V}_n(x) \leqslant \mathbb{V}_\infty(x)$. Let us now show the converse inequality $\mathbb{V}_\infty(x) \leqslant \liminf_{n\to\infty}\mathbb{V}_n(x)$ to have equality. By definition of the supremum, for all $\varepsilon > 0$, there exists a stopping time $T_\varepsilon \in \mathcal{M}_\infty$ such that

$$\mathbb{V}_\infty(x) = \sup_{\tau\in\mathcal{M}_\infty} \mathbb{E}_x[g(X_\tau)] \leqslant \mathbb{E}_x[g(X_{T_\varepsilon})] + \varepsilon.$$

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

By the dominated convergence theorem we have

$$\mathbb{E}_x[g(X_{T_\varepsilon})] = \lim_{n\to\infty} \mathbb{E}_x[g(X_{T_\varepsilon \wedge T_n})] \leqslant \liminf_{n\to\infty} \mathbb{V}_n(x).$$

Thus, for all $\varepsilon > 0$,

$$\mathbb{V}_\infty(x) \leqslant \liminf_{n\to\infty} \mathbb{V}_n(x) + \varepsilon,$$

and this completes the proof. □

4. Comparison between the tagged cell and the measure-valued process

In this section we investigate whether stopping a single well-chosen individual is equivalent to stopping the whole population. When dealing with a branching population, some important characteristics of the global population, e.g. laws of large numbers for functionals of the individuals, can be obtained by simply studying a suitably weighted random tagged lineage, by means of many-to-one formulas; see e.g. [3], [6], [12], and [16]. We prove that this property does not hold true for the optimal stopping problem. We use the simple toy example of cell division from Section 2.5 to show that stopping a suitably chosen tagged cell and stopping the whole population yield different value functions.

4.1. Tagged cell and many-to-one formula

Let us briefly describe the many-to-one formula and the definition of the tagged cell as presented in [12]. Heuristically, picking a branch uniformly at random along the genealogical tree describes the path of the so-called tagged cell. Many-to-one formulas state that the expectation of a functional of the tagged cell at some given time t is equal to the expectation of a corresponding suitably chosen functional of the whole population of cells present at time t.

More precisely, for any measurable positive function h and any $t \geqslant 0$, we have the following formula, commonly called the *many-to-one* formula:

$$\mathbb{E}_x\left[\mathrm{e}^{-rt}\int_E h(x)X_t(\mathrm{d}x)\right] = \sum_{i=1}^{n} \mathbb{E}_{x_i}\left[\frac{h(\chi_t)}{\chi_t}\right] \times x_i, \tag{4.1}$$

where $x = \mu = \sum_{i=1}^{n} \delta_{(x_i,0)}$, $(X_t)_{t\geqslant 0}$ is the measure-valued PDMP of Section 2.5.2, and $(\chi_t)_{t\geqslant 0}$ is the tagged process, representing the evolution of the size of the tagged cell, and r is the common growth rate of all cells. Its dynamics is a real-valued PDMP described in Section 2.5.1 but whose parameters are different from those of $(X_t)_{t\geqslant 0}$; see [3], [6], and [12] for details.

To compare the value functions of the tagged cell and the measure-valued process, we will impose the form of our reward functions based on (4.1). For the tagged cell, we choose a bounded non-negative function f continuous along the flow. For the process $(X_t)_t$, the reward function g has the form

$$g\colon \sum_{i=1}^{n} \delta_{(x_i,t)} \in E \mapsto \sum_{i=1}^{n} f(x_i)\,\mathrm{e}^{-rt},$$

where we used the time-augmented process defined in Section 2.5.3 to take the time dependence into account.

4.2. Comparison of the value functions

Numerically computing the value function $\mathbb{V}_N$ is very demanding, as for each iteration of the dynamic programming operators, one needs to compute the functions on the whole state

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

space E, at least at first sight. In fact, when dealing with the optimal stopping problem with horizon N, the dynamic programming recursion on functions $\mathbb{V}_n$ can be rewritten as a recursion on the random variables $\mathbb{V}_k(Z_{N-k})$. Still, the recursion is numerically intractable as one needs to compute conditional expectations at each step. In order to avoid such intricacies, we simply consider the optimal stopping problem with horizon $N = 1$ jump.

We thus have to compute

$$\mathbb{V}_0(Z_1) = g(Z_1),$$
$$\mathbb{V}_1(Z_0) = \sup_{t\geqslant 0}\{\mathbb{E}_{Z_0}[\mathbb{V}_0(Z_1)\mathbb{1}_{T_1\leqslant t}] + g\circ\Phi(Z_0,t)\times\mathbb{P}_{Z_0}(T_1 > t)\} \vee \mathbb{E}_{Z_0}[\mathbb{V}_0(Z_1)].$$

The only quantities to look at are $\mathbb{V}_0(Z_1) = g(Z_1)$ and $\mathbb{V}_1(Z_0) = \mathbb{V}_1(\mu)$. Moreover, this last expression is deterministic if we choose a deterministic Z_0. More specifically, we have

$$\mathbb{V}_0(Z_1) = \mathrm{e}^{-rT_1}\left[\sum_{i\neq I_1}(x_i\,\mathrm{e}^{rT_1} - \gamma)\mathbb{1}_{x_i<\gamma\,\mathrm{e}^{-rT_1}} + (x_{I_1}\,\mathrm{e}^{rT_1} - 2\gamma)\mathbb{1}_{x_{I_1}<2\gamma\,\mathrm{e}^{-rT_1}} + (n+1)\gamma\right]$$

and

$$\mathbb{V}_1(Z_0) = \sup_{t>0}\left\{\mathrm{e}^{-rt}\mathbb{E}_x\left[\left(\sum_{i\neq I_1}(x_i\,\mathrm{e}^{rT_1} - \gamma)\mathbb{1}_{x_i<\gamma\,\mathrm{e}^{-rT_1}} + (x_{I_1}\,\mathrm{e}^{rT_1} - 2\gamma)\mathbb{1}_{x_{I_1}<2\gamma\,\mathrm{e}^{-rT_1}} + (n+1)\gamma\right)\mathbb{1}_{T_1\leqslant t}\right]\right.$$
$$\left. + \left(\sum_i (x_i\,\mathrm{e}^{rt} - \gamma)\mathbb{1}_{x_i<\gamma\,\mathrm{e}^{-rt}} + n\gamma\right)\exp\left(-\sum(x_i^\alpha)\frac{\mathrm{e}^{rt\alpha}-1}{\alpha r}\right)\right\} \vee \mathbb{E}_x[\mathbb{V}_0(Z_1)],$$

where the first jump time T_1 is distributed as (2.1) and the random variable I_1 is the rank of the split cell. Its distribution is given by

$$\mathbb{P}_\mu(I_1 = j) = \frac{x_j^\alpha}{x_1^\alpha + \cdots + x_n^\alpha}.$$

We can numerically simulate $\mathbb{V}_0(Z_1)$. For $\mathbb{V}_1(\mu)$, we fix a (large enough) maximum time and discretize the time interval in order to compute (an approximation of) the supremum. In the same way, the value function for the tagged cell is given by

$$\mathcal{V}_1(x) = \sup_{t\leqslant 0}\left\{[(x\,\mathrm{e}^{rt} - \gamma)\mathbb{1}_{x\,\mathrm{e}^{rt}<\gamma} + \gamma]\exp\left(-\frac{x^\alpha}{\alpha r}(\mathrm{e}^{r\alpha t} - 1)\right)\right.$$
$$\left. + \mathbb{E}_x[((x\,\mathrm{e}^{rT_1}/2 - \gamma)\mathbb{1}_{x\,\mathrm{e}^{rT_1}<2\gamma} + \gamma)\mathbb{1}_{T_1\leqslant t}]\right\} \vee \mathbb{E}_x[\mathcal{V}_0(\mathcal{Z}_1)].$$

For the numerical results, we use the following parameters: $\alpha = 1$, $r = 2$, $\gamma = 1$. For the discretization step, we evaluate the maximum value of T_1 by the Monte Carlo simulations and we divide it by *nbpt*, which corresponds to the number of discretization points for the evaluation of the supremum. Here, $nbpt = 10\,000$ and the Monte Carlo sample size is $N = 100\,000$. We obtain the following approximation:

$$\mathcal{V}_1(3) = 1.0018 \pm 2.54\times 10^{-4},$$
$$\mathbb{V}_1(\delta_3) = 1.3447 \pm 3.6034\times 10^{-4}.$$

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press

Hence, even on this simple toy example where the cost functions g and h match and with a short one-step horizon, the optimal performance for the global population and the tagged cell differ.

This means that the many-to-one formula (4.1) holds for fixed times t but not for stopping times. Note in addition that stopping times for the processes $(X_t)_{t\geqslant 0}$ and $(\chi_t)_{t\geqslant 0}$ are not easily comparable. Intuitively, the set of stopping times for the measure-valued process $(X_t)_{t\geqslant 0}$ is richer than that of the tagged cell $(\chi_t)_{t\geqslant 0}$. Thus it is not surprising that $\mathbb{V}_1(\delta_3) > \mathcal{V}_1(3)$.

It will be the object of future work to design specific numerical approximations of the value function of the global measure-valued population.

Acknowledgement

This work was partially supported by Région Languedoc-Roussillon and FEDER under grant Chercheur(se)s d'Avenir, project PROMMECE.

References

[1] Bäuerle, N. and Rieder, U. (2011). *Markov Decision Processes with Applications to Finance* (Universitext). Springer, Heidelberg.
[2] Benaïm, M. and Lobry, C. (2016). Lotka–Volterra with randomly fluctuating environments or 'how switching between beneficial environments can make survival harder'. *Ann. Appl. Prob.* **26**, 3754–3785.
[3] Bertoin, J. (2006). *Random Fragmentation and Coagulation Processes* (Cambridge Studies in Advanced Mathematics **102**). Cambridge University Press, Cambridge.
[4] Buckwar, E. and Riedler, M. G. (2011). An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *J. Math. Biol.* **63**, 1051–1093.
[5] Campillo, F., Champagnat, N. and Fritsch, C. (2016). Links between deterministic and stochastic approaches for invasion in growth–fragmentation–death models. *J. Math. Biol.* **73**, 1781–1821.
[6] Cloez, B. (2017). Limit theorems for some branching measure-valued processes. *Adv. Appl. Prob.* **49**, 549–580.
[7] Cloez, B., Dessalles, R., Genadot, A., Malrieu, F., Marguet, A. and Yvinec, R. (2017). Probabilistic and piecewise deterministic models in biology. *ESAIM: Procs* **60**, 225–245.
[8] Costa, M. (2016). A piecewise deterministic model for a prey–predator community. *Ann. Appl. Prob.* **26**, 3491–3530.
[9] Davis, M. H. A. (1984). Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. R. Statist. Soc. Ser. B* **46**, 353–388.
[10] Davis, M. H. A. (1993). *Markov Models and Optimization* (Monographs on Statistics and Applied Probability **49**). Chapman & Hall, London.
[11] de Saporta, B., Dufour, F. and Zhang, H. (2015). *Numerical Methods for Simulation and Optimization of Piecewise Deterministic Markov Processes: Application to Reliability* (Mathematics and Statistics *Series*). Wiley-ISTE.
[12] Doumic, M., Hoffmann, M., Krell, N. and Robert, L. (2015). Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli* **21**, 1760–1799.
[13] Fournier, N. and Méléard, S. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Prob.* **14**, 1880–1919.
[14] Genadot, A. and Thieullen, M. (2014). Multiscale piecewise deterministic Markov process in infinite dimension: central limit theorem and Langevin approximation. *ESAIM Probab. Statist.* **18**, 541–569.
[15] Gugerli, U. S. (1986). Optimal stopping of a piecewise-deterministic Markov process. *Stochastics* **19**, 221–236.
[16] Guyon, J. (2007). Limit theorems for bifurcating Markov chains: application to the detection of cellular aging. *Ann. Appl. Prob.* **17**, 1538–1569.
[17] Kallenberg, O. (2017). *Random Measures, Theory and Applications* (Probability Theory and Stochastic Modelling **77**). Springer, Cham.
[18] Kolokoltsov, V. N. (2011). *Markov Processes, Semigroups and Generators* (Gruyter Studies in Mathematics **38**). Walter de Gruyter, Berlin.
[19] Riedler, M. G., Thieullen, M. and Wainrib, G. (2012). Limit theorems for infinite-dimensional piecewise deterministic Markov processes: applications to stochastic excitable membrane models. *Electron. J. Prob.* **17** (55), 48.
[20] Robert, L., Hoffmann, M., Krell, N., Aymerich, S., Robert, J. and Doumic, M. (2014). Division in *Escherichia coli* is triggered by a size-sensing rather than a timing mechanism. *BMC Biology* **12**, 17.

https://doi.org/10.1017/jpr.2020.18 Published online by Cambridge University Press