Approximation methods for piecewise deterministic Markov processes and their costs

Author(s):
Kritzer, Peter; Leobacher, Gunther; Szölgyenyi, Michaela; Thonhauser, Stefan

Approximation methods for piecewise deterministic Markov processes and their costs

Peter Kritzer, Gunther Leobacher, Michaela Szölgyenyi & Stefan Thonhauser

Published online: 09 Jan 2019.

Submit your article to this journal ☑

Article views: 149

View Crossmark data ☑

Taylor & Francis
Taylor & Francis Group

Approximation methods for piecewise deterministic Markov processes and their costs

Peter Kritzer[a], Gunther Leobacher[b], Michaela Szölgyenyi[c,d] and Stefan Thonhauser[e]

[a] Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Linz, Austria; [b] Institute for Mathematics and Scientific Computing, University of Graz, Graz, Austria; [c] Institute for Statistics, University of Klagenfurt, Klagenfurt, Austria; [d] Seminar for Applied Mathematics and RiskLab Switzerland, ETH Zurich, Zurich, Switzerland; [e] Institute for Statistics, Graz University of Technology, Graz, Austria

ABSTRACT

In this paper, we analyse piecewise deterministic Markov processes (PDMPs), as introduced in Davis (1984). Many models in insurance mathematics can be formulated in terms of the general concept of PDMPs. There one is interested in computing certain quantities of interest such as the probability of ruin or the value of an insurance company. Instead of explicitly solving the related integro-(partial) differential equation (an approach which can only be used in few special cases), we adapt the problem in a manner that allows us to apply deterministic numerical integration algorithms such as quasi-Monte Carlo rules; this is in contrast to applying random integration algorithms such as Monte Carlo. To this end, we reformulate a general cost functional as a fixed point of a particular integral operator, which allows for iterative approximation of the functional. Furthermore, we introduce a smoothing technique which is applied to the integrands involved, in order to use error bounds for deterministic cubature rules. We prove a convergence result for our PDMPs approximation, which is of independent interest as it justifies phase-type approximations on the process level. We illustrate the smoothing technique for a risk-theoretic example, and compare deterministic and Monte Carlo integration.

1. Introduction

Many models in risk theory can be formulated as piecewise deterministic Markov processes (PDMPs) – a general class of finite-variation sample path Markov processes introduced by Davis (1984). This applies, among others, to the classical Cramér–Lundberg model, the renewal risk models, and multi-portfolio models recently introduced by Albrecher & Lautscham (2015). Moreover, PDMPs are sufficiently general to allow for non-constant model parameters, i.e. quantities such as the hazard rate or the premium rate may be state dependent. Examples of PDMPs and their control in the field of insurance mathematics are, e.g. Dassios & Embrechts (1989), Embrechts & Schmidli (1994), Schäl (1998), Rolski (1999), Cai et al. (2009), Leobacher & Ngare (2016), and Eichler et al. (2017).

The general theory of PDMPs is well developed, see for example the monographs by Davis (1993), Jacobsen (2006), or Bäuerle & Rieder (2011) for general results on PDMPs and their optimal control. More specialised contributions to the control theory of PDMPs can be found in Davis (1993), Lenhart & Liaot (1985), Costa & Davis (1989), Dempster & Ye (1992), Almudevar (2001), Forwick et al. (2004),

---

CONTACT Stefan Thonhauser ✉ stefan.thonhauser@math.tugraz.at 🏢 Institute for Statistics, Graz University of Technology, Kopernikusgasse 24/III, 8010 Graz, Austria

Bäuerle & Rieder (2010), Costa & Dufour (2013), or Davis & Farid (1999) for viscosity solutions of associated Hamilton–Jacobi–Bellman equations, and Colaneri (2017) for a general comparison principle for solutions to control problems for PDMPs.

For the numerical treatment of (control) problems for PDMPs, however, only problem-specific solutions have been provided. A standard approach is to link expected values representing a quantity of interest in the problem to the solution of an associated integro-(partial) differential equation, see, e.g. Asmussen & Albrecher (2010). In only very few cases is it possible to derive an explicit solution to this integro-(partial) differential equation. Requiring an explicit solution typically restricts the complexity of the model significantly. One possibility is to solve the integro-(partial) differential equation numerically. This carries all the intricacies and difficulties of a combined numerical method for differential and integral equations. Alternatively one can apply crude Monte Carlo methods, see, e.g. Riedler (2013). Those methods, while robust, are limited in speed by the Monte Carlo convergence rate. Another – highly sophisticated – approach uses quantisation of the jump distribution, see de Saporta et al. (2016).

In this article we concentrate on particularly easy to implement methods similar to Monte Carlo. The aim is to adapt the problem in a way that also allows for deterministic numerical integration algorithms such as quasi-Monte Carlo (QMC). QMC has been applied successfully to problems in risk theory, see Tichy (1984), Coulibaly & Lefèvre (2008), Siegl & Tichy (2000), Albrecher & Kainhofer (2002), and Preischl et al. (2018). It should be noted that the finiteness of the total variation needed for the convergence estimate (Albrecher & Kainhofer 2002, Theorem 1) has not been proven.

We would like to highlight two features of our approach. Inspired by Albrecher & Kainhofer (2002), we reformulate a general cost functional as a fixed point of a particular integral operator, which allows for iterative approximation of the functional. In terms of numerical integration this means that we get a high-dimensional integration problem of fixed dimension, where the dimension is a multiple of the number of iterations. Having a fixed dimension is required for the application of standard QMC or other deterministic cubature rules.

The application of QMC requires some degree of regularity of the integrand. Only in rare cases these will be satisfied automatically. The examples from risk theory considered here lead to non-smooth integrands. For these situations, we introduce a smoothing technique which, in its simplest case, leads to C^2 integrands. From the earlier considerations, we obtain deterministic error bounds for those. We prove convergence in distribution of the 'smoothed processes' to the original ones, which implies convergence of the corresponding expected values for every initial value of the process. In Section 2.1 we even obtain uniform convergence with respect to the initial value in a particular setup from risk theory.

Our convergence result has an additional benefit for a typical situation in risk-theoretic modelling. In the literature on the analysis of ruin probabilities, or more generally, on Gerber–Shiu functions, the assumption of a claim size distribution of mixed exponential or phase-type form is quite common. Apart from the possibility to obtain explicit expressions for quantities of interest in such setups, this modelling approach is motivated by the fact that the class of phase-type distributions is dense in the class of distributions with support on $[0, \infty)$, see Rolski (1999, Theorem 8.2.3). Under mild assumptions on the claim size distribution we want to approximate, our convergence result applies and justifies the phase-type approximation procedure even on the process level. Furthermore, efficient and easy to implement numerical methods for the computation of important targets such as Gerber-Shiu functions and expected discounted future dividend payments of an insurance company are of particular importance when models become more general and hence also more complicated. This makes our contribution valuable from both the analytical and the numerical point of view.

We would like to emphasise that the methods presented here per se do not provide solutions to optimal control problems, which is the main application of PDMPs in risk theory. However, the integration algorithms as introduced here can be used in a policy iteration procedure for calculating costs associated with a fixed policy.

The paper is structured as follows. In Section 2 we recall the definition of a PDMP and provide some risk-theoretic examples. In Section 3 we derive the fixed point approach for valuation of a cost functional of a PDMP. Section 4 reviews deterministic numerical integration of possibly multivariate C^k functions. Subsequently, Section 5 is devoted to the aforementioned smoothing procedure, and presents a stability result. Section 6 contains an application of the smoothing to one of the risk-theoretic examples and a comparative study of deterministic and Monte Carlo integration for this example.

2. Piecewise deterministic Markov processes

In this section we first define PDMPs. Then we give a couple of examples of practical interest.

A PDMP is a continuous-time stochastic process with (possibly random) jumps, which follows a deterministic flow, e.g. the solution of an ordinary differential equation (ODE), between jump times. We will not give the most general definition of PDMPs here, but instead refer to the monograph by Davis (1993). For a subset A of $\mathbb{R}^d$ we denote by $A^\circ, \bar{A}$, and ∂A its interior, closure, and boundary, respectively. We write $\mathcal{B}(A)$ for the Borel σ-algebra on A.

Definition 2.1: Let $A \subseteq \mathbb{R}^d$. A function $\phi \colon A \times \mathbb{R} \to \mathbb{R}^d$ is called a *flow* on A, if

- ϕ is continuous,
- $\phi(x, 0) = x$ for all $x \in A$;
- for all $x \in A$ and all $s, t \in \mathbb{R}$ it holds that if $\phi(x, t) \in A$ and $\phi(\phi(x, t), s) \in A$ then $\phi(x, t + s) = \phi(\phi(x, t), s)$.

For fixed $x \in A$, let $\phi^{-1}(x, A) = \{t \in \mathbb{R} \colon \phi(x, t) \in A\}$. Then the function $\phi(x, \cdot) \colon \phi^{-1}(x, A) \to A$ is called a *trajectory* of the flow.

If ϕ is a flow on A, then we write $\partial_\phi^- A = \{x \in \partial A \colon \exists \varepsilon \in (0, \infty) \text{ such that } \forall t \in (0, \varepsilon) \colon \phi(x, t) \in A^\circ\}$ and $\partial_\phi^+ A = \{x \in \partial A \colon \exists \varepsilon \in (0, \infty) \text{ such that } \forall t \in (0, \varepsilon) \colon \phi(x, -t) \in A^\circ\}$.

Thus $\partial_\phi^- A$ consists of the points on the boundary of A from which the trajectory moves into A° immediately, and $\partial_\phi^+ A$ consists of the points on the boundary of A to which a trajectory moves from A° without passing other points on the boundary in-between. Furthermore, we write $\partial_\phi^1 A \colon = \partial_\phi^- A \backslash \partial_\phi^+ A$.

Remark 2.2: The classical example of a flow arises through ODEs. Let $g \colon \mathbb{R}^d \to \mathbb{R}^d$ be Lipschitz continuous. By the classical Picard–Lindelöf theorem on existence and uniqueness of solutions of ODEs we have that for every $x \in \mathbb{R}$ there exists a continuously differentiable function $\kappa \colon \mathbb{R} \to \mathbb{R}^d$ such that $\kappa(0) = x$ and $\kappa'(s) = g(\kappa(s))$ for all $s \in \mathbb{R}$. For $t \in \mathbb{R}$ we define $\phi(x, t) = \kappa(t)$. The function ϕ defines a flow on $\mathbb{R}^d$. If $A \subseteq \mathbb{R}^d$, then the restriction of ϕ to $A \times \mathbb{R}$ is a flow on A.

Definition 2.3: Let K be a finite set and let $d \colon K \to \mathbb{N}$ be a function which satisfies that, for every $k \in K$, $E_k \subseteq \mathbb{R}^{d(k)}$ and ϕ_k is a flow on E_k with $E_k = E_k^\circ \cup \partial_{\phi_k}^1 E_k$.

- (i) The *state space* $(E, \mathcal{E})$ of a PDMP is the measurable space defined by $E = \bigcup_{k \in K}(\{k\} \times E_k)$ and $\mathcal{E} = \sigma(\{\{k\} \times B \colon k \in K, B \in \mathcal{B}(E_k)\})$.
- (ii) The *flow* of a PDMP is defined by $\phi = \{\phi_k\}_{k \in K}$.
- (iii) The *active boundary* of the PDMP is defined by $\Gamma^* = \bigcup_{k=1}^K \partial_{\phi_k}^+ E_k$. Furthermore, we define a σ-algebra on $E \cup \Gamma^*$ by $\mathcal{E}^* = \sigma(\{\{k\} \times B \colon k \in K, B \in \mathcal{B}(E_k \cup \partial_{\phi_k}^+ E_k)\})$.
- (iv) The *jump intensity* λ of a PDMP is defined by a family of functions $\lambda = \{\lambda_k\}_{k \in K}$ with $\lambda_k \colon E_k \to [0, \infty)$ measurable and bounded for all $k \in K$.

(v) The *jump kernel Q* of a PDMP is a function $Q\colon \mathcal{E} \times (E \cup \Gamma^*) \to [0,1]$ such that $Q(A,\cdot)$ is $\mathcal{E}^*$-$\mathcal{B}([0,1])$ measurable for every $A \in \mathcal{E}$, and $Q(\cdot,x)$ is a probability measure on $(E,\mathcal{E})$ for every $x \in E$ with $Q(\{x\},x) = 0$.

We call the triple (ϕ,λ,Q) the *local characteristics* of a PDMP.

Given a state space $(E,\mathcal{E})$ and local characteristics (ϕ,λ,Q) of a PDMP we define the function $t^*\colon E \to [0,\infty]$ by

$$
t^*(k,y) = \begin{cases} \inf\{t > 0 : \phi_k(y,t) \in \partial^+_{\phi_k} E_k\} & \text{if } \exists t > 0 : \phi_k(y,t) \in \partial^+_{\phi_k} E_k, \\ \infty & \text{otherwise.} \end{cases}
$$

Definition 2.4: Let $(E,\mathcal{E})$ be a state space and let (ϕ,λ,Q) be local characteristics of a PDMP, let $x \in E$, and let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space. A *piecewise deterministic Markov process* starting in x is a stochastic process $X\colon [0,\infty) \times \Omega \to E$ which satisfies the following. There exists a sequence of random variables $(T_n)_{n\in\mathbb{N}}$ with $T_n \in [0,\infty]$ and $T_n \leq T_{n+1}$ a.s. and $\lim_{n\to\infty} T_n = \infty$ a.s. for all $n \in \mathbb{N}$ such that

(i) it holds $\mathbb{P}$-a.s. that $X_0 = x$,
(ii) for all $n \in \mathbb{N}$, $t \in [T_n, T_{n+1})$, and for $(k,y) \in E$ with $X_{T_n} = (k,y)$ it holds $\mathbb{P}$-a.s. that $X_t = \phi_k(y, t - T_n)$,
(iii) for all $s,t \in [0,\infty)$ it holds $\mathbb{P}$-a.s. that

$$
\mathbb{P}\big(T_{n+1} - T_n > t | X_s = (k,y) \quad \text{and} \quad T_n \leq s < T_{n+1}\big)
$$
$$
= \begin{cases} e^{-\int_0^t \lambda_k(\phi_k(y,\tau))\,d\tau} & \text{if } 0 < t < t^*(k,y), \\ 0 & \text{if } t \geq t^*(k,y), \end{cases}
$$

(iv) for all $n \in \mathbb{N}$ and all $A \in \mathcal{E}$ it holds $\mathbb{P}$-a.s. that

$$
\mathbb{P}\big(X_{T_{n+1}} \in A | X_{T_n-}\big) = Q(A, X_{T_n}).
$$

Theorem 2.5: *Let $(E,\mathcal{E})$ be a state space and let (ϕ,λ,Q) be local characteristics of a PDMP, let $x \in E$. There exist a probability space $(\Omega,\mathcal{F},\mathbb{P}_x)$ and a stochastic process $X\colon [0,\infty) \times \Omega \to E$ such that X is a PDMP starting in x with state space E and local characteristics (ϕ,λ,Q). Furthermore, X has the strong Markov property.*

Proof: The proof of Theorem 2.5 for a more general setup that also allows for the possibility of explosions and countable K can be found in Davis (1993, Section 2.25). ∎

Figure 1 illustrates a path of a PDMP.

Let $f\colon E \to \mathbb{R}$ be a function. For all $k \in K$ we denote by f_k the function $f_k\colon E_k \to \mathbb{R}$ which satisfies for all $x \in E_k$ that $f_k(x) = f(k,x)$. It is not hard to see that f is measurable if and only if f_k is measurable for every $k \in K$. We say that f is n-times continuously differentiable, if for every $k \in K$ there exists an open set $A_k \subseteq \mathbb{R}^{d(k)}$ with $E_k \subseteq A_k$ and an n-times continuously differentiable function $\hat{f}_k : A_k \to \mathbb{R}$ such that $f_k = \hat{f}_k|_{E_k}$. We write $C^n(E,\mathbb{R}^m)$ for the space of n-times differentiable functions on E and $C^n_b(E,\mathbb{R}^m)$ for the space of functions in $C^n(E,\mathbb{R}^m)$ for which all derivatives are bounded. Moreover, $C^n_0(E,\mathbb{R}^m)$ is the space of functions in $C^n_b(E,\mathbb{R}^m)$ for which all derivatives vanish at infinity.

Further, for $f\colon E \to \mathbb{R}$, a PDMP X, and $t \in (0,\infty)$ we write $\mathbb{E}(f(X_t)|X_0 = x) =: \mathbb{E}_x(f(X_t))$.

In the remainder of this section we provide some illustrative examples from risk theory. For other examples and applications in different fields we refer to Davis (1993), de Saporta et al. (2012), and Riedler (2013).
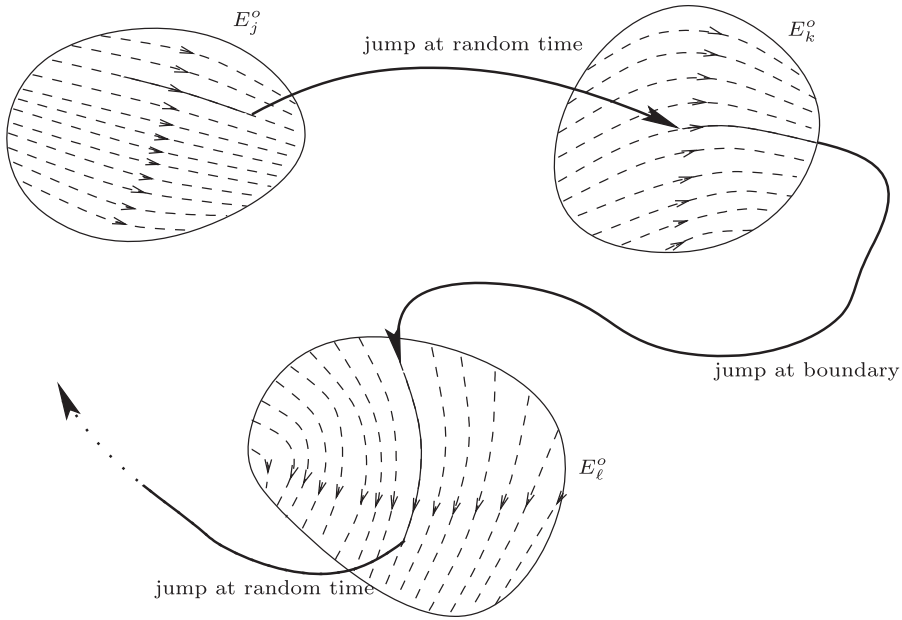
Figure 1. Illustration of a PDMP.

2.1. Examples

2.1.1. Classical Cramér–Lundberg model

Let $X = (X_t)_{t \geq 0}$ be a stochastic process given by

$$X_t = x + ct - S_t, \quad t \geq 0, \tag{1}$$

where x, $c \geq 0$, $N = (N_t)_{t \geq 0}$ is a homogeneous Poisson process with intensity $\lambda_N > 0$, $\{Y_i\}_{i \in \mathbb{N}}$ is a family of positive i.i.d. random variables with distribution function F_Y, and $S_t = \sum_{i=1}^{N_t} Y_i$ for all $t \geq 0$. A usual assumption in this kind of model is the independence of $\{Y_i\}_{i \in \mathbb{N}}$ and N. In risk theory the process X represents a standard model for the surplus of an insurance portfolio. A quantity of interest is the probability of X ever becoming negative, i.e. we are interested in $\mathbb{P}(\tau < \infty)$, where $\tau = \inf\{t \geq 0 : X_t < 0\}$. The model translates into a PDMP via

- $K = \{1, 2\}$,
- $E_1 = [0, \infty)$, $E_2 = (-\infty, 0)$,
- $\phi_1(y, t) = y + ct \forall y \in E_1$ and $\forall t \in \mathbb{R}$, $\phi_2(y, t) = y \forall y \in E_2$ and $\forall t \in \mathbb{R}$,
- $\lambda_1(y) = \lambda_N \forall y \in E_1$, $\lambda_2(y) = 0 \forall y \in E_2$.
- For $B_1 \in \mathcal{B}(E_1)$, $B_2 \in \mathcal{B}(E_2)$, and $B = (\{1\} \times B_1) \cup (\{2\} \times B_2)$,

$$Q(B, (1, y)) = \mathbb{P}(Y \in y - B_1) + \mathbb{P}(Y \in y - B_2)$$

for $y \in E_1$, and $Q(B, (2, y)) = \mathbb{P}(Y \in y - B_2)$,

where we have used the notation $y - B = \{y - y' : y' \in B\}$ for all $y \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R})$. For $y \in E_2$, any definition for Q will do, since the jump intensity is 0 there, but the above definition is provided for definiteness.

2.1.2. Cramér–Lundberg model with dividend payments

A classical modification of the model from Section 2.1.1 is the introduction of a dividend barrier at level $b > 0$. Then, once the surplus reaches the barrier, the incoming premium rate is immediately distributed as a dividend. Furthermore, if the process starts above b, the excess is distributed as a lump sum dividend, such that $X_{0+} = \min\{x, b\}$. A typical quantity of interest is the expected value of discounted future dividend payouts until ruin of the company, which is given by

$$V(x) = \begin{cases} \mathbb{E}_x \left(\int_0^\tau e^{-\delta t} c \mathbf{1}_{\{X_t = b\}} \, dt \right) & \text{if } x \le b, \\ x - b + \mathbb{E}_b \left(\int_0^\tau e^{-\delta t} c \mathbf{1}_{\{X_t = b\}} \, dt \right) & \text{if } x > b, \end{cases} \tag{2}$$

where $\delta > 0$ is a preference-based discount factor and $\tau = \inf\{t \ge 0 : X_t < 0\}$. The model translates into a PDMP via

- $K = \{1, 2, 3\}$,
- $E_1 = [0, b), E_2 = (-\infty, 0), E_3 = \{b\}$,
- $\phi_1(y, t) = y + ct \, \forall y \in E_1$ and $\forall t \in \mathbb{R}$, $\phi_2(y, t) = y \, \forall y \in E_2$ and $\forall t \in \mathbb{R}$, $\phi_3(y, t) = y \, \forall y \in E_3$ and $\forall t \in \mathbb{R}$,
- $\lambda_1(y) = \lambda_N \forall y \in E_1, \lambda_2(y) = 0 \, \forall y \in E_2, \lambda_3(y) = \lambda_N \, \forall y \in E_3$.
- For $B_k \in \mathcal{B}(E_k)$, $1 \le k \le 3$, and $B = (\{1\} \times B_1) \cup (\{2\} \times B_2) \cup (\{3\} \times B_3)$,

$$Q(B, (1, y)) = \mathbb{P}(Y \in y - B_1) + \mathbb{P}(Y \in y - B_2)$$

for $y \in E_1$, $Q(B, (2, y)) = \mathbb{P}(Y \in y - B_2)$ for $y \in E_2$, and

$$Q(B, (3, y)) = \mathbb{P}(Y \in y - B_1) + \mathbb{P}(Y \in y - B_2)$$

for $y \in E_3$. Finally, $Q(B, (1, y)) = \mathbf{1}_{B_3}(y)(3, y)$ for $y \in \partial^1_{\phi_1} E_1 = \{b\}$.

Note that only initial values $x \in (-\infty, b]$ translate to a viable initial value for the PDMP. However, this is sufficient for determining $V(x)$ for all $x \in \mathbb{R}$ via (2).

2.1.3. Cramér–Lundberg model with time dependent dividend barrier

In Albrecher & Kainhofer (2002) the model from Section 2.1.2 is further extended to include a time dependent barrier $b \colon [0, \infty) \to]0, \infty)$ of the form

$$b(t) = \left(b_0^m + \frac{t}{\alpha} \right)^{1/m},$$

where $\alpha, b_0 > 0$, $m > 1$. The quantity of interest is again the expected value of discounted future dividend payments until the time of ruin, i.e.

$$V(x) = \mathbb{E}_x \left(\int_0^\tau e^{-\delta t} (c - b_t) \mathbf{1}_{\{X_t = b_t\}} \, dt \right),$$

for $x \le b_0$, where again $\tau = \inf\{t \ge 0 : X_t < 0\}$ and $\delta > 0$ is a preference-based discount factor. The model translates into a PDMP via

- $K = \{1, 2, 3\}$,
- $E_1 = \{(s, y) \in \mathbb{R}^2 : 0 \le y < b(s)\}, E_2 = \{(s, y) \in \mathbb{R}^2 : y < 0\}, E_3 = \{(s, y) \in \mathbb{R}^2 : y = b(s)\}$,
- $\phi_1((s, y), t) = (s + t, y + ct) \forall (s, y) \in E_1$ and $\forall t \in \mathbb{R}$, $\phi_2((s, y), t) = (s + t, y) \, \forall y \in E_2$ and $\forall t \in \mathbb{R}$, $\phi_3((s, y), t) = (s + t, b(s + t)) \, \forall (s, y) \in E_3$ and $\forall t \in \mathbb{R}$,
- $\lambda_1(y) = \lambda_N \forall y \in E_1, \lambda_2(y) = 0 \, \forall y \in E_2, \lambda_3(y) = \lambda_N \, \forall y \in E_3$.

- For $B_k \in \mathcal{B}(E_k)$, $1 \leq k \leq 3$, and $B = (\{1\} \times B_1) \cup (\{2\} \times B_2) \cup (\{3\} \times B_3)$,

$$Q(B, (1, (s, y))) = \mathbb{P}(Y \in y - (\{s\} \times \mathbb{R}) \cap B_1) + \mathbb{P}(Y \in y - (\{s\} \times \mathbb{R}) \cap B_2)$$

for $(s, y) \in E_1$, $Q(B, (2, (s, y))) = \mathbb{P}(Y \in y - (\{s\} \times \mathbb{R}) \cap B_2)$ for $(s, y) \in E_2$, and

$$Q(B, (3, (s, y))) = \mathbb{P}(Y \in y - (\{s\} \times \mathbb{R}) \cap B_1) + \mathbb{P}(Y \in y - (\{s\} \times \mathbb{R}) \cap B_2)$$

for $(s, y) \in E_3$. Finally, $Q(B, (1, (s, y))) = \mathbf{1}_{B_3}((s, y))(3, (s, y))$ for $(s, y) \in \partial_{\phi_1}^1 E_1 = E_3$.

2.1.4. Cramér–Lundberg model with loan

In Dassios & Embrechts (1989) the model from Section 2.1.2 is modified such that the insurance company is not ruined when the surplus hits zero, but has the possibility to take up a loan at an interest rate $\rho > 0$. The time of ruin is given by $\tau = \inf\{t \geq 0 \colon X_t < -c/\rho\}$. The corresponding quantity of interest is

$$V(x) = \mathbb{E}_x \left(\int_0^\tau e^{-\delta t} c \mathbf{1}_{\{X_t = b\}} \, dt \right),$$

for $x \leq b$, where $\delta > 0$ is a preference-based discount factor. The model translates into a PDMP via

- $K = \{1, 2, 3, 4, 5\}$,
- $E_1 = [0, b)$, $E_2 = (-(c/\rho), 0)$, $E_3 = \{b\}$, $E_4 = (-\infty, -(c/\rho))$, $E_5 = \{-c/\rho\}$,
- $\phi_1(y, t) = y + ct \, \forall y \in E_1$ and $\forall t \in \mathbb{R}$, $\phi_2(y, t) = y \, \forall y \in E_2$ and $\forall t \in \mathbb{R}$, ϕ_3 is the flow of the ODE $z' = c + \rho z$ at $(y, t) \, \forall y \in E_3$ and $\forall t \in \mathbb{R}$, $\phi_4(y, t) = y \, \forall y \in E_4$ and $\forall t \in \mathbb{R}$, $\phi_5(y, t) = y \, \forall y \in E_5$ and $\forall t \in \mathbb{R}$,
- $\lambda_1(y) = \lambda_N \forall y \in E_1$, $\lambda_2(y) = \lambda_N \, \forall y \in E_2$, $\lambda_3(y) = \lambda_N \, \forall y \in E_3$, $\lambda_4(y) = 0 \, \forall y \in E_4$, $\lambda_5(y) = 0 \, \forall y \in E_5$.
- For $B_k \in \mathcal{B}(E_k)$, $1 \leq k \leq 5$, and $B = \bigcup_{k=1}^5 (\{k\} \times B_k)$,

$$Q(B, (1, y)) = \mathbb{P}(Y \in y - B_1) + \mathbb{P}(Y \in y - B_2) + \mathbb{P}(Y \in y - B_4)$$

for $y \in E_1$, $Q(B, (2, y)) = \mathbb{P}(Y \in y - B_2) + \mathbb{P}(Y \in y - B_4)$ for $y \in E_2$, and

$$Q(B, (3, y)) = \mathbb{P}(Y \in y - B_1) + \mathbb{P}(Y \in y - B_2)$$

for $y \in E_3$. Finally, $Q(B, (1, y)) = \mathbf{1}_{B_3}(y)(3, y)$ for $y \in \partial_{\phi_1}^1 E_1 = \{b\}$, and $Q(B, (2, y)) = \mathbf{1}_{B_2}(y)(1, y)$ for $y \in \partial_{\phi_2}^1 E_2 = \{0\}$.

2.1.5. Multidimensional Cramér–Lundberg model

In Albrecher & Lautscham (2015) a two-dimensional extension of the model in Section 2.1.2 is studied. The basis are independent surplus processes modelling two insurance portfolios $X_t^{(j)} = x^{(j)} + c^{(j)} t - S_t^{(j)}$, $j \in \{1, 2\}$, where $c^{(1)}, c^{(2)} \geq 0$ and $S^{(j)}$ are compound Poisson processes with intensities $\lambda^{(1)}$, $\lambda^{(2)}$ and jump size distributions $F_{Y^{(1)}}$, $F_{Y^{(2)}}$. Furthermore, $b^{(1)}, b^{(2)} \geq 0$ are barriers. As a new feature, the drift of the component at the barrier is added to the other component's drift, causing faster growth of the latter. Dividends are only paid when both surplus processes have reached their

individual barriers. We show how the model translates into a PDMP, namely

$$E_1 = \{(x^{(1)}, x^{(2)}) \in \mathbb{R}^2 : 0 \le x^{(1)} < b^{(1)}, 0 \le x^{(2)} < b^{(2)}\},$$

$$E_2 = \{(x^{(1)}, x^{(2)}) \in \mathbb{R}^2 : b^{(1)} = x^{(1)}, 0 \le x^{(2)} < b^{(2)}\},$$

$$E_3 = \{(x^{(1)}, x^{(2)}) \in \mathbb{R}^2 : 0 \le x^{(1)} < b^{(1)}, b^{(2)} = x^{(2)}\},$$

$$E_4 = \{(x^{(1)}, x^{(2)}) \in \mathbb{R}^2 : b^{(1)} = x^{(1)}, b^{(2)} = x^{(2)}\},$$

$$E_5 = \mathbb{R}^2 \setminus (E_1 \cup E_2 \cup E_3 \cup E_4).$$

The flow is given by

$$\phi_1(x, t) = x + \begin{pmatrix} c^{(1)} \\ c^{(2)} \end{pmatrix} t, \quad \phi_2(x, t) = x + \begin{pmatrix} 0 \\ c^{(1)} + c^{(2)} \end{pmatrix} t, \quad \phi_3(x, t) = x + \begin{pmatrix} c^{(1)} + c^{(2)} \\ 0 \end{pmatrix} t,$$

and $\phi_4(x, t) = \phi_5(x, t) = x$ for all $x \in \mathbb{R}^2$, $t \ge 0$. It remains to describe the jump behaviour. We get deterministic 'jumps' at the active boundaries of E_1, E_2, E_3 which do not manifest themselves as jumps of the process, i.e. $Q(A, (1, x)) = \mathbf{1}_A((2, x))$ for $(1, x) \in \partial^1_{\phi_1}(E_1)$ and similar for the other active boundaries. Since each surplus process is a compound Poisson process with drift, jumps in the components occur due to realisations of independent identically distributed exponential random variables (independence implies that mutual jumps occur with probability zero). The two-dimensional process thus jumps at the minimum of the individual jump times. This means that we have a constant jump intensity $\lambda_k = \lambda^{(1)} + \lambda^{(2)}$ for $k = 1, 2, 3, 4$, and $\lambda_5 = 0$. If a jump occurs at time $t \ge 0$, it happens with probability $\lambda^{(1)}/(\lambda^{(1)} + \lambda^{(2)})$ in the first surplus process with jump size distribution $F_{Y^{(1)}}$, and with probability $\lambda^{(2)}/(\lambda^{(1)} + \lambda^{(2)})$ in the second surplus process with jump size distribution $F_{Y^{(2)}}$. It remains to describe the jump kernel for the jumps from $x \in E$. To this end define, for $k_1, k_2 \in \{1, 2, 3, 4\}$ and $B \in \mathcal{B}(E_{k_2}) \subseteq \mathcal{B}(\mathbb{R}^2)$, and $(y^{(1)}, y^{(2)}) \in E_{k_1}$,

$$B^{(1)} = \{z^{(1)} \in \mathbb{R} : (z^{(1)}, z^{(2)}) \in B, z^{(2)} = y^{(2)}\},$$

$$B^{(2)} = \{z^{(2)} \in \mathbb{R} : (z^{(1)}, z^{(2)}) \in B, z^{(1)} = y^{(1)}\}.$$

Furthermore,

$$Q(\{k_2\} \times B, (k_1, y^{(1)}, y^{(2)})) = \frac{\lambda^{(1)}}{\lambda^{(1)} + \lambda^{(2)}} F_{Y^{(1)}}(y^{(1)} - B^{(1)}) + \frac{\lambda^{(2)}}{\lambda^{(1)} + \lambda^{(2)}} F_{Y^{(2)}}(y^{(2)} - B^{(2)}).$$

A quantity of interest in this model is again the expected value of discounted future dividend payments until the time of ruin of one of the portfolios,

$$V(x^{(1)}, x^{(2)}) = \mathbb{E}_{x^{(1)}, x^{(2)}} \left(\int_0^\tau e^{-\delta t} (c^{(1)} + c^{(2)}) \mathbf{1}_{E_4}(X_t^{(1)}, X_t^{(2)}) \, dt \right), \tag{3}$$

for $x^{(1)} \le b^{(1)}$, $x^{(2)} \le b^{(2)}$, with $\tau = \inf\{t \ge 0 : (X_t^{(1)}, X_t^{(2)}) \in E_5\}$, and $\delta > 0$ being a preference-based discount factor.

3. Iterated integrals and a fixed point approach

In this section we derive a method for numerical approximation of the quantities of interest appearing in the models introduced in the previous section. We rewrite the quantity of interest as a sum of integrals with fixed dimension and an error term that goes to zero exponentially fast with increasing dimension of the integral. This allows for the use of deterministic integration rules. The starting point

for the derivation of this integral representation is the observation that the quantity of interest is a fixed point of a certain integral operator associated to the PDMP.

Definition 3.1: Suppose there exists a set $K^c \subseteq K$ such that for all $k \in K^c$ it holds that $\lambda_k(x) = 0$, and $\phi_k(x, t) = x$ for all $x \in E_k$ and all $t \in \mathbb{R}$. We call $E^c := \bigcup_{k \in K^c} E_k$ a *cemetery* of the PDMP.

Definition 3.2: Let a PDMP be given and let $E^c \neq \emptyset$ be a cemetery of the PDMP. A *running reward function* $\ell \colon E \to \mathbb{R}$ is a measurable function satisfying $\ell|_{E^c} \equiv 0$. A *terminal cost function* $\Psi \colon E^c \to \mathbb{R}$ is a measurable function satisfying $\Psi|_{E \setminus E^c} \equiv 0$. The *cost functional* $V \colon E \to \mathbb{R}$ corresponding to E^c, ℓ, Ψ is defined by

$$V(x) = \mathbb{E}_x \left(\int_0^\tau e^{-\delta t} \ell(X_t) \, dt + e^{-\delta \tau} \Psi(X_\tau) \right), \tag{4}$$

where $\tau = \inf\{t \geq 0 \colon X_t \in E^c\}$.

Let T_1 be the first jump time. Equation (4) can be rewritten as follows,

$$
\begin{aligned}
V(x) = \mathbb{E}_x \Bigg[& \left(\int_0^{T_1} e^{-\delta t} \ell(\phi(x, t)) \, dt + \int_{T_1}^\tau e^{-\delta t} \ell(\phi(X_{T_1}, t - T_1)) \, dt + e^{-\delta \tau} \Psi(X_\tau) \right) \mathbf{1}_{\{T_1 < \tau\}} \\
& + \left(\int_0^\tau e^{-\delta t} \ell(\phi(x, t)) \, dt + e^{-\delta \tau} \Psi(\phi(x, \tau)) \right) \mathbf{1}_{\{\tau < T_1\}} \\
& + \left(\int_0^{T_1} e^{-\delta t} \ell(\phi(x, t)) \, dt + e^{-\delta T_1} \Psi(X_{T_1}) \right) \mathbf{1}_{\{T_1 = \tau\}} \Bigg].
\end{aligned}
$$

Since X is a PDMP and hence a strong Markov process, this yields $V = \mathcal{H} + \mathcal{G} V$ with $\mathcal{H} \colon E \to \mathbb{R}$, $\mathcal{G} \colon C^2(E, \mathbb{R}) \to \mathbb{R}$ defined by

$$
\begin{aligned}
\mathcal{H}(x) = \mathbb{E}_x \Bigg[& \left(\int_0^{T_1} e^{-\delta t} \ell(\phi(x, t)) \, dt \right) \mathbf{1}_{\{T_1 < \tau\}} \\
& + \left(\int_0^\tau e^{-\delta t} \ell(\phi(x, t)) \, dt + e^{-\delta \tau} \Psi(\phi(x, \tau)) \right) \mathbf{1}_{\{\tau < T_1\}} \\
& + \left(\int_0^{T_1} e^{-\delta t} \ell(\phi(x, t)) \, dt + e^{-\delta T_1} \Psi(X_{T_1}) \right) \mathbf{1}_{\{T_1 = \tau\}} \Bigg],
\end{aligned}
$$

$$\mathcal{G} V(x) = \mathbb{E}_x \left[e^{-\delta T_1} V(X_{T_1}) \mathbf{1}_{\{T_1 < \tau\}} \right]. \tag{5}$$

Recall that for every $t \geq 0$ it holds that $\mathbb{P}_x(T_1 > t) = \exp\left(-\int_0^t \lambda(\phi(x, s)) \, ds\right) =: 1 - F_W(t, x)$ and denote the corresponding density by f_W. With this, the function $\mathcal{H}$ and the operator $\mathcal{G}$ admit representations as integrals,

$$
\begin{aligned}
\mathcal{H}(x) = & \int_0^{t^*(x)} f_W(t, x) \left[\int_0^t e^{-\delta s} \ell(\phi(x, s)) \, ds + e^{-\delta t} \int_{E^c} \Psi(y) Q(dy, \phi(x, t)) \right] dt \\
& + (1 - F_W(t^*(x), x)) \left[\int_0^{t^*(x)} e^{-\delta s} \ell(\phi(x, s)) \, ds + e^{-\delta t^*(x)} \Psi(\phi(x, t^*(x))) \right],
\end{aligned}
$$

$$\mathcal{G} V(x) = \int_0^{t^*(x)} f_W(t, x) \, e^{-\delta t} \int_E V(y) Q(dy, \phi(x, t)) \, dt.$$

Note that $\mathcal{H}(x)$ corresponds to the expected discounted rewards collected before the first jump at time T_1 when starting in x. $\mathcal{G} V(x)$ represents the expected discounted rewards from time T_1 onwards

conditional on the event $\{X_{T_1} \notin E^c, X_0 = x\}$. Iterating the above steps $n \in \mathbb{N}$ times leads to

$$V(x) = \mathcal{G}^n V(x) + \sum_{i=0}^{n-1} \mathcal{G}^i \mathcal{H}(x). \tag{6}$$

Lemma 3.3: *Let $\Psi \colon E^c \to \mathbb{R}$ and $\ell \colon E \to \mathbb{R}$ be bounded, for all $k \in K$ assume that the functions λ_k are bounded by $C_\lambda \in (0, \infty)$, and for all $x \in E$ let $t^*(x) = \infty$. Then for all $x \in E$ and for all $n \in \mathbb{N}$ it holds that $|\mathcal{G}^n V(x)| \leq C_V (C_\lambda/(C_\lambda + \delta))^n$ and, in particular, it holds that $\lim_{n \to \infty} \mathcal{G}^n V(x) = 0$ uniformly in $x \in E$.*

Proof: The boundedness of ℓ and Ψ implies that also V is bounded by $C_V = (\|\ell\|_\infty/\delta) + \|\Psi\|_\infty$. Using the strong Markov property and Equation (5) we have by induction on n,

$$\begin{aligned}
\mathcal{G}^n V(x) &= \mathbb{E}_x \left[e^{-\delta T_1} \mathcal{G}^{n-1} V(X_{T_1}) 1_{\{T_1 < \tau\}} \right] \\
&= \mathbb{E}_x \left[e^{-\delta T_1} \mathbb{E}_{X_{T_1}} \left[e^{-\delta(T_n - T_1)} V(X_{T_n}) 1_{\{T_n < \tau\}} \right] 1_{\{T_1 < \tau\}} \right] \\
&= \mathbb{E}_x \left[\mathbb{E}_{X_{T_1}} \left[e^{-\delta T_n} V(X_{T_n}) 1_{\{T_n < \tau\}} 1_{\{T_1 < \tau\}} \right] \right] \\
&= \mathbb{E}_x \left[e^{-\delta T_n} V(X_{T_n}) \mathbf{1}_{\{\tau > T_n\}} \right],
\end{aligned} \tag{7}$$

where we used $1_{\{T_n < \tau\}} 1_{\{T_1 < \tau\}} = 1_{\{T_n < \tau\}}$ in the last equality. Recall that $\mathbb{P}(T_n - T_{n-1} > t \mid T_{n-1}, X_{T_{n-1}}) = \exp(-\int_0^t \lambda(\phi(s, X_{T_{n-1}})) \, ds) \geq \exp(-t \, C_\lambda)$. For every $n \in \mathbb{N}$ let $Z_n \sim$ Erlang (n, C_λ) be an Erlang-distributed random variable. Combining this with (7) we get that

$$\left| \mathcal{G}^n V(x) \right| \leq C_V \mathbb{E}_x \left[e^{-\delta T_n} \right] \leq C_V \mathbb{E} \left[e^{-\delta Z_n} \right] = C_V \left(\frac{C_\lambda}{C_\lambda + \delta} \right)^n.$$

The latter expression converges to zero as $n \to \infty$ uniformly in $x \in E$. ∎

Combining Lemma 3.3 with (6) results in the error estimate

$$\left| V(x) - \sum_{i=0}^{n-1} \mathcal{G}^i \mathcal{H}(x) \right| \leq C_V \left(\frac{C_\lambda}{C_\lambda + \delta} \right)^n. \tag{8}$$

Finally, we obtain the following representation,

$$\begin{aligned}
\mathcal{G}^{i-1} \mathcal{H}(x_0) &= \int_{t_1=0}^{t^*(x_0)} f_W(t_1, x_0) \, e^{-\delta t_1} \int_{x_1 \in E} \int_{t_2=0}^{t^*(x_1)} f_W(t_2, x_1) \, e^{-\delta t_2} \int_{x_2 \in E} \cdots \\
&\quad \times \int_{t_{i-1}=0}^{t^*(x_{k-2})} f_W(t_{i-1}, x_{i-2}) \, e^{-\delta t_{i-1}} \\
&\quad \int_{x_{i-1} \in E} \mathcal{H}(x_{i-1}) Q(dx_{i-1}, \phi(x_{i-2}, t_{i-1})) \, dt_{i-1} \cdots Q(dx_1, \phi(x_0, t_1)) \, dt_1 \\
&= \int_{t_1=0}^{t^*(x_0)} \int_{x_1 \in E} \cdots \int_{t_{i-1}=0}^{t^*(x_{i-2})} \int_{x_{i-1} \in E} \left(\prod_{j=1}^{i-1} f_W(t_j, x_{j-1}) \, e^{-\delta t_j} \right) \\
&\quad \mathcal{H}(x_{i-1}) Q(dx_{i-1}, \phi(x_{i-2}, t_{i-1})) \, dt_{i-1} \cdots Q(dx_1, \phi(x_0, t_1)) \, dt_1.
\end{aligned} \tag{9}$$

In (9) we denote by $\{t_j\}_{j \in \{1, \ldots, i-1\}}$ the family of inter-jump times and by $\{x_j\}_{j \in \{1, \ldots, i-1\}}$ the family of post-jump locations.

Remark 3.4: Solving the integral $\mathcal{G}^{i-1}\mathcal{H}(x_0)$ brings several advantages compared to a crude Monte Carlo approach. First, (9) is an integral with a fixed dimension. Hence, it can be approximated using deterministic integration rules like QMC, for which deterministic error bounds are available. Second, the bias of restricting oneself to a fixed number of jumps can be estimated uniformly in $x = x_0$ using the bias estimate in Lemma 3.3. Third, rare events like surviving a large number of jumps are – in this formulation – not rare in the sense that it is unlikely to draw such a realisation, which has the effect of importance sampling.

4. Cubature rules for C^κ-functions

In order to obtain convergence estimates for numerical integration methods such as QMC methods or other cubature rules, we need more regularity of the integrands than they admit in many practical applications. For example, we may need to bound a certain norm of the Hessian matrix of the integrand. In Section 5, we will rewrite the problem introduced in Section 3 so that the integrand is a function $f: [0,1]^d \to \mathbb{R}$ which satisfies $f \in C^2([0,1]^d)$, or more generally $f \in C^\kappa([0,1]^d)$ for some $\kappa \in \mathbb{N}$. We outline two different methods for treating such integrands f by cubature rules.

4.1. QMC methods

QMC methods are equal-weight cubature rules with M deterministically chosen integration nodes. Let the integrand $f: [0,1]^d \to \mathbb{R}$ satisfy $f \in C^2([0,1]^d)$. In order to obtain a convergence estimate for numerical integration of f using QMC, we require a so-called Koksma–Hlawka type inequality. The original Koksma–Hlawka inequality bounds the integration error of a QMC rule by the product of the variation of the integrand (in the sense of Hardy and Krause) and the so-called discrepancy of the integration node set (see, e.g. Niederreiter (1992, Chapter 2)). We remark, however, that we cannot easily apply the classical Koksma–Hlawka inequality in this paper, as we cannot rely on the integrands to have bounded variation in the sense of Hardy and Krause. Hence, we are going to resort to a variant of the Koksma–Hlawka inequality which was recently proven in Pausinger & Svane (2015). Let $Q_{M,d}(f) = 1/M \sum_{j=1}^{M} f(x_j)$ be a QMC rule using M integration nodes $x_1, \ldots, x_M \in [0,1)^d$. Then by Pausinger & Svane (2015, Theorem 3.12) we have

$$\left| \int_{[0,1]^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - Q_{M,d}(f) \right| \le \left(\sup_{\boldsymbol{x} \in [0,1]^d} f(\boldsymbol{x}) - \inf_{\boldsymbol{x} \in [0,1]^d} f(\boldsymbol{x}) + \frac{d}{16} M(f) \right) \mathrm{Disc}_{\mathrm{I}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M), \quad (10)$$

where $M(f) = \sup_{\boldsymbol{x} \in [0,1]^d} \|\mathrm{Hess}(f, \boldsymbol{x})\|$, $\mathrm{Hess}(f, \boldsymbol{x})$ is the Hessian matrix of f at $\boldsymbol{x}$, $\|\cdot\|$ denotes the usual operator norm, and where $\mathrm{Disc}_{\mathrm{I}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M)$ is the isotropic discrepancy of the integration node set,

$$\mathrm{Disc}_{\mathrm{I}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M) = \sup_{\substack{C \subseteq [0,1]^d \\ C \text{ convex}}} \left| \frac{1}{M} \sum_{j=1}^{M} \mathbf{1}_{\{x_j \in C\}} - \mu_d(C) \right|,$$

where μ_d denotes the Lebesgue measure on the $\mathbb{R}^d$. Now let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M \in [0,1]^d$. In Niederreiter (1992, Chapter 2) it is shown that

$$\mathrm{Disc}_{\mathrm{I}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M) \le 8d \left(\mathrm{Disc}_*(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M) \right)^{1/d},$$

where by $\mathrm{Disc}_*(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M)$ we denote the star discrepancy of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$, defined as

$$\mathrm{Disc}_*(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M) = \sup_{\boldsymbol{a} \in [0,1)^d} \left| \frac{1}{M} \sum_{j=1}^{M} \mathbf{1}_{\{x_j \in [\boldsymbol{0}, \boldsymbol{a})\}} - \mu_d([\boldsymbol{0}, \boldsymbol{a})) \right|,$$

where $[\boldsymbol{0}, \boldsymbol{a})$ denotes $[0, a_1) \times \cdots \times [0, a_d)$ for $\boldsymbol{a} = (a_1, \ldots, a_d)$. It is well known that common point sequences that are employed in QMC methods, such as Sobol' sequences or Halton sequences, have a

star discrepancy of order $(\log M)^d/M$ (and it is known that this order can, if at all, only be improved with respect to the exponent of the log-term). Hence, by using, e.g. Sobol' points in a QMC method for numerically integrating a C^2-function, we cannot expect an error that converges to zero faster than $(\log M)/M^{1/d}$.

As we shall see below, this order of magnitude can, with respect to the disadvantageous dependence on d, not be improved further for C^2-functions. However, there is room for improvement if we make additional smoothness assumptions on the integrand.

4.2. Product rules

In Hinrichs et al. (2017) it is shown that, by using products of Gauss rules, one can obtain the following result. Let $f : [0,1]^d \to \mathbb{R}$ be such that $f \in C^\kappa$ for some $\kappa \in \mathbb{N}$. Then, by using a product rule $Q_{G,\tilde{M},d}$ of $\tilde{M}$-point Gauss quadrature rules, one obtains

$$\left| \int_{[0,1]^d} f(\boldsymbol{x})d\boldsymbol{x} - Q_{G,\tilde{M},d}(f) \right| \leq c_\kappa d\tilde{M}^{-\kappa} \left\| f \right\|_{C^\kappa}, \quad \text{for } \tilde{M} \geq \kappa + 1, \tag{11}$$

where $c_\kappa = (\pi/2)(e/(6\sqrt{3}))^\kappa$, and where

$$\left\| f \right\|_{C^\kappa} = \max_{\substack{\boldsymbol{\beta} \in \mathbb{N}_0^d \\ \|\boldsymbol{\beta}\|_1 \leq \kappa}} \left\| D^{\boldsymbol{\beta}}(f) \right\|_{L_\infty},$$

where $D^{\boldsymbol{\beta}}$ denotes the (weak) partial derivative of order $\boldsymbol{\beta}$ for $\boldsymbol{\beta} \in \mathbb{N}_0^d$. A d-fold Gauss product rule as described above uses $M = \tilde{M}^d$ points in total, and hence yields a convergence order of $M^{-\kappa/d}$. It is known due to Bakhvalov (1959) that this convergence order is best possible. For the special case $\kappa = 2$, we only obtain a relatively small improvement over the bound implied by (10). However, there is an additional advantage in the bound (11). By requiring that the function f satisfies additional smoothness assumptions, namely that $f \in C^\kappa$ for some $\kappa \in \mathbb{N}$ which is possibly larger than 2, we obtain an improved convergence rate. Hence, we face a trade-off between imposing a higher degree of smoothness on the integrand f to obtain a higher accuracy in the quadrature rule, and the error we make by smoothing the integrand to that extent. It is therefore likely that the method needs to be fine-tuned on a case-by-case basis. In practice, product rules often cannot be applied, since, for example, for integrating a $d = 1024$-variate integrand using only $\tilde{M} = 2$ integration nodes per coordinate requires a point set consisting of $M = 2^{1024}$ integration nodes. To overcome the latter problem, it might be useful to apply the theory of weighted integration as introduced in Sloan & Woźniakowski (1998), possibly combined with truncation (see, e.g. Kritzer et al. (2016)) or multivariate decomposition methods (see, e.g. Kuo (2010)). A detailed analysis of these approaches applied to the present problem is left open for future research.

5. Smoothing of the integrand

The integrand in (9) is not necessarily a C^κ-function. Therefore, in this section we provide a technique for smoothing the integrand in order to apply convergence results for integration rules that are described in Section 4.

The piecewise construction of the process described in Definition 2.4 leads to the situation that $X_t = \phi(X_{T_{j-1}}, t - T_{j-1})$ for $t \in [T_{j-1}, T_j)$ is a function of $X_{T_{j-1}}$ and T_{j-1}. In particular, all subsequent pre-jump locations depend on all previous post-jump locations and jump times, via ϕ and λ. Consequently, regularity of the integrand depends on regularity of the flow ϕ and the intensity function λ. The analysis in this section is restricted to the case where the flow originates from autonomous ODEs, i.e. for all $k \in K$ there exist Lipschitz continuous functions $g_k : \mathbb{R}^{d(k)} \to \mathbb{R}^{d(k)}$

such that $(\partial/\partial t)\phi_k(y,t) = g_k(\phi_k(y,t))$. General results from the literature on ODEs, see, e.g. Grigorian (2009), yield that the derivatives $(\partial/\partial y)\phi_k$, $(\partial^2/\partial y^2)\phi_k$, $(\partial/\partial t)\phi_k$ can be described by so-called associated first- and second-order variational equations for which one requires g_k to be a C^2-function.

For the density f_W of the inter-jump times to be C^2 we need that $\lambda \in C^2(E, \mathbb{R})$. Also we need $\ell \in C_b^2(E, \mathbb{R})$, and $\Psi \in C_b^2(E, \mathbb{R})$ since they appear in the integral defining $\mathcal{H}$.

A serious problem with respect to smoothness arises if the PDMP model allows for jumps from the active boundary. Suppose $(k, y) \in E$ and $t^*(k, y) < \infty$. Then, conditional on $X_t = (k, y)$, the time of the next jump is distributed as $\min(T, t^*(k, y))$, where T has distribution function $F_T(t) = 1 - \exp(-\int_0^t \lambda_k(\phi_k(y, s))\, ds)$. But in general neither $t^*(k, y)$ nor $\min(T, t^*(k, y))$ will depend smoothly on y, even if λ_k has arbitrarily high regularity. We are not aware of a general remedy for this problem. However, for all PDMP models put forward in Section 2.1, the jumps from the active boundary do not constitute jumps of the original problem. In the following subsection we describe by example how PDMPs can be approximated by PDMPs that do not allow for jumps from the boundary.

Concerning the jump kernel Q, it is hard to state general sufficient regularity conditions. An exemplary favourable situation arises if the jump kernel Q admits a C^2-density f_Y in the sense that $Q(A, x) = \int_A f_Y(x_1, x)\, dx_1$ for all $A \in \mathcal{E}$ and all $x \in E$. In the one-dimensional examples from risk theory in Sections 2.1.1–2.1.4, this condition is satisfied and for the two-dimensional example in Section 2.1.5 we present a smoothing technique in Section 5.2.

5.1. Smoothing of the flow

Consider the example from Section 2.1.4 without dividend barrier. We can describe the problem alternatively with a state space consisting of three components:

- $K = \{1, 2, 3\}$,
- $E_1 = (-(c/\rho), \infty)$, $E_2 = (-\infty, -(c/\rho))$, $E_3 = \{-(c/\rho)\}$,
- ϕ_1 is determined by an autonomous ODE of the form $g_1 : \mathbb{R} \to \mathbb{R}$,

$$
g_1(y) = \begin{cases} c, & \text{if } y \in (0, \infty), \\ c + \rho y, & \text{if } y \in \left(-\dfrac{c}{\rho}, 0\right], \\ 0, & \text{if } y \in \left(-\infty, -\dfrac{c}{\rho}\right], \end{cases} \tag{12}
$$

for some $c > 0$, $\rho > 0$. The function ϕ_2 is given by $\phi_2(y, t) = y \; \forall y \in E_2$ and $\forall t \in \mathbb{R}$, and ϕ_3 by $\phi_3(y, t) = y \; \forall y \in E_3$ and $\forall t \in \mathbb{R}$,
- $\lambda_1(y) = \lambda_N \forall y \in E_1, \lambda_2(y) = 0 \; \forall y \in E_2, \lambda_3(y) = 0 \; \forall y \in E_3$.
- For $B = (\{1\} \times B_1) \cup (\{2\} \times B_2) \cup (\{3\} \times B_3) \in \mathcal{E}$,

$$
\begin{aligned}
Q(B, (1, y)) &= \mathbb{P}(Y \in y - B_1) + \mathbb{P}(Y \in y - B_2) + \mathbb{P}(Y \in y - B_3) && (\text{for } y \in E_1), \\
Q(B, (2, y)) &= \mathbb{P}(Y \in y - B_2) && (\text{for } y \in E_2), \\
Q(B, (3, y)) &= \mathbb{P}(Y \in y - B_2) + \mathbb{P}(Y \in y - B_3) && (\text{for } y \in E_3).
\end{aligned}
$$

Here, g_1 is not differentiable in 0. However, we may smoothen this discontinuity using a 'smoothened Heaviside function'. Note that $\Gamma^* = \emptyset$.

Definition 5.1: Let $\kappa \in \mathbb{N} \cup \{0\}$. We call a function $h \colon \mathbb{R} \to [0, 1]$ a C^κ-Heaviside function, if
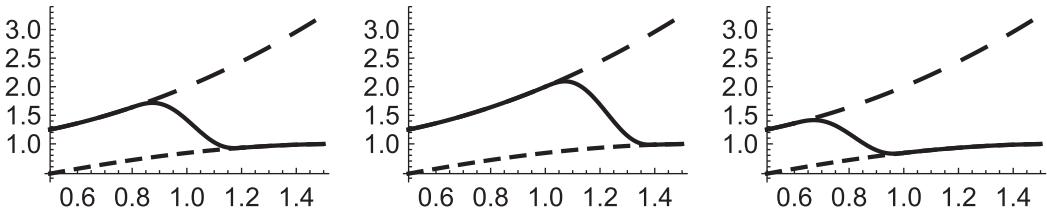
- $h(y) = 0$ for $y < -1$,
- $h(y) = 1$ for $y > 1$,

Figure 2. Illustration of smoothing a piecewise C^2-function with a discontinuity in $\xi = 1$.

- h is non-decreasing,
- $h(y) + h(-y) = 1$,
- h is κ-times continuously differentiable.

Lemma 5.2: *Let $\kappa \in \mathbb{N} \cup \{0\}$, and let $f: \mathbb{R} \to \mathbb{R}$ be a piecewise C^κ-function with discontinuity in $\xi \in \mathbb{R}$, i.e. let there exist C^κ-functions $f_1, f_2: \mathbb{R} \to \mathbb{R}$ with $f = f_1$ on $(-\infty, \xi)$ and $f = f_2$ on (ξ, ∞). Let h be a C^κ-Heaviside function. For every $\varepsilon > 0$ define $f^\varepsilon: \mathbb{R} \to \mathbb{R}$ by $f^\varepsilon(y) = f_1(y)h(y - \xi/\varepsilon) + f_2(y)h(-y - \xi/\varepsilon)$. Then,*

(i) $f^\varepsilon \in C^\kappa$ *for every $\varepsilon > 0$,*
(ii) $f^\varepsilon|_{\mathbb{R} \setminus (-\varepsilon, \varepsilon)} = f|_{\mathbb{R} \setminus (-\varepsilon, \varepsilon)}$ *for every $\varepsilon > 0$,*
(iii) *for all $y \in \mathbb{R} \setminus \{\xi\}$ it holds that $\lim_{\varepsilon \to 0+} f^\varepsilon(y) = f(y)$,*
(iV) *for all $\delta > 0$ it holds that $\lim_{\varepsilon \to 0+} \sup_{y \in \mathbb{R} \setminus (\xi - \delta, \xi + \delta)} |f^\varepsilon(y) - f(y)| = 0$.*

Proof: The elementary proof is left to the reader. ∎

There are various possible choices for the smoothing: from the left $f^{\varepsilon-}(y) = f_1(y)h(y - \xi + \varepsilon/\varepsilon) + f_2(y)h(-y - \xi + \varepsilon/\varepsilon)$ and from the right $f^{\varepsilon+}(y) = f_1(y)h(y - \xi - \varepsilon/\varepsilon) + f_2(y)h(-y - \xi - \varepsilon/\varepsilon)$. Figure 2 depicts these three possible smoothings for a function with a discontinuity in $\xi = 1$. A concrete example for a function h that satisfies the above requirements is given by

$$h(y) = \begin{cases} 0 & \text{if } y < -1, \\ \dfrac{1}{2} + \dfrac{15y}{16} - \dfrac{5y^3}{8} + \dfrac{3y^5}{16} & \text{if } y \in [-1, 1], \\ 1 & \text{if } y > 1. \end{cases} \tag{13}$$

For every $\varepsilon > 0$, a smoothed version of the function g_1 defined in (12) is given by

$$g_1^\varepsilon(y) = (c + \rho y) h\left(-\frac{y}{\varepsilon}\right) + c h\left(\frac{y}{\varepsilon}\right).$$

We can finally formulate a PDMP corresponding to the new model, where the flow has been smoothened,

- $K = \{1, 2, 3\}$,
- $E_1 = (-(c/\rho), \infty)$, $E_2 = (-\infty, -(c/\rho))$, $E_3 = \{-(c/\rho)\}$,
- $(\partial/\partial t)\phi_1^\varepsilon(y, t) = g_1^\varepsilon(\phi_1^\varepsilon(y, t)) \forall y \in E_1$ and $\forall t \in \mathbb{R}$, $\phi_k(y, t) = y \, \forall y \in E_k$ and $\forall t \in \mathbb{R}$, $k \in \{2, 3\}$;
- $\lambda_1(y) = \lambda_N \forall y \in E_1$, $\lambda_k(y) = 0 \, \forall y \in E_k$, $k \in \{2, 3\}$;
- for $B = (\{1\} \times B_1) \cup (\{2\} \times B_2) \cup (\{3\} \times B_3) \in \mathcal{E}$,

$$Q(B, (1, y)) = \mathbb{P}(Y \in y - B_1) + \mathbb{P}(Y \in y - B_2) + \mathbb{P}(Y \in y - B_3) \qquad \text{(for } y \in E_1\text{)},$$

$$Q(B, (2, y)) = \mathbb{P}(Y \in y - B_2) \qquad\qquad\qquad \text{(for } y \in E_2 \text{)} ,$$
$$Q(B, (3, y)) = \mathbb{P}(Y \in y - B_2) + \mathbb{P}(Y \in y - B_3) \qquad \text{(for } y \in E_3 \text{)} .$$

Note that $\Gamma^* = \emptyset$. Since the dividend barrier b is never reached, we also have to smoothen the reward function in a way that the region where dividends are paid can be reached, i.e. $\ell^\varepsilon(y) = c\, h(y - b + \varepsilon/\varepsilon)$. We will show convergence of the corresponding value functions in Section 6.

5.2. Smoothing of jump measures

We give an example for a class of jump measures that can be approximated by measures leading to C^2-integrands in (9).

Let $(E, \mathcal{E})$ be the state space of a PDMP and let (ϕ, λ, Q) be its local characteristics. Let the probability kernel Q satisfy the following assumption.

Assumption 5.3: We assume that

(i) there exists a positive integer n such that for every $k \in K$, and every $y \in E_k$, there exist sets $B_1(k, y), \ldots, B_n(k, y)$ such that
  (i) for every $j \in \{1, \ldots, n\}$ there exists $k_1 \in K$ such that $B_j(k, y) \subseteq E_{k_1}$,
  (ii) for every $j \in \{1, \ldots, n\}$ it holds that $\{(\bar{y}, z) \colon \bar{y} \in E_k, z \in B_j((k, \bar{y}))\}$ is a connected C^2-manifold,
(2) for every $k \in K$ and every $j \in \{1, \ldots, n\}$ the mapping from E_k to $\mathbb{R}$, $\bar{y} \mapsto Q(B_j((k, \bar{y}), x)$ is C^2,
(3) for all $x \in E$ it holds that $\sum_{j=1}^n Q(B_j(x), x) = 1$,
(4) for every $x \in E$ and every $j \in \{1, \ldots, n\}$ there exists a C^2-mapping $G_{j,x} \colon [0, 1]^{\dim(B_j)} \to B_j$ such that for all $A \in \mathcal{E}$ it holds that

$$Q(A \cap B_j, x) = \mu_{\dim(B_j)}(G_{j,x}^{-1}(A \cap B_j))Q(B_j, x),$$

  where μ_m denotes the m-dimensional Lebesgue measure,
(5) for every $k \in K$ and every $j \in \{1, \ldots, n\}$ the mapping from $E_k \times [0, 1]^{\dim(B_j)}$ to $\bigcup_{l \in K} E_l$, $(y, u) \mapsto G_{j,(k,y)}(u)$ is C^2.

Note that Assumption 5.3(1) implies that, for every $x \in E$, $B_j(x)$ is a C^2-manifold, and that for all $x_1 = (k_1, y_1), x_2 = (k_2, y_2) \in E$ with $k_1 = k_2$ we have $\dim B_j(x_1) = \dim B_j(x_2)$.

Under Assumption 5.3 we have for $x \in E$ and for $f \in C_b^2(E, \mathbb{R})$ that

$$\int_E f(y)Q(\mathrm{d}y, x) = \sum_{j=1}^n p_j(x) \int_{[0,1]^{\dim(B_j(x))}} f(G_{j,x}(u))\, \mathrm{d}u,$$

where $p_j(x) = Q(B_{k,j}, x)$ for all $x \in E$. For the integral in (9) this implies that we have iterated sums for each jump, which increases the complexity for large numbers of jumps. Instead, we may write the sum as an integral over $[0, 1]$,

$$\int_E f(y)Q(\mathrm{d}y, x) = \int_0^1 \sum_{j=1}^n \mathbf{1}_{[q_{k,j-1}(x), q_{k,j}(x))}(u_0) \int_{[0,1]^{\dim(B_j(x))}} f(G_{j,x}(u))\, \mathrm{d}u\, \mathrm{d}u_0,$$

where $q_0(x) = 0$ and $q_j(x) = p_1(x) + \cdots + p_j(x)$. However, with this 'trick' we have lost the property of the integrand being C^2. So, using again our smoothened Heaviside function $h \colon \mathbb{R} \to [0, 1]$, we can

smoothen the indicator functions,

$$
\int_E f(y) Q^\varepsilon (\mathrm{d}y, x)
$$

$$
= \int_0^1 \sum_{j=1}^n \left(h\left(\frac{u_0 - q_{j-1}(x)}{\varepsilon}\right) + h\left(\frac{q_j(x) - u_0}{\varepsilon}\right) \right) \int_{[0,1]^{\dim(B_j(x))}} f(G_{j,x}(u)) \, \mathrm{d}u \, \mathrm{d}u_0
$$

$$
= \int_0^1 \int_{[0,1]^{\dim(B_j(x))}} \sum_{j=1}^n \left(h\left(\frac{u_0 - q_{j-1}(x)}{\varepsilon}\right) + h\left(\frac{q_j(x) - u_0}{\varepsilon}\right) \right) f(G_{j,x}(u_1, \ldots, u_{\dim(B_j(x))})) \, \mathrm{d}u \, \mathrm{d}u_0 .
$$

This expression, considered as a function of x, is C^2 as it is a composition of C^2-functions.

Theorem 5.4: *In the setup of this section we have for all $f \in C_b^0(E, \mathbb{R})$ that*

$$
\lim_{\varepsilon \to 0} \int_E f(y) Q^\varepsilon (\mathrm{d}y, x) = \int_E f(y) Q(\mathrm{d}y, x).
$$

Proof: It holds that

$$
\left| \int_E f(y)(Q^\varepsilon (\mathrm{d}y, x) - Q(\mathrm{d}y, x)) \right|
$$

$$
= \left| \sum_{j=1}^n \int_0^1 \left(h\left(\frac{u_0 - q_{j-1}(x)}{\varepsilon}\right) + h\left(\frac{q_j(x) - u_0}{\varepsilon}\right) - 1_{[q_{j-1}(x), q_j(x))}(u_0) \right) \mathrm{d}u_0 \right.
$$

$$
\times \left. \int_{[0,1]^{\dim(B_j(x))}} f(G_{j,x}(u)) \, \mathrm{d}u \right|
$$

$$
\leq \sum_{j=1}^n \int_0^1 \left| h\left(\frac{u_0 - q_{j-1}(x)}{\varepsilon}\right) + h\left(\frac{q_j(x) - u_0}{\varepsilon}\right) - 1_{[q_{j-1}(x), q_j(x))}(u_0) \right| \mathrm{d}u_0
$$

$$
\times \int_{[0,1]^{\dim(B_j(x))}} \left| f(G_{j,x}(u)) \right| \mathrm{d}u .
$$

For our concrete example of h the first integral can be estimated by $\frac{5}{8}\varepsilon$. Thus

$$
\left| \int_E f(y)(Q^\varepsilon (\mathrm{d}y, x) - Q(\mathrm{d}y, x)) \right| \leq \frac{5\varepsilon n}{8} \|f\|_\infty ,
$$

yielding the statement of the theorem. ∎

Now, consider the example from Section 2.1.5. Here, a jump can be either a jump in x_1-direction or a jump in x_2-direction, i.e.

$$
X_{T_j} = \begin{cases} X_{T_j-} + (Y_1, 0) & \text{with probability } \dfrac{\lambda_1}{\lambda_1 + \lambda_2}, \\ X_{T_j-} + (0, Y_2) & \text{with probability } \dfrac{\lambda_2}{\lambda_1 + \lambda_2}. \end{cases}
$$

In this case we can find functions $G_1, G_2 : [0, 1] \to [0, \infty)$ such that $Y_1 \overset{d}{\sim} G_1(\Theta_1)$ and $Y_2 \overset{d}{\sim} G_2(\Theta_2)$ for uniform random variables Θ_1, Θ_2. Hence,

$$
\int_E f(y) Q(dy, (x_1, x_2)) \approx \int_0^1 \int_{[0,1]^2} h\Big(\varepsilon^{-1}\Big(\frac{\lambda_1}{\lambda_1 + \lambda_2} - u\Big)\Big) f\big(x_1 + G_1(\vartheta_1), x_2\big)
$$
$$
+ h\Big(\varepsilon^{-1}\Big(u - \frac{\lambda_1}{\lambda_1 + \lambda_2}\Big)\Big) f\big(x_1, x_2 + G_2(\vartheta_2)\big)\, d\vartheta_1\, d\vartheta_2\, du.
$$

Remark 5.5: If we consider, say, $i = 100$ in (9), then we get a very high number of terms to be summed in the integral. However, we always assume ε to be very small, in particular, we may assume that per jump at most two, and in most situations only one, of the terms $h(\varepsilon^{-1}(u - q_{j-1}(x))) + h(\varepsilon^{-1}(q_j(x) - u))$ are nonzero.

5.3. Convergence

In this section we prove a general convergence result for approximated versions of PDMPs with smoothing as above. We will exploit results on Feller processes presented in Kallenberg (2002, Chapter 19) and Ethier & Kurtz (1986, Chapters 4.2 and 4.8). For the remainder of this section we make the following assumptions:

  (i) $t^*(x) = \infty$ for all $x \in E$,
  (ii) $\lambda \in C_b(E, \mathbb{R})$,
  (iii) for all $f \in C_b(E)$ the mapping $x \mapsto \int_E f(\bar{x}) Q(d\bar{x}, x)$ is continuous.

With this, we can utilise the following theorem.

Theorem 5.6 (Davis 1993, Theorem 27.6): *If $t^*(x) = \infty$ for all $x \in E$ and for all $\lambda \in C_b(E, \mathbb{R})$, and if the mapping $x \mapsto \int_E f(y) Q(dy, x)$ is continuous for all $f \in C_b(E, \mathbb{R})$, then the PDMP is a Feller process.*

We give an example for a class of jump kernel which comprises the jump kernels of the one-dimensional examples in Section 2.1 and which satisfies (iii).

Example 5.7: Let $E_k \subseteq \mathbb{R}$ be an interval for every $k \in K$ and let f_Y be a bounded density function on $\mathbb{R}$. Furthermore, let, for every $x = (k, y) \in E$ and every $A \in \mathcal{E}$, $Q(A, (k, y)) = \sum_{j \in K} \int_{(y-A) \cap E_j} f_Y(\bar{y})\, d\bar{y}$. Then for every $f \in C_b(E, \mathbb{R})$ it holds that

$$
\left| \int_E f(x) Q(dx, (k, y_1)) - \int_E f(x) Q(dx, (k, y_2)) \right|
$$
$$
= \left| \sum_{j \in K} \int_{\mathbb{R}} \mathbf{1}_{E_j}(y_1 - \bar{y}) f_j(y_1 - \bar{y}) f_Y(\bar{y})\, d\bar{y} - \sum_{j \in K} \int_{\mathbb{R}} \mathbf{1}_{E_j}(y_2 - \bar{y}) f_j(y_2 - \bar{y}) f_Y(\bar{y})\, d\bar{y} \right|
$$
$$
\leq \sum_{j \in K} \left| \int_{\mathbb{R}} \mathbf{1}_{E_j}(y_1 - \bar{y}) f_j(y_1 - \bar{y}) f_Y(\bar{y})\, d\bar{y} - \int_{\mathbb{R}} \mathbf{1}_{E_j}(y_2 - \bar{y}) f_j(y_2 - \bar{y}) f_Y(\bar{y})\, d\bar{y} \right|
$$
$$
\leq \sum_{j \in K} \int_{\mathbb{R}} |\mathbf{1}_{E_j}(y_1 - \bar{y}) f_j(y_1 - \bar{y}) - \mathbf{1}_{E_j}(y_2 - \bar{y}) f_j(y_2 - \bar{y})| f_Y(\bar{y})\, d\bar{y}.
$$

Since, by assumption, all f_j are continuous and all E_j are intervals, it holds that $|\mathbf{1}_{E_j}(y_1 - \bar{y}) f_j(y_1 - \bar{y}) - \mathbf{1}_{E_j}(y_2 - \bar{y}) f_j(y_2 - \bar{y})|$ is bounded by $2\|f_j\|_\infty$ and goes to zero as $y_1 \to y_2$ for almost all $\bar{y}$.

Therefore, bounded convergence implies that the above sum converges to 0. From this the desired continuity follows.

The generator of X in the setup of the current section is given by

$$\mathcal{A}f(x) = \mathcal{X}f(x) + \lambda(x) \int_E (f(\bar{x}) - f(x))Q(\mathrm{d}\bar{x}, x), \quad x \in E, \tag{14}$$

where for $x = (k, y) \in E$ we define $\mathcal{X}f(x)$ by $(\mathcal{X}f)_k(y) = (\partial/\partial t)f_k(\phi_k(y, t))|_{t=0}$. Note that for $f \in C_b^1(E, \mathbb{R})$ this means $(\mathcal{X}f)(y) = g(y) \cdot \nabla f(y)$. So the domain $\mathcal{D}(\mathcal{A})$ of the generator consists of all functions in $C_b(E, \mathbb{R})$ which are continuously differentiable along the trajectories of the flow on all components, cf. Ethier & Kurtz (1986, p. 8), and $C_b^1(E, \mathbb{R}) \subseteq \mathcal{D}(\mathcal{A})$.

Definition 5.8 (Kallenberg 2002, Chapter 19): Let A be a closed linear operator with domain of definition $\mathcal{D}(A)$. A *core* for A is a linear subspace $D \subseteq \mathcal{D}(A)$ such that the restriction $A|D$ has closure A.

Proposition 5.9 (Kallenberg 2002, Proposition 19.9): *If $\mathcal{A}$ is the generator of a Feller semigroup $(P_t)_{t \geq 0}$, then any dense, $(P_t)_{t \geq 0}$-invariant subspace $D \subseteq \mathcal{D}(\mathcal{A})$ is a core for $\mathcal{A}$.*

Proposition 5.10: *Under the assumptions made in this section, and for $\mathcal{A}$ being defined as in (14), it is true that $C_b^\infty(E, \mathbb{R})$ is a core for $\mathcal{A}$.*

Proof: We certainly have that $C_b^\infty(E, \mathbb{R})$ is a dense subspace of $C_b(E, \mathbb{R})$. Furthermore, the transition semigroup satisfies $P_t : C_b(E, \mathbb{R}) \to C_b(E, \mathbb{R})$ for all $t \in [0, \infty)$, see (Davis 1993, p.76), since the PDMP is Feller by Theorem 5.6.

We have to prove that $C_b^\infty(E, \mathbb{R})$ is invariant under $(P_t)_{t \in [0,\infty)}$. We show this by proving that, for all $k \in \mathbb{N}$, $P_t C_b^k(E, \mathbb{R}) \subseteq C_b^k(E, \mathbb{R})$. For $k = 0$ this is just the Feller property. Since all derivatives are bounded in the sup-norm, differentiation and application of P_t commute, i.e. $(\partial^k/\partial x^k)P_t f = P_t(\partial^k/\partial x^k)f \in C_b(E, \mathbb{R})$ for all $k \in \mathbb{N}$. Consequently, $C_b^\infty(E, \mathbb{R})$ is a core for $\mathcal{A}$. ∎

Theorem 5.11 (Kallenberg 2002, Theorem 19.25): *Let X be a Feller process in E with semigroup $(P_t)_{t \geq 0}$ and generator $\mathcal{A}$ with domain $\mathcal{D}(\mathcal{A})$, and for all $n \in \mathbb{N}$ let X^n be Feller processes in E with semigroups $(P_t^n)_{t \geq 0}$ and generators $\mathcal{A}^n$ with domains $\mathcal{D}(\mathcal{A}^n)$. Let D be a core for $\mathcal{A}$. Then the following statements are equivalent:*

(i) *for every $f \in D$ there exists a sequence $(f^n)_{n \in \mathbb{N}}$ with $f^n \in \mathcal{D}(\mathcal{A}^n)$ for all $n \in \mathbb{N}$ and such that $f^n \to f$ and $\mathcal{A}^n f^n \to \mathcal{A}f$ uniformly as $n \to \infty$,*

(ii) *for all $t > 0$ we have $P_t^n \to P_t$ as $n \to \infty$ in the strong operator topology,*

(iii) *for every $f \in C_0(E, \mathbb{R})$ and every $t_0 \in (0, \infty)$ it holds that $P_t^n f \to P_t f$ as $n \to \infty$ uniformly for $t \in [0, t_0]$,*

(iv) *if $X_0^n \overset{d}{\to} X_0$ in E, then $X^n \overset{d}{\to} X$ in $D([0, \infty), E)$.*

Remark 5.12: The notion of weak convergence of processes in Item (iv) needs an explanation. Here, $D([0, \infty), E)$ is the space of càdlàg functions, equipped with the Skorokhod topology, see Ethier & Kurtz (1986, p. 118). With this topology, $D([0, \infty), E)$ is a Borel subset of a Polish space and for a sequence $(X^n)_{n \in \mathbb{N}}$ of $D([0, \infty), E)$-valued random variables (i.e. processes in E with càdlàg paths), and a $D([0, \infty), E)$-valued random variable X we have $X^n \overset{d}{\to} X$ if and only if $\lim_{n \to \infty} \mathbb{E}(F(X^n)) = \mathbb{E}(F(X))$ for all bounded Skorokhod continuous functions $D([0, \infty), E) \to \mathbb{R}$, see Kurtz & Protter (1996, Section 6) or Ethier & Kurtz (1986, Chapter 3). We do not wish to go into the details of the notion of Skorokhod continuity. It suffices to mention that from Kurtz & Protter (1996,

Section 8, Example 8.1) we know that for given continuous functions $f_1 : E \times [0, \infty) \to \mathbb{R}^d$ and $f_2 : [0, \infty) \to [0, \infty)$, and fixed $t \in [0, \infty)$, the following functionals exhibit this property:

$$F_1(\omega) = f_1(\omega(t), t) \quad (\text{for } \omega \in D([0, \infty), E)),$$

$$F_2(\omega) = \int_0^t f_2(t - s) f_1(\omega(s), s) \, ds \quad (\text{for } \omega \in D([0, \infty), E)).$$

Lemma 5.13: *Let $f : E \to \mathbb{R}$ be continuous and bounded, then the functional*

$$F_3(\omega) = \int_0^\infty e^{-\delta s} f(\omega(s)) \, ds \quad (\text{for } \omega \in D([0, \infty), E))$$

is Skorokhod continuous.

Proof: Let σ denote the Skorokhod metric on $D([0, \infty), E)$. Let $\varepsilon > 0$. There exists $t > 0$ such that $\int_t^\infty e^{-\delta s} \|f\|_\infty \, ds < \varepsilon/4$. By Skorokhod continuity of F_2 there exists an $\eta > 0$ such that for all $\omega_1, \omega_2 \in D([0, \infty), E)$ it holds that, if $\sigma(\omega_1, \omega_2) < \eta$ then $|\int_0^t e^{-\delta s} f(\omega_1(s)) \, ds - \int_0^t e^{-\delta s} f(\omega_2(s)) \, ds| < \varepsilon/2$. Therefore,

$$|F_3(\omega_1) - F_3(\omega_2)| = \left| \int_0^\infty e^{-\delta s} f(\omega_1(s)) \, ds - \int_0^\infty e^{-\delta s} f(\omega_2(s)) \, ds \right|$$

$$\leq \left| \int_0^t e^{-\delta s} f(\omega_1(s)) \, ds - \int_0^t e^{-\delta s} f(\omega_2(s)) \, ds \right| + 2 \int_t^\infty e^{-\delta s} \|f\|_\infty \, ds < \varepsilon.$$

■

We remark that a function $f : E \to \mathbb{R}$ is continuous if and only if $f_k : E_k \to \mathbb{R}$ is continuous for all k. In particular, every indicator function of a component $\{k\} \times E_k$ is continuous.

We are in the position to show that cost functionals indeed commute with weak limits of PDMPs.

Lemma 5.14: *Let X be a PDMP and $(X^n)_{n \in \mathbb{N}}$ be a sequence of PDMPs on the same state space E and with the same cemetery E^c, and let $\ell : E \to \mathbb{R}$ and $\Psi : E \to \mathbb{R}$ be a running reward function and a terminal cost function, respectively. Assume that both ℓ and Ψ are continuous and bounded. Assume further that $X_0^n = x$ for all $n \in \mathbb{N}$ and $X_0 = x$, and $X^n \overset{d}{\to} X$ in $D([0, \infty), E)$.*

Then

$$\mathbb{E}_x \left(\int_0^\tau e^{-\delta t} \ell(X_t^n) \, dt + e^{-\delta \tau} \Psi(X_\tau^n) \right) \to \mathbb{E}_x \left(\int_0^\tau e^{-\delta t} \ell(X_t) \, dt + e^{-\delta \tau} \Psi(X_\tau) \right)$$

as $n \to \infty$.

Proof: Recall that $\ell \equiv 0$ on E^c, and $\Psi \equiv 0$ on $E \backslash E^c$, so that $\int_0^\infty e^{-\delta s} \ell(\omega(s)) \, ds = \int_0^\tau e^{-\delta s} \ell(\omega(s)) \, ds$ and $\int_0^\infty \delta e^{-\delta s} \Psi(\omega(s)) \, ds = \int_\tau^\infty \delta e^{-\delta s} \Psi(\omega(s)) \, ds$. Thus by Lemma 5.13 the mappings $\omega \mapsto \int_0^\tau e^{-\delta s} \ell(\omega(s)) \, ds$ and $\omega \mapsto \int_\tau^\infty \delta e^{-\delta s} \Psi(\omega(s)) \, ds$ are Skorokhod continuous.

Moreover, if ω is a path of the PDMPs, then it holds that $\omega(s) = \omega(\tau)$ for all $s \geq \tau$, such that $\int_\tau^\infty \delta e^{-\delta s} \Psi(\omega(s)) \, ds = e^{-\delta \tau} \Psi(\omega(\tau))$. This completes the proof. ■

Also, finite time ruin probabilities, i.e. the probability of the PDMP reaching the cemetery before a given time horizon t, commute with weak limits, as we show next.

Lemma 5.15: *Let X be a PDMP and $(X^n)_{n \in \mathbb{N}}$ be a sequence of PDMPs on the same state space E and with the same cemetery E^c. Assume further that $X_0^n = x$ for all $n \in \mathbb{N}$ and $X_0 = x$, and $X^n \xrightarrow{d} X$ in $D([0, \infty), E)$.*

Then $\lim_{n \to \infty} \mathbb{P}_x(X_t^n \in E^c) = \mathbb{P}_x(X_t \in E^c)$ for all $t \geq 0$.

Proof: Consider a functional of the same form as F_1 in Remark 5.12, with $f_1 = \mathbf{1}_{E^c}$. Since the cemetery is the union of only entire $(\{k\} \times E_k)$, and is therefore a union of connected components of E, the indicator function of the cemetery is continuous. Therefore if we define $\psi(x, t) = \mathbb{P}_x(\tau \leq t) = \mathbb{P}_x(X_t \in E^c) = \mathbb{E}_x(\mathbf{1}_{E^c}(X_t))$ and $\psi^n(x, t) = \mathbb{P}_x(\tau^n \leq t) = \mathbb{P}_x(X_t^n \in E^c) = \mathbb{E}_x(\mathbf{1}_{E^c}(X_t^n))$, $n \in \mathbb{N}$, we have $\lim_{n \to \infty} \psi^n(x, t) = \psi(x, t)$ for all $x \in E$ and for all $t \geq 0$. ∎

The following theorem specifies conditions under which Theorem 5.11 is applicable in the PDMP setting.

Theorem 5.16: *Let X be a Feller PDMP with local characteristics (ϕ, λ, Q) and let X^n, $n \in \mathbb{N}$, be Feller PDMPs with local characteristics (ϕ^n, λ^n, Q^n). Further, let the following assumptions hold:*

(i) *$g^n \to g$ and $\lambda^n \to \lambda$ as $n \to \infty$, uniformly in $x \in E$,*
(ii) *for all $f \in C_b^\infty(E, \mathbb{R})$,*

$$\lim_{n \to \infty} \sup_{x \in E} \left| \int_E f(y) Q^n(dy, x) - \int_E f(y) Q(dy, x) \right| = 0, \tag{15}$$

(iii) *$X_0^n \xrightarrow{d} X_0$ in E.*

Then $X^n \xrightarrow{d} X$ in $D([0, \infty), E)$.

Proof: Let $\mathcal{D}(\mathcal{A}^n)$, $n \in \mathbb{N}$, and $\mathcal{D}(\mathcal{A})$ be the domains of the generators $\mathcal{A}^n$, $n \in \mathbb{N}$, and $\mathcal{A}$, corresponding to X^n and X, respectively. For $f^n \in \mathcal{D}(\mathcal{A}^n)$ we have

$$\mathcal{A}^n f^n(x) = \mathcal{X}^n f^n(x) + \lambda^n(x) \int_E (f^n(y) - f^n(x)) Q^n(x, dy),$$

$$(\mathcal{X}^n f^n)(x) = (g^n)(x) \cdot \nabla(f^n)(x).$$

By Proposition 5.10, $D = C_b^\infty(E, \mathbb{R})$ is a core for all generators involved. For every $f \in D$ we set $f^n = f$ for all $n \in \mathbb{N}$, such that trivially $f^n \to f$ as $n \to \infty$. Next, observe that we have for all $n \in \mathbb{N}$,

$$|\mathcal{A}^n f(x) - \mathcal{A}f(x)| \leq |g^n(x) \cdot \nabla f(x) - g(x) \cdot \nabla f(x)|$$

$$+ \left| \lambda^n(x) \int_E (f(y) - f(x)) Q^n(dy, x) - \lambda(x) \int_E (f(y) - f(x)) Q(dy, x) \right|$$

$$= |(g^n(x) - g(x)) \cdot \nabla f(x)| + \left| \lambda^n(x) \int_E (f(y) - f(x)) Q^n(dy, x) - \lambda(x) \int_E (f(y) - f(x)) Q(dy, x) \right|$$

$$\leq \|g^n - g\|_\infty \|\nabla f\|_\infty + \|f\|_\infty \left| \lambda^n(x) \int_E Q^n(dy, x) - \lambda(x) \int_E Q(dy, x) \right| \tag{16}$$

$$+ \left| \lambda^n(x) \int_E f(y) Q^n(dy, x) - \lambda(x) \int_E f(y) Q(dy, x) \right|. \tag{17}$$

Since Q^n, $n \in \mathbb{N}$, and Q are probability measures on $(E, \mathcal{B}(E))$, and since, by assumption, $g^n \to g$ and $\lambda^n \to \lambda$ uniformly in $x \in E$, the terms in (16) converge to zero. The term in (17) can be estimated as follows,

$$\left| \lambda^n(x) \int_E f(y) Q^n(dy, x) - \lambda(x) \int_E f(y) Q(dy, x) \right|$$

$$\leq \|\lambda^n\|_\infty \left| \int_E f(y) Q^n(dy, x) - \int_E f(y) Q(dy, x) \right| + \left| \int_E f(y) Q(dy, x) \right| \|\lambda^n - \lambda\|_\infty.$$

The latter expression tends to zero, since for all $x \in E$ it was assumed that (15) holds, and since $\lambda^n \to \lambda$ uniformly in $x \in E$. Thus, Item (i) of Theorem 5.11 holds. This implies that Item (iv) of Theorem 5.11 holds. The latter is equivalent to the assertion of this theorem. ∎

Remark 5.17: Note that in the Feller case we can move to another external state only due to a purely random jump, i.e. a jump determined by Q^n for $n \in \mathbb{N}$ or Q. Therefore, if we assume uniform convergence of the local characteristics across all state components, and in particular also $Q^n \to Q$ in the sense of (15), then the result of Theorem 5.16 still holds.

Since uniform convergence of the local characteristics and the assumption that $t^*(x) = \infty$ are essential in the proof of Theorem 5.16, we need an alternative argument for situations with an active boundary or for situations in which a smooth approximation fails. A prototypical univariate example for both cases is a drift of the form $g(x) = c \, \mathbf{1}_{\{x \leq b\}}$ for some $b \in \mathbb{R}$. Here one faces either a discontinuity or a subdivision of $\mathbb{R}$ into two state components, i.e. $\mathbb{R} = \{x \in \mathbb{R} : x \leq b\} \cup \{x \in \mathbb{R} : x > b\}$, with a continuous drift on each component. For a specific example, we find a method for dealing with this particular situation in the next section.

6. Application to the Cramér–Lundberg model with loan

In this section we apply our smoothing technique to the example presented in Section 2.1.4 and calculate the quantity of interest using different numerical integration methods. In this setup, ϕ_1 solves the ODE $(\partial / \partial t) \phi_1(y, t) = g_1(\phi_1(y, t)) \ \forall y \in E_1$ and $\forall t \in \mathbb{R}$, with

$$g_1(y) = \begin{cases} c & \text{if } y \in (0, \infty), \\ c + \rho y & \text{if } y \in \left(-\dfrac{c}{\rho}, 0 \right], \\ 0 & \text{if } y \in \left(-\infty, -\dfrac{c}{\rho} \right]. \end{cases}$$

In the setup of Section 2.1.4, the quantity of interest is the expected value of discounted future dividend payments until the time of ruin. The cemetery E^c is given by $E^c = (\{2\} \times E_2) \cup (\{3\} \times E_3)$, the running reward ℓ is given by $\ell_1 \equiv 0$, $\ell_4 \equiv c$, and the terminal cost is $\Psi(x) = 0$ for $x \in E^c$. For $x \in E$, $t \geq 0$, let

$$L(t, x) = \int_0^t e^{-\delta s} \ell(\phi(s, x)) \, ds.$$

Since g is not differentiable in 0 and $t^*(x) < \infty$ for all $x = (1, y)$ with $y \in E_1$, we replace g by a smoothed version and we also modify ℓ accordingly. For $\varepsilon > 0$ let

$$
g_1^\varepsilon(y) = \begin{cases}
c & \text{if } y \in (\varepsilon, b - \varepsilon], \\[2mm]
\dfrac{c(b-y)^3 \left(15\varepsilon(y - b) + 6(b - y)^2 + 10\varepsilon^2\right)}{\varepsilon^5} & \text{if } y \in (b - \varepsilon, b), \\[2mm]
c + \rho y & \text{if } y \in \left(-\dfrac{c}{\rho}, -\varepsilon\right), \\[2mm]
c + \dfrac{\rho(y + 3\varepsilon)(y - \varepsilon)^3}{16\varepsilon^3} & \text{if } y \in [-\varepsilon, \varepsilon], \\[2mm]
0 & \text{if } y \in \left(-\infty, -\dfrac{c}{\rho}\right] \cup [b, \infty).
\end{cases}
$$

Observe that $g_1^\varepsilon \in C^2(\mathbb{R})$, that $\lim_{y \nearrow b} g_1^\varepsilon(y) = 0$ and that $g_1^\varepsilon \geq 0$. For $\varepsilon > 0$ define the PDMP X^ε so that for all $y \in \mathbb{R}$ its flow $\phi_1^\varepsilon(\cdot, y)$ is the solution to the ODE $\frac{d}{dt}\phi_1^\varepsilon(t, y) = g^\varepsilon(\phi_1^\varepsilon(t, y))$ with $\phi_1^\varepsilon(0, x) = x$. Apart from that all specifications are the same as for the original PDMP X. In addition, we replace ℓ_1 by

$$
\ell_1^\varepsilon(y) = c\, h\left(\frac{y - b + \varepsilon}{\varepsilon}\right),
$$

where h can be chosen as in (13) and we define

$$
L^\varepsilon(t, x) = \int_0^t e^{-\delta s} \ell^\varepsilon(\phi^\varepsilon(s, x))\, ds.
$$

We aim at computing $\mathcal{G}^{i-1}\mathcal{H}$ for the smoothed process, in order to observe how (9) simplifies in this example. By the definition of the cemetery, $\mathcal{G}^{i-1}\mathcal{H}(x) = 0$ for all $x = (k, z) \in E$ with $k \in \{2, 3\}$. For $x = (1, z)$ with $z \in E_1$, any jumps bigger than $z + c/\rho$ lead to the cemetery, so we only need to integrate over jump sizes up to $z + c/\rho$. Thus, we get that

$$
\begin{aligned}
\mathcal{G}V(x) = \mathcal{G}V((1,z)) &= \int_0^\infty f_W(t, x)\, e^{-\delta t} \int_E V(y) Q(dy, \phi(x, t))\, dt \\
&= \int_0^\infty f_W(t, x)\, e^{-\delta t} \int_0^{z + c/\rho} V((1, z - y))\, dF_Y(y)\, dt.
\end{aligned}
$$

Moreover, since λ is constant on E_1 it holds for all $x = (1, z)$ with $z \in E_1$, $t \geq 0$ that $f_W(t, x) = \lambda_N e^{-\lambda_N t}$, where λ_N is as in Section 2.1.1. For $x = (1, z)$ with $z \in E_1$ we get

$$
\begin{aligned}
\mathcal{G}^{i-1}\mathcal{H}(x) = \int_{t_1 = 0}^\infty \lambda_N\, e^{-(\lambda_N + \delta)t_1} \int_{y_1 = 0}^{\chi_{1^-} + \frac{c}{\rho}} \cdots \int_{t_{i-1} = 0}^\infty \lambda_N\, e^{-(\lambda_N + \delta)t_{i-1}} \int_{y_{i-1} = 0}^{\chi_{(i-1)^-} + \frac{c}{\rho}} \\
\int_{t_i = 0}^\infty \lambda_N\, e^{-\lambda_N t_i} L^\varepsilon(t_i, \chi_{(i-1)})\, dt_i\, dF_Y(y_{i-1})\, dt_{i-1} \cdots dF_Y(y_1)\, dt_1,
\end{aligned} \tag{18}
$$

where the functions $\chi_{j^-}, \chi_j, j = 1, 2, \ldots, i - 1$ solve $\chi_{j^-} = \phi_1^\varepsilon(t_j, \chi_{j-1})$ and $\chi_j = \chi_{j^-} - y_j$.

Thus χ_{j^-} depends on $t_1, \ldots, t_j$ and $y_1, \ldots, y_{j-1}$, whereas χ_j depends on $t_1, \ldots, t_j$ and $y_1, \ldots, y_j$. However, this dependence has been suppressed in (18) for the sake of readability.

Assumption 6.1: The jump distribution admits a density $f_Y = F_Y'$, with $f_Y \in C_0^2$.

In what follows, suppose that Assumption 6.1 holds. A variable substitution $t_j = -\ln(v_j)$ and $y_j = (\chi_{j^-} + c/\rho)z_j$, where $v_j \in [0,1]$, $z_j \in [0,1]$, $\hat{\chi}_j(v_1, \ldots, v_j, z_1, \ldots, z_j) = \chi_{j^-}(t_1, \ldots, t_j, y_1, \ldots, y_j)$. We then put

$$v(v_1, \ldots, v_i, z_1, \ldots, z_{i-1}) = L^\varepsilon(-\ln(v_i), \hat{\chi}_{i-1}(v_1, \ldots, v_{i-1}, z_1, \ldots, z_{i-1})),$$

which leads to

$$\mathcal{G}^{i-1}\mathcal{H}(x) = \int_{v_1=0}^1 \cdots \int_{v_i=0}^1 \int_{z_1=0}^1 \cdots \int_{z_{i-1}=0}^1 \lambda_N^i \left[\prod_{j=1}^{i-1} v_j^{\delta+\lambda_N-1} \right] v_i^{\lambda_N-1} v(v_1, \ldots, v_i, z_1, \ldots, z_{i-1})$$

$$\times \left[\prod_{j=1}^{i-1} f_Y\left(z_j \left(\hat{\chi}_j + \frac{c}{\rho} \right) \right) \left(\hat{\chi}_j + \frac{c}{\rho} \right) dz_j \right] \prod_{j=1}^i dv_j. \tag{19}$$

Due to the recursive structure of the functions $\hat{\chi}_1, \hat{\chi}_2, \ldots, \hat{\chi}_{i-1}$, the Jacobi matrix of the substitution has lower triangular shape, such that its determinant is the product of the diagonal elements. For being able to reasonably apply (10) we need to bound the Hessian of the integrand. If for example the jump size distribution is the Gamma distribution with parameters $\alpha, \beta > 0$, i.e. $dF_Y(y) = (y^{\alpha-1}\beta^\alpha e^{-\beta y})/\Gamma(\alpha) \, dy$, then this boils down to the condition $\beta \geq 3$ and $\delta + \lambda > 3$, which implies that the integrand is bounded in $\mathbf{0}$. In the original problem statement this corresponds to an additional integrability condition on the jump size distribution. Finally, for $x \in E$ of the form $x = (4, b)$ we have

$$\mathcal{G}^{i-1}\mathcal{H}((4,b)) = \int_0^\infty \lambda \, e^{-\lambda t} e^{-\delta t} \int_0^{b+c/\rho} \mathcal{G}^{i-2}\mathcal{H}((1, b-y)) \, dF_Y(y) \, dt.$$

Remark 6.2: In Section 5.3 the stability, with respect to the smoothing parameter ε of the considered functional of the process, is dealt with in a fairly general setting. Unfortunately, because of the discontinuity of the drift g in the present example, we cannot achieve uniform convergence of the smoothed drift around the barrier level b, whereas point-wise convergence is achieved.

Theorem 6.3: *In the setup of this section, the following assertion holds true. There exists $C > 0$ such that $\|V^\varepsilon - V\|_\infty \leq C\varepsilon$.*

Proof: Recall that

$$V(x) = \mathbb{E}_x\left(\int_0^\tau e^{-\delta s} \ell(X_s) \, ds \right) \quad \text{and} \quad V^\varepsilon(x) = \mathbb{E}_x\left(\int_0^{\tau^\varepsilon} e^{-\delta s} \ell^\varepsilon(X_s^\varepsilon) \, ds \right),$$

where $\tau = \inf\{t \geq 0: X_t \in E^c\}$ and $\tau^\varepsilon = \inf\{t \geq 0: X_t^\varepsilon \in E^c\}$.

It is readily checked that $\sup_{y \in (-c/\rho, b-\varepsilon)} |g_1(y) - g_1^\varepsilon(y)| \leq 3\varepsilon\rho/16$ and that $|g_1(y_1) - g_1(y_2)| \leq \rho|y_1 - y_2|$ for all $y_1, y_2 \in (-c/\rho, b-\varepsilon)$. Hence we get from Kamke (1964, Theorem 8, p. 111) that

$$|\phi_1^\varepsilon(t, y) - \phi_1(t, y)| < \frac{3\varepsilon}{16}\left(e^{\rho t} - 1 \right)$$

for all $y \in (-c/\rho, b-\varepsilon)$ and for all $t \in [0, \min\{\theta_{b-\varepsilon}^\varepsilon, \tilde{\theta}_{b-\varepsilon}\}]$, where

$$\theta_{b-\varepsilon}^\varepsilon = \inf\{t \geq 0: \phi_1^\varepsilon(t, y) \geq b - \varepsilon\},$$

$$\tilde{\theta}_{b-\varepsilon} = \inf\{t \geq 0: \phi_1(t, y) \geq b - \varepsilon\},$$

$$\tilde{\theta}_b = \inf\{t \geq 0: \phi_1(t, y) \geq b\}.$$

Since g_1^ε and g_1 coincide on $(-c/\rho, b - \varepsilon) \backslash (-\varepsilon, \varepsilon)$ and $g_1^\varepsilon \geq g_1 \geq 0$, we can refine this estimate to get

$$|\phi_1^\varepsilon(t, y) - \phi_1(t, y)| < \frac{3\varepsilon}{16}\left(e^{\rho C(\varepsilon)} - 1\right),$$

for all $t \in [0, \min\{\theta_{b-\varepsilon}^\varepsilon, \tilde{\theta}_{b-\varepsilon}\}]$, where $C(\varepsilon) \in [0, \infty)$ is the time needed for the trajectory $\phi_1(\cdot, y)$ to cross $(-\varepsilon, \varepsilon)$. Note that $g_1^\varepsilon \geq g_1 \geq 0$ yields $\phi_1^\varepsilon(t, y) \geq \phi_1(t, y)$ for all $t \in [0, \tilde{\theta}_{b-\varepsilon}]$, and hence $\tilde{\theta}_{b-\varepsilon} \geq \theta_{b-\varepsilon}^\varepsilon$. For $t \geq \min\{\theta_{b-\varepsilon}^\varepsilon, \tilde{\theta}_{b-\varepsilon}\} = \theta_{b-\varepsilon}^\varepsilon$ we have by construction that $|\phi_1^\varepsilon(t, y) - \phi_1(t, y)| \leq \varepsilon$. In total we get

$$|\phi_1^\varepsilon(t, y) - \phi_1(t, y)| \leq \varepsilon \max\left(1, \frac{3}{16}\left(e^{\rho C(\varepsilon)} - 1\right)\right). \tag{20}$$

Since $\lim_{\varepsilon \to 0} C(\varepsilon) \to 0$, it holds that $|\phi_1^\varepsilon(t, y) - \phi_1(t, y)| \leq \varepsilon$ for sufficiently small $\varepsilon > 0$.

Recall that T_1 is the time of the first jump of X conditional on $X_0 = (1, y)$. Since the jump intensity is constant on E_1, T_1 is exponentially distributed with intensity λ_N. Hence, we can write

$$V((1, y)) = \mathbb{E}_{(1,y)}\left(L(T_1, (1, y)) + e^{-\delta T_1} V(X_{T_1})\right)$$
$$= \int_0^\infty \lambda_N e^{-\lambda_N s}\left(L(s, (1, y)) + e^{-\delta s}\int_E V(x_1)Q(dx_1, \phi_1(s, y))\right) ds$$
$$= \int_0^\infty \lambda_N e^{-\lambda_N s} L(s, (1, y)) ds + \int_0^\infty \lambda_N e^{-(\lambda_N + \delta)s}\int_E V(x_1)Q(dx_1, \phi_1(s, y)) ds,$$

and analogously for V^ε. We write $V((1, y)) = V_1(y)$ and $V^\varepsilon((1, y)) = V_1^\varepsilon(y)$ for $y \in E_1$. Therefore,

$$|V_1(y) - V_1^\varepsilon(y)| \leq \int_0^\infty \lambda_N e^{-\lambda_N s}|L(s, (1, y)) - L^\varepsilon(s, (1, y))| ds$$
$$+ \int_0^\infty \lambda_N e^{-(\lambda_N + \delta)s}\left|\int_E V(x_1)Q(dx_1, \phi_1(s, y)) - \int_E V^\varepsilon(x_1)Q(dx_1, \phi_1^\varepsilon(s, y))\right| ds$$
$$\leq \int_0^\infty \lambda_N e^{-\lambda_N s}|L(s, (1, y)) - L^\varepsilon(s, (1, y))| ds$$
$$+ \int_0^\infty \lambda_N e^{-(\lambda_N + \delta)s}\left|\int_E V(x_1)Q(dx_1, \phi_1(s, y)) - \int_E V(x_1)Q(dx_1, \phi_1^\varepsilon(s, y))\right| ds$$
$$+ \int_0^\infty \lambda_N e^{-(\lambda_N + \delta)s}\left|\int_E V(x_1)Q(dx_1, \phi_1^\varepsilon(s, y)) - \int_E V^\varepsilon(x_1)Q(dx_1, \phi_1^\varepsilon(s, y))\right| ds. \tag{21}$$

For $x = (1, y)$ and $t \geq 0$ it holds that $L^\varepsilon(s, x) = 0$ for $s \leq \theta_{b-2\varepsilon}^\varepsilon$, and

$$L^\varepsilon(s, x) = \int_0^s e^{-\delta r}\ell^\varepsilon(\phi^\varepsilon(r, x)) dr = c\int_0^s e^{-\delta r}h((\phi^\varepsilon(r, x) - b + \varepsilon)/\varepsilon) dr \leq c\int_{\theta_{b-2\varepsilon}^\varepsilon}^s e^{-\delta r} dr$$

for $s \geq \theta_{b-2\varepsilon}^\varepsilon$. On the other hand, we have that, for $x = (1, y)$ and $s \geq 0$, $L(s, x) = 0$ for $s \leq \tilde{\theta}_b$, and

$$L(t, x) = c\int_{\tilde{\theta}_b}^s e^{-\delta r} dr$$

for $s > \tilde{\theta}_b$. Using $\phi_1^\varepsilon(s, y) \geq \phi_1(s, y)$ for all $s \geq 0$, we get $\tilde{\theta}_b \geq \theta_{b-2\varepsilon}^\varepsilon$, such that

$$|L^\varepsilon(s, (1, y)) - L(s, (1, y))| = L^\varepsilon(s, (1, y)) - L(s, (1, y)) \leq c\int_{\theta_{b-2\varepsilon}^\varepsilon}^{\tilde{\theta}_b} e^{-\delta r} dr$$

for all $t \geq 0$. Hence,

$$\int_0^\infty \lambda_N \, e^{-\lambda_N s} |L(s, (1, y)) - L^\varepsilon(s, (1, y))| \, ds \leq c \int_{\theta_{b-2\varepsilon}^\varepsilon}^{\tilde{\theta}_b} e^{-\delta r} \, dr \leq c(\tilde{\theta}_b - \theta_{b-2\varepsilon}^\varepsilon).$$

Now $\tilde{\theta}_b - \theta_{b-2\varepsilon}^\varepsilon \leq (b - (b - 2\varepsilon - \varepsilon C_1(\varepsilon)))/c = \varepsilon(2 + C_1(\varepsilon))/c$, where $C_1(\varepsilon) = \max(1, \frac{3}{16}(e^{\rho C(\varepsilon)} - 1))$, see (20). With this, the first term in (21) can be estimated by

$$\int_0^\infty \lambda_N \, e^{-\lambda_N s} |L(s, (1, y)) - L^\varepsilon(s, (1, y))| \, ds \leq \varepsilon(2 + C_1(\varepsilon)). \tag{22}$$

Next, observe that (we remind the reader that the states $x \in E$ are denoted by $x = (k, y)$, which is why in the following the terms y_1, y_2 are not to be confused with the integration variables y_j used in and below (18)),

$$\left| \int_E V(x_1) Q(dx_1, (1, y_1)) - \int_E V(x_1) Q(dx_1, (1, y_2)) \right|$$

$$= \left| \int_0^{y_1 + c/\rho} V_1(y_1 - z) f_Y(z) \, dz - \int_0^{y_2 + c/\rho} V_1(y_2 - z) f_Y(z) \, dz \right|$$

$$\leq \left| \int_{\min(y_1, y_2) + c/\rho}^{\max(y_1, y_2) + c/\rho} V_1(z) f_Y(z) \, dz \right| \leq \|V_1\|_\infty \|f_Y\|_\infty |y_1 - y_2|.$$

Combining this with (20), we can estimate the second term in (21) by

$$\int_0^\infty \lambda_N \, e^{-(\lambda_N + \delta)s} \left| \int_E V(x_1) Q(dx_1, \phi_1(s, y)) - \int_E V^\varepsilon(x_1) Q(dx_1, \phi_1^\varepsilon(s, y)) \right| \, ds \tag{23}$$

$$\leq \frac{\lambda_N}{\lambda_N + \delta} \|V_1\|_\infty \|f_Y\|_\infty \sup_{s \geq 0} |\phi_1(s, y) - \phi_1^\varepsilon(s, y)| \leq \frac{\lambda_N}{\lambda_N + \delta} \|V_1\|_\infty \|f_Y\|_\infty \varepsilon \max$$

$$\times \left(1, \frac{3}{16} \left(e^{\rho C(\varepsilon)} - 1 \right) \right). \tag{24}$$

Furthermore, since

$$\left| \int_E V(x_1) Q(dx_1, (1, y_2)) - \int_E V^\varepsilon(x_1) Q(dx_1, (1, y_2)) \right| \leq \left\| V_1 - V_1^\varepsilon \right\|_\infty,$$

the third term in (21) can be estimated as follows,

$$\int_0^\infty \lambda_N \, e^{-(\lambda_N + \delta)s} \left| \int_E V(x_1) Q(dx_1, \phi_1^\varepsilon(s, y)) - \int_E V^\varepsilon(x_1) Q(dx_1, \phi_1^\varepsilon(s, y)) \right| \, ds$$

$$\leq \frac{\lambda_N}{\lambda_N + \delta} \|V_1 - V_1^\varepsilon\|_\infty. \tag{25}$$

Taking the supremum over $y \in E_1$ in (21) and using (22), (23), and (25) we obtain that

$$\|V_1 - V_1^\varepsilon\|_\infty \leq C\varepsilon + \frac{\lambda_N}{\lambda_N + \delta} \|V_1 - V_1^\varepsilon\|_\infty$$

for some constant C and for sufficiently small ε. Thus,

$$\frac{\delta}{\lambda_N + \delta} \|V_1 - V_1^\varepsilon\|_\infty \leq C\varepsilon,$$

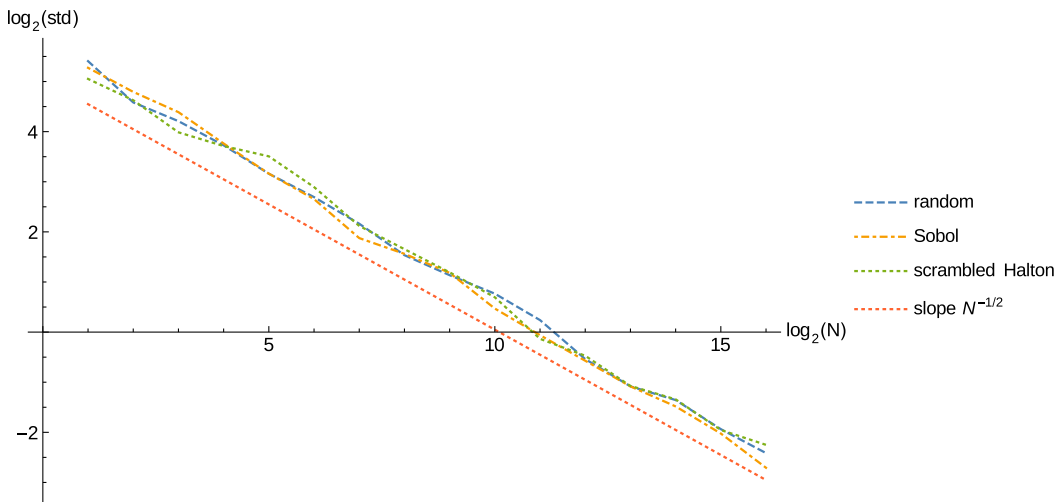which completes the proof.   ∎

Figure 3. The estimated standard deviation of the estimation.

6.1. Numerical experiment

We now solve the example presented above numerically. We set the following parameter values. The initial value of the PDMP $x_0 = 0$, the premium income rate $c = 5$, the credit rate $\rho = 0.05$, the intensity of the Poisson process $\lambda = 4$, the jump size distribution is for all $x \in [0, \infty)$ given by $F_Y(x) = 1 - e^{-\alpha x}$ with $\alpha = 1$, and the discount rate $\delta = 0.02$. With this, the optimal dividend threshold according to Dassios & Embrechts (1989) is $b = 3.24289$. Furthermore, we set the smoothing parameter $\varepsilon = 0.01$. For computing the flow it is enough to solve the corresponding ODE once and to store the solution for repeated use.

We implemented Monte Carlo (random), QMC with the Sobol' sequence (Sobol), and QMC with a scrambled version of the Halton sequence (scrambled Halton), where scrambling refers to a permutation of digits (see, e.g. Owen (2000)). The Sobol' point generator we used was taken from Frances Y. Kuo's homepage Kuo (n.d.) and is based on Joe & Kuo (2008).

The reference solution was calculated using Monte Carlo with $M = 5000 \cdot 2^{10} = 5120000$ sample paths and $d = 1024$, meaning that the maximum number of jumps we allow for is 512. In our plots we show the results plotted over an increasing number of integration nodes $M \in \{50 \cdot 2^j : 1 \leq j \leq 16\}$.

Figure 3 shows the estimated standard deviation (root mean square error) of the estimation, which is calculated by using 50 repetitions with randomly shifted versions of our integration nodes.

References

Albrecher H. & Kainhofer R. (2002). Risk theory with a nonlinear dividend barrier. *Computing* 68(4), 289–311.

Albrecher H. & Lautscham V. (2015). Dividends and the time of ruin under barrier strategies with a capital-exchange agreement. *Anales del Instituo de Actuarios Espanoles* 21(3), 1–30.

Almudevar A. (2001). A dynamic programming algorithm for the optimal control of piecewise deterministic Markov processes. *SIAM Journal on Control and Optimization* 40(2), 525–539.

Asmussen S. & Albrecher H. (2010). *Ruin probabilities*, 2nd ed. Advanced Series on Statistical Science and Applied Probability. Hackensack, NJ: World Scientific.

Bakhvalov N. S. (1959). On the approximate calculation of multiple integrals. *Vestnik MGU, Series Mathematical, Mechanics & Astronomy Physical Chemistry* 4, 3–18. In Russian.

Bäuerle N. & Rieder U. (2010). Optimal control of piecewise deterministic Markov processes with finite time horizon. *Modern Trends in Controlled Stochastic Processes: Theory and Applications* 123, 143.

Bäuerle N. & Rieder U. (2011). *Markov decision processes with applications to finance*. Heidelberg: Universitext, Springer.

Cai J., Feng R. & Willmot G. E. (2009). On the expectation of total discounted operating costs up to default and its applications. *Advances in Applied Probability* 41(2), 495–522.

Colaneri K., Eksi Z., Frey R. & Szölgyenyi M. (2017). Optimal liquidation under partial information with price impact. arXiv:1606.05079

Costa O. L. & Davis M. H. A. (1989). Impulse control of piecewise-deterministic processes. *Mathematics of Control, Signals, and Systems (MCSS)* 2(3), 187–206.

Costa O. L. & Dufour F. (2013). *Continuous average control of piecewise deterministic Markov processes*. New York: Springer.

Coulibaly I. & Lefèvre C. (2008). On a simple quasi-Monte Carlo approach for classical ultimate ruin probabilities. *Insurance: Mathematics and Economics* 42(3), 935–942.

Dassios A. & Embrechts P. (1989). Martingales and insurance risk. *Communications in Statistics. Stochastic Models* 5(2), 181–217.

Davis M. H. A. (1984). Piecewise-deterministic Markov processes: A general class of nondiffusion stochastic models. *Journal of the Royal Statistical Society Series B* 46(3), 353–388. With discussion.

Davis M. H. A. (1993). *Markov models and optimization*. Monographs on Statistics and Applied Probability. London: Chapman & Hall.

Davis M. H. A. & Farid M. (1999). Piecewise-deterministic processes and viscosity solutions. In W. M. McEneaney, G. George Yin, and Q. Zhang, eds., *Stochastic Analysis, Control, Optimization and Applications*. Boston: Springer. P. 249–268.

de Saporta B., Dufour F. & Zhang H. (2016). *Numerical methods for simulation and optimization of piecewise deterministic markov processes*. Mathematics and Statistics Series. London: ISTE; Hoboken, NJ: John Wiley & Sons, Inc.

de Saporta B., Dufour F., Zhang H. & Elegbede C. (2012). Optimal stopping for the predictive maintenance of a structure subject to corrosion. *Journal of Risk and Reliability* 226(2), 169–181.

Dempster M. A. H. & Ye J. J. (1992). Necessary and sufficient optimality conditions for control of piecewise deterministic Markov processes. *Stochastics: An International Journal of Probability and Stochastic Processes* 40(3–4), 125–145.

Eichler A., Leobacher G. & Szölgyenyi M. (2017). Utility indifference pricing of insurance catastrophe derivatives. *European Actuarial Journal* 7(2), 515–534.

Embrechts P. & Schmidli H. (1994). Ruin estimation for a general insurance risk model. *Advances in Applied Probability* 26(2), 404–422.

Ethier S. N. & Kurtz T. G. (1986). *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons, Inc.

Forwick L., Schäl M. & Schmitz M. (2004). Piecewise deterministic Markov control processes with feedback controls and unbounded costs. *Acta Applicandae Mathematica* 82(3), 239–267.

Grigorian A. (2009). Ordinary differential equation. Lecture notes. https://www.math.uni-bielefeld.de/ grigor/ odelec2009.pdf

Hinrichs A., Novak E., Ullrich M. & Woźniakowski H. (2017). Product rules are optimal for numerical integration in classical smoothness spaces. *Journal of Complexity* 38, 39–49.

Jacobsen M. (2006). *Point process theory and applications*. Probability and its Applications. Boston, MA: Birkhäuser Boston, Inc.

Joe S. & Kuo F. Y. (2008). Constructing Sobol' sequences with better two-dimensional projections. *SIAM Journal of Scientific Computation* 30, 2635–2654.

Kallenberg O. (2002). *Foundations of modern probability*, 2nd. Probability and its Applications (New York). New York: Springer-Verlag.

Kamke E. (1964). *Differentialgleichungen. I. Gewöhnliche differentialgleichungen*. 5th ed. Leipzig: Akademische Verlagsgesellschaft.

Kritzer P., Pillichshammer F. & Wasilkowski G. W. (2016). Very low truncation dimension for high dimensional integration under modest error demand. *Journal of Complexity* 35, 63–85.

Kuo F. Y. (n.d.). F. Y. Kuo's homepage. http://web.maths.unsw.edu.au/ ∼ fkuo/sobol/index.html. Last visited 14/12/2017.

Kuo F. Y., Sloan I. H., Wasilkowski G. W. & Woźniakowski H. (2010). Liberating the dimension. *Journal of Complexity* 26, 422–454.

Kurtz T. G. & Protter P. E. (1996). Weak convergence of stochastic integrals and differential equations I. In D. Talay and L. Tubaro, eds., *Probabilistic Models for Nonlinear Partial Differential Equations*. Berlin, Heidelberg: Springer.

Lenhart S. & Liaot Y. (1985). Integro-differential equations associated with optimal stopping time of a piecewise-deterministic process. *Stochastics: An International Journal of Probability and Stochastic Processes* 15(3), 183–207.

Leobacher G. & Ngare P. (2016). Utility indifference pricing of derivatives written on industrial loss indexes. *Journal of Computational and Applied Mathematics* 300, 68–82.

Niederreiter H. (1992). *Random number generation and quasi-Monte Carlo methods*. CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia: SIAM.

Owen A. B. (2000). Monte Carlo, quasi-Monte Carlo, and randomized quasi-Monte Carlo. In H. Niederreiter and J. Spanier, eds., *Monte Carlo and Quasi- Monte Carlo Methods 1998*. Springer. P. 86–97.

Pausinger F. & Svane A. M. (2015). A Koksma–Hlawka inequality for general discrepancy systems. *Journal of Complexity* 31, 773–793.

Preischl M., Thonhauser S. & Tichy R. F. (2018). Integral equations, quasi-monte carlo methods and risk modeling. In J. Dick, F. Y. Kuo and H. Woźniakowski, eds., *Contemporary Computational Mathematics – A Celebration of the 80th Birthday of Ian Sloan*, Vol. 1, 2. Cham: Springer. P. 1051–1074.

Riedler M. G. (2013). Almost sure convergence of numerical approximations for piecewise deterministic Markov processes. *Journal of Computational and Applied Mathematics* 239, 50–71.

Rolski T., Schmidli H., Schmidt V. & Teugels J. (1999). *Stochastic processes for insurance and finance*. Wiley Series in Probability and Statistics. New York: John Wiley & Sons.

Schäl M. (1998). On piecewise deterministic Markov control processes: Control of jumps and of risk processes in insurance. *Insurance: Mathematics and Economics* 22(1), 75–91.

Siegl T. & Tichy R. F. (2000). Ruin theory with risk proportional to the free reserve and securitization. *Insurance: Mathematics and Economics* 26(1), 59–73.

Sloan I. H. & Woźniakowski H. (1998). When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *Journal of Complexity* 14, 1–33.

Tichy R. F. (1984). Über eine zahlentheoretische Methode zur numerischen Integration und zur Behandlung von Integralgleichungen. *Osterreichische Akademie der Wissenschaften Mathematisch-Naturwissenschaftliche Klasse. Sitzungsberichte. Abteilung II* 193(4–7), 329–358.