GRADIENT ESTIMATION FOR SMOOTH STOPPING CRITERIA

BERND HEIDERGOTT 🆔,* *Vrije Universiteit Amsterdam*
YIJIE PENG,** *Peking University*

Abstract

We establish sufficient conditions for differentiability of the expected cost collected over a discrete-time Markov chain until it enters a given set. The parameter with respect to which differentiability is analysed may simultaneously affect the Markov chain and the set defining the stopping criterion. The general statements on differentiability lead to unbiased gradient estimators.

Keywords: Sensitivity analysis; Monte Carlo simulation; gradient estimation

2020 Mathematics Subject Classification: Primary 28A15
Secondary 60J20

1. Introduction

Consider a Markov chain $\{X_n\}$ with homogeneous transition kernel P_θ on a general state-space supported on $\mathbb{R}$. Moreover, we equip $\mathbb{R}$ with the usual topology, and, if not stated otherwise, measurability is understood with respect to the Borel field. Let $A_\theta \subset \mathbb{R}$ denote a measurable set, and let τ denote the first entry time into A_θ when starting the chain in x_0, i.e.,

$$\tau := \tau(x_0) = \inf\{n \geq 1 : X_n \in A_\theta\},$$

with $X_0 = x_0$, where we set $\tau = \infty$ if the set on the right-hand side is empty. For technical convenience we assume throughout the paper that $x_0 \notin A_\theta$. We consider the problem of estimating

$$\frac{d}{d\theta}\mathbb{E}\left[\sum_{i=0}^{\tau-1} h(X_i)\,\middle|\, X_0 = x_0\right],\tag{1}$$

for some measurable cost function $h: \mathbb{R} \mapsto \mathbb{R}$ and $x_0 \notin A_\theta$. Generally speaking, the problem arises naturally in analysing a system, modelled by P_θ, that is operating until a certain event occurs, here modelled by X_n entering A_θ, upon which some action is taken. For example, in inventory problems, orders are placed once the stock drops below a certain value; in maintenance, preventive maintenance is performed once the age of a component exceeds some threshold; and in service control, the speed of the server may be adjusted in case observed waiting times exceed some quality threshold.

In analysing a model we are interested not only in evaluating its performance but also in improving it. Let θ be a control parameter of the model. For example, in an inventory model θ may be the level that triggers the replenishing of stock. Changing the order level not only alters the cost, represented by h in the model, but may also have an impact on the demand, as having a low replenishing level may attract more demand. This illustrates that θ may simultaneously affect the boundary of A_θ, the instantaneous cost h, and the transition probability of the underlying process. In this paper, we will derive derivative expressions for this general case of θ influencing all aspects of the model. As we will show, the overall formulas can be translated into efficient unbiased simulation-based estimators.

With the advent of simulation-driven perturbation analysis techniques, such as perturbation analysis [7, 12, 13, 14, 24], the score function method [37], and weak differentiation [34], efficient unbiased estimators for the expression in (1) have been found for particular instances of the problem. More specifically, for applications of perturbation analysis to (s, S) inventory models with θ being either threshold s, S, or the span $S - s$, we refer to [5, 4]. Alternatively, a variant of the score function method, called the push-out score function, has been developed for this kind of threshold optimization in inventory models; see [35], as well as [36] for an early reference. For a perturbation analysis of a maintenance model where θ triggers (preventive) maintenance of system components whose age exceeds θ, we refer to [19, 16]. Finally, for applications to financial models, where θ is the value of a barrier in some option model, we refer to [15, 31, 17, 21].

General results for the case where A_θ and the instantaneous cost are independent on θ, and only the underlying stochastic process depends on θ, are provided in [22, 23]. Moreover, in the setting of Markov chains with discrete state-space, sensitivity analysis of the expected return time to a specified marked state i^* (also called time to absorption) with respect to the entries of the Markov transition probabilities is well-studied. Expressed in our setting, in these studies, we have $h = 1$, $A = \{i^*\}$ contains the marked state, and the transition probability of the random walk X_n depends on θ. See, for example, [8, 9] for details.

Before laying out the main steps of our approach for providing an unbiased estimator for the derivative in (1), we provide examples illustrating the relevance of the extension to the case A_θ. Without loss of generality, we consider θ to be scalar, as for the multidimensional case $\theta \in \mathbb{R}^k$, $k > 1$, estimation of a gradient can be reduced to that of estimating the partial derivatives separately.

Example 1. (Service quality.) Customers arrive at a service station according to a renewal point process. The interarrival times $\{I_n : n \in \mathbb{N}\}$ are independent and identically distributed (i.i.d.), with density $f_I(x)$ such that $0 < E[I_n] < \infty$ and $\mathbb{P}(I_n = 0) = 0$. Customers are served in order of arrival, and consecutive service times are i.i.d. random variables $\{S_n(\theta) : n \in \mathbb{N}\}$ with density $f_S(x; \theta)$, for $\theta \in \Theta \subset \mathbb{R}$, such that $0 < E[S_n(\theta)] < \infty$ and $\mathbb{P}(S_n(\theta) = 0) = 0$. Interarrival times and service times are assumed to be mutually independent. Consider the process of consecutive waiting times $\{X_n\}$, denoting the time that the corresponding customer has spent at the service station from arrival to beginning of service. The service system starts initially empty. Consecutive waiting times X_n follow the recursive relation

$$X_n = \max\{0, X_{n-1} + S_{n-1}(\theta) - I_n\}, \qquad n \geq 2, \tag{2}$$

and $X_1 = 0$. Letting $A_\theta = [\theta, \infty)$ and $h(X_i) \equiv 1$, with

$$\tau = \inf\{n : X_n \geq \theta\}, \qquad \text{for } \theta \geq 0, \tag{3}$$

we obtain from (1) the sensitivity of the expected index of the first customer who experiences a waiting time exceeding θ.

Example 2. (Investment.) For an investment strategy, a portfolio of m risky assets is considered. The value of asset i after the nth dividend date is denoted by $Z_n(i)$, and we let $X_n = (Z_n(1), \ldots, Z_n(m))$, $n \geq 0$. From a risk management perspective we are interested in the expected time until the value of the basket drops below θ:

$$\tau_\theta = \inf\left\{n : \sum_{i=0}^m Z_n(i) \leq \theta\right\} = \inf\{n : h(X_n) \leq \theta\},$$

for

$$h(X_n) = \sum_{i=0}^m Z_n(i).$$

The threshold θ may be used in an investment strategy where the investor is only inclined to restructure the portfolio once the value of the basket drops below θ.

Example 3. (Insurance.) Consider an insurance company where the nth claim W_n arrives at time t_n. For simplicity, we assume that there is a premium inflow of c per time unit. The capital at time t, denoted by $K(t)$, is then given by

$$K(t) = K_0 + ct - \sum_{i=0}^{N(t)-1} W_i,$$

where $\tau_0 = 0$, $N(t)$ is the number of claims incurred up to time t, and K_0 is the initial capital. The time until ruin is given by

$$\inf\{t : K(t) \leq 0\}.$$

The capital $K(t)$ increases between claims, and the only possible time epochs where $K(t)$ can fall below 0 are the jump epochs of $N(t)$. Hence, letting $X_n = (\eta_n, W_n, \hat{K}_n)$, with η_n the time between the nth claim and the $(n+1)$th, $\hat{K}_n$ the capital right after the nth claim, and $X_0 = (0, 0, K_0)$, ruin occurs for $\hat{K}_n < 0$. However, the insurance company will have to act before the zero-threshold is reached; in practice, the insurance company is interested in the time until the capital drops below a certain threshold θ. Let

$$\tau_\theta = \inf\{n : \hat{K}_n < \theta\}$$

denote the index of the claim causing the capital to drop below θ, and let ρ_n denote the nth jump time of $N(t)$; then ρ_{τ_θ} yields the time at which the capital drops below θ.

For our analysis, we take a Markov chain operator approach. In the following we focus on motivating our approach; formal definitions are postponed to Section 2.1. Specifically, let P_θ denote the Markov kernel of $\{X_n\}$. For analysing (1) we make use of the concept of a taboo kernel. More specifically, the taboo kernel of P_θ is a truncation of P_θ onto the complement of taboo set A_θ, defined by

$$[P_\theta]_{A_\theta}(B; x) = \begin{cases} P_\theta(B \cap A_\theta^c, x), & B \cap A_\theta^c \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

for all measurable sets B. In words, the taboo kernel avoids entering the set A_θ. When it is clear with respect to which set A_θ the taboo kernel is constructed, we simplify the notation by letting

$$\mathbf{T}_\theta(B; x) = [P_\theta]_{A_\theta}(B; x).$$

This implies

$$\int_\mathbb{R} h(y)\mathbf{T}_\theta(dy; x) = \mathbb{E}[h(X_{i+1})\mathbf{1}\{X_{i+1} \notin A_\theta\}|X_i = x],$$

provided the integral exits, where the transition from $x = X_i$ to X_{i+1} is governed by P_θ. Defining powers of transition kernels in the obvious way by $Q^n = Q^{n-1}Q$, with Q^0 denoting the identity, we have

$$\left(\mathbf{T}_\theta^n h\right)(x_0) = \mathbb{E}\left[\prod_{i=0}^n \mathbf{1}\{X_i \notin A_\theta\}h(X_n) \middle| X_0 = x_0\right], \quad n \geq 0, \tag{4}$$

for $x_0 \notin A_\theta$ and $\mathbf{T}_\theta^0$ being the identity operator. For technical convenience we include the indicator $\mathbf{1}\{X_0 \notin A_\theta\}$ in the indicator product; we may do this without loss of generality, thanks to our assumption that the initial state does not belong to the taboo set A_θ.

Since $x_0 \notin A_\theta$, we obtain for the random horizon summation

$$\mathbb{E}\left[\sum_{i=0}^{\tau-1} h(X_i) \middle| X_0 = x_0\right] = \mathbb{E}\left[\sum_{i=0}^\infty h(X_i)\prod_{j=0}^i \mathbf{1}\{X_j \notin A_\theta\} \middle| X_0 = x_0\right]$$

$$= \sum_{i=0}^\infty \int_\mathbb{R} h(x_i)\mathbf{T}_\theta^i(dx_i; x_0)$$

$$= (\mathbf{H}_\theta h)(x_0), \tag{5}$$

where

$$\mathbf{H}_\theta = \sum_{n=0}^\infty \mathbf{T}_\theta^n$$

is called the potential of the taboo kernel $\mathbf{T}_\theta$.

The paper is organized as follows. In Section 2 differentiation of the taboo kernel $\mathbf{T}_\theta$ is studied. The main result of the paper on the differentiability of the potential $\mathbf{H}_\theta$ is provided in Section 3. Section 4 is devoted to translating the overall formulas into gradient estimators. Section 5 addresses applications. We conclude by identifying topics of further research.

2. Differentiation of one-step taboo kernel

2.1. Transition kernels and norms

Let $(S, \mathcal{T})$ be a Polish measurable space, for $S \subset \mathbb{R}$. For ease of presentation we restrict our analysis to the one-dimensional case. Let $\mathcal{M}(S, \mathcal{T})$ denote the set of finite (signed) measures on $(S, \mathcal{T})$ and $\mathcal{M}_1(S, \mathcal{T})$ that of probability measures on $(S, \mathcal{T})$. Let $\eta : \mathbb{R} \to [1, \infty)$ be such that $\inf_x \eta(x) = 1$. The weighted supremum norm with respect to η of a cost function ψ is defined by

$$||\psi||_\eta = \sup_x \frac{|\psi(x)|}{\eta(x)},$$

and that of a (signed) measure $\mu \in \mathcal{M}(S, \mathcal{T})$ by

$$||\mu||_\eta = \int \eta(x)|\mu(dx)| = \int \eta(x)\mu^+(dx) + \int \eta(x)\mu^-(dx),$$

with μ^+, μ^- denoting the Hahn–Jordan decomposition; see [10]. Typically, μ has Lebesgue density f_μ and we have

$$||\mu||_\eta = \int \eta(x)|f_\mu(x)|dx.$$

Note that the property $\eta \geq 1$ implies

$$|\psi(x)| \leq ||\psi||_\eta \eta(x) \quad \text{for all } x;$$

moreover,

$$\left|\int \psi(x)f_\mu(x)dx\right| \leq \int |\psi(x)| |f_\mu(x)|dx \leq \int \left(\sup_x \frac{|\psi(x)|}{\eta(x)}\right)\eta(x)|f_\mu(x)|dx = ||\psi||_\eta||\mu||_\eta.$$

The mapping $P : \mathcal{T} \times S \to [0, 1]$ is called a (homogeneous) *transition kernel* on $(S, \mathcal{T})$ if (i) $P(\,\cdot\,; s) \in \mathcal{M}(S, \mathcal{T})$ for all $s \in S$, and (ii) $P(B;)$ is $\mathcal{T}$ measurable for all $B \in \mathcal{T}$. If, in the condition (i), $\mathcal{M}(S, \mathcal{T})$ can be replaced by $\mathcal{M}_1(S, \mathcal{T})$, then P is called a *Markov kernel* on $(S, \mathcal{T})$. Denote the set of transition kernels on $(S, \mathcal{T})$ by $\mathcal{K}(S, \mathcal{T})$ and the set of Markov kernels on $(S, \mathcal{T})$ by $\mathcal{K}_1(S, \mathcal{T})$. A transition kernel $P \in \mathcal{K}(S, \mathcal{T})$ with $0 < P(S; s) < 1$ for at least one $s \in S$ is called a *defective* Markov kernel.

Consider a family of Markov kernels $(P_\theta : \theta \in \Theta)$ on $(S, \mathcal{T})$, with $\Theta \subset \mathbb{R}$, and let $\mathcal{L}^1(P_\theta; \Theta) \subset \mathbb{R}^S$ denote the set of measurable mappings $g : S \to \mathbb{R}$ such that $\int_S |g(u)|P_\theta(du; s)$ is finite for all $\theta \in \Theta$ and $s \in S$. To simplify the notation, we set

$$(P_\theta g)(s) \triangleq \int_S g(u)P_\theta(du; s)$$

for $g \in \mathcal{L}^1(P_\theta; \Theta)$ and $s \in S$.

In what follows, we let Θ be an open neighbourhood of θ_0 and assume that $\mathcal{D} \subset \mathcal{L}^1(P_\theta; \Theta)$.

We call $P_\theta \in \mathcal{K}(S, \mathcal{T})$ differentiable with respect to θ for $\mathcal{D}$, or Θ-$\mathcal{D}$-differentiable for short, if $\partial_\theta P_\theta \in \mathcal{K}(S, \mathcal{T})$ exists such that for any $g \in \mathcal{D}$ and any $s \in S$,

$$\frac{d}{d\theta} \int_S g(u)P_\theta(du; s) = \int_S g(u)\partial_\theta P_\theta(du; s). \tag{6}$$

If the left-hand side of (6) equals zero for all $g \in \mathcal{D}$, then we say that $\partial_\theta P_\theta$ is not significant.

In the same vein, we call $P_\theta \in \mathcal{K}(S, \mathcal{T})$ differentiable with respect to x for $\mathcal{D}$, or X-$\mathcal{D}$-differentiable for short, if $\partial_x P_\theta \in \mathcal{K}(S, \mathcal{T})$ exists such that for any $g \in \mathcal{D}$ and any $s \in S$,

$$\frac{d}{dx} \int_S g(u)P_\theta(du; x) = \int_S g(u)\partial_x P_\theta(du; x). \tag{7}$$

If the left-hand side of (7) equals zero for all $g \in \mathcal{D}$, then we say that $\partial_x P_\theta$ is not significant.

Assume that P_θ has density $f_\theta(y; x)$, i.e., $P_\theta(B; x) = \int_B f_\theta(y; x)dy$ for all $x \in S$ and $B \in \mathcal{T}$, and suppose that the partial derivatives of $f_\theta(y; x)$ with respect to x and θ, denoted

respectively by $\partial_x f_\theta(y; x)$ and $\partial_\theta f_\theta(y; x)$, both exist. Provided that for all $x, y \in S$ it holds that $\partial_y f_\theta(y; x), \partial_x f_\theta(y; x), \partial_\theta f_\theta(y; x) > 0$ implies $f_\theta(y; x) > 0$, we let

$$\mathrm{SF}_X(y; x) := \mathrm{SF}_X^\theta(y; x) := \partial_x \log f_\theta(y; x) = \frac{\partial_x f_\theta(y; x)}{f_\theta(y; x)},$$

$$\mathrm{SF}_Y(y; x) := \mathrm{SF}_Y^\theta(y; x) := \partial_y \log f_\theta(y; x) = \frac{\partial_y f_\theta(y; x)}{f_\theta(y; x)},$$

and

$$\mathrm{SF}_\Theta(y; x) := \mathrm{SF}_\Theta^\theta(y; x) := \partial_\theta \log f_\theta(y; x) = \frac{\partial_\theta f_\theta(y; x)}{f_\theta(y; x)}.$$

The mapping $\mathrm{SF}_X(y; x)$ is called the *X-score function* acting on the conditioning state (and is explained in more detail in Remark 1 below), the mapping $\mathrm{SF}_Y(y; x)$ is called the *Y-score function* acting on the resulting state, and $\mathrm{SF}_\Theta(y; x)$ is called the Θ-*score function* acting on the system parameter. Under appropriate smoothness conditions,

$$\frac{d}{dx} \int_S g(u) P_\theta(du; x) = \int_S g(u) \partial_x P_\theta(du; x) = \int_S g(y) \mathrm{SF}_X(y; x) f_\theta(y; x) dy,$$

and

$$\frac{d}{d\theta} \int_S g(u) P_\theta(du; x) = \int_S g(u) \partial_\theta P_\theta(du; x) = \int_S g(y) \mathrm{SF}_\Theta(y; x) f_\theta(y; x) dy.$$

We revisit the sojourn time example put forward in Example 1. Denote the transition kernel of the waiting time chain by P_θ; i.e.,

$$P_\theta(B, x) = \mathbb{P}(X_{n+1}(\theta) \in B | X_n(\theta) = x)$$

for $x \geq 0$ and $B \subset (0, \infty)$ a (Borel) measurable set, or, more formally,

$$P_\theta(B, x) = \int_0^\infty \left(\int_0^{s+x} 1_{\{x+s-a \in B\}} f_I(a) da \right) f_S(s) ds,$$

and otherwise, for $B \in [0, \infty)$ with $0 \in B$,

$$P_\theta(B, x) = \int_0^\infty \left(\int_0^{s+x} 1_{\{x+s-a \in B\}} f_I(a) da \right) f_S(s) ds + \int_0^\infty \left(\int_{s+x}^\infty f_I(s) ds \right) f_S(a) da.$$

Inserting $B = (0, y]$ and differentiating with respect to y, we obtain the density of the continuous part of the transition kernel on $(0, \infty)$ as

$$f_\theta(y; x) = \int_0^\infty f_I(\max\{x + s - y, 0\}) f_S(s) ds,$$

$y > 0$, where we assume that $f_I(0) = 0$, and

$$\partial_x f_\theta(y; x) = \int_0^\infty 1_{\{x+s-y>0\}} \partial_z f_I(x + s - y) f_S(s) ds,$$

where $\delta_z f_I(x + s - y)$ is shorthand notation for the derivative of $f_I(z)$ with respect to z evaluated at $z = x + s - y$. If, for example, interarrival times are exponentially distributed with rate λ, i.e., $f_I(x) = \lambda e^{-\lambda x}$, $x \geq 0$, then, by $\delta_x f_I(x) = -\lambda f_I(x)$,

$$\partial_x f_\theta(y; x) = -\lambda \int_0^\infty 1_{\{x+s-y>0\}} f_I(x + s - y) f_S(s) ds = -\lambda f_\theta(y; x), \qquad y > 0,$$

so that in this case

$$\mathrm{SF}_X(y; x) = -\lambda. \tag{8}$$

Following the same line of argument, we have that

$$\mathrm{SF}_Y(y; x) = \lambda. \tag{9}$$

Furthermore, assume that $S(\theta)$ follows a log-normal distribution with location parameter θ and scale parameter σ:

$$f_S(s) = \frac{1}{s\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(s) - \theta)^2}{2\sigma^2}\right), \quad s > 0.$$

It is well known that the log-normal distribution provides a good fit to the length of calls in a service centre; see, for example, [6]. Then it is easily checked that

$$\partial_\theta f_S(s) = \frac{\log(s) - \theta}{\sigma^2} f_S(s),$$

and it follows straightforwardly that

$$\partial_\theta f(y; x) = \int_0^\infty \frac{\log(s) - \theta}{\sigma^2} f_I(\max\{x + s - y, 0\}) f_S(s) ds.$$

The above expression cannot be applied in a simulation setting as such. In the following we will therefore rewrite $\mathrm{SF}_\Theta(y; x)$ by conditioning on the interarrival times. More specifically, let z denote the interarrival time and write $f(y; x, z)$ for the density of X_{n+1} given $X_n = x$ and $I_{n+1} = z$; then

$$f(y; x, z) = f_S(y + z - x), \quad y > 0,$$

and

$$\partial_\theta f(y; x, z) = \frac{\log(y + z - x) - \theta}{\sigma^2} f(y; x, z) \quad \text{and} \quad \mathrm{SF}_\Theta(y; x, z) = \frac{\log(y + z - x) - \theta}{\sigma^2}.$$

When evaluating SF_Θ via simulation, we first observe X_n, X_{n+1}, I_{n+1}, and $S_n(\theta)$, and we then apply differentiation conditional on I_{n+1}, which gives

$$\mathrm{SF}_\Theta(X_{n+1}; X_n, I_{n+1}) = \frac{\log(S_n(\theta)) - \theta}{\sigma^2}, \quad X_{n+1} > 0.$$

In words, we introduce a sub-Markov chain, where, for given X_n, first the interarrival time is sampled, giving $X_n - I_{n+1}$, and then in a subsequent step the service time is sampled and added.

The η-norm is extended to transition kernels through the operator norm; that is, for $Q \in \mathcal{K}(\mathcal{S}, \mathcal{T})$ we have

$$\|Q\|_\eta = \sup_{x \in \mathbb{R}} \frac{1}{\eta(x)} \int \eta(y) |Q(dy, x)|.$$

Note that $||I||_\eta = 1$ for I being the identity operator. By algebra,

$$
\begin{aligned}
\left| \int g(y)Q(dy;x) \right| &\leq \int |g(y)| \, |Q(dy;x)| \\
&= \int \frac{|g(y)|}{\eta(y)} \eta(y) \, |Q(dy;x)| \\
&\leq \int \left(\sup_y \frac{|g(y)|}{\eta(y)} \right) \eta(y)|Q(dy;x)| \\
&\leq ||g||_\eta \, \eta(x) \sup_x \frac{1}{\eta(x)} \int \eta(y)|Q(dy;x)| \\
&= ||g||_\eta \, ||Q||_\eta \eta(x), \tag{10}
\end{aligned}
$$

for all x. In the following we let $\eta_\alpha(x) = \alpha^{|x|}$, for $x \in \mathbb{R}$, and for $x \in [0, \infty)$ it suffices to let $\eta(x) = \alpha^x$. For a polynomial function ψ, we have $\|\psi\|_{\eta_\alpha} < \infty$ for any $\alpha > 1$. In slight abuse of notation, we will simplify $|| \cdot ||_{\eta_\alpha}$ to $|| \cdot ||_\alpha$.

Example 4. Let $A_\theta = [0, \theta]$ and consider the taboo kernel $\mathbf{T}_\theta = [P_\theta]_{[0,\theta]}$, with P_θ the transition kernel of the waiting times introduced in Example 1. We denote the stability region of the queue by $\Theta = \{\theta > 0 : \mathbb{E}[S(\theta) - I] < 0\}$. Then

$$
\begin{aligned}
||\mathbf{T}_\theta||_\alpha &= \sup_{x \geq 0} \alpha^{-x} \mathbb{E}\left[\alpha^{S(\theta)-I+x} 1_{\{S(\theta)-I+x>\theta\}} \right] \\
&= \sup_{x \geq 0} \mathbb{E}\left[\alpha^{S(\theta)-I} 1_{S(\theta)-I+x>\theta} \right] \\
&= \mathbb{E}\left[\alpha^{S(\theta)-I} \right].
\end{aligned}
$$

For $[\theta_1, \theta_2] \subset \Theta$, there exists α^* such that $\mathbb{E}[\alpha^{S(\theta)-I}]$ has negative right-hand-side derivative at $\alpha = 1$ and is strictly convex as a mapping in α on $[1, \alpha^*)$, uniformly on $[\theta_1, \theta_2]$. This shows that there exists $\bar{\alpha} > 1$ such that

$$
\sup_{\theta \in [\theta_1, \theta_2]} ||\mathbf{T}_\theta||_{\bar{\alpha}} < 1.
$$

For details we refer to [26, 29].

Note that for given $S(\theta)$ and I, we obtain from the above argument that the expected accumulated cost exists and is finite for all h such that $||h||_\alpha < \infty$, i.e., for all cost functions h that are bounded by $c\bar{\alpha}^x$ or some finite constant c. Indeed, letting τ denote the first time a waiting time becomes smaller than θ, for $h \geq 0$ and $1 \leq \alpha \leq \bar{\alpha}$ we obtain, from $||\mathbf{T}_\theta||_\alpha < 1$ together with $||\mathbf{T}_\theta^0||_\alpha = ||\mathbf{T}_\theta||_\alpha^0 = 1$ and $||T_\theta^i||_\alpha \leq ||T_\theta||_\alpha^i$ for $i \geq 1$, that

$$
\begin{aligned}
\mathbb{E}\left[\sum_{i=0}^{\tau-1} h(X_i) \,\Big|\, X_0 = x \right] &= \sum_{i=0}^{\infty} \left(\mathbf{T}_\theta^i h \right)(x) \\
&\leq \sum_{i=0}^{\infty} ||\mathbf{T}_\theta||_\alpha^i ||h||_\alpha \alpha^x \\
&= \alpha^x \frac{||h||_\alpha}{1 - ||\mathbf{T}_\theta||_\alpha} \\
&< \infty,
\end{aligned}
$$

provided that $||h||_\alpha$ is finite.

2.2. Differentiability of the taboo kernel

We call a taboo kernel $\mathcal{D}$-*differentiable* if a linear operator $\left([P_\theta]_{A_\theta}\right)'$ exists such that for any $h \in \mathcal{D}$ it holds that

$$\frac{d}{d\theta}\mathbf{T}_\theta h = \frac{d}{d\theta}[P_\theta]_{A_\theta} h = \left([P_\theta]_{A_\theta}\right)' h.$$

If the above holds, then $\left([P_\theta]_{A_\theta}\right)'$ is the Θ-$\mathcal{D}$-derivative of $\mathbf{T}_\theta$ and we let $\mathbf{T}'_\theta := \left([P_\theta]_{A_\theta}\right)'$.

Notice that the derivative of the taboo kernel $[P_\theta]_{A_\theta}$ cannot be straightforwardly obtained in a pathwise sense through differentiating the density as the support of the taboo kernel A_θ is parametrized, which introduces discontinuities. In the same way, the derivative of the taboo kernel $[P_\theta]_{A_\theta}$ cannot be obtained by simply taking the taboo kernel of the measure-valued derivative of P_θ, as it is in general not true that $[\partial_\theta P_\theta]_{A_\theta}$ is the derivative of $[P_\theta]_{A_\theta}$; in formula, $[\partial_\theta P_\theta]_{A_\theta} \neq \left([P_\theta]_{A_\theta}\right)'$.

For the analysis provided in this section we require A_θ to depend on θ in a smooth way. This is expressed in the following.

Assumption 1. (Boundary parametrization.) *There exists a boundary mapping $g_\theta(x)$, such that*

- *for some set Ω, independent of θ, we have*

$$x \in A_\theta^c \quad \text{if and only if} \quad g_\theta(x) \in \Omega;$$

- *the partial derivatives of g_θ with respect to θ and x exist;*

- *the set $\{x : \partial_x g_\theta(x) = 0\}$ has Lebesgue measure zero, so that*

$$\delta_\theta(x) = -\frac{\partial_\theta g_\theta(x)}{\partial_x g_\theta(x)}$$

  is a differentiable and Lebesgue-almost-surely well-defined mapping.

As an illustration of Assumption 1, note that the stopping criterion in (3) is formalized by $\Omega = [0, 1)$ and

$$g_\theta(x) = \frac{x}{\theta} \quad \text{and} \quad \delta_\theta(x) = g_\theta(x).$$

In this paper, we show that $\mathbf{T}'_\theta$ can be defined via cost functions ψ from an appropriate class of functions, to be specified later, so that

$$\frac{\partial}{\partial\theta}(\mathbf{T}_\theta\psi)(x) = \frac{\partial}{\partial\theta}\int \psi(y)\mathbf{T}_\theta(dy, x) = \int \psi(y)\mathbf{T}'_\theta(dy, x) = \left(\mathbf{T}'_\theta\psi\right)(x),$$

where

$$\left(\mathbf{T}'_\theta\psi\right)(x) = \int_{A_\theta^c}\left(\psi(y)\partial_\theta f_\theta(y; x) + \partial_y(\psi(y)f_\theta(y; x)\delta_\theta(y))\right)dy. \tag{11}$$

Suppose that P_θ is differentiable in an appropriate way, e.g., weakly differentiable [20, 23], with derivative P'_θ; then

$$\left[P'_\theta\right]_{A_\theta}$$

would be the first part of the estimator representing the derivative of P_θ on the complement of the taboo set. In order to get a derivative for $\mathbf{T}_\theta$ rather than just P_θ, we need an additional

term. This compensator collects the state-space derivatives; if we denote this operator by $\mathbf{D}$, Equation (11) reads

$$\left(\mathbf{T}_\theta\right)' = \left[P'_\theta\right]_{A_\theta} + \mathbf{D}_\theta,$$

where

$$\left(\left[P'_\theta\right]_{A_\theta}\psi\right)(x) = \int_{A_\theta^c}\psi(y)\partial_\theta f_\theta(y;x)dy = \int_{A_\theta^c}\psi(y)\mathrm{SF}_\Theta(y;x)f_\theta(y;x)dy,$$

and

$$\left(\mathbf{D}_\theta\psi\right)(x) = \int_{A_\theta^c}\partial_y(\psi(y)f_\theta(y;x)\delta_\theta(y))dy.$$

The kernel $\mathbf{D}_\theta$ lives on the complement of the taboo set A_θ. We provide an explicit form of $\mathbf{D}_\theta$ in the following example.

Example 5. Consider the taboo set $A_\theta = (\theta, \infty)$; then we may let $\Omega = [0, 1]$ with $g_\theta(x) = x/\theta$, for $x \geq 0$. For $\psi(y) = y^p$, for some $p \geq 1$, we obtain

$$\mathbf{D}_\theta\psi(x) = \int_0^\theta\frac{y^p}{\theta}\left((p+1) + y\mathrm{SF}_Y(y;x)\right)f_\theta(y;x)dy.$$

Because of the inner differentiation ∂y under the integral in $\mathbf{D}_\theta$, the operator $\mathbf{D}_\theta$ splits naturally into two parts. Indeed, writing $\mathbf{D}_\theta$ in explicit form, we obtain the following for $\mathbf{T}'_\theta$:

$$
\begin{aligned}
\left(\mathbf{T}'_\theta\psi\right)(x) = {} & \left(\left[P'_\theta\right]_{A_\theta}\psi\right)(x) + \left(\mathbf{D}_\theta\psi\right)(x)\\
= {} & \underbrace{\int_{A_\theta^c}\psi(y)\mathrm{SF}_\Theta(y;x)f_\theta(y;x)dy}_{=:\,\mathbf{T}'_1}\\
& + \underbrace{\int_{A_\theta^c}\psi'(y)\,\delta_\theta(y)\,f_\theta(y;x)dy}_{=:\,\mathbf{T}'_2}\\
& + \underbrace{\int_{A_\theta^c}\psi(y)\underbrace{\left(\delta_\theta(y)\mathrm{SF}_Y(y;x) + \partial_y\delta_\theta(y)\right)}_{=:\,\nu_\theta(y;x)}f_\theta(y;x)dy}_{=:\,\mathbf{T}'_3}.
\end{aligned}
\tag{12}
$$

In words, the derivative of $\mathbf{T}_\theta$ splits into three parts: (i) the first part assesses the sensitivity of the Markov chain outside the taboo set A_θ; (ii) the second part measures the knock-on effect of continuing the process on states belonging to the boundary of A_θ, where the unperturbed process otherwise would have stopped; and (iii) the third part measures the effect perturbing θ has on the performance at the next step reached from the state x, where the perturbation is weighted by the impact of θ on the boundary. We explain the meaning of the second part in more detail in the following remark.

Remark 1. Let $\tilde{\psi}$ be some cost function and denote the one-step-ahead expected value by

$$\psi(x) = \left(\mathbf{T}_\theta\tilde{\psi}\right)(x).$$

Observe that under appropriate smoothness conditions

$$\psi'(x) = \partial_x(\mathbf{T}_\theta \psi)(x) = \partial_x \int_{A_\theta^c} \tilde{\psi}(y) f_\theta(y; x) dy$$

$$= \int_{A_\theta^c} \tilde{\psi}(y) \partial_x f_\theta(y; x) dy = \int_{A_\theta^c} \tilde{\psi}(y) \mathrm{SF}_X(y; x) f_\theta(y; x) dy,$$

which shows that differentiation of the cost function ψ can be captured by differentiation of the transition density, provided that ψ itself can be interpreted as a conditional expectation. This is exactly what happens when products of Markov kernel operators are differentiated. In the following we provide a more general discussion of this phenomenon. Note that, for $n \geq 1$,

$$(\mathbf{T}_\theta^n \psi)(y) = \mathbb{E}\left[\psi(X_n) \prod_{i=0}^{n} \mathbf{1}\{X_i \notin A_\theta\} \,\middle|\, X_0 = y \right]$$

is the conditional expectation of the reward after n transitions of the taboo chain. Interchanging differentiation and integration leads to

$$\partial_y (\mathbf{T}_\theta^n \psi)(y) = \partial_y (\mathbf{T}_\theta \mathbf{T}_\theta^{n-1} \psi)(y)$$

$$= \int_{A_\theta^c} (\mathbf{T}_\theta^{n-1} \psi)(z) \partial_y f_\theta(z; y) dz.$$

Using the X-score function, this derivative can easily be evaluated as follows:

$$\int_{A_\theta^c} (\mathbf{T}_\theta^{n-1} \psi)(y) \, \partial_y f_\theta(z; y) dz = \int_{A_\theta^c} (\mathbf{T}_\theta^{n-1} \psi)(z) \mathrm{SF}_X(z, y) f_\theta(z; y) dy.$$

In words, the differentiation in the part $\mathbf{T}_2'$ of the derivative representation in (12) relates to differentiating a conditional expectation with respect to the conditioning value, which can be dealt with by using the X-score function.

Differentiating a Markov chain with taboo set A_θ is related to the differentiation of a distribution with θ-dependent support. In the following, we compare the application of our approach to the case of a distribution with θ-dependent support.

Example 6. Let $A_\theta = (\theta, 1]$, for $0 < \theta < 1$, and $f_\theta(y; x) = 1$ on $[0, 1]$ and zero otherwise. The taboo kernel can now be written as the defective uniform distribution

$$(\mathbf{T}_\theta \psi)(x) = \int_0^\theta \psi(y) dy,$$

which is clearly the expected value with respect to the uniform distribution on $[0, \theta]$ scaled by θ. This is a classical example that cannot be handled by the score function, as the density $1_{[0,\theta]}(y)$ fails to be Lipschitz continuous; here it is well known that the derivative is given by

$$\frac{\partial}{\partial \theta}(\mathbf{T}_\theta \psi)(x) = \psi(\theta), \tag{13}$$

provided that ψ is continuous.

Let $g_\theta(y) = y/\theta$, $\Omega = [0, 1]$, and $f_\theta(y; x) = 1_{[0,1]}$. Noticing that $\partial_\theta f_\theta(y; x) = 0 = \partial_y f_\theta(y; x)$ and

$$\delta_\theta(x) = \frac{x}{\theta}, \qquad \partial_x \delta(x) = \frac{1}{\theta},$$

we obtain from (12) (where we set the score functions to zero)

$$\frac{\partial}{\partial \theta}(\mathbf{T}_\theta \psi)(x) = \int_0^\theta \big(\partial_y \psi(y)\delta_\theta(y) + \psi(y)\partial_y \delta_\theta(y)\big)dy. \tag{14}$$

For example, in the case of $\psi(y) = y^p$, for some $p > 1$, it is easily seen that

$$\frac{\partial}{\partial \theta}(\mathbf{T}_\theta \psi)(x) = \theta^p,$$

which indeed yields the true derivative.

The uniform distribution on $[0, \theta]$ can be retrieved by

$$\frac{1}{\theta}(\mathbf{T}_\theta \psi)(x) = \frac{1}{\theta} \int_0^\theta \psi(y)dy.$$

Note that, for ψ continuous, it holds that

$$\frac{\partial}{\partial \theta}\left(\frac{1}{\theta}\int_0^\theta \psi(y)dy\right) = \frac{\partial}{\partial \theta}\left(\frac{1}{\theta}(\mathbf{T}_\theta \psi)(x)\right) = \frac{1}{\theta}\left(\psi(\theta) - \frac{1}{\theta}\int_0^\theta \psi(y)ydy\right)$$

for all x, where the right-hand side is an expression for the measure-valued derivative of the uniform distribution. Comparing the above measure-valued derivative with our new one,

$$\frac{\partial}{\partial \theta}\left(\frac{1}{\theta}(\mathbf{T}_\theta \psi)(x)\right) = \frac{1}{\theta}\left(\int_0^\theta \big(\partial_y \psi(y)\delta_\theta(y) + \psi(y)\partial_y \delta_\theta(y)\big)dy - \frac{1}{\theta}\int_0^\theta \psi(y)ydy\right),$$

obtained via (14), it is apparent that the resulting estimators are different. However, both estimators have as a common part

$$-\frac{1}{\theta^2}\int_0^\theta \psi(y)dy.$$

The difference lies in the fact that the (positive part of the) measure-valued derivative concentrates mass on the boundary of the taboo set, whereas our new estimator leads to integration over the complement of the taboo set. Our approach requires ψ to be differentiable, whereas for measure-valued differentiation only continuity is required.

That our approach requires stronger conditions on the performance function than classical measure-valued differentiation stems from the fact that we develop a differentiation theory for Markov chains (in contrast to the differentiation of unconditional measures), which requires capturing the knock-on effect of a perturbation on the future development of the Markov chain; see Remark 1.

In the following we provide the theoretical development for the estimator in (11). We denote the Euclidean distance between y and $\bar{y}$ by $\lambda(y, \bar{y})$ and define the distance from y to a set A by $\lambda(y, A) = \inf\{\lambda(y, \bar{y}) : \bar{y} \in A\}$. Note that $\lambda(y, A) = 0$ in the case $y \in A$. Recall that Ω denotes the

base set from which A_θ^c is obtained via $A_\theta^c = \{x : g_\theta(x) \in \Omega\}$; see Assumption 1. We now define ϵ sets around the boundary $\partial\Omega = \overline{\Omega} \setminus \Omega^o$ through

$$\Omega_\epsilon := \Omega \cup \{y \in \Omega^c : \lambda(y, \partial\Omega) \le \epsilon\} \tag{15}$$

and

$$\Omega_{-\epsilon} := \{y \in \Omega : \lambda(y, \partial\Omega) \ge \epsilon\}.$$

We let

$$\chi_\epsilon(y) := \frac{\lambda(y, \Omega_{-\epsilon})}{\lambda(y, \Omega_\epsilon^c) + \lambda(y, \Omega_{-\epsilon})},$$

where, by definition, $\chi_\epsilon(y) = 0$ for $y \in \Omega_{-\epsilon}$, $\chi_\epsilon(y) = 1$ for $y \in \Omega_\epsilon^c$, and $\|\chi_\epsilon\|_\infty \le 1$ (where $\|\psi\|_\infty := \sup_{y \in \mathbb{R}} |\psi(y)|$ for $\psi : \mathbb{R} \mapsto \mathbb{R}$). It is easy to see that $\chi_\epsilon(y)$ converges to $\mathbf{1}\{y \in \Omega\}$, except for $y \in \partial\Omega$, as ϵ tends to zero. Let

$$\bar{\chi}_\epsilon(y) := \int_{y \in \mathbb{R}} \chi_\epsilon(z) \phi_\epsilon(y - z) dz,$$

where ϕ_ϵ is the density of a centred normal distribution with variance ϵ^2, i.e.,

$$\phi_\epsilon(y) = \frac{1}{\epsilon\sqrt{2\pi}} \exp\left(-\frac{y^2}{2\epsilon^2}\right).$$

In words, $\bar{\chi}_\epsilon(y)$ is a smoothed version of $\chi_\epsilon(y)$ using a standard normal density as mollifier. Smoothness of the normal density implies

$$\bar{\chi}_\epsilon'(y) = \int_{\mathbb{R}} \chi_\epsilon(z) \partial_y \phi_\epsilon(y - z) dz$$
$$= \frac{1}{\epsilon\sqrt{2\pi}} \int_{\mathbb{R}} \chi_\epsilon(z) \frac{(z - y)}{\epsilon^2} \exp\left(-\frac{(y - z)^2}{2\epsilon^2}\right) dz.$$

The following lemma summarizes properties of $\bar{\chi}_\epsilon$.

Lemma 1. *The functions $\{\bar{\chi}_\epsilon\}$ are smooth and satisfy*

$$\|\bar{\chi}_\epsilon\|_\infty < \infty \quad and \quad \|\bar{\chi}_\epsilon'\|_\infty < \infty.$$

Proof. The smoothness of $\bar{\chi}_\epsilon(y)$ comes directly from the construction of the convolution, and $\|\bar{\chi}_\epsilon\|_\infty \le \|\chi_\epsilon\|_\infty \le 1$, which straightforwardly leads to the smoothness and boundedness of $\bar{\chi}_\epsilon$. In addition,

$$\|\bar{\chi}_\epsilon'\|_\infty \le \frac{\|\chi_\epsilon\|_\infty}{\epsilon^3\sqrt{2\pi}} \int_{\mathbb{R}} |z - y| \exp\left(-\frac{(y - z)^2}{2\epsilon^2}\right) dz$$
$$= \frac{2\|\chi_\epsilon\|_\infty}{\epsilon^3\sqrt{2\pi}} \int_{z=0}^{\infty} e^{-z} dz \le \frac{2}{\epsilon^2\sqrt{2\pi}},$$

which leads to the conclusion. $\qquad\square$

We introduce the following transition kernels:

$$\widetilde{\mathbf{P}}_\theta(B; x) := \int_B \delta_\theta(y) f_\theta(y; x) dy, \quad \widehat{\mathbf{P}}_\theta(B; x) := \int_B (\mathrm{SF}_\Theta(y; x) + v_\theta(y; x)) f_\theta(y; x) dy,$$

$$\bar{\mathbf{P}}_\theta(B; x) := \int_B \mathrm{SF}_X(y; x) f_\theta(y; x) dy,$$

and

$$\mathbf{U}(B; x) := \int_B \sup_{\theta \in \Theta} f_\theta(y; x) dy.$$

Note that $\mathbf{U}$ will be used to establish bounds needed for applying the dominated convergence theorem. The kernels $\widetilde{\mathbf{P}}_\theta$, $\widehat{\mathbf{P}}_\theta$, and $\bar{\mathbf{P}}_\theta$ represent different parts of the gradient expression.

Assumption 2. (Boundary integration.) *Assume the following:*

(i) *The sets $\partial A_\theta = \overline{A}_\theta \setminus A_\theta^o$ and $\mathcal{N}_\theta = \{x \in \mathbb{R} : \partial_x g_\theta(x) = 0\}$ have Lebesgue measure zero. Moreover, we assume that $\mathcal{N}_\theta$ is measurable for all θ.*

(ii) *It holds for α that*

$$\lim_{y \to \pm \infty} \alpha^{2|y|} f_\theta(y; x) = 0.$$

(iii) *The operator $\mathbf{U}$ has bounded norm $\|\mathbf{U}\|_{\alpha^2} < \infty$.*

(iv) *The Lebesgue measure $\mathcal{L}(\cdot)$ of $A_\theta^\epsilon \setminus A_\theta$ converges uniformly to zero, i.e.,*

$$\lim_{\epsilon \to 0} \sup_{\theta \in \Theta} \mathcal{L}(A_\theta^\epsilon \setminus A_\theta) = 0,$$

where

$$A_\theta^\epsilon := \{x \in \mathbb{R} : g_\theta(x) \in \Omega_\epsilon\}.$$

(v) *We have*

$$\sup_{\theta \in \Theta} ||\mathrm{SF}_\Theta||_\alpha < \infty, \qquad \sup_{\theta \in \Theta} ||\mathrm{SF}_X||_\alpha < \infty,$$

$$\sup_{\theta \in \Theta} ||\nu_\theta||_\alpha < \infty, \qquad \sup_{\theta \in \Theta} ||\partial_\theta g_\theta||_\alpha < \infty, \qquad \sup_{\theta \in \Theta} ||\delta_\theta||_\alpha < \infty.$$

The following results are needed to prove unbiasedness of the derivative estimator. The technical conditions put forward in Assumption 2 guarantee that the mass of the set effected by changing the boundary of A_θ is sufficiently smooth; see the conditions (i) and (iv). The remaining conditions establish that all objects needed for the analysis are uniformly (in θ) well-defined in the norm sense. In the case when the process X_n exhibits some monotonicity in θ, these conditions reduce to assuming that the norms are well-defined for the 'worst' choice of θ.

Lemma 2. *Under the conditions (iii)–(v) in Assumption 2, we have*

$$\sup_{\theta \in \Theta} \|P'_\theta\|_\alpha < \infty, \qquad \sup_{\theta \in \Theta} \|\widetilde{\mathbf{P}}_\theta\|_\alpha < \infty, \qquad \sup_{\theta \in \Theta} \|\widehat{\mathbf{P}}_\theta\|_\alpha < \infty, \qquad \sup_{\theta \in \Theta} \|\bar{\mathbf{P}}_\theta\|_\alpha < \infty,$$

and

$$\lim_{\epsilon \to 0} \left\|[P_\theta]_{A_\theta^\epsilon} - [P_\theta]_{A_\theta}\right\|_\alpha = 0, \lim_{\epsilon \to 0} \sup_{\theta \in \Theta} \left\|[\widetilde{\mathbf{P}}_\theta]_{A_\theta^\epsilon} - [\widetilde{\mathbf{P}}_\theta]_{A_\theta}\right\|_\alpha = 0,$$

$$\lim_{\epsilon \to 0} \sup_{\theta \in \Theta} \left\|[\hat{\mathbf{P}}_\theta]_{A_\theta^\epsilon} - [\hat{\mathbf{P}}_\theta]_{A_\theta}\right\|_\alpha = 0.$$

Proof. By definition, (v) implies that for some finite constant c,

$$\sup_{\theta \in \Theta} |\mathrm{SF}_\Theta(y; x)| \le c\alpha^x \alpha^y \quad \text{for all } x, y.$$

This implies

$$\sup_{\theta \in \Theta} \left\| P_\theta' \right\|_\alpha \le \sup_x \frac{\int \alpha^y \big(\sup_{\theta \in \Theta} |\mathrm{SF}_\Theta(y;x)| f_\theta(y;x)\big) dy}{\alpha^x}$$

$$\le \sup_x \frac{\int c\alpha^{2y} \alpha^x \big(\sup_{\theta \in \Theta} f_\theta(y;x)\big) dy}{\alpha^x}$$

$$= \int c\alpha^{2y} \left(\sup_{\theta \in \Theta} f_\theta(y;x)\right) dy < \infty,$$

by (iii). The rest of the results in the first part of the conclusion can be proved similarly. By definition, for some finite constant c,

$$\sup_{\theta \in \Theta} |\delta_\theta(y)| \le c\alpha^y \quad \text{for all } y.$$

Then,

$$\sup_{\theta \in \Theta} \left\| \left[\widetilde{\mathbf{P}}_\theta\right]_{A_\theta^\epsilon} - \left[\widetilde{\mathbf{P}}_\theta\right]_{A_\theta} \right\|_\alpha \le 2 \sup_{\theta \in \Theta} \int_{A_\theta^\epsilon \setminus A_\theta} \alpha^y |\delta_\theta(y)| f_\theta(y;x) dy$$

$$\le 2 \sup_{\theta \in \Theta} \int_{A_\theta^\epsilon \setminus A_\theta} c\alpha^{2y} \left(\sup_{\theta \in \Theta} f_\theta(y;x)\right) dy.$$

From the conditions (iii) and (iv),

$$\lim_{\epsilon \to 0} \sup_{\theta \in \Theta} \int_{A_\theta^\epsilon \setminus A_\theta} c\alpha^{2y} \left(\sup_{\theta \in \Theta} f_\theta(y;x)\right) dy = 0,$$

by the absolute continuity of Lebesgue integral [39]. The rest of the results in the second part of the conclusion can be proved similarly. □

Recall the definition of the v-norm $|| \cdot ||_\alpha$ in (23). We denote the set of differentiable mappings with finite $|| \cdot ||_\alpha$-norm and finite derivative $|| \cdot ||_\alpha$-norm by

$$\mathcal{D} := \{\psi : \|\psi\|_\alpha < \infty \text{ and } \|\partial_x \psi(x)\|_\alpha < \infty\}.$$

The following theorem establishes our main differentiability result with respect to a single-step transition.

Theorem 1. *Under Assumption* 1 *and Assumption* 2, *it holds for all* $h \in \mathcal{D}$ *that*

$$\frac{d}{d\theta} \mathbf{T}_\theta h = \mathbf{T}_\theta' h,$$

with $\mathbf{T}_\theta'$ *as in* (11), *or, equivalently,*

$$\mathbf{T}_\theta' = \left[P_\theta'\right]_{A_\theta} + \mathbf{D}_\theta.$$

Proof. First, we have for any h in $\mathcal{D}$ that

$$\frac{d}{d\theta} \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) \, f_\theta(y;x) dy$$

$$= \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) \, \mathrm{SF}_\Theta(y;x) \, f_\theta(y;x) dy + \int_{\mathbb{R}} \bar{\chi}_\epsilon'(g_\theta(y)) \, \partial_\theta g_\theta(y) \, h(y) \, f_\theta(y;x) dy, \tag{16}$$

where the interchange of differentiation and integration in the second equality is justified by the dominated convergence theorem, using that

$$\int_{\mathbb{R}} \sup_{\theta \in \Theta} \left| \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) \, \mathrm{SF}_\Theta(y; x) \right| f_\theta(y; x) dy \leq ||\bar{\chi}_\epsilon||_\infty \|h\|_\alpha \sup_{\theta \in \Theta} ||\mathrm{SF}_\Theta||_\alpha \|\mathbf{U}\|_{\alpha^2} < \infty$$

and

$$\int_{\mathbb{R}} \sup_{\theta \in \Theta} \left| \bar{\chi}'_\epsilon(g_\theta(y)) \, \partial_\theta g_\theta(y) \, h(y) f_\theta(y; x) \right| dy \leq ||\bar{\chi}'_\epsilon||_\infty \|h\|_\alpha \sup_{\theta \in \Theta} ||\partial_\theta g_\theta||_\alpha \|\mathbf{U}\|_{\alpha^2} < \infty,$$

which are justified by Lemma 1 and Assumption 1. By the chain rule, we solve for $\bar{\chi}'_\epsilon(g_\theta(y))$ from

$$\partial_y \bar{\chi}_\epsilon(g_\theta(y)) = \bar{\chi}'_\epsilon(g_\theta(y)) \, \partial_y g_\theta(y),$$

which implies that for every $y \in \mathbb{R} \setminus \mathcal{N}_\theta$,

$$\bar{\chi}'_\epsilon(g_\theta(y)) = \left(\partial_y g_\theta(y) \right)^{-1} \partial_y \bar{\chi}_\epsilon(g_\theta(y)). \tag{17}$$

The value of the function on $\mathcal{N}_\theta$, which has zero Lebesgue measure by Assumption 2, does not affect the value of the integral. Inserting (17) into (16) yields

$$\int_{\mathbb{R}} \bar{\chi}'_\epsilon(g_\theta(y)) \, \partial_\theta g_\theta(y) \, h(y) f_\theta(y; x) dy$$

$$= \int_{\mathbb{R}} h(y) \, \partial_y \bar{\chi}_\epsilon(g_\theta(y)) \, \partial_\theta g_\theta(y) \left(\partial_y g_\theta(y) \right)^{-1} f_\theta(y; x) dy$$

$$= - \int_{\mathbb{R}} \partial_y \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) \, \delta_\theta(y) f_\theta(y; x) dy. \tag{18}$$

Letting

$$u = \bar{\chi}_\epsilon(g_\theta(y))$$

and

$$v = h(y) \, \delta_\theta(y) f_\theta(y; x)$$

and applying the formula for integration by parts, $- \int u' v = -uv + \int u v'$, to the right-hand side of (18) yields

$$- \int_{\mathbb{R}} \partial_y \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) \, \delta_\theta(y) f_\theta(y; x) dy$$

$$= - h(y) \bar{\chi}_\epsilon(g_\theta(y)) \, \delta_\theta(y) f_\theta(y; x) \big|_{y=-\infty}^{\infty} + \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, \partial_y(h(y) \, \delta_\theta(y) f_\theta(y; x)) \, dy.$$

By Assumption 1,

$$\left| h(y) \bar{\chi}_\epsilon(g_\theta(y)) \, \delta_\theta(y) f_\theta(y; x) \big|_{y=-\infty}^{\infty} \right| \leq \|\bar{\chi}_\epsilon\|_\infty \|h\|_\alpha \|\delta_\theta\|_\alpha \, \alpha^{2|y|} f_\theta(y; x) \Big|_{y=-\infty}^{\infty} = 0.$$

To summarize,

$$\frac{d}{d\theta} \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) f_\theta(y; x) dy$$

$$= \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) \, \mathrm{SF}_\Theta(y; x) f_\theta(y; x) dy$$

$$+ \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \left(h'(y) \, \delta_\theta(y) + h(y) \, \nu_\theta(y; x) \right) f_\theta(y; x) dy,$$

where the integrability of h' stems from the assumption that $h' \in \mathcal{D}$. To justify the interchange of derivative and ϵ-limit, we have

$$\lim_{\epsilon \to 0} \sup_{\theta \in \Theta} \left| \int_{\mathbb{R}} (\bar{\chi}_\epsilon(g_\theta(y)) - \mathbf{1}\{g_\theta(y) \in \Omega\}) \, \gamma_\theta(y; x) f_\theta(y; x) dy \right|$$

$$\leq 2\|h\|_\alpha \lim_{\epsilon \to 0} \sup_{\theta \in \Theta} \left\| [\widehat{\mathbf{P}}_\theta]_{A_\theta^\epsilon} - [\widehat{\mathbf{P}}_\theta]_{A_\theta} \right\|_\alpha + 2\|h'\|_\alpha \lim_{\epsilon \to 0} \sup_{\theta \in \Theta} \left\| [\widetilde{\mathbf{P}}_\theta]_{A_\theta^\epsilon} - [\widetilde{\mathbf{P}}_\theta]_{A_\theta} \right\|_\alpha = 0,$$

where

$$\gamma_\theta(y; x) := h'(y) \, \delta_\theta(y) + h(y)(\mathrm{SF}_\Theta(y; x) + \nu_\theta(y; x)),$$

and

$$\lim_{\epsilon \to 0} \left| \int_{\mathbb{R}} (\bar{\chi}_\epsilon(g_\theta(y)) - \mathbf{1}\{g_\theta(y) \in \Omega\}) \, h(y) f_\theta(y; x) dy \right| \leq 2\|h\|_\alpha \lim_{\epsilon \to 0} \left\| [P_\theta]_{A_\theta^\epsilon} - [P_\theta]_{A_\theta} \right\|_\alpha = 0.$$

Summarizing the above analysis, we have

$$\frac{d}{d\theta} \int_{\mathbb{R}} \mathbf{1}\{g_\theta(y) \in \Omega\} \, h(y) \, f_\theta(y; x) dy = \lim_{\epsilon \to 0} \frac{d}{d\theta} \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, h(y) \, f_\theta(y; x) dy$$

$$= \lim_{\epsilon \to 0} \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, \gamma_\theta(y; x) \, f_\theta(y; x) dy$$

$$= \int_{\mathbb{R}} \mathbf{1}\{g_\theta(y) \in \Omega\} \gamma_\theta \ (y; x) \, f_\theta(y; x) dy, \qquad (19)$$

which is equivalent to

$$\frac{d}{d\theta} \mathbf{T}_\theta h = \mathbf{T}'_\theta h.$$

This proves the theorem. □

Remark 2. The theory justifying the interchange of limit, derivative, and integral can be found in [38, 39]. If h is a constant, the estimator in the theorem goes back to the generalized likelihood ratio method in [32].

3. Differentiation of the potential kernel

In this section, we derive the measure-valued derivative of the taboo kernel in an n-step transition, which is used to establish the product rule for the differentiation operation defined by (11).

Theorem 2. *Under Assumption 1 and Assumption 2, if $\|P_\theta\|_\alpha < \infty$, then it holds for all $h \in \mathcal{D}$ that*

$$\frac{d}{d\theta} \mathbf{T}_\theta^n h = \sum_{k=0}^{n-1} \mathbf{T}_\theta^{n-k-1} \mathbf{T}'_\theta \mathbf{T}_\theta^k h, \quad n \geq 1.$$

Proof. For $k \geq 1$ we have

$$
\frac{d}{d\theta} \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \left([P_\zeta]_{A_\zeta}^k h\right)(y) f_\theta(y; x) dy \Big|_{\zeta=\theta}
$$

$$
= \underbrace{\int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \left([P_\zeta]_{A_\zeta}^k h\right)(y) \, \mathrm{SF}_\Theta(y; x) f_\theta(y; x) dy \Big|_{\zeta=\theta}}_{=:[a]}
$$

$$
+ \underbrace{\int_{\mathbb{R}} \bar{\chi}_\epsilon'(g_\theta(y)) \, \partial_\theta g_\theta(y) \left([P_\zeta]_{A_\zeta}^k h\right)(y) f_\theta(y; x) dy \Big|_{\zeta=\theta}}_{=:[b]},
$$

where the interchange of differentiation and integration is justified similarly as in Theorem [1] by noticing

$$
\left|\left([P_\theta]_{A_\theta}^k h\right)(y)\right| \leq \|h\|_\alpha \|P_\theta\|_\alpha^k \eta_\alpha(y) < \infty;
$$

see (10). Then,

$$
[b] = - \underbrace{\left([P_\theta]_{A_\theta}^k h\right)(y) \bar{\chi}_\epsilon(g_\theta(y)) \, \delta_\theta(y) f_\theta(y; x) \Big|_{y=-\infty}^{\infty}}_{=:[b_1]}
$$

$$
+ \underbrace{\int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \, \partial_y \left(\left([P_\theta]_{A_\theta}^k h\right)(y) \, \delta_\theta(y) f_\theta(y; x)\right) dy}_{=:[b_2]}.
$$

By Assumption [2],

$$
[b_1] \leq \|\bar{\chi}_\epsilon\|_\infty \|h\|_\alpha \|P_\theta\|_\alpha^k \|\delta_\theta\|_\alpha \, \alpha^{2|y|} f_\theta(y; x) \Big|_{y=-\infty}^{\infty} = 0.
$$

As for the term b_2,

$$
[b_2] = \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \left(\int_{z \in \mathbb{R}} \mathbf{1}\{g_\theta(z) \in \Omega\} \left([P_\zeta]_{A_\zeta}^{k-1} h\right)(z) \partial_y f_\zeta(z; y) dz \delta_\theta(y) \right.
$$

$$
\left. + \left([P_\zeta]_{A_\zeta}^k h\right)(y) \nu_\theta(y; x) \right) f_\theta(y; x) dy \Big|_{\zeta=\theta}.
$$

We now bound $[a] + [b_2]$, where we take $[b_2]$ as on the right-hand side above,

$$
\limsup_{\epsilon \to 0} \sup_{\zeta \in \Theta} \left| \int_{\mathbb{R}} \left(\bar{\chi}_\epsilon(g_\zeta(y)) - \mathbf{1}\{g_\zeta(y) \in \Omega\} \right) \tilde{\gamma}_{\zeta,\theta}(y; x) f_\zeta(y; x) dy \right|
$$

$$
\leq 2\|h\|_\alpha \|P_\theta\|_\alpha^{k-1} \left\|\bar{\mathbf{P}}_\theta\right\|_\alpha \limsup_{\epsilon \to 0} \sup_{\theta \in \Theta} \left\|[\widehat{\mathbf{P}}_\theta]_{A_\theta^\epsilon} - [\widehat{\mathbf{P}}_\theta]_{A_\theta}\right\|_\alpha
$$

$$
+ 2\|h'\|_\alpha \|P_\theta\|_\alpha^k \limsup_{\epsilon \to 0} \sup_{\theta \in \Theta} \left\|[\widetilde{\mathbf{P}}_\theta]_{A_\theta^\epsilon} - [\widetilde{\mathbf{P}}_\theta]_{A_\theta}\right\|_\alpha = 0,
$$

where

$$
\tilde{\gamma}_{\zeta,\theta}(y; x) := \int_{z \in \mathbb{R}} \mathbf{1}\{g_\zeta(z) \in \Omega\} \left([P_\theta]_{A_\theta}^{k-1} h\right)(z) \, \mathrm{SF}_X^\zeta(z; y) \, dz \delta_\zeta(y)
$$

$$
+ \left([P_\theta]_{A_\theta}^k h\right)(y) \left(\mathrm{SF}_\Theta^\zeta(y; x) + \nu_\zeta(y; x)\right),
$$

and

$$\lim_{\epsilon \to 0} \left| \int_{\mathbb{R}} (\bar{\chi}_\epsilon(g_\theta(y)) - \mathbf{1}\{g_\theta(y) \in \Omega\}) \left([P_\zeta]^k_{A_\zeta} h\right)(y) f_\theta(y; x) dy \right|_{\zeta=\theta}$$

$$\leq 2\|h\|_\alpha \|P_\theta\|^k_\alpha \lim_{\epsilon \to 0} \left\| [P_\theta]_{A^\epsilon_\theta} - [P_\theta]_{A_\theta} \right\|_\alpha = 0.$$

Following the line of argument put forward in the proof (see (19)), we arrive at

$$\frac{d}{d\theta} \int_{\mathbb{R}} \mathbf{1}\{g_\theta(y) \in \Omega\} \left([P_\zeta]^k_{A_\zeta} h\right)(y) f_\theta(y; x) dy \Big|_{\zeta=\theta}$$

$$= \lim_{\epsilon \to 0} \frac{d}{d\theta} \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \left([P_\zeta]^k_{A_\zeta} h\right)(y) f_\theta(y; x) dy \Big|_{\zeta=\theta}$$

$$= \lim_{\epsilon \to 0} \int_{\mathbb{R}} \bar{\chi}_\epsilon(g_\theta(y)) \tilde{\gamma}_{\theta,\theta}(y; x) f_\theta(y; x) dy$$

$$= \int_{\mathbb{R}} \mathbf{1}\{g_\theta(y) \in \Omega\} \tilde{\gamma}_{\theta,\theta}(y; x) f_\theta(y; x) dy.$$

The case of $k = 0$ has already been treated in Theorem 1.

By algebra,

$$\frac{1}{\Delta} \left([P_{\theta+\Delta}]^n_{A_{\theta+\Delta}} h - [P_\theta]^n_{A_\theta} h\right) = \sum_{k=0}^{n-1} [P_{\theta+\Delta}]^{n-k-1}_{A_{\theta+\Delta}} \frac{1}{\Delta} \left([P_{\theta+\Delta}]_{A_{\theta+\Delta}} - [P_\theta]_{A_\theta}\right) [P_\theta]^k_{A_\theta} h.$$

There exists a $\Delta_0 > 0$ small enough so that for every $|\Delta| < \Delta_0$, it holds that $\theta + \Delta \in \Theta$. For $|\Delta| < \Delta_0$, there exists $\zeta \in \Theta$ such that

$$\frac{1}{\Delta} \left([P_{\theta+\Delta}]_{A_{\theta+\Delta}} - [P_\theta]_{A_\theta}\right) [P_\theta]^k_{A_\theta} h = \left(\left[\widetilde{\mathbf{P}}_\zeta\right]_{A_\zeta} \left[\bar{\mathbf{P}}_\zeta\right]_{A_\zeta} + \left[\widehat{\mathbf{P}}_\zeta\right]_{A_\zeta} [P_\theta]_{A_\theta}\right) [P_\theta]^{k-1}_{A_\theta} h$$

by the mean value theorem. Notice that for $k \geq 1$

$$\sup_{\zeta \in \Theta} \left| \left(\left[\widetilde{\mathbf{P}}_\zeta\right]_{A_\zeta} \left[\bar{\mathbf{P}}_\zeta\right]_{A_\zeta} + \left[\widehat{\mathbf{P}}_\zeta\right]_{A_\zeta} [P_\theta]_{A_\theta}\right) [P_\theta]^{k-1}_{A_\theta} h(y) \right|$$

$$\leq \|h\|_\alpha \|P_\theta\|^{k-1}_\alpha \left(\sup_{\theta \in \Theta} \left\|\widetilde{\mathbf{P}}_\theta\right\|_\alpha \sup_{\theta \in \Theta} \left\|\bar{\mathbf{P}}_\theta\right\|_\alpha + \|P_\theta\|_\alpha \sup_{\theta \in \Theta} \left\|\widehat{\mathbf{P}}_\theta\right\|_\alpha\right) \alpha^y < \infty.$$

By the dominated convergence theorem,

$$\lim_{\Delta \to 0} \sum_{k=0}^{n-1} [P_{\theta+\Delta}]^{n-k-1}_{A_{\theta+\Delta}} \frac{1}{\Delta} \left([P_{\theta+\Delta}]_{A_{\theta+\Delta}} - [P_\theta]_{A_\theta}\right) [P_\theta]^k_{A_\theta} h = \sum_{k=0}^{n-1} [P_\theta]^{n-k-1}_{A_\theta} [P_\theta]_{A'_\theta} [P_\theta]^k_{A_\theta} h,$$

which leads to the conclusion. □

In order for $(\mathbf{H}_\theta h)(x_0)$ (the accumulated cost up to the time of hitting the set A_θ) to exist, the process X_n needs to exhibit some kind of drift towards A_θ. This is well known from the theory of geometric ergodicity, where the drift towards a recurrent set or state is used to construct a taboo kernel. We refer to [26] for details and to Example 4 for an illustration of this principle.

In words, $||\mathbf{T}_\theta||_\alpha < 1$ means that the set A_θ carries enough mass of P_θ so that removing A_θ from the state-space reduces the norm of P_θ to less than one.

Theorem 3. *Under Assumption 1 and Assumption 2, if*

$$\sup_{\theta \in \Theta} ||\mathbf{T}_\theta||_\alpha < 1,$$

then it holds for all $h \in \mathcal{D}$ that

$$\frac{d}{d\theta} \sum_{n=0}^{\infty} \mathbf{T}_\theta^n h = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \mathbf{T}_\theta^n \mathbf{T}_\theta' \mathbf{T}_\theta^k h,$$

or, equivalently, $\mathbf{H}_\theta$ is $\mathcal{D}$-differentiable with derivative

$$\mathbf{H}_\theta' = \mathbf{H}_\theta \mathbf{T}_\theta' \mathbf{H}_\theta.$$

Proof. By Assumption 2 we have

$$\sup_{\theta \in \Theta} ||\mathbf{T}_\theta'||_\alpha =: c < \infty.$$

Letting

$$\sup_{\theta \in \Theta} ||\mathbf{T}_\theta||_\alpha = \rho,$$

it follows from (10) and the assumption that $\rho < 1$ that

$$\sum_{n=1}^{\infty} \sup_{\theta \in \Theta} \left| \sum_{k=0}^{n-1} \mathbf{T}_\theta^{n-k-1} \mathbf{T}_\theta' \mathbf{T}_\theta^k h(x) \right| < \sum_{n=1}^{\infty} \rho^{n-1} c \alpha^x = \frac{c\rho}{1-\rho} \alpha^x. \tag{20}$$

The proof of the proposition above is a simple application of Theorem 2 and the dominated convergence theorem. □

In the case $A_\theta = A$, Theorem 3 recovers the known results for measure-valued differentiation and the score function; see [22] and the discussion therein.

4. Estimation procedures

In this section we translate the gradient expression put forward in Theorem 3 into simulation estimators. Let

$$g(x) = \left(\sum_{k=0}^{\infty} \mathbf{T}_\theta' \mathbf{T}_\theta^k \psi \right)(x), \tag{21}$$

where x denotes the initial value of the individual transition operators $\mathbf{T}_\theta' \mathbf{T}_\theta^k$.

Suppose that we evaluate $(\mathbf{T}_\theta' \mathbf{T}_\theta^k \psi)(X_i)$ at $X_i \notin A_\theta$. This kernel represents the event that we start in X_i and evaluate the expected value of ψ along a path that does not enter the taboo set, where the first transition of the path is governed by the derivative kernel $\mathbf{T}_\theta'$. Elaborating

on (12), we can replace $\mathbf{T}'_\theta$ by $\mathbf{T}_\theta$ with appropriate scaling via the score functions and the derivatives of δ_θ and ψ. This leads to the following estimator for $\left(\mathbf{T}'_\theta \mathbf{T}^k_\theta \psi\right)(X_i)$ at $X_i \notin A_\theta$:

$$\left(\mathbf{T}'_\theta \mathbf{T}^k_\theta \psi\right)(X_i) = \mathbb{E}\left[\phi_\psi(X_i; X_{i+1}, \ldots, X_{i+k+1}) \,\middle|\, X_i\right], \quad k \geq 1,$$

where

$$\phi_\psi(X_i; X_{i+1}, \ldots, X_{i+k+1}) =$$
$$\prod_{j=i}^{i+k+1} \mathbf{1}\{X_j \notin A_\theta\} \left(\underbrace{\psi(X_{i+1+k})\mathrm{SF}_\Theta(X_{i+1}; X_i)}_{\mathbf{T}'_1} + \underbrace{\psi(X_{i+1+k})\mathrm{SF}_X(X_{i+2}; X_{i+1})\delta_\theta(X_{i+1})}_{\mathbf{T}'_2} \right.$$
$$\left. + \underbrace{\psi(X_{i+1+k})\left(\mathrm{SF}_Y(X_{i+1}; X_i)\delta_\theta(X_{i+1}) + \partial_y \delta_\theta(X_{i+1})\right)}_{\mathbf{T}'_3} \right), \quad k \geq 1,$$

and, for $k = 0$, we can estimate $\mathbf{T}'_\theta \mathbf{T}^0_\theta \psi(X_i) = \mathbf{T}'_\theta \psi(X_i) = \mathbb{E}[\phi_\psi(X_i; X_{i+1})|X_i]$, for $X_i \notin A_\theta$, via

$$\phi_\psi(X_i; X_{i+1}) = \mathbf{1}\{X_i, X_{i+1} \notin A_\theta\} \left(\underbrace{\psi(X_{i+1})\mathrm{SF}_\Theta(X_{i+1}; X_i)}_{\mathbf{T}'_1} + \underbrace{\psi'(X_{i+1})\delta_\theta(X_{i+1})}_{\mathbf{T}'_2} \right.$$
$$\left. + \underbrace{\psi(X_{i+1})\left(\mathrm{SF}_Y(X_{i+1}; X_i)\delta_\theta(X_{i+1}) + \partial_y \delta_\theta(X_{i+1})\right)}_{\mathbf{T}'_3} \right).$$

Consider a cycle $\{X_0, X_1, \ldots, X_{\tau(x)-1}\}$ of the Markov chain, with $X_0 = x \notin A_\theta$, where we have expanded the notation for the first entrance time into A_θ, denoted so far by τ, by explicitly indicating the initial state. Starting the count from 0, the cycle length is $\tau(x)$. As the transitions from X_0 to X_1 up to the transition from $X_{\tau(x)-2}$ to $X_{\tau(x)-1}$ do not enter the taboo set, they may be considered as having been sampled via $\mathbf{T}_\theta$, whereas the transition from $X_{\tau(x)-1}$ to $X_{\tau(x)}$ cannot be considered as having been driven by $\mathbf{T}_\theta$, since with this transition the Markov chain enters the taboo set. This shows that there are $\tau(x) - 1$ taboo kernel transitions in a cycle of length $\tau(x)$, and the indices of the states from which a taboo kernel transition is executed run from 0 to $\tau(x) - 2$. Following the line of argument put forward in (5),

$$g(x) = \sum_{k=0}^{\infty} \left(\mathbf{T}'_\theta \mathbf{T}^k_\theta \psi\right)(x) = \sum_{k=0}^{\infty} \mathbb{E}\left[\prod_{i=0}^{k+1} \mathbf{1}\{X_i \notin A_\theta\}\phi_\psi(X_0; X_1, \ldots, X_{k+1}) \,\middle|\, X_0 = x\right]$$
$$= \mathbb{E}\left[\sum_{i=1}^{\tau(x)-1} \phi_\psi(X_0; X_1, \ldots, X_i) \,\middle|\, X_0 = x\right]$$

for $x \notin A_\theta$, where we set the sum to zero in the case where the lower bound is larger than the upper bound (which happens for realizations where $\tau(x) = 1$).

Treating the outer sum in the same way, we obtain from Theorem 3, with $x_0 \notin A_\theta$,

$$\frac{d}{d\theta} \sum_{n=0}^{\infty} \left(\mathbf{T}_\theta^n \psi\right)(x_0) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \left(\mathbf{T}_\theta^n \mathbf{T}_\theta' \mathbf{T}_\theta^k \psi\right)(x_0) \tag{22}$$

$$= \sum_{n=0}^{\infty} \left(\mathbf{T}_\theta^n g\right)(x_0)$$

$$= \sum_{n=0}^{\infty} \mathbb{E}\left[\prod_{i=0}^{n} \mathbf{1}\{X_i \notin A_\theta\} g(X_n) \,\middle|\, X_0 = x_0\right]$$

$$= \mathbb{E}\left[\sum_{i=0}^{\tau(x_0)-2} \sum_{k=i+1}^{\tau(x_0)-1} \phi_\psi(X_i; X_{i+1}, \ldots, X_k) \,\middle|\, X_0 = x_0\right].$$

Simulation offers the opportunity to first sample the time horizon $\tau(x)$. Then, re-initiating the random number stream, we can exactly repeat the stochastic experiment. Given $\tau(x)$, we can sample σ independently from everything else uniformly distributed on $\{0, 1, \ldots, \tau(x) - 2\}$. With this definition we can apply the randomization to the outer $\tau(x_0)$ sum in (22) and obtain the following alternative version of the estimator:

$$\left(\frac{d}{d\theta} \sum_{n=0}^{\infty} \mathbf{T}_\theta^n \psi\right)(x_0) = \mathbb{E}\left[(\tau(x_0) - 1) \sum_{k=\sigma+1}^{\tau(x_0)-1} \phi_\psi(X_\sigma; X_{\sigma+1}, \ldots, X_k) \,\middle|\, X_0 = x_0\right].$$

Notice that the above estimator contains several likelihood ratio terms, namely, $\mathrm{SF}_\Theta(y; x)$, $\mathrm{SF}_X(y; x)$ as a term of ν_θ, and $\mathrm{SF}_Y(y; x)$, which may lead to a large variance of the derivative estimator. Following the standard approach in measure-valued differentiation [20, 18, 23, 33, 34], we may apply the Hahn–Jordan decomposition,

$$\partial_\theta f_\theta(y; x) = c_\theta(x) \big(f_\theta^+(y; x) - f_\theta^-(y; x)\big),$$

where

$$f_\theta^+(y; x) := \left(\partial_\theta f_\theta(y; x)\right)^+ / c_\theta(x), \quad f_\theta^-(y; x) := \left(\partial_\theta f_\theta(y; x)\right)^- / c_\theta(x),$$

$$c_\theta(x) := \int_{\mathbb{R}} \left(\partial_\theta f_\theta(y; x)\right)^+ dy = \int_{\mathbb{R}} \left(\partial_\theta f_\theta(y; x)\right)^- dy,$$

by noticing that

$$\int_{\mathbb{R}} \left(\partial_\theta f_\theta(y; x)\right)^+ dy - \int_{\mathbb{R}} \left(\partial_\theta f_\theta(y; x)\right)^- dy = \partial_\theta \int_{\mathbb{R}} f_\theta(y; x) dy = 0.$$

The Hahn–Jordan decomposition is often not easy to sample from; typically one can find decompositions that are easier to sample from (see [20]).

We turn to the state derivative, i.e., the Y-score function $\mathrm{SF}_Y(y; x)$. We obtain that

$$\partial_y f_\theta(y; x) = \hat{c}_\theta(x) \big(\hat{f}^+(y; x) - \hat{f}^-(y; x)\big),$$

where

$$\hat{f}_\theta^+(y; x) := \left(\partial_y f_\theta(y; x)\right)^+ / \hat{c}_\theta(x), \quad \hat{f}_\theta^-(y; x) := \left(\partial_y f_\theta(y; x)\right)^- / \hat{c}_\theta(x),$$

$$\hat{c}_\theta(x) := \int_{\mathbb{R}} \left(\partial_y f_\theta(y; x)\right)^+ dy = \int_{\mathbb{R}} \left(\partial_y f_\theta(y; x)\right)^- dy,$$

by noticing

$$\int_{\mathbb{R}} \left(\partial_y f_\theta(y;x)\right)^+ dy - \int_{\mathbb{R}} \left(\partial_y f_\theta(y;x)\right)^- dy = f_\theta(y;x)\big|_{y=-\infty}^{\infty} = 0.$$

The estimator also requires differentiation of the conditional expectation of the future costs with respect to the initial value captured by $\mathrm{SF}_X(y;x)$. For this we construct the decomposition

$$\partial_x f_\theta(y;x) = \bar{c}_\theta(x) \left(\bar{f}_\theta^+(y;x) - \bar{f}_\theta^-(y;x)\right),$$

where

$$\bar{f}_\theta^+(y;x) := \left(\partial_x f_\theta(y;x)\right)^+ / \bar{c}_\theta(x), \quad \bar{f}_\theta^-(y;x) := \left(\partial_x f_\theta(y;x)\right)^- / \bar{c}_\theta(x),$$

$$\bar{c}_\theta(x) := \int_{\mathbb{R}} \left(\partial_x f_\theta(y;x)\right)^+ dy = \int_{\mathbb{R}} \left(\partial_x f_\theta(y;x)\right)^- dy,$$

by noticing that

$$\int_{\mathbb{R}} \left(\partial_x f_\theta(y;x)\right)^+ dy - \int_{\mathbb{R}} \left(\partial_x f_\theta(y;x)\right)^- dy = \partial_x \int_{\mathbb{R}} f_\theta(y;x) dy = 0.$$

Elaborating on the above decompositions, the score functions in the estimator can be replaced by appropriate splittings of the sample paths, where the positive and negative parts are given by the Markov kernels corresponding to the above densities. We refer to the literature on measure-valued differentiation for details.

Following the framework provided in [32], the results presented in this paper can be extended to the multidimensional case, where $S \subset \mathbb{R}^d$, for $d > 1$. Specifically, the weight supremum norm is extended to $\mathbb{R}^d$ by letting

$$\eta_\alpha(x_1, \ldots, x_d) = \alpha^{|x_1|} \cdots \alpha^{|x_d|}, \quad \alpha > 1. \tag{23}$$

5. Applications

Consider the following simple capital flow model for an unlimited liability company. Suppose at the nth period a company receives a random profit of W_n. If the company has a positive capital at the nth period, then (i) rX_n is given to the shareholder as a dividend, and (ii) the remaining capital is topped up by $M > 0$, and the excess of capital $\max(rX_n - M, 0)$ is extracted; otherwise, the company receives $-rX_n$ capital from the shareholders, for some $r \in (0, 1)$. The company has credit to borrow money up to $\theta \geq 0$, and if the company's capital is below $-\theta$, then it is ruined and no longer has any debt obligations. Then, with $X_0 = x$,

$$\tau := \tau(x) = \inf\{n \geq 1 : X_n + \theta \leq 0\},$$

and

$$X_n = (1 - r)X_{n-1} + W_n,$$

for $X_n \leq M$ and M otherwise. The cash flow of the company is

$$\sum_{n=0}^{\tau-1} rX_n,$$

where $X_0 = x_0$ gives the initial value. Let W_n, $n \in \mathbb{N}$, be a sequence of i.i.d. random variables that follow a normal distribution with mean $\mu = 0.1$ and variance $\sigma^2 = 1$, and set $x_0 = 0.3$,

$r = 0.3$, and $\theta = 0$. Let P_θ denote the transition kernel of the Markov chain X_n. Then $(P_\theta)(dy, x)$ has a transition density

$$f(y; x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(y - \mu - (1 - r)x)^2}{2\sigma^2}\right)$$

on $(-\infty, M]$ and a point mass

$$(P_\theta)(M, x) = \int_M^\infty f(y; x)dy.$$

We have $\Omega = (-\infty, 0]$, $g(x; \theta) = x + \theta$, $A_\theta = (-\infty, -\theta]$, and $\mathbf{T}_\theta = [P_\theta]_{(-\infty, -\theta]}$.

We are interested in estimating the sensitivity of the total cash flow until ruin,

$$h(\theta, x_0) = (\mathbf{H}_\theta \psi)(x) = \sum_{n=0}^\infty (\mathbf{T}_\theta^n \psi)(x) = \mathbb{E}\left[\sum_{n=0}^{\tau-1} rX_n \,\middle|\, X_0 = x_0\right].$$

We have

$$\|\mathbf{T}\|_\alpha = \sup_{x \leq M} \alpha^{-|x|} \left(\int_{-\theta}^M \alpha^{|y|} f(y; x)dy + \alpha^M \int_M^\infty f(y; x)dy\right),$$

and letting $\alpha = 1$ (which we may do as $|X_n| \leq M$ is bounded for $|\theta| < M$ and thus $\psi(X_n)$ also is) gives

$$\|\mathbf{T}\|_1 = \sup_{x \leq M} \int_{-\theta}^\infty f(y; x)dy = \mathbb{P}\big(\sigma Z + \mu + (1 - r)x > -\theta\big),$$

for Z a standard normal random variable. It follows that

$$\|\mathbf{T}\|_1 = \mathbb{P}\big(\sigma Z + \mu + (1 - r)M > -\theta\big) < 1$$

for any $\theta \leq 0$. Obviously, Assumptions 1 and 2 are satisfied, and by Theorem 3 we obtain

$$\partial_\theta h(\theta, x_0) = \big(\mathbf{H}_\theta \mathbf{T}_\theta' \mathbf{H}_\theta \psi\big)(x_0).$$

For ease of analysis, we let $M = \infty$ in the following. This is justified as $rX_n > M$ is a rare event for the numerical setting of the example, which can be easily justified numerically. Note that $\delta_\theta(x) = -1$, and

$$\mathrm{SF}_X(y; x) = \frac{1 - r}{\sigma^2}\left(y - \mu - (1 - r)x\right), \quad \mathrm{SF}_Y(y; x) = \frac{1}{\sigma^2}\left((1 - r)x + \mu - y\right).$$

After some algebra, we obtain for the unbiased estimator

$$-r(\tau - 1) + \frac{r}{\sigma^2}\left(\sum_{i=1}^{\tau-1} X_i\right)$$

$$\times \left\{\sum_{i=1}^{\tau-1}\left[(1 + (1 - r)^2)X_i - (1 - r)(X_{i-1} + X_{i+1}) - r\mu\right] + X_\tau - (1 - r)X_{\tau-1} - \mu\right\}.$$

TABLE 1. Derivative estimates for the expected cash flow in the ruin problem (mean ± standard error).

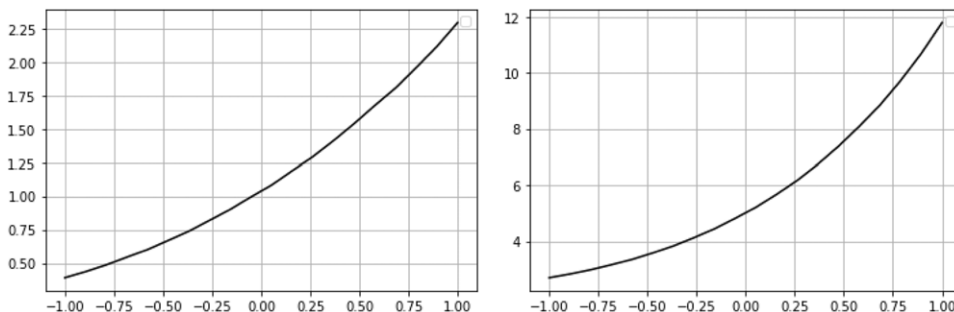|	this paper	FD(0.1)	FD(0.01)
$r = 0.3$, $n = 10^6$, mean ± std err:	1.24 ± 0.007	1.28 ± 0.04	1.20 ± 0.3
$r = 0.9$, $n = 10^8$, mean ± std err:	0.59 ± 0.009	0.53 ± 0.05	0.23 ± 0.5

FIGURE 1. Expected cash flow (left) and time to ruin (right) as a function of θ for $r = 0.3$.

We plot the total cash flow and the expected time to ruin in Figure 1. We compare the performance of our estimator to that of a finite-difference estimator (FD) defined by

$$\frac{1}{\delta} \left(\sum_{n=1}^{\tau_{\theta+\delta}-1} rX_n - \sum_{n=1}^{\tau_{\theta}-1} rX_n \right);$$

see [30]. We denote the FD with perturbation size 0.1 by FD(0.1), and the FD with perturbation size 0.01 by FD(0.01). The means and standard errors of the three derivative estimators are reported in Table 1, where we show results for $r = 0.3$ and additionally for $r = 0.9$, denoting by n the number of independent experiments used for the estimators.

In Table 1, we can see that the estimator in this paper leads to the lowest variance, while FD suffers from the bias–variance dilemma. In addition, the computation time of our estimator is about half that of FD, because the former only needs to simulate one sample path, while the latter needs to simulate two sample paths. The key to applying the common random numbers (CRN) technique is to synchronize the simulations of the two sample paths [28]. For the random horizon problem, it is not clear how to efficiently synchronize simulation of the sample paths under the perturbed parameter. Thus, it may not be easy to implement CRN to reduce the variance of FD. In this example, there is no need to calculate the Hahn–Jordan decomposition, but we note that one could split the random variables as follows: $\hat{c} = -1/\sqrt{2}$, $\bar{c} = (1 - r)/\sqrt{2\pi}$, and

$$\widehat{X}_n^+, \bar{X}_n^+ \sim \mu + (1 - r)X_{n-1} + Wei(2, \sqrt{2}\sigma), \quad \widehat{X}_n^-, \bar{X}_n^- \sim \mu + (1 - r)X_{n-1} - Wei(2, \sqrt{2}\sigma),$$

where $Wei(a, b)$ denotes a Weibull-distributed random variable with the density

$$\frac{a}{b} \left(\frac{x}{b} \right)^{a-1} e^{-(x/b)^a} \mathbf{1}\{x \geq 0\}.$$

6. Conclusion

In this paper, we have developed an approach for differentiating random horizon problems, where the parameter of interest may affect the stopping criterion as well as the transition dynamics of the underlying stochastic process. We discuss the multidimensional case and various implementations of the obtained derivative expressions.

Acknowledgements

Competing interests

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

References

[1] ASMUSSEN, S. *et al.* (2008). Asymptotic behavior of total times for jobs that must start over if a failure occurs. *Math. Operat. Res.* **33**, 932–944.

[2] ASMUSSEN, S., LIPSKY, L. AND THOMPSON, S. (2014). Checkpointing in failure recovery in computing and data transmission. In Analytical and Stochastic Modelling Techniques and Applications *(ASMTA 2014)*, eds B. Sericola, M. Telek and G. Horváth, Springer, Cham, pp. 253–272.

[3] AVRACHENKOV, K., PIUNOVSKIY, A. AND ZHANG, Y. (2015). Hitting times in Markov chains with restart and their application to network centrality. *Methodology Comput. Appl. Prob.* **20**, 1173–1188.

[4] BASHYAM, S. AND FU, M. (1994). Application of perturbation analysis to a class of periodic review (*s*, *S*) inventory systems. *Naval Res. Logistics* **41**, 47–80.

[5] BASHYAM, S. AND FU, M. (1998). Optimization of (*s*, *S*) inventory systems with random lead times and a service level constraint. *Manag. Sci.* **44**, 243–256.

[6] BROWN, L. *et al.* (2005). Statistical analysis of a telephone call center. *J. Amer. Statist. Assoc.* **100**, 36–50.

[7] CAO, X. (2007). *Stochastic Learning and Optimization: a Sensitivity-Based Approach.* Springer, New York.

[8] CASWELL, H. (2013). Sensitivity analysis of discrete Markov chains via matrix calculus. *Linear Algebra Appl.* **438**, 1727–1745.

[9] CASWELL, H. (2019). *Sensitivity Analysis: Matrix Methods in Demography and Ecology.* Springer, Cham.

[10] COHN, D. (1980). *Measure Theory.* Birkhäuser, Stuttgart.

[11] DEKKER, R. *et al.* (1998). Maintenance of light-standards—a case-study. *J. Operat. Res. Soc.* **49**, 132–143.

[12] FU, M. AND HU, J. Q. (1997). *Conditional Monte Carlo: Gradient Estimation and Optimization Applications.* Kluwer, Boston.

[13] FU, M. C. (2006). Gradient estimation. In *Handbooks in Operations Research and Management Science*, Vol. **13**, *Simulation*, eds S. Henderson and B. Nelson, North Holland, Amsterdam, pp. 575–616.

[14] GLASSERMAN, P. (1991). *Gradient Estimation via Perturbation Analysis.* Kluwer, Boston.

[15] GLASSERMAN P. (2004). *Monte Carlo Methods in Financial Engineering.* Springer, New York.

[16] HEIDERGOTT, B. (2001). A weak derivative approach to optimization of threshold parameters in a multi-component maintenance system. *J. Appl. Prob.* **38**, 386–406.

[17] HEIDERGOTT, B. (2001). Option pricing via Monte Carlo simulation: a weak derivative approach. *Prob. Eng. Inf. Sci.* **15**, 335–349.

[18] HEIDERGOTT, B. (2007). *Max-Plus Linear Stochastic Models and Perturbation Analysis.* Springer, New York.

[19] HEIDERGOTT, B. AND FARENHORST-YUAN, T. (2010). Gradient estimation for multicomponent maintenance systems with age-replacement policy. *Operat. Res.* **58**, 706–718.

[20] HEIDERGOTT, B., HORDIJK, A. AND WEISSHAUPT, H. (2008). Derivatives of Markov kernels and their Jordan decomposition. *J. Appl. Anal.* **14**, 13–26.

[21] HEIDERGOTT, B., LEAHU, H. AND VOLK-MAKAREWICZ, W. (2014) A smoothed perturbation analysis of Parisian options. *IEEE Trans. Automatic Control* **60**, 469–474.

[22] HEIDERGOTT, B. AND VÁZQUEZ-ABAD, F. (2006). Measure-valued differentiation for random horizon problems. *Markov Process. Relat. Fields* **12**, 509–536.

[23] HEIDERGOTT, B. AND VÁZQUEZ-ABAD, F. (2008). Measure-valued differentiation for Markov chains. *J. Optimization Theory Appl.* **136**, 187–209.

[24] HO, Y. C. AND CAO, X. (1991). *Perturbation Analysis of Discrete Event Dynamic Systems*. Kluwer, Boston.

[25] KALLENBERG O. (2001). *Foundations of Modern Probability*, 2nd edn. Springer, New York.

[26] KARTASHOV, N. (1996). *Strong Stable Markov Chains*. De Gruyter, Zeist.

[27] KULKARNI, G., NICOLA, V. AND TRIVEDI, S. (1987). The completion time of a job on multimode systems. *Adv. Appl. Prob.* **19**, 932–954.

[28] LAW, A. AND KELTON, D. (2000). *Simulation Modeling and Analysis*. McGraw-Hill, Boston.

[29] LEAHU, H. (2008). Measure-valued differentiations for finite products of measures. Doctoral Thesis, Vrije Universiteit Amsterdam.

[30] L'ECUYER, P. AND PERRON, G. (1994). On the convergence rates of IPA and FDC derivative estimators. *Operat. Res.* **42**, 643–656.

[31] LYUU, Y.-D. AND TENG, H.-W. (2011). Unbiased and efficient Greeks of financial options. *Finance Stoch.* **15**, 141–181.

[32] PENG, Y., FU, M. C., HU, J. Q. AND HEIDERGOTT, B. (2018). A new unbiased stochastic derivative estimator for discontinuous sample performances with structural parameters. *Operat. Res.* **66**, 487–499.

[33] PFLUG, G. (1992). Gradient estimates for the performance of Markov chains and discrete event processes. *Ann. Operat. Res.* **39**, 173–194.

[34] PFLUG, G. (1996). *Optimisation of Stochastic Models*. Kluwer, Boston.

[35] PFLUG, G. AND RUBINSTEIN, R. (2002). Inventory processes: quasi-regenerative property, performance evaluation, and sensitivity estimation via simulation. *Stoch. Models* **18**, 469–496.

[36] RUBINSTEIN, R. (1992). Sensitivity analysis of discrete event systems by the 'push out' method. *Ann. Operat. Res.* **39**, 229–250.

[37] RUBINSTEIN, R. AND SHAPIRO, A. (1993). *Discrete Event Systems: Sensitivity Analysis and Optimization by the Score Function Method*. John Wiley, Chichester.

[38] RUDIN, W. (1964). *Principles of Mathematical Analysis*. McGraw-Hill, New York.

[39] RUDIN, W. (1987). *Real and Complex Analysis*. McGraw-Hill, New York.