Measure–Valued Differentiation for Stationary Markov Chains

Bernd Heidergott[†]
Arie Hordijk[‡]
Heinz Weisshaupt[§]

24.07.2002

Abstract: We study general state-space Markov chains that depend on a parameter, say, θ. Sufficient conditions are established for the stationary performance of such a Markov chain to be differentiable with respect to θ. Specifically, we study the case of unbounded performance functions and thereby extend the result on weak differentiability of stationary distributions of Markov chains to unbounded mappings. The two main ingredients of our approach are (a) that we work within the framework of measure-valued differentiation (MVD) in order to study derivatives of unbounded performance functions, and (b) that we elaborate on normed ergodicity of Markov chains in order to establish the existence of the overall derivative expression. Our approach is not restricted to a particular estimation method. In fact, MVD expressions can be translated into various unbiased estimators. We illustrate our results with examples from queuing theory.

Keywords: measure-valued differentiation, perturbation analysis, normed ergodicity, Markov chains

[†]TU Eindhoven, Department of Mathematics and Computing Science
Division of Mathematics
P.O.Box 513, 5600 MB Eindhoven, the Netherlands
email: b.heidergott@tue.nl
[‡]Leiden University, Mathematical Institute
P.O.Box 9512, 2300 RA Leiden, the Netherlands
email: hordijk@math.leidenuniv.nl
[§]EURANDOM
P.O.Box 513, 5600 MB Eindhoven, the Netherlands
email: weisshaupt@eurandom.nl

1 Introduction

In recent years a great deal of attention has been devoted to the computation of derivatives of performance indicators in stochastic systems. More specifically, suppose that the system can be modeled by a (general state–space) Markov chain $\{X_\theta(n)\}$, depending on a (vector–valued) parameter $\theta \in \Theta$, and assume that the proces is ergodic for any $\theta \in \Theta$, that is, $X_\theta(n)$ converges, independent of the initial state, to a steady–state $X_\theta(\infty)$. We would like to compute the gradient of the expected value of the Markov chain in equilibrium, that is, $\nabla \mathbb{E}[X_\theta(\infty)]$. A typical example is the GI/G/1 queue where the distribution of the service times, or, of the interarrival times depends on a parameter, such as the mean. We may be interested in computing the sensitivity of the expected waiting time $\mathbb{E}[W_\theta]$ with respect to the parameter θ. Moreover, the computation of derivatives allows one to take an additional step and develop optimization procedures for the performance indicator of interest.

In general, however, closed–form expressions for the steady–state derivatives cannot be obtained, and one must resort to simulation methods. In addition, it is necessary to show consistency of such estimators, since the steady–state performance measure of the system under scrutiny is a limiting quantity and hence so is its gradient. Extra conditions that guarantee some type of uniform convergence, such as convexity are often imposed for that purpose.

A particularly neat situation occurs when the Markov chain $\{X_\theta(n)\}$ possesses a regenerative structure, that is, it restarts independent of the past whenever it hits a certain set α, called atom. If $\{X_\theta(n)\}$ is Harris ergodic with atom α and $\mathbb{E}[g(X_\theta(\infty))]$ is finite, then

$$\mathbb{E}[g(X_\theta(\infty))] \;=\; \frac{\mathbb{E}\left[\sum_{n=0}^{\tau_\theta - 1} g(X_\theta(n)) \,\middle|\, X(0) \in \alpha\right]}{\mathbb{E}[\tau_\theta]} \;,$$

where τ_θ denotes the first–entrance time of $X_\theta(n)$ into α. Unfortunately, the cycle time τ_θ typically depends on θ thus making differentiation of $\mathbb{E}[X_\theta(\infty)]$ a difficult task. However, the advent of *infinitesimal perturbation analysis* (IPA) allowed computing the sample gradient of $g(X_\theta(n))$, see [13, 6, 2], and, provided that the derivative process $\{\nabla X_\theta(n)\}$ regenerates at the same epochs as the original chain, it can be shown under some additional assumptions that

$$\nabla \mathbb{E}[g(X_\theta(\infty))] \;=\; \frac{\mathbb{E}\left[\sum_{n=0}^{\tau_\theta - 1} \nabla g(X_\theta(n)) \Big| \, X(0) \in \alpha\right]}{\mathbb{E}[\tau_\theta]} \;,$$

see [7, 8].

An alternative concept to sample–path differentiation is that of weak differentiation measures, as introduced by Pflug, see for example [18]. The concept of weak differentiation is to derive a (general state space) Markovian chain description of the system process such that the transition kernel, say $\mathbf{P}_\theta$, is differentiable as a function in θ. Then the derivative of the transition kernel can be represented as the difference between two transition kernels, say $\mathbf{P}_\theta^+$ and $\mathbf{P}_\theta^-$, i.e. $d\mathbf{P}_\theta/d\theta = \mathbf{P}_\theta^+ - \mathbf{P}_\theta^-$. Provided that the kernel is Harris ergodic with a regeneration set, say α, we estimate the gradient of the stationary costs as follows: We start the process in α and simulate the system process under $\mathbf{P}_\theta$ until it hits α. At a time, say k, we evaluate the *surrogate derivative* $\mathbf{D}(P_\theta, g, X_\theta(k))$ defined as follows: at $X_\theta(k)$ we split our sample path; we do this by performing this particular transition for one sub-path according to $\mathbf{P}_\theta^+$, whereas we perform it for the other sub-path according to $\mathbf{P}_\theta^-$; subsequently, we resume generating the transitions according to $\mathbf{P}_\theta$ until both (sub) paths hit α; the surrogate derivative $\mathbf{D}(P_\theta, g, \cdot)$ evaluates the difference between the g–performances evaluated for the variants of the processes. Summing the values of $\mathbf{D}(P_\theta, g, \cdot)$ over all τ_α states yields the desired gradient information, in formula:

$$\nabla \mathbb{E}[g(X_\theta(\infty))] \;=\; \frac{\mathbb{E}\left[\sum_{n=0}^{\tau_\theta - 1} \mathbf{D}(P_\theta, g, X_\theta(n)) \Big| \, X(0) \in \alpha\right]}{\mathbb{E}[\tau_\theta]} \;, \qquad (1)$$

see [18, 19]. While this approach does not suffer from the restriction that the derivative process has to regenerate at the same epochs as the Markov chain, weak differentiation is restricted to bounded performance functions, and extensions to more general classes of performance indicators are possible only in special cases, see, for example, [9]. For countable state–space, the derivative operator $\mathbf{D}(\cdot)$ is closely related to the deviation matrix, see [10].

In this paper we establish sufficient conditions for (1) to hold for unbounded mappings. To this end, we work within the framework of measure–valued differentiation (MVD), see [11, 12]. MVD extends the concept of weak differentiability, as introduced in [18, 19], so that performance measures out of a predefined class $\mathcal{D}$ can be handled, and thereby overcomes the restriction to bounded functions. As explained in [11], MVD implies no restriction to a particular estimation method and the estimator in (1) is only one possible translation of the measure–valued derivative of π into an estimator.

We will establish sufficient conditions for the stationary distribution π_θ to have a measure–valued derivative, called $\mathcal{D}$–derivative. Pflug shows in [19] that this holds true when $\mathcal{D}$ is the set of continuous bounded performance measures. As key result of this paper, we show that this statement extends to (more) general sets $\mathcal{D}$, where the main condition on the set $\mathcal{D}$ will be imposed by the ergodicity of the chain for mappings out of $\mathcal{D}$.

The paper is organized as follows. Section 2 introduces MVD. In Section 3 the main result of the paper is established, namely, that the stationary distribution of a $\mathcal{D}$–differentiable Markov chain is $\mathcal{D}$–differentiable. In Section 4, we provide a set of sufficient conditions, based on ergodicity concepts, for our main result to hold that can be verified in applications. In Section 5, we address gradient estimation and present some examples.

2 Background on MVD for Markov Chains

Let $(S, \mathcal{T})$ be a Polish measurable space. 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 $(S, \mathcal{T})$.

Definition 1 *The mapping $P : S \times \mathcal{T} \to [0, 1]$ is called a (homogeneous)* transition kernel *on $(S, \mathcal{T})$ if*

(a) $P(s; \cdot) \in \mathcal{M}(S, \mathcal{T})$ for all $s \in S$; and

(b) $P(\cdot; B)$ is $\mathcal{T}$ measurable for all $B \in \mathcal{T}$.

If, in condition (a), $\mathcal{M}(S, \mathcal{T})$ 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 $\mathcal{K}_1(S, \mathcal{T})$. Consider a family of Markov kernels $(P_\theta : \theta \in \Theta)$ on $(S, \mathcal{T})$, with $\Theta \subset \mathbb{R}$, and let $L^1(P_\theta; \Theta) \subset \mathbb{R}^S$ denote the set of measurable mappings $g : S \to \mathbb{R}$, such that $\int_S P_\theta(s; du)\, |g(u)|$ is finite for all $\theta \in \Theta$ and $s \in S$.

Definition 2 *Let Θ be an open neighborhood of θ_0. For $\theta \in \Theta$, let $P_\theta \in \mathcal{K}(S, \mathcal{T})$ and $\mathcal{D} \subset L^1(P_\theta; \Theta) \subset \mathbb{R}^S$. We call P_θ $\mathcal{D}$–*continuous at θ_0 *if, for any $g \in \mathcal{D}$ and any $s \in S$,*

$$\lim_{\Delta \to 0} \left| \int P_{\theta_0 + \Delta}(s; dz)\, g(z) \;-\; \int P_{\theta_0}(s; dz)\, g(z) \right| \;=\; 0 \, .$$

*Furthermore, we call P_θ $\mathcal{D}$–*Lipschitz continuous at θ_0 *if, for any $g \in \mathcal{D}$, a $K_g \in \mathcal{D}$ exists such that for any $\Delta \in \mathbb{R}$, with $\theta_0 + \Delta \in \Theta$:*

$$\left| \int P_{\theta_0 + \Delta}(\cdot; ds)\, g(s) \;-\; \int P_{\theta_0}(\cdot; ds)\, g(s) \right| \;\leq\; |\Delta|\, K_g \, .$$

We denote the set of bounded continuous mappings from S to $\mathbb{R}$ by $C_b(S)$. Let $P_\theta \in \mathcal{K}(S, \mathcal{T})$, for $\theta \in \Theta$. We call $\mathcal{D} \subset L^1(P_\theta; \Theta)$ a set of *test functions* for $(P_\theta : \theta \in \Theta)$ if for any $A \in \mathcal{T}$ its indicator function is in $\mathcal{D}$ and $C_b(S) \subset \mathcal{D}$.

Definition 3 *Let $\mathcal{D} \subset L^1(P_\theta; \Theta)$ be a set of test functions for $(P_\theta; \theta \in \Theta)$. We call $P_\theta \in \mathcal{K}(S, \mathcal{T})$* differentiable *at θ with respect to $\mathcal{D}$, or $\mathcal{D}$–*differentiable *for short, if for any $s \in S$ a $P'_\theta(s; \cdot) \in \mathcal{M}(S, \mathcal{T})$ exists, such that, for any $s \in S$ and any $g \in \mathcal{D}$*

$$\frac{d}{d\theta} \int_S P_\theta(s; du)\, g(u) \;=\; \int_S P'_\theta(s; du)\, g(u) \, . \tag{2}$$

If the left–hand side of equation (2) equals zero for all $g \in \mathcal{D}$, then we say that P'_θ is not significant.

If P_θ is $\mathcal{D}$–differentiable, then the measure–valued derivative $P'_\theta(\cdot; \cdot)$ is uniquely defined and again transition kernel on $(S, \mathcal{T})$, in formula: $P'_\theta(\cdot; \cdot) \in \mathcal{K}(S, \mathcal{T})$, see [11].

Example 1. A typical choice for $\mathcal{D}$ is the set of measurable bounded functions on S, denoted by $\mathcal{D}_b$ (it is easily checked that $\mathcal{D}_b$ is indeed a set of test functions).

In applications, to assume that the sample performance is bounded ($g \in \mathcal{D}_b$) is often too restrictive. A convenient set of performance functions is the set $\mathcal{D}^p$ of polynomially bounded performance functions defined by

$$\mathcal{D}^p \;=\; \left\{ g : S \to \mathbb{R} \;\middle|\; g(x) \leq \sum_{i=0}^{p} \kappa_i \, ||x||^i \;,\; \kappa_i \in \mathbb{R}, 0 \leq i \leq p \right\} \;,$$

for some $p \in \mathbb{N}$, where $||\cdot||$ denotes a norm on S (assuming that S is indeed equipped with a norm). Most cases of interest in applications fall within this setting. The set $\mathcal{D}^p$ is a set of test functions if and only if $\mathcal{D}^p \in L^1(P_\theta;\Theta)$, or, equivalently, if $\int_S P_\theta(s;du)||u||^p$ is finite for any $s \in S$ and $\theta \in \Theta$.

⁂

If P'_θ exists, then the fact that $P'_\theta(s;\cdot)$ fails to be a probability measure poses the problem of sampling from $P'_\theta(s;\cdot)$. For $s \in S$ fixed, we can represent $P'_\theta(s;\cdot)$ by its Hahn–Jordan decomposition as a difference between two probability measures. More precisely, for $s \in S$, let $([P'_\theta]^+(s;\cdot), [P'_\theta]^-(s;\cdot))$ denote the Hahn–Jordan decomposition of $P'_\theta(s;\cdot)$ and set

$$c_{P_\theta}(s) = [P'_\theta]^+(s;S) = [P'_\theta]^-(s;S) \tag{3}$$

and

$$P^+_\theta(s;\cdot) \;=\; \frac{[P'_\theta]^+(s;\cdot)}{c_{P_\theta}(s)}\,, \quad P^-_\theta(s;\cdot) \;=\; \frac{[P'_\theta]^-(s;\cdot)}{c_{P_\theta}(s)}\,,$$

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

$$\int_S P'_\theta(s;du)\,g(u) = c_{P_\theta}(s)\left(\int_S P^+_\theta(s;du)\,g(u) \;-\int_S P^-_\theta(s;du)\,g(u)\right)\,. \tag{4}$$

For the above line of argument we fixed s. For P^+_θ and P^-_θ to be Markov kernels, we have to consider P^+_θ and P^-_θ as functions in s and have to establish measurability of $P^+_\theta(\cdot;A)$ and $P^-_\theta(\cdot;A)$ for any $A \in \mathcal{T}$. This problem is equivalent to showing that $c_{P_\theta}(\cdot)$ in (3) is measurable as a mapping from S to $\mathbb{R}$. Unfortunately, only sufficient conditions are known, see [11]. For example, if S is finite, then measurability is guaranteed. In applications c_{P_θ} is calculated explicitly and its measurability is therefore established case by case. Specifically, in many examples that are of interest in applications, c_{P_θ} turns out to be a constant and measurability is thus guaranteed, see [11] for more details.

The Hahn–Jordan decomposition of P'_θ is not unique. To see this, choose $Q \in \mathcal{K}_1(S;\mathcal{T})$ so that $\int_S g(u)\,Q(s;du)$ is finite for any $g \in \mathcal{D}$ and $s \in S$. Set

$$\tilde{P}^+_\theta \;=\; \frac{1}{2}P^+_\theta + \frac{1}{2}Q\,, \quad \tilde{P}^-_\theta \;=\; \frac{1}{2}P^-_\theta + \frac{1}{2}Q\,.$$

Equation (4) implies for all $g \in \mathcal{D}$ and all $s \in S$ that

$$\frac{d}{d\theta}\int_S P_\theta(s;du)\,g(u) = 2c_{P_\theta}(s)\left(\int_S \tilde{P}^+_\theta(s;du)\,g(u) \;-\int_S \tilde{P}^-_\theta(s;du)\,g(u)\right)\,.$$

We now introduce the notion of $\mathcal{D}$–derivative, which extends the concept of a weak derivative.

Definition 4 *Let P_θ be $\mathcal{D}$-differentiable at θ. Any triple $(c_{P_\theta}(\cdot), P_\theta^+, P_\theta^-)$, with $P_\theta^\pm \in \mathcal{K}_1(S, \mathcal{T})$ and c_{P_θ} a measurable mapping from S to $\mathbb{R}$, that satisfies (4) is called a $\mathcal{D}$–*derivative *of P_θ. The kernel P_θ^+ is called the (normalized)* positive part *of ${P_\theta}'$ and P_θ^- is called the (normalized)* negative part *of ${P_\theta}'$; and $c_{P_\theta}(\cdot)$ is called the* normalizing factor.

We illustrate the concepts introduced above with a simple example.

Example 2. Let $P, Q \in \mathcal{K}_1(S, \mathcal{T})$ and set

$$P_\theta \;=\; \theta P \,+\, (1-\theta) Q \,, \quad \theta \in [0,1] \,.$$

Note that $P_\theta \in \mathcal{K}_1(S, \mathcal{T})$ for $\theta \in [0,1]$, and that $P_0 = Q$ and $P_1 = P$. Specifically, let $\mathcal{D}(P,Q) \;\triangleq\; L^1(P_\theta; \Theta)$ denote the set of measurable mappings $g : S \to \mathbb{R}$ such that both $\int_S P(s;du)\, g(u)$ and $\int_S Q(s;du)\, g(u)$ exist and are finite for any $s \in S$. For any $g \in \mathcal{D}(P,Q)$ and any $s \in S$, we now compute

$$\begin{aligned} \frac{d}{d\theta} \int_S P_\theta(s;du)\, g(u) &= \frac{d}{d\theta} \left(\theta \int_S P(s;du)\, g(u) \;+\; (1-\theta) \int_S Q(s;du)\, g(u) \right) \\ &= \int_S P(s;du)\, g(u) \;-\; \int_S Q(s;du)\, g(u) \,. \end{aligned}$$

Note that $\mathcal{D}(P,Q)$ is a set of test functions. Hence, P_θ is $\mathcal{D}(P,Q)$–differentiable with $\mathcal{D}(P,Q)$–derivative

$$\big(1 \,,\, P \,,\, Q \big) \,.$$

⁂

3 MVD of the Stationary Distribution

Let P_θ be ergodic and denote its unique invariant distribution by π_θ. Let $L^1(\pi_\theta)$ denote the set of measurable mappings $g : S \to \mathbb{R}$ such that $\int |g|\, d\pi_\theta$ is finite. We denote the ergodic projector of π_θ by Π_θ, that is, $\Pi_\theta : L^1(\pi_\theta) \to \mathbb{R}$ and, for any $g \in L^1(\pi_\theta)$, $\Pi_\theta g = \int g \, d\pi_\theta$. To simplify the notation, we set:

$$\mu\, g \;=\; \int g(s)\, \mu(ds) \,,$$

for $\mu \in \mathcal{M}(S, \mathcal{T})$, and

$$\Big(P_\theta \, g\Big)(s) \;=\; \int g(r)\, P_\theta(s; dr)\,,$$

for $P_\theta \in \mathcal{K}(S, \mathcal{T})$, provided that the expression exists. Note that $P_\theta g$ is a mapping from S to $\mathbb{R} \cup \{-\infty, \infty\}$. For any $\mu \in \mathcal{M}$ with $\mu(S) = 0$, we have the following rule of computation

$$\mu P_\theta^n \, g \;=\; \mu(P_\theta^n - \Pi_\theta)\, g\,, \quad n \in \mathbb{N}\,, \tag{5}$$

provided that the integrals exist. In what follows, we work locally and fix θ. With slight abuse of notation, we take Θ to be an open neighborhood of θ.

Theorem 1 *If*

(i) *P is $\mathcal{D}$–Lipschitz continuous at θ,*

(ii) *for any $\hat{\theta} \in \Theta$, if $h \in \mathcal{D}$ then $P_{\hat{\theta}} h \in \mathcal{D}$,*

(iii) *for any $h \in \mathcal{D}$ and any $\Delta \in \mathbb{R}$, with $\theta + \Delta \in \Theta$,*

$$\lim_{k \to \infty} (\pi_{\theta+\Delta} - \pi_\theta) P_\theta^k \, h \;=\; 0\,,$$

(iv) *for any $h \in \mathcal{D}$,*

$$(a) \qquad \sum_{n=0}^{\infty} |(P_\theta^n - \Pi_\theta)\, h| \;\in \mathcal{D}\,,$$

$$(b) \qquad \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h \;\in \mathcal{D}\,,$$

(v) *for any $h \in \mathcal{D}$ a finite number c_h exists such that*

$$\pi_{\hat{\theta}}\, |h| \;\le c_h\,, \quad \forall \hat{\theta} \in \Theta\,,$$

then π is $\mathcal{D}$–Lipschitz continuous at θ.

Moreover, if we assume, in addition to the above conditions, that P_θ is $\mathcal{D}$–differentiable at θ , then π_θ is $\mathcal{D}$–differentiable at θ with $\mathcal{D}$–derivative

$$\pi'_\theta \;=\; \pi_\theta \sum_{n=0}^{\infty} P'_\theta\, P_\theta^n\,,$$

or, equivalently,

$$\pi'_\theta \;=\; \pi_\theta P'_\theta \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\,.$$

Proof: Let I denote the unit operator, that is, $\pi_\theta I = \pi_\theta$. Simple algebraic calculation shows that

$$
\begin{aligned}
(\pi_{\theta+\Delta} - \pi_\theta)(I - P_{\theta+\Delta}) &= \pi_{\theta+\Delta} - \underbrace{\pi_{\theta+\Delta} P_{\theta+\Delta}}_{=\pi_{\theta+\Delta}} - \pi_\theta + \pi_\theta P_{\theta+\Delta} \\
&= \pi_\theta (P_{\theta+\Delta} - P_\theta) .
\end{aligned}
$$

Hence,

$$
(\pi_{\theta+\Delta} - \pi_\theta)(I - P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n = \pi_\theta (P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{k} P_\theta^n , \tag{6}
$$

with $P_\theta^0 = I$. By algebraic calculation,

$$
\begin{aligned}
&(\pi_{\theta+\Delta} - \pi_\theta)(I - P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n \\
&= (\pi_{\theta+\Delta} - \pi_\theta) \left\{ (I - P_\theta) \sum_{n=0}^{k} P_\theta^n + (P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n \right\} \\
&= (\pi_{\theta+\Delta} - \pi_\theta) \left\{ \sum_{n=0}^{k} P_\theta^n - \sum_{n=1}^{k+1} P_\theta^n + (P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n \right\} \\
&= (\pi_{\theta+\Delta} - \pi_\theta) \left\{ I - P_\theta^{k+1} + (P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n \right\} \\
&= \pi_{\theta+\Delta} - \pi_\theta - (\pi_{\theta+\Delta} - \pi_\theta) P_\theta^{k+1} + (\pi_{\theta+\Delta} - \pi_\theta)(P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n .
\end{aligned}
$$

Inserting the righthand side of the above equation into (6) yields:

$$
\begin{aligned}
\pi_\theta (P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{k} P_\theta^n = \pi_{\theta+\Delta} - \pi_\theta - (\pi_{\theta+\Delta} - \pi_\theta) P_\theta^{k+1} \\
+ (\pi_{\theta+\Delta} - \pi_\theta)(P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n .
\end{aligned} \tag{7}
$$

We now study the limit of the above expression as k tends off to ∞. Firstly, under assumption (iii), we have

$$
\lim_{k \to \infty} (\pi_{\theta+\Delta} - \pi_\theta) P_\theta^{k+1} h = 0 . \tag{8}
$$

As a second step, we show

$$
\lim_{k \to \infty} \pi_\theta (P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{k} P_\theta^n h = \pi_\theta (P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta) h , \tag{9}
$$

for $h \in \mathcal{D}$. By assumptions (ii) together with (v),

$$\pi_\theta(P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{k} P_\theta^n \, h$$

exists and is finite for any $k \geq 0$ and any $h \in \mathcal{D}$. Elaborating on (5), this implies that

$$\pi_\theta(P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{k} (P_\theta^n - \Pi_\theta) \, h$$

exists and is finite for any $k \geq 0$ and any $h \in \mathcal{D}$, where we use that fact that $\pi_\theta(P_{\theta+\Delta} - P_\theta)$ is a signed measure with total mass 0. By assumption (iv)(a), for any $h \in \mathcal{D}$,

$$\left| \sum_{n=0}^{k} (P_\theta^n - \Pi_\theta) \, h \right| \leq \sum_{n=0}^{k} |(P_\theta^n - \Pi_\theta) \, h| \ \in \mathcal{D} \, .$$

Moreover, by assumptions (ii) together with (v),

$$\pi_\theta(P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{\infty} |(P_\theta^n - \Pi_\theta) \, h|$$

exists and is finite, for any $h \in \mathcal{D}$. Hence, (9) follows from the dominated convergence theorem.

Finally, following the line of argument in the above second step, we show that

$$\begin{aligned} \lim_{k \to \infty} (\pi_{\theta+\Delta} - \pi_\theta)(P_\theta \ - \ P_{\theta+\Delta}) \sum_{n=0}^{k} P_\theta^n & \\ = \ (\pi_{\theta+\Delta} - \pi_\theta)(P_\theta \ - \ P_{\theta+\Delta}) \sum_{n=0}^{\infty} P_\theta^n \, . & \qquad (10) \end{aligned}$$

Taking the limit in (7) as k tends to ∞ we obtain from (8), (9) and (10), for $h \in \mathcal{D}$:

$$\begin{aligned} (\pi_{\theta+\Delta} - \pi_\theta) \, h \ &= \pi_\theta(P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta) \, h \\ &\quad - (\pi_{\theta+\Delta} - \pi_\theta)(P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta) \, h \, . \qquad (11) \end{aligned}$$

For $h \in \mathcal{D}$, condition (iv) (b) implies

$$\sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta) \, h \ =: \ \hat{h} \ \in \mathcal{D} \, ,$$

and $\mathcal{D}$–Lipschitz continuity of P_θ at θ implies

$$
\begin{aligned}
|\pi_\theta (P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h| &\leq \ \Delta \pi_\theta K_{\hat{h}} \\
&\leq \ \Delta c_{K_{\hat{h}}} \ < \infty \, ,
\end{aligned}
$$

where finiteness of $\pi_\theta K_{\hat{h}}$ is guaranteed by (v). From the same line of argument, we obtain that

$$
\begin{aligned}
&|(\pi_{\theta+\Delta} - \pi_\theta)(P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h| \\
&\qquad \leq (\pi_{\theta+\Delta} + \pi_\theta) |(P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h| \\
&\qquad \leq \Delta (\pi_{\theta+\Delta} + \pi_\theta) K_{\hat{h}} \\
&\qquad \leq \Delta 2 c_{K_{\hat{h}}} \ ,
\end{aligned}
$$

for any $h \in \mathcal{D}$. Because $\mathcal{D}$ is a set of test functions, the constant functions $\Delta c_{K_{\hat{h}}}, \Delta 2 c_{K_{\hat{h}}}$ as well as their sum lie in $\mathcal{D}$. From (11) it thus follows that π_θ is $\mathcal{D}$–Lipschitz continuous at θ. In particular, the Lipschitz factor is the constant function $3 c_{K_{\hat{h}}}$.

Starting point for the second part of the theorem is equation (11). From $\mathcal{D}$–Lipschitz continuity at θ of both π_θ and P_θ it follows that

$$
\lim_{\Delta \to 0} \frac{1}{\Delta} (\pi_\theta - \pi_{\theta+\Delta})(P_\theta - P_{\theta+\Delta}) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h \ = \ 0 \, ,
$$

for any $h \in \mathcal{D}$. Moreover, because P_θ is $\mathcal{D}$–Lipschitz at θ, assumption (v) yields

$$
\begin{aligned}
&\lim_{\Delta \to 0} \frac{1}{\Delta} \pi_\theta (P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h \\
&= \pi_\theta \left(\lim_{\Delta \to 0} \frac{1}{\Delta} (P_{\theta+\Delta} - P_\theta) \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h \right)
\end{aligned}
$$

and, by $\mathcal{D}$–differentiability of P_θ, this limit equals

$$
= \pi_\theta P_\theta' \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h \, .
$$

For $h \in \mathcal{D}$, we therefore obtain from (11)

$$
\pi_\theta P_\theta' \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta)\, h = \lim_{\Delta \to 0} \frac{1}{\Delta} (\pi_\theta \ - \ \pi_{\theta+\Delta})\, h \, .
$$

Using the fact the $P'_\theta(s;\cdot)$ is a signed measure with $P'_\theta(s;S)=0$, for any $s\in S$, it readily follows from (5) that

$$\pi_\theta \sum_{n=0}^{\infty} P'_\theta P^n_\theta \, h \;=\; \pi_\theta P'_\theta \sum_{n=0}^{\infty} (P^n_\theta - \Pi_\theta)\, h \;,$$

which concludes the proof. $\square$

Remark: If $h\in\mathcal{D}$ implies that $|h|\in\mathcal{D}$, then condition (iv) (a) already implies condition (iv) (b).

4 Ergodicity Framework

In this section, we provide sufficient conditions for conditions (ii) to (iv) to hold. Let $\mathbf{X}(\theta)=\{X_\theta(n)\}=\{X_\theta(s,n)\}$, for $\theta\in\Theta$, be the Markov Chain with initial state s and transition kernel P_θ, and set, for any $B\in\mathcal{T}$,

$$P^n_\theta(s,B) \;\triangleq\; P_\theta\,(s,n,B) = P(X_\theta(s,n)\in B)\;.$$

The joint state-space of $\mathbf{X}(\theta)$, $\theta\in\Theta$, is denoted by S. However, for any specific θ the chain $\mathbf{X}(\theta)$ may not be irreducible on S but only on a subset of S. For the following ergodicity analysis, we will require that the state-space is indeed irreducible and we denote by $S_\theta\subset S$ the class of states such that $\mathbf{X}(\theta)$ becomes irreducible as a Markov chain on S_θ. Furthermore, we denote by $\mathcal{T}_\theta$ the intersection of $\mathcal{T}$ and S_θ. Consequently, $(S_\theta,\mathcal{T}_\theta)$ is a measurable space for any $\theta\in\Theta$.

In Section 4.1, we discuss the general situation. Section 4.2 provides an alternative representation of the $\mathcal{D}$–derivative of π_θ for the situation where $\mathbf{X}(\theta)$ possesses an atom.

4.1 General Chains

The main technical conditions needed for the analysis in this section are introduced subsequently. We will use the following "Lyapunov function" condition:

(C1) There exists a function $g(s)\geq 0$, $s\in S_\theta$, such that any $\theta\in\Theta$

$$\mathbb{E}\big[g(X_\theta(s,m_\theta))\big] - g(s) \;\leq\; -\varepsilon \,+\, c\, I_{V_\theta}(s)\;, \tag{12}$$

for some $m_\theta \geq 1, \varepsilon > 0$ and $c < \infty$, where for some $d < \infty$

$$V_\theta \;=\; \{\, s \in S_\theta \;:\; g(s) \leq d \,\}$$

and $I_{V_\theta}(s) = 1$ if $s \in V_\theta$ and otherwise zero. Furthermore, we need the following Harris-type condition for the set V_θ.

(C2) For any $\theta \in \Theta$ there exist $n_\theta \geq 0$, $\phi_\theta(\cdot)$ a probability measure on $(S_\theta, \mathcal{T}_\theta)$, and $p_\theta \in (0,1)$ such that

$$\inf_{x \in V_\theta} P\big(X_\theta(x, n_\theta) \in B\big) \;\geq\; p_\theta\, \phi_\theta(B)\,, \tag{13}$$

for all $B \in \mathcal{B}$.

Under condition (**C1**), let

$$\xi_\theta(s) \;\triangleq\; g(X_\theta(s,1)) - g(s)\,, \quad s \in S\,,$$

and introduce the following condition:

(C3) The r.v. $\xi_\theta(s)$ is uniformly integrable (in s and θ) and there exist $\lambda > 0$ such that $\xi_\theta(s)\, e^{\lambda \xi_\theta(s)}$ are uniformly integrable (in s and θ).

Recall that uniform integrability of $\xi_\theta(s)$ in s and θ is defined as

$$\lim_{c \to \infty} \sup_{s,\theta} \int_{|t|>c} P(\xi_\theta(s) \in dt) \;=\; 0\,,$$

and similarly the uniform integrability of $\xi_\theta(s) e^{\lambda \xi_\theta(s)}$ requires that

$$\lim_{c \to \infty} \sup_{s,\theta} \int_{|t|>c} P(\xi_\theta(s) e^{\lambda \xi_\theta(s)} \in dt) \;=\; 0\,.$$

In order to establish conditions (ii) to (iv) in Theorem 1, we will work with normed ergodicity. Normed ergodicity dates back to the early eighties, see [4] for a first reference. It was originally used in analysis of Blackwell optimality; see [4], and [16] for a recent publication on this topic. Since then, it has been used in various forms under different names in many subsequent papers. In [14] it was shown for a countable Markov chain which may have one or several classes of essential states (a so-called multichained Markov chain), that normed ergodicity is equivalent to geometrical recurrence (for a similar result in Markov decision chains see [5]). Inspired by this result for a countable Markov chain a similar result was proved for a Harris chain in [17]. In this paper we use the recent results of [1].

Let $\mathcal{V}_v$ denote the Banach space of real-valued functions f on $\mathcal{S}$ with the finite v-norm

$$\|f\|_v = \sup_{s \in \mathcal{S}} \frac{|f(s)|}{|v(s)|}$$

and the associated operator norm for a linear operator, say $T : \mathcal{V}_v \longrightarrow \mathcal{V}_v$ is defined by

$$\|T\|_v = \sup_{\|f\|_v \leq 1} \|Tf\|_v \, .$$

For μ a (signed) measure the associated norm is

$$\|\mu\|_v = \sup_{\|f\|_v \leq 1} |\mu f| \, .$$

For our analysis, we choose v to be the following mapping:

$$v(s) \triangleq e^{\lambda g(s)} \, , \quad s \in \mathcal{S} \, , \tag{14}$$

for some positive λ, where g is defined in (**C1**).

Lemma 1 *The condition* (**C3**) *implies that for* λ *small enough*

$$\sup_{\theta \in \Theta} ||P_\theta||_v < \infty \, .$$

Proof: With

$$\xi_\theta(s) = g(X_\theta(s,1)) - g(s) \, , \quad s \in S \, ,$$

and

$$v(s) = e^{\lambda g(s)} \, , \quad s \in S \, ,$$

we find that

$$\begin{aligned} \|P_\theta\|_v &= \sup_s \frac{(P_\theta e^{\lambda g})(s)}{e^{\lambda g(s)}} \\ &= \sup_s \mathbb{E}\left[e^{\lambda(g(X_\theta(s,1))-g(s))}\right] \\ &= \sup_s \mathbb{E}\left[e^{\lambda \xi_\theta(s)}\right] \, . \end{aligned} \tag{15}$$

By (**C3**), $\xi_\theta(s)e^{\lambda\xi_\theta(s)}$ is uniformly integrable which implies

$$\sup_{s,\theta} \mathbb{E}\left[\left|\xi_\theta(s)e^{\lambda\xi_\theta(s)}\right|\right] < \infty\,.$$

For $s \in S$, we now write

$$\mathbb{E}\left[e^{\lambda\xi_\theta(s)}\right] = \mathbb{E}\left[e^{\lambda\xi_\theta(s)} I(\xi_\theta(s) \leq 1)\right] + \mathbb{E}\left[e^{\lambda\xi_\theta(s)} I(\xi_\theta(s) > 1)\right]\,.$$

The first term on the right hand side is bounded by e^λ and the second term is bounded by

$$\sup_{s,\theta} \mathbb{E}\left[\left|\xi_\theta(s)e^{\lambda\xi_\theta(s)}\right|\right] < \infty\,.$$

Inserting these bounds into (15) yields

$$\sup_\theta \|P_\theta\|_v \leq e^\lambda + \sup_{s,\theta} \mathbb{E}e^{\lambda\xi_\theta(s)} < \infty\,.$$

□

The following theorem follows from Theorems 3 and 5 of [1].

Theorem 2 *Conditions* (**C1**), (**C2**) *and* (**C3**) *imply that, for any $\theta \in \Theta$, there exist $c_\theta < \infty$ and $0 < \rho_\theta < 1$ such that for λ small enough*

$$\|P_\theta^n - \Pi_\theta\|_v \leq c_\theta \rho_\theta^n, \tag{16}$$

Note that this theorem implies that for any θ

$$\|\Pi_\theta\|_v < \infty.$$

Let

$$\mathcal{D}_v = \left\{g : S \to \mathbb{R} \,\middle|\, \exists r \in \mathbb{R} : |g(s)| \leq r \cdot v(s),\, s \in S\right\}.$$

In words, $\mathcal{D}_v$ is the set of mappings g from S to $\mathbb{R}$ that are bounded by $r \cdot v$ (for some finite number r). It is easily seen that Lemma 1 implies that any $g \in \mathcal{D}_v$ is integrable with respect to any P_θ, for $\theta \in \Theta$, or, more formally:

$$\mathcal{D}_v \subset L^1(P_\theta, \Theta)\,.$$

Theorem 3 *Let Θ denote an open neighborhood of θ. If*

(i) P_θ is $\mathcal{D}_v$–Lipschitz continuous in θ,

(ii) *conditions* (**C1**) *to* (**C3**) *are satisfied*

then π_θ *is* $\mathcal{D}_v$*–Lipschitz continuous.*

Moreover, if we assume, in addition to the above conditions, that P_θ *is* $\mathcal{D}_v$*–differentiable, then* π_θ *is* $\mathcal{D}_v$*–differentiable with* $\mathcal{D}_v$*–derivative*

$$\pi'_\theta \;=\; \pi_\theta \sum_{n=0}^{\infty} P'_\theta \, P^n_\theta \;,$$

or, equivalently,

$$\pi'_\theta \;=\; \pi_\theta P'_\theta \sum_{n=0}^{\infty} (P^n_\theta - \Pi_\theta) \;.$$

Proof: We will show that conditions (ii) to (iv) of Theorem 1 hold. Condition (ii) is a straightforward consequence of Lemma 1. Note that

$$\|P^n_\theta - \Pi_\theta\|_v \;\leq\; c_\theta \rho^n_\theta$$

implies that, for any $h \in \mathcal{D}_v$,

$$|(P^n_\theta - \Pi_\theta)h| \;\leq\; \bar{c}_\theta \rho^n_\theta \, v,$$

with

$$\bar{c}_\theta \triangleq c_\theta \, \|h\|_v \,.$$

Hence,

$$\sum_{n=0}^{\infty} |(P^n_\theta - \Pi_\theta)h| \;\leq\; \frac{\bar{c}_\theta}{1-\rho_\theta} \, v$$

and because $\frac{\bar{c}_\theta}{1-\rho_\theta} \, v \in \mathcal{D}_v$, this already implies that

$$\sum_{n=0}^{\infty} |(P^n_\theta - \Pi_\theta)h| \;\in \mathcal{D}_v \;.$$

Thus, (iv) (b) holds. Repeating the above argument without taking absolute values, or using the remark after Theorem 1, shows that (iv) (a) holds as well. With the relation (5) we have that

$$(\Pi_{\theta+\Delta} - \Pi_\theta) P^k_\theta h = (\Pi_{\theta+\Delta} - \Pi_\theta)(P^k_\theta - \Pi_\theta) h \;.$$

The condition (iii) then follows from (16) of Theorem 2.□

4.2 Chains with an Atom

Throughout this section, we assume that conditions (**C1**) – (**C3**) are satisfied. The setup is as in the previous section with the additional assumption that the chain possesses an atom, say α. The expression of the stationary distribution for a regenerative process is well-known (see [20])

$$\pi_\theta\, h \;=\; \frac{1}{\mathbb{E}[\tau_\theta(s)]}\mathbb{E}\left[\sum_{m=0}^{\tau_\theta(s)} h(X_\theta(s,m))\right]\,, \quad s\in\alpha\,,$$

where $\tau_\theta(s)$ is the recurrence time to the atom α. With the notation $(Q_\theta(s))f = \int_{\mathcal{S}\setminus\alpha} P_\theta(s,1,dy)f(y), s\in\mathcal{S}$ this can also be written as

$$\mathbb{E}\left[\sum_{m=0}^{\tau_\theta(s)} h(X_\theta(s,m))\right] \;=\; \sum_{m=0}^{\infty}(Q_\theta(s))^m\, h\,.$$

Hence,

$$\mathbb{E}[\tau_\theta(s)] \;=\; \sum_{m=0}^{\infty}(Q_\theta(s))^m\, e$$

and thus, provided that $s\in\alpha$,

$$\pi_\theta h = \frac{\sum_{m=0}^{\infty}(Q_\theta(s))^m\, h}{\sum_{m=0}^{\infty}(Q_\theta(s))^m\, e}\,,$$

and

$$\sum_{m=0}^{\infty}(Q_\theta(s))^m\,(I-\Pi_\theta)h(\alpha) = 0. \tag{17}$$

Using the taboo presentation of the deviation operator derived in [15] together with relation (5) the following representation of the derivative can be given

$$\pi_\theta' h = \pi_\theta(P_\theta')\sum_{n=0}^{\infty}(Q_\theta(s))^n\,(I-\Pi_\theta)h.$$

5 Gradient Estimation

In this section, we apply our result to gradient estimation. Section 5.1 establishes an interpretation of the expression for the measure-valued derivative of π in Theorem 1 in terms of stochastic processes. Illustrating examples are provided in Section 5.2.

5.1 The Process View

5.1.1 General Chains

Let P_θ be $\mathcal{D}$–differentiable with $\mathcal{D}$–derivative $(c_{P_\theta}, P_\theta^+, P_\theta^-)$ (see Definition 4), and let $X_\theta^\pm(s,n)$, with initial state s, evolve according to the kernel

$$P_\theta^\pm \, P_\theta^n \, ,$$

or, equivalently, for $n > 0$, let the transition from $X_\theta^\pm(s,n)$ to $X_\theta^\pm(s,n+1)$ be governed by P_θ, whereas the transition from $X_\theta^+(s,0)$ to $X_\theta^+(s,1)$ is governed by P_θ^+ and that from $X_\theta^-(s,0)$ to $X_\theta^-(s,1)$ by P_θ^-, respectively. Hence, while $n > 0$, both $X_\theta^+(s,n)$ and $X_\theta^-(s,n)$ are driven by the same Markov kernel, and, without loss of generality, we assume that the processes $\{X_\theta^\pm(s,n)\}$ are constructed using common random numbers. For $g \in \mathcal{D}$, we set

$$\mathbf{D}(P_\theta, g; s) \,=\, \mathbb{E}\left[c_{P_\theta}(s) \sum_{n=0}^{\infty} \big(g(X_\theta^+(s,n)) \,-\, g(X_\theta^-(s,n)) \big) \right] ,$$

or, equivalently,

$$\sum_{n=0}^{\infty} P_\theta' P_\theta^n g \,=\, \mathbf{D}(P_\theta, g) \, . \tag{18}$$

Let $\tau_\theta^\pm(s)$ denote the coupling time of $X_\theta^+(m,s)$ and $X_\theta^-(m,s)$, that is,

$$\tau_\theta^\pm(s) \,=\, \inf \big\{ m \in \mathbb{N} \,:\, X_\theta^+(s,m) = X_\theta^-(s,m) \big\} \, , \tag{19}$$

and $\tau_\theta^\pm(s) = \infty$ if the set on the righthand side of the above equation is empty. With this definition, for $g \in \mathcal{D}$,

$$\mathbf{D}(P_\theta, g; s) \,=\, \mathbb{E}\left[c_{P_\theta}(s) \sum_{n=0}^{\tau_\theta^\pm(s)} \big(g(X_\theta^+(s,n)) \,-\, g(X_\theta^-(s,n)) \big) \right] .$$

Let X_θ be distributed according to π_θ. Under the conditions in Theorem 1 it holds

$$\begin{aligned} \frac{d}{d\theta}\mathbb{E}[g(X_\theta)] &= \mathbb{E}\Big[\mathbf{D}(P_\theta, g; X_\theta) \Big] \\ &= \mathbb{E}\left[\mathbb{E}\left[c_{P_\theta}(X_\theta) \sum_{n=0}^{\tau_\theta^\pm(X_\theta)} \big(g(X_\theta^+(X_\theta,n)) \,-\, g(X_\theta^-(X_\theta,n)) \big) \middle| \, X_\theta \right] \right] , \end{aligned} \tag{20}$$

for any $g \in \mathcal{D}_v$.

Even if the stopping time $\tau_\theta^\pm$ is a.s. finite, it may be prohibitively large for the estimator in (20) to be of practical use. For this reason we introduce a truncated version of $\mathbf{D}(P_\theta, g; s)$: for $N > 0$, set

$$\mathbf{D}_N(P_\theta, g; s) \;=\; \sum_{n=0}^{N} P'_\theta P^n_\theta g \;=\; \mathbb{E}\left[c_{P_\theta}(s) \sum_{n=0}^{N} \big(g(X^+_\theta(s,n)) \;-\; g(X^-_\theta(s,n)) \big) \right] .$$

Conditions $(\mathbf{C1}) - (\mathbf{C3})$ imply that $\mathbf{D}_N(P_\theta, g; s)$ converges geometrically fast towards $\mathbf{D}(P_\theta, g; s)$. Hence, for N sufficiently large,

$$\begin{aligned} \frac{d}{d\theta}\mathbb{E}[g(X_\theta)] &\approx \mathbb{E}\Big[\mathbf{D}_N(P_\theta, g; X_\theta) \Big] \\ &= \mathbb{E}\left[\mathbb{E}\left[c_{P_\theta}(X_\theta) \sum_{n=0}^{N} \big(g(X^+_\theta(X_\theta,n)) \;-\; g(X^-_\theta(X_\theta,n)) \big) \middle| X_\theta \right] \right] , \end{aligned}$$

for any $g \in \mathcal{D}_v$.

Remark: Elaborating on (5), it holds that

$$P'_\theta \sum_{n=0}^{\infty} (P^n_\theta - \Pi_\theta) g \;=\; \mathbf{D}(P_\theta, g) \, , \quad g \in \mathcal{D}_v \, .$$

The operator $D := \sum_{n=0}^{\infty} (P^n_\theta - \Pi_\theta)$ is called *deviation operator* in the theory of Markov chains, and, elaborating on the deviation operator, the above equation reads

$$P'_\theta \, D \, g \;=\; \mathbf{D}(P_\theta, g) \, , \quad g \in \mathcal{D}_v \, ,$$

which extends a result on the relation between derivatives and the deviation operator in [10] to chains on a general state–space.

5.1.2 Chains with an Atom

We now turn to the situation where $\mathbf{X}(\theta)$ possesses an atom, denoted by $\alpha \in \mathcal{T}$. Let $X_\theta(\alpha, n)$ denote the Markov chain started in α and denote the first entrance time of the chain into α by $\tau_{\theta,\alpha}$. The expression of the stationary distribution for a chain with atom is well-known (see [20]) and we obtain from (20) for any $g \in \mathcal{D}_v$:

$$\frac{d}{d\theta}\mathbb{E}[g(X_\theta)] \;=\; \frac{1}{\mathbb{E}[\tau_{\theta,\alpha}]} \, \mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} \mathbf{D}(P_\theta, g; X_\theta(\alpha, n)) \right] .$$

The above estimator can be rewritten as follows. Denote by $\tau^{\pm}_{\theta,\alpha}(s)$ the first time that $X^+_\theta(s,m)$ and $X^-_\theta(s,m)$ simultaneously hit α:

$$\tau^{\pm}_{\theta,\alpha}(s) \,=\, \inf\left\{m \in \mathbb{N} \,:\, X^+_\theta(s,m) \in \alpha\,,\; X^-_\theta(s,m) \in \alpha\right\}.$$

Then, for any $g \in \mathcal{D}_v$:

$$\mathbf{D}(P_\theta, g; s) \,=\, \mathbb{E}\left[c_{P_\theta}(s) \sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(s)} \Big(g(X^+_\theta(s,m)) \,-\, g(X^-_\theta(s,m))\Big)\right]$$

and the overall estimator becomes

$$\frac{d}{d\theta}\mathbb{E}[g(X_\theta)] \,=\, \frac{1}{\mathbb{E}[\tau_{\theta,\alpha}]}\,\mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} \mathbf{D}(P_\theta, g; X_\theta(\alpha,n))\right]. \tag{21}$$

From a simulation point of view, formula (21) poses the problem of estimating the inner expected value. As we will show below the inner expected value can be avoided. To see this, we simplify the notation and set

$$H(s,m) \,=\, c_{P_\theta}(s)\big(g(X^+_\theta(s,m)) \,-\, g(X^-_\theta(s,m))\big). \tag{22}$$

With this notation the right hand side of (21) reads

$$\frac{1}{\mathbb{E}[\tau_{\theta,\alpha}]}\,\mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} \mathbb{E}\left[\sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(X_\theta(\alpha,n))} H(X_\theta(\alpha,n),m)\,\middle|\, X_\theta(\alpha,n)\right]\right].$$

Let $\mathcal{F}(\tau_{\theta,\alpha})$ denote the σ-field generated by $X_\theta(\alpha,0),\ldots,X_\theta(\alpha,\tau_{\theta,\alpha})$. By calculation,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} \mathbb{E}\left[\sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(X_\theta(\alpha,n))} H(X_\theta(\alpha,n),m)\,\middle|\, X_\theta(\alpha,n)\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} \mathbb{E}\left[\sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(X_\theta(\alpha,n))} H(X_\theta(\alpha,n),m)\,\middle|\, X_\theta(\alpha,n)\right]\,\middle|\,\mathcal{F}(\tau_{\theta,\alpha})\right]\right]
\end{aligned}$$

because $X_\theta(\alpha,n)$ is measurable with respect to $\mathcal{F}(\tau_{\theta,\alpha})$

$$\begin{aligned}
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} \sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(X_\theta(\alpha,n))} H(X_\theta(\alpha,n),m)\,\middle|\,\mathcal{F}(\tau_{\theta,\alpha})\right]\right] \\
&= \mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} \sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(X_\theta(\alpha,n))} H(X_\theta(\alpha,n),m)\right]
\end{aligned}$$

and inserting (22) yields

$$= \mathbb{E}\Bigg[\sum_{n=0}^{\tau_{\theta,\alpha}} c_{P_\theta}(X_\theta(\alpha,n)) \times$$
$$\sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(X_\theta(\alpha,n))} \Big(g(X^+_\theta(X_\theta(\alpha,n),m)) \,-\, g(X^-_\theta(X_\theta(\alpha,n),m)) \Big) \Bigg] . \quad (23)$$

Hence, (21) becomes

$$\frac{d}{d\theta}\mathbb{E}[g(X_\theta)] \,=$$
$$\frac{1}{\mathbb{E}[\tau_{\theta,\alpha}]}\,\mathbb{E}\Bigg[\sum_{n=0}^{\tau_{\theta,\alpha}} c_{P_\theta}(X_\theta(\alpha,n)) \sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(X_\theta(\alpha,n))} \Big(g(X^+_\theta(X_\theta(\alpha,n),m)) \,-\, g(X^-_\theta(X_\theta(\alpha,n),m)) \Big) \Bigg] .$$

We conclude this section by presenting an alternative representation of the above estimator. We define inhomogeneous Markov chains as follows. Let $X^+_\theta(n;m)$ be such that $X^+_\theta(n;m)$ starts in α and the first n transitions are performed according to Q_θ and the transition from $X^+_\theta(n;n)$ to $X^+_\theta(n;n+1)$ is generated according to P^+_θ and after $n+1$ the transition kernel is P_θ. Define $X^-_\theta(n;m)$ in the same vein. Note that on the event $\{\tau_\theta > n\}$

$$X^+_\theta(n;m) \,=\, X^-_\theta(n;m) \,=\, X_\theta(\alpha,m)\,, \quad m \le n\,. \qquad (24)$$

Denote by $\tau^{\pm}_{\theta,\alpha}(n)$ the first time that $X^+_\theta(n;m)$ and $X^-_\theta(n;m)$ simultaneously hit α:

$$\tau^{\pm}_{\theta,\alpha}(n) \,=\, \inf\big\{ m \ge n \,:\, X^+_\theta(n;m) \in \alpha\,,\, X^-_\theta(n;m) \in \alpha \big\}\,. \qquad (25)$$

The expression in (23) then is equal to

$$\mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} c_{P_\theta}(X_\theta(\alpha,n)) \sum_{m=0}^{\tau^{\pm}_{\theta,\alpha}(n)} \Big(g(X^+_\theta(n;m)) \,-\, g(X^-_\theta(n;m)) \Big) \right]$$
$$= \mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} c_{P_\theta}(X_\theta(\alpha,n)) \sum_{m=n+1}^{\tau^{\pm}_{\theta,\alpha}(n)} \Big(g(X^+_\theta(n;m)) \,-\, g(X^-_\theta(n;m)) \Big) \right] ,$$

where the last equality follows from (24). Hence,

$$\frac{d}{d\theta}\mathbb{E}[g(X_\theta)] \qquad (26)$$
$$= \frac{1}{\mathbb{E}[\tau_{\theta,\alpha}]}\,\mathbb{E}\left[\sum_{n=0}^{\tau_{\theta,\alpha}} c_{P_\theta}(X_\theta(\alpha,n)) \sum_{m=n+1}^{\tau^{\pm}_{\theta,\alpha}(n)} \Big(g(X^+_\theta(n;m)) \,-\, g(X^-_\theta(n;m)) \Big) \right] ,$$

with $g \in \mathcal{D}$.

5.2 Examples

Consider a single server queue with i.i.d. exponentially distributed service times with rate μ. Service times and interarrival times are independent and let the interarrival times be a sequence of i.i.d. random variables following a Cox distribution with rates η_j, $j = 1, 2$, and parameter θ, that is, the interarrival times consist with probability $1 - \theta$ of a single exponentially distributed stage with rate η_1, and a second stage with rate η_2 follows with probability θ. Let h_μ denote the density of the Exponential distribution with rate μ and write $\mathcal{E}_\mu$ for the distribution. Denoting the density of the sum of two independent exponentially distributed random variables with rate η_1 and η_2 by $h_{\eta_1,\eta_2}(x)$ and the corresponding distribution function by $\mathcal{E}_{(\eta_1,\eta_2)}$, the density of the interarrival times is given by

$$h_\theta(x) \,=\, (1-\theta)h_{\eta_1}(x) \,+\, \theta h_{\eta_1,\eta_2}(x)\,, \quad x \geq 0\,.$$

The parameter of interest is θ. Observe that, for $\theta = 1$, $h_\theta(x) = h_{\eta_1,\eta_2}(x)$ and the interarrival times follow a phase-type distribution, whereas, for $\theta = 0$, $h_\theta(x) = h_{\eta_1}(x)$ and the interarrival times follow an Exponential distribution.

5.2.1 Discrete State–Space

Let $X_\theta(n) = (X_\theta(1,n), X_\theta(2,n))$ be the state of the system, with $X_\theta(1,n) \in \mathbb{N}$ the total number of customers in the system, and $X_\theta(2,n) \in \{1,2\}$ the stage of the interarrival time. Let

$$P_\theta((k,i);(k',i')) \,=\, P(\, X_\theta(m+1) = (k',i') \,|\, X_\theta(m) = (k,i)\,)\,,$$

for $(k,i),(k',i') \in \mathbb{N} \times \{1,2\}$. Then, the probability that an arrival occurs is

$$P_\theta((k,1);(k+1,1)) \,=\, (1-\theta)\frac{\eta_1}{\eta_1 + \mu 1_{k>0}}\,,$$

$$P_\theta((k,2);(k+1,1)) \,=\, \frac{\eta_2}{\eta_2 + \mu 1_{k>0}}\,,$$

the probability that the state of the interarrival time jumps from stage 1 to 2 is

$$P_\theta((k,1);(k,2)) \,=\, \theta\frac{\eta_1}{\eta_1 + \mu 1_{k>0}}\,,$$

and the probability that a departure occurs is

$$P_\theta((k,1);(k-1,1)) \;=\; \frac{\mu}{\eta_1+\mu}\,,$$

$$P_\theta((k,2);(k-1,2)) \;=\; \frac{\mu}{\eta_2+\mu}\,.$$

Set $P = P_1$ and $Q = P_0$, then

$$P_\theta \;=\; \theta P \,+\, (1-\theta) Q\,. \tag{27}$$

For any θ the process is a discrete-time Markov chain which is irreducible, and hence any state is an atom. In particular, $(0,1)$ is an atom for each of the processes. Let us verify conditions (**C1**) to (**C3**). We assume that for any θ the process is ergodic, from well-known results of the $G/M/1$ queue (see [3]) this requires that the mean service time must be smaller than the mean interarrival time. Hence, for all $0 \le \theta \le 1$,

$$(1-\theta)\frac{1}{\eta_1} \,+\, \theta\left(\frac{1}{\eta_1}+\frac{1}{\eta_2}\right) \;>\; \frac{1}{\mu},$$

or

$$\frac{1}{\eta_1} > \frac{1}{\mu}$$

For the Lyapunov function we try the function from $S = \mathbb{N}_0 \times \{1,2\}$ to $\mathbb{R}_+$, which is linear in the number of customers, i.e.:

$$g(k,i) \;\triangleq\; c\,k, \tag{28}$$

for some $c > 0$ and $i = 1,2$. Then for $s = (k,1)$ with $k \ge 1$

$$\begin{aligned}
\mathbb{E}\big[g(X_\theta(s,1))\big] - g(s) &= P_\theta\big((k,1);(k+1,1)\big)\, c\,(k+1) \\
&\quad + P_\theta\big((k,1);(k,2)\,\big)\, c\,k \\
&\quad + P_\theta\big((k,1);(k-1,1)\,\big)\, c\,(k-1) \,-\, c\,k \\
&= (1-\theta)\frac{\eta_1}{\eta_1+\mu}\, c\,(k+1) \,+\, \theta\frac{\eta_1}{\eta_1+\mu} c\,k + \frac{\mu}{\eta_1+\mu}\, c\,(k-1) \,-\, c\,k \\
&< 0\,.
\end{aligned}$$

Note that $s = (k,2)$ is only a reachable state if $\theta > 0$, in this case we have similarly for $k \ge 1$,

$$\mathbb{E}\big[g(X_\theta(s,1))\big] - g(s) = c(k+1)\frac{\eta_2}{\eta_2+\mu} \,+\, c\,(k-1)\frac{\mu}{\eta_2+\mu} \,-\, c\,k\,,$$

which is smaller than 0 if

$$\frac{1}{\mu} < \frac{1}{\eta_2} .$$

We conclude that for any $c > 0$ the function g in (28) is a Lyapunov function with

$$V_\theta \;=\; \{(0,1),(0,2)\} \quad \text{ for } \;\theta > 0$$

and

$$V_\theta \;=\; \{(0,1)\} \quad \text{ for } \theta = 0 \,.$$

Without loss of generality, we take $c = 1$ in the Lyapunov function g. While the ergodicity condition is

$$\frac{1}{\eta_1} \;>\; \frac{1}{\mu} \,, \tag{29}$$

we also required above that

$$\frac{1}{\eta_2} \;>\; \frac{1}{\mu} \,. \tag{30}$$

Let us point out, without going into details, that we do not need relation (30) for satisfying the condition (**C1**) if we take m_θ sufficiently large in relation (12). Indeed, suppose that relation (29) is satisfied but relation (30) not, then

$$\eta_2 \;>\; \mu \;>\; \eta_1 \,. \tag{31}$$

In the long-run the fraction of states for which the interarrival process is in phase 1 and phase 2 is $f_1 \triangleq \frac{c}{\eta_1}$ and $f_2 \triangleq \frac{c\theta}{\eta_2}$, where $c \triangleq (\frac{1}{\eta_1} + \frac{\theta}{\eta_2})^{-1}$. By (31),

$$f_1 \;>\; f_2$$

and the derivations above give that

$$\mathbb{E}[g(X_\theta(s,1))] \;-\; g(s) \begin{cases} < 0 & \text{for } s = (k,1) \\ > 0 & \text{for } s = (k,2) \,. \end{cases} \tag{32}$$

Write $\mathbb{E}[g(X_\theta(s,t))] - g(s)$ as a telescope sum:

$$\mathbb{E}[g(X_\theta(s,t))] - g(s) \;=\; \mathbb{E}\left[\sum_{k=1}^{t}\Big(\mathbb{E}[g(X_\theta(s,k))\,|\,X_\theta(s,k-1)] \;-\; g(X_\theta(s,k-1))\Big)\right] .$$

Ergodicity implies that

$$\lim_{t\to\infty} \mathbb{E}\left[\sum_{k=1}^{t} \left(\mathbb{E}[g(X_\theta(s,k))] \,-\, g(X_\theta(s,k-1)) \right)\right]$$

exists. This together with (32), we find that the number of negative terms minus the number of positive terms in the right hand side tends to infinity as $t \to \infty$ with probability one. Hence, for t sufficiently large the left hand side is negative for any state s. Take $m_\theta = t$, then relation (12) is satisfied. We conclude that indeed the ergodicity condition is sufficient for condition (**C1**).

In an ergodic Markov chain with an atom the Harris condition is automatically fulfilled, which implies that condition (**C2**) is satisfied. It is straightforward to check the condition (**C3**). By Theorem 2 together with relation (14), we may choose the bounding function v as

$$v(s,i) \,\triangleq\, e^{\lambda s}\,, \quad (s,i) \in \mathcal{S}$$

for a sufficiently small λ, which gives

$$\mathcal{D}_v \,=\, \left\{ g : S \to \mathbb{R} \,\middle|\, \exists r \in \mathbb{R} : \, |g(s,i)| \,\le r \cdot e^{\lambda s},\, (s,i) \in S \right\}.$$

Note that Theorem 2 implies that

$$\mathcal{D}_\nu \subset L^1(P_\theta, \Theta) \cap L^1(\Pi_\theta, \Theta)\,.$$

Only the transition out of state $(k,1)$ does depend on θ. Specifically, for $g \in \mathcal{D}_\nu$, it holds:

$$\frac{d}{d\theta} \sum_{r \in S} P_\theta((k,1);r)\, g(r) \,=\, g(k,2)\, \frac{\eta_1}{\eta_1 + \mu 1_{k>0}} \,-\, g(k+1,1)\, \frac{\eta_1}{\eta_1 + \mu 1_{k>0}}\,,$$

whereas

$$\frac{d}{d\theta} \sum_{r \in S} P_\theta((k,2);r)\, g(r) \,=\, 0\,.$$

Set

$$P^+((k,1);(k,2)) \,=\, \frac{\eta_1}{\eta_1 + \mu 1_{k>0}} \,=\, P^-((k,1);(k+1,1))\,,$$

$$P^+((k,1);(k+1,1)) \,=\, 0 \,=\, P^-((k,1);(k,2))\,,$$

and for any other pair of states s, s' set $P_\theta^+(s;s') = P_\theta(s;s') = P_\theta^-(s;s')$. In words, under P^+ there are two possible events that can trigger the Markov chain to leave state $(k+1)$: a departure takes place, or the phase of the interarrival is increased. Under P^- the two possible events that can trigger the Markov chain the leave state $(k+1)$ are: a departure takes place, or an arrival occurs. Then, for any $g \in \mathcal{D}_\nu$ it holds that

$$\frac{d}{d\theta}\sum_{r\in S} P_\theta(s;r)\, g(r) \;=\; \sum_{r\in S} P^+(s;r)\, g(r) \;-\; \sum_{r\in S} P^-(s;r)\, g(r)\,, \quad s \in S\,,$$

which implies that $(1, P^+, P^-)$ is a $\mathcal{D}_\nu$–derivative of P_θ. Observe that $P^+ = P_1$ and $P^- = P_0$.

A quick way of obtaining the above result is as follows. We revisit the representation of P_θ as the mixture of the kernels P and Q in (27). For any $g \in \mathcal{D}_\nu$, it holds that $\int P(s;du)g(u)$ and $\int Q(s;du)g(u)$ exist and are finite. Hence, by Example 2, P_θ is $\mathcal{D}_\nu$-differentiable with $\mathcal{D}_\nu$-derivative $(1, P, Q)$, where $P^+ = P$ and $P^- = Q$.

Because the $\mathcal{D}_\nu$–derivative of P_θ is independent of θ, P_θ is $\mathcal{D}_\nu$–Lipschitz continuous at any $\theta \in [0,1]$. Hence, provided that $1/\eta_1 > 1/\mu$, Theorem 3 applies and the estimator in (20) with $\tau_{\bar{a}}^{\pm}$ as in (19) is unbiased for any $g \in \mathcal{D}_\nu$.

5.2.2 Continuous State–Space

Let $W_\theta(n)$ be the waiting time of the nth customer arriving to the system. We have $S = \mathbb{R}$ and we take the usual norm on $\mathbb{R}$ for $||\cdot||_S$. Let F_θ denote the distribution function of the interarrival times. Let $\{A_\theta(n)\}$ be the i.i.d. sequence of interarrival times and $\{S(n)\}$ the i.i.d. sequence of service times, respectively. Lindley's recursion yields:

$$W_\theta(n+1) = \max(\, W_\theta(n) + S(n) - A_\theta(n+1)\,,\, 0\,)\,, \quad n \ge 1\,,$$

and $W_\theta(1) = 0$. For $w > 0$, the transition kernel for the waiting times is given by

$$P_\theta(v;(0,w]) = \int_0^\infty \int_{\max(w+a-v,0)}^\infty G(ds)\, F_\theta(da)$$

and

$$P_\theta(v;\{0\}) = \int_0^\infty \int_0^{\max(a-v,0)} G(ds)\, F_\theta(da)$$

For the ergodicity of Markov chain $W_\theta(n)$, $0 \le \theta \le 1$, we need that (see [3])

$$\frac{1}{\mu} \; < \; \frac{1}{\eta_1} \, . \tag{33}$$

Note that $\mu^{-1} < \eta_1^{-1}$ implies $\mu^{-1} < \eta_1^{-1} + \eta_2^{-1}$. Since G and F_1 are both exponentially distributed we have

$$\int_0^\infty \int_0^a yG(dy)\, F_\theta(da) = \mu \left((1-\theta)\frac{1}{\eta_1} \, + \theta \left(\frac{1}{\eta_1} + \frac{1}{\eta_2} \right) \right) \, > \, 1$$

Choose c such that for all θ

$$\int_0^\infty \int_0^{\min(a,c)} yG(dy)\, F_\theta(da) \ge 1,$$

and choose the Lyapunov function such that for $v \ge 0$,

$$g(v) \; = \; v \quad \text{ for } \; v \in \mathbb{R}_+ \, .$$

Then, for $v \ge c$

$$\begin{aligned} \int P_\theta(v;dy)\, g(dy) &= \int_0^\infty \int_0^{\min(a,v)} (y - a + v)\, G(dy)\, F_\theta(da) \\ &\le v - 1 \le g(v) - 1. \end{aligned}$$

Hence, g satisfies condition (**C1**) with

$$V_\theta \; = \; [0, c] \, .$$

Also this Markov chain has an atom and it is straightforward to check conditions (**C2**) and (**C3**). The bounding function is

$$v(s) \; \triangleq \; e^{\lambda s} \, , \quad s \in \mathbb{R}_+ \, ,$$

which gives

$$\mathcal{D}_v \; = \; \left\{ g : S \to \mathbb{R} \,\middle|\, \exists r \in \mathbb{R} : \, |g(s)| \, \le r \cdot e^{\lambda s}, \, s \in S \right\} .$$

Note that Theorem 2 implies that

$$\mathcal{D}_v \subset L^1(P_\theta, \Theta) \cap L^1(\Pi_\theta, \Theta) \, .$$

Like the previous example, the kernel can be written as mixture of two kernels $P = P_1$ and $Q = P_0$:

$$P_\theta \; = \; \theta P + (1-\theta) Q \, , \quad \theta \in [0,1] \, .$$

Specifically, under P, the interarrival times are distributed according to $\mathcal{E}_{(\eta_1,\eta_2)}$, and under Q, the interarrival times have distribution $\mathcal{E}_{\eta_1}$. Hence, P_θ falls into the setup of Example 2 and it readily follows that P_θ is $\mathcal{D}_v$-differentiable with $\mathcal{D}_v$-derivative $(1, P, Q)$. Because the $\mathcal{D}_1$–derivative of P_θ is independent of θ, P_θ is $\mathcal{D}_1$–Lipschitz continuous at any $\theta \in [0,1]$. Hence, if the system is stable at $\theta = 1$ (see (33)), then Theorem 3 applies and the estimator in (26) with $\tau^{\pm}_{\theta,\alpha}$ as in (25) is unbiased for any $g \in \mathcal{D}_v$.

References

[1] A.A. Borovkov and A. Hordijk. On Normed Ergodicity of Markov Chains. *Technical Report MI 2000-40 Leiden University*, 2000.

[2] X. Cao. *Realization Probabilities: The Dynamics of Queueing Networks.* Springer, New–York, 1994.

[3] J.W. Cohen. *The Single Server Queue.* North_Holland, Amsterdam, 1969.

[4] R. Dekker and A. Hordijk. Average, sensitive and Blackwell optimal policies in denumerable Markov decision chains with unbounded rewards. *Math. Oper. Res.*, 13, 1998:395-421.

[5] R. Dekker, A.Hordijk and F.M. Spieksma. On the relation between recurrence and ergodicity properties in denumerable Markov decision chains. *Math. Oper. Res.*, 19, 1994:539-559.

[6] P. Glasserman. *Gradient Estimation via Perturbation Analysis.* Kluwer Academic Publishers, Boston, 1991.

[7] P. Glasserman. Stationary waiting time derivatives. *QUESTA*, 12, 1992:369–389.

[8] P. Glasserman. Regenerative derivatives of regenerative sequences. *Adv. Appl. Prob.*, 25, 1993:116–139.

[9] B. Heidergott. A weak derivative approach to optimization of threshold parameters in a multicomponent maintenance system. *J. Appl. Prob.*, 38, 2001:386–406.

[10] B. Heidergott and X. Cao. A note on the relation between weak derivatives and perturbation realization. *IEEE Trans. Aut. Control* (to appear).

[11] B. Heidergott and F. Vázquez-Abad. Measure–valued differentiation for stochastic processes: the finite horizon case. *EURANDOM report 2000-033*, 2000.

[12] B. Heidergott and F. Vázquez-Abad. Measure–valued differentiation for stochastic processes: the random horizon case. *GERAD G–2001–18*, 2000.

[13] Y. Ho and X. Cao. *Perturbation Analysis of Discrete Event Systems.* Kluwer Academic Publishers, Boston, 1991.

[14] A. Hordijk and F.M. Spieksma. On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. Appl. Prob.*, 24, 1992:343-376.

[15] A. Hordijk and F.M. Spieksma. A New Formula for the Deviation Matrix. Chapter 36 in *Probability, Statistics and Optimization* (ed. F.P. Kelly) Wiley, 1994.

[16] A. Hordijk and A.A. Yushkevich. Blackwell optimality in the class of all policies in Markov decision chains with a Borel state space and unbounded rewards. *Math. Meth. Oper. Res.*, 50, 1999:421-448.

[17] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* Springer, London, 1993.

[18] G. Pflug. Gradient estimates for the performance of Markov chains and discrete event processes. *Annals of Operations Research*, 39, 1992:173–194.

[19] G. Pflug. *Optimisation of Stochastic Models.* Kluwer Academic Publishers, Boston, 1996.

[20] S.M. Ross *Applied Probability Models with Optimization Applications.* Holden-Day, San Francisco, 1970.