Measure-Valued Differentiation for Random Horizon Experiments

Bernd Heidergott
Vrije Universiteit and Tinbergen Institute
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
e-mail: bheidergott@feweb.vu.nl

Arie Hordijk
Leiden University, Mathematical Institute
P.O. Box 9512, 2300 RA Leiden, The Netherlands
e-mail: hordijk@math.leidenuniv.nl

Abstract
We consider the accumulate costs over a cycle of a phase-homogeneous random walk. For this model we establish sufficient conditions for the existence of the derivative of the cycle cost and we establish an unbiased gradient estimator. The main stability condition for our analysis is that the expected cycle costs are finite. We thereby improve the results known in the literature so far, where usually finiteness of higher moments of the cycle length is assumed in order to establish unbiasedness of a particular gradient estimator.

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

1 Introduction
Let $X_\theta = \{X_\theta(n) : n \geq 0\}$ be a (general state-space) Markov chain, depending on a (vector-valued) parameter $\theta \in \Theta$ with $X_\theta(0) \in \alpha$, for some measurable set
A broad class of problems can be modeled by
\[
\mathbb{E} \left[ \frac{\tau_{\alpha, \theta}(s)}{s-1} \sum_{n=0}^{\tau_{\alpha, \theta}(s)-1} g(X_{\theta}(n)) \right],
\]
where \(\tau_{\alpha, \theta}\) denotes the first entrance time after \(n = 0\) of \(X_{\theta}\) into \(\alpha\) and \(g\) is some cost function.

In optimization and sensitivity analysis, one is interested in computing/estimating the derivative of the overall performance in (1) with respect to \(\theta\). It has already been observed in the literature that there are many situations where the derivative of (1) can be obtained from observing the process up to time \(\tau_{\alpha, \theta}\). However, for establishing unbiasedness of the estimator one usually requires that the second or third moment of \(\tau_{\alpha, \theta}\) has finite expected value. See, for example, [1, 3, 8].

In this paper, we take \(X_{\theta}\) to be a phase-homogenous random walk. In this case sufficient conditions for existence of the cycle performance in (1) can be obtained through the drift of \(X_{\theta}\). Will show in this paper that if \(g\) in (1) is bounded by a polynomial of degree \(p\), then the derivative of the cycle performance exists if the \((p+1)st\) moment of the drift is finite. For example, for \(g\) bounded, we require finiteness of the expected value of \(\tau_{\alpha, \theta}\). Since finiteness of the expected value of \(\tau_{\alpha, \theta}\) is already necessary for the cycle performance in (1) to exist, our analysis provides a set of minimal conditions for unbiasedness of a gradient estimator for the derivative of the cycle cost. To summarize, we show that for phase-homogenous random walks existence of the cycle performance already implies unbiasedness of the gradient estimator for bounded cost functions.

The paper is organized as follows. In Section 2, the basic concepts from Markov chain theory and measure-valued differentiation are introduced. In Section 3, phase-homogeneous random walks are introduced, which will be analyzed in this paper. The main technical analysis is provided in Section 4. Eventually, Section 5 illustrates our results with applications to the G/G/1 queue.

Editorial note: This paper provides minimal conditions for unbiasedness of gradient estimators and it is the first result of this type. In order to achieve this, a meticulous analysis had to be performed. The perseverance of the authors was put to a test during this process since it took more than three years to complete.
the analysis. We hope that the techniques developed in this paper will be fruitful
in further studies and an extension to generalized Jackson networks is topic of
further research.

2 Markov Chains and
Measure-Valued Differentiation

2.1 Basic Notations and Definitions

Let \((S, T)\) be a Polish measurable space. Let \(\mathcal{M}(S, T)\) denote the set of fi-
nite (signed) measures on \((S, T)\) and \(\mathcal{M}_1(S, T)\) that of probability mea-
ures on \((S, T)\). The mapping \(P : S \times T \to [0, 1]\) is called a (homogeneous) transi-
tion kernel on \((S, T)\) if (i) \(P(s; \cdot) \in \mathcal{M}(S, T)\) for all \(s \in S\); and (ii) \(P(\cdot; B)\)
is \(T\) measurable for all \(B \in T\). If, in condition (i), \(\mathcal{M}(S, T)\) can be replaced
by \(\mathcal{M}_1(S, T)\), then \(P\) is called a Markov kernel on \((S, T)\). Denote the set of
transition kernels on \((S, T)\) by \(K(S, T)\) and the set of Markov kernels on \((S, T)\)
by \(K_1(S, T)\). A transition kernel \(P \in K(S, T)\) with \(0 < P(s; S) < 1\) for at least
one \(s \in S\) is called defective.

Consider a family of Markov kernels \((P_\theta : \theta \in \Theta)\) on \((S, T)\), 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 \(\int_S P_\theta(s; du) |g(u)|\) is finite for all \(\theta \in \Theta\) and \(s \in S\). A kernel \(P_\theta\) is called
\(\mathcal{D}\)-preserving, with \(\mathcal{D} \subset \mathcal{L}^1(P_\theta; \Theta)\), if \(g \in \mathcal{D}\) implies \(\int_S P_\theta(\cdot; du)g(u) \in \mathcal{D}\). To
simplify the notation, we set
\[
(P_\theta g)(s) \triangleq \int_S P_\theta(s; du)g(u)
\]
for \(g \in \mathcal{L}^1(P_\theta; \Theta)\) and \(s \in S\).

For \(P_\theta \in K_1(S, T)\) and \(V \in T\), the taboo operator associated with \(P_\theta\) for
some taboo set \(V\) is defined as
\[
\forall g \in \mathcal{L}^1(P_\theta; \Theta) : \quad (\vee P_\theta g)(s) \triangleq \int_{u \not\in V} P_\theta(s; du)g(u)
\]
for \(s \in S\). Note that if \(P_\theta(s, V) > 0\) for some \(s \in S\), then \(\vee P_\theta\) is defective.

Taking \(\alpha = V\), the expression in (1) reads
\[
\mathbb{E} \left[ \sum_{n=0}^{\tau_{\alpha, s} - 1} g(X_\theta(n)) \right] = \sum_{n=0}^{\infty} \alpha^n P_\theta^n g
\]
provided that it exists. The operator

$$H_\theta \triangleq \sum_{n=0}^{\infty} \alpha P^n_\theta$$

is called the potential of $\alpha P_\theta$. Note that the potential of $\alpha P_\theta$ yields the distribution of a cycle of $X_\theta$, in formula,

$$\left( H_\theta g \right)(s) = \mathbb{E} \left[ \sum_{n=0}^{\tau_{\alpha,\theta}-1} g(X_\theta(n)) \bigg| X_\theta(0) = s \right],$$

for any $s \in S$ and for any $g$ for which (1) exists. Denoting by $e$ the mapping that maps any $s \in S$ onto 1, gives for any $s \in S$:

$$P_\theta(\tau_{\alpha,\theta} > n \mid X_\theta(0) = s) = (\alpha P^n_\theta e)(s) \quad (2)$$

and

$$\mathbb{E}[\tau_{\alpha,\theta} \mid X_\theta(0) = s] = \sum_{n=0}^{\infty} (\alpha P^n_\theta e)(s) = (H_\theta e)(s).$$

In what follows, we let $\Theta$ be an open neighborhood of $\theta_0$ and assume that $D \subset L^1(P_\theta; \Theta)$.

**Definition 1** We call $P_\theta \in K(S,T)$ differentiable at $\theta$ with respect to $D$, or $D$-differentiable for short, if $P'_\theta \in K(S,T)$ exists such that for any $g \in D$ and any $s \in S$:

$$\frac{d}{d\theta} \int_S P_\theta(s; du) g(u) = \int_S P'_\theta(s; du) g(u). \quad (3)$$

If the left-hand side of equation (3) equals zero for all $g \in D$, then we say that $P'_\theta$ is not significant.

We denote the set of bounded continuous mappings from $S$ to $\mathbb{R}$ by $C^b(S)$ and assume, unless stated otherwise, that $C^b(S) \subset D$. This implies that $P'_\theta$ in (3) is uniquely defined provided that $P_\theta$ is $D$-differentiable. For more details on measure-valued differentiation (MVD), we refer to [5, 10, 12, 13].

**Definition 2** Let $P_\theta$ be $D$-differentiable at $\theta$. Any triple $(cP_\theta(\cdot), P^+_\theta, P^-_\theta)$, with $P^\pm_\theta \in K_1(S,T)$ and $cP_\theta$ a measurable mapping from $S$ to $\mathbb{R}$ such that

$$\forall g \in D : \quad \int_S P'_\theta(s; du) g(u) = cP_\theta(s) \left( \int_S P^+_\theta(s; du) g(u) - \int_S P^-_\theta(s; du) g(u) \right)$$

is called a $D$-derivative of $P_\theta$. 4
Remark 1 If $P_\theta$ is $\mathcal{D}$-differentiable, so is $V_{P_\theta}$ provided that $V$ is independent of $\theta$. Moreover, if $(c_{P_\theta}(\cdot), P_\theta^+, P_\theta^-)$ is an instance of a $\mathcal{D}$-derivative for $P_\theta$, then an instance of a $\mathcal{D}$-derivative of $V_{P_\theta}$ is given by

$$
\left(c_{V_{P_\theta}}, V_{P_\theta}^+, V_{P_\theta}^-ight),
$$

with $c_{V_{P_\theta}}(s) = c_{P_\theta}(s)$ for $s \in S$.

Let $v : S \to \mathbb{R}$ be a measurable mapping such that

$$\inf_{s \in S} v(s) \geq 1. \quad (4)$$

The set $V$ of mappings from $S$ to $\mathbb{R}$ can be equipped with the so-called functional $v$-norm, where

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

For $\mu$ a (signed) measure the associated measure norm is

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

and for a kernel $P$ the associated operator norm reads

$$\|P\|_v = \sup_{s \in S} \sup_{\|f\|_v \leq 1} \left|\int f(z) P(s; dz)\right| |v(s)|. $$

If $g$ has finite $v$-norm, then $|g(s)| \leq c v(s)$ for any $s \in S$ and some finite constant $c$. Let $\mathcal{H}$ be an arbitrary set of measurable mappings and let $v \in L^1(P_\theta; \Theta)$. We denote the subset of $\mathcal{H}$ constituted out of the $v$-dominated functions by $(\mathcal{H}, v)$; in formula:

$$(\mathcal{H}, v) \triangleq \{ g \in \mathcal{H} : \|g\|_v < \infty \}. $$

Let $v_p(s) = \sum_{k=0}^p d_k |s|^k$ for finite constants $d_k \geq 0$, for $p \geq k > 0$ and $d_0 > 0$, then $(\mathcal{H}, p) \triangleq (\mathcal{H}, v_p)$ denotes the set of mappings $g \in \mathcal{H}$ that are bounded by a polynomial of degree $p$, that is, $g \in (\mathcal{H}, p)$ implies that

$$|g(s)| \leq \sum_{k=0}^p c_k |s|^k$$

for some finite constants $c_k \geq 0$, $k = 0, \ldots, p$.

We call $(\mathcal{H}, v)$ (resp. $(\mathcal{H}, p)$) Banach if $(\mathcal{H}, v)$ (resp. $(\mathcal{H}, p)$) is a Banach space with respect to the $v$-norm.
Let $P_\theta$ be $(\mathcal{H}, v)$-differentiable at $\theta \in \Theta$ with $(\mathcal{H}, v)$ Banach. Then, for any neighborhood $U = [\theta - \Delta, \theta + \Delta] \subset \Theta$ of $\theta \in \Theta$ a finite constant $M$ exists such that
\[ \forall |h| \leq \Delta : \|P_{\theta+h} - P_\theta\|_v \leq |h| M, \quad (5) \]
see [7]. In words, $P_\theta$ is locally $v$-norm Lipschitz. For a signed measure $\mu$ on $(S, \mathcal{S})$ we denote its positive part by $[\mu]^+$ and its negative part by $[\mu]^-$.

The absolute value of $\mu$, in symbols $|\mu|$, is defined by $|\mu| = [\mu]^+ + [\mu]^-$ and it holds that
\[ \forall |h| \leq \Delta : \int g(u)|P_{\theta+h} - P_\theta|(du) \leq ||P_{\theta+h} - P_\theta||_v v(s) \quad (6) \]
for all $g$ such that $\|g\|_v \leq 1$, see [7] for details.

For our analysis we require a set $\mathcal{D}$ of performance measures that satisfies the following conditions:

(i) There exists $v \in L^1(P_\theta; \Theta)$ such that $\mathcal{D}$ endowed with the $\| \cdot \|_v$-norm becomes a Banach space.

(ii) $P_\theta$ is $\mathcal{D}$-differentiable.

In the following we discuss typical examples for $\mathcal{D}$, where $C(S)$ denotes the set of all continuous mappings belonging to $L^1(P_\theta; \Theta)$.

- Let $\mathcal{H} = C(S)$ and let $v \in L^1(P_\theta; \Theta)$ such that $\mathcal{D}$ satisfies (4). Then, $\mathcal{D} = (\mathcal{H}, v)$ is the set of all continuous mappings bounded by $v$ up to multiplicative constant and $\mathcal{D}$ equipped with the $\| \cdot \|_v$-norm becomes the Banach space of continuous mappings with finite $v$-norm. In particular, for $v \equiv 1$, $\mathcal{D}$ becomes the set of bounded continuous mappings, denoted by $C^b(S)$.

- Let $\mathcal{H} = L^1(P_\theta; \Theta)$ and let $v \in L^1(P_\theta; \Theta)$ such that $\mathcal{D}$ satisfies (4). Then, $\mathcal{D} = (\mathcal{H}, v)$ is the set of all measurable mappings bounded by $v$ up to a multiplicative constant and $\mathcal{D}$ equipped with the $v$-norm becomes the Banach space of measurable mappings with finite $v$-norm.

---

One may slightly deviate from the requirement of continuity. For example, let $g$ be bounded by $v$ and denote the set of discontinuities of $g$ by $D_g$. Provided that $P_\theta(s, D_g) = 0$ for any $\theta \in \Theta$ and $s \in S$, we may consider $\mathcal{D} \cup \{g\}$ for our analysis.
The question whether (ii) is satisfied for $D$ depends of $P_{\theta}$. It may happen that $P_{\theta}$ is only $D$-differentiable for a particular choice of $D$. Roughly speaking, $D$-differentiability with respect to $D = (\mathcal{L}^1(P_\theta; \Theta), v)$ is the most restrictive condition, since it requires that indicator functions are differentiable. On the other hand, $D = C^b(S)$, that is, $D = (C(S), v \equiv 1)$ is the least restrictive choice for $D$, however, excluding the analysis of possibly unbounded cost functions. See the discussion in [13] for details.

2.2 Differentiating the Potential

Let $(P_\theta : \theta \in \Theta)$ be a collection of (possibly defective) Markov kernels on $(S, T)$. For example, $P_\theta$ may be obtained through $vP_\theta$ for $P_\theta \in K_1(S, T)$ and $V \in T$, as explained in the previous section. In this section, we will compute the derivative of the potential of $P_\theta$:

$$
\frac{d}{d\theta} \sum_{k=1}^{\infty} P_\theta^k g,
$$

where $g \in D$, for appropriately defined set of cost functions $D$. Note that since $g$ is independent of $\theta$ it holds that $dg/d\theta = 0$, which implies that

$$
\frac{d}{d\theta} \sum_{k=0}^{\infty} P_\theta^k g = \frac{d}{d\theta} \sum_{k=1}^{\infty} P_\theta^k g.
$$

To simplify the notation, we write $H_\theta$ for the potential of $P_\theta$ even though $P_\theta$ does not necessarily have to be a taboo kernel.

We will show that, under appropriate conditions, $H_\theta g$ is differentiable with derivative

$$
H_\theta P_\theta' H_\theta g.
$$

Starting point is a $D$-differentiable Markov kernel $P_\theta$. Let $\Theta_0 \triangleq (\theta_l, \theta_r) \subset \Theta$ be a neighborhood of $\theta$ such that $[\theta_l, \theta_r] \subset \Theta$. The following assumptions are needed for our analysis:

(A1) $\forall g \in D : P_\theta^0 g \in D$

(A2) $\forall \theta \in \Theta_0 \forall g \in D : P_\theta^n g \in D$ for all $n$.

The following theorem presents minimal conditions for (7) to hold.

Theorem 1 Let $P_\theta$ be $D$-differentiable, and assume that (A1) and (A2) hold. Let $g \in D$ and assume that $H_\theta g \in D$ for all $\theta \in \Theta_0$. Sufficient conditions for

$$
(H_\theta g)' = H_\theta P_\theta' H_\theta g,
$$
or, more explicitly,
\[
\frac{d}{d\theta} \sum_{k=1}^{\infty} P_{\theta}^k g = \sum_{k=1}^{\infty} P_{\theta}^k P_{\theta}' \sum_{l=1}^{\infty} P_{\theta}^l g
\]
are

\[
(B1) \lim_{\Delta \to 0} \frac{1}{\Delta} \left| H_{\theta}(P_{\theta+\Delta} - P_{\theta})H_{\theta}g - \Delta H_{\theta} P_{\theta}' H_{\theta}g \right| = 0
\]

\[
(B2) \lim_{\Delta \to 0} \frac{1}{\Delta} \left| H_{\theta}(P_{\theta+\Delta} - P_{\theta})(H_{\theta+\Delta} - H_{\theta})g \right| = 0.
\]

**Proof.** Note that by the conditions put forward in the theorem the expressions on the right-hand side in the statement of the theorem are well-defined.

By calculation,

\[
\frac{1}{\Delta} \left| \sum_{n=1}^{\infty} P_{\theta+\Delta}^n g - \sum_{n=1}^{\infty} P_{\theta}^n g - \Delta \sum_{k=1}^{\infty} P_{\theta}^k P_{\theta}' \sum_{l=1}^{\infty} P_{\theta}^l g \right| \leq \frac{1}{\Delta} \left| H_{\theta}(P_{\theta+\Delta} - P_{\theta})H_{\theta}g - \Delta H_{\theta} P_{\theta}' H_{\theta}g \right| + \frac{1}{\Delta} \left| H_{\theta}(P_{\theta+\Delta} - P_{\theta})(H_{\theta+\Delta} - H_{\theta})g \right|
\]

Hence, we arrive at

\[
\frac{1}{\Delta} \left| \sum_{n=1}^{\infty} P_{\theta+\Delta}^n g - \sum_{n=1}^{\infty} P_{\theta}^n g - \Delta \sum_{k=1}^{\infty} P_{\theta}^k P_{\theta}' \sum_{l=1}^{\infty} P_{\theta}^l g \right| \leq \frac{1}{\Delta} \left| H_{\theta}(P_{\theta+\Delta} - P_{\theta})H_{\theta}g - \Delta H_{\theta} P_{\theta}' H_{\theta}g \right| + \frac{1}{\Delta} \left| H_{\theta}(P_{\theta+\Delta} - P_{\theta})(H_{\theta+\Delta} - H_{\theta})g \right| = 0
\]

and by assumptions (B1) and (B2) the last terms tend to 0 and thus the expression in (8) tends to 0. ■

### 3 Phase-Homogeneous Random Walks on $\mathbb{R}_+$

We consider a collection of Markov chains on the positive half-line with jump variables $\xi_{\theta}(x)$ in state $x$. More specifically, for $x \in \mathbb{R}_+ \triangleq \{x \in \mathbb{R} : x \geq 0\}$ let

\[
\xi_{\theta}(x) = \max(\xi_{\theta} + x, 0) = x + \max(\xi_{\theta}, -x),
\]
with \( E[\xi_\theta] \) finite for any \( \theta \in \Theta \). Observe that the very definition of \( \xi_\theta(x) \) implies that

\[
(S1) \quad x \leq y \quad \Rightarrow \quad \xi_\theta(x) \leq \xi_\theta(y).
\]

**Example 1** Let \( W_\theta(n) \) be the waiting time of the \( n \)th customer in an \( G/G/1 \) queue. Let \( S_\theta \) be a sample of the service time and let \( A_\theta \) be an independent sample of the interarrival time. Then, provided that \( W_\theta(n) = w \), Lindley’s recursion yields:

\[
W_\theta(n + 1) = \max(w + S_\theta - A_\theta, 0).
\]

Denoting the drift by \( \xi_\theta = S_\theta - A_\theta \), the \((n + 1)\)st waiting time is obtained from

\[
W_\theta(n + 1) = \xi_\theta(w) = \max(\xi_\theta + w, 0) = \max(S_\theta - A_\theta + w, 0), \text{ where } W_\theta(n) = w.
\]

Let \( \mathcal{M} \) be the class of nonnegative monotone functions on \( S \). Standard are the following notions for stochastic comparison; see, for example, [11]. For stochastic variables \( X_1, X_2 \)

\[
X_1 \leq_{st} X_2 \iff E[f(X_1)] \leq E[f(X_2)] \text{ for } f \in \mathcal{M}.
\]

As shown in [11], \( X_1 \leq_{st} X_2 \) is equivalent to \( \bar{X}_1 \leq \bar{X}_2 \) a.s. for suitably chosen versions \( \bar{X}_1 \) and \( \bar{X}_2 \). A possibly defective transition kernel \( Q \) is monotone if

\[
Qf \in \mathcal{M} \text{ for } f \in \mathcal{M}.
\]

Let \( \Theta \) be a neighborhood of \( \theta \), and assume that

\[
\xi_\theta \leq_{st} \xi_\theta' \quad \text{for } \theta \leq \theta'.
\]

Identifying the random variables with an appropriate version that translates \( \leq_{st} \)-ordering to almost sure ordering, we assume that

\[
(S2) \quad \xi_\theta \leq \xi_\theta' \quad \text{for } \theta \leq \theta'
\]

with probability one. Hence, in the following we will work with random mappings \( \xi_\theta(x) \) that are a.s. monotone in both arguments, where \( (S1) \) is guaranteed by definition and \( (S2) \) is an assumption that has to be verified in applications.

By \( (S2) \) it then holds with probability one that

\[
\sup_{\theta \in \Theta_0} \xi_\theta = \xi_{\theta_0} \triangleq \xi
\]
and, since we have assumed that the expected value of the drift is finite, it follows that

$$E \left[ \sup_{\theta \in \Theta_0} \xi_{\theta} \right] = E[\xi]$$

(11)
is finite.

**Lemma 1** Suppose that (S2) holds. Then,

(i) for \(\theta' \geq \theta\)

$$\xi_{\theta'}(x) - \xi_{\theta}(x) \leq \xi_{\theta'} - \xi_{\theta}, \quad x \in S.$$  

(ii) for \(f \in \mathcal{M}\), it holds for \(\theta' \geq \theta\)

$$f(\xi_{\theta'}(x)) - f(\xi_{\theta}(x)) \leq f(x + \xi_{\theta'}) - f(x + \xi_{\theta}), \quad x \in S.$$  

**Proof.** We distinguish the following cases:

- If \(x + \xi_{\theta}, x + \xi_{\theta'} \geq 0\), then
  
  $$\max(x + \xi_{\theta'}, 0) - \max(x + \xi_{\theta}, 0) = \xi_{\theta'} - \xi_{\theta}.$$  

- If \(x + \xi_{\theta}, x + \xi_{\theta'} \leq 0\), then
  
  $$\max(x + \xi_{\theta'}, 0) - \max(x + \xi_{\theta}, 0) = 0.$$  

(12)

We have assumed \(\theta' \geq \theta\) and (S2) implies \(\xi_{\theta'} - \xi_{\theta} \geq 0\), and inserting the above inequality on the right-hand side of (12) yields

$$\max(x + \xi_{\theta'}, 0) - \max(x + \xi_{\theta}, 0) \leq \xi_{\theta'} - \xi_{\theta},$$

for \(x + \xi_{\theta}, x + \xi_{\theta'} \leq 0\).

- If \(x + \xi_{\theta} \leq 0 \leq x + \xi_{\theta'}\), then
  
  $$\max(x + \xi_{\theta'}, 0) - \max(x + \xi_{\theta}, 0) = x + \xi_{\theta'} - 0 \leq x + \xi_{\theta'} - x - \xi_{\theta} = \xi_{\theta'} - \xi_{\theta}.$$  

From the above discussion it follows that

$$\max(x + \xi_{\theta'}, 0) - \max(x + \xi_{\theta}, 0) \leq \xi_{\theta'} - \xi_{\theta}$$

with probability one, which concludes the proof of the first part.

For the proof of the second part of the lemma, we distinguish the following cases.
• If $x + \xi_\theta, x + \xi_{\theta'} \geq 0$, then
  
  $$f(\max(x + \xi_{\theta'}, 0)) - f(\max(x + \xi_\theta, 0)) = f(x + \xi_{\theta'}) - f(x + \xi_\theta).$$

• If $x + \xi_\theta, x + \xi_{\theta'} \leq 0$, then
  
  $$f(\max(x + \xi_{\theta'}, 0)) - f(\max(x + \xi_\theta, 0)) = 0.$$  \tag{13}

We have assumed $\theta' \geq \theta$ and (S2) implies $x + \xi_{\theta'} \geq x + \xi_\theta$. Hence, by the monotonicity of $f$
  
  $$f(x + \xi_{\theta'}) - f(x + \xi_\theta) \geq 0$$

and inserting the above inequality on the right-hand side of (13) yields
  
  $$f(\max(x + \xi_{\theta'}, 0)) - f(\max(x + \xi_\theta, 0)) \leq f(x + \xi_{\theta'}) - f(x + \xi_\theta)$$

for $x + \xi_\theta, x + \xi_{\theta'} \leq 0$.

• If $x + \xi_\theta \leq 0 \leq x + \xi_{\theta'}$, then monotonicity of $f$ implies $f(0) \geq f(x + \xi_\theta)$, which gives
  
  $$f(\max(x + \xi_{\theta'}, 0)) - f(\max(x + \xi_\theta, 0)) = f(x + \xi_{\theta'}) - f(0) \leq f(x + \xi_{\theta'}) - f(x + \xi_\theta).$$

From the above discussion it follows that
  
  $$f(\max(x + \xi_{\theta'}, 0)) - f(\max(x + \xi_\theta, 0)) \leq f(x + \xi_{\theta'}) - f(x + \xi_\theta)$$

with probability one. $\blacksquare$

On $\mathbb{R}_+$ we define a Markov kernel by
  
  $$P_\theta(x, B) \triangleq P(\xi_\theta(x) \in B) = \mathbb{E}[1_B(\max(\xi_\theta + x, 0))],$$  \tag{14}

where $x \in \mathbb{R}_+$ and $B$ a Borel-set. For given initial state $x_0$, the above Markov kernel defines a random walk on the positive half-line. The increment variable $\xi_\theta$ in (9) represents the drift of the random walk. Note that one typically assumes for stability that $\mathbb{E}[\xi] < 0$.

The Markov kernel $P_\theta$ defined in (14) enjoys the following properties.

**Lemma 2** The kernel is monotone that is

$$P_\theta f \in \mathcal{M} \text{ for } f \in \mathcal{M}.$$
If (S2) holds, then the kernel is monotone w.r.t. $\theta$ that is

$$P_\theta f \leq P_{\theta'} f \quad \text{for } \theta \leq \theta',$$

for any monotone integrable mapping $f$.

**Proof.** The definition of the kernel in (14) yields $(P_\theta f)(s) = \mathbb{E}[f(\xi_\theta(s))]$ and the first part of the lemma is a direct consequence of (S1), whereas the second part of the lemma follows directly from (S2).

**Lemma 3** Let $H \subset L^1(P_\theta; \Theta)$. Provided that $\mathbb{E}||\xi||^k$ is finite, for $1 \leq k \leq p$, it holds that $P_\theta$ satisfies (A2) for $\mathcal{D}$, where $\mathcal{D} = (H, p)$.

**Proof.** Note that for $g \in \mathcal{D}$ it holds for all $\theta \in \Theta_0$ that

$$|(P_\theta g)(x)| \leq c_0 + \sum_{k=1}^{p} c_k \mathbb{E}[\max(\xi_\theta + x, 0)]^k$$

\leq c_0 + \sum_{k=1}^{p} c_k \mathbb{E}[|\xi_\theta + x|^k]

\leq c_0 + \sum_{k=1}^{p} c_k \sum_{l=0}^{k} \binom{k}{l} x^l \mathbb{E}[|\xi_\theta|^{k-l}]

\leq c_0 + \sum_{k=1}^{p} c_k \sum_{l=0}^{k} \binom{k}{l} x^l \mathbb{E}[|\xi|^{k-l}].$$

Hence, provided that $\mathbb{E}[|\xi|^k]$ is finite for $1 \leq k \leq p$, it follows that $P_\theta g \in \mathcal{D}$ for $g \in \mathcal{D}$. Condition (A2) now follows from finite induction.

**Lemma 4** Let $\mathcal{D} = (\mathcal{H}, p)$, for $\mathcal{H} \subset L^1(P_\theta; \Theta)$. If $\xi_\theta$ has $\mathcal{D}$-derivative $(c_0, \xi_\theta^+, \xi_\theta^-)$, then $P_\theta$ is $\mathcal{D}$-differentiable and (A1) holds for $\mathcal{D}$.

**Proof.** Let $\mu_\theta$ denote the distribution of $\xi_\theta$ and let $\mu_\theta^\pm$ denote the distribution of $\xi_\theta^\pm$. The assumption that $\xi_\theta$ has a $\mathcal{D}$-derivative implies that

$$\forall g \in \mathcal{D} : \quad \frac{d}{d\theta} \int g(s) \mu_\theta(ds) = \int g(s) \mu_\theta'(ds),$$

with $\mu_\theta' = c_\theta(\mu_\theta^+ - \mu_\theta^-)$. Note that $g \in \mathcal{D} = (\mathcal{H}, p)$ implies that $g(\max(\xi_\theta + x, 0))$ as a function of $\xi_\theta$ lies in $\mathcal{D}$ as well. Hence, for $g \in \mathcal{D}$ it holds that

$$\frac{d}{d\theta} \int P_\theta(s; du) g(u) = \frac{d}{d\theta} \int g(\max(u + s, 0)) \mu_\theta(du)$$

$$= \int g(\max(u + s, 0)) \mu_\theta'(du).$$
For $A \in \mathcal{T}$, set

$$P'_\theta(s; A) \triangleq \int 1_A(\max(u + s, 0)) \mu'_\theta(du),$$

then

$$\forall g \in \mathcal{D} : \frac{d}{d\theta} \int P_\theta(s; du)g(u) = \int g(u)P'_\theta(s; du),$$

which establishes $\mathcal{D}$-differentiability of $P_\theta$.

We now show that (A1) holds for $\mathcal{D}$. If $\xi_\theta$ has $\mathcal{D}$-derivative $(c_\theta, \xi^+_\theta, \xi^-_\theta)$, then $\mathbb{E}[|\xi^+_\theta|^k]$ and $\mathbb{E}[|\xi^-_\theta|^k]$ are finite for $1 \leq k \leq p$. Since $g \in \mathcal{D}$ implies that $g(\max(\xi_\theta + x, 0))$ as a function of $\xi_\theta$ lies in $\mathcal{D}$ as well, $\mathcal{D}$-differentiability of $\xi_\theta$ yields

$$|(P'_\theta g)(x)| = c_\theta \left| \mathbb{E}[g(\max(\xi^+_\theta + x, 0))] - \mathbb{E}[g(\max(\xi^-_\theta + x, 0))] \right|$$

$$\leq c_\theta \sum_{k=0}^p d_k \mathbb{E}[|\xi^+_\theta + x|^k] + c_\theta \sum_{k=0}^p d_k \mathbb{E}[|\xi^-_\theta + x|^k]$$

$$\leq c_\theta \sum_{k=0}^p \sum_{l=0}^k \binom{k}{l} x^l (\mathbb{E}[|\xi^+_\theta|^{k-l}] + \mathbb{E}[|\xi^-_\theta|^{k-l}]).$$

Since $\mathbb{E}[|\xi^+_\theta|^k]$ are finite for $1 \leq k \leq p$, it follows that $P'_\theta g \in \mathcal{D}$. ■

### 4 Technical Analysis

Conditions (A1) and (A2) depend on the transition kernel alone and given the simple structure of the kernel, sufficient conditions for (A1) and (A2) to hold can be expressed in terms of the drift, see Lemma 3 and Lemma 4 above. In Section 4.1, we establish a bound for cost accumulated over a cycle. This result will be used in Section 4.2 to establish bounds on the effect of a perturbation of $\theta$ on the cost accumulated over a cycle. The overall result will be presented in Section 4.3.

#### 4.1 Lyapunov Conditions

**Lemma 5** Let $V \in \mathcal{T}$ and $g \in \mathcal{L}^1(\mathcal{P}_\theta; \Theta)$, with $g(s) \geq 0$ for all $s$. Suppose that there exists a Lyapunov function $g^\lambda \in \mathcal{L}^1(\mathcal{P}_\theta; \Theta)$ such that

$$g + (\nu P_\theta g^\lambda) \leq g^\lambda, \quad (15)$$
and moreover, suppose that for some $c$

$$\sup_{s \in V} g^\lambda(s) \leq c.$$ 

Then, for $N(s)$ the number of visits to the set $V$ provided that $X_\theta(0) = s$, it holds

$$\mathbb{E}_\theta \left[ \sum_{t=1}^\infty g(X_\theta(t)) \bigg| X_\theta(0) = s \right] \leq g^\lambda(s) + c \mathbb{E}_\theta[N(s)].$$

**Proof.** Let $\tau_1 < \tau_2 < \tau_3 < \ldots$ be the successive recurrence times to the set $V$, and let $\tau_0 = 0$. Then rewriting the expected costs with direct cost function $g$ over the infinite horizon in blocks over the periods between the successive recurrence times to the set $V$, we get

$$\mathbb{E}_\theta \left[ \sum_{t=1}^\infty g(X_\theta(t)) \bigg| X_\theta(0) = s \right] = \mathbb{E}_\theta \left[ \sum_{k=1}^\infty \sum_{t=\tau_{k-1}}^{\tau_k-1} g(X_\theta(t)) \bigg| X_\theta(0) = s \right]. \quad (16)$$

In the following we show that the expected costs over the first cycle are bounded by $g^\lambda$. Multiplying the Lyapunov inequality (15) form the left by $vP_{\theta}g$ gives

$$vP_{\theta}g + (vP_{\theta}^2g^\lambda) \leq vP_{\theta}g^\lambda.$$

Adding $g$ on both sides of the inequality and using the Laypunov inequality for the expression on the right-hand side of the inequality yields

$$g + vP_{\theta}g + vP_{\theta}^2g^\lambda \leq g^\lambda.$$

Repeating this argument $n$-times gives,

$$\sum_{l=0}^n vP_{\theta}^lg + vP_{\theta}^{n+1}g^\lambda \leq g^\lambda.$$

Since, $g^\lambda \geq g \geq 0$ we find by taking the limit as $n$ tends to infinity that

$$\sum_{l=0}^\infty vP_{\theta}^lg \leq g^\lambda. \quad (17)$$

Note that the $l$-th term in this sum is operator notation for the expected costs at time $l$ on the event that the first recurrence time $\tau_1 > l$, i.e.

$$(vP_{\theta}^lg)(s) = \mathbb{E}_\theta[g(X_\theta(l)) \mathbf{1}(\tau_1 > l) \mid X_\theta(0) = s],$$
where $\mathbf{1}(\tau_1 > l)$ is the indicator function of the event that the first recurrence time is larger than $l$. Hence,

$$
\sum_{l=0}^{\infty} (VP^l g)(s) = \sum_{l=0}^{\infty} \mathbb{E}_{\theta}[g(X_\theta(l)) \mathbf{1}(\tau_1 > l) \mid X_\theta(0) = s]
$$

$$
= \mathbb{E}_{\theta} \left[ \sum_{l=1}^{\tau_1 - 1} g(X_\theta(t)) \left| X_\theta(0) = s \right. \right]
$$

$$
\leq g^\lambda(s), \quad (18)
$$

where the last inequality follows from (17).

In a similar way, an upper bound for the expected costs over the $k$th cycle can be obtained. Indeed, by using the Markov property we find with $N(s)$ the number of recurrences to the set $V$ (note that we will allow that $N(s)$ is equal to infinity with positive probability, in this case the assertion is obvious true), for $k = 2, 3, \ldots$

$$
\mathbb{E}_{\theta} \left[ \sum_{t=\tau_{k-1}}^{\tau_{k-1} - 1} g(X_\theta(t)) \left| N(s) > k - 1, X_\theta(\tau_{k-1}) = s \right. \right] = \sum_{l=0}^{\infty} (VP^l g)(s) \leq g^\lambda(s),
$$

where the inequality follows from (17). Since, $\tau_{k-1}$ is a recurrence time to the set $V$ we have $X_\theta(\tau_{k-1}) \in V$ and, moreover, from the assumption $g^\lambda(X_\theta(\tau_{k-1})) \leq c$ we obtain for any $u \in V$:

$$
\mathbb{E}_{\theta} \left[ \sum_{t=\tau_{k-1}}^{\tau_{k-1} - 1} g(X_\theta(t)) \left| N(s) > k - 1, X_\theta(\tau_{k-1}) = u \right. \right] < c.
$$

By the strong Markov property this yields

$$
\mathbb{E}_{\theta} \left[ \sum_{t=\tau_{k-1}}^{\tau_{k-1} - 1} g(X_\theta(t)) \left| N(s) > k - 1, X_\theta(\tau_{k-1}) = u, X_\theta(0) = s \right. \right] < c, \quad \text{for all } u \in V.
$$

The above bound holds uniformly on $V$ and since $X_\theta(\tau_{k-1}) \in V$, we can disregard the condition $X_\theta(\tau_{k-1}) = u$ in the above bound which yields

$$
\mathbb{E}_{\theta} \left[ \sum_{t=\tau_{k-1}}^{\tau_{k-1} - 1} g(X_\theta(t)) \left| N(s) > k - 1, X_\theta(0) = s \right. \right] < c
$$

and we finally arrive at

$$
\mathbb{E}_{\theta} \left[ \sum_{t=\tau_{k-1}}^{\tau_{k-1} - 1} g(X_\theta(t)) \left| X_\theta(0) = s \right. \right] \leq cP_\theta(N(s) > k - 1 \mid X_\theta(0) = s). \quad (19)
$$
We now combine the above results in order to establish an upper bound for the overall expected costs as given on the right-hand side of (16):

\[
E_{\theta} \left[ \sum_{t=1}^{\infty} g(X_\theta(t)) \left| X_\theta(0) = s \right. \right] = E_{\theta} \left[ \sum_{k=1}^{\infty} \sum_{t=\tau_k-1}^{\tau_k-1} g(X_\theta(t)) \left| X_\theta(0) = s \right. \right] \\
\leq g^\lambda(s) + E_{\theta} \left[ \sum_{k=2}^{\infty} \sum_{t=\tau_k-1}^{\tau_k-1} g(X_\theta(t)) \left| X_\theta(0) = s \right. \right] \\
\leq g^\lambda(s) + c \sum_{k=2}^{\infty} P_\theta \left( N(s) > k - 1 \mid X_\theta(0) = s \right) \\
\leq g^\lambda(s) + c E_{\theta} \left[ N(s) \mid X_\theta(0) = s \right],
\]

which completes the proof. ■

Recall that we are interested in the cumulative cost until the Markov chains hits a predefined set \( \alpha \). We will apply the above lemma to the particular defective Markov transition kernel \( \alpha P_\theta \).

**Lemma 6** Let \( \alpha \subset V \) and for \( X_\theta(0) = s \) denote by \( N_\alpha(s) \) the number of visits to \( V \) without hitting \( \alpha \). Suppose that for \( x \notin V \)

\[
g + (V P_\theta g^1) \leq g^1,
\]

and

\[
c \triangleq \sup_{x \in V} (g(x) + (V P_\theta g^1(x))) < \infty
\]

then

\[
g^\lambda(x) \triangleq \begin{cases} 
  g^1(x) & \text{for } x \notin V \\
  c & \text{for } x \in V,
\end{cases}
\]

satisfies the conditions in Lemma 5 and it holds that

\[
E_{\theta} \left[ \sum_{t=1}^{\tau_{\alpha}-1} g(X_\theta(t)) \left| X_\theta(0) = s \right. \right] \leq g^\lambda(s) + c E_{\theta} \left[ N_\alpha(s) \right], \quad s \in S.
\]

**Proof.** The proof follows from Lemma 5 by taking with \( P_\theta = \alpha P_\theta \). ■

**Remark 2** For verifying inequality (15) for \( g^\lambda \), it is sufficient to check it for \( g^1 \) for \( x \notin V \) and to verify (21). In our applications we will verify (20) and (21). Note that if \( g^1 \) is bounded in absolute value by a polynomial of degree \( p \) then so is \( g^\lambda \).
4.2 Bounds on the Effect of a Finite Perturbation

Recall the definition of $\xi$ in (10). We say that a Lyapunov condition holds for $p$, with $p \geq 0$ if

- condition (S2) holds,
- it holds that $\mathbb{E}[\xi] < 0$ and
  $$\mathbb{E} \left[ |\xi|^{p+1} \right] < \infty,$$
- for each $x_0 \in \mathbb{R}^+$ it holds that $\sup_{x \in S} \mathbb{E}_{\theta}[N_\alpha(x_0,s)]$ is finite, where $N_\alpha(x_0,s)$ denotes the number of visits to $[0,x_0]$ with $X_\theta(0) = s$ and without hitting $\alpha$.

The Lyapunov condition allows to bound $H_\theta g(x)$ as a function in $x$. The precise statement is given in the following lemma.

**Lemma 7** Suppose that the Lyapunov condition holds for $p$, with $p \geq 0$ Let $\mathcal{H} \subset L^1(\mathcal{P}_\theta; \Theta)$, then for each $g \in (\mathcal{H}, p)$ a function $f \in (\mathcal{H}, p+1)$ exists such that

$$\sup_{\theta \in \Theta_0} H_\theta g \leq f.$$ 

**Proof.** Suppose that $g(x) \leq \sum_{k=0}^{p} c_k x^k$, then

$$H_\theta g \leq \sum_{k=0}^{p} c_k H_\theta g_k$$

with $g_k(x) = x^k$. Hence, it suffices to show the assertion for $g_k$. We will show it for $g(x) = x^p$ for $x \geq 0$, the proof for the other terms goes similarly. We try to satisfy the condition of Lemma 6 with the function $g^1(x) = c x^{p+1}$ and so we try to find $V \subset [0, \infty)$ such that

$$x^p + \int_{y \notin V} g^1(y) \mathbb{P}(\xi_\theta(x) \in dy) \leq c x^{p+1}.$$ 

Set $\hat{\xi}_\theta(x) \triangleq \max(\xi_\theta, -x)$, which gives $\hat{\xi}_\theta(x) + x = \xi_\theta(x)$. By computation,

$$x^p + \int_{y \notin V} c y^{p+1} \mathbb{P}(\xi_\theta(x) \in dy) \leq x^p + \int_{y \geq 0} c y^{p+1} \mathbb{P}(\xi_\theta(x) \in dy)$$

$$\leq x^p + \int_{y \geq 0} c(x+y)^{p+1} \mathbb{P}(\hat{\xi}_\theta(x) \in dy)$$

$$\leq x^p + c \sum_{k=0}^{p+1} (p+1)x^{p+1-k} \mathbb{E} \left[ \hat{\xi}_\theta(x)^k \right].$$

17
To simplify the notation, set
\[ h(x) \triangleq \sum_{k=2}^{p+1} (p+1)^{p+1-k} \left| E \left[ \hat{\xi}_\theta(x)^k \right] \right|. \]

With this notation, we obtain
\[ x^p + \int_{y \notin V} c y^{p+1} P(\xi_\theta(x) \in dy) \leq x^p + c \left( x^{p+1} + (p+1)x^p \left[ \hat{\xi}_\theta(x) \right] + h(x) \right). \] (22)

Recall that \( \hat{\xi}_\theta(x) = \max(\xi_\theta, -x) \), which implies \( \hat{\xi}_\theta(x) \approx \xi_\theta \) for \( x \) large. This motivates the following line of argument. Choose \( \varepsilon > 0 \) small enough such that
\[ E[|\xi| + \varepsilon] = E[\max(\xi_\theta, -x)] < \gamma \]

for some \( x_0' \) large enough such that
\[ E[|\xi| 1_{\xi < -x_0'}] < \varepsilon. \]

Then it holds for all \( x \geq x_0' \) that
\[ \forall \theta : \quad E[\hat{\xi}_\theta(x)] = E[\max(\xi_\theta, -x)] \leq E[\max(\xi, -x)] \leq \gamma. \]

Inserting this into (22) yields
\[ x^p + \int_{y \notin V} c y^{p+1} P(\xi_\theta(x) \in dy) \leq x^p + c \left( x^{p+1} + (p+1)x^p \gamma + h(x) \right) \]
\[ = cx^{p+1} + (1 + c(p+1)\gamma) x^p + ch(x). \] (23)

We now take \( c > 0 \) such that
\[ 1 + c(p+1)\gamma < 0 \]
and \( x_0' \) so large that for \( x \geq x_0' \)
\[ (1 + c(p+1)\gamma) x^p + ch(x) \leq 0. \] (24)

Inserting (24) into (23) yields
\[ x^p + \int_{y \notin V} c y^{p+1} P(\xi_\theta(x) \in dy) \leq cx^{p+1}. \]

This establishes for \( V = \{ x \mid x \leq x_0 \} \) the inequality
\[ \forall \theta : \quad x^p + \int_{y \notin V} g^1(y) P(\xi_\theta(x) \in dy) \leq cx^{p+1}. \]
By Lemma 6 we find a Lyapunov function of the type
\[ g_\lambda(x) = \begin{cases} \lambda x^{p+1} & \text{for } x \not\in V \\ c_0 & \text{for } x \in V, \end{cases} \]
where \( c_0 \) is defined as in (21). Hence, by Lemma 6 and our assumption that \( \sup_{s \in S, \theta \in \Theta} E_\theta[N_\alpha(x_0, s)] \) is finite, it follows that \( \sup_{\theta \in \Theta} H_\theta g \) is bounded by a polynomial of degree \( p + 1 \).

The following lemma provides a sufficient condition for \( \sup_{\theta \in \Theta} E_\theta[N_\alpha(x_0, s)] \) to be finite for any \( s \in S \).

**Lemma 8** Let \( \alpha = \{0\} \). Suppose that for each \( x_0 > 0 \)
\[ \mathbb{P}(\xi \leq -x_0) \triangleq p(x_0) > 0. \]
Then \( \sup_{s \in S, \theta \in \Theta} E_\theta[N_\alpha(x_0, s)] \) is finite.

**Proof.** For \( x_0 > 0 \) set \( V = \{x | x \geq x_0\} \). Note that for each \( x_0 > 0 \) there is \( p(x_0) > 0 \) such that
\[ \inf_{x \in V} \inf_{\theta \in \Theta} \mathbb{P}(X_\theta(t+1)=0 | X_\theta(t) = x, X_\theta = s) \]
\[ = \inf_{x \in V} \inf_{\theta \in \Theta} \mathbb{P}(\xi_\theta \leq -x) \]
\[ \geq \mathbb{P}(\xi \leq -x_0) = p(x_0) > 0, \]
for any \( s \in S \). In words, the probability that the process jumps from a state in \( V \) immediately to \( \alpha = \{0\} \) is at least \( p(x_0) \). A simple geometrical trial argument (with probability of success \( p(x_0) \)) then shows that
\[ \sup_{s \in S, \theta \in \Theta} E_\theta[N_\alpha(x_0, s)] \leq \frac{p(x_0)}{1 - p(x_0)}, \quad s \in S, \]
which completes the proof.

**Lemma 9** Let \( \alpha = \{0\} \). Let \( g \in L^1(P_\theta; \Theta) \cap M \) and assume that
(i) \( H_\theta g \) is bounded by a function \( f \) in a neighborhood \( \Theta_1 \) of \( \theta \):
\[ \forall \theta \in \Theta_1 : \quad h_\theta(x) \triangleq H_\theta g(x) \leq f(x) \]
(ii) \( (S2) \) holds.
Let \( \theta, \theta' \in \Theta_1 \), then for \( \theta \leq \theta' \)

\[
| (P_{\theta'} h_{\theta})(x) - (P_{\theta} h_{\theta})(x) | \leq \mathbb{E} \left[ f(\xi_{\theta'} - \xi_{\theta}) \right] \tag{25}
\]

and for \( \theta \geq \theta' \)

\[
| (P_{\theta'} h_{\theta})(x) - (P_{\theta} h_{\theta})(x) | \leq \mathbb{E} \left[ f(\xi_{\theta} - \xi_{\theta'}) \right] \; ; \tag{26}
\]

moreover,

\[
|h_{\theta'}(x) - h_{\theta}(x)| \leq \mathbb{E} \left[ |f(\xi_{\theta'} - \xi_{\theta})| \right] \sum_{n=0}^{\infty} P_{\theta}^n e. \tag{27}
\]

**Proof.** Since \( g \) is monotone, it follows from Lemma 2 by induction that

\[
h_{\theta} = \sum_{k=0}^{\infty} P_{\theta}^k g \in \mathcal{M}.
\]

The coupling of two processes with the same transition operator but different starting states, say \( x \) and \( x + y \) with \( y \geq 0 \) gives that

\[
h_{\theta}(x + y) \leq h_{\theta}(x) + h_{\theta}(y),
\]

since before absorbing in \( \alpha = \{0\} \) the state of the process with starting state \( x + y \) is always larger than that with starting state \( x \) and at absorption it is at most \( y \). With the monotonicity of \( h_{\theta}(x) \) in \( x \) we have the inequalities

\[
h_{\theta}(x) \leq h_{\theta}(x + y) \leq h_{\theta}(x) + h_{\theta}(y),
\]

and consequently

\[
h_{\theta}(x + y) - h_{\theta}(x) \leq h_{\theta}(y).
\]

Condition (S2) implies \( \xi_{\theta'}(x) - \xi_{\theta}(x) \geq 0 \) a.s. Substituting \( x + \xi_{\theta}(x) \) for \( x \) and \( (\xi_{\theta'}(x) - \xi_{\theta}(x)) \) for \( y \) in the above inequality yields

\[
h_{\theta}(x + \xi_{\theta'}(x)) - h_{\theta}(x + \xi_{\theta}(x)) \leq h_{\theta}(\xi_{\theta'}(x) - \xi_{\theta}(x))
\]

\[
\leq h_{\theta}(\xi_{\theta'} - \xi_{\theta}),
\tag{28}
\]

where the last inequality follows from Lemma 1 (i). For \( \theta \leq \theta' \) we have \( \xi_{\theta}(x) \leq_{st} \xi_{\theta'}(x) \) and, since \( h_{\theta} \in \mathcal{M} \), we obtain

\[
|P_{\theta'} h_{\theta} - P_{\theta} h_{\theta}|(x) = |\mathbb{E} \left[ h_{\theta}(x + \xi_{\theta'}(x)) - h_{\theta}(x + \xi_{\theta}(x)) \right]|
\]

\[
= \mathbb{E} \left[ h_{\theta}(x + \xi_{\theta'}(x)) - h_{\theta}(x + \xi_{\theta}(x)) \right]
\]

\[
\leq \mathbb{E} \left[ h_{\theta}(\xi_{\theta'} - \xi_{\theta}) \right].
\]
where the last inequality follows from (28). Since \( h_\theta(x) \leq f(x) \), we obtain
\[
|P_{\theta'} h_\theta - P_\theta h_\theta| \leq \mathbb{E}[f(\xi_{\theta'} - \xi_\theta)].
\] (29)

The proof for \( \theta \geq \theta' \) is similar.

The identity
\[
P^n_{\theta'} - P^n_\theta = \sum_{k=0}^{n-1} P^k_{\theta'} (P_{\theta'} - P_\theta) P^{n-k-1}_\theta
\]
can be proved by induction. This implies
\[
H_{\theta'} g - H_\theta g = \sum_{n=0}^{\infty} (P^n_{\theta'} - P^n_\theta) g
\]
\[
= \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} P^k_{\theta'} (P_{\theta'} - P_\theta) P^{n-k-1}_\theta g
\]
\[
= \sum_{n=0}^{\infty} P^n_{\theta'} (P_{\theta'} - P_\theta) \sum_{k=0}^{\infty} P^k_\theta g
\]
\[
= \sum_{n=0}^{\infty} P^n_{\theta'} (P_{\theta'} - P_\theta) h_\theta.
\]

Hence, together with (29) we find,
\[
|H_{\theta'} g - H_\theta g| \leq \mathbb{E}[|f(\xi_{\theta'} - \xi_\theta)|] \sum_{n=0}^{\infty} P^n_{\theta'} e.
\] (30)

Reversing the roles of \( \theta' \) and \( \theta \) proves the last assertion.

Lemma 10 Suppose that the Lyapunov condition holds for \( p \), with \( p \geq 0 \). Let \( \Theta_1 \subset \Theta \) such that, for \( l = 0, \ldots, p + 1 \),
\[
\sup_{\theta' \in \Theta_1} \left| \frac{d}{d\theta} \right|_{\theta=\theta'} \mathbb{E}[|\xi_{\theta'}|^l] \overset{\triangle}{=} a_l < \infty
\]
then for \( g \in (\mathcal{H}, p) \cap \mathcal{M} \) it holds for all \( \theta, \theta' \in \Theta_1 \)
\[
|(P_{\theta'} H_\theta g)(x) - (P_\theta H_\theta g)(x)| \leq c_0 + |\theta' - \theta| C_1,
\] (31)
where
\[
C_1 \overset{\triangle}{=} \sum_{k=1}^{p} c_k a_k.
\]

Proof. Let \( \theta' > \theta \). By Lemma 7 together with Lemma 9, it holds that
\[
|(P_{\theta'} H_\theta g)(x) - (P_\theta H_\theta g)(x)| \leq \mathbb{E}[f(\xi_{\theta'} - \xi_\theta)]
\]
with \( f(x) = \sum_{k=0}^{p+1} c_k x^k \). By calculation,

\[
E[|f(\xi_{\theta'} - \xi_\theta)|] \leq c_0 + \sum_{k=1}^{p+1} c_k E[|\xi_{\theta'} - \xi_\theta|^k].
\]

For \( \theta' > \theta \), it holds

\[
E[|\xi_{\theta'} - \xi_\theta|^k] = E[(\xi_{\theta'} - \xi_\theta)^k] \\
\leq E[\xi_\theta^k] - E[\xi_{\theta'}^k],
\]

where we use the fact that \( \xi_{\theta'} \geq \xi_\theta \) a.s., and, for \( \theta' < \theta \), we obtain

\[
E[|\xi_{\theta'} - \xi_\theta|^k] = E[(\xi_\theta - \xi_{\theta'})^k] \\
\leq E[\xi_\theta^k] - E[\xi_{\theta'}^k],
\]

where we use the fact that \( \xi_\theta \geq \xi_{\theta'} \) a.s. Combining the above inequalities we arrive at

\[
E[|\xi_{\theta'} - \xi_\theta|^k] \leq |E[\xi_{\theta'}^k] - E[\xi_\theta^k]|
\]

and using the fact that \( a_k \) is a Lipschitz-constant for \( E[\xi]^k \), which proves the claim for \( \theta' > \theta \). The proof for the case \( \theta' < \theta \) follows from the same reasoning.

\[\blacksquare\]

**Lemma 11** Let the Lyapunov condition be satisfied for \( p \), with \( p \geq 0 \), and let \( \Theta_1 \subset \Theta \) be a neighborhood of \( \theta \). If,

(i) \( \xi_\theta \) is \((\mathcal{H}, p + 1)\)-differentiable on \( \Theta_1 \)

(ii) for \( g \in (\mathcal{H}, p) \) it holds that

\[
\sup_x \sup_{\theta \in \Theta_1} \left| P_{\theta}^t H_\theta g(x) - P_{\theta'}^t H_{\theta'} g(x) \right| < \infty,
\]

then it holds for \( g \in (\mathcal{H}, p) \cap \mathcal{M} \) and \( \theta', \theta \in \Theta_1 \) that

\[
|(P_{\theta'} - P_{\theta}) H_\theta g(x)| \leq |\theta' - \theta| f_0(x)
\]

for \( f_1 \in (\mathcal{H}, 0) \), and

\[
|(H_{\theta'} g(x) - H_\theta g(x))| \leq |\theta' - \theta| f_1(x)
\]

for \( f_1 \in (\mathcal{H}, 1) \).
Proof. Let $H_{\theta}g = h_{\theta}$. We have assumed that $g \in (\mathcal{H}, p)$ and Lemma 7 yields $h_{\theta} \in (\mathcal{H}, p + 1)$. The second part of the lemma is a direct consequence of the first part. To see this, note that

$$h_{\theta'} - h_{\theta} = \sum_{n=0}^{\infty} P_{\theta}^{n}(P_{\theta'} - P_{\theta})h_{\theta},$$

see the proof of Lemma 9 for details. The first part of the lemma implies

$$|h_{\theta'} - h_{\theta}| \leq |\theta' - \theta| \sum_{n=0}^{\infty} P_{\theta}^{n} f_{0} = |\theta' - \theta| \mathcal{H}_{\theta} f_{0},$$

for $f_{0} \in (\mathcal{H}, 0)$. By Lemma 7, $H_{\theta} f_{0} \in (\mathcal{H}, 1)$, which concludes the proof of the second part of the lemma.

We now turn to the proof of the first part of the lemma. By (i), $P_{\theta}$ is $(\mathcal{H}, p + 1)$-differentiable, see Lemma 4. The Mean Value Theorem implies that

$$(P_{\theta'} - P_{\theta})h_{\theta}(x) = (\theta' - \theta)P_{\theta' + \delta(x)} h_{\theta}(x),$$

for $|\delta(x)| \leq |\theta' - \theta|$ for $|\theta' - \theta|$ sufficiently small.

Suppose that

$$\sup_{x} |P_{\theta}^{n} h_{\theta}(x)| = \infty.$$ 

Then, a sequence $(x_{k}, \Delta_{k})$ exists such that

$$\lim_{k \to \infty} \Delta_{k} = 0 \quad \text{and} \quad \lim_{k \to \infty} \Delta_{k} |P_{\theta}^{n} h_{\theta}(x_{k})| = \infty.$$  

(33)

Inserting this sequence into (32), we obtain

$$(P_{\theta + \Delta_{k}} - P_{\theta})h_{\theta}(x_{k}) = \Delta_{k} P_{\theta + \delta(x_{k})} h_{\theta}(x_{k}),$$

$$= \Delta_{k} P_{\theta} h_{\theta}(x_{k}) + \Delta_{k} (P_{\theta + \delta(x_{k})} h_{\theta}(x_{k}) - P_{\theta} h_{\theta}(x_{k})).$$

Note that for $k$ sufficiently large it holds that

$$|P_{\theta + \delta(x_{k})} h_{\theta}(x_{k}) - P_{\theta}^{n} h_{\theta}(x_{k})| \leq \sup_{x} \sup_{\theta' \in \Theta} |P_{\theta'}^{n} h_{\theta}(x) - P_{\theta}^{n} h_{\theta}(x)|.$$

Condition (ii) yields that

$$\sup_{x} \sup_{\theta' \in \Theta} |P_{\theta'}^{n} h_{\theta}(x) - P_{\theta}^{n} h_{\theta}(x)| < \infty,$$

which implies

$$\lim_{k \to \infty} \Delta_{k} (P_{\theta + \delta(x_{k})} h_{\theta}(x_{k}) - P_{\theta}^{n} h_{\theta}(x_{k})) = 0.$$
and we obtain from (33) that
\[
\limsup_{k \to \infty} (P_{\theta + \Delta} - P_{\theta}) h_{\theta}(x_k) \in \{\infty, -\infty\},
\]
which contradicts equation (31) in Lemma 10. Hence,
\[
\sup_x |P'_{\theta} h_{\theta}(x)| < \infty,
\]
and from condition (v) it thus follows that
\[
\sup_x |P'_{\hat{\theta}} h_{\theta}(x)| < \infty
\]
for any \(\hat{\theta}\) such that \(|\theta - \hat{\theta}| \leq \delta\) for \(\delta\) sufficiently small. We have thus shown that
\[
C \overset{\text{def}}{=} \sup_{\{\hat{\theta} : |\theta - \hat{\theta}| \leq \delta\}} \sup_x |P'_{\hat{\theta}} h_{\theta}(x)| < \infty.
\]
Applying the Mean Value Theorem now gives
\[
|(P_{\theta'} - P_{\theta}) h_{\theta}(x)| \leq |\theta' - \theta| C
\]
for \(|\theta' - \theta| \leq \delta\), which proves the claim.

\section{4.3 Main Result}

The analysis of the previous sections shows that if \(g\) is bounded by a polynomial of degree \(p\), then the cycle performance \(H_{\theta} g\) is monotone and bounded by a polynomial of degree \(p + 1\). Surprisingly enough, provided that \(g\) is monotone, multiplying the expected cycle cost by the weak derivative of the kernel reduces the order of the bound and \(P'_{\theta} H_{\theta} g\) is bounded by a constant, i.e., a polynomial of degree 0. Hence, finiteness of \(H_{\theta} g\) for any monotone cost function bounded by a polynomial of degree \(p\) implies that \(H_{\theta} P'_{\theta} H_{\theta} g\) exists. The precise technical analysis is put forward in the following theorem.

\textbf{Theorem 2} Let \((\mathcal{H}, p + 1)\) be Banach, for \(p \geq 0\), and \(\mathcal{H} \subset L^1(P_\theta; \Theta)\). Let \(\Theta_1 \subset \Theta\). Suppose that

(i) \((S2)\) holds,

(ii) \(E[\xi] < 0\) and \(E[|\xi|^l]\) finite for \(l = 1, \ldots, p + 1\),

24
(iii) suppose that for each $x_0 > 0$ it holds that \( \sup_{s \in S, \theta \in \Theta_1} \mathbb{E}[N_\alpha(x_0, s)] < \infty \),

(iv) \( \xi_\theta \) is \( (\mathcal{H}, p+1) \)-differentiable on \( \Theta_1 \) and for \( l = 1, \ldots, p + 1 \) and

\[
\sup_{\theta' \in \Theta_1} \left| \frac{d}{d\theta} \mathbb{E} \left[ (\xi_\theta)' \right] \right| < \infty,
\]

(v) for \( g \in (\mathcal{H}, p) \) it holds that

\[
\sup \sup_x \left| P_\theta' \xi_\theta g(x) - P_\hat{\theta}' \xi_\theta g(x) \right| < \infty,
\]

then it holds for any nonnegative and monotone \( g \in (\mathcal{H}, p) \) that

\[
\frac{d}{d\theta} \sum_{k=1}^\infty P_\theta^k g (x) = \sum_{k=1}^\infty P_\theta^k P_\theta' \sum_{l=1}^\infty P_\theta^l g \in (\mathcal{H}, 1).
\]

**Proof.** The proof follows from Theorem 1 provided that (A1), (A2) and (B1), (B2) hold.

The moment conditions on \( \xi \) imply (A2) for \( (\mathcal{H}, p+1) \), see Lemma 3. Lemma 4 yields \( (\mathcal{H}, p+1) \)-differentiability of \( P_\theta \) and establishes (A1) for \( (\mathcal{H}, p+1) \).

We now show (B1) for \( g \in (\mathcal{H}, p) \cap \mathcal{M} \). Let \( h_\theta \triangleq H_\theta g \). By Lemma 11, \( f_0 \in (\mathcal{H}, 0) \) exists such that

\[
\sup_{(\Delta, \theta + \Delta \in \Theta_1)} \frac{1}{|\Delta|} \left| (P_{\theta+\Delta} - P_\theta) h_\theta(x) \right| \leq f_0(x).
\]

We have already established that \( h_\theta \in (\mathcal{H}, p + 1) \) and that \( P_\theta \) is \( (\mathcal{H}, p + 1) \)-differentiable, which yields

\[
\lim_{\Delta \to 0} \frac{1}{|\Delta|} (P_{\theta+\Delta} - P_\theta) h_\theta = P_\theta' h_\theta.
\]

Applying the dominated convergence theorem then yields

\[
\lim_{\Delta \to 0} \frac{1}{|\Delta|} H_\theta (P_{\theta+\Delta} - P_\theta) h_\theta = H_\theta P_\theta' h_\theta,
\]

which establishes (B1). Moreover, by (34) together with (35) it follows from Lemma 7 that \( H_\theta P_\theta' h_\theta \in (\mathcal{H}, 1) \).

We now turn to condition (B2) for \( g \in (\mathcal{H}, p) \cap \mathcal{M} \). By Lemma 11,

\[
|h_{\theta+\Delta} - h_\theta| \leq |\Delta| f_1,
\]

for \( \Delta \) sufficiently small and \( f_1 \in (\mathcal{H}, 1) \). Recall that the positive part (of the Hahn-Jordan decomposition) of a signed measure \( \mu \) is denoted by \([\mu]^+\) and the
negative part by $[\mu^-]$. Note that

$$\left| \frac{1}{\Delta} (P_{\theta+\Delta} - P_\theta)(h_{\theta+\Delta} - h_\theta) \right|$$

$$\leq \frac{1}{\Delta} \cdot \left[ (P_{\theta+\Delta} - P_\theta)^+ \right] |h_{\theta+\Delta} - h_\theta| + \frac{1}{\Delta} \cdot \left[ (P_{\theta+\Delta} - P_\theta)^- \right] |h_{\theta+\Delta} - h_\theta|$$

$$\leq \left[ (P_{\theta+\Delta} - P_\theta)^+ \right] f_1 + \left[ (P_{\theta+\Delta} - P_\theta)^- \right] f_1$$

It holds that

$$\left[ (P_{\theta+\Delta} - P_\theta)^\pm \right] f_1 \leq \|P_{\theta+\Delta} - P_\theta\|_{f_1} f_1 \leq |\Delta| M f_1,$$

for some finite number $M$, see (5) together with (6). Hence,

$$\left| \frac{1}{\Delta} (P_{\theta+\Delta} - P_\theta)(h_{\theta+\Delta} - h_\theta) \right| \leq 2|\Delta| M f_1$$

and

$$\lim_{\Delta \to 0} \left| \frac{1}{\Delta} (P_{\theta+\Delta} - P_\theta)(h_{\theta+\Delta} - h_\theta) \right| = 0. \quad (36)$$

Note that

$$\left| \frac{1}{\Delta} (P_{\theta+\Delta} - P_\theta)(h_{\theta+\Delta} - h_\theta) \right| \leq \frac{1}{|\Delta|} \cdot |(P_{\theta+\Delta} - P_\theta)h_{\theta+\Delta}| + \frac{1}{|\Delta|} \cdot |(P_{\theta+\Delta} - P_\theta)h_\theta|$$

For $|\Delta|$ sufficiently small, Lemma 11 yields

$$\left| \frac{1}{\Delta} (P_{\theta+\Delta} - P_\theta)(h_{\theta+\Delta} - h_\theta) \right| \leq \tilde{f}_0,$$

for $\tilde{f}_0 \in (H, 0)$, which implies

$$H_\theta \left( \sup_{|\Delta|} \frac{1}{\Delta} (P_{\theta+\Delta} - P_\theta)(h_{\theta+\Delta} - h_\theta) \right) < \infty \quad (37)$$

By (36) together with (37), Condition (B2) follows from the Dominated Convergence Theorem, which completes the proof. ■

### 5 Applications

In this section, we apply our results to performance characteristics of the G/G/1 queue. In the first example, we study the dependence of the overflow probability of a certain level in a busy cycle with $\theta$ a parameter of the service time distribution. In the second example, we consider the same performance measure but this time for the M/G/1 queue where $\theta$ is the intensity of the arrival stream. We will base the analysis of the second example on the thinning model of Poisson processes.
5.1 Service time dependence on $\theta$

Let $W_{\theta}(n)$ be the waiting time of the $n$th customer in an $G/G/1$ queue. Let \{\{A(n)\}\} be the i.i.d. sequence of interarrival times with finite expected value and let \{\{S_{\theta}(n)\}\} the i.i.d. sequence of service times, respectively. We assume that the system is stable, i.e., $\sup_{\theta \in \Theta} E[S_{\theta}(1)] < E[A(1)]$. Set

$$\xi_{\theta}(n) \triangleq S_{\theta}(n) - A(n)$$

and

$$\xi_{\theta}(n, w) = \max(w + \xi_{\theta}(n), 0),$$

for $n \geq 1$. Lindley’s recursion yields:

$$W_{\theta}(n + 1) = \max(W_{\theta}(n) + \xi_{\theta}(n), 0) = \xi_{\theta}(n, W_{\theta}(n)), \quad n \geq 1,$$

and $W_{\theta}(1) = 0$. Let $\alpha = \{0\}$ denote the event that the waiting times regenerate.

We assume that $S_{\theta}(n)$ follows a Pareto ($\theta, 2$) distribution, i.e., $\mathbb{P}(S_{\theta}(n) > x) = \theta^2 (\theta + x)^{-2}$. Then, $E[S_{\theta}(n)] = \theta$, for any $n$, and the variance of $S_{\theta}(n)$ fails to exist. For $U$ uniformly distributed on $[0, 1]$, a sample of $S_{\theta}(n)$ can be obtained by the inverse probability function through $\theta((1 - U)^{-\frac{1}{2}} - 1)$. From this construction it follows that $S_{\theta}(n)$ is monotone with respect to $\theta$ which in turn implies (S2).

Let $f_{\theta,k}$ with

$$f_{\theta,k}(x) = k \theta^k (\theta + x)^{k+1}.$$

denote the density of the Pareto ($\theta,k$) distribution. Take as $\mathcal{D} = (L^1(P_{\theta}; \Theta), 1)$ the set of integrable measurable mappings. Then, the Pareto distribution is $\mathcal{D}$-differentiable and for $k = 2$ and $g \in \mathcal{D}$, differentiating with respect to $\theta$ yields,

$$\frac{d}{d\theta} \int g(x)f_{\theta,2}(x) \, dx = 4 \int g(x)\theta / (\theta + x)^3 \, dx - 6 \int g(x)\theta^2 / (\theta + x)^4 \, dx$$

$$= \frac{2}{\theta} \left( \int g(x)f_{\theta,2}(x) \, dx - \int g(x)f_{\theta,3}(x) \, dx \right), \quad (38)$$

for $g \in \mathcal{D}$. Hence, the Pareto ($\theta,2$) distribution has $\mathcal{D}$-derivative $(2/\theta, \text{Pareto}(\theta, 2), \text{Pareto}(\theta, 3))$. Note that the Pareto ($\theta,3$) distribution has finite first and second moment. Moreover, the positive part of the $\mathcal{D}$-derivative of the Pareto ($\theta,2$) distribution is the Pareto ($\theta,2$) distribution itself.
We now apply Theorem 2 to \( \mathcal{D} = (\mathcal{L}^1(P_\theta; \Theta), 0) \) (the set of bounded performance measures). Condition (ii) requires \( \mathbb{E}[|\xi|] \) to be finite and \( \mathbb{E}[\xi] < 0 \), which is satisfied since \( S_\theta(n) \) and \( A(n) \) have finite expected values and \( \sup_{\theta \in \Theta} \mathbb{E}[S_\theta(1)] < \mathbb{E}[A(1)] \), by assumption. Moreover, \( \mathbb{P}(A(n) > x) > 0 \) for any \( x \geq 0 \) implies condition (iii), see Lemma 8. Condition (iv) follows from (38). It remains to be shown that for \( g \in (\mathcal{H}, 0) \) it holds that

\[
\sup_{x} \sup_{\theta \in \Theta_1} \left| P'_\theta H_\theta g(x) - P'_\hat{\theta} H_\theta g(x) \right| < \infty.
\]

To see this note that \( P^{+}_\theta = P_\theta \) and that \( P^{-}_\theta \) is the transition kernel with a Pareto \((\theta, 3)\) distributed service time. Hence,

\[
|P'_\theta H_\theta g(x) - P'_\hat{\theta} H_\theta g(x)| \leq 2\theta |P_\theta H_\theta g(x) - P_\theta H_\theta g(x)| + 2\theta |(P^{-}_\theta H_\theta g(x) - P^{-}_\hat{\theta} H_\theta g(x)|
\]

A direct application of Lemma 10 yields

\[
\sup_{x} \sup_{\theta \in \Theta_1} \left| P_\theta H_\theta g(x) - P_\theta H_\theta g(x) \right| \leq C.
\]

Moreover, applying Lemma 10 to the version of the kernel with a Pareto \((\theta, 3)\) distributed service time yields

\[
\sup_{x} \sup_{\theta \in \Theta_1} \left| P^{-}_\theta H_\theta g(x) - P^{-}_\hat{\theta} H_\theta g(x) \right| \leq \hat{C}.
\]

Hence, for a suitable neighborhood of \( \theta \), Theorem 2 applies and we obtain for any nonnegative and monotone cost function \( g \) out of \( \mathcal{D}' \):

\[
\frac{d}{d\theta} \mathbb{E} \left[ \tau_{\alpha, \theta} - 1 \sum_{n=1}^{\infty} g(W_\theta(n)) \right] = \sum_{k=0}^{\infty} P^k_\theta P'_\theta \sum_{l=0}^{\infty} P^l_\theta g,
\]

where \( P_\theta \) denotes the taboo kernel of the waiting times with taboo set \( \alpha = \{0\} \).

In order to write the expression on the right-hand side in terms of random variables, we introduce the following variant of the waiting time sequence. For \( j \in \mathbb{N} \), set

\[
W^{-}_\theta (j; n+1) = \max(W^{-}_\theta (j; n) + S_\theta(n) - A(n), 0), \quad n \neq j,
\]

with \( W^{-}_\theta (j; 1) = 0 \), and for \( j = n \), let

\[
W^{-}_\theta (j; n+1) = \max(W^{-}_\theta (j; n) + S^{-}_\theta(n) - A(n), 0), \quad n \geq 1,
\]
where $S_\theta(n)$ follows a Pareto ($\theta,3$) distribution. We denote the first time that 
$\{W_\theta(n)\}$ and $\{W_\theta^-(j;n)\}$ simultaneously hit $\alpha$ by $r_\alpha^\pm(n)$. Then it holds that
\[
\frac{d}{d\theta} \mathbb{E} \left[ \sum_{n=1}^{r_{\alpha,\theta}^+} g(W_\theta(n)) \right] = \frac{2}{\theta} \mathbb{E} \left[ \sum_{j=1}^{r_{\alpha,\theta}^+} \sum_{n=j+1}^{r_{\alpha,\theta}^+} (g(W_\theta(n)) - g(W_\theta^-(j;n))) \right],
\]
for details see [6]. For example, letting, for some $b > 0$, $g_b(x) = 1$ for $x \geq b$ and zero otherwise, $H_\theta g_b$ is the expected number of overflows of level $b$ in a busy cycle. Note that, on the one hand, $S_\theta(n)$ fails to have a finite second moment which implies that $r_{\alpha,\theta}$ fails to have a finite second moment too. On the other hand, $r_{\alpha,\theta}$ needs to have a finite first moment for the cycle cost to exists. Hence, the key condition for applying Theorem 2 is the existence of the cycle cost.

### 5.2 Thinning of a Poisson Process

In this section we consider the waiting time of the $n$th customer in an $G/G/1$ queue in a slightly different setting. Let $\{A(n)\}$ be an i.i.d. sequence exponential distributed random variables with rate $\lambda$ constituting the interarrival times and let $\{S(n)\}$ be the i.i.d. sequence of service times, respectively. We introduce an i.i.d. sequence of $\{0,1\}$ random variables $\{\eta_\theta(n)\}$ with distribution $P(\eta_\theta = 1) = \theta = 1 - P(\eta_\theta = 0)$, for $\theta \in \Theta = [0,1]$. Set
\[
\xi_\theta(n) \triangleq \eta_\theta(n)S(n) - A(n) \quad \text{and} \quad \xi_\theta(n,w) = \max(w + \xi_\theta(n),0),
\]
for $n \geq 1$. Note that for $\theta \in [0,1]$ it holds
\[
\xi_\theta(n) \leq \xi_1(n) \triangleq S(n) - A(n),
\]
for $n \geq 1$, and that (S2) holds. Lindley’s recursion yields:
\[
W_\theta(n+1) = \max(W_\theta(n) + \eta_\theta(n)S(n) - A(n),0) = \xi_\theta(n,W_\theta(n)), \quad n \geq 1,
\]
and $W_\theta(1) = 0$. We assume that the system is stable for any $\theta \in [0,1]$, i.e., $\mathbb{E}[S_\theta(1)] < \mathbb{E}[A(1)]$.

The above model has the following interpretation. Customers arrive according to a Poisson-$\lambda$-process at the queue. An arriving customer is admitted to
the queue with probability $1 - \theta$. The total number of admitted customers out of the first $n$ arriving customers after the initial one is

$$m(n) \triangleq \sum_{k=1}^{n} \eta_{\theta}(k).$$

From this construction it follows that $W_{\theta}(n + 1)$ is the waiting time of the $(m(n) + 1)$st customer in a single-server queue with Poisson-$\lambda$-arrival stream. The above thinning of Poisson process yields again a Poisson process but with intensity $\lambda \theta$. Hence, $W_{\theta}(n + 1)$ can also be interpreted as the waiting time of the $(m(n) + 1)$st customer in single-server queue with Poisson-\(\lambda \theta\)-arrival stream.

Let $\alpha = \{0\}$ denote the event that the waiting times regenerate and take as

$$D = (L_{\theta}^{1}(P_{\theta}; \Theta), p).$$

Provided that $S(n)$ and $A(n)$ have finite $p$th moments it holds for any $g \in D$

$$\frac{d}{d\theta} \mathbb{E}[g(\xi_{\theta}(n))] = \frac{d}{d\theta} \left( \mathbb{E}[g(\xi_{\theta}(n))|\eta_{\theta}(n) = 1] \theta + \mathbb{E}[g(\xi_{\theta}(n))|\eta_{\theta}(n) = 0](1 - \theta) \right)$$

$$= \frac{d}{d\theta} \left( \mathbb{E}[g(S(n) - A(n))]|\theta + \mathbb{E}[g(-A(n))](1 - \theta) \right)$$

$$= \mathbb{E}[g(S(n) - A(n))]| - \mathbb{E}[g(-A(n))].$$

Since the right-hand side of the above equation is independent of $\theta$ and finite, it holds that

$$\sup_{\theta \in [0, 1]} \left| \frac{d}{d\theta} \mathbb{E}[g(\xi_{\theta}(n))] \right| < \infty.$$

It remains to be shown that for $g \in D$ it holds that

$$\sup_{x} \sup_{\theta \in \Theta} \left| P_{\theta}^{r}H_{\theta}g(x) - P_{\hat{\theta}}^{r}H_{\hat{\theta}}g(x) \right| < \infty.$$

To see this note that $P_{\theta}^{r} = (P_{\hat{\theta}} - P_{\theta})$ is independent of $\theta$, which gives for all $x \in \mathbb{R}_{+}$ and all $\theta, \hat{\theta} \in [0, 1]$ that

$$\left| P_{\theta}^{r}H_{\theta}g(x) - P_{\hat{\theta}}^{r}H_{\hat{\theta}}g(x) \right| = 0$$

From the above it is straightforward to see that Theorem 2 applies to any $g \in D$ provided that $S(n)$ and $A(n)$ have finite $(p + 1)$st moments. Specifically, for $j \in \mathbb{N}$, set

$$W_{\theta}^{+}(j; n + 1) = \max(W_{\theta}^{+}(j; n) + \eta_{\theta}S(n) - A(n), 0), \quad n \neq j,$$

with $W_{\theta}^{+}(j; 1) = 0$, and for $n = j$

$$W_{\theta}^{+}(j; n + 1) = \max(W_{\theta}^{+}(j; n) + S(n) - A(n), 0),$$
and define the ‘−’ version by

\[ W^-_\theta(j; n + 1) = \max(W^-_\theta(j; n) + \eta_0 S(n) - A(n) - 0), \quad n \geq 1, \]

with \( W^-_\theta(j; 1) = 0 \) and for \( j = n \)

\[ W^-_\theta(j; n + 1) = \max(W^-_\theta(j; n) - A(n) - 0). \]

We denote the first time that \( \{W^+_\theta(j; n)\} \) and \( \{W^-_\theta(j; n)\} \) simultaneously enter \( \alpha \) by \( \tau_{\alpha, \theta}(j) \), with \( \alpha = \{0\} \). Then, it holds that

\[
\frac{d}{d\theta} \mathbb{E} \left[ \sum_{n=1}^{\tau_{\alpha, \theta}-1} g(W^\theta(n)) \right] = \mathbb{E} \left[ \sum_{j=1}^{\tau_{\alpha, \theta}-1} \sum_{n=j+1}^{\tau_{\alpha, \theta}(j)-1} (g(W^\theta_{+}(j; n)) - g(W^\theta_{-}(j; n))) \right],
\]

for any \( g \in \mathcal{D}^p \), for details see [6].

For example, taking \( p = 0 \), it is sufficient that \( S(n) \) and \( A(n) \) have finite first moment. Again, our result applies in the case that \( S(n) \) has no finite second moment which in turn implies that \( \tau_{\alpha, \theta} \) has no finite second moment. In [8] finiteness of the second moment of \( \tau_{\alpha, \theta} \) is required which is an improvement on [1] where even the third moment has to be finite. For a first study of this problem we refer to [4].

For example, letting, for some \( b > 0 \), \( g_b(x) = 1 \) for \( x \geq b \) and zero otherwise,

\[ \mathbb{E} \left[ \sum_{n=1}^{\tau_{\alpha, \theta}-1} g_b(W^\theta(n)) \right] \]

is the expected number of overflows of level \( b \) in a busy cycle.

References


