

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CONVERGENCE RATES FOR STATIONARY DISTRIBUTIONS OF
SEMISTOCHASTIC PROCESSES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

JAMES BRODA
Norman, Oklahoma
2017

CONVERGENCE RATES FOR STATIONARY DISTRIBUTIONS OF
SEMISTOCHASTIC PROCESSES

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF MATHEMATICS

BY

Dr. Nikola Petrov, Chair

Dr. Alexander Grigo, Co-Chair

Dr. Stephen Ellis

Dr. John Albert

Dr. Ying Wang

© Copyright by JAMES BRODA 2017
All Rights Reserved.

Acknowledgements

I would like to thank my advisors, Nikola Petrov and Alex Grigo, for teaching me math and not letting me quit. The countless hours they spent with me went far above and beyond the call of duty. Their friendship has been as important as their knowledge of mathematics in helping me through graduate school.

I would also like to thank all of my friends and family for their support and patience. Many thanks to my mom, Danny, Caryn, Frank, Amber, Zara, Zoe, Zeah, Michael, and Jacob. Thank you to Fıstık for sharing both the sunny days and the rainy ones. Thank you to Papa Mike for being an inspiration and always encouraging me to follow my dreams.

I am also appreciative of the mathematics community at the University of Oklahoma. I am especially grateful to the Dynamical Systems group (past and present); Bryan, Lee, Sean, Dania, Garret, and Mahesh.

Finally, I would like to thank all of the teachers who shared with me their joy in learning; you are too many to name and I thank you all.

Contents

1	Introduction	1
1.1	An Ecological Problem	1
1.2	Semistochastic Processes	2
2	Stochastic Processes	8
2.1	General Stochastic Processes	8
2.2	Markov Processes and Martingales	10
2.3	Poisson Processes	12
3	Survival and Hazard Functions	16
3.1	Introduction to Survival Analysis	16
3.2	Semistochastic Survival Function and Memorylessness	18
4	Generators and Semigroups	21
4.1	Markov Semigroups and Infinitesimal Generators	21
4.2	Generator of the Semistochastic Process	24
4.3	Dissection of the Generator and Expansion of the Semigroup	28
5	Minorization and Convergence Rates	32
5.1	Discretization of Continuous Processes	32
5.2	Minorization, Drift, and Coupling	35
5.3	Convergence Rates for Semistochastic Processes	47
5.3.1	Examples	50
6	Carbon Content and other Scalar Models	56
6.1	Post-Disturbance Distributions	58
6.2	Exact Formula for Stationary Distributions	60
6.2.1	Examples	62

List of Figures

1.1	Simulation of a semistochastic process.	4
1.2	Schematic for pre- and post- disturbance levels.	5
1.3	Jump kernel.	6
5.1	On the construction of the minorizing measure in Theorem 5.5.	39
5.2	Plot of $(1 - \epsilon_{\Delta t})^{1/\Delta t}$ vs. Δt	52
5.3	Plots of $(1 - \epsilon_{\Delta t})^{\lfloor t/\Delta t \rfloor}$ vs. t for selected Δt	53
5.4	Plots of $(1 - \epsilon_{\Delta t})^{r/\Delta t}$ vs. Δt for selected κ	54
5.5	Plots of $(1 - \epsilon_{\Delta t})^{r/\Delta t}$ vs. κ for selected Δt	55

Abstract

The primary objects of study in this dissertation are semistochastic processes. The types of semistochastic processes we consider are continuous-time and continuous-state processes consisting of intervals of deterministic evolution punctuated by random disturbances of random severity. A natural question regarding such processes is whether they admit stationary distributions. While partial answers to this question exist in the literature, the primary aim of this dissertation is to supplement the criteria for existence with bounds on convergence rates. This requires careful analysis of the associated Markov semigroups and infinitesimal generators. We obtain our bounds on convergence rates by establishing minorization and drift conditions. Specific examples are considered in cases of bounded and unbounded state spaces. We also discuss a method of exact computation for the stationary distributions of a certain class of semistochastic processes. An important example to which we can apply our work concerns the modelling of the carbon content of an ecosystem.

Chapter 1

Introduction

1.1 An Ecological Problem

The research contained in this dissertation began due to an ecological problem: How does the carbon content of an ecosystem evolve due to randomly occurring catastrophes of random severity? The role of disturbances such as droughts, forest fires, and insect outbreaks on the dynamics of carbon has been discussed in [38], [32], [5], and [37]. In the absence of disturbances, the amount of carbon in an ecosystem increases naturally due to photosynthesis and eventually approaches the carrying capacity of the ecosystem. On occasion, however, an extreme event results in significant destruction of an ecosystem and consequently a drastic reduction in the amount of carbon stored in the ecosystem.

In order to model the carbon content of an ecosystem, semistochastic processes are studied in [24] and formulae derived for the densities of the corresponding stationary distributions. Given that a stationary distribution exists, one would like to know the rate at which the process approaches this equi-

librium. In this dissertation, we answer this question by establishing bounds on convergence rates; we also generalize the results from [24] to allow for the study of the dynamics of a broader range of semistochastic processes. To this end, we establish a more general notion of semistochastic processes, similar to what is studied by Gripenberg in [14], and extend the results from [24] in Chapter 6.

In this dissertation, we utilize purely probabilistic methods to establish bounds on convergence rates; consequently, our methods for determining convergence rates, are quite different from the methods used in [24] to develop exact formula for the stationary distributions. The methods we use have their origin in the study of discrete-time Markov chains and are based on establishing a combination of minorization and drift conditions. These approaches go back to Doeblin, and appear in various forms in [25], [29], and [36, 34]. While the problem of modelling the carbon content of an ecosystem is the original inspiration for this project, our work can be applied to any problem admitting a semistochastic model. Examples include problems from population dynamics, optimal harvesting, and evaluating the risk of developing rabies.

1.2 Semistochastic Processes

What follows is a brief introduction to the concept of a semistochastic process; a more detailed exposition of the general theory of stochastic processes is presented in Chapter 2. By semistochastic process we mean a continuous-time, continuous-state process $\{X_t\}$ which consists of intervals of deterministic evolution punctuated by random events. The random events we typically consider occur on time-scales much larger than the typical inter-event time, and can

be modeled as instantaneous events. These processes are also assumed to be doubly-stochastic in the sense that there is a random severity associated to each event as well as the random time at which it occurs. Consequently, these types of processes are quite different from other types of stochastic processes and can be used to model dynamical systems that lack conservation laws, see [13, 35]. Semistochastic processes do share some common features with what are typically referred to as stochastic clearing processes (see [39]). A clearing process, however, consists of epochs of random growth punctuated by instantaneous returns to the initial value once a critical threshold is reached. A semistochastic process replaces the random growth in a clearing process with deterministic growth and replaces the deterministic “clearing” with randomly occurring disturbances.

The operator-theoretic framework which we set up to study the dynamics of semistochastic processes applies equally well to both scalar- and vector-valued stochastic processes, but we restrict our attention to scalar processes when deriving bounds on convergence rates. In the scalar case we are thus interested in sample paths that are piecewise continuous, right-continuous, and have left-hand limits almost surely (*càdlàg*). We furthermore limit our attention to disturbances that on average correspond to a diminishing in value.

In the time between two consecutive disturbances $\{X(t)\}$ evolves deterministically, governed by the autonomous ordinary differential equation

$$\frac{d}{dt}x(t) = v(x(t)) . \tag{1.1}$$

To describe when the disturbances occur, one typically specifies a hazard function, which is a measure of the instantaneous rate of occurrence of the distur-

bances. If the disturbances come from a state-independent, time-homogeneous Poisson process then the resulting hazard function is constant and is equal to the standard intensity parameter for the underlying Poisson process. To allow for greater generality, we consider state-dependent hazard functions in this dissertation and define the corresponding hazard functions in subsequent sections.

A simulation of a one-dimensional semistochastic process with logistic deterministic growth and exponentially distributed inter-disturbance times can be seen in Figure 1.1.

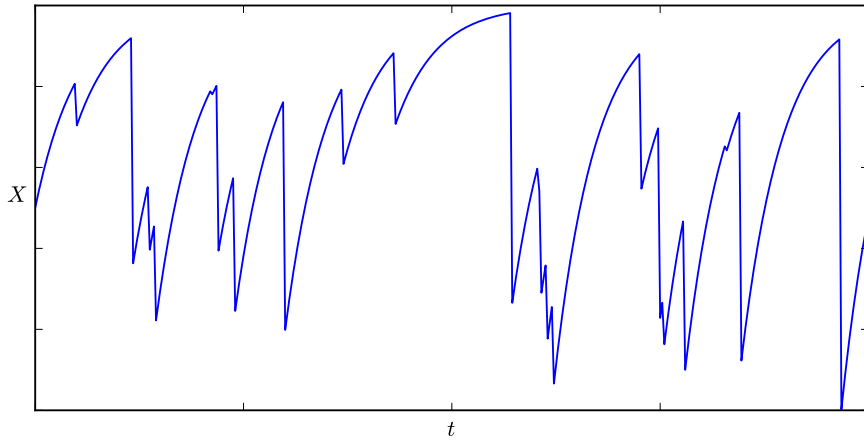


Figure 1.1: Simulation of a semistochastic process.

Our problem gains another element of randomness from the varying severity of the disturbances. In order to describe this severity, we introduce random variables Y_n^- and Y_n corresponding to the n^{th} pre- and post- disturbance values, respectively. If the n^{th} disturbance occurs at time T , then Y_n^- and Y_n are defined via

$$Y_n^- := \lim_{t \nearrow T} X(t), \quad Y_n := \lim_{t \searrow T} X(t) .$$

In the simplest case, we can model the severity by stipulating a multiplicative relation between Y_n^- and Y_n . An additional random variable, A_n is defined by setting

$$Y_n = A_n Y_n^- .$$

Alternatively, we can model the severity by specifying a function $P(x, A)$,

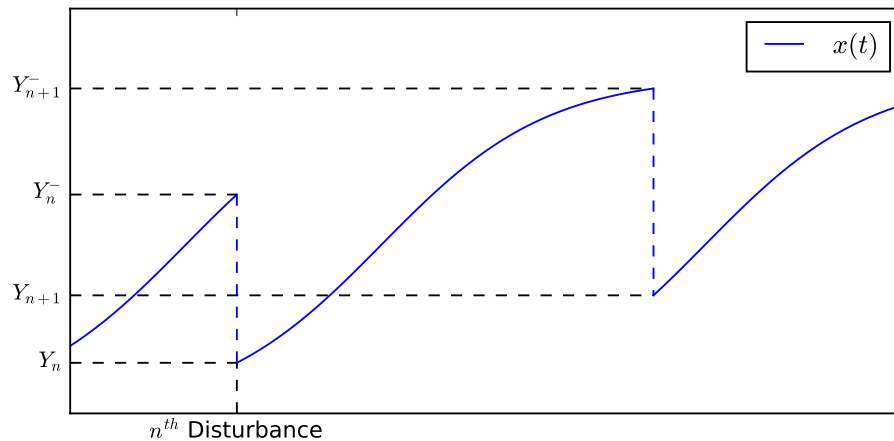


Figure 1.2: Schematic for pre- and post- disturbance levels.

which we call the *jump kernel* according to

$$P(x, A) = \mathbb{P}(Y_n \in A \mid Y_n^- = x) .$$

Having specified the types of processes we propose to study, we now mention some works that study similar processes, but usually under different assumptions or with different goals. The most common difference is due to the fact that most of the research on semistochastic processes is concerned with population dynamics, and demographers generally study processes with discrete state-spaces.

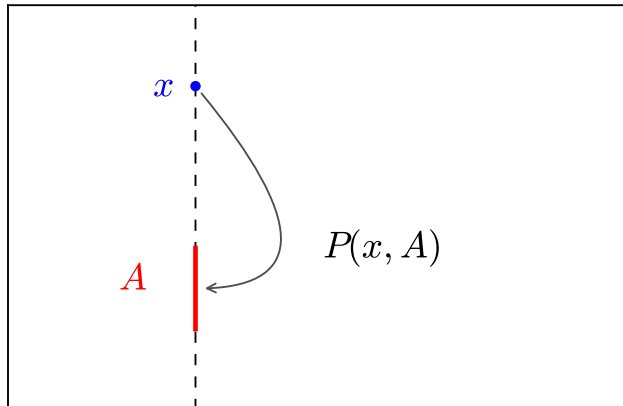


Figure 1.3: Jump kernel.

An interesting application of semistochastic processes is proposed by Bartoszyński in [3] to model the development of the rabies virus in an infected host. In this model, the population of the virus naturally decreases exponentially due to the immunological response of the host, but also has random upward jumps due to the viral life cycle. The state space of the model Bartoszyński constructs is discrete and the occurrence of jumps is allowed to depend on the current population.

Continuous-time and continuous state space processes subject to random catastrophes are studied by Gripenberg in [14]. Gripenberg derives an expression for stationary distributions using a limit theorem from [1] based on the concept of Harris recurrence. There is a connection between the type of recurrence condition that is established in [14] and the minorization conditions that we establish, however the issue of convergence rates is not addressed by Gripenberg.

Hanson and Ryan in [16] and [17] examine optimal harvesting problems of populations governed by similar processes with discrete state spaces, though

they do not allow for the randomization of the intensity of disturbances. They do, however, consider the possibility of populations experiencing both sudden decreases (jumps down) and sudden increases (jumps up). With slight modifications, the results we present can also be applied in these situations. Hanson and Tuckwell also study similar processes with discrete state spaces in [18, 19, 20], though their focus is generally on the computation of extinction times. The problem of determining extinction times in semistochastic models is addressed more recently by Cairns in [8].

Our present work builds upon the work of Leite, Petrov, and Weng [24] in which an integral equation for the stationary post-disturbance distribution of a continuous-time, continuous state space semistochastic process is presented. A framework for computing the overall stationary distributions for scalar-valued semistochastic processes with constant hazard functions is also established in [24]; we extend this result to state-dependent hazard functions in Chapter 6. The issue of convergence rates, however, is not addressed in [24], and except for those of Chapter 6, the methods used in this dissertation are rather distinct from those of [24].

Chapter 2

Stochastic Processes

What follows is a review of some of the basic definitions and properties related to stochastic processes in general, and Markov and Poisson processes in particular. For additional references see [21] or [31].

2.1 General Stochastic Processes

We start by defining some of the basic terminology from probability theory. A random experiment is any experiment that could in principle be repeated under the same conditions and has an uncertain outcome; the set of all elementary outcomes of the experiment is often called the *sample space*, typically denoted Ω . Subsets of the sample space are called *events*, and for certain events we can measure the probability of their occurring. Given events A and B , we denote the probability of A occurring as $\mathbb{P}(A)$ and the conditional probability of A given B as $\mathbb{P}(A|B)$. A probabilistic model is based on a triple $(\Omega, \mathcal{F}, \mathbb{P})$ consisting of a sample space Ω , a σ -algebra \mathcal{F} , and a probability measure \mathbb{P} . We denote the expectation of a random variable, X , by $\mathbb{E}[X]$. Our primary objects

of study are particular instances of stochastic processes, which we define now.

Definition 2.1. A *stochastic process* parameterized by T with state space \mathcal{X} is a collection

$$\{X(t) : t \in T\}$$

of \mathcal{X} -valued random variables on a sample space Ω . If T is the real line, or an interval on the real line, then it is said to be a continuous-time process. If T is a sequence of integers, it is called a discrete-time process.

An element of a stochastic processes can be written as either $X(t)$ or X_t and is a random variable for any fixed value of t . As a random variable $X(t)$ is a function on the sample space Ω , however, the convention is to suppress the dependence on ω . For a fixed outcome ω , the function $t \mapsto X(t)$ is referred to as a *realization* of the stochastic process. The semistochastic processes studied in this dissertation can be thought of as a special case of continuous-time stochastic processes with parameter $T = [0, \infty)$; we furthermore assume the state-space is topological. A probability measure π on (Ω, \mathcal{F}) is said to be an *invariant measure* or *stationary distribution* for the stochastic process $\{X(t)\}$ if π is invariant under the dynamics induced by $\{X(t)\}$. It is convenient to use the standard shorthand for conditional expectations,

$$\mathbb{E}_x[X_t] = \mathbb{E}[X_t | X_0 = x].$$

For any initial probability distribution μ , $\mathbb{E}_\mu[X_t]$ is defined in an analogous manner.

2.2 Markov Processes and Martingales

In order to define a Markov process, we first introduce the concept of a filtration. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a *filtration* is a family $\{\mathcal{F}_t : t \in T\}$ of σ -algebras such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s, t \in T$ with $s \leq t$. A filtration is *adapted* to a stochastic process $\{X(t) | t \in T\}$ if each $X(t)$ is \mathcal{F}_t -measurable. The *natural* filtration with respect to $\{X_t\}$ is defined by setting $\mathcal{F}_t = \sigma\{X_s^{-1}(B) : s \leq t, B \in \mathcal{B}\}$, with \mathcal{B} being the standard Borel σ -algebra on the state space. We are now ready to define the Markov property.

Definition 2.2. *Let $(\mathcal{X}, \mathcal{B})$ be a measure space and $X(t)$ an \mathcal{X} -valued stochastic process with filtration $\{\mathcal{F}_t\}$. The process $\{X(t)\}$ is said to have the **Markov property** if for any $B \in \mathcal{B}$ and for each $s, t \in T$ with $s \leq t$, we have*

$$\mathbb{P}(X_t \in B | \mathcal{F}_s) = \mathbb{P}(X_t \in B | X_s) .$$

A Markov process can be defined as a stochastic process possessing the Markov property with respect to its natural filtration. This property is commonly referred to as *memorylessness* due to the interpretation that the probability of arriving at some future state only depends on the current state. In case the parameter space T is discrete such a process is called a *Markov chain*. To describe the evolution of a Markov process, one can specify its transition kernel.

Definition 2.3. *Given a Markov process $\{X_t\}$ with state space $(\mathcal{X}, \mathcal{A})$, a map $Q : \mathcal{X} \times \mathcal{A} \mapsto [0, 1]$ is a **transition kernel** if*

- (i) *for each fixed $x \in \mathcal{X}$, the map $A \mapsto Q(x, A)$ is a probability measure on $(\mathcal{X}, \mathcal{A})$, and*

(ii) for each fixed $A \in \mathcal{A}$, the map $x \mapsto Q(x, A)$ is \mathcal{A} -measurable.

In the discrete-time case, the transition kernel has the natural interpretation that $Q(x, A)$ is the probability of moving from initial state $X_0 = x$ into A in one step. It is also important to understand the way in which Q interacts with real-valued functions and distributions on \mathcal{X} . We refer to a real-valued function $f(x)$ on \mathcal{X} as an *observable* and define the action of Q on observables via

$$[Qf](x) = \int Q(x, dy) f(y) .$$

We also define the left-action of Q on distributions via

$$[\mu Q](A) = \int \mu(dx) Q(x, A)$$

One can describe a stationary distribution π as one for which $\pi Q = \pi$. For Markov chains which are aperiodic and irreducible, stationary distributions are also limiting distributions, that is

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n \in A) = \pi(A)$$

for all measurable sets A . The types of stationary distributions we study in this dissertation are also limiting distributions in the above sense.

Another distinguished type of stochastic process is a martingale.

Definition 2.4. A stochastic process $\{X_t\}$ is called a ***martingale*** relative to a filtration $\{\mathcal{F}_t\}$ if

(i) $\{X_t\}$ is adapted to $\{\mathcal{F}_t\}$,

(ii) $\mathbb{E}[|X_t|] < \infty$ for all t ,

(iii) $\mathbb{P}(\mathbb{E}[X_t|\mathcal{F}_s] = X_s) = 1$ for all $s < t$.

Loosely speaking, martingales are models of fair games, where the expected value $\mathbb{E}[X_t]$ at any time t is always equal to the initial expected value $\mathbb{E}[X_0]$. We use a basic martingale to help establish drift conditions in Chapter 5.

2.3 Poisson Processes

A Poisson process is a type of stochastic process that is useful for modeling the times at which an event of interest occurs; we use properties of Poisson processes to derive an expression for the infinitesimal generator of a semistochastic process. Before defining a Poisson process, we introduce the more general notions of arrival and renewal processes.

An *arrival process* is an increasing sequence of random variables, $0 < \Theta_1 < \Theta_2 < \dots < \Theta_n$, where $\Theta_i < \Theta_j$ is meant to indicate that the random variable $\Theta_j - \Theta_i$ obtains only positive values. The random variables X_i are called *arrival epochs* and are meant to represent the times at which some repeating event occurs (arrives). Arrival processes are specified by the corresponding *interarrival time process* $\{T_n : n \in \mathbb{N}\}$ or alternatively by a *counting process* $\{N(t)|t \in [0, \infty)\}$ where each $N(t)$ measures the number of arrivals in the interval $[0, t)$.

A general *renewal process* is defined as an arrival process for which the interarrival times are independent and identically distributed random variables. A Poisson process is a special case of a renewal process that can be defined by either specifying the corresponding interarrival time distribution or by specifying the underlying counting process.

Definition 2.5. A **Poisson process** with rate parameter λ is a renewal process with interarrival times T_n that are independently and exponentially distributed with common density given by

$$f_T(t) = \lambda e^{-\lambda t}$$

for $t \geq 0$.

What distinguishes the Poisson process from arbitrary renewal processes is the memoryless property of the interarrival time random variables. A non-negative, non-deterministic random variable T is said to be *memoryless* if

$$\mathbb{P}(T > s + t | T > s) = \mathbb{P}(T > t) ,$$

For Poisson processes this is a simple consequence of the arithmetic of exponential distributions. An alternative characterization of a Poisson process as a counting process relies on the concepts of stationary and independent increments. In order to facilitate their definitions, let $\tilde{N}(s, t)$ denote the random variable $N(t) - N(s)$ which for $s < t$ counts the number of arrivals in the interval $(s, t]$.

Definition 2.6. A counting process $\{N(t)\}$ is said to have **independent increments** if for any finite sequence of times $t_1 < t_2 < t_3 \dots < t_n$ the random variables $N(t_1), \tilde{N}(t_1, t_2), \dots, \tilde{N}(t_{n-1}, t_n)$ are independent.

A counting process $\{N(t)\}$ is said to have **stationary increments** if for every $s, t \in (0, \infty)$ with $s < t$, the random variables $\tilde{N}(s, t)$ and $N(t - s)$ have the same distribution.

We are now ready to present the definition of a Poisson counting process.

Definition 2.7. A *Poisson counting process* $\{N(t)\}$ with rate parameter λ is a counting process with independent and stationary increments that has a probability mass function of the form

$$p_{N(t)}(n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}. \quad (2.1)$$

For each fixed t , the distribution in (2.1) is called a *Poisson distribution* and has mean value $\mathbb{E}[N_t] = \lambda t$. For a Poisson counting process with constant rate λ , we have the following estimates of arrival probabilities over a time interval Δt that is sufficiently close to zero.

$$\mathbb{P}(\text{one arrival}) = \lambda \Delta t + o(\Delta t)$$

$$\mathbb{P}(\text{no arrivals}) = 1 - \lambda \Delta t + o(\Delta t)$$

$$\mathbb{P}(\text{more than one arrival}) = o(\Delta t)$$

One can also define a non-homogeneous Poisson process with time-varying rate parameter $\Lambda(t)$.

Definition 2.8. A *non-homogeneous Poisson counting process*, $\{N(t)\}$ with intensity $\lambda(t)$ is a counting process with independent increments that satisfies the conditions:

(i) $N(0) = 0$,

(ii) For each $t > 0$, $N(t)$ has a Poisson distribution with mean

$$m(t) = \int_0^t \lambda(s) ds .$$

To facilitate the handling of non-homogeneous Poisson processes, we introduce concepts from Survival Analysis in the next chapter.

Chapter 3

Survival and Hazard Functions

In this chapter we review concepts from the field of Survival Analysis and define semistochastic versions of survival and hazard functions.

3.1 Introduction to Survival Analysis

Survival Analysis (or Reliability Analysis) is concerned with the time duration of some process before a random event (perhaps death or failure) occurs. Standard references include [11] and [27]. In the context of our semistochastic problem, we are concerned with the determination of the distribution of inter-disturbance times. First, we review some of the terminology commonly used in Survival Analysis and the corresponding mathematical definitions. We begin by considering a real-valued random variable, T , that records the time of occurrence of some event of interest.

Definition 3.1. *Given a random variable T with corresponding cumulative distribution function F_T , the **survival function**, $S(t)$, is defined as the prob-*

ablitic complement of F_T ,

$$S(t) = \mathbb{P}(T > t) = 1 - F_T(t) .$$

If T is a continuous random variable with density f_T , then $S(t)$ is given by

$$S(t) = \int_t^\infty f_T(s) ds . \tag{3.1}$$

The survival function thus measures the probability of arriving at time t with no occurrence of an event corresponding to the random variable T . We are also interested in the hazard function.

Definition 3.2. Given a random variable T that records the time of occurrence of some event of interest, the corresponding **hazard function**, $\lambda(t)$, is often defined as

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{P}(t < T \leq t + \Delta t)}{\Delta t} .$$

The hazard function is thus a measure of the instantaneous rate of occurrence of the events described by the random variable T . If T is a continuous random variable with density f_T , then we can give an alternative, and equivalent, though somewhat less transparent, representation of the $\lambda(t)$:

$$\lambda(t) = \frac{f_T(t)}{S(t)} .$$

Indeed, from (3.1) we have that $\frac{d}{dt} S(t) = -f_T(t)$, which leads to the differential equation

$$\lambda(t) = -\frac{d}{dt} \log(S(t)) .$$

Solving for $S(t)$ and using the implied initial data $S(0) = 1$ yields

$$S(t) = \exp \left(- \int_0^t \lambda(s) ds \right) . \quad (3.2)$$

This integral representation of the survival function is useful when studying a semistochastic process with a non-homogenous Poisson process driving the disturbances.

3.2 Semistochastic Survival Function and Memorylessness

For use in later sections, we define and investigate the relationships between semistochastic hazard functions, survival functions, and inter-disturbance time distributions for the semistochastic process $\{X_t\}$. In particular, we establish the memorylessness of the inter-disturbance time random variables T_n . For a semistochastic process, we have a deterministic flow $\phi^t(x)$ as well as a state-dependent rate parameter $\Lambda(x)$ that can be recast as a hazard function by composing with the flow,

$$\lambda(t, x_0) := \Lambda(\phi^t(x_0)) .$$

For fixed initial value x_0 we have a time-dependent hazard function as in the previous section. Referring to (3.2) and emphasizing the dependence on the initial value x_0 , our survival function is given by

$$S(t, x_0) = \exp \left(- \int_0^t \Lambda(\phi^s(x_0)) ds \right) \quad (3.3)$$

Let T denote the random variable recording the amount of time until the next disturbance. To show that T is memoryless requires the verification of

$$\mathbb{P}(T > t + s | T > s) = \mathbb{P}(T > t) \quad \text{for all } t, s > 0 .$$

Using the definition of conditional probability and the fact that $t + s > t$, this is equivalent to verifying that

$$\mathbb{P}(T > t + s) = \mathbb{P}(T > t) \mathbb{P}(T > s).$$

In terms of the semistochastic survival function $S(t, x)$ this is equivalent to

$$S(t + s, x_0) = S(s, \phi^t(x_0)) S(t, x_0) .$$

Indeed,

$$\begin{aligned} S(s, \phi^t(x_0)) S(t, x_0) &= \exp\left(-\int_0^s \Lambda(\phi^u(\phi^t(x_0))) du\right) \exp\left(-\int_0^t \Lambda(\phi^\tau(x_0)) d\tau\right) \\ &= \exp\left(-\int_0^s \Lambda(\phi^{u+t}(x_0)) d\tau\right) \exp\left(-\int_0^t \Lambda(\phi^\tau(x_0)) d\tau\right) \\ &= \exp\left(-\int_s^{t+s} \Lambda(\phi^\tau(x_0)) d\tau\right) \exp\left(-\int_0^t \Lambda(\phi^\tau(x_0)) d\tau\right) \\ &= \exp\left(-\int_s^t \Lambda(\phi^\tau(x_t)) d\tau - \int_0^t \Lambda(\phi^\tau(x_0)) d\tau\right) \\ &= \exp\left(-\int_0^{t+s} \Lambda(\phi^\tau(x_0)) d\tau\right) \\ &= S(t + s, x_0). \end{aligned}$$

Thus the fact that the semistochastic survival function is memoryless is essentially a consequence of semigroup property of the deterministic flow. The

significance of the memorylessness is that it allows us to study semistochastic process using tools designed for Markov processes.

Chapter 4

Generators and Semigroups

In this chapter we introduce the concepts of Markov semigroups and their associated infinitesimal generators. We develop a description of semistochastic processes in terms of their infinitesimal generators and establish a useful expansion formula for their semigroups.

4.1 Markov Semigroups and Infinitesimal Generators

In this section, we establish an operator-theoretic framework for studying a wide class of semistochastic processes. To understand the dynamics of a time-homogeneous Markov process, $\{X_t\}$, we study the associated Markov semigroup which we denote by \mathcal{U}^t .

Definition 4.1. *Let $\{X_t\}$ be a time-homogeneous Markov process with state space \mathcal{X} . The corresponding **Markov semigroup**, \mathcal{U}^t , is determined by its*

action on observables,

$$[\mathcal{U}^t f](x) = \mathbb{E}[f(X_t)|X_0 = x] ,$$

where f is any continuous and bounded function on \mathcal{X} that vanishes at infinity, we denote this function space by $C_0(\mathcal{X})$.

We can also specify the left-action of \mathcal{U}^t on a distribution μ by

$$[\mu \mathcal{U}^t](f) = \int [\mathcal{U}^t f](x) d\mu(x)$$

for all observables f . In this context, stationary distributions of a Markov process $\{X_t\}$ are characterized by their invariance under \mathcal{U}^t . Consequently, π being a stationary distribution is equivalent to $\pi \mathcal{U}^t = \pi$. If the Markov semigroup, \mathcal{U}^t , is strongly continuous on $C_0(\mathcal{X})$ (where we view $C_0(\mathcal{X})$ as a Banach space endowed with the supremum norm), then we can describe its infinitesimal generator, which we denote by \mathcal{L} , through its action on an appropriate subset of observables.

Definition 4.2. *Let $\{X_t\}$ be a Markov process with state space \mathcal{X} and semigroup, \mathcal{U}^t . The action of the infinitesimal generator of \mathcal{U}^t on observables is given by*

$$[\mathcal{L}f](x) = \lim_{t \searrow 0} \frac{\mathcal{U}^t f(x) - f(x)}{t} ,$$

where the domain of \mathcal{L} consists of all $f \in C_0(\mathcal{X})$ such that $\lim_{t \searrow 0} \frac{\mathcal{U}^t f(x) - f(x)}{t} \in C_0(\mathcal{X})$.

Under reasonable assumptions, the semistochastic processes which we con-

sider are actually Feller processes, and thus their associated semigroups is strongly continuous and hence has a corresponding infinitesimal generator. With regard to the potential restrictions on the domain of \mathcal{L} , these will be suppressed and the term observable is used both to represent an arbitrary element from the appropriate domain of either \mathcal{L} or \mathcal{U}^t as is appropriate.

Once one has access to the infinitesimal generator of a Markov process, then for any function f in the domain of \mathcal{L} , one can form a martingale with respect to the natural filtration for $\{X_t\}$ by defining

$$M_t := f(X_t) - f(X_0) - \int_0^t [\mathcal{L}f](X_s) ds \quad (4.1)$$

for $t \geq 0$. For this particular martingale, we have $\mathbb{E}[M_t | X_0] = 0$ for all $t \geq 0$. We make use of this basic martingale in the next chapter.

Further analysis of the infinitesimal generator and its adjoint gives rise to partial differential equations that provide another perspective for studying the underlying Markov process. Given \mathcal{U}^t and \mathcal{L} , if we fix an observable f and denote $[\mathcal{U}^t f](x)$ by $u(t, x)$, then $u(0, x) = f(x)$, and $u(t, x)$ must satisfy

$$\frac{\partial}{\partial t} u(t, x) = \lim_{\Delta t \searrow 0} \frac{[\mathcal{U}^{t+\Delta t} f](x) - [\mathcal{U}^t f](x)}{\Delta t} = [\mathcal{L}\mathcal{U}^t f](x) .$$

In other words, once the infinitesimal generator is known, one can in principle compute the action of the Markov semigroup on observables by solving the partial differential equation

$$\partial_t u = \mathcal{L}u \quad u(0, x) = f(x) . \quad (4.2)$$

On the other hand, once we compute the adjoint of the infinitesimal generator,

\mathcal{L}^* , we arrive at the partial differential equation

$$\partial_t \rho = \mathcal{L}^* \rho \quad \rho(0, x) = \rho_0(x) . \quad (4.3)$$

If $\rho_0(x)$ is the density for an initial probability distribution on the state space, then solving (4.3), one obtains the time-dependent density ρ_t . While solving either (4.3) or (4.2) would answer many of our questions regarding semistochastic processes, this is difficult to do in practice.

4.2 Generator of the Semistochastic Process

Before proceeding with the specification of the Markov semigroup and infinitesimal generator of interest, let us revisit the properties that characterize the semistochastic process $\{X(t)\}$, with corresponding state space \mathcal{X} . In the time between two consecutive disturbances $X(t)$ evolves deterministically, governed by the autonomous ordinary differential equation

$$\frac{d}{dt} x(t) = v(x(t)) . \quad (4.4)$$

We assume existence and uniqueness of solutions to (4.4), for example, by requiring that v be Lipschitz. For future reference, we now set up notation to refer to the solutions of the differential equation (4.4). The flow of (4.4) with initial condition $x(0) = x_0$ we denote by $\phi^t(x_0)$, and the time duration needed to deterministically evolve from x_0 to x_1 we denote by $\psi(x_0, x_1)$. Thus, in the absence of disturbances, we have

$$x_1 = \phi^t(x_0) \quad \iff \quad t = \psi(x_0, x_1) . \quad (4.5)$$

We assume that the occurrences of disturbances come from a state-dependent Poisson process with rate parameter $\Lambda(x)$; equivalently, for sufficiently small Δt

$$\mathbb{P}(\text{disturbance occurs in } (t, t + \Delta t]) = \Delta t \Lambda(X_t) .$$

Furthermore, to determine the severity of individual disturbances, we define a transition kernel, $P(x, dy)$, with the property that for any measurable subset of the state space B ,

$$\mathbb{P}(Y_n \in B | Y_n^- = x) = \int_B P(x, dy) , \quad (4.6)$$

where Y_n is the n^{th} post-disturbance random variable and Y_n^- is the n^{th} pre-disturbance random variable. For the semistochastic problem, we refer to the transition kernel P as a *jump kernel*. While it is difficult to develop an explicit formula for the semistochastic Markov semigroup, \mathcal{U}^t , there is a natural representation of the infinitesimal generator \mathcal{L} . To develop the expression for \mathcal{L} , we observe that for any fixed time duration Δt

$$\mathcal{U}^{\Delta t} f(x) - f(x) = \mathbb{E}[f(X_{\Delta t}) | X_0 = x] - f(x) .$$

If we let τ_1 denote the time of the first disturbance after time $t = 0$, and split the above equality into pre- and post- disaster time intervals, then

$$\begin{aligned} \mathcal{U}^{\Delta t} f(x) - f(x) &= \mathbb{E}[f(X_{\Delta t}) \mathbb{1}\{\Delta t < \tau_1\} | X_0 = x] \\ &\quad + \mathbb{E}[f(X_{\Delta t}) \mathbb{1}\{\Delta t \geq \tau_1\} | X_0 = x] - f(x) . \end{aligned}$$

We split the above equation into deterministic and random parts and approximate each separately. If Δt is sufficiently small, we can use the first-order Taylor expansion for the deterministic flow combined with the estimates for Poisson processes to arrive at

$$\begin{aligned}\mathbb{E}[f(X_{\Delta t}) \mathbb{1}\{\Delta t < \tau_1\} | X_0 = x] &= f(\phi^{\Delta t}(x))[1 - \Delta t\Lambda(x) + o(\Delta t)] \\ &= \Delta t\nabla f(x) \cdot v(x) - \Delta t f(x)\Lambda(x) + o(\Delta t) .\end{aligned}$$

To approximate the random part we take advantage of the fact that for small time intervals there is unlikely to be more than a single disturbance. Letting τ_2 denote the time at which the second disturbance occurs and using the standard errors for approximating solutions to ordinary differential equations we have

$$\begin{aligned}\mathbb{E}[f(X_{\Delta t})\mathbb{1}\{\Delta t \geq \tau_1\} | X_0 = x] &= \mathbb{E}[f(X_{\Delta t})\mathbb{1}\{\tau_1 \leq \Delta t < \tau_2\} | X_0 = x] + \mathbb{E}[f(X_{\Delta t})\mathbb{1}\{\Delta t \geq \tau_2\} | X_0 = x] \\ &= \mathbb{E}[f(X_{\Delta t})\mathbb{1}\{\tau_1 \leq \Delta t < \tau_2\} | X_0 = x] + o(\Delta t) \\ &= \mathbb{E}[f(\phi^{\Delta t - \tau_1}(y))\mathbb{1}\{\tau_1 \leq \Delta t \leq \tau_2\} | X_0 = x, X(\tau_1) = y] + o(\Delta t) \\ &= \mathbb{E}\left[\left(\int P(x, dy)f(y) + o(1)\right) \mathbb{1}\{\Delta t \geq \tau_2\} | X_0 = x\right] \\ &= \Delta t\Lambda(x) \int P(x, dy)f(y) + o(\Delta t) .\end{aligned}$$

Combining everything, we have

$$\mathcal{U}^{\Delta t} f(x) - f(x) = \Delta t \left(\nabla f(x) \cdot v(x) - \Lambda(x)f(x) + \Lambda(x) \int P(x, dy)f(y) + o(\Delta t) \right) .$$

Dividing by Δt and taking the limit as Δt approaches zero yields the equation for the infinitesimal generator:

$$[\mathcal{L}f](x) = \nabla f(x) \cdot v(x) - \Lambda(x)f(x) + \Lambda(x) \int P(x, dy) f(y) . \quad (4.7)$$

This can be rewritten by integrating the $\Lambda(x)f(x)$ term against the jump kernel $P(x, dy)$ to arrive at the standard form for the sum of a deterministic and pure jump process,

$$[\mathcal{L}f](x) = \nabla f(x) \cdot v(x) + \Lambda(x) \int P(x, dy) [f(y) - f(x)] .$$

We note for future use that if the state-space is one-dimensional and we have a multiplicative relationship between the pre- and post- disturbance values, then we can provide a slightly simpler representation for \mathcal{L} . If the multipliers are supported on the interval $[0, a]$ with common density $\rho(x, \alpha)$ for each fixed $x \in \mathcal{X}$, then $\int P(x, dy)f(y) = \int_{[0, a]} \rho(x, \alpha)f(\alpha x)d\alpha$, and we have

$$[\mathcal{L}f](x) = f'(x)v(x) + \Lambda(x) \int_0^a \rho(x, \alpha)[f(\alpha x) - f(x)]d\alpha . \quad (4.8)$$

In general, if we assume that the the jump kernel $P(x, dy)$ has a density $p(x, y)$,

$$P(x, dy) = p(x, y)dy ,$$

then we can determine the adjoint of the generator. We skip the computation here, but record the result for completeness. The action of the adjoint of the infinitesimal generator on densities, ϱ , of absolutely continuous measures is

given by

$$[\mathcal{L}^* \varrho](x) = -\operatorname{div}(\varrho(x)v(x)) + \int p(x, y) dy [\varrho(y)\Lambda(y) - \varrho(x)\Lambda(x)] .$$

While we will not pursue this further in this dissertation, it is possible that by studying \mathcal{L}^* in specific cases, one could determine the time-dependent density for the semistochastic process by solving (4.3).

4.3 Dissection of the Generator and Expansion of the Semigroup

In this section we develop a useful representation for the Markov semigroup of our stochastic process and from this derive an inequality which is used in our construction of minorizing measures. To achieve this, we start by separating the infinitesimal generator into the following components

$$\begin{aligned} [\mathcal{L}_0 f](x) &= \nabla f(x)v(x) - \Lambda(x)f(x) \\ [\mathcal{L}_1 f](x) &= \Lambda(x) \int P(x, dy) f(y) . \end{aligned}$$

Then $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1$ where \mathcal{L}_0 corresponds to deterministic evolution plus a loss term and \mathcal{L}_1 corresponds to a gain term. In what follows, we denote the sub-Markov semigroup generated by \mathcal{L}_0 as \mathcal{U}_0 . Recalling that $\phi^t(x)$ denotes the deterministic flow for the solution to the differential equation $x'(t) = v(x)$ with initial condition $x(0) = x$, we can readily compute

$$[\mathcal{U}_0^t f](x) = \exp\left(-\int_0^t \Lambda(\phi^s(x)) ds\right) f(\phi^t(x)) .$$

This can be verified directly:

$$\begin{aligned}
\frac{\partial}{\partial t}[\mathcal{U}_0^t f](x) &= \frac{\partial}{\partial t} \left(\exp \left(- \int_0^t \Lambda(\phi^s(x)) \, ds \right) f(\phi^t(x)) \right) \\
&= v(\phi^t(x)) \cdot \nabla f(\phi^t(x)) \exp \left(- \int_0^t \Lambda(\phi^s(x)) \, ds \right) \\
&\quad - \Lambda(\phi^t(x)) f(\phi^t(x)) \exp \left(- \int_0^t \Lambda(\phi^s(x)) \, ds \right) \\
&= [\mathcal{L}_0 \mathcal{U}_0^t f](x)
\end{aligned}$$

To streamline the exposition and to allow a more natural interpretation of \mathcal{U}_0 we remind the reader of the semistochastic survival function,

$$S(t, x) = \exp \left(- \int_0^t \Lambda(\phi^s(x)) \, ds \right) ,$$

which represents the conditional probability of starting at x and evolving deterministically for t units of time with no occurrence of a disturbance. With this in mind, the action of \mathcal{U}_0 on observables can be rewritten as

$$[\mathcal{U}_0^t f](x) = S(t, x) f(\phi^t(x)) .$$

In order to facilitate future estimates on the Markov semigroup, we develop an iterative scheme for computing \mathcal{U}^t .

Proposition 4.3. *Let \mathcal{U}^t be a strongly continuous Markov semigroup with infinitesimal generator \mathcal{L} and assume that \mathcal{L} can be decomposed as $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1$, with \mathcal{L}_0 generating the sub-Markov semigroup \mathcal{U}_0^t . Then the action of \mathcal{U}^t on*

an observable f can be decomposed into

$$[\mathcal{U}^t f](x) = [\mathcal{U}_0^t f](x) + \int_0^t [\mathcal{U}_0^{t-s} (\mathcal{L} - \mathcal{L}_0) \mathcal{U}^s f](x) ds .$$

Proof. Let $0 \leq s \leq t$, and recall that \mathcal{U}^0 and \mathcal{U}_0^0 are both identity operators. Then for any observable f ,

$$\begin{aligned} \int_0^t [\mathcal{U}_0^{t-s} (\mathcal{L} - \mathcal{L}_0) \mathcal{U}^s f](x) ds &= \int_0^t [(\mathcal{U}_0^{t-s} \mathcal{L} \mathcal{U}^s - \mathcal{U}_0^{t-s} \mathcal{L}_0 \mathcal{U}^s) f](x) ds \\ &= \int_0^t \left[\frac{d}{ds} (\mathcal{U}_0^{t-s} \mathcal{U}^s) f \right](x) ds \\ &= [(\mathcal{U}_0^0 \mathcal{U}^t - \mathcal{U}_0^t \mathcal{U}^0) f](x) \\ &= [\mathcal{U}^t f](x) - [\mathcal{U}_0^t f](x) \end{aligned}$$

Solving for \mathcal{U}^t above yields the result. □

Applying this expansion to the semistochastic generator given by (4.7) and recalling that $\mathcal{L} - \mathcal{L}_0 = \mathcal{L}_1$, we have

$$\begin{aligned} [\mathcal{U}^t f](x) &= [\mathcal{U}_0^t f](x) + \int_0^t [\mathcal{U}_0^{t-s} (\mathcal{L} - \mathcal{L}_0) \mathcal{U}^s f](x) ds \\ &= S(t, x) f(\phi^t(x)) \\ &\quad + \int_0^t ds S(t-s, x) \Lambda(\phi^{t-s}(x)) \int P(\phi^{t-s}(x), dy) [\mathcal{U}^s f](y) . \end{aligned}$$

We use this expansion to arrive at a useful inequality, which is the content of the following corollary

Corollary 4.4. *Let \mathcal{U}^t be a Markov semigroup with infinitesimal generator*

given by (4.7), then

$$[\mathcal{U}^t f](x) \geq \int_0^t ds S(t-s, x) \Lambda(\phi^{t-s}(x)) \int P(\phi^{t-s}(x), dy) S(s, x) f(\phi^s(y)) .$$

Proof. We discard the leading \mathcal{U}_0^t term from the expansion of the semigroup (which is positive) and replace the nested \mathcal{U}^s term with the smaller \mathcal{U}_0^s . \square

Corollary 4.4 is one of the key tools for establishing minorizations for semistochastic processes in the next section.

Chapter 5

Minorization and Convergence Rates

We begin this chapter by establishing a relationship between convergence rates for continuous-time Markov processes and their discretizations. We then investigate minorization and drift conditions for the purpose of estimating the rate of convergence of a semistochastic process towards its limiting distribution. The methods we use are purely probabilistic and are rooted in coupling or split-chain arguments (see [2], [28], [36], and [34]). We conclude this chapter by presenting our main result, which establishes lower bounds on the rates of convergence for both bounded and unbounded state semistochastic processes .

5.1 Discretization of Continuous Processes

We begin the discussion by describing a simple method for transforming continuous-time into discrete-time processes. We also establish an inequality for compar-

ing the rates of convergence of the discretized process to the original. To measure the distance between distributions, we use the total variation metric.

Definition 5.1. *Let μ_1 and μ_2 be distributions, then the **total variation distance** between μ_1 and μ_2 is given by*

$$d_{TV}(\mu_1, \mu_2) := \sup_A |\mu_1(A) - \mu_2(A)|,$$

where the supremum is taken over all measurable subsets $A \subset \mathcal{X}$.

It is also common to define the total variation distance between distributions to be $\sup_{|f|_\infty \leq 1} |\mu_1(f) - \mu_2(f)|$, which differs from our definition by a factor of 2. Our definition is, however, equivalent to

$$d_{TV}(\mu_1, \mu_2) = \sup_{0 \leq f(x) \leq 1} |\mu_1(f) - \mu_2(f)|.$$

Before constructing our discretization we review the action of the transition kernels of (discrete-time) Markov chains on observables. To this end, let $Q(x, dy)$ denote the one-step transition kernel for a Markov chain $\{X_n\}$, and let f be an observable, then

$$[Qf](x) := \mathbb{E}[f(X_1) | X_0 = x] = \int Q(x, dy) f(y).$$

And similarly for the n-step transition kernels,

$$[Q^n f](x) := \mathbb{E}[f(X_n) | X_0 = x] = \int Q^n(x, dy) f(y).$$

In the context of the continuous time processes studied in the previous chapter, we can form a discrete-time Markov chain by sampling at a fixed time interval,

Δt , and defining a discrete-time transition kernel by

$$Q = \mathcal{U}^{\Delta t} .$$

With this in mind, we now establish a useful lemma relating the convergence of a continuous-time process to the convergence of an associated discretization for a fixed value of Δt .

Lemma 5.2. *Let π denote the stationary distribution for a continuous-time Markov process $\{X_t\}$ with Markov semigroup \mathcal{U}^t and let Δt be some fixed positive time increment. If we set $Q = \mathcal{U}^{\Delta t}$, then for any initial distribution μ_0 of X_0 ,*

$$d_{TV}(\mu_0 \mathcal{U}^t, \pi) \leq d_{TV}(\mu_0 Q^n, \pi) ,$$

where $n = \lfloor t/\Delta t \rfloor$ is the greatest integer less than or equal to $t/\Delta t$.

Proof. Since $n = \lfloor t/\Delta t \rfloor$, we can represent t as

$$t = n\Delta t + \tau$$

for some $0 \leq \tau < \Delta t$. Then for any observable f with $0 \leq f(x) \leq 1$,

$$\begin{aligned} |\mu_0 \mathcal{U}^t f - \pi f| &= |\mu_0 \mathcal{U}^{n\Delta t} \mathcal{U}^\tau f - \pi f| \\ &= |\mu_0 \mathcal{U}^{n\Delta t} \mathcal{U}^\tau f - \pi \mathcal{U}^\tau f| \\ &\leq \sup\{|\mu_0 \mathcal{U}^{n\Delta t} g - \pi g| : |g|_\infty \leq 1\} \\ &= d_{TV}(\mu_0 Q^n, \pi) , \end{aligned}$$

where we have made use of the invariance of π , and the fact that if $0 \leq f(x) \leq 1$ then $0 \leq \mathcal{U}^t f \leq 1$ as well. Since the inequality holds for any f with $0 \leq f(x) \leq 1$, we conclude that

$$d_{TV}(\mu_0 \mathcal{U}^t, \pi) \leq d_{TV}(\mu_0 Q^n, \pi) .$$

□

5.2 Minorization, Drift, and Coupling

In this section we define both minorization and drift conditions for discrete-time Markov chains. We also present conditions whereby discretizations of semistochastic processes admit minorizations and satisfy drift conditions.

Definition 5.3. *A Markov chain with transition kernel Q on a state space \mathcal{X} is said to satisfy a **minorization condition** on a subset $A \subseteq \mathcal{X}$ if there is a probability measure η on \mathcal{X} , a positive integer n_0 , and a number $\epsilon > 0$ such that*

$$Q^{n_0}(x, B) \geq \epsilon \eta(B)$$

for all $x \in A$ and for any measurable set B .

This is equivalent to requiring for any nonnegative observable f , and for all $x \in A$,

$$[Q^{n_0} f](x) := \int Q^{n_0}(x, dy) f(y) \geq \epsilon \int f(y) d\eta(y) .$$

We also need the concept of a drift condition.

Definition 5.4. A Markov chain $\{X_n\}$ with transition kernel, Q on a state space \mathcal{X} is said to satisfy a **drift condition** if there is a nonnegative function $V : \mathcal{X} \mapsto \mathbb{R}_{\geq 0}$, a number $\beta < 1$, and some finite b such that

$$\mathbb{E}[V(X_1)|X_0 = x] \leq \beta V(x) + b$$

for all $x \in \mathcal{X}$.

To simplify the remaining exposition, we now assume that our state space is one-dimensional (though most of the arguments work for vector-valued state spaces as well). We first establish uniform minorization conditions for a class of semistochastic processes with bounded state spaces, and then establish minorization and drift conditions for semistochastic processes with unbounded state spaces.

Theorem 5.5. Let $\{X(t)\}$ be a scalar semistochastic process with bounded state space $\mathcal{X} = [0, k]$ for some $k < \infty$ and with infinitesimal generator given by (4.8) with $a = 1$ (disturbances only result in jumps down) so that the action of \mathcal{L} is given by

$$[\mathcal{L}f](x) = f'(x)v(x) + \Lambda(x) \int_0^1 \rho(x, \alpha)[f(\alpha x) - f(x)]d\alpha .$$

Fix a time interval Δt and let $Q := \mathcal{U}^{\Delta t}$ denote the discrete-time transition kernel for the corresponding uniform discretization of $\{X(t)\}$. Suppose moreover that the following conditions are met:

- (i) $0 < \lambda_* \leq \Lambda(x) \leq \lambda^* < \infty$ for all $x \in \mathcal{X}$, for some constants λ_* and λ^* ,
- (ii) $\rho(x, \alpha) \geq \rho_*$ for all $x \in \mathcal{X}$ and $\alpha \in [0, 1]$, for some constant $\rho_* > 0$,

(iii) the function v is non-negative, Lipschitz, and $v(0) \neq 0$ with $v(x) = 0$ for at most finitely many x .

then Q can be uniformly minorized. Specifically, for all $x \in \mathcal{X}$ and any non-negative observable f :

$$[Q f](x) \geq \epsilon \int f(z) d\eta(z),$$

with

$$\epsilon = \frac{\rho_* \Phi \lambda_* \exp(-\lambda^* \Delta t)}{k},$$

where we define Φ as

$$\Phi = \int_0^{\phi^{\Delta t}(0)} [\Delta t - \psi(0, z)] dz, \quad (5.1)$$

with ϕ and ψ defined as in (4.5), and η an absolutely continuous measure with Radon-Nikodym derivative given by

$$\frac{d\eta}{dz} = \frac{\Delta t - \psi(0, z)}{\Phi} \mathbb{1}_{\{0 \leq z \leq \phi^{\Delta t}(0)\}}.$$

Proof. Let ϕ and ψ be defined as in (4.5) and let f be a nonnegative observable, then applying the result from Corollary 4.4, we have

$$\begin{aligned} [Q f](x) &\geq \int_0^{\Delta t} ds S(\Delta t - s, x) \Lambda(\phi^{\Delta t - s}(x)) \\ &\quad \times \int_0^1 d\alpha \rho(x, \alpha) S(s, \alpha \phi^s(x)) f(\alpha \phi^s(x)), \end{aligned}$$

where

$$S(t, x) = \exp\left(-\int_0^t \Lambda(\phi^s(x)) ds\right).$$

Using the assumption that $0 < \Lambda(x) \leq \lambda^*$, we have

$$S(t, x) \geq \exp(-\lambda^* t) \quad \text{for all } x \in [0, k] .$$

Combining the above inequalities with the assumed bounds on $\rho(x, \alpha)$ and $\Lambda(x)$, we arrive at

$$[Qf](x) \geq \rho_* \lambda_* \exp(-\lambda^* \Delta t) \int_0^{\Delta t} ds \int_0^1 d\alpha f(\alpha \phi^s(x)) .$$

Changing variables according to

$$z = \alpha \phi^s(x) \quad dz = \phi^s(x) d\alpha$$

and interchanging the order of integration, we have

$$\begin{aligned} [Qf](x) &\geq \rho_* \lambda_* \exp(-\lambda^* \Delta t) \int_0^{\Delta t} ds \int_0^1 d\alpha f(\alpha \phi^s(x)) \\ &= \rho_* \lambda_* \exp(-\lambda^* \Delta t) \int_0^{\Delta t} ds (\phi^s(x))^{-1} \int_0^{\phi^s(x)} dz f(z) \\ &\geq \rho_* \lambda_* \exp(-\lambda^* \Delta t) \int_0^{\Delta t} ds k^{-1} \int_0^{\phi^s(0)} dz f(z) \\ &= \frac{\rho_* \lambda_*}{k} \exp(-\lambda^* \Delta t) \int_0^{\phi^{\Delta t}(0)} dz f(z) \int_{\psi(0, z)}^{\Delta t} ds \\ &= \frac{\rho_* \lambda_*}{k} \exp(-\lambda^* \Delta t) \int_0^{\phi^{\Delta t}(0)} f(z) [\Delta t - \psi(0, z)] dz \\ &> 0 . \end{aligned}$$

We have made use of both the monotonicity of $\phi^t(x)$, the boundedness of the state space (so that $k < \infty$), and the fact that $v(0) > 0$ in the preceding arguments to arrive at a positive lower bound for $[Qf](x)$ for all $x \in \mathcal{X}$. Thus,

we have established the uniform minorization

$$[Qf](x) \geq \epsilon \int f(z) d\eta(z),$$

with ϵ and η defined as in the statement of the theorem. □

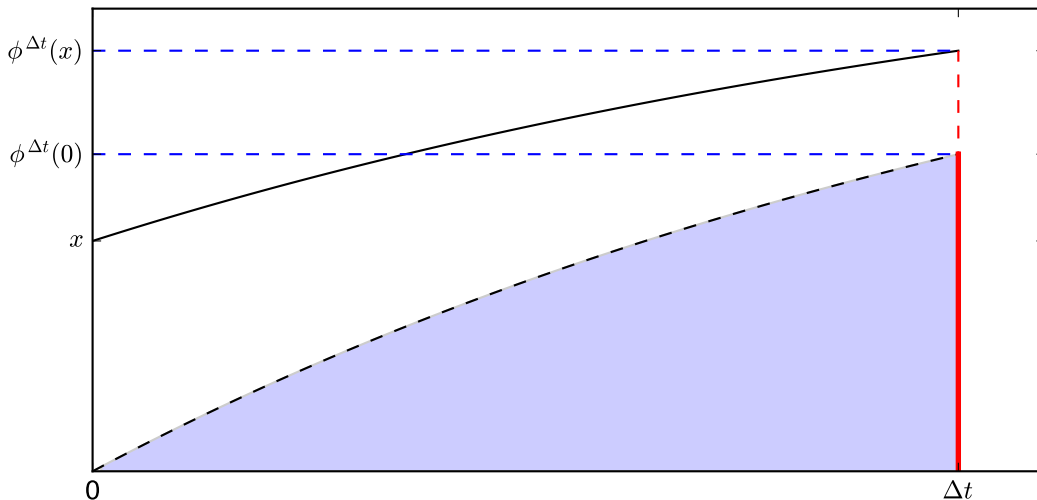


Figure 5.1: On the construction of the minorizing measure in Theorem 5.5.

While the restriction $v(0) > 0$ may seem unusual for biological models, it is a reasonable assumption for the carbon content problem since even in the event of a complete catastrophe, there is regrowth. The specific case of $v(x) = 1 - x$ with state space $\mathcal{X} = [0, 1]$ is considered in [24] as a model for carbon content in an ecosystem and meets all conditions of the theorem. If $v(0) = 0$ then one could determine a drift condition in a manner similar to what will be done in Theorem 5.8.

The significance of establishing a uniform minorization for a Markov chain is that this entails the uniform ergodicity of the process. Moreover, it can be proved that uniformly ergodic Markov chains possess unique stationary

distributions and converge to them exponentially fast; these results go back to Doeblin and can be found in Doob [12]; for additional discussion see [29, 36, 26]. One approach to proving these results is based on the idea of coupling, which we now explore in more detail. Given two random variables X and Y with respective distributions μ_X and μ_Y , then we have the inequality

$$d_{TV}(\mu_X, \mu_Y) \leq \mathbb{P}(X \neq Y) . \quad (5.2)$$

The significance of this inequality is that it bounds the total variation distance between the distributions of two random variables by the probability that they are not equal. To make use of this fact, consider a Markov chain $\{X_n\}$ on \mathcal{X} with transition kernel Q , initial distribution μ_0 , and stationary distribution π . We can construct a new Markov chain $\{X_n, Y_n\}$ on the product space $\mathcal{X} \times \mathcal{X}$ with the following properties

- (i) $X_0 \sim \mu_0$
- (ii) $Y_0 \sim \pi$
- (iii) $\mathbb{P}(X_{n+1} \in A | X_n) = Q(X_n, A)$
- (iv) $\mathbb{P}(Y_{n+1} \in A | Y_n) = Q(Y_n, A)$
- (v) There exists a random time T for which $X_n = Y_n$ for all $n \geq T$.

The time T is called the *coupling time* and represents a time after which the parallel processes $\{X_n\}$ and $\{Y_n\}$ “couple” and henceforth are equal. The significance of such a construction is that combined with (5.2) it produces a bound on the total variation distance between the n -step evolved distribution

μ_n and the stationary distribution π in terms of coupling times,

$$d_{TV}(\mu_n, \pi) \leq \mathbb{P}(T > n)$$

We will use the idea of coupling to present a simple proof of the following theorem.

Theorem 5.6. *If there exists an $n_0 \in \mathbb{N}$ such that the transition kernel Q of the Markov chain $\{X_n\}$ with state space \mathcal{X} satisfies*

$$Q^{n_0}(x, A) \geq \epsilon \eta(A)$$

for all $x \in \mathcal{X}$ and any measurable set A , then $\{X_n\}$ has a unique stationary distribution π , and for any initial distribution μ_0

$$d_{TV}(\mu_0 Q^n, \pi) \leq (1 - \epsilon)^{\lfloor n/n_0 \rfloor}$$

Proof. By appropriately redefining Q , it suffices to consider the particular case $n_0 = 1$. Let $\{X_n\}$ and $\{Y_n\}$ be independent Markov chains on \mathcal{X} with $X_0 \sim \mu_0$ and $Y_0 \sim \pi$. We can use the minorization to define the distributions for X_{n+1} and Y_{n+1} for all $n > 0$ according to the following scheme:

1. Flip an independent coin that has probability ϵ of landing heads.
2. If heads, then choose $x \in \mathcal{X}$ independently according to the distribution η and set $X_{n+1} = Y_{n+1} = x$.
3. If tails, then choose X_{n+1} and Y_{n+1} independently according to

$$\mathbb{P}(X_{n+1} \in A | X_n) = \frac{Q(X_n, A) - \epsilon \eta(A)}{1 - \epsilon},$$

$$\mathbb{P}(Y_{n+1} \in A | Y_n) = \frac{Q(Y_n, A) - \epsilon\eta(A)}{1 - \epsilon} .$$

The probabilities above are the residual probabilities from decomposing Q into

$$Q = \epsilon\eta + (1 - \epsilon) \left(\frac{Q - \epsilon\eta}{1 - \epsilon} \right) .$$

The construction furthermore guarantees that

$$\mathbb{P}(X_{n+1} \in A | X_n) = Q(X_n, A) \quad \text{and} \quad \mathbb{P}(Y_{n+1} \in A | Y_n) = Q(Y_n, A) .$$

The utility of this construction comes from the fact that we can use it to determine a bound on the coupling time. If we let T denote the time at which the first heads lands (which depends only on the minorization parameter ϵ), and define a new Markov chain $\{Z_n\}$ according to

$$Z_n = \begin{cases} Y_n & \text{for } n \leq T , \\ X_n & \text{for } n > T . \end{cases}$$

Then $\{Z_n\}$ and $\{X_n\}$ will couple with coupling time T . Moreover, due to the coin flip probabilities depending only on ϵ , we have

$$\mathbb{P}(T > n) = (1 - \epsilon)^n ,$$

and hence

$$d_{TV}(\mu_0 Q^n, \pi) \leq (1 - \epsilon)^n$$

□

If a uniform minorization cannot be achieved, then a combination of minorization and drift conditions can be used to compute convergence rates. One specific result is given below.

Theorem 5.7 (Theorem 12 in [36]). *Suppose a Markov chain $\{X_n\}$ with transition kernel Q and state space \mathcal{X} satisfies a drift condition, so that there exists a function $V : \mathcal{X} \mapsto \mathbb{R}_{\geq 0}$ for which*

$$\mathbb{E}[V(X_1) \mid X_0 = x] \leq \beta V(x) + b, \quad x \in \mathcal{X}$$

for some $\beta < 1$ and some $b < \infty$; and also satisfies the minorization condition

$$Q(x, A) \geq \epsilon \eta(A),$$

for some $\epsilon > 0$, some $\kappa > \frac{2b}{1-\beta}$, and for all $x \in V^{-1}[0, \kappa]$ and any measurable set A . Then the Markov chain has a unique stationary distribution π and for any $0 < r < 1$ and any $n \in \mathbb{N}$, we have for any initial distribution μ_0

$$d_{TV}(\mu_0 Q^n, \pi) \leq (1 - \epsilon)^{nr} + (\theta^{1-r} \Theta^r)^n \left(1 + \frac{b}{1 - \beta} + \mathbb{E}_{\mu_0}[V(X_0)] \right), \quad (5.3)$$

with

$$\theta = \frac{1 + 2b + \kappa\beta}{1 + \kappa} \quad \text{and} \quad \Theta = 1 + 2(\beta\kappa + b). \quad (5.4)$$

With this in mind, we now establish conditions under which an unbounded semistochastic process both satisfies a drift condition and can be minorized on an appropriate subset of its state space. If the jump kernel of our semistochastic process has the multiplicative form of (4.8) with $a = 1$, it will be convenient to introduce the quantity $\zeta(x)$ to denote the expected fractional

loss resulting from a single disturbance,

$$\zeta(x) := \int_0^1 \rho(x, \alpha) (1 - \alpha) \, d\alpha \in (0, 1) . \quad (5.5)$$

Thus larger values of $\zeta(x)$ correspond to an expectation of more severe disturbances and the limiting value $\zeta(x) = 0$ would result in purely deterministic growth.

Theorem 5.8. *Let $\{X(t)\}$ be a scalar semistochastic process with state space $\mathbb{R}_{\geq 0}$ and with infinitesimal generator given by (4.8) with $a = 1$ (disturbances only result in jumps down) so that the action of \mathcal{L} is given by*

$$[\mathcal{L}f](x) = f'(x)v(x) + \Lambda(x) \int_0^1 \rho(x, \alpha)[f(\alpha x) - f(x)]d\alpha .$$

Fix a time interval $\Delta t > 0$ and let $Q := \mathcal{U}^{\Delta t}$ denote the discrete-time transition kernel for the corresponding uniform discretization of $\{X(t)\}$. Suppose moreover that the following conditions are met:

- (i) $0 < \lambda_* \leq \Lambda(x) \leq \lambda^* < \infty$ for all $x \in \mathcal{X}$, for some constants λ_* and λ^* ,*
- (ii) $\rho(x, \alpha) \geq \rho_*$ for all $x \in \mathcal{X}$ and $\alpha \in [0, 1]$, for some constant $\rho_* > 0$,*
- (iii) $\zeta(x) \geq \zeta_*$ for all $x \in \mathcal{X}$, for some constant $\zeta_* > 0$,*
- (iv) the function v is Lipschitz, satisfies*

$$0 \leq v(x) \leq v^* = \text{const} , \quad v(0) \neq 0 , \quad (5.6)$$

and vanishes for at most finitely many x .

then Q satisfies a drift condition with respect to the identity map,

$$\mathbb{E}[X_1 | X_0 = x] \leq \beta V(x) + b,$$

for all $x \in \mathcal{X}$, with

$$\beta = e^{-\lambda_* \zeta_* \Delta t}, \quad b = \frac{v^*}{\lambda_* \zeta_*} (1 - e^{-\lambda_* \zeta_* \Delta t}),$$

Moreover, for any $\kappa < \infty$, Q admits the minorization

$$Q(x, B) \geq \epsilon \eta(B),$$

for any $x \in [0, \kappa]$ and all measurable $B \in \mathcal{X}$, with

$$\eta(B) = \int_0^{\phi^{\Delta t}(0)} \frac{\Delta t - \psi(0, z)}{\Phi} dz \mathbb{1}\{z \in B\}, \quad \epsilon = \frac{\rho_* \Phi \lambda_* \exp(-\lambda_* \Delta t)}{\kappa},$$

where Φ is defined by (5.1), and ϕ and ψ defined as in (4.5).

Proof. To establish the general drift condition, we will use the identity map $I : x \mapsto x$ for our drift function. The action of \mathcal{L} on I can be represented in terms of $\zeta(x)$,

$$\begin{aligned} [\mathcal{L}I](x) &= v(x) + \Lambda(x) \int_0^1 \rho(x, \alpha) [\alpha x - x] d\alpha \\ &= v(x) - \Lambda(x) \zeta(x) x. \end{aligned}$$

From the conditions on v and Λ , we thus have for all $x \in \mathcal{X}$,

$$[\mathcal{L}I](x) \leq v^* - \lambda_* \zeta_* x. \quad (5.7)$$

In order to establish bounds on $\mathbb{E}[I(X_{\Delta t})|X_0 = x]$, we will use the martingale introduced in (4.1) along with a standard Gronwall argument. From the martingale, we have for any $0 \leq t$,

$$\mathbb{E} \left[X_t - X_0 - \int_0^t [\mathcal{L}I](X_s) ds \mid X_0 = x \right] = 0$$

Setting $u(t) = \mathbb{E}[I(X_t)|X_0 = x] = \mathbb{E}[(X_t)|X_0 = x]$ and writing $[\mathcal{L}I](X_s)$ explicitly, we can rewrite the above equation for expectations as an integral equation

$$u(t) = u(0) + \int_0^t \mathbb{E} [v(X_s) - \Lambda(X_s)\zeta(X_s)X_s ds \mid X_0 = x] . \quad (5.8)$$

Our sample paths are all right continuous, and thus the right hand side of (5.8) can be differentiated with respect to t . Differentiating the integral equation and referencing (5.7) , we have

$$\begin{aligned} u'(t) &= \mathbb{E} [v(X_t) - \Lambda(X_t)\zeta(X_t)X_t \mid X_0 = x] \\ &\leq v^* - \lambda_* \zeta_* u(t) . \end{aligned}$$

Rearranging the above inequality and multiplying by the integrating factor $e^{\lambda_* \zeta_* t}$ gives

$$\frac{d}{dt} (e^{\lambda_* \zeta_* t} u(t)) \leq v^* e^{\lambda_* \zeta_* t} ,$$

which is equivalent to

$$\frac{d}{dt} \left(e^{\lambda_* \zeta_* t} u(t) - \frac{v^* e^{\lambda_* \zeta_* t}}{\lambda_* \zeta_*} \right) \leq 0 .$$

Since the expression in parentheses above is decreasing with t , it must obtain its minimum on $[0, \infty)$ at $t = 0$; recalling that $u(0) = x$, we have

$$e^{\lambda_* \zeta_* t} u(t) - \frac{v^*}{\lambda_* \zeta_*} e^{\lambda_* \zeta_* t} \leq x - \frac{v^*}{\lambda_* \zeta_*} .$$

Solving for $u(t)$, simplifying, and setting $t = \Delta t$ produces the desired drift condition,

$$\mathbb{E}[X_{\Delta t} | X_0 = x] \leq e^{-\lambda_* \zeta_* \Delta t} x + \frac{v^*}{\lambda_* \zeta_*} (1 - e^{-\lambda_* \zeta_* \Delta t})$$

The last thing to do is to minorize Q on $[0, \kappa]$ for any $\kappa < \infty$; but this can be done as in Theorem 5.5 for the bounded state-space. \square

5.3 Convergence Rates for Semistochastic Processes

We now report our results on convergence rates for two classes of semistochastic processes, one with bounded state space and one with unbounded.

Theorem 5.9. *Let $\{X_t\}$ be a semistochastic process with generator (4.8) on the state space $\mathcal{X} = [0, k]$, satisfying*

- (i) $0 < \lambda_* \leq \Lambda(x) \leq \lambda^* < \infty$ for all $x \in \mathcal{X}$, for some constants λ_* and λ^* ,
- (ii) $\rho(x, \alpha) \geq \rho_*$ for all $x \in \mathcal{X}$ and $\alpha \in [0, 1]$, for some constant $\rho_* > 0$,
- (iii) the function v is non-negative, Lipschitz, and $v(0) \neq 0$ with $v(x) = 0$ for at most finitely many x .

Then $\{X_t\}$ converges exponentially fast to its unique stationary distribution π .
 Namely, for any time increment $\Delta t > 0$, and any initial distribution μ_0 ,

$$d_{\text{TV}}(\mu_t, \pi) \leq (1 - \epsilon_{\Delta t})^{\lfloor t/\Delta t \rfloor} , \quad (5.9)$$

where

$$\epsilon_{\Delta t} := \frac{\rho_* \Phi \lambda_* \exp(-\lambda_* \Delta t)}{k} , \quad (5.10)$$

$$\Phi := \int_0^{\phi^{\Delta t}(0)} [\Delta t - \psi(0, z)] dz ,$$

ϕ and ψ defined in (4.5), and η an absolutely continuous measure with Radon-Nikodym derivative given by

$$\frac{d\eta}{dz} = \frac{\Delta t - \psi(0, z)}{\Phi} \mathbb{1}\{0 \leq z \leq \phi^{\Delta t}(0)\} . \quad (5.11)$$

Proof. First, we discretize the process by setting $Q = \mathcal{U}^{\Delta t}$, and apply Theorem 5.5 to obtain the uniform minorization

$$[Qf](x) \geq \epsilon_{\Delta t} \int f(y) d\eta(y)$$

for any observable f and all $x \in \mathcal{X}$ with $\epsilon_{\Delta t}$ with the measure η as in the statement of theorem. Next we apply Theorem 5.6 in conjunction with Lemma 5.2 to obtain

$$\begin{aligned} d_{\text{TV}}(\mu_t, \pi) &\leq d_{\text{TV}}(\mu_0 Q^{\lfloor t/\Delta t \rfloor}, \pi) \\ &\leq (1 - \epsilon_{\Delta t})^{\lfloor t/\Delta t \rfloor} , \end{aligned}$$

where μ_0 is an arbitrary initial distribution. □

In a similar manner, we arrive at bounds on the convergence rates for processes with unbounded states.

Theorem 5.10. *Let $\{X_t\}$ be a semistochastic process with generator (4.8) on the state space $\mathcal{X} = [0, \infty)$, satisfying*

- (i) $0 < \lambda_* \leq \Lambda(x) \leq \lambda^* < \infty$ for all $x \in \mathcal{X}$, for some constants λ_* and λ^* ,
- (ii) $\rho(x, \alpha) \geq \rho_*$ for all $x \in \mathcal{X}$ and $\alpha \in [0, 1]$, for some constant $\rho_* > 0$,
- (iii) $\zeta(x) \geq \zeta_*$ for all $x \in \mathcal{X}$, for some constant $\zeta_* > 0$,
- (iv) the function v is Lipschitz, satisfies

$$0 \leq v(x) \leq v^* = \text{const} , \quad v(0) \neq 0 ,$$

and vanishes for at most finitely many x .

Then $\{X_t\}$ has a unique stationary distribution π to which it converges at an exponential rate. Namely, for any initial distribution μ_0 and any $\Delta t > 0$, the estimate

$$d_{\text{TV}}(\mu_t, \pi) \leq \left(2 + \frac{b}{1 - \beta} + \mathbb{E}_{\mu_0}[X_0] \right) (1 - \epsilon_{\Delta t})^{r \lceil t/\Delta t \rceil} \quad (5.12)$$

holds with Φ given by (5.1),

$$\epsilon_{\Delta t, \kappa} := \frac{\rho_* \Phi \zeta_* \lambda_* \exp(-\lambda^* \Delta t)}{\kappa} ,$$

$$\beta := e^{-\lambda_* \zeta_* \Delta t} , \quad b := \frac{v^*}{\lambda_* \zeta_*} (1 - e^{-\lambda_* \zeta_* \Delta t}) ,$$

$$\theta := \frac{1 + 2b + \kappa\beta}{1 + \kappa} , \quad \Theta := 1 + 2(\beta\kappa + b) ,$$

$$r := \frac{\ln \theta}{\ln \frac{\theta(1-\epsilon_{\Delta t})}{\Theta}} = \frac{\ln \frac{1}{\theta}}{\ln \frac{1}{\theta} + \ln \Theta + \ln \frac{1}{1-\epsilon_{\Delta t}}} \in (0, 1) , \quad (5.13)$$

where κ can be chosen to be any number satisfying

$$\kappa > \frac{2b}{1-\beta} . \quad (5.14)$$

Proof. As in the proof of Theorem 5.9 we begin by discretizing the process by setting $Q = \mathcal{U}^{\Delta t}$, and then apply Theorem 5.8 to obtain both a drift condition with respect to the identity map and a minorization on $[0, \kappa]$. Next, we apply the result of Theorem 5.7 where we choose the value of r in such a way that the two terms in the right-hand side of (5.3) balance each other, which for large n gives us $(1-\epsilon)^r = \theta^{1-r}\Theta^r$, which gives the expression (5.13) for r . In particular, with this choice of r ,

$$(1-\epsilon)^{nr} + (\theta^{1-r}\Theta^r)^n \left(1 + \frac{b}{1-\beta} + \mathbb{E}_{\mu_0}[X_0] \right) = \left(2 + \frac{b}{1-\beta} + \mathbb{E}_{\mu_0}[X_0] \right) (1-\epsilon)^{nr}$$

for all n . We then apply Lemma 5.2 to obtain the desired bounds on the total variation distance between π and $\mu_t = \mu_0 \mathcal{U}^t$. \square

Remark 5.11. *Note that the rate of convergence in (5.12) depends on the choice of Δt and κ . To obtain tight bounds, one can choose values of Δt and κ that minimize $(1-\epsilon_{\Delta t})^{r/\Delta t}$, which can be done numerically as shown in the second example below.*

5.3.1 Examples

In these examples we illustrate how our results can be used in practice. In both examples we consider a semistochastic process $\{X_t\}$ with disturbances

corresponding only to jumps down. We further assume that the disturbances are generated by a Poisson process with constant rate parameter λ , and that the severity of disturbances is uniformly distributed. Equivalently, we are considering processes with infinitesimal generator given by

$$[\mathcal{L}f](x) = f'(x)v(x) + \lambda \int_0^1 [f(\alpha x) - f(x)]d\alpha .$$

We also demonstrate how one can optimize the relevant parameters to obtain tighter bound on rates of convergence.

Example 1. *In this example we consider a model of growth with saturation on $\mathcal{X} = [0, k]$:*

$$x'(t) = k - x , \quad k = \text{const} > 0 .$$

In this case (cf. (4.5)),

$$\phi^t(x) = k + (x - k)e^{-t} , \quad \psi(x_0, x) = \ln \frac{k - x_0}{k - x} .$$

From Theorem 5.9, for fixed Δt and arbitrary initial distribution μ_0 , the following bound holds

$$d_{\text{TV}}(\mu_0 \mathcal{U}^t, \pi) \leq (1 - \epsilon_{\Delta t})^{\lfloor t/\Delta t \rfloor}$$

(π is the unique stationary distribution). We have

$$\Phi = \int_0^{k(1-e^{-\Delta t})} \left(\Delta t - \ln \frac{k}{k-z} \right) dz = k(\Delta t + e^{-\Delta t} - 1) ,$$

$$\epsilon_{\Delta t} = \frac{\Phi \lambda e^{-\lambda \Delta t}}{k} = \lambda e^{-\lambda \Delta t} (\Delta t + e^{-\Delta t} - 1) .$$

For convergence rates, the quantity of interest is $(1 - \epsilon_{\Delta t})^{1/\Delta t}$ (cf. (5.9)). For concreteness, take $\lambda = 1$. In Figure 5.2, we plot $(1 - \epsilon_{\Delta t})^{1/\Delta t}$ as a function of Δt and observe that it exhibits a minimum at $\Delta t \approx 0.82$, for which $\epsilon_{\Delta t} \approx 0.115$. Choosing $\Delta t = 0.82$, we obtain that, for any initial distribution μ_0 ,

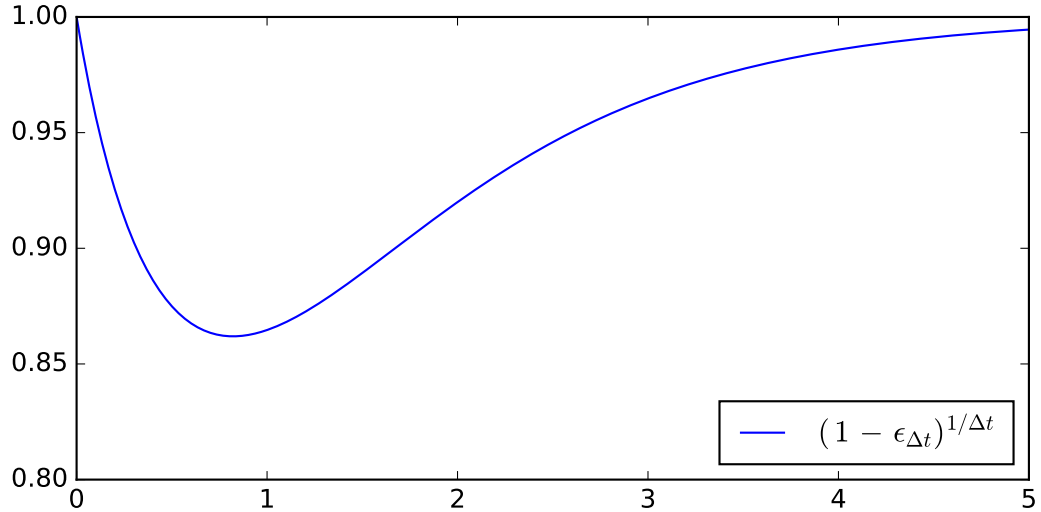


Figure 5.2: Plot of $(1 - \epsilon_{\Delta t})^{1/\Delta t}$ vs. Δt .

$$d_{\text{TV}}(\mu_t, \pi) \leq (1 - 0.115)^{\lfloor t/0.82 \rfloor} \leq 1.13 e^{-0.148t} .$$

For comparison, in Figure 5.3 we plot $(1 - \epsilon_{\Delta t})^{\lfloor t/\Delta t \rfloor}$ as a function of t for several values of Δt .

Example 2. Consider the case of constant growth rate on $\mathcal{X} = [0, \infty)$:

$$x'(t) = v = \text{const} > 0 .$$

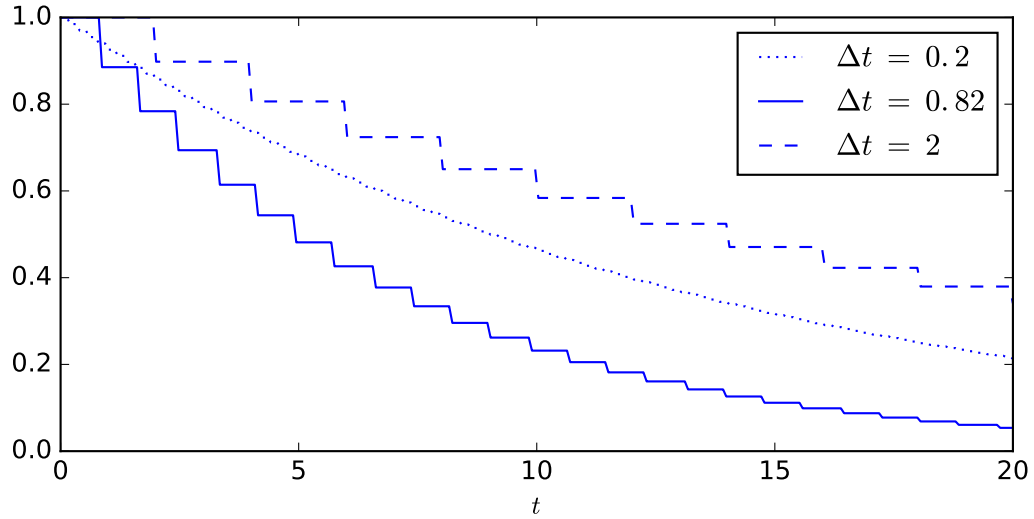


Figure 5.3: Plots of $(1 - \epsilon_{\Delta t})^{\lfloor t/\Delta t \rfloor}$ vs. t for selected Δt .

Our flow and time-duration functions are

$$\phi^t(x) = x + t, \quad \psi(x_0, x) = x - x_0.$$

From Theorem 5.10, for fixed $\Delta t > 0$, we can first establish a drift condition using the identity as our drift function. In this case the average fractional loss $\zeta(x) = \frac{1}{2}$ does not depend on x , so we can compute the expectation exactly,

$$\mathbb{E}[X_{\Delta t} | X_0 = x] = e^{-\lambda \Delta t / 2} + \frac{2v}{\lambda} (1 - e^{-\lambda \Delta t / 2}),$$

so the drift parameters are $\beta = e^{-\lambda \Delta t / 2}$, $b = \frac{2v}{\lambda} (1 - e^{-\lambda \Delta t / 2})$. To compute bounds on the convergence rates, we need to select a value $\kappa > \frac{2b}{1-\beta} = \frac{4v}{\lambda}$ for which we minorize the process on $[0, \kappa]$. We easily obtain $\Phi = v(\Delta t)^2$ and $\epsilon_{\Delta t, \kappa} = \frac{v(\Delta t)^2 \lambda e^{-\lambda \Delta t}}{\kappa}$, where we are emphasizing the dependence of ϵ on κ as

well as Δt . For θ and Θ (5.4) we obtain

$$\theta = \frac{1 + \frac{4v}{\lambda} + \left(\kappa - \frac{4v}{\lambda}\right) e^{-\lambda\Delta t/2}}{1 + \kappa}, \quad \Theta = 1 + \frac{4v}{\lambda} + \left(2\kappa - \frac{4v}{\lambda}\right) e^{-\lambda\Delta t/2};$$

in the expression for θ , note that the restriction on κ ensures the positivity of the exponential term in the numerator. Continuing the example with the specific values $v = 1$ and $\lambda = 2$, we obtain $\beta \approx 0.405$ and $b \approx 0.595$. We can then make appropriate choices for Δt and κ by minimizing the expression

$$(1 - \epsilon_{\Delta t, \kappa})^{\frac{r(\Delta t, \kappa)}{\Delta t}}$$

as illustrated in Figures 5.4 and 5.5.

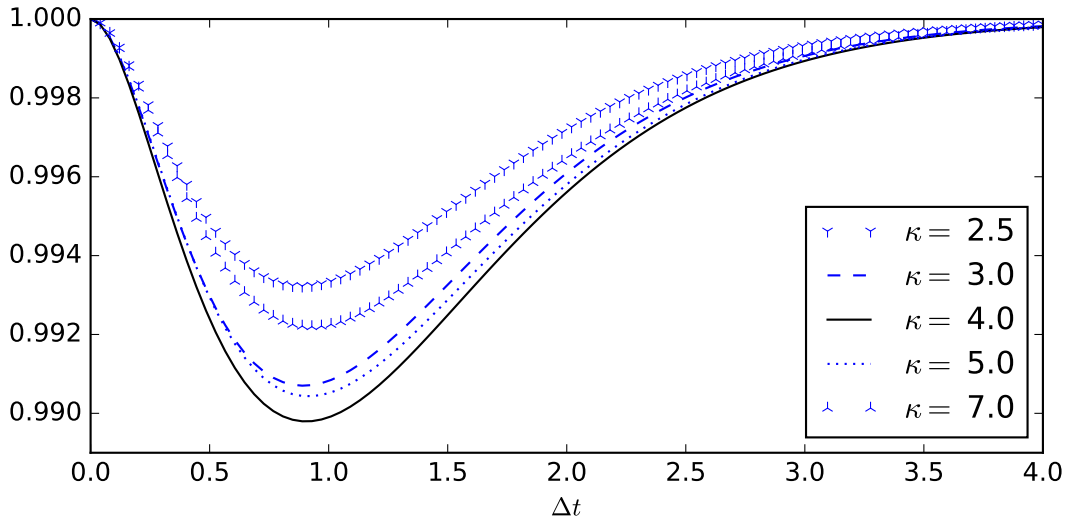


Figure 5.4: Plots of $(1 - \epsilon_{\Delta t})^{r/\Delta t}$ vs. Δt for selected κ .

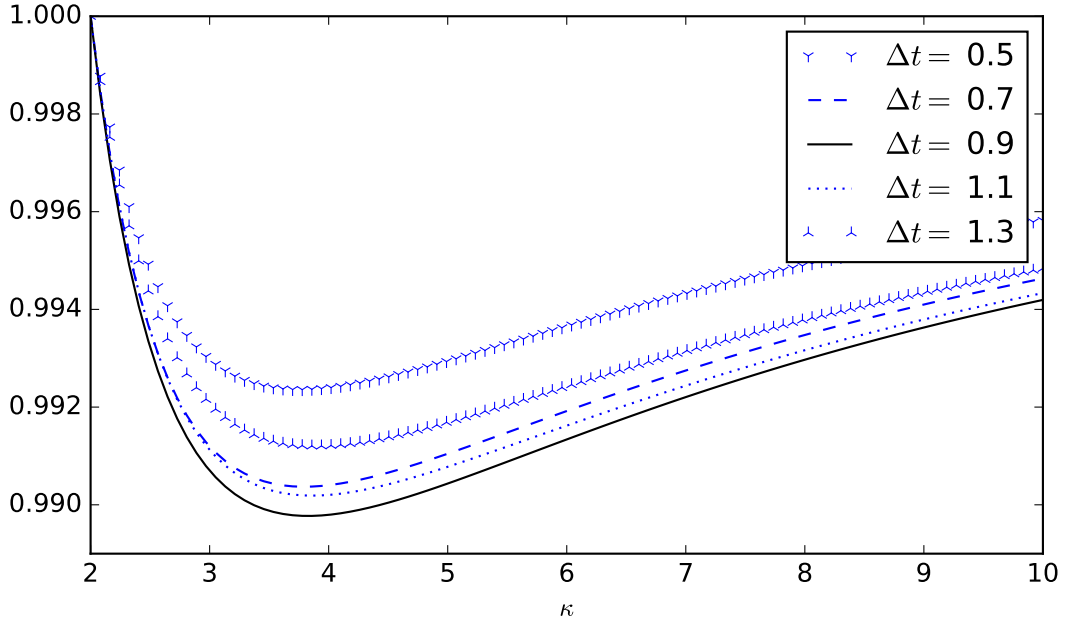


Figure 5.5: Plots of $(1 - \epsilon_{\Delta t})^{r/\Delta t}$ vs. κ for selected Δt .

Consequently, we choose $\Delta t = 0.904$, $\kappa = 3.83$, and r as in (5.13) to obtain

$$\begin{aligned} d_{\text{TV}}(\mu_t, \pi) &\leq C(1 - 0.070)^{r \lfloor t/0.904 \rfloor} \\ &\leq 1.02 C e^{-0.014t}, \end{aligned}$$

with $C = 3 + \mathbb{E}_{\mu_0}[X_0]$.

Chapter 6

Carbon Content and other Scalar Models

In this chapter we will report a formula that can be used to compute exactly the density for the stationary distribution of the semistochastic model for the original carbon content problem. Our arguments will be based on those of Leite, Petrov, and Weng in [24], however, we will generalize their results to allow for the rate of disturbances to be state-dependent by incorporating the semistochastic survival and hazard functions. We will also correct a slight averaging error in [24] and will conclude with a tabulation of the corrected results of the examples from [24].

Throughout this chapter we will consider a one-dimensional semistochastic process with infinitesimal generator acting on observables via

$$[\mathcal{L}f](x) = f'(x)v(x) + \Lambda(x) \int_0^1 [f(\alpha x) - f(x)]d\rho(\alpha) . \quad (6.1)$$

Thus we are considering epochs of deterministic growth interrupted by dis-

turbances that occur with a state-dependent hazard rate $\Lambda(x)$. Moreover, the severity of the disturbances is determined by the continuous random variables A_n which share the density $\rho(\alpha)$ obtain values in $[0, 1]$. As in previous chapters, we will denote the deterministic flow with initial condition $x(0) = x_0$ by $\phi^t(x_0)$, and the time duration needed to deterministically evolve from x_0 to x_1 by $\psi(x_0, x_1)$. Thus, in the absence of disturbances, we have

$$x_1 = \phi^t(x_0) \iff t = \psi(x_0, x_1) .$$

We will also need the semistochastic survival function, whose definition we recall

$$S(t, x) = \exp \left(- \int_0^t \Lambda(\phi^\tau(x)) \, d\tau \right) .$$

Then for T_n the random variable corresponding to the inter-disturbance time between the $(n-1)^{st}$ and n^{th} disturbances, we can express its density $f_{T_n}(t)$ in terms of the $(n-1)^{st}$ post-disturbance value x_{n-1} and the semistochastic survival and hazard functions

$$f_{T_n}(t) = S(t, x_{n-1})\Lambda(\phi^t(x_{n-1}))$$

Within this framework, if there exists an absolutely continuous stationary distribution, then we can express the corresponding density f_X in terms of the density of the asymptotic post-disturbance distribution f_Y . In Section 6.1 we will first derive an expression for the conditional density $f_{Y_{n+1}|Y_n}$ and from this an expression for f_Y . In Section 6.2 we will then provide a formula for computing the density of the stationary distribution in terms of f_Y .

6.1 Post-Disturbance Distributions

In this section we will study the densities of the post-disturbance distributions f_{Y_n} , corresponding to the n^{th} disturbance. If the corresponding random variables Y_n converge to a continuous random variable Y with density f_Y , then we can compute f_Y from the conditional p.d.f. $f_{Y_{n+1}|Y_n}$. We use essentially the same argument as in [24].

Theorem 6.1. *Let $\{X_t\}$ be a semistochastic process with state space $[0, k)$ and infinitesimal generator given by (6.1). Suppose further that $v(x)$ is non-negative, Lipschitz, and is zero at at most a countable number of points, then the conditional p.d.f. of the $(n + 1)^{\text{st}}$ post-disturbance level Y_{n+1} conditioned on the n^{th} post-disturbance level Y_n is given by*

$$f_{Y_{n+1}|Y_n}(x_{n+1}|x_n) = \int_{\max\{x_n, x_{n+1}\}}^k dx_{n+1}^- \frac{S(\psi(x_n, x_{n+1}^-), x_n) \Lambda(x_{n+1}^-)}{v(x_{n+1}^-) x_{n+1}^-} \rho\left(\frac{x_{n+1}}{x_{n+1}^-}\right).$$

If the post-disturbance levels Y_n tend asymptotically to some continuous random variable Y , then the p.d.f. f_Y satisfies

$$f_Y(y) = \int_0^k f_{Y_{n+1}|Y_n}(y|x) f_Y(x) dx,$$

as well as the non-negativity and normalization conditions:

$$f_Y \geq 0 \quad \text{and} \quad \int_0^k f_Y(x) dx = 1.$$

Proof. Let Δx_{n+1}^- be an infinitesimal positive increment, for the computations below, we will ignore terms higher than linear order in Δx_{n+1}^- . For $x_{n+1}^- \in$

$[x_n, k)$ and $x_{n+1}^- + \Delta x_{n+1}^- \in [x_n, k)$, we have

$$\begin{aligned}
& f_{Y_{n+1}^-|Y_n}(x_{n+1}^-|x_n) \Delta x_{n+1}^- \\
&= \mathbb{P}(Y_{n+1}^- \in (x_{n+1}^-, x_{n+1}^- + \Delta x_{n+1}^-] | Y_n = x_n) \\
&= \mathbb{P}(T_{n+1} \in (\psi(x_n, x_{n+1}^-), \psi(x_n, x_{n+1}^- + \Delta x_{n+1}^-)] | Y_n = x_n) \\
&= \mathbb{P}(T_{n+1} \in (\psi(x_n, x_{n+1}^-), \psi(x_n, x_{n+1}^- + \Delta x_{n+1}^-))) \\
&= \mathbb{P}\left(T_{n+1} \in \left(\psi(x_n, x_{n+1}^-), \psi(x_n, x_{n+1}^-) + \frac{\partial \psi}{\partial x_{n+1}^-}(x_n, x_{n+1}^-) \Delta x_{n+1}^-\right)\right) \\
&= \mathbb{P}\left(T_{n+1} \in \left(\psi(x_n, x_{n+1}^-), \psi(x_n, x_{n+1}^-) + \frac{\Delta x_{n+1}^-}{v(x_{n+1}^-)}\right)\right) \\
&= \frac{S(\psi(x_n, x_{n+1}^-), x_n) \Lambda(x_{n+1}^-)}{v(x_{n+1}^-)} \Delta x_{n+1}^-,
\end{aligned}$$

hence

$$f_{Y_{n+1}^-|Y_n}(x_{n+1}^-|x_n) = \frac{S(\psi(x_n, x_{n+1}^-), x_n) \Lambda(x_{n+1}^-)}{v(x_{n+1}^-)} \mathbb{1}_{[x_n, d)}(x_{n+1}^-). \quad (6.2)$$

Recalling the definition of the multiplier random variables A_n , we can also compute the conditional p.d.f. $f_{Y_{n+1}^-|Y_n^-}$,

$$\begin{aligned}
& f_{Y_{n+1}^-|Y_n^-}(x_{n+1}^-|x_{n+1}^-) \Delta x_{n+1}^- = \mathbb{P}(Y_{n+1}^- \in (x_{n+1}^-, x_{n+1}^- + \Delta x_{n+1}^-] | Y_{n+1}^- = x_{n+1}^-) \\
&= \mathbb{P}(A_{n+1} Y_{n+1}^- \in (x_{n+1}^-, x_{n+1}^- + \Delta x_{n+1}^-] | Y_{n+1}^- = x_{n+1}^-) \\
&= \mathbb{P}(A_{n+1} x_{n+1}^- \in (x_{n+1}^-, x_{n+1}^- + \Delta x_{n+1}^-)) \\
&= \mathbb{P}\left(A_{n+1} \in \left(\frac{x_{n+1}^-}{x_{n+1}^-}, \frac{x_{n+1}^- + \Delta x_{n+1}^-}{x_{n+1}^-}\right)\right) \\
&= \rho \left(\frac{x_{n+1}^-}{x_{n+1}^-}\right) \frac{\Delta x_{n+1}^-}{x_{n+1}^-},
\end{aligned}$$

dividing by Δx_{n+1} yields

$$f_{Y_{n+1}|Y_{n+1}^-}(x_{n+1}|x_{n+1}^-) = \frac{1}{x_{n+1}^-} \rho\left(\frac{x_{n+1}}{x_{n+1}^-}\right). \quad (6.3)$$

Combining (6.3) and (6.2) we can obtain the conditional p.d.f. $f_{Y_{n+1}|Y_n}$.

$$\begin{aligned} f_{Y_{n+1}|Y_n}(x_{n+1}|x_n) &= \int dx_{n+1}^- f_{Y_{n+1}|Y_{n+1}^-}(x_{n+1}|x_{n+1}^-) f_{Y_{n+1}^-|Y_n}(x_{n+1}^-|x_n) \\ &= \int_{\max\{x_n, x_{n+1}\}}^d dx_{n+1}^- \frac{S(\psi(x_n, x_{n+1}^-), x_n)}{v(x_{n+1}^-)x_{n+1}^-} \rho\left(\frac{x_{n+1}}{x_{n+1}^-}\right). \end{aligned}$$

As for the stationary p.d.f. of the post-disturbance level, we have

$$f_{Y_{n+1}}(x_{n+1}) = \int_0^k dx_n f_{Y_{n+1}|Y_n}(x_{n+1}|x_n) f_{Y_n}(x_n),$$

which reduces to (6.1) as $f_{Y_n} \rightarrow f_Y$ and $f_{Y_{n+1}} \rightarrow f_Y$.

□

6.2 Exact Formula for Stationary Distributions

In this section we establish a formula for the stationary p.d.f. f_X for $\{X_t\}$ in terms of the stationary post-disturbance distribution f_Y . Our arguments again follow those of [24], but we arrive at a slightly different result due to an averaging error in [24].

Theorem 6.2. *Under the assumptions of Theorem 6.1, the stationary p.d.f. f_X of $\{X_t\}$ is given by*

$$f_X(x) = \frac{\mathbb{1}_{(0,k)}(x)}{v(x)} \int_0^x dx_n f_Y(x_n) \int_{\psi(x_n, x)}^{\infty} d\tau S(\tau, x_n) \Lambda(\phi^\tau(x_n))$$

where f_Y is the stationary p.d.f. of the post-disturbance levels.

Proof. The strategy is to first compute the conditional p.d.f. $f_{X|Y_n, T_{n+1}}$ of X conditioned on the values of both the n^{th} post-disturbance level and the time interval T_{n+1} between the n^{th} and $(n+1)^{\text{st}}$ disturbance. We will set $Y_n = x_n$ and $T_{n+1} = \tau_{n+1}$. We will then consider an arbitrary value x^* between x_n and x_{n+1}^- , and an infinitesimal increment $\Delta x^* > 0$ be , so that

$$x_n \leq x^* < x^* + \Delta x^* < x_{n+1}^- .$$

For τ^* and $\tau^* + \Delta\tau^*$ the times between the moment of occurrence of the n^{th} disturbance and the moment when the level has values x^* and $x^* + \Delta x^*$, respectively we have

$$0 \leq \tau^* < \tau^* + \Delta\tau^* < \tau_{n+1} .$$

Moreover, we have $\frac{\Delta x^*}{\Delta\tau^*} = v(x^*)$, and consequently

$$f_{X|Y_n, T_{n+1}}(x^* | x_n, \tau_{n+1}) \Delta x^* = \Delta\tau^* = \frac{\Delta x^*}{v(x^*)} ,$$

equivalently,

$$f_{X|Y_n, T_{n+1}}(x | x_n, \tau_{n+1}) = \frac{\mathbb{1}_{(x_n, \phi_{\tau_{n+1}}(x_n)]}(x)}{v(x)} . \quad (6.4)$$

The stationary p.d.f. f_X can then be obtained from the conditional one (6.4) by averaging over Y_n and T_{n+1} ,

$$f_X(x) = \frac{\mathbb{1}_{(0, k)}(x)}{v(x)} \int_0^x dx_n f_Y(x_n) \int_{\psi(x_n, x)}^\infty d\tau S(\tau, x_n) \Lambda(\phi^\tau(x_n)) .$$

□

6.2.1 Examples

We will now tabulate corrected versions of the calculations presented in [24]. In the table, f_X is the density of the continuous random variable X corresponding to the stationary distribution of the semistochastic process with infinitesimal generator given by

$$[\mathcal{L}f](x) = f'(x)v(x) + \lambda \int_0^1 [f(\alpha x) - f(x)]d\alpha .$$

In particular, we are assuming that the severity of the disturbances is uniformly distributed, and that the hazard function is constant, which entails that the inter-disturbance times are exponentially distributed. For the cases of $v(x) = 1$ and $v(x) = 1 - x$, stationary densities exist for all values of the parameter λ . On the other hand, when $v(x) = x(1-x)$ (logistic growth), stationary densities exist only if $\lambda < 1$. In the logistic case with $\lambda \geq 1$, the stationary distribution is a delta distribution concentrated at $x = 0$ (this is an extinction level event).

$v(x)$	f_X	$\mathbb{E}[X]$	$\text{Var}[X]$
1	$\lambda^2 x e^{-\lambda x}$	$\frac{2}{\lambda}$	$\frac{2}{\lambda^2}$
$1 - x$	$(\lambda^2 + \lambda) x(1 - x)^{\lambda-1}$	$\frac{2}{\lambda + 2}$	$\frac{2\lambda}{(\lambda + 3)(\lambda + 2)^2}$
$x(1 - x)$	$\frac{\sin(\pi\lambda)}{\pi} \frac{(1 - x)^{\lambda-1}}{x^\lambda}$	$-(\lambda - 1)$	$\frac{-\lambda(\lambda - 1)}{2}$

Table 6.1: Statistics for densities of stationary distributions.

Bibliography

- [1] K. B. Athreya, D. McDonald, P. Ney. Limit theorems for semi-Markov processes and renewal theory for Markov chains. *The Annals of Probability* (1978): 788–797.
- [2] K. B. Athreya, P. Ney. A new approach to the limit theory of recurrent Markov chains. *Transactions of the American Mathematical Society* **245** (1978): 493–501
- [3] R. Bartoszyński. On the risk of rabies. *Mathematical Biosciences* **24:3–4** (1975) 355–377.
- [4] B. Beckage, W. J. Platt, L. J. Gross. Vegetation, fire, and feedbacks: a disturbance-mediated model of savannas. *The American Naturalist* **174:6** (2009) 805–818.
- [5] B. Bond-Lamberty, S. D. Peckham, D. E. Ahl, S. T. Gower. Fire as the dominant driver of central Canadian boreal forest carbon balance. *Nature* **450** (2007) 89–92.
- [6] P. J. Brockwell, J. Gani, S. I. Resnick. Birth, immigration and catastrophe process. *Advances in Applied Probability* **14:4** (1982) 709–731.
- [7] P. J. Brockwell, J. M. Gani, S. I. Resnick. Catastrophe processes with continuous state-space. *Australian Journal of Statistics* **25:2** (1983) 208–226.
- [8] B. J. Cairns. Evaluating the expected time to population extinction with semi-stochastic models. *Mathematical Population Studies* **16:3** (2009) 199–220.
- [9] J. S. Clark. Ecological disturbance as a renewal process: theory and application to fire history. *Oikos* **56:1** (1989) 17–30.

- [10] J. N. Corcoran, R. L. Tweedie. Perfect sampling from independent Metropolis-Hastings chains. *Journal of statistical planning and inference* **104**:2 (2002): 297–314.
- [11] D. R. Cox, D. Oakes. *Analysis of Survival Data*. Chapman and Hall, London, 1984.
- [12] J.I. Doob. *Stochastic Processes*. Wiley, New York, 1953.
- [13] M. H. A. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society B* **46**:3 (1984) 353–388.
- [14] G. Gripenberg. A stationary distribution for the growth of a population subject to random catastrophes. *Journal of Mathematical Biology* **17** (1983): 371–379.
- [15] G. Gripenberg. Extinction in a model for the growth of a population subject to catastrophes. *Stochastics: An International Journal of Probability and Stochastic Processes* **14**:2 (1985): 149–163.
- [16] F. B. Hanson, D. Ryan. Optimal harvesting with exponential growth in an environment with random disasters and bonanzas. *Mathematical Biosciences* **74**:1 (1985): 37–57.
- [17] F. B. Hanson, D. Ryan. Optimal harvesting of a logistic population in an environment with stochastic jumps. *Journal of Mathematical Biology* **24**:3 (1986): 259–277.
- [18] F. B. Hanson, H. C. Tuckwell. Persistence times of populations with large random fluctuations. *Theoretical Population Biology* **14**:1 (1978) 46–61.
- [19] F. B. Hanson, H. C. Tuckwell. Logistic growth with random density independent disasters. *Theoretical Population Biology* **19**:1 (1981) 1–18.
- [20] F. B. Hanson, H. C. Tuckwell. Population growth with randomly distributed jumps. *Journal of Mathematical Biology* **36**:2 (1997) 169–187.
- [21] S. Karlin. *A First Course in Stochastic Processes*. Academic Press, New York, 1966.
- [22] S. Karlin, H. M. Taylor. *A Second Course in Stochastic Processes*. Gulf Professional Publishing, Houston , 1981.
- [23] R. Lande. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist* **142**:6 (1993) 911–927.

- [24] M. C. A. Leite, N. P. Petrov, E. Weng. Stationary distributions of semistochastic processes with disturbances at random times and with random severity. *Nonlinear Analysis: Real World Applications* **13:2** (2012): 497–512.
- [25] S. P. Meyn, R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Annals of Applied Probability* **4** (1994):981–1011
- [26] S. P. Meyn, R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag, London, 1993.
- [27] T. Nakigawa. *Stochastic Processes: with Applications to Reliability Theory*. Springer Science and Business Media, London, 2011.
- [28] E. Nummelin. A splitting technique for Harris recurrent Markov chains. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **43** (1978): 309–318
- [29] E. Nummelin. General irreducible Markov chains and non-negative operators. Cambridge University Press, Cambridge, 1984.
- [30] A. G. Pakes, A. C. Trajstman, P. J. Brockwell. A stochastic model for a replicating population subjected to mass emigration due to population pressure. *Mathematical Biosciences* **45:1–2** (1979) 137–157.
- [31] E. Parzen. *Stochastic Processes*. Society for Industrial and Applied Mathematics, Philadelphia, 1999.
- [32] K. S. Pregitzer, E. S. Euskirchen. Carbon cycling and storage in world forests: biome patterns related to forest age. *Global Change Biology* **10:12** (2004) 2052–2077.
- [33] D. H. Reed, J. J. O’Grady, J. D. Ballou, R. Frankham. The frequency and severity of catastrophic die-offs in vertebrates. *Animal Conservation* **6:2** (2003) 109–114.
- [34] G. O. Roberts, J. S. Rosenthal. Quantitative bounds for convergence rates of continuous time Markov processes. *Electronic Journal of Probability* **1:9** (1996): 1–21
- [35] W.H. Romme, et al. Are large, infrequent disturbances qualitatively different from small, frequent disturbances? *Ecosystems* **1:6** (1998): 524–534.
- [36] J. S. Rosenthal. Minorization conditions and convergence rates for Markov chain Monte Carlo. *Journal of the American Statistical Association* **90:430** (1995): 558–566.

- [37] S. W. Running. Ecosystem disturbance, carbon, and climate. *Science* **321** (2008) 652–653.
- [38] P. E. Thornton, B. E. Law, H. L. Gholz, K. L. Clark, E. Falge, D. S. Ellsworth, A. H. Goldstein, R. K. Monson, D. Hollinger, M. Falk, J. Chen, J. P. Sparks. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. *Agricultural and Forest Meteorology* **113**:1 (2002) 185–222.
- [39] W. Whitt. The stationary distribution of a stochastic clearing process. *Operations Research* **29**:2 (1981) 294–308.