

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

McKinlay, Shaun Antony

Title:

From sojourn times and boundary crossings to iterated random functions

Date:

2015

Persistent Link:

<https://hdl.handle.net/11343/92277>

From Sojourn Times and Boundary Crossings to Iterated Random Functions

Shaun Antony McKinlay
ORCID iD: 0000-0003-4292-2591

Submitted in total fulfilment of the requirements of
the degree of Doctor of Philosophy

November 2015

*School of Mathematics and Statistics
The University of Melbourne*

Abstract

This thesis was motivated by related sojourn time and boundary crossing problems for the Brownian motion process. In Chapter 1, we discuss these motivating problems, and outline how they led naturally to the research in subsequent chapters. The available literature relevant to this thesis is reviewed in Chapter 2.

In Chapter 3, we extend the uniform law for sojourn times of processes with cyclically exchangeable increments to the case of random fields, with general parameter sets, that possess a suitable invariance property.

In Chapter 4, we consider several special cases of iterations of random i.i.d. linear functions with beta distributed fixed points that describe a nested interval scheme when iterated in a backward direction, and an ergodic Markov chain in the forward direction. We find these points by extending a result due to E. J. G. Pitman (1937), which allows us to solve the random equation by solving an equivalent random equation with a gamma distributed solution. In each case, these fixed points are the almost sure limit of the corresponding nested interval scheme and the stationary distribution of the corresponding Markov chain. Using our new technique, we also derive some new special cases of randomly weighted averages, and show how our results define some new stick-breaking constructions of beta distributed random variables.

In Chapter 5, we consider a class of discrete time Markov chains with state space $[0, 1]$ and the following dynamics. At each time step, first the direction of the next transition is chosen at random with probability depending on the current location. Then the length of the jump is chosen independently as a random proportion of the distance to the respective end point of the unit interval, the distributions of the proportions being fixed for each of the two directions. Under simple broad

conditions, we establish the ergodicity of such Markov chains and then derive closed form expressions for the stationary densities of the chains when the proportions are beta distributed with the first parameter equal to 1. Examples demonstrating the range of stationary distributions for processes described by this model are given, and an application to a robot coverage algorithm is discussed.

In Chapter 6, using our extension of the result due to Pitman (1937) from Chapter 4, we characterise the class of distributions of random stochastic matrices X with the property that the products $X(n)X(n-1)\cdots X(1)$ of i.i.d. copies $X(k)$ of X converge a.s. as $n \rightarrow \infty$ and the limit is Dirichlet distributed. This extends a result by Chamayou and Letac (1994) and is illustrated by several examples that are of interest in applications.

In Chapter 7, we analyse the stability of the boundary crossing probability for the multivariate Brownian motion process, with respect to small changes of the boundary. Under broad assumptions on the nature of the boundary, we obtain an analogue of the Borovkov and Novikov (2005) upper bound for the difference between boundary hitting probabilities for “close boundaries” in the univariate case. We also obtain upper bounds for the first boundary crossing time densities.

Declaration

This is to certify that:

1. the thesis comprises only my original work towards the PhD except where indicated in the Preface,
2. due acknowledgement has been made in the text to all other material used,
3. the thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Signed,

A handwritten signature in black ink, consisting of a stylized 'S' followed by a horizontal line and a small dot.

Shaun Antony McKinlay

Preface

The work contained in this thesis was carried out under the supervision of Professor Konstantin Borovkov at the School of Mathematics and Statistics, The University of Melbourne. This original results in this thesis are mostly contained in the published papers [BM12], [MB15], [M14], [MB16], and the submitted paper [M15].

Some of the text from the Introduction and Literature Review has already appeared in these papers. In particular, the first two paragraphs of Section 1.2.1 appear in [BM12], Section 1.2.3 is mostly contained in [MB15], and Section 2.5.4 is from [M14].

The work in Chapter 3 on the uniform law for sojourn measures of random fields was carried out in collaboration with K. Borovkov and also appear in [BM12]. In particular, the idea for the proof of the main result (presented as Theorem 3.1 below) is due to K. Borovkov. My contribution to that paper was 50%.

Some of the results in Chapter 4 on the beta distributed limits of iterated linear random functions appear in [M15]. This chapter also contains a result (presented as Lemma 4.1 below) from [M14].

Chapter 5 on the explicit form of the stationary distributions for a class of bounded Markov chains is mostly from the paper [MB16] that was co-authored by K. Borovkov. In particular, K. Borovkov contributed to the proof of Lemma 5.1 and to Example 5.3.1. My contribution to the paper was 80%.

The work on the limiting distributions of products of i.i.d. stochastic matrices (Chapter 6) also appears in [M14].

The work contained in Chapter 7 on boundary crossing probabilities for the multivariate Brownian motion process was carried out in collaboration with K. Borovkov, and is published in [MB15]. Specifically, K. Borovkov contributed several ideas to

the approach and methods that were used, and the proof of the Corollary 7.2 below. I also contributed to the approach and methods, and was primarily responsible for the technical lemmas. My contribution to the paper was 60%.

None of the work appearing in this thesis has been submitted for any other qualifications, nor was any part of it carried out prior to PhD candidature enrolment.

Acknowledgements

To my supervisor, Kostya Borovkov, thank you for your the countless hours you have spent with me during this project. Needless to say, this thesis would not have been possible without your support and encouragement over the years. Working with you has been as enjoyable as it has been challenging.

I would also like to thank Aihua Xia and Owen Jones for being on my supervisory panel, and for their integral role in the stimulating stochastic processes research group, and in particular for organising our weekly reading group.

To the administrative staff at the School of Mathematics and Statistics, and in particular Kirsten Hoak, thank you for your friendly and helpful assistance with the many administrative tasks over the years.

I am grateful to the Australian Government's Australian Postgraduate Award program, the ARC Centre of Excellence for Mathematics and Statistics of Complex Systems, the Maurice Belz Trust, The University of Melbourne, and ARC Discovery Projects DP120102398 and DP150102758 for providing funding for this research.

I am also grateful to the School of Mathematical Sciences at Queen Mary, University of London, for providing a visiting position during the latter stages of my research.

Lastly, I would like to thank to my family and friends, and especially my parents and Lucy.

Contents

List of Figures	ix
Notation	xi
1 Introduction	1
1.1 Preliminaries	4
1.1.1 Brownian motion and related processes	4
1.1.2 Distributions and special functions	6
1.1.3 Iterated random functions	10
1.2 The three motivating problems	11
1.2.1 An equidistribution theorem	11
1.2.2 The arcsine laws for Brownian motion	14
1.2.3 Boundary crossing approximations	21
2 Literature Review	23
2.1 Sojourn times for bridge processes	24
2.1.1 Lévy processes	24
2.1.2 Processes with exchangeable increments	24
2.1.3 Processes with cyclically exchangeable increments	25
2.2 Nested interval schemes on the unit interval	26
2.2.1 Split and choose the longest subinterval w.p. p	26
2.2.2 Other schemes	30
2.3 Markov chains on the unit interval	31
2.3.1 Stochastic give-and-take	31
2.3.2 A random donkey walk in $[0, 1]$	33

2.3.3	An extension of the random donkey walk in $[0, 1]$	36
2.4	Randomly weighted averages	37
2.5	Products of random stochastic matrices	41
2.5.1	2×2 stochastic matrices	41
2.5.2	$d \times d$ stochastic matrices	42
2.5.3	Random nested simplices	45
2.5.4	A random donkey walk in $\mathcal{S}_d$	47
2.6	Boundary crossing probabilities for Brownian motion	48
2.6.1	Exact solutions	49
2.6.2	Univariate approximation rates	49
3	The uniform law for sojourn measures of random fields	55
3.1	Introduction	56
3.2	The main result	57
3.3	Examples	59
4	Beta distributed limits of iterated linear random functions	62
4.1	Introduction	63
4.2	Main results	66
4.2.1	A general nested interval scheme	66
4.2.2	A relationship between beta and gamma distributed solutions to linear random equations	67
4.2.3	Some equivalent claims	70
4.2.4	Model 1	71
4.2.5	Model 2	79
4.3	Other applications	83
4.3.1	Randomly weighted averages	83
4.3.2	Stick-breaking constructions	87
5	Stationary distributions for a class of bounded Markov chains	90
5.1	Introduction	91
5.2	Main results	92
5.3	Examples	100

5.3.1	The case of polynomial $p(x)$	100
5.3.2	The case of piecewise constant $p(x)$	101
5.3.3	A robot coverage algorithm and random search	103
6	Products of i.i.d. stochastic matrices	108
6.1	Introduction	109
6.2	Main results	110
6.3	Proofs	112
6.4	Examples and applications	116
6.4.1	Random exchange models	116
6.4.2	Random nested simplices	120
6.4.3	Service networks with polling	121
7	Boundary crossing probabilities for multivariate Brownian motion	125
7.1	The main results	126
7.2	Proofs	129
	Bibliography	144

List of Figures

1.1	A sample path of the Empirical Process $E_n(t)$ for $n = 5$	12
3.1	A sample realisation of the field (3.9) with kernel (3.10), $d = 3$ and $n = 20$ (on the left pane) and the scaled histogram of a simulated sample of 10^5 independent copies of $F_\mu(X, X(a))$ for that field, a being the “North pole” (on the right pane).	60
5.1	Some examples of peaked stationary densities generated by piecewise constant $p(x)$ with $k = 2$	102
5.2	Some examples of bimodal stationary densities generated by piecewise constant $p(x)$ with $k = 2$	103
5.3	Some examples of trimodal stationary densities generated by piecewise constant $p(x)$, where $k = 6$, $\mathbf{p} := (p_1, \dots, p_6)$, $s_1 = 1/6$, $s_2 = 1/3$, $s_3 = 1/2$, $s_4 = 2/3$, $s_5 = 5/6$	104
5.4	The stationary density of the robot converge algorithm with p_i given by (5.23), $l_1 = r_1 = l_2 = r_2 = 3$, and $d_1 = d_2 = 1$	105

Notation

$\stackrel{a.s.}{=}$	Almost surely equality
$\stackrel{a.s.}{\rightarrow}$	Almost sure convergence
$\stackrel{d}{=}$	Equality in distribution
$A \sim F$	Random element A has distribution F , p. 7
$A^{(v)}$	Set A dilated (eroded) by $v > 0$ ($v < 0$), p. 126
A^c	Complement of set A
A^T	Transpose of matrix/vector A
$\ \cdot\ $	Euclidean norm
a.s.	Almost surely
$B_r(\boldsymbol{x})$	Open ball of radius r and centre $\boldsymbol{x}$, p. 126
$B(a, b)$	Beta function with parameters (a, b) , p. 8
$\{B_t\}$	Univariate Brownian bridge, p. 5
$\{\boldsymbol{B}_t\}$	m -dimensional Brownian bridge, p. 6
$Be(p)$	Bernoulli distribution with success probability p , p. 15
$\mathcal{B}(A)$	The class of Borel subsets of A
$\beta(a, b)$	Beta distribution with parameters (a, b) , p. 8
$D_{\boldsymbol{a}}$	Dirichlet distribution with parameter vector $\boldsymbol{a} \in \mathbb{R}_+^d$, p. 8
D_A	Dirichlet distribution with parameter matrix $A \in \mathcal{M}_{r,d}$, p. 9
δ_{ij}	Kronecker's delta, p. 122
∂A	Boundary of A , p. 21
$f \circ g$	Composition of functions f and g , $f(g(\cdot))$
$G_{(a_1, \dots, a_d)}$	The distribution $\Gamma_{a_1} \otimes \dots \otimes \Gamma_{a_d}$, p. 9
$\mathcal{G}$	The class of open sets $G \subset [0, T] \times \mathbb{R}^m$ containing $(0, \mathbf{0})$, p. 21
$\mathcal{G}_{K, \beta, \gamma}$	The subset of $\mathcal{G}$ satisfying [G1]–[G3], p. 127

$\{\gamma(t)\}$	The gamma process, p. 76
$\Gamma(a)$	Gamma function with parameter a , p. 7
Γ_a	Gamma distribution with shape parameter a and scale parameter 1, p. 7
H_t	The time t cross-section of the set $H \subset [0, T] \times \mathbb{R}^m$, p. 126
$\text{id}(\cdot)$	The identity function, p. 10
iff	If and only if
i.i.d.	Independent and identically distributed
$\mathbf{1}_A$	Indicator function of the event A , p. 12
$\mathbb{I}_n$	The n th interval of a nested interval scheme, p. 63
$\mathcal{K}_d$	The class of distributions on $\mathcal{P}_d$ satisfying [D1] – [D3] , p. 109
$\mathcal{L}(X)$	The distribution of X , p. 7
$\mathcal{L}(X) \otimes \mathcal{L}(Y)$	Law of (X, Y) for X and Y independent, p. 7
$\mathcal{M}_{r,c}$	Set of $r \times c$ matrices with non-negative entries, p. 6
$\mathcal{M}_{r,c}^+$	Set of $r \times c$ matrices with positive entries, p. 6
$\hat{\mu}$	The distribution of $(A, B) \in [0, 1]^2$, p. 18
$\mathbb{N}$	Set of positive integers
$\mathcal{P}_{r,c}$	Set of $r \times c$ stochastic matrices, p. 6
$\mathcal{P}_d$	Set of $d \times d$ stochastic matrices, p. 6
$\mathcal{P}_d^+$	Set of $d \times d$ stochastic matrices with positive entries, p. 6
$\sigma(X)$	σ -algebra generated by X
$\sigma(X_t, t \in T)$	σ -algebra generated by random variables X_t , $t \in T$
$P(G)$	Boundary non-crossing probability of G for $\{\mathbf{W}_t\}$, p. 21
$\Phi(x)$	Standard normal distribution function
$\bar{\Phi}(x)$	The tail of the standard normal distribution function
$\phi(x)$	Standard normal density function
$\{R_t\}$	The m -dimensional Bessel process, p. 5
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of positive real numbers
$\rho_h(A, B)$	Hausdorff distance between sets A and B , p. 127
$\rho_H(A, B)$	A “Hausdorff-type” distance between sets A and B , p. 127
$\mathcal{S}_d$	The standard d -dimensional simplex, p. 8
$\mathbb{S}^{n-1}$	The unit sphere in $\mathbb{R}^n$ with centre at $\mathbf{0}$

$T_r(\mathbf{x})$	The identity matrix with the r th row replace by vector $\mathbf{x}$, p. 122
$U[a, b]$	Uniform distribution on $[a, b]$
$\{W_t\}$	Univariate Brownian motion, p. 4
$\{\mathbf{W}_t\}$	m -dimensional Brownian motion, p. 5
$X(1, n)$	The right product of stochastic matrices $X(1) \cdots X(n)$, p. 43
$X(n, 1)$	The left product of stochastic matrices $X(n) \cdots X(1)$, p. 43
$X^{(i)}$	The i th row of matrix X , p. 9
$Z_{i\bullet}$	The i th row sum of matrix $(Z_{i,j})$, p. 7
$Z_{\bullet j}$	The j th column sum of matrix $(Z_{i,j})$, p. 7
$z_{\bullet}$	The sum of the components of a vector $(z_1, \dots, z_r)$, p. 7

Chapter 1

Introduction

This thesis presents the results of a research project motivated by three related sojourn time and boundary crossing problems for the Brownian motion process. A notable feature of our results is that the solution to one of these problems led us in a natural way to study random equations and iterated random functions.

Many of the results we present extend known univariate results to a multidimensional setting. In addition, we will often encounter beta distributions (and Dirichlet distributions in the multidimensional setting) as the laws of the positive (or negative) sojourn times for the Brownian motion process and associated conditioned processes such as the Brownian bridge, and as the limiting distribution in a number of special cases of iterated random functions.

The first two motivating problems concern two well-known results on the positive (or negative) sojourn time for the standard Brownian motion and bridge processes on the unit interval. In the first, we explore the essential properties of the Brownian bridge process that ensure that its sojourn time below the axis is uniformly distributed. This problem had actually been studied in several papers recently, where it was shown that this uniform law is a property common to processes with cyclically exchangeable increments (CEI) satisfying an additional condition ensuring that the time spent at zero has zero Lebesgue measure. We sought a multivariate extension of this result, obtaining a uniform law for sojourn measures of random fields with general parameter sets that possess a suitable analogue of the CEI property in the random fields setting.

Our second problem concerns the positive sojourn time of the standard Brownian motion process. In particular, we searched for a simple probabilistic argument relating that random time to the time of the last zero before time 1 (or, equivalently, the maximum up to time 1). Apparently the simplest argument was that used by P. Lévy [L39] in his original proof that the positive sojourn time has the arcsine law. That approach involved decomposing the positive sojourn time of the Brownian path into the segments before and after the last zero.

Our contributions originated by observing that Lévy's decomposition yields a random equation with an arcsine distributed solution. Extending a result due to E. J. G. Pitman [P37], we solve the random equation by solving an equivalent random equation with a gamma distributed solution. Observing that the backward iterations of the random equation describe a nested interval scheme, while forward iterations generate an ergodic Markov chain, we obtain that the limiting distributions of these processes is the arcsine distributed solution to the original random equation. Generalising the original random equation in several ways, and armed with our new technique, we prove a number of new results on the limiting distributions of such processes and generalisations thereof.

In our final problem, we consider boundary crossing probabilities for the multivariate Brownian motion process. The computation of these probabilities arise in a number of applied problems, most notably in mathematical finance, where such probabilities are used, for instance, in barrier options' pricing. However, even in the univariate case, closed form expressions for boundary crossing probabilities are only available in a few special cases. An alternative approach is to approximate the boundary crossing probability by computing that probability for another boundary for which the computation is feasible. To use the approximation, one must bound the difference between boundary crossing probabilities for the original boundary on the one hand side and the approximating boundary on the other. The main goal in the last problem was to provide such a bound.

The thesis is structured as follows. In the next section we introduce some key definitions and results required for a more detailed discussion of our motivating sojourn time and boundary crossing problems in Section 1.2.

In Chapter 2 we survey the available literature relevant to this thesis. We begin

by reviewing extensions of the uniform law for the sojourn time above the axis of the Brownian bridge process to wider classes of processes. In Section 2.2 we discuss several special cases of the nested interval schemes on the unit interval that have appeared previously, while some relevant special cases of Markov chains on the unit interval are discussed in Section 2.3. Sections 2.4 and 2.5 survey the literature on randomly weighted averages and products of i.i.d. random stochastic matrices that are related to the random equation from our second problem. Section 2.6 is devoted to the literature review related to our last problem. First we review exact solutions for boundary crossing probabilities for the multidimensional Brownian motion. Then we consider some approximation methods for the univariate Brownian motion that have appeared previously.

The original research presented in this thesis is contained in Chapters 3–7. In Chapter 3 we solve our first problem by providing an extension of the uniform law for the negative sojourn time of processes with cyclically exchangeable increments to the case of random fields.

Chapters 4–6 contain the results of research related to our second problem. In Chapter 4, we consider several special cases of a class of random equations that includes the original equation obtained from Lévy’s decomposition of the Brownian motion, and study the nested interval schemes and Markov chains generated in each case. In Chapter 5 we consider a more general version of the Markov chains generated by the forward iterations from Chapter 4, while in Chapter 6 we study the forward and backward products of i.i.d. random stochastic matrices that arise as the natural multivariate extension of the iterations from Chapter 4.

Chapter 7 is devoted to our third problem, where we provide a bound for the difference between boundary crossing probabilities of two “close boundaries” for the multivariate Brownian motion process.

The content of these chapters and how they arose from the motivating problems introduced above are discussed further in Section 1.2 below.

1.1 Preliminaries

In this section we briefly review the material from probability theory that will be fundamental to the sequel. This short review is intended to introduce the notation, objects and some key results that we will require. As such, we assume the reader is familiar with the fundamentals of probability theory and stochastic processes, stochastic calculus and the Brownian motion process (see e.g. [F71], [O07], and [MP10], respectively).

1.1.1 Brownian motion and related processes

A continuous time real-valued process $\{W_t\}_{t \geq 0}$ is called a Brownian motion started at $x \in \mathbb{R}$ if the following properties hold (see e.g. p.7 in [MP10]):

- (i) $W_0 = x$,
- (iii) for $0 \leq s \leq t$, $W_t - W_s$ is independent of $\sigma(W_u, u \leq s)$,
- (ii) for $0 \leq s \leq t$, $W_t - W_s$ is normally distributed with mean 0 and variance $t - s$,
- (iv) the sample paths of $\{W_t\}$ are continuous almost surely (a.s.).

We call the process $\{W_t\}_{t \geq 0}$ a standard Brownian motion if $x = 0$.

There is a vast literature devoted to the study of functionals of the Brownian motion process (see e.g. [BS02]). One important functional is the time that a standard Brownian motion first crosses an upper (or lower) linear boundary. The distribution and density of that time is known explicitly. Specifically, for $c \in \mathbb{R}$ and $\varepsilon, t, z > 0$, and denoting by Φ the standard normal distribution function, one has

$$\mathbb{P}\left(\sup_{0 \leq s \leq t} (W_s - cs) \leq \varepsilon\right) = \Phi(c\sqrt{t} + \varepsilon/\sqrt{t}) - e^{-2c\varepsilon} \Phi(c\sqrt{t} - \varepsilon/\sqrt{t}) \quad (1.1)$$

and

$$\mathbb{P}\left(\inf\{s : W_s - cs = z\} \in dt\right) = \frac{z}{\sqrt{2\pi t^3/2}} \exp\left(\frac{-(z + ct)^2}{2t}\right) dt \quad (1.2)$$

(see e.g. 1.1.4 on p.250 and 2.0.2 on p.295 of [BS02], respectively).

An important functional of the Brownian motion process is its running maximum $\{\overline{W}_t\}_{t \geq 0}$ given by

$$\overline{W}_t := \max_{0 \leq s \leq t} W_s, \quad t \geq 0. \quad (1.3)$$

We will similarly denote by $\{\overline{B}_t\}$ the maximum process of the Brownian bridge $\{B_t\}_{0 \leq t \leq l}$ which can be defined as follows. For $x, y \in \mathbb{R}$, the Brownian bridge from x to y of length $l > 0$ is the process $\{B_t\}_{0 \leq t \leq l}$ given by

$$B_t := W_t - \frac{t}{l}(W_l - y), \quad 0 \leq t \leq l. \quad (1.4)$$

Recall that the Brownian bridge can also be defined as the Brownian motion on $[0, l]$ conditioned on $W_0 = x$ and $W_l = y$. We call the process $\{B_t\}$ a standard Brownian bridge if $(x, y, l) = (0, 0, 1)$.

The last problem we consider in this thesis concerns the boundary crossing probabilities for the multidimensional Brownian motion process, which is defined as follows. Let $W^{(1)}, \dots, W^{(m)}$ be m independent Brownian motions started at $x_1, \dots, x_m$, respectively. The process $\{\mathbf{W}_t\}_{t \geq 0}$ defined by

$$\mathbf{W}_t := (W_t^{(1)}, \dots, W_t^{(m)}) \quad (1.5)$$

is called an m -dimensional Brownian motion process started at $\mathbf{x} = (x_1, \dots, x_m)$. If $\mathbf{x} = \mathbf{0} := (0, \dots, 0)$, we call it a standard m -dimensional Brownian motion.

A process closely related to the m -dimensional Brownian motion is the m -dimensional Bessel process $\{R_t\}_{t \geq 0}$ defined as the radial part of $\mathbf{W}_t$:

$$R_t := \sqrt{(W_t^{(1)})^2 + (W_t^{(2)})^2 + \dots + (W_t^{(m)})^2}, \quad t \geq 0. \quad (1.6)$$

It satisfies the stochastic differential equation (SDE)

$$dR_t = \frac{m-1}{2R_t} dt + dW_t, \quad t > 0, \quad R_0 = \|\mathbf{x}\| > 0, \quad (1.7)$$

where $\{W_t\}_{t \geq 0}$ is a standard univariate Brownian motion process (see e.g. p.148 in [O07]; we assume that the reader is familiar with the fundamentals of Itô stochastic calculus)

Similar to the above, we define the m -dimensional Brownian bridge $\{\mathbf{B}_t\}_{0 \leq t \leq l}$ from $\mathbf{x} \in \mathbb{R}^m$ to $\mathbf{y} \in \mathbb{R}^m$ of length l by

$$\mathbf{B}_t := \mathbf{W}_t - \frac{t}{l}(\mathbf{W}_l - \mathbf{y}), \quad 0 \leq t \leq l. \quad (1.8)$$

It is the solution the SDE

$$d\mathbf{B}_t = \frac{\mathbf{y} - \mathbf{B}_t}{l - t} dt + d\mathbf{W}_t, \quad 0 < t < l, \quad (1.9)$$

(see e.g. p.64 in [BS02]).

For $t > 0$, we define the time-reversed Brownian motion $\{\mathbf{X}_s\}_{0 \leq s \leq t}$ by

$$\mathbf{X}_s := \mathbf{W}_{t-s} - \mathbf{W}_t, \quad 0 \leq s \leq t.$$

Note that $\{\mathbf{X}_s\}_{0 \leq s \leq t}$ is a standard m -dimensional Brownian motion on $[0, t]$ since it is continuous and its finite dimensional distributions are Gaussian with the required covariances.

1.1.2 Distributions and special functions

For positive integers r and c , denote by $\mathcal{M}_{r,c}$ the set of $r \times c$ matrices $(p_{i,j})_{i=1}^r_{j=1}^c$ such that all $p_{i,j} \geq 0$, and by $\mathcal{P}_{r,c} \subset \mathcal{M}_{r,c}$ its subclass of matrices with $\sum_{j=1}^c p_{i,j} = 1$, $i = 1, \dots, r$. Let $\mathcal{M}_{r,c}^+$ and $\mathcal{P}_{r,c}^+$ be the subclasses of all positive matrices from $\mathcal{M}_{r,c}$ and $\mathcal{P}_{r,c}$, respectively (by a positive matrix/vector we mean a matrix/vector with all positive entries), and set $\mathbb{R}_+ := (0, \infty)$. Clearly, $\mathcal{P}_d := \mathcal{P}_{d,d}$ consists of all $d \times d$ stochastic matrices, with $\mathcal{P}_d^+ := \mathcal{P}_{d,d}^+$ its subclass of positive stochastic matrices. We similarly define $\mathcal{M}_d := \mathcal{M}_{d,d}$ and $\mathcal{M}_d^+ := \mathcal{M}_{d,d}^+$. For a matrix $(Z_{i,j}) \in \mathcal{M}_{r,c}$, we denote its row sums by

$$Z_{i\bullet} := \sum_{j=1}^c Z_{i,j}, \quad i = 1, \dots, r,$$

and its column sums by

$$Z_{\bullet j} := \sum_{i=1}^r Z_{i,j}, \quad j = 1, \dots, c.$$

Similarly, for a vector $(Y_1, \dots, Y_c)$, we denote the sum of its components by

$$Y_{\bullet} := \sum_{i=1}^c Y_i.$$

For a random element X , we denote its distribution by $\mathcal{L}(X)$ and write $X \sim F$ if $F = \mathcal{L}(X)$. If $Y \sim F_Y$ is independent of $X \sim F_X$, we write $F_X \otimes F_Y$ for the law of (X, Y) .

We denote by Γ_a the gamma distribution with scale parameter 1 and shape parameter $a > 0$, with density

$$\frac{u^{a-1} e^{-u}}{\Gamma(a)}, \quad u > 0,$$

where $\Gamma(a)$ is the gamma function. An important feature of the gamma distribution that we will use in this thesis is that its characteristic function has a very simple closed form. Specifically, for $X \sim \Gamma_a$, one has

$$\mathbb{E} e^{itX} = (1 - it)^{-a}, \quad t \in \mathbb{R}. \quad (1.10)$$

A well-known property of the gamma distribution is that the sums of independent gamma distributed random variables with the same scale parameter are also gamma distributed. In particular, if

$$(X_1, \dots, X_d) \sim \Gamma_{a_1} \otimes \dots \otimes \Gamma_{a_d} \quad \text{then} \quad X_{\bullet} \sim \Gamma_{a_{\bullet}}. \quad (1.11)$$

An important characterisation of the gamma distribution is as follows (cf. [L55]).

Theorem 1.1 ([L55]). *Suppose that $X_1, X_2 > 0$ are independent non-degenerate random variables. Then X_1 and X_2 are gamma distributed with a common scale parameter if and only if $X_1 + X_2$ is independent of $X_1/(X_1 + X_2)$.*

It follows that for $(X_1, X_2) \sim \Gamma_{a_1} \otimes \Gamma_{a_2}$, one has

$$X_1 \stackrel{d}{=} \frac{X_1}{X_1 + X_2}(\tilde{X}_1 + \tilde{X}_2), \quad (1.12)$$

where $(\tilde{X}_1, \tilde{X}_2)$ is an independent copy of (X_1, X_2) .

For parameters $a_1, a_2 > 0$, we denote by $\beta(a_1, a_2)$ the beta distribution with density

$$B(a_1, a_2)^{-1} u^{a_1-1} (1-u)^{a_2-1}, \quad 0 < u < 1,$$

where $B(a_1, a_2)$ is the beta function defined by

$$B(a_1, a_2) := \int_0^1 u^{a_1-1} (1-u)^{a_2-1} du.$$

A well-known relation between beta and gamma random variables is as follows. For $(X_1, X_2) \sim \Gamma_{a_1} \otimes \Gamma_{a_2}$, one has

$$\frac{X_1}{X_1 + X_2} \sim \beta(a_1, a_2).$$

For a positive vector $\mathbf{a} = (a_1, \dots, a_d)$, we denote by $D_{\mathbf{a}}$ the Dirichlet distribution on the standard simplex $\mathcal{S}_d := \mathcal{P}_{1,d}$ with density

$$\Gamma(\mathbf{a}_{\bullet}) \prod_{i=1}^d \frac{u_i^{a_i-1}}{\Gamma(a_i)}, \quad (u_1, \dots, u_d) \in \mathcal{S}_d,$$

with respect to $d^{-1/2}\nu$, where ν is the $(d-1)$ -dimensional Lebesgue measure on $\mathcal{S}_d$. Alternatively, one can represent this distribution as having density

$$\Gamma(\mathbf{a}_{\bullet}) \prod_{i=1}^d \frac{u_i^{a_i-1}}{\Gamma(a_i)}, \quad u_d := 1 - \sum_{i=1}^{d-1} u_i,$$

on the projection

$$\left\{ (u_1, \dots, u_{d-1}) \in \mathbb{R}_+^{d-1} : \sum_{i=1}^{d-1} u_i \leq 1 \right\}$$

of $\mathcal{S}_d$ on the $(d-1)$ -dimensional space spanned by the first $(d-1)$ coordinates, with respect to the $(d-1)$ -dimensional Lebesgue measure.

We denote by $G_{\mathbf{a}}$ the distribution $\Gamma_{a_1} \otimes \cdots \otimes \Gamma_{a_d}$ on $\mathcal{M}_{1,d}$. It follows from Theorem 1.1 that, for $(X_1, \dots, X_d) \sim \Gamma_{a_1} \otimes \cdots \otimes \Gamma_{a_d}$, one has that

$$X_{\bullet} \text{ is independent of } \left(\frac{X_1}{X_{\bullet}}, \dots, \frac{X_d}{X_{\bullet}} \right), \quad (1.13)$$

and so

$$(X_1, \dots, X_d) \stackrel{d}{=} \left(\frac{X_1}{X_{\bullet}}, \dots, \frac{X_d}{X_{\bullet}} \right) \tilde{X}_{\bullet}, \quad (1.14)$$

where $(\tilde{X}_1, \dots, \tilde{X}_d)$ is an independent copy of $(X_1, \dots, X_d)$.

For $A = (a_{i,j}) \in \mathcal{M}_{r,c}^+$, the Dirichlet distribution D_A on $\mathcal{P}_{r,c}$ is the law of the matrix $(X_{i,j}) \in \mathcal{P}_{r,c}$, such that

$$X^{(i)} := (X_{i,1}, \dots, X_{i,c}) \sim D_{(a_{i,1}, \dots, a_{i,c})}$$

and the row vectors $X^{(1)}, \dots, X^{(r)}$ are independent. Similarly, by G_A we denote the law of the matrix $(Z_{i,j}) \in \mathcal{M}_{r,c}$, such that $Z^{(i)} = (Z_{i,1}, \dots, Z_{i,c}) \sim G_{(a_{i,1}, \dots, a_{i,c})}$ and $Z^{(1)}, \dots, Z^{(r)}$ are independent.

In Chapter 4 we will encounter random variables that are products of independent beta and gamma distributed random variables. It turns out that the characteristic functions of such random variables can be written in terms of the Gauss hypergeometric function

$${}_2F_1(a, b; c; \eta) = B(b, c-b)^{-1} \int_0^1 u^{b-1} (1-u)^{c-b-1} (1-u\eta)^{-a} du, \quad (1.15)$$

for $\eta \in \mathbb{C}$, $\operatorname{Re}(c) > \operatorname{Re}(b) > 0$, and $|\arg(1-\eta)| < \pi$ (see e.g. 15.3.1 in [AS64]). This well-studied function is known to satisfy a large number of relations (see e.g. Chapter 15 in [AS64]). In particular, we will use the relations

$${}_2F_1(a, b; b; \eta) = (1-\eta)^{-a}, \quad (1.16)$$

$$(b-a) {}_2F_1(a, b; c; \eta) + a {}_2F_1(a+1, b; c; \eta) = b {}_2F_1(a, b+1; c; \eta), \quad (1.17)$$

and

$${}_2F_1(a, b; c; \eta) = (1-\eta)^{c-a-b} {}_2F_1(c-a, c-b; c; \eta), \quad (1.18)$$

see e.g. 15.1.8, 15.2.14, and 15.3.3 in [AS64], respectively.

1.1.3 Iterated random functions

Iterated random functions are often useful in the study of the stationary distributions of Markov chains, and can also be used to give bounds on the rates of convergence. The basic idea is that Markov chains can be constructed by iterating random functions on the state space E . More specifically, suppose there is a family $\{f_\theta : \theta \in \Theta\}$ of functions that map E to itself (not necessarily injective or bijective) with $f_\theta(x)$ jointly measurable in (θ, x) , and a probability distribution μ on Θ . An i.i.d. sequence $\theta_1, \theta_2, \dots$ with distribution μ generates an i.i.d. sequence of random functions $f_{\theta_1}, f_{\theta_2}, \dots$. The *forward iterations* of the mappings are given by the compositions

$$X_n(\cdot) := (f_{\theta_n} \circ f_{\theta_{n-1}} \circ \dots \circ f_{\theta_1})(\cdot), \quad n \geq 1, \quad X_0(\cdot) = id(\cdot), \quad (1.19)$$

while the *backward iterations* of the mappings are given by

$$Y_n(\cdot) := (f_{\theta_1} \circ f_{\theta_2} \circ \dots \circ f_{\theta_n})(\cdot), \quad Y_0(\cdot) = id(\cdot), \quad (1.20)$$

where $f \circ g(\cdot) := f(g(\cdot))$, and $id(\cdot)$ is the identity function. The theory of forward and backward iterations of i.i.d. random functions is well-established (see e.g. [CL91, DF99, L86]). Here we will only summarise the results relevant to this thesis.

It is clear that $X_n(\cdot)$ has the same distribution as $Y_n(\cdot)$ for each n . However, the properties of the forward and backward processes are very different. Under a broad contraction condition, for any fixed $x \in E$, the forward sequence $\{X_n(x)\}_{n \geq 0}$ is an ergodic Markov chain, while the backward sequence $\{Y_n(x)\}_{n \geq 0}$ converges a.s. and is not in general a Markov chain (see e.g. [DF99]).

The following theorem due to Letac [L86] gives a simple way to prove that the Markov chain $\{X_n(x)\}_{n \geq 0}$ is ergodic and to obtain its stationary distribution.

Theorem 1.2 ([L86]). *Assume that $Y := \lim_{n \rightarrow \infty} Y_n(x)$ a.s. exists and does not depend on x . Then the Markov chain $\{X_n(x)\}_{n \geq 0}$ is ergodic with stationary distribution $\mathcal{L}(Y)$.*

1.2 The three motivating problems

We will now discuss in more detail the sojourn time and boundary crossing problems for the Brownian motion process that motivated the work in this thesis.

1.2.1 An equidistribution theorem

Among the most interesting and important problems on the pathwise behaviour of random processes is the one on the distribution of the time spent by the trajectory of the process below (or above) a given (possibly random) line. The simplest question to be asked in that context is, of course, whether the trajectory of the process will cross that line at all. Historically, the first problem of that kind was perhaps the famous ballot theorem that effectively asserts that, for a simple symmetric random walk conditioned to be at point $-k < 0$ at time n , the probability that the trajectory will stay below zero on $(0, n]$ is k/n (for a historical background and the development of ballot theorems, see [T97]). That result is an immediate consequence of a simple invariance property of the random walk and can readily be extended to random walks with cyclicly exchangeable integer-valued jumps ≥ -1 (see e.g. Theorem 7.6.1 in [T97]).

It turned out that there are more interesting distributional results for sojourn times that are also simple consequences of some invariance properties of the processes in question. One of them is the uniform law for the sojourn time of the positive half-axis in a simple symmetric random walk conditioned to be at zero at the terminal time. That result is called the *Equidistribution Theorem* in Feller's textbook [F68] (cf. Section 9 on p.94) and is as follows. Let $n \geq 1$ be an integer. Consider all the polygonal paths with nodes at the points at (k, s_k) , $k = 0, \dots, 2n$ such that $s_0 = 0$, $s_k = s_{k-1} \pm 1$ for $k = 1, \dots, 2n-1$, and $s_{2n} = 0$. Then the number of such paths such that exactly $2k$ of the paths edges lie above the axis is independent of k .

The continuous time analog of the above result is the uniformity of the distribution of the time spent above zero by the standard Brownian bridge on $[0, 1]$, the result going back to [L39] (as discussed above, that paper also contained the arcsine law for the positive sojourn time of the Brownian motion on $[0, 1]$, see Section 1.2.2 below).

Theorem 1.3 ([L39]). *For a standard Brownian bridge $\{B_t\}_{t \in [0,1]}$ with associated maximum process $\{\overline{B}_t\}_{t \in [0,1]}$, the random variables*

$$\tau'_1 := \inf\{t : B_t = \overline{B}_1\} \quad \text{and} \quad \tau'_2 := \int_0^1 \mathbf{1}_{\{B_s > 0\}} ds$$

are both uniformly distributed on $[0, 1]$, where $\mathbf{1}_A$ denotes the indicator function for event A .

In recent years a number of papers appeared that are devoted to extending the uniform law for the positive sojourn time of the Brownian bridge process. That research culminated in the observation that the uniform law for the positive sojourn time is a property shared by bridge processes satisfying the CEI property (and satisfying an additional condition ensuring the the time spent at axis has zero Lebesgue measure). We will now present a rather nice example illustrating that observation for the empirical process $E_n(t)$ of n i.i.d. uniformly distributed points $U_1, \dots, U_n$ on $[0, 1]$ which is defined by

$$E_n(t) := \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{U_i \leq t\}} - t \right).$$

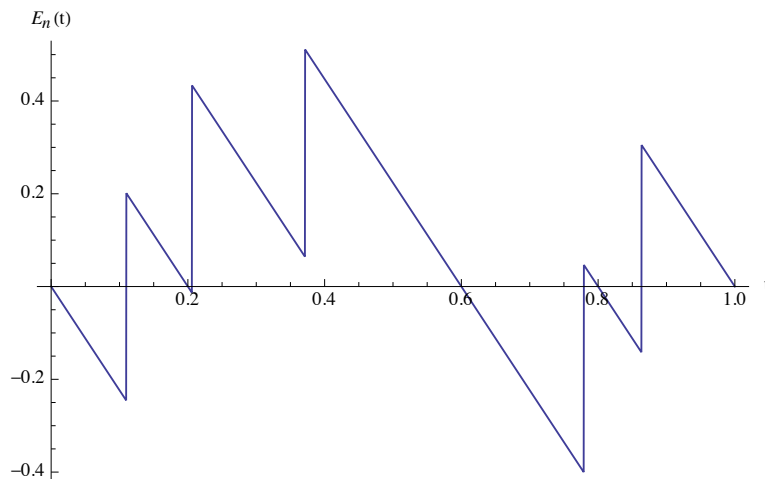


Figure 1.1: A sample path of the Empirical Process $E_n(t)$ for $n = 5$.

Direct verification that the time spent by E_n above zero is uniformly distributed is trivial for $n = 1$, takes some work when $n = 2$, and is a rather difficult computation

for $n = 3$.

However, since $E_n(0) = E_n(1) = 0$, one can view the process E_n as given on the unit circle $\mathbb{S}^1$ using, say, the natural complex number parametrisation

$$\tilde{E}_n(e^{2\pi it}) := E_n(t), \quad t \in [0, 1).$$

For a given trajectory E_n , denote by v the sojourn time above the axis. Then for that given realisation E_n , the set of times $\{t \in [0, 1] : E_n(t) > 0\}$ corresponds to those times for which the cyclically rotated process

$$\tilde{E}_n^t(e^{2\pi i \cdot}) := \tilde{E}_n(e^{2\pi i[\cdot + t(\bmod 1)]}) - \tilde{E}_n(e^{2\pi it})$$

spends less than v time units below zero. In particular, we have that

$$\nu \left\{ t \in [0, 1] : \int_0^1 \mathbf{1}_{\{\tilde{E}_n^t(e^{2\pi i u}) > \tilde{E}_n^t(0)\}} du < v \right\} = v,$$

where ν denotes the Lebesgue measure on $[0, 1]$. Since this equality holds for each fixed E_n , we obtain the desired result.

It is clear that essentially the same argument holds for the significantly larger class of CEI processes. The recent literature on the uniform law for CEI processes led us to consider the following problem.

Problem 1: Can the uniform law for CEI processes be extended to a multivariate setting?

It turned out that one can, in a rather natural way, extend the uniform law of CEI processes to the case of random fields. In particular, we prove a uniform law for sojourn measures of random fields possessing an invariance property that is similar to the CEI property in the univariate case.

In section Section 2.1 we review the literature devoted to the study of the uniform law for CEI processes. Our extension to the case of random fields as well as two interesting examples of random fields satisfying the required invariance property are presented in Chapter 3.

1.2.2 The arcsine laws for Brownian motion

The second problem we considered concerns the following celebrated results due to P. Lévy [L39].

Theorem 1.4. (Lévy's First Arcsine Law) *For a standard Brownian motion $\{W_t\}_{t \in [0,1]}$ with maximum $\overline{W}_1 = \max_{t \in [0,1]} W_t$, the random variables*

$$\tau_1 := \inf \left\{ t \in [0, 1] : W_t = \overline{W}_1 \right\} \quad \text{and} \quad \tau_2 := \sup \left\{ t \in [0, 1] : W_t = 0 \right\}$$

are both arcsine distributed.

Theorem 1.5. (Lévy's Second Arcsine Law) *For a standard Brownian motion $\{W_t\}_{t \in [0,1]}$, the positive sojourn time*

$$\tau_3 := \int_0^1 \mathbf{1}_{\{W_s > 0\}} ds$$

has the arcsine distribution.

The fact that the distributions of τ_1 and τ_2 coincide can be shown using a simple probabilistic argument that follows from another theorem due to P. Lévy.

Theorem 1.6. (Lévy's Theorem) *For a standard Brownian motion $\{W_t\}_{t \geq 0}$ with associated maximum process $\{\overline{W}_t\}_{t \geq 0}$, the process $\{Y_t\}_{t \geq 0}$ defined by*

$$Y_t := \overline{W}_t - W_t, \quad t \geq 0,$$

is a reflected Brownian motion.

Indeed, using Theorem 1.6, we easily obtain that

$$\begin{aligned} \mathbb{P}(\tau_2 \leq t) &= \mathbb{P}\left(\sup\{s \in [0, 1] : W_s = 0\} \leq t\right) \\ &= \mathbb{P}\left(\sup\{s \in [0, 1] : |W_s| = 0\} \leq t\right) \\ &= \mathbb{P}\left(\sup\{s \in [0, 1] : \overline{W}_s = W_s\} \leq t\right) \\ &= \mathbb{P}(\tau_1 \leq t). \end{aligned}$$

A common approach to prove Lévy's second arcsine law is to use a combinatorial argument to establish a link between the first maximum and the number of positive or negative terms for a simple random walk, and then use Donsker's invariance principle to extend this to the Brownian motion process (see e.g. p.137 in [MP10]). This apparent lack of a simple probabilistic argument led us to consider

Problem 2: Is there a simple probabilistic argument showing that τ_3 has the same distribution as either τ_1 or τ_2 ?

However, it turned out that there does exist a simple argument due to P. Lévy [L39] in his original proof that τ_3 has the arcsine law (see also Section 1 of [PY92] for a nice summary of Lévy's approach). That approach involved decomposing the positive sojourn time of the Brownian path into the segments before and after τ_2 . In particular, conditioning on $\tau_2 = t$, Lévy showed that the process up to time t (i.e. $\{W_s\}_{0 \leq s \leq t}$) is a Brownian bridge of length t . Observing that the positive sojourn time of the Brownian bridge is uniformly distributed (Theorem 1.3), and that the trajectory after the last zero is equally likely to be positive or negative independently of the trajectory segment prior to the last zero (see e.g. Section 3.1.2 in [MY06]), we have that

$$\begin{aligned}\tau_3 &= \int_0^1 \mathbf{1}_{\{W_t > 0\}} dt \\ &= \int_0^{\tau_2} \mathbf{1}_{\{W_t > 0\}} dt + \int_{\tau_2}^1 \mathbf{1}_{\{W_t > 0\}} dt \\ &\stackrel{d}{=} U\tau_2 + I_{1/2}(1 - \tau_2),\end{aligned}$$

where $(U, I_{1/2}) \sim U[0, 1] \otimes Be(1/2)$ is independent of τ_2 and $Be(p)$ denotes the Bernoulli distribution with success probability $p \in (0, 1)$. It is not hard to show that τ_3 is arcsine distributed by showing that the right hand side of the last equality above is arcsine distributed.

Note that the same approach can be used to show that $(\tau_3 | W_1 < 0) \sim \beta(1/2, 3/2)$ (and similarly that $(\tau_3 | W_1 > 0) \sim \beta(3/2, 1/2)$). Arguing as above, it suffices to show

that

$$U\tau_2 \sim \beta(1/2, 3/2)$$

for $(U, \tau_2) \sim U[0, 1] \otimes \beta(1/2, 1/2)$, which follows from the well-known fact that $AB \stackrel{d}{=} C$ for $(A, B) \sim \beta(a + b, c) \otimes \beta(a, b)$, and $C \sim \beta(a, b + c)$ (see e.g. p.55 in [GN04]).

Now consider the *random equation* in X :

$$X \stackrel{d}{=} UX + I_{1/2}(1 - X), \quad (1.21)$$

with $(U, I_{1/2}) \sim U[0, 1] \otimes Be(1/2)$ independent of the random variable X on the right hand side. We will say that a random equation in X has solution $X \sim F$ if, for $X \sim F$, the left and right hand sides of the equation are equal in distribution.

A number of our original contributions originated with the observation that the above implies that $X \sim \beta(1/2, 1/2)$ satisfies the random equation (1.21). This observation raised the following rather natural questions.

- (i) Is the solution to (1.21) unique in distribution? More generally, if one considers the random equation

$$X \stackrel{d}{=} AX + B(1 - X), \quad (A, B) \in [0, 1]^2 \text{ is independent of } X, \quad (1.22)$$

under what conditions on (A, B) will the solutions X be unique in distribution?

- (ii) What conditions on (A, B) would ensure that the solutions to (1.22) are beta distributed?
- (iii) Does equation (1.22) arise in other applications?
- (iv) Are there meaningful and useful multivariate extensions of equation (1.22) ?

First we will address question (i). Rearranging (1.21) we obtain

$$X \stackrel{d}{=} (A - B)X + B,$$

which is a special case of the *perpetuity equation*

$$V \stackrel{d}{=} DV + C, \quad (D, C) \in \mathbb{R}^2 \text{ is independent of } V. \quad (1.23)$$

It turns out that the uniqueness of solutions to the perpetuity equation (1.23) is related to the asymptotic, as $n \rightarrow \infty$, behaviour of the iterations

$$V_n = D_n V_{n-1} + C_n, \quad n \geq 1, \quad (1.24)$$

where (D_n, C_n) are i.i.d. $\mathbb{R}^2$ -valued random vectors with the same distribution as (D, C) . The key reference on the recursions (1.24) is [V79], where sufficient conditions for the existence and uniqueness of the limiting distribution of V_n as $n \rightarrow \infty$ are given. In particular, if the condition

$$\sum_{k=1}^n \log |D_k| \xrightarrow{d} -\infty, \quad n \rightarrow \infty, \quad (1.25)$$

is met, then Lemma 1.5 in [V79] implies that any solution V of (1.23) is unique in distribution and V_n converges in distribution to V for all V_0 .

Another notable paper is [K73], where a multidimensional version of (1.24) is studied (i.e. C_n and D_n are $d \times d$ matrices, and V_n is a d -dimensional vector). A number of studies have also considered special cases of (1.24), where the stationary distribution is found by solving random equation (1.23) (see [CL91, V79] and references therein), however there are relatively few special cases where explicit formulae for the stationary distributions are known. The processes we consider in Chapter 4 are generated by several special cases of perpetuity equation (1.23) that admit closed form solutions that have not appeared in the literature.

Returning to our original equation (1.22), this implies that if the condition

$$\mathbb{P}(|A - B| = 1) < 1 \quad (1.26)$$

holds (which is equivalent to $\mathcal{L}(A, B)$ not being concentrated on $\{(1, 0), (0, 1)\}$ and which ensures that (1.25) holds for $D_n = A_n - B_n$), then the solution X of (1.22) is

unique (in distribution). Further, starting from a fixed $X_0 = x \in [0, 1]$ and setting

$$X_n := (A_n - B_n)X_{n-1} + B_n, \quad n \geq 1, \quad (1.27)$$

where (A_n, B_n) are i.i.d. $[0, 1]^2$ -valued random vectors with distribution $\hat{\mu} := \mathcal{L}(A, B)$, one has that $X_n \xrightarrow{d} X$ as $n \rightarrow \infty$ for all $x \in [0, 1]$.

Note that the so defined process $\{X_n\}$ can also be written as the forward iteration $X_{n+1}(\cdot) = F_{n+1}(X_n(\cdot))$ of i.i.d. random mappings

$$F_n := f_{A_n, B_n}, \quad n \geq 1, \quad (1.28)$$

which are random elements of the family

$$\mathcal{F} = \{f_{a,b}(x) = ax + b(1-x) : (a, b) \in [0, 1]^2\} : \quad (1.29)$$

$$X_n(\cdot) := F_n \circ F_{n-1} \circ \cdots \circ F_1(\cdot), \quad n \geq 1, \quad X_0(\cdot) = \text{id}(\cdot), \quad (1.30)$$

where X_n from (1.27) is given by $X_n(x)$ from (1.30). Similarly, the backward iterations $Y_{n+1}(\cdot) = Y_n(F_{n+1}(\cdot))$ of these mappings are given by

$$Y_n(\cdot) := F_1 \circ F_2 \circ \cdots \circ F_n(\cdot), \quad n \geq 1, \quad Y_0(\cdot) = \text{id}(\cdot), \quad (1.31)$$

where clearly $X_n(\cdot)$ has the same distribution as $Y_n(\cdot)$ for each n . However, the properties of the forward and backward processes are very different. Under a broad contraction condition, for any fixed $x \in [0, 1]$, the forward sequence $\{X_n(x)\}_{n \geq 0}$ is an ergodic Markov chain, while the ranges of the backward mappings (1.31) is a sequence of nested intervals with lengths that converge to zero a.s., and for a fixed $x \in [0, 1]$, is not in general a Markov chain (see e.g. [DF99]). In addition, the limiting distributions for the two sequences coincide due to the following contraction principle (cf. Theorem 1.2 above). If

$$\lim_{n \rightarrow \infty} Y_n(x) \xrightarrow{a.s.} Y, \quad Y \text{ does not depend on } x, \quad (1.32)$$

which follows from condition (1.26), then the Markov chain $\{X_n(x)\}_{n \geq 0}$ is ergodic

with stationary distribution given by the law of Y .

Further, observe that the right hand side of the random equation (1.22) has the form of the *randomly weighted average* Z of random variables A and B by W , defined by

$$Z := AW + B(1 - W), \quad (A, B) \text{ is independent of } W. \quad (1.33)$$

One can consider a more general random weighting

$$Z = A_1W_1 + A_2W_2 + \cdots + A_dW_d, \quad d \geq 2, \quad (1.34)$$

where the random weight vector $(W_1, \dots, W_d) \in \mathcal{S}_d$ (cf. p.8 above) is independent of $(A_1, \dots, A_d)$.

In Chapter 4, we consider several special cases of such Markov chains and nested interval schemes related to the random equation (1.22). In particular, we establish the equivalence between beta distributed solutions to random equation (1.22) and gamma distributed solutions to another random equation. This proposition allows us to easily establish the explicit form of the beta distributed solutions to (1.22) in a number of special cases. We also provide some new examples of beta distributed randomly weighted averages Z defined by (1.33) and (1.34).

It turns out that, under certain conditions on $\mathcal{L}(A, B)$, the forward process (1.30) is a “donkey walk” on $[0, 1]$, the name originating from the well-known story of Buridan’s donkey (attributed to the French scientist and philosopher Jean Buridan (1295–1358)), where an equally hungry and thirsty donkey that is halfway between a food and water dies from indecision (see e.g. [SP00, L02]). In order to survive, the donkey must move, and the model we will now describe can be thought of as a donkey moving randomly towards food or water. Starting from a point $x \in [0, 1]$ and for some fixed $p \in (0, 1)$, the donkey moves left the distance Lx with probability p , or right the distance $R(1 - x)$ with probability $1 - p$, where L and R are $[0, 1]$ -valued random variables. The donkey continues moving in this way, each new step being independent of the past.

In Chapter 5, we consider an extension of this donkey walk, where the probability

p is a function of the donkey's current location (see e.g. [P09, RL10]) . Under simple broad conditions, we establish the ergodicity of this “extended donkey walk” and then derive closed form expressions for the stationary densities of the chains when the proportions L and R have beta distribution $\beta(1, a)$ with second parameter positive.

Lastly, observe that the forward and backward process are also related to products of i.i.d. 2×2 random stochastic matrices since the processes can be written equivalently as

$$(X_n(x), 1 - X_n(x)) = (x, 1 - x) \begin{pmatrix} A_1 & 1 - A_1 \\ B_1 & 1 - B_1 \end{pmatrix} \cdots \begin{pmatrix} A_n & 1 - A_n \\ B_n & 1 - B_n \end{pmatrix}, \quad (1.35)$$

and

$$(Y_n(x), 1 - Y_n(x)) = (x, 1 - x) \begin{pmatrix} A_n & 1 - A_n \\ B_n & 1 - B_n \end{pmatrix} \cdots \begin{pmatrix} A_1 & 1 - A_1 \\ B_1 & 1 - B_1 \end{pmatrix}, \quad (1.36)$$

respectively. If the contraction condition (1.26) holds, then (1.32) also holds, and so it is clear that for the left product we have that

$$\begin{pmatrix} A_n & 1 - A_n \\ B_n & 1 - B_n \end{pmatrix} \cdots \begin{pmatrix} A_1 & 1 - A_1 \\ B_1 & 1 - B_1 \end{pmatrix} \xrightarrow{a.s.} \begin{pmatrix} Y & 1 - Y \\ Y & 1 - Y \end{pmatrix} \quad (1.37)$$

as $n \rightarrow \infty$, where $Y := \lim_{n \rightarrow \infty} Y_n(x)$ a.s., and for the right product we have that

$$\begin{pmatrix} A_1 & 1 - A_1 \\ B_1 & 1 - B_1 \end{pmatrix} \cdots \begin{pmatrix} A_n & 1 - A_n \\ B_n & 1 - B_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} X & 1 - X \\ X & 1 - X \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} Y & 1 - Y \\ Y & 1 - Y \end{pmatrix}$$

as $n \rightarrow \infty$, where X satisfies the random equation (1.22) that can be equivalently written as

$$(X, 1 - X) \stackrel{d}{=} (X, 1 - X) \begin{pmatrix} A & 1 - A \\ B & 1 - B \end{pmatrix}. \quad (1.38)$$

Clearly, the left and right products above are special cases of the left and right products of i.i.d. $d \times d$ stochastic matrices $M(1), M(2), \dots$. We study these products in the general case when $d \geq 2$ in Chapter 6. In particular, we obtain an equivalence

relation similar to that presented in Chapter 4, which allows us to characterise the class of distributions of random stochastic matrices M with the property that the products $M(n)M(n-1)\cdots M(1)$ of i.i.d. copies $M(k)$ of M converge a.s. as $n \rightarrow \infty$ and the limit is Dirichlet distributed.

1.2.3 Boundary crossing approximations

Let $\mathbf{W} = \{\mathbf{W}_t = (W_t^{(1)}, \dots, W_t^{(m)})\}_{t \geq 0}$ be the standard m -dimensional Brownian motion process, $\mathbf{W}_0 = \mathbf{0}$. For a fixed $T < \infty$, let $\mathcal{G}$ be the class of open sets $G \subset (0, T) \times \mathbb{R}^m$ (the first component representing time), with $(0, \mathbf{0}) \in \partial G$, where ∂G denotes the boundary of G .

In a number of applied problems (one notable example being barrier options' pricing), one needs to compute the probability

$$P(G) := \mathbb{P}((t, \mathbf{W}_t) \in G, t \in (0, T)) \quad (1.39)$$

for $\mathbf{W}$ to stay within a given set G in the time-space during the time interval $(0, T)$. It is well known that that can be done by solving the respective boundary value problem for the heat equation in m dimensions (see e.g. Section 4.3C in [KS98] for a discussion of the univariate case and [PW08] for an efficient numerical scheme for computing $P(G)$ for cylindric sets G).

In the case of general $G \in \mathcal{G}$, a possible approach to approximate evaluation of $P(G)$ in nontrivial univariate cases is to approximate G with another set $\tilde{G} \in \mathcal{G}$ for which the computation of $P(\tilde{G})$ is tractable. For instance, when

$$G = \{(t, x) : t \in (0, T), g_-(t) < x < g_+(t)\}, \quad (1.40)$$

where $g_-(t) < g_+(t)$ are smooth enough continuous functions, one could use a $\tilde{G}$ of the same nature but with piece-wise linear boundaries $\tilde{g}_\pm$ approximating $g_\pm$, respectively. Recall that, for such boundaries, the problem of calculating $P(\tilde{G})$ reduces (by conditioning on the process' values at the times where the boundaries' have their "junction points") to calculating the values of multivariate normal CDFs.

To justify the use of such approximations, however, one must provide bounds for

the approximation error $|P(G) - P(\tilde{G})|$. In the univariate case, rather tight bounds of such type were obtained for the one-dimensional Brownian motion [BN05] and then extended to time-homogeneous univariate diffusions process [DB08]. The third problem we will consider in this thesis is an extension of the outlined approximation approach to the multivariate case.

Problem 3: Provide a bound for the approximation error $|P(G) - P(\tilde{G})|$ for the case of general $G \in \mathcal{G}$.

The above Problem 3 is the subject of Chapter 7, where under broad assumptions on G , we obtain the desired bound for $|P(G) - P(\tilde{G})|$. We also obtain upper bounds for the first boundary crossing time densities.

Chapter 2

Literature Review

In this chapter we briefly review the approaches and results found in the literature that are relevant to the problems considered in this thesis. We present each topic in chronological order so that the progression is clear to the reader. In some cases, we will present proofs to illustrate the standard techniques used in the literature. We do this to demonstrate how our approach differs from those standard techniques.

The notation used in this chapter is not intended to carry over to subsequent chapters, unless referred to specifically at a later point.

2.1 Sojourn times for bridge processes

We begin by reviewing the literature concerning extensions of Lévy's uniform law for the positive (or negative) sojourn time of the Brownian bridge process to wider classes of processes.

2.1.1 Lévy processes

Let $X = \{X_t\}_{t \geq 0}$, $X_0 = 0$, be a process with stationary independent increments (i.e. a Lévy process).

The uniform law for the positive sojourn time of the Brownian bridge process was extended first to the case of Lévy bridges $X^{(t)} := \{X_t | X_1 = 0\}_{0 \leq t \leq 1}$ by Gettoor and Fitzsimmons in [FG95]. That is, they showed that the occupation time $\int_0^1 \mathbf{1}_{\{X^{(s)} > 0\}} ds$ is uniformly distributed on $[0, 1]$ (see Theorem 3.1 in [FG95]).

The proof of Theorem 3.1 in [FG95] is rather long and so we will not review it here. However, it is based on the application of Fourier-Laplace transforms and an entire function argument that was used in Kac's derivation of the arcsine law for symmetric stable processes (see [K51]).

2.1.2 Processes with exchangeable increments

Let $X = \{X_t\}_{t \in [0,1]}$, $X_0 = 0$, be a process with *exchangeable increments* with trajectories from the space $D[0, 1]$ of right-continuous, left-limited functions on $[0, 1]$. That is, processes $\{X_t\}$ such that, for each n , one has

$$\left\{ X_{\frac{k}{n}} - X_{\frac{k-1}{n}} \right\}_{1 \leq k \leq n} \stackrel{d}{=} \left\{ X_{\frac{\sigma(k)}{n}} - X_{\frac{\sigma(k)-1}{n}} \right\}_{1 \leq k \leq n}$$

for every permutation σ of $\{1, 2, \dots, n\}$.

In [K96], necessary and sufficient conditions are obtained for the sojourn time of $Y_t := X_t - tX_1$ below zero to be uniformly distributed. The key results in [K96] relevant to this thesis are formulated for an arbitrary measurable function $f(t)$, $0 \leq t < 1$. Introduce the *sojourn function* of f by

$$F(f, x) := \int_0^1 \mathbf{1}_{\{f(t) \leq x\}} dt, \quad x \in \mathbb{R}. \quad (2.1)$$

The value $F(X, x)$ is the random duration of time the process X spent at or below the level x during the time interval $[0, 1]$.

For $u \in [0, 1)$ and function f , let

$$f_{(u)}(t) := f((t + u) \bmod 1) - f(u). \quad (2.2)$$

That is, the graph of $f_{(u)}$ is obtained by cutting the graph of f at the point $t = u$, interchanging the two segments and translating the result to get $f_{(u)}(0) = 0$.

A crucial result in [K96] that is relevant to our study is

Lemma 2.1. ([K96]) *Let f have continuous sojourn function, and $U \sim U[0, 1]$. Then*

$$\mathbb{P}(F(f_{(U)}, 0) \leq p) = p, \quad p \in (0, 1).$$

That is, $F(f_{(U)}, 0) \sim U[0, 1]$.

The above Lemma 2.1 is applied in [K96] to processes $\{X_t\}$, $X_0 = 0$ with trajectories in $D[0, 1]$ having exchangeable increments. In particular, [K96] uses a representation theorem to give necessary and sufficient conditions for such processes to have continuous sojourn distributions (see Lemma 1.2 in [K96]), and for $F(Y, 0) \sim U[0, 1]$, where $Y_t := X_t - tX_1$, $0 \leq t < 1$ (see Lemma 1.3(a) in [K96]).

It turns out that the argument from Knight [K96] also extends (with conditions) to any measurable bridge process with cyclically exchangeable increments (or, equivalently, any measurable stationary periodic processes on $\mathbb{R}_+$), as discussed in the next section.

2.1.3 Processes with cyclically exchangeable increments

For $X = \{X_t\}_{0 \leq t \leq 1}$, we say that X has cyclically exchangeable increments if, for any $u \in [0, 1]$, the process

$$X_{(u)}(t) := \begin{cases} X_{u+t} - X_u + X_0 & \text{for } 0 \leq t < 1 - u, \\ X_1 - X_u + X_{u+t-1} & \text{for } 1 - u \leq t \leq 1, \end{cases} \quad (2.3)$$

has the same distribution as X . Note that the transformations $X_{\langle u \rangle}$ defined above and $X_{(u)}$ defined by Knight in [K96] are equivalent in the case $X_0 = X_1 = 0$.

It appears that Kallenberg [K99] was the first to note that the argument from Knight [K96] also extends to measurable bridge process with cyclically exchangeable increments as follows.

Theorem 2.2. ([K99], [CHY01]) *If X has cyclically exchangeable increments, $X(0) = X(1) = 0$ and $F(X, \cdot)$ is continuous a.s., then $F(X, 0) \sim U[0, 1]$.*

2.2 Nested interval schemes on the unit interval

In this section, we will review the existing literature on the limiting distributions of interval splitting and nested interval schemes.

2.2.1 Split and choose the longest subinterval w.p. p

Consider the following interval splitting scheme. Choose a uniformly distributed *splitting point* S_1 in $[0, 1]$, and select the larger of the subintervals $[0, S_1]$, $[S_1, 1]$ with probability p , and the smaller with probability $1 - p$. Continuing this procedure in the same way independently of the chosen subinterval, we obtain a random sequence $\{S_n\}_{n \geq 1}$ that under broad conditions converges a.s. as $n \rightarrow \infty$ to a random variable which we denote by Y_p .

The above described scheme was apparently first studied in the papers [CLZ81, CGZ84]. The former presented two proofs of the fact that $Y_{1/2} \sim \beta(1/2, 1/2)$, and the latter showed that $Y_1 \sim \beta(2, 2)$ and also derived several properties of $\mathcal{L}(Y_0)$.

We will now proceed to prove each of these theorems. Our purpose is to illustrate two of the standard techniques used in the literature as they differ from the techniques we will employ to prove several similar results in Chapter 4.

Theorem 2.3. ([CLZ81]) *The random variable $Y_{1/2}$ has law $\beta(1/2, 1/2)$.*

Proof. Let $m(t) := \mathbb{E}e^{tY_{1/2}}$ be the moment generating function of $Y_{1/2}$. Conditioning

on S_1 , we have

$$\begin{aligned} m(t) &= \mathbb{E}\mathbb{E}(e^{tY_{1/2}}|S_1) = \int_0^1 \mathbb{E}(e^{tY_{1/2}}|S_1 = x) dx \\ &= \frac{1}{2} \int_0^1 m(tx) dx + \frac{1}{2} \int_0^1 e^{tx} m(t(1-x)) dx. \end{aligned} \quad (2.4)$$

Setting $m^{(k)}(t) := \frac{d^k}{dt^k} m(t)$, $k \geq 1$, we obtain

$$\begin{aligned} m^{(1)}(t) &= \frac{1}{2} \int_0^1 x m^{(1)}(t) dx + \frac{1}{2} \int_0^1 (1-x) e^{tx} m^{(1)}(t(1-x)) dx \\ &\quad + \frac{1}{2} \int_0^1 x e^{tx} m(t(1-x)) dx, \end{aligned}$$

by differentiating each side of (2.4). Using induction, we obtain

$$m^{(k)}(t) = \frac{1}{2} \int_0^1 x^k m^{(k)}(t) dx + \frac{1}{2} \sum_{j=0}^k \binom{k}{j} \int_0^1 (1-x)^j x^{k-j} e^{tx} m^{(j)}(t(1-x)) dx.$$

Therefore, since $\mathbb{E}(Y_{1/2}^k) = m^{(k)}(0)$ and $m^{(0)}(0) = m(0) = 1$, we have

$$\begin{aligned} \mathbb{E}(Y_{1/2}^k) &= \frac{1}{2} m^{(k)}(0) \int_0^1 x^k dx + \frac{1}{2} \sum_{j=0}^k \binom{k}{j} m^{(j)}(0) B(k-j+1, j+1) \\ &= \frac{1}{2k} \sum_{j=0}^{k-1} m^{(j)}(0) = \frac{1}{2k} \frac{2k-1}{2(k-1)} \sum_{j=0}^{k-2} m^{(j)}(0) = \frac{\prod_{j=1}^{k-1} (2j+1)}{\prod_{j=1}^k 2j} = \frac{(2k)!}{(\prod_{j=1}^k 2j)^2} \\ &= 2^{-2k} \binom{2k}{k} \end{aligned}$$

for $k \geq 1$, where the right hand side of the last equality is equal to $\mathbb{E}(Z^k)$ for $Z \sim \beta(1/2, 1/2)$. We conclude that $Y_{1/2} \sim \beta(1/2, 1/2)$ since $Y_{1/2}$ and Z are bounded with identical moments. $\square$

Theorem 2.4. ([CGZ84], [DLS86]) *The random variable Y_1 has law $\beta(2, 2)$.*

Proof. We begin by deriving the density of Y_1 on the interval $(0, 1/2]$. Let $F(y)$ be the distribution function of Y_1 . Conditioning on S_1 , for $y \in (0, 1/2]$, we have

$$F(y) = \left(\int_0^y + \int_y^{1/2} + \int_{1/2}^1 \right) \mathbb{P}(Y_1 \leq y | S_1 = x) dx$$

$$= \int_0^y F\left(\frac{y-x}{1-x}\right) dx + \int_{1/2}^1 F\left(\frac{y}{x}\right) dx.$$

Since by symmetry $F(y) = 1 - F(1-y)$, $y \in \mathbb{R}$, one has $F\left(\frac{y-x}{1-x}\right) = 1 - F\left(\frac{1-y}{1-x}\right)$, and therefore

$$\begin{aligned} F(y) &= y - \int_0^y F\left(\frac{1-y}{1-x}\right) dx + \int_{1/2}^1 F\left(\frac{y}{x}\right) dx \\ &= y - (1-y) \int_{1-y}^1 \frac{F(z)}{z^2} dz + y \int_y^{2y} \frac{F(z)}{z^2} dz. \end{aligned}$$

Differentiating the above equation yields

$$\begin{aligned} f(y) := \frac{d}{dy} F(y) &= 1 + \int_{1-y}^y \frac{F(z)}{z^2} dz + \int_y^{2y} \frac{F(z)}{z^2} dz \\ &\quad + \frac{F(2y)}{2y} - \frac{F(y)}{y} - \frac{F(1-y)}{1-y}, \quad y \in (0, 1/2), \end{aligned}$$

and differentiating once again we obtain the differential equation

$$\frac{d}{dy} f(y) = \frac{f(2y)}{y} - \frac{f(y)}{y} + \frac{f(y)}{1-y}, \quad y \in (0, 1/2). \quad (2.5)$$

For any constant $c \in \mathbb{R}$, $cy(1-y)$ is a solution to (2.5). Therefore $f(y) = cy(1-y)$, $y \in (0, 1/2)$ for some constant c . Since $f(y) = f(1-y)$, $y \in (0, 1)$, and $\int_0^1 f(y) dy = 1$, we conclude that $f(y) = 6y(1-y)$, $y \in (0, 1)$, which is the density of the $\beta(2, 2)$ distribution. $\square$

In the remark following the above proof of Theorem 2.4 in [CGZ84], the authors note that R. Pyke had shown that

$$Y_1 \stackrel{d}{=} (1 - U/2)Y_1 + (U/2)I_{1/2}, \quad (2.6)$$

where on the right hand side $(Y_1, U, I_{1/2}) \sim \beta(2, 2) \otimes U[0, 1] \otimes Be(1/2)$. Although no derivation was given in [CGZ84], presumably Pyke had observed that for

$$Z := \begin{cases} Y_1 S_1, & S_1 > 1/2, \\ S_1 + Y_1(1 - S_1), & S_1 \leq 1/2, \end{cases}$$

one has that

$$Y_1 \stackrel{d}{=} Z. \quad (2.7)$$

Indeed, conditioned on S_1 , the so defined Z is the limiting location for the sequence of shrinking intervals, which by construction has the same distribution as Y_1 .

It is precisely the above observation that was used in [DLS86] as the basis of an independent proof of Theorem 2.4. After establishing that Y_1 satisfies the random equation (2.7) via the above argument (which is their Lemma 2), they show in their Theorem 2 that the $\beta(2, 2)$ distribution satisfies (2.7). The proof of Theorem 2.4 is completed by using random equation (2.7) to uniquely determine the moments of Y_1 in their Theorem 3.

Another study considering the limiting distribution of Y_p is [O88], where simulations were used to test whether the limiting distribution of Y_p is beta distributed. That study indicated that the limiting distributions were not beta distributed unless $p \in \{1/2, 1\}$. Indeed, that observation was proved in [DLS86].

Theorem 2.5. ([DLS86]) *The random variable Y_p is beta distributed if and only if $p \in \{1/2, 1\}$.*

The next paper to consider the properties of the distribution of Y_p was apparently [JK95]. In that paper, the authors observe that Y_p is a mixture of the distributions of the random variables WY_p , $W + (1 - W)Y_p$, $(1 - W)Y_p$ and $1 - W + WY_p$ with probabilities $\frac{1}{2}(1 - p)$, $\frac{1}{2}p$, $\frac{1}{2}p$ and $\frac{1}{2}(1 - p)$, respectively, where $W \sim U[0, 1/2]$ is independent of Y_p . Therefore, setting $\mu_r := \mathbb{E}Y_p^r$, we have, for $r \in \mathbb{N}$,

$$\begin{aligned} \mu_r &= \frac{1}{2}(1 - p)\mathbb{E}[(WY_p)^r] + \frac{1}{2}p\mathbb{E}[(W + (1 - W)Y_p)^r] \\ &\quad + \frac{1}{2}p\mathbb{E}[(1 - W)Y_p^r] + \frac{1}{2}(1 - p)\mathbb{E}[(1 - W + WY_p)^r] \\ &= \frac{1}{2}(1 - p)\mathbb{E}[W^r]\mu_r + \frac{1}{2}p\mathbb{E}[W^r] + \frac{1}{2}p \sum_{j=1}^r \binom{r}{j} \mathbb{E}[W^{r-j}(1 - W)^j]\mu_j \\ &\quad + \frac{1}{2}p\mathbb{E}[(1 - W)^r]\mu_r + \frac{1}{2}(1 - p)\mathbb{E}[(1 - W)^r] + \frac{1}{2}(1 - p) \sum_{j=1}^r \binom{r}{j} \mathbb{E}[(1 - W)^{r-j}W^j]\mu_j \\ &= \mu_r((1 - p)M_{r,0} + pM_{0,r}) + \frac{1}{2}(pM_{r,0} + (1 - p)M_{0,r}) \end{aligned}$$

$$+ \frac{1}{2} \sum_{j=1}^{r-1} \binom{r}{j} (pM_{r-j,j} + (1-p)M_{j,r-j})\mu_j,$$

where $M_{a,b} := \mathbb{E}[W^a(a-W)^b]$. We conclude that

$$\mu_r = \frac{pM_{r,0} + (1-p)M_{0,r} + \sum_{j=1}^{r-1} \binom{r}{j} (pM_{r-j,j} + (1-p)M_{j,r-j})\mu_j}{2(1 - (1-p)M_{r,0} + pM_{0,r})}. \quad (2.8)$$

2.2.2 Other schemes

The following nested interval scheme was introduced in the elegant note [K88]. It is an extension of the scheme considered in [CLZ81] (cf. Theorem 2.3). Let $[A_n, B_n]$ be the random subinterval of $[0, 1]$ corresponding to the n th stage of the process ($n \geq 0$), and set $[A_0, B_0] := [0, 1]$. The scheme consists of choosing $(n+1)$ st interval independently from the n th interval by the following rule using a fixed number $k \in \mathbb{N}$. Take k i.i.d. points $U_1^{(n)}, \dots, U_k^{(n)}$ uniformly distributed on $[A_n, B_n]$, and let $C_n := U_{(1)}^{(n)}$, $D_n := U_{(k)}^{(n)}$, where $U_{(1)}^{(n)} \leq \dots \leq U_{(k)}^{(n)}$ are the order statistics of $U_1^{(n)}, \dots, U_k^{(n)}$. Then the $(n+1)$ st interval is taken to be $[C_n, B_n]$, $[A_n, D_n]$ or $[C_n, D_n]$ independently at each stage with probabilities p, q, r , respectively, with $p+q+r=1$. Clearly A_n, B_n converge a.s. to a random point Z . The distribution of Z is given by the main result in [K88].

Theorem 2.6. ([K88]) *The random variable Z has the $\beta(k(p+r), k(q+r))$ distribution.*

The proof of this theorem in [K88] is similar to the proof of Theorem 2.3 above, and so we will not repeat it in full here. However, rather than using the moment generating function $m(t) = \mathbb{E}e^{tZ}$ to compute the moments, the author of [K88] twice differentiates a relation for $m(t)$ similar to (2.4) above to obtain a differential equation for $m(t)$. The proof is completed by showing that the moment generating function of the $\beta(k(p+r), k(q+r))$ distribution is the only solution to the differential equation satisfying $m(t) \rightarrow 1$ as $t \rightarrow 0$.

The random variable Z was also considered in Section 4 of [JK95]. Similar to the their derivation of the moments of Y_p in Section 2.2.1 above, the authors of [JK95] use the same approach to derive the moments of Z , thus providing an alternative

proof of Theorem 2.6 above.

More recently, the following size-biased interval splitting scheme was considered in [GH06]. Take a point U_n uniformly distributed on the n th interval $[A_n, B_n]$, and set $L_n := U_n - A_n$, $R_n := B_n - U_n$. Choose the $(n+1)$ st interval to be $[A_n, U_n]$ with probability $P_1 := L_1^c/(L_1^c + R_1^c)$, and $[U_n, B_n]$ with probability $1 - P_1$, for a fixed $c \in \mathbb{R}$. Clearly $[A_n, B_n]$ converges almost surely to a random point Z_c in $[0, 1]$. The main result in [GH06] relevant to this thesis is as follows.

Theorem 2.7. ([GH06]) *For $r \in \mathbb{N}$, $c \in \mathbb{R}$, one has*

$$\mu_r := \mathbb{E}(Z_c^r) = \frac{\sum_{j=0}^{r-1} \binom{r}{j} \mu_j a_c(k, n-k)}{2(1 - a_c(r, 0))},$$

where

$$a_c(j, k) := 2 \int_0^1 \frac{u^{c+j}(1-u)^k}{u^c + (1-u)^c} du, \quad j, k \in \mathbb{N}.$$

Using Theorem 2.7, the authors of [GH06] prove the following theorem.

Theorem 2.8. ([GH06]) *One has*

- (i) $Z_0 \sim \beta(1/2, 1/2)$,
- (ii) $Z_1 \sim \beta(1, 1)$,
- (iii) $Z_\infty \sim \beta(2, 2)$.

Remark 2.1. Note that parts (i) and (iii) of Theorem 2.8 are Theorem 2.3 and Theorem 2.4, respectively.

2.3 Markov chains on the unit interval

In this section, we will discuss the existing literature relevant to our study of discrete time homogeneous Markov chains with state space $[0, 1]$.

2.3.1 Stochastic give-and-take

The first model we consider originated from the following deterministic process called the game of give-and-take by C. C. Li in [L61]. At the start of the process, two players

(call them players 1 and 2, resp.) each have positive amounts x_0 and y_0 of capital, respectively, where $x_0 + y_0 = 1$ for convenience. In the first step, player 1 gives player 2 the proportion $1 - a$ of her capital x_0 , and player 2 gives the proportion b of her capital y_0 to player 1. Continuing the process using the same rules of exchange, at the $(n + 1)$ st step, we have

$$\begin{aligned}x_{n+1} &= ax_n + b(1 - x_n), \\y_{n+1} &= 1 - x_{n+1},\end{aligned}$$

for $n = 0, 1, \dots$. It is not hard to show that if $a \in (0, 1)$, one has $x_n \rightarrow b/(a + b - 1)$ as $n \rightarrow \infty$.

The stochastic give-and-take model introduced in the pioneering paper [DR63] is a modification of the give-and-take model when the proportions that the players exchange at each step is random. More specifically, let A_n (resp. B_n) be the random proportion that player 1 (resp. player 2) retains at the $(n + 1)$ st step of the process, and let X_n (resp. Y_n) be the capital that player 1 (resp. player 2) holds at the n th step. Then

$$\begin{aligned}X_{n+1} &= A_n X_n + B_n(1 - X_n), \\Y_{n+1} &= 1 - X_{n+1},\end{aligned}$$

for $n \geq 0$. In addition, it is assumed that the random vectors $(A_0, B_0), (A_1, B_1), \dots$ are i.i.d., $\mathbb{P}(|A_1 - B_1| = 1) < 1$, X_0 and $\{(A_n, B_n)\}_{n \geq 0}$ are independent. In that case $\{X_n\}_{n \geq 0}$ is a Markov chain, and the following result holds.

Theorem 2.9. ([DR63]) *The Markov chain $\{X_n\}_{n \geq 0}$ is ergodic.*

Several examples of stationary (or limiting) distributions of $\{X_n\}$ are given in [DR63] for special cases of distribution of (A_0, B_0) . The main example relevant to this thesis is as follows.

Theorem 2.10. ([DR63]) *Let $(A_0, B_0) \sim \beta(a, b) \otimes \beta(b, a)$ for some $a, b > 0$. Then $X_n \xrightarrow{d} X \sim \beta(a + b, a + b)$.*

2.3.2 A random donkey walk in $[0, 1]$

We consider in this section the previous literature devoted to the study of the random donkey walk introduced in [DF99, SP00] (see Section 1.2.2). We describe these models in terms of a particle (the “donkey”) at location $X_n \in [0, 1]$ after $n \geq 0$ steps, and moving according to the following rule. If $X_n = x$, we first choose a direction of travel, either left to the interval $[0, x]$, or right to the interval $[x, 1]$, where the choice does not depend on x (we consider an extended version of this model where the choice can depend on x in the next section). Secondly, we move to a random (not necessarily uniformly distributed) point in the interval corresponding to the chosen direction.

The first paper to consider special cases of the above described class of Markov chains was apparently [DF99]. The motivating example in that study is as follows. If the chain is at x in $(0, 1)$, it chooses one of the two intervals $(0, x)$ or $(x, 1)$ with equal probability $1/2$, and then moves independently to a random y uniformly distributed in the chosen interval. Therefore, the transition density is

$$k(x, y) = \begin{cases} \frac{1}{2x}, & 0 < y < x < 1, \\ \frac{1}{2(1-x)}, & 0 < x < y < 1. \end{cases}$$

The ergodicity of this chain is established in the rather general Theorem 1 in [DF99], which considers Markov chains constructed by iterating random functions. The authors of [DF99] identify the stationary distribution by guessing it has a continuously differentiable density $f(x)$. In that case we have

$$f(y) = \int_0^1 k(x, y) f(x) dx = \frac{1}{2} \int_0^y \frac{f(x)}{1-x} dx + \frac{1}{2} \int_y^1 \frac{f(x)}{x} dx.$$

Differentiating the above equation yields

$$\frac{f'(y)}{f(y)} = \frac{1}{2} \left(\frac{1}{1-y} - \frac{1}{y} \right), \quad (2.9)$$

and therefore $f(y) = \frac{\pi}{\sqrt{y(1-y)}}$, $0 < y < 1$. The authors’ remark that the same argument shows that choosing $(0, x)$ with probability p and $(x, 1)$ with probability

$1 - p$ leads to a $\beta(1 - p, p)$ stationary distribution. That is, the following theorem holds.

Theorem 2.11. ([DF99, SP00]) *Consider a Markov chain with the following dynamics. If the chain is at $x \in (0, 1)$, with probability p , $1 - p$ it moves to a uniformly distributed point in $(0, x)$, $(x, 1)$, respectively. The process continues in the same way, each step being independent of the past. Then the Markov chain is ergodic, with stationary distribution $\beta(1 - p, p)$.*

The above theorem was also proved independently in [SP00] using the method of moments. That same paper also considered two other special cases of random donkey walks. The first special case had a deterministic rule for the direction chosen, and a random rule for the distance travelled. In particular, the following theorem was proved.

Theorem 2.12. ([SP00]) *Suppose from its initial position $X_0 \in (0, 1)$, the particle moves alternately in the direction of 0, of 1, of 0, of 1, etc, where the length of the move is chosen uniformly from the corresponding interval. Then $\{X_n\}_{n \geq 0}$ does not converge, however each of the subsequences $\{X_{2n+1}\}_{n \geq 0}$ and $\{X_{2n}\}_{n \geq 0}$ converges in distribution:*

$$X_{2n+1} \xrightarrow{d} X \sim \beta(1, 2) \text{ and } X_{2n} \xrightarrow{d} X' \sim \beta(2, 1) \quad (2.10)$$

as $n \rightarrow \infty$.

The second special case considered in [SP00] had a random rule for the direction chosen, and a deterministic rule for the distance travelled. In particular, the following theorem was proved.

Theorem 2.13. ([SP00]) *Suppose from its initial position $X_0 \in (0, 1)$, the particle moves in either direction with equal probability $1/2$, and moves to the middle of the corresponding interval. Then*

$$X_n \xrightarrow{d} X \sim U[0, 1] \text{ as } n \rightarrow \infty. \quad (2.11)$$

The paper [BW02] was apparently the next to study special cases of random donkey walks. In particular, that paper considered the following extension of the

model described in Theorem 2.12 above. Starting from a uniformly distributed point $X_0 \in [0, 1]$, the particle moves in the first step towards 0 or 1 with probability p , $1 - p$, respectively. The particle then moves a distance chosen at random from the corresponding interval. In subsequent steps, if the particle moved towards 1 at step $n \geq 1$ (i.e. $X_n > X_{n-1}$), then X_{n+1} is chosen uniformly from $[0, X_n]$, otherwise X_{n+1} is chosen uniformly from $[X_n, 1]$. Then, similar to Theorem 2.12 above, the process $\{X_n\}_{n \geq 0}$ does not converge, however the following result holds for subsequences $\{X_{2n+1}\}_{n \geq 0}$ and $\{X_{2n}\}_{n \geq 0}$.

Theorem 2.14. ([BW02]) *For the Markov chain described above, $\{X_{2n}\}_{n \geq 0}$ converges in distribution to a random variable having the density*

$$2(1 - p + (2p - 1)u), \quad 0 \leq u \leq 1,$$

and $\{X_{2n+1}\}_{n \geq 0}$ converges in distribution to a random variable having the density

$$2(p + (1 - 2p)u), \quad 0 \leq u \leq 1.$$

The most recent paper to consider special cases of random donkey walks is apparently [SP08]. In that paper, two special cases are considered, the first of which is the following extension of the model presented in Theorem 2.11 above. Let F_L and F_R be two fixed distributions on $[0, 1]$. Given the chain value $X_n = x \in [0, 1]$ at time n , the next jump of the process is to the left with probability p or to the right with probability $1 - p$. If the jump is to the left, its length is given by an independent random proportion $L_{n+1} \sim F_L$ of the length of the interval $[0, x]$. Otherwise, the chain jumps to the right for a distance given by an independent random proportion $R_{n+1} \sim F_R$ of the length of the interval $[x, 1]$. Clearly, the chain considered in Theorem 2.11 is a special case of this model when $F_L = F_R = U[0, 1]$. The following theorem is an extension of the result for that model to the case when $F_L = F_R = \beta(1, b)$, $b > 0$.

Theorem 2.15. ([SP08]) *For the Markov chain described above, let $F_L = F_R = \beta(1, b)$ for some $b > 0$. Then $X_n \xrightarrow{d} X \sim \beta(b(1 - p), bp)$ as $n \rightarrow \infty$.*

The second special case considered in [SP08] is as follows.

Theorem 2.16. ([SP08]) *For the Markov chain described above, let $p = 1/2$ and $F_L = F_R = \beta(a, a + 1)$ for some $a > 0$. Then $X_n \xrightarrow{d} X \sim U[0, 1]$ as $n \rightarrow \infty$.*

2.3.3 An extension of the random donkey walk in $[0, 1]$

In this section, we will consider the following random donkey walk, where the rule determining the direction of travel depends on the particle's current location. As above, let F_L and F_R be two fixed distributions on $[0, 1]$, and let $p : [0, 1] \rightarrow [0, 1]$ be a fixed measurable function. Given the chain value $X_n = x \in [0, 1]$ at time n , the next jump of the process is to the left with probability $p(x)$ or to the right with probability $1 - p(x)$. As above, if the jump is to the left, its length is given by an independent random proportion $L_{n+1} \sim F_L$ of the length of the interval $[0, x]$. Otherwise, the chain jumps to the right for a distance given by an independent random proportion $R_{n+1} \sim F_R$ of the length of the interval $[x, 1]$.

The first paper to consider a special case of the above described chain was apparently [BW02]. The authors of that paper describe their model as follows. Starting from a random point $X_0 \sim U[0, 1]$, in consecutive iterations, choose with probability $p \in (0, 1]$ a uniformly distributed point in longer of two subintervals of $[0, 1]$ and with probability $1 - p$ a uniformly distributed point in shorter one. Clearly, the so described model corresponds to the extended random donkey walk with $F_L = F_R = U[0, 1]$ and

$$p(x) = \begin{cases} p, & 0 \leq x < 1/2, \\ 1 - p, & 1/2 \leq x \leq 1. \end{cases}$$

Theorem 2.17. ([BW02]) *In the scheme described above, the limit in distribution of $\{X_n\}_{n \geq 0}$ exists, and its density f_p has the form*

$$f_p(x) = \begin{cases} C(1 - x)^{-p}x^{-(1-p)}, & 0 < x \leq 1/2, \\ Cx^{-p}(1 - x)^{-(1-p)}, & 1/2 < x < 1, \end{cases}$$

where C is a normalising constant such that $\int_0^1 f_p(x) dx = 1$.

The case when $F_L = F_R$ has support on the whole $[0, 1]$ and $p(x) = x$ was

apparently the next model to be studied in the literature in [P09]. The first two theorems in that paper show that the so described Markov chain $\{X_n\}_{n \geq 0}$ is ergodic and converges in distribution to the stationary distribution of the chain. The final theorem in that paper is as follows.

Theorem 2.18. ([P09]) *In the scheme described above, let $p(x) = x$, and $F_L = F_R = \beta(1, b)$, $b > 0$. Then $X_n \xrightarrow{d} X \sim \beta(b, b)$ as $n \rightarrow \infty$.*

The most recent paper to consider special cases of the extended random donkey walk is [RL10]. In that paper, the authors derive the form of the stationary density (when it exists) for the chain $\{X_n\}_{n \geq 0}$ when $F_L = F_R = U[0, 1]$. Their main result is as follows.

Theorem 2.19. ([RL10]) *Suppose that for a given function $p(\cdot)$ the Markov chain $\{X_n\}_{n \geq 0}$, $F_L = F_R = U[0, 1]$ described above has a stationary distribution, and denote its density by $f(x)$. Then the explicit form of $f(x)$ is given by*

$$f(x) = C \exp\left(\int \frac{x - p(x)}{x(1 - x)} dx\right),$$

where C is a normalising constant such that $\int_0^1 f(x) dx = 1$.

2.4 Randomly weighted averages

Let A and B be two independent random variables. In [V87], Van Assche introduced the random variable Z uniformly distributed between A and B with distribution

$$\mathbb{P}(Z \leq x | A = a, B = b) = \begin{cases} 1 & \text{for } x \geq \max(a, b) \\ 0 & \text{for } x < \min(a, b) \\ \frac{x - \min(a, b)}{|b - a|} & \text{for } \min(a, b) \leq x < \max(a, b). \end{cases} \quad (2.12)$$

Using Stieltjes transforms, Van Assche proved the following characterisation theorem.

Theorem 2.20. ([V87]) *Let A and B be i.i.d. random variables on $[0, 1]$, and let Z be uniformly distributed between A and B . Then $Z \sim U[0, 1]$ iff $A \sim \beta(1/2, 1/2)$.*

The above theorem means that choosing a random variable uniformly on a random interval with arcsine distributed end points is the same as choosing a random variable uniformly on $[0, 1]$.

A more basic approach was taken by Johnson and Kotz in [JK90a], who employ the more direct method of moments to generalise Theorem 2.20. The following excerpt from [JK90a] is indicative of several results in this thesis: “The cases studied in this article serve as examples demonstrating that utilisation of first principles can sometimes be more advantageous than more sophisticated techniques, such as those employed in [V87]”.

The key to the approach in [JK90a] was the realisation that the definition of Z with distribution (2.12) is equivalent to

$$Z := \frac{1}{2}(A + B) + \left(W - \frac{1}{2}\right) |A - B|,$$

where $W \sim U[0, 1]$ is independent of (A, B) . Then, if $A \stackrel{d}{=} B$ (and therefore $(A - B)$ has a distribution symmetrical about 0), one has

$$Z \stackrel{d}{=} \frac{1}{2}(A + B) + \left(W - \frac{1}{2}\right) (A - B) \tag{2.13}$$

$$= AW + B(1 - W), \tag{2.14}$$

where (2.13) follows from

$$W - 1/2 \stackrel{d}{=} 1/2 - W. \tag{2.15}$$

Johnson and Kotz then note that (2.14) has the same distribution as Z for any random variable W satisfying (2.15), not only when it is uniformly distributed on $[0, 1]$.

From (2.14), setting $\mu_r := \mathbb{E}[Z^r]$, $\nu_r := \mathbb{E}[A^r] = \mathbb{E}[B^r]$, $r \in \mathbb{N}$, we have

$$\mu_r = \sum_{j=0}^r \binom{r}{j} \mathbb{E}[W^j (1 - W)^{r-j}] \nu_j \nu_{r-j} \tag{2.16}$$

since W , A , and B are mutually independent. Using relation (2.16), Johnson and Kotz derive several relations for the variance and kurtosis of Z as a function of the moments of $\mathcal{L}(A)$ in the special cases where $\mathcal{L}(W)$ is $U[0, 1]$ and $\beta(a, a)$ for $a > 0$, and when W satisfies (2.15) with A and B normally distributed.

Johnson and Kotz then turn their attention to deriving the distribution of Z using representation (2.14) for several special cases of distributions of W , A , and B . We summarise their results in the following theorem.

Theorem 2.21. ([JK90a]) *Suppose W , A , and B are mutually independent, and Z is the random variable with distribution defined by (2.14). Then the following claims hold true.*

- (i) *If W , A , and B have distribution $U[0, 1]$, then Z has density*

$$f(u) = -2(1-u)\ln(1-u) - 2u\ln u, \quad u \in (0, 1).$$

- (ii) *If $W = 1/2$, and A and B have distribution $U[0, 1]$, then Z has the symmetric triangular distribution on $[0, 1]$, with density*

$$f(u) = \begin{cases} 4u, & 0 \leq u \leq 1/2, \\ 4(1-u), & 1/2 < u \leq 1. \end{cases} \quad (2.17)$$

- (iii) *If W has density (2.17), and A and B have distribution $U[0, 1]$, then Z has distribution function*

$$F(u) = 1 - 4u^2\ln u - 4(1-u)^2\ln 2 - 4(1-u) + 6(1-u)^2, \quad u \in (0, 1).$$

- (iv) *If A and B are Cauchy distributed, and W has support on $[0, 1]$, then Z is Cauchy distributed.*

- (v) *If A and B are normally distributed, then*

$$\mathbb{P}(Z \leq u|W) = \Phi\left(\frac{u}{\sqrt{W^2 + (1-W)^2}}\right),$$

where the unconditional distribution of Z is obtained by taking expectations of the above relation. In particular, if $W \sim U[0, 1]$, then

$$\mathbb{P}(Z \leq u) = \int_0^1 \Phi\left(\frac{u}{\sqrt{w^2 + (1-w)^2}}\right) dw.$$

Johnson and Kotz next observe that if W is symmetric around $1/2$ (i.e. satisfies (2.15)), then $\mathbb{E}[W^r] = \mathbb{E}[(1-W)^r]$, and therefore relation (2.16) can be written in the form

$$\mu_r = 2\mathbb{E}[W^r]\nu_r + (\text{function of } \nu_1, \nu_2, \dots, \nu_{r-1}). \quad (2.18)$$

If the moments of W and Z are known, then the moments of A can be computed successively by (2.18). Therefore, the distribution of A is characterised by that of Z when the distribution of A is determined by its moments. This is true, in particular, when the support of A is bounded. That is, the following holds true.

Theorem 2.22. ([JK90a]) *Suppose W , A , and B are mutually independent, $A \stackrel{d}{=} B$, W satisfies (2.15), and the support of A is bounded. Then the distribution of A is characterised by that of Z .*

In a final remark in [JK90a], Johnson and Kotz observe that their method naturally extends to random weighted averages of more than two independent random variables $X_1, X_2, \dots, X_n$, given by

$$Z = W_1X_1 + W_2X_2 + \dots + W_dX_d, \quad d \geq 2, \quad (2.19)$$

where the random weight vector $(W_1, \dots, W_d) \in \mathcal{S}_d$ is independent of $(X_1, \dots, X_d)$

Recently, several related papers have considered the generalisation (2.14) suggested in [JK90a] (see [SH09], [SR12], [H12], [HHM13], [RS13], [R14]). Similar to that of Van Assche in [V87], the approach taken in these studies is to obtain a relation between the Stieltjes transforms of the weighted average Z and $X_1, \dots, X_d$.

2.5 Products of random stochastic matrices

In this section we will briefly review the literature on the left and right products of i.i.d. random stochastic matrices that arise naturally from our Second Problem (cf. (1.35) and (1.36) in the Introduction).

2.5.1 2×2 stochastic matrices

Let $\{(A_n, B_n)\}_{n \geq 1}$ be independent random vectors taking values in $[0, 1]^2$ with common distribution function $F(x, y) = \mathbb{P}(A_i \leq x, B_i \leq y)$. Each random vector (A_i, B_i) defines the random 2×2 stochastic matrix

$$\begin{pmatrix} A_i & 1 - A_i \\ B_i & 1 - B_i \end{pmatrix}, \quad i = 1, 2, \dots$$

Since the class of stochastic matrices is closed under multiplication, the left product

$$\begin{pmatrix} X_n & 1 - X_n \\ Y_n & 1 - Y_n \end{pmatrix} := \begin{pmatrix} A_n & 1 - A_n \\ B_n & 1 - B_n \end{pmatrix} \begin{pmatrix} A_{n-1} & 1 - A_{n-1} \\ B_{n-1} & 1 - B_{n-1} \end{pmatrix} \dots \begin{pmatrix} A_1 & 1 - A_1 \\ B_1 & 1 - B_1 \end{pmatrix} \quad (2.20)$$

is also a random stochastic matrix. Let $F^{(n)}(x, y) = \mathbb{P}(X_n \leq x, Y_n \leq y)$. The earliest results on the properties of the product (2.20) were perhaps due to U. Grenader in the treatise [G63], who showed that if F is not concentrated on $\{(1, 0), (0, 1)\}$, then $F^{(n)}$ converges to limit distribution concentrated on the diagonal $\{(x, x) : 0 \leq x \leq 1\}$ as $n \rightarrow \infty$. In that case, we will denote by G the distribution function of the first element of the limiting random stochastic matrix. That is, $G(x) := \mathbb{P}(X \leq x)$, where

$$\begin{pmatrix} X_n & 1 - X_n \\ Y_n & 1 - Y_n \end{pmatrix} \xrightarrow{a.s.} \begin{pmatrix} X & 1 - X \\ X & 1 - X \end{pmatrix}. \quad (2.21)$$

Some early results on the limiting distribution of the left product (2.20) as $n \rightarrow \infty$ are due to M. Rosenblatt in [R71], who showed that if F assigns zero probability on the diagonal set

$$\{(x, y) \in [0, 1]^2 : x = y\},$$

then

$$G(t) = \int_{0 \leq y < x \leq 1} G\left(\frac{t-y}{x-y}\right) dF(x, y) + \int_{0 \leq x < y \leq 1} \left(1 - G\left(\frac{t-y}{x-y}\right)\right) dF(x, y). \quad (2.22)$$

Unfortunately, there seems to be no general method for solving integral equation (2.22). However, the following two examples were given in [R71]. Suppose

$$\begin{pmatrix} A_n & 1 - A_n \\ B_n & 1 - B_n \end{pmatrix} = \begin{pmatrix} 1/3 & 2/3 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 1 & 0 \\ 1/3 & 2/3 \end{pmatrix} \text{ with probability } 1/2,$$

then G is the Cantor distribution on $[0, 1]$. However, if

$$\begin{pmatrix} A_n & 1 - A_n \\ B_n & 1 - B_n \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix} \text{ with probability } 1/2,$$

then G the uniform distribution on $[0, 1]$. The above two examples demonstrate that it is not even clear when the distribution G is discrete, continuous singular or absolutely continuous.

Some generalisations of Rosenblatt's examples were given by T.-C. Sun in [S75], who considered distributions concentrated at two points in the unit square, showing that the limit is often a singular distribution or a uniform distribution.

The first study to consider an absolutely continuous F with support on the whole unit square was apparently [V86]. In that paper, W. Van Assche used integral equation (2.22) and the properties of some special function to derive the distribution of the limit G when A_i and B_i are i.i.d. with the symmetric beta distribution. In particular, he proved the following theorem.

Theorem 2.23. ([V86]) *Suppose $F(x, y) = F(x)F(y)$ and F is the $\beta(a, a)$ distribution for some $a > 0$. Then G is the $\beta(2a, 2a)$ distribution.*

2.5.2 $d \times d$ stochastic matrices

Several studies have also considered the left and right products of i.i.d. copies $\{X(n)\}_{n \geq 1}$ of a random $d \times d$ matrices X with non-negative entries (not necessarily

stochastic). We denote the left products by

$$X(n, 1) := X(n)X(n-1) \cdots X(1), \quad (2.23)$$

which clearly has the same distribution as the right product

$$X(1, n) := X(1)X(2) \cdots X(n), \quad (2.24)$$

and therefore any assertions concerning the distribution of (2.23) as $n \rightarrow \infty$ apply to that of (2.24) as well.

The asymptotic behaviour of the left products (2.23) was apparently first studied by Bellman [B54], who showed, under certain conditions, that

$$\lim_{n \rightarrow \infty} n^{-1} \mathbb{E} \log X(n, 1)_{i,j}$$

exists, where $X(n, 1)_{i,j}$ is the entry in the i th row and j th column of $X(n, 1)$. This result was later strengthened to almost sure convergence in [FK60], where the behaviour of the norms $\|X(n, 1)\| := \max_{i \leq d} \sum_{j=1}^d |X(n, 1)_{i,j}|$ was studied.

Several equivalent conditions for convergence in distribution of the right products (2.24) without normalisation were given in [KS84]. See also [BL85] and [M00] for surveys on the properties of the products (2.23) and (2.24).

The study of the left products (2.23) in the case when the random matrix X is stochastic (i.e., where all the row sums of X equal one) was apparently initiated by Rosenblatt in [R65], who applied existing results for compact semigroups. Recent results on the infinite product of deterministic and random stochastic matrices by B. Touri can be found in [T12].

Since the class of stochastic matrices is a semigroup under matrix multiplication, the products (2.23) and (2.24) generate left and right random walks on stochastic matrices by

$$n \mapsto X(n, 1) \quad (2.25)$$

and

$$n \mapsto X(1, n), \quad (2.26)$$

respectively.

The right random walk (2.26) is central to several related problems involving distributed averaging. These include distributed computation, distributed optimization, distributed estimation, and distributed coordination (see [T12] and references therein). They are also related to certain Markov processes called Potlatch models. These models were introduced in [HL81] and [LS81], and are described in their simplest form in [KS84]. We will discuss these models further in Section 6.4.1. A comprehensive reference for random walks on stochastic matrices and more general semigroups is [HM11].

The study of the properties and distribution of the limit of the left products of i.i.d. copies $X(1), X(2), \dots$ of X given by

$$\hat{X} := \lim_{n \rightarrow \infty} X(n)X(n-1) \cdots X(1) \quad a.s., \quad (2.27)$$

was apparently initiated in [VKJ93]. In particular, the main result of that paper is the following extension of Theorem 2.23 (although the authors of [VKJ93] were apparently unaware of W. Van Assche's paper [V86]).

Theorem 2.24. ([VKJ93]) *Let $Y^{(i)}(n) := (Y_{i,1}(n), \dots, Y_{i,d}(n))$ ($i = 1, \dots, d$, $n = 1, 2, \dots$) be i.i.d. random vectors and*

$$Y(n) := \begin{pmatrix} Y^{(1)}(n) \\ \vdots \\ Y^{(d)}(n) \end{pmatrix}.$$

Then the limit distribution of each row of $Y(n)Y(n-1) \cdots Y(1)$ is $D_{(ad, ad, \dots, ad)}$ if and only if the distribution of each $Y^{(i)}(n)$ is $D_{(a, a, \dots, a)}$.

A novel feature of Theorem 2.24 is that it gives necessary conditions on a random stochastic matrix X with i.i.d. rows for the product (2.27) to converge to a Dirichlet distributed limit with all parameters equal.

In the paper [CL94], the authors consider the more general case when the rows of X are independent and Dirichlet distributed, but not necessarily i.i.d. In particular, they consider the case when $X \sim D_A$, $A = (\alpha_{i,j}) \in \mathcal{M}_{r,c}^+$, in which case the following

theorems hold true.

Theorem 2.25. ([CL94]) *If $(\mathbf{Y}, X) \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{r\bullet})} \otimes D_A$, then $\mathbf{Y}X \sim D_{(\alpha_{\bullet 1}, \dots, \alpha_{\bullet c})}$.*

Theorem 2.26. ([CL94]) *Let $r = c = d$, $X \sim D_A$, and*

$$(\alpha_{1\bullet}, \dots, \alpha_{d\bullet}) = (\alpha_{\bullet 1}, \dots, \alpha_{\bullet d}), \quad (2.28)$$

If $\{X_n\}_{n \geq 1}$ is a sequence of i.i.d. random variables with same distribution as X , then

$$\lim_{n \rightarrow \infty} X_n X_{n-1} \cdots X_1$$

exists a.s. with all rows identical to some $\mathbf{Y} \in \mathcal{S}_d$, $\mathbf{Y} \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$. Furthermore, if $\mathbf{Y}_1$ is a random vector taking values in $\mathcal{S}_d$ that is independent of X , then $\mathbf{Y}_1 X \stackrel{d}{=} \mathbf{Y}_1$ iff $\mathbf{Y}_1 \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$.

2.5.3 Random nested simplices

Consider the following special case of the nested interval process introduced by Kennedy in [K88] (cf. Theorem 2.6 above). Choose two points independently and uniformly distributed on the interval $[0, 1]$. These random points are the extreme points of a random interval. Consider two new points independently and uniformly distributed on this random interval, and so on. The nested intervals of this sequence shrink to a random point Y with distribution $\beta(2, 2)$, corresponding the special case of Theorem 2.6 above when $r = 1$, $k = 2$.

The problem considered by Letac and Scarsini in [LS98] is an extension to any dimension of the above process. They consider sequences of nested simplices with dimension $d-1$ ($d \geq 2$), the above described process corresponding to the case $d = 2$. Then, as above, if the extreme points of each simplex are uniformly distributed in the previous simplex of the sequence, the sequence converges to a random point $\mathbf{Y} \in \mathcal{S}_d$ with distribution $D_{(d, \dots, d)}$.

To describe the above model more precisely, we will first recall some definitions from [LS98]. A real $(d-1)$ -dimensional *affine space* E is a set such that there exists a real $(d-1)$ -dimensional linear space $\overline{E}$ and a map $E \times E \rightarrow \overline{E}$ denoted by $(\mathbf{a}, \mathbf{b}) \mapsto \vec{\mathbf{a}\mathbf{b}}$ with the following two properties.

(i) For all $\mathbf{a} \in E$, the map $\mathbf{b} \mapsto \vec{\mathbf{a}\mathbf{b}}$ is a bijection from E to $\overline{E}$.

(ii) For all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in E$, one has $\vec{\mathbf{a}\mathbf{c}} = \vec{\mathbf{a}\mathbf{b}} + \vec{\mathbf{b}\mathbf{c}}$.

An *affine frame* of E is a set $(\mathbf{a}_1, \dots, \mathbf{a}_d)$ of d points from E such that

$$(\vec{\mathbf{a}_1\mathbf{a}_2}, \vec{\mathbf{a}_1\mathbf{a}_3}, \dots, \vec{\mathbf{a}_1\mathbf{a}_d})$$

is a linear basis of $\overline{E}$. Taking $(x_1, \dots, x_d) \in \mathbb{R}^d$ with $x_1 + \dots + x_d = 1$, there exists a unique $\mathbf{a} \in E$ such that for a fixed $\mathbf{O} \in E$ one has

$$\vec{\mathbf{O}\mathbf{a}} = x_1 \vec{\mathbf{O}\mathbf{a}_1} + \dots + x_d \vec{\mathbf{O}\mathbf{a}_d}.$$

In particular, $\mathbf{a}$ does not depend on the particular choice of $\mathbf{O}$, and therefore $\mathbf{a}$ is also written as

$$\mathbf{a} = x_1 \mathbf{a}_1 + \dots + x_d \mathbf{a}_d,$$

where $(x_1, \dots, x_d)$ are called the *barycentric coordinates* of $\mathbf{a}$ with respect to the affine frame $(\mathbf{a}_1, \dots, \mathbf{a}_d)$. If they are all non-negative (i.e. $(x_1, \dots, x_d) \in \mathcal{S}_d$), then $\mathbf{a}$ is in the convex hull of $(\mathbf{a}_1, \dots, \mathbf{a}_d)$.

Now return to the model of Letac and Scarsini [LS98] introduced above. They consider a real $(d - 1)$ -dimensional affine space E and a convex hull of d points $\mathbf{p}_1, \dots, \mathbf{p}_d$ in E which forms a simplex S . Choosing d points $\mathbf{p}_1(1), \dots, \mathbf{p}_d(1)$ in S and taking their convex hull yields a new simplex $S(1) \subset S$. Similarly, choosing points $\mathbf{p}_1(2), \dots, \mathbf{p}_d(2)$ in $S(1)$ to obtain $S(2)$ spanned by them, and continuing in this fashion, we obtain a sequence $\{S(n)\}_{n \geq 1}$ of nested simplices.

Suppose that, in the above procedure, the points are chosen at random using the following mechanism. Let $\{X(n)\}_{n \geq 1}$ be a sequence of i.i.d. elements from the class of stochastic matrices $\mathcal{P}_d$ with distribution D_A , A being from the class $\mathcal{M}_d$ of $d \times d$ matrices with non-negative entries, and such that condition (2.28) holds. Then, for each $n = 1, 2, \dots$ and $k = 1, \dots, d$, the point $\mathbf{p}_k(n)$ is specified by its barycentric coordinates given by $X^{(k)}(n)$ with respect to $(\mathbf{p}_1(n-1), \dots, \mathbf{p}_d(n-1))$. As shown in Theorem 2.2 in [LS98], the so defined simplices $S(n)$ shrink to a random point

$\mathbf{Y} \in E$ as $n \rightarrow \infty$:

$$\mathbf{p}_k(n) \xrightarrow{a.s.} \mathbf{Y}, \quad k = 1, \dots, d, \quad (2.29)$$

where the barycentric coordinates of $\mathbf{Y}$ with respect to $(\mathbf{p}_1, \dots, \mathbf{p}_d)$ are Dirichlet distributed with parameter $(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})$.

2.5.4 A random donkey walk in $\mathcal{S}_d$

The well-known story of Buridan's donkey (cf. Section 2.3.2) motivated the authors of [SP00] to consider the following model. Let $\mathbf{p}_1, \dots, \mathbf{p}_d$ be $d \geq 2$ given points in the plane. Starting at a point $\mathbf{R}(0)$ inside the convex hull of $(\mathbf{p}_1, \dots, \mathbf{p}_d)$ at time 0, at time $td + r$, $t \geq 0$, $r \in \{1, \dots, d\}$, the donkey moves from the point $\mathbf{R}(td + r - 1)$ to a point $\mathbf{R}(td + r)$ which is uniformly distributed on the straight line segment connecting the points $\mathbf{R}(td + r - 1)$ and $\mathbf{p}_r$. As noted in [L02], although it is easy to prove existence of the stationary distributions for the d homogeneous Markov chains $\{\mathbf{R}(td + r)\}_{t \geq 0}$, $r = 1, \dots, d$, their computation for $d > 2$ is a difficult problem (the authors of [SP00] focus on the case where $d = 2$, cf. Section 2.3.2 above).

A modification of that scheme in which the donkey travels inside a simplex S with vertices $\mathbf{p}_1, \dots, \mathbf{p}_d$ in $\mathbb{R}^{d-1}$ by choosing on step $td + r$ its new position at random inside the convex hull spanned by its current location and vectors from the set $\{\mathbf{p}_1, \dots, \mathbf{p}_d\} \setminus \{\mathbf{p}_r\}$ was considered in [L02]. The main result of [L02] establishing the form of the stationary distribution of the thus modified donkey walk was actually proved in that paper for yet another version of the model. Namely, rather than the current position of the donkey determining the range for the next step, one instead takes a point with barycentric coordinates (with respect to the affine frame $(\mathbf{p}_1, \dots, \mathbf{p}_d)$) taken from rows of i.i.d. random stochastic matrices.

Specifically, let $X := (X_{i,j})$ be a random element of $\mathcal{P}_d$, and $\{X(k)\}_{k \geq 0}$ be a sequence of i.i.d. random matrices with distribution $\mathcal{L}(X)$. The donkey's position at time n is specified by the vector $\mathbf{B}(n) = (B_1(n), \dots, B_d(n)) \in \mathcal{S}_d$ of its barycentric coordinates with respect to the affine frame $(\mathbf{p}_1, \dots, \mathbf{p}_d)$. Starting at a non-random point $\mathbf{B}(0)$, for $r = 1, \dots, d$ and $k = 0, 1, \dots$, given $\mathbf{B}(kd + r - 1) = (x_1, \dots, x_d)$,

one has

$$\mathbf{B}(kd + r) = (x_1, \dots, x_{r-1}, 0, x_{r+1}, \dots, x_d) + x_r(X_{r,1}(k), \dots, X_{r,d}(k)). \quad (2.30)$$

In other words, setting $\tilde{\mathbf{p}}_r := \sum_{j=1}^d X_{r,j}(k) \mathbf{p}_j$, the location of the donkey at time $kd + r$ will have the same barycentric coordinates in the frame $(\mathbf{p}_1, \dots, \mathbf{p}_{r-1}, \tilde{\mathbf{p}}_r, \mathbf{p}_{r+1}, \mathbf{p}_d)$ as those for the donkey's location at time $kd + r - 1$ in the original frame $(\mathbf{p}_1, \dots, \mathbf{p}_d)$. It is clear that each of the sequences $\{\mathbf{B}(kd + r)\}_{k \geq 0}$, $r = 1, \dots, d$, forms a homogeneous Markov chain.

The following assertion was actually proved in [L02] instead of the stated-there Theorem 3 concerning the “true donkey walk” discussed at the beginning of this section.

Theorem 2.27. ([L02]) *Let $A = (\alpha_{i,j}) \in \mathcal{M}_d^+$ satisfy (2.28), and $X \sim D_A$. Then, for $r = 1, \dots, d$, the stationary distribution of the r th homogeneous Markov chain $\{\mathbf{B}(td + r)\}_{t \geq 0}$ given by (2.30) is Dirichlet with parameters $(\beta_{r,1}, \dots, \beta_{r,d})$ defined by*

$$\beta_{r,j} := \begin{cases} \sum_{i=j}^r \alpha_{i,j} & \text{for } j \leq r, \\ \sum_{i=j}^{d+r} \alpha_{i,j} & \text{for } j > r, \end{cases} \quad (2.31)$$

with the convention that $\alpha_{i,j} = \alpha_{i,j'}$ when $j = j' \pmod{d}$.

2.6 Boundary crossing probabilities for Brownian motion

In this section we briefly review the literature concerning boundary crossing probabilities for Brownian motion that is most relevant to this thesis. First we consider some special cases in which exact solutions for the multidimensional Brownian motion exist, and then discuss the approximation methods for the univariate Brownian motion which we will extend to the multivariate setting.

2.6.1 Exact solutions

Closed form expression for the probability $P(G)$ (cf. (1.39)) for $\mathbf{W}$ to stay within a given set in the time-space G during the time interval $(0, T)$ is only available in a few special cases. Although not discussed here, there is vast literature devoted to different approaches to computing boundary crossing probabilities and first hitting time densities for the univariate Brownian motion. For a recent bibliography of the published work on that topic see e.g. [K08].

Much less has been done in the multivariate case. A number of studies have considered the probability

$$p_{\mathbf{x}}(C) := \mathbb{P}(\mathbf{x} + \mathbf{W}_t \in C, t \in (0, T)) = P((0, T) \times (C - \mathbf{x})), \quad \mathbf{x} \in C,$$

when $C \subset \mathbb{R}^m$ is a cone, the simplest form of which is defined as follows. For $\mathbf{y} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$, let $\theta(\mathbf{y})$ be the angle between $\mathbf{y}$ and the coordinate vector $(1, 0, \dots, 0) \in \mathbb{R}^m$. A cone of angle $\alpha \in (0, \pi)$ is defined as $\mathcal{C}_\alpha := \{\mathbf{y} \in \mathbb{R}^m : 0 < \theta(\mathbf{y}) < \alpha\}$.

It was apparently F. Spitzer who was the first to consider the probability $p_{\mathbf{x}}(\mathcal{C}_\alpha)$ in the two-dimensional case. In [S58], he gave an integral transform for the function $p_{\mathbf{x}}(\mathcal{C}_\alpha)$. This probability was later computed explicitly in [I85].

More recently, the case $m \geq 3$ has been considered for “generalised cones” defined as follows. If D is a proper open connected subset of the unit sphere $\mathbb{S}^{m-1}$ in $\mathbb{R}^m$, the generalised cone $\mathcal{C}_D$ generated by D is the set of all rays emanating from the origin $\mathbf{0}$ and passing through D . Under some technical restrictions on D , a representation for $p_{\mathbf{x}}(\mathcal{C}_D)$ as an infinite series involving confluent hypergeometric functions and eigenfunctions of the Laplace-Beltrami operator on $\mathbb{S}^{m-1}$ was given in [D87]. This result was later strengthened in [BS97], where the same analytic formula was shown to hold for a larger class of generalized cones. An alternative technique based on the reflection principle was used in [M11] to compute $p_{\mathbf{x}}(C)$ in case of “wedges” C .

2.6.2 Univariate approximation rates

As discussed in the Section 1.2.3, due to the lack of closed-form expressions for the probability $P(G)$ in the general case, finding approximate solutions and their

approximation rates is an important alternative approach. The literature to date on this approximation method has focused on the univariate case, which we will briefly review here. In particular, we will concentrate on the methods used by Borovkov and Novikov in [BN05], as we use a similar approach to prove an analogue of their main result (Theorem 1) in the case of the multivariate Brownian motion (see Chapter 7). For the remainder of this section, we will denote by $W = \{W_t\}_{t \geq 0}$ the standard univariate Brownian motion process.

Recall that, in the univariate case, piecewise linear boundaries are commonly used to approximate general boundaries as the boundary crossing probabilities in that case are relatively easy to compute. An explicit formula for the boundary crossing probability

$$\bar{P}(\tilde{G}) := 1 - P(\tilde{G})$$

in the univariate one-sided case where

$$\tilde{G} = \{(t, x) : t \in (0, T), x < \tilde{g}(t)\}, \quad \tilde{g}(0) > 0, \quad (2.32)$$

and $\tilde{g}(t)$ is a piecewise linear function on $[0, T]$ was given in [WP97]. In that case, there are $0 = t_0 < t_1 < \dots < t_n = T$ such that $\tilde{g}(t)$ is linear on each of the intervals $[t_{j-1}, t_j]$, $j = 1, 2, \dots, n$. The approach in [WP97] is based on the observation that the event that W does not cross $\tilde{g}(t)$ on the interval $[0, T]$ is the intersection of n conditionally independent (given the values of W at the points t_{j-1}) events that W does not cross $\tilde{g}(t)$ on $(t_{j-1}, t_j]$. In particular, writing $\tilde{g}_j = \tilde{g}(t_j)$, they show that

$$\bar{P}(\tilde{G}) = 1 - P(\tilde{G}) = 1 - \mathbb{E}h(W_{t_1}, \dots, W_{t_n}; \tilde{g}), \quad (2.33)$$

where

$$h(x_1, \dots, x_n; \tilde{g}) = \prod_{j=1}^n \mathbf{1}_{\{x_j < \tilde{g}_j\}} \left[1 - \exp \left(-\frac{2(\tilde{g}_{j-1} - x_{j-1})(\tilde{g}_j - x_j)}{t_j - t_{j-1}} \right) \right].$$

In the case of general continuous boundaries $g(t)$, the authors of [WP97] then show that for a sequence of piecewise linear functions $\tilde{g}^{(n)}(t)$ with $\tilde{g}^{(n)}(t) \rightarrow g(t)$ as $n \rightarrow \infty$ uniformly in $[0, T]$, and denoting by $\tilde{G}_n$ the set (2.32) with upper barrier $\tilde{g}^{(n)}(t)$, one

has that

$$\overline{P}(G) = \lim_{n \rightarrow \infty} \overline{P}(\tilde{G}_n). \quad (2.34)$$

As noted by the authors of [WP97], it is clear that one can obtain an upper (resp. lower) bound for the probability $\overline{P}(G)$ by computing the probability (2.33) for a dominating (resp. dominated) piecewise linear boundary for $\tilde{g}(t)$.

A generalisation of (2.33) to the case of two-sided piecewise linear boundaries was given in [NFK99]. In that case, $\tilde{G}$ is given by

$$\tilde{G} = \{(t, x) : t \in (0, T), \tilde{g}_-(t) < x < \tilde{g}_+(t)\}, \quad (2.35)$$

where for $0 = t_0 < t_1 < \dots < t_n = T$, the functions $\tilde{g}_\pm(t)$ are linear on each of the intervals $[t_{j-1}, t_j]$, $j = 1, 2, \dots, n$. For general boundaries $g_\pm$ and $K \in \mathbb{R}$, let

$$P(g_-, g_+, K) := \mathbb{P}(g_-(s) < W_s < g_+(s), 0 \leq s \leq T; W_T > K).$$

Then, as shown in [NFK99], one has that

$$P(\tilde{g}_-, \tilde{g}_+, K) = \mathbb{E} \left(\mathbf{1}_{\{W_T > K\}} \prod_{i=0}^{n-1} p(i, \tilde{g}_-, \tilde{g}_+ | W_{t_i}, W_{t_{i+1}}) \right),$$

where

$$p(i, \tilde{g}_-, \tilde{g}_+ | x, y) := \mathbb{P}(\tilde{g}_-(s) < W_s < \tilde{g}_+(s), t_i \leq s \leq t_{i+1} | W_{t_i} = x, W_{t_{i+1}} = y).$$

The authors of [NFK99] then give a simple bound for the approximation error $|P(g_-, g_+, K) - P(\tilde{g}_-, \tilde{g}_+, K)|$ in terms of the distance

$$\int_0^T \left[\frac{d}{ds} (\tilde{g}(s) - g(s)) \right]^2 ds.$$

If one also has $g_\pm(it/n) = \tilde{g}_\pm(it/n)$, $0 \leq i \leq n$, and $g_\pm$ are twice continuously differentiable, then following rate of convergence is given in [NFK99]:

$$|P(g_-, g_+, K) - P(\tilde{g}_-, \tilde{g}_+, K)| \leq \text{const} \sqrt{\frac{\ln(n)}{n^3}}.$$

In [PW01], the results obtained for one-sided boundaries in [WP97] are extended to the two-sided case. The authors of [PW01] first present a formula similar to (2.33) that holds for two-sided piecewise linear approximating boundaries $\tilde{g}_\pm$. It is then observed that the convergence in probability (2.34) holds for sequences $\{\tilde{G}_n\}_{n \geq 1}$ defined by the right hand side of (2.35) with upper and lower boundaries $\tilde{g}_\pm^{(n)}$, and such that $\tilde{g}_\pm^{(n)}(t) \rightarrow g_\pm(t)$ as $n \rightarrow \infty$ uniformly on $[0, T]$. That paper also gives an asymptotic bound on the approximation error. Specifically, assuming that $g_\pm$ are twice continuously differentiable, $g''_\pm(0) \neq 0$, with $g''_\pm$ having at most finitely many roots in $(0, T]$, the authors of [PW01] propose a rule for choosing partition points for the sequence of piecewise linear approximation upper and lower boundaries $\{\tilde{g}_\pm^{(n)}\}_{n \geq 1}$. Under that choice, for

$$\Delta_n := |P(G) - P(\tilde{G}_n)|,$$

one has

$$\limsup_{n \rightarrow \infty} n^2 \Delta_n \leq A,$$

where A is a constant that depends on the boundaries $g_\pm$ and the partition rule.

To implement the approximation procedure in [PW01], one must approximate G with $\tilde{G}_n$ for relatively small n (so that the computation of $P(\tilde{G}_n)$ is feasible). In that case, we require an exact (rather than asymptotic) upper bound for Δ_n . In [BN05], the authors obtain the desired exact upper bound for Δ_n of the form C/n^2 (for constant C), that holds under more general conditions on the boundaries $g_\pm$, and for arbitrary partitions of $[0, T]$ of rank $O(n^{-1})$.

The key assumption used by the authors of [BN05] was a Lipschitz condition on the boundary. Denoting by Lip_K , $K > 0$, the class of functions $g : [0, T] \mapsto \mathbb{R}$ satisfying

$$|g(t+h) - g(t)| \leq Kh, \quad 0 \leq t < t+h \leq T,$$

they prove the following result.

Theorem 2.28. ([BN05]) *Let $g_\pm \in \text{Lip}_K$ and consider functions $f_\pm$ on $[0, T]$ such*

that $\|g_{\pm} - f_{\pm}\| \leq \varepsilon$ for some $\varepsilon > 0$. Then

$$\|P(-\infty, g_+) - P(-\infty, f_+)\| \leq (2.5K + 2T^{-1/2}) \varepsilon$$

and

$$\|P(g_-, g_+) - P(f_-, f_+)\| \leq (5K + 4T^{-1/2}) \varepsilon.$$

The same bounds also hold for the differences

$$\|P(-\infty, g_+; B) - P(-\infty, f_+; B)\| \quad \text{and} \quad \|P(g_-, g_+; B) - P(f_-, f_+; B)\|,$$

where, for a Borel set B ,

$$P(g_-, g_+; B) := \mathbb{P}(g_-(t) < W_t < g_+(t), t \in [0, T]; W_T \in B).$$

We will now give a sketch of the proof of Theorem 2.28 given in [BN05], as we will use a similar argument in Chapter 7 to obtain analogue of that theorem in the multivariate case. First, we note that

$$\begin{aligned} P(g_-, g_+ + \varepsilon) - P(g_-, g_+) &= \mathbb{P}\left[0 \leq \sup_{0 \leq t \leq 1} (W_t - g_+(t)) < \varepsilon, \inf_{0 \leq t \leq 1} (W_t - g_-(t)) > 0\right] \\ &\leq \mathbb{P}\left[0 \leq \sup_{0 \leq t \leq 1} (W_t - g_+(t)) < \varepsilon\right] \\ &= P(-\infty, g_+ + \varepsilon) - P(-\infty, g_+) =: D_\varepsilon(g_+) \\ &\leq \sup_{f \in \text{Lip}_K} D_\varepsilon(f) =: D_\varepsilon. \end{aligned}$$

Using the same argument, one can show that

$$P(g_- - \varepsilon, g_+ + \varepsilon) - P(g_-, g_+ + \varepsilon) \leq D_\varepsilon,$$

and therefore

$$\begin{aligned} 0 &\leq P(g_- - \varepsilon, g_+ + \varepsilon) - P(g_-, g_+) \\ &= [P(g_- - \varepsilon, g_+ + \varepsilon) - P(g_-, g_+ + \varepsilon)] + [P(g_-, g_+ + \varepsilon) - P(g_-, g_+)] \\ &\leq 2D_\varepsilon. \end{aligned}$$

Similarly,

$$0 \leq P(g_-, g_+) - P(g_- + \varepsilon, g_+ - \varepsilon) \leq 2D_\varepsilon,$$

and using the last two inequalities and the relation

$$P(g_+ + \varepsilon, g_+ - \varepsilon) \leq P(f_-, f_+) \leq P(g_- - \varepsilon, g_+ + \varepsilon),$$

we conclude that

$$|P(g_-, g_+) - P(f_-, f_+)| \leq 2D_\varepsilon.$$

It remains to obtain a bound for $D_\varepsilon(g)$.

Lemma 2.29. ([BN05]) *Let g be a function on $[0, 1]$ with $g(0) \geq 0$, such that, for some $K^\pm \geq 0$,*

$$-K^-h \leq g(t+h) - g(t) \leq K^+h, \quad 0 \leq t < t+h \leq 1.$$

Then

$$D_\varepsilon(g) \leq (2K^+ + 0.5K^- + 2)\varepsilon, \quad \varepsilon > 0.$$

Chapter 3

The uniform law for sojourn measures of random fields

In this chapter we present a rather natural extension of the uniform law for sojourn times of processes with cyclically exchangeable increments (cf. Theorem 2.2) to the case of random fields with general parameter sets. Our approach is based on a suitably defined extension of the cyclically exchangeable increments property to the random fields setting. One can obtain the uniform law for sojourn times of processes with cyclically exchangeable increments as a special case of our more general Theorem 3.1 below.

3.1 Introduction

Let $X = \{X(t)\}_{0 \leq t \leq 1}$ be a measurable real-valued stochastic process. In what follows, we will be interested in situations where $X(0) = X(1)$, in which case the definition of $X_{\langle u \rangle}$ (cf. (2.3)) simplifies to

$$X_{\langle u \rangle}(t) := X(u + t \pmod{1}) - X(u) + X(0), \quad t \in [0, 1].$$

In that case, it is more convenient to view the process X as given on the unit circle $\mathbb{S}^1$ using, say, the natural complex number parametrisation

$$\tilde{X}(e^{2\pi it}) := X(t), \quad t \in [0, 1].$$

Moreover, assuming integrability of X , observe that the cyclic exchangeability of increments now translates into the invariance, with respect to rotations of the parametric set $\mathbb{S}^1$, of the distribution of the process $\{\tilde{X}^0(u) : u \in \mathbb{S}^1\}$ given by

$$\tilde{X}^0(e^{2\pi it}) := X(t) - \int_0^1 X(s) ds, \quad t \in [0, 1].$$

Now the point $t = 0$ ceases to be special since, for any fixed $a \in \mathbb{S}^1$, the distribution of the “time” spent by $\tilde{X}^0$ below the level $\tilde{X}^0(0)$ coincides with that of the time spent by $\tilde{X}^0$ below the level $\tilde{X}^0(a)$, and in view of the result of Theorem 2.2, that distribution is uniform.

The objective of the present chapter is to demonstrate how the above result can, in a natural (and rather elementary) way, be extended to the random fields setting where the parametric set T of the field $\{X(t)\}_{t \in T}$ is endowed with a finite measure μ . We show that, for a fixed $a \in T$, provided that a certain invariance property is satisfied, the μ -measure of the subset of T on which the values of X do not exceed that of $X(a)$, is also uniformly distributed. In a sense, this result is a “randomised” version of the well-known fact that, for a random variable ξ with continuous distribution function H , one has $H(\xi) \sim U(0, 1)$.

3.2 The main result

Let $(T, \mathcal{T}, \mu)$ be a measure space with a finite measure and $\{X(t)\}_{t \in T}$ a real-valued measurable random field on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with parameter set T . Without losing generality, we will assume that $\mu(T) = 1$. Further, let

$$F_\mu(X, x) := \mu(t \in T : X(t) \leq x), \quad x \in \mathbb{R},$$

be the μ -measure of the parameter values t for which $X(t)$ was at or below the level x .

Now suppose that G is a family of measurable transformations of T , endowed with a σ -algebra $\mathcal{G}$ containing cylinders with bases from $\mathcal{T}$, i.e. such that

$$\{g \in G : g(t) \in B\} \in \mathcal{G} \quad \text{for any } t \in T, B \in \mathcal{T}.$$

Denote by $X_g(\cdot) := X(g(\cdot))$ the composition of the random field X and argument transformation $g \in G$, and by $\mu_g(\cdot) := \mu(g(\cdot))$ a similar transformation of the measure μ .

A natural extension of Theorem 2.2 to this setting is given by

Theorem 3.1. *Suppose that $X_g \stackrel{d}{=} X$ and $\mu_g = \mu$ for any $g \in G$, and that there exists a probability measure ν on $(G, \mathcal{G})$ such that, for some fixed $a \in T$,*

$$\nu(g \in G : g(a) \in B) = \mu(B), \quad B \in \mathcal{T}. \quad (3.1)$$

If $F_\mu(X, \cdot)$ is continuous a.s. then $F_\mu(X, X(a)) \sim U[0, 1]$.

Remark 3.1. In the context of the situation discussed in the Introduction, we had $T = \mathbb{S}^1$, μ was the uniform distribution on $\mathbb{S}^1$, the family G consisted of all rotations of $\mathbb{S}^1$ (with $X_{\langle u \rangle}$ corresponding to “rotating” the field X by the angle $2\pi u$), ν was the Haar measure on G . Clearly, property (3.1) held true for any $a \in T$.

There is a discrete parameter set analog of the above theorem. Prior to stating it, observe that, in the case of a finite parameter set T , we can always take $\mathcal{T} = 2^T$ and that, assuming that μ is the uniform distribution on T , we automatically have

$\mu_g = \mu$ for any transformation g of T . Moreover, in that case G is also necessarily finite, so one can always take $\mathcal{G} = 2^G$.

The assertion has the following form.

Theorem 3.2. *Suppose that $X_g \stackrel{d}{=} X$ for any $g \in G$, $\text{card}(T) = N < \infty$, $\mu(\{t\}) = 1/N$, $t \in T$, and that there exists a probability measure ν on G such that, for some fixed $a \in T$, relation (3.1) holds. If $\mathbb{P}(X(s) = X(t)) = 0$ for any $s \neq t$, then $F_\mu(X, X(a))$ follows the uniform distribution on $\{1, 2, \dots, N\}$.*

The demonstration of the above result uses the same type of argument as that of Theorem 3.1 and is actually even simpler, and so we will only present here the proof of the latter.

Proof of Theorem 3.1. For $p \in (0, 1)$, introduce

$$\begin{aligned} q_X(p) &:= \sup\{x \in \mathbb{R} : F_\mu(X, x) \leq p\}, \\ Q_X(p) &:= \{t \in T : X(t) \leq q_X(p)\}. \end{aligned}$$

By the continuity assumption,

$$F_\mu(X, q_X(p)) \equiv \mu(Q_X(p)) = p, \quad (3.2)$$

and it is not hard to see that, for a fixed $t \in T$, one has

$$\{F_\mu(X, X(t)) \leq p\} = \{t \in Q_X(p)\}. \quad (3.3)$$

Observe also that, due to the invariance of μ under g , we have

$$F_\mu(X_g, x) = \mu(t : X_g(t) \leq x) = \mu_g(t : X_g(t) \leq x) = \mu(s : X(s) \leq x) = F_\mu(X, x)$$

for any $g \in G$, $x \in \mathbb{R}$, and so $q_X(p) = q_{X_g}(p)$, $p \in (0, 1)$. Therefore, for any $p \in (0, 1)$, $g \in G$,

$$a \in Q_{X_g}(p) \quad \text{iff} \quad g(a) \in Q_X(p). \quad (3.4)$$

Now consider the product probability space $(\Omega \times G, \mathcal{F} \otimes \mathcal{G}, \mathbb{P}^* = \mathbb{P} \times \nu)$, and let γ be the second coordinate random variable on the space, so that γ is a random

element of G independent of X and following the distribution ν . Then we clearly have $X_\gamma \stackrel{d}{=} X$, and so, for any $p \in (0, 1)$,

$$\begin{aligned} \mathbb{P}(F_\mu(X, X(a)) \leq p) &= \mathbb{P}^*(F_\mu(X_\gamma, X_\gamma(a)) \leq p) \\ &= \mathbb{P}^*(a \in Q_{X_\gamma}(p)) \end{aligned} \tag{3.5}$$

$$= \mathbb{P}^*(\gamma(a) \in Q_X(p)) \tag{3.6}$$

$$\begin{aligned} &= \mathbb{E}^* \mathbb{P}^*(\gamma(a) \in Q_X(p) \mid X) \\ &= \mathbb{E} \mu(Q_X(p)) \end{aligned} \tag{3.7}$$

$$= p, \tag{3.8}$$

where (3.5), (3.6), (3.7) and (3.8) follow from (3.3), (3.4), (3.1) and (3.2), respectively. The theorem is proved.

3.3 Examples

Example 3.1. Let $U_1, \dots, U_n$ be an i.i.d. sample of points uniformly distributed over $\mathbb{S}^{d-1}$, $d \geq 2$, $K : \mathbb{R}^d \rightarrow \mathbb{R}$ be a measurable function, and set

$$X(t) := \sum_{k=1}^n K(t - U_k), \quad t \in \mathbb{S}^{d-1}. \tag{3.9}$$

Take $G := SO(d)$, the special orthogonal group on $\mathbb{R}^d$, and ν the Haar measure on it. Then it is clear that, for a nice enough K , the field X will satisfy the conditions of Theorem 3.1 with μ being the uniform distribution on $\mathbb{S}^{d-1}$ (due to symmetry).

To graphically illustrate this example, assume that $d = 3$ and let

$$K(x) := 1 - (\|x\|/2)^{1/10}, \quad x \in \mathbb{R}^3, \tag{3.10}$$

where $\|\cdot\|$ is the Euclidean norm. It is clear that the conditions of Theorem 3.1 are met with the above choice of K .

One can think of X as describing the relief (the surface altitude relative to a fixed reference level, i.e. the surface of a sphere) of a “random planet” with n

mountains that have their summits at the sample points U_k and possess symmetric shape features. A simulation of such a surface, with $n = 20$, is depicted on the left pane in Fig. 1.

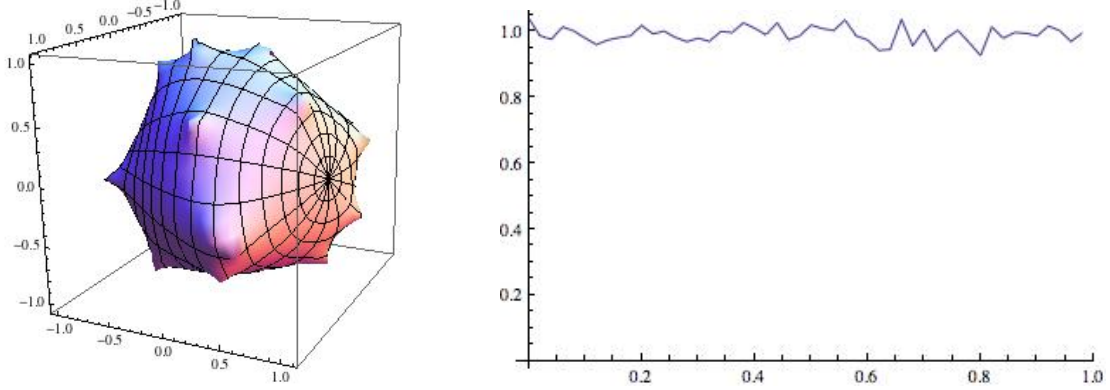


Figure 3.1: A sample realisation of the field (3.9) with kernel (3.10), $d = 3$ and $n = 20$ (on the left pane) and the scaled histogram of a simulated sample of 10^5 independent copies of $F_\mu(X, X(a))$ for that field, a being the “North pole” (on the right pane).

The right pane in Fig. 1 depicts the scaled histogram (with 50 bins) for an i.i.d. sample of 10^5 copies of $F_\mu(X, X(a))$ for that field, a being the “North pole” (on the “coordinate sphere” $\mathbb{S}^2$). The values of $F_\mu(X, X(a))$ were computed using antithetic Monte Carlo methods. The simulation confirms the assertion of Theorem 3.1 that the area on the coordinate sphere where the terrain is below the height of the relief at the North pole is uniformly distributed.

Example 3.2. Consider a rectangular array of random variables $\{X(t)\}_{t \in T}$ with $T = \{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$ for some $n, m \in \mathbb{N}$, such that $\mathbb{P}(X(s) = X(t)) = 0$ for any $s \neq t$. In other words, X is a random $m \times n$ matrix whose entries are distinct a.s. We will also assume that the distribution of the matrix is invariant w.r.t. cyclic permutations of its rows and columns.

A real-life example of such a situation is provided by an $n \times n$ Latin square experimental design, which is obtained by choosing at random a square from a given isotopy class of $n \times n$ Latin squares (or, to simplify the procedure, starting with a given $n \times n$ Latin square and then cyclicly permuting its rows and columns at random), and then assigning the treatment factor levels to the respectively labelled elements of the array.

Introduce G as the collection of all compositions of the form $g_r \circ g_c$, where g_r is a cyclic permutation of the rows, and g_c a cyclic permutation of the columns of T . Then clearly $X_g \stackrel{d}{=} X$ for any $g \in G$, and the uniform distribution ν on G will have the desired property (3.1) for any fixed $a \in T$ (recall that, in the discrete case, we assume that μ is uniform on T).

Now we see that all the conditions of Theorem 3.2 are met, and therefore, say, the number of the entries of the matrix X that don't exceed the first entry in its first row (i.e. the random variable $\text{card}(t \in T : X(t) \leq X((1, 1)))$) has the uniform distribution on $\{1, 2, \dots, nm\}$.

Chapter 4

Beta distributed limits of iterated linear random functions

In this chapter we consider several special cases of iterations of random i.i.d. linear functions that describe nested interval schemes when iterated in a backward direction, and ergodic Markov chains in the forward direction. All these cases will have beta distributed fixed points, and we find these fixed points using an extension of a result due to E. J. G. Pitman [P37] to solve for the beta distributed solution of a random equation by reducing the latter to a random equation with a related gamma distributed solution. In each case, these fixed points are the almost sure limits of the corresponding nested interval schemes and the stationary distributions of the corresponding Markov chains. We will also present some apparently new examples of beta distributed randomly weighted averages and stick-breaking schemes with beta distributed limits.

4.1 Introduction

In this chapter, we study the classes of interval-valued processes and ergodic Markov chains generated by the family $\mathcal{F}$ (cf. (1.29)) of linear random functions. Our approach is centred around the following two key observations. Firstly, every process in one class has a corresponding “dual” process in the other class, with the same limiting distribution (where the backward iterations converge a.s. and the forward iterations converge in distribution). This observation enables us to immediately obtain previously unknown limiting distributions of several processes of interest due to their duality with well studied processes appearing in the literature.

Our second key observation is that one can show the limit is beta distributed by using a related random equation with a gamma distributed solution, allowing for simple analytic proofs. In addition, this random equation generates a corresponding ergodic Markov chain with a gamma distributed stationary distribution.

These two key observations form the systematic approach used in this chapter, whereby we prove assertions concerning the limiting distributions of random processes we consider and provide solutions to two random equations by solving only one of them.

We can interpret the forward process (1.30) in terms of the movement of a particle in $[0, 1]$, where X_n is the location of the particle at time n , and the particle moves from X_{n-1} to X_n by (1.27). In particular, we will consider those models where, for each n , at least one of the following holds: one has $A_n = 1$ (which would imply a move to the right) or $B_n = 0$ (implying a move to the left). See Section 2.3 for a review of the literature concerning models of this type (where, at each step, a particle at $X_n \in [0, 1]$ moves at time $n + 1$ either left to a random point in $[0, X_n]$, or right to a random point in $[X_n, 1]$) having in most cases beta distributed stationary distributions.

We will interpret the backwards process (1.31) as generating a sequence of nested intervals $\mathbb{I}_0 = [0, 1] \supset \mathbb{I}_1 \supset \mathbb{I}_2 \supset \dots$ given by the ranges of the corresponding backward mappings $Y_0, Y_1, Y_2, \dots$ (i.e. Y_n maps $[0, 1]$ onto $\mathbb{I}_n$), and consider only the case when $|\mathbb{I}_n| \rightarrow 0$ as $n \rightarrow \infty$ (i.e. when F_1 is a contraction).

In some of the cases we consider, the nested interval scheme generated by the

backward iteration (1.31) is an interval splitting scheme of the following type. Choose a random (not necessarily uniformly distributed) *splitting point* S_1 in $[0, 1]$ and select according to a given (possibly random) rule one of the subintervals $[0, S_1]$, $[S_1, 1]$. Continuing this procedure in the same way independently on the chosen subinterval, we obtain a random sequence $\{S_n\}_{n \geq 1}$. When the law of the splitting point S_1 is not concentrated on $\{0, 1\}$, the length of the chosen interval tends to zero a.s. as $n \rightarrow \infty$, and therefore S_n converges a.s. to a random variable Y . In several cases considered previously, this limiting random variable Y turns out to be beta distributed (see Section 2.2.1 for a review of the literature concerning these interval splitting schemes). We call this scheme the CGZ model after Chen, Goodman and Zame who first studied it in [CGZ84] under the conditions that $S_n \sim U[0, 1]$ and the largest subinterval is selected with probability p at each step.

Most of the nested interval schemes considered in the literature can be classified using a convenient notation of the form

$$\langle L_N, L'_{M'} | L''_{M''} \rangle,$$

where L, L', L'', M', M'' represent letters and N is a natural number which have the following meaning.

Imagine each step in a nested interval scheme $\mathbb{I}_0 = [0, 1] \supset \mathbb{I}_1 \supset \mathbb{I}_2 \supset \dots$ is generated by the following two stage process. In the first stage of step n we choose a fixed (one and the same for all n) number of subintervals of $\mathbb{I}_{n-1}$, $n \geq 1$. In the second stage, we select one of the chosen subintervals to be $\mathbb{I}_n$.

We set $L := S$ if the intervals from the first stage can be constructed by *splitting* the interval $\mathbb{I}_{n-1}$, $n \geq 1$, into two or more subintervals (i.e. they form a partition of $\mathbb{I}_{n-1}$), otherwise we set $L := G$ for *general*. The number N is equal to the number of intervals generated in the first stage. We set $L' := D$ if the rule for the first stage is *deterministic*, or $L' := R$ if the rule is *random*. The subscript M' of the second component is equal to I if the rule is the same (i.i.d. after scaling and translation) for each stage of the process, otherwise $M' := G$ for *general*. Similarly, in the final component we set $L'' := D$ (resp. R) if the rule for the second stage is deterministic (resp. random), and $M'' := I$ if the rule is the same (i.i.d.) for each stage of the

process, or $M'' := G$ otherwise.

We will mostly be interested in nested interval schemes with $M' = M'' = I$ since in that case the scheme is completely determined by the rule for choosing the first interval $\mathbb{I}_1$ from $[0, 1]$ (e.g. the CLZ model is an $\langle S_2, R_I | R_I \rangle$ scheme in the general case).

To illustrate our classification, consider the Kennedy's nested interval scheme (cf. Section 2.2.2) in the case $p, q, r > 0$. For $k > 3$ the above scheme is an $\langle G_3, R_I | R_I \rangle$ nested interval scheme, whereas for $k = 3$ it is an $\langle S_3, R_I | R_I \rangle$ scheme. Note that interval schemes can admit different classifications within the introduced notation. For example, when $p = 1$, Kennedy's scheme can be viewed as either $\langle S_3, R_I | D_I \rangle$ or $\langle G_1, R_I | D_I \rangle$, where in the latter we only considered the middle interval $[C_n, D_n]$, rather than choosing the middle interval deterministically from the three intervals in the former. Despite that, the classification may still be useful for understanding of and distinguishing between the schemes we will consider.

The remainder of the chapter is structured as follows. In Section 4.2, we present our main results, beginning in Section 4.2.1 with a description of the general nested interval scheme generated by the backward iteration (1.31). We then devote Section 4.2.2 to proving an equivalence result for beta distributed solutions to (1.22) and gamma distributed solutions to a related random equation. In Section 4.2.3 we introduce several equivalent claims concerning limiting distributions for processes and solutions to random equations we consider.

In Sections 4.2.4 and 4.2.5 we use the approach discussed above to find the explicit form for the limiting distributions of nested interval schemes and Markov chains on $[0, 1]$ in two broad special cases. In the first one, at each step either $A_n = 1$ or $B_n = 0$. In the second more general case, the distribution of (A_n, B_n) can have support on an arbitrary subset of the unit square $[0, 1]^2$, with most cases considered being closely related to the beta distribution. We conclude this chapter by showing how the results obtained in Section 4.2 also imply some apparently new examples of beta distributed randomly weighted averages in Section 4.3.1, and stick-breaking schemes with beta distributed limits in Section 4.3.2.

4.2 Main results

We start by describing the general nested interval scheme generated by the backward iteration (1.31).

4.2.1 A general nested interval scheme

Recall that $F_1, F_2, \dots$ are the random functions from class $\mathcal{F}$ which are generated by a sequence of i.i.d. vectors $(A_1, B_1), (A_2, B_2), \dots$ with distribution $\hat{\mu}$ (see p.18). Plainly,

$$F_1(x) = (A_1 - B_1)x + B_1$$

maps $[0, 1]$ onto $\mathbb{I}_1 = [\min(A_1, B_1), \max(A_1, B_1)]$. So the n th backward iteration from (1.31) maps $[0, 1]$ onto the n th interval of the following nested interval scheme. Starting with the unit interval $[0, 1]$, select the interval $\mathbb{I}_1$ with end points A_1 and B_1 . For the second step, noting that the mapping F_2 is linear, we see that the end points of the second nested interval $\mathbb{I}_2 \subset \mathbb{I}_1$ are given by

$$\bar{A}_2 := Y_2(1) = (A_1 - B_1)A_2 + B_1 \quad \text{and} \quad \bar{B}_2 := Y_2(0) = (A_1 - B_1)B_2 + B_1.$$

It follows that

$$\begin{aligned} \mathcal{L} \left(\frac{(\bar{A}_2 - B_1, \bar{B}_2 - B_1)}{A_1 - B_1} \middle| A_1 - B_1 > 0 \right) &= \mathcal{L}((A_1, B_1)), \\ \mathcal{L} \left(\frac{(\bar{A}_2 - A_1, \bar{B}_2 - A_1)}{B_1 - A_1} \middle| A_1 - B_1 < 0 \right) &= \mathcal{L}((1 - A_1, 1 - B_1)). \end{aligned} \tag{4.1}$$

This is, if $A_1 > B_1$, after translating and scaling the real line so that $\mathbb{I}_1$ turns into $[0, 1]$, the law of the end points of the interval $\mathbb{I}_2$ is the same as that of the end points of $\mathbb{I}_1$. Otherwise, if $A_1 < B_1$, the law of $(\bar{A}_2, \bar{B}_2)$ is the same (after translating and scaling) as that of $(1 - A_1, 1 - B_1)$. Similarly, the end points of the n th interval are given by $\bar{A}_n := Y_n(1)$ and $\bar{B}_n := Y_n(0)$, with

$$\mathcal{L} \left(\frac{(\bar{A}_n - \bar{B}_{n-1}, \bar{B}_n - \bar{B}_{n-1})}{\bar{A}_{n-1} - \bar{B}_{n-1}} \middle| \prod_{j=1}^{n-1} (A_j - B_j) > 0 \right) = \mathcal{L}((A_1, B_1)), \tag{4.2}$$

$$\mathcal{L} \left(\frac{(\bar{A}_n - \bar{A}_{n-1}, \bar{B}_n - \bar{A}_{n-1})}{\bar{B}_{n-1} - \bar{A}_{n-1}} \middle| \prod_{j=1}^{n-1} (A_j - B_j) < 0 \right) = \mathcal{L}((1 - A_1, 1 - B_1)).$$

We will mostly be interested in special cases where one of the following conditions is met:

[M1] $A > B$ a.s.,

[M2] $\hat{\mu} = \hat{\mu}' := \mathcal{L}((1 - A, 1 - B))$,

[M3] $\hat{\mu}' = \mathcal{L}((B, A))$.

If at least one of [M1]–[M3] holds the nested interval schemes generated by (1.31) satisfy the following i.i.d. property.

[I] For all $n \geq 1$, the conditional distribution of

$$\frac{\mathbb{I}_n - \min \mathbb{I}_{n-1}}{|\mathbb{I}_{n-1}|}$$

given $\mathbb{I}_{n-1}$ does not depend on $\mathbb{I}_{n-1}$ and is one and the same for all $n \geq 1$.

4.2.2 A relationship between beta and gamma distributed solutions to linear random equations

For a positive vector $\mathbf{t} \in \mathbb{R}_+^d$, let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d) \sim G_{\mathbf{t}}$. Then we have from Lukacs's characterisation of the gamma distribution (cf. Theorem 1.1) that

$$\xi_{\bullet} \text{ is independent of } \left(\frac{\xi_1}{\xi_{\bullet}}, \dots, \frac{\xi_d}{\xi_{\bullet}} \right), \quad (4.3)$$

and so

$$(\xi_1, \dots, \xi_d) \stackrel{d}{=} \left(\frac{\xi_1}{\xi_{\bullet}}, \dots, \frac{\xi_d}{\xi_{\bullet}} \right) \tilde{\xi}_{\bullet}, \quad (4.4)$$

where $(\tilde{\xi}_1, \dots, \tilde{\xi}_d)$ is an independent copy of $\boldsymbol{\xi}$.

A remarkable observation made by E. J. G. Pitman [P37] is the following extension of (4.3). Let, as above, $\boldsymbol{\xi} \sim G_{\mathbf{t}}$, and $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a *scale independent*

measurable function, i.e., for any $a \neq 0$,

$$f(ax_1, \dots, ax_d) \equiv f(x_1, \dots, x_d), \quad (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Then the random variable $f(\xi_1, \dots, \xi_d)$ is independent of $\xi_\bullet$.

The following lemma is an extension of that observation to random functions.

Lemma 4.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(E, \mathcal{E})$ a measurable space, and $X : \Omega \rightarrow E$ an E -valued random variable. Suppose $H : \mathbb{R}^r \times E \rightarrow \mathbb{R}_+$ is jointly measurable and*

$$H(ay_1, \dots, ay_r, x) = H(y_1, \dots, y_r, x)$$

holds for all $a \neq 0$, $(y_1, \dots, y_r) \in \mathbb{R}^r$ and $x \in E_1 \subset E$, where $\mathbb{P}(X \in E_1) = 1$. If $\mathbf{V} = (V_1, \dots, V_r) \sim G_{\mathbf{t}}$, $\mathbf{t} \in \mathbb{R}_+^r$, is independent of X , then $V_\bullet$ is independent of $H(\mathbf{V}, X)$.

Proof. Denote by $\phi(s, u)$, $(s, u) \in \mathbb{R}_+^2$ the joint Laplace transform of $V_\bullet$ and $H(\mathbf{V}, X)$. Then

$$\begin{aligned} \phi(s, u) &= \mathbb{E} e^{-s(V_1 + \dots + V_r) - uH(V_1, \dots, V_r, X)} \\ &= C_{\mathbf{t}} \int_0^\infty \dots \int_0^\infty v_1^{t_1-1} \dots v_r^{t_r-1} e^{-(1+s)\sum v_k} \mathbb{E} e^{-uH(v_1, \dots, v_r, X)} dv_1 \dots dv_r, \end{aligned}$$

where $C_{\mathbf{t}}^{-1} = \prod_{k=1}^r \Gamma(t_k)$. Making the substitution $y_k = (1+s)v_k$ and observing that $H(v_1, \dots, v_r, X) \stackrel{d}{=} H(y_1, \dots, y_r, X)$ by virtue of scale independence, we have that

$$\phi(s, u) = \frac{C_{\mathbf{t}}}{(1+s)^{\sum t_k}} \int_0^\infty \dots \int_0^\infty y_1^{t_1-1} \dots y_r^{t_r-1} e^{-\sum y_k} \mathbb{E} e^{-uH(y_1, \dots, y_r, X)} dy_1 \dots dy_r,$$

which is the product of two functions, one depending on s , and the other on u . Therefore $V_\bullet$ and $H(\mathbf{V}, X)$ are independent as claimed. $\square$

We will now use Lemma 4.1 to prove the following proposition that is fundamental to our approach as it provides a simple way to show that solutions to (1.22) are beta distributed.

Proposition 4.2. *Random equation (1.22) has solution $X \sim \beta(a, b)$ for some values $a, b > 0$ iff*

$$V_1 \stackrel{d}{=} AV_1 + BV_2, \quad \text{where } V_1, V_2, (A, B) \text{ are mutually independent,} \quad (4.5)$$

holds for $(V_1, V_2) \sim \Gamma_a \otimes \Gamma_b$.

Proof. Suppose that relation (4.5) holds for $(V_1, V_2) \sim \Gamma_a \otimes \Gamma_b$. Then

$$AV_1 + BV_2 = \frac{AV_1 + BV_2}{V_1 + V_2}(V_1 + V_2) \stackrel{d}{=} \frac{AV_1 + BV_2}{V_1 + V_2} \tilde{V}, \quad (4.6)$$

where $\tilde{V} \sim \Gamma_{a+b}$ is independent of (A, B, V_1, V_2) , the equality in distribution following from Lemma 4.1. Similarly, we have that

$$V_1 = \frac{V_1}{V_1 + V_2}(V_1 + V_2) \stackrel{d}{=} \frac{V_1}{V_1 + V_2} \tilde{V}, \quad (4.7)$$

where $\tilde{V} \sim \Gamma_{a+b}$ is independent of (V_1, V_2) . Taking logarithms of the right-hand sides of (4.6) and (4.7), we obtain from (4.5) that

$$\ln \left(\frac{AV_1 + BV_2}{V_1 + V_2} \right) + \ln(\tilde{V}) \stackrel{d}{=} \ln \left(\frac{V_1}{V_1 + V_2} \right) + \ln(\tilde{V}). \quad (4.8)$$

Denoting by ψ , φ and χ the characteristic functions of the first, second (and fourth), and third terms in (4.8), respectively, we have

$$\psi(u)\varphi(u) = \chi(u)\varphi(u).$$

It is not hard to show that $\varphi(u) \neq 0$. Therefore $\psi \equiv \chi$, and so

$$\frac{V_1}{V_1 + V_2} \stackrel{d}{=} \frac{AV_1 + BV_2}{V_1 + V_2}.$$

We conclude that random equation (1.22) has solution $Y \stackrel{d}{=} V_1/(V_1 + V_2) \sim \beta(a, b)$.

Conversely, suppose that random equation (1.22) has solution $Y \sim \beta(a, b)$. Following the same steps as above in reverse order, we conclude that (4.5) holds for $(V_1, V_2) \sim \Gamma_a \otimes \Gamma_b$. Proposition 4.2 is proved. $\square$

Proposition 4.2 relates the original sequences of iterations of functions on $[0, 1]$ to the following forward and backward iterations defined on $\mathbb{R}_+$. For the family $\{g_{a,c}(x') = ax' + c : (a, c) \in [0, 1] \times \mathbb{R}_+\}$ and an i.i.d. sequence $\{(A_n, B_n, V_n)\}_{n \geq 1}$ of random vectors with distribution $\hat{\mu} \otimes \Gamma_b$, set $G_n := g_{A_n, B_n V_n}$. Denote the forward iterations of $\{G_n\}_{n \geq 1}$ by

$$X'_n(\cdot) := G_n \circ G_{n-1} \circ \cdots \circ G_1(\cdot), \quad n \geq 1, \quad X'_0(\cdot) = \text{id}(\cdot), \quad (4.9)$$

and its backward iterations by

$$Y'_n(\cdot) := G_1 \circ G_2 \circ \cdots \circ G_n(\cdot), \quad n \geq 1, \quad Y'_0(\cdot) = \text{id}(\cdot). \quad (4.10)$$

Starting from a fixed $x' > 0$ and writing $X'_n = X'_n(x')$, the forward process (4.9) is given by

$$X'_n := A_n X'_{n-1} + B_n V_n, \quad n \geq 1. \quad (4.11)$$

This is also a special case of recursion (1.24), and so X'_n converges in distribution to a unique (in distribution) X' when condition (1.25) is satisfied. In our case, since $A_1 \in [0, 1]$, this conditions is satisfied when

$$\mathbb{P}(A_1 = 1) < 1. \quad (4.12)$$

4.2.3 Some equivalent claims

For $(a, b) \in \mathbb{R}_+^2$, we introduce the following claims to be used in the statements of assertions in the remainder of this chapter:

[A1] Suppose $V_2 \sim \Gamma_b$. Then $V_1 \sim \Gamma_a$ is the unique solution to (4.5).

[A2] $X \sim \beta(a, b)$ is the unique solution to (1.22).

[A3] For $Y_n(x)$ given by (1.31), as $n \rightarrow \infty$, one has $Y_n(x) \xrightarrow{a.s.} Y \sim \beta(a, b)$ uniformly in $x \in [0, 1]$.

[A4] For $X_n(x)$ being the location of the particle after the n th step of the Markov chain (1.30) starting from $x \in [0, 1]$, one has $X_n(x) \xrightarrow{d} X \sim \beta(a, b)$ as $n \rightarrow \infty$.

[A5] For $X'_n(x')$ being the location of the particle after the n th step of the Markov chain (4.9) starting from $x' \in \mathbb{R}_+$, one has $X'_n(x') \xrightarrow{d} X' \sim \Gamma_a$ as $n \rightarrow \infty$.

The next proposition shows that to establish [A1]–[A5], it suffices to just prove [A1] (or [A2]).

Proposition 4.3. *Suppose that the distribution $\hat{\mu}$ of (A_1, B_1) is such that the convergence conditions (1.26), (4.12) hold, and at least one of [A1] or [A2] holds true. Then all assertions [A1]–[A5] hold true.*

Proof. Since the convergence conditions (1.26), (4.12) are satisfied, any solutions to random equations (1.22) and (4.5) are unique. Therefore, by Proposition 4.2, [A1] holds if and only if [A2] holds. Further, the law of V_1 from [A1] is the stationary distribution of the Markov chain from [A5], and the law of X from [A2] is the stationary distribution of the Markov chain from [A4] which coincides with the law of the limiting location in the nested interval scheme from [A3] by the contraction principle in [CL91]. $\square$

Now consider several classes of distributions $\hat{\mu}$ on $[0, 1]^2$ satisfying convergence conditions (1.26), (4.12), with beta distributed solutions to random equation (1.22). Uniqueness of the solutions to (1.22) and (4.5) is implied by the convergence conditions.

4.2.4 Model 1

In this model we assume that for some fixed $p \in (0, 1)$, we have

$$(A, B) \stackrel{d}{=} I_p(1 - L, 0) + (1 - I_p)(1, R), \quad (4.13)$$

where $I_p \sim Be(p)$ is independent of the random vector $(L, R) \in (0, 1)^2$ and $Be(p)$ denotes the Bernoulli distribution with success probability p . Clearly this model satisfies the convergence conditions (1.26) and (4.12).

This model has two important features. Firstly, since $A > B$ a.s., the nested interval scheme generated by (1.31) satisfies the independence property [I] (see p.67). Secondly, the representation (4.13) allows for a simple description of the nested

interval scheme generated by (1.31), the Markov chain on $[0, 1]$ given by (1.27), and the Markov chain on $\mathbb{R}_+$ given by (4.11). These processes are described respectively in (1A), (1B), and (1C) below. In particular, the nested interval scheme ($\langle G_2, R_I | R_I \rangle$ in the general case) described by (1A) and the Markov chain described by (1B) are dual in the sense of the contraction principle (Proposition 1) in [CL91].

- (1A) Starting with the unit interval, we choose the two subintervals $[0, 1 - L]$ and $[R, 1]$, i.e. by moving the right (resp. left) end point of the interval to the left (resp. right) a distance given by the proportion L (resp. R) of the interval. Next, select the interval $[0, 1 - L]$ with probability p , or $[R, 1]$ with probability $1 - p$, continuing independently in the same way on the chosen subinterval. If $1 - L \stackrel{d}{=} R$, then this is the CGZ model ($\langle S_2, R_I | R_I \rangle$ in the general case).
- (1B) A particle starts from a point $x \in [0, 1]$ and moves left the distance Lx with probability p , or right the distance $R(1 - x)$ with probability $1 - p$. The process continues in this way, each new step being independent of the past.
- (1C) A particle starts from a point $x' \in \mathbb{R}_+$ and moves left the distance Lx' with probability p , or right the distance RV_1 with probability $1 - p$, where $V_1 \sim \Gamma_a$ for some $a > 0$. The process continues in this way, each new step being independent of the past.

Clearly, assertions [A3]–[A5] give the limiting distributions of the processes described in (1A)–(1C), respectively.

To introduce our first special case of Model 1, consider the following $\langle S_2, D_I | R_I \rangle$ version of the CGZ model. Fix a $p \in (0, 1)$ and, starting with the unit interval, choose the interval $[0, p]$ with probability p , or $[p, 1]$ with probability $1 - p$. The next nested interval is obtained in the same way from the first chosen interval, and so on. It is clear that this scheme can be generated by a single uniform random variable $U \sim U[0, 1]$ as follows. On the first step, of the two intervals $[0, p]$ and $(p, 1]$, we choose the interval containing the value of U . At the next step, we partition the chosen interval into two with the same ratio of lengths, and again choose the subinterval containing the value of U , and so on. It is obvious from

this representation that the limiting location of the nested intervals has distribution $U[0, 1]$.

Thus [A2] holds for this scheme. It follows from Proposition 4.3 that [A1] will also hold. That fact, however, admits an elementary direct proof as well.

Theorem 4.4. *Under assumption (4.13), let $L = (1 - p)$, $R = p$. Then [A1] holds for $a = b = 1$.*

Proof. Since the V_i 's are independent exponential when $a = b = 1$, one has

$$\begin{aligned}\mathbb{E}e^{it(AV_1 + BV_2)} &= p\mathbb{E}e^{itpV_1} + (1 - p)\mathbb{E}e^{it(V_1 + pV_2)} \\ &= p(1 - itp)^{-1} + (1 - p)(1 - it)^{-1}(1 - itp)^{-1} = (1 - it)^{-1},\end{aligned}$$

which is the characteristic function of V_1 as required. $\square$

Note that assertion [A4] under the conditions of Theorem 4.4 extends Theorem 2 in [SP00] (which we presented as Theorem 2.13 above), where it was proved in the special case $p = 1/2$. Likewise, assertion [A3] in this case when interpreted for the products of 2×2 stochastic matrices (cf. (1.37)) extends the result of example on p.160 of [R71] that also considered the special case when $p = 1/2$ (cf. Section 2.5.1).

The next case extends the situation considered in [SP08] where assertion [A4] from Theorem 4.5 below was first proved (cf. Theorem 2.15), extending that assertion in the case $z = 1$ from [SP00] and [DF99] (cf. Theorem 2.11).

Theorem 4.5. *Under assumption (4.13), let $L \stackrel{d}{=} R \sim \beta(1, z)$ for a fixed $z > 0$. Then [A1]–[A5] holds for $(a, b) = ((1 - p)z, pz)$.*

Proof. According to Proposition (4.3) it suffices to show that [A2] holds. Clearly, random equation (1.22) can be written as

$$X \stackrel{d}{=} (1 - C)X + CD, \tag{4.14}$$

where X is independent of

$$(C, D) = (1 - A + B, B/(1 - A + B)) = \begin{cases} (L, 0), & I_p = 1, \\ (R, 1), & I_p = 0, \end{cases} \tag{4.15}$$

and (C, D) has distribution

$$(C, D) \sim \beta(1, z) \otimes Be(1 - p). \quad (4.16)$$

Therefore, it suffices to show that $X \sim \beta((1-p)z, pz)$ is the unique solution to (4.14) with (C, D) distributed as per (4.16). Uniqueness follows from Lemma 1.5 in [V79] since condition (1.25) is satisfied for an i.i.d. sequence with the same distribution as $(1 - C)$. The fact that $X \sim \beta((1-p)z, pz)$ is the solution follows from Section 3 in [S94] (see also Proposition 1 in [SP08] where it was shown that (4.14) holds for $X \sim \beta((1-p)z, pz)$ using the method of moments). $\square$

Remark 4.1. Observe that, for $z = 1$, assertion [A3] under the conditions of Theorem 4.5 gives the limit distribution of the following $\langle S_2, R_I | R_I \rangle$ version of the CGZ model that was considered in [K93]. Choose a uniformly distributed point U_1 in $[0, 1]$ and select (independently of U_1) the subinterval $[0, U_1]$ with probability p , or $[U_1, 1]$ with probability $1 - p$. Continuing independently in the same way on the chosen subinterval, the limiting location of this interval splitting scheme has distribution $\beta(1 - p, p)$.

One can also show that Theorem 4.5 holds by proving assertion [A1] analytically using characteristic functions. In doing so now, we will establish the limiting distribution in two cases that have not been considered previously in the literature, by using the well-known properties of the Gauss hypergeometric function (cf. (1.15)).

Indeed, for $(V_1, V_2) \sim \Gamma_z \otimes \Gamma_y$ with $z = (1 - p)v$, $y = pv$, $L \stackrel{d}{=} R \sim \beta(1, v)$ and (A, B) given by (4.13), we have

$$\begin{aligned} & \mathbb{E}e^{it(AV_1 + BV_2)} \\ &= p\mathbb{E}e^{it(1-L)V_1} + (1-p)\mathbb{E}e^{it(V_1 + RV_2)} \\ &= \frac{y}{y+z} \int_0^1 \mathbb{E}(e^{it(1-L)V_1} | 1-L=u) \mathbb{P}((1-L) \in du) \\ & \quad + \frac{z}{y+z} \mathbb{E}e^{itV_1} \int_0^1 \mathbb{E}(e^{itRV_2} | R=u) \mathbb{P}(R \in du) \\ &= \frac{y}{y+z} \int_0^1 (1-itu)^{-z} u^{y+z-1} B(y+z, 1)^{-1} du \end{aligned}$$

$$\begin{aligned}
& + \frac{z}{y+z} (1-it)^{-z} \int_0^1 (1-itu)^{-y} (1-u)^{y+z-1} B(1, y+z)^{-1} du \\
& = \frac{y}{y+z} {}_2F_1(z, y+z; y+z+1; it) + \frac{z}{y+z} (1-it)^{-z} {}_2F_1(y, 1; y+z+1; it)
\end{aligned} \tag{4.17}$$

$$= \frac{y}{y+z} {}_2F_1(z, y+z; y+z+1; it) + \frac{z}{y+z} {}_2F_1(z+1, y+z; y+z+1; it) \tag{4.18}$$

$$= {}_2F_1(z, y+z+1; y+z+1; it) \tag{4.19}$$

$$= (1-it)^{-z}, \tag{4.20}$$

where (4.18), (4.19), and (4.20) follow from (1.18), (1.17), and (1.16), respectively. The right hand side of (4.20) is the characteristic function of V_1 , and so [A1] holds as claimed.

We will now use the expression (4.17) and the well-known identity

$${}_2F_1(a, b; c; \eta) = {}_2F_1(b, a; c; \eta) \tag{4.21}$$

to establish the limiting distribution in our next case. Indeed, using (4.21) it is clear that expression (4.17) is equal to

$$\frac{y}{y+z} {}_2F_1(z, y+z; y+z+1; it) + \frac{z}{y+z} (1-it)^{-z} {}_2F_1(1, y; y+z+1; it). \tag{4.22}$$

Hence assertion [A1] holds under assumption (4.13) with $L \sim \beta(1, y+z)$, $R \sim \beta(y, z+1)$, $p = y/(y+z)$, and $(V_1, V_2) \sim \Gamma_y \otimes \Gamma_1$. We have proved the following theorem.

Theorem 4.6. *Under assumption (4.13), let $L \sim \beta(1, y+z)$, $R \sim \beta(y, z+1)$, and $p = y/(y+z)$, $y, z > 0$. Then [A1]–[A5] hold for $(a, b) = (z, 1)$.*

We can derive yet another case from (4.22) as follows. Recalling that (4.20) is equal to (4.22), we have

$$\begin{aligned}
(1-it)^{-z} &= \frac{y}{y+z} {}_2F_1(z, y+z; y+z+1; it) + \frac{z}{y+z} (1-it)^{-z} {}_2F_1(1, y; y+z+1; it) \\
&= \frac{y}{y+z} (1-it)^{1-z} {}_2F_1(1, y+1; y+z+1; it) \\
&\quad + \frac{z}{y+z} (1-it)^{-z} {}_2F_1(1, y; y+z+1; it),
\end{aligned}$$

the second equality following from relations (4.21) and 15.3.3 in [AS64]. Therefore

$$(1 - it)^{-1} = \frac{y}{y + z} {}_2F_1(1, y + 1; y + z + 1; it) + \frac{z}{y + z} (1 - it)^{-1} {}_2F_1(1, y; y + z + 1; it).$$

The last equality implies that assertion [A1] holds under assumption (4.13) with $L \sim \beta(z, y + 1)$, $R \sim \beta(y, z + 1)$, $p = y/(y + z)$, and $(V_1, V_2) \sim \Gamma_1 \otimes \Gamma_1$. The following theorem is proved.

Theorem 4.7. *Under assumption (4.13), let $L \sim \beta(z, y + 1)$, $R \sim \beta(y, z + 1)$, and $p = y/(y + z)$, $y, z > 0$. Then [A1]–[A5] hold for $a = b = 1$.*

Remark 4.2. Note that assertion [A4] under the conditions of Theorem 4.7 is an extension of Proposition 2 in [SP08] (cf. Theorem 2.16 above), where it was proved in the special case $y = z$. In that case, we can derive an interesting distributional relation involving mixtures of gamma distributions from which we will obtain an apparently new explicit form solution to the perpetuity equation as a special case.

Indeed, using the equivalent random equation (4.14), we have from assertion [A2] under conditions of Theorem 4.7 that

$$U \stackrel{d}{=} U(1 - C) + I_{1/2}C,$$

for $(U, C, I_{1/2}) \sim U[0, 1] \otimes \beta(y, y + 1) \otimes Be(1/2)$. Therefore, for $(V_1, V_2) \sim \Gamma_{y/2} \otimes \Gamma_{y/2+1}$

$$\begin{aligned} U(\tilde{V}_1 + \tilde{V}_2) &\stackrel{d}{=} \frac{UV_2 + I_{1/2}V_1}{V_1 + V_2}(\tilde{V}_1 + \tilde{V}_2) \\ &\stackrel{d}{=} \frac{UV_2 + I_{1/2}V_1}{V_1 + V_2}(V_1 + V_2) = UV_2 + I_{1/2}V_1, \end{aligned} \quad (4.23)$$

where $(\tilde{V}_1, \tilde{V}_2)$ is an independent copy of (V_1, V_2) , and the second equality in distribution follows from Lemma 4.1.

Let $\{\gamma(t)\}_{t \geq 0}$, $\gamma(0) = 0$, be a gamma process with independent increments such that $\gamma(t) - \gamma(s) \sim \Gamma_{t-s}$ for $0 \leq s < t$. Suppose also that $\{\gamma_1(\cdot)\}$, $\{\gamma_2(\cdot)\}, \dots$ are independent copies of the process $\{\gamma(t)\}_{t \geq 0}$, $U \sim U[0, 1]$, and $J_1, J_2, \dots$ are i.i.d. with distribution $Be(1/2)$. Assuming independence of all the processes and random

variables introduced above, we can rewrite random equation (4.23) as

$$U\gamma(y+1) \stackrel{d}{=} U\gamma(y/2+1) + J_1\gamma_1(y/2).$$

By repeated substitution, we obtain

$$U\gamma(y+1) \stackrel{d}{=} U\gamma(1) + \sum_{k=1}^{\infty} J_k\gamma_k(y/2^k),$$

where the series converges a.s. by Kolmogorov's two series theorem (see e.g. p.386 in [S95]). By conditioning on $\{J_k\}$ we see that

$$\sum_{k=1}^{\infty} J_k\gamma_k(y/2^k) \stackrel{d}{=} \gamma_1\left(y \sum_{k=1}^{\infty} \frac{J_k}{2^k}\right) \stackrel{d}{=} \gamma_1(yU_1),$$

where $U_1 \sim U[0, 1]$ is independent of $(\{\gamma(\cdot)\}, \{\gamma_1(\cdot)\}, U)$. So we obtained the following curious relation:

$$U\gamma(y+1) \stackrel{d}{=} U\gamma(1) + \gamma_1(yU_1). \quad (4.24)$$

In particular, for $y = 1$, the random variable

$$U\gamma(1) + \gamma_1(U_1)$$

is exponentially distributed, and therefore $V \sim U[0, 1]$ is the unique solution to the perpetuity equation (1.23) with $(D, C) = (U, \gamma_1(U_1))$.

One can also show that (4.24) holds by computing the Laplace transform (or characteristic function) of each side. A search of the literature yielded only one paper [J05] that considered the mixture $\gamma_1(yU_1)$. In that paper the explicit form for the density of $\gamma(U)/\gamma(1)$ was given.

The last special case of Model 1 we consider is as follows.

Theorem 4.8. *Under assumption (4.13), let $2L \stackrel{d}{=} 2R \sim \beta(1, z)$, $z > 0$, and $p = 1/2$. Then [A1]–[A5] hold for $a = b = z + 1$.*

Proof. In view of Proposition 4.3, we will only prove [A1]. For $M \sim \beta(1, z)$ independent on (V_1, V_2) , the characteristic function of the right-hand side of (4.5) with

$a = b = z + 1$ is

$$\begin{aligned}\mathbb{E}e^{it(AV_1 + BV_2)} &= \frac{1}{2}\mathbb{E}e^{it(1-M/2)V_1} + \frac{1}{2}\mathbb{E}e^{it(V_1 + MV_2/2)} \\ &= \frac{1}{2}\int_0^1 \frac{(1-u)^{z-1}z}{(1-(1-u/2)it)^{z+1}} du + \frac{1}{2}\mathbb{E}e^{itV_1}\mathbb{E}e^{itMV_2/2}.\end{aligned}$$

The last term above is equal to $(1-it)^{-z-1}(1-it/2)^{-1}/2$ due to the well-known property of “beta-gamma algebra” that for $(X, Y) \sim \beta(a, b) \otimes \Gamma_{a+b}$, one has $XY \sim \Gamma_a$ (cf. (1.12)). The integral in the second last term can be computed and is equal to $(1-it)^{-z}(1-it/2)^{-1}$. Therefore, the characteristic function of $AV_1 + BV_2$ is equal to

$$\frac{(1-it)^{-z}}{2(1-it/2)} + \frac{(1-it)^{-z-1}}{2(1-it/2)} = (1-it)^{-z-1}$$

which means that equation (4.5) is satisfied. $\square$

Assertion [A3] under conditions of Theorem 4.8 gives the limiting distribution of an $\langle S_2, R_I | R_I \rangle$ version of the CGZ model where the splitting point has density

$$f(u) = \begin{cases} z(1-2u)^{z-1}, & 0 < u < 1/2, \\ z(2u-1)^{z-1}, & 1/2 \leq u < 1, \\ 0, & \text{elsewhere,} \end{cases} \quad (4.25)$$

and the larger of the intervals is chosen at each step. In particular, the following corollary holds true.

Corollary 4.9. *Starting with the unit interval, choose a random splitting point S_1 with density (4.25) and select the larger of the intervals $[0, S_1], [S_1, 1]$. Continue independently in this way on the chosen subinterval. Then the limiting location of the scheme has distribution $\beta(z+1, z+1)$.*

Note that Corollary 4.9 extends the results in [CGZ84] and [DLS86], where its claim was proved for $z = 1$ (cf. Theorem 2.4 above).

4.2.5 Model 2

Now consider a more general model where the support of $\hat{\mu}$ can be an arbitrary (closed) subset of the whole unit square $[0, 1]^2$. In this more general model, the nested interval scheme, Markov chain $\{X_n\}$ on $[0, 1]$ given by (1.27), and Markov chain $\{X'_n\}$ on $\mathbb{R}_+$ given by (4.11) are described respectively in (2A), (2B), and (2C) below, with the limiting distributions given by assertions [A3]–[A5], respectively.

(2A) Starting with the unit interval $\mathbb{I}_0 = [0, 1]$, the interval $\mathbb{I}_1$ is given by $[\min(A_1, B_1), \max(A_1, B_1)]$ with $(A_1, B_1) \sim \hat{\mu}$ (cf. Section 4.2.1). For $n \geq 1$, the $(n + 1)$ st interval is chosen from $\mathbb{I}_n$ using the distribution

$$\hat{\mu}_{n+1} = \begin{cases} \hat{\mu}_n & \text{if } A_n > B_n, \\ \hat{\mu}'_n & \text{otherwise,} \end{cases}$$

shifted and scaled to $\mathbb{I}_n^2$ (see [M2] on p.67), with $\hat{\mu}_0 := \hat{\mu}$. Recall that if either of the conditions [M1]–[M3] on $\hat{\mu}$ is satisfied, then the independence property [I] holds true.

(2B) A particle starts from a point $x \in [0, 1]$ and moves to the point $A_1x + B_1(1 - x)$. The process continues in this way, each new step being independent of the past.

(2C) A particle starts from a point $x' \in \mathbb{R}_+$ and moves to the point $A_1x' + B_1V_1$ for $V_1 \sim \Gamma_a$, for some $a > 0$. The process continues in this way, each new step being independent of the past.

First we will consider a special case of Model 2 discussed in the pioneering work of DeGroot and Rao [DR63]. The result of their Example 6 (which we presented as Theorem 2.10 above) coincides with assertion [A4] in the following theorem.

Theorem 4.10. *If $(A, B) \sim \beta(w, y) \otimes \beta(y, z)$ for some $w, y, z > 0$, then [A1]–[A5] hold for $(a, b) = (w + y, y + z)$.*

Proof of Theorem 4.10. As usual, will only prove [A1]. By (1.12) we have that $AV_1 \sim \Gamma_w$ and $BV_2 \sim \Gamma_y$. Hence, by independence, the sum $AV_1 + BV_2$ has the same distribution as $V_1 \sim \Gamma_{w+y}$, as required. $\square$

In particular, if $w = z$, then [M3] holds, and so assertion [A3] under the conditions of Theorem 4.10 gives the limiting location of the $\langle G_1, R_I | D_I \rangle$ nested interval scheme described in the following corollary.

Corollary 4.11. *Starting with the unit interval, choose two independent points $A_1 \sim \beta(z, y)$ and $B_1 \sim \beta(y, z)$ in $[0, 1]$, and select the interval with these end points. Continue independently in this way on the chosen subinterval. Then the limiting location of the scheme has distribution $\beta(y + z, y + z)$.*

Corollary 4.12. *If $A \sim \beta(w, y)$ and $B \equiv 1$, then assertions [A1]–[A5] hold for $(a, b) = (w + y, y)$.*

Proof. To prove [A1], set $B \equiv 1$ in the proof of Theorem 4.10. □

Remark 4.3. The interval-valued process (2A) generated by $\hat{\mu}$ from Theorem 4.12 is the following $\langle S_2, R_G | D_G \rangle$ interval splitting model that apparently has not appeared previously in the literature. The random rule represented by R_G prescribes to choose the splitting point S_n with the shifted and scaled distribution $\beta(w, y)$ (resp. $\beta(y, w)$) on $\mathbb{I}_{n-1}$ if n is odd (resp. even). The deterministic rule represented by D_G prescribes to select the right (resp. left) interval if n is odd (resp. even). Then, by assertion [A3] under the conditions of Theorem 4.12, the limiting location of this scheme has law $\beta(w + y, y)$.

Remark 4.4. Now consider the process $\bar{X} := \{\bar{X}_n\}_{n \geq 0}$ defined in terms of the forward iteration (1.27) as follows. If n is even set $\bar{X}_n := X_n$, otherwise set $\bar{X}_n := 1 - X_n$. Then $\bar{X}$ is described as follows. A particle starts from a fixed $\bar{X}_0 = x \in [0, 1]$ and moves in alternating directions at each step. The particle first moves left to $\bar{X}_1$ with the shifted and scaled distribution $\beta(y, w)$ on $[0, \bar{X}_0]$. In the next step, the particle moves right to $\bar{X}_2$ with the shifted and scaled distribution $\beta(w, y)$ on $[\bar{X}_0, 1]$, and then continues this cycle independently of the past. We have from assertion [A4] under the conditions of Theorem 4.12 that the distribution of $\bar{X}_{2n}$ converges weakly to $\beta(w + y, y)$ and that of $\bar{X}_{2n+1}$ converges weakly to $\beta(w, w + y)$ as $n \rightarrow \infty$. Although this result extends Theorem 1 in [SP00] (cf. Theorem 2.12 above) where the special case when $w = y = 1$ was first proved, it is not a new one. It is a special case of a more general model considered on p.2 in [L02].

It turns out we can extend Corollary 4.12 when $w = y = 1$ as we will show in our next theorem.

Theorem 4.13. *If $(A, B) \sim U[p, 1] \otimes Be(1 - p)$, $p \in (0, 1)$, then [A1]–[A5] hold for $(a, b) = (2, 1)$.*

Proof. As usual, we will only prove [A1]. The characteristic function for the first term of the right hand side of (4.5) is given by

$$\mathbb{E}e^{itAV_1} = \frac{1}{1-p} \int_p^1 \frac{du}{(1-itu)^2} = \frac{1}{(1-it)(1-ipt)}.$$

Similarly, for the last term in (4.5) we have that

$$\mathbb{E}e^{itBV_2} = p + \frac{1-p}{1-it} = \frac{1-ipt}{1-it}. \quad (4.26)$$

Therefore, by independence, the characteristic function of the right hand side of (4.5) is $(1-it)^{-2}$. Assertion [A1] holds, and hence the theorem is proved. $\square$

Our next result is a simple consequence of the following lemma.

Lemma 4.14. *Suppose $(A, B, V_1, V_2) \sim U[0, y/(w+y)] \otimes \beta(w, y) \otimes \Gamma_2 \otimes \Gamma_{w+y+1}$ for $w, y > 0$. Then*

$$AV_1 + BV_2 \sim \Gamma_{w+1}. \quad (4.27)$$

Proof. The characteristic function for the first term on the left hand side of (4.27) is given by

$$\mathbb{E}e^{itAV_1} = \frac{w+y}{y} \int_0^{y/(w+y)} \frac{du}{(1-itu)^2} = \left(1 - \frac{ity}{w+y}\right)^{-1}.$$

The characteristic function of the term BV_2 is

$$\mathbb{E}e^{itBV_2} = B(w, y)^{-1} \int_0^1 \frac{u^{w-1}(1-u)^{y-1}}{(1-itu)^{w+y+1}} du = (1-it)^{-w-1} \left(1 - \frac{ity}{w+y}\right).$$

Therefore, by independence, we obtain that the characteristic function of the left hand side of (4.27) is $(1-it)^{-w-1}$, as required. $\square$

Corollary 4.15. *If $(A, B) \sim U[0, y/(y+1)] \otimes \beta(1, y)$, then [A1]–[A5] hold for $(a, b) = (2, y+2)$.*

Proof. Assertion [A1] follows by setting $w = 1$ in Lemma 4.14. $\square$

Theorem 4.16. *If $(A, B) = (A', A'B')$ for $(A', B') \sim \beta(w+y, y) \otimes \beta(y, w)$, then [A1]–[A5] hold for $a = b = w+y$.*

Proof. Now we will only prove [A2]. For $(V_1, V_2, V_3) \sim \Gamma_y \otimes \Gamma_w \otimes \Gamma_y$, one has that

$$(A', B') \stackrel{d}{=} \left(\frac{V_1 + V_2}{V_1 + V_2 + V_3}, \frac{V_1}{V_1 + V_2} \right).$$

Therefore it suffices to show that

$$X \stackrel{d}{=} \frac{V_2 X + V_1}{V_1 + V_2 + V_3},$$

which by Lemma 4.1 is equivalent to showing that

$$X(V_1 + V_2 + V_3) \stackrel{d}{=} XV_2 + V_1.$$

The last relation follows from Theorem 2 (B) in [D98], as required. $\square$

Remark 4.5. Clearly $A > B$ a.s. under the conditions of Theorem 4.16, and so setting

$$(C, D) := (1 - A + B, B/(1 - A + B)),$$

it follows that $X \sim \beta(w+y, w+y)$ satisfies perpetuity equation (4.14) for $(C, D) \sim \beta(2y, w) \otimes \beta(y, y)$.

Remark 4.6. Note that assertion [A3] under the conditions of Theorem 4.16 gives the limiting distribution of the following apparently new $\langle G_1, R_I | D_I \rangle$ nested interval scheme. The random rule represented by R_I prescribes to choose the interval with right end point $A'_1 \sim \beta(w+y, y)$, and left end point with distribution $\beta(y, w)$ shifted and scaled to $[0, A'_1]$. The deterministic rule represented by D_I prescribes to select $\mathbb{I}_1$ as the interval generated from the first stage. We continue independently in the same way on $\mathbb{I}_1$ ad infinitum.

In the special case where w and y are integers, the above nested interval scheme corresponds to choosing an interval with end points taken from the uniform order statistics $U_{(1)}, U_{(2)}, \dots, U_{(d)}$ as follows. For $k = 1, 2, \dots, \lfloor d/2 \rfloor$, if one chooses $\mathbb{I}_1 = [U_{(k)}, U_{(d-k+1)}]$ and continues in the same way on $\mathbb{I}_1$ independently of the past, then the limiting location of this scheme has distribution $\beta(d - k + 1, d - k + 1)$. This is an extension of Kennedy's scheme in [K88] (cf. Section 2.2.2 above), where the above claim was proved for $k = 1$.

4.3 Other applications

In this section we will apply the methods and results presented above to two closely related areas of research. The first deals with randomly weighted averages introduced by Van Assche in [V87] (see Section 2.4). The second is concerned with stick-breaking constructions of beta distributed random variables.

4.3.1 Randomly weighted averages

Recall from Van Assche [V87] that choosing a random variable uniformly on a random interval with arcsine distributed end points is the same as choosing a random variable uniformly on $[0, 1]$ (cf. Theorem 2.20 above). In this Section, we will apply the results of Section 4.2 to provide several examples of triples $(X, A, B) \in [0, 1]^3$ satisfying this property, i.e., the result of choosing a random variable with distribution $\mathcal{L}(X)$ on a random interval with end points having distribution $\mathcal{L}((A, B))$ is the same as choosing a random variable with distribution $\mathcal{L}(X)$ on $[0, 1]$. We will call a random vector (X, A, B) with this property an *invariant triple*. Clearly, the example from Van Assche [V87] coincides with the invariant triple

$$(X, A, B) \sim U[0, 1] \otimes \beta(1/2, 1/2) \otimes \beta(1/2, 1/2).$$

Also recall the more basic approach taken by Johnson and Kotz in [JK90a] which culminated in the observation that the random variable Z given by (2.14) defines a randomly weighted average when A and B are i.i.d. However, it is clear that (2.14) defines a randomly weighted average in the more general case when A and B are

dependent and not identically distributed, the only required condition is that the symmetric condition (2.15) holds for W . It follows that for X satisfying random equation (1.22) and the symmetric condition (2.15), the random vector (X, A, B) is an invariant triple. We will now present several examples of these invariant triples, each following from assertion **[A2]** under the conditions of a theorem from Section 4.2.

Example 4.1. For (A, B) under the conditions of Theorem 4.4 with $p = 1/2$, and $X \sim U[0, 1]$ independent of (A, B) , the random vector (X, A, B) is an invariant triple.

Example 4.2. For (A, B) under the conditions of Theorem 4.5 with $p = 1/2$, and $X \sim \beta(z/2, z/2)$, $z > 0$, independent of (A, B) , the random vector (X, A, B) is an invariant triple.

Example 4.3. For (A, B) under the conditions of Theorem 4.7 with $p = y/(y + z)$, $y, z > 0$, and $X \sim U[0, 1]$ independent of (A, B) , the random vector (X, A, B) is an invariant triple.

Example 4.4. For (A, B) under the conditions of Theorem 4.8 with $p = 1/2$, and $X \sim \beta(z + 1, z + 1)$, $z > 0$, independent of (A, B) , the random vector (X, A, B) is an invariant triple.

Our next example appeared recently in [R14] and was proved using Stieltjes transforms. A comment by the authors of [JK90a] that first principles can sometimes be more advantageous than more sophisticated techniques, such as the method of Stieltjes transforms in [V87], is also relevant here. Indeed, the next example follows from the implied assertion **[A2]** under the conditions of Theorem 4.10, and that assertion follows from elementary “beta-gamma algebra”.

Example 4.5. For (A, B) under the conditions of Theorem 4.10 with $z = w$, and $X \sim \beta(w + y, w + y)$ independent of (A, B) , the random vector (X, A, B) is an invariant triple.

Example 4.6. For $(A, B) = (A', A'B')$ for $(A', B') \sim \beta(x + y, y) \otimes \beta(y, x)$, and $X \sim \beta(x + y, x + y)$ independent of (A, B) , the random vector (X, A, B) is an invariant triple.

We will now consider the generalisation (2.19) proposed by Johnson and Kotz [JK90a]. As discussed in Section 2.4, several recent papers have considered the randomly weighted average Z given by (2.19) (cf. [SH09], [SR12], [H12], [HHM13], [RS13], [R14]). These papers present several examples of random variables Z defined by (2.19) that are beta distributed. In some instances, the special cases considered are actually a result of the more general Theorem 1 in [CL94] (cf. Theorem 2.25 above). In particular, by setting $c = 2$ in that theorem, we obtain the following corollary.

Corollary 4.17 ([CL94]). *Suppose $(X_1, \dots, X_d) \sim \beta(a_1, b_1) \otimes \dots \otimes \beta(a_d, b_d)$ is independent of $(W_1, \dots, W_d) \sim D_{(a_1+b_1, \dots, a_d+b_d)}$. Then, for Z given by (2.19), one has that $Z \sim \beta(a_\bullet, b_\bullet)$.*

In our next theorem, we will show that if the parameters $(a_1, \dots, a_d, b_1, \dots, b_d)$ satisfy $a_1 b_\bullet = b_1 a_\bullet$, then there is another Dirichlet law, such that for a vector $(W_1, \dots, W_d)$ with that law one still has that $Z \sim \beta(a_\bullet, b_\bullet)$ as in Corollary 4.17.

Theorem 4.18. *Suppose $(X_1, \dots, X_d) \sim \beta(a_1, b_1) \otimes \dots \otimes \beta(a_d, b_d)$, $a_1 b_\bullet = b_1 a_\bullet$, is independent of $(W_1, \dots, W_d) \sim D_{(a_1+b_1+1, a_2+b_2, \dots, a_d+b_d)}$. Then, for Z given by (2.19), one has that $Z \sim \beta(a_\bullet, b_\bullet)$.*

Proof. We begin by presenting a distributional equality that we will use to prove our claim. Recall from the proof of Theorem 4.5 that random equation (4.14) with (C, D) given by (4.16) has unique solution $X \sim \beta((1-p)z, pz)$. It follows from Lemma 4.1 that, for $(Y_1, Y_2, I_{1-p}) \sim \Gamma_z \otimes \Gamma_1 \otimes Be(1-p)$, one has

$$X(Y_1 + Y_2) \stackrel{d}{=} XV_1 + I_{1-p}Y_2 \stackrel{d}{=} Y_3 + I_{1-p}Y_2, \quad (4.28)$$

where $Y_3 \sim \Gamma_{(1-p)z}$ (by (1.12)) is independent of (I_{1-p}, Y_2) , and (Y_1, Y_2) is independent of X .

It is clear from Lemma 4.1 that to prove Theorem 4.18 it suffices to show that

$$\tilde{Z}V_\bullet \stackrel{d}{=} \sum_{i=1}^d V_i X_i \quad (4.29)$$

for $(\tilde{Z}, V_1, \dots, V_d) \sim \beta(a_\bullet, b_\bullet) \otimes \Gamma_{a_1+b_1+1} \otimes \Gamma_{a_2+b_2} \otimes \dots \otimes \Gamma_{a_d+b_d}$ and $(V_1, \dots, V_d)$ independent of $(X_1, \dots, X_d)$. Now the left hand side of (4.29) has the same form as the left hand side of (4.28) for $z = a_\bullet + b_\bullet$ and $p = b_\bullet/(a_\bullet + b_\bullet)$, and we conclude that

$$\tilde{Z}V_\bullet \stackrel{d}{=} Y_3 + I_{a_\bullet/(a_\bullet+b_\bullet)}Y_2, \quad (4.30)$$

where $(Y_2, Y_3, I_{a_\bullet/(a_\bullet+b_\bullet)}) \sim \Gamma_{a_\bullet} \otimes \Gamma_1 \otimes Be(a_\bullet/(a_\bullet + b_\bullet))$. For the right hand side of (4.29), it is clear from (1.12) that $\sum_{i=2}^d V_i X_i \sim \Gamma_{a_\bullet - a_1}$, and $V_1 X_1$ has the same form as the left hand side of (4.28) for $z = a_1 + b_1$ and $p = b_1/(a_1 + b_1)$. Hence

$$\sum_{i=1}^d V_i X_i \stackrel{d}{=} Y_2 + I_{a_1/(a_1+b_1)}Y_3, \quad (4.31)$$

where $(Y_2, Y_3, I_{a_1/(a_1+b_1)}) \sim \Gamma_{a_\bullet} \otimes \Gamma_1 \otimes Be(a_1/(a_1 + b_1))$.

Since the right hand sides of (4.30) and (4.31) are equal in law if and only if $a_1/(a_1 + b_1) = a_\bullet/(a_\bullet + b_\bullet)$ which holds iff $a_1 b_\bullet = b_1 a_\bullet$, the theorem is proved. $\square$

In particular, Theorem 3.1 (b) in [H12] is a special case of our Theorem 4.18.

Our next theorem is a generalisation of Example 2 in [SR12], where it was proved in the case when $a_1 = a_2 = \dots = a_d = 1$.

Theorem 4.19. *Suppose*

$$(X_1, \dots, X_d) \sim \beta(a_1 + 1/2, a_1 + 1/2) \otimes \dots \otimes \beta(a_d + 1/2, a_d + 1/2)$$

is independent of $(W_1, \dots, W_d) \sim D_{(a_1, \dots, a_d)}$. Then Z given by (2.19) is distributed according to $\beta(a_\bullet + 1/2, a_\bullet + 1/2)$.

Proof. As in the proof of Theorem 4.18, it suffices to show that

$$\tilde{Z}V_\bullet \stackrel{d}{=} \sum_{i=1}^d X_i V_i, \quad (4.32)$$

for $(\tilde{Z}, V_1, \dots, V_d) \sim \beta(a_\bullet + 1/2, a_\bullet + 1/2) \otimes \Gamma_{a_1} \otimes \dots \otimes \Gamma_{a_d}$, and $(V_1, \dots, V_d)$ independent of $(X_1, \dots, X_d)$. The characteristic function of the left hand side of (4.32)

is

$$\begin{aligned}
\mathbb{E}e^{itZV_\bullet} &= B(a_\bullet + 1/2, a_\bullet + 1/2)^{-1} \int_0^1 (1 - itu)^{-a_\bullet} u^{a_\bullet - 1/2} (1 - u)^{a_\bullet - 1/2} \\
&= {}_2F_1(a_\bullet, a_\bullet + 1/2; 2a_\bullet + 1; it) \\
&= \left(\frac{2}{1 + \sqrt{1 - it}} \right)^{2a_\bullet}, \tag{4.33}
\end{aligned}$$

where (4.33) follows from 15.1.13 in [AS64] (cf. (1.16)). Since the terms on the right hand side of (4.32) are independent of each other, the characteristic function of the sum is the product of the characteristic functions of all the terms. The characteristic function of the i th term is given by

$$\begin{aligned}
\mathbb{E}e^{itX_iV_i} &= B(a_i + 1/2, a_i + 1/2)^{-1} \int_0^1 (1 - itu)^{-a_i} u^{a_i - 1/2} (1 - u)^{a_i - 1/2} \\
&= {}_2F_1(a_i, a_i + 1/2; 2a_i + 1; it) \\
&= \left(\frac{2}{1 + \sqrt{1 - it}} \right)^{2a_i},
\end{aligned}$$

and so the characteristic function of the right hand side of (4.32) is given by the right hand side of (4.33), as required. $\square$

4.3.2 Stick-breaking constructions

In this section, we will show that some special cases in Section 4.2 define stick-breaking constructions of beta distributed random variables. We will begin by reviewing the stick-breaking process. A *stick-breaking process* $\{Y_n\}_{n \geq 1}$ is generated by a sequence of $[0, 1]$ -valued random variables $A_1, A_2, \dots$ by

$$Y_1 = A_1, \quad Y_n = (1 - A_1) \cdots (1 - A_{n-1})A_n, \quad n \geq 2. \tag{4.34}$$

That is, $\{Y_n\}_{n \geq 1}$ is generated by recursively breaking off random fractions A_n of an interval as follows. Starting with the unit interval, break off the random fraction A_1 from the left end of the interval and set $Y_1 := A_1$ to be the left-end point of the remaining interval. This process is repeated by breaking off the random fraction A_2

from the left end of the remaining stick of length $1 - A_1$, where $Y_2 := A_2(1 - A_1)$ is again the left-end point of the remaining interval. Continuing in the same way on the remaining stick, we obtain (4.34), where clearly $\sum_{i=1}^{\infty} Y_i \stackrel{a.s.}{=} 1$ when the A_k 's are not identically equal to zero. In particular, when the A_n 's are i.i.d. with law $\beta(1, z)$, the random vector $(Y_1, Y_2, \dots)$ has the so-called GEM (Griffiths-Engen-McCloskey) distribution with parameter $z > 0$. Letting $\tilde{Y}_1 \geq \tilde{Y}_2 \geq \dots$ be the ranked values of the Y_n 's, the random vector $(\tilde{Y}_1, \tilde{Y}_2, \dots)$ has the well-known Poisson-Dirichlet distribution with parameter $z > 0$.

Now suppose that the A_n 's are i.i.d. Then using the approach of Sethuraman in [S94], one can construct a random variable X on $[0, 1]$ by taking random proportions $W_1, W_2, \dots$ of the intervals $Y_1, Y_2, \dots$ by

$$\begin{aligned}
X &:= \sum_{i=1}^{\infty} W_i Y_i \\
&= \sum_{i=1}^{\infty} W_i A_i \sum_{j=1}^{i-1} (1 - A_j) \\
&= W_1 A_1 + (1 - A_1) \sum_{i=2}^{\infty} W_i A_i \sum_{j=2}^{i-1} (1 - A_j) \\
&\stackrel{d}{=} W_1 A_1 + (1 - A_1) X.
\end{aligned} \tag{4.35}$$

That is, X satisfies random equation (4.14) for $(C, D) = (A_1, W_1)$, and so any random variable satisfying (4.14) has a stick-breaking construction by (4.35). We will now give four examples of these stick-breaking constructions, each following from the implied assertion [A2] under the conditions of theorems from Section 4.2 (by applying transformation (4.15) to the random variables (A, B) from that assertion).

Our first example is the well-known stick-breaking construction of $X \sim \beta(a, b)$ with arbitrary $a, b > 0$ that follows from the work of Sethuraman in [S94] (see also Section 2.1 in [PZWGC10]). The result is implied by assertion [A2] of Theorem 4.5.

Example 4.7. The law of the random variable defined by (4.35) is $\beta(a, b)$ when $(A_1, W_1), (A_2, W_2), \dots$ are i.i.d. with law $\beta(1, a + b) \otimes Be(a/(b + b))$, and $Y_1, Y_2, \dots$ are defined by (4.34).

Our second example follows from Theorem 4.6 when $y = z$.

Example 4.8. For all $a > 0$, the law of the random variable defined by (4.35) is $U[0, 1]$ when $(A_1, W_1), (A_2, W_2), \dots$ are i.i.d. with law $\beta(a, a + 1) \otimes Be(1/2)$, and $Y_1, Y_2, \dots$ are defined by (4.34).

We note that the fact that $X \sim U[0, 1]$ satisfies (4.14) for $(C, D) = (A_1, W_1) \sim \beta(a, a + 1) \otimes Be(1/2)$ was first shown in the proof of Proposition 2 in [SP08].

Our next example is an apparently new stick-breaking construction for a symmetric beta distributed random variable. The result is implied by assertion [A2] of Theorem 4.8.

Example 4.9. The random variable defined by (4.35) is $\beta(a + 1, a + 1)$ distributed when $(2A_1, W_1), (2A_2, W_2), \dots$ are i.i.d. with law $\beta(1, a) \otimes Be(1/2)$, and $Y_1, Y_2, \dots$ are defined by (4.34).

Our final example which follows from assertion [A2] of Theorem 4.16 also appears to be a new stick-breaking construction for a symmetric beta distributed random variable.

Example 4.10. The random variable defined by (4.35) is $\beta(a + b, a + b)$ distributed when $(A_1, W_1), (A_2, W_2), \dots$ are i.i.d. with law $\beta(2b, a) \otimes \beta(b, b)$, and $Y_1, Y_2, \dots$ are defined by (4.34).

Chapter 5

Stationary distributions for a class of bounded Markov chains

In this chapter we consider an extended version of the donkey walk on $[0, 1]$ from Chapter 4, where the direction of the next move is chosen at random with probability depending on the current location (see Section 2.3.3 for a review of the relevant literature on models of this type). Under simple broad conditions, we establish the ergodicity of such Markov chains and then derive closed form expressions for the stationary densities of the chains when the proportions are beta distributed with the first parameter equal to 1. Examples demonstrating the range of stationary distributions for processes described by this model are given, and an application to a robot coverage algorithm is discussed.

5.1 Introduction

This chapter is mostly devoted to deriving explicit formulae for the stationary densities for the following extended version donkey walk on $[0, 1]$ that appear in some interesting applications (see e.g. Section 4 in [IN03], Section 5 in [RL10], and Section 5.3 below). Recall that the chain dynamics are as follows. Let F_L and F_R be two fixed distributions on $[0, 1]$, $p : [0, 1] \rightarrow [0, 1]$ a fixed measurable function. Given the chain value $X_n = x \in [0, 1]$ at time n , the next jump of the process is to the left with probability $p(x)$ or to the right with probability $1 - p(x)$. If the jump is to the left, its length is given by an independent random proportion $L_{n+1} \sim F_L$ of the length of the interval $[0, x]$. Otherwise, the chain jumps to the right for a distance given by an independent random proportion $R_{n+1} \sim F_R$ of the length of the interval $[x, 1]$.

That is, the evolution of our Markov chain $X = \{X_n\}_{n \geq 0}$ is given by

$$X_{n+1} = X_n - X_n L_{n+1} I_{n+1} + (1 - X_n) R_{n+1} (1 - I_{n+1}), \quad n = 0, 1, \dots, \quad (5.1)$$

where $X_0 = x_0 \in [0, 1]$, $I_{n+1} := \mathbf{1}_{\{U_{n+1} < p(X_n)\}}$, and $\{L_n\}_{n \geq 1}$, $\{R_n\}_{n \geq 1}$ and $\{U_n\}_{n \geq 1}$ are independent sequences of i.i.d. random variables such that $L_n \sim F_L$, $R_n \sim F_R$, and $U_n \sim U[0, 1]$, the uniform law on $[0, 1]$.

The above model was apparently first introduced in Section 2.1 of [DF99] for $p(x) \equiv p$ and $F_L = F_R = U[0, 1]$ (see Theorem 2.11 above). The case of non-constant $p(x)$ was also discussed, but not pursued in [DF99]. As discussed in the Sections 2.3.2 and 2.3.2, further special cases of that model for various choices of $p(x)$ and distributions F_L , F_R were considered in [SP00, BW02, SP08, P09, RL10]. In particular, [P09] dealt with the case when $p(x) \equiv x$ and $F_L = F_R = \beta(1, z)$, $z > 0$. It was shown in [P09] that in that case X was ergodic with stationary law $\beta(z, z)$ (see Theorem 2.18 above). The case when $F_L = F_R = U[0, 1]$ and $p(x)$ is piecewise continuous was studied in [RL10].

In this chapter, in the case where $F_L = \beta(1, l)$, $F_R = \beta(1, r)$ for some $l, r > 0$, we derive the stationary density for piecewise continuous $p(x)$ satisfying a natural condition that ensures ergodicity. In particular, for $F_L = F_R = \beta(1, z)$, $z > 0$, and

linear $p(x) = cx + (1 - b)(1 - x)$, $b, c \in (0, 1]$, the Markov chain X is ergodic with stationary distribution $\beta(bz, cz)$ (see Section 5.3.1 below). We find that many of the existing results on the form of the stationary density for the above model are special cases of our more general Theorem 5.3 below. We also show how the same approach can be used to compute (at least, numerically) the stationary distribution when $F_L = \beta(l_1, l_2)$, $F_R = \beta(r_1, r_2)$, $l_1, r_1 \in \mathbb{N}$, $l_2, r_2 > 0$, by solving a two-point boundary value problem for a system of $l_1 + r_1$ ordinary differential equations.

One of the main reasons for considering such more general models is that the class of their stationary laws is far richer than in the special case $l = r = 1$. In particular, a larger variety of unimodal and multimodal stationary densities can be obtained, including densities with arbitrarily high and “sharp” peaks at any given locations in the sample space (by choosing large enough l and r and an appropriate p). We will use this feature to generalise the robot coverage algorithm from [RL10] (see Section 5.3.3 below).

The Markov chain X is also a special case of the so-called give-and-take model that was introduced, in its deterministic form, in the context of human genetics in [L61] and then studied in [DR63] (cf. Section 2.3.1 above). The Markov chain X is a version of that model where at each step only one player exchanges her fortune with the other. At time $n \geq 0$, if $I_{n+1} = 0$, player 1 (whose fortune is given by X_n) receives a proportion R_{n+1} of player 2’s fortune (which is given by $1 - X_n$), otherwise if $I_{n+1} = 1$, player 2 receives a proportion L_{n+1} of player 1’s fortune.

In Section 5.2, we establish the ergodicity of the Markov chain (5.1) under simple broad conditions and derive the form of its stationary density in the ergodic case when $F_L = \beta(1, l)$, $F_R = \beta(1, r)$, $l, r > 0$. Some examples (extending results from [BW02, SP08, P09, RL10]) are presented in Section 5.3.

5.2 Main results

First we show that the following three conditions imply the ergodicity of the Markov chain X . They are by no means necessary for ergodicity, but are quite suitable for our model in the context of the present study.

[E1] For some $\delta \in (0, 1/2)$,

$$\sup_{x \in [0, \delta]} \max\{p(x), 1 - p(1 - x)\} =: 1 - \varepsilon < 1.$$

[E2] For some $\delta \in (0, 1/2)$,

$$\max\{F_L(1 - \delta), F_R(1 - \delta)\} =: 1 - \eta < 1.$$

[E3] There exist $\delta \in (0, 1/2)$ and non-empty interval $(s, t) \subset (\delta, 1 - \delta)$, such that F_L and F_R have densities f_L and f_R on the intervals $(1 - t - \delta, 1 - s)$ and $(s - \delta, t)$, respectively. Moreover, for

$$g(z) := \min \left\{ \inf_{y \in [1 - \delta, 1]} f_L\left(\frac{y - z}{y}\right), \inf_{y \in [0, \delta]} f_R\left(\frac{z - y}{1 - y}\right) \right\}, \quad z \in (s, t),$$

one has

$$\gamma := \int_s^t g(z) dz > 0.$$

Lemma 5.1. *If the Markov chain X given by (5.1) satisfies conditions [E1]–[E3] with a common δ , then X is ergodic.*

Proof. Let $\tau_B(x) := \inf\{k \geq 1 : X_k(x) \in B\}$, where $X_n(x)$ denotes the value of the chain after n steps when $X_0 = x$, and $B \in \mathcal{B}([0, 1])$ (the Borel subsets of $[0, 1]$). It follows from Theorems 1.3, 2.1 in [B98] that X will be ergodic provided that there exists a $V \in \mathcal{B}([0, 1])$, a probability measure φ on $([0, 1], \mathcal{B}([0, 1]))$, and a $q > 0$ such that

$$(I') \quad \mathbb{P}(\tau_V(x) < \infty) = 1, \quad \forall x \in [0, 1],$$

$$(I) \quad \sup_{x \in V} \mathbb{E}\tau_V(x) < \infty,$$

$$(II) \quad \mathbb{P}(X_1(x) \in B) \geq q\varphi(B), \quad \forall x \in V, \quad \forall B \in \mathcal{B}([0, 1])$$

(note that condition (II) implies aperiodicity of X).

Setting $V := [0, \delta] \cup [1 - \delta, 1]$ we have by condition [E2] that

$$\begin{aligned} \mathbb{P}(X_1(x) \in V) &= \mathbb{P}(X_1(x) \in V | I_1 = 1)p(x) + \mathbb{P}(X_1(x) \in V | I_1 = 0)(1 - p(x)) \\ &\geq \mathbb{P}(L_1 > 1 - \delta)p(x) + pr(R_1 > 1 - \delta)(1 - p(x)) \end{aligned}$$

$$\geq \eta$$

for all $x \in [0, 1]$. Therefore, the standard argument yields that, for any $x_0 \in [0, 1]$ and $n \geq 1$,

$$\begin{aligned} \mathbb{P}(\tau_V(x_0) > n) &= \mathbb{P}\left(\bigcap_{k=1}^n \{X_k(x_0) \notin V\}\right) \\ &= \underbrace{\int_{V^c} \cdots \int_{V^c}}_{n-1} \left[\prod_{k=1}^{n-1} \mathbb{P}(X_1(x_{k-1}) \in dx_k) \right] \mathbb{P}(X_1(x_{n-1}) \notin V) \\ &\leq (1 - \eta) \underbrace{\int_{V^c} \cdots \int_{V^c}}_{n-2} \left[\prod_{k=1}^{n-2} \mathbb{P}(X_1(x_{k-1}) \in dx_k) \right] \mathbb{P}(X_1(x_{n-2}) \notin V) \\ &\leq (1 - \eta)^n. \end{aligned}$$

Hence (I') and (I) hold true.

To show that (II) is also satisfied, let

$$\varphi(B) := \frac{1}{\gamma} \int_{B \cap (s, t)} g(y) dy, \quad B \in \mathcal{B}([0, 1]). \quad (5.2)$$

Using conditions **[E1]** and **[E3]**, we have, for $B \in \mathcal{B}([0, 1])$ and $x \in [0, \delta]$,

$$\begin{aligned} \mathbb{P}(X_1(x) \in B) &\geq \mathbb{P}(X_1(x) \in B \cap (s, t)) \\ &= (1 - p(x)) \int_{B \cap (s, t)} \frac{1}{1 - x} f_R\left(\frac{y - x}{1 - x}\right) dy \\ &\geq \varepsilon \int_{B \cap (s, t)} g(y) dy \\ &= \varepsilon \gamma \varphi(B). \end{aligned}$$

A similar argument shows that the above lower bound $\varepsilon \gamma \varphi(B)$ also holds when $x \in [1 - \delta, 1]$. Therefore (II) is met with $q = \varepsilon \gamma$ and φ defined by (5.2). The lemma is proved. $\square$

Now we turn to deriving closed form expressions for stationary distributions in the ergodic case when both F_L and F_R are absolutely continuous, with densities

denoted by $f_L(x)$ and $f_R(x)$, $x \in (0, 1)$, respectively. Since the transition probability of X now has density

$$f(x, y) := \mathbb{P}(X_1(x) \in dy)/dy = \begin{cases} \frac{1-p(x)}{1-x} f_R\left(\frac{y-x}{1-x}\right), & 0 \leq x < y < 1, \\ \frac{p(x)}{x} f_L\left(\frac{x-y}{x}\right), & 0 < y < x \leq 1, \end{cases} \quad (5.3)$$

the stationary distribution Π of X (when it exists) will also be absolutely continuous with density $\pi(x)$ satisfying the usual integral equation

$$u(y) = \int_0^1 u(x) f(x, y) dx, \quad 0 < y < 1. \quad (5.4)$$

The existence of the stationary density is obvious from the standard relations

$$\begin{aligned} \Pi(B) &= \int_0^1 \mathbb{P}(X_1(x) \in B) \Pi(dx) = \int_0^1 \left[\int_B f(x, y) dy \right] \Pi(dx) \\ &= \int_B \left[\int_0^1 f(x, y) \Pi(dx) \right] dy = \int_B \pi(y) dy, \quad B \in \mathcal{B}([0, 1]), \end{aligned}$$

where $\pi(y) := \int_0^1 f(x, y) \Pi(dx)$ and the second last equality follows from Fubini's theorem.

The case when $f_L(x) \equiv f_R(x) \equiv 1$, $x \in (0, 1)$, was studied in [RL10]. In that case, one can differentiate integral equation (5.4) to obtain a simple separable differential equation that is easily solved to give a closed form for the stationary density $f(x)$ (coinciding with our f from (5.15) below with $l = r = 1$).

We extend this result to the case where $f(x, y)$ has the semidegenerate form: for some $N, M \geq 1$,

$$f(x, y) = \begin{cases} \sum_{i=1}^N a_i(y) b_i(x), & 0 \leq x < y < 1, \\ \sum_{j=1}^M c_j(y) d_j(x), & 0 < y \leq x \leq 1, \end{cases} \quad (5.5)$$

and the factors satisfy the following conditions:

[C1] all $a_i(y), c_j(y)$ are continuous on $(0, 1)$;

[C2] all $b_i(x)$ are piecewise continuous on $[0, 1)$;

[C3] all $d_j(x)$ are piecewise continuous on $(0, 1]$.

The next assertion shows that the stationary distribution of the Markov chain, with transition density (5.5), solves a two-point boundary value problem for a system of $N + M$ ordinary differential equations. Our result is an extension of Theorem 3.1 in [G72] that was proved under the more restrictive assumption that all $a_i(y)$, $b_i(x)$, $c_j(y)$, and $d_j(x)$ in (5.5) are continuous on $[0, 1]$.

Theorem 5.2. *Suppose that $f(x, y)$ has the semidegenerate form (5.5) satisfying [C1]–[C3] and let (5.4) have a solution $u(y)$ that is integrable on $(0, 1)$. Then the following claims are true.*

(i) *Any such solution $u(y)$ is continuous on $(0, 1)$.*

(ii) *If*

$$\alpha_i(y) := \int_0^y b_i(x)u(x) dx, \quad i = 1, 2, \dots, N, \quad (5.6)$$

$$\beta_j(y) := \int_y^1 d_j(x)u(x) dx, \quad j = 1, 2, \dots, M, \quad (5.7)$$

then

$$u(y) = \sum_{i=1}^N a_i(y)\alpha_i(y) + \sum_{j=1}^M c_j(y)\beta_j(y), \quad 0 < y < 1. \quad (5.8)$$

Let S be the union of finite discontinuity sets for $b_i(x)$, $i = 1, 2, \dots, N$, and $d_j(x)$, $j = 1, 2, \dots, M$. Then

$$\alpha'_i(y) = b_i(y)u(y), \quad y \in (0, 1) \setminus S, \quad (5.9)$$

$$-\beta'_j(y) = d_j(y)u(y), \quad y \in (0, 1) \setminus S, \quad (5.10)$$

$$\alpha_i(0+) = 0, \quad i = 1, 2, \dots, N, \quad \beta_j(1-) = 0, \quad j = 1, 2, \dots, M. \quad (5.11)$$

(iii) *Conversely, let $\alpha_i(y)$, $i = 1, 2, \dots, N$, and $\beta_j(y)$, $j = 1, 2, \dots, M$, be continuous solutions of (5.8)–(5.11) such that $u(y)$ is integrable on $(0, 1)$. Then $u(y)$ given by (5.8) is a solution of (5.4).*

Remark 5.1. There appears to be a typo in Theorem 3.1 in [G72], which contains $c_j(x)$ in the expressions on the right hand sides of (5.7) and (5.10) instead of $d_j(x)$, as above.

Proof. (i) We have

$$\begin{aligned} u(y) &= \int_0^1 u(x)f(x,y) dx \\ &= \sum_{i=1}^N a_i(y) \int_0^y b_i(x)u(x) dx + \sum_{j=1}^M c_j(y) \int_y^1 d_j(x)u(x) dx. \end{aligned} \quad (5.12)$$

Each $a_i(y)$ and $c_j(y)$ is continuous on $(0, 1)$ by [C1], and so the continuity of $u(y)$, $y \in (0, 1)$, follows since each integral on the right hand side above is absolutely continuous as $u(x)$ is integrable on $(0, 1)$ and, by [C2] and [C3], for any $\varepsilon > 0$, $b_i(x)$ and $d_j(x)$ are bounded on $[0, 1 - \varepsilon]$ and $[\varepsilon, 1]$, respectively.

(ii) As the right-hand side of (5.12) coincides with that of (5.8) in view of (5.6) and (5.7), relation (5.8) is obvious. Differentiating (5.6) and (5.7) at $y \in (0, 1) \setminus S$ (which is possible as the integrands are continuous at such y), we obtain (5.9) and (5.10), respectively. The boundary conditions (5.11) follow from the definitions (5.6), (5.7).

(iii) Now suppose that $\alpha_i(y)$, $i = 1, 2, \dots, N$, and $\beta_j(y)$, $j = 1, 2, \dots, M$, are continuous solutions to (5.8)–(5.11). Since, as above, for any $\varepsilon > 0$, $b_i(x)$ and $d_j(x)$ are bounded on $[0, 1 - \varepsilon]$ and $[\varepsilon, 1]$, respectively, and $u(y)$ is integrable on $(0, 1)$, we have from (5.9)–(5.11) and the assumed continuity that the functions $\alpha_i(y)$ and $\beta_j(y)$ are given by the right hand sides of (5.6) and (5.7), respectively. Substituting these representations into (5.8) shows that $u(y)$ satisfies (5.4) with $f(x, y)$ given by (5.5). The theorem is proved. $\square$

Theorem 5.2 allows us to easily derive the form of the stationary distribution when

$$F_L = \beta(1, l), \quad F_R = \beta(1, r), \quad l, r > 0. \quad (5.13)$$

Indeed, in this case $f_L(y) = l(1 - y)^{l-1}$, $f_R(y) = r(1 - y)^{r-1}$, $y \in (0, 1)$, so that the transition density (5.3) for the chain has the semidegenerate form (5.5) with

$N = M = 1$ and

$$a_1(y) = r(1-y)^{r-1}, \quad b_1(x) = \frac{1-p(x)}{(1-x)^r}, \quad c_1(y) = ly^{l-1}, \quad d_1(x) = \frac{p(x)}{x^l}. \quad (5.14)$$

Theorem 5.3. *Assume that $p(x)$ is piecewise continuous on $[0, 1]$ and satisfies [E1], and that (5.13) holds true. Then X is ergodic with stationary density*

$$\pi(x) = Cx^l \left(\frac{r}{1-x} + \frac{l}{x} \right) \exp \left(-r \int_{1/2}^x \frac{p(t)}{1-t} dt - l \int_{1/2}^x \frac{p(t)}{t} dt \right), \quad 0 < x < 1, \quad (5.15)$$

where $C > 0$ is a normalising constant such that $\int_0^1 \pi(x) dx = 1$.

Proof. It follows from the assumptions that X also satisfies [E2] and [E3] (all with a common δ) and so is ergodic by Lemma 5.1. Since the transition probabilities have densities (5.3), X has a stationary density π .

Clearly, the set S of discontinuity points of the factors b_i and d_j (see Theorem 5.2) coincides here with the finite set of all values x and $1-x$ such that x is a discontinuity point of $p(\cdot)$, and the factors (5.14) satisfy conditions [C1]–[C3].

Substituting (5.14) into (5.9)–(5.10) with $u(y)$ given by (5.8) and letting $\alpha(y) := \alpha_1(y)$, $\beta(y) := \beta_1(y)$, we obtain

$$\begin{aligned} \alpha'(y)(1-y)^r &= (1-p(y)) (r(1-y)^{r-1}\alpha(y) + ly^{l-1}\beta(y)), \quad y \in (0, 1) \setminus S, \\ -\beta'(y)y^l &= p(y) (r(1-y)^{r-1}\alpha(y) + ly^{l-1}\beta(y)), \quad y \in (0, 1) \setminus S. \end{aligned} \quad (5.16)$$

Adding the equations yields

$$\alpha'(y)(1-y)^r - r(1-y)^{r-1}\alpha(y) = \beta'(y)y^l + ly^{l-1}\beta(y), \quad y \in (0, 1) \setminus S,$$

which is equivalent to $(\alpha(y)(1-y)^r)' = (\beta(y)y^l)'$. Integrating and assuming that α and β are continuous on $(0, 1)$, we conclude that

$$\alpha(y)(1-y)^r = \beta(y)y^l + C_1, \quad y \in (0, 1), \quad (5.17)$$

where the boundary condition $\alpha(0+) = 0$ ensures that $C_1 = 0$.

Substituting (5.17) into the second relation in (5.16), we obtain the separable

differential equation

$$\beta'(y) = -\beta(y)p(y) \left(\frac{r}{1-y} + \frac{l}{y} \right), \quad y \in (0, 1) \setminus S,$$

with the general solution

$$\beta(y) = C_2 \exp \left(-r \int_{1/2}^y \frac{p(t)}{1-t} dt - l \int_{1/2}^y \frac{p(t)}{t} dt \right), \quad y \in (0, 1), \quad (5.18)$$

where C_2 is a constant. Now from (5.17) we also have

$$\alpha(y) = C_2 \frac{y^l}{(1-y)^r} \exp \left(-r \int_{1/2}^y \frac{p(t)}{1-t} dt - l \int_{1/2}^y \frac{p(t)}{t} dt \right), \quad y \in (0, 1). \quad (5.19)$$

We see that both $\alpha(y)$ and $\beta(y)$ are continuous and $u(y)$ given by (5.8) is integrable on $(0, 1)$ indeed. It follows from (5.14), (5.18), (5.19) and Theorem 5.2 (iii) that the right hand side of (5.15) is a solution to integral equation (5.4) and so is equal to the stationary density of X when $C > 0$ is chosen so that $\int_0^1 \pi(x) dx = 1$. $\square$

Remark 5.2. Note that the above approach can be used to compute (at least, numerically) the stationary density of X when $p(x)$ satisfies [E1], while $F_L = \beta(l_1, l_2)$, $l_1 \in \mathbb{N}$, $l_2 > 0$, and $F_R = \beta(r_1, r_2)$, $r_1 \in \mathbb{N}$, $r_2 > 0$. Indeed, in this case we have

$$f_L \left(\frac{x-y}{x} \right) = \frac{1}{B(l_1, l_2)} \left(\frac{x-y}{x} \right)^{l_1-1} \left(\frac{y}{x} \right)^{l_2-1}, \quad 0 < y < x \leq 1,$$

and

$$f_R \left(\frac{y-x}{1-x} \right) = \frac{1}{B(r_1, r_2)} \left(\frac{y-x}{1-x} \right)^{r_1-1} \left(\frac{1-y}{1-x} \right)^{r_2-1}, \quad 0 \leq x < y < 1.$$

Since $l_1, r_1 \in \mathbb{N}$, the middle factors on the right-hand sides are polynomials in y with coefficients that are functions of x . Therefore we see that the transition probabilities (5.3) are semidegenerate with $N = r_1$, $M = l_1$. It remains to solve the two-point boundary value problem for the system of $l_1 + r_1$ ordinary differential equations.

One could also use the same approach to compute the stationary density when $p(x)$ satisfies [E1] and the distributions F_L, F_R are finite mixtures of $\beta(1, z)$ with

different z -values or, more generally, are of the form

$$f_L(x) = \sum_{j=1}^M \mu_j (1-x)^{l_j-1}, \quad \sum_{j=1}^M \mu_j / l_j = 1, \quad \mu_j \in \mathbb{R}, \quad l_j > 0, \quad j = 1, \dots, M,$$

and

$$f_R(x) = \sum_{i=1}^N \lambda_i (1-x)^{r_i-1}, \quad \sum_{i=1}^N \lambda_i / r_i = 1, \quad \lambda_i \in \mathbb{R}, \quad r_i > 0, \quad i = 1, \dots, N.$$

In this case, the transition density (5.3) is also semidegenerate, and it remains to solve the two-point boundary value problem for the system of $N + M$ ordinary differential equations.

5.3 Examples

5.3.1 The case of polynomial $p(x)$

Suppose that the function p is polynomial: for a $k \in \mathbb{N}$, one has

$$p(x) = \sum_{n=0}^k p_n x^n = \sum_{n=0}^k q_n (x-1)^n,$$

where, of course, $p_n = p^{(n)}(0)/n!$ and $q_n = p^{(n)}(1)/n!$, $n \geq 0$, and $p(x) \geq 0$ for $x \in [0, 1]$. Assuming that $p_0 < 1$ and $q_0 > 0$ to ensure that condition **[E1]** is satisfied, and that (5.13) holds true, we see that the conditions of Theorem 5.3 are met, and so the Markov chain X is ergodic. A straightforward computation of the integral in (5.15) shows that X has stationary density of the form

$$\pi(x) = C x^{l(1-p_0)-1} (1-x)^{rq_0-1} (l + (r-l)x) \exp \left(r \sum_{n=1}^k \frac{q_n}{n} (x-1)^n - l \sum_{n=1}^k \frac{p_n}{n} x^n \right),$$

$0 < x < 1$.

In particular, if $p(x)$ is linear: $p(x) = cx + (1-b)(1-x)$, where $b, c \in (0, 1]$, and $l = r = z$:

$$F_L = F_R = \beta(1, z), \quad z > 0, \tag{5.20}$$

we immediately obtain that X has stationary distribution $\beta(bz, cz)$. This is a direct extension of Theorem 3 in [P09] and Proposition 1 in [SP08] (stated above as Theorems 2.18 and 2.16), where the special cases $p(x) \equiv x$, and $p(x) \equiv p \in (0, 1)$, respectively, were considered.

5.3.2 The case of piecewise constant $p(x)$

Next we consider the case of piecewise constant $p(x)$, under the assumption (5.20). We will show, in particular, that in that case one can obtain multimodal stationary densities, with modes located at the discontinuity points of $p(x)$.

Suppose that

$$p(x) = p_i \in [0, 1], \quad s_{i-1} \leq x < s_i, \quad i = 1, \dots, k,$$

where $k \in \mathbb{N}$, $0 = s_0 < s_1 < \dots < s_k = 1$, and $p_1 < 1$, $p_k > 0$. Then $p(x)$ satisfies [E1], and Theorem 5.3 implies that the Markov chain X has stationary density

$$\pi(x) = C_i x^{z(1-p_i)-1} (1-x)^{zp_i-1}, \quad s_{i-1} \leq x < s_i, \quad i = 1, \dots, k, \quad (5.21)$$

where $C_1, \dots, C_k$ are positive constants. That is, the (continuous) stationary density of X is “glued” of pieces of different beta densities on disjoint intervals (s_{i-1}, s_i) .

To find the constants C_i , we note that, by the continuity of $\pi(x)$,

$$C_i = C_{i-1} \left(\frac{s_{i-1}}{1-s_{i-1}} \right)^{z(p_i-p_{i-1})} = C_1 \prod_{j=1}^{i-1} \left(\frac{s_j}{1-s_j} \right)^{z(p_{j+1}-p_j)}, \quad 2 \leq i \leq k. \quad (5.22)$$

Using notation $B_x(a, b) := \int_0^x t^{a-1} (1-t)^{b-1} dt$, $x \in [0, 1]$, $a, b > 0$, for the incomplete beta function, we obtain from the relation $\int_0^1 \pi(x) dx = 1$, (5.21) and (5.22) that

$$\begin{aligned} 1 &= \sum_{i=1}^k C_i [B_{s_i}(z(1-p_i), zp_i) - B_{s_{i-1}}(z(1-p_i), zp_i)] \\ &= C_1 \sum_{i=1}^k [B_{s_i}(z(1-p_i), zp_i) - B_{s_{i-1}}(z(1-p_i), zp_i)] \prod_{j=1}^{i-1} \left(\frac{s_j}{1-s_j} \right)^{z(p_{j+1}-p_j)}, \end{aligned}$$

thus yielding C_1 as the reciprocal of the sum on the right-hand side, the values of $C_2, \dots, C_k$ being now given by (5.22). Note that the assertion of Theorem 1 in [BW02] (stated above as Theorem 2.17) is a special case of the above general formula, corresponding to $(k, z, p_1, p_2, s_1) = (2, 1, 0, 1, 1/2)$.

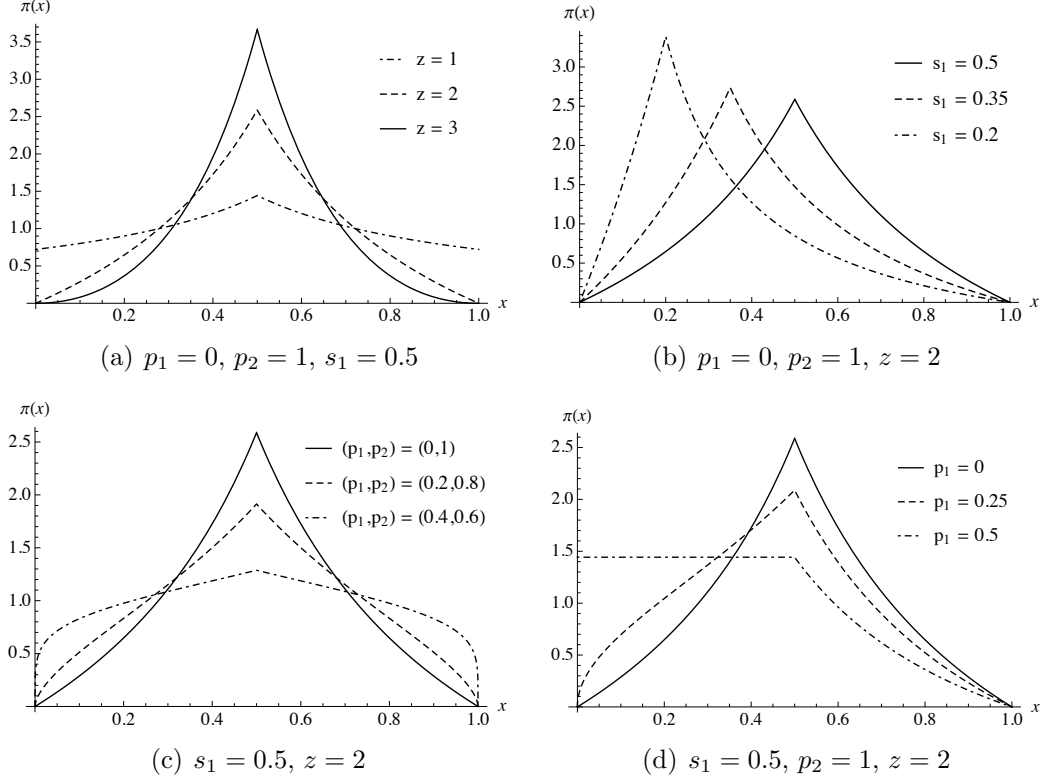


Figure 5.1: Some examples of peaked stationary densities generated by piecewise constant $p(x)$ with $k = 2$.

Some examples of peaked stationary densities of the form (5.21) with various choices of s_1 and $(k, z, p_1, p_2) = (2, 1, 0, 1)$ were given in [RL10]. That, of course, is a very special case producing a rather limited range of peaked stationary densities (see Fig. 3 in [RL10]). In the more general case where the parameter z can assume any positive value, one can create a much richer variety of (arbitrarily high-) peaked stationary densities, see Fig. 5.1 for some examples of peaked stationary densities generated by piecewise constant $p(x)$ with $k = 2$. Moreover, such functions $p(x)$ can also generate bimodal stationary densities (see Fig. 5.2), while models with $k > 2$ can have more general multimodal stationary densities (see Fig. 5.3 for examples of multimodal densities corresponding to models with $k = 6$).

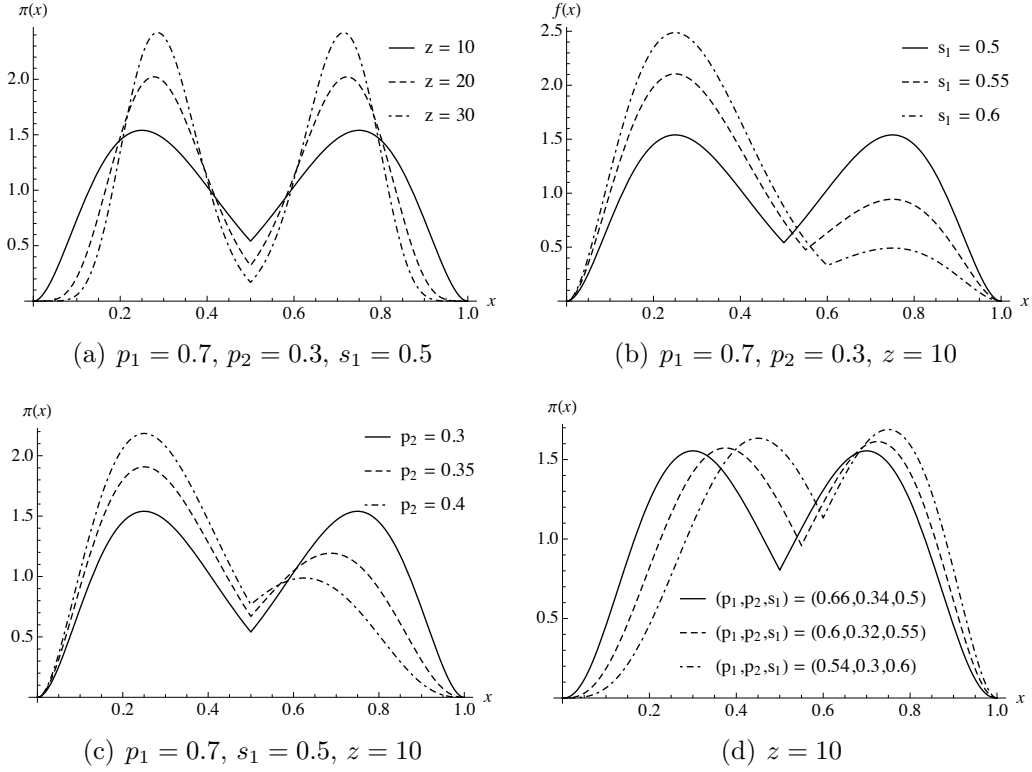


Figure 5.2: Some examples of bimodal stationary densities generated by piecewise constant $p(x)$ with $k = 2$.

5.3.3 A robot coverage algorithm and random search

In this example, we discuss a generalisation of the robot coverage algorithm suggested in Section 5 of [RL10], where the following scenario was considered.

Suppose a robot is moving periodically in a rectangular room of size $d_1 \times d_2$. At each location the robot stops, a measurement is taken, and then the robot moves to the next location according to some rule. The objective of the rule is to ensure the measurements cover the whole room, with certain areas in the room given higher priority. To achieve that, one can make the robot move according to a $D := [0, d_1] \times [0, d_2]$ -valued Markov chain $\{(X_{1,n}, X_{2,n})\}_{n \geq 0}$, with a given stationary density (that will specify the degree of attention the robot will be paying to different areas of the room).

Given that the robot is at $(x_1, x_2) \in D$ at time $n \geq 0$, at time $n + 1$ it moves to its next location in D according to the following algorithm.

Step 1: For two given measurable functions $p_i : [0, d_i] \rightarrow [0, 1]$, $i = 1, 2$, the i th

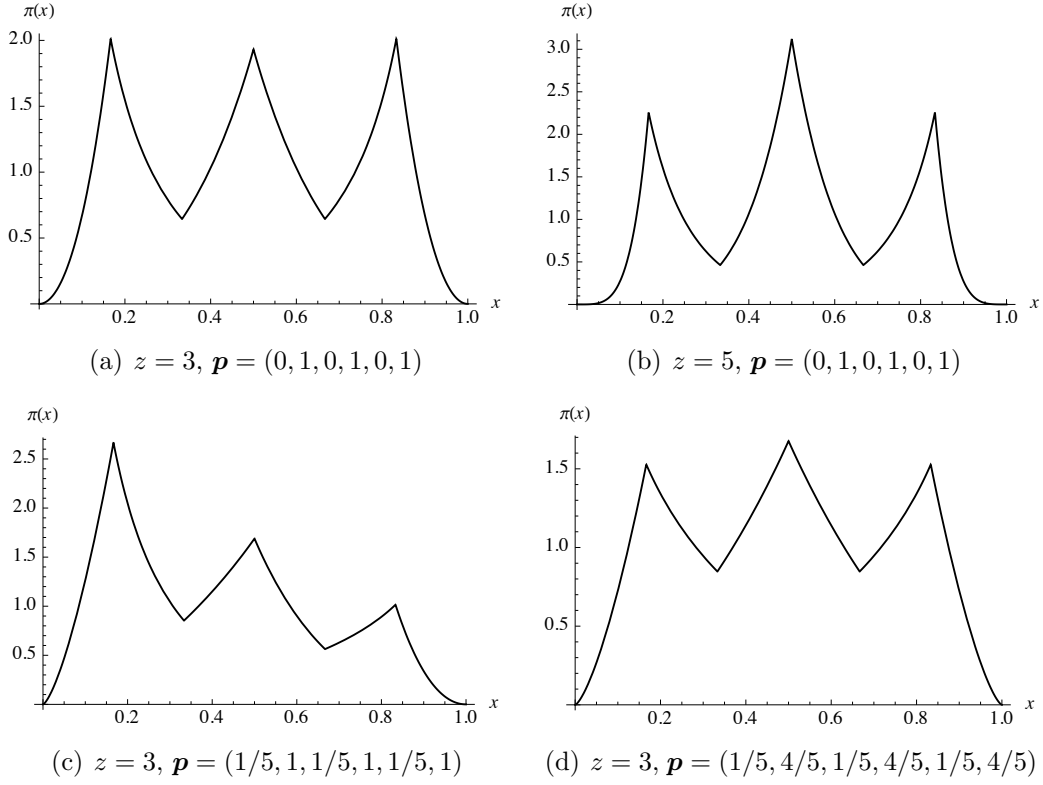


Figure 5.3: Some examples of trimodal stationary densities generated by piecewise constant $p(x)$, where $k = 6$, $\mathbf{p} := (p_1, \dots, p_6)$, $s_1 = 1/6$, $s_2 = 1/3$, $s_3 = 1/2$, $s_4 = 2/3$, $s_5 = 5/6$.

component of the displacement vector is negative with probability $p_i(x_i)$ (and positive with probability $1 - p_i(x_i)$), $i = 1, 2$, the signs of the two components being independent of each other.

Step 2: The distances Δ_i to be travelled in the i th dimensions, $i = 1, 2$, are selected at random as follows:

$$\Delta_i := \begin{cases} R_{i,n}(d_i - x_i) & \text{if moving in positive direction along axis } i, \\ (1 - L_{i,n})x_i & \text{if moving in negative direction along axis } i, \end{cases}$$

where $L_{i,n} \sim \beta(1, l_i)$, $R_{i,n} \sim \beta(1, r_i)$, $i = 1, 2$, $n \in \mathbb{N}$, are all independent of each other and of the choices made at Step 1.

Clearly, the stationary density on D of the robot location process $\{(X_{1,n}, X_{2,n})\}_{n \geq 0}$ is the product of the stationary densities of the component processes $\{X_{1,n}\}_{n \geq 0}$ and

$\{X_{2,n}\}_{n \geq 0}$. The algorithm suggested in [RL10] used indicator functions p_i and uniformly distributed $L_{i,n}$, $R_{i,n}$ only, so that the above version allows one to design much more general “preference functions” (i.e., stationary densities for the Markov chain describing the robot’s movement) for the coverage algorithm.

Now suppose, as it was done in the example in Section 5 of [RL10], that there is a single point of interest at $(y_1, y_2) := (0.2d_1, 0.5d_2)$. Setting

$$p_i(x_i) := \mathbf{1}_{\{x_i > y_i\}}, \quad i = 1, 2, \quad (5.23)$$

we obtain a stationary density on D with a single peaked mode at the point (y_1, y_2) when $l_1, r_1, l_2, r_2 \geq 1$. See Fig. 5.4 below for a plot of the stationary density when $l_1 = r_1 = l_2 = r_2 = 3$ and $d_1 = d_2 = 1$.

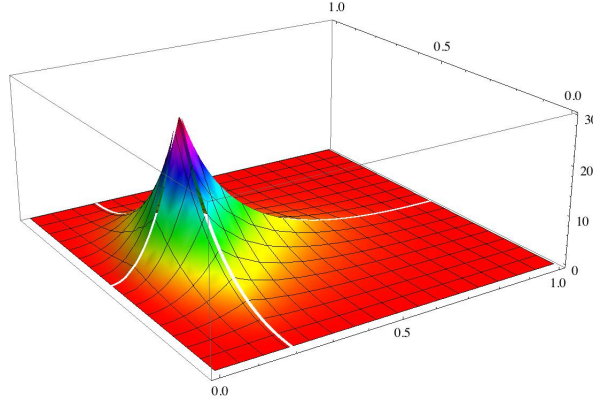


Figure 5.4: The stationary density of the robot converge algorithm with p_i given by (5.23), $l_1 = r_1 = l_2 = r_2 = 3$, and $d_1 = d_2 = 1$.

Remark 5.1. To design a robot coverage algorithm for a general bounded connected region $D \subset \mathbb{R}^d$, $d \geq 2$, with piecewise smooth boundary (rather than just a rectangle) and an arbitrary given “preference function”, one can use the Markov chain model suggested in [B94] for simulating random vectors with given densities on such sets. That model can also be used for constructing coverage algorithms on the boundaries of the sets.

We also note that the above algorithm can be modified to adapt to the level of measurements, which will basically turn it into a random sequential search algorithm. More specifically, suppose that our robot measures a scalar quantity $g(\mathbf{x})$

depending on the location $\mathbf{x}$ in the search space $D = [0, 1]^d$, $d \geq 1$. The goal is to find $\mathbf{x}^* := \arg \max_{\mathbf{x} \in D} g(\mathbf{x})$ of the global maximum of the objective function $g : D \rightarrow \mathbb{R}$. What distinguishes this setup from the usual optimisation problem is that one now aims to minimise not the amount of computation required to find a satisfactory approximation to the maximum point but, rather, the *distance traveled by the robot* in the process.

In a typical sequential random search algorithm (see e.g. Chapter 1 in [Z03]), given the current “best-found position” $\mathbf{y}$ (with the largest value of g among all the points of D visited so far), a new candidate point $\mathbf{x}$ is generated at random according to a distribution depending on the current position and the “past search history”. If $g(\mathbf{x}) > g(\mathbf{y})$ then we move to the new position $\mathbf{x}$, otherwise a new candidate point is generated, according to the same distribution (or its modification). It is well understood that it is important to incorporate a “systematic search-domain reduction into random optimisation” procedure (see e.g. [SL92]). In classical implementations of the search procedure where the new candidate point is sampled from the uniform distribution on a sphere or cube centred at $\mathbf{y}$ (see e.g. [SC76, LJ73]), this is achieved, roughly speaking, by “shrinking” the size of the respective set (sphere or cube) at an exponential rate. Alternatively, one can try to achieve basically the same effect by changing the shape of the sampling distribution (akin to changing the “temperature” in simulated annealing). That can be achieved using our results on the “peaked shape” of the stationary distribution of Markov chains.

One can construct such an algorithm as follows. Fix a value $v \in [0, 1/2]$ that will specify our function $p(x)$, and choose a sequence $\{z_n > 0\}_{n \geq 0}$, $z_n \uparrow \infty$ as $n \rightarrow \infty$.

Step 0: Initialise algorithm parameters: an initial point $\mathbf{X}_0 = (X_{0,1}, \dots, X_{0,d})$ and the iteration index $n := 0$. Set $\mathbf{Y} := \mathbf{X}_0$.

Step 1: Generate $\mathbf{X}_{n+1}$, of which the components $X_{n+1,j}$ are obtained from the respective values of $X_{n,j}$ according to transitions in d independent Markov chains of the form (5.1) with $F_L = F_R = \beta(1, z_n)$, $p(x) = 1 - v + (2v - 1)\mathbf{1}_{\{x < Y_j\}}$, $j = 1, \dots, d$.

Step 2: If $g(\mathbf{X}_{n+1}) > g(\mathbf{Y})$ then set $\mathbf{Y} := \mathbf{X}_{n+1}$ to update the best-found point.

Step 3: Set $n := n + 1$ and go to Step 1.

The procedure continues until a suitable stopping criterion is satisfied (e.g., the total travel distance reaches a prescribed level etc.).

Chapter 6

Products of i.i.d. stochastic matrices

In this chapter, we study the products of i.i.d. stochastic matrices that arise as the natural multivariate extension of the iterations of linear random functions considered in Chapter 4. In particular, we will obtain a characterisation of the class of distributions of random stochastic matrices X with the property that the products $X(n)X(n-1)\cdots X(1)$ of i.i.d. copies $X(k)$ of X converge a.s. as $n \rightarrow \infty$ and the limit is Dirichlet distributed. Our characterisation follows from a multivariate analogue of Proposition 4.2 and extends results by Chamayou and Letac [CL94] (cf. Theorems 2.25 and 2.26 above). We will also present several examples that are of interest in applications.

6.1 Introduction

Let X be a random $d \times d$ matrix with non-negative entries, and $\{X(n)\}_{n \geq 1}$ be an i.i.d. sequence of random matrices with the same distribution as X . In this chapter, we study the limit of the left and right products given by $X(n, 1)$ and $X(1, n)$, respectively (see (2.23) and (2.24) above).

Recall that Chamayou and Letac [CL94] studied the left products (2.23) for random stochastic matrices X satisfying the following conditions:

[R1] The rows of X are independent.

[R2] The rows of X are Dirichlet distributed.

[R3] Letting $(\alpha_{i,1}, \dots, \alpha_{i,d})$ be the Dirichlet parameters of the i th row of X , we have $\alpha_{i\bullet} \equiv \sum_{j=1}^d \alpha_{i,j} = \sum_{j=1}^d \alpha_{j,i} \equiv \alpha_{\bullet i}$ for $i = 1, \dots, d$.

It is shown in [CL94] (cf. Theorem 2.26 above) that the above conditions are sufficient to ensure that:

[D1] The products $X(n, 1)$ converge a.s. to some random matrix $\hat{X}$ as $n \rightarrow \infty$.

[D2] The limit $\hat{X}$ has identical rows a.s.

[D3] The rows of $\hat{X}$ are Dirichlet distributed.

Unfortunately, conditions **[R1]**–**[R3]** are quite restrictive. Surprisingly, it turns out assertions **[D1]**–**[D3]** remain true under much broader conditions, and that none of the conditions **[R1]**–**[R3]** is necessary.

Denote by $\mathcal{K}_d$ the class of all distributions of a random $d \times d$ stochastic matrix X such that **[D1]**–**[D3]** hold. In this chapter, we extend the result in [CL94] by providing a characterisation theorem (Theorem 6.4) for the class $\mathcal{K}_d$. We find that $\mathcal{K}_d$ is much broader than the class of distributions satisfying **[R1]**–**[R3]**. In addition, once the Dirichlet parameters of the limiting distribution are known, it is elementary to verify if a given distribution belongs to the class $\mathcal{K}_d$.

Chamayou and Letac's result above relies on Theorem 2.26 that admits a rather elegant proof suggested by S. Lauritzen and also presented in [CL94]. The authors

of [CL94] pointed out, using Erdős's remark on the existence of God's book for the best proofs, that S. Lauritzen's proof could be taken from that book. The proof of our Theorem 6.3 which contains Theorem 6.1 as a special case, is also rather simple and concise. Similar to S. Lauritzen's proof, it is based on an insightful observation (Lemma 4.1 above) that extends the well-known result by E. J. G. Pitman on scale independent functions of gamma distributed random variables.

The Dirichlet distribution appearing in the limit [D3] is widely used in statistical applications. These include the modelling of compositional data (see e.g. [A86]), Bayesian analysis, statistical genetics, and nonparametric inference (see [NTT11] for a recent survey on the Dirichlet distribution and its applications).

The main results are presented in Section 6.2, while their proofs are presented in Section 6.3. Examples and applications to random exchange models, random nested simplices, and a service network with polling are given in Section 6.4.

6.2 Main results

Let $A = (\alpha_{i,j}) \in \mathcal{M}_{r,c}^+$. We will now restate for convenience the two main theorems in [CL94] (presented as Theorems 2.25 and 2.26 in Chapter 2, see p. 45).

Theorem 6.1. ([CL94]) *If $(\mathbf{Y}, X) \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{r\bullet})} \otimes D_A$, then $\mathbf{Y}X \sim D_{(\alpha_{\bullet 1}, \dots, \alpha_{\bullet c})}$.*

Theorem 6.2. ([CL94]) *If $r = c = d$, $X \sim D_A$, and*

$$(\alpha_{1\bullet}, \dots, \alpha_{d\bullet}) = (\alpha_{\bullet 1}, \dots, \alpha_{\bullet d}), \quad (6.1)$$

then $\mathcal{L}(X) \in \mathcal{K}_d$, and $\widehat{X}^{(1)} \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$, where $\widehat{X}$ is the a.s. limit in [D1]. Furthermore, if $\mathbf{Y}$ is a random vector taking values in $\mathcal{S}_d$ that is independent of X , then $\mathbf{Y}X \stackrel{d}{=} \mathbf{Y}$ iff $\mathbf{Y} \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$.

Theorem 6.2 is actually a simple consequence of Theorem 6.1. Likewise, our extension of Theorem 6.2 (Theorem 6.4 below) is based on the following result that, as we show at the end of this section, implies Theorem 6.1.

Theorem 6.3. *Let $\mathbf{t} = (t_1, \dots, t_r) \in \mathbb{R}_+^r$ and $\mathbf{s} = (s_1, \dots, s_c) \in \mathbb{R}_+^c$ with $t_\bullet = s_\bullet$. Suppose X is a random element of $\mathcal{M}_{r,c}$ independent of both $\mathbf{Y} \sim D_{\mathbf{t}}$ and $\mathbf{V} \sim G_{\mathbf{t}}$.*

Then

$$\mathbf{Y}X \sim D_s \quad \text{iff} \quad \mathbf{V}X \sim G_s. \quad (6.2)$$

To state the next theorem, we will need the following conditions on the random matrix $X \in \mathcal{P}_d$.

[L1] There exists a vector $\mathbf{t} \in \mathbb{R}_+^d$ such that, for a random vector $\mathbf{V} \sim G_{\mathbf{t}}$ independent of X , one has $\mathbf{V}X \stackrel{d}{=} \mathbf{V}$.

[L2] For an i.i.d. sequence $\{X(n)\}_{n \geq 1}$ with $X(1) \stackrel{d}{=} X$, there exists an $m < \infty$ such that $\mathbb{P}(X(m, 1) \in \mathcal{P}_d^+) > 0$.

Remark 6.1. Note that, unlike the conditions of Theorem 6.2 (under which [L1] holds), condition [L1] does not mean that X must be positive. In Section 6.4, we provide examples where [L1] is met, but most of the entries of X are zeros.

Theorem 6.4. (i) *Relation*

$$\mathcal{L}(X) \in \mathcal{K}_d \quad (6.3)$$

holds iff [L1] and [L2] are met for X .

(ii) *If (6.3) holds, then $\widehat{X}^{(1)} \sim D_{\mathbf{t}}$, where the vector $\mathbf{t}$ is the same as in [L1], and if $\mathbf{Y}$ is a random element of $\mathcal{S}_d$ independent of X , then $\mathbf{Y}X \stackrel{d}{=} \mathbf{Y}$ iff $\mathbf{Y} \stackrel{d}{=} \widehat{X}^{(1)}$.*

We conclude this section by showing that the assertions of Theorems 6.1 and 6.2 do follow from those of our Theorems 6.3 and 6.4. Let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d) \sim G_{\mathbf{t}}$ for some $\mathbf{t} \in \mathbb{R}_+^d$. Then it is well-known that (see for instance formula 2.1.2 in [NTT11])

$$\left(\frac{\xi_1}{\xi_{\bullet}}, \dots, \frac{\xi_d}{\xi_{\bullet}} \right) \sim D_{\mathbf{t}}. \quad (6.4)$$

We will now use (6.4) and relation (4.4) to show that Theorem 6.1 follows from Theorem 6.3. Let $A = (\alpha_{i,j}) \in \mathcal{M}_{r,c}^+$ and $(\mathbf{V}, Z) \sim G_{(\alpha_{1\bullet}, \dots, \alpha_{r\bullet})} \otimes G_A$. Then we have from (6.4) that

$$X := \begin{pmatrix} \frac{Z_{1,1}}{Z_{1\bullet}} & \dots & \frac{Z_{1,c}}{Z_{1\bullet}} \\ \vdots & \ddots & \vdots \\ \frac{Z_{r,1}}{Z_{r\bullet}} & \dots & \frac{Z_{r,c}}{Z_{r\bullet}} \end{pmatrix} \sim D_A, \quad (6.5)$$

and

$$\mathbf{Y} := \left(\frac{V_1}{V_\bullet}, \dots, \frac{V_r}{V_\bullet} \right) \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{r\bullet})}. \quad (6.6)$$

Therefore

$$(\mathbf{V}, X) \sim G_{(\alpha_{1\bullet}, \dots, \alpha_{r\bullet})} \otimes D_A, \quad (6.7)$$

and

$$(\mathbf{Y}, X) \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{r\bullet})} \otimes D_A. \quad (6.8)$$

Now

$$\mathbf{V}X = \sum_{k=1}^r \left(\frac{Z_{k,1}}{Z_{k\bullet}}, \dots, \frac{Z_{k,c}}{Z_{k\bullet}} \right) V_k, \quad (6.9)$$

and relations (4.4) and (6.7) imply that the random vector on the right hand side of (6.9) is distributed as

$$\sum_{k=1}^r (Z_{k,1}, \dots, Z_{k,c}) = (Z_{\bullet 1}, \dots, Z_{\bullet c}) \sim G_{(\alpha_{\bullet 1}, \dots, \alpha_{\bullet c})}. \quad (6.10)$$

It follows from Theorem 6.3 that for $\mathbf{Y}$ satisfying (6.8), one has $\mathbf{Y}X \sim D_{(\alpha_{\bullet 1}, \dots, \alpha_{\bullet c})}$, thus establishing the claim of Theorem 6.1.

To see that Theorem 6.2 follows from Theorems 6.3 and 6.4, we suppose that $\{X(n)\}_{n \geq 1}$ are i.i.d. with law D_A , and A satisfies (6.1). By Theorem 6.1, if $\mathbf{Y} \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$ is independent of $X(1)$, then $\mathbf{Y}X(1) \sim D_{(\alpha_{\bullet 1}, \dots, \alpha_{\bullet d})}$, and Theorem 6.3 implies that $\mathbf{V}X(1) \stackrel{d}{=} \mathbf{V}$ for $\mathbf{V} \sim G_{(\alpha_{\bullet 1}, \dots, \alpha_{\bullet d})}$ independent of $X(1)$. Thus [L1] holds with $\mathbf{t} = (\alpha_{\bullet 1}, \dots, \alpha_{\bullet d})$. Since $\mathbb{P}(X(1) \in \mathcal{P}_d^+) = 1$, [L2] is satisfied as well, and the assertion of Theorem 6.2 follows by applying Theorem 6.4.

6.3 Proofs

We will begin with the following obvious consequence of Lemma 4.1.

Corollary 6.5. *Let $\mathbf{V} = (V_1, \dots, V_r) \sim G_{\mathbf{t}}$ and $X = (X_{i,j}) \in \mathcal{M}_{r,c}$ be random elements independent of each other, X having positive row sums a.s. Define the*

function $H(v_1, \dots, v_r, X) = (H_1(v_1, \dots, v_r, X), \dots, H_c(v_1, \dots, v_r, X)) \in \mathbb{R}_+^c$ by

$$H_j(v_1, \dots, v_r, X) := \sum_{i=1}^r \frac{v_i X_{i,j}}{v_\bullet}, \quad 1 \leq j \leq c. \quad (6.11)$$

Then $H(\mathbf{V}, X)$ is independent of $V_\bullet$.

Proof of Theorem 6.3. We will proceed in the same way as in the proof of Proposition 4.2. Suppose the right relation in (6.2) holds, i.e., $\mathbf{V}X \sim G_s$ for $\mathbf{V} \sim G_t$ independent of X . Then, by Corollary 6.5, for the scale independent function H defined in (6.11), the random vector $H(V_1, \dots, V_r, X) \equiv \left(\frac{V_1}{V_\bullet}, \dots, \frac{V_r}{V_\bullet}\right) X$ is independent of $V_\bullet$, and therefore

$$\mathbf{V}X = \left(\frac{V_1}{V_\bullet}, \dots, \frac{V_r}{V_\bullet}\right) X V_\bullet \stackrel{d}{=} \left(\frac{V_1}{V_\bullet}, \dots, \frac{V_r}{V_\bullet}\right) X \tilde{V}_\bullet, \quad (6.12)$$

where $(\tilde{V}_1, \dots, \tilde{V}_r) \sim G_t$ is independent of $(\mathbf{V}, X)$. Since $\mathbf{V}X \sim G_s$, for $\mathbf{Z} := (Z_1, \dots, Z_c) \sim G_s$, one has

$$\mathbf{V}X \stackrel{d}{=} \mathbf{Z} \stackrel{d}{=} \left(\frac{Z_1}{Z_\bullet}, \dots, \frac{Z_c}{Z_\bullet}\right) \tilde{Z}_\bullet, \quad (6.13)$$

$(\tilde{Z}_1, \dots, \tilde{Z}_c)$ being an independent copy of $\mathbf{Z}$ (cf. (4.4)).

Equating the logarithms of the components of the vectors on the right hand sides of (6.12) and (6.13), we obtain that

$$\begin{aligned} & \left(\ln \left(\frac{\sum_{i=1}^r V_i X_{i,1}}{V_\bullet} \right), \dots, \ln \left(\frac{\sum_{i=1}^r V_i X_{i,c}}{V_\bullet} \right) \right) + \ln(\tilde{V}_\bullet) \mathbf{e}_c \\ & \stackrel{d}{=} \left(\ln \left(\frac{Z_1}{Z_\bullet} \right), \dots, \ln \left(\frac{Z_c}{Z_\bullet} \right) \right) + \ln(\tilde{Z}_\bullet) \mathbf{e}_c, \end{aligned} \quad (6.14)$$

where $\mathbf{e}_c := (1, \dots, 1) \in \mathbb{R}^c$.

Since $t_\bullet = s_\bullet$, one has $\tilde{V}_\bullet \stackrel{d}{=} \tilde{Z}_\bullet$ as both follow $\Gamma_{t_\bullet} = \Gamma_{s_\bullet}$, and so letting ψ , φ and χ denote the characteristic functions of the first, second (and fourth), and third terms in (6.14), respectively, we have

$$\psi(u_1, \dots, u_c) \varphi(u_1, \dots, u_c) = \chi(u_1, \dots, u_c) \varphi(u_1, \dots, u_c).$$

Noting that

$$\begin{aligned}\varphi(u_1, \dots, u_c) &= \mathbb{E} e^{iu_\bullet \ln(V_\bullet)} = \frac{1}{\Gamma(t_\bullet)} \int_0^\infty e^{iu_\bullet \ln x} x^{t_\bullet-1} e^{-x} dx \\ &= \frac{1}{\Gamma(t_\bullet)} \int_0^\infty x^{iu_\bullet+t_\bullet-1} e^{-x} dx = \frac{\Gamma(t_\bullet + iu)}{\Gamma(t_\bullet)} \neq 0\end{aligned}$$

for $t_\bullet > 0$, we conclude that $\psi \equiv \chi$ and therefore the respective random vectors have a common distribution. Hence one has

$$\left(\frac{V_1}{V_\bullet}, \dots, \frac{V_r}{V_\bullet} \right) X = \left(\frac{\sum_{i=1}^r V_i X_{i,1}}{V_\bullet}, \dots, \frac{\sum_{i=1}^r V_i X_{i,c}}{V_\bullet} \right) \stackrel{d}{=} \left(\frac{Z_1}{Z_\bullet}, \dots, \frac{Z_c}{Z_\bullet} \right) \sim D_s \quad (6.15)$$

from (6.4). Observing that the left hand side of (6.15) has the form $\mathbf{Y}X$ for $\mathbf{Y} \sim D_t$ independent of X (cf. (6.4)), we have that $\mathbf{Y}X \sim D_s$, and so the left relation in (6.2) holds.

Conversely, suppose that $\mathbf{Y}X \sim D_s$ for $\mathbf{Y} \sim D_t$ independent of X . Using the same steps as above, but following them in the reverse order, one can easily conclude that $\mathbf{V}X \sim G_s$ for $\mathbf{V} \sim G_t$ independent of X . Theorem 6.3 is proved. $\square$

To prove Theorem 6.4, we need the following simple extension of Proposition 2.2 from [CL94].

Proposition 6.6. *Let $X(1)$ be a random element of $\mathcal{P}_d$ satisfying [L2], and $\{X(n)\}_{n \geq 1}$ be an i.i.d. sequence. Then there exists a random element $\mathbf{W}$ of $\mathcal{S}_d$ such that*

$$X(n, 1) \xrightarrow{a.s.} \mathbf{e}_d^T \mathbf{W} \quad (6.16)$$

as $n \rightarrow \infty$, where $\mathbf{e}_d^T$ denotes the transpose of $\mathbf{e}_d$. Furthermore, if $\mathbf{Y}$ is a random element of $\mathcal{S}_d$ that is independent of $X(1)$, then $\mathbf{Y}X(1) \stackrel{d}{=} \mathbf{Y}$ iff $\mathbf{Y} \stackrel{d}{=} \mathbf{W}$.

Proof. By Proposition 2.2 in [CL94], there exists a random element $\mathbf{W}$ of $\mathcal{S}_d$ such that

$$X(nm, 1) \xrightarrow{a.s.} \mathbf{e}_d^T \mathbf{W} \quad (6.17)$$

as $n \rightarrow \infty$, where $m < \infty$ is from [L2]. For $k \in (nm, (n+1)m) \cap \mathbb{N}$, one has

$$\begin{aligned} X(k, 1) &= X(k, nm+1)X(nm, 1) \\ &= X(k, nm+1)\mathbf{e}_d^T \mathbf{W} + X(k, nm+1)(X(nm, 1) - \mathbf{e}_d^T \mathbf{W}) \\ &= \mathbf{e}_d^T \mathbf{W} + o(1) \quad a.s. \end{aligned}$$

as $n \rightarrow \infty$, since $X\mathbf{e}_d^T = \mathbf{e}_d^T$ for any $X \in \mathcal{P}_d$. Hence, the limit in (6.16) exists and coincides with that in (6.17). The second part follows immediately from the proof of Proposition 2.2 in [CL94]. $\square$

Proof of Theorem 6.4. (i) Suppose [L1] and [L2] hold. Then, by Proposition 6.6, there exists a random element $\mathbf{W}$ of $\mathcal{S}_d$ such that

$$X(n, 1) \xrightarrow{a.s.} \widehat{X} \equiv \mathbf{e}_d^T \mathbf{W}$$

as $n \rightarrow \infty$, and therefore [D1] and [D2] hold. From condition [L1] and Theorem 6.3, it follows that $\mathbf{Y}X(1) \stackrel{d}{=} \mathbf{Y}$ for $\mathbf{Y} \sim D_{\mathbf{t}}$ independent of $X(1)$. By Proposition 6.6, the law of $\mathbf{Y}$ coincides with the distribution of the limit $\widehat{X}^{(i)}, i = 1, 2, \dots, d$, i.e.

$$\widehat{X}^{(1)} = \widehat{X}^{(2)} = \dots = \widehat{X}^{(d)} \stackrel{d}{=} \mathbf{Y} \sim D_{\mathbf{t}}, \quad (6.18)$$

and so [D3] holds. We have proved that $\mathcal{L}(X) \in \mathcal{K}_d$.

Conversely, let $\mathcal{L}(X) \in \mathcal{K}_d$. Observe that if $X(m, 1) \in \mathcal{P}_d^+$ for some $m \in \mathbb{N}$, then $X(m+n, 1) \in \mathcal{P}_d^+$ for all $n \in \mathbb{N}$. Then since $X(m, 1)$ converges a.s. to $\widehat{X} \in \mathcal{P}_d^+$, there is an $m < \infty$ such that $X(m, 1) \in \mathcal{P}_d^+$ a.s., and so [L2] holds. Furthermore, clearly

$$\widehat{X} = \lim_{n \rightarrow \infty} X(n, 1) = \lim_{n \rightarrow \infty} X(n, 2)X(1) = \widehat{X}'X(1),$$

where $\widehat{X}' \stackrel{d}{=} \widehat{X}$ is independent of $X(1)$. In particular, one has $\widehat{X}^{(1)}X(1) \stackrel{d}{=} \widehat{X}^{(1)} \sim D_{\mathbf{t}}$ for some $\mathbf{t} \in \mathbb{R}_+^d$. Applying Theorem 6.3, we have that $\mathbf{V}X(1) \stackrel{d}{=} \mathbf{V}$ for $\mathbf{V} \sim G_{\mathbf{t}}$, and so [L1] holds.

(ii) If (6.3) holds, then by (i) we have that [L1] and [L2] are satisfied, in which case we have (6.18), where $\mathbf{t}$ is the same as in [L1]. Furthermore, for $\mathbf{Y}$ a random element of $\mathcal{S}_d$ that is independent of X , one has by Proposition 6.6 that $\mathbf{Y}X \stackrel{d}{=} \mathbf{Y}$

iff $\mathbf{Y} \stackrel{d}{=} \widehat{X}^{(1)}$, where $\widehat{X}^{(1)} \sim D_t$ by (6.18). □

6.4 Examples and applications

6.4.1 Random exchange models

Consider the following random exchange model, which is a discrete time analogue of certain continuous time Markov processes, called Potlatch models (see for instance [HL81] and [LS81]). Suppose we have $d < \infty$ bins labelled by numbers $1, 2, \dots, d$, that hold amounts $q_k(n)$, $k = 1, \dots, d$, of a homogeneous commodity at times $n = 0, 1, 2, \dots$, respectively. The dynamics of the model is as follows: at time $n \geq 1$, the vector $\mathbf{q}(n-1) := (q_1(n-1), \dots, q_d(n-1))$ changes to $\mathbf{q}(n) := \mathbf{q}(n-1)X(n)$, where $\{X(n)\}_{n \geq 1}$ are i.i.d. random elements of $\mathcal{P}_d$ with distribution $\mathcal{L}(X)$, which generates the Markov chain

$$\mathbf{q}(n) = \mathbf{q}(0)X(1, n), \quad n \geq 1. \quad (6.19)$$

Clearly $\sum_{k=1}^d \mathbf{q}_k(n) = \sum_{k=1}^d \mathbf{q}_k(0)$ for all n , and without loss of generality, we assume that this quantity is equal to one.

As noted in [CL94], it is easy to prove that Markov chain (6.19) has a stationary distribution when $\mathcal{L}(X) \in \mathcal{K}_d$. Indeed, defining the random maps $K_n : \mathcal{S}_d \rightarrow \mathcal{S}_d$ by $K_n(\mathbf{y}) := \mathbf{y}X(n)$, one has that, for any $\mathbf{y} \in \mathcal{S}_d$,

$$K_1 \circ \dots \circ K_n(\mathbf{y}) = \mathbf{y}X(n, 1) \xrightarrow{a.s.} \widehat{X}^{(1)} \text{ as } n \rightarrow \infty.$$

Applying Proposition 1 in [CL91], we obtain that $\mathcal{L}(\widehat{X}^{(1)})$ is the stationary distribution of Markov chain (6.19).

The random exchange model (6.19) is a higher dimension analogue of the stochastic give-and-take model studied in [DR63] and introduced in its deterministic form in the context of human genetics in [L61] (c.f. Section 2.3.1 above). This model is applicable to the study of dynamical systems (see for instance [BJM10]), e.g. it provides a model of a closed economy, where agents exchange real wealth at each step in a random manner.

Now we consider three special cases of random stochastic matrices X with $\mathcal{L}(X) \in \mathcal{K}_d$. The purpose of these examples is to demonstrate that none of the conditions [D1]–[D3] from [CL94] is necessary for $\mathcal{L}(X) \in \mathcal{K}_d$. We will also discuss the respective special cases of random exchange model (6.19). Our first example leads to a generalisation of the result in [CL94] to the case of “extended” Dirichlet distributions, where the parameters of the distribution are permitted to be zero. Our second example demonstrates that the rows of X need not be independent.

Example 6.1. For a vector $\mathbf{a} \in \mathcal{M}_{1,d}$, we set $D_{\mathbf{a}}$ to be the weak limit of the distribution $D_{\mathbf{a}+\epsilon}$ as $\epsilon \downarrow 0$, with the usual interpretation of the sum $\mathbf{a} + \epsilon \mathbf{e}_d$ with $\mathbf{a} \in \mathbb{R}^d$ and $\epsilon \in \mathbb{R}$. In other words, the components of $\mathbf{Y} \sim D_{\mathbf{a}}$ that correspond to zero components of $\mathbf{a}$ are identically zero, whereas the subvector of $\mathbf{Y}$ consisting of the components Y_j of that random vector that correspond to $a_j > 0$ is a usual Dirichlet distributed vector. Likewise, for a matrix $A \in \mathcal{M}_{r,c}$, we set D_A to be the weak limit of the distribution of $D_{A+\epsilon}$ as $\epsilon \downarrow 0$. We define the distributions $G_{\mathbf{a}}$ and G_A for $\mathbf{a} \in \mathcal{M}_{1,d}$ and $A \in \mathcal{M}_{r,c}$ in a similar way.

Let $A = (\alpha_{i,j}) \in \mathcal{M}_d$ with $\alpha_{i\bullet} = \alpha_{\bullet i} > 0$ for $i = 1, \dots, d$. If $X \sim D_A$ and X satisfies [L2], then it turns out that $\mathcal{L}(X) \in \mathcal{K}_d$ and $\widehat{X}^{(1)} \sim D_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$, and so $D_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$ is the stationary distribution of Markov chain (6.19).

Indeed, by Theorem 6.4 it suffices to show that $\mathbf{V}X \stackrel{d}{=} \mathbf{V}$ for $\mathbf{V} \sim G_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$ independent of X . To do that, let $(\mathbf{V}, Z) \sim G_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})} \otimes G_A$, and define $X \in \mathcal{P}_d$ by (6.5). Then $\mathbf{V}$ is independent of X , and $\mathbf{V}X \sim G_{(\alpha_{1\bullet}, \dots, \alpha_{d\bullet})}$ by (6.9) and (6.10), as required.

In particular, we have obtained the stationary distribution for the following simple model, which, to the best of the author’s knowledge, has not previously been studied: at time $n \geq 1$, an independent uniform proportion of the commodity previously held in bin k , $k = 1, 2, \dots, d$, is shifted to the (neighbouring) bin $k+1 \pmod{d}$.

d). In this case vector $\mathbf{q}(n)$ is defined by (6.19) with

$$X := \begin{pmatrix} U_1 & 1 - U_1 & 0 & \cdots & 0 \\ 0 & U_2 & 1 - U_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_{d-1} & 1 - U_{d-1} \\ 1 - U_d & 0 & \cdots & 0 & U_d \end{pmatrix}, \quad (6.20)$$

where the U_k , $k = 1, \dots, d$, are i.i.d. $U[0, 1]$ -distributed random variables. Observing that X defined by (6.20) satisfies **[L2]** for $m = d - 1$, we have from above that $\hat{X}^{(1)} \sim D_{(2, \dots, 2)}$, and so $D_{(2, \dots, 2)}$ is the stationary distribution of Markov chain (6.19) with $X(n)$ distributed as (6.20).

Example 6.2. In this example, we consider a random stochastic matrix X with all $d \geq 2$ rows dependent, which generates Markov chain (6.19) with the same stationary distribution as that corresponding to the random stochastic matrix (6.20). The behaviour of this model is controlled by the decisions of a “leader” as follows. At time $n \geq 1$, the “leader” shifts a uniform proportion of the commodity held in bin 1 to bin 2. If the proportion shifted is greater than $1/2$, then no other shifts occur in the system at time n . However, if the proportion shifted is less than or equal to $1/2$, then all the commodity previously held in bin k , $k = 2, 3, \dots, d$, is shifted to the (neighbouring) bin $k + 1 \pmod{d}$. In this model, the random vector $\mathbf{q}(n)$ is given by (6.19), with the random element $X \in \mathcal{P}_d$ defined by

$$X := \begin{pmatrix} U & 1 - U & 0 & \cdots & 0 \\ 0 & I & 1 - I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I & 1 - I \\ 1 - I & 0 & \cdots & 0 & I \end{pmatrix}, \quad (6.21)$$

where $U \sim U[0, 1]$, and $I := \mathbf{1}_{\{U > 1/2\}}$.

We will show that $\mathcal{L}(X) \in \mathcal{K}_d$ and $\hat{X}^{(1)} \sim D_{(2, \dots, 2)}$. By Theorem 6.4, it suffices to show that **[L1]** and **[L2]** hold for X , $\mathbf{t} = (2, \dots, 2)$ being the vector from **[L1]**. It is

not hard to directly verify that X defined by (6.21) satisfies **[L2]** for $m = 2d - 2$. Now let $\mathbf{V} \sim G_{(2, \dots, 2)}$ be independent of X , and denote by $\varphi(u_1, \dots, u_d)$ the characteristic function of $\mathbf{V}X$. Then, setting $u_{d+1} := u_1$, we have

$$\begin{aligned}
\varphi(u_1, \dots, u_d) &= \mathbb{E} \exp \left\{ i(u_1 U + u_2(1 - U))V_1 + i \sum_{j=2}^d (u_j I + u_{j+1}(1 - I))V_j \right\} \\
&= \mathbb{E} \exp \left\{ i \sum_{j=2}^d u_{j+1} V_j \right\} \int_0^{1/2} \frac{ds}{[1 - i(u_1 s + u_2(1 - s))]^2} \\
&\quad + \mathbb{E} \exp \left\{ i \sum_{j=2}^d u_j V_j \right\} \int_{1/2}^1 \frac{ds}{[1 - i(u_1 s + u_2(1 - s))]^2} \\
&= \left(\frac{(1 - iu_2)^2}{\prod_{k=1}^d (1 - iu_k)^2} \right) \left(\frac{i}{(1 - iu_2)(2i + u_1 + u_2)} \right) \\
&\quad + \left(\frac{(1 - iu_1)^2}{\prod_{k=1}^d (1 - iu_k)^2} \right) \left(\frac{i}{(1 - iu_1)(2i + u_1 + u_2)} \right) \\
&= \prod_{k=1}^d (1 - iu_k)^{-2},
\end{aligned}$$

which is the characteristic function of $\mathbf{V}$ as well, and so we conclude that $\mathbf{V}X \stackrel{d}{=} \mathbf{V}$. Therefore **[L1]** holds with $\mathbf{t} = (2, \dots, 2)$, as required.

Example 6.3. In our final example, we will consider the following modification of the model from our last example. At time $n \geq 1$, a uniform proportion of the commodity held in bin 1 is shifted to bin 2 as in the previous model, but, independently of that proportion, with probability $1/2$ all the commodity previously held in bin k , $k = 2, 3, \dots, d$, is shifted to the (neighbouring) bin $k + 1 \pmod{d}$. The only difference in this model is that the shift from bin 1 to bin 2 is independent of the shifts from bin k to bin $k + 1 \pmod{d}$. In this model, the random vector $\mathbf{q}(n)$ is given by (6.19), with $X \in \mathcal{P}_d$ given by (6.21) for $(U, I) \sim U[0, 1] \otimes Be(1/2)$.

We claim that $\mathcal{L}(X) \in \mathcal{K}_d$ and $\widehat{X}^{(1)} \sim D_{(1/2, \dots, 1/2)}$. By Theorem 6.4 it suffices to show that **[L1]** and **[L2]** hold for X , with $\mathbf{t} = (1/2, \dots, 1/2)$ being the vector from **[L1]**. As in Example 1, clearly **[L2]** holds for $m = d - 1$. Now let

$$\mathbf{V} = (V_1, \dots, V_d) \sim G_{(1/2, \dots, 1/2)}$$

be independent of X , and denote by $\phi(u_1, \dots, u_d)$ the characteristic function of $\mathbf{V}X$. Then, setting $u_{d+1} := u_1$, we have that

$$\begin{aligned}
& \phi(u_1, \dots, u_d) \\
&= \mathbb{E} \exp \left\{ i(u_1 U + u_2(1 - U))V_1 + i \sum_{j=2}^d (u_j I + u_{j+1}(1 - I))V_j \right\} \\
&= \mathbb{E} \exp \left\{ i \sum_{j=2}^d (u_j I + u_{j+1}(1 - I))V_j \right\} \int_0^1 \mathbb{E} \exp \{ [i(u_1 t + u_2(1 - t))V_1] \} dt \\
&= \frac{1}{2} \left(\mathbb{E} \exp \left\{ i \sum_{j=2}^d u_{j+1} V_j \right\} + \mathbb{E} \exp \left\{ \sum_{j=2}^d u_j V_j \right\} \right) \int_0^1 \frac{dt}{\sqrt{1 - i(u_1 t + u_2(1 - t))}} \\
&= \left(\frac{\sqrt{1 - iu_1} + \sqrt{1 - iu_2}}{2 \prod_{k=1}^d \sqrt{1 - iu_k}} \right) \left(\frac{2\sqrt{1 - iu_1} - 2\sqrt{1 - iu_2}}{i(u_2 - u_1)} \right) \\
&= \prod_{k=1}^d \frac{1}{\sqrt{1 - iu_k}}
\end{aligned}$$

is the characteristic function of $\mathbf{V}$, and we conclude that $\mathbf{V}X \stackrel{d}{=} \mathbf{V}$. Therefore [L1] holds with $\mathbf{t} = (1/2, \dots, 1/2)$ as required.

6.4.2 Random nested simplices

The study of random triangles and, in particular, nested sequences of random triangles, has been extensive in the probabilistic literature (see for instance [LS98] and references therein). In one such study [LS98], the authors use Theorem 1.2 in [CL94] (see Theorem 6.2 above) to prove that the limiting distribution of certain random nested simplices is Dirichlet (cf. Section 2.5.3). We will briefly recall the setup from [LS98]. Let $\mathbf{p}_1, \dots, \mathbf{p}_d$ be d points in a real $(d - 1)$ -dimensional affine space E , and take their convex hull to obtain a simplex S . Now suppose that $X(1) \sim D_A$, $A = (\alpha_{i,j}) \in \mathcal{M}_d^+$ satisfying

$$(\alpha_{1\bullet}, \dots, \alpha_{d\bullet}) = (\alpha_{\bullet 1}, \dots, \alpha_{\bullet d}),$$

and one chooses d points $\mathbf{p}_1(1), \dots, \mathbf{p}_d(1)$ in S with barycentric coordinates given by $X^{(1)}(1), \dots, X^{(d)}(1)$ with respect to $(\mathbf{p}_1, \dots, \mathbf{p}_d)$, taking their convex hull to obtain a

new simplex $S(1) \subset S$. Continuing this processes independently on $S(1)$, and so on, the so defined simplices shrink a.s. to a random point $\mathbf{Y} \in E$ as $n \rightarrow \infty$ (cf. (2.29) on p. 47). We extend this result as follows.

Theorem 6.7. *If $X(1)$ satisfies [L1] and [L2], then (2.29) holds true. Furthermore, the barycentric coordinates of $\mathbf{Y}$ with respect to $(\mathbf{p}_1, \dots, \mathbf{p}_d)$ have distribution $D_{\mathbf{t}}, \mathbf{t}$ being the vector from [L2].*

Proof. The proof of Theorem 6.7 basically repeats that of Theorem 2.2 from [LS98], except we use Theorem 6.4 in place of Theorem 6.2 above. For $\mathbf{B}(0) \in \mathcal{S}_d$, define $\mathbf{Z}(n) \in S(n)$ as the point with barycentric coordinates $\mathbf{B}(0)$ with respect to the affine frame $(\mathbf{p}_1(n), \dots, \mathbf{p}_d(n))$, and denote by $\mathbf{B}(n)$ the barycentric coordinates of $\mathbf{Z}(n)$ with respect to the affine frame $(\mathbf{p}_1, \dots, \mathbf{p}_d)$. Then, as shown in the proof of Theorem 2.2 from [LS98], one has

$$\mathbf{B}(n) = \mathbf{B}(0)X(n, 1), \quad n = 0, 1, \dots$$

Now suppose $X(1)$ satisfies [L1] and [L2]. By Theorem 6.4, $\mathcal{L}(X(1)) \in \mathcal{K}_d$ with $\widehat{X}^{(1)} = (Y_1, \dots, Y_d) \sim D_{\mathbf{t}}, \mathbf{t}$ being the vector from [L2]. Then clearly $\mathbf{B}(n) \xrightarrow{a.s.} \widehat{X}^{(1)}$ as $n \rightarrow \infty$ for any $\mathbf{B}(0)$. In particular, for any $k = 1, \dots, d$, we can set

$$\mathbf{B}(0) := (0, \dots, 0, 1, 0, \dots, 0),$$

where 1 is at the k th place, to obtain that $\mathbf{p}_k(n) \xrightarrow{a.s.} \mathbf{Y}$ as $n \rightarrow \infty$, where $\mathbf{Y}$ has barycentric coordinates $(Y_1, \dots, Y_d)$ with respect to $(\mathbf{p}_1, \dots, \mathbf{p}_d)$, as required. $\square$

6.4.3 Service networks with polling

Recall the random donkey walk in $\mathcal{P}_d$ that was introduced in [L02] (cf. Section 2.5.4). This model actually describes the progression of a customer in the following discrete-time closed network with polling. Suppose there are d buffer nodes in the network, accessed in cyclic order by a single server, with customers never leaving the network. At time $n = 1$, the server accesses node 1 and moves customers from that node to other nodes in the network randomly, with transition probabilities taken from

the first row of a random $X(1) \in \mathcal{P}_d$. At time $n \leq d$, the server accesses node n , and similarly moves customers from that node to other nodes with transition probabilities taken from the n th row of $X(1)$. The same procedure is repeated in a cyclical fashion, with the server accessing node 1 at time $n = d + 1$, and so on, with transition probabilities taken from the rows of independent random elements $X(j) \stackrel{d}{=} X(1)$ for the j th cycle. In this formulation, the vector $\mathbf{B}(n)$ represents the distribution of customers in the network after n steps, with initial distribution $\mathbf{B}(0)$.

Markov chains with the cyclic property that each of the sequences $\{\mathbf{B}(td+r)\}_{t \geq 0}$, $r = 1, \dots, d$, forms a homogeneous Markov chain, were also considered in [D98], where several examples of such Markov chains were provided.

For a matrix $X = (X_{i,j}) \in \mathcal{M}_d$ and $r = 1, 2, \dots, d$, set

$$(T_r(X))_{i,j} := (\delta_{ij} + X_{r,j}\delta_{ri} - \delta_{rj}\delta_{ij}), \quad 1 \leq i, j \leq d,$$

δ_{ij} being Kronecker's delta (i.e., $T_r(X)$ is the $d \times d$ identity matrix with the r th row replaced by the r th row of X).

An extension of Theorem 2.27 is given by the following theorem.

Theorem 6.8. *Let X be a random element of $\mathcal{P}_d$ such that $T_1(X) \cdots T_d(X)$ satisfies condition [L2]. Let $V = (V_{i,j}) \sim G_B$ be a random matrix independent of X , where $B \in \mathcal{M}_d^+$ has entries $\beta_{i,j}$ given by (2.31). Furthermore, suppose that*

$$V^{(r)} \stackrel{d}{=} V^{(r-1)}T_r(X), \quad 1 \leq r \leq d, \quad (6.22)$$

where $V^{(0)} := V^{(d)}$.

Then, for $r = 1, \dots, d$, the stationary distribution of the r th homogeneous Markov chain $\{\mathbf{B}(td+r)\}_{t \geq 0}$ is Dirichlet with parameters $(\beta_{r,1}, \dots, \beta_{r,d})$.

Proof. Let $\{X(t)\}_{t \geq 0}$ be i.i.d. with law $\mathcal{L}(X)$, and fix $r \in \{1, \dots, d\}$. Then

$$R(t) := \begin{cases} T_1(X(t)) \cdots T_d(X(t)) & \text{for } r = d, \\ T_{r+1}(X(t)) \cdots T_d(X(t))T_1(X(t)) \cdots T_r(X(t)) & \text{for } r < d, \end{cases}$$

$t = 0, 1, \dots$, are i.i.d. random elements in $\mathcal{P}_d$, and

$$\mathbf{B}(td + r) = \mathbf{B}(r)R(1, t),$$

where $R(1, t)$ is given by (2.24). Clearly, $R(1)$ satisfies **[L2]** since $T_1(X) \cdots T_d(X)$ satisfies **[L2]**, and so by Proposition 6.6, there exists a random vector $\mathbf{W} \in \mathcal{S}_d$ such that

$$R(t, 1) \xrightarrow{a.s.} \mathbf{e}_d^T \mathbf{W}$$

as $t \rightarrow \infty$. Using Proposition 1 in [CL91] applied to the space $\mathcal{S}_d$ and to the random mappings from $\mathcal{S}_d$ to $\mathcal{S}_d$ defined by $\mathbf{x} \mapsto \mathbf{x}Z(t)$, $t \geq 0$, it follows that $\{\mathbf{B}(td + r)\}_{t \geq 0}$ has a unique stationary distribution.

Now $V^{(r)} \stackrel{d}{=} V^{(r)}R(1)$ by (6.22), and therefore Theorem 6.4 implies that

$$\widehat{R}^{(1)} \sim D_{(\beta_{r,1}, \dots, \beta_{r,d})}.$$

Hence the stationary distribution of $\{\mathbf{B}(td + r)\}_{t \geq 0}$ is $D_{(\beta_{r,1}, \dots, \beta_{r,d})}$, as required. $\square$

The following remark shows that our Theorem 6.8 is indeed an extension of Theorem 2.27.

Remark 6.1. Let $A = (\alpha_{i,j}) \in \mathcal{M}_d$ be such that $\alpha_{i\bullet} = \alpha_{\bullet i} > 0$ for $i = 1, \dots, d$, and $B = (\beta_{i,j}) \in \mathcal{M}_d^+$, where the numbers $\beta_{i,j}$ are given by (2.31). Let $(V, Z) \sim G_B \otimes G_A$, and $X \sim D_A$ defined by (6.5) be such that $T_1(X^{(1)}) \cdots T_d(X^{(d)})$ satisfies **[L2]**. Then, for $r = 1, \dots, d$ and using the convention that $r - 1 = d$ for $r = 1$, one has

$$\begin{aligned} V^{(r-1)}T_r(X) &= (V_{r-1,1}, \dots, V_{r-1,r-1}, 0, V_{r-1,r+1}, \dots, V_{r-1,d}) + V_{r-1,r} \left(\frac{Z_{r,1}}{Z_{r\bullet}}, \dots, \frac{Z_{r,d}}{Z_{r\bullet}} \right) \\ &\stackrel{d}{=} (V_{r-1,1}, \dots, V_{r-1,r-1}, 0, V_{r-1,r+1}, \dots, V_{r-1,d}) + (Z_{r,1}, \dots, Z_{r,d}), \end{aligned}$$

where the equality in distribution holds by (4.4) and the fact that $V_{r-1,r} \stackrel{d}{=} Z_{r\bullet}$, as both follow $\Gamma_{\beta_{r-1,r}} = \Gamma_{\alpha_{r\bullet}}$. It remains to observe that $\beta_{r,r} = \alpha_{r,r}$ and, for $j \neq r$, $\beta_{r,j} = \beta_{r-1,j} + \alpha_{r,j}$, hence the vector in the last line above is distributed as $V^{(r)}$.

The conditions of Theorem 6.8 are met, and we conclude that, for all $r = 1, \dots, d$, the homogeneous Markov chain $\{\mathbf{B}(td + r)\}_{t \geq 0}$ has stationary distribution

$D_{(\beta_{r,1},\dots,\beta_{r,d})}$. In particular, for X defined by (6.20), one has $\beta_{r,j} = 1$ for $j \neq r + 1 \pmod{d}$, and $\beta_{r,j} = 2$ for $j = r + 1 \pmod{d}$.

Chapter 7

Boundary crossing probabilities for multivariate Brownian motion

In this chapter we provide a solution to Problem 3 from the Introduction by giving an upper bound for the difference between boundary hitting probabilities for the multivariate Brownian motion process for two different boundaries. Our bound holds under broad assumptions about the nature of the boundary, including the Lipschitz condition (in a Hausdorff-type metric) on its time cross-sections. As an auxiliary result, we also give upper bounds for the first boundary crossing time densities.

7.1 The main results

Let $\mathbf{W} = \{\mathbf{W}_t = (W_t^{(1)}, \dots, W_t^{(m)})\}_{t \geq 0}$ be the standard m -dimensional Brownian motion process, $\mathbf{W}_0 = \mathbf{0}$. For a fixed $T < \infty$, let $\mathcal{G}$ be the class of open sets $G \subset [0, T] \times \mathbb{R}^m$ (the first component representing time) with $(0, \mathbf{0}) \in \partial G$.

As discussed in the Introduction, it is important in a number of applied problems to compute the probability $P(G)$ (cf. (1.39)) for $\mathbf{W}$ to stay within G during the time interval $(0, T)$.

The aim of this chapter is to extend the approximation approach outlined in the Introduction to the multivariate case and provide bounds for approximation errors. The results in this chapter are also of interest for the theory of boundary value problems for parabolic partial differential equations.

First we will introduce some notation which we will need to define a Hausdorff-type metric that proved to be the most natural one for measuring the closeness of the sets G and $\tilde{G}$ in the context of our problem.

For $H \subset [0, T] \times \mathbb{R}^m$, let

$$H_t := \{\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m : (t, \mathbf{x}) \in H\}, \quad t \in [0, T], \quad (7.1)$$

be the time t section of H . For $A \subset \mathbb{R}^m$ and $\mathbf{x} \in \mathbb{R}^m$, let $\rho(\mathbf{x}, A) := \inf_{\mathbf{y} \in A} \|\mathbf{x} - \mathbf{y}\|$ be the distance from $\mathbf{x}$ to the set A , $\|\cdot\|$ being the Euclidian norm in $\mathbb{R}^m$, A^c be the complement of A , and let

$$A^{(v)} := \begin{cases} \{\mathbf{x} \in \mathbb{R}^m : \rho(\mathbf{x}, A) < v\}, & v > 0, \\ \{\mathbf{x} \in \mathbb{R}^m : \rho(\mathbf{x}, A^c) \leq -v\}^c, & v \leq 0, \end{cases}$$

denote the dilation (case $v > 0$) and erosion (case $v < 0$), respectively, of the set A by the set $B_{|v|}(\mathbf{0})$, where $B_r(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^m : \|\mathbf{x} - \mathbf{y}\| < r\}$ stands for the open ball of radius $r > 0$ with centre at $\mathbf{x} \in \mathbb{R}^m$.

For $H \subset [0, T] \times \mathbb{R}^m$, let

$$H^{(v)} := \{(t, \mathbf{x}) : t \in [0, T], \mathbf{x} \in H_t^{(v)}\}, \quad v \in \mathbb{R}.$$

The Hausdorff distance between sets $A, \tilde{A} \subset \mathbb{R}^m$ is defined by

$$\rho_h(A, \tilde{A}) := \inf\{\varepsilon > 0 : A \subset \tilde{A}^{(\varepsilon)} \text{ and } \tilde{A} \subset A^{(\varepsilon)}\}.$$

It will be convenient for us to use the metric

$$\rho_H(A, \tilde{A}) := \max\{\rho_h(A, \tilde{A}), \rho_h(A^c, \tilde{A}^c)\}.$$

For positive numbers K, β, γ , introduce the class $\mathcal{G}_{K, \beta, \gamma} \subset \mathcal{G}$ of sets G satisfying the following conditions on their cross-sections.

[G1] The following Lipschitz condition holds:

$$\rho_H(G_s, G_t) \leq K(t - s), \quad 0 < s < t < T.$$

[G2] For any $t \in (0, T)$ and $\mathbf{g} \in \partial G_t$, there exists a $\mathbf{y} \in \mathbb{R}^m$ such that $B_\beta(\mathbf{y}) \subset G_t^c$ and $\mathbf{g} \in \partial B_\beta(\mathbf{y})$.

[G3] For any $t \in (0, T)$, there exists a $v_0 > 0$ such that

$$\mathbb{E}(1 + \|\mathbf{W}_t\|; 0 < \rho(\mathbf{W}_t, G_t^c) < v) < \gamma v, \quad 0 < v < v_0.$$

Condition [G1] is a multivariate version of the Lipschitz condition on the boundaries $g_\pm$ of the curvilinear strip (1.40) in the univariate case. It is easy to see that [G1] implies that ∂G_t is a Lipschitz continuous function of t in Hausdorff metric, but the converse is not true. It turns out though that, in the multivariate case, Lipschitz continuity of ∂G_t in the ordinary Hausdorff metric ρ_h will not suffice for deriving the desired bounds.

Likewise, using just the Hausdorff metric $\rho_h(G_s, G_t)$ for measuring the rate of change of the sections G_t with time, also proved to be inadequate. For example, imagine a “Swiss cheese”-type set, e.g. the unit disc with a large number of small holes in it. The Hausdorff distance between that set and the disc can be very small, but a change of the time section from that disc at time t to the “Swiss cheese set” at time $t + h$ can clearly lead to a rather high probability of the first hitting time

of ∂G during $(t, t + h)$, making it impossible to obtain necessary for the derivation of our results bounds for the first hitting time density (and leading to impassable obstacles in other elements of the proof as well).

The role of **[G2]** is to ensure that the section boundaries ∂G_t do not have “tight outside folds”. Together with **[G1]**, it means that, for any fixed $t \in (0, T)$, as the point $\mathbf{x}$ approaches the boundary ∂G_t , the probability that the Brownian motion starting at $\mathbf{x}$ at time t will not leave G during the residual time interval (t, T) decays as a linear function of $\rho(\mathbf{x}, \partial G_t)$, and it is exactly that rate of decay that is needed for our main result (7.2). The above-mentioned linear bound follows from the fact that the desired probability will not exceed the probability of hitting a cone with the axis parallel to the time axis and the base of radius β on the left, which is disjoint with G and which, due to conditions **[G1]** and **[G2]**, one can “attach” to the boundary ∂G at a point with the time component equal to t and space component $\rho(\mathbf{x}, \partial G_t)$ -distant from $\mathbf{x}$. The respective bound for the cone is obtained in Lemma 7.6 below.

Condition **[G3]** may be a bit more technical and less transparent. Roughly speaking, it means that the boundary ∂G_t “cannot have too many folds”. It is only used in the proof of Proposition 7.4 establishing an upper bound for the first hitting time density. The condition enables one to “translate” the “local bounds” on the conditional probability that the hitting time of ∂G belongs to $(t, t + h)$ given the value of $\mathbf{W}_t$ to the required bound for the unconditional probability of the hitting time being in that interval.

The main result of this chapter is the following bound.

Theorem 7.1. *If $G \in \mathcal{G}_{K, \beta, \gamma}$, then there exists a $c = c(K, \beta, \gamma) < \infty$ such that*

$$P(G^{(\varepsilon)}) \leq P(G) + c\varepsilon \text{ for any } \varepsilon > 0. \quad (7.2)$$

It turns out that, in the important special case of convex cross-sections G_t , conditions **[G2]** and **[G3]** are superfluous. While that **[G2]** holds in this case for arbitrary large $\beta > 0$ is obvious, the validity of **[G3]** is not hard to establish using Cauchy’s surface area formula that implies that the volume of the portion of the integration region $\{\mathbf{x} : 0 < \rho(\mathbf{x}, G_t^c) < v\}$ that lies in the spherical layer $\{\mathbf{x} : k \leq \|\mathbf{x}\| < k + 1\}$ cannot exceed cvk^{m-1} , as demonstrated in the proof of the following

assertion.

Corollary 7.2. *Assume that $G \in \mathcal{G}$ satisfies [G1] and G_t is convex for any time $t \in (0, T)$. Then G also satisfies [G2] with any $\beta > 0$ and [G3] for some $\gamma < \infty$, and so the bound from Theorem 7.1 holds true.*

The next result is a trivial consequence of Theorem 7.1. We state it here because it is the natural multivariate extension of the main bound from [BN05] (see Theorem 2.28 above).

Corollary 7.3. *Suppose $G \in \mathcal{G}_{K,\beta,\gamma}$. For any $\varepsilon > 0$, if sets $G', G'' \in \mathcal{G}$ are such that $G \subset G' \subset G^{(\varepsilon)}$, $G \subset G'' \subset G^{(\varepsilon)}$, then*

$$|P(G') - P(G'')| < c\varepsilon \tag{7.3}$$

for some constant $c = c(K, \beta, \gamma) < \infty$.

Remark 7.1. The form of the statement in the above assertion is somewhat different from the one in the univariate case where we basically estimated the difference $P(G^{(\varepsilon)}) - P(G^{(-\varepsilon)})$. This is because the multivariate situation is noticeably more complicated. In particular, in $m \geq 2$ dimensions, for a set $G \in \mathcal{G}_{K,\beta,\gamma}$ it is not necessarily true that $G^{(-\varepsilon)} \in \mathcal{G}_{K,\beta,\gamma}$, even if we allow the parameters of the class $\mathcal{G}_{K,\beta,\gamma}$ in the last instance to be different from those for the one containing G . One implication of that observation is that, without some additional restrictive assumptions, the estimation of $P(G^{(\varepsilon)}) - P(G^{(-\varepsilon)})$ becomes then impossible. On the other hand, the framework of our Theorem 7.1 is quite simple and appears to be the most natural in the multivariate setup.

7.2 Proofs

Without loss of generality, we can (and will) assume in this section that $T = 1$.

The proof is based on the same idea as in [BN05] and uses the total probability formula representation (7.5) below for the difference between the probabilities $P(G^{(\varepsilon)})$ and $P(G)$. This difference is just the probability that the Brownian motion

leaves the set G but does not leave the slightly bigger $G^{(\varepsilon)}$. To estimate that difference, one needs bounds for both the integrand and the “integrator” in the integral on the right-hand side of (7.5).

First we will deal with the latter, obtaining in Proposition 7.4 an upper bound for the density of the first hitting time of ∂G (of which the derivation relies on Lemmata 7.5–7.9). For the former, one can expect that the integrand is $O(\varepsilon)$ as $\varepsilon \rightarrow 0$, with the coefficient of ε in the bound increasing as t approaches 1, as less and less time will be remaining for the Brownian motion to leave $G^{(\varepsilon)}$. The desired bound for the integrand will follow from Lemma 7.10 which we use to bound the tail of the distribution of the time the Brownian motion starting at time t at a boundary point of G_t hits a cone located outside $G^{(\varepsilon)}$ and “attached” to $\partial G^{(\varepsilon)}$ at a point with the time component equal to t . As we already pointed out, conditions [G1] and [G2] make it possible to use that result to bound the probability of not hitting $\partial G^{(\varepsilon)}$ during the residual time interval $(t, 1)$.

For a measurable $H \subset [0, 1] \times \mathbb{R}^m$, let

$$\tau(H) := \inf\{t > 0 : (t, \mathbf{W}_t) \in \partial H\}, \quad (7.4)$$

setting $\tau(H) := 1$ when $\mathbf{W}_t \notin \partial H_t$, $t \in (0, 1)$. Letting $\tau := \tau(G)$, $\tau^{(\varepsilon)} := \tau(G^{(\varepsilon)})$, we have from the Markov property of the Brownian motion that, for $\varepsilon > 0$,

$$D_\varepsilon(G) := P(G^{(\varepsilon)}) - P(G) = \int_{(0,1)} \mathbb{P}(\tau^{(\varepsilon)} = 1 | \tau = t) \mathbb{P}(\tau \in dt). \quad (7.5)$$

The following proposition, establishing absolute continuity of τ and providing upper bounds for its density, is of independent interest.

Proposition 7.4. *The random variable τ has density p on $(0, 1)$ satisfying*

$$p(t) \leq 8m^2\gamma \begin{cases} \sqrt{\frac{1}{\pi t}} + \frac{m-1}{2\beta-Kt} + 2K + \frac{2}{t}, & t \in (0, \min\{\beta/K, 1\}), \\ \sqrt{\frac{K}{\pi\beta}} + \frac{\beta+2}{2t-\beta/K} + \frac{m-1}{\beta} + K, & t \in [\min\{\beta/K, 1\}, 1). \end{cases}$$

To prove the proposition, set $\tau_t := \inf\{s > t : (s, \mathbf{W}_s) \in \partial G\}$ for $t \in (0, 1)$ and

note that, for $0 < h < 1 - t$,

$$\begin{aligned}\mathbb{P}(\tau \in (t, t+h)) &= \int_{G_t} \mathbb{P}(\tau \in (t, t+h) | \mathbf{W}_t = \mathbf{z}) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}) \\ &= \int_{G_t} \mathbb{P}(\tau > t | \mathbf{W}_t = \mathbf{z}) \mathbb{P}(\tau_t < t+h | \mathbf{W}_t = \mathbf{z}) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}).\end{aligned}\quad (7.6)$$

Next we will bound the two conditional probabilities appearing as factors in the integrand on the right hand side of (7.6). It will be convenient to use the notation

$$r(\mathbf{z}) := \rho(\mathbf{z}, \partial G_t) \quad (7.7)$$

(for a fixed t). The following lemma gives a bound for the first factor.

Lemma 7.5. *For $t \in (0, 1)$, one has*

$$\frac{\mathbb{P}(\tau > t | \mathbf{W}_t = \mathbf{z})}{2r(\mathbf{z})} \leq \begin{cases} \sqrt{\frac{1}{\pi t}} + \frac{2(\|\mathbf{z}\| + r(\mathbf{z}))}{t} + \frac{m-1}{2\beta-Kt} + 2K, & t < \beta/K, \\ \sqrt{\frac{K}{\pi\beta}} + \frac{\|\mathbf{z}\| + r(\mathbf{z}) + \beta/2}{t - \beta/(2K)} + \frac{m-1}{\beta} + K, & t \geq \beta/K. \end{cases}$$

The proof of Lemma 7.5 uses our next three lemmata. The main idea is that the desired conditional boundary non-hitting probability cannot exceed the probability of not hitting a simple “more remote” set, namely, a cone “attached from the outside” to the “nearest” to $\mathbf{z}$ point of ∂G_t , with the axis parallel to the time axis and such that its base is on the right. For that simpler boundary, the desired bound is obtained in Lemma 7.6. This is done by reducing the problem to a univariate straight line boundary crossing probability by a diffusion process, namely, by the radial process of the m -dimensional Brownian bridge. To obtain the desired bound for the latter problem, we use the comparison Lemma 7.7 and a simple reference process (7.13), which is just a univariate arithmetic Brownian motion. For that process, the respective bound is given in Lemma 7.8.

For $u, v > 0$, introduce the (possibly truncated) cones

$$C(v, u) := \{(s, \mathbf{x}) \in [0, v] \times \mathbb{R}^m : \|\mathbf{x}\| \leq u - Ks\}, \quad (7.8)$$

$$C^*(v, u) := \{(s, \mathbf{x}) \in [0, v] \times \mathbb{R}^m : |x_i| \leq (u - Ks)/\sqrt{m}, i = 1, \dots, m\}.$$

Clearly, $C^*(v, u) \subset C(v, u)$.

We will slightly abuse notation by denoting by $\mathbb{P}_{\mathbf{x}}$ the distribution on the canonical space corresponding to the Brownian motion process that starts at the point $\mathbf{W}_0 = \mathbf{x} \in \mathbb{R}^m$ and keeping the notation $\tau(H)$ for the stopping time (7.4) for that process.

Lemma 7.6. *For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ with $\|\mathbf{x}\| > \beta$, we have*

$$\frac{\mathbb{P}_{\mathbf{x}}(\tau(C(t, \beta)) > u | \mathbf{W}_t = \mathbf{y})}{\|\mathbf{x}\| - \beta} \leq \begin{cases} \sqrt{\frac{2}{\pi u}} + 2 \left(\frac{2(\|\mathbf{y}\| - \beta)}{t} + \frac{m-1}{2\beta - Kt} + 2K \right)^+, & u \leq t/2, t < \beta/K, \\ \sqrt{\frac{2}{\pi u}} + 2 \left(\frac{\|\mathbf{y}\| - \beta/2}{t - \beta/(2K)} + \frac{m-1}{\beta} + K \right)^+, & u \leq \beta/(2K), t \geq \beta/K, \end{cases}$$

where $x^+ := \max\{0, x\}$.

Note that the above upper bounds agree at $t = \beta/K$.

To prove Lemma 7.6, we will require the following two additional lemmas. For a univariate process $X = \{X_t\}_{t \geq 0}$ and $x \in \mathbb{R}$, set

$$\eta_x(X) := \inf\{t \geq 0 : X_t = x\}. \quad (7.9)$$

Lemma 7.7. *Let $\{W_t\}_{t \geq 0}$ be the standard univariate Brownian motion given on a filtered probability space, $\{Y_t\}_{t \geq 0}$ a continuous adapted process on the same space. Let $X_t^{(1)}$ and $X_t^{(2)}$ be strong unique solutions of the stochastic differential equations (SDEs)*

$$dX_t^{(i)} = a_i(t, X_t^{(i)}, Y_t)dt + dW_t, \quad X_0^{(i)} = x_0, \quad i = 1, 2,$$

where a_i are continuous. Suppose that, for a given $l < x_0$, one has $a_1(t, x, y) < a_2(t, x, y)$ for all $(t, x) \in [0, \infty) \times (l, \infty)$, $y \in \mathbb{R}$. Then $X_t^{(1)} < X_t^{(2)}$ a.s. for all $t \in (0, \eta_l(X^{(1)}))$.

The proof of Lemma 7.7 below follows the argument proving a somewhat weaker assertion of Lemma 4 on p.120 of [GS72].

Proof. Define the continuously differentiable function

$$\Delta(t) := X_t^{(2)} - X_t^{(1)} = \int_0^t (a_2(s, X_s^{(2)}, Y_s) - a_1(s, X_s^{(1)}, Y_s)) ds, \quad t < \eta_l.$$

Then, for all points $t < \eta_l$ with $\Delta(t) = 0$, we have that $X_t^{(1)} = X_t^{(2)}$, and so at these points

$$\Delta'(t) = a_2(t, X_t^{(2)}, Y_t) - a_1(t, X_t^{(1)}, Y_t) > 0.$$

In particular, we have $\Delta(0) = 0$, $\Delta'(0+) > 0$. Therefore we can find a $\delta > 0$ such that $\Delta(t) > 0$ for all $0 < t \leq \delta$.

Now suppose the set $\{t \in (0, \eta_l) : \Delta(t) = 0\}$ is not empty. Then for $t_1 := \inf\{t \in (0, \eta_l) : \Delta(t) = 0\}$ we have $t_1 > 0$, $\Delta(t_1) = 0$, $\Delta'(t_1) > 0$, and so there exists a $\delta_1 > 0$ such that $\Delta(t) < 0$ for $t \in [t_1 - \delta_1, t_1]$. Therefore $\Delta(t)$ changes signs on the interval $[\delta, t_1 - \delta_1]$, i.e., it takes on the value zero there, which contradicts the definition of t_1 . We conclude that $\{t \in (0, \eta_l) : \Delta(t) = 0\}$ is empty a.s., and since $\Delta(t) > 0$ for sufficiently small t , $\Delta(t) > 0$ for all $t \in (0, \eta_l)$ as required. $\square$

Recall that $\{W_t\}_{t \geq 0}$ is the standard univariate Brownian motion process.

Lemma 7.8. *For $c \in \mathbb{R}$ and $\varepsilon > 0$,*

$$\mathbb{P}\left(\sup_{0 \leq s \leq t} (W_s - cs) < \varepsilon\right) \leq \varepsilon \left(\sqrt{\frac{2}{\pi t}} + 2c^+\right), \quad t > 0.$$

Proof. The probability on the left hand side above is known explicitly (cf. (1.1)): denoting by Φ the standard normal distribution function,

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \leq s \leq t} (W_s - cs) < \varepsilon\right) &= \Phi(c\sqrt{t} + \varepsilon/\sqrt{t}) - e^{-2c\varepsilon} \Phi(c\sqrt{t} - \varepsilon/\sqrt{t}) \\ &\leq \Phi(c\sqrt{t} + \varepsilon/\sqrt{t}) - e^{-2c^+\varepsilon} \Phi(c\sqrt{t} - \varepsilon/\sqrt{t}) \\ &\leq \Phi(c\sqrt{t} + \varepsilon/\sqrt{t}) - \Phi(c\sqrt{t} - \varepsilon/\sqrt{t}) + (1 - e^{-2c^+\varepsilon}) \\ &\leq \sup_{x \in \mathbb{R}} \Phi'(x) \times 2\varepsilon/\sqrt{t} + 2c^+\varepsilon \\ &= \varepsilon \left(\sqrt{\frac{2}{\pi t}} + 2c^+\right). \end{aligned}$$

$\square$

Proof of Lemma 7.6. Fix a $t \in (0, 1)$ and let $\mathbf{B} = \{\mathbf{B}_s = (B_s^{(1)}, \dots, B_s^{(m)})\}_{0 \leq s \leq t}$ be an m -dimensional Brownian bridge process starting at $\mathbf{x} \in \mathbb{R}^m$ at time 0 and ending at $\mathbf{y} \in \mathbb{R}^m$ at time t .

In order to use Lemma 7.7, we will now derive an SDE for the radial process $S_s := \|\mathbf{B}_s\|$ of $\mathbf{B}$. Recall from (1.9) that $\mathbf{B}$ satisfies the SDE

$$d\mathbf{B}_s = \frac{\mathbf{y} - \mathbf{B}_s}{t - s} ds + d\mathbf{W}_s, \quad 0 < s < t, \quad (7.10)$$

By Itô's formula, the squared radial process has stochastic differential

$$dS_s^2 = 2 \sum_{i=1}^m B_s^{(i)} dB_s^{(i)} + m ds, \quad 0 < s < t. \quad (7.11)$$

Setting $\xi_s := \mathbf{B}_s/S_s$, we have $\|\xi_s\| \equiv 1$ and therefore

$$\xi_s(\mathbf{y}) := \xi_s \mathbf{y}^T \leq \|\mathbf{y}\|, \quad (7.12)$$

where $\mathbf{y}^T$ denotes the transpose of $\mathbf{y}$. Then, for $0 < s < t$, we have from (7.10) and (7.11) that

$$\begin{aligned} dS_s^2 &= 2 \sum_{i=1}^m B_s^{(i)} \left(\frac{y_i - B_s^{(i)}}{t - s} ds + dW_s^{(i)} \right) + m ds \\ &= 2 \left(\frac{\mathbf{B}_s \mathbf{y}^T - S_s^2}{t - s} + \frac{m}{2} \right) ds + 2 \mathbf{B}_s d\mathbf{W}_s^T \\ &= 2 \left(\frac{S_s \xi_s(\mathbf{y}) - S_s^2}{t - s} + \frac{m}{2} \right) ds + 2 S_s \xi_s d\mathbf{W}_s^T \\ &= 2 \left(\frac{S_s \xi_s(\mathbf{y}) - S_s^2}{t - s} + \frac{m}{2} \right) ds + 2 S_s d\widetilde{W}_s, \end{aligned}$$

where $\{\widetilde{W}_t\}_{t \geq 0}$ is a standard univariate Brownian motion, and the last equality follows from Theorem 8.4.2 in [O07].

Using the above SDE for $\{S_s^2\}$ and Itô's formula with $f(x) = \sqrt{x}$, we have

$$\begin{aligned} dS_s &= f'(S_s^2) dS_s^2 + \frac{1}{2} f''(S_s^2) (dS_s^2)^2 \\ &= \frac{1}{2S_s} \left[2 \left(\frac{S_s \xi_s(\mathbf{y}) - S_s^2}{t - s} + \frac{m}{2} \right) ds + 2 S_s d\widetilde{W}_s \right] - \frac{1}{8S_s^3} (2S_s)^2 ds \\ &= \left(\frac{\xi_s(\mathbf{y}) - S_s}{t - s} + \frac{m - 1}{2S_s} \right) ds + d\widetilde{W}_s, \quad 0 < s < t. \end{aligned}$$

Now introduce, for a fixed $a < \|\mathbf{x}\|$ and $t_0 \in (0, t)$, the reference process

$$\bar{S}_s := \|\mathbf{x}\| + \left(\frac{\|\mathbf{y}\| - a}{t - t_0} + \frac{m-1}{2a} \right) s + \widetilde{W}_s, \quad s \geq 0, \quad (7.13)$$

Since $\|\mathbf{y}\| \geq \xi_s(\mathbf{y})$ by (7.12), Lemma 7.7 implies that, for all $s \in [0, \min\{t_0, \eta_a(S)\}]$, one has $\bar{S}_s \geq S_s$ a.s.

Consider first the case $t \geq \beta/K$ and set $t_0 := \beta/(2K)$, $a := \beta/2$. Then, for all $u \leq t_0$, one has

$$\begin{aligned} & \mathbb{P}_{\mathbf{x}}(\tau(C(t, \beta)) > u | \mathbf{W}_t = \mathbf{y}) \\ &= \mathbb{P}\left(\inf_{0 \leq s \leq u} (S_s - \beta + Ks) > 0\right) \\ &\leq \mathbb{P}\left(\inf_{0 \leq s \leq u} (\bar{S}_s - \beta + Ks) > 0\right) \\ &= \mathbb{P}\left[\inf_{0 \leq s \leq u} \left(\|\mathbf{x}\| + \left(\frac{\|\mathbf{y}\| - \beta/2}{t - \beta/(2K)} + \frac{m-1}{\beta}\right)s + W_s - \beta + Ks\right) > 0\right] \\ &= \mathbb{P}\left[\sup_{0 \leq s \leq u} \left(W_s - \left(\frac{\|\mathbf{y}\| - \beta/2}{t - \beta/(2K)} + \frac{m-1}{\beta} + K\right)s\right) < \|\mathbf{x}\| - \beta\right] \\ &\leq (\|\mathbf{x}\| - \beta) \left[\sqrt{\frac{2}{\pi u}} + 2 \left(\frac{\|\mathbf{y}\| - \beta/2}{t - \beta/(2K)} + \frac{m-1}{\beta} + K \right)^+ \right] \end{aligned}$$

by Lemma 7.8.

Now consider the case $t < \beta/K$ and set $t_0 := t/2$, $a := \beta - Kt/2$. Then, for all $u \leq t_0$, we have

$$\begin{aligned} & \mathbb{P}_{\mathbf{x}}(\tau(C(t, \beta)) > u | \mathbf{W}_t = \mathbf{y}) \\ &\leq \mathbb{P}\left(\inf_{0 \leq s \leq u} (\bar{S}_s - \beta + Ks) > 0\right) \\ &= \mathbb{P}\left[\inf_{0 \leq s \leq u} \left(\|\mathbf{x}\| + \left(\frac{\|\mathbf{y}\| - \beta + Kt/2}{t/2} + \frac{m-1}{2\beta - Kt}\right)s + W_s - \beta + Ks\right) > 0\right] \\ &= \mathbb{P}\left[\sup_{0 \leq s \leq u} \left(W_s - \left(\frac{2(\|\mathbf{y}\| - \beta)}{t} + \frac{m-1}{2\beta - Kt} + 2K\right)s\right) < \|\mathbf{x}\| - \beta\right] \\ &\leq (\|\mathbf{x}\| - \beta) \left[\sqrt{\frac{2}{\pi u}} + 2 \left(\frac{2(\|\mathbf{y}\| - \beta)}{t} + \frac{m-1}{2\beta - Kt} + 2K \right)^+ \right]. \end{aligned}$$

Lemma 7.6 is proved. □

Proof of Lemma 7.5. Fix $t \in (0, 1)$ and $\mathbf{z} \in G_t$. Reversing the time for the conditional Brownian motion process, we have, for any fixed $t' \in (0, t)$,

$$\begin{aligned}\mathbb{P}(\tau > t | \mathbf{W}_t = \mathbf{z}) &= \mathbb{P}(\mathbf{W}_s \in G_s, s \in (0, t) | \mathbf{W}_t = \mathbf{z}) \\ &= \mathbb{P}_{\mathbf{z}}(\mathbf{W}_s \in G_{t-s}, s \in (0, t) | \mathbf{W}_t = \mathbf{0}) \\ &\leq \mathbb{P}_{\mathbf{z}}(\mathbf{W}_s \in G_{t-s}, s \in (0, t') | \mathbf{W}_t = \mathbf{0}).\end{aligned}\quad (7.14)$$

One can clearly choose a $\mathbf{b} \in \partial G_t$ such that $\rho(\mathbf{z}, \mathbf{b}) = r(\mathbf{z})$. Then by condition [G2] there exists a ball $B_\beta(\mathbf{c}) \subset G_t^c$ with $\mathbf{b} \in \partial B_\beta(\mathbf{c})$. Using $B_\beta(\mathbf{c})$ as the base for the cone

$$C := (0, \mathbf{c}) + C(t, \beta),$$

it follows from Lipschitz condition [G1] that $G_{t-s} \subset C_s^c$, $s \in [0, t]$. Therefore,

$$\begin{aligned}\mathbb{P}_{\mathbf{z}}(\mathbf{W}_s \in G_{t-s}, s \in (0, t') | \mathbf{W}_t = \mathbf{0}) &\leq \mathbb{P}_{\mathbf{z}}(\mathbf{W}_s \in C_s^c, s \in (0, t') | \mathbf{W}_t = \mathbf{0}) \\ &= \mathbb{P}_{\mathbf{z}-\mathbf{c}}(\tau(C(t, \beta)) > t' | \mathbf{W}_t = -\mathbf{c}).\end{aligned}\quad (7.15)$$

Since $\|-\mathbf{c}\| \leq \|\mathbf{z}\| + r(\mathbf{z}) + \beta$, we immediately obtain the bounds stated in Lemma 7.5 from Lemma 7.6 with

$$t' = u := \begin{cases} t/2, & t < \beta/K, \\ \beta/(2K), & t \geq \beta/K. \end{cases}$$

□

Now we will turn to bounding the second factor in the integrand on the right hand side of (7.6). Recall the definition (7.7) of $r(\mathbf{z})$ (see p. 131).

Lemma 7.9. *For and fixed $t \in (0, 1)$, $\mathbf{z} \in G_t$, one has*

$$\mathbb{P}(\tau_t < t + h | \mathbf{W}_t = \mathbf{z}) \leq 2m \exp\left(\frac{r(\mathbf{z})K}{m} - \frac{r(\mathbf{z})^2}{2hm}\right), \quad h \in (0, \min\{r(\mathbf{z})/K, 1 - t\}).$$

Proof. We will again use cones and reduction to univariate boundary crossing problems. For h from the specified interval, setting $\bar{\Phi}(x) := 1 - \Phi(x)$, we have

$$\mathbb{P}(\tau_t < t + h | \mathbf{W}_t = \mathbf{z}) \leq \mathbb{P}(\tau(C(h, r(\mathbf{z}))) < h)$$

$$\begin{aligned}
&\leq \mathbb{P}(\tau(C^*(h, r(\mathbf{z}))) < h) \\
&\leq 2m \mathbb{P}\left(\sup_{0 \leq s \leq h} (W(s) + Ks/\sqrt{m}) \geq r(\mathbf{z})/\sqrt{m}\right) \\
&= 2m \int_0^h \frac{r(\mathbf{z})/\sqrt{m}}{\sqrt{2\pi}s^{3/2}} \exp\left(\frac{-(r(\mathbf{z})/\sqrt{m} - Ks/\sqrt{m})^2}{2s}\right) ds \\
&\quad (7.16)
\end{aligned}$$

$$\begin{aligned}
&= r(\mathbf{z}) \sqrt{\frac{2m}{\pi}} e^{r(\mathbf{z})K/m} \int_0^h s^{-3/2} \exp\left(-\frac{r(\mathbf{z})^2}{2sm} - \frac{K^2s}{2m}\right) ds \\
&\leq r(\mathbf{z}) \sqrt{\frac{2m}{\pi}} e^{r(\mathbf{z})K/m} \int_0^h s^{-3/2} \exp\left(-\frac{r(\mathbf{z})^2}{2sm}\right) ds \\
&= 4m e^{r(\mathbf{z})K/m} \bar{\Phi}(r(\mathbf{z})/\sqrt{hm}) \\
&\quad (7.17)
\end{aligned}$$

$$\leq 2m \exp\left(\frac{r(\mathbf{z})K}{m} - \frac{r(\mathbf{z})^2}{2hm}\right), \quad (7.18)$$

where (7.16) follows from Kendall's formula (see (1.2)), (7.17) follows by making the substitution $u = r(\mathbf{z})/\sqrt{sm}$, and (7.18) follows by using the bound $\bar{\Phi}(x) \leq \frac{1}{2}e^{-x^2/2}$, $x > 0$. The lemma is proved. $\square$

Proof of Proposition 7.4. Suppose that $t < \beta/K$. Then from (7.6) and the bounds derived in Lemmata 7.5, 7.9, for $0 < h < 1 - t$ we have

$$\begin{aligned}
\mathbb{P}(\tau \in (t, t+h)) &= \int_{G_t} \mathbb{P}(\tau > t | \mathbf{W}_t = \mathbf{z}) \mathbb{P}(\tau_t < t+h | \mathbf{W}_t = \mathbf{z}) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}) \\
&\leq 4m \int_{G_t} \left(\sqrt{\frac{1}{\pi t}} + \frac{2(\|\mathbf{z}\| + r(\mathbf{z}))}{t} + \frac{m-1}{2\beta - Kt} + 2K \right) \\
&\quad \times r(\mathbf{z}) \exp\left(\frac{r(\mathbf{z})K}{m} - \frac{r(\mathbf{z})^2}{2hm}\right) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}) \\
&= 4m \left(\left(\sqrt{\frac{1}{\pi t}} + \frac{m-1}{2\beta - Kt} + 2K \right) I_1 + \frac{2}{t} (I_2 + I_3) \right), \quad (7.19)
\end{aligned}$$

where

$$\begin{aligned}
I_1 &:= \int_{G_t} r(\mathbf{z}) \exp\left(\frac{r(\mathbf{z})K}{m} - \frac{r(\mathbf{z})^2}{2hm}\right) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}), \\
I_2 &:= \int_{G_t} r(\mathbf{z}) \|\mathbf{z}\| \exp\left(\frac{r(\mathbf{z})K}{m} - \frac{r(\mathbf{z})^2}{2hm}\right) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}), \\
I_3 &:= \int_{G_t} r(\mathbf{z})^2 \exp\left(\frac{r(\mathbf{z})K}{m} - \frac{r(\mathbf{z})^2}{2hm}\right) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}).
\end{aligned}$$

Set

$$Z := r(\mathbf{W}_t) \mathbf{1}_{\{\mathbf{W}_t \in G_t\}},$$

where $\mathbf{1}_E$ is the indicator of event E . Then, for $u(x) := xe^{xK/m - x^2/(2hm)}$, we have

$$I_1 = \mathbb{E}Z \exp\left\{\frac{ZK}{m} - \frac{Z^2}{2hm}\right\} = \int_0^\infty u(x) d\nu(x),$$

where $\nu(x)$ is the distribution function of Z . Integrating by parts and using the bound $\nu(x) < \gamma x$, $x \in (0, v_0)$, from [G3], we obtain

$$\begin{aligned} I_1 &= [\nu(x)u(x)]_0^\infty - \int_0^\infty \nu(x) du(x) = - \int_0^\infty \nu(x) du(x) \\ &= \int_0^\infty \nu(x) \left(\frac{x^2}{hm} - \frac{xK}{m} - 1 \right) \exp\left(\frac{xK}{m} - \frac{x^2}{2hm} \right) dx \\ &< \frac{1}{hm} \int_0^\infty \nu(x) x^2 \exp\left(\frac{xK}{m} - \frac{x^2}{2hm} \right) dx = \frac{1}{hm} \left(\int_0^{v_0} + \int_{v_0}^\infty \right) (\cdots) dx \\ &< \frac{1}{hm} \left(\gamma \int_0^\infty x^3 \exp\left(\frac{xK}{m} - \frac{x^2}{2hm} \right) dx + \int_{v_0}^\infty x^2 \exp\left(\frac{xK}{m} - \frac{x^2}{2hm} \right) dx \right) \\ &= \gamma hm \int_0^\infty s^3 e^{sK} \sqrt{h/m - s^2/2} ds + \sqrt{hm} \int_{v_0/\sqrt{hm}}^\infty s^2 e^{sK} \sqrt{h/m - s^2/2} ds \\ &= 2\gamma hm + o(h), \quad h \searrow 0, \end{aligned}$$

where the second last relation follows by making the substitution $s = x/\sqrt{hm}$.

Using [G3] and following the same steps as above, we conclude that

$$I_2 < 2\gamma hm + o(h), \quad h \searrow 0.$$

Finally, it is even simpler to show that $I_3 = o(h)$ as $h \searrow 0$.

Then, from (7.19), we have

$$\mathbb{P}(\tau \in (t, t+h)) < 8m^2\gamma h \left(\sqrt{\frac{1}{\pi t}} + \frac{m-1}{2\beta - Kt} + 2K + \frac{2}{t} \right) + o(h).$$

It follows that τ has an absolutely continuous distribution specified by a density p

satisfying

$$p(t) \leq 8m^2\gamma \left(\sqrt{\frac{1}{\pi t}} + \frac{m-1}{2\beta - Kt} + 2K + \frac{2}{t} \right), \quad t \in (0, \min\{\beta/K, 1\}).$$

Now consider the case when $t \geq \beta/K$. Then, from (7.6) and the bounds derived in Lemmata 7.5, 7.9, we have

$$\begin{aligned} \mathbb{P}(\tau \in (t, t+h)) &= \int_{G_t} \mathbb{P}(\tau > t | \mathbf{W}_t = \mathbf{z}) \mathbb{P}(\tau_t < t+h | \mathbf{W}_t = \mathbf{z}) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}) \\ &\leq 4m \int_{G_t} \left(\sqrt{\frac{K}{\pi\beta}} + \frac{\|\mathbf{z}\| + r(\mathbf{z}) + \beta/2}{t - \beta/(2K)} + \frac{m-1}{\beta} + K \right) \\ &\quad \times r(\mathbf{z}) \exp\left(\frac{r(\mathbf{z})K}{m} - \frac{r(\mathbf{z})^2}{2hm}\right) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z}) \\ &= 4m \left[\left(\sqrt{\frac{K}{\pi\beta}} + \frac{\beta}{2t - \beta/K} + \frac{m-1}{\beta} + K \right) I_1 + \frac{I_2 + I_3}{t - \beta/(2K)} \right]. \end{aligned}$$

Therefore, for $t \in [\beta/K, 1)$,

$$\mathbb{P}(\tau \in (t, t+h)) \leq 8m^2h\gamma \left(\sqrt{\frac{K}{\pi\beta}} + \frac{\beta+2}{2t - \beta/K} + \frac{m-1}{\beta} + K \right).$$

As above, it follows that τ has density p satisfying

$$p(t) \leq 8m^2\gamma \left(\sqrt{\frac{K}{\pi\beta}} + \frac{\beta+2}{2t - \beta/K} + \frac{m-1}{\beta} + K \right), \quad t \in [\min\{\beta/K, 1\}, 1).$$

Proposition 7.4 is proved. □

As we pointed out in the outline of the idea of the proof of Theorem 7.1, we will also use the following lemma that provides a bound for the integrand on the right hand side of (7.5).

Lemma 7.10. *For $\mathbf{x} \in \mathbb{R}^m$ and $0 < r < \|\mathbf{x}\|$, we have*

$$\frac{\mathbb{P}_{\mathbf{x}}(\tau(C(t, r)) > t)}{2(\|\mathbf{x}\| - r)} \leq \begin{cases} \sqrt{\frac{1}{\pi t}} + \frac{m-1}{2r-Kt} + K, & t < r/K, \\ \sqrt{\frac{K}{\pi r}} + \frac{m-1}{r} + K, & t \geq r/K. \end{cases}$$

Proof. Similarly to the proof of Lemma 7.6, we will reduce the problem to a univariate straight line boundary crossing problem and then use the comparison Lemma 7.7 to obtain the desired bound for the latter problem.

Denote by $R = \{R_s\}_{s \geq 0}$ an m -dimensional Bessel process started at $\|\mathbf{x}\|$ at time 0. Since R is defined by the radial part of $\mathbf{W}_t$ (cf. (1.6)), one has that

$$\mathbb{P}_{\mathbf{x}}(\tau(C(t, r)) > t) = \mathbb{P}\left(\inf_{0 \leq s \leq t} (R_s - r + Ks) > 0\right). \quad (7.20)$$

As is well-known (cf. (1.7)), R satisfies the SDE

$$dR_s = \frac{m-1}{2R_s} ds + d\widetilde{W}_s, \quad s > 0, \quad R_0 = \|\mathbf{x}\|,$$

$\{\widetilde{W}_s\}_{s \geq 0}$ being a standard univariate Brownian motion process.

Consider first the case $t < r/K$ and let

$$\overline{R}_s := \|\mathbf{x}\| + \frac{m-1}{2r-Kt}s + \widetilde{W}_s, \quad s \in [0, t/2].$$

Then, by Lemma 7.7, we have $\overline{R}_s \geq R_s$ a.s. for all $s \in [0, \min\{t/2, \eta_{r-Kt/2}(R)\}]$ (cf. (7.9)). From here, (7.20) and Lemma 7.8, one has

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}(\tau(C(t, r)) > t) &\leq \mathbb{P}\left(\inf_{0 \leq s \leq t/2} (\overline{R}_s - r + Ks) > 0\right) \\ &= \mathbb{P}\left(\inf_{0 \leq s \leq t/2} \left(\|\mathbf{x}\| + \frac{m-1}{2r-Kt}s + \widetilde{W}_s - r + Ks\right) > 0\right) \\ &= \mathbb{P}\left[\sup_{0 \leq s \leq t/2} \left(\widetilde{W}_s - \left(\frac{m-1}{2r-Kt} + K\right)s\right) < \|\mathbf{x}\| - r\right] \\ &\leq 2(\|\mathbf{x}\| - r) \left(\sqrt{\frac{1}{\pi t}} + \frac{m-1}{2r-Kt} + K\right). \end{aligned}$$

Now consider the case $t \geq r/K$ and let

$$\overline{R}_s := \|\mathbf{x}\| + \frac{m-1}{r}s + \widetilde{W}_s, \quad s \in [0, r/(2K)].$$

Then, by Lemma 7.7, we have $\overline{R}_s \geq R_s$ a.s. for all $s \in [0, \min\{r/(2K), \eta_{r/2}(R)\}]$,

and so from (7.20) and Lemma 7.8, a similar derivation yields the bound

$$\mathbb{P}_{\mathbf{x}}(\tau(C(t, r)) > t) \leq 2(\|\mathbf{x}\| - r) \left(\sqrt{\frac{K}{\pi r}} + \frac{m-1}{r} + K \right),$$

as required. $\square$

Now we can complete the proof of Theorem 7.1. The integrand on the right hand side of (7.5) has the form

$$\begin{aligned} \mathbb{P}(\tau^{(\varepsilon)} = 1 | \tau = t) &= \int_{\partial G_t} \mathbb{P}(\tau^{(\varepsilon)} = 1, \mathbf{W}_t \in d\mathbf{z} | \tau = t) \\ &= \int_{\partial G_t} \mathbb{P}_{\mathbf{z}}(\mathbf{W}_s \in G_{t+s}^{(\varepsilon)}, s \in (0, 1-t)) \mathbb{P}(\mathbf{W}_t \in d\mathbf{z} | \tau = t). \end{aligned} \quad (7.21)$$

For any $\mathbf{z} \in \partial G_t$, by [G2] there is a point $\mathbf{y}$ such that $B_\beta(\mathbf{y}) \subset G_t^c$ and $\mathbf{z} \in \partial B_\beta(\mathbf{y})$. Clearly, $B_{\beta_\varepsilon}(\mathbf{y}) \subset (G_t^{(\varepsilon)})^c$, where $\beta_\varepsilon := \beta - \varepsilon$ (we assume without loss of generality that $\varepsilon < \beta/2$), and so $G_t^{(\varepsilon)} \subset (B_{\beta_\varepsilon}(\mathbf{y}))^c$. By condition [G1], we then also have

$$G_{t+s}^{(\varepsilon)} \subset G_t^{(\varepsilon+Ks)} \subset (B_{\beta_\varepsilon-Ks}(\mathbf{y}))^c = ((0, \mathbf{y}) + C(1-t, \beta_\varepsilon))_s^c, \quad s \leq \frac{\beta_\varepsilon}{K},$$

using notation (7.8). Therefore, since $\|\mathbf{z} - \mathbf{y}\| - \beta_\varepsilon = \varepsilon$, one has from Lemma 7.10 that

$$\begin{aligned} \mathbb{P}_{\mathbf{z}}(\mathbf{W}_s \in G_{t+s}^{(\varepsilon)}, s \in (0, 1-t)) &\leq \mathbb{P}_{\mathbf{z}-\mathbf{y}}(\tau(C(1-t, \beta_\varepsilon)) > 1-t) \\ &\leq 2\varepsilon \begin{cases} \sqrt{\frac{1}{\pi(1-t)}} + \frac{m-1}{2\beta_\varepsilon-K(1-t)} + K, & 1-t < \beta_\varepsilon/K, \\ \sqrt{\frac{K}{\pi\beta_\varepsilon}} + \frac{m-1}{\beta_\varepsilon} + K, & 1-t \geq \beta_\varepsilon/K. \end{cases} \end{aligned}$$

Now it follows from (7.5) and (7.21) that

$$\begin{aligned} D_\varepsilon(G) &\leq 2\varepsilon \int_0^{1-\beta_\varepsilon/K} \left(\sqrt{\frac{K}{\pi\beta_\varepsilon}} + \frac{m-1}{\beta_\varepsilon} + K \right) p(t) dt \\ &\quad + 2\varepsilon \int_{1-\beta_\varepsilon/K}^1 \left(\sqrt{\frac{1}{\pi(1-t)}} + \frac{m-1}{2\beta_\varepsilon-K(1-t)} + K \right) p(t) dt \end{aligned}$$

$$\leq 2\varepsilon \left(\sqrt{\frac{K}{\pi\beta_\varepsilon}} + \frac{m-1}{\beta_\varepsilon} + K \right) + 2\varepsilon \int_0^1 \frac{p(t) dt}{\sqrt{\pi(1-t)}}.$$

Here

$$\begin{aligned} \int_0^1 \frac{p(t) dt}{\sqrt{\pi(1-t)}} &\leq \sqrt{\frac{2}{\pi}} + \int_{1/2}^1 \frac{p(t) dt}{\sqrt{\pi(1-t)}} \\ &\leq \sqrt{\frac{2}{\pi}} + 8m^2\gamma\sqrt{\frac{2}{\pi}} \left(\sqrt{\frac{K}{\pi\beta}} + 2(\beta+2) + \frac{m-1}{\beta} + K \right) =: c^*, \end{aligned}$$

where the second inequality follows by assuming without loss of generality that $\beta/K < 1/2$ and applying the second bound from Proposition 7.4. Recalling that $\beta_\varepsilon \geq \beta/2$, we conclude that

$$D_\varepsilon(G) \leq 2\varepsilon \left(c^* + \sqrt{\frac{2K}{\pi\beta}} + \frac{2(m-1)}{\beta} + K \right).$$

Theorem 7.1 is proved. □

Proof of Corollary 7.2. Condition **[G2]** is clearly satisfied (with arbitrary large $\beta > 0$) due to the convexity of G_t , so we only need to verify **[G3]**.

We can assume without loss of generality that there exists a $\delta > 0$ such that $B_\delta(\mathbf{0}) \subset (\partial G)_0$ (for otherwise it is easy to see that, in view of **[G1]**, one has $P(G) = 0$ and the whole problem becomes trivial).

Consider first the case $t > t_0 := \delta/(2K)$ (assuming that $t_0 < 1$).

Introduce the sequence of spherical layers $C_k := B_k(\mathbf{0}) \setminus B_{k-1}(\mathbf{0})$, $k = 1, 2, \dots$. As the cross-section G_t is convex, it follows from Cauchy's surface area formula (see e.g. Theorem 5.5.2 on p.56 in [KR97]) that the $((m-1)$ -dimensional) surface area of $\partial(G_t \cap B_k(\mathbf{0}))$ does not exceed the surface area of $B_k(\mathbf{0})$ which is equal to $k^{m-1}\omega_{m-1}$, where ω_{m-1} is the unit sphere area.

Therefore, again using the convexity of G_t , it follows that, for any $\varepsilon > 0$, the volume of $V_k := (G_t \setminus G_t^{(-\varepsilon)}) \cap C_k$ does not exceed $\varepsilon k^{m-1}\omega_{m-1}$. As the maximum value of the density of $\mathbf{W}_t$ on V_k less than or equal to the density's value on $\partial B_{k-1}(\mathbf{0})$,

we conclude that

$$\begin{aligned}
\mathbb{E}(1 + \|\mathbf{W}_t\|; \mathbf{W}_t \in G_t \setminus G_t^{(-\varepsilon)}) &= \sum_{k=1}^{\infty} \mathbb{E}(1 + \|\mathbf{W}_t\|; \mathbf{W}_t \in V_k) \\
&\leq \frac{\varepsilon \omega_{m-1}}{(2\pi t)^{m/2}} \sum_{k=1}^{\infty} (1+k) k^{m-1} e^{-(k-1)^2/2t} \\
&\leq \gamma \varepsilon
\end{aligned}$$

for some $\gamma = \gamma(m, t_0) < \infty$, as we only consider $t \in (t_0, 1)$.

Now turn to the case $t \in (0, t_0)$. By [G1], for such t one has $B_{\delta/2}(\mathbf{0}) \subset G_t$. Choosing $v_0 := \delta/4$ (which we can assume to be less than one without loss of generality), one can employ the same argument as above but with the spherical layers $C_k := B_{k+\delta/4}(\mathbf{0}) \setminus B_{k-1+\delta/4}(\mathbf{0})$, $k = 1, 2, \dots$, making use of the observation that the maximum value of the density of $\mathbf{W}_t$ on the “innermost layer” C_1 is bounded for $t \in (0, t_0)$. The corollary is proved. $\square$

Bibliography

- [AS64] ABRAMOWITZ, M. AND STEGUN, I. (1964). *Handbook of Mathematical Functions*. Dover, New York.
- [A86] AITCHISON, J. (1986). *The Statistical analysis of compositional data*. Chapman & Hall, London.
- [AKV12] AMBRUS, G., KEVEI, P. AND VIGH, V. (2012). The diminishing segment process. *Statist. Probab. Lett.* **82**, 191–195.
- [BS97] BAÑUELOS, R. AND SMITS, R. G. (1997). Brownian motion in cones. *Probab. Th. Rel. Fields* **108**, 299–319.
- [B54] BELLMAN, R. (1954). Limit theorems for non-commutative operators. I. *Duke Math. J.* **21**, 491–500.
- [BW02] BIALKOWSKI, M. AND WESOŁOWSKI, J. (2002). Asymptotic behavior of some random splitting schemes. *Probab. Math. Statist.* **22**, 181–191.
- [BS02] BORODIN, A. N. AND SALMINEN, P. (2002). *Handbook of Brownian motion—Facts and formulae*, 2nd edn. Birkhauser, Basel.
- [B98] BOROVKOV, A. A. (1998). *Ergodicity and Stability of Stochastic Processes*. Wiley, New York.
- [B94] BOROVKOV, K. A. (1994). On simulation of random vectors with given densities in regions and on their boundaries. *J. Appl. Probab.* **31**, 205–220.
- [BM12] BOROVKOV, K. AND MCKINLAY, S. (2012). The uniform law for sojourn measures of random fields. *Statist. Probab. Lett.* **82**, 1745–1749.

- [BN05] BOROVKOV, K. AND NOVIKOV, A. (2005). Explicit bounds for approximation rates of boundary crossing probabilities for the wiener process. *J. Appl. Prob.* **42**, 82–92.
- [BL85] BOUGEROL, P. AND LACROIX, J. (1985). *Products of random matrices with applications to Schrödinger operators*. Birkhäuser, Boston.
- [BJM10] BRUNEAU, L., JOYE, A. AND MERKLI, M. (2010). Infinite products of random matrices and repeated interaction dynamics. *Ann. Inst. Henri Poincaré Probab. Stat.* **46**, 442–464.
- [CL91] CHAMAYOU, J.-F. AND LETAC, G. (1991). Explicit stationary distributions for compositions of random functions and products of random matrices. *J. Theoret. Probab.* **4**, 3–36.
- [CL94] CHAMAYOU, J.-F. AND LETAC, G. (1994). A transient random walk on stochastic matrices with Dirichlet distributions. *Ann. Probab.* **22**, 424–430.
- [C00] CHAUMONT, L. (2000). An extension of Vervaat’s transformation and its consequences. *J. Theor. Prob.* **13**, 259–277.
- [CHY01] CHAUMONT, L., HOBSON, D.G. AND YOR, M. (2001). Some consequences of the cyclic exchangeability property for exponential functionals of Lévy processes. *Sém. Prob. XXXV*, **15**, 334–347.
- [CGZ84] CHEN, R., GOODMAN, R. AND ZAME, A. (1984). Limiting distributions of two random sequences. *J. Multivariate Anal.* **14**, 221–230.
- [CLZ81] CHEN, R., LIN, E. AND ZAME, A. (1981). Another arc sine law. *Sankhyā A* **43**, 371–373.
- [D87] DEBLASSIE, D. (1987). Exit times from cones in $\mathbb{R}^n$ of Brownian motion. *Probab. Th. Rel. Fields* **74**, 1–29.
- [DR63] DEGROOT, M. H. AND RAO, M. M. (1963). Stochastic give-and-take. *J. Math. Anal. Appl.* **7**, 489–498.

- [DLS86] DEVROYE, L., LETAC, G. AND SESHADRI, V. (1986). The limit behavior of an interval splitting scheme. *Statist. Probab. Lett.* **4**, 183–186.
- [DF99] DIACONIS, P. AND FREEDMAN, D. (1999). Iterated random functions. *SIAM Review* **41**, 45–76.
- [DNPR01] DI NARDO, E., NOBILE, A. G., PIROZZI, E. AND RICCIARDI, L. M. (2001). A computational approach to first-passage-time problems for Gauss-Markov processes. *Adv. Appl. Probab.* **33**, 453–482.
- [DB08] DOWNES, A. AND BOROVKOV, K. (2008). First passage densities and boundary crossing probabilities for diffusion processes. *Meth. Comp. Appl. Prob.* **10**, 621–644.
- [D98] DUFRESNE, D. (1998). Algebraic properties of Beta and Gamma distributions, and applications. *Adv. in Appl. Math.* **20**, 285–299.
- [F68] FELLER, W. (1968). *An introduction to probability theory. Vol. 1. 3rd edn.* Wiley, New York.
- [F71] FELLER, W. (1971). *An introduction to probability theory. Vol. 2. 2nd edn.* Wiley, New York.
- [FG95] FITZSIMMONS, P. J. AND GETOOR, R. K. (1995). Occupation time distributions for Lévy bridges and excursions. *Stoch. Proc. Appl.* **58**, 73–89.
- [FK60] FURSTENBERG, H. AND KESTEN, H. (1960). Products of random matrices. *Ann. Math. Statist.* **31**, 457–469.
- [GH06] GHORBEL, M. AND HUILLET, T. (2006). Interval Partitioning Arising from Size-Biased Picking. *Comm. Statist. Theory Methods* **35**, 1173–1187.
- [GS72] GIHMAN, I. I. AND SKOROHOD, A. V. (1972). *Stochastic differential equations*, Springer-Verlag.
- [G72] GOLBERG, M. A. (1978). Boundary and initial-value methods for solving Fredholm equations with semidegenerate kernels. *J. Optim. Theory Appl.* **24**, 89–131.

- [G63] GRENANDER, U. (1963). *Probabilities on Algebraic Structures*. Wiley, New York.
- [GN04] GUPTA, A. K. AND NADARAJAH, S. (2004). *Handbook of Beta Distribution and its Applications* Dekker, New York,
- [HM11] HÖGNÄS, G. AND MUKHERJEA, A. (2011). *Probability measures on semi-groups, 2nd edn*. Springer, New York.
- [HL81] HOLLEY, R. AND LIGGETT, T., M. (1981). Generalized potlatch and smoothing processes. *Z. Wahrsch. verw. Geb.* **55**, 165–196.
- [H12] HOMEI, H. (2012). Randomly weighted averages with beta random proportions. *Statist. Probab. Lett.* **82**, 1515–1520.
- [HHM13] HOMEI, H., HASANZADEH, L., MINA, G. (2013). Randomly Weighted Averages on Order Statistics. *Appl. Math. (Irvine)* **4**, 1340–1346.
- [IN03] IACUS, S. M., NEGRI, I. (2003). Estimating unobservable signal by Markovian noise induction. *Stat. Methods Appl.* **12**, 153–167.
- [I85] IYENGAR, S. (1985). Hitting lines with two-dimensional Brownian motion. *SIAM J. Math. Anal.* **45**, 983–989.
- [JK90a] JOHNSON, N. AND KOTZ, S. (1990). Randomly Weighted Averages: Some Aspects and Extensions *The American Statistician* **44**, 245–249.
- [JK90b] JOHNSON, N. AND KOTZ, S. (1990). Use of moments in deriving distributions and some characterizations *Math. Sci.* **15**, 42–52.
- [JK95] JOHNSON, N. AND KOTZ, S. (1995). Use of moments in studies of limit distributions arising from iterated random subdivisions of an interval. *Statist. Probab. Lett.* **24**, 111–119.
- [J05] JONES, M. C. (2005). The Uniform Mixture of Generalized Arc-Sine Distributions. *J. Korean Statist. Soc.* **34**, 35–38.
- [K51] KAC, M. (1951). On some connections between probability theory and differential and integral equations. *Proc. 2nd Berkeley Symp. Prob. Statist.* 189–215.

- [K08] KAHALE, N. (2008). Analytic crossing probabilities for certain barriers by Brownian motion. *Ann. Appl. Probab.* **18**, 1424–1440.
- [K99] KALLENBERG, O. (1999). Ballot theorems and sojourn laws for stationary processes. *Ann. Prob.* **27**, 2011–2019.
- [KS98] KARATZAS, I. AND SHREVE S. E. (1998). *Brownian motion and stochastic calculus*. Springer-Verlag.
- [K88] KENNEDY, D. (1988). A note on stochastic search methods for global optimization. *Adv. in Appl. Probab.* **20**, 476–478.
- [K93] KEROV, S. V. (1993). Transition probabilities for continual young diagrams and the Markov moment problem. *Funct. Anal. Appl.* **27**, 104–117.
- [K73] KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.
- [KS84] KESTEN, H. AND SPITZER, F. (1984). Convergence in distribution of products of random matrices. *Z. Wahrsch. verw. Geb.* **67**, 363–386.
- [KR97] KLAIN, D. A. AND ROTA, G.-C. (1997). *Introduction to geometric probability*. Cambridge University Press.
- [K96] KNIGHT, F. B. (1996). The uniform law for exchangeable and Lévy process bridges. *Hommage à P. A. Meyer et J. Neveu, Astérisque* **236**, 171–188.
- [KD04] KOTZ, S. AND VAN DORP, J. R. (2004). *Beyond beta: Other families of distributions with bounded support and applications*. World Scientific, Singapore.
- [L58] LAMPERTI, J. (1958). An occupation time theorem for a class of stochastic processes. *Trans. Amer. Math. Soc.* **88**, 380–387.
- [L86] LETAC, G. (1986). A contraction principle for certain Markov chains and its applications. *Contemp. Math.* **50**, 263–273.
- [L02] LETAC, G. (2002). Donkey walk and Dirichlet distributions. *Statist. Probab. Lett.* **57**, 17–22.

- [LS98] LETAC, G. AND SCARSINI, M. (1998). Random nested tetrahedra. *Adv. in Appl. Probab.* **30**, 619–627.
- [L39] LÉVY, P. (1939). Sur certains processus stochastiques homogènes. *Compositio Math.* **7**, 283–339.
- [L61] LI, C. (1961). *Human Genetics*. McGraw-Hill, New York.
- [LS81] LIGGETT, T. M. AND SPITZER, F. (1981). Ergodic theorems for coupled random walks and other systems with locally interacting components. *Z. Wahrsch. verw. Geb.* **56**, 443–468.
- [L55] LUKACS, E. (1955). A characterisation of the Gamma distribution. *Ann. Math. Statist.* **26**, 319–324.
- [LJ73] LUUS, R. AND JAAKOLA, T.H.I. (1973). Optimization by direct search and systematic reduction of the size of search region. *Amer. Inst. Chemical Eng. J.* **19**, 760–766.
- [MY06] MANSUY R. AND YOR M. (2006). *Random times and enlargements of filtrations in Brownian setting*. Springer-Verlag, Berlin.
- [M14] MCKINLAY, S. (2014). A characterisation of transient random walks on stochastic matrices with Dirichlet distributed limits. *J. Appl. Prob.* **51**, 542–555.
- [M15] MCKINLAY, S. (2015). On beta distributed limits of iterated linear random functions. Available at <http://arxiv.org/abs/1508.00600>. Preprint.
- [MB15] MCKINLAY, S. AND BOROVKOV, K. (2015). On approximation rates for boundary crossing probabilities for the multivariate Brownian motion process. *Commun. Stoch. Anal.* **9**, 267–282.
- [MB16] MCKINLAY, S. AND BOROVKOV, K. (2016). On explicit form of the stationary distributions for a class of bounded Markov chains. *J. Appl. Prob.* **53**, 231–243.

- [MP10] MORTERS, P. AND PERES, Y. (2010). *Brownian motion*. Cambridge University Press, Cambridge.
- [M11] MUIRHEAD, S. (2011). *Financial Geometry: Pricing multi-asset barrier options using the generalised reflection principle*. Honours Thesis, The University of Melbourne.
- [M00] MUKHERJEA, A. (2000). *Topics in products of random matrices*. Tata Institute of Fundamental Research, Mumbai.
- [N01] NEININGER, R. (2001). Rates of convergence for products of random stochastic 2×2 matrices. *J. Appl. Prob.* **38**, 799–806.
- [NTT11] NG, K. W., TIAN, G.-L. AND TANG, M.-L. (2011). *Dirichlet and Related Distributions: Theory, Methods and Applications*. John Wiley & Sons, Ltd., Chichester, UK.
- [NFK99] NOVIKOV, A., FRISHLING, V., AND KORDZAKHIA, N. (1999). Approximations of boundary crossing probabilities for a Brownian motion. *J. Appl. Prob.* **36**, 1019–1030.
- [O88] O’CONNOR, C. A., HOOK, A. AND ELMAGHRABY A. S. (1988). Simulations of random divisions of an interval *ACM SIGSIM Simulation Digest* **19**, Issue 3, 12–14.
- [O07] ØKSENDAL, B. (2007). *Stochastic differential equations*, 6th Ed. Springer.
- [P09] PACHECO-GONZALEZ, C. G. (2009). Ergodicity of a bounded Markov chain with attractiveness towards the centre. *Statist. Probab. Lett.* **79**, 2177–2181.
- [PZWGC10] PAISLEY, J., ZAAS, A., WOODS, C. W., GINSBURG, G. S., AND CARIN, L. (2010). A stick-breaking construction of the beta process. In *Proc. International Conference on Machine Learning*. Haifa, Israel.
- [PW08] PATIE, P. AND WINTER, C. (2008). First exit time probability for multi-dimensional diffusions: A PDE-based approach. *J. Comput. Appl. Math.* **222**, 42–53.

- [P37] PITMAN, E. J. G. (1937). The “closest” estimates of statistical parameters. *Math. Proc. Cambridge Philos. Soc.* **33**, 212–222.
- [PY92] PITMAN, J. AND YOR, M. (1992). Arcsine laws and interval partitions derived from a stable subordinator. *Proc. London Math. Soc.* **65**, 326–356.
- [PW01] PÖTZELBERGER, K. AND WANG L. (2001). Boundary crossing probability for Brownian motion. *J. Appl. Probab.* **38**, 152–164.
- [P09] PREATER, J. (2009). A species of voter model driven by immigration. *Statist. Probab. Lett.* **79**, 2131–2137.
- [RS95] RACHEV, S. AND SAMORODNITSKY, G. (1995). Limit laws for a stochastic process and random recursion arising in probabilistic modelling. *Adv. Appl. Prob.* **27**, 185–202.
- [RL10] RAMLI, M. A. AND LENG, G. (2010). The stationary probability density of a class of bounded Markov processes. *Adv. Appl. Probab.* **42**, 986–993.
- [R14] ROOZEGAR, R. (2014). Certain Family of Weighted Averages with Beta Random Weights. *J. Stat. Pro. Lett.* **1**, 19–22.
- [RS13] ROOZEGAR, R. AND SOLTANI, A. R. (2013). Weighted averages with random proportions that are jointly uniformly distributed over the unit simplex. *J. Stat. Comput. Simul.* , to appear.
- [R65] ROSENBLATT, M. (1965). Products of independent and identically distributed stochastic matrices *J. Math. Anal. Appl.* **11**, 1–10.
- [R71] ROSENBLATT, M. (1971). *Markov Processes. Structure and Asymptotic Behaviour*. Springer-Verlag, Berlin.
- [S94] SETHURAMAN, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica* **4**, 639–650.
- [SC76] SCHRACK, G. AND CHOIT, M. (1976). Optimized relative step size random searches. *Mathem. Progr.* **10**, 230–244.

- [SL92] SPAANS, R. AND LUUS, R. (1992). Importance of search-domain reduction in random optimization *J. Optim. Theor. Appl.* **75**, 635–638.
- [S95] SHIRYAEV, A. N. (1995). *Probability*, 2nd edn. Springer-Verlag, New York.
- [SH09] SOLTANI, A. R. AND HOMEI, H. (2009). Weighted averages with random proportions that are jointly uniformly distributed over the unit simplex. *Statist. Probab. Lett.* **79**, 1215–1218.
- [SR12] SOLTANI, A. R. AND ROOZEGAR, R. (2012). On distribution of randomly ordered uniform incremental weighted averages: Divided difference approach. *Statist. Probab. Lett.* **82**, 1012–1020.
- [S58] SPITZER, F. (1958). Some theorems concerning 2-dimensional brownian motion. *Trans. Amer. Math. Soc.* **87**, 187–197.
- [S12] STOYANOV, J. (2012). Inference problems involving moment determinacy distributions. *Comm. Statist. Theory Methods* **41**, 2864–2878.
- [SP08] STOYANOV, J. AND PACHECO-GONZALEZ, C. G. (2008). A class of Markov chains with beta ergodic distributions. *Math. Sci.* **33**, 110–119.
- [SP00] STOYANOV, J. AND PIRINSKY, C (2000). Random motions, classes of ergodic Markov chains and beta distributions. *Statist. Probab. Lett.* **50**, 293–304.
- [S75] SUN, T.-C. (1975). Limits of convolutions of probability measures on the set of 2×2 stochastic matrices. *Bull. Inst. Math. Sinica* **3**, 235–248.
- [T97] TAKÁCS, L. (1997). On the ballot theorems. In: *Advances in combinatorial methods and applications to probability and statistics*, N. Balakrishnan, ed., Birkhäuser, Boston, MA, 97–114.
- [T12] TOURI, B. (2012). *Product of random stochastic matrices and distributed averaging*. Springer, Berlin.
- [V86] VAN ASSCHE, W. (1986). Products of 2×2 stochastic matrices with random entries. *J. Appl. Probab.* **23**, 1019–1024.

- [V87] VAN ASSCHE, W. (1987). A random variable uniformly distributed between two independent random variables. *Sankhyā A* **49**, 207–211.
- [V79] VERVAAT, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Prob.* **11**, 750–783.
- [VKJ93] VOLODIN, N. A., KOTZ, S. AND JOHNSON. N. L. (1993). Use of moments in distribution theory: a multivariate case. *J. Multivariate Anal.* **46**, 112–119.
- [WP97] WANG L. AND PÖTZELBERGER. K. (1997). Boundary crossing probabilities for Brownian motion and general boundaries. *J. Appl. Probab.* **34**, 54–65.
- [Z03] ZABRINSKI, Z. B. (2003). *Stochastic adaptive search for global optimization*. Kluwer, Boston.