EXIT PROBABILITIES OF CONSTRAINED SIMPLE RANDOM WALKS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

KAMİL DEMİRBERK ÜNLÜ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
FINANCIAL MATHEMATICS

SEPTEMBER 2018

Approval of the thesis:

EXIT PROBABILITIES OF CONSTRAINED SIMPLE RANDOM WALKS

submitted by **KAMİL DEMİRBERK ÜNLÜ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Financial Mathematics Department, Middle East Technical University** by,

Prof. Dr. Ömür Uğur
Director, Graduate School of **Applied Mathematics** __________

Prof. Dr. Sevtap Selçuk Kestel
Head of Department, **Financial Mathematics** __________

Assoc. Prof. Dr. Ali Devin Sezer
Supervisor, **Financial Mathematics, METU** __________

Examining Committee Members:

Prof. Dr. Fatih Tank
Insurance and Actuarial Science, Ankara University __________

Assoc. Prof. Dr. Ali Devin Sezer
Financial Mathematics, METU __________

Prof. Dr. Sevtap Selçuk Kestel
Financial Mathematics, METU __________

Assoc. Prof. Dr. Ceylan Talu Yozgatlıgil
Statistics, METU __________

Assoc. Prof. Dr. Özge Sezgin Alp
Accounting and Financial Management, Başkent University __________

Date: __________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: KAMİL DEMİRBERK ÜNLÜ

Signature :

ABSTRACT

EXIT PROBABILITIES OF CONSTRAINED SIMPLE RANDOM WALKS

Ünlü, Kamil Demirberk
Ph.D., Department of Financial Mathematics
Supervisor : Assoc. Prof. Dr. Ali Devin Sezer

September 2018, 79 pages

Consider a nearest neighbor stable two dimensional random walk X constrained to remain on the positive orthant. X is assumed stable, i.e., its average increment points toward the origin. X represents the lengths of two queues (or two stacks in computer science applications) working in parallel. The probability p_n that the sum of the components of this random walk reaches a high level n before the random walk returns to the origin is a natural performance measure, representing the probability of a buffer overflow in a busy cycle. The stability of the walk implies that p_n decays exponentially in n. Let Y be the same constrained random walk as X, but constrained only on its second component and the jump probabilities on its first component reversed. The present thesis shows that one can approximate p_n with the probability that components of Y ever equal each other, with exponentially decaying relative error, if X starts from an initial point with nonzero first component. We further construct a class of Y-harmonic functions from single and conjugate points on a characteristic surface, with which the latter probability can be either computed perfectly in some cases, or approximated with bounded relative error in general. We provide numerical examples showing the effectiveness of the computed approximations and indicate possible applications of our results in finance and insurance.

Keywords: Approximation of probabilities of rare events, exit probabilities, constrained random walks, queueing systems, shared memory management, large deviations, credit risk, modeling of insurance and financial systems

ÖZ

KISITLI BASİT RASTGELE YÜRÜYÜŞLERİN ÇIKIŞ OLASILIKLARI

Ünlü, Kamil Demirberk
Doktora, Finansal Matematik Bölümü
Tez Yöneticisi : Doç. Dr. Ali Devin Sezer

Eylül 2018 , 79 sayfa

X iki boyutta en yakın komşularına geçerek hareket eden, pozitif koordinat düzlemine kısıtlı bir rastgele yürüyüş olsun. Bu yürüyüşün dengeli olduğu farz edilsin yani, artışlarının ortalaması orijin ($(0, 0)$ noktası) yönünde olsun. X, iki paralel kuyruk sistemindeki kuyruk uzunluklarını veya bilgisayar biliminde iki yığının uzunluğunu temsil etmektedir. p_n, rastgele yürüyüşün her iki bileşeninin toplamının orijine geri dönmeden n gibi büyük bir değere ulaşması olasılığını göstersin. p_n olasılığı bu sistemler için doğal bir performans ölçüsüdür ve yoğun bir döngüde taşma olasılığını ifade etmektedir. Rastgele yürüyüşün dengeli olmasından dolayı p_n olasılığı, n arttıkça üstel hızla sıfıra yakınsar. Y, X ile aynı özelliklere sahip fakat sadece ikinci bileşeni kısıtlı ve birinci bileşeninin artış olasılıkları yer değiştirmiş iki boyutlu rastgele yürüyüş olsun. Bu tez, X'in başlangıç noktasının birinci bileşeni 0 dan farklı seçildiğinde, p_n olasılığının, Y rastgele yürüyüşünün p_n'e karşılık gelen bir olasılığıyla yaklaşık olarak hesaplanabildiğini ve bu yaklaşık hesapta göreli hatanın üstel hızla sıfıra yakınsadığını göstermektedir. Ayrıca bu tezde bir karakteristik yüzey üzerindeki tek ve eşlenik noktalardan yola çıkarak Y-harmonik fonksiyonları oluşturulmuş ve bu fonksiyonlar kullanılarak Y'nin p_n'e karşılık gelen olasılığı, bazı durumlarda mükemmel şekilde ve genel olarak üstten sınırlı göreceli hata ile yaklaşık olarak hesaplanmıştır. Yapılan hesaplamaların etkinliğini gösteren sayısal örnekler verilmiş ve bu hesaplamaların finans ve sigortacılık sektörlerindeki olası uygulamalarından bahsedilmiştir.

Anahtar Kelimeler: Nadir olayların olasılıklarının yaklaşık hesabı, çıkış olasılıkları,

kısıtlı rastgele yürüyüş, kuyruk sistemleri, paylaşımlı bellek yönetimi, büyük sapmalar, kredi riski, sigorta ve finansal sistemlerin modellenmesi

To My Family

ACKNOWLEDGMENTS

Firstly, I would like to express my gratitude to my supervisor Assoc. Prof. Dr. Ali Devin Sezer for his continuous support, patience, motivation, and immense knowledge. It was a real privilege and an honour to work with him.

I am very thankful to my thesis progress committee members, Assoc. Prof. Dr. Ali Devin Sezer, Assoc. Prof. Dr. Ceylan Talu Yozgatlıgil, Prof. Dr. Sevtap Selçuk Kestel and Prof. Dr. Gerhard-Wilhelm Weber for their guiding comments and suggestions.

I would like to thank the thesis defense committee members, Prof. Dr. Fatih Tank, Assoc. Prof. Dr. Ali Devin Sezer, Prof. Dr. Sevtap Selçuk Kestel, Assoc. Prof. Dr. Ceylan Talu Yozgatlıgil, Assoc. Prof. Dr. Özge Sezgin Alp, Prof. Dr. Mehmet Yılmaz and Prof. Dr. Sibel Tarı for their suggestions, corrections and comments.

I would like to express my thanks to TÜBİTAK (Scientific and Technological Research Council of Turkey) for its financial support during my graduate education.

Special thanks to my wife Eda for her support, sacrifice and endless love. Also, I would like to thank my family, my mother Döne, my father Taner and my brother Asilcan for their continuous support, love and patience.

Furhermore, I would like to thank my friends and colleagues Fatma Başoğlu Kabran, Mustafa Hilmi Pekalp and Yasin Okkaoğlu for their friendship, support and help.

Lastly, I would like to thank all the members of the Institute of Applied Mathematics at METU and Department of Statistics at Ankara University who helped me directly or indirectly to accomplish this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

$\mathbb{C}$	Complex Numbers
I.S.	Importance Sampling
i.i.d.	Independent and Identically Distributed
$\mathbb{R}$	Real Numbers
$\mathfrak{R}(z)$	Real part of a complex number z
$\mathbb{Z}$	Integers
$\mathbb{Z}_+$	Positive Integers

CHAPTER 1

INTRODUCTION

A random walk is said to be constrained if its dynamics force it to remain in a given set. Constrained random walks can be used to model a great variety of systems from queueing systems, proccesses using a computer's memory, computer networks to networks of companies. The simplest constrained random walk arises from the modeling of a single queue depicted in Figure 1.1.

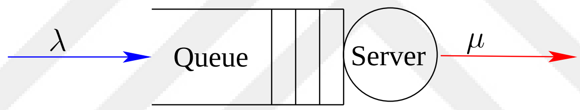

Figure 1.1: The $M/M/1$ queue

This is a system modeling the arrival of customers to a system to receive service; the arrival is Poisson with rate λ, the service rate is exponential with rate μ. There is no loss of generality in assuming that the rates add to 1; if not, one can renormalize them without changing any of the quantities that we will be working with, for this reason we will always assume that the rates add to 1. If we observe this system at arrivals and service completions we can model it using a constrained random walk S on the positive integers (see Figure 1.2). A jump to the right (occurring with probability λ) represents a customer arrival and a jump to the left (occurring with probability μ) represents a service completion; the walk represents the number customers at jump k. The constraining boundary for S is 0. When S hits 0 it cannot jump to the left anymore, i.e., when the system is empty no service can happen. In this discussion we have used the random walk S as a model of a queue; it can also be used to model the reserves or the equity of a single company, for example, as in [67, Chapter 5]. In that interpretation, hitting 0 corresponds to the default of the company.

Figure 1.2: The simple random walk

A stability condition associated with S is the following: $\lambda < \mu$, i.e., the system serves on average faster than the arrival rate. Stability is a desirable property, most systems are designed to be stable so that they can perform predictably and reliably. Throughout this thesis we will work with stable processes under assumptions similar to this.

S is mathematically very simple because it is single dimensional. Multidimensional models constrained random walks arise as models of systems with multiple components. The simplest of these are corresponds to two tandem queues (Figure 1.3) and two parallel queues (Figure 1.4).

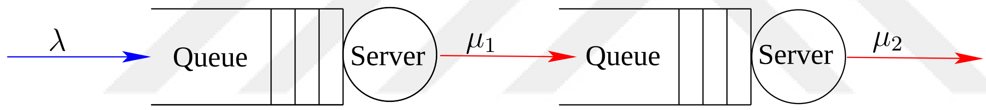

Figure 1.3: Two tandem network

In the two tandem queues there are two servers. Customer (or jobs / packages / claims depending on the application) arrive to the first server according to a Poisson process with rate λ then receive service at server 1 and wait for service in the second queue to receive service at server 2. The service times are exponentially distributed with rates μ_1 and μ_2, respectively.

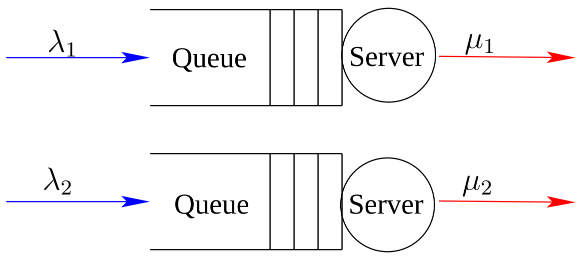

Figure 1.4: Two parallel queues

In two parallel queues, customers arrive at queues 1 and 2 with rates λ_1 and λ_2, then they are serviced at Server 1 and Server 2 with the rates μ_1 and μ_2.

The state of these systems at the k^{th} jump of the system are represented by a two dimensional constrained random walk X on the positive orthant $\mathbb{Z}_+^2$; the first [second] component represents the number of jobs in the first [second] queue. The

$$\partial_1 = \{x \in \mathbb{Z}^2 : x(1) = 0, x(2) > 0\}, \partial_2 = \{x \in \mathbb{Z}^2 : x(2) = 0, x(1) > 0\}$$

serve as the constraining boundaries; the meaning of the constraints is the same as in the case of a single queue: when a queue is empty no service can happen in that queue. The dynamics of the tandem walk is given in Figure 1.5; the dynamics of the parallel walk is given in Figure 1.6. The constrained random walk corresponding to the two parallel queues is also called a *simple* constrained random walk. In this thesis we will focus on this random walk.

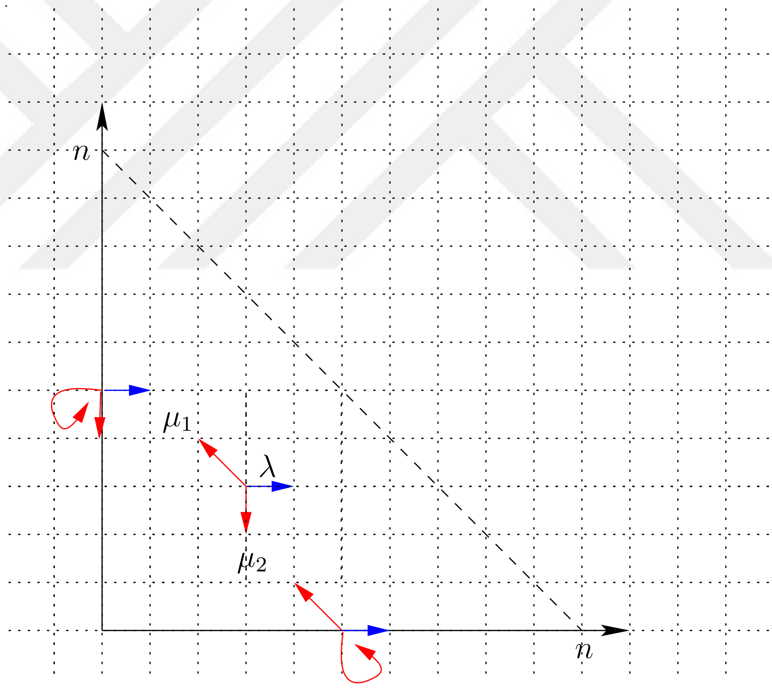

Figure 1.5: Two dimensional constrained random walk

The stability assumption for the tandem walk is $\lambda < \mu_1, \mu_2$; the stability assumption for the simple random walk is $\lambda_i < \mu_i$, $i = 1, 2$.

Define the stopping times

$$\tau_n \doteq \inf\{k > 0 : X_k(1) + X_k(2) = n\},$$

i.e., the first time the sum of the components of the walk X equals n. A natural performance measure associated with the walk X is the following probability:

$$p_n(x) \doteq P_x(\tau_n < \tau_0).$$

This probability can be represented geometrically as follows. Define the region

$$A_n = \left\{x \in \mathbb{Z}_+^2 : x(1) + x(2) \leq n\right\}, \tag{1.1}$$

and its boundary

$$\partial A_n = \left\{x \in \mathbb{Z}_+^2 : x(1) + x(2) = n\right\}. \tag{1.2}$$

p_n is then the probability that X hits 0 before hitting the diagonal line ∂A_n.

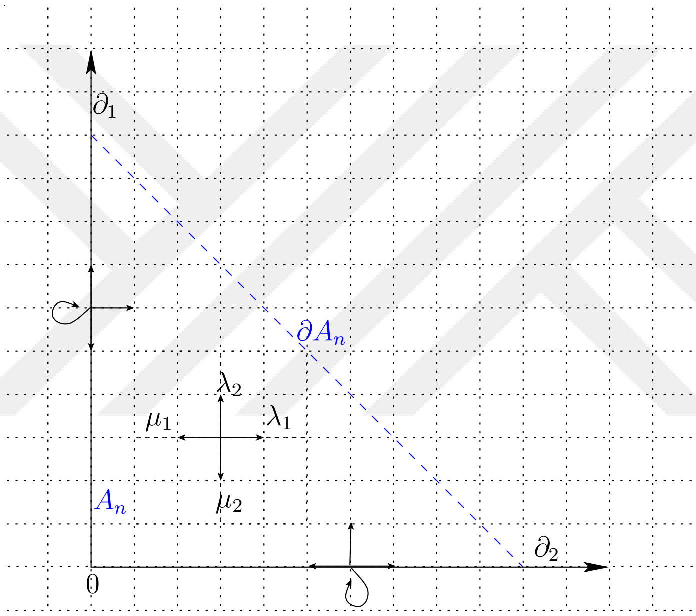

Figure 1.6: Dynamics of X, ∂A_n and A_n

Its stability implies that X moves, on average, towards the origin. This means that these systems work in cycles restarting each time X hits 0. The probability p_n has the following practical interpretation: suppose that the customers (or packets / jobs etc.) arriving at the system is stored in a joint buffer of size n. p_n is the probability that a cycle starting from point x ends with a buffer overflow, i.e., system failure.

In single dimension (i.e., for the constrained random walk S of Figure 1.2) the computation of p_n is trivial, one can find explicit formulas to it by solving the single dimensional recursive equation that it satisfies. As soon as one goes to two dimensions, the problem turns out to be difficult and its solution has received considerable attention over the last three decades. Even the effective computation of p_n via simulation proved difficult [33]. Asymptotically optimal importance sampling algorithms for it were constructed in [25]; more details about this study can be found in Chapter 5; see also [25] for further references on importance sampling for this probability. Recently [64] developed approximation formulas for p_n and proved that the relative error of these formulas converge to 0 exponentially *for the two tandem walk. The goal of this thesis to extend these formulas to the simple random walk and prove that the relative error for these formulas also go to* 0 *for the simple random walk.*

Although the main approach of the present work is parallel to that of [64, 65], many new challenges appear in the treatment of the two parallel queues and new ideas and techniques are required to meet these challenges; there are also differences in the assumptions made and the results obtained. We make a detailed comparison of these two cases in Section 6.1 of the Conclusion (Chapter 6).

In the rest of this introduction we explain the approach of this thesis to the approximation of the probability p_n and lay out our main results. This is done in Section 1.2. The further notation needed for this is given in the next section. The plan for the rest of the thesis is given in Section 1.3.

1.1 Definitions

Let us begin with a formal definition of the simple random walk. For this we need the constraining map π: $X \in \partial_i$ when ith queue is empty. Since server can only serves

$$\pi(x, v) \doteq \begin{cases} v, & \text{if } x + v \in \mathbb{Z}_+^2, \\ 0, & \text{otherwise,} \end{cases}$$

π constraints the random walk X on positive quadrant, $\mathbb{Z}_+^2$. Let I_k be an independent and identically distributed (iid) sequence taking values in $\{(1,0), (-1,0), (0,1), (0,-1)\}$. X, the constrained random walk, can be written as

$$X_0 = x \in \mathbb{Z}_+^2, \quad X_{k+1} \doteq X_k + \pi(X_k, I_k), k = 1, 2, 3, ...$$

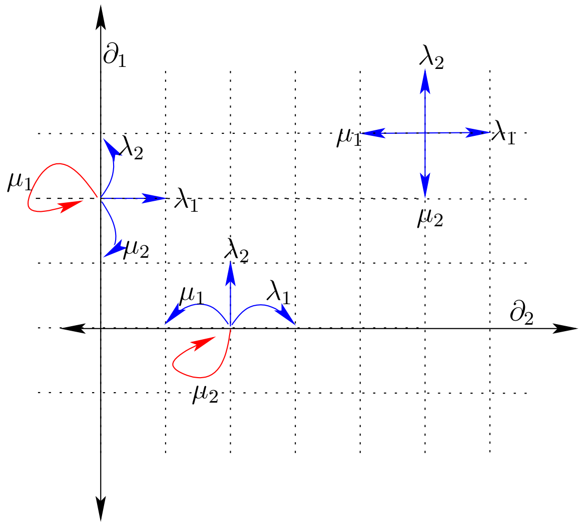

Figure 1.7: The increments of X and the reflecting boundaries

The utilization rates of the nodes are:

$$\rho_i = \frac{\lambda_i}{\mu_i}, i = 1, 2.$$

We assume that X is stable, i.e.,

$$\rho_1 < 1 \text{ and } \rho_2 < 1.$$

The system utilization rate:

$$r = \frac{\lambda_1 + \lambda_2}{\mu_1 + \mu_2}$$

plays a central role in our analysis. Without loss of generality we can assume

$$\rho_2 \leq r \leq \rho_1. \tag{1.3}$$

If this doesn't hold, we can rename the nodes so that this holds.

We will make two further technical assumptions:

$$\rho_1 \neq \rho_2, \quad \frac{r^2}{\rho_2} < 1. \tag{1.4}$$

The first of these is needed in the construction of the Y-harmonic functions in Section 2.1, see (2.6). The second is useful both in the computation of $P_y(\tau < \infty)$ (see the proof of Proposition 2.21) and in the limit analysis (see the proof of Proposition 2.8). We further comment on these assumptions in the Conclusion (Section 6.2).

Define the linear transformation

$$\mathcal{I} \doteq \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

and the affine transformation

$$T_n = ne_1 + \mathcal{I},$$

where (e_1, e_2) is the standard basis for $\mathbb{R}^2$. Furthermore, define the constraining map

$$\pi_1(x, y) = \begin{cases} y, & \text{if } x + y \in \mathbb{Z} \times \mathbb{Z}_+, \\ 0, & \text{otherwise.} \end{cases}$$

Define Y to be a constrained random walk on $\mathbb{Z} \times \mathbb{Z}_+$ with increments

$$J_k \doteq \mathcal{I} I_k, \tag{1.5}$$

$$Y_{k+1} = Y_k + \pi_1(Y_k, J_k).$$

Y has the same increments as X, but the probabilities of the increments e_i and $-e_i$ are reversed.

Define

$$B \subset \mathbb{Z} \times \mathbb{Z}_+,$$

$$B \doteq \{y : y(1) = y(2)\},$$

and the hitting time

$$\tau \doteq \inf \{k : Y_k \in B\}.$$

A function h on $\mathbb{Z} \times \mathbb{Z}_+$ is said to be Y-harmonic if

$$\mathbb{E}_y[h(Y_1)] = h(y), y \in \mathbb{Z} \times \mathbb{Z}_+.$$

1.2 Summary of our analysis

[64, Proposition 3.1] asserts, in a more general framework than the model given above, that for any $y \in \mathbb{Z}_+^2$, $y(1) > y(2)$, $P_{T_n(y)}(\tau_n < \tau_0) \to P_y(\tau < \infty)$. The approximation idea connecting these two probabilities is shown in Figure 1.8: by applying T_n, we move the origin of the coordinate system to $(n, 0)$ and take limits, which leads to the limit problem of computing $P_y(\tau < \infty)$ where the limit Y process is the same process as X (observed from the point $(n, 0)$) but not constrained on ∂_1.

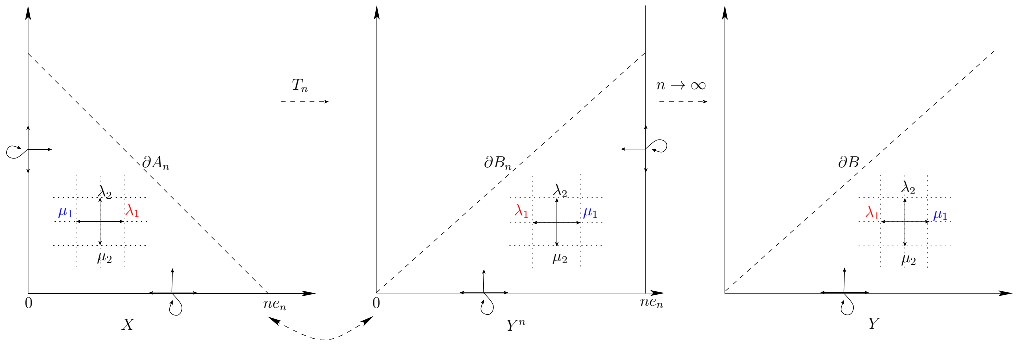

Figure 1.8: Transformations and the limit problem

A more interesting convergence analysis is when the initial point is given in x coordinates. A convergence analysis from this point of view has only been performed so far for the two tandem queues in [64, 65]. The goal of the present work is to extend this analysis to two parallel queues. Our main result is the following theorem:

Theorem 1.1. *For any* $x \in \mathbb{R}^2_+$, $x(1) + x(2) < 1$, $x(1) > 0$, *there exists* $C_7 > 0$ *and* $N > 0$ *such that*

$$\frac{|P_{x_n}(\tau_n < \tau_0) - P_{T_n(x_n)}(\tau < \infty)|}{P_{x_n}(\tau_n < \tau_0)} < e^{-C_7 n},$$

for $n > N$, *where* $x_n = \lfloor xn \rfloor$.

Thus, as n increases $P_{T_n(x_n)}(\tau < \infty)$ approximates $P_{x_n}(\tau_n < \tau_0)$ very well (with exponentially decaying relative error in n) if $x(1) > 0$. In the tandem case there is a simple explicit formula for $P_y(\tau < \infty)$. In the parallel walk case a simple explicit formula exists under the additional condition

$$\rho_1 \rho_2 = r^2. \tag{1.6}$$

The formula for $P_y(\tau < \infty)$ under this condition is

$$P_y(\tau < \infty) = r^{y(1)-y(2)} + \frac{(1-r)r}{r - \rho_2}\left(\rho_1^{y(1)} - r^{y(1)-y(2)}\rho_1^{y(2)}\right). \tag{1.7}$$

1.3 Plan of the thesis

Chapter 2 devoted to the approximation of p_n. The formula (1.7) is derived in Proposition 2.21 and is based on the class of Y-harmonic functions constructed in Section 2.1 from single and conjugate points on a characteristic surface associated with Y. A generalization of (1.7) can be used to find upper and lowerbounds for $P_y(\tau < \infty)$ when (1.6) doesn't hold, see Propositions 2.22 and 2.25. Subsection 2.6.1 illustrates how one can use these results to construct finer approximations of $P_y(\tau < \infty)$ with diminishing relative error using superposition of Y-harmonic functions defined by single and conjugate points on the characteristic surface.

Define the stopping times

$$\sigma_1 = \inf\{k : X_k \in \partial_1\}, \quad \bar{\sigma}_1 = \inf\{k : T_n(Y_k) \in \partial_1\}. \tag{1.8}$$

If we set the initial position of Y to $Y_0 = T_n(X_0)$, we have

$$\{\tau_n < \tau_0\} \cap \{\tau_n < \sigma_1 \wedge \tau_0\} = \{\tau < \infty\} \cap \{\tau < \bar{\sigma}_1 < \infty\}.$$

The main argument in the proof of Theorem 2.19 is this: most of the probability of the events $\{\tau_n < \tau_0\}$ and $\{\tau < \infty\}$ come from the events $\{\tau_n < \sigma_1 \wedge \tau_0\}$ and $\{\tau < \bar{\sigma}_1 < \infty\}$ respectively, if the initial position X_0 of X is away from ∂_1. The full implementation of this argument will require the following steps:

1. Construction of Y-harmonic functions, $Y - z$ harmonic functions and bounds on $\mathbb{E}_y[z^\tau 1_{\{\tau<\infty\}}]$ for $z > 1$ (Sections 2.1 and 2.2); (this step is mathematically one of the most novel aspects of the thesis, see the comparison in Section 6.1),
2. Large deviations (LD) analysis of $P_{x_n}(\tau_n < \tau_0)$ (Section 2.3),
3. LD analysis of $P_{x_n}(\sigma_1 < \tau_n < \tau_0)$ (Section 2.4),
4. LD analysis of $P_{x_n}(\bar{\sigma}_1 < \tau < \infty)$ (Subsection 2.4.1).

These steps are put together in Section 2.5. Section 2.6 treats the problem of computing $P_y(\tau < \infty)$ from the Y-harmonic functions of Section 2.1.

In computer science the constrained simple random walks are used to model multiple stacks running on a joint memory; the time τ_n represents a memory overflow. A quantity that is tightly connected to p_n that has been studied in the computer science context is the following expectation:

$$\mathbb{E}_x[\max(X_{\tau_n}(1), X_{\tau_n}(2))]. \tag{1.9}$$

This is the expectation of the size of the largest stack when a memory overflow occurs. The analysis of this expectation has received wide attention since [41] where the problem is introduced. In Chapter 3 we point out how the methods and approximations methods we develop for p_n can be used to analyze this expectation and we provide a numerical example that the approach does give accurate approximations for a range of initial points x. A rigorous analysis of this approximation, including proofs of convergence, remains for future work. Chapter 5 gives a literature review of some of the works treating the expectation above.

In insurance and finance applications constraints can represent dividend payments or short selling restrictions; hitting boundaries can represent default or total reserves hitting a threshold. Chapter 4 gives two examples from insurance and finance demonstrating these applications.

Chapter 5 is devoted to the literature on analysis of p_n. Firstly we summarize [25, 33, 36, 60], which are works on importance sampling and large deviations analysis of constrained random walks with iid increments arising from queueing networks. We further review [64, 65] whose transformation and approximation techniques we apply in this thesis to the calculation of p_n for the simple random walk. Lastly we consider the works [28] and [68], which study the expecation (1.9) under various assumptions on the jump distribution of the underlying walk.

Finally, Chapter 6, the Conclusion of our thesis, compares the analyses of the tandem and paralel constrained random walks and discusses directions for future research.

CHAPTER 2

APPROXIMATION OF THE EXIT PROBABILITIES

2.1 Harmonic functions of Y

Following [64], introduce the interior characteristic polynomial of Y:

$$\boldsymbol{p}(\beta,\alpha) \doteq \lambda_1 \frac{1}{\beta} + \mu_1\beta + \lambda_2 \frac{\alpha}{\beta} + \mu_2 \frac{\beta}{\alpha}.$$

and characteristic polynomial of Y on ∂_2:

$$\boldsymbol{p}_1(\beta,\alpha) \doteq \lambda_1 \frac{1}{\beta} + \mu_1\beta + \lambda_2 \frac{\alpha}{\beta} + \mu_2.$$

As in [64], we will construct Y-harmonic functions from solutions of $\boldsymbol{p} = 1$; the set of all solutions of this equation defines the characteristic surface

$$\mathcal{H} \doteq \{(\beta,\alpha) \in \mathbb{C}^2 : \boldsymbol{p}(\beta,\alpha) = 1\},$$

define, similarly, the characteristic surface for ∂_2:

$$\mathcal{H}_1 \doteq \{(\beta,\alpha) \in \mathbb{C}^2 : \boldsymbol{p}_1(\beta,\alpha) = 1\},$$

Multiplying both sides of $\boldsymbol{p} = 1$ by α transforms it to the quadratic equation

$$\alpha\left(\lambda_1 \frac{1}{\beta} + \mu_1\beta - 1\right) + \lambda_2 \frac{\alpha^2}{\beta} + \mu_2\beta = 0, \tag{2.1}$$

Define

$$\boldsymbol{\alpha}(\beta,\alpha) \doteq \frac{1}{\alpha}\frac{\beta^2}{\rho_2}; \tag{2.2}$$

if for a fixed β, α_1 and α_2 are distinct roots of (2.1), they will satisfy

$$\alpha_2 = \boldsymbol{\alpha}(\beta,\alpha_1),$$

by simple algebra; we will call the points $(\beta, \alpha_1) \in \mathcal{H}$ and $(\beta, \alpha_2) \in \mathcal{H}$ arising from such roots conjugate. Following [64] we refer to the function $\boldsymbol{\alpha}$ as the conjugator. An example of two conjugate points for the real section of the characteristic surface $\mathcal{H}$ for $\lambda_1 = 0.15$, $\lambda_2 = 0.2$, $\mu_1 = 0.25$, $\mu_2 = 0.4$; the end points of the dashed line are an example of a pair of conjugate points (β, α_1) and (β, α_2). Each such pair defines a Y-harmonic function, see Proposition 2.3, are shown in Figure 2.1.

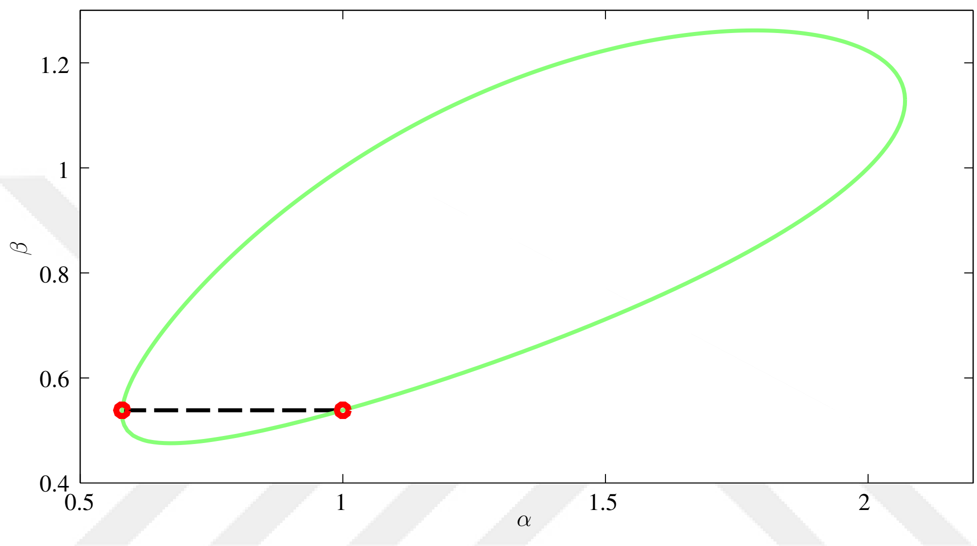

Figure 2.1: The real section of the characteristic surface $\mathcal{H}$

For any point $(\beta, \alpha) \in \mathcal{H}$ define the following $\mathbb{C}$-valued function on $\mathbb{Z}^2$:

$$z \mapsto [(\beta, \alpha), z], z \in \mathbb{Z}^2,$$

$$[(\beta, \alpha), z] \doteq \beta^{z(1)-z(2)} \alpha^{z(2)}.$$

Lemma 2.1. *$[(\beta, \alpha), \cdot]$ is Y-harmonic on $\mathbb{Z} \times \mathbb{Z}_+ - \partial_2$ when $(\beta, \alpha) \in \mathcal{H}$. In addition $x \mapsto [(\beta, \alpha), T_n(x)]$, $x \in \mathbb{Z}_+^2$, is X-harmonic on $\mathbb{Z}_+^2 - \partial_1 \cup \partial_2$.*

Proof. As in [64], the first claim follows from the definitions involved:

$$\begin{aligned}\mathbb{E}_z[(\beta,\alpha),Z_1] &= \mu_1\beta^{z(1)-z(2)+1}\alpha^{z(2)} + \lambda_2\beta^{z(1)-z(2)-1}\alpha^{z(2)+1}\\ &\quad + \lambda_1\beta^{z(1)-z(2)-1}\alpha^{z(2)} + \mu_2\beta^{z(1)-z(2)+1}\alpha^{z(2)-1}\\ &= \beta^{z(1)-z(2)}\alpha^{z(2)}\left(\lambda_1\frac{1}{\beta} + \mu_1\beta + \lambda_2\frac{\alpha}{\beta} + \mu_2\frac{\beta}{\alpha}\right)\\ &= \beta^{z_1-z_2}\alpha^{z_2}\boldsymbol{p}(\beta,\alpha) = [(\beta,\alpha),z].\end{aligned}$$

and the second claim follows from the first and the fact that $J_k = \mathcal{I}I_k$ (see (1.5)). □

Define

$$C(\beta,\alpha) \doteq \left(1-\frac{\beta}{\alpha}\right), (\beta,\alpha)\in\mathbb{C}^2, \alpha\neq 0.$$

Proceeding parallel to [64], one can define the following class of Y-harmonic functions from the functions $[(\beta,\alpha),\cdot]$:

Proposition 2.2. *Suppose* $(\beta,\alpha)\in\mathcal{H}\cap\mathcal{H}_1$. *Then* $[(\beta,\alpha),\cdot]$ *is* Y*-harmonic.*

Proof. Lemma 2.1 says that for $(\beta,\alpha)\in\mathcal{H}$, $[(\beta,\alpha),\cdot]$ is Y-harmonic on $\mathbb{Z}\times\mathbb{Z}_+ - \partial_2$. An argument parallel to the proof of Lemma 2.1 given as:

$$\begin{aligned}\mathbb{E}_z[(\beta,\alpha),Z_1] &= \mu_1\beta^{z(1)-z(2)+1}\alpha^{z(2)} + \lambda_2\beta^{z(1)-z(2)-1}\alpha^{z(2)+1}\\ &\quad + \lambda_1\beta^{z(1)-z(2)-1}\alpha^{z(2)} + \mu_2\beta^{z(1)-z(2)}\alpha^{z(2)}\\ &= \beta^{z(1)-z(2)}\alpha^{z(2)}\left(\lambda_1\frac{1}{\beta} + \mu_1\beta + \lambda_2\frac{\alpha}{\beta} + \mu_2\right)\\ &= \beta^{z_1-z_2}\alpha^{z_2}\boldsymbol{p}_1(\beta,\alpha) = [(\beta,\alpha),z].\end{aligned}$$

$[(\beta,\alpha),\cdot]$ is Y-harmonic on ∂_2 when $(\beta,\alpha)\in\mathcal{H}_1$. These two facts imply the statement of the proposition. □

The next proposition gives us another class of Y-harmonic functions constructed from conjugate points on $\mathcal{H}$, it is a special case of [64, Proposition 4.9]:

Proposition 2.3. *Suppose* $(\beta,\alpha_1)\neq(\beta,\alpha_2)$*, are conjugate points on* $\mathcal{H}$*. Then*

$$h_\beta \doteq C(\beta,\alpha_2)[(\beta,\alpha_1),\cdot] - C(\beta,\alpha_1)[(\beta,\alpha_1),\cdot]$$

is Y-harmonic.

For sake of completeness and easy reference, let us reproduce the argument given in the proof of [64, Proposition 4.9]:

Proof. That h_β is Y-harmonic on $\mathbb{Z} \times \mathbb{Z}_+ - \partial_2$ follows from Lemma 2.1. For $y \in \partial_2$ a direct computation gives:

$$
\begin{aligned}
&\mathbb{E}_y[[(\beta, \alpha_i), y + \pi_1(y, J_1)]] - [(\beta, \alpha_i), y] \\
&= \mu_1 \beta^{y(1)+1} + \lambda_1 \beta^{y(1)-1} + \lambda_2 \beta^{y(1)-1}\alpha + \mu_2 \beta^{y(1)} - \beta^{y(1)} \\
&= \beta^{y(1)} \left(\mu_1 \beta + \lambda_1 \frac{1}{\beta} + \lambda_2 \frac{\alpha}{\beta} + \mu_2 - 1 \right) \\
&= \beta^{y(1)} \left(\mu_1 \beta + \lambda_1 \frac{1}{\beta} + \lambda_2 \frac{\alpha}{\beta} + \mu_2 \frac{\beta}{\alpha} - \mu_2 \frac{\beta}{\alpha} + \mu_2 - 1 \right) \\
&= \beta^{y(1)} \left(1 - \mu_2 \frac{\beta}{\alpha} + \mu_2 - 1 \right) \\
&= \beta^{y(1)} \left(\mu_2 \left(1 - \frac{\beta}{\alpha} \right) \right) \\
&= C(\beta, \alpha_i) \beta^{y(1)}.
\end{aligned}
$$

It follows that

$$
\mathbb{E}_y[h_\beta(y + \pi_1(y, J_1))] - h_\beta(y) = C(\beta, \alpha_1) C(\beta, \alpha_2)(\beta^{y(1)} - \beta^{y(1)}) = 0,
$$

i.e., h_β is Y-harmonic on ∂_2 as well. □

The intersection of $\mathcal{H}$ and $\mathcal{H}_1$ consists of the points $(0, 0)$, $(1, 1)$ and (ρ_1, ρ_1). The last of these gives us our first nontrivial loglinear Y-harmonic function:

Lemma 2.4. $[(\rho_1, \rho_1), \cdot]$ *is Y-harmonic.*

The proof follows from Proposition 2.2 and the fact that $(\rho_1, \rho_1) \in \mathcal{H} \cap \mathcal{H}_1$. Fixing $\beta \in \mathbb{C}$ and solving (2.1) gives us the two conjugate points corresponding to β. It is also natural to start the computation from a fixed α and find its β and its conjugate. For this, one rewrites $\boldsymbol{p} = 1$, now as a polynomial in β:

$$
\left(\mu_1 + \frac{\mu_2}{\alpha} \right) \beta^2 - \beta + \lambda_1 + \lambda_2 \alpha = 0. \tag{2.3}
$$

For α fixed, the roots of (2.3) are

$$\beta_1(\alpha) = \frac{1+\sqrt{\Delta(\alpha)}}{2\left(\frac{\mu_2}{\alpha}+\mu_1\right)}, \quad \beta_2(\alpha) = \frac{1-\sqrt{\Delta(\alpha)}}{2\left(\frac{\mu_2}{\alpha}+\mu_1\right)}, \tag{2.4}$$

where

$$\Delta(\alpha) = 1 - 4\left(\frac{\mu_2}{\alpha}+\mu_1\right)(\lambda_1+\lambda_2\alpha),$$

and for $z \in \mathbb{C}$, $\sqrt{z}$ is the square root of z satisfying $\Re(\sqrt{z}) \geq 0$.

The function $y \mapsto P_y(\tau < \infty)$ takes the value 1 on ∂B; therefore, of special significance to us is the solution of (2.3) with $\alpha = 1$. The roots (2.4) for $\alpha = 1$ are

$$\beta_1(1) = \frac{1+\sqrt{(2\lambda_1+2\lambda_2-1)^2}}{2(\mu_1+\mu_2)} = \frac{2(\lambda_1+\lambda_2)}{2(\mu_1+\mu_2)} = r,$$

$$\beta_2(1) = \frac{1-\sqrt{(2\lambda_1+2\lambda_2-1)^2}}{2(\mu_1+\mu_2)} = \frac{2(1-\lambda_1-\lambda_2)}{2(\mu_1+\mu_2)} = 1.$$

That $r \leq \rho_1 < 1$ implies $C(r,1) = (1-r) \neq 0$. The assumption $\rho_1 \neq \rho_2$ implies

$$C(r, \boldsymbol{\alpha}(r,1)) = 1 - \frac{r}{\boldsymbol{\alpha}(r,1)} = 1 - r\frac{\rho_2}{r^2} = 1 - \frac{\rho_2}{r} \neq 0.$$

Therefore, by Proposition 2.3, the root $\beta_1 = r$ above defines the Y-harmonic function

$$\begin{aligned} h_r &= C(r, \boldsymbol{\alpha}(r,1))[(r,1),\cdot] - C(r,1)[(r,\boldsymbol{\alpha}(r,1)),\cdot] \\ &= (1-\rho_2/r)[(r,1),\cdot] - (1-r)[(r,r^2/\rho_2),\cdot]. \end{aligned}$$

For this function to be useful in our analysis, we need $r^2/\rho_2 < 1$ (see Proposition 2.20), therefore, we assume:

$$\frac{r^2}{\rho_2} < 1. \tag{2.5}$$

The scalar multiple of h_r is frequently used in the calculations, therefore, we will denote it in bold thus:

$$\boldsymbol{h}_r = \frac{1}{1-\rho_2/r}h_r = [(r,1),\cdot] - \frac{1-r}{1-\rho_2/r}[(r,r^2/\rho_2),\cdot]; \tag{2.6}$$

the assumption $\rho_1 \neq \rho_2$ ensures that the denominator $1-\rho_2/r$ is nonzero.

2.2 Laplace transform of τ

To bound approximation errors we will have to argue that we can truncate time without losing much probability. For this, it will be useful to know that there exists $z > 1$ such that

$$\mathbb{E}_y \left[z^\tau 1_{\{\tau<\infty\}} \right] < \infty. \tag{2.7}$$

In [64, 25], bounds similar to this are obtained using large deviations arguments, which are based on the ergodicity of the underlying chain. In [59], again a similar bound is obtained invoking the geometric ergodicity of the underlying process. The process underlying (2.7) is not stationary. For this reason, these arguments do not immediately generalize to the analysis of (2.7). To prove the existence of $z > 1$ such that (2.7) holds, we will extend the characteristic surface an additional dimension to include a new parameter; points on the generalized surface will correspond to discounted (in our case we are in fact interested in inflated costs) expected cost functions of the process Y, i.e., points on this surface will give us functions of the form

$$\mathbb{E}_y \left[z^\tau g(\tau) 1_{\{\tau<\infty\}} \right].$$

We will use these functions to find our desired z.

2.2.1 $1/z$-level characteristic surfaces and Y-z-harmonic functions

The development in this subsection is parallel to Section 2.1 with an additional variable $z \in \mathbb{C}$. A function h on $\mathbb{Z} \times \mathbb{Z}_+$ is said to be $Y - z$-harmonic if

$$z\mathbb{E}_y[h(Y_1)] = h(y), y \in \mathbb{Z} \times \mathbb{Z}_+.$$

Let as before $\boldsymbol{p}$ denote the characteristic polynomial of Y; the set of all solutions of the equation $z\boldsymbol{p} = 1$ defines the $1/z$-level characteristic surface

$$\mathcal{H}^z \doteq \{(\beta, \alpha) \in \mathbb{C}^2 : z\boldsymbol{p}(\beta, \alpha) = 1\}.$$

Similarly, define

$$\mathcal{H}_1^z \doteq \{(\beta, \alpha) \in \mathbb{C}^2 : z\boldsymbol{p}_1(\beta, \alpha) = 1\},$$

the $1/z$-level characteristic surface on ∂_1. These surfaces reduce to the ordinary characteristic surfaces when $z = 1$.

Multiplying both sides of $z\boldsymbol{p} = 1$ by $\frac{\alpha}{z}$ transforms it to the quadratic (in α) equation

$$\alpha\left(\lambda_1\frac{1}{\beta} + \mu_1\beta - \frac{1}{z}\right) + \alpha^2\frac{\lambda_2}{\beta} + \mu_2\beta = 0,$$

whose discriminant is

$$\Delta_z(\beta) \doteq \left(\lambda_1\frac{1}{\beta} + \mu_1\beta - \frac{1}{z}\right)^2 - 4\lambda_2\mu_2.$$

Let $\boldsymbol{\alpha}$ be the conjugator defined in (2.2). If $(\beta, \alpha_1) \in \mathcal{H}_z$ and $\alpha \neq 0$ then $(\beta, \alpha_2, z) \in \mathcal{H}_z$ for $\alpha_2 = \boldsymbol{\alpha}(\beta, \alpha_1)$; if $\Delta_z(\beta) \neq 0$ (β, α_1) and (β, α_2) will be distinct points on $\mathcal{H}_z$ and we will call them conjugate.

Lemma 2.5. $[(\beta, \alpha), \cdot]$ *is* $Y - z$ *harmonic on* $\mathbb{Z} \times \mathbb{Z}_+ - \partial_2$ *when* $(\beta, \alpha) \in \mathcal{H}_z$. *In addition* $x \mapsto [(\beta, \alpha), T_n(x)]$, $x \in \mathbb{Z}_+^2$, *is* $X - z$*-harmonic on* $\mathbb{Z}_+^2 - \partial_1 \cup \partial_2$.

Proof. The proof is parallel to that of Lemma 2.1.

$$\begin{aligned} z\mathbb{E}_y[(\beta, \alpha), Y_1] &= z\mu_1\beta^{y(1)-y(2)+1}\alpha^{y(2)} + z\lambda_2\beta^{y(1)-y(2)-1}\alpha^{y(2)+1} \\ &\quad + z\lambda_1\beta^{y(1)-y(2)-1}\alpha^{y(2)} + z\mu_2\beta^{y(1)-y(2)+1}\alpha^{y(2)+1} \\ &= \beta^{y(1)-y(2)}\alpha^{y(2)}\left(z\mu_1\beta + z\lambda_2\frac{\alpha}{\beta} + z\lambda_1\frac{1}{\beta} + z\mu_2\beta\alpha\right) \\ &= \beta^{y(1)-y(2)}\alpha^{y(2)}z\boldsymbol{p}(\beta, \alpha) = [(\beta, \alpha), y]. \end{aligned}$$

□

Define

$$C_z(\beta, \alpha) \doteq z\left(1 - \frac{\beta}{\alpha}\right), (\beta, \alpha) \in \mathbb{C}^2, \alpha \neq 0.$$

Parallel to Section 2.1, the above definitions give us the following class of $Y - z$-harmonic functions.

Proposition 2.6. *Suppose* $(\beta, \alpha) \in \mathcal{H}^z \cap \mathcal{H}_1^z$. *Then* $[(\beta, \alpha), \cdot]$ *is* $Y - z$*-harmonic.*

Proof. Lemma 2.5 says that for $(\beta, \alpha) \in \mathcal{H}^z$, $[(\beta, \alpha), \cdot]$ is $Y - z$-harmonic on $\mathbb{Z} \times$

${\mathbb{Z}^2}_+ - \partial_2$. An argument parallel to the proof of Lemma 2.5 given as:

$$\begin{aligned} z\mathbb{E}_y[(\beta,\alpha),Y_1] &= z\mu_1\beta^{y(1)-y(2)+1}\alpha^{y(2)} + z\lambda_2\beta^{y(1)-y(2)+1}\alpha^{y(2)-1} \\ &+ z\lambda_1\beta^{y(1)-y(2)-1}\alpha^{y(2)} + z\mu_2\beta^{y(1)-y(2)}\alpha^{y(2)} \\ &= \beta^{y(1)-y(2)+1}\alpha^{y(2)}\left(z\mu_1\beta + z\lambda_2\frac{\beta}{\alpha} + z\frac{\lambda_1}{\beta} + z\mu_2\right) \\ &= \beta^{y(1)-y(2)}\alpha^{y(2)} z\boldsymbol{p}_1(\beta,\alpha) = [(\beta,\alpha),y]. \end{aligned}$$

□

Proposition 2.7. *Suppose* $(\beta,\alpha_1) \neq (\beta,\alpha_2)$, *are conjugate points on* $\mathcal{H}^z$. *Then*

$$h_{z,\beta} \doteq C_z(\beta,\alpha_2)[(\beta,\alpha_1),\cdot] - C_z(\beta,\alpha_1)[(\beta,\alpha_2),\cdot] \tag{2.8}$$

is $Y - z$*-harmonic.*

Proof. That $h_{z,\beta}$ is $Y - z$-harmonic on $\mathbb{Z}\times\mathbb{Z}_+ - \partial_2$ follows from Lemma 2.5. For $y \in \partial_2$ a direct computation gives:

$$\begin{aligned} & z\mathbb{E}_y[[(\beta,\alpha_i), y + \pi_1(y,J_1)]] - [(\beta,\alpha_i),y] \\ &= z\mu_1\beta^{y(1)+1} + z\lambda_1\beta^{y(1)-1} + z\lambda_2\beta^{y(1)-1}\alpha + z\mu_2\beta^{y(1)} - z\beta^{y(1)} \\ &= \beta^{y(1)}\left(z\mu_1\beta + z\lambda_1\frac{1}{\beta} + z\lambda_2\frac{\alpha}{\beta} + z\mu_2 - 1\right) \\ &= \beta^{y(1)}\left(z\mu_1\beta + z\lambda_1\frac{1}{\beta} + z\lambda_2\frac{\alpha}{\beta} + z\mu_2\frac{\beta}{\alpha} - z\mu_2\frac{\beta}{\alpha} + z\mu_2 - 1\right) \\ &= \beta^{y(1)}\left(1 - z\mu_2\frac{\beta}{\alpha} + z\mu_2 - 1\right) \\ &= z\beta^{y(1)}\left(\mu_2\left(1 - \frac{\beta}{\alpha}\right)\right) \\ &= zC(\beta,\alpha_i)\beta^{y(1)}. \end{aligned}$$

It follows that

$$\mathbb{E}_y[h_{z,\beta}(y + \pi_1(y,J_1)] - h_\beta(y) = zC(\beta,\alpha_1)C(\beta,\alpha_2)(\beta^{y(1)} - \beta^{y(1)}) = 0,$$

i.e., $h_{z,\beta}$ is Y-harmonic on ∂_2 as well. □

2.2.2 Existence of the Laplace transform of τ

We next use the $Y - z$ harmonic functions constructed in Propositions 2.6 and 2.7 to get our existence result.

Proposition 2.8. *There exist $z_0 > 1$ and C_1 such that*

$$\mathbb{E}_y[z_0^\tau 1_{\{\tau<\infty\}}] < C_1, \tag{2.9}$$

for all $y \in \mathbb{Z} \times \mathbb{Z}_+, y(1) \geq y(2)$.

Proof. Let us first prove the following: if we can find, for some $z_0 > 1$ and $C_1 > 0$, a Y-z_0 harmonic function h satisfying $h(y) \geq 1$ on ∂B and $C_1 > h \geq 0$ on B we are done. The reason is as follows: that h is $Y - z_0$-harmonic and the optional sampling theorem imply that $h(Y_{\tau\wedge n})z_0^{\tau\wedge n}$ is a martingale. It follows that

$$h(y) = \mathbb{E}_y[h(Y_{\tau\wedge n})z_0^{\tau\wedge n}],$$

for $y \in B$. Decompose the last expectation to $\{\tau \leq n\}$ and $\{\tau > n\}$:

$$h(y) = \mathbb{E}_y[h(Y_\tau)z_0^\tau 1_{\{\tau\leq n\}}] + \mathbb{E}_y[h(Y_n)z_0^n 1_{\{\tau>n\}}]].$$

That $h \geq 0$ on B implies

$$h(y) \geq \mathbb{E}_y[h(Y_{\tau\wedge n})z_0^\tau 1_{\{\tau\leq n\}}].$$

Now $\lim_{n\to\infty} h(Y_\tau)z_0^\tau 1_{\{\tau\leq n\}} = h(Y_\tau)z_0^\tau 1_{\{\tau<\infty\}}$. This and Fatou's lemma imply

$$h(y) \geq \mathbb{E}_y[h(Y_\tau)z_0^\tau 1_{\{\tau<\infty\}}]$$

Finally, $h \geq 1$ on ∂B and $h \leq C_1$ give (2.9). To get our desired h we start from the points $(r, 1)$ and (ρ_1, ρ_1) on $\mathcal{H}$. The first point gives us the root $(1, r)$ of the equation

$$z\boldsymbol{p}(\beta, 1) = 1.$$

That

$$\begin{aligned}
\frac{\partial z\boldsymbol{p}(\beta,\alpha)}{\partial\beta}|_{(1,r,1)} &= z\left(-\lambda_1\frac{1}{\beta^2} + \mu_1 - \lambda_2\alpha\frac{1}{\beta^2} + \mu_2\frac{1}{\alpha}\right)|_{(1,r,1)} \\
&= -\lambda_1\frac{1}{r^2} + \mu_1 - \lambda_2\frac{1}{r^2} + \mu_2 \\
&= -\frac{1}{r^2}(\lambda_1 + \lambda_2) + \frac{1}{r}(\lambda_1 + \lambda_2) \\
&= (\lambda_1 + \lambda_2)\frac{r-1}{r^2} \neq 0.
\end{aligned}$$

and the implicit function theorem gives us a differentiable function $\boldsymbol{\beta}_1$ on an open interval I_1 around $z = 1$ that satisfies

$$z\boldsymbol{p}(\boldsymbol{\beta}_1(z), 1) = 1, z \in I_1, \boldsymbol{\beta}_1(1) = r.$$

The conjugate of $(\boldsymbol{\beta}_1, 1)$ on $\mathcal{H}^z$ is $(\boldsymbol{\beta}_1, \boldsymbol{\alpha}(\boldsymbol{\beta}_1, 1))$ (whenever possible, we will omit the z variable and simply write $\boldsymbol{\beta}_1$; similarly, we will write $\boldsymbol{\alpha}$ for $\boldsymbol{\alpha}(\boldsymbol{\beta}_1, 1)$). These points give us the $Y - z$-harmonic function

$$\begin{aligned}\boldsymbol{h}_z &= C(\boldsymbol{\beta}_1, \boldsymbol{\alpha})[(\boldsymbol{\beta}_1, 1), \cdot] - C(\boldsymbol{\beta}_1, 1)[(\boldsymbol{\beta}_1, \boldsymbol{\alpha}), \cdot] \\ &= \left(1 - \frac{\rho_2}{\boldsymbol{\beta}_1}\right)[(\boldsymbol{\beta}_1, 1), \cdot] - (1 - \boldsymbol{\beta}_1)[(\boldsymbol{\beta}_1, \boldsymbol{\alpha}), \cdot] \\ &= [(\boldsymbol{\beta}_1, 1), \cdot] - \frac{1 - \boldsymbol{\beta}_1}{1 - \rho_2/\boldsymbol{\beta}_1}[(\boldsymbol{\beta}_1, \boldsymbol{\alpha}), \cdot].\end{aligned}$$

That $\boldsymbol{\alpha}(\boldsymbol{\beta}_1(1), 1) = 0 < r^2/\rho_2 < 1$ (Assumption 2.5)implies that $0 < \boldsymbol{\alpha}(\boldsymbol{\beta}_1(z), 1) < 1$ if we choose $z > 1$ close enough to 1. $\boldsymbol{h}_z$ will almost serve as our h, except that it does take negative values on a small section of B. To get a positive function we will add to $\boldsymbol{h}_z$ a constant multiple of the $Y - z$-harmonic function defined by the point on $\mathcal{H}^z \cap \mathcal{H}_2^z$ that is the continuation of (ρ_1, ρ_1) on $\mathcal{H}$. This point is $(\boldsymbol{\beta}_2(z), \boldsymbol{\beta}_2(z))$ where $\boldsymbol{\beta}_2(z)$ is the root of the the equation

$$\left(\frac{\lambda_1}{\beta} + \mu_1\beta + \lambda_2 + \mu_2\right) = \frac{1}{z};$$

satisfying $\boldsymbol{\beta}_2(1) = \rho_1$. The implicit function theorem (or direct calculation) shows that $\boldsymbol{\beta}_2$ is smooth in an open interval I_2 containing 1. Now $(\boldsymbol{\beta}_2, \boldsymbol{\beta}_2) \in \mathcal{H}^z \cap \mathcal{H}_2^z$ and Proposition 2.6 imply that $[(\boldsymbol{\beta}_2, \boldsymbol{\beta}_2), \cdot] \geq 0$ is a $Y - z$ harmonic function. Now define

$$h' \doteq \boldsymbol{h}_z + C_0[(\boldsymbol{\beta}_2, \boldsymbol{\beta}_2), \cdot].$$

By its definition h' is $Y - z$ harmonic. We would like to choose C_0 large enough so that h' is bounded below by 1 on ∂B and is nonnegative on B. By our assumption (1.3), $\boldsymbol{\beta}_1(1) = r < \boldsymbol{\beta}_2(1) = \rho_1$; therefore, for $z > 1$ close enough to 1, we will still have $\boldsymbol{\beta}_1(z) < \boldsymbol{\beta}_2(z)$; let us assume that I_1 and I_2 are tight enough that this holds. By definition,

$$h'(y) = \boldsymbol{\beta}_1^{y(1)-y(2)}\left(1 - \frac{1 - \boldsymbol{\beta}_1}{1 - \rho_2/\boldsymbol{\beta}_1}\boldsymbol{\alpha}^{y(2)}\right) + C_0\boldsymbol{\beta}_2^{y(1)-y(2)}\boldsymbol{\beta}_2^{y(2)}.$$

$\boldsymbol{\beta}_2 > \boldsymbol{\beta}_1$ implies that h' takes its most negative value for $y(1) = y(2)$, i.e., on ∂B and if we can choose $C_0 > 0$ so that h is nonnegative on ∂B, it will be so on all of B. On ∂B, h' reduces to

$$1 - \frac{1-\boldsymbol{\beta}_1}{1-\rho_2/\boldsymbol{\beta}_1}\boldsymbol{\alpha}^{y(2)} + C_0\boldsymbol{\beta}_2^{y(2)}.$$

If $\boldsymbol{\alpha} \le \boldsymbol{\beta}_2$, then setting $C_0 = \frac{\boldsymbol{\beta}_1 - 1}{1-\rho_2/\boldsymbol{\beta}_1}$ would imply $h' \ge 1$ on ∂B; $\boldsymbol{\alpha}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in (0,1)$ imply

$$h' \le 1 + C_0;$$

then, h' can serve as our desired $Y - z$ harmonic function h with $C_1 = 1 + C_0$.

Now let us consider the case $\boldsymbol{\alpha} > \boldsymbol{\beta}_2$:ordinary calculus implies that if we choose C_0 large enough we can make the minimum $m_0 < 0$ over $y(2) > 0$ of

$$-\frac{1-\boldsymbol{\beta}_1}{1-\rho_2/\boldsymbol{\beta}_1}\boldsymbol{\alpha}^{y(2)} + C_0\boldsymbol{\beta}_2^{y(2)};$$

arbitrarily close to 0; then choosing $h = \frac{1}{1+m_0}h'$ gives us a $Y - z$ harmonic function that satisfies $h \ge 1$ on ∂B, $h \ge 0$ on B and $h \le C_1$ on B where $C_1 = \frac{1}{1+m_0}(1 + C_0)$. □

We now use (2.9) to derive an upper bound on the probability that τ is finite but takes a too long time:

Proposition 2.9. *For any $\delta > 0$, there exists $C_2 > 0$ such that*

$$P_y(nC_2 < \tau < \infty) \le e^{-\delta n}, \tag{2.10}$$

for any $y \in \mathbb{Z} \times \mathbb{Z}_+$, $y(1) > y(2)$ and $n > 1$.

Proof. Let $z_0 > 1$ and C_1 be as in (2.9). For any $A > 0$, Chebyshev's inequality gives

$$\begin{aligned} P_y(nC_2 < \tau < \infty) &= P_y(z_0^{nC_2} < z_0^{\tau} < \infty) \\ &\le \mathbb{E}_y[z_0^{\tau} < \infty] z_0^{-nC_2} \\ &\le e^{-n(C_2\log(z_0) - \log(C_1)/n)}. \end{aligned}$$

Choosing $C_2 = \left(\delta + \frac{\log(C_1)}{n}\right) / \log(z_0)$ gives (2.10). □

2.3 LD limit for $P_x(\tau_n < \tau_0)$

Define

$$V(x) \doteq \log \rho_1(x(1)-1) \wedge \log(r)(x(1)+x(2)-1) \wedge \log \rho_2(x(2)-1).$$

Assumption (1.3) implies

$$-\log(\rho_2)(1-x(2)) \geq -\log(r)(1-(x(1)+x(2))),$$

and therefore

$$V(x) = \log(r)(x(1)+x(2)-1) \wedge \log \rho_1(x(1)-1). \tag{2.11}$$

The level curves of V for for $\lambda_1 = 0.2$, $\lambda_2 = 0.1$, $\mu_1 = 0.3$, $\mu_2 = 0.4$ are shown in Figure 2.2.

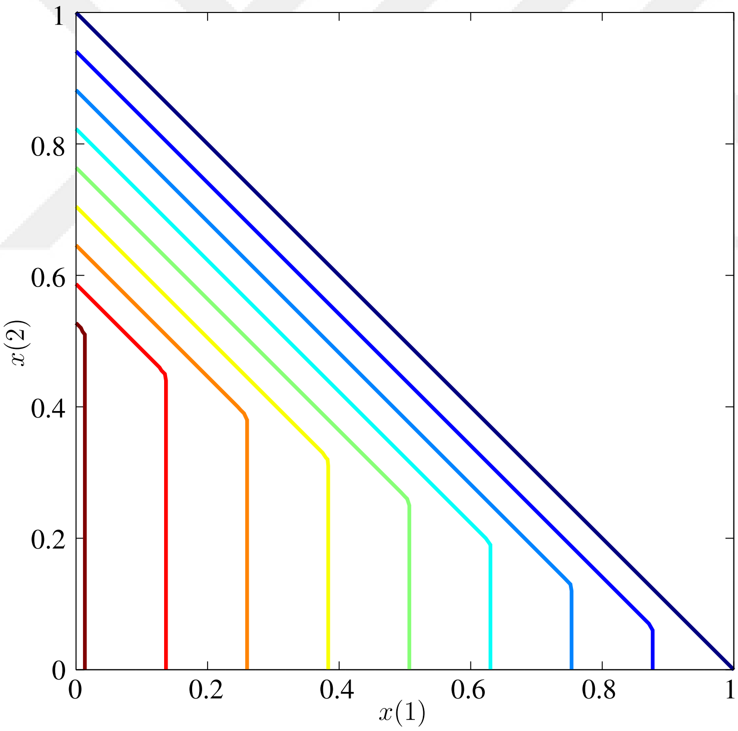

Figure 2.2: Level curves of V for $\lambda_1 = 0.2$, $\lambda_2 = 0.1$, $\mu_1 = 0.3$, $\mu_2 = 0.4$

The goal of this section is to prove:

Theorem 2.10. *V is the LD limit of $P_x(\tau_n < \tau_0)$, i.e.,*

$$\lim -\frac{1}{n} \log P_{\lfloor nx \rfloor}(\tau_n < \tau_0) = V(x). \tag{2.12}$$

for $x(1)+x(2)<1$, $x\in\mathbb{R}^2_+$.

Proof. Propositions 2.14 and 2.16 state

$$\liminf -\frac{1}{n}\log P_{\lfloor nx\rfloor}(\tau_n<\tau_0)\geq V(x). \tag{2.13}$$

and

$$\limsup -\frac{1}{n}\log P_{\lfloor nx\rfloor}(\tau_n<\tau_0)\leq V(x). \tag{2.14}$$

These imply (2.12). ☐

Next two subsections prove (2.13) and (2.14). To prove the first, we will proceed parallel to [25, 60, 64] and construct a sequence of supermartingales M^n starting from a subsolution of a limit Hamilton Jacobi Bellman (HJB) equation associated with the problem. To prove the bound (2.14) we will directly construct a sequence of subharmonic functions of the process X.

2.3.1 LD lowerbound for $P_x(\tau_n<\tau_0)$

For $a\subset\{0,1,2\}$ define the Hamiltonian function

$$H_a(q)\doteq -\log\left(\sum_{v\in a^c}p(v)e^{-\langle q,v\rangle}+\sum_{v\in\mathcal{V},v(i)\geq 0,i\in a}p(v)e^{-\langle q,v\rangle}+\sum_{v\in\mathcal{V},v(i)<0,i\in a}p(v)\right).$$

We will denote $H_\emptyset$ by H. H_a is convex in q. For $x\in\mathbb{R}^2_+$, define

$$\boldsymbol{b}(x)\doteq\{i:x(i)=0\}.$$

Following [23] one can represent V as the value function of a continuous time deterministic control problem; the HJB equation associated with this control problem is

$$H_{\boldsymbol{b}(x)}(DV(x))=0; \tag{2.15}$$

a function $W\in C^1$ is said to be a classical subsolution of (2.15) if

$$H_{\boldsymbol{b}(x)}(DV(x))\geq 0; \tag{2.16}$$

supersolutions are defined by replacing $\geq$ in (2.16) with $\leq$.

To prove (2.13) will proceed parallel to [64, Section 7]: find an upperbound on $P_x(\tau_n < \tau_0)$ by constructing a supermartingale associated with the process X. To construct our supermartingale we will proceed parallel to [25, 60] and use a subsolution of (2.15), i.e., a solution of (2.16).

Define

$$r_0 \doteq (0,0), \;\; r_1 \doteq \log(\rho_1)(1,0), \;\; r_2 \doteq \log(r)(0,1), \;\; r_3 \doteq \log(r)(1,1) \tag{2.17}$$

and

$$\begin{aligned}
&\tilde{V}_0(x,\epsilon) \doteq -\log(\rho_1) - 3\epsilon, \quad \tilde{V}_1(x,\epsilon) \doteq -\log(\rho_1) + \langle r_1, x\rangle - 2\epsilon, \\
&\tilde{V}_2(x,\epsilon) \doteq -\log(r) + \langle r_2, x\rangle - \epsilon, \\
&\tilde{V}_3(x,\epsilon) \doteq -\log(r) + \langle r_3, x\rangle,
\end{aligned}$$

and

$$\tilde{V}(x,\epsilon) \doteq \bigwedge_{i=0}^{3} \tilde{V}_i(x,\epsilon). \tag{2.18}$$

A direct calculation gives

Lemma 2.11. *The gradients defined in* (2.17) *satisfy*

$$\begin{aligned}
&H(r_0) = H_1(r_0) = H_2(r_0) = 0, \quad H(r_1) = H_2(r_1) = 0, \\
&H(r_2) > 0, H_1(r_2) > 0, \qquad H(r_3) = 0.
\end{aligned}$$

Proof.

$$\begin{aligned}
H(r_0) &= -\log\left(\lambda_1 e^0 + \lambda_2 e^0 + \mu_1 e^0 + \mu_2 e^0\right) = -\log(1) = 0. \\
H_1(r_0) &= -\log\left(\lambda_1 e^0 + \lambda_2 e^0 + \mu_1 e^0 + \mu_2\right) = -\log(1) = 0. \\
H_2(r_0) &= -\log\left(\lambda_1 e^0 + \lambda_2 e^0 + \mu_1 + \mu_2 e^0\right) = -\log(1) = 0. \\
H(r_1) &= -\log\left(\lambda_1 e^{-\log(\rho_1)} + \lambda_2 e^0 + \mu_1 e^{\log(\rho_1)} + \mu_2 e^0\right) \\
&= -\log(\lambda_1 \frac{1}{\rho_1} + \lambda_2 + \mu_1\rho_1 + \mu_2) \\
&= -\log(\mu_1 + \lambda_2 + \lambda_1 + \mu_2) = -\log(1) = 0. \\
H(r_2) &= -\log\left(\lambda_1 e^0 + \lambda_2 e^{-r} + \mu_1 e^0 + \mu_2 e^r\right) \\
&= -\log(\lambda_1 + \mu_1 - r(\lambda_2 - \mu_2)) > 0.
\end{aligned}$$

$$
\begin{aligned}
H_2(r_2) &= -\log\left(\lambda_1 e^0 + \lambda_2 e^{-r} + \mu_1 e^0 + \mu_2\right) \\
&= -\log(\lambda_1 + \mu_1 + \mu_2 + \frac{\lambda_2}{r}) = -\log(1 - \lambda_2(1 - \frac{1}{r})) > 0. \\
H_2(r_1) &= -\log\left(\lambda_1 e^{-\log(\rho_1)} + \lambda_2 e^0 + \mu_1 e^{\log(\rho_1)} + \mu_2\right) \\
&= -\log(\lambda_1 \frac{1}{\rho_1} + \lambda_2 + \mu_1 \rho_1 + \mu_2) \\
&= -\log(\mu_1 + \lambda_2 + \lambda_1 + \mu_2) = -\log(1) = 0. \\
H(r_3) &= -\log\left(\lambda_1 e^{-\log(r)} + \lambda_2 e^{-\log(r)} + \mu_1 e^{\log(r)} + \mu_2 e^{\log(r)}\right) \\
&= -\log(\lambda_1 \frac{1}{r} + \lambda_2 \frac{1}{r} + \mu_1 r + \mu_2 r) \\
&= -\log((\lambda_1 + \lambda_2)\frac{1}{r} + (\mu_1 + \mu_2) r) = -\log(1) = 0.
\end{aligned}
$$

□

The 0-level curve of the Hamiltonians is shown in Figure 2.3 and the gradients r_i is shown in Figure

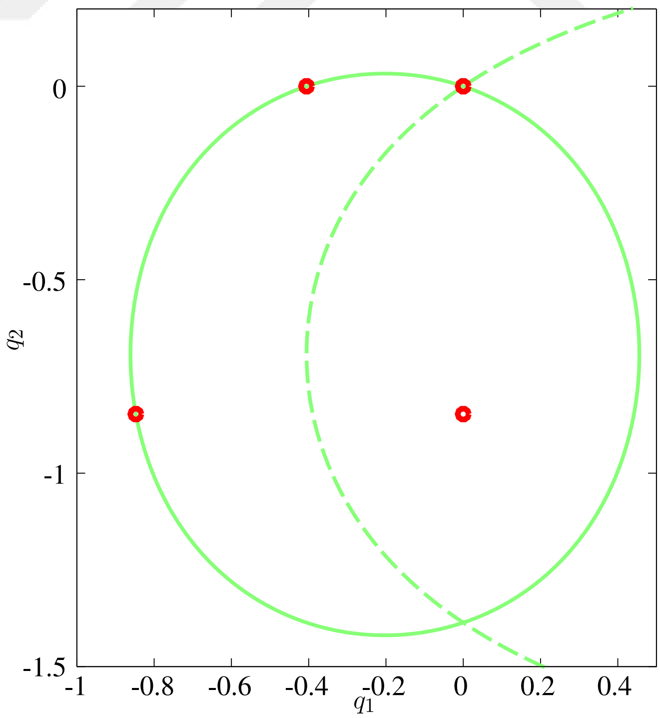

Figure 2.3: The 0-level curves of H and H_1 (dashed line)

Define

$$C_3 \doteq -\frac{1}{\log(r)}, C_4 \doteq -\log(\rho_1)C_3 = \frac{\log(\rho_1)}{\log(r)}. \tag{2.19}$$

$\log(r) < \log(\rho_1) < 0$ implies

$$1 > C_4 > 0. \tag{2.20}$$

By equating $\tilde{V}_2(x, \epsilon)$ and $\tilde{V}_3(x, \epsilon)$

$$-\log(r) + \log(r)x_2 - \epsilon = -\log(r) + \log(r)x_1 + \log(r)x_2,$$

we find the first component of the intersection point as $x_1 = -\frac{1}{\log(r)}\epsilon$, then we equate $\tilde{V}_1(x, \epsilon)$ and $\tilde{V}_3(x, \epsilon)$

$$-\log(\rho_1) + \log(\rho_1)C_3 - 2\epsilon = -\log(r) + \log(r)C_3 + \log(r)x_2,$$

to find the last component of the intersection point $x_2 = 1-\frac{\log(\rho_1)}{\log(r)}+C_3\left(1 + \frac{\log(\rho_1)}{\log(r)}\right)\epsilon$. Thus, the functions $\tilde{V}_i$, $i = 1, 2, 3$ meet at

$$x^* = (C_3\epsilon, 1 - C_4 + C_3(1 + C_4)\epsilon)\,; \tag{2.21}$$

i.e.,

$$\begin{aligned}
\tilde{V}_1(x^*) &= -\log(\rho_1) + \langle r_1, x^* \rangle - 2\epsilon \\
&= -\log(\rho_1) + \log(\rho_1)\left(-\frac{1}{\log(r)}\epsilon\right) - 2\epsilon \\
&= -\log(\rho_1) - (2 + C_4)\epsilon. \\
\tilde{V}_2(x^*) &= -\log(r) + \langle r_2, x^* \rangle - \epsilon \\
&= -\log(r) + \log(r)\left(1 - \frac{\log(\rho_1)}{\log(r)} - \frac{1}{\log(r)}\left(1 + \frac{\log(\rho_1)}{\log(r)}\right)\epsilon\right) - \epsilon \\
&= -\log(r) + \log(r) - \log(\rho_1) - \frac{1}{\log(r)}\log(r)\epsilon - \frac{1}{\log(r)}\log(\rho_1)\epsilon - \epsilon \\
&= -\log(\rho_1) - \epsilon - \frac{\log(\rho_1)}{\log(r)}\epsilon - \epsilon \\
&= -\log(\rho_1) - (2 + C_4)\epsilon.
\end{aligned}$$

$$
\begin{aligned}
\tilde{V}_3(x^*) &= -\log(r) + \langle r_2, x^* \rangle \\
&= -\log(r) - \log(r)\frac{1}{\log(r)}\epsilon \\
&+ \log(r)\left(1 - \frac{\log(\rho_1)}{\log(r)} - \frac{1}{\log(r)}\left(1 + \frac{\log(\rho_1)}{\log(r)}\right)\epsilon\right) \\
&= -\log(r) + \epsilon + \log(r) - \log(\rho_1) - \epsilon + \frac{1}{\log(r)}\log(\rho_1)\epsilon \\
&= -\log(\rho_1) - (2 + C_4)\epsilon.
\end{aligned}
$$

$$
\tilde{V}_1(x^*) = \tilde{V}_2(x^*) = \tilde{V}_3(x^*) = -\log(\rho_1) - (2 + C_4)\,\epsilon. \tag{2.22}
$$

We assume that $\epsilon > 0$ is small enough so that x^* satisfies

$$
x^*(1), x^*(2) > 0, x^*(1) + x^*(2) < 1.
$$

$\tilde{V}(\cdot, \epsilon)$ equals $\tilde{V}_i(\cdot, \epsilon)$ in the region

$$
R_i \doteq \{x \in \mathbb{R}^2 : \tilde{V}(x, \epsilon) = \tilde{V}_i(x, \epsilon)\},
$$

these regions are shown in Figure 2.4. As in [25, 60], we will mollify $\tilde{V}(x, \epsilon)$ with

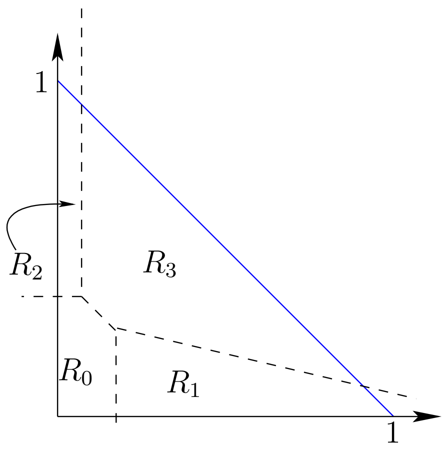

Figure 2.4: Regions R_i

$$
\eta(x) \doteq 1_{|x| \leq 1}(|x|^2 - 1)^2, x \in \mathbb{R}^2, C_5 \doteq \int_{\mathbb{R}^2} \eta(x)dx, \eta_\delta(x) \doteq \frac{1}{\delta^2 C_5}\eta(x/\delta), \delta > 0.
$$

to get our smooth subsolution of (2.15):

$$
V(x, \epsilon) \doteq \int_{\mathbb{R}^2} \tilde{V}(x + y, \epsilon)\eta_{0.5C_3\epsilon}(y)dy, \tag{2.23}
$$

Lemma 2.12. *The function $V(x,\epsilon)$ of* (2.23) *satisfies* (2.16) *and*

$$\left|\frac{\partial^2 V(\cdot,\epsilon)}{\partial x_i \partial x_j}\right| \leq C_6/\epsilon; \tag{2.24}$$

where C_6 is independent of x. Furthermore,

$$V(x,\epsilon) \leq \epsilon \textit{ for } x(1)+x(2)=1, x\in\mathbb{R}^2_+. \tag{2.25}$$

Proof. The proof is parallel to that of [60, Lemma 2.3.2]. $\tilde{V}$ is the minimum of four affine functions and hence is Lipschitz continuous and has a bounded (piecewise constant) gradient almost everywhere. This implies

$$\begin{aligned} DV(x,\varepsilon) &= \int_{\mathbb{R}^2} D\tilde{V}(x+y,\epsilon)\eta_{0.5C_3\epsilon}(y)dy \\ &= \sum_{i=0}^{3} w_i(x)D\tilde{V}_i, = \sum_{i=1}^{3} w_i(x)r_i; \end{aligned} \tag{2.26}$$

where

$$w_i(x) = \int_{\mathbb{R}^2} \eta_{0.5C_3\epsilon}(x)1_{R_i}(x)dx.$$

This shows that $V(\cdot,\varepsilon)\in C^1$. To show

$$H_{\boldsymbol{b}(x)}(DV(\cdot,\epsilon)) \geq 0, \tag{2.27}$$

one considers $x\in\mathbb{R}^{2o}_+ \doteq \{x\in\mathbb{R}^2_+, x(1),x(2)>0\}$, $x\in\partial_1$ and $x\in\partial_2$ separately. We will provide the details only for the first two. For $x\in\mathbb{R}^{2o}_+$,

$$H_{\boldsymbol{b}(x)}(DV(\cdot,\epsilon)) = H(DV(\cdot,\epsilon)).$$

By Lemma 2.11 we know that all r_i satisfy $H(r_i)\geq 0$. That H is a convex function and Jensen's inequality imply that the $DV(x,\epsilon)=\sum_{i=0}^{3} w_i(x)r_i$ satisfies $H(DV(x,\epsilon))\geq 0$; this proves (2.27) for $x\in\mathbb{R}^{2o}_+$.

We know by (2.20) that $C_4\in(0,1)$; therefore, by (2.22)

$$\begin{aligned} \tilde{V}_0(\cdot,\epsilon) &= -\log(\rho_1) - 3\epsilon < -\log(\rho_1) - (2+C_4)\epsilon \\ &= \tilde{V}_1(x^*,\epsilon) = \tilde{V}_2(x^*,\epsilon) = \tilde{V}_3(x^*,\epsilon). \end{aligned}$$

This implies that the region R_0 intersects all of the R_1, R_2 and R_3 and in particular that the strip $\{x \in \mathbb{R}^2_+ : x(1) < C_3\epsilon\}$ lies in $R_0 \cup R_2$. Then, for $x \in \partial_1$, the ball $B(x, C_3\epsilon/2)$ lies completely in $R_0 \cup R_2$, which implies

$$DV(x,\epsilon) = w_0(x)r_0 + w_1(x)r_2, w_0(x) + w_1(x) = 1.$$

By Lemma 2.11 we know that $H_1(r_0) = 0$ and $H_1(r_2) \geq 0$. These and the convexity of H_i imply (2.27) for $x \in \partial_1$.

The bound (2.24) follows from the Lipschitz continuity of $\tilde{V}(\cdot,\epsilon)$, differentiation under the integral sign in (2.23) and bounds on the first derivative of η. Finally, (2.25) follows from

$$\tilde{V}_3(x,\epsilon) \leq \frac{\epsilon}{\sqrt{2}}, \text{ for any } x \in B(y, \epsilon C_3/2),\ y \text{ such that } y(1) + y(2) = 1, y \in \mathbb{R}^2_+,$$

□

To get our upperbound on the probability $P_x(\tau_n < \tau_0)$ we define

$$M_k^{(n,\epsilon)} \doteq e^{-nV(X_k/n,\epsilon) - \frac{C_6 k}{n\epsilon}},$$

where C_6 is as in (2.24).

Lemma 2.13. *$M^{(n,\epsilon)}$ is a supermartingale.*

Proof. The Markov property of X implies that it suffices to show

$$\begin{aligned} \mathbb{E}_x\left[M_1^{(n,\epsilon)}\right] &\leq e^{-nV(x/n,\epsilon)} \\ \mathbb{E}_x\left[M_1^{(n,\epsilon)} e^{nV(x/n,\epsilon)}\right] &\leq 1 \\ -\log\left(\mathbb{E}_x\left[M_1^{(n,\epsilon)} e^{nV(x/n,\epsilon)}\right]\right) &\geq 0. \end{aligned} \tag{2.28}$$

The expression on the left equals

$$-\log\left(\sum_{v:v(i)\geq 0, i\in \boldsymbol{b}(x)} e^{-n(V((x+v)/n,\epsilon) - V(x/n,\epsilon))} p(v) + \sum_{v:v(i)=-1, i\in\boldsymbol{b}(x)} p(v)\right) + \frac{C_6}{n\epsilon}. \tag{2.29}$$

A Taylor expansion and the bound (2.24) imply

$$|(V((x+v)/n,\epsilon) - V(x/n,\epsilon)) - \langle DV(x), v/n\rangle| \leq \frac{C_6}{n^2\epsilon}.$$

Then, the expression in (2.29) is bounded below by

$$-\log\left(\sum_{v:v(i)\geq 0, i\in \boldsymbol{b}(x)} e^{-\langle DV(x),v\rangle}p(v)+\sum_{v:v(i)=-1,i\in\boldsymbol{b}(x)} p(v)\right)-\frac{C_6}{n\epsilon}+\frac{C_6}{n\epsilon}.$$

The log term above equals $H_{\boldsymbol{b}(x)}(DV(x,\epsilon))$, which by Lemma 2.12 is nonnegative. This proves (2.28). □

Proposition 2.14. *Let $x\in\mathbb{R}^2_+$ with $x(1)+x(2)<1$, $x_n=\lfloor nx\rfloor$ and let V be as in* (2.11). *Then for any $\varepsilon>0$ there exists an integer N such that for $n>N$*

$$P_{x_n}(\tau_n<\tau_0)\leq e^{-n(V(x)-\varepsilon)}. \tag{2.30}$$

In particular,

$$\liminf -\frac{1}{n}\log P_{\lfloor nx\rfloor}(\tau_n<\tau_0)\geq V(x). \tag{2.31}$$

The proof is parallel to that of [65, Proposition 4.3].

Proof. The inequality (2.31) follows from (2.30) upon taking limits. The rest of the proof focuses on (2.30). Let $\epsilon_n>0$ be a sequence satisfying $\epsilon_n\to 0$ and $\epsilon_n n\to\infty$. Let $\tau_{0,n}=\tau_n\wedge\tau_0$. The optional sampling theorem ([27, Theorem 5.7.6]) applied to the supermartingale $M_k=M_k^{(n,\epsilon_n)}$ at time $\tau_{0,n}$ gives

$$\mathbb{E}_{x_n}\left[M_{\tau_{0,n}}\right]\leq M_0=e^{-nV(x_n/n,\epsilon_n)}.$$

Restricting the expectation on the left to $\{\tau_n<\tau_0\}$ makes it smaller:

$$\mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}}M_{\tau_n}\right]\leq e^{-nV(x_n/n,\epsilon_n)}.$$

Expanding M_{τ_n} using its definition gives

$$\mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}}e^{-nV(X_{\tau_n}/n,\epsilon_n)}e^{\frac{-C_6\tau_n}{n\epsilon_n}}\right]\leq e^{-nV(x_n/n,\epsilon_n)}.$$

$X_{\tau_n}\in\partial A_n$ and the bound (2.25) reduce the last display to

$$\mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}}e^{\frac{-C_6\tau_n}{n\epsilon_n}}\right]\leq e^{-nV(x_n/n,\epsilon_n)}e^{n\epsilon_n}. \tag{2.32}$$

By the definitions involved we have

$$\lim_{n\to\infty}V(x_n,\epsilon_n)=V(x).$$

This, $n\epsilon_n \to \infty$ and taking the $\liminf -\frac{1}{n}\log$ of both sides in (2.32) gives

$$\liminf_{n\to\infty} -\frac{1}{n}\log \mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}} e^{\frac{-C_6\tau_n}{n\epsilon_n}}\right] \geq V(x). \tag{2.33}$$

Now suppose that (2.30) doesn't hold, i.e., there exists $\varepsilon > 0$ and a sequence n_k such that

$$P_{x_{n_k}}(\tau_n < \tau_0) > e^{-n_k(V(x)-\varepsilon)}, \tag{2.34}$$

for all k; we pass to this subsequence and omit the subscript k. [60, Theorem A.1.1] implies that there is a $C_2 > 0$ such that

$$P_{x_n}(\tau_{0,n} > nC_2) \leq e^{-n(V(x)+1)}, \tag{2.35}$$

for n large. Then

$$\begin{aligned}
\mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}} e^{-\frac{C_6\tau_n}{n\epsilon_n}}\right] &\geq \mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}} e^{-\frac{C_6\tau_n}{n\epsilon_n}} 1_{\{\tau_{0,n}\leq nC_2\}}\right] \\
&\geq e^{\frac{-C_6C_2}{n\epsilon_n}n}\mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}}1_{\{\tau_{0,n}\leq nC_2\}}\right] \\
&\geq e^{\frac{-C_6C_2}{n\epsilon_n}n}\left(P_{x_n}(\tau_n<\tau_0) - P_{x_n}(\tau_{0,n} > nC_2)\right) \\
&\geq e^{\frac{-C_6C_2}{n\epsilon_n}n}\left(e^{-n(V(x)-\varepsilon)} - e^{-(V(x)+1)n}\right).
\end{aligned}$$

Now taking $\limsup -\frac{1}{n}\log$ of both sides gives

$$\limsup_{n\to\infty} -\frac{1}{n}\log \mathbb{E}_{x_n}\left[1_{\{\tau_n<\tau_0\}} e^{\frac{-C_6\tau_n}{n\epsilon_n}}\right] \leq V(x) - \varepsilon.$$

This contradicts (2.33). Therefore, the assumption (2.34) is false and there does exist $N > 0$ such that (2.30) holds for $n > N$. This finishes the proof of this proposition.

□

2.3.2 LD upperbound for $P_x(\tau_n < \tau_0)$

The LD upperbound corresponds (because of the $-\log$ transform) to a lowerbound on the probability $P_x(\tau_n < \tau_0)$. To get a lower bound on this probability, it suffices to have a submartingale of X with the right values when X hits $\partial A_n \cup \{0\}$. As opposed to the analysis of the previous section (where we constructed a supermartingale from a subsolution to a limit HJB equation), one can directly construct a subharmonic function of X to get the desired submartingale. The next proposition gives this explicit subharmonic function. In its proof the following fact will be useful: if g_1 and

g_2 are subharmonic functions of X at a point x, then so is $g_1 \vee g_2$, this follows from the definitions involved.

Proposition 2.15.

$$f_n(x) \doteq {\rho_1}^{n-x(1)} \vee r^{(n-x(1))-x(2)} \vee \rho_1^{n-1}$$

is a subharmonic function of X on $A_n - \partial A_n$

Proof. We note

$$\rho_1^{n-x(1)} = \rho_1^{(n-x(1))-x(2)} \rho_1^{x(2)} = [(\rho_1, \rho_1), T_n(x)].$$

Furthermore, $(\rho_1, \rho_1) \in \mathcal{H}$. It follows from these and Lemma 2.5 that $x \mapsto \rho_1^{n-x(1)}$ is X-harmonic for $x \in \mathbb{Z}_+^{2o} \doteq \mathbb{Z}_+^2 - \{\partial_1 \cup \partial_2\}$. A parallel argument proves the same for $x \mapsto r^{(n-x(1))-x(2)}$. The constant function $x \mapsto \rho_1^{n-1}$ is trivially X-harmonic for all $x \in \mathbb{Z}_+^2$. It follows that their maximum, f_n is subharmonic on $\mathbb{Z}_+^{2o}$.

It remains to prove that f_n is subharmonic on ∂_1 and ∂_2. $f_n(x) = r^{(n-x(1))-x(2)} \vee \rho_1^{n-1}$ for $x \in \partial_1 \cup \{x \in \mathbb{Z}_+^2, x(1) = 1\}$. Both $x \mapsto r^{(n-x(1))-x(2)}$ and $x \mapsto \rho_1^{n-1}$ are X-harmonic on ∂_1. It follows from these that f_n is subharmonic on ∂_1.

For $x \in \partial_2 \cap \{x \in \mathbb{Z}_2^+ : x(1) < n\}$ we have $f_n(x) = \rho_1^{n-x(1)} \vee \rho_1^{n-1}$. By Lemma 2.4 and by fact that $I_k = \mathcal{I} J_k$ we know $x \mapsto \rho_1^{n-x(1)}$ is harmonic on ∂_2; the same trivially holds for $x \mapsto \rho_1^{n-1}$; therefore, $x \mapsto \rho_1^{n-x(1)} \vee \rho_1^{n-1}$ is subharmonic on $\partial_2 \cap \{x \in \mathbb{Z}_2^+ : x(1) < n\}$ Furthermore, by definition $f_n(x) \geq \rho_1^{n-x(1)} \vee \rho_1^{n-1}$. These imply that f_n is subharmonic on $\partial_2 \cap \{x \in \mathbb{Z}_2^+ : x(1) < n\}$. The last three paragraphs together imply the statement of the proposition. ☐

Proposition 2.16.

$$P_x(\tau_n < \tau_0) \geq f_n(x) - f_n(0) \tag{2.36}$$

and in particular

$$\limsup -\frac{1}{n} \log P_{\lfloor nx \rfloor}(\tau_n < \tau_0) \leq V(x), \tag{2.37}$$

for $x \in \mathbb{R}_+^2$, $x(1) + x(2) < 1$, $x(1) > 0$.

Proof. By Proposition 2.15, we know that f_n is a subharmonic function of X. It follows that $f(X_n)$ is a submartingale. This and the optional sampling theorem imply:

$$\begin{aligned} f_n(x) &\leq \mathbb{E}_x[f_n(\tau_n \wedge \tau_0)] \\ f_n(x) &\leq P_x(\tau_n < \tau_0)(1 - f_n(0)) + f_n(0) \\ &\leq P_x(\tau_n < \tau_0) + f_n(0), \end{aligned}$$

where we have used $f_n(x) = 1$ for $x \in \partial A_n$; this gives (2.36). Taking $-\frac{1}{n}\log$ of both sides and applying $\limsup$ gives (2.37). □

2.4 LD limit of $P_x(\sigma_1 < \tau_n < \tau_0)$

To implement the argument given in the introduction we need an LD lowerbound for the probability

$$P_x(\sigma_1 < \tau_n < \tau_0). \tag{2.38}$$

We will obtain the desired bound through a subsolution of the limit HJB equation associated with X. This is parallel to the construction given in [65, Proposition 4.3] and the argument of Section 2.3.1. The main difference is in the construction of the subsolution. Bounding (2.38) requires a subsolution consisting of two pieces, one piece for before σ_1 and one for after. For the first piece we need the following additional root of the limit Hamiltonian:

$$r_4 \doteq (\log(\rho_1/r), \log(r))\,. \tag{2.39}$$

Now define

$$\begin{aligned} \tilde{V}_4(x,\epsilon) &\doteq -\log(r) + \langle r_4, x\rangle, \\ \tilde{V}(0,x,\epsilon) &\doteq \bigwedge_{i\in\{0,2,4\}} \tilde{V}_i(x,\epsilon), \\ \tilde{V}(1,x,\epsilon) &\doteq \tilde{V}(x,\epsilon) = \bigwedge_{i=0}^{3} \tilde{V}_i(x,\epsilon), \end{aligned} \tag{2.40}$$

and

$$\begin{aligned} V_\sigma(0,x) &\doteq \tilde{V}(0,x,0) = (-\log(\rho_1)) \wedge (-\log(r) + \langle r_4, x\rangle), \\ V_\sigma(1,x) &\doteq \tilde{V}(1,x,0) = (-\log(\rho_1) + \langle r_1, x\rangle) \wedge (-\log(r) + \langle r_3, x\rangle). \end{aligned}$$

where the vectors r_i are as in (2.17).

Now define the smoothed subsolution:

$$V(x,\epsilon,i) \doteq \int_{\mathbb{R}^2} \tilde{V}(x+y,\epsilon,i)\eta_{0.5C_3\epsilon}(y)dy, i = 0,1. \tag{2.41}$$

The function $\tilde{V}(0,\cdot,\cdot)$ is obtained from $\tilde{V}(1,\cdot,\cdot)$ by striking out $\tilde{V}_1$ from the minimum and replacing $\tilde{V}_3$ with $\tilde{V}_4$. In particular, the components $\tilde{V}_0$ and $\tilde{V}_2$ are common to both $\tilde{V}(1,\cdot,\cdot)$ and $\tilde{V}(0,\cdot,\cdot)$; this ensures that these functions overlap around an open region along ∂_1, which implies in particular that

$$V(1,x,\epsilon) = V(0,x,\epsilon), \tag{2.42}$$

for $x \in \partial_1$.

Remark 2.1. *The condition* (2.42) *allows one to think of* $V(\cdot,\cdot,\cdot)$ *as a subsolution of the HJB equation on a manifold; the manifold consists of two copies of* $\mathbb{R}^2_+$*, glued to each other along* $\{x \in \mathbb{R}^2_+, x(1) = 0\}$.

We use $V(\cdot,\cdot,\cdot)$ to construct the supermartingale

$$M_k^{(n,\epsilon,\sigma)} \doteq e^{-nV(X_k/n,\epsilon,1_{\{k<\sigma_1\}}) - \frac{C_6 k}{n\epsilon}},$$

where C_6/ϵ is an upperbound on the second derivative of $V(\cdot,\cdot,\cdot)$, which can be obtained by an argument parallel to the one used in the proof of (2.24) of Lemma 2.12. The main difference from subsection 2.3.1 is that the smooth subsolution has an additional parameter i to keep track of whether X has touched ∂_1; this appears as the $1_{\{k<\sigma_1\}}$ term in the definition of the supermartingale $M^{(n,\epsilon,\sigma)}$. A three stage version of this argument appears in [65, Proposition 4.3] to bound another related probability arising from the analysis of the two dimensional tandem random walk.

The main result of this section is the following:

Proposition 2.17. *For any* $\epsilon > 0$*, there exists* $N > 0$ *such that*

$$P_{x_n}(\sigma_1 < \tau_n < \tau_0) \le e^{-n(V_\sigma(0,x)-\epsilon)}, \tag{2.43}$$

for $n > N$*, where* $x_n = \lfloor nx \rfloor$, $0 < x(1) + x(2) < 1$, $x \in \mathbb{R}^2_+$.

Proof. Parallel to the proof of Proposition 2.14, we choose a sequence $\epsilon_n \to 0$ with $n\epsilon_n \to \infty$; (2.44) follows from an application of the optional sampling theorem to the supermartingale $M^{(n,\epsilon_n,\sigma)}$ and the bound (2.35). □

2.4.1 LD limit for $P_x(\bar{\sigma}_1 < \tau < \infty)$

For this subsection and the next section it will be convenient to express the Y process in x coordinates, we do this by setting, $\bar{X}_k \doteq T_n(Y_k)$; $\bar{X}_k$ has the following dynamics:

$$\bar{X}_{k+1} = \bar{X}_k + \pi_1(\bar{X}_k, I_k).$$

$\bar{\sigma}_1$ of (1.8) in terms of $\bar{X}$ is $\bar{\sigma}_1 = \inf\{k : \bar{X}_k \in \partial_1\}$.

The processes $\bar{X}$ and X have the same dynamics except that $\bar{X}$ is not constrained on ∂_1. By definition, $\bar{X}_0 = X_0$. Note the following: $\bar{X}$ hits $\{x \in \mathbb{Z} \times \mathbb{Z}_+ : x(1) + x(2) = n\}$ exactly when Y hits $\{y \in \mathbb{Z} \times \mathbb{Z}_+ : y(1) = y(2)\}$; i.e., if we define

$$\bar{\tau}_n \doteq \inf\{k : \bar{X}_k(1) + \bar{X}_k(2) = n\},$$

then $\tau = \bar{\tau}_n$.

Proposition 2.18. *For any $\epsilon > 0$, there exists $N > 0$ such that*

$$P_{x_n}(\bar{\sigma}_1 < \bar{\tau}_n < \infty) \leq e^{-n(V_\sigma(0,x) - \epsilon)}, \tag{2.44}$$

for $n > N$, where $x_n = \lfloor nx \rfloor$, $x(1) + x(2) < 1$, $x \in \mathbb{R} \times \mathbb{R}_+$

Proof. The two stage subsolution $V(\cdot, \cdot, \cdot)$ of (2.41) is a subsolution for the $\bar{X}$ process as well because, $\bar{X}$ has identical dynamics as X with one less constraint. Therefore, the proof of Proposition 2.17 applies verbatim to the current setup with one change: in the proof of (2.44) we truncate time with the bound (2.35) for τ_n. We replace this with the corresponding bound (2.10) for τ. □

2.5 Completion of the limit analysis

We now combine Propositions 2.16, 2.17 and 2.18 to get the main approximation result of this work:

Theorem 2.19. *For any $x \in \mathbb{R}^2_+$, $x(1) + x(2) < 1$, $x(1) > 0$, there exists $C_7 > 0$ and $N > 0$ such that*

$$\frac{|P_{x_n}(\tau_n < \tau_0) - P_{T_n(x_n)}(\tau < \infty)|}{P_{x_n}(\tau_n < \tau_0)} = \frac{|P_{x_n}(\tau_n < \tau_0) - P_{x_n}(\bar{\tau}_n < \infty)|}{P_{x_n}(\tau_n < \tau_0)} < e^{-C_7 n},$$

for $n > N$, where $x_n = \lfloor xn \rfloor$.

That $P_{x_n}(\bar{\tau}_n < \infty) = P_{T_n(x_n)}(\tau < \infty)$ follows from the the definitions in subsection 2.4.1.

Proof. The definitions (2.11) and (2.40) imply that

$$2C_7 = V_\sigma(x,0) - V(x) > 0,$$

for $x \in \mathbb{R}^2_+$, $x(1) + x(2) < 1$, $x(1) > 0$. Choose $\epsilon < C_7$. The processes X and $\bar{X}$ follow exactly the same path until they hit ∂_1. It follows that

$$|P_{x_n}(\tau_n < \tau_0) - P_{x_n}(\bar{\tau}_n < \infty)| \leq P_{x_n}(\sigma_1 < \tau_n < \tau_0) + P_{x_n}(\bar{\sigma}_1 < \bar{\tau}_n < \infty). \quad (2.45)$$

By Propositions 2.17, 2.18 and 2.16 there exists $N > 0$ such that

$$P_{x_n}(\sigma_1 < \tau_n < \tau_0) + P_{x_n}(\bar{\sigma}_1 < \bar{\tau}_n < \infty) \leq e^{-n(V_\sigma(0,x)-\epsilon/2)}, \quad (2.46)$$

and

$$P_x(\tau_n < \tau_0) \geq e^{-n(V(x)-\epsilon/2)}, \quad (2.47)$$

for $n > N$. The bounds (2.45), (2.46) and (2.47) give

$$\frac{|P_{x_n}(\tau_n < \tau_0) - P_{x_n}(\bar{\tau}_n < \infty)|}{P_{x_n}(\tau_n < \tau_0)} < e^{-nC_7},$$

for $n > N$. □

2.6 Computation of $P_y(\tau < \infty)$

Theorem 2.19 tells us that $P_y(\tau < \infty)$, $y = T_n(x_n)$, approximates $P_{x_n}(\tau_n < \tau_0)$ with exponentially decaying relative error for $x(1) > 0$. To complete our analysis, it remains to compute $P_y(\tau < \infty)$. As a function of y, $P_y(\tau < \infty)$ is a Y-harmonic function. Furthermore, it is ∂B-determined, i.e., it has the representation

$$y \to \mathbb{E}[g(Y_\tau)1_{\{\tau<\infty\}}],$$

for some function g on ∂B (for $y \mapsto \partial P_y(\tau < \infty)$, g equals 1 identically). We will try to compute $P_y(\tau < \infty)$ as a superposition of the Y-harmonic functions expounded in Section 2.1; because $P_y(\tau < \infty)$ is 1 for $y \in \partial B$, we would like the superposition to be as close to 1 as possible on ∂B. We have two classes of Y-harmonic functions given in Propositions 2.2 (constructed from a single point on $\mathcal{H}$) and 2.3 (constructed from conjugate points on $\mathcal{H}$). The first class gives us only one nontrivial Y-harmonic function, computed in Lemma 2.4: $h_{\rho_1} = [(\rho_1, \rho_1), \cdot]$. Remember that we have assumed $\boldsymbol{\alpha}(r, 1) = r^2/\rho_2 < 1$. This implies that, among the functions in the second class, the most relevant for the computation of $P_y(\tau < \infty)$ is

$$\boldsymbol{h}_r = \frac{1}{1-\rho_2/r} h_r = [(r, 1), \cdot] - \frac{1-r}{1-\rho_2/r}[(r, \rho_2/r^2), \cdot],$$

because this Y-harmonic function exponentially converges to 1 for $y = (k, k) \in \partial B$ and $k \to \infty$. A simple criterion to check whether a Y-harmonic function of the form $\sum_{i=1}^{I} c_i[(\beta_i, \alpha_i)]$ is ∂B-determined is given in [64]:

Proposition 2.20. *A Y-harmonic function of the form $\sum_{i=1}^{I} c_i[(\beta_i, \alpha_i)]$ is ∂B determined if $|\beta_i| < 1$ and $|\alpha_i| \leq 1$, $i = 1, 2, 3, ..., I$.*

Our main theoretical result on $P_y(\tau < \infty)$ arises from a linear combination of h_{ρ_1} and h_r:

Proposition 2.21. *If*

$$\rho_2\rho_1 = r^2, \tag{2.48}$$

then

$$P_y(\tau < \infty) = \boldsymbol{h}_r(y) + \frac{1-r}{1-\rho_2/r} h_{\rho_1}(y), \tag{2.49}$$

for $y \in B$.

Proof. The right side of (2.49) is Y-harmonic by construction. Furthermore, $\rho_2\rho_1 = r^2$ implies $\boldsymbol{h}_r(y) + \frac{1-r}{1-\rho_2/r} h_{\rho_1}(y) = 1$ for $y \in \partial B$. Therefore, to prove (2.49) it suffices to prove that

$$\boldsymbol{h}_r + \frac{1-r}{1-\rho_2/r} h_{\rho_1} \tag{2.50}$$

is ∂B-determined. For this we will use Proposition 2.20; in the present case, the β_i are $\rho_1, r < 1$ and the α_i are 1 and $\rho_1 \leq 1$. It follows that (2.50) is ∂B-determined. ☐

If (2.48) doesn't hold, i.e., if $r^2 \neq \rho_1\rho_2$ then one can proceed in several ways. As a first step, one can use the functions h_r and h_{ρ_1} to construct lower and upper bounds on $P_y(\tau < \infty)$:

Proposition 2.22. *There exists positive constants c_0, c_1, and C_8*

$$P_y(\tau < \infty) \leq h^{a,0}(y) \leq C_8 P_y(\tau < \infty); \tag{2.51}$$

where

$$h^{a,0} = c_0 \boldsymbol{h}_r + c_1 h_{\rho_1}. \tag{2.52}$$

In particular, $h^{a,0}$ approximates $P_y(\tau < \infty)$ with bounded relative error.

Proof. If $\rho_1 > r^2/\rho_2$, one can set $c_0 = 1$ and $c_1 = \frac{1-r}{1-\rho_2/r}$ since, for these values

$$h^{a,0} = \boldsymbol{h}_r + \frac{1-r}{1-\rho_2/r} h_{\rho_1} \geq 1 \tag{2.53}$$

on ∂B. Both $h^{a,0}$ and $y \mapsto P_y(\tau < \infty)$; this and (2.53) imply $h^{a,0}(y) \geq P_y(\tau < \infty)$ for $y \in B$. To get the second bound on (2.51) set

$$C_8 = 1 + \frac{1-r}{1-\rho_2/r} \max_{x \geq 0} \left[\rho_1^x - \left(\frac{r^2}{\rho_2} \right)^x \right]. \tag{2.54}$$

With this choice of C_8 we get the second bound in (2.51) on ∂B; that both $y \mapsto P_y(\tau < \infty)$ and $h^{a,0}$ are ∂B-determined implies the same bound on all of B. If $\rho_1 < r^2/\rho_1$, first choose C_0 so that

$$1 + \min_{x \geq 0} \left[C_0 \rho_1^x - \frac{1-r}{1-\rho_2/r} \left(\frac{r^2}{\rho_2} \right)^x \right] \geq 1/2. \tag{2.55}$$

Then

$$h^{a,0}(y) = 2h_r(y) + 2C_0 h_{\rho_1}(y) \geq 1,$$

for $y \in \partial B$, from which the first bound in (2.51) follows. To get the second bound, set

$$C_8/2 = 1 + \max_{x \geq 0} \left[C_0 \rho_1^x - \frac{1-r}{1-\rho_2/r} \left(\frac{r^2}{\rho_2} \right)^x \right],$$

and proceed as above. $\square$

Our choice of the constant $1/2$ in (2.55) is arbitrary, any value between $(0, 1)$ would suffice for the argument. Therefore, the constants c_0 and c_1 are not unique and they can be optimized to reduce relative error.

Proposition 2.23. *For $x \in \mathbb{R}_+^2$, $x(1) + x(2) < 1$, $x(1) > 0$, $x_n = \lfloor nx \rfloor$, and for n large, $h^{a,0}$ of* (2.52) *evaluated at $T_n(x_n)$ approximates $P_{x_n}(\tau_n < \tau_0)$ with bounded relative error.*

Proof. We know by Theorem 2.19 that, for $x \in \mathbb{R}_+^2$, $x(1) + x(2) < 1$ and $x(1) > 0$, $P_{T_n(x_n)}(\tau < \infty)$ approximates $P_{x_n}(\tau_n < \tau_0)$ with vanishing relative error. On the other hand, the above Proposition tells us that $h^{a,0}$ of (2.52) approximates $P_y(\tau < \infty)$ with bounded relative error. These imply that $h^{a,0}(T_n(x))$ approximates $P_{x_n}(\tau_n < \tau_0)$ with bounded relative error. □

Proposition 2.7 gives not one but a one-complex-parameter family of Y-harmonic functions. A natural question is whether one can obtain finer approximations of $P_y(\tau < \infty)$ than what $h^{a,0}$ provides. In this, we need ∂B-determined Y-harmonic functions. The next proposition (an adaptation of [64, Proposition 4.13] to the current setting) identifies a class of these which are naturally suitable for the approximation of $P_y(\tau < \infty)$.

Proposition 2.24. *There exists $0 < R < 1$ such that for all $\alpha \in \mathbb{C}$ with $R < |\alpha| \leq 1$, $\max(|\beta_1(\alpha)|, |\boldsymbol{\alpha}(\beta_1(\alpha), \alpha)|) < 1$; in particular $h_{\beta_1(\alpha)}$ is ∂B-determined.*

Proof. We know by [64, Proposition 4.7] that $|\beta_1(\alpha)| \leq r < 1$ for all $|\alpha| = 1$. Then

$$|\boldsymbol{\alpha}(\beta_1(\alpha), \alpha)| = \left| \frac{\beta_1(\alpha)^2}{\alpha \rho_2} \right| \leq \frac{r^2}{\rho_2} < 1,$$

where the last inequality is the assumption (2.5). The functions β_1 and $\boldsymbol{\alpha}$ are continuous; it follows that the inequality above holds also for $R < |\alpha| \leq 1$ if $R < 1$ is sufficiently close to 1. That h_{β_1} is ∂B-determined follows from these and Proposition 2.20. □

We can now use as many of the ∂B-determined Y-harmonic functions identified in Propositions 2.3 and 2.24 as we like to construct finer approximations of $P_y(\tau < \infty)$. Once the approximation is constructed upperbounds on its relative error can be computed from the maximum and the minimum of the approximation on ∂B- as was done in the proof of Proposition 2.22:

Proposition 2.25. *Let R be as in Proposition 2.20. For $c_k \in \mathbb{C}$ and $R < |\alpha_k| \leq 1$ $k = 0, 1, 2, ..., K$ define*

$$h^{a,K} = \Re(h^{a*,K}), h^{a*,K} = \boldsymbol{h}_r + c_0 h_{\rho_1} + \sum_{i=1}^{K} c_k h_{\beta_1(\alpha_k)}. \tag{2.56}$$

Then $h^{a,K}$ is Y-harmonic and ∂B-determined. Furthermore, for

$$c^* = \max_{y \in \partial B} |h^{a*,K} - 1| < \infty, \tag{2.57}$$

$h^{a*,K}$ approximates $P_y(\tau < \infty)$ with relative error bounded by c^*.

Proof. We know by Propositions 2.2 and 2.3 that $h^{a*,K}$ is Y-harmonic. That $R < |\alpha_k| \leq 1$ and Proposition 2.20 imply that $h^{a*,K}$ is also ∂B-determined, i.e.,

$$h^{a*,K}(y) = \mathbb{E}_y[h^{a*,K}(Y_\tau) 1_{\{\tau<\infty\}}].$$

Taking the real part of both sides gives:

$$h^{a,K}(y) = \mathbb{E}_y[h^{a,K}(Y_\tau) 1_{\{\tau<\infty\}}], \tag{2.58}$$

i.e, $h^{a,K}$ is Y-harmonic and ∂B-determined. That $c^* < \infty$ follows from

$$\max |\beta_1(\alpha_k), \boldsymbol{\alpha}(\beta_1(\alpha_k), \alpha_k)| < 1.$$

(see Proposition 2.20).

The inequality

$$1 - c^* < h^{a,K}(k, k) < 1 + c^*, \tag{2.59}$$

follows from (2.57), $|\Re(z) - 1| \leq |z - 1|$ for any $z \in \mathbb{C}$. It follows from (2.59) and (2.58) that

$$(1 - c^*)\mathbb{E}_y[1_{\{\tau<\infty\}}] \leq h^{a,K}(y) \leq (1 + c^*)\mathbb{E}_y[1_{\{\tau<\infty\}}]$$
$$(1 - c^*)P_y(\tau < \infty) \leq h^{a,K}(y) \leq (1 + c^*)P_y(\tau < \infty),$$

This implies that $h^{a,K}$ approximates $P_y(\tau < \infty)$ with relative error bounded by c^*. □

One of the key aspects of Proposition 2.25 is that it shows us how to compute an upper bound on the relative error of an approximation of the form (2.56) from the values it takes on ∂B. We can use this to choose the α_k and the c_k to reduce relative error, the next subsection illustrates this procedure.

2.6.1 Finer approximations when $r^2 \neq \rho_1\rho_2$

To illustrate how one can use approximations of the form (2.56) to improve on the approximation provided by Proposition 2.22, let us assign values to the parameters λ_i and μ_i satisfying the assumptions (1.3), (1.4):

$$\lambda_1 = 0.1, \mu_1 = 0.2, \lambda_2 = 0.2, \mu_2 = 0.5;$$

for these choice of parameters we have

$$r = \frac{\lambda_1 + \lambda_2}{\mu_1 + \mu_2} = \frac{3}{7}.$$

We note $r^2 = 9/49 \neq 1/5 = \rho_1\rho_2$; therefore, we don't have an explicit formula for $P_y(\tau < \infty)$. But Proposition 2.22 implies that

$$h^{a,0}(y) = \boldsymbol{h}_r + \frac{1-r}{1-\rho_2/r} h_{\rho_1} \tag{2.60}$$

approximates $P_y(\tau < \infty)$ with relative error bounded by

$$C_8 - 1 = \frac{1-r}{1-\rho^2/r}\left(\rho_1^{x^*} - \alpha_2^{x^*}\right) = 0.3607,$$

$$\alpha_2 = \frac{r^2}{\rho_2}, \quad x^* = \log(\log(\rho_1)/\log(\alpha_2))/(\log(\alpha_1) - \log(\rho_1));$$

where C_8 is computed as in (2.54). Then, by Theorem 2.19, $h^{a,0}(T_n(x_n))$ approximates $P_{x_n}(\tau_n > \tau_0)$ with relative error converging to a level bounded by $C_8 - 1 = 0.3607$.

We can reduce this error by using further Y-harmonic functions given by Propositions 2.3 and 2.24 and constructing an approximation of the form (2.56). We note $\beta_1(0.7) = 0.34610$ and therefore, by an argument parallel to the proof of Proposition 2.24, we infer that $|\beta_1(\alpha)| \leq 0.34619$, $|(\beta_1(\alpha), \alpha)| < 1$ for $|\alpha| = 0.7$. Thus, $h_{\beta_1(\alpha)}$ is Y-harmonic and ∂B-determined for all $|\alpha| = 0.7$, and we can use this class of functions in improving our approximation of $P_y(\tau < \infty)$. Let us begin with using $K = 3$ additional Y-harmonic functions of this form in our approximation: for the α's let us take

$$\alpha_{1,j} = 0.7e^{i\frac{j}{4}2\pi}, j \in \{1, 2, 3 = K\}.$$

The resulting harmonic functions are

$$h_{\beta_1(\alpha_{1,j})}, j \in \{1, 2, 3 = K\}.$$

(see (2.4) and (2.8)).

Our approximation $h^{a,K}$ will be of the form

$$h^{a,K} = \Re(h^{a*,K}), \quad h^{a*,K} = \boldsymbol{h}_r + c_{1,0}h_{\rho_1} + \sum_{j=1}^{K} c_{1,j}h_{\beta_1(\alpha_{1,j})}.$$

One can choose the coefficients $c_{1,j}$, $j \in \{0,1,2,3=K\}$ in a number of ways, for example, by minimizing L_p errors. Here we will proceed in the following simple way: the ideal situation would be $h^{a,K}(y) = 1$ for all $y \in \partial B$, which would mean $h^{a,K}(y) = P_y(\tau < \infty)$, but this will not hold in general. We will instead require that this identity holds for $y = (k,k)$, $k = 0,1,2,3$. This leads to the following four dimensional linear equation:

$$1 = h^{a,K}(k,k) = \boldsymbol{h}_r((k,k)) + c_{1,0}h_{\rho_1}(k,k) + \sum_{j=1}^{3} c_{1,j}h_{\beta_1(\alpha_{1,j})}(k,k), \tag{2.61}$$

$k = 0,1,2,3$; Solving (2.61) gives

$$c_{1,0} = 7.80744 - 0.12974i, \quad c_{1,1} = -0.25880 + 1.46155i,$$
$$c_{1,2} = -0.26358 - 0.01349i, \quad c_{1,3} = 0.17597 + 0.01433i.$$

Once the approximation is computed, following Proposition 2.25 one can easily compute its relative error in approximating $P_y(\tau < \infty)$ by computing

$$\max_{y\in\partial B} |h^{a*,K}(y) - 1|.$$

That $\max_{j=1,2,..,K}(r, \rho_1, r^2/\rho_1, |\alpha_{1,j}|, |\beta_1(\alpha_{1,j}), \boldsymbol{\alpha}(\beta_1(\alpha_{1,j}), \alpha_{1,j})) < 1$ implies $\text{argmax}_{y\in\partial B}|h^{a*,K}(y) - 1|$ is finite. For $h^{a*,K}$ computed above, the maximizer turns out to be $y^* = (4,4) = (K+1, K+1)$ and the maximum approximation error is

$$c^* = \max_{y\in\partial B} |h^{a*,K}(y) - 1| = |h^{a,K}((y^*)) - 1| = 0.17764. \tag{2.62}$$

The graph of the approximation error $|h^{a*,K}(y) - 1|$, $y \in \partial B$ is shown in Figure 2.5.

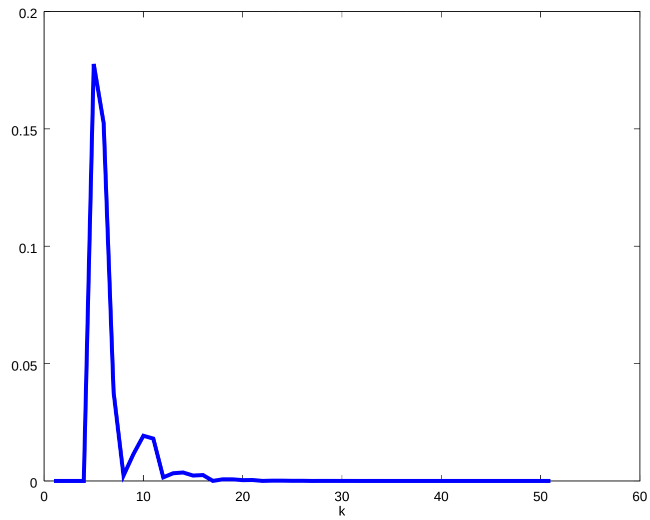

Figure 2.5: $|h^{a*,K}(y) - 1|$ as a function of $y = (k, k)$

By Proposition 2.25

$$\left| \frac{P_y(\tau < \infty) - h^{a,0}(y)}{P_y(\tau < \infty)} \right| \le c^*.$$

Theorem 2.19 now implies that $h^{a,K}(T_n(x_n))$ approximates $P_{x_n}(\tau_n < \tau_0)$ with relative error bounded by $c^* = 0.17764$ for n large. Therefore, in improving our approximation from $h^{a,0}$ of (2.60) to $h^{a,3}$ by adding three Y-harmonic functions of the form $h_{\beta_1(\alpha_{1,j})}$ to the approximating basis, the relative error decreases from $c_0^* = 0.3607$ to $c_0^* = 0.17764$. Figure 2.6 shows the level curves of $-\frac{1}{n}\log h^{a,3}(T_n(x))$ and $-\frac{1}{n}\log P_x(\tau_n < \tau_0)$ (the latter computed numerically via iteration of the harmonic equation satisfied by $P_x(\tau_n < \tau_0)$) for $n = 60$; the level curves overlap completely except along ∂_1, as suggested by our analysis.

To illustrate how the approximation error decreases when K increases, let us repeat the computation above with $K = 20$. The resulting maximum relative error turns out to be:

$$c^* = \max_{y \in \partial B} |h^{a*,20}(y) - 1| = |h^{a*,20}((21, 21)) - 1| = 1.6211 \times 10^{-3}.$$

The probability $P_{(4,0)}(\tau_{60} < \tau_0)$, computed numerically, equals 4.6658×10^{-17}, the best approximation of this quantity computed above is $h^{a,20}(56, 0) = 5.2 \times 10^{-17}$. The discrepancy arises from the proximity of $(4, 0)$ to ∂_1. As we move away from the ∂_1, these quantities get closer $P_{(10,0)}(\tau_{60} < \tau_0) = 3.3303 \times 10^{-15}$, $h^{a,20}(50, 0) =$

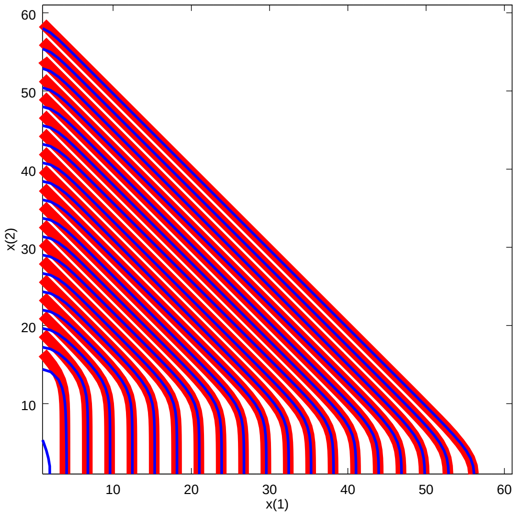

Figure 2.6: Level curves of $-\frac{1}{n}\log h^{a,0}(T_n(x))$ and $-\frac{1}{n}\log P_x(\tau_n < \tau_0)$, $n = 60$

3.3358×10^{-15}, compatible with the maximum relative error computed above. Figure 2.7 shows $K \mapsto h^{a,K}(50,0)$ and $P_{(10,0)}(\tau_{60} < \tau_0)$ (the flat line), drawn at 10^{-15} scale.

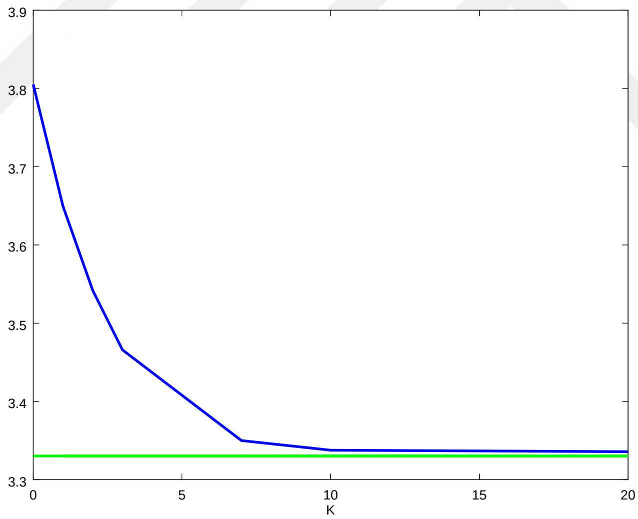

Figure 2.7: $K \mapsto h^{a,K}(50,0)$ and $P_{(10,0)}(\tau_{60} < \tau_0)$

CHAPTER 3

APPROXIMATION OF THE GREATEST COMPONENT OF THE EXIT POINT

A well studied quantity in computer science is the following expectation:

$$\mathbb{E}_x[\max(X_{\tau_n}(1), X_{\tau_n}(2))], \tag{3.1}$$

this is the expectation of the greatest component of the exit point. If X models two stacks working in parallel on a shared memory, this is the length of the longest stack at the time of overflow. The study of this expectation was introduced in [41, section 2.2.2, exercise 13]; a great amount of literature exists on this expecation, see [16, 28, 34, 46, 47, 68]. Some of these works are reviewed in Chapter 5. To the best of our knowledge the problem of approximating this expectation, in the general case (i.e., when the jump probabilities are arbitrarily chosen) remains an open problem. In the present chapter we discuss how the approach of the present work can be used to construct approximations in the stable case.

The Markov property of X implies

$$\begin{aligned}&\mathbb{E}_x[\max(X_{\tau_n}(1), X_{\tau_n}(2))]\\ &\quad = P_x(\tau_0 < \tau_n)\mathbb{E}_0[\max(X_{\tau_n}(1), X_{\tau_n}(2))] + \mathbb{E}_x[1_{\{\tau_n < \tau_0\}} \max(X_{\tau_n}(1), X_{\tau_n}(2))].\end{aligned}$$

The same idea applied at $x = (0, 0)$ gives

$$\begin{aligned}&\mathbb{E}_0[\max(X_{\tau_n}(1), X_{\tau_n}(2))]\\ &\quad = P_0(\tau_0 < \tau_n)\mathbb{E}_0[\max(X_{\tau_n}(1), X_{\tau_n}(2))] + \mathbb{E}_0[1_{\{\tau_n < \tau_0\}} \max(X_{\tau_n}(1), X_{\tau_n}(2))].\end{aligned}$$

Therefore, the computation of the expecation (3.1) can be reduced to the computation of $P_x(\tau_n < \tau_0)$ and $\mathbb{E}_x[1_{\{\tau_n<\tau_0\}} \max(X_{\tau_n}(1), X_{\tau_n}(2))]$. The previous chapters were devoted to the approximation of $P_x(\tau_n < \tau_0)$; in this chapter we would like to point out how the approach of the earlier chapters can be used to approximate

$$f_n(x) \doteq \mathbb{E}_x[1_{\{\tau_n<\tau_0\}} \max(X_{\tau_n}(1), X_{\tau_n}(2))]. \tag{3.2}$$

f_n takes the value

$$f_n(x) = \max(x_1, x_2)$$

for $x \in \partial A_n$. Note that this is not constant on ∂A_n; it depends both on x and n (e.g., $f_n((n, 0)) = n$.). This is in contrast to $P_x(\tau_n < \tau_0)$ which takes the constant value 1 on ∂A_n. The graph of f_n on ∂A_n is shown in Figure 3.1. This change leads to several

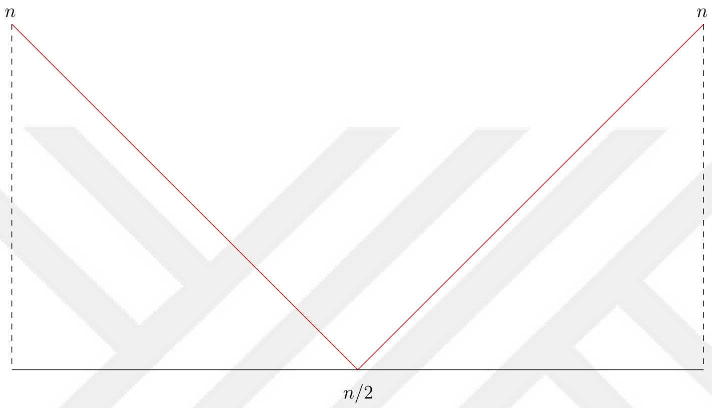

Figure 3.1: f_n on ∂A_n

difficulties in applying our approach to the approximation of (3.2):

1. in the approximation of f_n, there is no fixed limit problem; the limit problem depends on n,

2. the value of f_n on ∂A_n gives the limit boundary condition on ∂B only up to (n, n); there is no obvious extension of f_n to all of ∂B- the simplest extension would be to assign the value 0 for $(k, k) \in \partial B$ for $k > n$,

3. the function f_n is piecewise linear on ∂A_n; none of the Y-harmonic functions we have constructed in the earlier chapters have linear behavior on ∂A_n: we either have constant values or exponential growth or decay.

We propose to deal with these issues as follows:

1. construct a Y-harmonic function h_l that equals approximately function $(y_1, y_2) \mapsto y_2$ on ∂B,

2. approximate $f_n(T_n(y))$ on ∂B with $(y_1, y_2) \mapsto n - y_1$,

3. approximate $(y_1, y_2) \mapsto n - y_1$ on ∂B, using h_l and the Y-harmonic functions constructed in the earlier chapters.

The functions $(y_1, y_2) \mapsto n - y_1$ and $f_n(T_n(y))$ equal each other only up to $(n/2, n/2)$ on ∂B (See Figure 3.2). Therefore, the resulting approximation will be good only for $y \in \partial B$ with $y_1 = y_2 \leq n/2$, and, we expect the resulting approximation, when extended to all of A_n, to be good only for points away from the upper half of ∂A_n (i.e., for points x below the diagonal line $x(1) = x(2)$; this indeed turns out to be the case, see Figure 3.3).

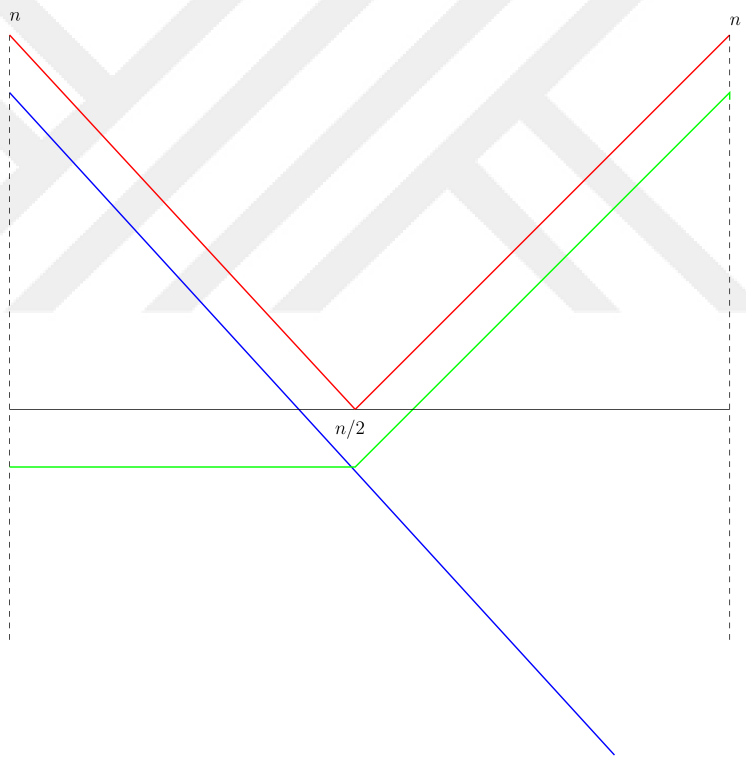

Figure 3.2: The approximation and the error

The next section constructs a Y-harmonic function having linear behavior on ∂B. In Section 3.2 we use this function and the Y-harmonic functions of the previous chapters to construct an approximation of f_n; as expected, the approximation performs well away from the upper section of ∂A_n.

3.1 Harmonic function resulting from diferentiation

Remember that our Y-harmonic function are of the form

$$\begin{aligned} h_\beta(y) &= \beta(\alpha)^{y_1-y_2}\left(C(\beta(\alpha),\boldsymbol{\alpha}(\alpha))\alpha^{y_2} - C(\beta(\alpha),\alpha)\boldsymbol{\alpha}(\alpha)^{y_2}\right) \\ &= \beta(\alpha)^{y_1-y_2}[(\beta,\alpha),(y_2,y_2)], \end{aligned}$$

where (β,α) and $(\beta,\boldsymbol{\alpha})$ are conjugate points on the characteristic surface. A natural way to introduce linear (in y) terms into this function is to differentiate it with respect to α. This is what we will do. Let an overbar denote partial differentiation with respect to α: e.g., $\bar{\boldsymbol{\alpha}}$ denotes $\frac{\partial\boldsymbol{\alpha}}{\partial\alpha}$. The chain rule gives

$$\begin{aligned} \frac{d}{d\alpha}h_\beta &= (y_1-y_2)\beta^{y_1-y_2-1}\bar{\beta}[(\beta,\alpha),(y_2,y_2)] \\ &\quad + \beta^{y_1-y_2}\left(\left(\frac{\partial C}{\partial\beta}\bar{\beta} + \bar{C}\left(\bar{\boldsymbol{\alpha}} + \frac{\partial\boldsymbol{\alpha}}{\partial\beta}\bar{\beta}\right)\right)\alpha^{y_2} + y_2 C\alpha^{y_2-1}\right. \\ &\quad \left. - \left(\frac{\partial C}{\partial\beta}\bar{\beta} + \bar{C}\right)\boldsymbol{\alpha}^{y_2} - C\boldsymbol{\alpha}^{y_2-1}y_2\left(\bar{\boldsymbol{\alpha}} + \frac{\partial\boldsymbol{\alpha}}{\partial\beta}\bar{\beta}\right)\right). \end{aligned}$$

To obtain linear behaviour on ∂B, it suffices to set $\alpha = 1$; the last display reduces to the following for $\alpha = 1$ on ∂B, i.e., for $y_1 = y_2$:

$$\begin{aligned} h_l \doteq \frac{d}{d\alpha}h_{\beta(1)} &= \left(\left(\frac{\partial C}{\partial\beta}\bar{\beta} + \bar{C}\left(\bar{\boldsymbol{\alpha}} + \frac{\partial\boldsymbol{\alpha}}{\partial\beta}\bar{\beta}\right)\right) + y_2 C\right. \\ &\quad \left. - \left(\frac{\partial C}{\partial\beta}\bar{\beta} + \bar{C}\right)\boldsymbol{\alpha}^{y_2} - C\boldsymbol{\alpha}^{y_2-1}y_2\left(\bar{\boldsymbol{\alpha}} + \frac{\partial\boldsymbol{\alpha}}{\partial\beta}\bar{\beta}\right)\right); \end{aligned} \tag{3.3}$$

note that we have a y_2C term here, i.e., a term with linear growth. This was our aim. In the next subsection we compute all of the partial derivatives listed above and simplify them. These will be used in the numerical study of Section 3.2.

3.1.1 Simplification of the partial derivatives

Formulas for the derivates appearing in the last display are

$$\bar{C} = \frac{\partial C}{\partial \alpha} = \mu_2 \frac{\beta}{\alpha^2}, \quad \frac{\partial C}{\partial \beta} = -\frac{\mu_2}{\alpha} \quad \bar{\boldsymbol{\alpha}} = -\frac{\beta^2}{\rho_2 \alpha^2}, \quad \frac{\partial \boldsymbol{\alpha}}{\partial \beta} = \frac{2\beta}{\rho_2 \alpha},$$

$$\bar{\beta} = -\frac{1}{2}\left(\frac{\bar{\Delta}}{2\sqrt{\Delta}}\left(\frac{\mu_2}{\alpha} + \mu_1\right)^{-1} + \left(1 - \Delta^{1/2}\right)\left(\frac{\mu_2}{\alpha} + \mu_1\right)^{-2} \mu_2 \alpha^{-2}\right),$$

$$\bar{\Delta} = 4\frac{\mu_2}{\alpha^2}(\lambda_1 + \lambda_2 \alpha) - 4\lambda_2\left(\frac{\mu_2}{\alpha} + \mu_1\right).$$

Define $\lambda_s = \lambda_1 + \lambda_2$, $\mu_s = \mu_1 + \mu_2$ and $r = \lambda_s/\mu_s$. Substituting $\alpha = 1$ in these give

$$\begin{aligned}
\Delta(1) &= 1 - 4(\mu_2 + \mu_1)(\lambda_1 + \lambda_2) = 1 - 4\mu_s\lambda_s \\
&= 1 - 4\mu_s(1 - \mu_s) = 1 - 4\mu_s + \mu_s^2 \\
&= (2\mu_s - 1)^2 = (2\mu_s - \lambda_s - \mu_s)^2 \\
&= (\mu_s - \lambda_s)^2.
\end{aligned}$$

$$\begin{aligned}
\beta(1) &= \frac{1 - \sqrt{\Delta(1)}}{2(\mu_1 + \mu_2)} = \frac{1 - \sqrt{(\mu_s - \lambda_s)^2}}{2(\mu_1 + \mu_2)} \\
&= \frac{1 - \mu_s + \lambda_s}{2(\mu_1 + \mu_2)} = \frac{(1 - \mu_1 - \mu_2) + (\lambda_1 + \lambda_2)}{2(\mu_1 + \mu_2)} \\
&= \frac{(\lambda_1 + \lambda_2) + (\lambda_1 + \lambda_2)}{2(\mu_1 + \mu_2)} = \frac{\lambda_1 + \lambda_2}{\mu_1 + \mu_2} \\
&= r.
\end{aligned}$$

$$\begin{aligned}
\bar{\beta}(1) &= -\frac{1}{2}\left(\frac{4(\mu_2\lambda_s - \lambda_2\mu_s)}{2(\mu_s - \lambda_s)\mu_s} - (1 - \mu_s + \lambda_s)\frac{\mu_2}{\mu_s^2}\right) \\
&= -\left(\frac{\mu_2\lambda_s - \lambda_2\mu_s}{(\mu_s - \lambda_s)\mu_s} - \frac{\lambda_s\mu_s}{\mu_s^2}\right) = -\frac{1}{\lambda_s}\left(\frac{\mu_2\lambda_s - \lambda_2\mu_s}{\mu_s - \lambda_s} - r\mu_2\right) \\
&= -\frac{\mu_2}{\mu_s}\left(\frac{r\mu_s - \rho_2\mu_s}{\mu_s - \lambda_s} - r\right) = -\frac{\mu_2}{\mu_2}\left(\frac{\mu_s(r - \rho_2)}{\mu_s - \lambda_s}\right) \\
&= -\frac{\mu_2}{\mu_s}\left(\frac{r - \rho_2}{1 - r} - r\right).
\end{aligned}$$

$$\bar{C} = \mu_2 r, \frac{\partial C}{\partial \beta} = -\mu_2, \bar{\boldsymbol{\alpha}} = -\frac{r^2}{\rho_2},$$

$$\frac{\partial \boldsymbol{\alpha}}{\partial \beta} = \frac{2r}{\rho_2}, \bar{\Delta} = 4(\mu_2\lambda_s - \lambda_2\mu_s).$$

The first term in equation 3.3 we have $\frac{\partial C}{\partial \beta}(r, \boldsymbol{\alpha})$ and $\bar{C}$ which are evaluted at $\alpha = 1$ as

$$\bar{C}(r, \boldsymbol{\alpha}) = \mu_2 \frac{r}{(\frac{r}{\alpha\rho_2})^2} = \mu_2 \frac{r}{\frac{r^4}{\rho_2^2\alpha^2}} = \mu_2 \frac{\rho_2^2\alpha^2}{r^3}\Big|_{\alpha=1} = \mu_2 \frac{\rho_2^2}{r^3},$$

$$\frac{\partial C}{\partial \beta}(r, \boldsymbol{\alpha}) = -\frac{\mu_2}{\frac{r^2}{\alpha\rho_2}} = -\frac{\mu_2\rho_2\alpha}{r^2}\Big|_{\alpha=1} = -\frac{\lambda_2}{r^2}.$$

The first term in 3.3 is $\left(\frac{\partial C}{\partial \beta}\bar{\beta} + \bar{C}\left(\bar{\boldsymbol{\alpha}} + \frac{\partial \boldsymbol{\alpha}}{\partial \beta}\bar{\beta}\right)\right)$ which equals to

$$\begin{aligned}
&-\frac{\lambda_2}{r^2}\bar{\beta}(1) + \mu_2 \frac{\rho_2^2}{r^3}\left(-\frac{r^2}{\rho_2} + \frac{2r}{\rho_2}\bar{\beta}(1)\right) = -\frac{\lambda_2}{r^2}\bar{\beta}(1) + \mu_2 \frac{\rho_2^2}{r^3}\left(-r^2 + 2r\bar{\beta(1)}\right) \\
&= -\frac{\lambda_2}{r^2}\bar{\beta}(1) + \frac{\lambda_2}{r^3}(-r^2 + 2r\bar{\beta}(1)) = -\frac{\lambda_2}{r^2}\bar{\beta}(1) - \frac{\lambda_2 r^2}{r^3} + \frac{2\lambda_2\bar{\beta}(1)r}{r^3} \\
&= -\frac{\lambda_2\bar{\beta}(1)}{r^2} - \frac{\lambda_2}{r} + \frac{2\lambda_2\bar{\beta}(1)}{r^2} = \frac{\lambda_2\bar{\beta}(1)}{r^2} - \frac{\lambda_2}{r} = \frac{\lambda_2}{r}\left(\frac{\bar{\beta}(1)}{r} - 1\right).
\end{aligned}$$

The second term is

$$y_2 C(\beta, \boldsymbol{\alpha}) = y_2\mu_2(1 - \frac{\beta}{\boldsymbol{\alpha}}) = y_2\mu_2\left(1 - \frac{\beta}{\frac{\beta^2}{\alpha\rho_2}}\right) = y_2\mu_2\left(1 - \frac{\alpha\rho_2}{\beta}\right)\Big|_{\alpha=1} = y_2\mu_2\left(1 - \frac{\rho_2}{r}\right).$$

The third term is

$$-\left(\frac{\partial C}{\partial \beta}(r,1)\bar{\beta}(1) + \bar{C}(r,1)\right)\boldsymbol{\alpha}^{y_2} = -(-\mu_2\bar{\beta}(1) + \mu_2 r)\boldsymbol{\alpha}^{y_2} = \mu_2(\bar{\beta}(1) - r)\left(\frac{r^2}{\rho_2}\right).$$

The last term is

$$-C(\beta, \alpha)\boldsymbol{\alpha}^{y_2-1}y_2\left(\bar{\boldsymbol{\alpha}} + \frac{\partial \boldsymbol{\alpha}}{\partial \beta}\bar{\beta}\right) = \frac{\mu_2}{\rho_2}r(1-r)\left(r - 2\bar{\beta}(1)\right)\boldsymbol{\alpha}^{y_2-1}y_2.$$

Then overall, our function at the boundary is:

$$\frac{d}{d\alpha}[(r,1),(y_2,y_2)] = \frac{\lambda_2}{r}\left(\frac{\bar{\beta}(1)}{r} - 1\right) + y_2\mu_2\left(1 - \frac{\rho_2}{r}\right) + \left(\mu_2\bar{\beta}(1) - \mu_2 r\right)\boldsymbol{\alpha}^{y_2}$$
$$+ \frac{\mu_2}{\rho_2} r(1-r)(r - 2\bar{\beta}(1))\boldsymbol{\alpha}^{y_2-1}y_2.$$

3.2 Numerical example

In this section we construct, using the Y-harmonic function h_l of 3.3 and the Y-harmonic function h_β of the previous chapters to construct an approximation of the expectation $f_n(x)$ of (3.2) using the approach outlined at the beginning of this chapter.

Our approximation of $f_n(x)$ will be of the form $h^{exp,K}(T_n(x))$ where

$$h^{exp,K} = \Re(h^{exp*,K}) = n\boldsymbol{h}_r - \frac{1}{1-\rho_2/r}h_l + c_0 h_{\rho_1} + \sum_{k=1}^{K} c_k h_{\beta_1(\alpha_k)}$$

where

$$\alpha_k = Re^{ij/K2\pi}, 0 < R < 1,$$

and the coefficient of h_l is chosen so that the y_2 term in $-\frac{1}{1-\rho_2/r}h_l$ has a -1 coeffient on ∂B. The key difference between $h^{a,K}$ (used in the approximation of $P_x(\tau < \infty)$) and $h^{exp,K}$ above is the new harmonic function h_l, which introduces the y_2 term on the boundary ∂B.

Our approach to the choice of the constants c_0, and c_k will be the same as in the previous section: choose them so that $h^{exp,K}(y_2, y_2) = n - y_2$ for $y_2 \leq K$. This gives a $(K+1)\times(K+1)$ linear system to be solved for c_0, c_1,...,c_K.

To see how well this approximation works, let us try it on a numerical example. For parameter values we use the same parameter values used in the previous chapters: $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\mu_1 = 0.2$, $\mu_2 = 0.5$; as we know from Chapter 2, $R = 0.7$ ensures that all $h_{\beta_1(\alpha_k)}$ are ∂B-determined with $|\beta_1(\alpha_k)| < 1$; for K we take $K = 20$, i.e., we use additional 20 Y-harmonic functions of the form $h_{\beta_1(\alpha)}$ where $|\alpha| = 0.7$. For n we take 60.

Figure 3.3 show the level curves of $-\frac{1}{n}\log f_n$ and $-\frac{1}{n}\log h^{exp,K}(T_n(\cdot))$(thin, curving away from the x_2-axis near the top corner); to compute f_n numerically, we iterate the harmonic equation satisfied by f_n until it converges, for $n = 60$, less than 1000 iterations guarantee convergence. The level curves overlap away from the upper half of A_n (above the line $x_1 = x_2$); this is as expected: by construction, f_n and $h^{exp,K}(T_n(\cdot)$ overlap only upto $(n/2, n/2)$ on on ∂A_n (see Figure 3.2).

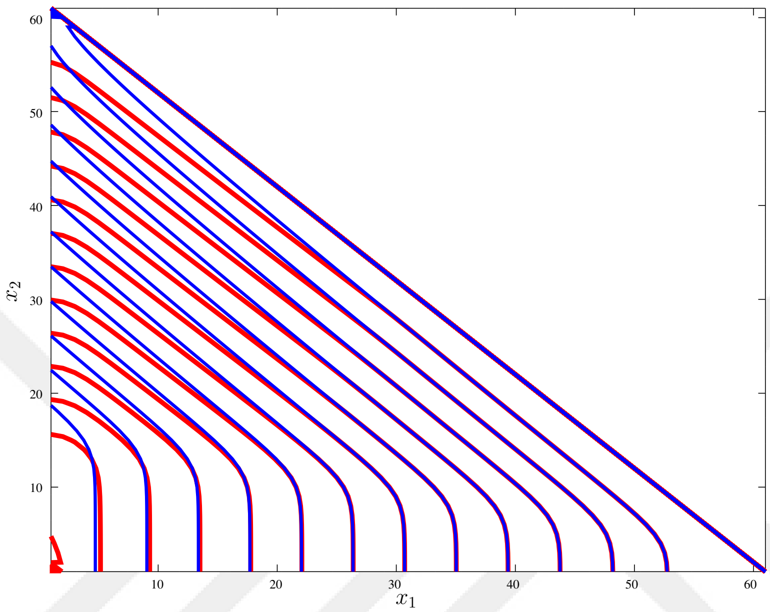

Figure 3.3: Level curves of $-\frac{1}{n}\log f_n$ and $-\frac{1}{n}\log h^{exp,K}(T_n(\cdot))$

The graphs of $\log(f_n)$ and $\log(h^{expk,K}(T_n(\cdot))$ along $x_2 = 0$ is given in Figure 3.4; as is clear from the figure, $\log(h^{expk,K}(T_n(\cdot))$ provides an excellent approximation of $\log(f_n)$.

As x_2 increases the quality of the approximation deteriorates for smaller values of x_1, see Figure 3.5, which is as expected.

Giving a theoretical proof of the goodness of these approximations and their extension to $x(2) > x(1)$ remain for future work.

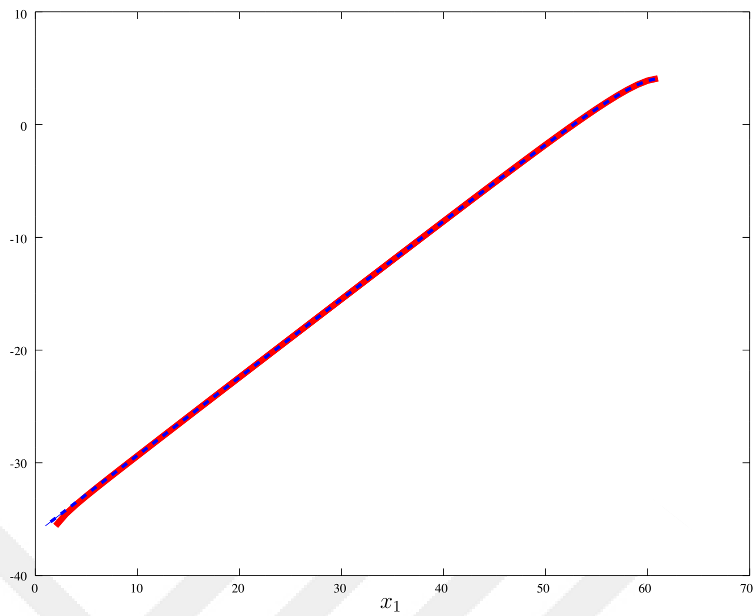

Figure 3.4: $\log(h^{exp,K}(n - x_1, x_2))$ (dashed line) and $\log(f_n)$ on $x_2 = 0$

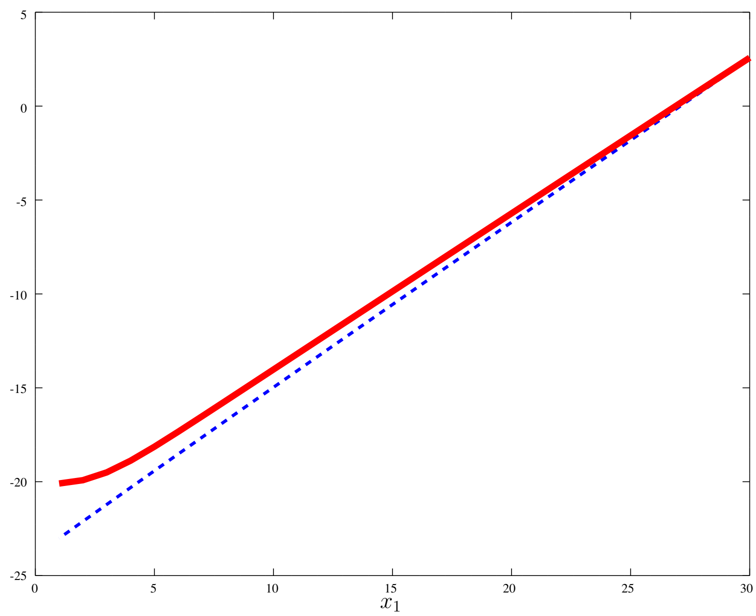

Figure 3.5: $\log(h^{exp,K}(n - x_1, x_2))$ (dashed line) and $\log(f_n)$ on $x_2 = 30$

CHAPTER 4

APPLICATIONS TO INSURANCE AND FINANCE

In the last decade research has emerged modeling systems of companies; this corresponds to multidimensional processes, see e.g. , [9, 63] and references in these works. Constraints to these proceses can model dividend payment [32] or restrictions on short selling. Therefore, constrained random walks can also serve as models of financial systems. In this chapter we provide two examples giving financial interpretation to our results.

4.1 Probability of low total reserves of an insurance system

The simplest possible insurance system consists of of two insurance companies A and B. One can model the reserves of these two companies as a process X in $\mathbb{R}^2$, the first [second] component representing the reserves of A [B]. X is a discrete time process, each step representing a single period (a month, quarter, year, etc.). A common simplification [67, Chapter 5] is to assume the reserves to be integer valued (where one unit can represent, e.g., total premiums collected in a period) in which case we get a process in $\mathbb{Z}^2$; in the rest of this chapter we will make this assumption; yet another common simplication ([67, Chapter 5]) is to assume that statistically, the quarters represented by the increments of X are independent of each other. This leads to a random walk model for X. In general, the distribution of the increments will depend on great number of things: the revenues of the companies, the regulatory environment within which they operate, agreements with them etc. so in general it can take arbitrary values, an example is given in Figure 4.1.

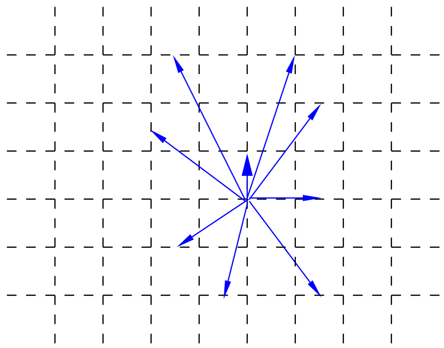

Figure 4.1: Possible increments of X can in general be arbitrary

So far X is unconstrained. Constraints on X can be introduced to model dividend payments; suppose that A [B] pays dividends of one unit whenever its reserves hit k_1 [k_2] at the end of a period; these payments correspond to contraints on the lines $\{x : x(i) = k_i\}$, $i = 1, 2$. We implement these constraints using the constraining map

$$\pi(x, v) \doteq \begin{cases} v, & \text{if } x(i) + v(i) < k_i, i = 1, 2, \\ 0, & \text{otherwise,} \end{cases}$$

and X is defined as the constrained process

$$X_k = X_{k-1} + \pi(X_k, I_k),$$

where I_k denotes the iid increments of the process X. The constraining boundaries of X are $\{x : x(i) = k_i\}$, $i = 1, 2$; these boundaries play the same role as the boundaries ∂_i did in the earlier chapters.

The stability of X in this context means that in each period both companies make money on average, i.e., both components of $\mathbb{E}[I_k]$ are strictly positive. When X is stable, i.e., when both companies are profitable on average, (k_1, k_2) becomes a recurrent point and X moves in cycles that restart each time it hits (k_1, k_2). As we recall (k_1, k_2) plays the same role that the origin $(0, 0)$ shown in the earlier chapters. Under this stability assumption a natural question about this system is as follows: what is the probability that in a cycle the system's total reserves go below a given thershold n. Let τ_n denote the first time $X(1) + X(2)$ equals n and let τ_0 denote first time it hits the point (k_1, k_2); the probability of total reserves going below n is represented by the probability

$$P_x(\tau_n < \tau_0); \tag{4.1}$$

geometrically this is the probability of X hitting the line $\{x : x(1)+x(2) = n\}$ before hitting the point (k_1, k_2), see Figure 4.2.

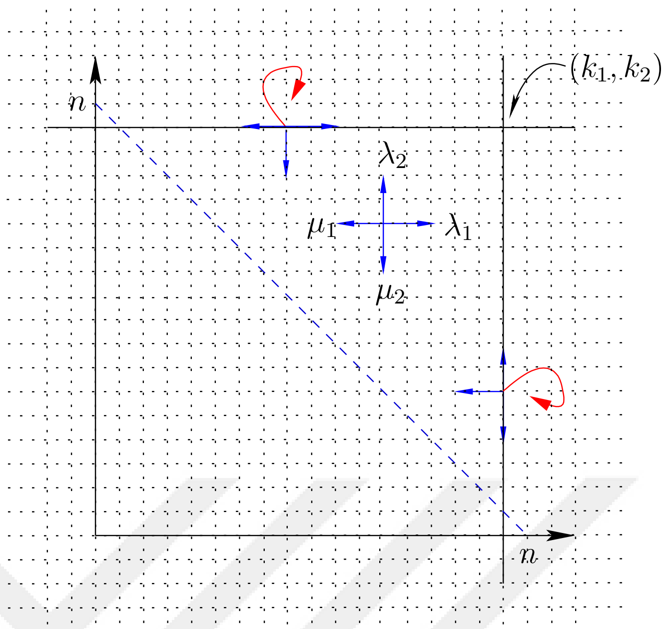

Figure 4.2: Constrained Random Walk for Insurance Model

If the increments take values in $\{(1, 0), (-1, 0), (0, 1), (0, -1)\}$ with probabilities μ_1, λ_1, μ_2, λ_2 the probability (4.1) is exactly the probability we have approximated in Theorem 2.19. This gives one possible application / interpretation of our results in insurance. The same framework can also be used to study credit risk; the next section discusses this possibility.

4.2 Application to Credit Risk

Consider the same setup as above, i.e., two companies A and B whose equity is modeled by the two dimensional random walk X. As before, let ∂_i denote the coordinate axes and σ_i the first time X hits ∂_i. In the present framework, σ_1 models the default time of company A and σ_2 models the default time of company B (see Figure 4.3) $\sigma_1 \wedge \sigma_2$ models the first time one of the companies in the system go bankrupt. If there are no dividend payments,

$$P_x(\sigma_1 \wedge \sigma_2 < \infty) \tag{4.2}$$

is the probability that this system ever goes bankrupt. The stability assumption implies that this probability is strictly less than 1, i.e., the problem is nontrivial. $P_x(\sigma_1 \wedge \sigma_2 < \infty)$ is of the same type as $P_y(\tau < \infty)$ for which we have developed formulas in Section 2.6 of Chapter 2. The main difference is that the exit boundaries are different; in Section 2.6 the exit boundary ∂B is $\{y : y(1) = y(2)\}$ whereas in (4.2) the exit boundary is $\partial_1 \cup \partial_2$. We think that the type of harmonic functions constructed in Section 2.6 can be extended to develop formulas for (4.2) for the exit boundary $\partial_1 \cap \partial_2$. This is an interesting research problem for future work.

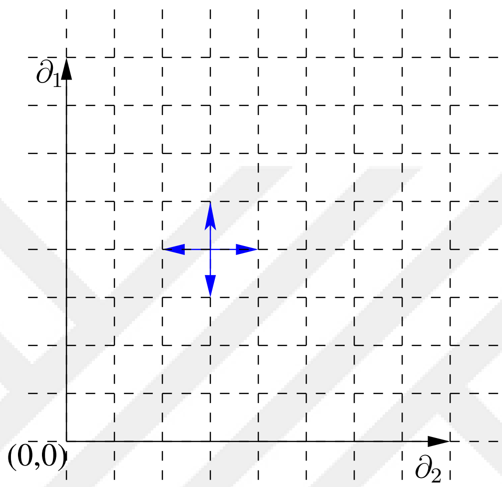

Figure 4.3: X hitting $\partial_1 \cup \partial_2$ is the first default time

Similar to the application above, allowing A and B to pay dividends whenever their equity hits levels k_1 and k_2 respectively corresponds to introducing reflecting boundaries at$\{x : (i) = k_i\}$ (see Figure 4.4).

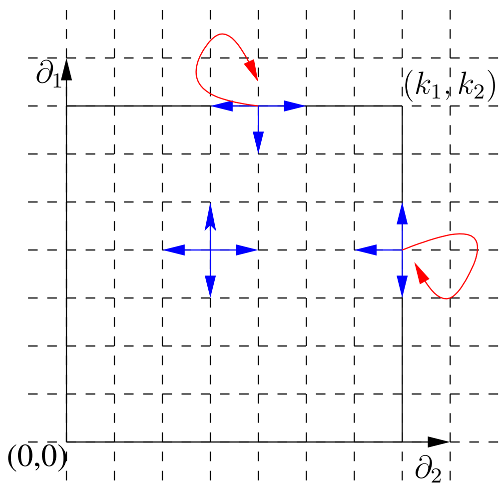

Figure 4.4: Constrained Simple Random Walk for Finance Application

Then once again, the stability assumption implies that X travels in cycles which restart every time it hits (k_1, k_2), then the probability

$$P_x(\sigma_1 \wedge \sigma_2 < \tau_0)$$

is the probability that the system goes bankrupt in the current cycle. This is exactly a problem of the type that we have studied in Chapter 2; the main difference once again being the different geometry of the exit boundary (rectangular for the current case, triangular in Chapter 2). As also pointed out in [64], we think that the analysis of the triangular case can be extended to the rectangular case using similar structures such as characteristic surfaces and conjugate points on them; an implementation of these ideas also remain for future work.

CHAPTER 5

LITERATURE REVIEW

There is a wide literature on the approximation of the probability $P_x(\tau_n < \tau_0)$ and other expectations and probabilities of constrained processes using many techniques including large deviations analysis and simulation, see, [1, 2, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 26, 28, 29, 30, 31, 33, 34, 35, 37, 38, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 61, 62, 66, 68].

In this section we review [25, 33, 36, 60, 64, 65, 28, 68]; for further literature review we refer the reader to [64].

[33] evaluates the performance of an IS estimator of rare event in tandem Jackson network and computes the large deviations decay rate of p_n for the initial points $(1, 0)$. A popular heuristic in importance sampling is to use a static change of measure implied by large deviations analysis (in the approximation of p_n for two tandem random walk, this corresponds to interchanging λ with the smaller of the μ_i). It was observed in this work that this heuristic can perform very poorly in the estimation of p_n in multidimensional constrained random walks.

The large deviations result in [33] is as follows: for a d-dimensional constrained random walk representing a Jackson network, we have

$$\lim_{n\to\infty} -\frac{1}{n}\log p_n((1,0,...,0)) = \bigwedge_{i=1}^{d} -\log \rho_i, \tag{5.1}$$

where p_n is as before the probability $P_{(1,0,\ldots,0)}(\tau_n < \tau_0)$, where

$$\rho_i = [\lambda(I-P)^{-1}]_i/\mu_i,$$

P is the routing matrix of the network and λ is the vector of arrival rates. To establish this result, [33] proves

$$b_1 \rho_*^n n^{-1} \leq p_n \leq b_2 \rho_*^n (n+1)^d, \rho_* = \max_j \rho_j,$$

using a theorem from [3], which claims

$$\pi(\partial A_n) = c_0^{-1} \mathbb{E}_0 \left(\int_0^{\tau_0} 1_{\{X_t \in \partial A_n\}} dt \right),$$

where π is the stationary measure of π.

[36] computes the large deviation rate function for constrained random walks

The prior works that thesis has greatest connection to are [25, 60, 64, 65]. [25, 60] shows how to use, design and implement IS algorithms for the simulation of the probability p_n when the starting point x is near the origin. An Their methodology is based on subsolutions of the HJB equation arising from the limit analysis of the second moment of the IS estimator. This equation turns out to be the same that appears in the large deviations analysis of p_n. We have used the simple random walk version of this equation in Chapter 2, see (2.15) in our analysis of the relative error of the approximation as well as obtaining LD lowerbounds on the probability p_n.

The IS estimator of p_n is of the following form:

$$\widehat{p}_n = 1_{A_n} \prod_{k=0}^{T_n - 1} \frac{\Theta(Y(k+1))}{\bar{\Theta}^n(Y(k+1)|X^n(k))}.$$

The goal here is to choose the IS measure $\bar{\Theta}$ so that the second moment of this estimator is minimized, this leads to the optimization problem

$$V_n(x) = \inf \mathbb{E}^{\mathbb{P}}(\widehat{p}_n) = \inf \mathbb{E}^{\mathbb{P}} \left(1_{A_n} \prod_{k=0}^{T_n - 1} \frac{\Theta(Y(k+1))}{\bar{\Theta}^n(Y(k+1)|X^n(k))} \right).$$

This function satisfies a dynamics programming equation, which , for x in the interior,

$$V_n(x) = \inf \sum_{i=0}^{2} V_n \left(x + \frac{1}{n} v_i \right) \frac{\bar{\Theta}(v_i)}{\Theta(v_i)} \Theta(v_i).$$

Applying $-\frac{1}{n} \log$ to this equation and taking limits reduces this equation to (2.15). [25, 60] uses subsolutions to this limit equation to construct its asymptotically optimal IS algorithms.

The present work is a continuation / extension of [65], which treats the approximation of $P_x(\tau_n < \tau_0)$ for the two dimensional tandem walk; [65] derives the following explicit approximate formula for this probability:

$$\begin{aligned} P_x(\tau_n < \tau_0) &\approx P_y(\tau < \infty) \\ &= \rho_2^{y(1)-y(2)} + \frac{\mu_2 - \lambda}{\mu_2 - \mu_1}\rho^{y(1)-y(2)}\rho_1^{y(2)} + \frac{\mu_2 - \lambda}{\mu_1 - \mu_2}\rho^{y(1)-y(2)}\rho_1^{y(2)}. \end{aligned}$$

Section 6.1 gives a detail comparison of the results of this thesis and [65] covering the tandem walk.

[28] studies the expectation

$$\mathbb{E}_0[\max(X_{\tau_n}(1), X_{\tau_n}(2))],$$

for the simple constrained random walk using combinatorics techniques to count the trajectories of the walker; note that this work focuses on the initial condition $x = 0$. [28] transforms the walk in a triangle to a square, than the walk in a square reduces to a 1 dimensional simple random walk. For these techniques to work [28] focuses on two particular cases of the jump probabilities. In the first case authors set $p = \frac{1}{4}$, thus the probability of each increments equal to one over four, that is

$$\begin{aligned} P(I_i = (1,0)) = P(I_i = (0,1)) = \frac{1}{4}, \\ P(I_i = (-1,0)) = P(I_i = (0,-1)) = \frac{1}{4}, \end{aligned}$$

this is called as "metastable case". The second case is the "contracting case", where the jump probabilities are taken to be

$$\begin{aligned} P(I_i = (1,0)) = p < P(I_i = (-1,0)) = q, \\ P(I_i = (0,1)) = p < P(I_i = (0,-1)) = q. \end{aligned} \tag{5.2}$$

where $p \in \left(0, \frac{1}{4}\right)$. In this case the probability of a positive increment is less than the probability of a negative increment. Note that in both cases the jump probabilities are symmetric, i.e., there is no distinction between the queues. [28] proves that

$$A(p,n) = \mathbb{E}_0[\max(X_{\tau_n}(1), X_{\tau_n}(2))] \sim n0.67526...,$$

for the metastable case and

$$A(p,n) \sim \frac{3}{4}n,$$

for the contracting case.

[28] uses the variable m (rather than n as we do above) to denote buffer size (or equivalently the position of the exit boundary). In the rest of our review of [28] we will also use m to denote buffer size. Very briefly, the reductions mentioned above are performed as follows. Γ_m represents the graph on the integer interval $[1, ..., m]$ with forward edges $\{(x, x+1)|0 \leq x < m\}$, backward edges $\{(x, x-1)|0 < x \leq m\}$ and loops {(0,0) ;(m,m) }.

The set of all edges is

$$\{(x, x+1)|0 \leq x < m\} \cup \{(x, x-1)|0 < x \leq m\} \cup \{(0,0); (m,m)\}.$$

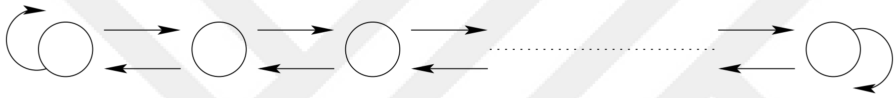

Figure 5.1: The graph Γ_m

$U_{k,n,t}$ denotes the set of paths of length n from vertex 0 to k in one-dimensional random walk containing t loops. $U_k(x, u)$ denotes the generating function of number of paths.

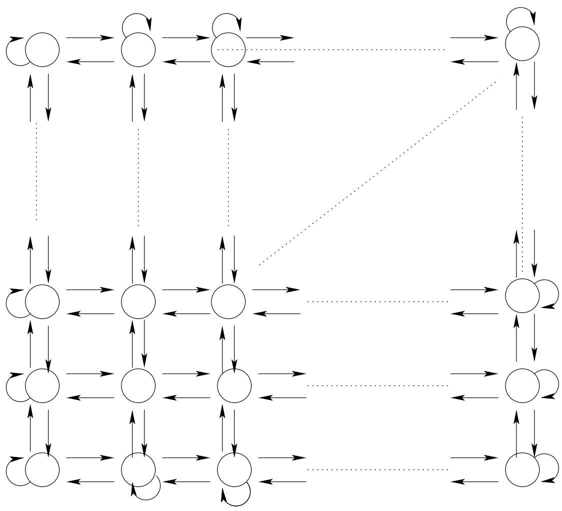

Figure 5.2: The multi-graph Λ_m

Λ_m shows the multi-graph whose vertices is $\{[x_1, x_2] : 0 \leq x_1, x_2 \leq m\}$ and the set of edges

$$\{([x_1,x_2],[|x_1+1|^-,x_2])\}\cup\{([x_1,x_2],[|x_1-1|^+,x_2])\}\cup$$
$$\{([x_1,x_2],[x_1,|x_2+1|^-])\}\cup\{([x_1,x_2],[x_1,|x_2-1|^+])\},$$

where $x\in\mathbb{Z}^2$ and

$$|x-1|^+=\begin{cases}x-1, & \text{if } x-1>0\\ 0, & \text{otherwise}\end{cases}$$

$$|x+1|^-=\begin{cases}x+1, & \text{if } m>x+1\\ m, & \text{otherwise}\end{cases}$$

$\theta_{k_1,k_2,n,t}$ denotes the set of paths of length n from the point $[0,0]$ to the point$[k_1,k_2]$ in two dimensional random walk containing t loops; $\hat{\theta}_{k_1,k_2}(x,u)$ denotes the exponential generating function of numbers of paths

$T_{k_1,k_2,n,t}$ denotes set of paths of length n from the point $[0,0]$ to the point$[k_1,k_2]$ in two dimensional random walk, which t are loops with $k_1+k_2<m$. The generating function of this paths is denoted by $T_{k_1,k_2}(x,u)$.

Lastly, π_{k_1,k_2} denotes the probability of hitting the absorbing barrier (k_1,k_2) where $k_1+k_2=m$. [28] estimates this probability by counting the trajectories of the random walk:

1. Estimate the generating function of triangular walks from walks in square lattice by using the reflection principle i.e.,

   $$T_{k_1,k_2}(x,u)=Q_{k_1,k_2}(x,u)-Q_{m-k_2,m-k_1}(x,u),$$

   when $k_1+k_2<m$.

   $T_{k_1,k_2}(x,u)=x(Q_{k_1-1,k_2}(x,u)+Q_{k_1,k_2-1}(x,u)-Q_{k_1+1,k_2}(x,u)+Q_{k_1,k_2+1}(x,u)),$

   when $k_1+k_2=m$.

2. Estimate the generating function of two dimensional counting walks in a square i.e.,

   $$\widehat{Q}_{k_1,k_2}(x,u)=\sum_{n,t\geq 0}Q_{k_1,k_2,n,t}u^t\frac{x^n}{n!}=\widehat{U}_{k_1}(x,u)\widehat{U}_{k_2}(x,u),$$

where

$$\widehat{U}_k(x,u) = \sum_{n,t\geq 0} U_{k,n,t} u^t \frac{x^n}{n!}.$$

3. Transform counting of walks in a square to shuffles of 1 dimensional walks over an interval i.e.,

$$U_k(x,u) = \sum_{n,t\geq 0} U_{k,n,t} x^n u^t.$$

4. Express the generating functions of the walks over an integer in terms of partial fraction decompostion i.e.,

$$U_k(x,u) = -\sum \frac{c_k(\varphi)}{(1-2x\cos\varphi)},$$

where

$$c_k(\varphi) = \begin{cases} 2\sin\varphi \frac{\sin(m-k+1)\varphi - u\sin(m-k)\varphi}{(m+2)\cos(m+2)\varphi - 2(m+1)u\cos(m+1)\varphi + mu^2\cos m\varphi}, & \text{if } \varphi \neq 0 \\ -\frac{1}{m+1}, & \text{otherwise} \end{cases}$$

The general probability of reaching absorbing state is given as

$$\pi_{k_1,k_2} = \left(\frac{p}{q}\right)^{\frac{m}{2}} T_{k_1,k_2}\left((pq)^{\frac{1}{2}}\left(\frac{q}{p}\right)^{\frac{1}{2}}\right),$$

for the metastable case the limiting distribution is

$$\lim_{m\to\infty} m\pi_{\lambda m,(1-\lambda)m} = 4\left(\frac{1}{2} + f(\lambda) + f(1-\lambda)\right);$$

where

$$f(x) = \sum_{n\geq 1} cosh\pi x \frac{\sinh n\pi x}{\sinh n\pi}.$$

The expected size of the largest stack can now be approximated by integrating with respect to f. For the contracting case it is shown that the probability of absorption converges to the uniform distribution:

$$\lim_{m\to\infty} m\pi_{\lambda m,(1-\lambda)m} = 1.$$

where $\lambda \in [\alpha,\beta]$ and $[\alpha,\beta]$ is a subinterval of $[0,1]$. Again, the expected size of the largest stack is now found by integrating with respect to the uniform distribution.

An early reference on the approximation of the expectation $\mathbb{E}_0[\max(X_{\tau_n}(1), X_{\tau_n}(2))]$, is [68] treating the symmetric case (5.2) where the jump probabilities of the first and second dimension are the same. As opposed to [28], [68] considers the unstable case where $p > q$ and also focuses on the initial condition $x = 0$.

The main idea in [68] is that after the first $\log m$ step, the probability of returning to the one of the reflecting boundaries is asymptotically 0, thus the analysis reduces to simple random walk in 2 dimensions with one absorbing barrier at $x(1) + x(2) = n$. Details are given below.

The approximation result in [68] is the following

$$A(p, n) = \frac{n}{2} + \sqrt{\frac{n}{2\pi(4p-1)}} + O\left(\frac{\log(n)}{\sqrt{n}}\right). \tag{5.3}$$

[68] proves (5.3) by 1) showing that one may ignore the constraints on the dynamics of X and 2) computing the same quantity for the unconstrained process (this is [68, Lemma 1]).

The first step of the argument has itself two steps: 1) a case covering when X_0 is sufficiently away from the constraining boundaries (this is [68, Lemma 2]) and 2) and argument when $X_0 = 0$; the second part runs the process for t number of steps and just uses the law of large numbers (with large deviations estimates) to show that the process will be in a region where the first part works (this is [68, Lemma 3]).

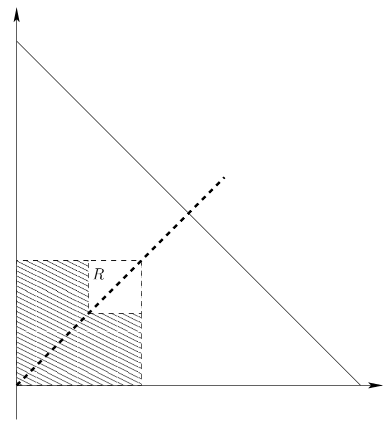

Figure 5.3: The region R

Define the region $R \subset \mathbb{Z}_+^2 : [\lceil \lambda' \log n \rceil, \lceil \lambda \log n \rceil + 1]^2$, where $0 < \lambda' < \lambda$ are constants that are chosen based on the parameter p and $\epsilon > 0$ is a small number, again chosen depending on the parameter p; the region R is shown in Figure 5.3; $\lambda' < \lambda$; so for n large enough R is nonempty and grows linearly with $\log(n)$. The dashed line in Figure 5.3 shows the average dynamics of X.

Let us give an summary of the above argument using Figure 5.3:

1. Use large deviations bounds to show that starting with $X_0 = 0$, with high probability X will be in the rectangle R at time $t = \lfloor \lambda \log(n) \rfloor + 1$. The contribution of paths with $X_t \notin R$ to $\mathbb{E}_0[|\max(X_\tau(1), X_\tau(2))|]$ is negligable.

2. For $X_0 = a \in R$, as far as the computation of $\mathbb{E}_x[|\max(X_\tau(1), X_\tau(2))|]$ is concerned, X can be replaced with the unconstrained process X'.

These two steps give the approximation (5.3).

CHAPTER 6

CONCLUSION

6.1 Comparison with the tandem case

This section compares the analysis and results of the current work to those of [64, 65] treating the approximation of the probability $P_x(\tau_n < \tau_0)$ for the constrained random walk representing two tandem queues, which has the increments $(1, 0)$, $(-1, 1)$ and $(0, -1)$. The main idea is the same for both walks: i.e., approximation of $P_x(\tau_n < \tau_0)$ by $P_y(\tau < \infty)$ and computing/ approximating the latter via harmonic functions constructed out of single and conjugate points on the characteristic surface. However, the assumptions, the results and the analysis manifest nontrivial differences. Let us begin with the assumptions:

Assumption $r^2/\rho_2 < 1$ In the tandem case $\beta_1(1) = \rho_2$ and the conjugate point of $(\rho_2, 1)$ is (ρ_2, ρ_1), therefore, the stability assumption automatically implies $\boldsymbol{\alpha}(r, 1) < 1$. For the parallel case, $\boldsymbol{\alpha}(r, 1)$ can indeed be greater than 1 if r and ρ_1 are close and ρ_2 is small; we therefore explicitly assume $r^2/\rho_2 < 1$. This assumption appears in two places: 1) in the convergence analysis, in the derivation of the bound (2.9) and 2) in the computation of $P_y(\tau < \infty)$ in Section 2.6. We think that the use of the assumption $r^2/\rho_2 < 1$ in the first case can be removed without much change from the arguments of the present and earlier works; the details remain for future work. We think that the computation of $P_y(\tau < \infty)$ when $r^2/\rho_2 > 1$ presents genuine difficulties, the treatment of which also remains for future work. Next we point out the differences in results:

Region where $P_y(\tau < \infty)$ is a good approximation for $P_x(\tau_n < \tau_0)$ That the tandem walk involves no jumps of the form $(-1,0)$ implies that $P_{T_n(x_n)}(\tau < \infty)$ provides an approximation of $P_{x_n}(\tau_n < \tau_0)$ with exponentially decaying relative error for all x away from 0; in contrast, the presence of the jump $(-1,0)$ in the parallel case, implies that the same approximation works only away from ∂_1 for the parallel walk case treated in the present work. This difference shows itself in the proofs of exponential decay of relative error, too, this is discussed below.

Explicit formula for $P_y(\tau < \infty)$ In the case of the tandem walk, the probability $P_y(\tau < \infty)$ can be explicitly represented as a linear combination of the harmonic functions h_{ρ_1} and h_{ρ_2} for all stable parameter values as long as $\mu_1 \neq \mu_2$; in the parallel case this only happens when $r^2 = \rho_1\rho_2$ (see Proposition 2.21). When $r^2 \neq \rho_1\rho_2$, h_r and h_{ρ_1} can only provide an approximation of $P_y(\tau < \infty)$ with bounded relative error (Proposition 2.22). This relative error can be reduced by adding into the approximation further ∂B-determined Y-harmonic functions (Proposition 2.25 and subsection 2.6.1).

The changes in argument from the tandem walk to the parallel walk are as follows:

Analysis of $P_x(\tau_n < \tau_0)$ In prior works [25, 60, 61, 64] the LD analysis of $P_x(\tau_n < \tau_0)$ and similar quantities are based on sub and supersolutions of the limit HJB equation, similar to the analysis given in subsection 2.3.1. In the present work, a novelty is the use of explicit subharmonic functions (Proposition 2.15) of the constrained random walk X in the proof of the upperbound Proposition 2.16.

Analysis of $P_x(\sigma_1 < \tau_n < \tau_0)$ The probability corresponding to $P_x(\sigma_1 < \tau_n < \tau_0)$ in the tandem case is $P_x(\sigma_1 < \sigma_{1,2} < \tau_n < \tau_0)$. For the proof of the exponential decay of the relative error, we need upperbound on these probabilities. Both papers develop these upperbound from subsolutions to a limit HJB equation. The subsolution consists of three pieces (one for each of the stopping times σ_1, $\sigma_{1,2}$ and τ_n) for the tandem walk, and two pieces for the parallel walk (one for each of the times σ_1 and τ_n). In the tandem case, the pieces of the subsolution are constructed from the subsolution for the probability $P_x(\tau_n < \tau_0)$, whereas in the parallel case a new piece is introduced based on the gradient r_4 of (2.39).

Analysis of $P_x(\bar{\sigma}_1 < \tau < \infty)$ The probability corresponding to $P_x(\bar{\sigma}_1 < \tau < \infty)$ in the tandem case is $P_x(\bar{\sigma}_1 < \bar{\sigma}_{1,2} < \tau < \infty)$. The special nature of the tandem walk allowed us to find upperbounds on this probability from the explicit formula we have for $P_y(\tau < \infty)$; this significantly simplified the analysis of the tandem walk case. For the parallel walk, we extended the analysis of $P_x(\sigma_1 < \tau_n < \tau_0)$, based on subsolutions, to $P_x(\bar{\sigma}_1 < \tau < \infty)$. In this, the most significant novelty is the analysis given Section 2.2, where we prove the existence of $z > 1$ such that $\mathbb{E}_z[z^\tau 1_{\{\tau<\infty\}}] < \infty$. For this, we introduce what we call $Y - z$-harmonic functions and provide methods of construction of classes of them from points on $1/z$-level characteristic surfaces, which are generalizations of characteristic surfaces.

6.2 Conclusion

The probability $P_y(\tau < \infty)$ approximates $P_x(\tau_n < \tau_0)$ well when x is away from ∂_1; as noted in the previous section, this is in contrast to the tandem case, where the approximation is good away from the origin. How can one extend the approximation to the region along ∂_1? A natural idea, already pointed out in [64] is to repeat the same analysis, but this time taking the corner $(0, n)$ as the origin of the Y process, i.e., to use the change of coordinate $y = T_n(x) = (x(1), n - x(2))$ to construct the Y process. Numerical calculations indicate that the resulting approximation will be accurate (i.e., exponentially decaying relative error) along ∂_1 between the points $(0, n)$ and $(0, \lfloor(1 - C_4)n\rfloor)$ (see (2.19) for the definition of C_4). We believe that arguments and computations parallel to the ones given in the present work would imply these results; the details are left for future work. We think that the extension of the approximation to the region along the line segment between $(0, 0)$ and $(0, \lfloor(1 - c_1)n\rfloor)$ requires further ideas and computations.

We expect the analysis linking $P_x(\tau_n < \tau_0)$ to $P_y(\tau < \infty)$ when $\rho_1 = \rho_2$ to be parallel to the analysis given in the current work. For the computation of $P_y(\tau < \infty)$, when $\rho_1 = \rho_2$, the case $\lambda_1 = \lambda_2$, $\mu_1 = \mu_2$ appears to be particularly simple. In this case, upon taking limits in (2.49) one obtains

$$P_y(\tau < \infty) = r^{y(1)-y(2)} + (1-r)r^{y(1)}(y(1) - y(2)),$$

where $r = \rho_1 = \rho_2$. A complete analysis of the computation of $P_y(\tau < \infty)$ when $\rho_1 = \rho_2$ remains for future work.

In subsection 2.6.1, the computation of $P_y(\tau < \infty)$ when $r^2 \neq \rho_1\rho_2$ proceeds as follows: 1) we first construct a candidate approximation $h^{a,K} = \mathfrak{R}(h^{a*,K})$ of $P_y(\tau < \infty)$ 2) we find an upperbound on the relative error of the approximation by finding the maximum of $|h^{a*,K} - 1|$ on ∂B. A natural question is the following: given a relative error bound, can we know apriori that an approximation having that maximum relative error can be constructed? If that is possible, how many Y-harmonic functions of the form given by Proposition 2.3 would we need? To answer these questions require a fine understanding of the functional analytic properties of the span of the ∂B-determined Y-harmonic functions given by Propositions 2.2 and 2.3. This appears to be a difficult problem because the functions given in these propositions don't have simple geometric properties, such as the orthogonality of the Fourier basis in L^2. A study of this problem remains for future work.

The exact formula for $P_y(\tau < \infty)$ for the tandem case has a remarkable extension to d dimensions; this is derived in [64] and is based on harmonic-systems, a concept defined in that work. We think that it is also possible, in the case of parallel queues, to obtain nontrivial harmonic systems in higher dimensions. A complete characterization of such systems and the question of under what conditions they would give a rich class of Y-harmonic functions to approximate $P_y(\tau < \infty)$ also remain challenging problems for future research.

REFERENCES

[1] M. Alanyali and B. Hajek, On large deviations in load sharing networks, Annals of Applied Probability, pp. 67–97, 1998.

[2] D. Aldous, *Probability approximations via the Poisson clumping heuristic*, volume 77, Springer Science & Business Media, 2013.

[3] S. Asmussen, *Applied Probability and Queues*, Wiley , New York., 1987.

[4] S. Asmussen, *Applied probability and queues*, volume 51, Springer Science & Business Media, 2008.

[5] S. Asmussen and P. W. Glynn, *Stochastic simulation: algorithms and analysis*, volume 57, Springer Science & Business Media, 2007.

[6] R. Atar and P. Dupuis, Large deviations and queueing networks: methods for rate function identification, Stochastic processes and their applications, 84(2), pp. 255–296, 1999.

[7] J. Blanchet, Optimal sampling of overflow paths in Jackson networks, Mathematics of Operations Research, 38(4), pp. 698–719, 2013.

[8] J. Blanchet, P. Glynn, and K. Leder, On lyapunov inequalities and subsolutions for efficient importance sampling, ACM Transactions on Modeling and Computer Simulation (TOMACS), 22(3), p. 13, 2012.

[9] J. Blanchet, J. Li, and Y. Shi, Stochastic risk networks: Modeling, analysis and efficient Monte Carlo, 2015.

[10] J. Blanchet and M. Mandjes, Rare event simulation for queues, Rare Event Simulation Using Monte Carlo Methods, pp. 87–124, 2009.

[11] J. H. Blanchet, K. Leder, and P. W. Glynn, Efficient simulation of light-tailed sums: an old-folk song sung to a faster new tune..., in *Monte Carlo and Quasi-Monte Carlo Methods 2008*, pp. 227–248, Springer, 2009.

[12] P.-T. D. Boer and V. F. Nicola, Adaptive state-dependent importance sampling simulation of markovian queueing networks, European Transactions on Telecommunications, 13(4), pp. 303–315, 2002.

[13] A. A. Borovkov and A. A. Mogul'skii, Large deviations for markov chains in the positive quadrant, Russian Mathematical Surveys, 56(5), pp. 803–916, 2001.

[14] J. Collingwood, R. D. Foley, and D. R. McDonald, Networks with cascading overloads, in *Proceedings of the 6th International Conference on Queueing Theory and Network Applications*, pp. 33–37, ACM, 2011.

[15] F. Comets, F. Delarue, and R. Schott, Distributed algorithms in an ergodic markovian environment, Random Structures & Algorithms, 30(1-2), pp. 131–167, 2007.

[16] F. Comets, F. Delarue, and R. Schott, Large deviations analysis for distributed algorithms in an ergodic markovian environment, Applied Mathematics and Optimization, 60(3), pp. 341–396, 2009.

[17] M. A. Crane and D. L. Iglehart, Simulating stable stochastic systems, i: General multiserver queues, Journal of the ACM (JACM), 21(1), pp. 103–113, 1974.

[18] J. G. Dai and M. Miyazawa, Reflecting brownian motion in two dimensions: Exact asymptotics for the stationary distribution, Stochastic Systems, 1(1), pp. 146–208, 2011.

[19] P.-T. De Boer, Analysis of state-independent importance-sampling measures for the two-node tandem queue, ACM Transactions on Modeling and Computer Simulation (TOMACS), 16(3), pp. 225–250, 2006.

[20] T. Dean and P. Dupuis, Splitting for rare event simulation: A large deviation approach to design and analysis, Stochastic processes and their applications, 119(2), pp. 562–587, 2009.

[21] A. Dieker and M. Mandjes, On asymptotically efficient simulation of large deviation probabilities, Advances in applied probability, 37(2), pp. 539–552, 2005.

[22] T. J. W. I. R. C. R. Division, V. Anantharam, P. Heidelberger, and P. Tsoucas, *Analysis of rare events in continuous time Markov chains via time reversal and fluid approximation*, 1990.

[23] P. Dupuis and R. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, John Wiley & Sons, New York, 1997.

[24] P. Dupuis and R. S. Ellis, The large deviation principle for a general class of queueing systems. i, Transactions of the American Mathematical Society, 347(8), pp. 2689–2751, 1995.

[25] P. Dupuis, A. D. Sezer, and H. Wang, Dynamic importance sampling for queueing networks, Annals of Applied Probability, 17(4), pp. 1306–1346, 2007.

[26] P. Dupuis and H. Wang, Importance sampling, large deviations, and differential games, Stochastics: An International Journal of Probability and Stochastic Processes, 76(6), pp. 481–508, 2004.

[27] R. Durrett, *Probability: theory and examples*, Cambridge university press, 2010.

[28] P. Flajolet, The evolution of two stacks in bounded space and random walks in a triangle, in *International Symposium on Mathematical Foundations of Computer Science*, pp. 325–340, Springer, 1986.

[29] R. D. Foley and D. R. McDonald, Constructing a harmonic function for an irreducible nonnegative matrix with convergence parameter r> 1, Bulletin of the London Mathematical Society, 44(3), pp. 533–544, 2012.

[30] R. D. Foley, D. R. McDonald, et al., Large deviations of a modified jackson network: Stability and rough asymptotics, The Annals of Applied Probability, 15(1B), pp. 519–541, 2005.

[31] M. R. Frater, T. M. Lennon, and B. D. Anderson, Optimally efficient estimation of the statistics of rare events in queueing networks, IEEE Transactions on Automatic Control, 36(12), pp. 1395–1405, 1991.

[32] H. U. Gerber, On the probability of ruin in the presence of a linear dividend barrier, Scandinavian Actuarial Journal, 1981(2), pp. 105–115, 1981.

[33] P. Glasserman and S.-G. Kou, Analysis of an importance sampling estimator for tandem queues, ACM Transactions on Modeling and Computer Simulation (TOMACS), 5(1), pp. 22–42, 1995.

[34] N. Guillotin-Plantard and R. Schott, *Dynamic random walks: Theory and applications*, Elsevier, 2006.

[35] P. Heidelberger, S. Juneja, and P. Shahabuddin, Efective bandwidth and fast simulation of ATM intree networks, Performance Evaluation, 20, pp. 45–66, 1994.

[36] I. Ignatiouk-Robert, Large deviations of Jackson networks, Annals of Applied Probability, pp. 962–1001, 2000.

[37] I. Ignatiouk-Robert, C. Loree, et al., Martin boundary of a killed random walk on a quadrant, The Annals of Probability, 38(3), pp. 1106–1142, 2010.

[38] I. Ignatyuk, V. A. Malyshev, and V. Scherbakov, Boundary effects in large deviation problems, Russian Mathematical Surveys, 49(2), pp. 41–99, 1994.

[39] S. Juneja and V. Nicola, Efficient simulation of buffer overflow probabilities in Jackson networks with feedback, ACM Transcations on Modeling and Computer Simulation, 15, pp. 281–315, 2005.

[40] S. Juneja and P. Shahabuddin, Rare-event simulation techniques: an introduction and recent advances, Handbooks in operations research and management science, 13, pp. 291–350, 2006.

[41] D. E. Knuth, *Art of Computer Programming Volume 1: Fundamental Algorithms*, Addison-Wesley Publishing Company, 1972.

[42] M. Kobayashi and M. Miyazawa, Revisiting the tail asymptotics of the double qbd process: refinement and complete solutions for the coordinate and diagonal directions, in *Matrix-analytic methods in stochastic models*, pp. 145–185, Springer, 2013.

[43] D. P. Kroese and V. F. Nicola, Efficient simulation of a tandem jackson network, ACM Transactions on Modeling and Computer Simulation (TOMACS), 12(2), pp. 119–141, 2002.

[44] I. Kurkova and V. Malyshev, Martin boundary and elliptic curves, Markov Process. Related Fields, 4(2), pp. 203–272, 1998.

[45] G. Louchard and R. Schott, Probabilistic analysis of some distributed algorithms, Random Structures & Algorithms, 2(2), pp. 151–186, 1991.

[46] G. Louchard, R. Schott, M. Tolley, and P. Zimmermann, Random walks, heat equation and distributed algorithms, Journal of Computational and Applied Mathematics, 53(2), pp. 243–274, 1994.

[47] R. S. Maier, Colliding stacks: A large deviations analysis, Random Structures & Algorithms, 2(4), pp. 379–420, 1991.

[48] R. S. Maier, Large fluctuations in stochastically perturbed nonlinear systems: Applications in computing, arXiv preprint chao-dyn/9305009, 1993.

[49] D. McDonald, Asymptotics of first passage times for random walk in an orthant, Annals of Applied Probability, pp. 110–145, 1999.

[50] D. I. Miretskiy, W. R. W. Scheinhardt, and M. R. H. Mandjes, *State-dependent importance sampling for a Jackson tandem network*, Centrum voor Wiskunde en Informatica, 2008.

[51] M. Miyazawa, Tail decay rates in double qbd processes and related reflected random walks, Mathematics of Operations Research, 34(3), pp. 547–575, 2009.

[52] M. Miyazawa, Light tail asymptotics in multidimensional reflecting processes for queueing networks, Top, 19(2), pp. 233–299, 2011.

[53] P. Ney and E. Nummelin, Markov additive processes i. eigenvalue properties and limit theorems, The Annals of Probability, pp. 561–592, 1987.

[54] V. F. Nicola and T. S. Zaburnenko, Efficient importance sampling heuristics for the simulation of population overflow in jackson networks, ACM Transactions on Modeling and Computer Simulation (TOMACS), 17(2), p. 10, 2007.

[55] S. Parekh and J. Walrand, A quick simulation method for excessive backlogs in networks of queues, IEEE Transactions on Automatic Control, 34(1), pp. 54–66, 1989.

[56] R. Randhawa and S. Juneja, Combining importance sampling and temporal difference control variates to simulate markov chains, ACM Transactions on Modeling and Computer Simulation (TOMACS), 14(1), pp. 1–30, 2004.

[57] A. Ridder, Importance sampling algorithms for first passage time probabilities in the infinite server queue, European Journal of Operational Research, 199(1), pp. 176–186, 2009.

[58] G. Rubino and B. Tuffin, *Rare event simulation using Monte Carlo methods*, John Wiley & Sons, 2009.

[59] L. Setayeshgar and H. Wang, Efficient importance sampling schemes for a feed-forward network, ACM Transactions on Modeling and Computer Simulation (TOMACS), 23(4), p. 21, 2013.

[60] A. D. Sezer, *Dynamic Importance Sampling for Queueing Networks, Ph.D. Thesis*, Brown University Division of Applied Mathematics, 2005.

[61] A. D. Sezer, Importance sampling for a markov modulated queuing network, Stochastic Processes and Their Applications, 119(2), pp. 491–517, 2009.

[62] A. D. Sezer, Asymptotically optimal importance sampling for Jackson networks with a tree topology, Queueing Systems, 64(2), pp. 103–117, 2010.

[63] A. D. Sezer, Modeling of an insurance system and its large deviations analysis, Journal of Computational and Applied Mathematics, 235(3), pp. 535–546, 2010.

[64] A. D. Sezer, Exit probabilities and balayage of constrained random walks, arXiv preprint arXiv:1506.08674, 2015.

[65] A. D. Sezer, Approximation of excessive backlog probabilities of two tandem queues, arXiv preprint arXiv:1801.04674, 2018.

[66] A. Shwartz and A. Weiss, *Large deviations for performance analysis*, Stochastic Modeling Series, Chapman & Hall, London, 1995, ISBN 0-412-06311-5, Queues, Communications, and Computing, With an appendix by Robert J. Vanderbei.

[67] Y.-K. Tse, *Nonlife actuarial models: Theory, methods and evaluation*, Cambridge University Press, 2009.

[68] A. C. Yao, An analysis of a memory allocation scheme for implementing stacks, SIAM Journal on Computing, 10(2), pp. 398–403, 1981.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Ünlü, Kamil Demirberk
E-mail:kdunlu@ankara.edu.tr
Phone: +90 312 212 67 20 / 1473
Fax: +90 312 223 32 02

EDUCATION

Degree	**Institution**	**Year of Graduation**
Ph.D.	METU, Financial Mathetmatics	2018
M.S.	METU, Financial Mathetmatics	2012
B.S.	Cankaya University, International Trade	2010
B.S.	Cankaya University, Business Management	2009
High School	TED Ankara College	2004

PROFESSIONAL EXPERIENCE

Year	**Place**	**Enrollment**
2018-present	Ankara University, Department of Statistics	Lecturer
2015-2018	Ankara University, Department of Statistics	Specialist