Biomolecular Motors and Diffusion Ratchets

by
John Fricks

A dissertation submitted to the faculty of University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research.

Chapel Hill
2004

Approved by:

Amarjit Budhiraja, Advisor

Tim Elston, Reader

Chuanshu Ji, Reader

M. Ross Leadbetter, Reader

Gordon Simons, Reader

ABSTRACT
JOHN FRICKS: Biomolecular Motors and Diffusion Ratchets.
(Under the direction of Amarjit Budhiraja.)

Biomolecular motors are proteins, or structures of multiple proteins, that play a central role in accomplishing mechanical work in the interior of a cell. While chemical reactions fuel this work, it is not exactly known how this chemical-mechanical conversion occurs. Recent advances in molecular biological techniques have enabled at least indirect observations of molecular motors which in turn have led to significant research in the mathematical modeling of these motors in the hope of shedding light on the underlying mechanisms involved in intracellular transport. Kinesin which moves along microtubules that are spread throughout the cell is a prime example of the type of motor that will be discussed in this thesis. On one end of the motor, there are twin heads that move step by step on the microtubule. The other end consists of a long amino acid chain which attaches itself to cargo that must be transported. The motion is linked to the presence of a chemical, ATP, but how the ATP is involved in motion is not clearly understood.

One commonly used model for Kinesin in the biophysics literature is the Brownian ratchet mechanism. In this thesis, a precise mathematical formulation of a Brownian ratchet (or more generally a diffusion ratchet) will be given via an infinite system of stochastic differential equations with reflection. It will be proved that this formulation arises in the weak limit of a natural discrete space pure jump Markov process that is used to describe motor dynamics in the literature. Using renewal theory it will be shown that the asymptotic velocity of the motor exists in an **almost sure** sense. We will also establish a functional central limit theorem in order to quantify fluctuations about the asymptotic velocity. This result will yield the effective diffusivity of the motor. Numerical techniques will be provided to compute asymptotic quantities such as asymptotic velocity, effective diffusivity, and the randomness parameter for this

model and other closely related models.

Linearly progressive biomolecular motors often carry cargos via an elastic linkage. A two-dimensional coupled stochastic dynamical system will be introduced to model the dynamics of the motor-cargo pair. Weak convergence results will be established in order to relate the models with the natural discrete space pure jump Markov model for the dynamics. Using ergodic theory for Markov processes, it will be shown that the asymptotic velocity exists in the sense of convergence in probability. Numerical results for computing the associated asymptotic quantites will be presented. Frequently in experiments, the motor is too small to be tracked; only the cargo which is much larger can be dynamically observed. Filtering algorithms to infer the position of the motor and to estimate model parameters based on cargo observations will be presented.

ACKNOWLEDGMENTS

I would especially like to thank Professor Budhiraja for his efforts and encouragment.

Contents

List of Figures

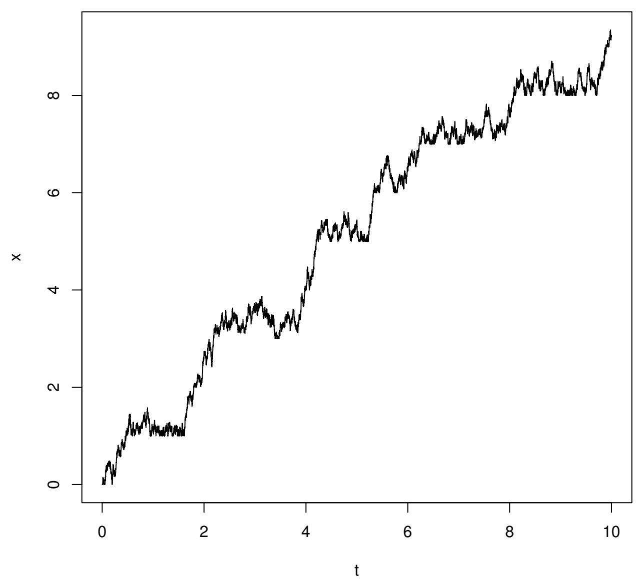
List of Tables

Chapter 1

Introduction

Biomolecular motors are proteins, or structures consisting of multiple proteins, that play a central role in accomplishing mechanical work in the interior of a cell. Frequently, the exact nature of the chemical-mechanical energy conversion is not completely understood. However, recent advances in molecular biological techniques have enabled *in vitro* observations of molecular motors and/or their cargos which in turn have led to significant research in the mathematical modeling of these motors in the hope of shedding light on the underlying mechanisms involved in intracellular transport. One commonly studied model for the dynamics of a molecular motor is the *Brownian Ratchet* model.

In a Brownian ratchet, a particle representing the biological motor diffuses between equally spaced barriers. When the particle encounters the barrier from the left it is free to pass through; however, it is instantaneously reflected back when it encounters the barrier from the right. Hence, the ratchet mechanism has the effect of introducing a positive drift to the dynamics of the particle. In practice, one is interested in gaining information about the asymptotic velocity of the motor, first passage times, locations and distances between barrier sites, parameters of the governing diffusion, etc. The goal of this thesis is to address such questions for a motor moving according to a Brownian ratchet mechanism, using the theory of reflected diffusion processes. A sample path of a Brownian ratchet process with a drift coefficient 0 and diffusion coefficient 1, with barriers which are unit distance apart, starting at 0, is given below.



In the general theory that we will develop, the drift and diffusion coefficients may be state dependent; namely, between the barriers the process will evolve according to a diffusion process. In order to distinguish from the constant drift and diffusion coefficient case, we will refer to the above mechanism as a *Diffusion Ratchet*.

Biological motors are responsible for intracellular transport of cargos, such as large protein molecules, to locations in the cell where they are needed. Unlike the ratchet process which models the dynamics of the motor, the process representing the cargo has no reflecting barriers since the cargo is floating somewhat freely in the cell, attached to the motor via a protein strand. However, there is interaction between the two processes. The farther the cargo is behind the motor the greater the forward drift for the cargo and the greater the backward drift of the motor. This thesis will investigate the pair of coupled stochastic processes, $(X(\cdot), Y(\cdot))$, which represent the cargo and the motor. We begin in Chapter 2 with relevant biological background and mathematical preliminaries.

In Chapter 3, we will provide a rigorous definition of a diffusion ratchet as a stochastic process with paths in $C([0,\infty) : I\!R_+)$ (the space of continuous functions from $[0,\infty)$ to $I\!R_+$). The process is described via a system of stochastic differential equations with reflection (See Definition 3.1.2). In order to justify the name *Diffusion Ratchet* for the constructed process, denoted hereafter by $X(\cdot)$, we begin by considering the natural discrete analog of a diffusion ratchet mechanism. Namely, we consider a particle moving on a discrete lattice in $I\!R_+$ which is free to move to the left or right when it is on a non-barrier site; however, it cannot move to the left (or does so with a "negligible" probability) when it is on a ratchet site. We will show that, after suitable scaling of time and space, the Markov chain described by the particle dynamics converges weakly in $D([0,\infty) : I\!R_+)$ (the space of functions from $[0,\infty)$ to $I\!R_+$ which are right continuous and have left limits with the usual Skorokhod topology) to $X(\cdot)$.

As was mentioned above, frequently the motor is pulling a cargo which is undergoing a diffusive motion, and its dynamics are coupled to that of the "ratcheted" motor. Since the motor is moving along a straight track ($I\!R_+$), which without loss of generality can be taken to be the x-axis, only the dynamics of the x-coordinate of the location of the cargo is coupled with the motor dynamics. Denote the location of the motor and the x-coordinate of the cargo at time instant t as $X(t)$ and $Y(t)$, respectively. In section 3.2, we will present a rigorous definition for the stochastic process, $(X(\cdot), Y(\cdot))$, with paths in $C([0,\infty) : I\!R_+ \times I\!R)$ given via an infinite system of partially reflected two dimensional diffusion processes. As was done in the case of a motor alone, in order to justify the definition, we will introduce a discrete analogue where the motor process is following the "ratchet" behavior described above, but the cargo process is free to step left or right at every lattice site. The jump rates for both the motor and cargo can be state dependent. We will show that the pure jump Markov process described by this discrete analogue after suitable scaling will converge weakly in $D([0,\infty) : I\!R_+ \times I\!R)$ to $(X(\cdot), Y(\cdot))$.

In Chapter 4, we will consider two asymptotic quantities that have been of interest to experimental researchers, asymptotic velocity and effective diffusivity. A natural definition of asymptotic velocity is $\lim_{t\to\infty} \frac{X(t)}{t}$. In the case where the motor is not pulling any cargo and the coefficients of the diffusion are periodic, we will show using renewal theory that $\frac{X(t)}{t}$ converges almost surely to a deterministic quantity which is independent of the initial location of the motor. This limiting quantity, denoted by ν, represents the asymptotic velocity of the motor. We will, in fact, establish a functional central limit theorem stating that, if $\xi_n(t) \doteq \frac{X(nt)-\nu nt}{\sqrt{nt}}$, then $\xi_n(\cdot)$ converges weakly in $C([0,\infty) : I\!R)$ to $\sigma W(\cdot)$ where W is a standard Wiener process. The constant σ^2 is referred to as the effective diffusivity of the motor. We will give explicit representations for the asymptotic velocity and the effective diffusivity in terms of the moments of first hitting times of certain reflected diffusion processes. In the special case when the drift and diffusion coefficients are constants, we will give exact values of the quantities in terms of the model parameters.

In Chapter 5, we will undertake the study of asymptotic velocity for the case where the motor is pulling a cargo. The key observation which made the analysis in Chapter 4 possible was that if σ_j denotes the first time $X(t)$ reaches the level jL, where L is the distance between barrier sites, then, due to the periodicity of the coefficients $\{\sigma_{j+1} - \sigma_j\}_{j \in I\!N_0}$ are independent, identically distributed random variables. For the two dimensional coupled system considered in Section 3.2, this property does not hold, since the distribution of $\sigma_{j+1} - \sigma_j$ depends on $X(\sigma_j) - Y(\sigma_j)$ which in general is not independent of $X(\sigma_k) - Y(\sigma_k)$ for $k \leq j$. Due to this difficulty, we will consider the special case where the interaction between cargo and motor is modeled via a linear spring. Denoting the length of the elastic linkage, $X(t)-Y(t)$, by $Z(t)$, we will show in Chapter 5 that the Markov process $\Pi(t) \doteq (\lfloor X(t)/L \rfloor L, Z(t))$ has a unique, invariant probability measure. This result will show that $\frac{1}{t}\int_0^t Z(s)ds$ converges in probability to a deterministic quantity which is independent of the initial conditions, and $\frac{Z(t)}{t}$ converges in probabiltiy to zero. This will then enable us to prove that $\frac{X(t)}{t}$ converges in probability to a deterministic quantity. This result will show that the asymptotic velocity exists and will give a representation of the quantity in terms of the invariant distribution of Π_t.

Using results from Chapter 4, one can compute asymptotic velocity and effective diffusivity explicitly for the case where the motor is not pulling any cargo, and the coefficients in the model are constant by using exact formulas for the Laplace transform of the distribution of certain hitting times. However, in general one needs to develop numerical methods for the computation of these asymptotic quantities. In Chapter 6, we will consider several numerical approaches to this problem. We will begin by addressing the one-dimensional cargoless case and present three numerical schemes for the computation of the moments of hitting times functionals. The first is a straightforward path simulation based method. Namely, we simulate a large number of trajectories of the diffusion process and determine the corresponding hitting times. The moments of hitting times then can be approximated by the corresponding empirical moments. The second technique is to devise a suitable Markov chain approximation for the reflected diffusion process and then use the transition matrix of the chain to approximate the probability distribution and, hence the moments, of the hitting time. Finally, the third method is a linear programming method introduced in [10] which is based on the martingale characterization of a Markov process via its generator. In section 6.2.1 of Chapter 6 we will consider the computation of the asymptotic parameters in the case when the motor is pulling a cargo. We will study a numerical scheme that has been recently introduced for the one dimensional cargoless setting in [23]. We will extend this method to the more complex setting of the coupled motor/cargo system and present numerical results for the computation of asymptotic velocity and effective diffusivity of the motor. The work in this section is joint with Prof. Tim Elston.

One major difficulty in the study of intracellular transport mechanisms is that the biological motors are frequently too small to be observed directly. However, these motors are often pulling cargos which are significantly larger than the motor and technical advances have allowed observations of cargos pulled by the unobserved biomolecular motors. Thus, it becomes of great interest to investigate whether the cargo observations can enable us to gain information on the motor dynamics. In practice, one will be interested in inferring about the current location of the motor based on current and past observations and estimating model parameters using observational data. In Chapter 7, using techniques from non-linear filtering theory we will address such state tracking and parameter estimation problems. In sections 7.1.1 and 7.1.2, we will provide two algorithms for approximating the conditional mean of the motor location from the cargo observations data. The first method will make use of explicit formulas for transition probability functions and numerical integration techniques. The second method is base on particle filters which are very flexible and easy to adapt for a broad family of models. In section 7.1.3, we will compare the two methods using simulated data. Clearly, one can obtain better estimates for state $X(t)$ if one can use the entire observation trajectory $\{Y(s) : 0 \leq s \leq T\}$ rather than just the current and past observations. In section 7.2, we will present a particle system algorithm for computing the conditional mean of the state given the whole observation trajectory. Finally, in section 7.3, we will consider the problem of parameter estimation. We will consider two methodologies, the first one is a simple Bayesian technique based on the augmentation of the state vector to include the parameter. The second method is a form of maximum likelihood estimation that has been introduced by Hurzeler and Kunsch [11]. We will conclude this section by comparing the two methods for simulated data.

Chapter 2

Biological and Mathematical Backround

2.1 Biomolecular Motors

The interior of a cell is a busy place. "Instructions" from DNA need to be extracted and distributed to the organelles that build proteins. Proteins then need to be transported to where they are needed. Cells must convert chemical energy to mechanical energy to accomplish these and other tasks. Biomolecular motors are one part of the picture. These biomolecular motors are proteins that accomplish a great deal of the mechanical work of the cell. Frequently, the exact mechanism of the chemical-mechanical energy conversion is not completely understood. It is hoped that mathematical models may help resolve competing explanations of these mechanisms.

An important aspect of recent research has been the observability of the motors as they function in the cell. This is being done through new methods such as laser traps which allow manipulation and observation of the motors (or objects attached to the motors). Laser traps also allow a constant (or near constant) load to be applied to the motor. Improved light microscopy has also played an important role in being able to observe cell activity. These advances in the observation capabilities have provided an impetus to research in the mathematical modeling for the dynamics of molecular motors.

Kinesin is perhaps the prime example of the type of motor that will be studied in this thesis. Kinesin moves along microtubules that are spread throughout the cell. On one end of the motor, there are twin heads that move step by step on the microtubule. The other end consists of a long amino acid chain which attaches itself to cargo that must be transported. The motion is linked to the presence of a chemical–ATP, but it is not clearly understood how the ATP is involved in motion. Moreover, even the size of each step is difficult to determine, and the way in which the motor interacts with the microtubule is not completely known. Since there is a stepping motion of the heads, point processes have been used to model the behavior of Kinesin [18]. An alternative model views one physical step of the motor as the end result of several chemical steps and transitions between these steps are modeled via a Markov jump process [22]. These Markov jump models typically ignore the physical dynamics that must occur in reality between the natural periodic step on the microtubules.. In the present work, we will primarily focus on yet another commonly used model to describe the behavior of Kinesin, namely the Brownian ratchet model [7]. In this model, the physical steps correspond to the location of the barriers of the ratchet, and the physical dynamics between the ratchet sites are modeled via reflected diffusion processes.

There are many other types of motors; some of which can be modeled via a Brownian ratchet. Many of these are linear motors similar to Kinesin moving along some “track” . An important example is RNA polymerase which moves along a DNA molecule during transcription facilitating the transfer of “information” from the DNA to other parts of the cell [24].

Another example is Myosin, a larger cousin of Kinesin which is found in muscle cells. Myosin slides along actin filaments to perform microscopic tasks similar to Kinesin. By changing the shape of individual cells, Myosin is the muscle’s source of contraction and expansion when taken in aggregate [3] .

A polymerization ratchet illustrates one of the appeals of a Brownian ratchet model. A cargo, such as a protein, sits on the tip of a polymer chain. Since the protein is being bombarded by smaller particles, such as water molecules, it diffuses, moving slightly away and back to the tip of the chain. When the protein moves far enough from the tip of the chain for another monomer to attach itself, then the protein has been permanently displaced the length of one monomer. The benefit to the cell is obvious. The cell has captured energy that is inherent to its environment while applying a rather small amount of energy itself (only the energy required to bond the monomer). So, the Brownian Ratchet conforms to a guiding principle of Biology which is that evolution has produced highly efficient mechanisms for using energy [19] .

The rotary motor which drives the flagella of the *E.coli* bacteria is still another prominent example [7]. This motor uses an ion stream (produced by the difference in electrical charge from one side of a membrane to another) to produce a rotary action.

The Brownian ratchet model for the motor with an elastic linkage to the cargo is one possible explanation for the relatively tiny motors's ability to spin the relatively large flagella.

2.2 Mathematical Preliminaries

The central mathematical element in this project is the theory of reflected diffusion processes. The stochastic process that models a ratchet motion behaves as a reflected diffusion after it has passed one barrier but has yet to reach the next. For this reason, much of the analysis of the behavior of the ratchet process can be reduced to that of reflected diffusions. The study of reflected processes traces its origin to a paper by Skorokhod in 1961 [20]. The question asked by Skorokhod in this paper is the following: Is there a $C([0,\infty) : \mathbb{R}_+)$ valued stochastic process which evolves according to a stochastic differential equation (SDE)

$$dX(t) = b(X(t),t)dt + a(X(t),t)dW_t, \tag{2.1}$$

when $X(t)$ is in $(0,\infty)$ and is instantaneously reflected back when it is about to exit $[0,\infty)$? Here a and b are functions from $\mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ satisfying the usual Lipschitz and growth conditions. This work led to the formulation of what is now called the Skorokhod Problem which is defined as follows.

For a Polish space S, denote by $D([0,\infty) : S)$ the space of functions from $[0,\infty)$ to S which are right continuous and have left limits. This space will be endowed with the usual Skorokhod topology. If $S = I\!R$, we will sometime write the above space as $D[0, \infty)$.

Definition 2.2.1 (The Skorodhod problem in 1-D). *Let $x(\cdot) \in D([0,\infty) : I\!R)$. We say that a pair of trajectories $z(\cdot), l(\cdot) \in D([0,\infty) : I\!R_+)$ solve the Skorokhod Problem for $x(\cdot)$ if*

1. $z(t) = x(t) + l(t)$ for all $t \in [0,\infty)$.

2. $l(0) = 0$.

3. $l(t)$ is increasing, and $l(t)$ increases only when $z(t) = 0$, i.e

$$l(t) = \int_{[0,t]} 1_{\{z(s)=0\}} dl(s).$$

It is easy to see that one solution of the Skorokhod problem posed by $x(\cdot)$ is given by $l(t) \doteq -0 \wedge \inf_{0 \le s \le t} x(s) \ = 0 \vee \sup_{0 \le s \le t} x(s)^-$ and $z(t) \doteq x(t) + l(t)$. In fact, one can show that the above pair is the unique solution of the Skorokhod problem posed by $x(\cdot)$ [8]. In view of this uniqueness property one can define the Skorokhod map:

$$\Gamma : D([0,\infty) : I\!R) \to D([0,\infty) : I\!R_+)$$

as $\Gamma(x(\cdot)) = z(\cdot)$, where $(z(\cdot), z(\cdot) - x(\cdot))$ is the unique solution of Skorokhod problem posed by $x(\cdot)$. We refer to $z(\cdot)$ as the constrained or reflected version of $x(\cdot)$. Using this map, one can give a positive answer to the question posed by Skorokhod, and the reflected process is the unique solution of the stochastic integral equation

$$X(t) = \Gamma(X(0) + \int_0^{\cdot} b(X(s), s) ds + \int_0^{\cdot} a(X(s), s) dW_s)(t). \tag{2.2}$$

The extension of Skorokhod problem to higher dimensions has been studied by various authors (cf. [4], [1], [9], [21], and [25]). In higher dimensions, the shape of the region to which the process is constrained could be quite complicated. Also, the direction of the reflection back into the region is something that needs to be specified. It is often taken to be the inward normal to the boundary, but it may be defined differently. These differing directions of reflection emerge naturally from applications. There are a number of general results in this area that allow one to say which types of regions and directions of reflections lead to a unique solution to the Skorokhod problem. See [4], [16], [9].

In the next section we will rigorously define a diffusion ratchet using the theory of reflected stochastic differential equations(RSDE) (See Definition 3.1.2). In doing so, the following existence and uniqueness result will be crucially used. For the rest of this work $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ will denote a filtered probability space satisfying the usual hypothesis.

Theorem 2.2.2 (Existence and Uniqueness of the Skorokhod problem in 1-D). *Let $b(\cdot), a(\cdot)$ be Lipschitz continuous. Let $W(\cdot)$ be a standard Wiener martingale on some filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})$. Then for every $x \in I\!R_+$, there exists a unique, continuous, $\mathcal{F}_t$ adapted process $X(\cdot)$ such that*

$$\int_0^T (|b(X(s))| + a^2(X(s)))ds < \infty, \ a.e. \ \forall\, T > 0$$

and the following integral equation holds.

$$X(t) = \Gamma\left(x + \int_0^{\cdot} b(X(s))ds + \int_0^{\cdot} a(X(s))dW(s)\right)(t). \tag{2.3}$$

Moreover, the following representation for $X(\cdot)$ holds.

$$X(t) = x + \int_0^t b(X(s))ds + \int_0^t a(X(s))dW(s) + l(t) \tag{2.4}$$

where

$$l(t) = -\left(0 \wedge \inf_{0 \le s \le t}\left(x + \int_0^s b(X(u))du + \int_0^s a(X(u))dW(u)\right)\right). \tag{2.5}$$

For the proof of the above result we refer the reader to [4] where a much more general case is treated. The above RSDE will play a central role in our definition of the diffusion ratchet. Roughly speaking, a diffusion ratchet will be obtained by pasting together a sequence $\{X^{(i)}(\cdot)\}$ of reflected diffusion processes, where the i-th process has the reflection barrier at iL instead of 0 and its initial condition is $X^{(i)}(0) = iL$, where L denotes the distance between ratchet sites. As outlined in Chapter 1, we will show that this process arises as the weak limit, of a natural discrete time analog of the ratchet mechanism, with a suitable scaling of time and space. One of the key ingredients to the proof of this weak convergence result is the following Aldous-Kurtz criterion for tightness of processes with paths in $D[0,\infty)$. For a proof we refer the reader to Billingsley [2].

Theorem 2.2.3 (Aldous -Kurtz Tightness Criterion). *Let Λ be an index set. Consider a collection of processes $\{x^\gamma, \gamma \in \Lambda\}$ defined on suitable probability spaces $(\Omega^\gamma, \mathcal{F}^\gamma, P^\gamma)$ and taking values in $D[0,\infty)$. Assume that for each rational $t \in [0,\infty)$ and $\delta > 0$ there exists a compact set $K_{t,\delta} \subset I\!R$ such that $\sup_{\gamma\in\Gamma} P^\gamma\{x^\gamma(t) \in K^c_{t,\delta}\} \le \delta$. Define $\mathcal{F}^\gamma_t$ to be the σ-algebra generated by $\{x^\gamma_s, s \le t\}$. Let $\mathcal{T}^\gamma_T$ be the set of $\mathcal{F}^\gamma_t$-stopping times which are less than or equal to T with probability one, and assume for each $T \in [0,\infty)$*

$$\lim_{\delta\to 0} \sup_{\gamma\in\Gamma} \sup_{\tau\in\mathcal{T}^\gamma_T} E^\gamma(1 \wedge |x^\gamma(\tau+\delta) - x^\gamma(\tau)|) = 0. \tag{2.6}$$

Then $\{x^\gamma, \gamma \in \Gamma\}$ is tight.

Another key step in our proof of the weak convergence result is to split up the re-scaled discrete time process into a sum of two processes, one of which will converge to a process which is absolutely continuous to Lebesgue measure and another which converges to an Itô integral. In proving the second convergence, we will need the following well known characterization of the Wiener Process [13].

Theorem 2.2.4 (A Characterization of Wiener Process). *For $f \in C_0^2(I\!R)$ and continuous, $\mathcal{F}_t$-adapted process $W(t)$ define*

$$M_f(t) = f(W(t)) - f(0) - \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(W(s))ds. \tag{2.7}$$

W(t) is a Wiener process if and only if $M_f(t)$ is a $\mathcal{F}_t$ local martingale.

As stated earlier in the introduction, one of the prime objects of interest is the asymptotic velocity of the motor. Intuitively, asymptotic velocity is closely connected to the amount of time a process takes to travel a given distance. Thus, it becomes important to study the time the process takes to travel from one barrier to the next. In the following theorem, we show that for this random time all moments are finite. In order to cover the coupled motor-cargo problem, we will prove this result for a slightly more general model; namely, for the case where $X^x(\cdot)$ is given as

$$X^x(t) = \Gamma\Big(x + \int_0^\cdot b(s,\omega)ds + \int_0^\cdot a(s,\omega)dW(s)\Big)(t) \tag{2.8}$$

where $b(s,\omega)$ and $a(s,\omega)$ are progressively measurable functions such that $\sup_{s,\omega}(|b(s,\omega)| + |a(s,\omega)|) < \infty$ and $\inf_{s,\omega}|a(s,\omega)| \geq m > 0$.

Theorem 2.2.5. *Let $X^x(t)$ be the process given via (2.8). Let $L \in (0, \infty)$ and define*

$$\tau^x \doteq \inf\{t : X^x(t) \notin [0, L)\}. \tag{2.9}$$

Then there exists an $\epsilon > 0$ such that $Ee^{\tau^x u} < \infty$ for all $-\epsilon < u < \epsilon$.

In order to prove the theorem, we begin with the following lemma.

Lemma 2.2.6.

$$\sup_{x \in [0,L]} P[\tau^x > 1] < 1 \tag{2.10}$$

Proof. We will argue by contradiction. Suppose that

$$\sup_{x \in [0,L]} P[\tau^x > 1] = 1. \tag{2.11}$$

Then, there is a sequence $\{x_n\} \in [0, L]$ such that $P[\tau^{x_n} > 1] \to 1$. Since $[0, L]$ is in a compact set, there is a convergent subsequence $\{x'_n\}$ which converges to some $x \in [0, L]$. We know that $P[\tau^{x'_n} > 1] \to 1$. Now, if $P[\tau^y > 1]$ is a continuous function in y, then $P[\tau^{x'_n} > 1] \to P[\tau^x > 1]$ which implies that $P[\tau^x > 1] = 1$. This in particular says that $P(X^x(\frac{1}{2}) \in [0, L]) = 1$ which is clearly impossible in view of the uniform non-degeneracy of the diffusion coefficient. Thus we have a contradiction, which proves (2.10). Therefore it suffices to show that $P[\tau^y > 1]$ is a continuous function of y. Note that $P[\tau^y > 1] = P(Z(y) < L)$, where $Z(y) \doteq \sup_{0 \le s \le 1} X^y(s)$. Note that $Z(y_n) \to Z(y)$ in probability as $y_n \to y$. Finally, observing that the distribution of $Z(y)$ is absolutely continuous with respect to the Lebesgue measure on $[0, \infty)$, we have that $P(Z(y_n) < L) \to P(Z(y) < L)$ as $y_n \to y$. This proves that $P[\tau^y > 1]$ is a continuous function of y. ■

Proof of Theorem 2.2.5. Let $\alpha \doteq \sup_{x\in[0,L]} P[\tau^x > 1]$. From Lemma 2.2.6 we have that $\alpha \in (0,1)$. Now, fix $x \in [0,L]$ and suppress it from the notation.

$$\begin{aligned} P[\tau > n] &= E\Big(I_{[\tau>n]} I_{[\tau>n-1]}\Big) \\ &= E\Big(E[I_{[\tau>n]}|\mathcal{F}_{n-1}] I_{[\tau>n-1]}\Big) \\ &\leq \alpha P[\tau > n-1] \end{aligned}$$

By induction, $P[\tau > n] \leq \alpha^n$. Now we consider the the Laplace transform, $Ee^{\tau u}$. Clearly it is finite for $u \leq 0$. So without loss of generality consider $u \geq 0$. Then

$$\begin{aligned} \int_0^\infty e^{\tau u} dP &= \int_0^\infty P[e^{\tau u} > s] ds \\ &= \int_0^\infty P[\tau > \frac{ln(s)}{u}] ds \\ &= u \int_0^\infty P[\tau > v] e^{uv} dv \\ &\leq u \sum_{k=0}^\infty P[\tau > k] e^{u(k+1)} \\ &\leq u \sum_{k=0}^\infty \alpha^k e^{u(k+1)} \\ &= ue + ue^u \sum_{k=1}^\infty (\alpha e^u)^k \end{aligned}$$

The series in the last expression on the right will converge if $\alpha e^u < 1$, i.e. $u < -ln(\alpha)$. Since $\alpha < 1$ this last quantity is greater than zero. Thus the result follows on taking $\epsilon = -ln(\alpha)$. ∎

Chapter 3

Diffusion Ratchets

In this chapter, we begin by defining a natural discrete state space model for the biological motor. Then, we will define a diffusion ratchet via a countable sequence of reflecting diffusion processes, where the i-th process, $X^{(i)}(\cdot)$, in the sequence is a reflected diffusion with the reflecting barrier and initial condition as iL. Denote by $\tau^{(i)}$ the first time the i-th process hits $(i+1)L$. The diffusion ratchet will be a $C([0,\infty) : I\!R_+)$ valued process constructed by patching together the paths of $\{X^{(i)}(t) : 0 \leq t \leq \tau^{(i)}\}_{i\geq 0}$. Then we will show that on each of the intervals $[iL, (i+1)L)$ the discrete space model converges weakly to the i-th reflected diffusion process. We will finally show that the mapping from the countable number of processes to the single "patched up" process is continuous, and thus prove that the discrete space model converges to the diffusion ratchet.

We will then study the more complicated setting of an interdependent motor/cargo system. As in the case of a motor only, we will first describe the dynamics via a discrete space Markov pure jump process. We will again consider a diffusion approximation and obtain in the limit a two dimensional coupled stochastic dynamical system. The first component of the system will evolve as a diffusion ratchet; however, its dynamics will depend on the "cargo process" which will evolve as a diffusion that has no reflection sites.

3.1 Dynamics of the Motor: Jump Markov Processes and Diffusion Approximations

We consider a particle, representing a biological motor moving on a track positioned along the x-axis. Ratchet sites are located on the track at equally spaced intervals of length L. When the particle is at a "non-ratchet" site it can move either to the left or to the right. However, when the particle is at a ratchet site, it can move only to the right. First we need to define a lattice on which the discrete space model will be defined. Let $I\!N' = \{\frac{m}{L}, m \in I\!N_0\}$. For a fixed $n \in I\!N'$, we will consider a discrete space model where the particle takes steps of size $\frac{1}{n}$, and thus the state space of the particle position is $S_n \doteq \{\frac{j}{n}, j \in I\!N_0\}$. This discretization ensures that the ratchet sites are on the lattice.

The precise dynamics of the particle can be described as follows. Let $X_n(t)$ denote the position of the particle at time t. Given that $X_n(t) = x$, the waiting time to the next jump has an exponential distribution with rate $\lambda_n(x)$. If $x = iL$ for some $i \in I\!N_0$, then the particle moves to the site on the right (i.e. $x + \frac{1}{n}$) with probability $p_n(x)$ and remains on the original site with probability $(1 - p_n(x))$. On the other hand, if $x \in S_n \setminus \{iL; i \in I\!N_0\}$, the particle moves to the right with probability $p_n(x)$ and to the site on the left with probability $(1 - p_n(x))$. We assume that the rate of jump to the right:

$$\lambda_n(x)p_n(x) = n^2\left(\alpha(x) + \frac{b^+(x)}{n}\right) \tag{3.1}$$

where as the rate of jump to the left (or at ratchet site: rate of remaining at the original site) is

$$\lambda_n(x)(1 - p_n(x)) = n^2\left(\alpha(x) + \frac{b^-(x)}{n}\right) \tag{3.2}$$

Thus

$$\lambda_n(x) = n^2\left(2\alpha(x) + \frac{b^+(x) + b^-(x)}{n}\right) \tag{3.3}$$

and

$$p_n(x) = \frac{n\alpha(x) + b^+(x)}{2n\alpha(x) + b^+(x) + b^-(x)} \tag{3.4}$$

We next introduce the diffusion ratchet $X(t)$, which is a stochastic process with continuous sample paths given as follows. Roughly speaking, $X(\cdot)$ behaves like a reflecting diffusion when it lies in the interval $[iL, (i+1)L)$; $i \in \mathbb{N}_0$, with iL acting as the reflecting barrier. Let $\{W^{(i)}\}_{i\in\mathbb{N}_0}$ be a sequence of independent standard Brownian motions given on some probability space $(\Omega, \mathcal{F}, P)$. Denote by $\mathcal{D}_i$ the subset of $\mathcal{D}([0,\infty) : \mathbb{R})$ defined as

$$\mathcal{D}_i \doteq \{x \in \mathcal{D}([0,\infty) : \mathbb{R}) | x(0) = iL\}.$$

Also, let

$$\hat{\mathcal{D}}_i \doteq \{x \in \mathcal{D}\big([0,\infty) : [iL,\infty)\big) | x(0) = iL\}.$$

Let $\Gamma_i : \mathcal{D}_i \to \hat{\mathcal{D}}_i$ be the Skorokhod map defined as:

$$\Gamma_i(x)(t) \doteq x(t) - \inf_{0 \le s \le t} \big(x(s) - iL\big). \tag{3.5}$$

Let $X^{(i)}(\cdot)$ be the unique strong solution of the integral equation:

$$X^{(i)}(t) = \Gamma_i \left(iL + \int_0^{\cdot} b(X^{(i)}(s))ds + \int_0^{\cdot} a(X^{(i)}(s))dW^{(i)}(s) \right)(t),\ t \in (0, \infty), \tag{3.6}$$

where in addition to the Lipschitz condition on the coefficients, we assume that there exist b^*, a_* and a^* in $I\!R$ such that for all $x \in I\!R^+$

$$|b(x)| \le b^* \text{ and } 0 < a_* \le a(x) \le a^*. \tag{3.7}$$

Next, for $i \in I\!N_0$, define stopping times $\tau^{(i)}$ as

$$\tau^{(i)} \doteq \inf\{t : X^{(i)}(t) \ge (i+1)L\}. \tag{3.8}$$

Also set $\sigma^{(0)} = 0$ and define

$$\sigma^{(i)} \doteq \tau^{(i-1)} + \sigma^{(i-1)},\ \ i \ge 1. \tag{3.9}$$

The following lemma will guarantee that the diffusion ratchet we will construct has paths in the space $C([0, \infty) : I\!R_+)$.

Lemma 3.1.1. *For all $i \in I\!N_0$, $P(0 < \sigma^{(i)} < \infty) = 1$ and $\sigma^{(i)} \to \infty$ almost surely, as $i \to \infty$.*

Proof. In order to show $P(0 < \sigma^{(i)}) = 1$ it suffices to show that $P(0 < \tau^{(0)}) = 1$. Note that $P(X^{(0)}(0) = 0) = 1$, and so $P(X^{(0)}(0) = L) = 0$. The continuity of sample paths of $X^{(0)}(\cdot)$ then implies that and $P(0 < \tau^{(0)}) = 1$.

Next we need to show that $P(\sigma^{(i)} < \infty) = 1$. For this, it suffices to show that $P(\tau^{(j)} < \infty) = 1$ for all j. However, this is an immediate consequence of Theorem 2.2.5 which in fact says that $E\tau^{(j)} < \infty$. This shows that $P(0 < \sigma^{(i)} < \infty) = 1$ for all i.

For the second part of the lemma, we will first show that there exists $\delta, \epsilon \in (0, \infty)$ such that

$$\inf_{j \in I\!N_0} P(\tau^{(j)} > \delta) \; > \; \epsilon. \tag{3.10}$$

Let $\epsilon \in (0, 1)$ be arbitrary. Define

$$U^{(i)}(u) = iL + \int_0^u b(X^{(i)}(s))ds + \int_0^u a(X^{(i)}(s))dW^{(i)}(s), \;\; u \in [0, \infty). \tag{3.11}$$

Note that for $\delta > 0$

$$\begin{aligned} P(\tau^{(j)} \leq \delta) &= P(\sup_{0 \leq s \leq \delta} |X^{(j)}(s) - jL| \geq L) \\ &\leq P(\sup_{0 \leq s \leq \delta} |U^{(j)}(s) - jL| \geq \frac{L}{2}) \\ &\leq 2\frac{E(\sup_{0 \leq s \leq \delta} |U^{(j)}(s) - jL|)}{L} \\ &\leq \frac{C\delta^{1/2}}{L}, \end{aligned}$$

for a universal constant C, where the last step follows on using (3.7). Now (3.10) follows on choosing δ small enough so that $\frac{C\delta^{1/2}}{L} < (1 - \epsilon)$. Finally, using the Borel-Cantelli lemma we have that the $\tau^{(j)} \geq \delta$ for infinitely many j; therefore, the sum is infinite almost surely. ■

We are now ready to define the diffusion ratchet process.

Definition 3.1.2 (Diffusion Ratchet). *Let, for $i \in I\!N_0$, $X^{(i)}, \tau^{(i)}, \sigma^{(i)}$ be defined via (3.6), (3.8) and (3.9), respectively. Define the stochastic process $X(\cdot)$ with paths in $C([0,\infty) : I\!R_+)$ as follows.*

$$X(t) \doteq X^{(i)}(t - \sigma^{(i)}); \quad t \in [\sigma^{(i)}, \sigma^{(i+1)}), \quad i \in I\!N_0.$$

We will refer to $X(\cdot)$ as the diffusion ratchet process.

Note that the diffusion ratchet has the desired properties; namely, after the process has reached iL and before it hits $(i+1)L$, it behaves like a reflected diffusion with reflecting barrier at iL, drift coefficient $b(\cdot)$ and diffusion coefficient $a(\cdot)$.

The following is the main result of this section. Let α, b^+, b^- be as in (3.2) and (3.1). Set $b(x) = b^+(x) - b^-(x)$ and $\alpha(x) = \frac{a^2(x)}{2}$.

Theorem 3.1.3. *The sequence $X_n(\cdot)$ converges weakly to $X(\cdot)$, in $\mathcal{D}([0,\infty) : I\!R_+)$, as $n \to \infty$.*

The proof of the above theorem is rather long and, therefore, before proceeding with the proof we outline the key steps involved.

We define the family of processes $\{\tilde{X}_n^{(i)}(\cdot), \tilde{U}_n^{(i)}(\cdot)\}_{i \in I\!N_0, n \in I\!N'}$ and stopping times $\{\tau_n^{(i)}\}$

as follows. For fixed $i \in I\!N_0$ and $n \in I\!N'$,

$$\begin{aligned}
\tilde{X}_n^{(i)}(t) &\doteq \tilde{U}_n^{(i)}(t) \doteq iL \text{ for } 0 \le t < \beta_n^1 \\
\tilde{U}_n^{(i)}(t) &\doteq \tilde{X}_n^{(i)}(\beta_n^j -) + \varphi_n^j \text{ for } \beta_n^j \le t < \beta_n^{j+1} \\
\tilde{X}_n^{(i)}(t) &\doteq iL + (\tilde{U}_n^{(i)}(t) - iL)^+ \text{ for } \beta_n^j \le t < \beta_n^{j+1} \\
\tau_n^{(i)} &\doteq \inf\{t : \tilde{X}_n^{(i)}(t) = (i+1)L\}
\end{aligned}$$

where $\beta_n^j \doteq \sum_{k=0}^{j} \rho_n^k$ with $(\rho_n^0, \varphi_n^0) \doteq (0,0)$ and (ρ_n^k, φ_n^k) is successively defined by

$$\begin{aligned}
P\left[\rho_n^{k+1} > s, \varphi_n^{k+1} \right. &= \left. \frac{1}{n} \middle| \rho_n^j, \varphi_n^j, j \le k \right] = e^{-\lambda_n(\tilde{X}_n^{(i)}(\beta_n^k -))s} p_n(\tilde{X}_n^{(i)}(\beta_n^k -)) \\
P\left[\rho_n^{k+1} > s, \varphi_n^{k+1} \right. &= \left. -\frac{1}{n} \middle| \rho_n^j, \varphi_n^j, j \le k \right] = e^{-\lambda_n(\tilde{X}_n^{(i)}(\beta_n^k -))s} (1 - p_n(\tilde{X}_n^{(i)}(\beta_n^k -)))
\end{aligned}$$

Let $\sigma_n^{(i)} \doteq \sum_{j=0}^{i-1} \tau_n^{(i)}$, $i \ge 1$, and $\sigma_n^{(0)} \doteq 0$. Define

$$\hat{X}_n(t) = \tilde{X}_n^{(i)}(t - \sigma_n^{(i)}); \quad t \in [\sigma_n^{(i)}, \sigma_n^{(i+1)}); \quad i \in I\!N_0.$$

Note that, by construction, $\hat{X}_n(\cdot)$ has the same law as $X_n(\cdot)$. So, if we show that $\hat{X}_n(\cdot) \Rightarrow X(\cdot)$, then we have proven the theorem. Next, let

$$\mathcal{X}_0 \doteq \mathcal{D}([0,\infty) : I\!R_+) \times [0,\infty],$$

where $[0,\infty]$ denotes the one point compactification of $I\!R_+$. Let $\mathcal{X} \doteq \mathcal{X}_0^{\otimes\infty}$. We will endow $\mathcal{X}$ with the usual topology and consider the Borel σ-field $\mathcal{B}(\mathcal{X})$. Then for each

n,

$$Z_n \doteq \{(\tilde{X}_n^{(i)}, \tau_n^{(i)})\}_{i \in I\!N_0} \quad \text{and} \quad Z \doteq \{(X^{(i)}, \tau^{(i)})\}_{i \in I\!N_0} \tag{3.12}$$

take values in $\mathcal{X}$. Define

$$\tilde{\mathcal{X}} \doteq \{(x_i, \beta_i)_{i \in I\!N_0} \in \mathcal{X} | 0 < \beta_i < \infty \text{ and } x_i \in A \ \forall\, i \text{ and } \sum_{i=0}^{j} \beta_i \to \infty \text{ as } j \to \infty\}$$

where

$$\begin{aligned} A &= \{\phi(\cdot) \in C([0,\infty) : I\!R_+) : \forall \delta > 0, \exists \text{ some } t' \in [\tau(\phi(\cdot)), \tau(\phi(\cdot)) + \delta] \\ &\quad \text{such that } \phi(t') > L\}, \end{aligned} \tag{3.13}$$

and

$$\tau(\phi(\cdot)) \doteq \inf\{t : \phi(t) \geq L\}. \tag{3.14}$$

For each i, $X^{(i)} \in A$, since the diffusion coefficient $a(x)$ is uniformly non-degenerate. We see from this fact and from Lemma 3.1.1, that

$$P(Z \in \tilde{\mathcal{X}}) = 1. \tag{3.15}$$

The key step in the proof of Theorem 3.1.3 is to establish the weak convergence of $\{Z_n\}$ to Z. The first part of the proof will be to show that each component of $\{Z_n\}$ converges to each component of Z. So, for fixed i we will show that

$$(\tilde{X}_n^{(i)}, \tau_n^{(i)}) \Rightarrow (X^{(i)}, \tau^{(i)}). \tag{3.16}$$

As a reminder, unless necessary, we will suppress i in the notation to follow. In order to prove (3.16), we will proceed as follows: First, define

$$N(t) = \max\{m : \frac{m}{n^3} \leq t\}.$$

Let

$$\Delta_j \tilde{U}_n = \tilde{U}_n\left(\frac{j+1}{n^3}\right) - \tilde{U}_n\left(\frac{j}{n^3}\right),$$

and

$$\tilde{w}_n^{(i)}(t) = \sum_{\ell=0}^{N(t)-1} \frac{[\Delta_\ell \tilde{U}_n^{(i)} - E_\ell^n \Delta_\ell \tilde{U}_n^{(i)}]}{a(\tilde{X}_n^{(i)}(\frac{\ell}{n^3}))}.$$

We will show that $\{(\tilde{X}_n, \tilde{U}_n, \tilde{w}_n)\}$ is tight, for which we need only to show that the marginals are tight. The tightness of $\tilde{U}_n$ and $\tilde{w}_n$ will be shown by using Theorem 2.2.3 and the tightness of $\tilde{X}_n$ will be an immediate consequence of the continuity of the Skorokhod map. The next step will be the identification of the weak limits. Let (X, U, W) denote a weak subsequential limit. In order to identify U, we begin by showing that W is a brownian motion with respect to $\sigma\{X(s), U(s), W(s), s \leq t\}$. Then, we will break up the process $\tilde{U}_n$ into a sum of two terms, one of which will be a Riemann sum that will converge to the absolutely continuous portion of the process in (3.11) and another which will converge to the Itô integral. To show the convergence to the Itô integral, we will use Theorem 2.2.4 of the previous section. Using this, we will show that the limit of $\tilde{U}_n$ solves (3.11). Recalling that $\tilde{X}_n = \Gamma(\tilde{U}_n)$, once the limit of $\tilde{U}_n$ has been identified we will have verified that the weak limit X solves (3.6). Since the solution to such an equation is unique, this uniquely identifies the limit of $\tilde{X}_n$. We will complete the proof of (3.16) by proving that the τ_n converge weakly to τ which is done by showing that the hitting time is a continuous functional (on a certain set of measure one) of the corresponding diffusion process and appealing to the continuous mapping theorem.

To help in the identification of the limits, we will first establish that locally–for small time steps–the discrete space model will have approximately the same mean and variance as the continuous process. This will aid in a number of future calculations. Let E_j^n be expectation conditioned on $\mathcal{F}(\tilde{X}_n(\frac{i}{n^3}), \tilde{U}_n(\frac{i}{n^3}), i \leq j)$ and E_x^n be the expectation conditioned on $\tilde{U}_n$ having initial state x (which will be of the form iL. Recall that $\lambda(\cdot)$ is $O(n^2)$) we have that for $\tilde{U}_n$ the probability of a jump in a small time interval (of size $\frac{1}{n^3}$) is

$$P\{\Delta_j \tilde{U}_n = \frac{1}{n} \text{ or } \Delta_j \tilde{U}_n = -\frac{1}{n}\} = \frac{1}{n^3}\left(\lambda_n\big(X_n\left(\frac{j}{n^3}\right)\big)\right) + o\left(\frac{1}{n}\right) \quad \to 0, \text{ as } n \to \infty.$$

Local Consistency Conditions Let $x' = \tilde{X}_n(\frac{j}{n^3})$.

$$E_j^n \Delta_j \tilde{U}_n = b(x')\frac{1}{n^3} + o\left(\frac{1}{n^3}\right) \tag{3.17}$$

and

$$E_j^n (\Delta_j \tilde{U}_n - E_j^n \Delta_j \tilde{U}_n)^2 = a^2(x')\frac{1}{n^3} + o\left(\frac{1}{n^3}\right) \tag{3.18}$$

A consequence of these local consistency conditions is the following lemma.

Lemma 3.1.4. $\{(\tilde{X}_n, \tilde{U}_n, \tilde{w}_n)\}$ *is tight.*

Proof. We will first prove the tightness of $\{\tilde{U}_n\}$ by using Theorem 2.2.3. The proof of tightness of $\tilde{w}_n$ is very similar and therefore is omitted.

Using (3.17) and (3.18), we can establish the following inequality.

$$\begin{aligned}
E_x^n|\tilde{U}_n(t)-x|^2 &= E_x^n\Bigg|\sum_{i=0}^{N(t)-1}[E_i^n\Delta_i\tilde{U}_n+(\Delta_i\tilde{U}_n-E_i^n\Delta_i\tilde{U}_n)] \\
&\quad +[E_{N(t)}^n\Delta'\tilde{U}_n+(\Delta'\tilde{U}_n-E_{N(t)}^n\Delta'\tilde{U}_n)]\Bigg|^2 \\
&\leq 2E_x^n\Bigg(\Bigg|\sum_{i=0}^{N(t)-1}E_i^n\Delta_i\tilde{U}_n+E_{N(t)}^n\Delta'\tilde{U}_n\Bigg|^2 \\
&\quad +\Bigg|\sum_{i=0}^{N(t)-1}(\Delta_i\tilde{U}_n-E_i^n\Delta_i\tilde{U}_n)+(\Delta'\tilde{U}_n-E_{N(t)}^n\Delta'\tilde{U}_n)\Bigg|^2\Bigg)
\end{aligned}$$

where $\Delta'\tilde{U}_n=\tilde{U}_n(t)-\tilde{U}_n(\frac{N(t)}{n^3})$ which leads to

$$\begin{aligned}
E_x^n|\tilde{U}_n(t)-x|^2 &= E_x^n\Bigg|\sum_{i=0}^{N(t)-1}[E_i^n\Delta_i\tilde{U}_n+(\Delta_i\tilde{U}_n-E_i^n\Delta_i\tilde{U}_n)] \\
&\leq 2E_x^n\Bigg|\sum_{i=0}^{N(t)-1}\left(b(\tilde{X}_n(\frac{i}{n^3}))\frac{1}{n^3}+o(\frac{1}{n^3})\right) \\
&\quad +b(\tilde{X}_n(\frac{N(t)}{n^3}))(t-\frac{N(t)}{n^3})+o(\frac{1}{n^3})\Bigg|^2 \\
&\quad +2E_x^n\sum_{i=0}^{N(t)-1}\left(a^2(\tilde{X}_n(\frac{i}{n^3}))\frac{1}{n^3}+o(\frac{1}{n^3})\right) \\
&\quad +2E_x^n\left(a^2(\tilde{X}_n(\frac{N(t)}{n^3}))(t-\frac{N(t)}{n^3})+o(\frac{1}{n^3})\right) \\
&\leq 2\left|Kt+N(t)o(\frac{1}{n^3})\right|^2+2\left(K^2t+N(t)o(\frac{1}{n^3})\right),
\end{aligned}$$

where K is the bound for the maximum of $|b^*|$, $|b_*|$ and $|a^*|$.

This shows that the first condition in Theorem 2.2.3 is satisfied.

For the second condition in Theorem 2.2.3, fix $T>0$ and take an arbitrary stopping time ς s.t. $\varsigma \leq T$ $w.p.1$. Note that

$$\begin{aligned} E^n_x(1 \wedge |\tilde{U}_n(\varsigma+\delta') - \tilde{U}_n(\varsigma)|) &\leq (E^n_x|\tilde{U}_n(\varsigma+\delta') - \tilde{U}_n(\varsigma)|^2)^{1/2} \\ &\leq \left(2\left|K\delta' + N(\delta')o\left(\frac{1}{n^3}\right)\right|^2 \right. \\ &\quad \left. +2\left(K^2\delta' + N(\delta')o\left(\frac{1}{n^3}\right)\right)\right)^{\frac{1}{2}} \end{aligned}$$

Since the right side above converges to 0 as $\delta' \to 0$ and $n \to \infty$, we have that the second condition in Theorem 2.2.3 holds. So, the sequence of processes $\{\tilde{U}_n\}$ is tight. The tightness of $\{\tilde{X}_n\}$ is now an immediate consequence of the fact that $\tilde{X}_n = \Gamma(\tilde{U}_n)$ and that Γ is a continuous map. ∎

The next step in the proof is the identification of the weak limits of the sequence $\{(\tilde{X}_n, \tilde{U}_n, \tilde{w}_n)\}$.

Lemma 3.1.5. *Let $W^{(i)}$ be a Wiener process defined on some probability space $(\Omega, \mathcal{F}, P)$ and let $(X^{(i)}, U^{(i)})$ be defined via (3.6) and (3.11), respectively. Then $(\tilde{X}_n, \tilde{U}_n, \tilde{w}_n) \equiv (\tilde{X}^{(i)}_n, \tilde{U}^{(i)}_n, \tilde{w}^{(i)}_n)$ converges weakly to $(X^{(i)}, U^{(i)}, W^{(i)})$, as $n \to \infty$.*

Proof. In what follows, we will suppress i from the notation as needed. We know that there exists convergent subsequences for the measures induced by the $\{(\tilde{X}_n, \tilde{U}_n, \tilde{w}_n)\}$ processes. So we need to establish the following: For any weakly convergent subsequence, the weak limit of $(\tilde{X}^{(i)}_{n'}, \tilde{U}^{(i)}_{n'}, \tilde{w}^{(i)}_{n'})$, denoted by $(\tilde{X}^{(i)}, \tilde{U}^{(i)}, \tilde{w}^{(i)})$ has the same law as $(X^{(i)}, U^{(i)}, W^{(i)})$ We will use n for n' to simplify notation. First we will use Theorem 2.2.4 to show that $\tilde{w}^{(i)}$ a $\mathcal{F}_t$-Wiener process where

$\mathcal{F}_t \doteq \sigma\{\tilde{X}^{(i)}(s), \tilde{U}^{(i)}(s), \tilde{w}^{(i)}(s), s \le t\}$. By construction of $\tilde{w}_n(t)$ and $\tilde{U}_n(t)$

$$|\tilde{w}_n^{(i)}(t) - \tilde{w}_n^{(i)}(t-)| \le \frac{2}{n \times a_*} \tag{3.19}$$

for every t, and therefore the jumps of $\tilde{w}_n^{(i)}$ converge to zero uniformly in t. From Theorem 13.4 of [2], we now have that $\tilde{w}^{(i)}$ has continuous paths almost surely. In a similar way, it is seen that $\tilde{U}^{(i)}$ and $\tilde{X}^{(i)}$ a.s. have continuous paths as well.

Fix $t \ge 0, \zeta > 0, q < \infty, t_i \in [0,t]$ with $t_{i+1} > t_i$ for $i \in \{0, ..., q\}$, and let $f \in C_0^2(\mathbb{R})$. Then,

$$\begin{aligned} f(\tilde{w}_n(t+\zeta)) \quad - \quad & f(\tilde{w}_n(t)) - \int_t^{t+\zeta} \frac{1}{2}\frac{\partial^2}{\partial x^2} f(\tilde{w}_n(s))ds \\ = \quad & \sum_{i=N(t)+1}^{N(t+\zeta)-1} [f(\tilde{w}_n(\frac{(i+1)}{n^3})) - f(\tilde{w}_n(\frac{i}{n^3}))] \\ & - \sum_{i=N(t)+1}^{N(t+\zeta)-1} \frac{1}{2} f_{xx}(\tilde{w}_n(\frac{i}{n^3}))\frac{1}{n^3} + \epsilon_n, \end{aligned}$$

and so

$$\begin{aligned} f(\tilde{w}_n(t+\zeta)) \quad - \quad & f(\tilde{w}_n(t)) - \int_t^{t+\zeta} \frac{1}{2}\frac{\partial^2}{\partial x^2} f(\tilde{w}_n(s))ds \\ = \quad & \sum_{i=N(t)}^{N(t+\zeta)-1} f_x(\tilde{w}_n(\frac{i}{n^3}))\frac{[\Delta_i \tilde{U}_n - E_i^n \Delta_i \tilde{U}_n]}{a(\tilde{X}_n(\frac{i}{n^3}))} \\ & + \frac{1}{2} \sum_{i=N(t)}^{N(t+\zeta)-1} f_{xx}(\tilde{w}_n(\frac{i}{n^3}))\frac{[\Delta_i \tilde{U}_n - E_i^n \Delta_i \tilde{U}_n]^2}{a^2(\tilde{X}_n(\frac{i}{n^3}))} \\ & - \frac{1}{2} \sum_{i=N(t)}^{N(t+\zeta)-1} f_{xx}(\tilde{w}_n(\frac{i}{n^3}))\frac{1}{n^3} + \epsilon_n. \end{aligned}$$

Here, $E^n|\epsilon_n| \to 0$ as $n \to \infty$.

Thus if H is a bounded, continuous function from $I\!R^{3\times q}$ to $I\!R$, then

$$\begin{aligned} & E\bigg| H(\tilde{X}_n(t_i), \tilde{U}_n(t_i), \tilde{w}_n(t_i), 1 \le i \le q) \\ \times & \left[f(\tilde{w}_n(t+\zeta)) - f(\tilde{w}_n(t)) - \int_t^{t+\zeta} \frac{1}{2}\frac{\partial^2}{\partial x^2} f(\tilde{w}_n(s))ds \right] \bigg| \\ \le & K'(E^n|\epsilon_n|) \quad \to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.20}$$

Since $(\tilde{X}_n, \tilde{U}_n, \tilde{w}_n) \Rightarrow (\tilde{X}, \tilde{U}, \tilde{w})$, using the Skorokhod representation theorem, we can assume without loss of generality that $(\tilde{X}_n, \tilde{U}_n, \tilde{w}_n) \to (\tilde{X}, \tilde{U}, \tilde{w})$, almost surely. Thus taking a limit as $n \to \infty$ on the left side of (3.20) we have via an application of the dominated convergence theorem that

$$E\bigg| H(\tilde{X}(t_i), \tilde{U}(t_i), \tilde{w}(t_i), 1 \le i \le q) \times \left[f(\tilde{w}(t+\zeta)) - f(\tilde{w}(t)) - \int_t^{t+\zeta} \frac{1}{2}\frac{\partial^2}{\partial x^2} f(\tilde{w}(s))ds \right] \bigg| = 0. \tag{3.21}$$

Now, let $\mathcal{F}_t = \sigma\{\tilde{X}(s), \tilde{U}(s), \tilde{w}(s), s \le t\}$ which is generated by functions of the form of $H(\cdot)$. Thus,

$$E\left[\left(f(\tilde{w}(t+\zeta)) - f(\tilde{w}(t)) - \int_t^{t+\zeta} \frac{1}{2}\frac{\partial^2}{\partial x^2} f(\tilde{w}(s))ds \right) I_F \right] = 0 \tag{3.22}$$

for $F \in \mathcal{F}_t$, and so

$$E\left[f(\tilde{w}(t+\zeta)) - f(\tilde{w}(t)) - \int_t^{t+\zeta} \frac{1}{2}\frac{\partial^2}{\partial x^2} f(\tilde{w}(s))ds \bigg| \mathcal{F}_t \right] = 0 \quad w.p.1 \tag{3.23}$$

We conclude $f(\tilde{w}(t)) - f(\tilde{w}(0)) - \int_0^t \frac{1}{2}\frac{\partial^2}{\partial x^2} f(\tilde{w}(s))ds$ is an $\mathcal{F}_t$-martingale for all $f \in C_0^2(I\!R)$. So, by Theorem 2.2.4 $\tilde{w}(\cdot)$ is an $\mathcal{F}_t$-Wiener process.

Now, we will identify the limit process $(\tilde{X}, \tilde{U})$. For each $\delta > 0$ and $t \in [j\delta, (j+1)\delta)$; define

$$\tilde{U}_n^\delta(t) \doteq \tilde{U}_n(j\delta), \quad \tilde{U}^\delta(t) \doteq \tilde{U}(j\delta), \quad \tilde{X}_n^\delta(t) \doteq \tilde{X}_n(j\delta), \quad \tilde{X}^\delta(t) \doteq \tilde{X}(j\delta).$$

Let $N^\delta(t) = \max\{m : m\delta \le t\}$. Using the definition of $w(\cdot)$ and the local consistency properties (3.17), (3.18), we have that

$$\tilde{U}_n^\delta(t) - x = \int_0^t b(\tilde{X}_n^\delta(s))ds + \sum_{j=0}^{N^\delta(t)-1} a(\tilde{X}_n^\delta(\delta i))[w_n(\delta(i+1)) - w_n(\delta i)] + \epsilon_n^{\delta,t},$$

where $E|\epsilon_n^{\delta,t}| \to 0$ as $\delta \to 0$, uniformly in n and t in any bounded interval. Since $(\tilde{X}_n, \tilde{U}_n) \to (\tilde{X}, \tilde{U})$, we have that $(\tilde{X}_n^\delta, \tilde{U}_n^\delta) \to (\tilde{X}^\delta, \tilde{U}^\delta)$ with probability one in $D[0,\infty)$. Thus letting $n \to \infty$ in the above display, we have

$$\tilde{U}^\delta(t) - x = \int_0^t b(\tilde{X}^\delta(s))ds + \sum_{j=0}^{N^\delta(t)-1} a(\tilde{X}^\delta(\delta i))[\tilde{w}(\delta(i+1)) - \tilde{w}(\delta i)] + O(\delta) + \epsilon^{\delta,t} \quad (3.24)$$

where $E|\epsilon^{\delta,t}| \to 0$ as $\delta \to 0$. Note that for each j, $\tilde{U}(j\delta)$ and $\tilde{X}(j\delta)$ are independent of all of the random variables $\{\tilde{w}(s) - \tilde{w}(j\delta), s \ge j\delta\}$. So, we have

$$\tilde{U}^\delta(t) - x = \int_0^t b(\tilde{X}^\delta(s))ds + \int_0^t a(\tilde{X}^\delta(s))d\tilde{w}(s) + \overline{\epsilon_{\delta,t}} \quad (3.25)$$

where $E|\overline{\epsilon_{\delta,t}}| \to 0$ as $\delta \to 0$. By the continuity of $a(\cdot)$ and $\tilde{X}(\cdot)$, we have that

$$\int_0^t a(X^\delta(s))d\tilde{w}(s) \to \int_0^t a(X(s))d\tilde{w}(s), \quad (3.26)$$

in probability, as $\delta \to 0$. Similarly, $\int_0^t b(\tilde{X}^\delta(s))ds \to \int_0^t b(\tilde{X}(s))ds$. Furthermore, since

$\tilde{X}_n = \Gamma(\tilde{U}_n)$ for all n, we have that $\tilde{X} = \Gamma(\tilde{U})$. Therefore, the limit process $(\tilde{X}, \tilde{U})$ solves

$$\begin{aligned}\tilde{U}(t) &= x + \int_0^t b(\tilde{X}(s))ds + \int_0^t a(\tilde{X}(s))d\tilde{w}(s) \\ \tilde{X}(t) &= \Gamma(\tilde{U})(t),\end{aligned}$$

By strong (and therefore weak) uniqueness of the solution to the above equation, we now have that $(\tilde{X}, \tilde{U}, \tilde{w})$ has the same probability law as $(X^{(i)}, U^{(i)}, W^{(i)})$. This proves the lemma. ■

In order to complete the proof of (3.16) we now need to show the convergence of the stopping times to the corresponding limits. This convergence is an immediate consequence of the following lemma and the continuous mapping theorem.

Lemma 3.1.6. *Let $\tau(\cdot)$ be defined via (3.14). Let $\phi \in A$ and $\{\phi_n\}$ be a sequence in $D([0,\infty) : I\!R_+)$ such that $\phi_n \to \phi$. Then $\tau(\phi_n) \to \tau(\phi)$.*

Proof. Note that since ϕ is continuous, ϕ_n converges to ϕ, uniformly on compacts. First we will show

$$\liminf_{n\to\infty} \tau(\phi_n) \geq \tau(\phi). \tag{3.27}$$

We will argue via contradiction, suppose that (3.27) is false. Then, $\tau' \doteq \liminf_{n\to\infty} \tau(\phi_n) < \tau(\phi)$. So, there is a subsequence $\{\tau(\phi_{n'})\}$ that converges to τ'. For any $\delta > 0$ and sufficiently large N, $\{\tau(\phi_{n'})\} \in [\tau' - \delta, \tau' + \delta]$ for all $n' \geq N$. Let δ be small enough so that this interval does not include $\tau(\phi(\cdot))$. There exists an $\epsilon > 0$ such that for all t in this interval, $\phi(t) \leq L - \epsilon$ since $t < \tau(\phi(\cdot))$. For some t in this interval $\phi_{n'}(t) = L$. So,

$\sup_{t\in[\tau'-\delta,\tau'+\delta]} |\phi_{n'}(t) - \phi(t)| \geq \epsilon$ for all $n' > N$. This is a contradiction to the fact that $\phi_n(\cdot)$ converges to $\phi(\cdot)$, uniformly on compacts.

Next we show that

$$\tau'' \doteq \limsup_{n\to\infty} \tau(\phi_n(\cdot)) \leq \tau(\phi(\cdot)). \tag{3.28}$$

Once again, we proceed via contradiction. Assume that (3.28) is false, i.e. $\tau'' > \tau(\phi)$ Again, there is a subsequence $\{\tau(\phi_{n''})\}$ that converges to τ''. Now, if we look at the interval $[\tau(\phi(\cdot)), \tau(\phi(\cdot)) + \frac{\tau(\phi(\cdot))-\tau''}{2}]$, then for sufficiently large N $\tau(\phi_{n''}(\cdot))$ are to the right of this interval for all $n'' > N$. Thus, the corresponding $\phi_{n''}(t)$ are less than L for t in the interval. By our assumption, there is a t' in this interval such that $\phi(t') > L$. This contradicts the fact that $\phi_n(\cdot) \to \phi(\cdot)$. This proves (3.28). Combining (3.27) and (3.28) we have the result. ■

An immediate consequence of the above results is the following corollary.

Corollary 3.1.7. *Let $\{Z_n\}$ and Z be defined via (3.12). The sequence $\{Z_n\}$ of $\mathcal{X}$ valued random elements converges weakly to Z.*

Proof. From Lemma 3.1.5 and Lemma 3.1.6 we have that

$$(X_n^{(i)}, \tau_n^{(i)}) \equiv (X_n^{(i)}, \tau(X_n^{(i)})) \Rightarrow (X^{(i)}, \tau(X^{(i)})) \equiv (X^{(i)}, \tau^{(i)}).$$

So, every component of Z_n converges to the respective component of Z. This proves the result. ■

The final step in the proof of Theorem 3.1.3 is the following lemma. We begin with the following notation. For $z = \{x^{(i)}, \tau^{(i)}\}_{i\geq 0} \in \mathcal{X}$, define $J_z : I\!N_0 \to I\!N_0$ as follows. $J_z(0) \doteq 0$ and for $k \geq 1$, $J(k) \doteq \inf\{i > J(k-1) : \tau^{(i)} \neq 0\}$. Also define $x_z \in \mathcal{D}([0,\infty) : I\!R_+)$ as:

$$x_z(t) \doteq \begin{cases} x^{(J(k))}(t - \sigma^{(J(k-1))}) & \text{if } t \in [\sigma^{(J(k-1))}, \sigma^{(J(k))});\quad k \in I\!N \\ 0 & \text{if } t \geq \sigma^{\infty}, \end{cases} \tag{3.29}$$

where for $i \geq 1$ $\sigma^{(i)} = \sum_{j=1}^{i} \tau^{(j)}$, $\sigma^{(0)} \doteq 0$ and $\sigma^{\infty} \doteq \lim_{j\to\infty} \sigma^{(j)}$.

Lemma 3.1.8. *Let $\Psi : \mathcal{X} \to \mathcal{D}([0,\infty) : I\!R_+)$ be defined as $\Psi(z) = x_z$, $z = \{x^{(i)}, \tau^{(i)}\}_{i\geq 0}$, where x_z is given via (3.29). Then Ψ is continuous at every $z \in \tilde{\mathcal{X}}$.*

Proof: Take a sequence $\{z_n\} \in \mathcal{X}$ which converges to $z \in \tilde{\mathcal{X}}$. We will show that $\Psi(z_n) \to \Psi(z)$. It suffices to show that $\Psi(z_n)(t) \to \Psi(z)(t)$ uniformly on $[\sigma^{(i-1)}, \sigma^{(i)}]$ for all $i \in I\!N$. In order to prove this statement we will show the following two results.

(i)

$$\Psi(z_n)(t) \to \Psi(z)(t), \text{ uniformly for } t \in [\sigma^{(i-1)} + \delta, \sigma^{(i)} - \delta] \tag{3.30}$$

for any δ s.t. $0 < \delta < \frac{\sigma^i - \sigma^{i-1}}{2}$ and $i \in I\!N$.

(ii)

$$\Psi(z_n)(t) \to \Psi(z)(t), \text{ uniformly for } t \in [\sigma^{(i-1)} - \delta, \sigma^{(i-1)} + \delta] \tag{3.31}$$

for any δ s.t. $0 < \delta < (\sigma^{i-1} - \sigma^{i-2}) \wedge (\sigma^i - \sigma^{i-1})$ and $i \geq 2$.

Clearly, the result follows once (i) and (ii) are proven.

Now we prove (i). Note that since $\sigma_n^j \to \sigma^j$ for all j as $n \to \infty$, we can find $N \in I\!N$ s.t. for all $n \geq N$, $\sigma_n^{(j)} \in [\sigma^{(j)} - \delta, \sigma^{(j)} + \delta]$ for $j = i, i-1$. So, for t in the fixed interval

$[\sigma^{(i-1)} + \delta, \sigma^{(i)} - \delta]$,

$$\begin{aligned} \Psi(z_n)(t) &= x_n^{(i)}(t - \sigma_n^{(i-1)}) \\ \Psi(z)(t) &= x^{(i)}(t - \sigma^{(i-1)}). \end{aligned}$$

For simplicity, we will re-parameterize with $s = t - \sigma^{(i-1)}$ and set $\varsigma_n = \sigma^{(i-1)} - \sigma_n^{(i-1)}$. Note that $|\varsigma_n| \leq \delta$.) Thus, for s in $C = [\delta, \tau^{(i)} - \delta]$,

$$\begin{aligned} \Psi(z_n)(s + \sigma^{(i-1)}) &= x_n^{(i)}(s + \varsigma_n) \\ \Psi(z)(s + \sigma^{(i-1)}) &= x^{(i)}(s). \end{aligned}$$

So, we need to show that $x_n(s + \varsigma_n) \to x(s)$ uniformly on C (suppressing the i). We know that $x(s)$ is uniformly continuous on $[0, \tau]$, and that $x_n(s) \to x(s)$ uniformly on compact intervals since $x(s)$ is continuous and $x_n(s)$ converges to $x(s)$ in the Skorokhod space. Now,

$$|x(s) - x_n(s + \varsigma_n)| \leq |x(s) - x(s + \varsigma_n))| + |x(s + \varsigma_n) - x_n(s + \varsigma_n)| \tag{3.32}$$

for every $s \in C$. Thus,

$$\begin{aligned} \sup_{s \in C} |x(s) - x_n(s + \varsigma_n)| &\leq \sup_{s \in C} |x(s) - x(s + \varsigma_n))| + \sup_{s \in C} |x(s + \varsigma_n) - x_n(s + \varsigma_n)| \\ &\leq \sup_{s \in C} |x(s) - x(s + \varsigma_n))| + \sup_{s \in [0, \tau]} |x(s) - x_n(s)|. \end{aligned}$$

Now, the right hand side goes to zero, since x is uniformly continuous, $\varsigma_n \to 0$ and $x_n \to x$ uniformly. This proves (i).

Now we consider (ii). Let us denote the interval $[\sigma^{(i-1)} - \delta, \sigma^{(i-1)} + \delta]$ by K. Let $s = t - \sigma^{(i-1)}$ and $\varsigma_n = \sigma_n^{(i-1)} - \sigma^{(i-1)}$, and again ensure large enough n to guarantee $\sigma_n^{(i)}$ are within δ of the limit. Thus,

$$\begin{aligned}
\sup_{s\in K} |\Psi(z_n)(s) - \Psi(z)(s)| &\le \sup_{s\in K} |\Psi(z_n)(s) - \Psi(z)(t - \varsigma_n)| \\
&\quad + \sup_{s\in K} |\Psi(z)(s - \varsigma_n) - \Psi(z)(s)| \\
&\le \sup_{r\in[\sigma^{(i-1)}-2\delta, \sigma^{(i-1)}]} |x_n^{(i-1)}(r) - x(r)| \vee \sup_{r\in[0,2\delta]} |x_n^{(i)}(r) - x(r)| \\
&\quad + \sup_{s\in K} |\Psi(z)(s - \varsigma_n) - \Psi(z)(s)|.
\end{aligned}$$

Once more the uniform convergence of $x_n^{(i)}$ to $x^{(i)}$ shows that the right side of the expression above converges to 0 as $n \to \infty$. This proves (ii) and hence the result. ■

Proof of Theorem 3.1.3 .

The proof is an immediate consequence of Corollary 3.1.7 and Lemma 3.1.8 on observing that $X_n = \Psi(Z_n)$ and $X = \Psi(Z)$. ■

3.2 The Coupled System: Motor and Cargo

We will begin by introducing the natural discrete state Markov pure jump process $(X_n(t), Y_n(t))$ describing the dynamics of the Motor-Cargo pair. The state space for this Markov process will be $\overline{S_n} \doteq S_n \times \tilde{S}_n$, where for $n \in I\!N'$, $S_n \doteq \{\frac{j}{n} : j \in I\!N_0\}$ and $\tilde{S}_n \doteq \{\frac{j}{n} : j \in \boldsymbol{Z}\}$, where $I\!N'$ is as in the last section. Once more, ratchet sites are located on the track at equally spaced intervals of length L; when the motor is at a "non-ratchet" site it can either move to the left or to the right in steps of size $\frac{1}{n}$.

However, when the motor is at a ratchet site, it can only move to the right. The cargo, on the other hand, is free to move to the left or to the right, at every site.

The dynamics of the Markov process will be described via jump rates

$$\lambda_n^X(x,y) = n^2\left(2\alpha_1(x,y) + \frac{b_1^+(x,y)+b_1^-(x,y)}{n}\right) \tag{3.33}$$

$$\lambda_n^Y(x,y) = n^2\left(2\alpha_2(x,y) + \frac{b_2^+(x,y)+b_2^-(x,y)}{n}\right) \tag{3.34}$$

for each component of the process and via probabilities of jumps to the right:

$$p_n^X(x,y) = \frac{n\alpha_1(x,y)+b_1^+(x,y)}{2n\alpha_1(x,y)+b_1^+(x,y)+b_1^-(x,y)} \tag{3.35}$$

$$p_n^Y(x,y) = \frac{n\alpha_2(x,y)+b_2^+(x,y)}{2n\alpha_2(x,y)+b_2^+(x,y)+b_2^-(x,y)} \tag{3.36}$$

In the above equations α_i and β_i for $i = 1,2$ are bounded functions from $I\!R_+ \times I\!R$ to $I\!R$. Also, it is assumed that $\inf_{x,y}\alpha_i(x,y) > 0$ for $i = 1,2$. We define $\{(X_n, Y_n)\}$ to be a sequence of processes that have laws described as follows. Let $(X_n(t), Y_n(t))$ describe the position of the system at time t. Given that the process is at (x,y) at time t, the waiting time to the next transition is exponentially distributed with rate $\lambda_n^Y(x,y) + \lambda_n^X(x,y)$. At that time instant, there is a transition in the y dimension with probability $\frac{\lambda_n^Y(x,y)}{\lambda_n^X(x,y)+\lambda_n^Y(x,y)}$; otherwise, there is a transition in the x dimension. If the transition is in the y dimension, the y component increases by $\frac{1}{n}$ with probability $p_n^Y(x,y)$ and decreases with probability $1 - p_n^Y(x,y)$. Similarily, if the transition is in the x component, the x moves to the right with probability $p_n^X(x,y)$ and to the left with probability $1 - p_n^X(x,y)$ unless previously the x component was at a ratchet site.

If the x component was at a ratchet site, then the x increases with probability $p_n^X(x,y)$ and remains at the ratchet site with probability $1 - p_n^X(x,y)$.

An explicit way to construct (X_n, Y_n) is as follows: Let $\{(\tilde{\mathcal{Y}}_1^n(k), \tilde{\mathcal{Y}}_2^n(k)) \equiv \tilde{\mathcal{Y}}^n(k), k = 0, 1, ...\}$ be a $\overline{S_n}$-valued Markov chain with initial distribution $\delta_{(0,y_0)}$ and transition function

$$
\begin{aligned}
\mu_n\Big((x,y),\Lambda\Big) &\doteq \frac{\lambda_n^Y(x,y)p_n^Y(x,y)}{\lambda_n^X(x,y)+\lambda_n^Y(x,y)}\delta_{(x,y+1/n)}(\Lambda) \\
&+ \frac{\lambda_n^Y(x,y)(1-p_n^Y(x,y))}{\lambda_n^X(x,y)+\lambda_n^Y(x,y)}\delta_{(x,y-1/n)}(\Lambda) \\
&+ \frac{\lambda_n^X(x,y)p_n^X(x,y)}{\lambda_n^X(x,y)+\lambda_n^Y(x,y)}\delta_{(x+1/n,y)}(\Lambda) \\
&+ \frac{\lambda_n^X(x,y)(1-p_n^X(x,y))}{\lambda_n^X(x,y)+\lambda_n^Y(x,y)}\delta_{(\psi_n(x),y)}(\Lambda), \Lambda \in \overline{S_n},
\end{aligned}
$$

where $\psi_n(x)$ is x if $\frac{x}{L}$ is in $I\!N_0$ and is $x - 1/n$ otherwise.

Let $\{\Delta_i\}_{i\geq 0}$ be $i.i.d.$ $Exp(1)$ which are also independent of $\{\tilde{\mathcal{Y}}^n(k)\}_{k\in I\!N_0}$. Then

$$
\begin{aligned}
(X_n(t),Y_n(t)) &\doteq \tilde{\mathcal{Y}}^n(0) \text{ for } 0 \leq t \leq \frac{\Delta_0}{\lambda_n^X(\tilde{\mathcal{Y}}^n(0))+\lambda_n^Y(\tilde{\mathcal{Y}}^n(0))} \\
&\quad \tilde{\mathcal{Y}}^n(k) \text{ for } \sum_{j=0}^{k-1}\frac{\Delta_j}{\lambda_n^X(\tilde{\mathcal{Y}}^n(j))+\lambda_n^Y(\tilde{\mathcal{Y}}^n(j))} \leq t \leq \sum_{j=0}^{k}\frac{\Delta_j}{\lambda_n^X(\tilde{\mathcal{Y}}^n(j))+\lambda_n^Y(\tilde{\mathcal{Y}}^n(j))}
\end{aligned}
\tag{3.37}
$$

A somewhat more convenient representation (in distribution) for (X_n, Y_n) is given as follows. Let

$$
\lambda \doteq \sup_n \sup_{(x,y)\in I\!R_+\times I\!R} \frac{\lambda_n^X(x,y)+\lambda_n^Y(x,y)}{n^2}.
$$

Define a transition function $\mu'_n : (I\!\!R_+ \times I\!\!R, \mathcal{B}(I\!\!R_+ \times I\!\!R)) \to [0,1]$ as

$$\begin{aligned}
\mu'_n\big((x,y),\Lambda\big) &= \left(1 - \frac{\lambda_n^X(x,y) + \lambda_n^Y(x,y)}{n^2\lambda}\right)\delta_{(x,y)}(\Lambda) \\
&+ \frac{\lambda_n^X(x,y)}{n^2\lambda}\Big[p_n^X(x,y)\delta_{(x+1/n,y)}(\Lambda) \\
&+ (1 - p_n^X(x,y))\delta_{(\psi_n(x),y)}(\Lambda)\Big] \\
&+ \frac{\lambda_n^Y(x,y)}{n^2\lambda}\Big[p_n^Y(x,y)\delta_{(x,y+1/n)}(\Lambda) \\
&+ (1 - p_n^Y(x,y))\delta_{(x,y-1/n)}(\Lambda)\Big]
\end{aligned}$$

Now, let $\mathcal{Y}^n(k) \equiv (\mathcal{Y}_1^n(k), \mathcal{Y}_2^n(k))$ be a Markov chain with transition function μ'. Let $V(t)$ be an independent Poisson process with rate 1. Then

$$(X_n(t), Y_n(t))_{t\le 0} \stackrel{\mathcal{L}}{=} (\mathcal{Y}_1^n(V(\lambda n^2 t)), \mathcal{Y}_2^n(V(\lambda n^2 t)))_{t\le 0}$$

The following result will be key in the proof of our main result in this section

Proposition 3.2.1 (Proposition 4.5 of [14]). (X_n, Y_n) *converges weakly in* $\mathcal{D}([0,\infty) : I\!\!R_+ \times I\!\!R)$ *iff* $(X_n^*(\cdot), Y_n^*(\cdot)) \doteq (\mathcal{Y}_1^n(\lfloor \lambda n^2 \cdot \rfloor), \mathcal{Y}_2^n(\lfloor \lambda n^2 \cdot \rfloor))$ *converges weakly in* $\mathcal{D}([0,\infty) : I\!\!R_+ \times I\!\!R)$ *in which case the limits are the same.*

Our next step will be to study the diffusion limit of the above Markov chain, as $n \to \infty$. In the limit, one would expect to obtain a diffusion ratchet, representing the dynamics of the biological motor, which is coupled with an unconstrained diffusion process representing the cargo. More precisely, we would like to prove that as $n \to \infty$, (X_n, Y_n) converges weakly, in $D([0,\infty) : I\!\!R_+ \times I\!\!R)$ to the process (X, Y), with paths in $C([0,\infty) : I\!\!R_+ \times I\!\!R)$, given as the solution of the following equations.

Definition 3.2.2 (Motor and Cargo).

$$\begin{cases} \textit{For } t \in & [0,\infty) \\ Y^{(i)}(t) & = Y^{(i)}(0) + \int_0^t b_2(X^{(i)}(s),Y^{(i)}(s))ds + \int_0^t a_2(X^{(i)}(s),Y^{(i)}(s))dB^{(i)}(s), \\ & \\ X^{(i)}(t) & = \Gamma_i\left(iL + \int_0^\cdot b_1(X^{(i)}(s),Y^{(i)}(s))ds + \int_0^\cdot a_1(X^{(i)}(s),Y^{(i)}(s))dW^{(i)}(s)\right)(t), \\ & \\ \tau^{(i)} & \doteq \inf\{t : X^{(i)}(t) = (i+1)L\}, \quad \sigma^{(i)} \doteq \tau^{(i-1)} + \sigma^{(i-1)} \\ & \\ \textit{For } t \in & [\sigma^{(i)}, \sigma^{(i+1)}), \quad i \in I\!N_0 \\ X(t) & \doteq X^{(i)}(t-\sigma^{(i)}) \\ Y(t) & \doteq Y^{(i)}(t-\sigma^{(i)}) \\ Y^{(i)}(0) & \doteq Y^{(i-1)}(\tau^{(i)}) \textit{ for } i \geq 1 \\ Y^{(0)}(0) & \doteq y_0 \end{cases} \tag{3.38}$$

where $W^{(i)}(\cdot)$ and $B^{(i)}(\cdot)$ are sequences of independent Wiener processes defined on some probability space $(\Omega,\mathcal{F},P)$, Γ_i is defined via (3.5). It is assumed that for all $(x,y) \in I\!R_+ \times I\!R$, $b_i(x,y)$ is uniformly bounded for $i=1,2$ and $0 < a_* \leq a_i(x,y) \leq a^*$ for $i=1,2$. Also, $b_i(\cdot,\cdot)$ and $a_i(\cdot,\cdot)$ are assumed be globally Lipschitz continuous.

To ensure that the above definition of the stochastic process (X,Y) is well-defined for all $t \in [0,\infty)$, we have the following lemma. For some of the proofs and calculations, we need to define an “unreflected” form of $X^{(i)}$

$$U^{(i)}(t) = iL + \int_0^t b_1(X^{(i)}(s),Y^{(i)}(s))ds + \int_0^t a_1(X^{(i)}(s),Y^{(i)}(s))dW^{(i)}(s). \tag{3.39}$$

Lemma 3.2.3. *For all $i \in I\!N_0$, $P(0 < \sigma^{(i)} < \infty) = 1$ and $\sigma^{(i)} \to \infty$ almost surely, as $i \to \infty$.*

Proof. Proof of the first part of the lemma is identical to that of Lemma 3.1.1. For the second part we will show that $\exists \delta, \epsilon \in (0,\infty)$ s.t. $P[\tau^{(j)} \geq \delta | \mathcal{F}_{j-1}] > \epsilon$, for all $j = 1, 2, ...$ where $\mathcal{F}_j = \sigma\{X(s), Y(s) : s \leq \sigma^{(j)}\}$. Now,

$$\begin{aligned}
P[\tau^{(j)} \leq \delta | \mathcal{F}_{j-1}] &= P[\sup_{0\leq s\leq \delta} |X^{(j)}(s) - jL| \geq L | \mathcal{F}_{j-1}] \\
&\leq P[\sup_{0\leq s\leq \delta} |U^{(j)}(s) - jL| \geq \frac{L}{2} | \mathcal{F}_{j-1}] \\
&\leq 2\frac{E[\sup_{0\leq s\leq\delta} |U^{(j)}(s) - jL| \Big| \mathcal{F}_{j-1}]}{L} \\
&\leq \frac{C\delta^{1/2}}{L},
\end{aligned}$$

where in the last step we have used the boundedness of coefficients b_1 and a_1. Select δ sufficiently small so that

$$\sum_{j=1}^{\infty} P[\tau^{(j)} \geq \delta | \mathcal{F}_{j-1}] = \infty \text{ a.s.}$$

Then from the Borel-Cantelli Lemma (see Corollary 3.2 of Chapter 4 of [5], we have that $P[\tau^{(j)} \geq \delta$ for infinitely many $j] = 1$. This proves the result. ■

We are now ready to prove the main result of this section. Let b_i be as above, and $\alpha_i(x,y) = \frac{a_i(x,y)}{2}$ for $i = 1, 2$. Recall the definition of (X_n, Y_n) given in (3.37).

Theorem 3.2.4. *The sequence (X_n, Y_n) converges weakly to (X, Y), in $\mathcal{D}([0,\infty) : \mathbb{R}_+ \times \mathbb{R})$, as $n \to \infty$.*

To prove the above theorem, we will show that (X_n^*, Y_n^*) converges weakly in $\mathcal{D}([0,\infty) : \mathbb{R}_+ \times \mathbb{R})$ as $n \to \infty$. The theorem will then follow from Proposition 3.2.1. Define recursively a family of processes $\{(\tilde{X}_n^{(i)}, \tilde{U}_n^{(i)}, \tilde{Y}_n^{(i)})\}_{i\in\mathbb{N}_0, n\in\mathbb{N}'}$ and stopping times

$\{\tau_n^{(i)}\}_{i\in\mathbb{N}_0, n\in\mathbb{N}'}$ as follows. For $i = 0$,

$$(\tilde{X}_n^{(i)}(t), \tilde{U}_n^{(i)}(t), \tilde{Y}_n^{(i)}(t)) \doteq \tilde{\Xi}_n^{(i)}(\lfloor n^2\lambda t\rfloor)$$

where $\{\Xi_n^{(i)}(k)\}_{k\in\mathbb{N}_0}$ is a discrete space Markov chain with state space $G_n \equiv S_n \times \tilde{S}_n \times \tilde{S}_n$ and transition function

$$\begin{aligned}
\hat{\mu}_n^{(i)}\big((x,u,y),\Lambda\big) &= \left(1 - \frac{\lambda_n^X(x,y) + \lambda_n^Y(x,y)}{n^2\lambda}\right)\delta_{(x,u,y)}(\Lambda) \\
&+ \frac{\lambda_n^X(x,y)}{n^2\lambda}\Big[p_n^X(x,y)\delta_{(x+1/n,u+1/n,y)}(\Lambda) \\
&+ (1 - p_n^X(x,y))\delta_{iL+(x-1/n-iL)^+,u-1/n,y)}(\Lambda)\Big] \\
&+ \frac{\lambda_n^Y(x,y)}{n^2\lambda}\Big[p_n^Y(x,y)\delta_{(x,u,y+1/n)}(\Lambda) \\
&+ (1 - p_n^Y(x,y))\delta_{(x,u,y-1/n)}(\Lambda)\Big],
\end{aligned} \tag{3.40}$$

where the initial condition of the above Markov chain is $\delta_{(0,0,y_0)}$. Also, define

$$\tau_n^{(i)} = \inf\{t : \tilde{X}_n^{(i)}(t) = (i+1)L\} \tag{3.41}$$

Having defined $(\tilde{X}_n^{(j)}(t), \tilde{U}_n^{(j)}(t), \tilde{Y}_n^{(j)}(t))$ for $j = 1,\ldots,i-1$, define for $j = i$,

$$(\tilde{X}_n^{(i)}(t), \tilde{U}_n^{(i)}(t), \tilde{Y}_n^{(i)}(t)) \doteq \tilde{\Xi}_n^{(i)}(\lfloor n^2\lambda t\rfloor)$$

where $\{\Xi_n^{(i)}(k)\}_{k\in\mathbb{N}_0}$ is a discrete space Markov chain with state space G_n and with transition function $\hat{\mu}_n^{(i)}$ defined via (3.40). The initial condition of this Markov chain is $\delta_{(iL,iL,\tilde{Y}_n^{(i-1)}(\tau_n^{(i-1)})}$ where $\tau_n^{(i-1)}$ is defined via (3.41). Set $\sigma_n^{(i)} = \sum_{j=0}^{i-1}\tau_n^{(j)}$.

Now, define

$$(\hat{X}_n, \hat{Y}_n)(t) = (\tilde{X}_n^{(i)}, \tilde{Y}_n^{(i)})(t - \sigma_n^{(i)}); \;\; t \in [\sigma_n^{(i)}, \sigma_n^{(i+1)}); \;\; i \in I\!N_0.$$

Note that, by construction, $(\hat{X}_n, \hat{Y}_n)$ has the same law as (X_n^*, Y_n^*). So, if we show that $(\hat{X}_n, \hat{Y}_n) \Rightarrow (X, Y)$, then we have proven the theorem. Next, let

$$\mathcal{X}_0' \doteq \mathcal{D}([0,\infty) : I\!R_+) \times \mathcal{D}([0,\infty) : I\!R_+) \times [0,\infty],$$

where $[0,\infty]$ denotes the one point compactification of $I\!R_+$. Let $\mathcal{X}' \doteq \mathcal{X}_0'^{\otimes\infty}$. We will endow $\mathcal{X}'$ with the usual topology and consider the Borel σ-field $\mathcal{B}(\mathcal{X}')$. Then for each n,

$$Z_n(i) \doteq (\tilde{X}_n^{(i)}, \tilde{Y}_n^{(i)}, \tau_n^{(i)}) \quad \text{and} \quad Z(i) \doteq (X^{(i)}, Y^{(i)}, \tau^{(i)}) \tag{3.42}$$

take values in $\mathcal{X}'$.

Defining

$$\begin{aligned} \tilde{\mathcal{X}} &\doteq \{(x_i, y_i, \beta_i)_{i \in I\!N_0} \in \mathcal{X} | 0 < \beta_i < \infty, x_i \in A, y_i \in C([0,\infty) : I\!R) \\ &\qquad\qquad \forall\, i \text{ , and } \textstyle\sum_{i=0}^{j} \beta_i \to \infty \text{ as } j \to \infty\} \end{aligned} \tag{3.43}$$

where A is as in (3.13) and $\tau(\phi(\cdot))$ is defined as (3.14). For each i, $X^{(i)} \in A$, the coefficients are bounded and the diffusion coefficient $a_1(x, y)$ is uniformly non-degenerate. We see from this fact and Lemma 3.2.3 that

$$P(Z \in \tilde{\mathcal{X}}) = 1. \tag{3.44}$$

The key step in the proof of Theorem 3.2.4 is to establish the weak convergence of $\{Z_n\}$ to Z. Observing that, conditioned on $\tilde{Y}_n^{(i-1)}(\tau_n^{(i-1)})$, $Z_n^{(i)}$ is independent of $\{Z_n^{(j)}, j < i\}$ and $(\tilde{X}_n^{(i-1)}(\tau_n^{(i-1)}), \tilde{Y}_n^{(i-1)}(\tau_n^{(i-1)})) = (\tilde{X}_n^{(i)}(0), \tilde{Y}_n^{(i)}(0))$, we have that it suffices to show that

$$(\tilde{X}_n^{(i)}, \tilde{Y}_n^{(i)}, \tau_n^{(i)}) \Rightarrow (X^{(i)}, \tilde{Y}^{(i)}, \tau^{(i)}), \text{ for all } i. \tag{3.45}$$

We will proceed inductively.

Proposition 3.2.5. *Assume that for some $i \in I\!N$ that $Z_n(i-1) \Rightarrow Z(i-1)$, then $Z_n(i) \Rightarrow Z(i)$.*

Proof. Using the Skorokhod representation theorem we can assume without loss of generality that $Z_n(i-1) \to Z(i-1)$ almost surely. Note that since $\tilde{Y}^{(i-1)}(\cdot)$ has continuous paths almost surely we have that $\tilde{Y}_n^{(i-1)}(\cdot) \to \tilde{Y}^{(i-1)}(\cdot)$ uniformly on compact intervals. Also, since $\tau_n^{(i-1)} \to \tau^{(i-1)}$ and $\tau^{(i-1)} < \infty$ a.s., we have that $\tilde{Y}_n^{(i-1)}(\tau_n^{(i-1)}) \to \tilde{Y}^{(i-1)}(\tau^{(i-1)})$. Noting that $\tilde{Y}_n^{(i)}(0) = \tilde{Y}_n^{(i-1)}(\tau_n^{(i-1)})$, we have that $\{\tilde{Y}_n^{(i)}(0)\}_{n \in I\!N'}$ is a tight family. Since $\tilde{X}_n^{(i)}(0) = \tilde{U}_n^{(i)}(0) = iL$, $\{\tilde{X}_n^{(i)}(0), \tilde{U}_n^{(i)}(0\}_{n \in I\!N'}$ is a tight family. In what follows once more, we will suppress i from the notation when needed. Define

$$N(t) = \max\{m : \frac{m}{n^2\lambda} \leq t\} \tag{3.46}$$

and

$$\begin{aligned}
\Delta_j \tilde{Y}_n &= \tilde{Y}_n\left(\frac{j+1}{n^2\lambda}\right) - \tilde{Y}_n\left(\frac{j}{n^2\lambda}\right) \\
\Delta' \tilde{Y}_n &= \tilde{Y}_n(t) - \tilde{Y}_n\left(\frac{N(t)}{n^2\lambda}\right) \\
\Delta_j \tilde{U}_n &= \tilde{U}_n\left(\frac{j+1}{n^2\lambda}\right) - \tilde{U}_n\left(\frac{j}{n^2\lambda}\right) \\
\Delta' \tilde{U}_n &= \tilde{U}_n(t) - \tilde{U}_n\left(\frac{N(t)}{n^2\lambda}\right).
\end{aligned}$$

Here E^n_j be expectation conditioned on $\mathcal{F}(\tilde{X}_n(s), \tilde{Y}_n(s), \tilde{U}_n(s), s \leq \frac{j}{n^2\lambda})$.

Define

$$\tilde{w}^{(i)}_n(t) = \sum_{\ell=0}^{N(t)-1} \frac{[\Delta_\ell \tilde{U}^{(i)}_n - E^n_\ell \Delta_\ell \tilde{U}^{(i)}_n]}{a_1(\tilde{X}^{(i)}_n(\frac{\ell}{n^2\lambda}), \tilde{Y}^{(i)}_n(\frac{\ell}{n^2\lambda}))}.$$

and

$$\tilde{B}^{(i)}_n(t) = \sum_{\ell=0}^{N(t)-1} \frac{[\Delta_\ell \tilde{Y}^{(i)}_n - E^n_\ell \Delta_\ell \tilde{Y}^{(i)}_n]}{a_2(\tilde{X}^{(i)}_n(\frac{\ell}{n^2\lambda}), \tilde{Y}^{(i)}_n(\frac{\ell}{n^2\lambda}))}.$$

We will next show that $\{(\tilde{X}_n, \tilde{U}_n, \tilde{Y}_n, \tilde{w}_n, \tilde{B}_n)\}$ is tight, for which it suffices to show that the marginals are tight. The tightness of $\tilde{U}_n$, $\tilde{Y}_n$, $\tilde{w}_n$, and $\tilde{B}_n$ will be shown by using Theorem 2.2.3, and the tightness of $\tilde{X}_n$ will be an immediate consequence of the continuity of the Skorokhod map.

Using the above definitions, one can easily establish the following local consistency conditions which will be used repeatedly in what follows and will confirm at least on an intuitive level that the $\tilde{U}_n(\cdot)$ processes approximate an unreflected version of X, and $\tilde{Y}_n(\cdot)$ will approximate Y.

Local Consistency Conditions

Let $x' = \tilde{X}_n(\frac{j}{n^2\lambda}), y' = \tilde{Y}_n(\frac{j}{n^2\lambda})$.

$$\begin{aligned} E_j^n \Delta_j \tilde{U}_n &= b_1(x', y')\frac{1}{n^2\lambda} + O\left(\frac{1}{n^3}\right) \\ E_j^n \Delta_j \tilde{Y}_n &= b_2(x', y')\frac{1}{n^2\lambda} + O\left(\frac{1}{n^3}\right) \\ E_j^n (\Delta_j \tilde{U}_n - E_j^n \Delta_j \tilde{U}_n)^2 &= a_1^2(x', y')\frac{1}{n^2\lambda} + O\left(\frac{1}{n^3}\right) \\ E_j^n (\Delta_j \tilde{Y}_n - E_j^n \Delta_j \tilde{Y}_n)^2 &= a_2^2(x', y')\frac{1}{n^2\lambda} + O\left(\frac{1}{n^3}\right) \end{aligned}$$

and

$$E_j^n (\Delta_j \tilde{U}_n - E_j^n \Delta_j \tilde{U}_n)(\Delta_j \tilde{Y}_n - E_j^n \Delta_j \tilde{Y}_n) = O\left(\frac{1}{n^3}\right)$$

where $O(\frac{1}{n^3})$ is a quantity which is bounded in absolute value by $\frac{C}{n^3}$ where C is a universal constant.

Using these local consistency conditions, we can establish the following inequality in a straightforward manner

$$E^n |\tilde{U}_n(t) - \tilde{U}_n(0)|^2 \leq 2 \left| Kt + N(t)O(\frac{1}{n^3}) \right|^2 + 2\left(K^2 t + N(t)O(\frac{1}{n^3}) \right), \qquad (3.47)$$

where K is the bound for the maximum of $|b_1^+|$, $|b_1^-|$ and $|a_1|$. Observing that $N(t) \leq n^2\lambda(t+1)$, we have that the first condition in Theorem 2.2.3 is satisfied.

For the second condition of Theorem 2.2.3, fix $T > 0$ and take an arbitrary stopping time ς s.t. $\varsigma \leq T$ $w.p.1$. Note that for $\delta' > 0$

$$
\begin{aligned}
E_x^n(1 \wedge |\tilde{U}_n(\varsigma + \delta') - \tilde{U}_n(\varsigma)|) &\leq (E_x^n|\tilde{U}_n(\varsigma + \delta') - \tilde{U}_n(\varsigma)|^2)^{1/2} \\
&\leq \left(2\left|K\delta' + N(\delta')O\left(\frac{1}{n^3}\right)\right|^2\right. \\
&\quad \left.+2\left(K^2\delta' + N(\delta')O\left(\frac{1}{n^3}\right)\right)\right)^{\frac{1}{2}},
\end{aligned}
$$

where K is a universal constant. Since the right side above converges to 0 as $\delta' \to 0$ and $n \to 0$, we have that the second condition in Theorem 2.2.3 holds. This shows that $\tilde{U}_n$ is tight. The tightness of $\tilde{X}_n$ is now an immediate consequence of the fact that $\tilde{X}_n = \Gamma(\tilde{U}_n)$ and that Γ is a continuous map. Proof of tightness of $\{\tilde{Y}_n, w_n, B_n\}$ is very similar. In particular, note that (3.47) follows for $\tilde{U}_n$ replaced by $\tilde{Y}_n$ in exactly the same manner. This along with the already proved tightness of $\{\tilde{Y}_n(0)\}$ gives that the first condition in Theorem 2.2.3 is satisfied. The remainder of the proof is virtually identical to that of $\tilde{U}_n$.

Denote the measure induced by $(\tilde{X}_n, \tilde{U}_n, \tilde{Y}_n, \tilde{w}_n, \tilde{B}_n)$ by Q_n. The above tightness shows that every subsequence of Q_n admits a convergent sequence. So, in order to prove the result it suffices to show that for any weakly convergent subsequence, the weak limit of $(\tilde{X}_{n'}^{(i)}, \tilde{U}_{n'}^{(i)}, \tilde{Y}_{n'}^{(i)}, \tilde{w}_{n'}^{(i)}, \tilde{B}_{n'}^{(i)})$, denoted by $(\tilde{X}^{(i)}, \tilde{U}^{(i)}, \tilde{Y}^{(i)}, \tilde{W}^{(i)}, \tilde{B}^{(i)})$ has the same law as $(X^{(i)}, U^{(i)}, Y^{(i)}, W^{(i)}, B^{(i)})$. We will use n for n' to simplify notation. First we will use Theorem 2.2.4 to show that the weak limit of $\tilde{w}_{n'}^{(i)}$ and $\tilde{B}_{n'}^{(i)}$ are independent Wiener process. Fix $t \geq 0, \zeta > 0, q < \infty, t_i \in [0, t]$ with $t_{i+1} > t_i$ for $i \in \{0, ..., q\}$, and let $f \in C_0^2(I\!R \times I\!R)$. Then,

$$\begin{aligned}
f((\tilde{w}_n, \tilde{B}_n)(t+\zeta)) &- f((\tilde{w}_n, \tilde{B}_n)(t)) - \frac{1}{2}\int_t^{t+\zeta} \sum_{i=1}^{2} f_{x_i x_i}((\tilde{w}_n, \tilde{B}_n)(s))ds \\
&= \sum_{i=N(t)+1}^{N(t+\zeta)-1} [f((\tilde{w}_n, \tilde{B}_n)(\frac{(i+1)}{n^2\lambda})) - f((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))] \\
&\quad + [f((\tilde{w}_n, \tilde{B}_n)(\frac{N(t)+1}{n^2\lambda})) - f((\tilde{w}_n, \tilde{B}_n)(t))] \\
&\quad + [f((\tilde{w}_n, \tilde{B}_n)(t+\zeta) - f((\tilde{w}_n, \tilde{B}_n)(\frac{N(t+\zeta)}{n^2\lambda}))] \\
&\quad - \frac{1}{2}\sum_{i=N(t)+1}^{N(t+\zeta)-1} \sum_{i=1}^{2} f_{x_i x_i}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{1}{n^2\lambda} + \epsilon_n,
\end{aligned}$$

where $E^n|\epsilon_n| \to 0$ as $n \to \infty$. Therefore,

$$\begin{aligned}
f((\tilde{w}_n, \tilde{B}_n)(t+\zeta)) &- f((\tilde{w}_n, \tilde{B}_n)(t)) - \frac{1}{2}\int_t^{t+\zeta} \sum_{i=1}^{2} f_{x_i x_i}((\tilde{w}_n, \tilde{B}_n)(s))ds \\
&= \sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_1}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{[\Delta_i \tilde{U}_n - E_i^n \Delta_i \tilde{U}_n]}{a_1((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))} \\
&\quad + \sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_2}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{[\Delta_i \tilde{Y}_n - E_i^n \Delta_i \tilde{Y}_n]}{a_2((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))} \\
&\quad + \frac{1}{2}\sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_1 x_2}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{[\Delta_i \tilde{U}_n - E_i^n \Delta_i \tilde{U}_n]}{a_1((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))}\frac{[\Delta_i \tilde{Y}_n - E_i^n \Delta_i \tilde{Y}_n]}{a_2((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))} \\
&\quad + \frac{1}{2}\sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_2 x_1}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{[\Delta_i \tilde{U}_n - E_i^n \Delta_i \tilde{U}_n]}{a_1((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))}\frac{[\Delta_i \tilde{Y}_n - E_i^n \Delta_i \tilde{Y}_n]}{a_2((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))} \\
&\quad + \frac{1}{2}\sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_1 x_1}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{[\Delta_i \tilde{U}_n - E_i^n \Delta_i \tilde{U}_n]^2}{a_1^2((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))} \\
&\quad - \frac{1}{2}\sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_1 x_1}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{1}{n^2\lambda} \\
&\quad + \frac{1}{2}\sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_2 x_2}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{[\Delta_i \tilde{Y}_n - E_i^n \Delta_i \tilde{Y}_n]^2}{a_2^2((\tilde{X}_n, \tilde{Y}_n)(\frac{i}{n^2\lambda}))} \\
&\quad - \frac{1}{2}\sum_{i=N(t)}^{N(t+\zeta)-1} f_{x_2 x_2}((\tilde{w}_n, \tilde{B}_n)(\frac{i}{n^2\lambda}))\frac{1}{n^2\lambda} \\
&\quad + \epsilon_n
\end{aligned}$$

Thus if H is a bounded, continuous function from $I\!R^{5\times q}$ to $I\!R$, then

$$\begin{aligned} & E\bigg|H(\tilde{X}_n(t_i),\tilde{U}_n(t_i),\tilde{Y}_n(t_i),\tilde{w}_n(t_i),\tilde{B}_n(t_i),1\le i\le q) \\ \times\ & \bigg[f((\tilde{w}_n,\tilde{B}_n)(t))-\frac{1}{2}\int_t^{t+\zeta}\sum_{i=1}^{2}f_{x_ix_i}((\tilde{w}_n,\tilde{B}_n)(s))ds\bigg]\bigg| \\ \le\ & K'(E^n|\epsilon_n|) \quad \to 0 \ \text{ as } \ n\to\infty. \end{aligned} \tag{3.48}$$

Since $(\tilde{X}_n,\tilde{U}_n,\tilde{Y}_n,\tilde{w}_n,\tilde{B}_n) \Rightarrow (\tilde{X},\tilde{U},\tilde{Y},\tilde{W},\tilde{B})$, using the Skorokhod representation theorem, we can assume without loss of generality that $(\tilde{X}_n,\tilde{U}_n,\tilde{Y}_n,\tilde{w}_n,\tilde{B}_n) \to (\tilde{X},\tilde{U},\tilde{Y},\tilde{W},\tilde{B})$, almost surely. Thus taking limit as $n\to\infty$ on the left side of (3.48) we have via an application of dominated convergence theorem that

$$\begin{aligned} & E\bigg|H(\tilde{X}(t_i),\tilde{U}(t_i),\tilde{Y}(t_i),\tilde{W}(t_i),\tilde{B}(t_i),1\le i\le q) \\ \times\ & \bigg[f((\tilde{W},\tilde{B})(t+\zeta))-f((\tilde{W},\tilde{B})(t))-\frac{1}{2}\int_t^{t+\zeta}\sum_{i=1}^{2}f_{x_ix_i}f((\tilde{W},\tilde{B})(s))ds\bigg]\bigg| \qquad (3.49) \\ & = 0. \qquad (3.50) \end{aligned}$$

Now, let $\mathcal{F}_t = \sigma\{\tilde{X}(s),\tilde{U}(s),\tilde{Y}(s),\tilde{W}(s),\tilde{B}(s), s\le t\}$ which is generated by functions of the form of $H(\cdot)$. Thus,

$$E\bigg[\bigg(f((\tilde{W},\tilde{B})(t+\zeta))-f((\tilde{W},\tilde{B})(t))-\frac{1}{2}\int_t^{t+\zeta}\sum_{i=1}^{2}f_{x_ix_i}f((\tilde{W},\tilde{B})(s))ds\bigg)I_F\bigg]=0 \tag{3.51}$$

for $F\in\mathcal{F}_t$, and so

$$E\bigg[f((\tilde{W},\tilde{B})(t+\zeta))-f((\tilde{W},\tilde{B})(t))-\frac{1}{2}\int_t^{t+\zeta}\sum_{i=1}^{2}f_{x_ix_i}f((\tilde{W},\tilde{B})(s))ds\bigg|\mathcal{F}_t\bigg]=0 \quad w.p.1 \tag{3.52}$$

Thus, $f((\tilde{W},\tilde{B})(t+\zeta)) - f((\tilde{W},\tilde{B})(t)) - \frac{1}{2}\int_t^{t+\zeta}\sum_{i=1}^2 f_{x_i x_i} f((\tilde{W},\tilde{B})(s))ds$ is an $\mathcal{F}_t$-martingale for all $f \in C_0^2(I\!R)$. Next, by construction of $\tilde{w}_n(t)$, $\tilde{U}_n(t)$, $\tilde{B}_n(t)$,$\tilde{Y}_n(t)$

$$|\tilde{w}_n(t) - \tilde{w}_n(t-)| \leq \frac{2}{n \times a_*} \tag{3.53}$$

$$|\tilde{B}_n(t) - \tilde{B}_n(t-)| \leq \frac{2}{n \times a_*} \tag{3.54}$$

for every t and therefore the jumps of $\tilde{w}_n(\cdot)$ and $\tilde{B}_n(\cdot)$ converge to zero uniformly in t. From Theorem 13.4 of [2], we now have that $\tilde{W}$ and $\tilde{B}$ have continuous paths almost surely. So, by Theorem 2.2.4 $\tilde{W}(\cdot)$ and $\tilde{B}(\cdot)$ are $\mathcal{F}_t$-Wiener processes.

Now, we will identify the limit process $(\tilde{X},\tilde{U},\tilde{Y})$. For each $\delta > 0$ and $t \in [j\delta,(j+1)\delta)$, define

$$\tilde{U}_n^\delta(t) \doteq \tilde{U}_n(j\delta) \qquad \tilde{U}^\delta(t) \doteq \tilde{U}(j\delta) \tag{3.55}$$

$$\tilde{X}_n^\delta(t) \doteq \tilde{X}_n(j\delta) \qquad \tilde{X}^\delta(t) \doteq \tilde{X}(j\delta) \tag{3.56}$$

$$\tilde{Y}_n^\delta(t) \doteq \tilde{Y}_n(j\delta) \qquad \tilde{Y}^\delta(t) \doteq \tilde{Y}(j\delta). \tag{3.57}$$

Also, define $N^\delta(t) = \max\{m : m\delta \leq t\}$ Using the definition of $\tilde{w}_n(\cdot)$ and $\tilde{B}_n(\cdot)$ and the local consistency properties (3.17), (3.18), we have that

$$\begin{aligned}
\tilde{U}_n^\delta(t) - \tilde{U}_n(0) &= \int_0^t b_1(\tilde{X}_n^\delta(s),\tilde{Y}_n^\delta(s))ds \\
&+ \sum_{j=0}^{N(t)-1} a_1(\tilde{X}_n^\delta(\delta i),\tilde{Y}_n^\delta(\delta i))[\tilde{w}_n(\delta(i+1)) - \tilde{w}_n(\delta i)] + \epsilon_n^{\delta,t} \\
\tilde{Y}_n^\delta(t) - \tilde{Y}_n(0) &= \int_0^t b_2(\tilde{X}_n^\delta(s),\tilde{Y}_n^\delta(s))ds \\
&+ \sum_{j=0}^{N(t)-1} a_2(\tilde{X}_n^\delta(\delta i),\tilde{Y}_n^\delta(\delta i))[\tilde{B}_n(\delta(i+1)) - \tilde{B}_n(\delta i)] + {\epsilon'}_n^{\delta,t}
\end{aligned}$$

where $E|\epsilon_n^{\delta,t}| \to 0$ and $E|{\epsilon'}_n^{\delta,t}| \to 0$ as $\delta \to 0$, uniformly in n and t in any bounded interval. Since $(\tilde{X}_n, \tilde{U}_n, \tilde{Y}_n) \to (\tilde{X}, \tilde{U}, \tilde{Y})$, we have that

$$(\tilde{X}_n^\delta(\cdot), \tilde{U}_n^\delta(\cdot), \tilde{Y}_n^\delta(\cdot)) \to (\tilde{X}^\delta(\cdot), \tilde{U}^\delta(\cdot), \tilde{Y}^\delta(\cdot))$$

with probability one in the D-space. Letting $n \to \infty$ in the above display, we have

$$\begin{aligned}
\tilde{U}^\delta(t) - \tilde{U}(0) &= \int_0^t b_1(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))ds \\
&+ \sum_{j=0}^{N(t)-1} a_1(\tilde{X}^\delta(\delta i), \tilde{Y}^\delta(\delta i))[\tilde{W}(\delta(i+1)) - \tilde{W}(\delta i)] + O(\delta) + \epsilon^{\delta,t} \\
\tilde{Y}^\delta(t) - \tilde{Y}(0) &= \int_0^t b_2(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))ds \\
&+ \sum_{j=0}^{N(t)-1} a_2(\tilde{X}_n^\delta(\delta i), \tilde{Y}^\delta(\delta i))[\tilde{B}(\delta(i+1)) - \tilde{B}(\delta i)] + O(\delta) + {\epsilon'}^{\delta,t}
\end{aligned}$$

where $E|\epsilon^{\delta,t}| \to 0$ and $E|{\epsilon'}^{\delta,t}| \to 0$ as $\delta \to 0$. So, using the boundedness of the coefficients we have

$$\begin{aligned}
\tilde{U}^\delta(t) - \tilde{U}(0) &= \int_0^t b_1(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))ds \\
&+ \int_0^t a_1(\tilde{X}^\delta(s), \tilde{Y}_n^\delta(s))d\tilde{W}(s) + \overline{\epsilon_{\delta,t}} \\
\tilde{Y}^\delta(t) - \tilde{Y}(0) &= \int_0^t b_2(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))ds \\
&+ \int_0^t a_2(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))d\tilde{B}(s) + \overline{{\epsilon'}_{\delta,t}}
\end{aligned}$$

where $E|\overline{\epsilon_{\delta,t}}| \to 0$and $E|\overline{{\epsilon'}_{\delta,t}}| \to 0$ as $\delta \to 0$. By the continuity of $a_i(\cdot), i = 1, 2,$ $\tilde{X}(\cdot)$, and $\tilde{Y}(\cdot)$, we have that

$$\int_0^t a_1(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))d\tilde{W}(s) \rightarrow \int_0^t a_1(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))d\tilde{W}(s) \tag{3.58}$$

$$\int_0^t a_2(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))d\tilde{B}(s) \rightarrow \int_0^t a_2(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))d\tilde{B}(s) \tag{3.59}$$

in probability, as $\delta \to 0$. Similarly, $\int_0^t b_i(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))ds \rightarrow \int_0^t b_i(\tilde{X}^\delta(s), \tilde{Y}^\delta(s))ds$ for $i = 1, 2$. Furthermore, since $\tilde{X}_n = \Gamma(\tilde{U}_n)$ for all n, we have that $\tilde{X} = \Gamma(\tilde{U})$. Therefore, the limit process $(\tilde{X}, \tilde{U})$ solves

$$\begin{aligned} \tilde{U}(t) &= \tilde{U}(0) + \int_0^t b_1(\tilde{X}(s), \tilde{Y}(s))ds + \int_0^t a_1(\tilde{X}(s), \tilde{Y}(s))d\tilde{W}(s) \\ \tilde{X}(t) &= \Gamma(\tilde{U})(t), \end{aligned}$$

and $\tilde{Y}$ solves

$$\tilde{Y}(t) = \tilde{Y}(0) + \int_0^t b_2(\tilde{X}(s), \tilde{Y}(s))ds + \int_0^t a_2(\tilde{X}(s), \tilde{Y}(s))d\tilde{B}(s) \tag{3.60}$$

Noting that $\tilde{Y}_n^{(i)}(0) = \tilde{Y}_n^{(i-1)}(\tau_n^{(i-1)})$ and recalling that $\tilde{Y}_n^{(i-1)}(\tau_n^{(i-1)}) \rightarrow \tilde{Y}^{(i-1)}(\tau^{(i-1)})$, we have $\tilde{Y}^{(i)}(0) = \tilde{Y}^{(i-1)}(\tau^{(i-1)})$.

By strong (and weak) uniqueness of the solution to the above equation, we now have that $(\tilde{X}, \tilde{U}, \tilde{W}, \tilde{Y}, \tilde{B})$ has the same probability law as $(X^{(i)}, U^{(i)}, W^{(i)}, Y^{(i)}, B^{(i)})$. Finally note that $X^{(i)} \in A$ a.s., where A is defined in 3.13. Thus, we have on using Lemma 3.1.6 that

$$Z_n(i) \rightarrow Z(i)$$

This proves Proposition 3.2.5. ■

Notice that since $\tilde{Y}_n^{(0)}(0) = y_0$ for all n, one can show exactly as in the proof of Proposition 3.2.5 that $Z_n(0) \Rightarrow Z(0)$. An immediate consequence of the above results is the following corollary.

Corollary 3.2.6. *Let $\{Z_n\}$ and Z be defined via 3.42. Then, the sequence $\{Z_n\}$ of $\mathcal{X}$ valued random elements converges weakly to Z.*

The final step in the proof of Theorem 3.2.4 is the following lemma. We begin with some notation. For $z = \{x^{(i)}, y^{(i)}, \tau^{(i)}\}_{i\geq 0} \in \mathcal{X}$, define $(x_z, y_z) \in \mathcal{D}([0,\infty) : I\!R_+ \times I\!R)$ as:

$$\begin{aligned} x_z(t) &\doteq x^{(i)}(t - \sigma^{(i)}) \\ y_z(t) &\doteq y^{(i)}(t - \sigma^{(i)}) \text{ if } \ t \in [\sigma^{(i)}, \sigma^{(i+1)}); \ \ i \in I\!N_0 \end{aligned}$$

where $\sigma^{(i)} = \sum_{j=1}^{i} \tau^{(j)}$ for $i \geq 1$, $\sigma^{(0)} \doteq 0$.

Lemma 3.2.7. *Let $\Psi : \mathcal{X} \to \mathcal{D}([0,\infty) : I\!R_+ \times I\!R)$ be defined as $\Psi(z) = (x_z, y_z)$, $z = \{x^{(i)}, y^{(i)}, \tau^{(i)}, y'^{(i)}\}_{i\geq 0}$, where (x_z, y_z) is given via (3.61). Then Ψ is continuous at every $z \in \tilde{\mathcal{X}}$.*

Proof. Take a sequence $\{z_n\} \in \mathcal{X}$ which converges to $z \in \tilde{\mathcal{X}}$. We will show that $\Psi(z_n) \to \Psi(z)$. It suffices to show that $\Psi(z_n)(t) \to \Psi(z)(t)$ uniformly on $[\sigma^{(i-1)}, \sigma^{(i)}]$ for all $i \in I\!N$. In order to prove this statement we will show the following two results.

(i)

$$\Psi(z_n)(t) \to \Psi(z)(t), \text{ uniformly for } \ t \in [\sigma^{(i-1)} + \delta, \sigma^{(i)} - \delta] \tag{3.61}$$

for any δ s.t. $0 < \delta < \frac{\sigma^i - \sigma^{i-1}}{2}$ and $i \in I\!N$.

(ii)

$$\Psi(z_n)(t) \to \Psi(z)(t), \text{ uniformly for } t \in [\sigma^{(i-1)} - \delta, \sigma^{(i-1)} + \delta] \tag{3.62}$$

for any δ s.t. $0 < \delta < (\sigma^{(i-1)} - \sigma^{(i-2)}) \wedge (\sigma^{(i)} - \sigma^{(i-1)})$ and $i \geq 2$.

Clearly, the result follows once (i) and (ii) are proven.

Now we prove (i). Note that since $\sigma_n^{(j)} \to \sigma^{(j)}$ for all j as $n \to \infty$, we can find $N \in I\!N$ s.t. for all $n \geq N$, $\sigma_n^{(j)} \in [\sigma^{(j)} - \delta, \sigma^{(j)} + \delta]$ for $j = i, i-1$. So, for t in the fixed interval $[\sigma^{(i-1)} + \delta, \sigma^{(i)} - \delta]$

$$\begin{aligned} \Psi(z_n)(t) &= (x_n^{(i)}, y_n^{(i)})(t - \sigma_n^{(i-1)}) \\ \Psi(z)(t) &= (x^{(i)}, y^{(i)})(t - \sigma^{(i-1)}) \end{aligned}$$

For simplicity, we will reparameterize with $s = t - \sigma^{(i-1)}$ and set $\varsigma_n = \sigma^{(i-1)} - \sigma_n^{(i-1)}$. Note that $|\varsigma_n| \leq \delta$.) Thus, for s in $C = [\delta, \tau^{(i)} - \delta]$,

$$\begin{aligned} \Psi(z_n)(s + \sigma^{(i-1)}) &= (x_n^{(i)}, y_n^{(i)})(s + \varsigma_n) \\ \Psi(z)(s + \sigma^{(i-1)}) &= (x^{(i)}, y^{(i)})(s) \end{aligned}$$

So, we need to show that $(x_n, y_n)(s+\varsigma_n) \to (x, y)(s)$ uniformly on C (suppressing the i). We know that $(x, y)(s)$ is uniformly continuous on $[0, \tau]$, and that $(x_n, y_n)(s) \to (x, y)(s)$ uniformly on compact intervals since $(x, y)(s)$ is continuous and $(x_n, y_n)(s)$ converges to $(x,y)(s)$ in the Skorokhod space.[2] Now,

$$\|(x,y)(s)-(x_n,y_n)(s+\varsigma_n)\| \leq \|(x,y)(s)-(x,y)(s+\varsigma_n))\|+\|(x,y)(s+\varsigma_n)-(x_n,y_n)(s+\varsigma_n)\| \tag{3.63}$$

for every $s \in C$. Thus,

$$\begin{aligned}
\sup_{s\in C}\|(x,y)(s)-(x_n,y_n)(s+\varsigma_n)\| &\leq \sup_{s\in C}\|(x,y)(s)-(x,y)(s+\varsigma_n))\| \\
&\quad +\sup_{s\in C}\|(x,y)(s+\varsigma_n)-(x_n,y_n)(s+\varsigma_n)\| \\
&\leq \sup_{s\in C}\|(x,y)(s)-(x,y)(s+\varsigma_n))\| \\
&\quad + \sup_{s\in[0,\tau]}\|(x,y)(s)-(x_n,y_n)(s)\|
\end{aligned}$$

Now, the right hand side goes to zero, since (x,y) is uniformly continuous, $\varsigma_n \to 0$ and $(x_n,y_n) \to (x,y)$ uniformly. This proves (i).

Now we consider (ii). Let us denote the interval $[\sigma^{(i-1)}-\delta, \sigma^{(i-1)}+\delta]$ by K. Let $s = t - \sigma^{(i-1)}$ and $\varsigma_n = \sigma_n^{(i-1)} - \sigma^{(i-1)}$, and again ensure large enough n to guarantee $\sigma_n^{(i)}$ are within δ of the limit.

$$\begin{aligned}
\sup_{s\in K}\|\Psi(z_n)(s)-\Psi(z)(s)\| &\leq \sup_{s\in K}\|\Psi(z_n)(s)-\Psi(z)(t-\varsigma_n)\| \\
&\quad +\sup_{s\in K}\|\Psi(z)(s-\varsigma_n)-\Psi(z)(s)\| \\
&\leq \sup_{r\in[\sigma^{(i-1)}-2\delta,\sigma^{(i-1)}]}\|(x_n^{(i-1)},y_n^{(i-1)})(r)-(x,y)(r)\| \\
&\quad \vee \sup_{r\in[0,2\delta]}\|(x_n^{(i)},y_n^{(i)})(r)-x(r)\| \\
&\quad +\sup_{s\in K}\|\Psi(z)(s-\varsigma_n)-\Psi(z)(s)\|
\end{aligned}$$

Once more the uniform convergence of $(x_n^{(i)}, y_n^{(i)})$ to $(x^{(i)}, y^{(i)})$ shows that the right side of the expression above converges to 0 as $n \to \infty$. This proves (ii) and hence the result.

■

Proof of Theorem 3.2.4 .

The proof is an immediate consequence of Corollary 3.2.6, Lemma 3.2.7, Proposition 3.2.1, and the continuous mapping theorem on observing that $(\hat{X}_n, \hat{Y}_n) = \Psi(Z_n)$, $(X, Y) = \Psi(Z)$, and $(\hat{X}_n, \hat{Y}_n) \stackrel{\mathcal{L}}{=} (X_n^*, Y_n^*)$. ■

Chapter 4

Asymptotics of a Diffusion Ratchet

An important physical feature of the molecular motors that are being modeled is their velocity which is defined to be

$$V(t) = \frac{X(t)}{t} \tag{4.1}$$

where $X(t)$ is the diffusion ratchet introduced in Section 3.1. We are interested in studying the asymptotic velocity, namely we will like to consider the asymptotics of $V(t)$ as $t \to \infty$. We will restrict our attention to the case where the drift and diffusion coefficients are periodic with period L. Such an assumption is quite natural in many situations because of the cyclic mechano-chemical steps involved in motor transport. The periodic case includes, as a special case, the commonly studied situation where the drift and diffusion coefficients are constant.

Our work has been motivated by work of Elston and Peskin [7] where the authors study the steady state velocity for an "imperfect ratchet". An imperfect ratchet is defined in terms of a tilted potential, namely, away from a barrier there is a negative linear potential or, in our terminology, a negative constant drift which represents the load F_l. Near the lower barrier, the drift changes in a continuous manner to a positive value $F_0 - F_l$ which represents the "ratchet strength". Thus if F_0 is large enough and the transition from negative drift to positive drift is rapid enough, it has the effect of almost instantaneously reflecting the motor at the lower barrier. However, in an imperfect ratchet model, irrespective of the value of F_0, the motor can, with a positive probability, cross the barrier to the left. This probability, of course, becomes smaller and smaller as F_0 increases. In [7], starting with the Fokker-Planck equation for a diffusion with constant drift and diffusion coefficients, and imposing periodicity conditions consistent with the barrier locations, the authors write a PDE for the transition density of the process. They then investigate the steady-state density of the "imperfect ratchet" by setting the partial derivative of the density with respect to t to zero. Using this new equation, the steady-state velocity is then calculated. The velocity here is defined to be the "net probability flux" of the steady-state distribution times the period (distance between barriers). The "net probability flux" can be thought of as the "asymptotic probability" that a particle moves across a certain area per time period. From these results they take formal limit of the steady state velocity as F_0 goes to infinity. This limit, in [7], is defined to be the steady state velocity of the "perfect ratchet". Intuitively, this last limit should correspond to the asymptotic velocity (i.e. $\lim_{t\to\infty} \frac{X(t)}{t}$) of the diffusion ratchet constructed in Section 2.

In our terminology, the model considered in [7] corresponds to a diffusion ratchet with constant mean (which we will call μ here) and constant diffusion parameter (which we will call σ). For $\mu \neq 0$ they find the steady state velocity of the "perfect ratchet" to be

$$\frac{D}{L}\frac{\omega_l^2}{e^{\omega_l}-1-\omega_l} \tag{4.2}$$

where $D = \frac{\sigma^2}{2}$ and $\omega_l = -\frac{2\mu}{\sigma^2}L$, and when $\mu = 0$ the steady state velocity is obtained to be

$$\frac{2D}{L}. \tag{4.3}$$

In this section we will show that for the periodic ratchet case, asymptotic velocity, i.e. $\lim_{t\to\infty}\frac{X(t)}{t}$ exists, almost surely. We will also obtain an explicit expression for the asymptotic velocity in the constant coefficient case and show that it equals the value obtained via steady state analysis undertaken in [7] (See Theorem 4.1.2). In order to capture the fluctuations of the velocity about its asymptotic value, we will prove a functional central limit theorem. This result also enables us to identify the effective diffusivity of the motor.

We begin with the following periodicity assumption.

Condition 4.0.8. *Assume that the coefficients $a, b : I\!R^+ \to I\!R$ satisfy*

$$a(x+L) = a(x), \;\; b(x+L) = b(x), \;\; \forall x \in I\!R^+.$$

4.1 Asymptotic Velocity

Theorem 4.1.1. *Suppose that the coefficients a, b satisfy Lipschitz continuirty, (3.7) and Condition 4.0.8. Let $X(\cdot)$ be the diffusion ratchet defined in Definition 3.1.2. Then $\frac{X(t)}{t}$ converges almost surely to $\frac{L}{E\tau^{(0)}}$, where $\tau^{(0)}$ is defined in (3.8).*

Proof. Let $X^{(i)}(\cdot)$ be given by (3.6). Condition 4.0.8 implies that $\tau^{(i)} \doteq \inf\{t :$

$X^{(i)}(t) \geq (i+1)L\}$ is an i.i.d. sequence. From Theorem 2.2.5 we have that $E\tau^{(i)} < \infty$. Next, let

$$n_t \doteq \inf\{m : \sum_{i=0}^{m-1} \tau^{(i)} \geq t\}. \tag{4.4}$$

Then

$$\begin{aligned} \frac{X(t)}{t} &= \frac{\sum_{i=0}^{n_t-1} X^{(i)}(\tau^{(i)}) + \epsilon_t}{t} \\ &= \frac{n_t L}{t} + \frac{\epsilon_t}{t} \end{aligned}$$

where $0 \leq \epsilon_t < L$. Next note that $n.$ is a renewal process. Thus by the renewal theorem (Theorem 5.2.1 of [6])

we have that, as $t \to \infty$, $\frac{n_t}{t} \to \frac{1}{E\tau^{(0)}}$, almost surely. Also, the second term above clearly goes to zero. This completes the proof. ■

We will now calculate the asymptotic velocity of a diffusion ratchet with constant coefficients.

Theorem 4.1.2. *Suppose that $b(x) = \mu$ and $a(x) = \sigma$ for all $x \in I\!R_+$. Then*

$$\lim_{t\to\infty} \frac{X(t)}{t} = \begin{cases} \frac{D}{L}\frac{\omega_l^2}{e^{\omega_l}-1-\omega_l} & \text{if } \mu \neq 0 \\ \\ \frac{2D}{L}. & \text{if } \mu = 0, \end{cases} \tag{4.5}$$

where $D = \frac{\sigma^2}{2}$ and $\omega_l = -\frac{2\mu}{\sigma^2}L$.

Proof. In view of Theorem 4.1.1, in order to calculate the asymptotic velocity , we only need to calculate $E\tau^{(0)}$. To do this we will use the Laplace transform, $\phi(\lambda) \doteq$

$E_0[e^{-\lambda\tau^{(0)}}]$. From Chapter 5 of [8], this Laplace transform is given as

$$\phi(\lambda) = \frac{\alpha + \beta}{\beta e^{-\alpha L} + \alpha e^{\beta L}} \tag{4.6}$$

where

$$\alpha \equiv \alpha(\lambda) = \frac{\sqrt{\mu^2 + 2\sigma^2\lambda}}{\sigma^2} + \frac{\mu}{\sigma^2}$$
$$\beta \equiv \beta(\lambda) = \frac{\sqrt{\mu^2 + 2\sigma^2\lambda}}{\sigma^2} - \frac{\mu}{\sigma^2}$$

Note that

$$\alpha'(\lambda) = \beta'(\lambda) = \frac{1}{\sqrt{\mu^2 + 2\sigma^2\lambda}} \tag{4.7}$$

Now,

$$\phi'(\lambda) = \frac{(\alpha' + \beta')(\beta e^{\alpha L} + \alpha e^{-\beta L}) - (\alpha + \beta)(\beta' e^{-\alpha L} - L\alpha'\beta e^{-\alpha L} + \alpha' e^{\beta L} + L\alpha\beta' e^{\beta L})}{(\beta e^{-\alpha L} + \alpha e^{\beta L})^2} \tag{4.8}$$

Note that,

$$E\tau^{(0)} = -\phi'(0).$$

First, we will evaluate $\phi'(0)$ with $\mu > 0$. We then get

$$\alpha(0) = \frac{2\mu}{\sigma^2}, \quad \beta(0) = 0, \quad \alpha'(0) = \beta'(0) = \frac{1}{\mu}.$$

which yields

$$\phi'(0) = \frac{\frac{2}{\mu}\frac{2\mu}{\sigma^2} - \frac{2\mu}{\sigma^2}\left(\frac{1}{\mu}e^{\frac{-2\mu}{\sigma^2}L} + \frac{1}{\mu} + \frac{2\mu}{\sigma^2}\frac{1}{\mu}L\right)}{\left(\frac{2\mu}{\sigma^2}\right)^2} \tag{4.9}$$

Simplifying gives

$$\phi'(0) = \left(1 - \frac{2\mu}{\sigma^2}L - e^{\frac{-2\mu}{\sigma^2}L}\right)\frac{\sigma^2}{2\mu^2}. \tag{4.10}$$

When $\mu < 0$, we have that

$$\alpha(0) = 0, \quad \beta(0) = \frac{2|\mu|}{\sigma^2}, \quad \alpha'(0) = \beta'(0) = \frac{1}{|\mu|}$$

which leads to the expression (4.10) for $\phi'(0)$. So,

$$\begin{aligned} \frac{L}{E\tau^{(0)}} &= \frac{L}{-\phi'(0)} \\ &= \frac{L}{e^{\frac{-2\mu}{\sigma^2}L} - 1 + \frac{2\mu}{\sigma^2}L} \\ &= \frac{L^2\left(\frac{4\mu^2}{\sigma^4}\right)}{L\frac{2}{\sigma^2}\left(e^{\frac{-2\mu}{\sigma^2}L} - 1 + \frac{2\mu}{\sigma^2}L\right)} \end{aligned}$$

Substituting $D = \frac{\sigma^2}{2}$ and $\omega_l = -\frac{2\mu}{\sigma^2}L$, we have the result for the case $\mu \neq 0$.

We now have consider the case when μ is 0. In this case,

$$\phi(\lambda) = \frac{2}{e^{-\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L} + e^{\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L}} \tag{4.11}$$

and

$$\phi'(\lambda) = \frac{-2\left(e^{\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L} - e^{-\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L}\right)\frac{L}{\sqrt{2\sigma^2\lambda}}}{\left(e^{-\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L} + e^{\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L}\right)^2} \tag{4.12}$$

which we rewrite (using a Taylor expansion) as

$$\phi'(\lambda) = \frac{-2\left(2\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L + o(\sqrt{\lambda})\right)\frac{L}{\sqrt{2\sigma^2\lambda}}}{\left(e^{-\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L} + e^{\frac{\sqrt{2\sigma^2\lambda}}{\sigma^2}L}\right)^2} + o(\lambda) \quad \text{as } \lambda \to 0 \tag{4.13}$$

Now, taking a limit, we have

$$\lim_{\lambda \searrow 0} \phi'(\lambda) = \frac{-L^2}{\sigma^2} \tag{4.14}$$

Using this expression and the fact that $D = \frac{\sigma^2}{2}$, we have that

$$\frac{L}{E\tau^{(0)}} = \frac{L}{\frac{2L^2}{\sigma^2}} = \frac{2D}{L} \tag{4.15}$$

This completes the proof. ■

The above theorem gives an explicit form for the asymptotic velocity for the relatively simple case of constant drift and diffusion coefficients. In the general situation of a state varying drift and diffusion, explicit calculation of the Laplace transform of $\tau^{(0)}$ becomes untenable. This difficulty also arises when one considers the most basic model of a motor which is pulling a cargo. In Chapter 6, we will explore various numerical methods for approximating the asymptotic velocities for situations where exact calculations are not possible.

4.2 Functional Central Limit Theorem

In this section, we prove a functional central limit theorem for the asymptotic velocity of a diffusion ratchet. Recall that we denote a diffusion ratchet as $X(t)$ and denote its "velocity" as $V_t = \frac{X(t)}{t}$. The following theorem gives a functional central limit result for fluctuations of $V(t)$ from the asymptotic velocity, $\frac{L}{\mu}$. This result allows us to identify the effective diffusivity of the ratchet which combined with the asymptotic velocity can be connected to the "randomness parameter".

Theorem 4.2.1 (FCLT for Asymptotic Velocity).

$$\frac{1}{\sqrt{n}}\left(X(n\cdot) - \frac{L}{\mu}\cdot\right) \text{ converges weakly to } \frac{L\sigma}{\mu^{3/2}}B(\cdot)$$

in $C([0,\infty) : I\!R)$ as $n \to \infty$, where $\mu = E\tau^{(0)}$, $\sigma^2 = Var(\tau^{(0)})$, and $\{B(t)\}$ is a standard Brownian motion.

The above Theorem yields the effective diffusivity of the biological motor as$\frac{L^2\sigma^2}{\mu^3}$.

Proof. Similar to other proofs in this chapter, we will proceed by reducing $X(t)$ to a renewal process plus an error. We can write

$$X(t) = L\nu(t) + \epsilon_t \tag{4.16}$$

where $\nu(t) = \max\{k : \sum_{i=0}^{k-1} \tau^{(i-1)} \leq t\}$ and $\epsilon_t = \lfloor \frac{X(t)}{L} \rfloor L$. Thus,

$$X(nt) = L\nu(nt) + \epsilon_{nt}. \tag{4.17}$$

Centering and normalizing, we obtain

$$\frac{\mu^{3/2}}{L\sigma\sqrt{n}}\left(X(nt)-\frac{Lnt}{\mu}\right)=\frac{\nu(nt)-\frac{nt}{\mu}}{\sigma\mu^{-3/2}\sqrt{n}}+\frac{\epsilon_{nt}}{L\sigma\mu^{-3/2}\sqrt{n}} \tag{4.18}$$

Since $0\leq\epsilon_{nt}<L$, $\sup_{0\leq s\leq t}\left|\frac{\epsilon_{ns}}{L\sigma\mu^{-3/2}\sqrt{n}}\right|\to 0$ for all t with probability one as $n\to\infty$. The first term on the right side is a standard Brownian motion by Theorem 14.6 of [2]. This proves the result. ■

Chapter 5

Asymptotic Velocity of a Biomolecular Motor Pulling a Cargo

In Chapter 3, we presented a diffusion ratchet model for the dynamics of a biological motor pulling a cargo. The dynamics were described via a system of coupled reflected stochastic differential equations. In this chapter, we will consider a special case where the linkage between the motor and cargo is given by a linear spring. More precisely, letting X(t) and Y(t) represent the location of the motor and the cargo, respectively; at time t the dynamics of $(X(t), Y(t))$ is given as in Definition 3.2.2 with $b_1(x, y)$ replaced by $-\beta_1(x-y)$ and $b_2(x, y)$ replaced by $-\beta_2(y-x)$, where $\beta_i \in (0, \infty), i = 1, 2$. Furthermore, for the sake of simplicity we take $a_1(x, y) \equiv a_1$ and $a_2(x, y) \equiv a_2$ where $a_1, a_2 \in (0, \infty)$ are constants. A somewhat more convenient and explicit way to denote the dynamics in this case is as follows.

$$
\begin{aligned}
X(0) &= x \\
Y(0) &= z + x \\
dX(t) &= -\beta_1(X(t) - Y(t))dt + a_1 dW_1(t) + dl(t) \\
dY(t) &= -\beta_2(Y(t) - X(t))dt + a_2 dW_2(t) \\
x &\in [0, L), z \in I\!R,
\end{aligned}
$$

where

$$
\begin{aligned}
l(t) &= \sum_{j=1}^{J(t)-1} l_j + L_{J(t)-1}(t) \\
J(t) &= \inf\{i : \sigma_i \geq t\} \\
\sigma_i &= \inf\{t : X(t) \geq (i+1)L\} \text{ for } i \in I\!N, \sigma_0 \doteq 0 \\
l_j &= L_{j-1}(\sigma_j) - L_{j-1}(\sigma_{j-1})
\end{aligned}
$$

with

$$
L_j(t) = -\inf_{\sigma_j \leq s \leq t} \left(\left(-\beta_1 \int_{\sigma_j}^{s} Z(u)du + W_1(s) - W_1(\sigma_j) \right) \bigwedge 0 \right) I_{[\sigma_j, \infty)}(t) \tag{5.1}
$$

and

$$
Z(t) \doteq X(t) - Y(t).
$$

Notice that if $X(t)$ is greater than $Y(t)$ then there is a positive drift in the cargo dynamics and a negative drift in the motor dynamics. The reverse is true when $Y(t)$

is greater than $X(t)$. Thus one expects that the distribution of $Z(t) \equiv X(t) - Y(t)$ converges to a stationary distribution as $t \to \infty$. The evolution of $Z(t)$ is described by the following equation:

$$dZ(t) = -\beta Z(t)dt + \sigma dW(t) + dl(t); Z(0) = z, \tag{5.2}$$

where $\beta \doteq \beta_1 + \beta_2$, $\sigma \doteq \sqrt{a_1^2 + a_2^2}$, and

$$W(\cdot) \doteq \frac{a_1^2}{\sqrt{a_1^2 + a_2^2}} W_1(\cdot) + \frac{a_2^2}{\sqrt{a_1^2 + a_2^2}} W_2(\cdot)$$

is a standard Brownian motion.

In this chapter we will show that the asymptotic velocity of the motor is well-defined; namely, the quantity $\frac{X(t)}{t}$ converges in probability to a deterministic value. Note that

$$\begin{aligned} \frac{X(t)}{t} &= \frac{Y(t)}{t} - \frac{Z(t)}{t} & (5.3) \\ &= \frac{y}{t} - \frac{\beta_2}{t} \int_0^t Z(s)ds + a_2 \frac{W_2(t)}{t} - \frac{Z(t)}{t}. & (5.4) \end{aligned}$$

We will show that $\frac{Z(t)}{t} \to 0$ in probability and $\frac{1}{t} \int_0^t Z(s)ds$ converges in probability to a deterministic constant μ as $t \to \infty$. This prove the main result of this chapter, namely,

$$\text{l.i.p.}_{t \to \infty} \frac{X(t)}{t} = -\beta_2 \mu,$$

where l.i.p. denotes limit in probability. The key lemma in the proof of the result is the following.

Lemma 5.0.2. *There exist $c, \alpha, b, \kappa, \delta_0 \in (0, \infty)$ such that for all $z \in I\!R, x \in [0, L)$ and* *all* $\delta_0 \leq \Delta$,

$$E \sup_{0 \leq s \leq \Delta} |Z(s) - z|^2 \leq C\Delta(1 + \Delta z^2) \tag{5.5}$$

and

$$E_{x,z} Z^2(\Delta) - z^2 \leq -\alpha \Delta z^2 + b I_{\{|z| < \kappa\}} \tag{5.6}$$

where $E_{x,z}$ *refers to expectation with respect to probability measure under which* $X(0) = x$ *and* $Y(0) = x + z$ *a.s.*

Remark Denoting $U(t) \doteq L \lfloor \frac{X(t)}{L} \rfloor$, we see that $(U(t), Z(t))$ is a Markov process with respect to $\mathcal{F}_t \doteq \sigma\{X(s), Y(s) : s \leq t\}$. We will prove Lemma 5.0.2 later in the chapter. We begin by observing the following important consequence of the Lemma.

Lemma 5.0.3. *For all* $\Delta \leq \delta_0$ *and* $\pi \equiv (x, z) \in [0, L) \times \mathbb{R}$, *there exist* d_1, d_2 *(possibly depending on* Δ*) such that*

$$\limsup_{n \to \infty} E_{x,z}(Z^2(n\Delta)) < d_1 |\pi|^2 + d_2.$$

Proof. Let $\Pi(t) \doteq (U(t), Z(t))$. Then observing that the first component of $\Pi(t)$ takes values in a compact set $[0, L]$. We have from Lemma 5.0.2 that for some $\tilde{b}, \tilde{\kappa} \in (0, \infty)$,

$$E_\pi |\Pi(n\Delta)|^2 - |\pi|^2 \leq -\alpha \Delta |\pi|^2 + \tilde{b} I_{\{|\pi| \leq \tilde{\kappa}\}}, \tag{5.7}$$

for all $\pi \in [0, L] \times \mathbb{R}$. Noting that 5.7 implies the condition (V2) (from page 262 of [17]), we have from Theorem 12.3.4 of [17] that the Markov chain $\{\Pi(n\Delta)\}_{n \geq 1}$ has at least one invariant measure. Due to the non-degeneracy of the diffusion coefficients a_1, a_2, this Markov chain, $\{\Pi(n\Delta)\}_{n \geq 1}$, is ψ-irreducible in the sense of Section 4.2 of [17] for $\psi = \lambda_1 \otimes \lambda_2$ where λ_1 is the normalized Lebesgue measure on $[0, L]$ and λ_2 is the standard normal measure on $I\!R$. Furthermore the chain is strongly aperiodic in the sense of Section 5.4.3 of [17]. This shows that the chain, $\{\Pi(n\Delta)\}_{n\geq 1}$, has a unique invariant measure. Denote this measure by ν_Δ.

From (5.7), we have that condition (V3) from page 337 of [17] is satisfied with $f(\pi) = \frac{\alpha}{2}\Delta|z|^2 + 1, \pi \equiv (x,z)$. From Theorem 14.2.6 and Proposition 14.3.1 [17], we then obtain

$$E_{x,z}(Z(n\Delta)^2) \to \int_{[0,L]\times I\!R} z^2 \nu_\Delta(dx,dz) \text{ as } n \to \infty. \tag{5.8}$$

Furthermore form Theorem 14.2.3(i) of [17] and the representation for the invariant measure in Theorem 10.0.1 of [17], we obtain

$$\int_{[0,L]\times I\!R} z^2 \nu_\Delta(dx,dz) \leq \frac{2|\pi|^2}{\alpha\Delta} + C_3, \tag{5.9}$$

where $C_3 \in (0,\infty)$ is independent of Δ. Combining (5.8) and (5.9), we obtain the result. ■

Lemma 5.0.4. *For all $(x,z) \in [0,L] \times I\!R$,*

$$\frac{Z(t)}{t} \to 0$$

in $P_{x,z}$-probability as $t \to \infty$.

Proof. Note that

$$\frac{Z(t)}{t} = \frac{Z(\Delta\lfloor \frac{t}{\Delta}\rfloor)}{t} + \frac{Z(t) - Z(\Delta\lfloor \frac{t}{\Delta}\rfloor)}{t}$$

Thus,

$$\limsup_{t\to\infty} E_{x,z}\left|\frac{Z(t)}{t}\right| \le \limsup_{t\to\infty}\frac{E_{x,z}|Z(\Delta\lfloor\frac{t}{\Delta}\rfloor)|}{t} + \limsup_{t\to\infty}\frac{E_{x,z}|Z(t)-Z(\Delta\lfloor\frac{t}{\Delta}\rfloor)|}{t}.$$

From Lemma 5.0.3, we have that the first term above equals zero. Also, from Lemma 5.0.2, we have that

$$\limsup_{t\to\infty}\frac{E_{x,z}|Z(t)-Z(\Delta\lfloor\frac{t}{\Delta}\rfloor)|}{t} \le \limsup_{t\to\infty}\frac{\sqrt{C\Delta(1+\Delta E|Z(\Delta\lfloor\frac{t}{\Delta}\rfloor)|^2}}{t}.$$

Applying Lemma 5.0.3 again, we see that the expression on the right side equals 0. This proves the result. ■

Proposition 5.0.5. *The Markov process* $\{\Pi(t)\}_{t\ge 0}$ *admits a unique invariant measure:* ν.

Proof. It suffices to show that the family of probability measures, $\{\nu_t, t \ge 0\}$, where

$$\nu_t(A) \doteq \frac{1}{t}\int_0^t P_{x,z}[Z(s)\in A]ds, A \in \mathcal{B}(I\!\!R), \tag{5.10}$$

is tight. We will show that

$$\sup_t \int_{[0,L]\times I\!\!R} |\tilde{z}|^2 d\nu_t(d\tilde{x}, d\tilde{z}) < \infty. \tag{5.11}$$

This clearly will prove the required tightness and hence prove the result. Note that

$$\int_{[0,L]\times I\!\!R} |\tilde{z}|^2 d\nu_t(d\tilde{x}, d\tilde{z}) = \frac{1}{t}\int_0^t E_{x,z}|Z(s)|^2 ds. \tag{5.12}$$

Also, note that

$$E_{x,z}|Z(s)|^2 = E_{x,z}|Z(\Delta\lfloor\frac{s}{\Delta}\rfloor)|^2 + E_{x,z}|Z(s) - Z(\Delta\lfloor\frac{s}{\Delta}\rfloor)|^2 \tag{5.13}$$

$$\leq E_{x,z}|Z(\Delta\lfloor\frac{s}{\Delta}\rfloor)|^2 + C\Delta(1 + \Delta E|Z(\Delta\lfloor\frac{t}{\Delta}\rfloor)|^2 \tag{5.14}$$

$$\leq C\Delta + (1 + C\Delta^2)(d_1|z|^2 + d_2), \tag{5.15}$$

where the first inequality follows from Lemma 5.0.2 while the second follows from Lemma 5.0.3. Substituting the above inequality into 5.12, we get

$$\sup_t \int_{[0,L]\times I\!R} |\tilde{z}|^2 d\nu_t(d\tilde{x}, d\tilde{z}) \leq C\Delta + (1 + C\Delta^2)(d_1|z|^2 + d_2) \tag{5.16}$$

$$< \infty. \tag{5.17}$$

This proves the result. ■

We now come to the main result of the chapter.

Theorem 5.0.6. *For all* $(x, z) \in [0, L] \times I\!R$*,*

$$\frac{X(t)}{t} \to -\beta_2 \int_{[0,L]\times I\!R} z\nu(dx, dz)$$

in $P_{x,z}$ *probability as* $t \to \infty$*; where* ν *is as in Proposition 5.0.5.*

Proof. Observing that

$$\frac{X(t)}{t} = \frac{Y(t)}{t} - \frac{Z(t)}{t} \tag{5.18}$$

$$= \frac{y}{t} - \frac{\beta_2}{t}\int_0^t Z_s ds + \frac{a_2 W_2(t)}{t} - \frac{Z(t)}{t}, \tag{5.19}$$

we have from Lemma 5.0.4 that

$$\limsup_{t\to\infty} E\left|\frac{X(t)}{t} + \beta_2 \int_{[0,L]\times I\!R} z\nu(du,dz)\right| = \beta_2 \limsup_{t\to\infty} E_{u,z}\left|\frac{1}{t}\int_0^t Z(s)ds - \int_{[0,L]\times I\!R} z\nu(du,dz)\right|$$

From the ergodicity proven in Proposition (5.0.5), we have for all $k \in I\!N$

$$E_{u,z}\left|\frac{1}{t}\int_0^t f_k(Z(s))ds - \int_{[0,L]\times I\!R} f_k(z)\nu(du,dz)\right| \to 0$$

as $t \to \infty$, where $f_k(z) = (z \vee k) \wedge (-k)$. The result will follow from the dominated convergence theorem if we have that

$$\sup_t \frac{1}{t}\int_0^t E_{x,z}|Z(s)|ds < \infty.$$

However, the above is an immediate consequence of (5.11) which was established in the proof of Proposition 5.0.5. This proves the result. ■

We will now provide the proof of Lemma 5.0.2 which was critically used in the proof of the above theroem. In what follows $c_1, c_2, ...$ will denote generic constants whose values change from one proof to the next. We begin by obtaining some preliminary bounds.

In order to bound $EZ^2(\Delta) - z^2$, we first obtain a bound on the "push" term as follows:

$$l(t) \leq \beta_1 t \sup_{0\leq u\leq t}|Z(u)| + a_1 \sum_{j=1}^{\infty} \underbrace{\sup_{\sigma_{j-1}\wedge t\leq s\leq \sigma_j \wedge t}|W_1(s) - W_1(\sigma_{j-1})|}_{M_j(t)}. \quad (5.20)$$

Writing $l(t)$ as

$$l(t) = \underbrace{l(t) - \beta_1 t \sup_{0\le u\le t} |Z(u)|}_{\tilde{l}(t)} + \beta_1 \underbrace{t \sup_{0\le u\le t} |Z(u)|}_{l^*(t)}, \tag{5.21}$$

we have from (5.20) that

$$\tilde{l}(t) \le a_1 \sum_{j=1}^{\infty} M_j(t). \tag{5.22}$$

The above bound enables us to prove the following inequality.

Lemma 5.0.7.

$$E\tilde{l}^2(t) \le ct(\sqrt{E[X^2(t)]} + 1). \tag{5.23}$$

Proof. Observe that

$$E\tilde{l}^2(t) \le a_1^2 \sum_{j=1}^{\infty} EM_j^2(t) + 2a_1^2 \sum_{k=1}^{\infty} \sum_{i=k+1}^{\infty} EM_i(t)M_k(t). \tag{5.24}$$

From Burkholder-Grundy-Davis inequalities, we have that

$$EM_j^2(t) \le c_1 E(\sigma_j \wedge t - \sigma_{j-1} \wedge t), \tag{5.25}$$

where c_1 is a universal constant.

This immediately yields that

$$\sum_{j=1}^{\infty} EM_j^2(t) \le c_1 E \sum_{j=1}^{\infty} (\sigma_j \wedge t - \sigma_{j-1} \wedge t) = c_1 t. \tag{5.26}$$

Let $\mathcal{G}_t = \sigma\{W_i(s) : s \le t, i = 1,2\}$, and set $\mathcal{F}_k \doteq \mathcal{G}_{\sigma_k}$. Then,

$$
\begin{aligned}
E[\sum_{i=k+1}^{\infty} M_i(t)|\mathcal{F}_k] &= E[\sum_{i=k+1}^{J(t)} M_i(t)|\mathcal{F}_k] \\
&\le E[\left(\sum_{i=k+1}^{\infty} M_i^2(t)\right)^{1/2} J(t)^{1/2}|\mathcal{F}_k] \\
&\le \left(E[\sum_{i=k+1}^{\infty} M_i^2(t)|\mathcal{F}_k]\right)^{1/2} (E[J(t)|\mathcal{F}_k])^{1/2} \\
&\le \sqrt{c_1}\sqrt{t}\sqrt{E[J(t)|\mathcal{F}_k]}, \qquad (5.27)
\end{aligned}
$$

where in the last step we have once more used the Burkholder-Grundy-Davis inequalities. An immediate consequence of the above inequality is the following:

$$
\begin{aligned}
\sum_{k=1}^{\infty}\sum_{i=k+1}^{\infty} E[M_k(t)M_i(t)] &= E\sum_{k=1}^{\infty} M_k(t)E[\sum_{i=k+1}^{\infty} M_i(t)|\mathcal{F}_k] \\
&\le \sqrt{c_1}\sqrt{t}E\sum_{k=1}^{J(t)} M_k(t)\sqrt{E[J(t)|\mathcal{F}_k]} \\
&\le \sqrt{c_1}\sqrt{t}E\left(\left(\sum_{k=1}^{\infty} M_k^2(t)\right)^{1/2}\left(\sum_{k=1}^{\infty} 1_{\{k\le J(t)\}}E[J(t)|\mathcal{F}_k]\right)^{1/2}\right) \\
&\le \sqrt{c_1}\sqrt{t}\left(E\sum_{k=1}^{\infty} M_k^2(t)\right)^{1/2}\left(E\sum_{k=1}^{\infty} 1_{\{k\le J(t)\}}E[J(t)|\mathcal{F}_k]\right)^{1/2} \\
&\le c_1 t\left(E\sum_{k=1}^{\infty} 1_{\{k\le J(t)\}}E[J(t)|\mathcal{F}_k]\right)^{1/2}, \qquad (5.28)
\end{aligned}
$$

where we have used (5.27) in the first inequality and (5.26) in the fourth inequality. On observing that the event $\{k \le J(t)\}$ is $\mathcal{F}_k$-measurable. We rewrite the right side of

(5.28) as

$$\begin{aligned} c_1 t \left(E \sum_{k=1}^{\infty} 1_{\{k \le J(t)\}} E[J(t)|\mathcal{F}_k] \right)^{1/2} &= c_1 t \left(E \sum_{k=1}^{\infty} E[J(t) 1_{\{k \le J(t)\}} |\mathcal{F}_k] \right)^{1/2} \\ &= c_1 t \left(E \sum_{k=1}^{\infty} J(t) 1_{\{k \le J(t)\}} \right)^{1/2} \\ &= c_1 t \sqrt{E J^2(t)}. \end{aligned} \tag{5.29}$$

Substituting (5.26), (5.28), and (5.29) into (5.24), we have

$$E\tilde{l}^2(t) \le c_2 (t + 2t\sqrt{E[J^2(t)]}).$$

Finally, observing that $J(t) \le \frac{X(t)+L}{L}$, we obtain from the above inequality that

$$E\tilde{l}^2(t) \le ct(\sqrt{E[X^2(t)]} + 1),$$

for a suitable constant c. ■

Next we obtain a bound on $E(X^2(t))$.

Lemma 5.0.8. *There exists a* $\gamma_1 \in (0, \infty)$ *such that for all* $t \ge 0$*,*

$$E(X(t)^2) \le \gamma_1 (1 + t^2 + t^2 E \sup_{0 \le u \le t} |Z(u)|^2). \tag{5.30}$$

Proof. Recall that for $t \in [0, \infty)$,

$$X(t) = x - \beta_1 \int_0^t Z(s) ds + a_1 W_1(s) + l(t).$$

Thus recalling that $x \in [0, L]$, we have from Lemma 5.0.7 that

$$\begin{aligned} EX^2(t) &\leq c_1[L^2 + \beta_1^2 t^2 E \sup_{0\leq u\leq t} |Z(u)|^2 + t + t(\sqrt{E[X^2(t)]} + 1)] \\ &\leq c_2[1 + t + t^2 E \sup_{0\leq u\leq t} |Z(u)|^2 + t\sqrt{EX^2(t)}]. \end{aligned} \tag{5.31}$$

Set

$$d_1 = c_2[1 + t + t^2 E \sup_{0\leq u\leq t} |Z(u)|^2] \tag{5.32}$$

$$d_2 = c_2 t \tag{5.33}$$

$$EX^2(t) \doteq \alpha^2. \tag{5.34}$$

Then (5.31) can be rewritten as $\alpha^2 \leq d_1 + d_2\alpha$. This immediately yields that

$$\alpha^2 \leq c_3(1 + t + t^2 E \sup_{0\leq u\leq t} |Z(u)|^2).$$

Hence, the result is shown. ■

Combining (5.0.7) and (5.0.8), we have that.

$$\begin{aligned} E\tilde{l}^2(t) &\leq t\sqrt{\gamma_1}(\sqrt{1 + t^2 + t^2 E \sup_{0\leq u\leq t} |Z(u)|^2} + 1) \\ &\leq \gamma_2(t + t^2 + t^2\sqrt{E \sup_{0\leq u\leq t} |Z(u)|^2}), \end{aligned} \tag{5.35}$$

or a suitable $\gamma_2 \in (0, \infty)$. Now, we proceed to the proof of Lemma 5.0.2.

Proof. From (5.2), we have via an application of Itô's formula

$$EZ^2(t \wedge \tau_k) = z^2 - 2\beta E \int_0^{t\wedge\tau_k} Z^2(s)ds + 2E \int_0^{t\wedge\tau_k} Z(s)dl(s) + E(t \wedge \tau_k), \tag{5.36}$$

where $\tau_k \doteq \inf\{t : |Z(s)| \geq k\}$. Note that

$$\begin{aligned}\left|\int_0^{t\wedge\tau_k} Z(s)d\tilde{l}(s)\right| &\leq \left(\sup_{0\leq s\leq t}|Z(s)|\right)l(t)\\ &= (\sup_{o\leq s\leq t}|Z(s)|)[\tilde{l}(t)+\beta_1 l^*(t)], \qquad (5.37)\end{aligned}$$

where the last equality follows from (5.21). Recalling the definition of $l^*(s)$, we have

$$E[(\sup_{0\leq s\leq t} Z(s))l^*(s)] \leq tE\sup_{0\leq u\leq t}|Z(u)|^2. \qquad (5.38)$$

Also, using (5.35), we have

$$\begin{aligned}E\sup_{0\leq u\leq t}|Z(u)|\tilde{l}(t) &\leq \sqrt{E(\sup_{0\leq u\leq t}|Z(u)|^2)}\sqrt{E\tilde{l}^2(t)}\\ &\leq c_1\sqrt{E(\sup_{0\leq u\leq t}|Z(u)|^2)}\left(\sqrt{t}+t+t\left(E(\sup_{0\leq u\leq t}|Z(u)|^2)\right)^{\frac{1}{4}}\right)\\ &= c_1(\sqrt{t}(E(\sup_{0\leq u\leq t}|Z(u)|^2))^{\frac{1}{2}}+t(E(\sup_{0\leq u\leq t}|Z(u)|^2))^{\frac{1}{2}}\\ &+ t(E(\sup_{0\leq u\leq t}|Z(u)|^2))^{\frac{3}{4}}). \qquad (5.39)\end{aligned}$$

Now fix $\Delta > 0$, and let $\alpha_\Delta \doteq E\sup_{0\leq u\leq\Delta}|Z(u)|^2$. From (5.36), (5.37), (5.38), and (5.39), we then have that

$$\begin{aligned}EZ^2(\Delta\wedge\hat{\tau}_k) - z^2 &\leq -2\beta\Delta E\inf_{0\leq u\leq\Delta}|Z(u)|^2 + 2\beta_1\Delta\alpha_\Delta\\ &+2c_2\sqrt{\Delta}\sqrt{\alpha_\Delta}(1+\sqrt{\Delta}+\sqrt{\Delta}\,(\alpha_\Delta)^{\frac{1}{4}}) + \Delta. \qquad (5.40)\end{aligned}$$

Define $\tilde{Z}(s) \doteq Z(s) - z$. Then applying Itô's formula we have

$$\tilde{Z}^2(t) = -2\beta \int_0^t \tilde{Z}(s)Z(s)ds + 2\int_0^t \tilde{Z}(s)d\tilde{W}(s) + 2\int_0^t \tilde{Z}(s)dl(s) + t.$$

Thus,

$$\begin{aligned}(1-2\beta\Delta)\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2 \;\; &\le \;\; 2\sup_{0\le s\le\Delta}|\int_0^s \tilde{Z}(u)d\tilde{W}(u)| \\ &\quad +2\sup_{0\le s\le\Delta}|\int_0^s \tilde{Z}(u)dl(u)| + \Delta + 2\Delta\beta|z|\left(\sup_{0\le u\le\Delta}|\tilde{Z}(u)|\right).\end{aligned}$$

Taking expectations and using the above bound above, we have

$$\begin{aligned}(1-2\beta\Delta)E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2 \;\; &\le \;\; 4\sqrt{\Delta}\sqrt{E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2} \\ &+ \;\; 2E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|l(\Delta) \\ &+ \;\; \Delta + 2\Delta\beta|z|E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|. \qquad (5.41)\end{aligned}$$

Furthermore, from (5.21) we have

$$\begin{aligned}E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|l(\Delta) \;\; &\le \;\; E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|\tilde{l}(\Delta) \\ &+ \;\; \beta_1\Delta E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2 + \beta_1\Delta|z|E\sup_{0\le s\le\Delta}|\tilde{Z}(u)|. \qquad (5.42)\end{aligned}$$

Next, using (5.35) we get

$$\begin{aligned} E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|\tilde{l}(\Delta) &\le (E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2)^{\frac{1}{2}}(E\tilde{l}^2(\Delta))^{\frac{1}{2}} \\ &\le c_3(E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2)^{\frac{1}{2}}(\Delta+\Delta^2+\Delta^2\sqrt{E\sup_{0\le u\le\Delta}|Z(u)|^2})^{\frac{1}{2}} \\ &\le c_4(E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2)^{\frac{1}{2}}(\Delta+\Delta^2+\Delta^2\sqrt{E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2}+|z|\Delta^2)^{\frac{1}{2}} \\ &\le c_5(E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2)^{\frac{1}{2}}(\Delta^{\frac{1}{2}}+\Delta+\Delta(E\sup_{0\le u\le\Delta}|\tilde{Z}(u)|^2)^{\frac{1}{4}}+\sqrt{|z|}\Delta). \end{aligned} \tag{5.43}$$

Letting $\Theta = E\sup_{0\le s\le\Delta}|\tilde{Z}(u)|^2$, we have from (5.41), (5.42), and (5.43) that

$$\begin{aligned} (1-(2\beta+2\beta_1)\Delta)\Theta &\le 2\sqrt{\Delta}\sqrt{\Theta}+(2\beta+2\beta_1)\Delta|z|\sqrt{\Theta} \\ &\quad +2c_5\sqrt{\Theta}(\sqrt{\Delta}+\Delta+\Delta\Theta^{\frac{1}{4}}+\Delta\sqrt{|z|})+\Delta. \end{aligned}$$

Dividing by $\sqrt{\Theta}$,

$$\begin{aligned} (1-(2\beta+2\beta_1)\Delta)\sqrt{\Theta} &\le 2\sqrt{\Delta}+(2\beta+2\beta_1)\Delta|z| \\ &\quad +2c_5(\sqrt{\Delta}+\Delta+\Delta\Theta^{\frac{1}{4}}+\Delta\sqrt{|z|})+\frac{\Delta}{\sqrt{\Theta}} \\ &\le 2\sqrt{\Delta}+2c_6(\sqrt{\Delta}+\Delta+\Delta(1+\Theta^{\frac{1}{2}}) \\ &\quad +\Delta(1+|z|))+\frac{\Delta}{\sqrt{\Theta}} \\ &\le 2\sqrt{\Delta}+\beta\Delta|z|+2c_6\sqrt{\Delta}+2c_6\Delta \\ &\quad +2c_6\Delta+2c_6\Delta\Theta^{\frac{1}{2}}+2c_6\Delta+2c_6\Delta|z|+\frac{\Delta}{\sqrt{\Theta}}. \end{aligned}$$

Subtracting and absorbing constants, we have for $\Delta < 1$,

$$(1-(2\beta+2\beta_1-2C)\Delta)\sqrt{\Theta} \le c_7\sqrt{\Delta}+c_7\Delta|z|+\frac{\Delta}{\sqrt{\Theta}}.$$

Now choosing δ_0 sufficiently small, we obtain that for all $\Delta < \delta_0$,

$$\sqrt{\Theta} \le c_8\sqrt{\Delta}(1+\sqrt{\Delta}|z|)+\frac{\Delta}{\sqrt{\Theta}}.$$

Thus,

$$\Theta \le c_8\sqrt{\Delta}(1+\sqrt{\Delta}|z|)\sqrt{\Theta}+\Delta.$$

Using the quadratic formula, we obtain

$$\Theta \le C\Delta(1+\Delta z^2), \tag{5.44}$$

for a suitable constant C. This proves (5.5) of Lemma 5.0.2. Next, noting that $|Z(u)| = |\tilde{Z}(u)+z|$, we have

$$Z^2(u) = \tilde{Z}^2(u)+z^2+2z\tilde{Z}(u).$$

Thus,

$$E\inf_{0\le s\le\Delta}|Z(u)|^2 \ \ge\ z^2-2|z|E\sup_{0\le s\le\Delta}|\tilde{Z}(u)| \tag{5.45}$$

$$\ge\ z^2-2|z|\sqrt{\Theta}. \tag{5.46}$$

Combining (5.40), (5.44), and (5.46)

$$\begin{aligned} EZ^2(\Delta \wedge \tau_k) - z^2 &\leq -2\beta\Delta(z^2 - 2|z|\sqrt{\Theta}) + 2\beta_1\Delta(z^2 + \Theta + 2|z|\sqrt{\Theta}) \\ &+2c_2\sqrt{\Delta}\sqrt{2z^2 + 2\theta}(1 + \sqrt{\Delta} + \sqrt{\Delta}\left(2z^2 + 2\Theta\right)^{\frac{1}{4}}) + \Delta. \end{aligned}$$

Recalling the definition of Θ we have

$$\begin{aligned} EZ^2(\Delta \wedge \tau_k) - z^2 &\leq -2\beta_2\Delta z^2 + c_9|z|\Delta^{\frac{3}{2}}(1 + \sqrt{\Delta}|z|) + c_9\Delta(1 + \Delta z^2) \\ &+c_9\sqrt{\Delta}(|z| + c_9\sqrt{\Delta}(1 + \sqrt{\Delta}|z|) + c_9\Delta(|z| + \sqrt{\Delta}(1 + \sqrt{\Delta}|z|)) \\ &+c_9\Delta(|z|^{\frac{3}{2}} + \Delta(1 + \Delta z^2) + 1). \end{aligned}$$

Taking $k \to \infty$ we then have

$$\begin{aligned} EZ^2(\Delta) - z^2 &\leq -2\beta_2\Delta z^2 + c_{10}(z^2 + 1)\Delta^2 \\ &+ c_{10}(|z|^{\frac{3}{2}} + |z| + 1)\Delta^{\frac{3}{2}} + c_{10}(|z| + 1)\Delta + c_{10}|z|\Delta^3 + c_{10}|z|\Delta^{\frac{1}{2}}. \end{aligned}$$

The result now follows, since for sufficiently large z and sufficiently small Δ the first term on the right side of the above expression will dominate. ■

Chapter 6

Numerical Methods for Computing Asymptotic Parameters

In the Chapters 4, using renewal theory we proved the existence of asymptotic velocity of the motor and by a functional central limit theorem identified its effective diffusivity. We also gave explicit representation of these asymptotic quantities in terms of moments of hitting times for certain reflected diffusions. For the case where the drift and diffusion coefficients are constants, on can give exact formulas in terms of model parameters for the moments and the asymptotic quantities of interest can be determined exactly. Moreover in general, explicit calculations are not possible and thus one has to result to numerical methods.

As noted above, the key quantities of interest are expressed in terms of the moments of hitting times of certain diffusions. In this chapter we will present some numerical methods for the computation of these moments. The first is a straightforward path simulation based method. Namely, we simulate a large number of trajectories of the process and determine the corresponding hitting times. The moments of hitting times can then be approximated by the corresponding empirical moments. The second technique is to devise a suitable Markov chain approximation for the reflected diffusion process and then use the transition matrix of the chain to approximate the probability distribution and, hence the moments, of the hitting time. Finally, the third method is a linear programming method introduced in [10] which is based on the martingale characterization of a Markov process via its generator.

In section 6.2.1 of this chapter we will consider the computation of the asymptotic parameters in the case when the motor is pulling a cargo. We will study a numerical scheme that has been recently introduced for the one dimensional cargoless setting in [23]. We will extend this method to the more complex setting of the coupled motor/cargo system and present numerical results for the computation of asymptotic velocity and effective diffusivity of the motor. The work in this last section is joint with Prof. Tim Elston.

6.1 The One-dimensional Case: Motor Without a Cargo

It was seen in Theorem 4.1.2 that if the drift and diffusion coefficients are constant, one can do various moment calculations for hitting times explicitly. However, in general exact calculations are not possible, and thus one needs to resort to numerical approximations. There are three principle numerical methods we will explore for use with the diffusion ratchet. Each method will be briefly discussed, then we will move on to detailed discussions and computational results for each. There will be special attention paid to calculations concerning hitting times of reflected diffusion processes. This is important because of the relationship, proved in Theorem 4.1.1, between hitting times and asymptotic velocity.

6.1.1 Path Simulation.

The Euler method or some higher order method for the numerical solution of diffusion processes can be adapted to simulate approximate paths of reflected diffusion processes. This method is simple to program and is easily adapted to obtain results for hitting times and other functionals of the process. However, it is computationally intensive, since many simulations of the path must be performed and averaged over to get relatively stable results. Moreover, the results would differ if the numerical procedure is repeated (with a new random seed). Error estimates are difficult to obtain.

6.1.2 Markov Chain Approximation.

It is well known that a reflected diffusion process can be approximated in law by a suitably scaled discrete time- discrete state Markov chain [15]. This means that the expected value of a continuous (in $D([0,\infty) : \mathbb{R}_+)$) functional of the linearly interpolated discrete time Markov chain will converge to the expected value of the corresponding functional of the limit diffusion process as the parameters in the time and state discretization approach their limits. So, in practice one can do calculations on an appropriately chosen Markov chain (discrete in time and state) to obtain numerical approximations for the expected value of functionals of the limit diffusion process. In particular, this method can be used for numerical calculations of the expected hitting times for reflected diffusion processes. One drawback is that in this method one needs to calculate the whole probability distribution of the Markov chain, at each time instant, even though one is interested in the expected value of a very specific functional of the Markov chain. In practice, this means that the numerical procedure requires repeated multiplication of large matrices. On the plus side, this is a method well-suited for a bounded state space, which is the situation that we have in the computation of the expected hitting time. Another benefit is that it is often possible to obtain a rigorous evaluation of the procedure such as rates of convergence and error bounds. Furthermore, since one computes exactly the expected values for the approximating Markov chain, unlike path simulation, there is no random number generation and thus repeated calculations will yield identical results.

6.1.3 Linear Programming.

The third method is a rather novel approach to the calculation of moments of hitting times for Markov processes[10]. In this approach one considers the occupation measure (until the hitting time) of the Markov process and formulates various linear conditions that are satisfied by its moments as constraints in a linear program. The conditions satisfied by the moments are obtained using the martingale characterization of a Markov process via its generater. Furthermore, one can introduce the Hausdorff moment conditions, which are satisfied by moments of any probability measure, as additional constraints in the linear program. The key observation is that, under suitable conditions, the moments of the hitting time can be expressed in terms of the moments of the occupation measure. Thus, in particular, one can use the expected hitting time as the objective function and a subset of the linear moment conditions as constraints to formulate a linear programming problem. By maximizing and minimizing the objective function, under the constraints, one obtains upper and lower bounds for the expected hitting times. If one uses only a few moment conditions as constraints, the interval determined by the lower and upper bound is rather large and thus not very informative about the true value of the expected hitting time. However, in many applications, the interval is seen to shrink dramatically as additional moment conditions are introduced as constraints in the linear program.

We now discuss each of the methods in a greater detail.

1. The Path Simulation technique is appealing on intuitive grounds. We want to simulate

$$X(t) = \Gamma\left(x_0 + \int_0^{\cdot} b(X(s))ds + \int_0^{\cdot} a(X(s))dW_s\right)(t).x \tag{6.1}$$

We will use an Euler type method where we simply start at x_0, choose a small $\delta > 0$ and simulate an approximation for the value of X at δ. Then using this estimate for $X(\delta)$, we approximate the value of X at 2δ and so on. We will denote the simulated sequence by $\{Z_i\}_{i\geq 1}$, where Z_i is the simulated approximation of $X(i\delta)$. The equations used to generate $\{Z_i\}_{i\geq 1}$ are as follows.

$$\begin{aligned} Z_0 &= x_0 \\ Z_{i+1} &= (Z_i + b(Z_i)\delta + a(Z_i)N(0,\delta))^+ \end{aligned}$$

where $N(0,\delta)$ is a simulated normal random variable, independent of previous variables, with mean 0 and variance δ. We will simulate until $Z_i \geq L$. This yields a simulated approximate value for τ as $\delta \times i$, where i is the time step at which Z exits $[0, L)$. Path simulation can be improved by adapting many of the other classical methods for solving ODEs. Clearly, one must perform many calculations to get a single simulation from the approximate hitting time distribution. By repeating these steps a large number of times one can obtain an approximation for the mements of τ.

2. In the Markov Chain approximation method, one first writes down a discrete time-discrete state Markov chain which well approximates the probability law of $X(\cdot \wedge \tau)$, where $\tau \doteq \inf\{t : X(t) = L\}$ and $X(\cdot)$ is given by (6.1). This approximating chain, denoted by $\{Z_i^n\}_{i\geq 1}$, where n is the discretization parameter, will have the state space: $\{\frac{j}{n} : j \in \mathbb{N}_0, j \leq nL\}$. The transition probabilities of the approximating chain are given as follows: For $z = \frac{j}{n}$, $j \in \{1, 2, \cdots, L(n-1)\}$,

$$\begin{aligned} p_n(z, z + \frac{1}{n}) &= \frac{\frac{a^2(z)}{2} + \frac{b^+(z)}{n}}{S} \\ p_n(z, z - \frac{1}{n}) &= \frac{\frac{a^2(z)}{2} + \frac{b^-(z)}{n}}{S} \\ p_n(z, z) &= 1 - p_n(z, z + \frac{1}{n}) - p_n(z, z - \frac{1}{n}), \end{aligned} \tag{6.2}$$

where $S \doteq \sup_{x\in[0,L)}(\frac{a^2(x)}{2}+|b(x)|)$. For the lowest state the transition probabilities are given as

$$\begin{aligned} p_n(0, 0+\frac{1}{n}) &= \frac{\frac{a^2(z)}{2}+\frac{b^+(z)}{n}}{S} \\ p_n(0,0) &= 1-p_n(0,0+\frac{1}{n}) \end{aligned} \tag{6.3}$$

The highest state is absorbing and so returns to itself with probability one, namely

$$p_n(L,L) = 1. \tag{6.4}$$

Adapting techniques from chapter 5 of [15], one can show that if one linearly interpolates the above Markov chain with time intervals of length $\Delta t_n \doteq \frac{1}{n^2 S}$ then the resulting continuous time process converges weakly to $X(\cdot \wedge \tau)$. From this it follows that if $\tau_n \doteq \inf\{j : Z_j^n = L\}$ then

$$\frac{1}{n^2 S}E(\tau_n) \to E(\tau).$$

Thus one can approximate $E(\tau)$ by computing $E(\tau_n)$. Next note that

$$\begin{aligned} E(\tau_n) &= \sum_{i=1}^{\infty} P(\tau_n > i) \\ &= \sum_{i=1}^{\infty}(1-P(\tau_n \le i)) \\ &= \sum_{i=1}^{\infty}(1-P(Z_i^n = L)), \end{aligned}$$

where the last step follows on recalling that L is an absorbing state for the Markov chain. Thus one can find $E(\tau_n)$ by computing $P(Z_i^n = L)$ for $i \ge 1$. In practice, of course, one will truncate the above infinite sum to finitely many terms. Since Z_i^n is a finite state Markov chain, one can compute $P(Z_i^n = L)$ by multiplying the probability law vector of Z_{i-1}^n by the transition probability matrix defined by (6.2), (6.3), (6.4). By suitable modifications, this method can be used to approximate higher moments of τ as well.

3. Finally, we describe the linear programming method. We begin by observing that (6.1) can be rewritten as

$$X(t) = x_0 + \int_0^t b(X(s))ds + \int_0^t a(X(s))dW_s + \ell(t), \tag{6.5}$$

where $\ell(t)$ is an increasing process satisfying

$$\ell(t) = \int_0^t 1_{\{X(s)=0\}} d\ell(s).$$

Let τ be as before, i.e., $\inf\{t : X(t) = L\}$. Let $f \in C^2(I\!R_+)$. Then an application of Itô's formula gives

$$f(X(t \wedge \tau)) = f(x_0) + \int_0^{t\wedge\tau} Af(X(s))ds + \int_0^{t\wedge\tau} f'(Z(s))dW(s) + f'(0)\ell(t \wedge \tau), \tag{6.6}$$

where A is the infinitesimal generator of the diffusion given as follows.

$$Af(x) \doteq \frac{1}{2}a^2(x)\frac{\partial^2 f}{\partial x^2} + b(x)\frac{\partial f}{\partial x}.$$

Taking expected value in (6.6) we have that

$$Ef(X(t \wedge \tau)) = f(x_0) + E \int_0^{t\wedge\tau} Af(X(s))ds + f'(0)E\ell(t \wedge \tau). \tag{6.7}$$

Recalling that $E(\tau) < \infty$, we have taking limit as $t \to \infty$ that

$$f(L) = Ef(X(\tau)) = f(x_0) + E \int_0^{\tau} Af(X(s))ds + f'(0)E\ell(\tau). \tag{6.8}$$

Now define a finite measure μ_0 on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ as follows.

$$\mu_0(B) \doteq E\left[\int_0^{\tau} I_B(X(s))ds \right], \quad B \in \mathcal{B}(\mathbb{R}_+).$$

Note that $\mu_0(B)$ is the expected amount of time the process spends in set B before the hitting time τ, and $\int_{E_0} \mu_0(dy) = E\tau$. In terms of μ_0, (6.8) can be rewritten as

$$\int_{[0,L]} Af(x)\mu_0(dx) + f(x) - f(L) + f'(0)\vartheta, \tag{6.9}$$

where $\vartheta \doteq E\ell(\tau)$. Suppose now that the coefficients $a(x)$ and $b(x)$ are polynomials of the form $\sum_{j=0}^{p} a_j x^j$, $\sum_{j=0}^{p} b_j x^j$ for $x \in [0, L]$. In practice, one can always approximate general coefficients a and b by such monomials. If we take f in (6.9) to be a monomial then the above equation can be rewritten in the form

$$\sum_{i=0}^{n} c_i m_i + \vartheta\ d + \kappa = 0 \qquad \textbf{(M)}$$

where

$$m_k = \int x^k \mu_0(dx)$$

for $k = 0, 1, 2, ...$ and n, c_i, d, κ are constants which depend on f and can be written explicitly. Note that $m_0 = E\tau$.

Equation (**M**) gives a set of linear conditions which the parameters $\{\{m_i\}_{i\geq 1}, \vartheta\}$ must satisfy. One can obtain additional constraints on these parameters on considering the Hausdorff moment conditions for the sequence $\{m_i\}$ which say that

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} \frac{m_{j+k}}{L^{j+k}} \geq 0. \qquad \textbf{(H)}$$

One can now formulate the following linear programming problems.

$$\begin{cases} \textbf{maximize} \quad m_0 \\ \\ \textbf{subject to:} \quad \textbf{(M)} \text{ and } \textbf{(H)}. \end{cases} \tag{6.10}$$

$$\begin{cases} \textbf{minimize} \quad m_0 \\ \\ \textbf{subject to:} \quad \textbf{(M)} \text{ and } \textbf{(H)}. \end{cases} \tag{6.11}$$

Note that **(M)** and **(H)** contain infinitely many linear constraints and in practice one takes a finite sub-collection (say, of cardinality k) of these constraints. The value of the two linear programs, denoted by $\overline{m}_k$ and $\underline{m}_k$, respectively, must satisfy the condition

$$\overline{m}_k \;\; \geq \;\; E(\tau) \;\; \geq \;\; \underline{m}_k.$$

As k becomes larger and larger, these upper and lower bounds for $E(\tau)$ become tighter and tighter yielding good approximations for $E(\tau)$.

The following are some sample calculation using the above three methods. They are the estimated expected value for the hitting time for $L = 2$ and a variety of drift and diffusion coefficents. These are compared with the true expected value obtained from Theorem 4.1.2.

sigma	mean	true value	simulation	Markov	lp lower	lp upper
1	0	4	4.4302	3.9596	4	4
1	1	1.50916	1.5984	1.4899	1.50916	1.50916
1	2	0.875042	0.9104	0.8647	0.874934	0.875115
1	3	0.611111	0.6404	0.6041	0.610081	0.612025
1	-1	24.7991	29.7683	22.749	24.799	24.7992
1	-2	371.495	-	-	369.484	374.446
1	-3	9041.21	-	-	7231.72	19592.1
1	6	0.319444	29.7683	0.3158	0.305782	0.332666
2	1	0.735759	0.8270	0.7267	0.735759	0.735759
2	-1	1.43656	1.7643	1.4091	1.43656	1.43656

There are a few important things to notice. First, the linear programming approach gives very good results for most of the parameter values. Though performance depreciates for large negative drift with the number of constraints held constant. This can be remedied by increasing the number of constraints. For instance in the case of sigma equal to 1 and drift of -3, the interval changes from (7231.72, 19592.1) to (9037.2, 9058.67) when the number of constraints is increased from 194(the number used for the other cases) to 320. However, this increase rather dramatically increases the computation time. For a diffusion parameter of 1 and drift of -2 and -3, the simulation and Markov chain approaches become unwieldy. The first method is poor because of the length of time on average it takes to reach a level of 2 for these negative drifts. The calculations were done by attempting to take 5,000 paths with a time discretization of 0.005. Even attempting to simulate 500 paths at a discretization of 0.5 took several hours using matlab on a Sun Ultra 10 workstation for the case of sigma equal to 1 and mean of -2. The resulting approximation was more than an order of magnitude away from the true value. For the Markov chain approach one needs to multiply the transition matrix until the chain is mostly absorbed at the level L. The following calculations were done with a space discretization of 21 states (states a distance of 0.2 apart from 0 to 2). Again, this takes a tremendous number of calculations when the drift is negative. For positive drift, 100,000 transitions were sufficient for the chain to be absorbed. For negative drift, a half million transitions were necessary to be "mostly absorbed", i.e. the chain is in the absorption state with a probability greater than 0.95.

6.2 Motor and Cargo

In the previous section, we introduced a number of methods to calculate moments of the hitting times for a model of the motor alone. From results of Chapter 4, it follows that these moment calculations are sufficient for calculating both asymptotic velocity and the effective diffusivity. However, as seen Chapter 5, for the coupled motor/cargo system one dose not have the simplifying renewal structure that was critical in obtaining the results of Chapter 4. However from Chapter 5, we know that the Markov Process $\Pi_t = (\lfloor \frac{X(t)}{L} \rfloor L, Z(t))$ is ergodic. In this section we will exploit the ergodicity of Π_t to design a numerical scheme for the computation of asymptotic velocity and effective diffusivity of the coupled system. Such a scheme for the one-dimensional cargoless situation was introduced by Wang, Elston, and Peskin [23].

6.2.1 The Wang, Elston, and Peskin Method

In a recent paper [23], Wang, Elston, and Peskin introduced a numerical method to calculate asymptotic velocity and effective diffusivity for biomolecular motors which are not pulling a cargo. The method critically exploits the periodicity in the dynamics of the motor. The first step in this method is to approximate the underlying diffusion by a suitable pure jump process. Let $\Delta > 0$ be the space discretization parameter. We assume that $\Delta = \frac{L}{n}$ for some $n \in \mathbb{N}$. We will denote the approximating jump Markov process by X_n. The state space of $\{X_n\}$ is given as

$$\{x_i^j = (i-1)\Delta + jL, 1 \leq i \leq N, j \in \boldsymbol{Z}\},$$

where $N \doteq \frac{L}{\Delta}$. Note that one can write $X_n(t)$ as $\Delta U_n(t) + LV_n(t)$, where

$$V_n(t) = \lfloor \frac{X_n(t)}{L} \rfloor$$

and

$$U_n(t) = \frac{1}{\Delta}(X_n(t) - LV_n(t)).$$

Also, observe that $(U_n(t), V_n(t))$ is a Markov process with state space $S_1 \times I\!N_0$, where

$$S_1 = \{0, 1, 2, ..., N-1\}.$$

Let $p_i(j,t) \doteq P[U_n(t) = i-1, V_n(t) = j]$ for $i = 1, ..., N, j \in I\!N_0$. By convention, $p_i(-1,t) \equiv 0, i = 1, ..., N$. Then, one has the following Kolmogorov equations for $(p_i(j,t))_{i,j}$:

$$\frac{d\vec{p}(j,t)}{dt} = R\vec{p}(j,t) + R_+\vec{p}(j-1,t) \tag{6.12}$$

where

$$\vec{p}(j,t) = \begin{pmatrix} p_1(j,t) \\ p_2(j,t) \\ . \\ . \\ . \\ p_N(j,t) \end{pmatrix}. \tag{6.13}$$

We take $p_0(0,0) = 1$. If we denote the forward jump rate from state i to $i+1$ as F_i and the backward jump rate from state i to $i-1$ as B_i, then we may describe the matrices R and R_+ as follows. R is a tridiagonal matrix with

$$R_{i,i} = -(F_i + B_i) \ ; \ R_{i-1,i} = B_i \ ; \ R_{i+1,i} = F_i \tag{6.14}$$

for $i = 2, 3, ..., N-1$, and

$$R_{1,1} = -F_1 \; ; \; R_{2,1} = F_1 \tag{6.15}$$

$$R_{N,N} = -(F_N + B_N) \; ; \; R_{N-1,N} = B_{N-1}. \tag{6.16}$$

R_+ is a matrix of all zeros except that $(R_+)_{1,N} = F_N$. Note that the transitions have been decomposed into those that remain in a particular period (corresponding to R), and those that will move the motor forward one period (corresponding to R_+).

In [23], equations (6.13) are used to explicitly derive formulas for the asymptotic velocity and the effective diffusivity. The formula for asymptotic velocity is

$$v = L \sum_{i=1}^{N} [R_+ p_s]_i. \tag{6.17}$$

where p_s is the solution to the equation

$$(R + R_+) p_s = 0$$

subject to the constraint $\sum_{i=1}^{N} (p_s)_i = 1$. The vector p_s is the stationary distribution of the Markov chain $\{U_n(t)\}$. Thus, the formula can be interpreted as the proportion of time in the "top" state of the period times the rate of going forward to the next period. The formula for effective diffusivity is

$$d = \frac{L^2}{2} \sum_{i=1}^{N} [R_+ p_s + 2R_+ r]_i \tag{6.18}$$

where r is the solution to

$$(R_+ + R)r = \left(\sum_{i=1}^{N}[R_+p_s]_i - (R_+)\right)p_s, \tag{6.19}$$

subject to the constraint $\sum_{i=1}^{N} r_i = 0$.

We now present the extension of this method to the two-dimensional coupled motor/cargo case. The crucial step in the one dimensional version of this method was to split the dynamics into that of $U(t)$ and $V(t)$ i.e., the steps within the period and steps between the periods. We will do something similar for the two dimensional model. We will begin as before by approximating $(X(t), Y(t))$ by a suitable jump Markov process, denoted by $(X_n(t), Y_n(t))$. $X_n(\cdot)$ will have, as before, state space

$$\{x_i^j = (i-1)\Delta + jL, 1 \leq i \leq N, j \in I\!N\}$$

For $Y_n(\cdot)$, we will use spatial discretization of the form $\{k\Delta | k \in \boldsymbol{Z}\}$. Let $Z'_n(t) \doteq \frac{1}{\Delta}\left(Y_n(t) - \lfloor\frac{X_n(t)}{L}\rfloor\right)$, i.e. $Z'_n(t)$ represents the distance from the cargo to the last barrier crossed. Also, let

$$\begin{aligned} V_n(t) &= \lfloor\frac{X_n(t)}{L}\rfloor \\ U_n(t) &= \frac{1}{\Delta}\left(X_n(t) - \lfloor\frac{X_n(t)}{L}\rfloor\right). \end{aligned}$$

Then, $(U_n(t), V_n(t), Z'_n(t))$ is a three dimensional Markov process in the state space $S_1 \times I\!N_0 \times \boldsymbol{Z}$. Next, note that even though strictly speaking Z'_n is supported on an unbounded set, one expects that due to the elastic linkage between the motor and cargo

$Z'_n(t)$ will, with large probability, stay in a bounded set. Thus for numerical purposes, we will assume that $Z'_n(t)$ takes values in $S_2 \doteq \{-k_1N+1, -k_1N+2, ..., k_2N\}$

Let $M = |S_2|, i.e. M = k_1N + k_2N$. Denote for $k \in S_1$, $j \in I\!N_0$, and $l \in \{1, ..., M\}$. Also, denote $P[U_n(t) = j, V_n(t) = k, Z'_n(t) = l - k_1N]$ by $p_{k,l}(j,t)$. Letting

$$\vec{p}(j,t) = \begin{pmatrix} p_{1,1}(j,t) \\ p_{2,1}(j,t) \\ . \\ . \\ . \\ p_{N,1}(j,t) \\ . \\ . \\ . \\ p_{1,M}(j,t) \\ p_{2,M}(j,t) \\ . \\ . \\ . \\ p_{N,M}(j,t) \end{pmatrix}, \quad (6.20)$$

we have the following Kolmogorov forward equations for $\vec{p}(j,t)$

$$\frac{d\vec{p}(j,t)}{dt} = R\vec{p}(j,t) + R_+\vec{p}(j-1,t). \quad (6.21)$$

where the matrix R is now given as follows The central difference in the two dimensional case is that when a transition from one period to the next occurs. The new R matrix is as follows:

$$\begin{matrix} R(1) & \hat{B}(2) & 0 & 0 & 0 & 0 \\ \hat{F}(1) & R(2) & \hat{B}(3) & 0 & .... & .... \\ 0 & \hat{F}(2) & R(3) & \hat{B}(4) & .... & 0 \\ ... & ... & ... & .... & .... & \hat{B}(M) \\ 0 & .... & ... & 0 & \hat{F}(M-1) & R(M) \end{matrix}$$

In the above, $R(l)$ is a matrix of a similar form to R as in the one dimensional case, but where the forward and backward transitions for the motor depend on the state of the cargo, l. More precisely $R(l)$ is and $N \times N$ matrix with

$$R_{i,i}(l) = -(F_i(l) + B_i(l)) \; ; \; R_{i,i-1}(l) = B_i(l) \; ; \; R_{i,i+1}(l) = F_i(l) \tag{6.22}$$

for $i = 2, 3, ..., N-1$, and

$$R_{1,1}(l) \quad = \quad -F_1(l) \; ; \; R_{2,1} = F_1(l) \tag{6.23}$$

$$R_{N,N}(l) \quad = \quad -(F_N(l) + B_N(l)) \; ; \; R_{N-1,N}(l) = B_{N-1}(l). \tag{6.24}$$

In addition, the matrix $\hat{F}(l)$ is a $N \times N$ matrix with $(\hat{F}(l))_{i,j} = 0$ for $i \neq j$ and $(\hat{F}(l))_{i,j} \doteq F_{li}$, where F_{li} is the rate at which $Z'{}_n(t)$ jumps to the right given $(U_n(t-), V_n(t-), Z'_n(t-)) = (i, k, l)$. $\hat{B}(l)$ are defined in a similar way.

The matrix R_+ will be of the form

$$\begin{array}{ccccccc}
0 & ... & ... & 0 & 0 & .. & 0 \\
... & ... & ... & ... & ... & ... & ... \\
R_+(1) & ... & 0 & 0 & & ... & 0 \\
... & ... & ... & ... & ... & ... & ... \\
0 & ... & R_+(M-N-2) & 0 & 0 & ... & 0 \\
0 & ... & 0 & R_+(M-N-1) & 0 & ... & 0 \\
0 & ... & 0 & 0 & R_+(M-N) & ... & R_+(M)
\end{array}$$

where $R_+(l)$ is similar in form to R_+ but where the forward rates are determined by the state of the cargo, l. More precisely, $R_+(l)$ is a matrix of all zeros except that $(R_+(l))_{1,N} = F_{lN}$, where F_{lN} is defined as above. To understand the unusual structure of R_+ recall that $Z'_n(t)$ which is the position of the cargo minus the last barrier site that is crossed. Recall also that the sub-matrix, $R_+(l)$ represent the transitions of the motor from one period into the next . Thus, when the process corresponding to the motor moves up one period, the process, $Z'_n(t)$, needs to be instantly shifted by one period.

Using the above forward equations, we can derive the following expressions for asymptotic velocity and effective diffusivity:

$$v = L \sum_{l=1}^{M} \sum_{i=1}^{N} [R_+ p_s]_{i,l}, \tag{6.25}$$

where p_s is the solution of

$$(R + R_+)p_s = 0$$

subject to the constraint $\sum_{l=1}^{M} \sum_{i=1}^{N} (p_s)_{i,l} = 1$, and

$$d = \frac{L^2}{2} \sum_{l=1}^{M} \sum_{i=1} N[R_+ p_s + 2R_+ r]_{i,l}, \tag{6.26}$$

where r is the solution to

$$(R_+ + R)r = \left(\sum_{l=1}^{M} \sum_{i=1}^{N} [R_+ p_s]_{i,l} - R_+ \right) p_s, \tag{6.27}$$

subject to the constraint $\sum_{l=1}^{M} \sum_{i=1}^{N} r_{i,l} = 0$.

The above method is very general and can be used for many models that have some sort of a periodic structure in the motor dynamics. This method was used for the model in Chapter 5 with the following parameters: $a_1 = 1$, $a_2 = 0.1$, $L = 0.5$, and $\beta_1 = \beta_2 = \beta$.

Table 6.1: Numerical Results

β	asymptotic velocity	effective diffusivity
0	1.9231	0.3210
0.5	1.0529	0.2102
0.8	0.6817	0.1464
1	0.4949	0.1099
1.5	0.2027	0.0465
2	0.0758	0.0167
2.5	0.0281	0.0057
3	0.0116	0.0023
3.5	0.0059	0.0016
5	0.0024	0.0018
7.5	0.0017	0.0020
10	0.0015	0.0021
15	0.0015	0.0022

One thing to notice is the decrease in velocity as the elasticity increases. Also, the calculations stabilize to fixed values at higher levels of β. Intuitively, the motor and cargo begin to act as one unit since they are so tightly coupled. With path simulation, it is virtually impossible to get stable asymptotic calculations for such a tightly coupled system. In addition, these calculations can be done relatively quickly since they involves only solving for a large, but reasonable linear system.

Chapter 7

Filterting, Smoothing, and Parameter Estimation

One of the main difficulties in the study of biomolecular motors is that they are too small to be observed directly. Frequently, these motors are transporting a cargo via an elastic linkage, which is significantly larger than the motor. Recent advances in molecular biological techniques have enabled scientists to observe such cargos dynamically over time. Thus, it is of great interest to investigate whether one can infer properties of the dynamics of the motors from the observed dynamics of the cargo. In practice, one would be interested in inferring about the current location of the motor from the past and current observations. One would also like to obtain parameter estimates of various coefficients in the model. In this chapter, we will use techniques from non-linear filtering and parameter estimation under partial observations to address these problems for the coupled model of motor and cargo which was introduced in chapter 4. Recall that the location process for the motor/cargo pair: $(X(t), Y(t))$ describes a two dimensional Markov process with state space $\mathbb{R}_+ \times \mathbb{R}$. In this chapter, we will take $\beta_1 = \beta_2 \doteq \beta$, and σ_x, σ_y will be denoted σ_1, σ_2, respectively. Also, recall that the stochastic integral equation for the cargo dynamics is given as

$$Y(t) = y + \int_0^t \beta(X(s) - Y(s))ds + \sigma_2 W(t) \tag{7.1}$$

The goal of this chapter is to use the observed values of $\{Y(t)\}_{t\geq 0}$ to infer about the motor location $X(t)$ and to estimate model parameters such as β. We will begin in Section 7.1 by considering the problem of computing the conditional distribution of $X(t)$ given past and current observations, i.e. $\{Y(s) : s \leq t\}$. This conditional distribution which captures all the information on $X(t)$ which can be extracted from observed data and other available statistical information is called the non-linear filter. Note that once the filter is available, one can obtain the estimate of the motor state at any given time by computing the mean of the non-linear filter.

We will consider two methods for computing the non-linear filter; the first method will make use of explicit formulas for transition probability function and numerical integration techniques. The second method is based on particle filters which are very flexible and easy to adapt for a broad family of models. We will conclude the section by making numerical comparisons of the performance of the two methods on simulated data.

In theory, one can get a better estimate for $\{X(t)\}_{0\leq t\leq T}$, if one is able to use the whole observation trajectory, $\{Y(t)\}_{0\leq t\leq T}$, rather than just the past and current observations, $\{Y(s)\}_{0\leq s\leq t}$, The conditional distribution of $X(t)$ given $\{Y(t)\}_{0\leq t\leq T}$, is referred to as the non-linear smoother. Computing a non-linear smoother is numerically much more intensive than computing the non-linear filter. Furthermore, due to the ratchet mechanism of the model, it seems likely that current observations do not carry too much information on the state location far back in time. Nevertheless, we will consider, briefly, the non-linear smoother in Section 7.2. It will be observed that non-linear smoother does not perform too much better that the non-linear filter.

Finally, in Section 7.3, we will consider the problem of parameter estimation. We will consider two methodologies, the first one is a simple Bayesian technique based on the augmentation of the state vector to include the parameter. The second method is a form of maximum likelihood estimation that has been introduced by Hurzeler and Kunsch [11]. We will conclude this section by comparing the two methods for simulated data.

7.1 Filtering

In practice, the measurements are digitized; so they are collected only at discrete time points $0, \Delta, 2\Delta, ...$ for fixed $\Delta > 0$. Thus, it is natural to consider the discretized version of the system defined in (5.1) which is described as follows:

$$
\begin{aligned}
Y_{i+1} &= Y_i + \beta(X_{i+1} - Y_{i+1})\Delta + \xi_{i+1} && (7.2) \\
X_{i+1} &= \Psi(X_i, Y_i, \eta_{i+1}) \\
&= (X_i + \beta(Y_i - X_i)\Delta + \eta_{i+1}) \vee \left(\lfloor \frac{X_i}{L} \rfloor L \right) && (7.3)
\end{aligned}
$$

where $\xi_{i+1} \sim N(0, \sigma_2^2\Delta)$ and $\eta_{i+1} \sim N(0, \sigma_1^2\Delta)$. Denote the conditional distribution of X_i given $\{Y_j, j \leq i\}$ by F_i. Then a straightforward calculation gives the following recursive formula for coupling F_i,

$$F_{i+1}(dx) = Cq(Y_{i+1}|x, Y_i) \int_{\tilde{x} \in I\!R_+} \gamma(dx|\tilde{x}, Y_i) F_i(d\tilde{x}) \tag{7.4}$$

where C is a normalizing constant, and γ is the transition cumulative distribution of the next state, X_{i+1}, given the current state of the system (X_i, Y_i), i.e.

$$\begin{aligned}
\gamma(x|z, y) = P[X_{i+1} \leq x | X_i = z, Y_i = y] &= \Phi(x) \text{ for } x \geq \lfloor \frac{z}{L} \rfloor L \\
&= 0 \text{ for } x < \lfloor \frac{z}{L} \rfloor L,
\end{aligned} \tag{7.5}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a $N(z + \beta(y - z)\Delta, \sigma_1^2\Delta)$ random variable, and $q(y|x, z)$ is the conditional density of Y_{i+1} given $X_{i+1} = x$ and $Y_i = z$ which is the density of a $N\left(\frac{z+\beta x\Delta}{1+\beta\Delta}, \frac{\sigma_2^2\Delta}{(1+\beta\Delta)^2}\right)$ random variable. Now, we will consider two methods for obtaining approximations for F_i using the recursions in (7.4).

7.1.1 Method 1: Approximating Filter via Numerical Integrations.

To approximate the integral in (7.4), we choose a discrete grid in the state space $I\!R_+$. A general problem in the numerical integration method is that it is $O(g^2N)$, where g is the size of the grid and N is the number of time points. However, using the special features of the dynamics, one can significantly reduce the computational burden. Due to the ratchet mechanism in the dynamics of $\{X_n\}$, the support of the measure, F_i, keeps shifting to the right as i increases. Indeed, given that $X_i \geq jL$, the conditional distribution of X_{i+1} has support $[jL, \infty)$. Thus, one can reduce computation by using fewer grid points but suitably adapting the grid as i changes. We denote the grid points at time instant i by $\{x_i^1, x_i^2, ..., x_i^k\}$. Denote the cumulative distribution function corresponding to the measure $F_i(dx)$ by $F(x)$, i.e. $F_i(x) = \int_0^x F_i(dy)$. Note that

$$F_{i+1}(x) = C \int_0^x \left(\int_0^\infty q(Y_{i+1}|\overline{x}, Y_i) \int_{\tilde{x} \in I\!R_+} \gamma(d\overline{x}|\tilde{x}, Y_i) \right) F_i(d\tilde{x}).$$

One has the following recursive way for approximately computing $\Theta_i \doteq \{F_i(x_i^j)\}_{j=1}^k$. Having computed Θ_m for $m = 0, 1, ..., i-1$, we obtain Θ_i via the formula

$$F_{i+1}(x_{i+1}^k) \approx C \sum_{k'<k} \left(\sum_{j=1}^{k-1} q(Y_{i+1}|x_{i+1}^{k'}, Y_i)[\gamma(x_{i+1}^{k'+1}, x_i^j, y_i) - \gamma(x_{i+1}^{k'}, x_i^j, y_i)][F_i(x_i^{j+1}) - F_i(x_i^j)] \right). \tag{7.6}$$

In practice, we use the same grid for a fixed number of time steps following which the grid is shifted to the right. The key idea is that once it has become very likely that the ratchet process has passed a given barrier site, there is no longer any need to calculate on that portion of the grid. Our rule of thumb is to fix a small $\epsilon \in (0, 1)$ and place the grid Θ_{i+1} on $[mL, \infty)$ for the computation of F_{i+1} where m is such that $F_i(mL) < \epsilon$. Performence of this algorithm is discussed in Section 7.1.3.

7.1.2 Method 2: Particle Filter.

Particle filters are a very versatile and adaptive set of ideas for computing the nonlinear filter. The basic idea in obtaining these filters is to replace the expected values and integrals such as in (7.4) by Monte Carlo sample averages. We describe now one of the basic forms of particle filters which will be used in our application. Fix the size of the Monte Carlo samples to be p. Denote by F^p_{n-1} the approximation of F_{n-1} based on the Monte Carlo sampling. Then, F^p_{n-1} is described via points $X^{(j)}_{n-1}; j = 1, ..., p$, and their respective weights, $w^{(j)}_{n-1}; j = 1, ..., p$. The weights are non-negative and sum to one, and F^p_{n-1} is given in the following expression:

$$F^p_{n-1}(dx) \doteq \sum_{j=1}^{p} w^{(j)}_{n-1} \delta_{X^{(j)}_{n-1}}(dx).$$

Having computed the approximate filter F^p_i at time points $i = 1, ..., n-1$, we obtain F^p_n as follows:

1. Propagate the particles using the evolution equation (7.3), i.e. obtain p zero mean, normal random numbers with variance $sigma^2_1\Delta$, and denote them by $\{\eta^{(j)}_n\}^p_{j=1}$. Let

$$X^{(j)}_n \doteq \psi(X^{(j)}_{n-1}, Y_{n-1}, \eta^{(j)}_n).$$

2. Readjust the weights of the particle to account for the new information obtained via Y_n as follows:

$$w^{(j)}_n = C w^{(j)}_{n-1} q(Y_n | X^{(j)}_n, Y_{n-1}),$$

where C is a normalizing constant. The aproximation F^p_n is now described via $\{X^{(j)}_n, w^{(j)}_n\}^p_{j=1}$.

The particle filter described above suffers from numerical instabilities because it leads to the weights of only a few particles dominate the overall expectations. In order to remedy that problem, one resamples from the weighted distribution after every few steps. Resampling also introduces unwanted extra randomness into the algorithm. Thus, one must be careful not to resample too frequently. In practice, we fix a resampling interval of length l; namely, after every l steps, we add a third step to the algorithm as follows.

3. Resample p points with replacement from the distribution F_n^p. Call the resampled points once again $\{X_n^{(j)}\}_{j=1}^p$. Reset the weights to be $w_n^{(j)} = \frac{1}{p}; j = 1, ..., p.$.

For the application under consideration particle filters are particularly useful since the particles of low probability are automatically eliminated, avoiding the issue of grid shifting that was necessary with method one.

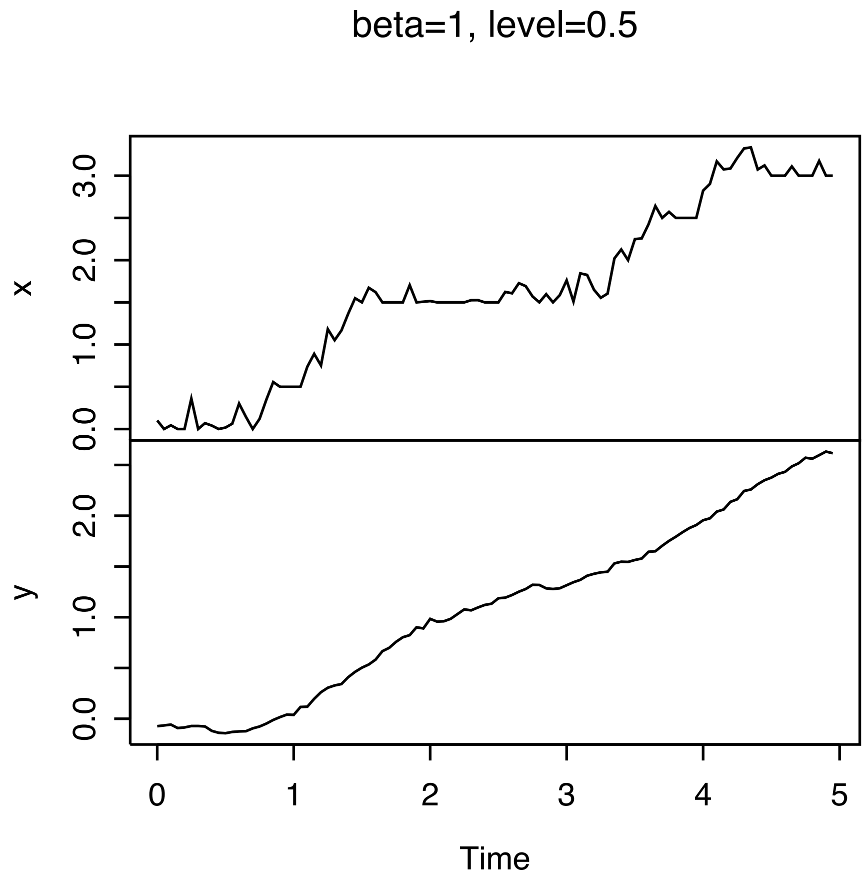
7.1.3 Numerical Comparison

In this section, we will make numerical comparisons of performance of the two methods introduced in Sections 7.1.1 and 7.1.2. The initial distributions for X_0 will be uniform on $[0, L)$ with $\sigma_1 = 1, \sigma_2 = 0.1$ and L equal to 0.5. In the simulation, Y_0 is set to be -0.5, and X_0 is set to 0. We will simulate (X_i, Y_i) paths for $N = 100$ time steps. A typical simulation of (X_i, Y_i) for $\beta = 1$ path is shown in Figure 7.1. Denote the filter obtained by the ith method, $k = 1, 2$, at time intervals i by F_i^k. For comparison purposes we used the root mean square error for each of the methods, defined as $\sqrt{\sum_{i=1}^N (x_i - \hat{x}_i^k)^2/N}$, where $\hat{x}_i^k \doteq \int x F_i^k(dx), k = 1, 2$. For computing the filter using method one, then one needs to decide the number of grid points, g. Results are shown in Table 7.1 for $g = 20, 40, 80, 120$. For using method two, we need to decide on the number of particles, p, to be used. Note that one crude estimate of the motor location is the cargo location. The results are given in Table 7.1 for $p = 50, 100, 200, 300$ It is seen that for the simulated path that using this estimate yields a root mean square error of 1.358. Table 7.1 shows that both method one and method two yield significantly better root mean square errors. Though a direct comparison between g and p cannot be made, it is clear that method one is more computationally intensive that the particle method, and computation using 300 particles is faster that the computation with method one using 120 grid points. We also observe that the particle filter stabilizes at around 100 paths, and the performance is not significantly changed by further increasing the number of particles.

In our numerical results, it was also observed that method two did qualitatively better in tracking the path of the motor than method one. This is illustrated in Figures 7.2 and 7.3 for a coarse grid/few particles example and a fine grid/many particles example, respectively.

One expects that as β becomes larger, the position of the cargo becomes a better estimate for the motor location since a large β corresponds to a tighter linkage between the motor and cargo. We studied the case when $\beta = 2$ which yielded a root mean square error of 0.845717, when the cargo location was used as an estimate. This was better than the results obtained for method one, even for the case of $g = 120$. However, the particle filter seems to do much better than both of these other estimates. Results are seen in Table 7.2. A simulation and results from the two methods are given in Figure 7.4.

Figure 7.1: Simulation



Finally, in Table 7.3, we summarize the performance of the different methods for ten different realizations of motor/cargo trajectories. We use $g = 120$ and $p = 300$ for this calculation. It is seen that with few exceptions the particle filter consistently outperforms both method one and the cargo as an estimator.

Figure 7.2: Comparison of Methods One and Two with Coarsest Grid, $\beta = 1$

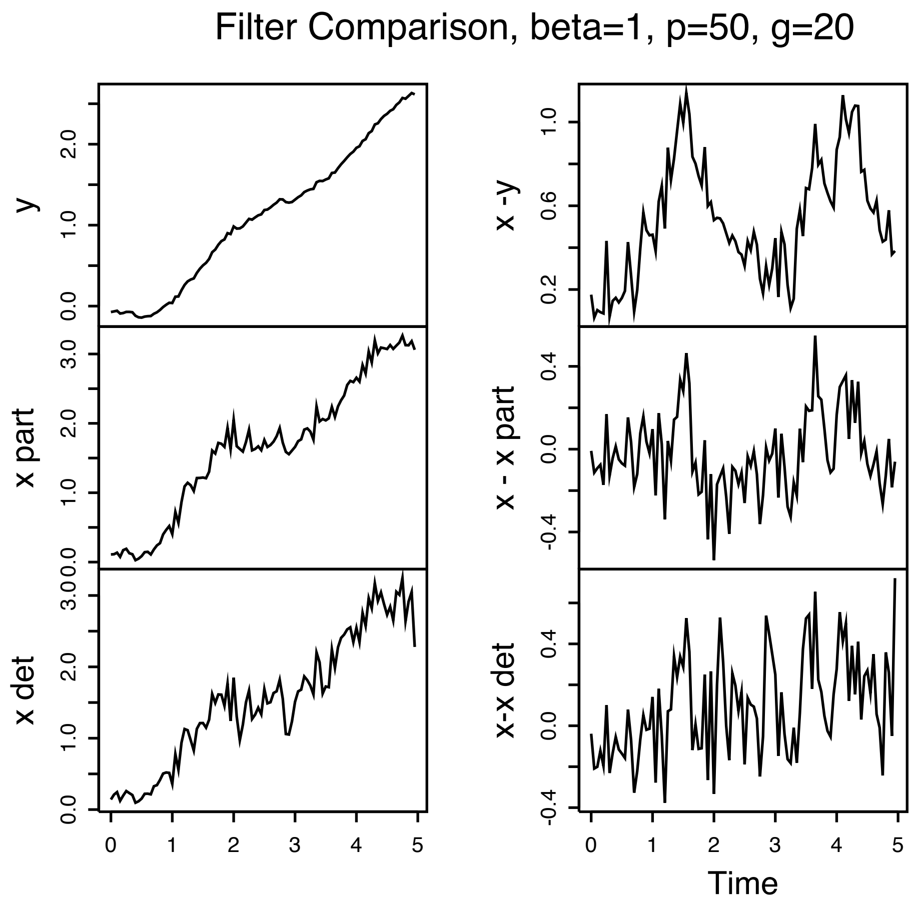


Table 7.1: Comparison for $\beta = 1$.

Method Two		Method One	
Num of Particles	Root Sq Error	Num of Grid Pts	Root Sq Error
50	0.4406443	20	0.5878180
100	0.3980596	40	0.5263298
200	0.3898426	80	0.5051189
300	0.3975316	120	0.4986271

Figure 7.3: Comparison of Methods One and Two with Finest Grid, $\beta = 1$

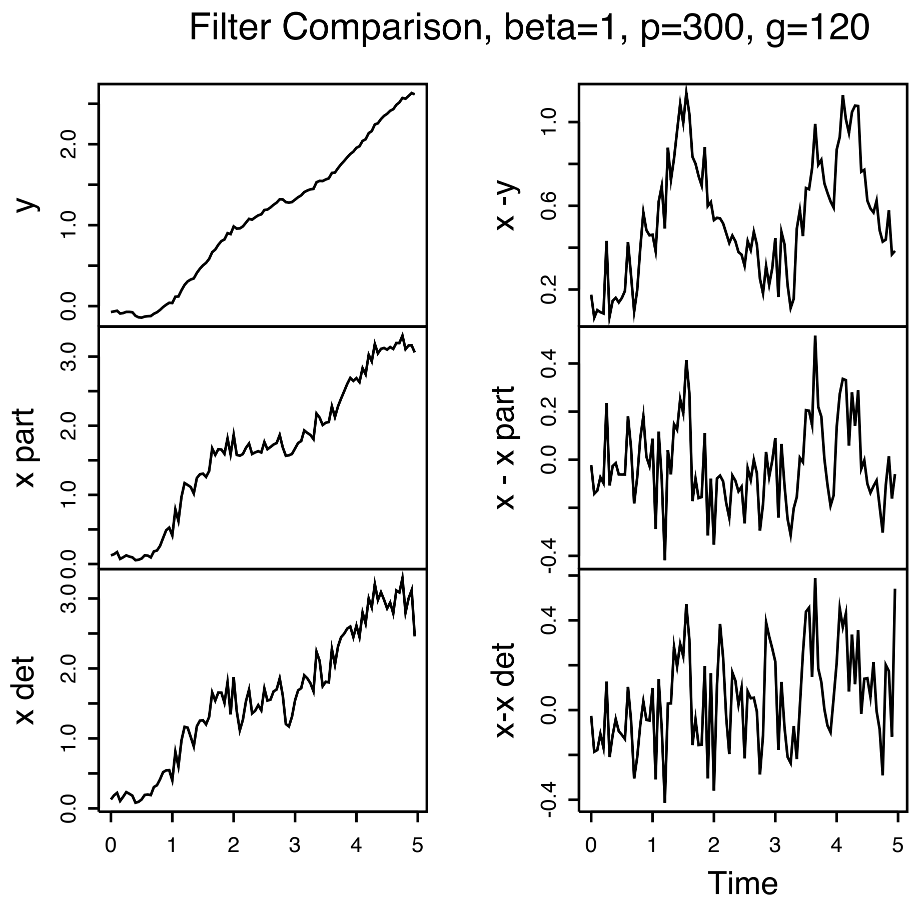


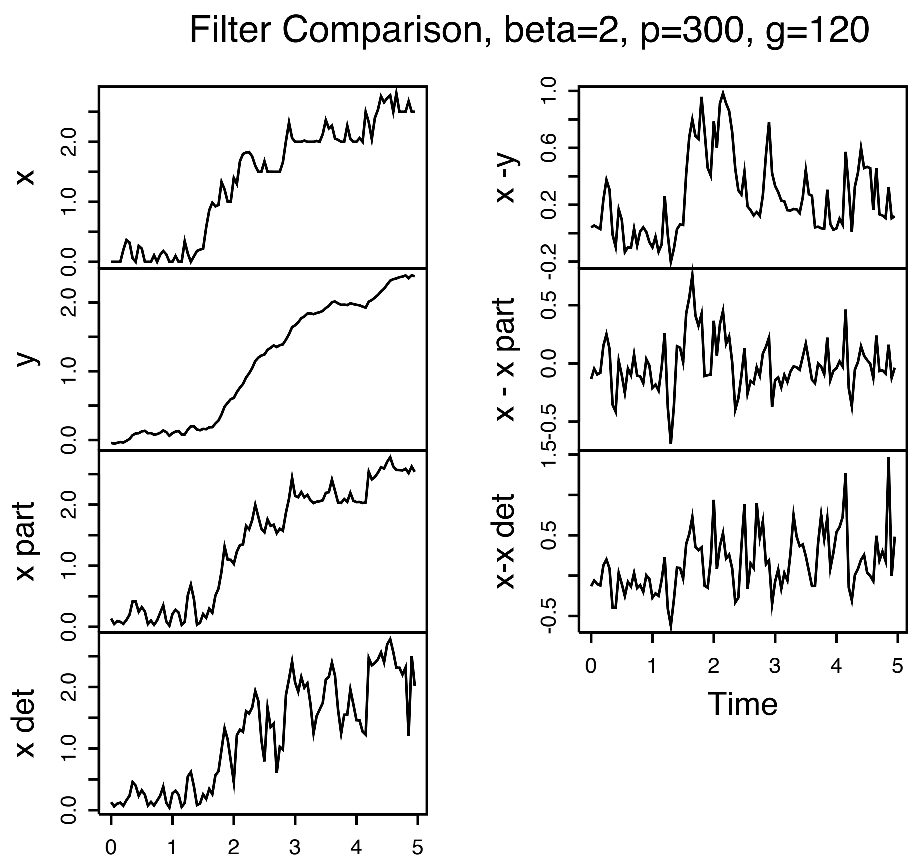
Table 7.2: Comparison for $\beta = 2$.

Method Two		Method One	
Num of Particles	Root Sq Error	Num of Grid Pts	Root Sq Error
50	0.5137411	20	1.0424467
100	0.4870611	40	0.9458944
200	0.4917597	80	0.9000441
300	0.5064459	120	0.8877230

Table 7.3: Comparisons over Multiple Runs

$\beta = 0.5$			$\beta = 1$			$\beta = 2$		
Meth 2	Meth 1	Cargo	Meth 2	Meth 1	Cargo	Meth 2	Meth 1	Cargo
0.6863396	0.6881833	2.5620726	0.6298464	0.7364986	1.2508070	0.5267787	1.1327014	1.4306390
0.7070487	0.9523465	3.1489532	0.6214509	0.8348315	1.1901541	0.4458444	0.5204378	0.5984934
0.6968931	0.8122405	3.1461678	0.5390040	0.7075981	1.3011144	0.4873355	1.3787924	1.2683393
0.7249526	0.9078173	2.6951204	0.5730014	0.7003326	1.4665928	0.4894461	0.7689176	0.8164092
0.6875986	0.7638412	2.2772461	0.6168053	0.8771855	2.0260558	0.5749243	0.8449805	1.1452869
0.7151371	0.7507665	1.8813909	0.4793599	0.7472386	1.9094463	0.5478237	0.8109069	0.8834096
0.9662632	1.1374571	3.1486008	0.6071232	0.5827366	0.7382102	0.3801122	2.0151536	1.0252873
0.7611157	0.7434068	2.8914932	0.6084563	0.7584619	1.4908681	0.4771855	1.0259939	0.6531618
0.6971034	0.6588305	1.8992784	0.4988502	0.7321273	1.1658740	0.4979003	1.5592010	1.1228489
0.8139982	0.8266756	1.6787565	0.5269929	0.8952635	1.6792052	0.5070097	0.6500806	1.0650676

Figure 7.4: Comparison of Methods One and Two, $\beta = 2$



7.2 Particle Smoothing

In applications where on-line tracking is not of main concern, one can obtain better state estimates by computing the conditional distribution of X_i given $Y_1, ..., Y_N$ rather than the filter, i.e. the conditional distribution of X_i given $Y_1, ..., Y_i$, where N is the total number of time steps. This former conditional distribution is referred to as the non-linear smoother. In this section, we will describe a particle system based algorithm for approximating the non-linear smoother which was introduced in [11]. The key step in the algorithm is the computation of the particle cloud $\{X_{0:N}^{(j)}\}_{j=1}^{p}$, where $X_{0:N}^{(j)}$ represents the sequence $\{X_1^{(j)}, X_2^{(j)}, ..., X_N^{(j)}\}$. The empirical distribution of the particle cloud:

$$F_{sm}(dx_0, ..., dx_n) \doteq \frac{1}{N} \sum_{j=1}^{p} \delta_{X_{0:N}^{(j)}}(dx_0, ..., dx_N)$$

is the estimate of the nonlinear smoother; namely the conditional distribution of $X_{0:N} = (X_0, ..., X_N)$ given $Y_{0:N} = (Y_0, ..., Y_N)$. The particle cloud is obtained by first computing the filter estimate and then conducting a backward sweep in the following manner.

1. Compute particle filter at time instants $0, 1, ..., N$. Suppose that they are given by particles $\{\hat{X}_i^{(j)}\}_{j=1}^{p}, i = 0, 1, ..., N$ with associated weights $w_i^{(j)}; j = 1, ..., p, i = 0, 1, ..., N$.

2. The smoother at time instant N equals the filter at time instant N, i.e. $\{X_i^{(j)}\}_{j=1}^{p} = \{\hat{X}_i^{(j)}\}_{j=1}^{p}$.

3. Suppose now that $\{X_{i+1}^{(j)}\}_{j=1}^{p}$ has been obtained. The smoother at time instant i

is obtained by an accept/reject scheme as follows.

(a) Sample one particle, X^c, from $\frac{1}{p}\sum_{j=1}^{p}\delta_{\hat{X}_{i-1}^{(j)}}$.

(b) We then propagate this particle by one step using equation (7.5). i.e. we chose a $N(0, \sigma_1^2\Delta)$ random number and denote it by η^c. Let

$$\hat{X}^c \doteq \psi(X^c, Y_{n-1}, \eta^c).$$

(c) For $x \in (0, \infty)$, $z \doteq \lfloor \frac{x}{L} \rfloor$, and let μ be a probability measure on $[zL, \infty)$, given as

$$\mu(A) = \frac{1}{2}\delta_{zL}(A) + \frac{1}{2}\int_A e^{-(u-zL)}du, A \in \mathcal{B}([zL, \infty)).$$

Define $\hat{\gamma}(\tilde{x}, x, Y_i)$ as the Radon-Nikodym derivative of $\gamma(d\tilde{x}, x, Y_i)$ with respect to μ. Let $S = \sup_{\tilde{x},x}\hat{\gamma}(\tilde{x}, x, Y_i)$.

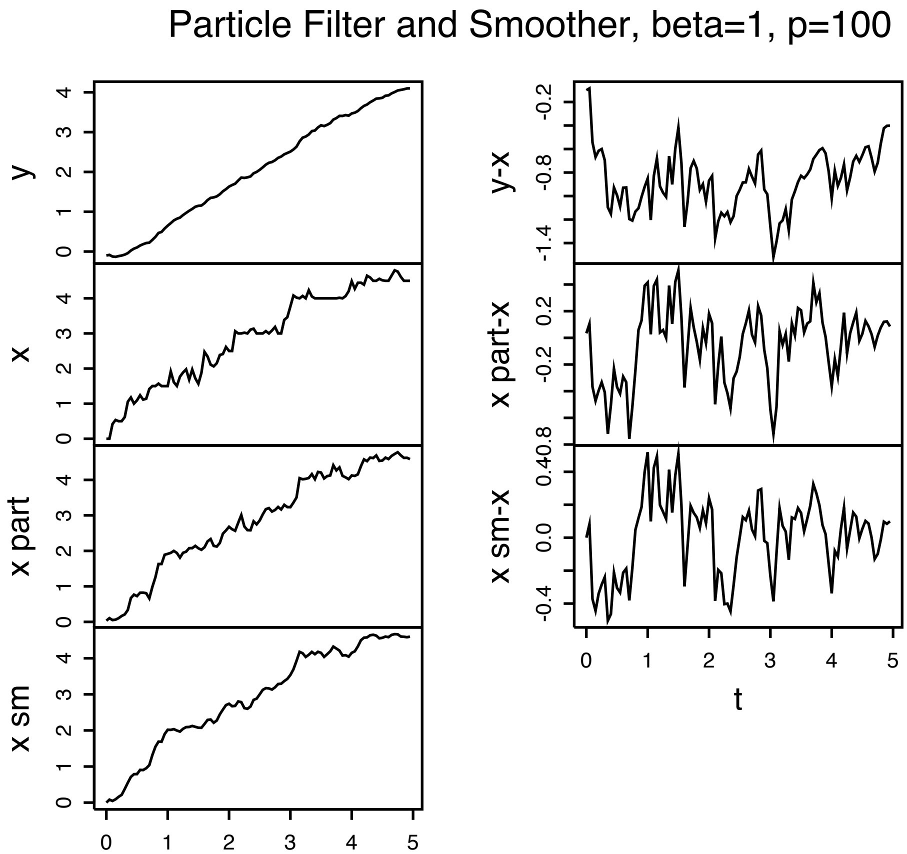
(d) Generate a uniform $[0, S]$ random variable, U. We accept $\tilde{X}^c$ and denote it by $X_i^{(j)}$ if U is less than $q(Y_i|\hat{X}^c, Y_{i-1})\hat{\gamma}(\hat{X}^c, X_{i+1}^{(j)}, Y_i)$. Otherwise, we reject $\tilde{X}^c$ and repeat starting with (3a).

4. The accept/reject scheme is conducted for each $j = 1, 2, ..., p$

5. We now have the smoother for time step i, $\frac{1}{p}\sum_{j=1}^{p}\delta_{X_i^{(j)}}$. We return to 3 with $i = i - 1$. Stop when $i = 0$.

We applied the smoothing algorithm for the application at hand for model parameter values as in Section 7.1.3 and with $\beta = 1$. Due to the ratchet mechanism of the

Figure 7.5: Example of Particle Smoother



dynamics, we expect that future observations do not contribute too much information on past state values. It was observed that the smoothing algorithm gave a root mean square error of 0.23476 where as the particle filter algorithm of Section 7.1.2 yielded a root mean square error of 0.28157. For this simulation using the cargo as an estimator would have yielded an error of 0.9053. Figure 7.5 shows a typical simulated motor/cargo paths, and the corresponding estimates using filter and smoothing approximations.

7.3 Parameter Estimation

Non-linear filtering theory gives methodologies for the computation of state estimates for the observation trajectory for the situation where all the model parameters are specified. However, in practice one rarely has access to the value of these parameters and thus they need to be estimated before filtering algorithms can be used. For the model in Section 7.1, there are various parameters that can be unknown such as σ_1, σ_2, L, and β. In this section, we will focus on the case when the only unknown parameter in the model is β. One can derive analogous methodologies for estimation of other parameters as well. We will consider two methods, the first is a simple Bayesian technique which treats the unknown parameters as an unobserved state and the second is a form of likelihood maximization. These are described in sections 7.3.1 and 7.3.2, respectively. A comparison of the two methods is made in section 7.3.3.

7.3.1 Method One: Dynamic Parameter Model

The basic idea is to propose a dynamic model for the parameter of the form

$$\hat{\beta}_i = \hat{\beta}_{i-1} + \epsilon_i$$

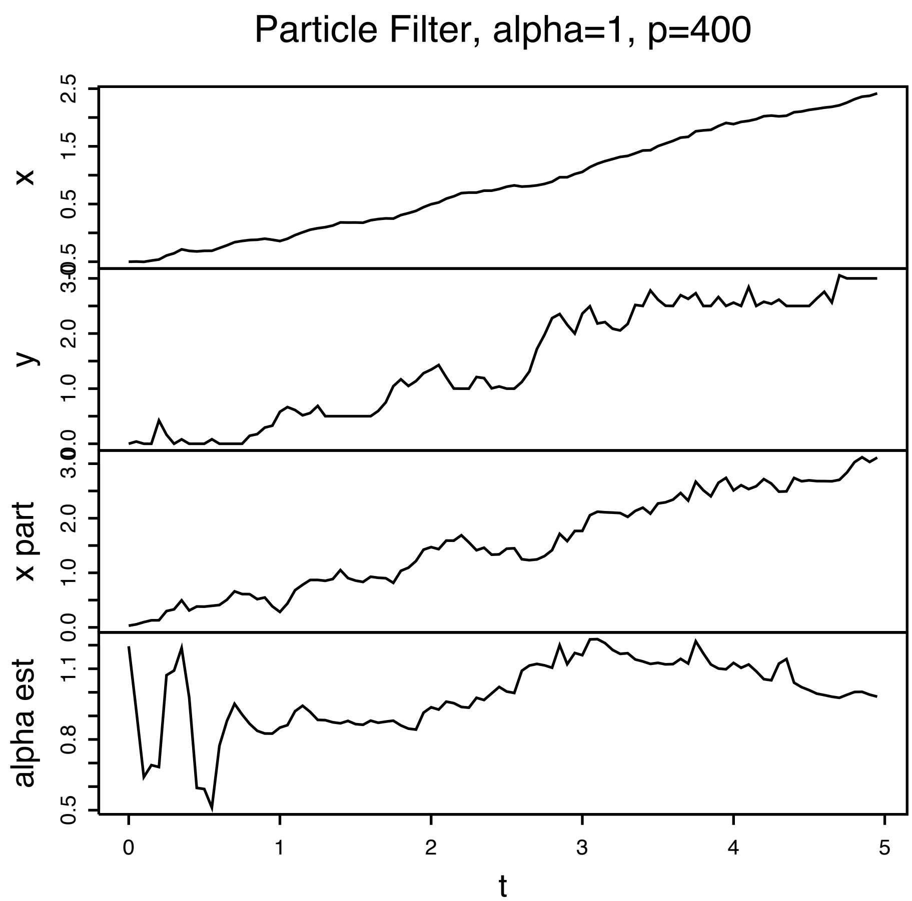
where ϵ_i is a sequence of zero mean independent normal random variables with variance $\frac{1}{i^2}$, and $\hat{\beta}_0$ is a normal random variable with mean equal to apriori guess of the parameter. The variance is taken to decrease as i increases so as to avoid excessive influence of observations far in the past. We now consider the augmented Markov process: $(X_i, \hat{\beta}_i, Y_i)$ where $\{Y_i\}$ is the observed component and $(X_i, \hat{\beta}_i)$ is the unobserved component of the Markov process. Particle filters as in Section 7.1 are then used to compute the conditional mean of $\hat{\beta}_i$ given $Y_0, Y_1, ..., Y_i$ for $i = 0, 1, ..., N$. The conditional mean of $\hat{\beta}_N$ given $Y_0, Y_1, ..., Y_N$ gives the estimated value of the parameter based on the whole observed sequence $Y_0, Y_1, ..., Y_N$. In practice, it is observed that one may obtain more stable estimates by not taking the last estimate in the sequence, but by averaging over the last portion of the sequence.

Figure 7.6 gives a simulation of the motor/cargo path which is generated from the model with $\beta = 1$, and the corresponding estimate of $\hat{\beta}$ that is obtained dynamically over time. In this computation, the number of particles, p, was taken to be 400 and number of time steps $N = 100$. We took the initial distribution of β_0 to be $N(\beta + 0.2, 0.1)$ where β is the true parameter. Averaging over the last twenty observations yields and estimate of β to be 1.01536. To assess the variability of the estimate over different motor/cargo path, we computed estimates for twenty different paths and $\beta = 0.5, 1, 2$. Results are summarized in table 7.4.

Table 7.4: Parameter Estimation for Different Levels of β

β	average estimate	sample s.d.
0.5	0.6141	0.2017
1	1.102	0.3485
2	2.029	0.3436

Figure 7.6: Estimation of β using Bayesian Techniques



7.3.2 Method Two: Stochastic Maximum Likelihood.

Let $p_{y_{1:N}|\beta}(\cdot)$ denote the density function of $Y_1, ..., Y_N$ when the parameter value is β. The central object used in parameter estimate via likelihood maximization is the log likelihood function.:

$$L(\beta) \doteq \log p_{y_{1:N}|\beta}(Y_{1:N})$$

The goal is to find the value of β with maximizes the above quantity for the given observation vector $Y_{1:N}$. In this section, we will follow an algorithm which has been introduced by Hürzeler and Kunsch [12]. The key step in the algorithm is to obtain a particle system approximation for the non-linear smoother for one fixed value of parameter $\beta = \beta_0$. Based on this one can compute $L(\beta) - L(\beta_0)$ for a range of values of β which are close to β_0 in the following manner. Observe that

$$L(\theta) - L(\theta_0) = \log \frac{p_{y_{1:N}|\beta}(Y_{1:N})}{p_{y_{1:N}|\beta_0}(Y_{1:N})} \tag{7.7}$$

$$= \log \left(E \left[\frac{p_{x_{1:N},y_{1:N}|\beta}(Y_{1:N})}{p_{x_{1:N},y_{1:N}|\beta_0}(Y_{1:N})} |Y_{1:N}, \theta_0 \right] \right) \tag{7.8}$$

The last quantity can be approximated via the particle smoother $\frac{1}{p}\sum_{j=1}^{p} \delta_{X^{(j)}_{1:N}}$ as follows:

$$L(\beta) - L(\beta_0) \approx \log \left(\frac{1}{p} \sum_{j=1}^{p} \frac{\nu(X_0, Y_0|\beta)}{\nu(X_0, Y_0|\beta_0)} \right. \tag{7.9}$$

$$\cdot \left. \prod_{i=1}^{N} \frac{d\gamma(X_i^{(j)}|X_{i-1}^{(j)}, Y_{i-1}, \beta) q(Y_i|X_i^{(j)}, Y_{i-1}, \beta)}{d\gamma(X_i^{(j)}|X_{i-1}^{(j)}, Y_{i-1}, \beta) q(Y_i|X_i^{(j)}, Y_{i-1}, \beta_0)} \right) \tag{7.10}$$

So, by calculating one smoother one may calculate the likelihood for a range of parameter values. This approximation becomes poor when β is too far from β_0, and in this case a new smoother must be produced. Thus in practice, we start with an initial guess for the parameter and then use (7.8) to fine the maximizer of $L(\beta)$ in the neighborhood of $L(\beta_0)$. When the approximation depends on too few of the particles, we take the current maximum as the current guess for β. We then recalculate the smoother at the current guess, and again begin to maximize the new approximate likelihood. We stop when we find a maximum.

Due to the calculation of possibly multiple smoothers, this method is very computationally expensive. Indeed, in order to avoid the convergence of the scheme to local maxima one needs to start with a skeleton of parameter values approximating the parameter space rather than a single value. Even though a lot of computation can be parallelized, the algorithm is extremely expensive.

7.3.3 Numerical Example

We will now compare the two methods described in Sections 7.3.2 and 7.3 above for parameter estimation. We will use β at levels $0.5, 1, 2$ for simulating the model. The Bayesian estimate is calculated with a filter of size $p = 300$. The maximum likelihood estimator is calculated using a smoother of size $p = 100$. For the first method we use the initial distribution of the parameter as $N(\beta_0, 0.1)$ and for the second method use the initialization for likelihood maximization as β_0.

Table 7.5: Comparison, $\beta = 0.5$

β_0	Likelihood Estimate	Bayesian Estimate
0.3	0.39	0.519
0.5	0.62	0.618
0.7	0.74	0.510
1	1.1	0.435

Poor performance of the likelihood methods can be attributed to the fact that in its current form the algorithm tends to converge to a local maximum and thus sometimes fails to find a global maximum. Furthermore, in this method one computes

Table 7.6: Comparison, $\beta = 1$

β_0	Likelihood Estimate	Bayesian Estimate
0.7	1.27	0.991
1	1.3	1.010
1.2	1.48	1.062
1.5	1.59	0.959

Table 7.7: Comparison, $\beta = 2$

β_0	Likelihood Estimate	Bayesian Estimate
1.7	2.45	2.159
2	2.63	1.710
2.2	2.41	1.870
2.5	2.62	2.213

the whole smoother for a fixed (incorrect) value which could be numerically quite unstable if the starting parameter value is far from the true parameter. In contrast, the advantage of the Bayesian method is that it continuously and hence dynamically corrects for a poor parameter value and improves the filtering procedure while it is on going. This simultaneous estimation of parameter and filtering seems to yield good results in practice.

Bibliography

[1] R. Anderson and S. Orey. Small random perturbations of dynamical systems with reflecting boundaries. *Nagoya Math J*, 60:189–216, 1976.

[2] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley & Sons Inc, New York, second edition, 1999.

[3] Steven Block. Fifty ways to love your lever: Myosin motors. *Cell*, 87:151–157, October 1996.

[4] Paul Dupuis and Hitoshi Ishii. On Lipschitz continuity of the solution mapping to the Skorokhod problem with applications. *Stochastics and Stochastics Reports*, 35:31–62, 1991.

[5] Richard Durrett. *Probability: Theory and Examples*. Wadsworth, Inc, Belmont, CA, 1991.

[6] Richard Durrett. *Essentials of Stochastic Processes*. Springer, New York, 1999.

[7] Timothy C. Elston and Charles S. Peskin. The role of protein flexibility in molecular motor function: Coupled diffusion in a tilted periodic potential. *SIAM Journal of Applied Math.*, 60(3):842–867, 2000.

[8] J. Michael Harrison. *Brownian Motion and Stochastic Flow Systems*. Krieger Publishing Company, Malabar, Florida, 1985.

[9] J.M. Harrison and M.I. Reiman. Reflected brownian motion on an orthant. *The Annals of Prabability*, 9:302–308, 1981.

[10] Kurt Helmes, Stefan Rohl, and Richard Stockbridge. Computing moments of the exit time distribution for Markov processes by linear programming. *Operations Research.*, 49(4):516–530, July-August 2001.

[11] Markus Hurzeler and Hans Kunsch. Monte carlo approximations for general state-space models. *Journal of Computational and Graphical Statistics*, 7(2):175–193, 1998.

[12] Markus Hurzeler and Hans Kunsch. *Sequential Monte Carlo Methods in Practice*, chapter Approximating and Maximising the Likelihood for a General State-Space Model. Springer-Verlag, NY, 2001.

[13] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, New York, second edition, 1991.

[14] Thomas Kurtz. *Approximation of Population Processes.* SIAM, 1981.

[15] Harold J Kushner and Paul Dupuis. *Numerical Methods for Stochastic Control Problems in Continous Time.* Springer, New York, second edition, 2001.

[16] P.L. Lions and A.S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, 37:511–537, 1984.

[17] S P Meyn and R L Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, New York, 1993.

[18] A. Mogilner, T. Elston, H. Wang, and G. Oster. *Computational Cell Biology*, chapter Molecular motors: examples. Springer, NY, 2002.

[19] A. Mogilner, T. Elston, H. Wang, and G. Oster. *Computational Cell Biology*, chapter Molecular motors: theory. Springer, NY, 2002.

[20] A. V. Skorokhod. Stochastic equations for diffusion processes in bounded region. *Theory of Probability and its Applications.*, VI:264–274, 1961.

[21] H. Tanaka. Stochastic differential equations with reflecting boundary conditions in convex regions. *Hiroshoma Math J*, 9:163–177, 1979.

[22] David Brian Walton. *Analysis of Single-Molecule Kinesin Assay Data by Hidden Markov Model Filtering.* PhD thesis, University of Arizona, 2002.

[23] Hongyun Wang, Charles S. Peskin, and Timothy C. Elston. A robust numerical algorithm for studying biomolecular transport processes. *Journal of Theoretical Biology.*, 221:491–511, 2003.

[24] Michelle D. Wang and all. Force and velocity measured for single molecules of RNA polymerase. *Science*, 282:902–907, 30 October 1998.

[25] R. Williams and S.R.S. Varadhan. Brownian motion in a wedge with oblique reflection. *Comm. Pure Appl. Math.*, 24:147–225, 1985.