UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

STRONG CONVERGENCE OF A MILSTEIN SCHEME FOR A CEV-LIKE SDE AND SOME CONTRIBUTIONS TO THE ANALYSIS OF THE STOCHASTIC MORRIS-LECAR NEURON MODEL

TESIS PARA OPTAR AL GRADO DE DOCTOR EN
CIENCIAS DE LA INGENIERIA MENCIÓN MODELACIÓN MATEMÁTICA

HÉCTOR CRISTIAN OLIVERO QUINTEROS

PROFESOR GUÍA:
JOAQUÍN FONTBONA TORRES

MIEMBROS DE LA COMISIÓN:
DANIEL REMENIK ZISIS
JAIME SAN MARTÍN ARISTEGUI
SOLEDAD TORRES DÍAZ

SANTIAGO DE CHILE
2016

RESUMEN DE LA TESIS
PARA OPTAR AL GRADO DE DOCTOR EN
CIENCIAS DE LA INGENIERIA
MENCIÓN MODELACIÓN MATEMÁTICA
POR: HÉCTOR OLIVERO QUINTEROS
FECHA: 10/11/2016
PROF. GUÍA: SR. JOAQUÍN FONTBONA TORRES

CONVERGENCIA FUERTE DE UN ESQUEMA DE MILSTEIN PARA UNA EDE DE TIPO CEV Y ALGUNAS CONTRIBUCIONES AL ANÁLISIS DEL MODELO DE NEURONAS DE Morris-Lecar ESTOCÁSTICO

Desde muy temprano en el desarrollo de la teoría de procesos estocásticos ha existido un creciente interés por aplicar sus herramientas en diferentes contextos; ya en el año 1900 Bachelier creó un modelo de movimiento Browniano para describir el mercado de acciones en París [7], y desde entonces el rango de aplicaciones del modelamiento estocástico ha seguido creciendo y hoy en día incluye desde economía hasta biología.

Esta tesis tiene dos partes, cada una de ellas dedicada al estudio de un modelo estocástico diferente. En la primera se estudia la aproximación numérica de la solución de una ecuación diferencial estocástica con aplicaciones en finanzas. Mientras que en la segunda se estudia un modelo estocástico para neuronas con énfasis en los “comportamientos asintóticos”.

La primera parte de esta tesis se organiza como sigue. En el Capítulo 2 se presenta una breve introducción a los métodos clásicos de aproximación de soluciones de ecuaciones diferenciales estocásticas y se recuerdan sus propiedades de convergencia. Luego, en el Capítulo 3 se estudia un esquema numérico para aproximar las soluciones de

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma|X_s|^\alpha dW_s.$$

Esta ecuación se puede ver como la generalización del modelo CIR para tasas de interés y tiene un gran rango de aplicaciones en finanzas. El principal resultado de este capítulo es la convergencia fuerte con tasa 1 del esquema numérico estudiado a la solución exacta de la ecuación. Este capítulo está basado en un trabajo conjunto con Mireille Bossy [16], el cual ha sido aceptado para su publicación en la revista Bernoulli.

En la segunda parte de esta tesis se estudia el modelo de Morris-Lecar para una red de neuronas. El principal objetivo es estudiar el comportamiento del sistema cuando el tiempo o el número de neuronas se va a infinito. Sin embargo, antes de abordar esas temáticas, se discuten dos versiones estocásticas para el modelo de Morris-Lecar, y la relación entre ellas. Los resultados principales de esta parte de la tesis son la caracterización del comportamiento límite, en intervalos de tiempo finito, para una red de neuronas cuando el número de neuronas diverge a infinito y un resultado de sincronización para una red finita de neuronas cuando el tiempo diverge a infinito.

RESUMEN DE LA TESIS
PARA OPTAR AL GRADO DE DOCTOR EN
CIENCIAS DE LA INGENIERIA
MENCIÓN MODELACIÓN MATEMÁTICA
POR: HÉCTOR OLIVERO QUINTEROS
FECHA: 22/11/2016
PROF. GUÍA: SR. JOAQUÍN FONTBONA TORRES

STRONG CONVERGENCE OF A MILSTEIN SCHEME FOR A CEV-LIKE SDE AND SOME CONTRIBUTIONS TO THE ANALYSIS OF THE STOCHASTIC MORRIS-LECAR MODEL

From very early in the development of the theory of stochastic processes there has been an increasing interest to apply its tools in different contexts; already in 1900 Bachelier created a model of Brownian motion to describe the stock market in Paris [7], and since then, the range of application of stochastic modeling has keep growing and nowadays includes from economy to biology.

This thesis has two parts, each of them dedicated to the study of a different stochastic model. In the first part, we study the numerical approximation of the solutions of a stochastic differential equation with applications in finance. In the second part, we study a stochastic model for neurons with emphasis in "asymptotic behaviors".

The first part of this thesis goes as follows: in Chapter 2 we present a short introduction to the most classical methods to approximate solutions of stochastic differential equations, and we recall their convergence properties. Next, in Chapter 3 we study a numerical scheme to approximate the solutions of

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma|X_s|^\alpha dW_s.$$

This equation can be seen as a generalization of the CIR model for interest rates, and it has a great range of application in finance. The main result of this chapter is the strong convergence at rate 1 of the studied numerical scheme to the exact solution. This chapter is based on a joint work with Mireille Bossy [16], which has been accepted for publication in the Bernoulli Journal.

In the second part of this thesis we study the stochastic version of the Morris-Lecar model for neurons. Our interest is the asymptotic behaviors of the system, meaning, the limit behaviors when the time or the number of neurons goes to infinity. Nevertheless, before addressing these issues, we discuss two different stochastic versions of the Morris-Lecar model, and the connection between them. The main results of this part are the relation between this models, the limit behavior on a finite time interval for both of the stochastic models when the number of neurons goes to infinity and a synchronization result for a finite number of neurons when the time goes to infinite.

Along the years I was working in this thesis, life has happened, and three good, hardworking, and beloved men have left us. One of them, knew me my hole life, with the other two, we did not have the time to known each other well enough, but both gave me so significant people to my life: my beloved wife and my dearest friend.

TO THE MEMORY
OF

ALFONSO SABINO CAYUL CALFUQUIR, MY FATHER IN LAW,

ANDRÉS ITURRIAGA NAVARRO, MY BEST FRIEND'S FATHER,

AND

JORGE LUIS PALMA OLIVERO, MY DEAR COUSIN.

Acknowledgements

I want to start my acknowledgments by expressing my gratitude to the several institutions and programs without their financial support this thesis would not exist. I am deeply grateful to the *Proyecto Mecesup UCH0607*, the *Dirección de Postgrado y Postítulo de la Vicerrectoría de Asuntos Académicos de la Universidad de Chile*, the *Instituto Francés de Chile - Embajada de Francia en Chile*, the research center INRIA Sophia Antipolis, the Center for Mathematical Modeling CMM, and the *Núcleo Milenio* Stochastic Models of Complex and Disordered Systems: PROYECTO NUCLEO MILENIO NC 130062.

I also want to express my gratitude to my advisor Joaquín Fontbona. We have come a long way together. He was one of my Professors when I started my undergraduate studies, and since then, he has not been only a great teacher, but also a role model about what means to be a mathematician. I have many many things to thank to Joaquín, but over all, I want to thanks the trust he placed in me and his unbounded patience.

During this thesis, I had the privilege and pleasure to work also with Mireille Bossy. Since I met her, she has been so generous and patient, so supporting and kind, she has taught me so much. I am so deeply grateful to her.

I also want to thank also to Professors: Soledad Torres, Daniel Remenik and Jaime San Martin for their kindness in reviewing this thesis.

I want to thank to all the members of the Mathematical Engineering Department (DIM) at University of Chile. Academics, administrative staff, classmates. I have been a student in the department for ten years, and I have to say that I have always felt welcomed at DIM, I have always felt part of it, and that has been because all the wonderful people conforming it.

I want to thank to my friends Anne, Clarita, María José, Natalia, Roberto, Sebastián and Víctor for their friendship along the time we have shared.

I want to thanks to my extended family, all those people who love me and support me no matters what.

I want to thanks to my dear friend Paula Rodriguez. Paula, you have been an amazing friend for over than ten years, I am really grateful for the gift of your friendship.

I want to thanks to my dear friend Andrés Iturriaga. My dear Itukun, It makes me very happy and very proud to be part of your life. Thanks you for being my friend.

I also want to thanks to my parents for their infinite love. Mamá, Papá you are always in my heart.

Finally, I want to thank to my wife, María. My beloved María, without you I would never had got here. There are occasions when sentiments and emotions are far too much to be put it in words, so let me hold your hand and love you for ever.

Contents

Chapter 1

Introduction

"Y el azar se le iba enredando poderoso, invencible."

Causas y azares,
Silvio Rodriguez.

In this thesis we address the study of two stochastic models, the first one from finance and the second one from neuroscience. Although this two models are quite different, and we study them from very different perspectives, they are both described in terms of stochastic differential equations.

In the next section we give an intuitive and rough approximation to stochastic differential equations, we do not aim to give a compressive and precise introduction to the subject, but to show through a simple example how stochastic differential equations can be included in modeling. For a complete and detailed introduction to the the theory of stochastic calculus and stochastic differential equations see [44], [52] or [69].

1.1 Stochastic Differential Equations

From a very intuitive perspective, stochastic differential equations (SDE) driven by a Brownian motion can be seen as the formalization of the idea of a dynamical system evolving according to an ordinary differential equation (ODE) subject to random perturbations. For example, if we are modeling the instantaneous interest rate X_t, it is known that X is a process attracted to a long term value, let us say $b > 0$, the simpler model we can propose is

$$dX_t = a(b - X_t)dt, \tag{1.1.1}$$

with $a > 0$. This equation is linear with constant coefficients, so it has an explicit solution given by

$$X_t = X_0 e^{-at} + (1 - e^{-at})b,$$

and effectively, as the time goes to infinity, X_t goes to b. Even more, X_t is positive for all times, which is an important property of the instantaneous interest rate. But, what happened if we want to add a random perturbation to the equation, the reason to do this could be take into account some measurement errors, or maybe the previous knowledge of X tell us that in its nature there is some intrinsic randomness.

Inspired in the central limit theorem, we could assume that the random perturbation we want to add is a normal random variable with zero mean (in average the measurement errors cancel with each other), a variance increasing in time (the errors are cumulative), also it is desirable certain independence between the errors between $[t_1, t_2]$ and the errors between $[t_3, t_4]$ if $t_3 \geq t_2$. The mathematical formalization of this idea is the Brownian motion.

The Brownian motion should be the most famous continuous time Markov process. Own its name to the Scottish botanist Robert Brown, who observed in the motion of a pollen grain suspended in water in 1827. About 70 years later, Einstein in [26] explain the Brownian motion as the consequence of the interaction between the pollen grain and the water molecules. And twenty years later, Wiener [80, 81] gave a rigorous proof of the existence of the process. The Brownian motion or Wiener process satisfies

1. $W_0 = 0$ a.s.,
2. $W_t - W_s \sim N(0, t - s)$,
3. $\forall 0 < t_1 < t_2 < \ldots < t_n$, the random variables $W_{t_1}, W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent.

Let us go back to the equation (1.1.1), we can add the noise given by a Browian motion multiply for a parameter $\varepsilon > 0$ and obtain

$$dX_t = a(b - X_t)dt + \varepsilon dW_t, \tag{1.1.2}$$

Give a full meaning to this equation is not easy, as a function of t, the trajectories of the Brownian motion are nowhere differentiable, and then what does it mean dW_t? The answer for this question was given by Itô in [45], and according with Itô's theory the equation (1.1.2) has for solution

$$X_t = X_0 e^{-at} + (1 - e^{-at})b + \varepsilon e^{-at} \int_0^t e^{as} dW_s, \tag{1.1.3}$$

which is the well known Orstein-Uhlenbeck process. In [46], Itô established that under a Lipschitz condition for the coefficients a SDE like (1.1.2) has a unique solution.

In [77] Vasicek proposed to use the Orstein-Uhlenbeck process (1.1.3) as a model for the instantaneous interest rate. However, it can be shown that this process has normal law, and then, there is a positive probability of X_t being negative, regardless the values of X_0 and b. A solution to this problem was given by Cox, Ingersoll and Ross in [20], where they propose to use a Feller diffusion [31] as model for the interest rate, that is

$$dX_t = a(b - X_t)dt + \varepsilon\sqrt{X_t}dW_t. \tag{1.1.4}$$

Although there is not a closed formula for the solution of (1.1.4), it can be shown that the mean of this process is attracted to b, the process is always non negative, and under the condition $2ab > \varepsilon^2$, the process is strictly positive.

1.2 Part I: A Milstein Scheme for a CEV like SDE

Numerical Schemes for SDE

In the previous section we have seen how easy is to find a SDE without an explicit solution. So, it is very natural to ask for numerical methods that allows to approximate the value of a process defined though a SDE at some given time. The first answer to this question, is the natural extension of the Euler scheme to the stochastic setting. Let us recall that for a ODE of the form

$$dX_t = b(X_t)dt, \quad X_0 = x_0,$$

the Euler scheme given the temporal grid $0 = t_0 < t_1 < t_2 < \ldots < t_n = T$ is given by

$$\begin{aligned} X_{t_0} &= x_0 \\ X_{t_{n+1}} &= X_{t_n} + b(X_{t_n})(t_{n+1} - t_n). \end{aligned}$$

Then for the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0, \tag{1.2.1}$$

the Euler-Maruyama scheme [57] will be

$$\begin{aligned} X_{t_0} &= x_0 \\ X_{t_{n+1}} &= X_{t_n} + b(X_{t_n})(t_{n+1} - t_n) + \sigma(X_{t_n})(W_{t_{n+1}} - W_{t_n}). \end{aligned} \tag{1.2.2}$$

Two natural question arise immediately:

1. Does the numerical scheme converge to the exact solution of the (2.1.1) when the time grid becomes finer and finer?

2. Do we know the speed of this convergence?

This questions have not absolute answer and will depend on the hypothesis we assume about b and σ. If both functions are Lipschitz, the numerical scheme effectively converges to the exact solution when the step size of the grid, that is $\Delta t = \max\{t_{i+1} - t_i\}$, goes to zero. Even more, the convergence occurs with rate $\sqrt{\Delta t}$. This means, that when the step size of the grid is divided by four, the approximation error of the scheme is reduced by half.

Although the Euler-Maruyama scheme has in his favor to be quite economic in its hypotheses, has the downside of its rate of convergence being $\sqrt{\Delta t}$. We will see later that the rate of convergence can be improved to Δt, by considering the Milstein scheme, but the price to pay is that b and σ needs to be twice continuously differentiable with bounded derivatives.

Numerical Schemes for a CEV like SDE

The main objective of this part of the thesis is to study a numerical scheme for the equation

$$dX_t = b(X_t)dt + \sigma|X_t|^{\alpha}dW_t, \tag{1.2.3}$$

where $\sigma > 0$ and $\alpha \in [1/2, 1)$. It is can be shown that the solutions of this equation are almost surely positive.

The importance of equation (1.2.3) comes from its applications, mainly in finance. For example, this equation is used to model instantaneous interest rate (for $\alpha = 1/2$ and $b(x) = a - bx$, equation (1.2.3) becomes the CIR equation (1.1.4)), or to model stochastic volatility. An application outside the field of finance comes from fluid mechanics, where is used to model the instantaneous turbulent frequency (see [25]). All this potential applications has motivated a great interest in find numerical schemes to approximate the solution of equation (1.2.3). See for example [1–3], [10], [14], [39] and [50, 51].

Notice that in this case the function $\sigma(x) = \sigma|x|^{\alpha}$, so it is not Lipchitz, and then the convergence of the Euler-Maruyama scheme is not obvious. Even worse, we can not ensure that the scheme will preserve the positiveness of the initial condition. Indeed,

$$X_{t_1} = X_{t_0} + b(X_{t_0})(t_1 - t_0) + \sigma X_{t_0}^{\alpha}(W_{t_1} - W_{t_0}),$$

and $W_{t_1} - W_{t_0}$ has normal law with zero mean and variance equal to $t_1 - t_0$, so there is positive probability of X_{t_1} being negative.

A very simple solution to this problem is symmetrize the scheme, that is

$$\begin{aligned} X_{t_0} &= x_0 \\ X_{t_{n+1}} &= |X_{t_n} + b(X_{t_n})(t_{n+1} - t_n) + \sigma X_{t_n}^{\alpha}(W_{t_{n+1}} - W_{t_n})|. \end{aligned} \tag{1.2.4}$$

Following this strategy Bossy and Diop in [14], proved the convergence at rate $\sqrt{\Delta t}$ of the symmetrized Euler scheme to the exact solution under some hypothesis on b and σ.

The main result of the first part of this thesis will be the convergence at rate Δt of a symmetrized Milstein scheme that we will define later.

Summary of Part I

The first part of this thesis is composed by two chapters. In Chapter 2 we present the two more classical and standard numerical schemes for SDEs: The Euler-Maruyama scheme and The Milstein scheme. In the case of the Euler-Maruyama scheme we present the Theorem that characterize the convergence of the scheme towards the exact solution without proof. In the case of the Milstein scheme we give a proof because it will serve us as a guide in the proof of the main theorem of the first part of the thesis.

In Chapter 3 we present the main result of this part of the thesis which is the strong convergence at rate Δt of a symmetrized version of the Miltein scheme. The structure of this chapter is the following: First we introduce the scheme and enunciate the main Theorem. Then, we state several technical Lemmas, use them to prove the main Theorem, after that we include a section with numerical experiments that allows to test our scheme, and compare it with other known schemes in practice. Next we discuss the main conclusions of this part of the thesis, to end the chapter with the proof of the technical Lemmas.

1.3 Part II: Stochastic Morris-Lecar Model

Before to pass to the mathematical model we study in the second part of this thesis, let us introduce some of the language that we will need.

Some basic elements from neuroscience

This short section does not intend to be a complete and precise introduction to neuroscience. We just present a few elements needed to follow the mathematical developments we will do later. Most of the material has been taken from Chapter 1 in [27].

The basic structure of a neuron is showed in Figure 1.3.1a. A neuron have a *soma* which is the processing center of the neuron, *dendrites* which are the input lines of a neuron, and an *axon* which serves to connect with other neurons.

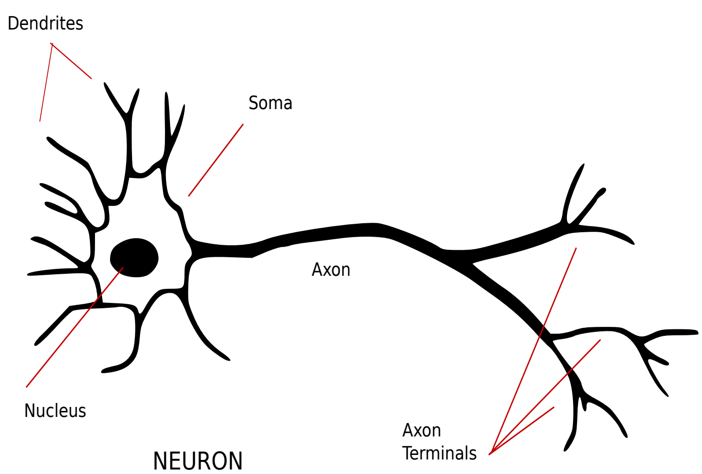
Like any other cells, neurons are enclosed by a membrane that separate them from the medium, and produce a difference in the concentration of ions between the inside and the outside. As a result, there is a difference of electrical potential between the two sides of the membrane, which we call the *membrane potential* or *membrane voltage*. If there is no external stimulus, the membrane will be at it *resting potential*. However, neurons are *excitable cells*[1], meaning that they are capable of change its membrane potencial by allowing the pass of ions through *channels* embedded in their membrane, as we see in Figure 1.3.1b.

[1]Muscle cells are also excitable cells.

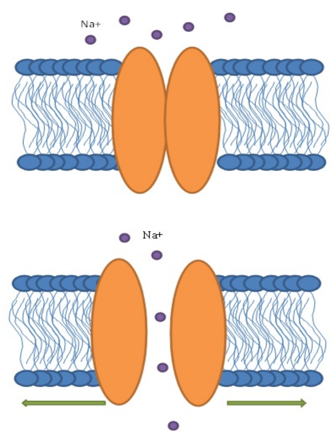
The ion channels are selective, in the sense that each channel allows the pass of a specific ions. This motivate us to talk about Calcium channels, Potassium Channels, Sodium channels, etc. Additionally, the ion channels are *voltage-gated* , which means that the channels will open and close depending on the voltage membrane.

The main characteristic of neurons, is their ability to propagate information. This propagation is in the form of a transient electrical signal called *action potential* or *spike*. An action potential correspond to an abrupt change in the neuron's membrane potential, which start when the membrane potential reach certain threshold, and then propagates through the axon to other neurons. Once a neuron has spiked, can not do it again immediately. This is called te *refractory period*.

A very important feature about the action potential is that the amplitud and speed of it does not depend on the stimulus. If the stimulus, let us say an input current, is not enough to raise the voltage over to the threshold mentioned before, the neuron will not spike. On the other hand, once the membrane voltage has reached the aforementioned threshold, the resulting spike will have the same amplitud and speed no matter by how much the threshold was exceeded. This implies that the information contained in the stimulus will be codified in the frequency of the spikes.



(a) Basic structure of a neuron.



(b) Ion Channel.

Figure 1.3.1: Basis structure of a Neuron and Ion Channels. [2]

Neurons communicate with each other through *synapses*. Synapses can be chemical or electrical. In a chemical synapses there exist a small separation, call the synaptic cleft, between the axon of the presynaptic neuron and the dendrites of the postsynaptic one. To communicate an action potential to the postsynaptic neuron, the presynaptic one will release chemical neurotransmitters to the synaptic cleft. As a result of the reception of

[2]This images have been taken from Wikimedia Commons.
Left Image: By derivative work: Notjim (talk)Neurone.png: Looxix at fr.wikipedia - Neurone.png, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=4824168
Right image: By Tom Mork - Own work, GFDL, https://commons.wikimedia.org/w/index.php?curid=19402370

this neurotransmitters, the membrane voltage of the postsynaptic neuron will change. If the change in the membrane potential of the postsynaptic neuron is big enough, an action potential will be generated by the postsynaptic neuron and transmitted to its neighbors. In Figure 1.3.2a we display a chemical synapse.

On the other hand, in a electrical synapses there exist a intercellular channel that allows the permanent flux of ion between the inside of both neurons. We display an electrical synapse in Figure 1.3.2b.

(a) Chemical Synapse. Neuron A release neurotransmitters into the synaptic cleft.

(b) Electrical Synapse. Neurons A and B are connected by a gap junctions that allow the permanent flux of ions between them.

Figure 1.3.2: Different type of synapses. [3]

Deterministic Mathematical Models

There exist three families of models to describe the dynamic of a neuron: Binary models, Leaky integrate-and-fire models and dynamical models. We are interested in the third type. For details in binary models see [58], and for a review on Leaky integrate-and-fire models see [18].

[3]This images have been taken from Wikimedia Commons.
Left Image: By vectorization: Mouagip (talk)`Synapse_diag1.png`: Drawn by fr:Utilisateur:DakeCorrections of original PNG by en:User:NretsThis vector graphics image was created with Adobe Illustrator. - `Synapse_diag1.png`, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=11438067
Right image: CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=282849

The dynamical models have the advantage of being much more realistic, although they are much more difficult to analyze. In this models, the membrane potential evolution is completely describe through a dynamical system. For example the famous Hodgking Huxley model [41] is given by

$$\begin{aligned} CdV_t &= I^{ext} - g_{\rm K} n^4 (V - V_{\rm K}) - g_{\rm Na} m^3 h (V - V_{\rm Na}) - g_{\rm L}(V - V_{\rm L}) \\ dm_t &= \rho_m(V_t)(1 - m_t) - \zeta_m(V_t) m_t \\ dn_t &= \rho_n(V_t)(1 - n_t) - \zeta_n(V_t) n_t \\ dh_t &= \rho_h(V_t)(1 - h_t) - \zeta_h(V_t) h_t, \end{aligned}$$

where V is the membrane potential, n is the proportion of potassium ion channels and m y h are the proportion of two different type of sodium channels.

Due to the nonlinearities in the dynamic for V, and the high dimensionality of the Hodgkin Huxley model, is quite standard to consider a simplification of the model, being the FitzHugh Nagumo the more standard one.

In the FitzHugh Nagumo model a neuron is describe by two variables, the voltage V and a recovery variable w. There is no direct interpretation to the recovery variable w, so instead of using this model, we will consider the Morris-Lecar model for neurons[4] [63] which is given by

$$\begin{aligned} CdV_t &= I^{ext} - g_{\rm K} n (V - V_{\rm K}) - g_{\rm Ca} m (V - V_{\rm Ca}) - g_{\rm L}(V - V_{\rm L}) \\ dm_t &= \rho_m(V_t)(1 - m_t) - \zeta_m(V_t) m_t \\ dn_t &= \rho_n(V_t)(1 - n_t) - \zeta_n(V_t) n_t, \end{aligned} \tag{1.3.1}$$

where in this case, m represent the proportion of open Calcium channels. Notice that there is a dimension reduction in this model compared with the one from Hodgkin and Huxley, and the dynamics for V is “more linear”.

Stochastic mathematical models

Nowadays there is enough evidence to say that there is an intrinsic randomness in neurons. In fact, it is known that each ion channel behaves as a two state continuous time Markov process, that jumps from close to open with rate $\rho_x(V_t)$, and from open to close with rate $\zeta_x(V_t)$. In this way the more natural stochastic model that can be proposed for a neuron is the Hybrid model, proposed in [67].

In the hybrid model the voltage of the neuron evolves according to a continuous dynamic, as in the Morris-Lecar model, but the proportion of open channels are jump processes taking values in the set $\{1/N_c, \dots, 1 - 1/N_c, 1\}$, where N_c is the number of channels in the neuron.

[4] Originally, Morris and Lecar propose their model for the voltage oscillations of a muscle fiber of a barnacle. Later has become a popular model for neurons.

Although the hybrid model has the advantage of being very physiologically meaningful, it has the downside of been very expensive to simulate. This motivate the approximation of the hybrid model by a diffusive model. Inspired in the Langevin approximation of the hybrid model in [8] is proposed a diffusive model for the dynamics of the ions channels.

In this part of the thesis we will study the stochastic versions of the Morris-Lecar model: the hybrid and the diffusive one.

The mean field approach

Our main interest are not single neurons. Our main interest are networks of N interacting neurons, where N is large. How large? In a human brain there are around $85,000,000,000$ of neurons[5], each of them interacting in mean field way with approximately $10,000$ neighbors[6].

Given a system of interacting particles (in our case neurons), the interaction is of mean field type, if the aggregate effect of the system over a single particle depends only on the empirical measure of the whole particle system. For instance, if we denote by $b(X^{(i)}, X^{(j)})$ the interaction between the particle i and the particle j, the dynamics of the i-th particle can be of the form

$$dX_t^{(i)} = F(X_t^{(i)})dt + G(X_t^{(i)})dW_t + \frac{1}{N}\sum_j b(X_t^{(i)}, X_t^{(j)}). \tag{1.3.2}$$

This kind of system was introduced by Kac [49] to study the Boltzman equation in statistical mechanics. The key point of (1.3.2) is that, under suitable hypotheses, when the number N of interacting particles diverges to infinity, finitely many particles start to behave as independent copies of the process given by

$$\begin{aligned} d\bar{X}_t &= F(\bar{X}_t)dt + G(\bar{X}_t)dW_t + \int b(\bar{X}_t, y)\mu_t(dy), \\ \mu_t &= \text{law}(\bar{X}_t), \end{aligned} \tag{1.3.3}$$

or to be precise, for any $k \in \mathbb{N}$, when the number of particles goes to infinity, the law of k fixed particles tends to $P^{\otimes k}$, where P is the law of the solution of (1.3.3). This property is called propagation of chaos. For further background on this topic see [60] and [74].

In this part of the thesis, we are going to assume that the interaction between neurons is of mean field type. This assumption as been made by several authors in the last years. See for example [8], [15], [23], [29], [33] and [62].

[5] See [6].
[6] See [27, p.7]

Our goal

The main objective of this part of the thesis is to study networks of interacting neurons. We are interested in the asymptotic behavior of the neurons when either the time or the size of the network goes to infinity.

Summary of Part II

The second part of this thesis is composed by a single chapter. The first section is devoted to a small introduction that complement what we have said here. In the following section, we discuss some modeling issues for the stochastic version of the Morris-Lecar model for a single neuron. We continue in the next section with models for a network of neurons interacting under chemical and electrical synapses. Next we pass to the first asymptotic limit, and we prove the propagation of chaos property for finite time windows. In the following section, we discuss the behavior of a network of finite size when the time goes to infinity. In particular, we prove an interesting result on synchronization. We finish the chapter with a summary of the main conclusions of our work.

Part I

Strong Convergence of the Symetrized Milsteim Scheme

Chapter 2

Numerical Schemes for Stochastic Differential Equations

2.1 Introduction

Let us consider the following time homogeneous Stochastic Differential Equation (SDE for short)

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{2.1.1}$$

where W is a standard d-dimensional Brownian motion over a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ which satisfies the usual hypothesis, (ie) the filtration is right continuous and complete. The functions $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ will be regular enough so (2.1.1) has a unique strong solution, which is true for example if b and σ are Lipschitz functions with linear growth. For details see [44], [52], [65], [69].

Even when (2.1.1) is well posed, there are only a few cases when the solution for this equation can be analytically computed, and in most of the situations some approximating scheme is needed. In the rest of this chapter we will present two classical approximation schemes and their main features.

Before presenting the schemes, let us introduce some notation that we will use for the rest of Part I of the thesis. Let $x_0 > 0$, $T > 0$, and $N \in \mathbb{N}$. We define the constant step size $\Delta t = T/N$ and the time grid $\{t_k\}_{k=0}^N$, with $t_k = k\Delta t$. We introduce also the function

$$\begin{aligned} \eta &: \quad [0,T] \to \{t_0, \ldots, t_N\} \\ \eta(t) &= \sup_{k\in\{1,\ldots,N\}}\{t_k : t_k \leq t\}, \end{aligned}$$

in words, to a given t, η assigns the element of the grid which is just to the left of t.

2.2 The Euler-Maruyama Scheme

The Euler-Maruyama Scheme (EMS), is the natural extension of the Euler scheme for ODE to the stochastic setting. It was proposed by Maruyama in [57], and has the advantage of being applicable for a wide range of SDE. On the downside, as we see next, it rate of convergence is only 1/2.

The EMS $(\widehat{X}_{t_k}, k = 0, \dots, N)$ is given by

$$\widehat{X}_{t_k} = \begin{cases} X_0, \text{ for } k = 0, \\ \widehat{X}_{t_{k-1}} + b(\widehat{X}_{t_{k-1}})\Delta t + \sigma(\widehat{X}_{t_{k-1}})(W_{t_k} - W_{t_{k-1}}), \text{ for } k = 1, \dots, N. \end{cases}$$

In the following, we use the time continuous version of the EMS, $(\widehat{X}_t, 0 \leq t \leq T)$ given by

$$\widehat{X}_t = \hat{X}_{\eta(t)} + b(\hat{X}_{\eta(t)})(t - \eta(t)) + \sigma(\hat{X}_{\eta(t)})(W_t - W_{\eta(t)}), \tag{2.2.1}$$

from where is easy to conclude

$$\widehat{X}_t = X_0 + \int_0^t b(\hat{X}_{\eta(s)})ds + \int_0^t \sigma(\hat{X}_{\eta(s)})dW_s. \tag{2.2.2}$$

This last integral representation is quite useful to prove the following

Proposition 2.2.1. *Suppose that the functions b and σ are globally Lipschitz. Let $p \geq 1$ be an integer such that $\mathbb{E}\left[\|X_0\|^{2p}\right] < \infty$. Then there exists an increasing function K such that, for any integer $N \geq 1$,*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_t - \widehat{X}_t\|^{2p}\right] \leq K(T)\Delta t^p.$$

A sketch of the proof of this Proposition can be found in [75], and a complete proof in [30].

2.3 The Milstein Scheme

The Milstein scheme was introduced by Milstein in [61] for one dimensional SDEs having smooth diffusion coefficient. Introducing an appropriated correction term, this scheme has better convergence rate for the strong error than the classical Euler-Maruyama scheme (see Proposition 2.3.2 bellow).

With the previous notations, the Milstein Scheme (MS) $(\tilde{X}_{t_k}, k = 0, \dots, N)$ is given by

$$\tilde{X}_{t_k} = \begin{cases} X_0, \text{ for } k = 0, \\ \tilde{X}_{t_{k-1}} + b(\tilde{X}_{t_{k-1}})\Delta t + \sigma(\tilde{X}_{t_{k-1}})(W_{t_k} - W_{t_{k-1}}) \\ \qquad +\frac{1}{2}\sigma(\tilde{X}_{t_{k-1}})\sigma'(\tilde{X}_{t_{k-1}})\left[(W_{t_k} - W_{t_{k-1}})^2 - \Delta t\right], \text{ for } k = 1, \dots, N. \end{cases}$$

In the following, we use the time continuous version of the MS, $(\tilde{X}_t, 0 \leq t \leq T)$ satisfying

$$\begin{aligned}\tilde{X}_t = \tilde{X}_{\eta(t)} + b(\tilde{X}_{\eta(t)})(t-\eta(t)) + \sigma(\tilde{X}_{\eta(t)})(W_t - W_{\eta(t)}) \\ + \frac{1}{2}\sigma(\tilde{X}_{\eta(t)})\sigma'(\tilde{X}_{\eta(t)})\left[(W_t - W_{\eta(t)})^2 - (t-\eta(t))\right],\end{aligned} \tag{2.3.1}$$

Just as in the case of the EMS, the MS can also be represented as a semimartingale

$$\tilde{X}_t = X_0 + \int_0^t b(\tilde{X}_{\eta(s)})ds + \int_0^t \left[\sigma(\tilde{X}_{\eta(s)}) + \sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_s - W_{\eta(s)})\right] dW_s. \tag{2.3.2}$$

Using this integral representation and Itô's calculus, is quite standard to prove the following Lemma, which proof we omit.

Lemma 2.3.1. *Suppose that the functions b and σ are globally Lipschitz and have linear growth. Then for any $p \geq 1$*

$$\mathbb{E}(\sup_{0\leq t\leq T} \tilde{X}_t^{2p}) \leq C(1 + \mathbb{E}(X_0^{2p})), \tag{2.3.3}$$

and

$$\sup_{t\in[0,T]} \mathbb{E}\left[|\tilde{X}_t - \tilde{X}_{\eta(t)}|^{2p}\right] \leq K(T)\Delta t^p. \tag{2.3.4}$$

Proposition 2.3.2. *Suppose that the functions b and σ are twice continuously differentiable with bounded derivatives. Let $p \geq 1$ be an integer such that $\mathbb{E}|X_0|^{4p} < \infty$. Then there exists an increasing function K such that, for any integer $N \geq 1$,*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|X_t - \tilde{X}_t|^{2p}\right] \leq K(T)\Delta t^{2p}.$$

Since the next chapter deals with the extension of the Milstein scheme to a specific SDE, we include the proof of this proposition because it will guide us in the proof of the main result of the next chapter.

Proof. Let us consider $\tilde{\mathcal{E}}_t = X_t - \tilde{X}_t$. Then

$$d\tilde{\mathcal{E}}_t = [b(X_t) - b(\tilde{X}_{\eta(t)})]dt + \left[\sigma(X_t) - \sigma(\tilde{X}_{\eta(t)}) - \sigma(\tilde{X}_{\eta(t)})\sigma'(\tilde{X}_{\eta(t)})(W_t - W_{\eta(t)})\right]dW_t.$$

So, by Itô's Lemma, we have for all $p \geq 1$

$$\begin{aligned}\mathbb{E}(\tilde{\mathcal{E}}_t^{2p}) =& 2p\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-1}[b(X_s) - b(\tilde{X}_{\eta(s)})]\right) ds \\ &+ p(2p-1)\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-2}\left[\sigma(X_s) - \sigma(\tilde{X}_{\eta(s)}) - \sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_s - W_{\eta(s)})\right]^2\right) ds.\end{aligned}$$

If, in the second term on the right-hand side we add and subtract $\sigma(\tilde{X}_s)$, we obtain the following bound for the $2p$-th moment of the error:

$$\begin{aligned}\mathbb{E}(\tilde{\mathcal{E}}_t^{2p}) \leq & C\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-1}[b(X_s)-b(\tilde{X}_{\eta(s)})]\right)ds + C\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-2}\Big[\sigma(X_s)-\sigma(\tilde{X}_s)\Big]^2\right)ds \\ & + C\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-2}\Big[\sigma(\tilde{X}_s)-\sigma(\tilde{X}_{\eta(s)})-\sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_s-W_{\eta(s)})\Big]^2\right)ds.\end{aligned}$$

Applying the Lipschitz property in the second term of the right-hand side, and Young's inequality to the third, we obtain

$$\begin{aligned}\mathbb{E}(\tilde{\mathcal{E}}_t^{2p}) \leq & C\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-1}[b(X_s)-b(\tilde{X}_{\eta(s)})]\right)ds + C\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p}\right)ds \\ & + C\int_0^t \mathbb{E}\left(\Big[\sigma(\tilde{X}_s)-\sigma(\tilde{X}_{\eta(s)})-\sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_s-W_{\eta(s)})\Big]^{2p}\right)ds.\end{aligned}$$

Adding and subtracting $b(\tilde{X}_s)$, and using the Lipschitz property of b we get

$$\begin{aligned}\mathbb{E}(\tilde{\mathcal{E}}_t^{2p}) \leq & C\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-1}[b(\tilde{X}_s)-b(\tilde{X}_{\eta(s)})]\right)ds + C\int_0^t \mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p}\right)ds \\ & + C\int_0^t \mathbb{E}\left(\Big[\sigma(\tilde{X}_s)-\sigma(\tilde{X}_{\eta(s)})-\sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_s-W_{\eta(s)})\Big]^{2p}\right)ds.\end{aligned} \tag{2.3.5}$$

Let us assume for now

$$\mathbb{E}\left(\Big[\sigma(\tilde{X}_s)-\sigma(\tilde{X}_{\eta(s)})-\sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_s-W_{\eta(s)})\Big]^{2p}\right) \leq C\Delta t^{2p}, \tag{2.3.6}$$

and

$$\mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-1}[b(\tilde{X}_s)-b(\tilde{X}_{\eta(s)})]\right) \leq \sup_{u\leq s}\mathbb{E}(\tilde{\mathcal{E}}_u^{2p}) + C\Delta t^{2p}. \tag{2.3.7}$$

Then, the bound (2.3.5) becomes

$$\mathbb{E}(\tilde{\mathcal{E}}_t^{2p}) \leq C\int_0^t \sup_{u\leq s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)ds + C\Delta t^{2p}.$$

Since the right side of this last bound is increasing, we can take supremum on the left side to obtain

$$\sup_{u\leq t}\mathbb{E}(\tilde{\mathcal{E}}_u^{2p}) \leq C\int_0^t \sup_{u\leq s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)ds + C\Delta t^{2p},$$

from where, thanks to Gronwall's Lemma we can conclude

$$\sup_{u\leq t}\mathbb{E}(\tilde{\mathcal{E}}_u^{2p}) \leq C\Delta t^{2p}.$$

It only remains to prove (2.3.6) and (2.3.7). We start with (2.3.6). To do so, we apply the Itô formula to $\sigma(\tilde{X}_s)$. We have

$$\begin{aligned}\sigma(\tilde{X}_s) - \sigma(\tilde{X}_{\eta(s)}) &= \int_{\eta(s)}^{s} \sigma'(\tilde{X}_u) d\tilde{X}_u + \frac{1}{2}\int_{\eta(s)}^{s} \sigma''(\tilde{X}_u) d\langle \tilde{X}\rangle_u \\ &= \int_{\eta(s)}^{s} \sigma'(\tilde{X}_u) b(\tilde{X}_{\eta(s)}) du \\ &\quad + \int_{\eta(s)}^{s} \sigma'(\tilde{X}_u)\left[\sigma(\tilde{X}_{\eta(s)}) + \sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})\right] dW_u \\ &\quad + \frac{1}{2}\int_{\eta(s)}^{s} \sigma''(\tilde{X}_u)\left[\sigma(\tilde{X}_{\eta(s)}) + \sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})\right]^2 du.\end{aligned}$$

So,

$$\begin{aligned}&\mathbb{E}\Big(\Big[\sigma(\tilde{X}_s) - \sigma(\tilde{X}_{\eta(s)}) - \sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_s - W_{\eta(s)})\Big]^{2p}\Big) \\ &\quad \le C\mathbb{E}\left(\left[\int_{\eta(s)}^{s} \sigma(\tilde{X}_{\eta(s)})\left[\sigma'(\tilde{X}_u) - \sigma'(\tilde{X}_{\eta(s)})\right] dW_u\right]^{2p}\right) \\ &\quad\quad + C\mathbb{E}\left(\left[\int_{\eta(s)}^{s} \sigma'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)}) dW_u\right]^{2p}\right) \\ &\quad\quad + C\mathbb{E}\left(\left[\int_{\eta(s)}^{s} \sigma'(\tilde{X}_u) b(\tilde{X}_{\eta(s)}) du\right]^{2p}\right) \\ &\quad\quad + C\mathbb{E}\left(\left[\int_{\eta(s)}^{s} \sigma''(\tilde{X}_u)\left[\sigma(\tilde{X}_{\eta(s)}) + \sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})\right]^2 du\right]^{2p}\right).\end{aligned} \tag{2.3.8}$$

We will show now that each term in right side of the last bound is itself bounded by $C\Delta t^{2p}$. For the first and the second term we will use the Burkholder-Davis-Gundy inequality (see [52, p. 166]), which says that there exists a constant C_p depending only on p such that

$$\begin{aligned}&\mathbb{E}\left(\left[\int_{\eta(s)}^{s} \sigma(\tilde{X}_{\eta(s)})\left[\sigma'(\tilde{X}_u) - \sigma'(\tilde{X}_{\eta(s)})\right] dW_u\right]^{2p}\right) \\ &\qquad\qquad \le C_p\mathbb{E}\left(\left[\int_{\eta(s)}^{s} \sigma(\tilde{X}_{\eta(s)})^2\left[\sigma'(\tilde{X}_u) - \sigma'(\tilde{X}_{\eta(s)})\right]^2 du\right]^{p}\right).\end{aligned}$$

To continue with the computation recall that for $q \ge 1$, by Jensen's inequality

$$\left(\int_a^b f(s)ds\right)^p \le (a-b)^{p-1}\int_a^b f(s)^p ds, \tag{2.3.9}$$

so, thanks to the Lipschitz property of σ',

$$\begin{aligned}&\mathbb{E}\left(\left[\int_{\eta(s)}^{s} \sigma(\tilde{X}_{\eta(s)})\left[\sigma'(\tilde{X}_u) - \sigma'(\tilde{X}_{\eta(s)})\right] dW_u\right]^{2p}\right) \\ &\qquad\qquad \le C\Delta t^{p-1}\int_{\eta(s)}^{s} \mathbb{E}\left(\sigma(\tilde{X}_{\eta(s)})^{2p}\left[\tilde{X}_u - \tilde{X}_{\eta(s)}\right]^{2p}\right) du.\end{aligned}$$

From here, thanks to Cauchy Schwartz inequality, (2.3.3), and (2.3.4) we conclude

$$\mathbb{E}\left(\left[\int_{\eta(s)}^{s}\sigma(\tilde{X}_{\eta(s)})\left[\sigma'(\tilde{X}_u)-\sigma'(\tilde{X}_{\eta(s)})\right]dW_u\right]^{2p}\right)\leq C\Delta t^{2p}.$$

With the same arguments we used to bound the first term in (2.3.8), we can tackle the second one

$$\begin{aligned}\mathbb{E}&\left(\left[\int_{\eta(s)}^{s}\sigma'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_u-W_{\eta(s)})dW_u\right]^{2p}\right)\\&\leq\mathbb{E}\left(\left[\int_{\eta(s)}^{s}\sigma'(\tilde{X}_u)^2\sigma(\tilde{X}_{\eta(s)})^2\sigma'(\tilde{X}_{\eta(s)})^2(W_u-W_{\eta(s)})^2du\right]^{p}\right)\\&\leq C\Delta t^{p-1}\int_{\eta(s)}^{s}\mathbb{E}\left(\sigma(\tilde{X}_{\eta(s)})^{2p}(W_u-W_{\eta(s)})^{2p}\right)du,\end{aligned}$$

where, in the last line, we have used that σ' is bounded. To finish this computation we have to recall that the increments of the brownian motion are independent and normally distributed, so

$$\mathbb{E}\left(\left[\int_{\eta(s)}^{s}\sigma'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(s)})\sigma'(\tilde{X}_{\eta(s)})(W_u-W_{\eta(s)})dW_u\right]^{2p}\right)\leq C\Delta t^{2p}.$$

To complete the proof of (2.3.6), it only remains to bound the third and fourth terms in (2.3.8). To this end, we just have to apply (2.3.9) and recall the finiteness of the moments of the MS given by (2.3.3) in Lemma 2.3.1.

Now we are going to prove (2.3.7). To do so, first we introduce the notation

$$\tilde{\Sigma}_u=\sigma(X_u)-\sigma(\tilde{X}_{\eta(u)})[1+\sigma'(\tilde{X}_{\eta(s)})(W_u-W_{\eta(s)})],$$

for the difference of the diffusion coefficient of X and $\tilde{X}$.

If we apply Itô's Lemma to the function $f(x,y)=x^{2p-1}(b(y_0)-b(y))$ and the vector of semimartingales $(\tilde{\mathcal{E}},\tilde{X})$ between $\eta(s)$ and s we obtain

$$\begin{aligned}\mathbb{E}&\left(\tilde{\mathcal{E}}_s^{2p-1}[b(\tilde{X}_{\eta(s)})-b(\tilde{X}_s)]\right)\\&=C_p\mathbb{E}\int_{\eta(s)}^{s}\tilde{\mathcal{E}}_u^{2p-1}\left[b'(\tilde{X}_u)b(\tilde{X}_u)+\frac{1}{2}b''(\tilde{X}_u)\sigma^2(\tilde{X}_{\eta(s)})[1+\sigma'(\tilde{X}_{\eta(s)})(W_u-W_{\eta(s)})]\right]du\\&\quad+C_p\mathbb{E}\int_{\eta(s)}^{s}\tilde{\mathcal{E}}_u^{2p-2}[b(X_u)-b(\tilde{X}_{\eta(s)})]\left[b(\tilde{X}_u)-b(\tilde{X}_{\eta(s)})\right]du\\&\quad+C_p\mathbb{E}\int_{\eta(s)}^{s}\tilde{\mathcal{E}}_u^{2p-3}[b(\tilde{X}_{\eta(s)})-b(\tilde{X}_u)]\tilde{\Sigma}_u^2du\\&\quad+C_p\mathbb{E}\int_{\eta(s)}^{s}\tilde{\mathcal{E}}_u^{2p-2}\tilde{\Sigma}_u\left(b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1+\sigma'(\tilde{X}_{\eta(s)})(W_u-W_{\eta(s)})]\right)du\\&=:I_1+I_2+I_3+I_4.\end{aligned}$$

By the finiteness of the moments of $\tilde{X}$, the boundedness of b' and b'', the linear growth of b and σ, and the finiteness of the moments of a normal random variable, we can apply Holder's inequality and get

$$\begin{aligned} I_1 \leq & \int_{\eta(s)}^{s} \left[\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p} \\ & \times \left[\mathbb{E}\left(\left((b'\cdot b)(\tilde{X}_u) + \frac{1}{2}b''(\tilde{X}_u)\sigma^2(\tilde{X}_{\eta(s)})\left[1+\sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})\right]^2\right)^{2p}\right)\right]^{1/2p} du \\ \leq & \, C\left[\sup_{u\leq s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p} \Delta t. \end{aligned}$$

For the bound of I_2, if we add and subtract $b(\tilde{X}_u)$ inside of the first parenthesis, since b is Lipschitz, we have

$$\begin{aligned} I_2 \leq & C\int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p-1}|\tilde{X}_u - \tilde{X}_{\eta(s)}|\right) du + C\int_{\eta(s)}^{s} \mathbb{E}\left(\mathcal{E}_u^{2p-2}|\tilde{X}_u - \tilde{X}_{\eta(u)}|^2\right) du \\ \leq & C\left[\sup_{u\leq s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p} \Delta t^{3/2} + C\left[\sup_{u\leq s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/p} \Delta t^2. \end{aligned}$$

To bound I_3, we notice that

$$\tilde{\Sigma}_u^2 \leq 2\left(\sigma(X_u) - \sigma(\tilde{X}_u)\right)^2 + 2\left(\sigma(\tilde{X}_u) - \sigma(\tilde{X}_{\eta(u)}) - \sigma(\tilde{X}_{\eta(u)})\sigma'(\tilde{X}_{\eta(u)})(W_u - W_{\eta(u)})\right)^2,$$

and then thanks to the Lipschitz property of b and σ, we obtain

$$\begin{aligned} I_3 \leq & \, C\int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p-1}|\tilde{X}_{\eta(s)} - \tilde{X}_u|\right) du \\ & + C\int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p-3}|\tilde{X}_{\eta(s)} - \tilde{X}_u|\left(\sigma(\tilde{X}_u) - \sigma(\tilde{X}_{\eta(u)}) - \sigma(\tilde{X}_{\eta(u)})\sigma'(\tilde{X}_{\eta(u)})(W_u - W_{\eta(u)})\right)^2\right) du. \end{aligned}$$

For the first term in the right-hand side we have

$$\begin{aligned} \int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p-1}|\tilde{X}_{\eta(s)} - \tilde{X}_u|\right) du \leq & \int_{\eta(s)}^{s} \left[\mathbb{E}\left(\mathcal{E}_u^{2p}\right)\right]^{1-\frac{1}{2p}} \left[\mathbb{E}\left(|\tilde{X}_{\eta(s)} - \tilde{X}_u|^{2p}\right)\right]^{\frac{1}{2p}} \\ \leq & C\left[\sup_{u\leq s}\mathbb{E}\left(\mathcal{E}_u^{2p}\right)\right]^{1-1/2p} \Delta t^{3/2}, \end{aligned}$$

due to the bound (2.3.4). For the second term, we apply Hölder's inequality to get

$$\begin{aligned}
&\int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p-3}|\tilde{X}_{\eta(s)} - \tilde{X}_u|\left(\sigma(\tilde{X}_u) - \sigma(\tilde{X}_{\eta(u)}) - \sigma(\tilde{X}_{\eta(u)})\sigma'(\tilde{X}_{\eta(u)})(W_u - W_{\eta(u)})\right)^2\right) du \\
&\quad \leq \int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)^{1-3/2p} \mathbb{E}\left(|\tilde{X}_{\eta(s)} - \tilde{X}_u|^{2p}\right)^{1/2p} \\
&\qquad \times \mathbb{E}\left(\left(\sigma(\tilde{X}_u) - \sigma(\tilde{X}_{\eta(u)}) - \sigma(\tilde{X}_{\eta(u)})\sigma'(\tilde{X}_{\eta(u)})(W_u - W_{\eta(u)})\right)^{2p}\right)^{1/p} du \\
&\quad \leq C\left[\sup_{u\leq s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-3/2p} \Delta t^{7/2}.
\end{aligned}$$

To summarize

$$I_3 \leq C\left[\sup_{u\leq s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p} \Delta t^{3/2} + \left[\sup_{u\leq s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-3/2p} \Delta t^{7/2}.$$

Now we bound I_4

$$\begin{aligned}
I_4 &\leq \int_{\eta(s)}^{s} \mathbb{E}\left|\tilde{\mathcal{E}}_u^{2p-2}\tilde{\Sigma}_u\left(b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1 + \sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})]\right)\right| du \\
&\leq C\int_{\eta(s)}^{s} \mathbb{E}\left|\tilde{\mathcal{E}}_u^{2p-1}\left(b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1 + \sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})]\right)\right| du \\
&\quad + \int_{\eta(s)}^{s} \mathbb{E}\left|\tilde{\mathcal{E}}_u^{2p-2}\left(\sigma(\tilde{X}_u) - \sigma(\tilde{X}_{\eta(u)}) - \sigma(\tilde{X}_{\eta(u)})\sigma'(\tilde{X}_{\eta(u)})(W_u - W_{\eta(u)})\right)\right. \\
&\qquad\qquad \left. \times \left(b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1 + \sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})]\right)\right| du.
\end{aligned}$$

Again, by Holder's inequality the first term in the right side is bounded by

$$\begin{aligned}
&\int_{\eta(s)}^{s} \mathbb{E}\left|\tilde{\mathcal{E}}_u^{2p-1}\left(b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1 + \sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})]\right)\right| du \\
&\quad \leq \int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)^{1-1/2p} \mathbb{E}\left(\left[b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1 + \sigma'(\tilde{X}_{\eta(s)})(W_u - W_{\eta(s)})]\right]^{2p}\right)^{1/2p} du \\
&\quad \leq C\left[\sup_{u\leq s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p} \Delta t,
\end{aligned}$$

thanks to the boundedness of b' and σ', the linear growth of σ and the finiteness of the moments of a normal random variable. For the second term in the right side of the bound for I_4, again by Holder's inequality

$$\begin{aligned}
&\int_{\eta(s)}^{s} \mathbb{E}\left|\tilde{\mathcal{E}}_u^{2p-2}\left(\sigma(\tilde{X}_u)-\sigma(\tilde{X}_{\eta(u)})-\sigma(\tilde{X}_{\eta(u)})\sigma'(\tilde{X}_{\eta(u)})(W_u-W_{\eta(u)})\right)\right. \\
&\qquad\qquad \left.\times\left(b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1+\sigma'(\tilde{X}_{\eta(s)})(W_u-W_{\eta(s)})]\right)\right| du \\
&\quad\le \int_{\eta(s)}^{s} \mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)^{1-1/p} \mathbb{E}\left(\left[\sigma(\tilde{X}_u)-\sigma(\tilde{X}_{\eta(u)})-\sigma(\tilde{X}_{\eta(u)})\sigma'(\tilde{X}_{\eta(u)})(W_u-W_{\eta(u)})\right]^{2p}\right)^{1/2p} \\
&\qquad\qquad \times\mathbb{E}\left(\left[b'(\tilde{X}_u)\sigma(\tilde{X}_{\eta(u)})[1+\sigma'(\tilde{X}_{\eta(s)})(W_u-W_{\eta(s)})]^{2p}\right]\right)^{1/2p} du \\
&\quad\le C\left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/p}\Delta t^2,
\end{aligned}$$

So

$$I_4 \le C\left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p}\Delta t + C\left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/p}\Delta t^2.$$

Putting all the last calculations in we find

$$\begin{aligned}
\mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-1}[b(\tilde{X}_{\eta(s)})-b(\tilde{X}_s)]\right) &\le C\left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p}\Delta t + C\left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/2p}\Delta t^{3/2} \\
&\quad + C\left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-1/p}\Delta t^2 + \left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right]^{1-3/2p}\Delta t^{7/2}.
\end{aligned}$$

Applying Young's Inequality in all terms in the right, we get

$$\mathbb{E}\left(\tilde{\mathcal{E}}_s^{2p-1}[b(\tilde{X}_{\eta(s)})-b(\tilde{X}_s)]\right) \le C\left[\sup_{u\le s}\mathbb{E}\left(\tilde{\mathcal{E}}_u^{2p}\right)\right] + C\Delta t^{2p}.$$

□

Chapter 3

Strong convergence of the symmetrized Milstein scheme for some CEV-like SDEs

Based on a joint work with
Mireille Bossy [16], accepted for
publication in Bernoulli Journal.

3.1 Introduction and main result

As we saw in the previous chapter, under suitable hypotheses the Milstein Scheme has a better convergence rate than the classical Euler Maruyama Scheme. This well-know fact produces remarks on blogs, internet forums, and software packages that sometimes recommend to use the Milstein scheme for constant elasticity of variance (CEV) models in finance, or its extension with stochastic volatility as SABR model, (see e.g Delbaen and Shirakawa [24] and Lions and Musiela [55] for a discussion on the (weak) existence of such models); CEV are popular stochastic volatility models of the form

$$dX_t = \mu X_t dt + \sigma X_t^{\gamma} dW_t$$

with $0 < \gamma < 1$. But the interesting fact in this story is that the rate of convergence of the Milstein scheme, for such family of processes with $0 < \gamma < 1$ is not yet well studied, to the best of our knowledge.

In this chapter we establish a rate of convergence result for a symmetrized version of the Milstein scheme applied to the solution of the one dimensional SDE

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma |X_s|^{\alpha} dW_s, \tag{3.1.1}$$

where $x_0 > 0$, $\sigma > 0$ and $\frac{1}{2} \leq \alpha < 1$. Of course Equation (3.1.1) does not satisfies the hypothesis to apply the classical result of Milstein [61]. In particular, the diffusion coefficient is only Hölder continuous whereas the classical hypothesis is to have a $\mathcal{C}^2$ diffusion coefficient.

The main picture of our convergence rate result is that Milstein scheme stays of order one in the case of Equation (3.1.1), but some attention must be paid to the values of $b(0)$, α and σ.

There exist in the literature other strategies for the discretization of the solution to (3.1.1). There are some results based on the Lamperti transformation of the equation, for example, by Alfonsi [1,3], and by Chassagneux, Jacquier and Mihaylov [19]. And also, there some are results where the equation (3.1.1) is discretized directly, as in Berkaoui, Bossy and Diop [10] or in Kahl and Jackël [50]. In the numerical experiments section, we compare the symmetrized Milstein scheme with a selection of schemes proposed in the aforementioned references. We also experiment the symmetrized Milstein scheme in a multilevel Monte Carlo application and we compare with other schemes.

In the whole chapter, we work under the following basis-hypothesis:

Hypothesis 3.1.1. *The power parameter α in the diffusion coefficient of Equation (3.1.1) belongs to $[\frac{1}{2}, 1)$. The drift coefficient b is Lipschitz with constant $K > 0$, and is such that $b(0) > 0$.*

Hypothesis 3.1.1 is a classical assumption to ensure a unique strong solution valued in $\mathbb{R}^+$. We assume it in all the forthcoming results of the chapter, without recall it explicitly. To state the convergence result (see Theorem 3.1.6), another Hypothesis 3.1.5 will be added and discussed, that in particular constrains the values α, $b(0)$ and σ.

3.1.1 The symmetrized Milstein scheme

To complete our task we follow the ideas of Berkaoui, Bossy and Diop in [10] who analyze the rate of convergence of the strong error for the symmetrized Euler scheme applied to Equation (3.1.1). Although, whereas they utilize an argument of change of time, we consider first a weighted $L^p(\Omega)$-error for which we prove a convergence result, and then we utilize this result to prove the convergence of the actual $L^p(\Omega)$-error.

We consider $x_0 > 0$, $T > 0$, and $N \in \mathbb{N}$. We define the constant step size $\Delta t = T/N$ and $t_k = k\Delta t$. Over this discretization of the interval $[0, T]$ we define the Symmetrized Milstein Scheme (SMS) $(\overline{X}_{t_k}, k = 0, \ldots, N)$ by

$$\overline{X}_{t_k} = \begin{cases} x_0, \text{ for } k = 0, \\ \left| \overline{X}_{t_{k-1}} + b(\overline{X}_{t_{k-1}})\Delta t + \sigma \overline{X}^{\alpha}_{t_{k-1}}(W_{t_k} - W_{t_{k-1}}) + \dfrac{\alpha\sigma^2}{2}\overline{X}^{2\alpha-1}_{t_{k-1}}\left[(W_{t_k} - W_{t_{k-1}})^2 - \Delta t\right] \right|, \\ \qquad \text{for } k = 1, \ldots, N. \end{cases}$$

In the following, we use the time continuous version of the SMS, $(\overline{X}_t, 0 \le t \le T)$ satisfying

$$\begin{aligned} \overline{X}_t = \Big| \overline{X}_{\eta(t)} &+ b(\overline{X}_{\eta(t)})(t - \eta(t)) + \sigma \overline{X}^{\alpha}_{\eta(t)}(W_t - W_{\eta(t)}) \\ &+ \frac{\alpha\sigma^2}{2}\overline{X}^{2\alpha-1}_{\eta(t)}\left[(W_t - W_{\eta(t)})^2 - (t - \eta(t))\right] \Big|, \end{aligned} \tag{3.1.2}$$

where $\eta(t) = \sup_{k \in \{1,\ldots,N\}}\{t_k : t_k \le t\}$. We also introduce the increment process $(\overline{Z}_t, 0 \le t \le T)$ defined by

$$\begin{aligned} \overline{Z}_t = \overline{X}_{\eta(t)} &+ b(\overline{X}_{\eta(t)})(t - \eta(t)) + \sigma \overline{X}^{\alpha}_{\eta(t)}(W_t - W_{\eta(t)}) \\ &+ \frac{\alpha\sigma^2}{2}\overline{X}^{2\alpha-1}_{\eta(t)}\left[(W_t - W_{\eta(t)})^2 - (t - \eta(t))\right], \end{aligned} \tag{3.1.3}$$

so that $\overline{X}_t = |\overline{Z}_t|$. Thanks to Tanaka's Formula, the semi-martingale decomposition of $\overline{X}_t$ is given by

$$\begin{aligned} \overline{X}_t = x_0 &+ \int_0^t \operatorname{sgn}(\overline{Z}_s) b(\overline{X}_{\eta(s)})ds + L^0_t(\overline{X}) \\ &+ \int_0^t \operatorname{sgn}(\overline{Z}_s)\left[\sigma \overline{X}^{\alpha}_{\eta(s)} + \alpha\sigma^2 \overline{X}^{2\alpha-1}_{\eta(s)}(W_s - W_{\eta(s)})\right] dW_s \end{aligned} \tag{3.1.4}$$

where $\operatorname{sgn}(x) = 1 - 2\mathbb{1}_{[x \le 0]}$.

Moment upper bound estimations for X and $\overline{X}$

We summarize some facts about the process $(X_t, 0 \le t \le T)$, the proofs of which can be found in Bossy and Diop [14].

Lemma 3.1.2. *For any $q \ge 1$, there exists a positive constant C depending on q, but also on the parameters $b(0)$, K, σ, α and T such that, for any $x_0 > 0$,*

$$\mathbb{E}\Big[\sup_{0 \le t \le T} X_t^{2q}\Big] \le C(1 + x_0^{2q}). \tag{3.1.5}$$

When $\frac{1}{2} < \alpha < 1$, for any $q > 0$,

$$\sup_{0 \le t \le T} \mathbb{E}\left[X_t^{-q}\right] \le C(1 + x_0^{-q}). \tag{3.1.6}$$

When $\alpha = \frac{1}{2}$, for any q such that $1 < q < \frac{2b(0)}{\sigma^2} - 1$,

$$\sup_{0 \le t \le T} \mathbb{E}\left[X_t^{-q}\right] \le C x_0^{-q}. \tag{3.1.7}$$

Lemma 3.1.3. *Let $(X_t, 0 \leq t \leq T)$ be the solution of (3.1.1) with $\frac{1}{2} < \alpha < 1$. For all $\mu \geq 0$, there exists a positive constant $C(T, \mu)$, increasing in μ and T, depending also on b, σ, α and x_0 such that*

$$\mathbb{E} \exp \left(\mu \int_0^T \frac{ds}{X_s^{2(1-\alpha)}} \right) \leq C(T, \mu). \tag{3.1.8}$$

When $\alpha = \frac{1}{2}$, the inequality (3.1.8) holds if $b(0) > \frac{\sigma^2}{2}$ and $\mu \leq \frac{\sigma^2}{8}(\frac{2b(0)}{\sigma^2} - 1)^2$.

Notice that the condition $b(0) > \sigma^2/2$ is also imposed by the Feller test in the case $\alpha = \frac{1}{2}$ for the strict positivity of X, that allows to rewrite Equation (3.1.1) as

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma \sqrt{X_s} dW_s.$$

Using the semimartingale representation (3.1.4), we prove the following Lemma regarding the existence of moments of any order for $\overline{X}_t$.

Lemma 3.1.4. *For any $q \geq 1$, there exists a positive constant C depending on q, but also on the parameters $b(0)$, K, σ, α and T such that for any $x_0 > 0$,*

$$\mathbb{E}\Big[\sup_{t \in [0,T]} \overline{X}_t^{2q} \Big] \leq C(1 + x_0^{2q}).$$

The proof of this lemma is based on the Lipschitz property of b and classical combination of Itô formula and Young Inequality. For the sake of completeness, we give a short proof in the Appendix.

3.1.2 Strong rate of convergence

The main result of this works is the strong convergence at rate one of the SMS $\overline{X}$ to the exact process X. The convergence holds in L^p for $p \geq 1$. To state it, we add to Hypothesis 3.1.1 the following.
For any x in $\mathbb{R}^+$, we denote $\lceil x \rceil$ the rounded up integer.

Hypothesis 3.1.5.

(i) Let $p \geq 1$. To control the $L^p(\Omega)$-norm of the error, if $\alpha > \frac{1}{2}$ we assume $b(0) > 2\alpha(1 - \alpha)^2\sigma^2$. Whereas for $\alpha = \frac{1}{2}$ we assume $b(0) > 3(2\lceil p \vee 2 \rceil + 1)\sigma^2/2$.

(ii) The drift coefficient b is of class $\mathcal{C}^2(\mathbb{R})$, and b'' has polynomial growth.

We now state our main theorem. To lighten the notation, we consider for $\alpha \in (\frac{1}{2}, 1)$

$$\begin{cases} b_\sigma(\alpha) := b(0) - 2(1-\alpha)^2 \alpha \sigma^2, \\ K(\alpha) := K + \dfrac{\alpha\sigma^2}{2}(2\alpha - 1)[2(1-\alpha)]^{-\frac{2(1-\alpha)}{2\alpha-1}} \end{cases} \tag{3.1.9}$$

and we extend this definitions to $\alpha = \frac{1}{2}$ taking limits. So, $b_\sigma(\frac{1}{2}) = \lim_{\alpha\to\frac{1}{2}} b_\sigma(\alpha) = b(0)-\sigma^2/4$, and $K(\frac{1}{2}) = \lim_{\alpha\to\frac{1}{2}} K(\alpha) = K$. Notice that $\lim_{\alpha\to 1} K(\alpha) = K + \sigma^2/2$, and since $K(\alpha)$ is continuous on $(\frac{1}{2}, 1)$, we have that $K(\alpha)$ is bounded. This is especially important in the definition of $\Delta_{\max}(\alpha)$ bellow, because tells us that $\alpha \mapsto \Delta_{\max}(\alpha)$ is strictly positive and bounded on $[\frac{1}{2}, 1)$.

Theorem 3.1.6. *Assume Hypotheses 3.1.1 and 3.1.5. Define a maximum step size $\Delta_{\max}(\alpha)$ as*

$$\Delta_{\max}(\alpha) = \frac{x_0}{(1-\sqrt{\alpha})b_\sigma(\alpha)} \wedge \begin{cases} \dfrac{1}{4\alpha K(\alpha)}, & \text{for } \alpha \in (\frac{1}{2}, 1) \\ \dfrac{1}{4K} \wedge x_0, & \text{for } \alpha = \frac{1}{2}. \end{cases} \tag{3.1.10}$$

Let $(X_t, 0 \le t \le T)$ be the process defined on (3.1.1) and $(\overline{X}_t, 0 \le t \le T)$ the symmetrized Milstein scheme given in (3.1.2). Then for any $p \ge 1$ that allows Hypotheses 3.1.5, there exists a constant C depending on p, T, $b(0)$, α, σ, K, and x_0, but not on Δt, such that for all $\Delta t \le \Delta_{\max}(\alpha)$,

$$\sup_{0\le t\le T} \left(\mathbb{E}\left[|X_t - \overline{X}_t|^p\right]\right)^{\frac{1}{p}} \le C\Delta t. \tag{3.1.11}$$

About Hypothesis 3.1.5. Notice that for $\alpha > \frac{1}{2}$, Assumption (i) does not depend on p and becomes easier to fulfill as α increases. On the other hand, for $\alpha = \frac{1}{2}$, Assumption (i) depends on p in a unpleasant manner. However, as we will see later in Section 3.4 (see Table 3.4.1), this kind of dependence in p is expected, and similar conditions are asked in the literature for other approximation schemes in order to obtain similar rate of convergence results.

Also, notice that (i) is a sufficient condition: in the numerical experiments we still observe a rate of convergence of order one for parameters that do not satisfy it, but we also observe that for parameters such that $b(0) \ll \sigma^2$, although the convergence occurs, it does in a sublinear fashion.

On the other hand, Assumption (ii) is the classical requirement for the strong convergence of the Milstein scheme. As we will see later in the proof of the main theorem, with the help of the ItÃť formula, this hypothesis let us conclude that

$$\mathbb{E}\left[|X_s - \overline{X}_s|^{2p-1}\big(b(X_{\eta(s)}) - b(X_s)\big)\right] \le C\left(\sup_{u\le s}\mathbb{E}\left[|X_u - \overline{X}_u|^{2p}\right] + \Delta t^{2p}\right)$$

instead of

$$\mathbb{E}\left[|X_s - \overline{X}_s|^{2p-1}\big(b(X_{\eta(s)}) - b(X_s)\big)\right] \le C\left(\sup_{u\le s}\mathbb{E}\left[|X_u - \overline{X}_u|^{2p}\right] + \Delta t^{p}\right),$$

which is the classical bound obtained for the Euler-Maruyama scheme under a Lipschitz condition for a drift b.

The rest of this chapter is organized as follow. In Section 3.2 we state some preliminary results on the scheme which will be building blocks in the proof of Theorem 3.1.6. Section 3.3 is devoted to the proof of the convergence rate. The main idea is first to introduce a weight process in the $L^{4p}(\Omega)$-error. Wet get the rate of convergence for this weighted error process, and we use this intermediate bound to control the $L^{2p}(\Omega)$-error, from where we finally control the $L^p(\Omega)$-error. Also, as a byproduct we obtain the order one convergence of the Projected Milstein Scheme (see 3.3.3). In section 3.4, we display some numerical experiments to show the effectiveness of the theoretical rate of convergence of the scheme, but also to test Hypotheses 3.1.5-(i) on a set of parameters. In this section we also shows how the inclusion of the SMS in a Multilevel Monte Carlo framework could help to optimize the computational time of weak approximation of assets valuation. Finally, in Section 3.5 we present the proof of the preliminary results on the scheme.

3.2 Some preliminary results for $\overline{X}$

This short section is devoted to state some results about the behavior of $\overline{X}$, their proofs are postponed to Section 3.5. All these results hold under Hypothesis 3.1.5-(i) which is in fact stronger than what we need here. So, we present the next lemmas with their minimal hypotheses (still assuming Hypothesis 3.1.1).

Lemma 3.2.1 (Local error). *For any $x_0 > 0$, for any $p \geq 1$, there exists a positive constant C, depending on p, T, the parameters of the model $b(0)$, K, σ, α, but not on Δt such that*

$$\sup_{0\leq t\leq T} \mathbb{E}\left[|\overline{X}_t - \overline{X}_{\eta(t)}|^{2p}\right] \leq C\Delta t^p.$$

By construction the scheme $\overline{X}$ is nonnegative, but a key point of the convergence proof resides in the analysis of the behavior of $\overline{X}$ or $\overline{Z}$ visiting the point 0. The next Lemma shows that although $\overline{Z}_t$ is not always positive, the probability of $\overline{Z}_t$ being negative is actually very small under suitable hypotheses.

Lemma 3.2.2. *For $\alpha \in [\frac{1}{2}, 1)$, if $b(0) > 2\alpha(1-\alpha)^2\sigma^2$, and $\Delta t \leq \frac{1}{2K(\alpha)}$, then there exists a positive constant γ, depending on the parameters of the model, but not on Δt, such that*

$$\sup_{k=0,\dots,N-1} \mathbb{P}\left(\inf_{t_k < s \leq t_{k+1}} \overline{Z}_s \leq 0\right) \leq \exp\left(-\frac{\gamma}{\Delta t}\right). \tag{3.2.1}$$

In particular,

$$\sup_{0\leq t\leq T} \mathbb{P}\left(\overline{Z}_t \leq 0\right) \leq C\exp\left(-\frac{\gamma}{\Delta t}\right). \tag{3.2.2}$$

To prove Lemma 3.2.2, it is necessary to establish before the following one, which although technical, gives some intuition about the difference between the SMS and the Symmetrized Euler scheme presented in [10].

Lemma 3.2.3. *For $\alpha \in [\frac{1}{2}, 1)$, if $b(0) > 2\alpha(1-\alpha)^2\sigma^2$, we set $\bar{x}(\alpha) = \frac{b_\sigma(\alpha)}{K(\alpha)} > 0$. Then for all $t \in [0, T]$, and for all $\rho \in (0, 1]$,*

$$\mathbb{P}\left[\overline{Z}_t \leq (1-\rho) b_\sigma(\alpha)\Delta t,\ \overline{X}_{\eta(t)} < \rho \bar{x}(\alpha)\right] = 0.$$

Roughly speaking, from this lemma we see that when $\overline{Z}_{\eta(t)} > 0$, $\overline{Z}_t$ becomes negative only when

$$|\overline{Z}_t - \overline{Z}_{\eta(t)}| > \rho \bar{x}(\alpha).$$

But observe that only the left-hand side of this inequality depends on Δt, and its expectation decreases to zero proportionally to $\sqrt{\Delta t}$, according to Lemma 3.2.1.

Now imposing Δt small enough, we prove an explicit bound for the local time moment of $\overline{X}$.

Lemma 3.2.4. *For $\alpha \in [\frac{1}{2}, 1)$, if $b(0) > 2\alpha(1-\alpha)^2\sigma^2$ and $\Delta t \leq \frac{1}{2K(\alpha)} \wedge \frac{x_0}{(1-\sqrt{\alpha})b_\sigma(\alpha)}$, then there exist positive constants C and $\gamma > 0$ depending on α, $b(0)$, K, and σ but not in Δt such that*

$$\mathbb{E}\left(L_T^0(\overline{X})^2\right) \leq C \frac{1}{\sqrt{\Delta t}} \exp\left(\frac{-\gamma}{2\Delta t}\right).$$

We end this section with another key preliminary result, which is the convergence rate of order 1 for the *corrected* local error. Although the classical local error is of order 1/2, as stated in Lemma 3.2.1, the local error seen by the diffusion coefficient function, corrected with the Milstein term stays of order 1.

Lemma 3.2.5 (Corrected local error process)**.** *Let us fix $p \geq 1$, and $\alpha \in [\frac{1}{2}, 1)$. For $\alpha > \frac{1}{2}$, assume $b(0) > 2\alpha(1-\alpha)^2\sigma^2$, whereas for $\alpha = \frac{1}{2}$, assume $b(0) > 3(2p+1)\sigma^2/2$. Then, there exists $C > 0$, depending on the parameters of the model but not in Δt, such that for all $\Delta t \leq \Delta_{\max}(\alpha)$, the Corrected Local Error satisfies*

$$\sup_{0 \leq t \leq T} \mathbb{E}\left[\left|\sigma\overline{X}_t^\alpha - \sigma\overline{X}_{\eta(t)}^\alpha - \alpha\sigma^2\overline{X}_{\eta(t)}^{2\alpha-1}(W_t - W_{\eta(t)})\right|^{2p}\right] \leq C\Delta t^{2p}.$$

3.3 Proof of the main Theorem 3.1.6

The proof of Theorem 3.1.6 is built in several steps. First, we work with the $L^{2p}(\Omega)$-norm of the error, for $p \geq 1$, then at the last step of the proof we go back to the $L^p(\Omega)$-norm for $p \geq 1$.

In what follows we denote

$$\mathcal{E}_t := \overline{X}_t - X_t$$

and

$$\Sigma_t := \operatorname{sgn}(\overline{Z}_t)\left[\sigma\overline{X}_{\eta(t)}^\alpha + \alpha\sigma^2\overline{X}_{\eta(t)}^{2\alpha-1}(W_t - W_{\eta(t)})\right] - \sigma X_t^\alpha$$

so that

$$d\mathcal{E}_t = \big(\operatorname{sgn}(\overline{Z}_t)b(\overline{X}_{\eta(t)}) - b(X_t)\big)\,dt + dL^0_t(\overline{X}) + \Sigma_t dW_t.$$

Also, to make the notation lighter, we will denote the Corrected Local Error by

$$D_t(\overline{X}) := \sigma\overline{X}^{\alpha}_t - \sigma\overline{X}^{\alpha}_{\eta(t)} - \alpha\sigma^2\overline{X}^{2\alpha-1}_{\eta(t)}(W_t - W_{\eta(t)}).$$

3.3.1 The Weighted Error

Before to prove the main theorem, we establish in the Proposition 3.3.1 the convergence of a weighted error. For $p \geq 1$, let us consider $(\beta_t, 0 \leq t \leq T)$, defined by

$$\beta_t = 2p\|b'\|_\infty + 2p(4p-1) + \frac{8\alpha^2 p(4p-1)\sigma^2}{X_t^{2(1-\alpha)}}, \tag{3.3.1}$$

and the Weight Process $(\Gamma_t, 0 \leq t \leq T)$ defined by

$$\Gamma_t = \exp\left(-\int_0^t \beta_s ds\right). \tag{3.3.2}$$

The Weight Process is adapted, almost surely positive, and bounded by 1. Its paths are non increasing and hence has bounded variation, and also satisfies

$$d\Gamma_t = -\beta_t\Gamma_t dt.$$

The process $(\Gamma_t)_{t\geq 0}$ can be seen as an integrating factor in the sense of linear first order ODE (see for example [73]). When we apply the Itô's Lemma to $\Gamma_t^2\mathcal{E}_t^{4p}$, instead of $\mathcal{E}_t^{4p}$ alone, we can remove a very annoying term that appears in the righthand side upper bound (see the proof of Lemma 3.3.3). The exponential weight in a $L^p(\Omega)$-norm is a useful tool to obtain a priori upper bound. As an example, for the existence and uniqueness of the solution of a Backward SDE, it is introduced a norm with exponential weight, such that the operator associated to the BSDE is contractive under this new norm (see proof of Theorem 1.2 in [68]). In the same way, here we introduce this exponential weight to get the following a priori error bound, which will allow us to prove Theorem 3.1.6.

Proposition 3.3.1 (Weighted Error)**.** *Under the hypotheses of Theorem 3.1.6, for $p \geq 1$ and $\alpha \in [\frac{1}{2}, 1)$, there exists a constant C not depending on Δt such that for all $\Delta t \leq \Delta_{\max}(\alpha)$*

$$\sup_{0\leq t\leq T} \mathbb{E}\left[\Gamma_t^2\mathcal{E}_t^{4p}\right] \leq C\Delta t^{4p}. \tag{3.3.3}$$

Remark 3.3.2.

(i) Since we are going to work first with the $L^{2p}(\Omega)$-norm of the error, when $\alpha = \frac{1}{2}$, Hypothesis 3.1.5-(i) becomes

$$b(0) > \frac{3(2[2p \vee 2] + 1)\sigma^2}{2} = \frac{3(4p+1)\sigma^2}{2},$$

in particular

$$\left(\frac{2b(0)}{\sigma^2} - 1\right) > 12p + 2. \tag{3.3.4}$$

(ii) From Lemma 3.1.2, the process β has polynomial moments of any order for $\alpha > \frac{1}{2}$, and when $\alpha = \frac{1}{2}$, there exists C such that $\mathbb{E}[\beta_t^q] < C$, for all $1 < q < 2b(0)/\sigma^2 - 1$. From the previous point in this Remark, it follows that the process β has moments at least up to order $12p + 2$.

(iii) From the definition β in (3.3.1), and due to Lemma 3.1.3, there exists a constant C such that $\mathbb{E}[\Gamma_T^{-q}] < C$ for all $q > 0$ when $\alpha > \frac{1}{2}$, whereas for $\alpha = \frac{1}{2}$, the q-th negative moment of the weight process is finite, as soon as

$$2p(4p-1)\sigma^2\, q \leq \frac{\sigma^2}{8}\left(\frac{2b(0)}{\sigma^2} - 1\right)^2.$$

Notice that, thanks to point (i) in this Remark, a sufficient condition such that this last inequality holds is

$$16p(4p-1)\, q \leq (12p+2)^2.$$

We cut the proof of Proposition 3.3.1 in two technical lemmas.

Lemma 3.3.3. *Under the hypotheses of Theorem 3.1.6, for $p \geq 1$ and $\alpha \in [\frac{1}{2}, 1)$, there exists a constant C not depending on Δt such that for all $\Delta t \leq \Delta_{\max}(\alpha)$*

$$\mathbb{E}[\Gamma_t^2 \mathcal{E}_t^{4p}] \leq 4p \int_0^t \mathbb{E}\left[\Gamma_s^2 \mathcal{E}_s^{4p-1}\left(b(X_{\eta(s)}) - b(X_s)\right)\right] ds + 4p\|b'\|_\infty \int_0^t \sup_{0\leq u\leq s} \mathbb{E}[\Gamma_u^2 \mathcal{E}_u^{4p}] ds + C\Delta t^{4p}. \tag{3.3.5}$$

Lemma 3.3.4. *Under the hypotheses of Theorem 3.1.6, for $p \geq 1$ and $\alpha \in [\frac{1}{2}, 1)$, there exists a constant C not depending on Δt such that for all $\Delta t \leq \Delta_{\max}(\alpha)$, and for any $s \in [0, T]$*

$$|\mathbb{E}[\Gamma_s^2 \mathcal{E}_s^{4p-1}(b(X_{\eta(s)}) - b(X_s))]| \leq C \sup_{u\leq s} \mathbb{E}[\Gamma_u^2 \mathcal{E}_u^{4p}] + C\Delta t^{4p}. \tag{3.3.6}$$

As we will see soon, we prove (3.3.6) with the help of the Itô's formula applied to b, and here is where we need b of class $\mathcal{C}^2$ required in Hypotheses 3.1.5-(ii).

Proof of Proposition 3.3.1. Thank to Lemmas 3.3.3 and 3.3.4, we have

$$\mathbb{E}[\Gamma_t^2 \mathcal{E}_t^{4p}] \leq C \int_0^t \sup_{u\leq s} \mathbb{E}[\Gamma_u^2 \mathcal{E}_u^{4p}] ds + C\Delta t^{4p}.$$

and since the right-hand side is increasing, it follows

$$\sup_{s\leq t}\mathbb{E}[\Gamma_s^2\mathcal{E}_s^{4p}]\leq C\int_0^t\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]ds+C\Delta t^{4p},$$

from where we conclude the result thanks to Gronwall's Inequality. □

Now we present the proof of the technical lemmas.

Proof of Lemma 3.3.3. By the integration by parts formula,

$$\begin{aligned}\mathbb{E}[\Gamma_t^2\mathcal{E}_t^{4p}]=\ &4p\mathbb{E}\left[\int_0^t\Gamma_s^2\mathcal{E}_s^{4p-1}\left\{\operatorname{sgn}(\overline{Z}_s)b(\overline{X}_{\eta(s)})-b(X_s)\right\}ds\right]\\&+2p(4p-1)\mathbb{E}\left[\int_0^t\Gamma_s^2\mathcal{E}_s^{4p-2}\Sigma_s^2ds\right]\\&+4p\mathbb{E}\left[\int_0^t\Gamma_s^2\mathcal{E}_s^{4p-1}dL_s^0(\overline{X})\right]-\mathbb{E}\left[\int_0^t2\beta_s\Gamma_s^2\mathcal{E}_s^{4p}ds\right].\end{aligned}$$

Thanks to Lemma 3.2.4 and the control in the moments of the exact process in Lemma 3.1.2 we have

$$\mathbb{E}\left[\int_0^t\Gamma_s^2\mathcal{E}_s^{4p-1}dL_s^0(\overline{X})\right]\leq\mathbb{E}\left[\int_0^t|X_s^{4p-1}|dL_s^0(\overline{X})\right]\leq\sqrt{\mathbb{E}\left[\sup_{0\leq s\leq T}X_s^{8p-2}\right]\mathbb{E}\left[L_T^0(\overline{X})^2\right]}\leq C\Delta t^{4p}.$$

On the other hand, with $\operatorname{sgn}(x)=1-2\mathbb{1}_{\{x<0\}}$, calling $\Delta W_s=W_s-W_{\eta(s)}$, we get for all $0\leq s\leq t$

$$\Sigma_s^2\leq2\left[\sigma X_s^\alpha-\sigma\overline{X}_s^\alpha\right]^2+2\left[\sigma\overline{X}_s^\alpha-\sigma\overline{X}_{\eta(s)}^\alpha-\alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s\right]^2+R_s^\Sigma\mathbb{1}_{\{\overline{Z}_s<0\}},$$

where we put aside all the terms multiplied by $\mathbb{1}_{\{\overline{Z}_s<0\}}$ in

$$\begin{aligned}R_s^\Sigma:=&4\left[\sigma\overline{X}_{\eta(s)}^\alpha+\alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s\right]\\&\times\left\{\sigma\overline{X}_{\eta(s)}^\alpha+\alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s+\sigma X_s^\alpha-\sigma\overline{X}_s^\alpha+\sigma\overline{X}_s^\alpha-\sigma\overline{X}_{\eta(s)}^\alpha-\alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s\right\}.\end{aligned}$$

So, from the previous computations, the Lipschitz property of b, and Young's Inequality, we conclude

$$\begin{aligned}\mathbb{E}[\Gamma_t^2\mathcal{E}_t^{4p}] \leq 4p\int_0^t \mathbb{E}\left[\Gamma_s^2\mathcal{E}_s^{4p-1}\left(b(X_{\eta(s)})-b(X_s)\right)\right]ds \\ +4p\|b'\|_\infty\int_0^t \mathbb{E}\left[\Gamma_{\eta(s)}^2\mathcal{E}_{\eta(s)}^{4p}\right]ds + 4p\|b'\|_\infty\int_0^t \mathbb{E}[\Gamma_s^2\mathcal{E}_s^{4p}]ds \\ +4p(4p-1)\mathbb{E}\left[\int_0^t \Gamma_s^2\mathcal{E}_s^{4p-2}\left(\sigma X_s^\alpha-\sigma\overline{X}_s^\alpha\right)^2ds\right] \\ +(4p-2)(4p-1)\mathbb{E}\left[\int_0^t \Gamma_s^2\mathcal{E}_s^{4p}ds\right] \\ +2(4p-1)\int_0^t \mathbb{E}[D_s(\overline{X})^{4p}]ds \\ -\mathbb{E}\left[\int_0^t 2\beta_s\Gamma_s^2\mathcal{E}_s^{4p}ds\right]+\int_0^t \mathbb{E}\left[R_s\mathbb{1}_{\{\overline{Z}_s<0\}}\right]ds + C\Delta t^{4p}.\end{aligned}$$

where $R_s = 4p(4p-1)\mathcal{E}_s^{4p-2}R_s^\Sigma + 8p\mathcal{E}_s^{4p-1}b(\overline{X}_{\eta(s)})$, and from Lemma 3.2.2 we have

$$\mathbb{E}\left[R_s\mathbb{1}_{\{\overline{Z}_s<0\}}\right] \leq C\Delta t^{4p}.$$

Since $\Delta t \leq \Delta_{\max}(\alpha)$ and Remark 3.3.2-(i), we can apply Lemma 3.2.5 so $\mathbb{E}[D_s(\overline{X})^{4p}] \leq C\Delta t^{4p}$. Introducing these estimations in the previous computations, we have

$$\begin{aligned}\mathbb{E}[\Gamma_t^2\mathcal{E}_t^{4p}] \leq & 4p\int_0^t \mathbb{E}\left[\Gamma_s^2\mathcal{E}_s^{4p-1}\left(b(X_{\eta(s)})-b(X_s)\right)\right]ds \\ &+ 4p\|b'\|_\infty\int_0^t \mathbb{E}\left[\Gamma_{\eta(s)}^2\mathcal{E}_{\eta(s)}^{4p}\right]ds \\ &+ \left(4p\|b'\|_\infty + (4p-2)(4p-1)\right)\mathbb{E}\left[\int_0^t \Gamma_s^2\mathcal{E}_s^{4p}ds\right] \\ &+ 4p(4p-1)\mathbb{E}\left[\int_0^t \Gamma_s^2\mathcal{E}_s^{4p-2}\left(\sigma X_s^\alpha-\sigma\overline{X}_s^\alpha\right)^2ds\right] \\ &- \mathbb{E}\left[\int_0^t 2\beta_s\Gamma_s^2\mathcal{E}_s^{4p}ds\right] + C\Delta t^{4p}.\end{aligned}$$

Now we use the particular form of the weight process. Since for all $\frac{1}{2}\leq\alpha\leq 1$, for all $x\geq 0$, $y\geq 0$,

$$|x^\alpha-y^\alpha|(x^{1-\alpha}+y^{1-\alpha}) \leq 2\alpha|x-y|, \tag{3.3.7}$$

we have

$$\mathbb{E}\int_0^t \Gamma_s^2\mathcal{E}_s^{4p-2}\left(\sigma X_s^\alpha-\sigma\overline{X}_s^\alpha\right)^2ds \leq \mathbb{E}\int_0^t \Gamma_s^2\mathcal{E}_s^{4p}\frac{4\alpha^2\sigma^2}{X_s^{2(1-\alpha)}}ds,$$

and then, from the definition of β in (3.3.1), we conclude

$$\mathbb{E}[\Gamma_t^2\mathcal{E}_t^{4p}] \leq 4p\int_0^t \mathbb{E}\left[\Gamma_s^2\mathcal{E}_s^{4p-1}\left(b(X_{\eta(s)})-b(X_s)\right)\right]ds + 4p\|b'\|_\infty\int_0^t \mathbb{E}\left[\Gamma_{\eta(s)}^2\mathcal{E}_{\eta(s)}^{4p}\right]ds + C\Delta t^{4p}.$$

from where

$$\mathbb{E}[\Gamma_t^2\mathcal{E}_t^{4p}] \leq 4p\int_0^t \mathbb{E}\left[\Gamma_s^2\mathcal{E}_s^{4p-1}\left(b(X_{\eta(s)}) - b(X_s)\right)\right]ds + 4p\|b'\|_\infty\int_0^t \sup_{0\leq u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]ds + C\Delta t^{4p}.$$

□

Proof of Lemma 3.3.4. By integration by parts

$$\begin{aligned}\mathbb{E}\left[\Gamma_s^2\mathcal{E}_s^{4p-1}(b(X_{\eta(s)}) - b(X_s))\right] &= -\mathbb{E}\int_{\eta(s)}^s 2\Gamma_u\mathcal{E}_u^{4p-1}(b(X_{\eta(s)}) - b(X_u))\beta_u\Gamma_u du\\ &\quad + \mathbb{E}\int_{\eta(s)}^s \Gamma_u^2 d\left(\mathcal{E}_u^{4p-1}(b(X_{\eta(s)}) - b(X_u))\right).\end{aligned} \tag{3.3.8}$$

Applying Hölder's Inequality to the first term in the right-hand side we have

$$\begin{aligned}&\left|\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2\mathcal{E}_u^{4p-1}[b(X_{\eta(s)}) - b(X_u)]\beta_u du\right|\\ &\qquad\leq \int_{\eta(s)}^s \left(\mathbb{E}\left[\Gamma_u^2\mathcal{E}_u^{4p}\right]\right)^{1-\frac{1}{4p}}\left(\mathbb{E}\left[|b(X_{\eta(s)}) - b(X_u)|^{4p}\beta_u^{4p}\right]\right)^{\frac{1}{4p}}du.\end{aligned}$$

Recalling the Remark 3.3.2-(ii), we have that $\mathbb{E}[\beta_u^{8p}]$ is finite, so applying Lemma 3.2.1,

$$\mathbb{E}\left[|b(X_{\eta(s)}) - b(X_u)|^{4p}\beta_u^{4p}\right] \leq \sqrt{\mathbb{E}\left[|b(X_{\eta(s)}) - b(X_u)|^{8p}\right]\mathbb{E}[\beta_u^{8p}]} \leq C\Delta t^{2p}.$$

Then,

$$\left|\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2\mathcal{E}_u^{4p-1}[b(X_{\eta(s)}) - b(X_u)]\beta_u du\right| \leq C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t^{3/2}.$$

Applying the Itô's Formula to the second term in the right-hand side of (3.3.8), and taking expectation we get

$$\begin{aligned}&\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2 d\Big(\mathcal{E}_u^{4p-1}[b(X_{\eta(s)}) - b(X_u)]\Big)\\ &\quad= -\sigma\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2\mathcal{E}_u^{4p-1}\left(b'(X_u)b(X_u) + \frac{\sigma^2}{2}b''(X_u)X_u^{2\alpha}\right)du\\ &\quad\quad+(4p-1)\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2\mathcal{E}_u^{4p-2}(b(X_{\eta(s)}) - b(X_u))\left\{\operatorname{sgn}(\overline{Z}_u)b(\overline{X}_{\eta(s)}) - b(X_u)\right\}du\\ &\quad\quad+\frac{\sigma^2}{2}(4p-1)(4p-2)\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2\mathcal{E}_u^{4p-3}(b(X_{\eta(s)}) - b(X_u))\Sigma_u^2 du\\ &\quad\quad-(4p-1)\sigma^2\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2\mathcal{E}_u^{4p-2}b'(X_u)X_u^\alpha\Sigma_u du\\ &\quad\quad+\frac{(4p-1)}{2}\mathbb{E}\int_{\eta(s)}^s \Gamma_u^2\mathcal{E}_u^{4p-2}(b(X_{\eta(s)}) - b(X_u))dL_u^0(\overline{X})\\ &\quad=: I_1 + I_2 + I_3 + I_4 + I_5.\end{aligned} \tag{3.3.9}$$

By the finiteness of the moment of X, the linear growth of b, and the polynomial growth of b'', applying Holder's inequality, we have

$$\begin{aligned}|I_1| &\leq C\int_{\eta(s)}^{s}\left(\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\left(\mathbb{E}\left[|b'(X_u)b(X_u)+\frac{\sigma^2}{2}b''(X_u)X_u^{2\alpha}|^{4p}\right]\right)^{\frac{1}{4p}}du\\ &\leq C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t.\end{aligned}$$

For the bound of I_2, since b is Lipschitz, and $\operatorname{sgn}(x)=1-2\mathbb{1}_{\{x<0\}}$, we have

$$\begin{aligned}|I_2| \leq& C\int_{\eta(s)}^{s}\mathbb{E}\left[\Gamma_u^2\mathcal{E}_u^{4p-2}|X_{\eta(s)}-X_u||\overline{X}_{\eta(s)}-\overline{X}_u|\right]du\\ &+C\int_{\eta(s)}^{s}\mathbb{E}\left[\Gamma_u^2\mathcal{E}_u^{4p-1}|X_{\eta(s)}-X_u|\right]du+\int_{\eta(s)}^{s}\mathbb{E}|R_u^{(2)}\mathbb{1}_{\{\overline{Z}_u<0\}}|du\\ \leq& C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{2p}}\Delta t^2+C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t^{3/2}+C\Delta t^{4p}\end{aligned}$$

Where again, all the terms multiplied by $\mathbb{1}_{\{\overline{Z}_u<0\}}$ are putted in the rest $R_u^{(2)}$, and the expectation of the product is bounded with Lemma 3.2.2.

In a similar way for the bound of I_3, decomposing Σ_u with $\operatorname{sgn}(x)=1-2\mathbb{1}_{\{x<0\}}$,

$$\begin{aligned}|I_3| \leq& C\int_{\eta(s)}^{s}\mathbb{E}\left[\Gamma_u^2\mathcal{E}_u^{4p-3}|X_{\eta(s)}-X_u|D_s(\overline{X})^2\right]du\\ &+C\int_{\eta(s)}^{s}\mathbb{E}\left[\Gamma_u^2\mathcal{E}_u^{4p-3}|X_{\eta(s)}-X_u|\left(\sigma\overline{X}_u-\sigma X_u^{\alpha}\right)^2\right]du+\int_{\eta(s)}^{s}\mathbb{E}|R_u^{(3)}\mathbb{1}_{\{\overline{Z}_u<0\}}|du.\end{aligned}$$

For the first term in the right-hand side we have

$$\begin{aligned}\mathbb{E}\left[\Gamma_u^2\mathcal{E}_u^{4p-3}|X_{\eta(s)}-X_u|D_s(\overline{X})^2\right] &\leq \left(\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{3}{4p}}\left(\mathbb{E}[|X_{\eta(s)}-X_u|^{4p}]\right)^{\frac{1}{4p}}\left(\mathbb{E}[D_s(\overline{X})^{4p}]\right)^{\frac{1}{2p}}\\ &\leq\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{3}{4p}}\Delta t^{5/2},\end{aligned}$$

due to the bound for the increments of the exact process and Lemma 3.2.5. For the second term, applying (3.3.7), and noting that from Remark 3.3.2-(ii), the exact process has negative moments up to of order $12p+2$

$$\begin{aligned}&\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p-3}|X_{\eta(s)}-X_u|(\sigma\overline{X}_u-\sigma X_u^{\alpha})^2]\\ &\leq C\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p-1}|X_{\eta(s)}-X_u|\frac{1}{X_u^{2(1-\alpha)}}]\\ &\leq C\left(\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\left(\mathbb{E}[|X_{\eta(s)}-X_u|^{8p}]\right)^{1/8p}\left(\mathbb{E}\left[\frac{1}{X_u^{16(1-\alpha)p}}\right]\right)^{1/8p}\\ &\leq\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t^{1/2}.\end{aligned}$$

We control the third term in the right-hand side in the bound for $|I_3|$ using again Lemma 3.2.2, so

$$|I_3| \leq C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{3}{4p}}\Delta t^{7/2} + \left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t^{3/2} + C\Delta t^{4p}.$$

Now we bound $|I_4|$.

$$\begin{aligned}|I_4| \leq{}& C\int_{\eta(s)}^{s}\mathbb{E}|\Gamma_u^2\mathcal{E}_u^{4p-2}D_u(\overline{X})b'(X_u)X_u^{\alpha}|du\\ &+C\int_{\eta(s)}^{s}\mathbb{E}|\Gamma_u^2\mathcal{E}_u^{4p-2}(\sigma\overline{X}_u-\sigma X_u^{\alpha})b'(X_u)X_u^{\alpha}|du+\int_{\eta(s)}^{s}\mathbb{E}|R_u^{(4)}\mathbb{1}_{\{\overline{Z}_u<0\}}|du.\end{aligned}$$

We control the first term in the right-hand side using HÃűlder's inequality, Lemma 3.2.5 and the control in the moments of the exact process for all $0\leq u\leq s$

$$\begin{aligned}\mathbb{E}|\Gamma_u^2\mathcal{E}_u^{4p-2}D_u(\overline{X})b'(X_u)X_u^{\alpha}| &\leq (\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}])^{1-\frac{1}{2p}}(\mathbb{E}[D_s(\overline{X})^{4p}])^{\frac{1}{4p}}(\mathbb{E}[b'(X_u)^{4p}X_u^{4p\alpha}])^{\frac{1}{4p}}\\ &\leq C\left(\sup_{u\leq s}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{2p}}\Delta t.\end{aligned}$$

For the second term in the right-hand side of the bound for $|I_4|$, we use one more time (3.3.7), and the existence of negative moments of the exact process X, and then

$$\begin{aligned}&\mathbb{E}|\Gamma_u^2\mathcal{E}_u^{4p-2}\left[\sigma\overline{X}_u-\sigma X_u^{\alpha}\right]b'(X_u)X_u^{\alpha}|\\ &\qquad\leq C\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p-1}\frac{1}{X_u^{(1-\alpha)}}b'(X_u)X_u^{\alpha}]\leq C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}.\end{aligned}$$

To control the third term in the right-hand side of the bound for $|I_4|$ we use Lemma 3.2.2 just as before. So

$$|I_4| \leq C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{2p}}\Delta t^{2} + C\left(\sup_{u\leq s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t + C\Delta t^{4p}.$$

Finally,

$$\begin{aligned}|I_5| &= \frac{(4p-1)}{2}\mathbb{E}\int_{\eta(s)}^{s}\Gamma_u^2X_u^{4p-2}|b(X_{\eta(s)})-b(X_u)|dL_u^0(\overline{X})\\ &\leq C\mathbb{E}\left[\sup_{u\leq s}\left[1+X_u^{4p-1}\right]L_T^0(\overline{X})\right]\\ &\leq C\sqrt{\mathbb{E}\left[\sup_{u\leq s}\left[1+X_u^{4p-1}\right]^2\right]}\sqrt{\mathbb{E}[L_T^0(\overline{X})^2]}\leq C\Delta t^{4p},\end{aligned}$$

the last inequality comes from Lemmas 3.1.2 and 3.2.4.

Putting all the last calculations in (3.3.8) we obtain

$$\begin{aligned}\left|\mathbb{E}[\Gamma_s^2\mathcal{E}_s^{4p-1}(b(X_{\eta(s)})-b(X_s))]\right| \le C\left(\sup_{u\le s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t^{3/2} + C\left(\sup_{u\le s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{4p}}\Delta t \\ + C\left(\sup_{u\le s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{1}{2p}}\Delta t^{2} + C\left(\sup_{u\le s}\mathbb{E}[\Gamma_u^2\mathcal{E}_u^{4p}]\right)^{1-\frac{3}{4p}}\Delta t^{7/2} \\ + C\Delta t^{4p}.\end{aligned}$$

Applying Young's Inequality in all terms in the right, we get the desired inequality (3.3.6). □

Remark 3.3.5. Proceeding as in the Proof of Lemma 3.3.4, if $p=1$ or $p\ge 3/2$, is not difficult to prove

$$\left|\mathbb{E}\left[\mathcal{E}_s^{2p-1}(b(X_{\eta(s)})-b(X_s))\right]\right| \le C\sup_{u\le s}\mathbb{E}[\mathcal{E}_u^{2p}] + C\Delta t^{2p}. \tag{3.3.10}$$

Notice that (3.3.10) is similar to (3.3.6) with the process $\Gamma=1$. The restriction on p comes from the following observation. Applying the Itô's Lemma to the function $(x,y)\mapsto x^{2p-1}y$ for $p\ge 1$, with the couple of processes $(\mathcal{E}_\cdot, b(X_{\eta(\cdot)})-b(X_\cdot))$, in the analogous of the identity (3.9) with $\Gamma=1$, it will appear a term of the form

$$\mathbb{E}\int_{\eta(s)}^{s}\mathcal{E}_u^{2p-3}[b(X_{\eta(s)})-b(X_u)]\Sigma_u^2du.$$

The restriction $p\ge 3/2$ avoids the situation where $2p-3$ is negative.

3.3.2 Proof of Theorem 3.1.6

We start by controlling the L^{2p}-error. First, assume $p=1$ or $p\ge 3/2$. By the Itô's formula we have

$$\begin{aligned}\mathbb{E}[\mathcal{E}_t^{2p}] &= 2p\mathbb{E}\int_0^t \mathcal{E}_s^{2p-1}\left\{\mathrm{sgn}(\overline{Z}_s)b(\overline{X}_{\eta(s)})-b(X_s)\right\}ds \\ &\quad + 2p\mathbb{E}\int_0^t\mathcal{E}_s^{2p-1}dL_s^0(\overline{X}) + p(2p-1)\mathbb{E}\int_0^t\mathcal{E}_s^{2p-2}\Sigma_s^2ds.\end{aligned}$$

As we have seen before, $\mathbb{E}\int_0^t\mathcal{E}_s^{2p-1}dL_s^0(\overline{X})\le C\Delta t^{2p}$, and $\mathrm{sgn}(x)=1-2\mathbb{1}_{\{x<0\}}$, so

$$\begin{aligned}\mathbb{E}[\mathcal{E}_t^{2p}] \le 2p\mathbb{E}\int_0^t\mathcal{E}_s^{2p-1}\left[b(\overline{X}_{\eta(s)})-b(X_{\eta(s)})\right]ds \\ + 2p\mathbb{E}\int_0^t\mathcal{E}_s^{2p-1}\left[b(X_{\eta(s)})-b(X_s)\right]ds \\ + 8p(2p-1)\mathbb{E}\int_0^t\mathcal{E}_s^{2p-2}\left[\sigma\overline{X}_s^\alpha-\sigma X_s^\alpha\right]^2ds \\ + 8p(2p-1)\mathbb{E}\int_0^t\mathcal{E}_s^{2p-2}D_s(\overline{X})^2ds + \mathbb{E}\int_0^t R_s\mathbb{1}_{\{\overline{Z}_s<0\}}ds + C\Delta t^{2p},\end{aligned} \tag{3.3.11}$$

where $R_s = 4p(2p-1)\mathcal{E}_s^{2p-2}\left[\sigma\overline{X}_{\eta(s)}^{\alpha} + \alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\left(W_s - W_{\eta(s)}\right)\right] + 8p\mathcal{E}_s^{2p-1}b(\overline{X}_{\eta(s)})$. If we use the Lipschitz property of b, and Young's inequality in the first term in the right of (3.3.11), Lemma 3.2.5 in the fourth one, and Lemma 3.2.2 in the fifth one, we have

$$\begin{aligned}\mathbb{E}[\mathcal{E}_t^{2p}] \leq C\int_0^t \sup_{u\leq s}\mathbb{E}[\mathcal{E}_u^{2p}]ds + 2p\int_0^t \mathbb{E}[\mathcal{E}_s^{2p-1}(b(X_{\eta(s)}) - b(X_s))]ds \\ + 8p(2p-1)\int_0^t \mathbb{E}[\mathcal{E}_s^{2p-2}(\sigma\overline{X}_s^{\alpha} - \sigma X_s^{\alpha})^2]ds + C\Delta t^{2p}.\end{aligned} \tag{3.3.12}$$

And, according to Remark 3.3.5, we have

$$\left|\mathbb{E}[\mathcal{E}_s^{2p-1}(b(X_{\eta(s)}) - b(X_s))]\right| \leq C\sup_{u\leq s}\mathbb{E}[\mathcal{E}_u^{2p}] + C\Delta t^{2p}.$$

On the other hand, using again (3.3.7), we have

$$\mathbb{E}[\mathcal{E}_s^{2p-2}(\sigma\overline{X}_s^{\alpha} - \sigma X_s^{\alpha})^2] \leq C\mathbb{E}[\mathcal{E}_s^{2p}X_s^{-2(1-\alpha)}] = C\mathbb{E}[\Gamma_s\mathcal{E}_s^{2p}X_s^{-2(1-\alpha)}\Gamma_s^{-1}],$$

and applying Cauchy-Schwartz inequality,

$$\mathbb{E}[\Gamma_s\,\mathcal{E}_s^{2p}\frac{1}{X_s^{2(1-\alpha)}}\Gamma_s^{-1}] \leq (\mathbb{E}[\Gamma_s^2\mathcal{E}_s^{4p}])^{\frac{1}{2}}\left(\mathbb{E}[\frac{1}{X_s^{4(1-\alpha)}}\Gamma_s^{-2}]\right)^{\frac{1}{2}}.$$

The first term in the right-hand side is the weight error controlled by Proposition 3.3.1. To control the second one, let us recall Remark 3.3.2. For $\alpha > \frac{1}{2}$, the exact process and the weight process Γ have negative moments of any order, therefore the second term in the last inequality is bounded by a constant. On the other hand, for $\alpha = \frac{1}{2}$ we need a finer analysis. From the second point in Remark 3.3.2, the $12p+2$-th negative moment of the exact process is finite, and since

$$16p(4p-1)\,2\frac{6p+1}{6p} < (12p+2)^2,$$

according with the third point of Remark 3.3.2, the $2(6p+1)/6p$-th negative moment of the weight process Γ is also finite. Therefore, when $\alpha = \frac{1}{2}$,

$$\mathbb{E}\left[\frac{1}{X_s^{4(1-\alpha)}}\Gamma_s^{-2}\right] = \mathbb{E}\left[\frac{1}{X_s^2}\Gamma_s^{-2}\right] \leq \left(\mathbb{E}\left[\frac{1}{X_s^{12p+2}}\right]\right)^{\frac{1}{6p+1}}\left(\mathbb{E}\left[\Gamma_s^{-2\frac{6p+1}{6p}}\right]\right)^{\frac{6p}{6p+1}} \leq C,$$

and then in any case

$$\mathbb{E}\left[\Gamma_s\mathcal{E}_s^{2p}\frac{1}{X_s^{2(1-\alpha)}}\Gamma_s^{-1}\right] \leq C\Delta t^{2p}.$$

Introducing all the last computations in (3.3.12) we get

$$\mathbb{E}[\mathcal{E}_t^{2p}] \leq C\int_0^t\left(\sup_{u\leq s}\mathbb{E}\left[\mathcal{E}_u^{2p}\right]\right)ds + C\Delta t^{2p}.$$

Since the right-hand side is increasing, thanks to Gronwall's Inequality we have, for $p = 1$ or $p \geq 3/2$,

$$\sup_{0\leq t\leq T} \left(\mathbb{E}[\mathcal{E}_t^{2p}]\right)^{\frac{1}{2p}} \leq C\Delta t.$$

We extend to $p \in (1, 3/2)$, thanks to Jensen's inequality

$$(\mathbb{E}[\mathcal{E}_t^{2p}])^{\frac{1}{2p}} \leq (\mathbb{E}[\mathcal{E}_t^{\frac{6}{2}}])^{\frac{2}{6}} \leq C\Delta t,$$

and we conclude that $(\mathbb{E}[\mathcal{E}_t^{2p}])^{\frac{1}{2p}} \leq C\Delta t$ for all $p \geq 1$ satisfying Remark 3.3.2-(i).

Now we control the L^p-error. For $\alpha = \frac{1}{2}$ and $p \geq 2$, denoting $p' = \frac{p}{2} \geq 1$,

$$\sup_{0\leq t\leq T} \left(\mathbb{E}\left[\mathcal{E}_t^{p}\right]\right)^{\frac{1}{p}} = \sup_{0\leq t\leq T} \left(\mathbb{E}[\mathcal{E}_t^{2p'}]\right)^{\frac{1}{2p'}} \leq C\Delta t.$$

Hypothesis 3.1.5-(i) gives

$$b(0) > \frac{3(2p+1)\sigma^2}{2} = \frac{3(4p'+1)\sigma^2}{2}.$$

Since (3.3.4) in Remark 3.3.2 is satisfied, we can control the $L^{2p'}(\Omega)$-norm of the error and then

$$\sup_{0\leq t\leq T} (\mathbb{E}[\mathcal{E}_t^{p}])^{\frac{1}{p}} = \sup_{0\leq t\leq T} \left(\mathbb{E}[\mathcal{E}_t^{2p'}]\right)^{\frac{1}{2p'}} \leq C\Delta t.$$

If $p \in [1, 2)$, Hypothesis 3.1.5-(i) is $b(0) > 15\frac{\sigma^2}{2}$, which is enough to bound the $L^2(\Omega)$-norm of the error, and then from Jensen's inequality

$$\sup_{0\leq t\leq T} (\mathbb{E}[\mathcal{E}^{p}])^{\frac{1}{p}} \leq \sup_{0\leq t\leq T} \left(\mathbb{E}[\mathcal{E}_t^{2}]\right)^{\frac{1}{2}} \leq C\Delta t.$$

The case $\alpha > 1/2$ is easier. Since the Hypothesis in the parameters for this case does not depend on p, we can conclude for any $p \geq 1$ from Jensen's inequality and the control for the $L^{2p}(\Omega)$-norm of the error.

Remark 3.3.6. Let us mention an example of extension of our convergence result, based on simple transformation method: consider the 3/2-model, namely the solution of

$$r_t = r_0 + \int_0^t c_1 r_s (c_2 - r_s) ds + \int_0^t c_3\, r_s^{3/2} dW_s.$$

Applying the ItÃt's Formula to $v_t = f(r_t)$, with $f(x) = x^{-1}$, we have

$$v_t = v_0 + \int_0^t \left(c_1 + c_3^2 - c_1 c_2 v_s\right) ds + \int_0^t c_3\, v_s^{1/2} dB_s,$$

where $B_s = -W_s$ is a Brownian motion. We can approximate v with the SMS $\bar{v}$, and then define $\bar{r}_t := 1/\bar{v}_t$. Then we can deduce the strong convergence with rate one of $\bar{r}_t$ to r_t from our previous results.

Transformation methods can be used in a more exhaustive manner, in the context of CEV-like SDEs and we refer to [19] for approximation results and examples, using this approach.

3.3.3 Strong Convergence of the Projected Milstein Scheme

The Projected Milstein Scheme (PMS) is defined by $\widehat{X}_0 = x_0$, and

$$\widehat{X}_{t_k} = \left(\widehat{X}_{t_{k-1}} + b(\widehat{X}_{t_{k-1}})\Delta t + \sigma \widehat{X}^{\alpha}_{t_{k-1}}(W_{t_k} - W_{t_{k-1}}) + \frac{\alpha\sigma^2 \widehat{X}^{2\alpha-1}_{t_{k-1}}}{2}\left[(W_{t_k} - W_{t_{k-1}})^2 - \Delta t\right]\right)^+,$$

where for all $x \in \mathbb{R}$, $(x)^+ = \max(0, x)$. The continuous time version of the (PMS) is given by

$$\widehat{X}_t = \left(\widehat{X}_{\eta(t)} + b(\widehat{X}_{\eta(t)})\Delta t + \sigma \widehat{X}^{\alpha}_{\eta(t)}(W_t - W_{\eta(t)}) + \frac{\alpha\sigma^2 \widehat{X}^{2\alpha-1}_{\eta(t)}}{2}\left[(W_t - W_{\eta(t)})^2 - \Delta t\right]\right)^+. \quad (3.3.13)$$

Notice that for all $t \in [0, T]$, $0 \leq \widehat{X}_t \leq \overline{X}_t$, then the positive moments of the PMS are bounded (see Lemma 3.1.4).

To obtain a strong convergence rate for the PMS, we first show that the PMS and the SMS coincide with a large probability.

Lemma 3.3.7. *Let us consider the stopping time* $\tau = \inf\{s \geq 0 : \overline{X}_s \neq \widehat{X}_s\}$. *Assume that* $b(0) > 2\alpha(1-\alpha)^2\sigma^2$ *and* $\Delta t \leq 1/(2K(\alpha))$. *Then for any* $p \geq 1$,

$$\mathbb{P}\left(\tau \leq T\right) \leq C\Delta t^p.$$

Proof. Notice that τ is almost surely strictly positive because both schemes start from the same deterministic initial condition x_0. On the other hand

$$\mathbb{P}\left(\tau \leq T\right) = \sum_{i=0}^{N-1} \mathbb{P}\left(\tau \in (t_k, t_{k+1}]\right),$$

and according to Lemma 3.2.2

$$\mathbb{P}\left(\tau \in [t_i, t_{i+1})\right) = \mathbb{P}\left(\inf_{t_k < s \leq t_{k+1}} \overline{Z}_s \leq 0, \overline{X}_{t_k} = \widehat{X}_{t_k}\right) \leq \mathbb{P}\left(\inf_{t_k < s \leq t_{k+1}} \overline{Z}_s \leq 0\right) \leq \exp\left(-\frac{\gamma}{\Delta t}\right).$$

So,

$$\mathbb{P}\left(\tau \leq T\right) \leq \frac{T}{\Delta t}\exp\left(-\frac{\gamma}{\Delta t}\right).$$

Since for any $p \geq 1$, there exists a constant C_p such that $\exp(-\gamma/\Delta t)/\Delta t \leq C_p \Delta t^p$, we have

$$\mathbb{P}\left(\tau \leq T\right) \leq C\Delta t^p.$$

□

Corollary 3.3.8. *Assume Hypotheses 3.1.1 and 3.1.5. Consider a maximum step size* $\Delta_{\max}(\alpha)$ *defined in* (3.1.10). *Let* $(X_t, 0 \leq t \leq T)$ *be the process defined on* (3.1.1) *and* $(\widehat{X}_t, 0 \leq t \leq T)$ *the Projecter Milstein scheme given in* (3.3.13). *Then for any* $p \geq 1$ *that allows Hypotheses 3.1.5, there exists a constant* C *depending on* p, T, $b(0)$, α, σ, K, *and* x_0, *but not on* Δt, *such that for all* $\Delta t \leq \Delta_{\max}(\alpha)$,

$$\sup_{0 \leq t \leq T}\left(\mathbb{E}[|X_t - \widehat{X}_t|^p]\right)^{\frac{1}{p}} \leq C\Delta t. \quad (3.3.14)$$

Proof. Notice that for all $t \in [0,T]$, with $\tau = \inf\{s \geq 0 : \overline{X}_s \neq \widehat{X}_s\}$,

$$\begin{aligned}\mathbb{E}[|\widehat{X}_t - X_t|^p] &= \mathbb{E}[|\widehat{X}_t - X_t|^p \mathbb{1}_{\{\tau \leq T\}}] + \mathbb{E}[|\widehat{X}_t - X_t|^p \mathbb{1}_{\{\tau > T\}}] \\ &\leq \sqrt{\mathbb{E}[|\widehat{X}_t - X_t|^{2p}]\mathbb{P}(\tau \leq T)} + \mathbb{E}[|\overline{X}_t - X_t|^p] \leq C\Delta t^p\end{aligned}$$

where the last inequality comes from Lemma 3.3.7 and Theorem 3.1.6. □

3.4 Numerical Experiments and Conclusion

We start this section with the analysis of two numerical experiments. The first one aims to study empirically the strong rate of convergence of the SMS in comparison with other schemes proposed the literature. The second one aims to study the impact of including the SMS in a Multilevel Monte Carlo application.

3.4.1 Empirical study of the strong rate of converge

In this experiment we compute the error of the schemes as a function of the step size Δt for different values of the parameters α and σ.

For $\alpha > \frac{1}{2}$ we compare the SMS with the Symmetrized Euler Scheme (SES) introduced in [10], and with the Balanced Milstein Scheme (BMS) presented in [51]. Whereas for $\alpha = \frac{1}{2}$, in addition to the aforementioned schemes, we will also compare SMS with the Modified Euler Scheme (MES) proposed in [19], and with the Alfonsi Implicit Scheme (AIS) proposed in [1].

Let us first, shortly review those different schemes.

Alfonsi Implicit Scheme (AIS). Proposed in [1], the AIS can be applied to equation (3.1.1) when the drift is a linear function. A priori, the AIS can be applied for $\alpha \in [\frac{1}{2}, 1)$, but is relevant to observe that only when $\alpha = \frac{1}{2}$, the AIS is in fact an explicit scheme (also know as drift-implicit square-root Euler approximations) whereas in any other case is not. This implies that in order to compute the AIS for $\alpha > \frac{1}{2}$, at each time step it is necessary to solve numerically a non linear equation. This extra step in the implementation of the scheme brings questions about the impact of the error of this subroutine in the error of the scheme, and about the computing performance of the scheme. Since this questions are beyond the scope of the present work, we include the AIS in the comparison only in the Cox–Ingersol–Ross (CIR) case (linear drift and $\alpha = \frac{1}{2}$). In this context, the AIS can be use only if $\sigma^2 > 4b(0)$, for other values of the parameters the AIS is not defined. In terms of convergence, when $\alpha = \frac{1}{2}$, according to Theorem 2 in Alfonsi [3], the AIS converges in the $L^p(\Omega)$-norm, for $p \geq 1$, at rate Δt when $(1 \vee 3p/4)\sigma^2 < b(0)$. When $\alpha > \frac{1}{2}$, the AIS (see Section 3 of [3]) converge as soon as $b(0) > 0$, at rate Δt to the exact solution.

Balanced Milstein Scheme (BMS). The BMS was introduced by Kahl and Schurz in [51], and although its convergence it is not proven for Equation (3.1.1) (see Remark 5.12 in [51]), numerical experiments shows a competitive behavior (see [50]). Also, the BMS can be easily implemented for $\alpha \in [\frac{1}{2}, 1)$, so we decide to include it in our numerical comparison.

Modified Euler Scheme (MES). Introduced in [19], the MES can be applied to the Equation (3.1.1) for $\alpha \in [\frac{1}{2}, 1)$ when the drift has the form $b(x) = \mu_1(x) - \mu_2(x)x$ for μ_1 and μ_2 suitables functions. The rate of convergence in the $L^1(\Omega)$-norm of the MES depends on the parameters. For $\alpha = \frac{1}{2}$ the rate is 1 if σ^2 is big enough compared with $b(0)$, and it is $\rho < 1$ in other case. When $\alpha > \frac{1}{2}$, the MES converges at rate 1 as soon as $b(0) > 0$ (see Proposition 4.1 in [19]). When $\alpha > \frac{1}{2}$, the implementation of the MES requires some extra tuning which is not explicitly given in [19] (see Remark 5.1), so we implement the MES only for $\alpha = \frac{1}{2}$.

Symmetryzed Euler Scheme (SES). The SES, introduced in [10], is an explicit scheme which can be apply to the equation (3.1.1) for $\alpha \in [\frac{1}{2}, 1)$ and any Lipschitz drift function b. It has the weakest hypothesis over b of all the schemes discussed in this chapter. If $\alpha = \frac{1}{2}$, according to Theorem 2.2 in [10], the rate of convergence of the SES is $\sqrt{\Delta t}$ under suitable conditions for $b(0)$, σ^2 and K. When $\alpha > \frac{1}{2}$ the SES converge at rate $\sqrt{\Delta t}$ as soon as $b(0) > 0$ (see Theorem 2.2 in [10]).

We summarize the theoretical analysis of the schemes above in Table 3.4.1 for $\alpha = \frac{1}{2}$, and in Table 3.4.2 for $\alpha > \frac{1}{2}$.

Simulation setup

In our simulations we consider a time horizon $T = 1$, and $x_0 = 1$. In order to include as many schemes as possible we consider for all simulations a linear drift

$$b(x) = 10 - 10x.$$

To measure the error of each scheme, we estimate its $L^1(\Omega)$-norm for which a theoretical rate is proposed for all the selected schemes.

Let $\mathbb{E}|\mathcal{E}_T^{\text{SMS}}|$, $\mathbb{E}|\mathcal{E}_T^{\text{BMS}}|$, $\mathbb{E}|\mathcal{E}_T^{\text{SES}}|$, $\mathbb{E}|\mathcal{E}_T^{\text{MES}}|$, and $\mathbb{E}|\mathcal{E}_T^{\text{AIS}}|$ be the $L^1(\Omega)$-norm of the error for the SMS, BMS, SES, MES and AIS respectively. To estimate these quantities, we consider as a reference solution the AIS approximation for $\Delta t = \Delta_{\max}(\alpha)/2^{12}$ when $\alpha = \frac{1}{2}$, and the SMS for $\Delta t = \Delta_{\max}(\alpha)/2^{12}$ when $\alpha > \frac{1}{2}$. Then for each

$$\Delta t \in \left\{ \frac{\Delta_{\max}(\alpha)}{2^n}, n = 1, \ldots 9 \right\},$$

Scheme	Norm	Drift	Convergence's Condition	Theoretical rate
SMS	$L^p, p \geq 1$	b Lipschitz, $b \in \mathcal{C}^2$ b'' with polynomial growth	$b(0) > 3\left(2[p \vee 2] + 1\right)\frac{\sigma^2}{2}$	1
AIS [3]	L^p, $p \in [1, \frac{4b(0)}{3\sigma^2})$	$b(x) = a - bx$	$b(0) > (1 \vee \frac{3}{4}p)\sigma^2$	1
BMS				undetermined
MES [19]	L^1	$b(x) = \mu_1(x) - \mu_2(x)x$ $\mu_i \in \mathcal{C}_b^2 \cap \mathcal{C}_b^0,\ \mu_1 \geq 0$ $\mu_1' \leq 0,\ \mu_2' \geq 0$	$b(0) > \frac{5\sigma^2}{2}$	1
			$b(0) > \frac{3\sigma^2}{2}$	$\frac{1}{2}$
			$b(0) > \sigma^2$	$\left(\frac{1}{6}, \frac{1}{2} - \frac{\sigma^2}{2b(0)+\sigma^2}\right)$
SES [10]	$L^p, p \geq 1$	b Lipschitz	$b(0) > \left[\sqrt{\frac{8}{\sigma^2}\mathcal{K}(\frac{p}{2} \vee 1)} + 1\right]\frac{\sigma^2}{2}$, $\mathcal{K}(q) = K(16q-1)$ $\vee 4\sigma^2(8p-1)^2$	$\frac{1}{2}$

Table 3.4.1: Summary of the condition over the parameters for the convergence of the different schemes for $\alpha = \frac{1}{2}$.

Scheme	Norm	Drift	Convergence's Condition	Theoretical rate
SMS	$L^p, p \geq 1$	b Lipschitz, $b \in \mathcal{C}^2$ b'' with polynomial growth	$b(0) > 2\alpha(1-\alpha)^2\sigma^2$	1
AIS	$L^p, p \in [1, \frac{4b(0)}{3\sigma^2})$	$b(x) = a - bx$	$b(0) > 0$	1
BMS				undetermined
MES	L^1	$b(x) = \mu_1(x) - \mu_2(x)x$ $\mu_i \in \mathcal{C}_b^2 \cap \mathcal{C}_b^0,\ \mu_1 \geq 0$ $\mu_1' \leq 0,\ \mu_2' \geq 0$	$b(0) > 0$	1
SES	$L^p, p \geq 1$	b Lipschitz	$b(0) > 0$	$\frac{1}{2}$

Table 3.4.2: Summary of the condition over the parameters for the convergence of the different schemes when $\alpha > \frac{1}{2}$.

we estimate $\mathbb{E}|\mathcal{E}_T^{\cdots}|$ by computing 5×10^4 trajectories of the corresponding scheme, and comparing them with the reference solution. The results of these simulations are reported in Figures 3.4.1 ($\alpha = \frac{1}{2}$) and 3.4.2 ($\alpha > \frac{1}{2}$). The graphs plot the $\text{Log}\mathbb{E}|\mathcal{E}_T^{\cdots}|$ in terms of the $\text{Log}\Delta t$, and we have added the plot of the identity map to serve as reference for rate of order 1. The schemes with a slope smaller than the slope of the reference line have an order of convergence smaller than one. To obtain a more quantitative comparison of the schemes, we also perform a regression analysis on the model

$$\log(\mathbb{E}|\mathcal{E}_T^{\cdots}|) = \rho \log(\Delta t) + C.$$

Notice that $\hat{\rho}$, the estimated value for ρ, corresponds to the empirical rate of convergence of the different schemes. We present the result of this regression analysis in Tables 3.4.3 and 3.4.4.

Empirical results for $\alpha = \frac{1}{2}$. Figure 3.4.1 and Table 3.4.3 present the result for the CIR case. From Table 3.4.1, we observe that we can distinguish five cases for the parameters.

The first case ($\sigma^2 = 1$) is such that $b(0) > 6\sigma^2$: the SMS, the AIS, and the MES have a theoretical rate of convergence equal to Δt, whereas the SES has a theoretical rate of convergence equal to $\sqrt{\Delta t}$. In Figure 3.4.1a, we observe that the graphs of the SMS, the AIS, the BMS, and the MES seem parallel to the reference line, which is expected, while the SES has a smaller slope. This is also confirms in the first line of the Table 3.4.3, where we observe that the empirical rates of convergence are close to the theoretical ones. Notice that the BMS has a competitive empirical rate of convergence, although the theoretical one is not known.

σ^2	Observed $L^1(\Omega)$ convergence rate $\hat{\rho}$ (and its R^2 value)									
	SMS		AIS		BMS		MES		SES	
	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)
1	0.996	(99.9%)	1.006	(99.9%)	1.006	(99.9%)	0.996	(99.9%)	0.594	(99.3%)
4	0.997	(99.9%)	1.005	(99.9%)	1.004	(99.9%)	0.996	(99.9%)	0.534	(99.8%)
6.25	0.998	(99.9%)	1.004	(99.9%)	1.000	(99.9%)	0.994	(99.9%)	0.524	(99.9%)
9	0.998	(99.9%)	1.002	(99.9%)	0.986	(99.9%)	0.789	(99.9%)	0.516	(99.9%)
36	0.641	(99.7%)	0.628	(99.8%)	0.454	(99.3%)	0.358	(99.4%)	0.472	(99.9%)

Table 3.4.3: Empirical rate of convergence $\hat{\rho}$ for the $L^1(\Omega)$-error of the schemes when $\alpha = \frac{1}{2}$ for different values of σ^2.

The second case ($\sigma^2 = 4$) is such that $b(0) \in (5\frac{\sigma^2}{2}, 6\sigma^2)$, now only the AIS and the MES have a theoretical rate of convergence equal to Δt. However, how we can see in Figure 3.4.1b the SMS still shows a linear behavior in this case. Recall that the condition over the parameters is a sufficient condition and we believe that could be improved. Notice that also the BMS shows a linear behavior. In the second line of Table 3.4.3, we can observe that empirical rates of convergence are close to one for all the scheme but the SES.

In Figure 3.4.1c, we illustrate the third case ($\sigma^2 = 6.25$) and then $b(0) \in (3\sigma^2/2, 5\sigma^2/2)$. In this case, only the AIS has a theoretical rate of convergence equal to one. For the MES is $\sqrt{\Delta t}$, but in the graphics we still observe a linear behavior for the MES, and also for the SMS and the BMS. This is confirm in the third line of Table 3.4.3.

The fourth case is ($\sigma^2 = 9$) and $b(0) \in (\sigma^2, 3\sigma^2/2)$, which we display in Figure 3.4.1d. For this values of the parameters the theoretical rate of convergence is known only for the AIS and the MES. Nevertheless, we observe in the graphs and in the fourth line of Table 3.4.3 that all the schemes seems to reach their optimal convergence rates.

Finally, the fifth case is ($\sigma^2 = 36$) and then $b(0) < \sigma^2$. In this case all the schemes have a sublinear behavior as we can see in Figure 3.4.1d and the fifth line of Table 3.4.3. This case illustrate the necessity of some condition over the parameters of the model to obtain the optimal rate of convergence for the SMS.

Empirical results for $\alpha > \frac{1}{2}$. In Figure 3.4.2 and Table 3.4.4 we present the results of the simulations for $\alpha = 0.6$, and $\alpha = 0.7$.

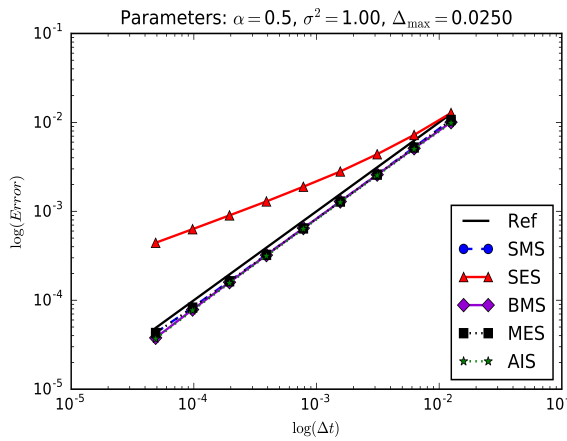
In these cases, it can be observed in numerical experiments that the MES needs smaller Δt to achieve its theoretical order one convergence rate, unless one tunes the projection operator in the manner of Remark 5.1 in [19]. Since this tuning is not explicitly given we do not include the MES in these simulations.

Parameters		Observed $L^1(\Omega)$ convergence rate $\hat{\rho}$ (and its R^2 value)					
α	σ^2	SMS		BMS		SES	
		$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)
0.6	49	0.9819	(99.9%)	0.7296	(99.1%)	0.5273	(99.8%)
	53.29	0.9766	(99.9%)	0.7788	(99.3%)	0.5133	(99.9%)
	144	0.6609	(98.9%)	0.4336	(97.3%)	0.5074	(99.9%)
0.7	64	1.004	(99.9%)	0.9022	(99.7%)	0.5242	(99.8%)
	81	0.9991	(99.9%)	0.8813	(99.7%)	0.5327	(99.7%)
	225	0.9146	(99.7%)	0.6497	(97.6%)	0.6410	(99.2%)

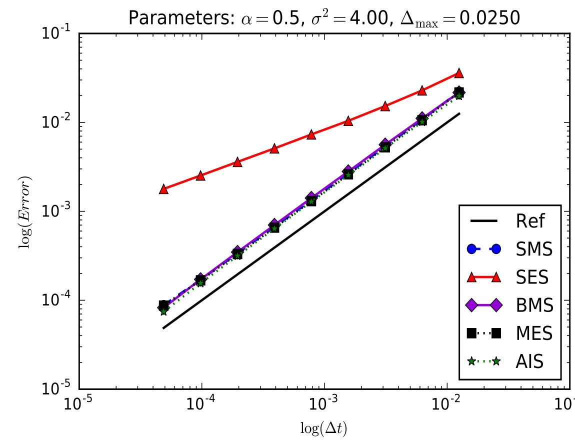
Table 3.4.4: Empirical rate of convergence $\hat{\rho}$ for the $L^1(\Omega)$-error, when $\alpha > \frac{1}{2}$ for different values of α and σ^2.

We have observed in the numerical experiments three cases for the parameters. The first one is when $b(0) > 2\alpha(1-\alpha)^2\sigma^2$ ($\sigma^2 = 49$ and $\sigma^2 = 64$). In this case, Theorem 3.1.6 holds and we observe the order one convergence (see Figures 3.4.2a, 3.4.2b, and first and fourth row in Table 3.4.4). The second case is when the parameters do not satisfy $b(0) > 2\alpha(1-\alpha)^2\sigma^2$ ($\sigma^2 = 53.29$ and $\sigma^2 = 81$), and then we can not apply Theorem 3.1.6, but in the numerical simulations we still observe the order one convergence (see Figures 3.4.2c, 3.4.2d, and second and fifth row in Table 3.4.4). Finally the third case, is when $\sigma \gg b(0)$, and then we do not observe a linear convergence anymore (see Figures 3.4.2e, 3.4.2f, and third and six row in Table 3.4.4). Notice that in the three cases the SMS performs better than the BMS, specially when σ^2 grows. (see Table 3.4.4).

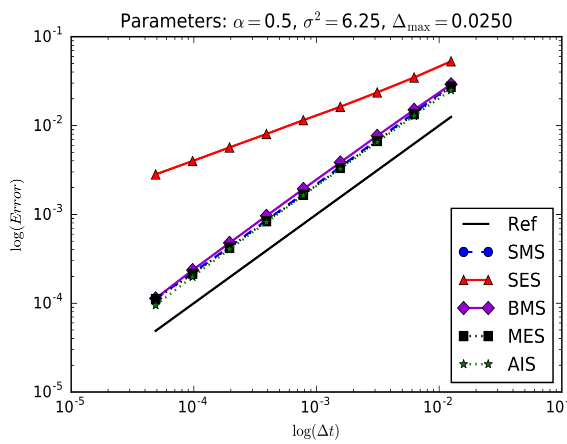
The second and third case show us that some restriction has to be impose on the parameters to observe the convergence of order one. But our restriction, although sufficient, it seems to be too strong, specially for α close to one.



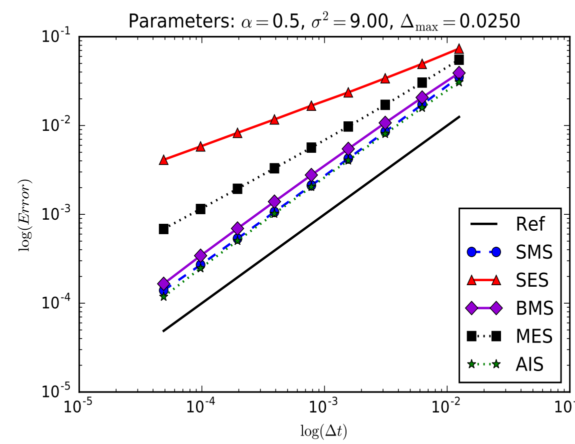
(a) Parameters in case 1: $b(0) > 6\sigma^2$.



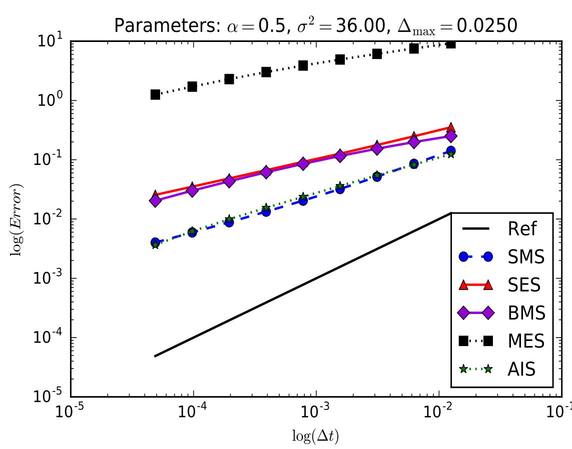
(b) Parameters in case 2: $b(0) \in (5\sigma^2/2, 6\sigma^2)$.



(c) Parameters in case 3: $b(0) \in (3\sigma^2/2, 5\sigma^2/2)$.

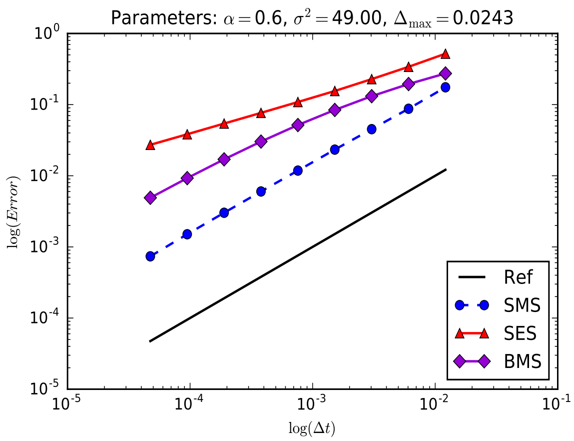


(d) Parameters in case 4: $b(0) \in (\sigma^2, 3\sigma^2/2)$.

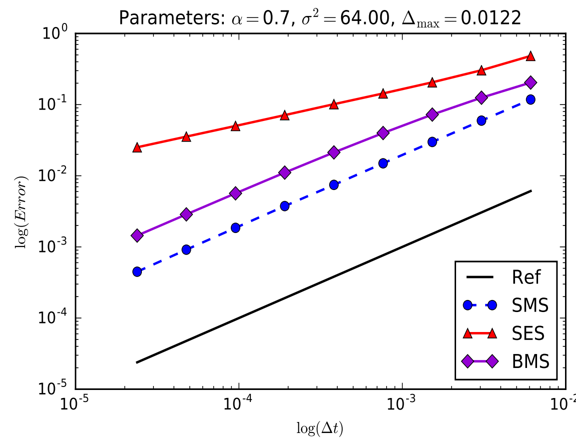


(e) Parameters in case 5: $b(0) < \sigma^2$.

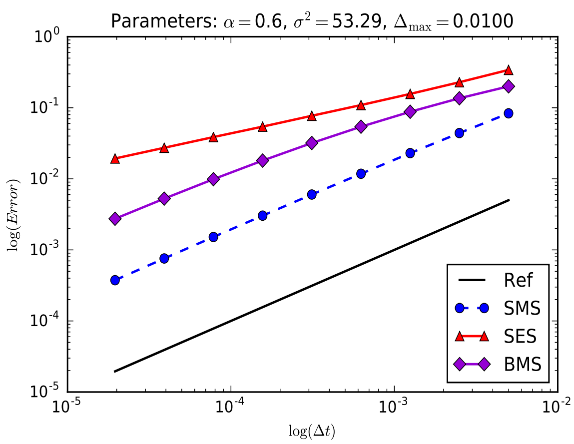
Figure 3.4.1: Step size Δt versus the estimated $L^1(\Omega)$-strong error for the CIR Process (Log-Log scale). The identity map serves as a reference line of rate one.



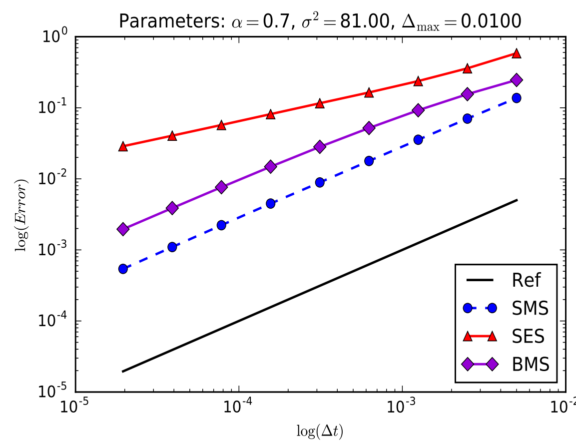
(a) Parameters in case 1: $\alpha = 0.60$.



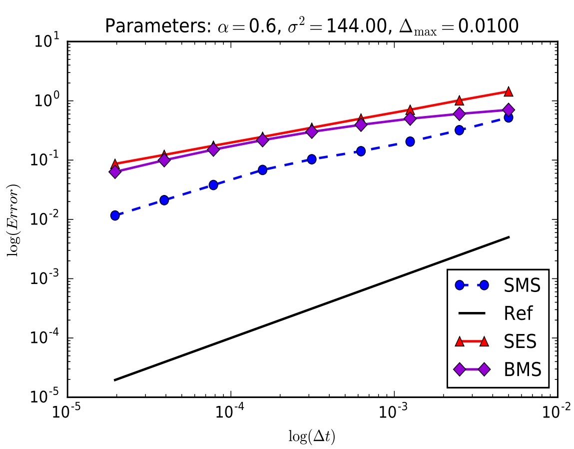
(b) Parameters in case 1: $\alpha = 0.70$.



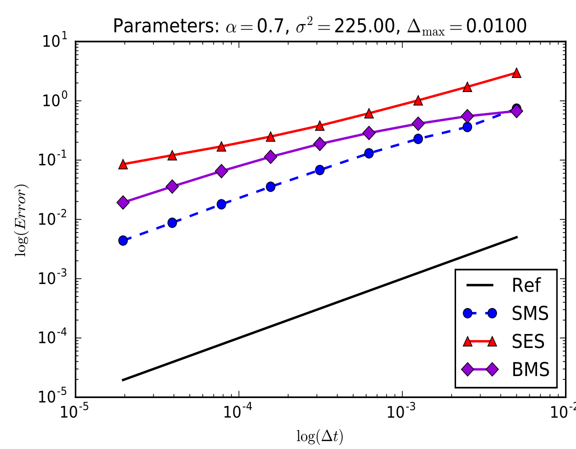
(c) Parameters in case 2: $\alpha = 0.60$.



(d) Parameters in case 2: $\alpha = 0.70$.



(e) Parameters in case 3: $\alpha = 0.60$



(f) Parameters in case 3: $\alpha = 0.70$

Figure 3.4.2: Step size Δt versus the estimated $L^1(\Omega)$-error for $\alpha > \frac{1}{2}$ and different values for σ^2 (Log-Log scale). The identity map serves as a reference line of rate one. .

3.4.2 Application of the SMS in Multilevel Monte Carlo

We continue this section by testing the SMS in the context of a multilevel Monte Carlo application widely used nowadays in computational finance (see e.g. [40] and references therein). Multilevel Monte Carlo is an efficient technique introduced by Giles [35] to decrease the computational complexity of an estimator combining Monte Carlo simulation and time discretization scheme for a given threshold in the accuracy. For details we refer to [34, 35, 40].

For this experiment, we consider the classical but non trivial test-case of the Zero Coupon Bound (ZCB) pricing of maturity T,

$$B(0,T) = \mathbb{E}\left[\exp\Big(-\int_0^T r_s ds\Big)\right],$$

under the hypothesis that the short interest rate dynamics $(r_t, t \geq 0)$ is modeled with a CIR process ($\alpha = \frac{1}{2}$ and $b(x) = a - bx$) :

$$dr_t = (a - br_t)dt + \sigma\sqrt{r_t}dW_t.$$

In this context, the price of the ZCB admits a wellknown closed-form solution given by (see e.g [20, 54])

$$B(0,T) = A(T)e^{-B(T)r_0},$$

where r_0 is the initial value of the interest rate, and for $\lambda = \sqrt{b^2 + 2\sigma^2}$

$$A(T) = \left[\frac{2\lambda e^{(b+\lambda)T/2}}{(\lambda+b)(e^{\lambda T}-1)+2\lambda}\right]^{\frac{2a}{\sigma^2}}, \quad B(T) = \frac{2(e^{\lambda T}-1)}{(\lambda+b)(e^{\lambda T}-1)+2\lambda}.$$

Let $\mathbb{E}\widehat{B}(\Delta t_{(l)})$ a discrete-time weak approximation of $B(0,T)$ with the time step $\Delta t_{(l)}$. We consider the L-level Monte Carlo estimator :

$$\widehat{Y}_T = \frac{1}{N_0}\sum_{i=1}^{N_0}\widehat{B}^{(i)}(\Delta t_{(0)}) + \sum_{l=1}^{L}\frac{1}{N_l}\sum_{i=1}^{N_l}\Big(\widehat{B}^{(i)}(\Delta t_{(l)}) - \widehat{B}^{(i)}(\Delta t_{(l-1)})\Big).$$

For a targeted mean-square error ϵ^2 on the computation of the quantity $B(0,T)$

$$\mathbb{E}[(\widehat{Y}_T - B(0,T))^2] = \mathcal{O}(\epsilon^2),$$

one can choose the following a priori parametrization of the MLMC method in order to minimize the computational time (complexity) (see [34, 35, 40]): we use the estimation $L = \dfrac{\log\epsilon^{-1}}{\log 2}$; from one level to the next one the time step is divided by 2, $\Delta t_{(l)} = \frac{1}{2^{(l+1)}}$; the number of trajectories to simulate is estimated with Giles formula [35]

$$N_l = \frac{2}{\epsilon^2}\sqrt{V_l\Delta t_l}\left(\sum_{l=0}^{L}\sqrt{V_l/\Delta t_{(l)}}\right),$$

with $V_l = \overline{\mathbb{V}ar}\Big(\widehat{B}^{(1)}(\Delta t_{(l)}) - \widehat{B}^{(1)}(\Delta t_{(l-1)})\Big)$. As an estimator for the bias variance, the strong rate of convergence of the discretisation scheme enters as a key ingredient in the N_l a priori estimation. A scheme with a reduced strong bias will then allow a smaller N_l. We apply the MLMC computation for the SMS, the PMS, the AIS and the BMS.

We summarize the results of the performance comparison between the four schemes in Table 3.4.5. The computation have been run using a initial interest rate $r_0 = 1$, the maturity of the bond $T = 1$, the drift parameters $a = b = 10$, the volatility $\sigma = 1$. For the MLMC simulation, we fix a minimum number of trajectories equal to 500, and a minimal number of levels equal to 6.

In Table 3.4.5, we give the measures of the CPU time for a set of three decreasing targeted errors, as long as the effective measured error and the total number of simulated trajectories. As expected, the required threshold error has been reached by the MLMC strategy. As also expected (see Giles [34]), Milstein schemes perform better than their Euler versions. Finally, as ϵ decreases, the SMS clearly performs better than his PMS version.

ϵ = 1.0e-03 ($L = 9$, $\Delta t_{(L)} = 1/2^{10}$)	SMS	PMS	AIS	BMS
CPU time ($N_0 + \cdots + N_L$) (observed error)	0.2304 (792 651) (1.970e-05)	0.2657 (950 838) (3.347e-04)	0.264 (990 769) (3.132e-04)	0.274 (992 432) (3.292e-04)
ϵ = 1.0e-04 ($L = 13$, $\Delta t_{(L)} = 1/2^{14}$)	SMS	PMS	AIS	BMS
CPU time ($N_0 + \cdots + N_L$) (observed error)	16.871 (56 229 224) (4.870e-05)	20.843 (70 876 600) (1.091e-04)	17.311 (73 824 621) (9.538e-06)	16.95 (73 668 115) (2.203e-06)
ϵ = 1.0e-05 ($L = 16$, $\Delta t_{(L)} = 1/2^{17}$)	SMS	PMS	AIS	BMS
CPU time ($N_0 + \cdots + N_L$) (observed error)	1589.6 (5 531 879 264) (4.752e-06)	1910.7 (6 913 546 698) (5.889e-06)	1576.4 (7 368 734 119) (3.912e-06)	1540.2 (7 333 474 098) (5.653e-07)

Table 3.4.5: CPU time to achieve the target error for the different schemes. The observed error is $|\widehat{Y}_T - B(0,T)|$.

3.4.3 Conclusion

In this chapter we have recovered the classical rate of convergence of the Milstein scheme in a context of non smooth diffusion coefficient, although we have to impose some restrictions over the parameters of the SDE (3.1.1) to ensure the theoretical order one of convergence. Typically, if the quotient $b(0)/\sigma^2$ is big enough we will observe the optimal convergence rate.

In the numerical simulations we have observed that, despite the fact it is necessary to impose some restriction over the parameters of SDE (3.1.1) to obtain the order one convergence, Hypothesis 3.1.5 seems to be not optimal, specially for $\alpha = \frac{1}{2}$. Also, through numerical simulations, we have observed that the use of SMS could improve the computation times in a Multilevel Monte Carlo framework, at least as well as the (CIR specialized) one-order schemes.

Although our result seems more restrictive in term of hypotheses on the set of parameters, in particular if we compare SMS with Lamperti's transformation-based schemes (see the recent works in [3] and [19], the SMS can be applied to a more general class of drifts functions and in various contexts. It is thus a useful complement of the existing literature.

3.5 Proofs for preliminary lemmas

3.5.1 On the Positive Moments of the SMS

Proof of Lemma 3.1.4. Let us recall the notations $\Delta s = s - \eta(s)$, and $\Delta W_s = W_s - W_{\eta(s)}$. Let us define $\tau_m = \inf\{t \geq 0 : \overline{X}_t \geq m\}$. Then by ItÃť's Formula, Young's inequality and the Lipschitz property of b, we have

$$\begin{aligned}\mathbb{E}[\overline{X}_{t\wedge\tau_m}^{2p}] \leq x_0^{2p} + C\mathbb{E}\left[\int_0^{t\wedge\tau_m} \overline{X}_s^{2p} + C + \overline{X}_{\eta(s)}^{2p} ds\right] \\ + C\mathbb{E}\left[\int_0^{t\wedge\tau_m} \left(\sigma\overline{X}_{\eta(s)}^{\alpha} + \alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s\right)^{2p} ds\right].\end{aligned} \tag{3.5.1}$$

From the definition of $\overline{X}$, a straightforward computation shows that for all $s \in [0,t]$ almost surely

$$\overline{X}_s^{2p} \leq C\left(1 + \overline{X}_{\eta(s)}^{2p} + \Delta W_s^{\frac{2p}{1-\alpha}} + \left(\Delta W_s^2 - \Delta s\right)^{\frac{2p}{2(1-\alpha)}}\right).$$

Putting this in (3.5.1), we have

$$\begin{aligned}
\mathbb{E}[\overline{X}^{2p}_{t\wedge\tau_m}] &\le x_0^{2p} + C\mathbb{E}\left[\int_0^{t\wedge\tau_m} 1 + \overline{X}^{2p}_{\eta(s)} + \left(\sigma\overline{X}^{\alpha}_{\eta(s)} + \alpha\sigma^2\overline{X}^{2\alpha-1}_{\eta(s)}\Delta W_s\right)^{2p} ds\right] \\
&\quad + C\mathbb{E}\left[\int_0^{t\wedge\tau_m} \Delta W_s^{\frac{2p}{1-\alpha}} + \left(\Delta W_s^2 - (s-\eta(s))\right)^{\frac{2p}{2(1-\alpha)}} ds\right] \\
&\le x_0^{2p} + C\mathbb{E}\left[\int_0^{t\wedge\tau_m} 1 + \overline{X}^{2p}_{\eta(s)} + \overline{X}^{2p\alpha}_{\eta(s)} + \overline{X}^{2p(2\alpha-1)}_{\eta(s)}\Delta W_s^{2p} ds\right] \\
&\quad + C\int_0^T \mathbb{E}\left[\Delta W_s^{\frac{2p}{1-\alpha}}\right] + \mathbb{E}\left[\left(\Delta W_s^2 - \Delta s\right)^{\frac{2p}{2(1-\alpha)}}\right] ds.
\end{aligned}$$

Since $\alpha \in [\frac{1}{2}, 1)$ we have $\overline{X}^{2p\alpha}_{\eta(s)} \le 1 + \overline{X}^{2p}_{\eta(s)}$, and then, using Young's Inequality and the finiteness of the moments of Gaussian random variables, we conclude

$$\mathbb{E}[\overline{X}^{2p}_{t\wedge\tau_m}] \le Cx_0^{2p} + C\mathbb{E}\left[\int_0^{t\wedge\tau_m} \overline{X}^{2p}_{\eta(s)} ds\right] \le Cx_0^{2p} + C\int_0^t \sup_{u\le s}\mathbb{E}[\overline{X}^{2p}_{u\wedge\tau_m}]ds.$$

Since the right-hand side is increasing, we can take supremum in the left-hand side and from here, applying Gronwall's inequality, and taking $m\to\infty$ we get

$$\sup_{t\le T}\mathbb{E}[\overline{X}^{2p}_t] \le Cx_0^{2p}.$$

From here, following standard argument using Burkholder-Davis-Gundy inequality we can conclude on Lemma 3.1.4. □

3.5.2 On the Local Error of the SMS

Proof of Lemma 3.2.1. From the definition of $\overline{X}$, and the algebraic inequality for positive real numbers $(a_1 + \ldots + a_n)^p \le n^p(a_1^p + \ldots + a_n^p)$ we have

$$\begin{aligned}
\left|\overline{X}_t - \overline{X}_{\eta(t)}\right|^{2p} \le 3^{2p}\Big(& b(\overline{X}_{\eta(t)})^{2p}(t-\eta(t))^{2p} + \sigma^{2p}\overline{X}^{2\alpha p}_{\eta(t)}(W_t - W_{\eta(t)})^{2p} \\
&+ \frac{\alpha^{2p}\sigma^{4p}}{2^{2p}}\overline{X}^{(2\alpha-1)2p}_{\eta(t)}\left[(W_t - W_{\eta(t)})^2 - (t-\eta(t))\right]^{2p}\Big).
\end{aligned}$$

Thanks to the linear growth of b, Lemma 3.1.4 and the properties of the Brownian Motion it is quite easy to conclude the existence of a constant C such that

$$\mathbb{E}\left[|\overline{X}_t - \overline{X}_{\eta(t)}|^{2p}\right] \le C\Delta t^p,$$

from where the result follows. □

3.5.3 On the Probability of SMS being close to zero

From $b_\sigma(\alpha)$ and $K(\alpha)$ defined in (3.1.9), let us recall the notation

$$\bar{x}(\alpha) := \frac{b_\sigma(\alpha)}{K(\alpha)}$$

introduced in Lemma 3.2.3. As $b_\sigma(\alpha) > 0$ under Hypothesis 3.1.5-(i), $\bar{x}(\alpha)$ is bounded away from 0. In particular,

$$\lim_{\alpha\to\frac{1}{2}} \bar{x}(\alpha) = \frac{(b(0) - \sigma^2/4)}{K}, \quad \text{whereas} \quad \lim_{\alpha\to 1} \bar{x}(\alpha) = \frac{b(0)}{K + \sigma^2/2}.$$

Proof of Lemma 3.2.3. Denoting $\Delta W_s = (W_s - W_{\eta(s)})$, and $\Delta s = s - \eta(s)$, we have for all $s \in [0, T]$,

$$\overline{Z}_s = \frac{\alpha\sigma^2}{2}\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s^2 + \sigma\overline{X}_{\eta(s)}^{\alpha}\Delta W_s + \overline{X}_{\eta(s)} + \left(b(\overline{X}_{\eta(s)}) - \frac{\alpha\sigma^2}{2}\overline{X}_{\eta(s)}^{2\alpha-1}\right)\Delta s.$$

From the Lipschitz property of b and the following bound for any $x > 0$

$$x^{2\alpha-1} \leq 4(1-\alpha)^2 + (2\alpha-1)[2(1-\alpha)]^{-\frac{2(1-\alpha)}{2\alpha-1}} x, \tag{3.5.2}$$

we have

$$\overline{Z}_s \geq \frac{\alpha\sigma^2}{2}\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s^2 + \sigma\overline{X}_{\eta(s)}^{\alpha}\Delta W_s + \overline{X}_{\eta(s)} + \left(b_\sigma(\alpha) - K(\alpha)\overline{X}_{\eta(s)}\right)\Delta s. \tag{3.5.3}$$

So,

$$\begin{aligned}
&\mathbb{P}\Big[\overline{Z}_s \leq (1-\rho)b_\sigma(\alpha)\Delta s,\ \overline{X}_{\eta(s)} < \rho\bar{x}(\alpha)\Big] \\
&\quad\leq \mathbb{P}\Big[\frac{\alpha\sigma^2}{2}\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s^2 + \sigma\overline{X}_{\eta(s)}^{\alpha}\Delta W_s + \overline{X}_{\eta(s)} + \left(\rho b_\sigma(\alpha) - K(\alpha)\overline{X}_{\eta(s)}\right)\Delta s \leq 0,\ \overline{X}_{\eta(t)} < \rho\bar{x}(\alpha)\Big].
\end{aligned}$$

From the independence of ΔW_s with respect to $\mathcal{F}_{\eta(s)}$, if we denote by $\mathcal{N}$ a standard Gaussian variable, we have

$$\begin{aligned}
&\mathbb{P}\Big[\overline{Z}_s \leq (1-\rho)b_\sigma(\alpha)\Delta s,\ \overline{X}_{\eta(s)} < \rho\bar{x}(\alpha)\Big] \\
&\quad\leq \mathbb{E}\Big[\mathbb{P}\Big(\frac{\alpha\sigma^2}{2}x^{2\alpha-1}\Delta s\mathcal{N}^2 + \sigma\sqrt{\Delta s}x^\alpha\mathcal{N} + x + [\rho b_\sigma(\alpha) - K(\alpha)x]\,\Delta s \leq 0\Big)\Big|_{x=\overline{X}_{\eta(s)}} \mathbb{1}_{\{\overline{X}_{\eta(s)}<\rho\bar{x}(\alpha)\}}\Big].
\end{aligned}$$

Notice that in the right-hand side we have a quadratic polynomial of a standard Gaussian random variable. Let us compute its discriminant:

$$\begin{aligned}
\Delta(x,\alpha) &= \sigma^2 x^{2\alpha}\Delta s - 2\alpha\sigma^2 x^{2\alpha-1}\Delta s\left(x + \left(\rho b_\sigma(\alpha) - K(\alpha)x\right)\Delta s\right) \\
&= -(2\alpha-1)\sigma^2 x^{2\alpha}\Delta s - 2\alpha\sigma^2 x^{2\alpha-1}\Delta s^2\left(\rho b_\sigma(\alpha) - K(\alpha)x\right).
\end{aligned}$$

Since $b_\sigma(\alpha) > 0$, we have $\Delta(x, \alpha) < 0$ for all $\alpha \in [\frac{1}{2}, 1)$, and $x \leq \rho\bar{x}(\alpha)$. So, for all $\Delta s \leq \Delta t$ we have

$$\mathbb{P}\left[\overline{Z}_s \leq (1-\rho)b_\sigma(\alpha)\Delta s,\ \overline{X}_{\eta(s)} < \rho\bar{x}(\alpha)\right] = 0,$$

taking $\Delta s = \Delta t$ we conclude on the Lemma. □

Proof of Lemma 3.2.2. We have

$$\mathbb{P}\left(\inf_{t_k\le s<t_{k+1}} \overline{Z}_s \le 0\right) = \mathbb{P}\left(\inf_{t_k\le s<t_{k+1}} \overline{Z}_s \le 0, \overline{X}_{t_k} \ge \bar{x}(\alpha)\right) + \mathbb{P}\left(\inf_{t_k\le s<t_{k+1}} \overline{Z}_s \le 0, \overline{X}_{t_k} < \bar{x}(\alpha)\right) \tag{3.5.4}$$

We start with the second term in the right hand of the last inequality. By continuity of the path of $\overline{Z}$ and Lemma 3.2.3, we have

$$\mathbb{P}\left(\inf_{t_k\le s<t_{k+1}} \overline{Z}_s \le 0, \overline{X}_{t_k} < \bar{x}(\alpha)\right) = \sum_{s\in\mathbb{Q}\cap(t_k,t_{k+1}]} \mathbb{P}\left(\overline{Z}_s \le 0, \overline{X}_{t_k} < \bar{x}(\alpha)\right) = 0.$$

On the other hand, from 3.5.3 we have

$$\overline{Z}_s \ge \sigma\overline{X}^{\alpha}_{\eta(s)}\Delta W_s + (1 - K(\alpha)\Delta s)\,\overline{X}_{\eta(s)} + b_\sigma(\alpha)\Delta s. \tag{3.5.5}$$

Then

$$\begin{aligned}
&\mathbb{P}\Big(\inf_{t_k\le s<t_{k+1}} \overline{Z}_s \le 0,\ \overline{X}_{t_k} \ge \bar{x}(\alpha)\Big)\\
&\le \mathbb{P}\Big(\inf_{t_k< s\le t_{k+1}} \frac{\overline{X}^{1-\alpha}_{\eta(s)}}{\sigma} + \frac{\left(b_\sigma(\alpha) - K(\alpha)\overline{X}_{\eta(s)}\right)\Delta s}{\sigma\overline{X}^{\alpha}_{\eta(s)}} + \Delta W_s \le 0,\ \ \overline{X}_{t_k} \ge \bar{x}(\alpha)\Big)\\
&= \mathbb{E}\Big[\psi(\overline{X}_{t_k})\mathbb{1}_{\{\overline{X}_{t_k}\ge\bar{x}(\alpha)\}}\Big],
\end{aligned}$$

where the last equality holds thanks to the Markov Property of the Brownian motion, for

$$\psi(x) = \mathbb{P}\Big(\inf_{0<u\le\Delta t} \frac{x^{1-\alpha}}{\sigma} + \frac{b_\sigma(\alpha) - K(\alpha)x}{\sigma x^{\alpha}}u + B_u \le 0\Big),$$

where (B_t) denotes a Brownian Motion independent of (W_t).

If $(B^{\mu}_t, 0 \le t \le T)$ is a Brownian motion with drift μ, starting at y_0, then for all $y \le y_0$, we have (see [13]):

$$\mathbb{P}\left(\inf_{0<s\le t} B^{\mu}_s \le y\right) = \frac{1}{2}\operatorname{erfc}\left(\frac{y_0 - y}{\sqrt{2t}} + \frac{\mu\sqrt{t}}{\sqrt{2}}\right) + \frac{1}{2}\exp\left(-2\mu(y_0 - y)\right)\operatorname{erfc}\left(\frac{y_0 - y}{\sqrt{2t}} - \frac{\mu\sqrt{t}}{\sqrt{2}}\right), \tag{3.5.6}$$

where for $z \in \mathbb{R}$, $\operatorname{erfc} z = \sqrt{2/\pi}\int_{\sqrt{2}z}^{\infty} \exp\left(-u^2/2\right) du$. In our case $y_0 = x^{1-\alpha}/\sigma$, $\mu = (b_\sigma(\alpha) - K(\alpha)x)/\sigma x^{\alpha}$, and $y = 0$. Then

$$\psi(x) = \frac{1}{2}\operatorname{erfc}\left(\frac{[(1 - K(\alpha)\Delta t)x + b_\sigma(\alpha)\Delta t]}{\sqrt{2\Delta t}\sigma x^{\alpha}}\right) + \frac{1}{2}\exp\left(-\frac{2[b_\sigma(\alpha) - K(\alpha)x]x}{\sigma^2 x^{2\alpha}}\right)\operatorname{erfc}\left(\frac{x - [b_\sigma(\alpha) - K(\alpha)x]\Delta t}{\sqrt{2\Delta t}\sigma x^{\alpha}}\right).$$

Since $\Delta t \leq 1/2K(\alpha)$, for any $x \geq \bar{x}(\alpha)$, the arguments in the erfc function in the last equality are both positives, and then recalling that for all $z > 0$ $\mathrm{erfc}(z) \leq \exp(-z^2)$, we obtain

$$\begin{aligned}\psi(x) &\leq \frac{1}{2}\exp\left(-\frac{[(1-K(\alpha)\Delta t)x + b_\sigma(\alpha)\Delta t]^2}{2\Delta t\sigma^2 x^{2\alpha}}\right)\\ &\quad + \frac{1}{2}\exp\left(-\frac{2[b_\sigma(\alpha)-K(\alpha)x]x}{\sigma^2 x^{2\alpha}}\right)\exp\left(-\frac{[x-[b_\sigma(\alpha)-K(\alpha)x]\Delta t]^2}{2\Delta t\sigma^2 x^{2\alpha}}\right)\\ &\leq \exp\left(-\frac{[(1-K(\alpha)\Delta t)x + b_\sigma(\alpha)\Delta t]^2}{2\Delta t\sigma^2 x^{2\alpha}}\right).\end{aligned}$$

So, for all $x \geq \bar{x}(\alpha)$

$$\psi(x) \leq \exp\left(-\frac{(1-K(\alpha)\Delta t)^2 x^{2(1-\alpha)}}{2\sigma^2\Delta t}\right).$$

Then

$$\begin{aligned}\mathbb{P}\Big(\inf_{t_k\leq s<t_{k+1}} \overline{Z}_s \leq 0,\ \overline{X}_{t_k} \geq \bar{x}(\alpha)\Big) &\leq \mathbb{E}\Big[\exp\left(-\frac{(1-K(\alpha)\Delta t)^2\overline{X}_{t_k}^{2(1-\alpha)}}{2\sigma^2\Delta t}\right)\mathbb{1}_{\{\overline{X}_{t_k}\geq\bar{x}(\alpha)\}}\Big]\\ &\leq \exp\left(-\frac{(1-K(\alpha)\Delta t)^2\bar{x}(\alpha)^{2(1-\alpha)}}{2\sigma^2\Delta t}\right),\end{aligned}$$

and finally, choosing $\gamma = \bar{x}(\alpha)^{2(1-\alpha)}/8\sigma^2$, we get

$$\mathbb{P}\Big(\inf_{t_k\leq s<t_{k+1}} \overline{Z}_s \leq 0\Big) \leq \exp\Big(-\frac{\gamma}{\Delta t}\Big).$$

□

3.5.4 On the Local Time of the SMS at Zero

The Stopping Times $(\Theta_\alpha, \frac{1}{2} \leq \alpha < 1)$

In what follows, we consider

$$\Theta_\alpha = \inf\left\{s > 0 : \overline{X}_s < (1-\sqrt{\alpha})b_\sigma(\alpha)\Delta t\right\}. \tag{3.5.7}$$

Lemma 3.5.1. *Assume $b(0) > 2\alpha(1-\alpha)^2\sigma^2$, and $\Delta t \leq 1/(2K(\alpha)) \wedge x_0/[(1-\sqrt{\alpha})b_\sigma(\alpha)]$. Then there exists a positive constant γ depending on α, $b(0)$, K and σ but not on Δt such that*

$$\mathbb{P}(\Theta_\alpha \leq T) \leq \frac{T}{\Delta t}\exp\left(\frac{-\gamma}{\Delta t}\right). \tag{3.5.8}$$

Proof. First, notice that the condition $\Delta t < x_0/[(1-\sqrt{\alpha})b_\sigma(\alpha)]$ ensures that the stopping time Θ_α is almost surely strictly positive.

To enlighten the notation along this proof, let us call $l_\sigma(\alpha) := (1-\sqrt{\alpha})b_\sigma(\alpha)$, and $\zeta_k = \inf_{t_k<s\leq t_{k+1}} \overline{Z}_s$. We split the proof in three steps.

Step 1. Let us prove that for a suitable function $\psi : \mathbb{R} \to [0,1]$ and the set $A_k = \{\overline{X}_{t_k} > \bar{x}(\alpha)\sqrt{\alpha}\} \in \mathcal{F}_{t_k}$:

$$\mathbb{P}\left(\Theta_\alpha \le T\right) \le \sum_{k=0}^{N-1} \mathbb{E}\left(\psi(\overline{X}_{t_k})\mathbb{1}_{A_k}\right) \tag{3.5.9}$$

Indeed,

$$\mathbb{P}\left(\Theta_\alpha \le T\right) \le \sum_{k=0}^{N-1} \mathbb{P}\left(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} > l_\sigma(\alpha)\Delta t\right).$$

But, for each $k = 0, \dots, N-1$

$$\begin{aligned}\mathbb{P}\left(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} > l_\sigma(\alpha)\Delta t\right) &= \mathbb{P}\left(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} > l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} < \bar{x}(\alpha)\sqrt{\alpha}\right)\\ &+ \mathbb{P}\left(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} > l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} \ge \bar{x}(\alpha)\sqrt{\alpha}\right).\end{aligned}$$

Since $l_\sigma(\alpha)\Delta t = (1-\sqrt{\alpha})b_\sigma(\alpha)\Delta t \le \bar{x}(\alpha)\sqrt{\alpha}$, we have

$$\begin{aligned}\mathbb{P}\Big(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} > l_\sigma(\alpha)\Delta t,\ \overline{X}_{t_k} < \bar{x}(\alpha)\sqrt{\alpha}\Big)&\\ &\le \mathbb{P}\left(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} < \bar{x}(\alpha)\sqrt{\alpha}\right)\\ &\le \sum_{s \in \mathbb{Q}\cap(t_k,t_{k+1}]} \mathbb{P}\left(\overline{Z}_s \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} < \bar{x}(\alpha)\sqrt{\alpha}\right) = 0,\end{aligned}$$

thanks to Lemma 3.2.3. On the other hand, we have

$$\begin{aligned}&\mathbb{P}\Big(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \overline{X}_{t_k} > l_\sigma(\alpha)\Delta t,\ \overline{X}_{t_k} \ge \bar{x}(\alpha)\sqrt{\alpha}\Big)\\ &\quad = \mathbb{P}\left(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} \ge \bar{x}(\alpha)\sqrt{\alpha}\right)\\ &\quad \le \mathbb{P}\Big(\inf_{t_k < s \le t_{k+1}} \frac{\overline{X}^{1-\alpha}_{\eta(s)}}{\sigma} + \frac{\left(b_\sigma(\alpha) - K(\alpha)\overline{X}_{\eta(s)}\right)\Delta s}{\sigma \overline{X}^{\alpha}_{\eta(s)}} + \Delta W_s \le \frac{l_\sigma(\alpha)\Delta t}{\sigma\overline{X}^{\alpha}_{\eta(s)}},\ \ \overline{X}_{t_k} \ge \bar{x}(\alpha)\sqrt{\alpha}\Big)\\ &\quad = \mathbb{E}\Big[\psi(\overline{X}_{t_k})\mathbb{1}_{\{\overline{X}_{t_k} > \bar{x}(\alpha)\sqrt{\alpha}\}}\Big],\end{aligned}$$

where the inequality comes from (3.5.3), and the last equality holds thanks to the Markov Property of the Brownian motion, for

$$\psi(x) = \mathbb{P}\Big(\inf_{0<u\le\Delta t} \frac{x^{1-\alpha}}{\sigma} + \frac{b_\sigma(\alpha) - K(\alpha)x}{\sigma x^\alpha}u + B_u \le \frac{l_\sigma(\alpha)\Delta t}{\sigma x^\alpha}\Big),$$

where (B_t) denotes a Brownian Motion independent of (W_t). Summarizing

$$\mathbb{P}\left(\zeta_k \le l_\sigma(\alpha)\Delta t,\ \ \overline{X}_{t_k} > l_\sigma(\alpha)\Delta t\right) \le \mathbb{E}\Big[\psi(\overline{X}_{t_k})\mathbb{1}_{\{\overline{X}_{t_k} > \bar{x}(\alpha)\sqrt{\alpha}\}}\Big],$$

and we have (3.5.9) for $A_k = \left\{\overline{X}_{t_k} > \bar{x}(\alpha)\sqrt{\alpha}\right\}$.

Step 2. Let us prove that for all $x \geq \bar{x}(\alpha)\sqrt{\alpha}$:

$$\psi(x) \leq \exp\left(-\frac{(1-K(\alpha)\Delta t)^2(\bar{x}(\alpha)\sqrt{\alpha})^{2(1-\alpha)}}{2\sigma^2\Delta t}\right). \tag{3.5.10}$$

Applying again (3.5.6), we have

$$\begin{aligned}\psi(x) &= \frac{1}{2}\operatorname{erfc}\left(\frac{[(1-K(\alpha)\Delta t)x+\sqrt{\alpha}b_\sigma(\alpha)\Delta t]}{\sqrt{2\Delta t}\sigma x^\alpha}\right)\\ &\quad+\frac{1}{2}\exp\left(-\frac{2[b_\sigma(\alpha)-K(\alpha)x][x-(1-\sqrt{\alpha})b_\sigma(\alpha)\Delta t]}{\sigma^2 x^{2\alpha}}\right)\\ &\quad\times\operatorname{erfc}\left(\frac{1}{\sqrt{2\Delta t}\sigma x^\alpha}\left[(1+K(\alpha)\Delta t)x-(2-\sqrt{\alpha})b_\sigma(\alpha)\Delta t\right]\right)\\ &=: A(x)+B(x).\end{aligned}$$

Since $\Delta t \leq 1/(2K(\alpha))$, and $\operatorname{erfc}(z) \leq \exp(-z^2)$ for all $z > 0$ we have

$$A(x) \leq \frac{1}{2}\exp\left(-\frac{\left[(1-K(\alpha)\Delta t)x+\sqrt{\alpha}b_\sigma(\alpha)\Delta t\right]^2}{2\sigma^2\Delta t x^{2\alpha}}\right) \leq \frac{1}{2}\exp\left(-\frac{(1-K(\alpha)\Delta t)^2 x^{2(1-\alpha)}}{2\sigma^2\Delta t}\right).$$

On the other hand, for $x \geq \bar{x}(\alpha)\sqrt{\alpha}$, and $\Delta t \leq 1/(2K(\alpha))$, it follows

$$x > (2-\sqrt{\alpha})b_\sigma(\alpha)\Delta t/(1+K\Delta t),$$

so the argument of the function erfc in B is positive, and then

$$\begin{aligned}B(x) &\leq \frac{1}{2}\exp\left(-\frac{2[b_\sigma(\alpha)-K(\alpha)x][x-(1-\sqrt{\alpha})b_\sigma(\alpha)\Delta t]}{\sigma^2 x^{2\alpha}}\right)\\ &\quad\times\exp\left(-\frac{\left[(1+K(\alpha)\Delta t)x-(2-\sqrt{\alpha})b_\sigma(\alpha)\Delta t\right]^2}{2\Delta t\sigma^2 x^{2\alpha}}\right)\\ &= \frac{1}{2}\exp\left(-\frac{\left[(1-K(\alpha)\Delta t)x+\sqrt{\alpha}b_\sigma(\alpha)\Delta t\right]^2}{2\Delta t\sigma^2 x^{2\alpha}}\right)\\ &\leq \frac{1}{2}\exp\left(-\frac{(1-K(\alpha)\Delta t)^2 x^{2(1-\alpha)}}{2\sigma^2\Delta t}\right).\end{aligned}$$

So

$$\psi(x) = A(x)+B(x) \leq \exp\left(-\frac{(1-K(\alpha)\Delta t)^2 x^{2(1-\alpha)}}{2\sigma^2\Delta t}\right),$$

and since the right-hand side is decreasing on x, we have (3.5.10).

Step 3. Let us conclude. Putting together (3.5.9) and (3.5.10) we have

$$\begin{aligned}\mathbb{P}\left(\Theta_\alpha \leq T\right) &\leq \sum_{k=0}^{N-1} \mathbb{E}\left(\psi(\overline{X}_{t_k})\mathbb{1}_{\{\overline{X}_{t_k}>\bar{x}(\alpha)\sqrt{\alpha}\}}\right) \\ &\leq \sum_{k=0}^{N-1} \exp\left(-\frac{(1-K(\alpha)\Delta t)^2(\bar{x}(\alpha)\sqrt{\alpha})^{2(1-\alpha)}}{2\sigma^2\Delta t}\right)\mathbb{P}\left(\overline{X}_{t_k}>\bar{x}(\alpha)\sqrt{\alpha}\right) \\ &\leq \frac{C}{\Delta t}\exp\left(-\frac{\gamma}{\Delta t}\right)\end{aligned}$$

with $\gamma = (\bar{x}(\alpha)\sqrt{\alpha})^{2(1-\alpha)}/(8\sigma^2)$. □

Proof of Lemma 3.2.4. From (3.1.4), standard arguments show that $\mathbb{E}[L_T^0(\overline{X})^4] \leq C(T)$. On the other hand, thanks to Corollary VI.1.9 on Revuz and Yor [72, p. 212], we have almost surely

$$L^0_{T\wedge\Theta_\alpha}(\overline{X}) = \lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_0^{T\wedge\Theta_\alpha}\mathbb{1}_{[0,\varepsilon)}(\overline{X}_s)d\langle\overline{X}\rangle_s = 0,$$

because for $\varepsilon < (1-\sqrt{\alpha})b_\sigma(\alpha)\Delta t$, and $s \leq T\wedge\Theta_\alpha$, $\mathbb{1}_{[0,\varepsilon)}(\overline{X}_s) = 0$. a.s. Now, since

$$L^0_T(\overline{X}) = L^0_T(\overline{X})\mathbb{1}_{\{\Theta_\alpha<T\}} + L^0_{T\wedge\Theta_\alpha}(\overline{X})\mathbb{1}_{\{T\leq\Theta_\alpha\}} = L^0_T(\overline{X})\mathbb{1}_{\{\Theta_\alpha<T\}},$$

we can conclude that

$$\mathbb{E}[L^0_T(\overline{X})^2] = \mathbb{E}[L^0_T(\overline{X})^2\mathbb{1}_{\{\Theta_\alpha<T\}}] \leq \sqrt{\mathbb{E}[L^0_T(\overline{X})^4]\mathbb{P}\left(\Theta_\alpha<T\right)} \leq C\sqrt{\frac{1}{\Delta t}\exp\left(-\frac{\gamma}{\Delta t}\right)}.$$

□

3.5.5 On the negative moments of the stopped increment process $(\overline{Z}_{t\wedge\Theta_\alpha})$

To prove Lemma 3.2.5 (see section 3.5.6 below), we need to control the negative moments of the stopped increment process $\{\overline{Z}_{t\wedge\Theta_\alpha}\}_{0\leq t\leq T}$. This is the object of the following lemmas, that can be summarize in the following

Lemma 3.5.2. *Let $q \geq 1$. Let Θ_α be the stopping time defined in (3.5.7). Let us assume $\Delta t \leq \Delta_{\max}(\alpha)$. Moreover, let us assume $b(0) > 2\alpha(1-\alpha)^2$ when $\alpha \in (\frac{1}{2}, 1)$, and $b(0) > \frac{3}{2}\sigma^2(q+1)$ when $\alpha = \frac{1}{2}$. Then there exists a constant C depending on $b(0)$, σ, α, T and q but not on Δt, such that*

$$\forall t \in [0,T], \quad \mathbb{E}\left[\overline{Z}^{-q}_{t\wedge\Theta_\alpha}\right] \leq C\left(1+\frac{1}{x_0^q}\right).$$

Existence of Negative Moments. Case $\alpha = \frac{1}{2}$

The proof of the existence of Negative Moments of $\overline{Z}_{t\wedge\Theta_\alpha}$ has two parts. First we study the quotient $\overline{X}_{\eta(s)}/\overline{Z}_s$, and then we proof the main result of the section.

Lemma 3.5.3. *For $\alpha = \frac{1}{2}$, and $\Delta t \leq 1/(4K) \wedge x_0$ we have*

$$\sup_{0\leq s\leq T} \mathbb{P}\left(\overline{Z}_s \leq \frac{\overline{X}_{\eta(s)}}{2}\right) \leq C\Delta t^{\frac{15}{8\sigma^2} b_\sigma(1/2)}. \tag{3.5.11}$$

To prove this lemma, we need the following auxiliary result which is a straightforward adaptation of the Lemma 3.6 in [14].

Lemma 3.5.4. *Assume Hypothesis 3.1.1 holds, and $b(0) > \sigma^2/4$. Assume also that $\Delta t \leq 1/(4K) \wedge x_0$. Then, for any $\gamma \geq 1$ there exists a constant C depending on the parameters $b(0)$, K, σ, x_0, T, and also on γ, such that*

$$\sup_{k=0,\dots,N} \mathbb{E}\exp\left(-\frac{\overline{X}_{t_k}}{\gamma\sigma^2\Delta t}\right) \leq C\left(\frac{\Delta t}{x_0}\right)^{\frac{2}{\sigma^2}b_\sigma(1/2)\left(1-\frac{1}{2\gamma}\right)}.$$

Proof. First, from the definition of $\overline{X}_{t_k}$ we have

$$\overline{X}_{t_k} \geq \overline{X}_{t_{k-1}} + (b_\sigma(1/2) - K\overline{X}_{t_{k-1}})\Delta t + \sigma\sqrt{\overline{X}_{t_{k-1}}}\left(W_{t_k} - W_{t_{k-1}}\right),$$

then

$$\mathbb{E}\exp\left(-\mu_0\overline{X}_{t_k}\right) \leq \mathbb{E}\exp\Big(-\mu_0\Big[\overline{X}_{t_{k-1}} + (b_\sigma(1/2) - K\overline{X}_{t_{k-1}})\Delta t + \sigma\sqrt{\overline{X}_{t_{k-1}}}\left(W_{t_k} - W_{t_{k-1}}\right)\Big]\Big),$$

where $\mu_0 = 1/\gamma\sigma^2\Delta t$. From here, just as in Lemma 3.6 in [14], we conclude

$$\mathbb{E}\exp\left(-\mu_0\overline{X}_{t_k}\right) \leq \exp\left(-\mu_0 b_\sigma(1/2)\Delta t\right)\mathbb{E}\exp\left(-\mu_0\overline{X}_{t_{k-1}}\left[1 - K\Delta t - \frac{\sigma^2\Delta t}{2}\mu_0\right]\right). \tag{3.5.12}$$

Then if we introduce the same sequence $(\mu_j)_j \geq 0$ of Lemma 3.6 in [14], given by

$$\mu_j = \begin{cases} \frac{1}{\gamma\sigma^2\Delta t}, & j = 0,\\ \mu_{j-1}\left[1 - K\Delta t - \frac{\sigma^2\Delta t}{2}\mu_{j-1}\right], & j \geq 1. \end{cases}$$

We can repeat the proof in [14] and find out that if $\Delta t \leq 1/(2K)$ then, the sequence $(\mu_j)_j \geq 0$ is nonnegative, decreasing and satisfies the following bound

$$\mu_j \geq \mu_1\left(\frac{1}{1 + \frac{\sigma^2}{2}\Delta t(j-1)\mu_0}\right) - K\left(\frac{\Delta t(j-1)\mu_0}{1 + \frac{\sigma^2}{2}\Delta t(j-1)\mu_0}\right),\ \forall j \geq 1.$$

On the other hand making the same calculations to obtain (3.5.12) we can get for any $j \in \{0, \dots, k-1\}$,

$$\mathbb{E}\exp\left(-\mu_j \overline{X}_{t_{k-j}}\right) \leq \exp\left(-\mu_j b_\sigma(1/2)(\tfrac{1}{2})\Delta t\right) \mathbb{E}\exp\left(-\mu_j \overline{X}_{t_{k-j-1}}\left[1 - K\Delta t - \frac{\sigma^2 \Delta t}{2}\mu_{j+1}\right]\right),$$

from where, by an induction argument we have

$$\mathbb{E}\left(-\mu_0 \overline{X}_{t_k}\right) \leq \exp\left(-b_\sigma(1/2)\sum_{j=0}^{k-1}\mu_j \Delta t\right)\exp\left(x_0 \mu_k\right).$$

From here, and the bound for the sequence $(\mu_j)_j \geq 0$, we have

$$\mathbb{E}\left(-\mu_0 \overline{X}_{t_k}\right) \leq C\left(\frac{\Delta t}{x_0}\right)^{\frac{2b_\sigma(1/2)}{\sigma^2}\left(1-\frac{1}{2\gamma}\right)}.$$

From where we see immediately

$$\sup_{k=0,\dots,N}\mathbb{E}\exp\left(-\frac{\overline{X}_{t_k}}{\gamma\sigma^2\Delta t}\right) \leq C\left(\frac{\Delta t}{x_0}\right)^{\frac{2b_\sigma(1/2)}{\sigma^2}\left(1-\frac{1}{2\gamma}\right)}. \quad \square$$

Proof of Lemma 3.5.3. We start by proving

$$\sup_{0\leq s\leq T}\mathbb{P}\left(\overline{Z}_s \leq \frac{\overline{X}_{\eta(s)}}{2}\right) \leq \sup_{k=0,\dots,N}\mathbb{E}\exp\left(-\frac{\overline{X}_{t_k}}{\gamma\sigma^2\Delta t}\right). \tag{3.5.13}$$

Indeed, if we call $\Delta s = s - \eta(s)$, and $\Delta W_s = (W_s - W_{\eta(s)})$, then

$$\begin{aligned}
\mathbb{P}\left(\overline{Z}_s \leq \frac{\overline{X}_{\eta(s)}}{2}\right) &\leq \mathbb{P}\left(\sigma\sqrt{\overline{X}_{\eta(s)}}\Delta W_s + b_\sigma(1/2)\Delta s + (1-K\Delta s)\overline{X}_{\eta(s)} \leq \frac{\overline{X}_{\eta(s)}}{2}\right) \\
&\leq \mathbb{E}\left[\mathbb{P}\left(\frac{\Delta W_s}{\sqrt{\Delta s}} \leq \frac{b_\sigma(1/2)\Delta s + (\frac{1}{2} - K\Delta s)\overline{X}_{\eta(s)}}{\sigma\sqrt{\overline{X}_{\eta(s)}}}\middle|\mathcal{F}_{\eta(s)}\right)\right] \\
&\leq \mathbb{E}\exp\left(-\frac{(b_\sigma(1/2)\Delta s + (\frac{1}{2} - K\Delta s)\overline{X}_{\eta(s)})^2}{2\sigma^2\Delta s\overline{X}_{\eta(s)}}\right) \\
&\leq \mathbb{E}\exp\left(-\frac{(1-2K\Delta t)^2\overline{X}_{\eta(s)}}{8\sigma^2\Delta t}\right).
\end{aligned}$$

From here, the bound (3.5.13) follows easily, and then we conclude using Lemma 3.5.4. $\square$

Lemma 3.5.5. *Let $\Theta_{\frac{1}{2}}$ be the stopping time defined in (3.5.7), and $q \geq 1$. If $\Delta t \leq \Delta_{\max}(1/2)$, and*

$$b(0) > \frac{3}{2}\sigma^2(q+1). \tag{3.5.14}$$

Then there exists a constant C depending on $b(0)$, σ, α, T and q but not on Δt, such that

$$\forall t \in [0,T], \quad \mathbb{E}\left[\overline{Z}^{-q}_{t\wedge\Theta_{\frac{1}{2}}}\right] \leq C\left(1 + \frac{1}{x_0^q}\right).$$

Proof. Let us call $\Delta W_s := (W_s - W_{\eta(s)})$, and $\Delta s := (s - \eta(s))$. By Ito's formula

$$\begin{aligned}\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_{\frac{1}{2}}}^{-q}\right] =& \frac{1}{x_0^q} - q\,\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{b(\overline{X}_{\eta(s)})}{\overline{Z}_s^{q+1}} ds\right] \\ &+ \frac{q(q+1)}{2}\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{1}{\overline{Z}_s^{q+2}}\left(\sigma\sqrt{\overline{X}_{\eta(s)}} + \frac{\sigma^2}{2}\Delta W_s\right)^2 ds\right].\end{aligned} \tag{3.5.15}$$

But,

$$\left(\sigma\sqrt{\overline{X}_{\eta(s)}} + \frac{\sigma^2}{2}\Delta W_s\right)^2 \leq \sigma^2\overline{X}_{\eta(s)} + \sigma^2\overline{Z}_s, \quad \mathbb{P}-\text{a.s.} \tag{3.5.16}$$

Indeed,

$$\begin{aligned}\left(\sigma\sqrt{\overline{X}_{\eta(s)}} + \frac{\sigma^2}{2}\Delta W_s\right)^2 =& \sigma^2\left(\overline{X}_{\eta(s)} + \sigma\sqrt{\overline{X}_{\eta(s)}}\Delta W_s + \frac{\sigma^2}{4}\Delta W_s\right) \\ =& \sigma^2\left(\overline{X}_{\eta(s)} + \overline{Z}_s - (\overline{X}_{\eta(s)} + b(\overline{X}_{\eta(s)})\Delta s - \frac{\sigma^2}{4}\Delta s)\right).\end{aligned}$$

But, thanks to the Lipschitz property of b,

$$\begin{aligned}\overline{X}_{\eta(s)} + b(\overline{X}_{\eta(s)})\Delta s - \frac{\sigma^2}{4}\Delta s \geq& \overline{X}_{\eta(s)} + \left(b(0) - K\overline{X}_{\eta(s)}\right)\Delta s - \frac{\sigma^2}{4}\Delta s \\ =& b_\sigma(1/2)\Delta s + (1 - K\Delta s)\overline{X}_{\eta(s)} \geq 0,\end{aligned}$$

since $\Delta s \leq \Delta t \leq 1/(2K)$, and $b_\sigma(1/2) > 0$. So we have (3.5.16). Introducing (3.5.16) in (3.5.15), and using $b(x) \geq b(0) - Kx$, we have

$$\begin{aligned}\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_{\frac{1}{2}}}^{-q}\right] \leq& \frac{1}{x_0^q} - q\,\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{b(0)}{\overline{Z}_s^{q+1}} ds\right] + qK\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{\overline{X}_{\eta(s)}}{\overline{Z}_s^{q+1}} ds\right] \\ &+ \frac{q(q+1)}{2}\sigma^2\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{1}{\overline{Z}_s^{q+2}}\left\{\overline{X}_{\eta(s)} + \overline{Z}_s\right\} ds\right].\end{aligned} \tag{3.5.17}$$

Since

$$\frac{\overline{X}_{\eta(s)}}{\overline{Z}_s} \leq \frac{\overline{X}_{\eta(s)}}{\overline{Z}_s}\mathbb{1}_{\{\overline{Z}_s \leq \overline{X}_{\eta(s)}/2\}} + 2,$$

and applying Hölder's Inequality for some $\varepsilon > 0$, we have

$$\begin{aligned}\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_{\frac{1}{2}}}^{-q}\right] \leq& \frac{1}{x_0^q} - q\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{b(0)}{\overline{Z}_s^{q+1}} ds\right] + 2qK\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{1}{\overline{Z}_s^{q}} ds\right] \\ &+ \frac{3q(q+1)}{2}\sigma^2\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{1}{\overline{Z}_s^{q+1}} ds\right] \\ &+ \frac{C}{\Delta t^{q+2}}\int_0^T \left(\mathbb{E}[\overline{X}_{\eta(s)}^{1/\varepsilon}]\right)^\varepsilon \mathbb{P}\left(\overline{Z}_s \leq \overline{X}_{\eta(s)}/2\right)^{1-\varepsilon} ds.\end{aligned}$$

Since $b(0) > 3\sigma^2(q+1)/2$, we have $15b_\sigma(1/2)/8\sigma^2 > 2q+2$, so choosing $\varepsilon = q/(2q+2)$, and applying Lemma 3.5.3 we have

$$\mathbb{P}\left(\overline{Z}_s \leq \overline{X}_{\eta(s)}/2\right)^{1-\varepsilon} \leq C\Delta t^{q+2},$$

and then

$$\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_{\frac{1}{2}}}^{-q}\right] \leq \frac{1}{x_0^q} + 2qK\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{1}{\overline{Z}_s^q} ds\right] + q\left(\frac{3(q+1)}{2}\sigma^2 - b(0)\right)\mathbb{E}\left[\int_0^{t\wedge\Theta_{\frac{1}{2}}} \frac{1}{\overline{Z}_s^{q+1}} ds\right] + C.$$

Since from the Hypotheses, the third term in the right-hand side is negative, we can conclude thanks to Gronwall's Lemma. □

Existence of Negative Moments. Case $\alpha > \frac{1}{2}$

Lemma 3.5.6. *For $\alpha \in (\frac{1}{2}, 1)$, if $b(0) > 2\alpha(1-\alpha)^2\sigma^2$ and $\Delta t \leq 1/(4\alpha K(\alpha))$, there exists $\gamma > 0$ such that*

$$\sup_{0\leq s\leq T} \mathbb{P}\left(\overline{Z}_s \leq \left(1 - \frac{1}{2\alpha}\right)\overline{X}_{\eta(s)}\right) \leq \exp\left(-\frac{\gamma}{\Delta t}\right). \tag{3.5.18}$$

Proof. Let us call $\Delta W_s := (W_s - W_{\eta(s)})$, $\Delta s := (s - \eta(s))$, and

$$q(\overline{X}_{\eta(s)}, \Delta W_s) = \frac{\alpha\sigma^2}{2}\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s^2 + \sigma\overline{X}_{\eta(s)}^{\alpha}\Delta W_s + \frac{\overline{X}_{\eta(s)}}{2\alpha} + \left(b_\sigma(\alpha) - K(\alpha)\overline{X}_{\eta(s)}\right)\Delta s.$$

Notice that for fix $x \in \mathbb{R}$, $q(x, \cdot)$ is a quadratic polynomial. Using (3.5.3), we have

$$\mathbb{P}\Big(\overline{Z}_s \leq \left(1 - \frac{1}{2\alpha}\right)\overline{X}_{\eta(s)}\Big) \leq \mathbb{P}\Big(q(\overline{X}_{\eta(s)}, \Delta W_s) \leq 0, \overline{X}_{\eta(s)} \leq \bar{x}(\alpha)\Big)$$
$$+ \mathbb{P}\Big(q(\overline{X}_{\eta(s)}, \Delta W_s) \leq 0, \overline{X}_{\eta(s)} \geq \bar{x}(\alpha)\Big),$$

where recall, $\bar{x}(\alpha) = b_\sigma(\alpha)/K(\alpha)$. But

$$\mathbb{P}\Big[q(\overline{X}_{\eta(s)}, \Delta W_s) \leq 0, \overline{X}_{\eta(s)} \leq \bar{x}(\alpha)\Big] = \mathbb{E}\left[\mathbb{P}\left(q(x, \sqrt{\Delta s}\mathcal{N}) \leq 0\right)\Big|_{x=\overline{X}_{\eta(s)}} \mathbb{1}_{\{\overline{X}_{\eta(s)} \leq \bar{x}(\alpha)\}}\right],$$

where $\mathcal{N}$ stands for a standard Gaussian random variable. As in the Lemma 3.2.3, we have a quadratic polynomial in $\mathcal{N}$, its discriminant is

$$\begin{aligned}\Delta &= \sigma^2 x^{2\alpha}\Delta s - 2\alpha\sigma^2 x^{2\alpha-1}\Delta s\left[\frac{x}{2\alpha} + (b_\sigma(\alpha) - K(\alpha)x)\Delta s\right] \\ &= -2\alpha\sigma^2 x^{2\alpha-1}\Delta s^2\left(b_\sigma(\alpha) - K(\alpha)x\right),\end{aligned}$$

so if $x \leq \bar{x}(\alpha)$, $\Delta < 0$ and the quadratic form in $\mathcal{N}$ has not real roots, and in particular is non negative almost surely. Then

$$\mathbb{P}\Big(q(\overline{X}_{\eta(s)}, \Delta W_s) \leq 0, \overline{X}_{\eta(s)} \leq \bar{x}(\alpha)\Big) = 0.$$

On the other hand,

$$\mathbb{P}\Big(q(\overline{X}_{\eta(s)},\Delta W_s)\leq 0,\overline{X}_{\eta(s)}\geq \bar{x}(\alpha)\Big)$$
$$\leq \mathbb{E}\left[\mathbb{P}\left(\mathcal{N}\leq -\frac{b_\sigma(\alpha)\Delta s+\left(\frac{1}{2\alpha}-K(\alpha)\Delta s\right)x}{\sigma x^\alpha\sqrt{\Delta s}}\right)\Bigg|_{x=\overline{X}_{\eta(s)}}\mathbb{1}_{\left\{\overline{X}_{\eta(s)}\geq\bar{x}(\alpha)\right\}}\right],$$

and since $\Delta t\leq 1/(4\alpha K(\alpha))$ we can apply the exponential bound for Gaussian tails and get

$$\mathbb{P}\Big(q(\overline{X}_{\eta(s)},\Delta W_s)\leq 0,\overline{X}_{\eta(s)}\geq \bar{x}(\alpha)\Big)$$
$$\leq \mathbb{E}\left[\exp\left(-\frac{\left(\frac{1}{2\alpha}-K(\alpha)\Delta s\right)^2x^{2(1-\alpha)}}{\sigma^2\Delta s}\right)\Bigg|_{x=\overline{X}_{\eta(s)}}\mathbb{1}_{\left\{\overline{X}_{\eta(s)}\geq\bar{x}(\alpha)\right\}}\right].$$

We conclude by taking $\gamma=\bar{x}(\alpha)^{2(1-\alpha)}/(16\sigma^2)$. □

Lemma 3.5.7. *Let Θ_α be the stopping time defined in (3.5.7). Let us assume for $\alpha\in(\frac{1}{2},1)$, $b(0)>2\alpha(1-\alpha)^2$, and $\Delta t\leq\Delta_{\max}(\alpha)$, then for all $q\geq 1$, there exists a constant C depending on $b(0)$, σ, α, T and p but not on Δt, such that*

$$\forall t\in[0,T],\quad \mathbb{E}\left[\overline{Z}_{t\wedge\Theta_\alpha}^{-q}\right]\leq C\left(1+\frac{1}{x_0^q}\right).$$

Proof. Let us call $\Delta W_s:=W_s-W_{\eta(s)}$. By Ito's formula and the Lipschitz property of b,

$$\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_\alpha}^{-q}\right]\leq\frac{1}{x_0^q}-q\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{b(0)}{\overline{Z}_s^{q+1}}ds\right]+qK\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{\overline{X}_{\eta(s)}}{\overline{Z}_s^{q+1}}ds\right]$$
$$+\frac{q(q+1)}{2}\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{1}{\overline{Z}_s^{q+2}}\left(\sigma\overline{X}_{\eta(s)}^\alpha+\alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s\right)^2ds\right].\quad(3.5.19)$$

Following the same ideas to prove (3.5.16), for all $s\in[0,t]$ we can easily prove that almost surely

$$\left(\sigma\overline{X}_{\eta(s)}^\alpha+\alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\Delta W_s\right)^2\leq\sigma^2\overline{X}_{\eta(s)}^{2\alpha}+2\alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}\overline{Z}_s.$$

Introducing this bound in the previous inequality, we have

$$\begin{aligned}\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_\alpha}^{-q}\right]\leq&\frac{1}{x_0^q}-q\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{b(0)}{\overline{Z}_s^{q+1}}ds\right]+qK\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{\overline{X}_{\eta(s)}}{\overline{Z}_s^{q+1}}ds\right]\\&+\frac{q(q+1)}{2}\sigma^2\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{1}{\overline{Z}_s^{q+2}}\left\{\overline{X}_{\eta(s)}^{2\alpha}+2\alpha\overline{X}_{\eta(s)}^{2\alpha-1}\overline{Z}_s\right\}ds\right].\end{aligned}\quad(3.5.20)$$

since for $r\in\{1,2\alpha-1,2\alpha\}$,

$$\left(\frac{\overline{X}_{\eta(s)}}{\overline{Z}_s}\right)^r\leq\left(\frac{\overline{X}_{\eta(s)}}{\overline{Z}_s}\right)^r\mathbb{1}_{\{\overline{Z}_s\leq\overline{X}_{\eta(s)}(1-\frac{1}{2}\alpha)\}}+\left(\frac{2\alpha}{2\alpha-1}\right)^r.$$

we get

$$\begin{aligned}\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_\alpha}^{-q}\right] \leq &\frac{1}{x_0^q} - q\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{b(0)}{\overline{Z}_s^{q+1}}ds\right] + \frac{2\alpha}{2\alpha-1}qK\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{1}{\overline{Z}_s^{q}}ds\right]\\ &+\frac{q(q+1)}{2}\sigma^2\frac{(2\alpha)^{2\alpha+1}}{(2\alpha-1)^{2\alpha}}\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{1}{\overline{Z}_s^{q+2(1-\alpha)}}ds\right]\\ &+C\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\left\{\frac{\overline{X}_{\eta(s)}}{\overline{Z}_s^{q+1}}+\frac{\overline{X}_{\eta(s)}^{2\alpha}}{\overline{Z}_s^{q+2}}+\frac{\overline{X}_{\eta(s)}^{2\alpha-1}}{\overline{Z}_s^{q+1}}\right\}\mathbb{1}_{\{\overline{Z}_s\leq\overline{X}_{\eta(s)}(1-\frac{1}{2}\alpha)\}}ds\right].\end{aligned}$$

The last term in the previous inequality is bounded because of the definition of Θ_α and the Lemma 3.5.6. Indeed,

$$\begin{aligned}&\mathbb{E}\Big[\int_0^{t\wedge\Theta_\alpha}\left\{\frac{\overline{X}_{\eta(s)}}{\overline{Z}_s^{q+1}}+\frac{\overline{X}_{\eta(s)}^{2\alpha}}{\overline{Z}_s^{q+2}}+\frac{\overline{X}_{\eta(s)}^{2\alpha-1}}{\overline{Z}_s^{q+1}}\right\}\mathbb{1}_{\{\overline{Z}_s\leq\overline{X}_{\eta(s)}(1-\frac{1}{2}\alpha)\}}ds\Big]\\ &\qquad\leq\frac{C}{\Delta t^{q+2}}\int_0^T\sqrt{\mathbb{E}\left[\left(\overline{X}_{\eta(s)}+\overline{X}_{\eta(s)}^{2\alpha}+\overline{X}_{\eta(s)}^{2\alpha-1}\right)^2\right]\mathbb{P}\left(\overline{Z}_s\leq\overline{X}_{\eta(s)}\Big(1-\frac{1}{2\alpha}\Big)\right)}ds\\ &\qquad\leq\frac{C}{\Delta t^{q+2}}\exp\left(\frac{\gamma}{\Delta t}\right)\leq C.\end{aligned}$$

So, (3.5.20) becomes

$$\begin{aligned}\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_\alpha}^{-q}\right] \leq &\frac{1}{x_0^q} - q\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{b(0)}{\overline{Z}_s^{q+1}}ds\right] + \frac{2\alpha}{2\alpha-1}qK\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{1}{\overline{Z}_s^{q}}ds\right]\\ &+\frac{q(q+1)}{2}\sigma^2\frac{(2\alpha)^{2\alpha+1}}{(2\alpha-1)^{2\alpha}}\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{1}{\overline{Z}_s^{q+2(1-\alpha)}}ds\right]+C.\end{aligned}\tag{3.5.21}$$

But, for any $A_1, A_2 > 0$, the mapping $z \mapsto \frac{A_1}{z^{q+2(1-\alpha)}} - \frac{A_2}{z^{q+1}}$ is bounded, and (3.5.21) becomes

$$\mathbb{E}\left[\overline{Z}_{t\wedge\Theta_\alpha}^{-q}\right] \leq \frac{1}{x_0^q} + 2qK\mathbb{E}\left[\int_0^{t\wedge\Theta_\alpha}\frac{1}{\overline{Z}_s^{q}}ds\right] + C,$$

from where we can conclude applying Gronwall's Lemma. □

3.5.6 On the corrected local error process

Proof of Lemma 3.2.5. Let us recall the notation in the proof of the main Theorem

$$D_s(\overline{X}) := \sigma\overline{X}_s^\alpha - \sigma\overline{X}_{\eta(s)}^\alpha - \alpha\sigma^2\overline{X}_{\eta(s)}^{2\alpha-1}(W_s - W_{\eta(s)}),\tag{3.5.22}$$

and also introduce

$$S_{u\wedge\Theta_\alpha}(\overline{X}) := \sigma\overline{X}_{\eta(s\wedge\Theta_\alpha)}^\alpha + \alpha\sigma^2\overline{X}_{\eta(s\wedge\Theta_\alpha)}^{2\alpha-1}(W_{u\wedge\Theta_\alpha} - W_{\eta(s\wedge\Theta_\alpha)}),$$

and $\Delta W_s := (W_s - W_{\eta(s)})$.

Using Lemma 3.5.1, and the finiteness of the moments of D, is easy to prove

$$\mathbb{E}\left[D_s(\overline{X})^{2p}\right] \leq C\mathbb{E}\left[D_{s\wedge\Theta_\alpha}(\overline{X})^{2p}\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}}\right] + C\Delta t^{2p}.$$

Then we only have to prove

$$\mathbb{E}\left[D_{s\wedge\Theta_\alpha}(\overline{X})^{2p}\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}}\right] \leq C\Delta t^{2p}. \tag{3.5.23}$$

Notice that $\overline{X}_{s\wedge\Theta_\alpha} = \overline{Z}_{s\wedge\Theta_\alpha}$, so

$$D_{s\wedge\Theta_\alpha}(\overline{X})\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}} = \left\{\sigma\overline{Z}^\alpha_{s\wedge\Theta_\alpha} - \sigma\overline{X}^\alpha_{\eta(s\wedge\Theta_\alpha)} - \alpha\sigma^2\overline{X}^{2\alpha-1}_{\eta(s\wedge\Theta_\alpha)}\Delta W_{s\wedge\Theta_\alpha}\right\}\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}}.$$

Then applying ItÃt's Formula to the function $\sigma|x|^\alpha$ which is $\mathcal{C}^2$ for $x \geq C\Delta t$, we have

$$\begin{aligned}
D_{s\wedge\Theta_\alpha}(\overline{X})\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}} =& \Bigg\{ \int_{\eta(s\wedge\Theta_\alpha)}^{s\wedge\Theta_\alpha} \left(\frac{\alpha\sigma}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}} - \frac{\alpha\sigma}{\overline{X}^{1-\alpha}_{\eta(s\wedge\Theta_\alpha)}}\right)\sigma\overline{X}^\alpha_{\eta(s\wedge\Theta_\alpha)}dW_u \\
&+ \int_{\eta(s\wedge\Theta_\alpha)}^{s\wedge\Theta_\alpha} \frac{\alpha^2\sigma^3\overline{X}^{2\alpha-1}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}}\Delta W_{u\wedge\Theta_\alpha}dW_u \\
&+ \int_{\eta(s\wedge\Theta_\alpha)}^{s\wedge\Theta_\alpha} \frac{\alpha\sigma}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}}b(\overline{X}_{\eta(s\wedge\Theta_\alpha)})du \\
&- \int_{\eta(s\wedge\Theta_\alpha)}^{s\wedge\Theta_\alpha} \frac{1}{2}\frac{\alpha(1-\alpha)\sigma}{\overline{Z}^{2-\alpha}_{u\wedge\Theta_\alpha}}S_{u\wedge\Theta_\alpha}(\overline{X})^2du\Bigg\}\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}} \\
=:& J_1 + J_2 + J_3 - J_4.
\end{aligned} \tag{3.5.24}$$

Notice that on the event $\{\eta(s) \leq \Theta_\alpha\}$ we have $\eta(s) = \eta(s\wedge\Theta_\alpha)$, and then

$$\mathbb{E}[|J_1|^{2p}] = \mathbb{E}\left[\left|\int_{\eta(s)}^{s\wedge\Theta_\alpha}\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}}\left(\frac{\alpha\sigma}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}} - \frac{\alpha\sigma}{\overline{X}^{1-\alpha}_{\eta(s\wedge\Theta_\alpha)}}\right)\sigma\overline{X}^\alpha_{\eta(s\wedge\Theta_\alpha)}dW_u\right|^{2p}\right].$$

By the Burkholder-Davis-Gundy inequality, there exists a constant C_p depending only on p such that

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{\eta(s)}^{s\wedge\Theta_\alpha}\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}}\left(\frac{\alpha\sigma}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}} - \frac{\alpha\sigma}{\overline{X}^{1-\alpha}_{\eta(s\wedge\Theta_\alpha)}}\right)\sigma\overline{X}^\alpha_{\eta(s\wedge\Theta_\alpha)}dW_u\right|^{2p}\right] \\
&\qquad \leq (\alpha\sigma^2)^{2p}C_p\mathbb{E}\left[\left(\int_{\eta(s)}^{s\wedge\Theta_\alpha}\left(\frac{\overline{X}^{1-\alpha}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}\overline{X}^{1-\alpha}_{\eta(s\wedge\Theta_\alpha)}}\right)^2\overline{X}^{2\alpha}_{\eta(s\wedge\Theta_\alpha)}\mathbb{1}_{\{\Theta_\alpha \geq \eta(s)\}}du\right)^p\right],
\end{aligned}$$

observing that the integrand in the right-hand side is positive. And we have

$$\mathbb{E}[|J_1|^{2p}] \leq (\alpha\sigma^2)^{2p} C_p \mathbb{E}\left[\left(\int_{\eta(s)}^{s}\left(\frac{\overline{X}^{1-\alpha}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}}\right)^2 \overline{X}^{4\alpha-2}_{\eta(s\wedge\Theta_\alpha)} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} du\right)^p\right]$$
$$\leq C\mathbb{E}\left[\left(\int_{\eta(s)}^{s}\left(\frac{\left(\overline{X}^{1-\alpha}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}\right)}{\overline{Z}^{1-\alpha}_{u\wedge\Theta_\alpha}}\frac{\left(\overline{X}^{\alpha}_{\eta(s\wedge\Theta_\alpha)} + \overline{Z}^{\alpha}_{u\wedge\Theta_\alpha}\right)}{\overline{Z}^{\alpha}_{u\wedge\Theta_\alpha}}\right)^2 \overline{X}^{4\alpha-2}_{\eta(s\wedge\Theta_\alpha)} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} du\right)^p\right].$$

But for $x, y \geq 0$, and $\beta \in [0, \frac{1}{2})$ it holds $|x^\beta - y^\beta|(x^{1-\beta} + y^{1-\beta}) \leq 2|x - y|$, so

$$\mathbb{E}[|J_1|^{2p}] \leq C\mathbb{E}\left[\left(\int_{\eta(s)}^{s}\left(\overline{X}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}_{u\wedge\Theta_\alpha}\right)^2 \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{\overline{X}^{4\alpha-2}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{2}_{u\wedge\Theta_\alpha}} du\right)^p\right]$$
$$\leq C\Delta t^{p-1}\int_{\eta(s)}^{s}\mathbb{E}\left[\left(\overline{X}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}_{u\wedge\Theta_\alpha}\right)^{2p} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{\overline{X}^{2p(2\alpha-1)}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{2p}_{u\wedge\Theta_\alpha}}\right] du.$$

Let $a > 1$. Thanks to Hölder's inequality we have

$$\mathbb{E}\left[\left(\overline{X}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}_{u\wedge\Theta_\alpha}\right)^{2p} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{\overline{X}^{2p(2\alpha-1)}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{2p}_{u\wedge\Theta_\alpha}}\right]$$
$$\leq \left(\mathbb{E}\left[\left[\overline{X}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}_{u\wedge\Theta_\alpha}\right]^{\frac{2ap}{(a-1)}} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}}\right]\right)^{1-1/a}\left(\mathbb{E}\left[\frac{\overline{X}^{2ap(2\alpha-1)}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{2ap}_{u\wedge\Theta_\alpha}}\right]\right)^{1/a}.$$

We use Lemma 3.2.1 to bound the Local Error of the scheme

$$\mathbb{E}\left[\left(\overline{X}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}_{u\wedge\Theta_\alpha}\right)^{\frac{2ap}{(a-1)}} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}}\right] \leq C\Delta t^{\frac{ap}{(a-1)}},$$

On the other hand, when $\alpha > \frac{1}{2}$, we have control of any negative moment of $\overline{Z}_{u\wedge\Theta_\alpha}$, so

$$\mathbb{E}\left[\frac{\overline{X}^{4ap(2\alpha-1)}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{2ap}_{u\wedge\Theta_\alpha}}\right] \leq \sqrt{\mathbb{E}\left[\overline{X}^{2ap(2\alpha-1)}_{\eta(s\wedge\Theta_\alpha)}\right]\mathbb{E}\left[\frac{1}{\overline{Z}^{4ap}_{u\wedge\Theta_\alpha}}\right]} \leq C,$$

whereas when $\alpha = \frac{1}{2}$, we choose $a > 1$, such that $2b(0)/\sigma^2 > 3(2ap + 1)$, so we have control of the $2ap$-th negative moment of $\overline{Z}_{u\wedge\Theta_\alpha}$. And then

$$\mathbb{E}\left[\frac{\overline{X}^{4ap(2\alpha-1)}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{2ap}_{u\wedge\Theta_\alpha}}\right] = \mathbb{E}\left[\frac{1}{\overline{Z}^{2ap}_{u\wedge\Theta_\alpha}}\right] \leq C.$$

So, in any case we have

$$\mathbb{E}\left[\left(\overline{X}_{\eta(s\wedge\Theta_\alpha)} - \overline{Z}_{u\wedge\Theta_\alpha}\right)^{2p} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{\overline{X}^{2p(2\alpha-1)}_{\eta(s\wedge\Theta_\alpha)}}{\overline{Z}^{2p}_{u\wedge\Theta_\alpha}}\right] \leq C\Delta t^p.$$

And then we can conclude $\mathbb{E}[|J_1|^{2p}] \leq C\Delta t^{2p}$.

Using the same arguments for $\mathbb{E}[|J_2|^{2p}]$, we have

$$\begin{aligned}
\mathbb{E}[|J_2|^{2p}] &\leq C_p \mathbb{E}\left[\left(\int_{\eta(s)}^{s\wedge\Theta_\alpha} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{\alpha^2\sigma^6 \overline{X}_{\eta(s)}^{2(2\alpha-1)}}{\overline{Z}_u^{2(1-\alpha)}} \Delta W_u^2 du\right)^p\right] \\
&\leq C\Delta t^{p-1} \int_{\eta(s)}^{s} \mathbb{E}\left[\mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{\overline{X}_{\eta(s)}^{2(2\alpha-1)p}}{\overline{Z}_{u\wedge\Theta_\alpha}^{2(1-\alpha)p}} \Delta W_{u\wedge\Theta_{\frac{1}{2}}}^{2p}\right] du \\
&\leq C\Delta t^{p-1} \int_{\eta(s)}^{s} \sqrt{\mathbb{E}\left(\frac{\overline{X}_{\eta(s)}^{4(2\alpha-1)p}}{\overline{Z}_{u\wedge\Theta_\alpha}^{4(1-\alpha)p}}\right)} \sqrt{\mathbb{E}\left(\mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \Delta W_{u\wedge\Theta_{\frac{1}{2}}}^{4p}\right)} du \leq C\Delta t^{2p}.
\end{aligned}$$

To bound $\mathbb{E}[|J_3|^{2p}]$ we proceed as follows

$$\begin{aligned}
\mathbb{E}[|J_3|^{2p}] &= \mathbb{E}\left[\left(\int_{\eta(s)}^{s\wedge\Theta_\alpha} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{\alpha\sigma}{\overline{Z}_{u\wedge\Theta_\alpha}^{1-\alpha}} b(\overline{X}_{\eta(s\wedge\Theta_\alpha)}) du\right)^{2p}\right] \\
&\leq (\alpha\sigma)^{2p} \Delta t^{2p-1} \int_{\eta(s)}^{s} \mathbb{E}\left(\frac{1}{\overline{Z}_{u\wedge\Theta_\alpha}^{2(1-\alpha)p}} b\left(\overline{X}_{\eta(s\wedge\Theta_\alpha)}\right)^{2p}\right) du \\
&\leq (\alpha\sigma)^{2p} \Delta t^{2p-1} \int_{\eta(s)}^{s} \mathbb{E}\left(\frac{1}{\overline{Z}_{u\wedge\Theta_\alpha}^{2p}}\right)^{1-\alpha} \mathbb{E}\left(b\left(\overline{X}_{\eta(s\wedge\Theta_\alpha)}\right)^{\frac{2p}{\alpha}}\right)^{\alpha} du \\
&\leq C\Delta t^{2p}.
\end{aligned}$$

Finally for $\mathbb{E}[|J_4|^{2p}]$ we consider first $\alpha > \frac{1}{2}$. In this case we have control of any negative moment of $\overline{Z}_{u\wedge\Theta_\alpha}$. So proceeding as before

$$\begin{aligned}
\mathbb{E}[|J_4|^{2p}] &= \mathbb{E}\left[\left(\int_{\eta(s)}^{s\wedge\Theta_\alpha} \mathbb{1}_{\{\Theta_\alpha\geq\eta(s)\}} \frac{1}{2} \frac{\alpha(1-\alpha)\sigma}{\overline{Z}_{u\wedge\Theta_\alpha}^{2-\alpha}} S_{u\wedge\Theta_\alpha}(\overline{X})^2 du\right)^{2p}\right] \\
&\leq C\Delta t^{2p-1} \int_{\eta(s)}^{s} \mathbb{E}\left(\frac{1}{\overline{Z}_{u\wedge\Theta_\alpha}^{2p(2-\alpha)}} S_{u\wedge\Theta_\alpha}(\overline{X})^{4p}\right) du \\
&\leq C\Delta t^{2p}.
\end{aligned}$$

The case $\alpha = \frac{1}{2}$ is a little more delicate. Let us recall the identity used in the proof of (3.5.16)

$$\left(\sigma \overline{X}_{\eta(s\wedge\Theta_{\frac{1}{2}})}^{1/2} + \frac{\sigma^2}{2} \Delta W_{u\wedge\Theta_{\frac{1}{2}}}\right)^2 = \sigma^2 \overline{Z}_{u\wedge\Theta_{\frac{1}{2}}} - \sigma^2 \left(b(\overline{X}_{\eta(s\wedge\Theta_{\frac{1}{2}})}) - \frac{\sigma^2}{4}\right) (u \wedge \Theta_{\frac{1}{2}} - \eta(s \wedge \Theta_{\frac{1}{2}})),$$

so, we have from the definition of $\Theta_{\frac{1}{2}}$

$$S_{u\wedge\Theta_{\frac{1}{2}}}(\overline{X})^{4p} = \left(\sigma\overline{X}^{\alpha}_{\eta(s\wedge\Theta_{\frac{1}{2}})} + \frac{\sigma^2}{2}\Delta W_{u\wedge\Theta_{\frac{1}{2}}}\right)^{4p} \leq C\left(\overline{Z}^{2p}_{u\wedge\Theta_{\frac{1}{2}}} + \left(b(\overline{X}_{\eta(s\wedge\Theta_{\frac{1}{2}})}) - \frac{\sigma^2}{4}\right)^{2p}\Delta t^{2p}\right)$$
$$\leq C\left(1 + \left(b(\overline{X}_{\eta(s\wedge\Theta_{\frac{1}{2}})}) - \frac{\sigma^2}{4}\right)^{2p}\right)\overline{Z}^{2p}_{u\wedge\Theta_{\frac{1}{2}}}.$$

Then

$$\begin{aligned}\mathbb{E}[|J_4|^{2p}] &\leq C\Delta t^{2p-1}\int_{\eta(s)}^{s}\mathbb{E}\left[\frac{1}{\overline{Z}^{3p}_{u\wedge\Theta_\alpha}}\left(\sigma\overline{X}^{\alpha}_{\eta(s\wedge\Theta_{\frac{1}{2}})} + \frac{\sigma^2}{2}\Delta W_{u\wedge\Theta_{\frac{1}{2}}}\right)^{4p}\right]du \\ &\leq C\Delta t^{2p-1}\int_{\eta(s)}^{s}\mathbb{E}\left[\frac{1}{\overline{Z}^{p}_{u\wedge\Theta_\alpha}}\left(1 + \left(b(\overline{X}_{\eta(s\wedge\Theta_{\frac{1}{2}})}) - \frac{\sigma^2}{4}\right)^{2p}\right)\right]du \\ &\leq C\Delta t^{2p-1}\int_{\eta(s)}^{s}\sqrt{\mathbb{E}\left(\frac{1}{\overline{Z}^{2p}_{u\wedge\Theta_\alpha}}\right)}\sqrt{\mathbb{E}\left(1 + \left(b(\overline{X}_{\eta(s\wedge\Theta_{\frac{1}{2}})}) - \frac{\sigma^2}{4}\right)^{4p}\right)}du \\ &\leq C\Delta t^{2p}.\end{aligned}$$

So for every $\alpha \in [\frac{1}{2}, 1)$, $\mathbb{E}[J_4^{2p}] \leq C\Delta t^{2p}$, from where we conclude on the Lemma. □

Part II

Stochastic Morris Lecar Model

Chapter 4

Stochastic Morris-Lecar Model for Neurons

Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02)

4.1 Introduction

The study of the brain can be trace back to the ancient greeks, but it was in the middle of the 20th century when the Mathematical Neuroscience started. Probably, the breaking point was in 1952 with the work of Alan Lloyd Hodgkin and Andrew Fielding Huxley [41]. In this seminal work, they explain through a nonlinear system of ordinary differential equations, the ionic mechanisms underlying the propagation of action potentials in the squid giant axon.

Although Hodgkin Huxley model is highly validated and physiologically meaningful, it has the drawback of being mathematically complex. This fact has motivated the development of others neuron models that aim to be mathematically simpler, but still representative of the phenomena under study. Examples of those alternative models are the FitzHugh Nagumo model [32,64], and the Morris-Lecar model [63]. We have chosen to focus in the Morris-Lecar model because is simpler than the Hodgkin Huxley model, but it is still consistent with the *kinetic formalism*. In addition, most of our results can be extended to the Hodgking Huxley model.

Despite of the accuracy of the aforementioned deterministic models, nowadays there is enough evidence showing that there is an intrinsic randomness in the behavior of the neurons. Even more, some researchers claim that this *stochastic facilitation* is essential for the functioning of them (See discussion in [37]).

Then, the question remains. How to incorporate randomness to the deterministic models? In this Chapter we follow the ideas of Pakdaman et al. in [67] and we consider an *hybrid model*, that is, a continuous process representing the voltage of the membrane of a neuron, coupled with two jump processes representing the proportion of open ion channels. Since the hybrid process can be very expensive to simulate, we also consider a *diffusive approximation* for the hybrid model proposed in [8, 15]. Notice that we are interested in (large) networks of neurons, so the simulation cost for each neuron can not be too high.

Our main goal in this Chapter is to study the *asymptotic behaviors* of a network of interacting neurons, (i.e.) the behavior of the network when either the number of neurons or the time goes to infinity.

This chapter is organized as follows. First in Section 4.2 we discuss some modeling issues for a single neuron. We recall the original deterministic Morris-Lecar model and then we introduce two stochastic versions of it: the hybrid model and its diffusive approximation. In Section 4.3 we incorporate interaction to the models discussed in the previous section, and we prove the well posedness of these interacting systems and that its solution are uniformly bounded on time. In Section 4.4 we study the behavior of a network of neurons in a finite time window when the number of neurons goes to infinity. We show for both stochastic models that the propagation of chaos property holds. That is, fixed finitely many neurons are asymptotically independent when the number of neurons goes to infinity. In Section 4.5, we discussed about the longtime behavior of the particle system and of the nonlinear process. Finally in Section 4.6 we summarize the conclusions of this Chapter.

4.2 Single Neuron Model

4.2.1 Deterministic Model for a Single Neuron

In [63], Morris and Lecar propose the following model for the voltage oscillations of a muscle fiber of a barnacle.

$$
\begin{aligned}
C_0\frac{dV_t}{dt} &= I - g_{\mathrm{Ca}}m_t(V_t - V_{\mathrm{Ca}}) - g_{\mathrm{K}}n_t(V_t - V_{\mathrm{K}}) - g_{\mathrm{L}}(V - V_{\mathrm{L}})\\
\frac{dm_t}{dt} &= \lambda_m(V_t)[m_\infty(V_t) - m_t]\\
\frac{dn_t}{dt} &= \lambda_n(V_t)[n_\infty(V_t) - n_t],
\end{aligned}
\tag{4.2.1}
$$

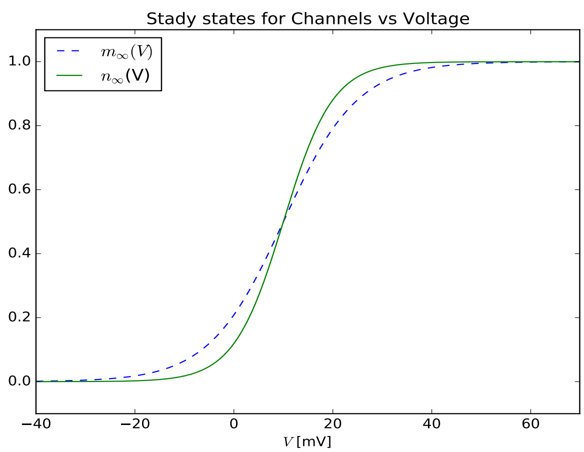
where I stands for the input current, m represents the proportion of open channels for Calcium ions, n represents the proportion of open channels for Potasium ions, and the functions m_∞, λ_m, n_∞ and λ_n are given by

$$\begin{aligned} m_\infty(v) &= \frac{1}{2}\left\{1+\tanh\left(\frac{v-V_1}{V_2}\right)\right\}, \qquad \lambda_m(v) = \bar{\lambda}_m \cosh\left(\frac{v-V_1}{2V_2}\right) \\ n_\infty(v) &= \frac{1}{2}\left\{1+\tanh\left(\frac{v-V_3}{V_4}\right)\right\}, \qquad \lambda_n(v) = \bar{\lambda}_n \cosh\left(\frac{v-V_3}{2V_4}\right), \end{aligned} \tag{4.2.2}$$

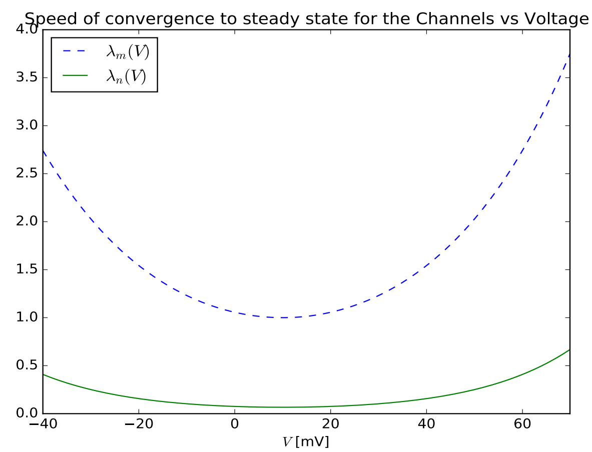
where the constants λ_m, λ_n, V_1, V_2, V_3, V_4, are chosen by fitting some experimental data. In Table 4.2.1 are shown the values obtained by Morris and Lecar [63] for these constants.

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
g_{Ca}	4	V_{Ca}	100	V_1	10	V_3	10
g_{K}	8	V_{K}	-70	V_2	15	V_4	10
g_{L}	2	V_{L}	-50	$\bar{\lambda}_m$	1	$\bar{\lambda}_n$	1/15
C_0	20						

Table 4.2.1: Example of values for the parameters of the model. Taken from [63].



(a) Voltage dependent steady states for ion channels

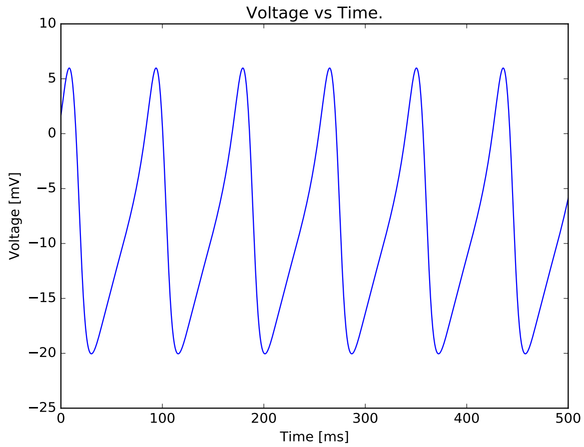


(b) Voltage dependent speed of convergence towards the steady state $\lambda_x(V)$ for $x = m, n$. Notice that $\min\{\lambda_x(V) : V \in \mathbb{R}\} > 0$.

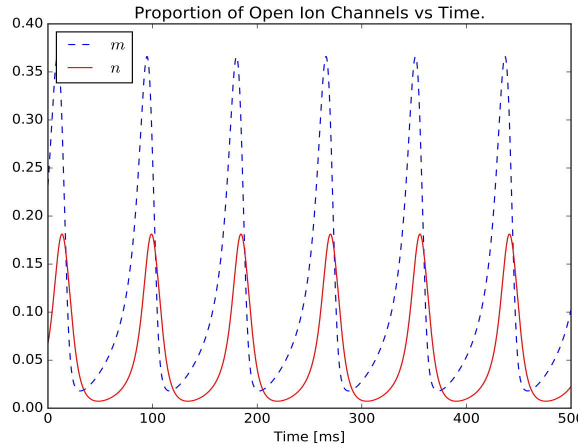
Figure 4.2.1: Auxiliary functions of Morris-Lecar Model

In Figure 4.2.2 we observe an example of the trajectories of the Morris-Lecar model for the parameters in Table 4.2.1, and input current $I = 65$ [µA].

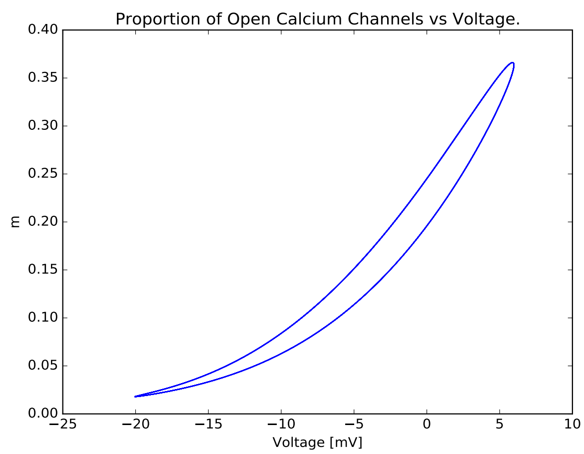
The system (4.2.1) can exhibit several behaviors depending on the value of the parameters. Some examples of the response of the system for different values of the input current I are displayed in Figure 4.2.3. Further analysis on this model from the dynamical system point of view can be found in [27, p. 59], in [56], and in [71].



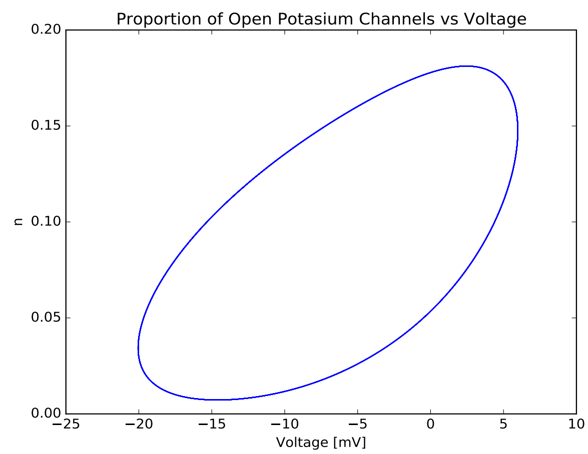
(a) Voltage.



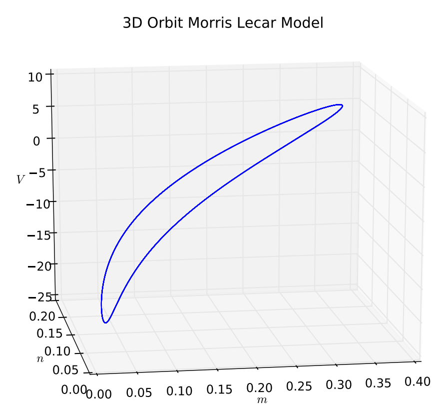
(b) Proportions of open ion channels.



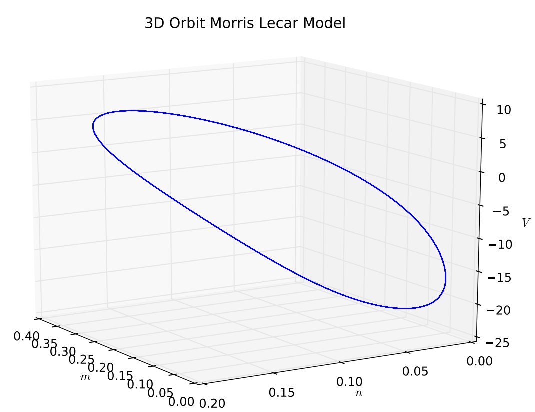
(c) Orbit open Calcium channels vs Voltage



(d) Orbit open Potasium channels vs Voltage

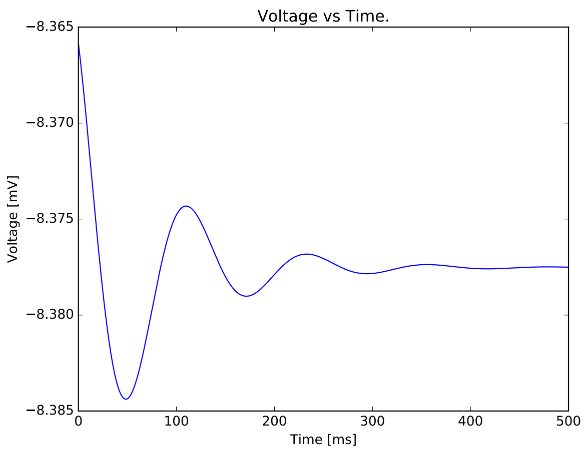


(e) 3D Orbit: View 1

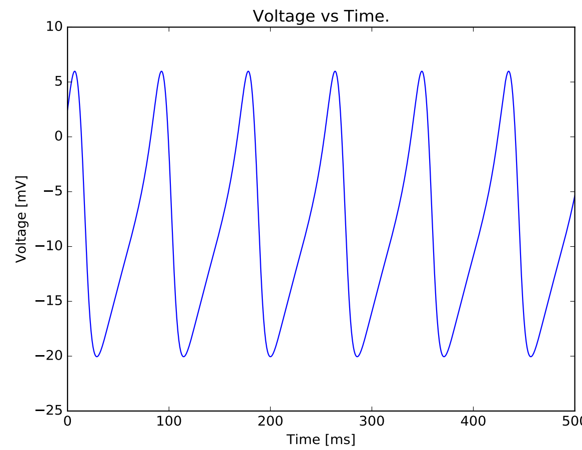


(f) 3D Orbit: View 2

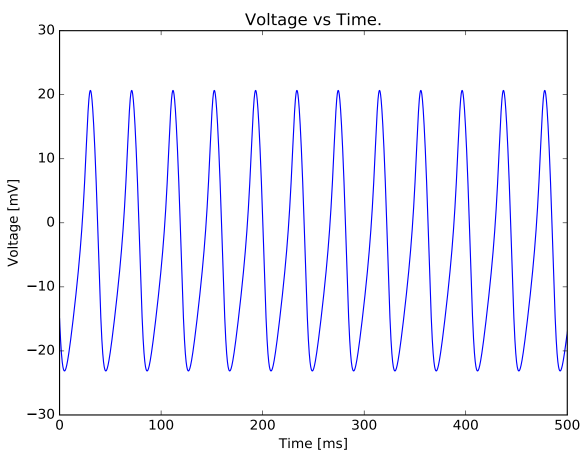
Figure 4.2.2: Example of the trajectories of the Morris-Lecar Model.



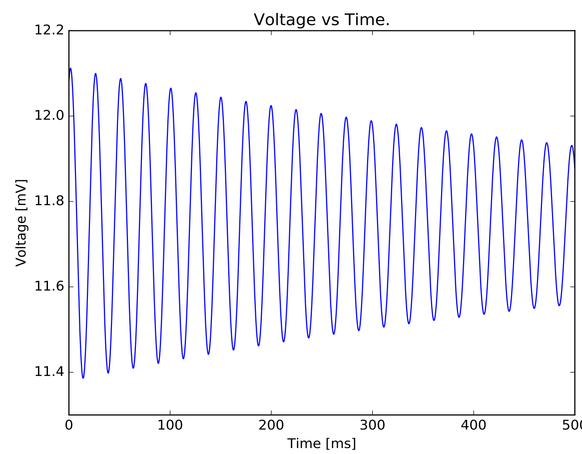
(a) $I = 61$ [µA]. Low frequency, damped oscillations.



(b) $I = 65$ [µA]. Low frequency oscillations



(c) $I = 100$ [µA]. High frequency oscillations



(d) $I = 310$ [µA]. High frequency, damped oscillations

Figure 4.2.3: Different responses of the system for different values of the input current I. Other parameters are fixed according to Table 4.2.1.

In the following, we will consider that the membrane capacitance $C_0 = 1$, or in other words we normalize the equation for the voltage. Also, for $x = n, m$, we will use the more modern notation:

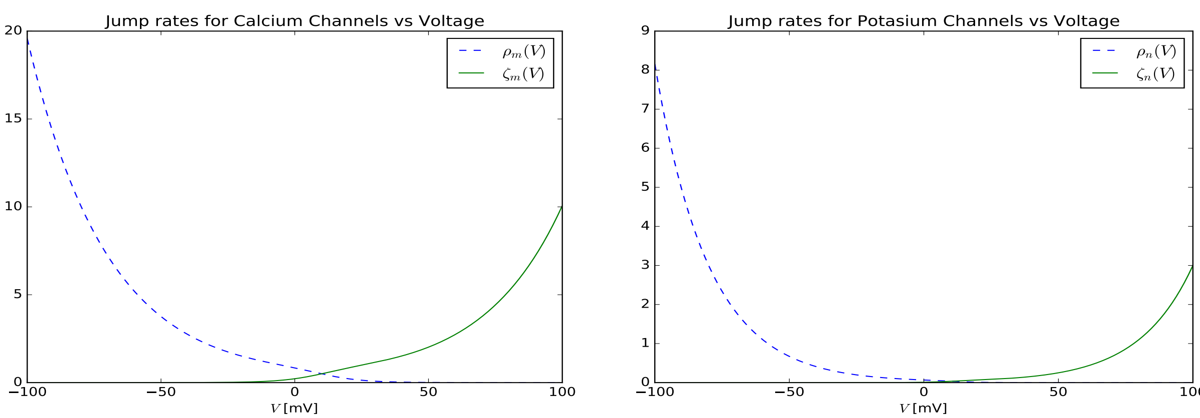
$$\rho_x(V_t) = \lambda_x(V_t)x_\infty(V_t),$$
$$\zeta_x(V_t) = \lambda_x(V_t)[1 - x_\infty(V_t)],$$

thus, the model for a single neuron is written as

$$\begin{aligned}\frac{dV_t}{dt} &= I - g_{\mathrm{Ca}}m_t(V_t - V_{\mathrm{Ca}}) - g_{\mathrm{K}}n_t(V_t - V_{\mathrm{K}}) - g_{\mathrm{L}}(V_t - V_{\mathrm{L}})\\ \frac{dm_t}{dt} &= \rho_m(V_t)(1 - m_t) - \zeta_m(V_t)m_t\\ \frac{dn_t}{dt} &= \rho_n(V_t)(1 - n_t) - \zeta_n(V_t)n_t.\end{aligned} \tag{4.2.3}$$

In Figure 4.2.4 we show the functions ρ and ζ for m and n. Notice that for $x = m, n$, $\min\{\rho_x + \zeta_x\} > 0$.

Remark 4.2.1. In what follows we will assume without further comments that the functions ρ_x and ζ_x are bounded and Lipschitz. Even if this is not globally true for the explicit formulas that define them, we will see that independently of the properties of ρ_x and ζ_x, the voltage coordinate is bounded, that is, there exists some compact interval $[-V_\infty^{\max}, V_\infty^{\max}]$ from where V_t does not escape. So we only care about ρ_x and ζ_x in this compact interval, where these functions are uniformly bounded and Lipschitz.



(a) Rate functions for Calcium channels. (b) Rate functions for Potasium channels.

Figure 4.2.4: Rate functions for ion channels. Notice that $\min\{\rho_x + \zeta_x\} > 0$.

4.2.2 The Hybrid Model for a Single Neuron

Although the model (4.2.3) mimics with great precision the data obtained experimentally (see [63]), it somehow ignores the intrinsic randomness present in the evolution of the ionic channels. To take into account this random nature, we consider the hybrid model proposed by Pakdaman, Thieullen & Wainrib in [67].

For $N_c \in \mathbb{N}$, the hybrid model $X_t^{N_c} = (V_t^{N_c}, m_t^{N_c}, n_t^{N_c})$, will be given by

$$\frac{dV_t^{N_c}}{dt} = I - g_{\text{Ca}} m_t^{N_c}(V_t^{N_c} - V_{\text{Ca}}) - g_{\text{K}} n_t^{N_c}(V_t^{N_c} - V_{\text{K}}) - g_{\text{L}}(V_t^{N_c} - V_{\text{L}}), \tag{4.2.4}$$

and $m_t^{N_c}, n_t^{N_c}$ Markov Processes taking values in

$$E_{N_c} = \left\{0, \frac{1}{N_c}, \dots, \frac{N_c - 1}{N_c}, 1\right\}.$$

In this model we consider a single neuron having N_c channels of each type. The variable $V_t^{N_c}$ will represent the voltage of the membrane cell, whereas $m_t^{N_c}$ and $n_t^{N_c}$ will represent the proportion of open Calcium and Potassium channels respectively.

The process $x_t^{N_c}$, for $x = m, n$, is built by considering that each of the N_c x-type channels behaves independently from the other as a time continuous Markov process which changes from close to open at rate $\rho_x(V_t^{N_c})$, and from open to close at a rate $\zeta_x(V_t^{N_c})$.

Let us denote

$$\begin{aligned} \lambda_0^x(V, u) &= N_c(1 - u)\rho_x(V), \\ \lambda_1^x(V, u) &= N_c u \zeta_x(V). \end{aligned} \tag{4.2.5}$$

Hence the jump rate of x^{N_c}, at state (V, u), is given by

$$\lambda^x(V, u) = \lambda_0^x(V, u) + \lambda_1^x(V, u),$$

and the transition probabilities of x^{N_c} are given by

$$\mathbb{P}\left(x_{t+h}^{N_c} - x_t^{N_c} = \frac{1}{N_c} \middle| (V_t^{N_c}, x_t^{N_c}) = (V, u) \right) = \lambda_0^x(V, u)h + o(h),$$

$$\mathbb{P}\left(x_{t+h}^{N_c} - x_t^{N_c} = -\frac{1}{N_c} \middle| (V_t^{N_c}, x_t^{N_c}) = (V, u) \right) = \lambda_1^x(V, u)h + o(h),$$

$$\mathbb{P}\left(x_{t+h}^{N_c} - x_t^{N_c} = 0 \middle| (V_t^{N_c}, x^{N_c} t) = (V, u) \right) = 1 - \lambda^x(V, u)h + o(h),$$

and for any other $r \in \mathbb{R}$

$$\mathbb{P}\left(x_{t+h}^{N_c} - x_t^{N_c} = r \middle| (V_t^{N_c}, x^{N_c} t) = (V, u) \right) = o(h).$$

Notice that the process X^{N_c} is an example of a Piecewise Deterministic Markov Process (PDMP). See [22] and [47] for general background.

The infinitesimal generator $\mathcal{L}$ of the hybrid process is given by

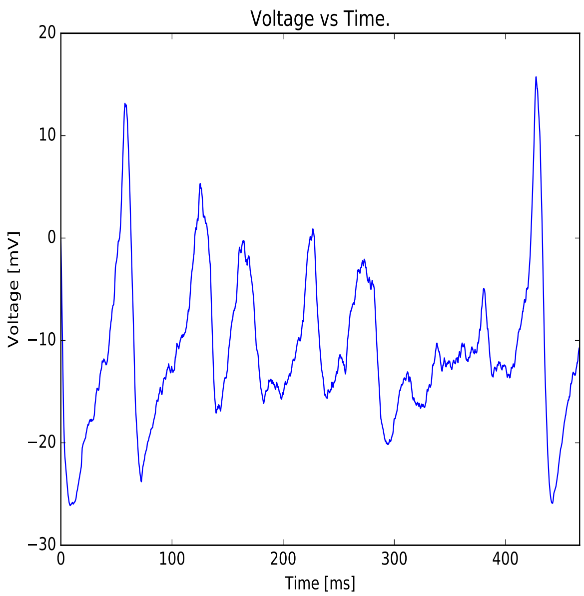
$$\begin{aligned}\mathcal{L}f(V,m,n) = [I - g_{\text{Ca}}m(V - V_{\text{Ca}}) - g_{\text{K}}n(V - V_{\text{K}}) - g_{\text{L}}(V - V_{\text{L}})]\,\frac{\partial f}{\partial V}(V,m,n)\\ + \lambda_0^m(V,m,n)\left[f\left(V, m + \frac{1}{N_c}, n\right) - f\left(V,m,n\right)\right]\\ + \lambda_1^m(V,m,n)\left[f\left(V, m - \frac{1}{N_c}, n\right) - f\left(V,m,n\right)\right]\\ + \lambda_0^n(V,m,n)\left[f\left(V, m, n + \frac{1}{N_c}\right) - f\left(V,m,n\right)\right]\\ + \lambda_1^n(V,m,n)\left[f\left(V, m, n - \frac{1}{N_c}\right) - f\left(V,m,n\right)\right].\end{aligned} \tag{4.2.6}$$

The fact that this dynamics defines globally a unique PDMP is discussed later.

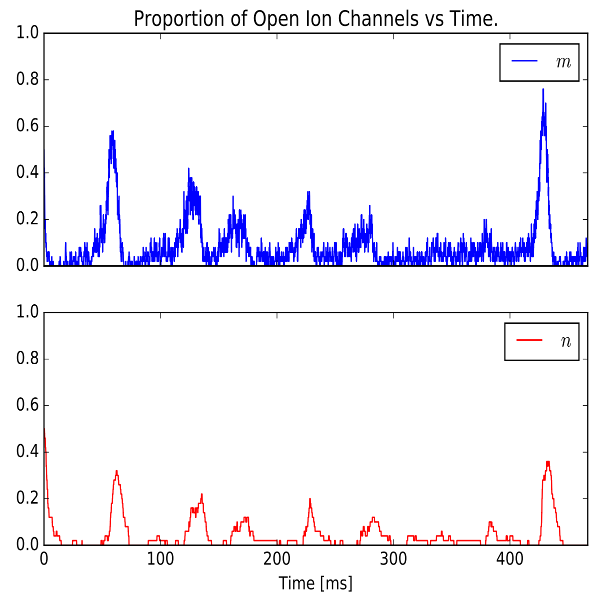
In Figures 4.2.5 and 4.2.6 we display some trajectories of the hybrid model for different values of the number of channels $N_c = 50, 500, 1000, 10000$, and in Table 4.2.2 we report the computation time of each simulation. Notice that the computation time seems to grow linearly with the number of channels. This can be a limitation for the use of this model since in a neuron the order of magnitud of the number of ion channels is in the tens of thousands (See [17]).

N_c	Computation Time [s]
50	43.54
500	280.04
1000	576.87
10000	5586.90

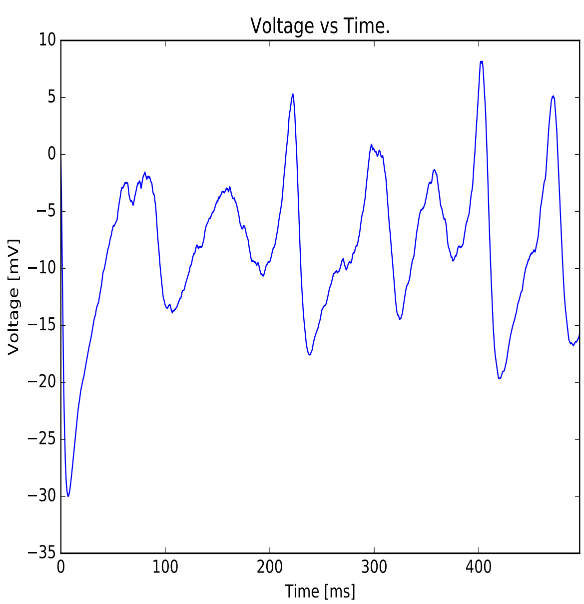
Table 4.2.2: Computation time in a standard laptop for the simulation of the hybrid model for one neuron for different values of N_c. Simulations coded in Python3.



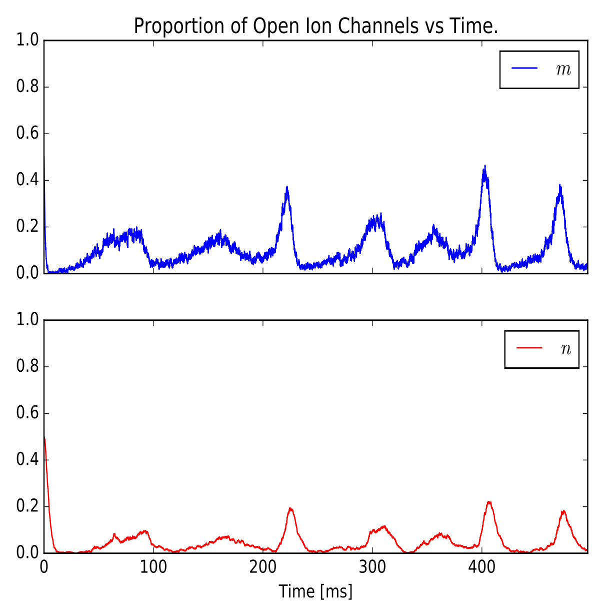
(a) Voltage.



(b) Proportions of open ion channels.

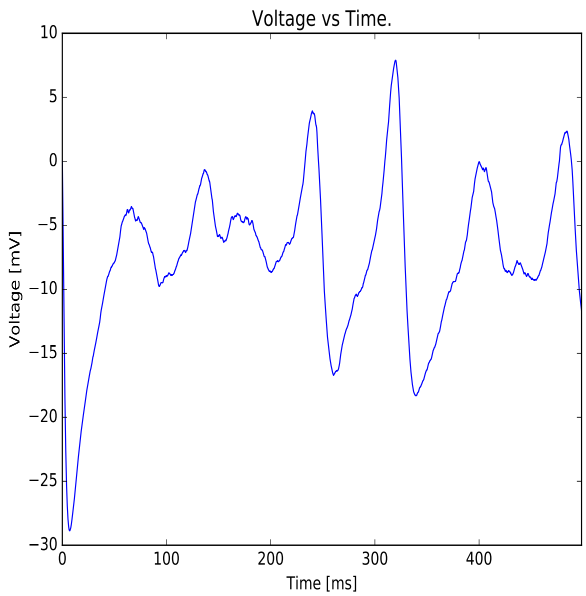


(c) Voltage.

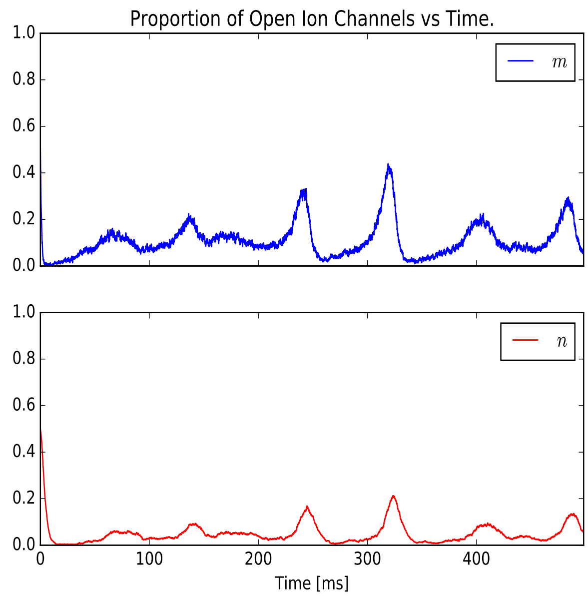


(d) Proportions of open ion channels.

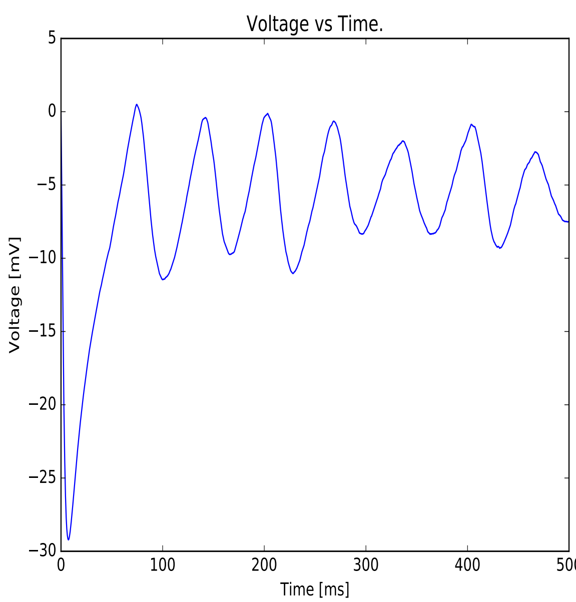
Figure 4.2.5: Examples of trajectories for the Stochastic Morris-Lecar Model for different number of ion channels. First row $N_c = 50$. Second row $N_c = 500$.



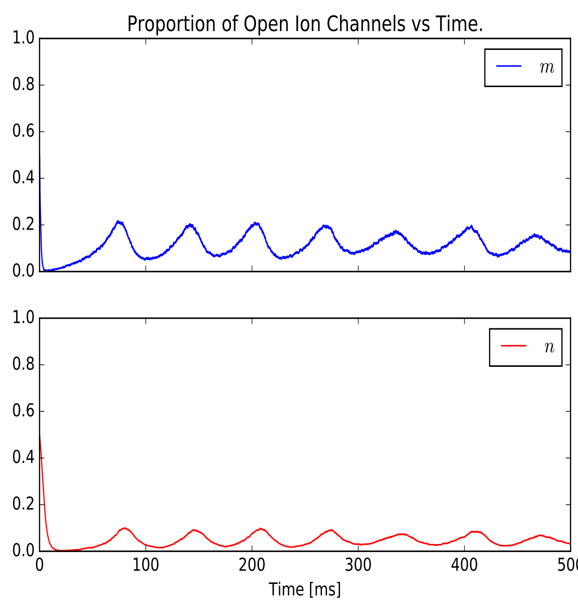
(a) Voltage.



(b) Proportions of open ion channels.



(c) Voltage.



(d) Proportions of open ion channels.

Figure 4.2.6: Examples of trajectories for the Stochastic Morris-Lecar Model for different number of ion channels. First row $N_c = 1000$. Second row $N_c = 10000$.

The Integral Representation of the Proportion Processes

For $x = m, n$, let $\mathcal{N}_0^x, \mathcal{N}_1^x$ be two independent Poisson Measures on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensities $d\lambda \otimes ds$, then the discrete components of the hybrid model can be represented as

$$\begin{aligned} x_t^{N_c} = x_0^{N_c} &+ \int_0^{t+} \int_0^{\infty} \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_0^x(V_{s-}^{N_c}, x_{s-}^{N_c})\}} \mathcal{N}_0^x(d\lambda, ds) \\ &- \int_0^{t+} \int_0^{\infty} \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_1^x(V_{s-}^{N_c}, x_{s-}^{N_c})\}} \mathcal{N}_1^x(d\lambda, ds), \end{aligned}$$

where the jump rates λ_1^x and λ_0^x are given in (4.2.5). If we call $\tilde{\mathcal{N}}_0^x$ and $\tilde{\mathcal{N}}_1^x$ the compensated Poisson Measures, we have

$$\begin{aligned} x_t^{N_c} &= x_0^{N_c} + \int_0^{t} \int_0^{\infty} \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_0^x(V_{s-}^{N_c}, x_{s-}^{N_c})\}} d\lambda ds - \int_0^{t} \int_0^{\infty} \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_1^x(V_{s-}^{N_c}, x_{s-}^{N_c})\}} d\lambda ds \\ &+ \int_0^{t+} \int_0^{\infty} \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_0^x(V_{s-}^{N_c}, x_{s-}^{N_c})\}} \tilde{\mathcal{N}}_0(d\lambda, ds) - \int_0^{t+} \int_0^{\infty} \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_1^x(V_{s-}^{N_c}, x_{s-}^{N_c})\}} \tilde{\mathcal{N}}_1^x(d\lambda, ds) \\ &= x_0^{N_c} + \frac{1}{N_c} \int_0^t \lambda_0^x(V_s^{N_c}, x_s^{N_c}) - \lambda_1^x(V_s^{N_c}, x_s^{N_c}) ds + R_t^x + Q_t^x, \end{aligned}$$

with R^x and Q^x orthogonal martingales, that is, they don't have simultaneous jumps. Let us call $M^x = Q^x + R^x$, then we have the following semimartingale decomposition for $x_t^{N_c}$,

$$x_t^{N_c} = x_0^{N_c} + \int_0^t (1 - x_s^{N_c}) \rho_x(V_s^{N_c}) - x_s^{N_c} \zeta_x(V_s^{N_c}) ds + M_t^x. \tag{4.2.7}$$

Summarizing, the hybrid model for a single neuron is given by

$$\begin{aligned} V_t^{N_c} &= V_0^{N_c} + \int_0^t I - g_{\mathrm{L}}(V_s^{N_c} - V_{\mathrm{L}}) - g_{\mathrm{Ca}} m_t^{N_c}(V_s^{N_c} - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_t^{N_c}(V_s^{N_c} - V_{\mathrm{K}}) ds \\ m_t^{N_c} &= m_0^{N_c} + \int_0^t (1 - m_s^{N_c}) \rho_m(V_s^{N_c}) - m_s^{N_c} \zeta_m(V_s^{N_c}) ds + M_t^m. \\ n_t^{N_c} &= n_0^{N_c} + \int_0^t (1 - n_s^{N_c}) \rho_n(V_s^{N_c}) - n_s^{N_c} \zeta_n(V_s^{N_c}) ds + M_t^n. \end{aligned} \tag{4.2.8}$$

Remark 4.2.2. The well posedness of the equation (4.2.8) is not obvious, but is not difficult to prove either. We will do the construction later in Lemma 4.3.4.

In the next result, we show that, in any finite time window, if the jump rates are given by (4.2.5), the martingale part of the processes x^{N_c} goes to zero as the number of ion channels goes to infinity. This tells us that as the number of ion channels increase the dynamic of the hybrid process becomes closer to the dynamic of the deterministic Morris-Lecar model. In fact, in Proposition 4.2.4 below we prove the convergence of the hybrid model to the deterministic one (4.2.3).

Lemma 4.2.3. *Let $T > 0$. For $x = m, n$*

$$\mathbb{E}\left[\sup_{t\leq T}|M_t^x|\right] \leq \sqrt{\frac{(\|\rho_x\|_\infty + \|\zeta_x\|_\infty)\, T}{N_c}}.$$

Proof. By Doob's Inequality

$$\mathbb{E}\left[\sup_{t\leq T}(M_t^x)^2\right] \leq 4\mathbb{E}\left[(M_T^x)^2\right] = 4\mathbb{E}\left[\langle M^x\rangle_T\right].$$

On the other hand, $\langle M^x\rangle_T = \langle Q^x\rangle_T + \langle R^x\rangle_T$, and by equality (3.9) in page 62 of Ikeda & Watanabe [44]

$$\begin{aligned}\langle M^x\rangle_T &= \int_0^T\int_0^\infty\left(\frac{1}{N_c}\mathbb{1}_{\{\lambda\leq\lambda_0^{N_c}(V_{s-}^{N_c},x_{s-}^{N_c}))\}}\right)^2 d\lambda ds + \int_0^T\int_0^\infty\left(\frac{-1}{N_c}\mathbb{1}_{\{\lambda\leq\lambda_1^{N_c}(V_{s-}^{N_c},x_{s-}^{N_c}))\}}\right)^2 d\lambda ds\\ &= \frac{1}{N_c^2}\int_0^T \lambda_0^{N_c}(V_s^{N_c},x_s^{N_c})) + \lambda_1^{N_c}(V_s^{N_c},x_s^{N_c}))ds\\ &= \frac{1}{N_c}\int_0^T (1-x_s^{N_c})\rho_x(V_s^{N_c}) + x_s^{N_c}\zeta_x(V_s^{N_c})ds\\ &\leq \frac{1}{N_c}\int_0^T |1-x_s^{N_c}||\rho_x(V_s^{N_c})| + |x_s^{N_c}||\zeta_x(V_s^{N_c})|ds.\end{aligned}$$

Hence,

$$\langle M^x\rangle_T \leq \frac{(\|\rho_x\|_\infty + \|\zeta_x\|_\infty)\, T}{N_c},$$

from the results follows using Burkholder-Davis-Gundy inequality. □

Proposition 4.2.4. *Given $T > 0$, let $X_t = (V_t, m_t, n_t)$ be the solution of (4.2.3), and $X_t^{N_c} = (V_t^{N_c}, m_t^{N_c}, n_t^{N_c})$ be the solution of (4.2.8). Then for all $T > 0$*

$$\lim_{N_c\to\infty}\mathbb{E}\left(\sup_{t\leq T}\|X_t^{N_c} - X_t\|_1\right) = 0.$$

Proof.

$$\begin{aligned}|V_t^{N_c} - V_t| &\leq \int_0^t g_{\mathrm{L}}|V_s^{N_c} - V_s| + g_{\mathrm{Ca}}|m_s^{N_c}||V_s^{N_c} - V_s| + g_{\mathrm{K}}|n_s^{N_c}||V_s^{N_c} - V_s|ds\\ &\quad + \int_0^t g_{\mathrm{Ca}}|V_s - V_{\mathrm{Ca}}||m_s^{N_c} - m_s| + g_{\mathrm{K}}|V_s - V_{\mathrm{K}}||n_s^{N_c} - n_s|ds\\ &\leq C\int_0^t \sup_{u\leq s}\|X_u^{N_c} - X_u\|_1 ds.\end{aligned}$$

On the other hand for $x = m, n$

$$\begin{aligned}|x_t^{N_c} - x_t| &\le \int_0^t (1 - x_s^{N_c})|\rho_x(V_s^{N_c}) - \rho_x(V_s)| + x_s^{N_c}|\zeta_x(V_s^{N_c}) - \zeta_x(V_s)|ds. \\ &\quad + \int_0^t \rho_x(V_s)|x_s^{N_c} - x_s| + \zeta_x(V_s)|x_s^{N_c} - x_s|ds + |M_t^x| \\ &\le C \int_0^t |V_s^{N_c} - V_s| + |x_s^{N_c} - x_s|ds + |M_t^x|.\end{aligned}$$

Adding up these computations, we conclude

$$\|X_t^{N_c} - X_t\|_1 \le C \int_0^t \sup_{u \le s} \|X_u^{N_c} - X_u\|_1 ds + \sup_{s \le t} |M_s^m| + \sup_{s \le t} |M_s^n|,$$

and since the right-hand side is increasing, yields

$$\sup_{s \le t} \|X_s^{N_c} - X_s\|_1 \le C \int_0^t \sup_{u \le s} \|X_u^{N_c} - X_u\|_1 ds + \sup_{s \le t} |M_s^m| + \sup_{s \le t} |M_s^n|.$$

Taking expectation and applying Proposition 4.2.3, we conclude

$$\mathbb{E}\left(\sup_{s \le t} \|X_s^{N_c} - X_s\|_1\right) \le C \int_0^t \mathbb{E}\left(\sup_{u \le s} \|X_u^{N_c} - X_u\|_1\right) ds + \frac{C}{\sqrt{N_c}},$$

from where we conclude that

$$\mathbb{E}\left(\sup_{t \le T} \|X_t^{N_c} - X_t\|_1\right) \le \frac{Ce^{CT}}{\sqrt{N_c}}.$$

□

Remark 4.2.5. In [67] the authors obtain the convergence of the hybrid model to the deterministic one (4.2.3) in probability.

4.2.3 Diffusive Approximation of the Hybrid Model and Perturbed Hybrid Model

In [67], the authors obtained a functional central limit theorem, which allows them to propose a diffusive approximation (or Langevin approximation) of the hybrid system (4.2.8) given by

$$\begin{aligned}V_t^* &= V_0^* + \int_0^t I - g_{\mathrm{Ca}} m_s (V_s^* - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_s (V_s^* - V_{\mathrm{K}}) - g_{\mathrm{L}}(V_s^* - V_{\mathrm{L}}) ds \\ x_t^* &= x_0^* + \int_0^t \rho_x(V_s^*)(1 - x_s^*) - \zeta_x(V_s^*) x_s^* ds \\ &\quad + \frac{1}{\sqrt{N_c}} \int_0^t \sqrt{\rho_x(V_s^*)(1 - x_s^*) + \zeta_x(V_s^*) x_s^*} dW_s^x, \quad x = m, n,\end{aligned} \tag{4.2.9}$$

where N_c is the number of ion channels. Nevertheless, this equation raises two complex issues. First, it is not clear in general that this equation is well-posed. Notice that the argument inside of the square root can become negative if $x_s^* \notin [0,1]$. Secondly, Pakdaman et al. remark that the hybrid model and its Langeving approximation can have quite different qualitative behavior (See [66]).

For the well posedness of the diffusion, Baladron et al. in [8] proposed to incorporate a cut-off function χ in the dynamics to ensure the permanence of the proportion processes m y n in the interval $[0,1]$.

For the qualitative differences between the hybrid process and its Langevin approximation there is not a satisfactory solution. Nevertheless, considering the elevated computational cost of simulating the hybrid model, is interesting and useful to consider the diffusive model, especially in a context of large network of neurons. This approach has been followed by several authors in the literature see [21] and the references therein.

The diffusive model proposed in [8] is given by

$$\begin{aligned}
V_t &= V_0 + \int_0^t I - g_{\mathrm{Ca}} m_s (V_s - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_s (V_s - V_{\mathrm{K}}) - g_{\mathrm{L}}(V_s - V_{\mathrm{L}}) ds \\
x_t &= x_0 + \int_0^t \rho_x(V_s)(1 - x_s) - \zeta_x(V_s) x_s ds \\
&\quad + \int_0^t \sigma \sqrt{|\rho_x(V_s)(1 - x_s) + \zeta_x(V_s) x_s|}\, \chi(x_s) dW_s^x, \quad x = m, n.
\end{aligned} \tag{4.2.10}$$

A rigorous proof of the existence and uniqueness of the solution of this equation is given in [15].

In Figure 4.2.7 we present some sample trajectories of the solution of (4.2.10) for different levels of noise. For the simulation we considered the cut-off function

$$\chi(x) = 0.1 \exp\left(\frac{-0.5}{1 - (2x - 1)^2}\right),$$

and the Euler Maruyama scheme.

In what follows, we will show that the model (4.2.10) can be obtained as the limit in distribution of a perturbed version of the hybrid model (4.2.8). This is interesting because at some level it explains the origin of the differences between the hybrid model and its diffusive approximation.

Let us consider the perturbed hybrid model $\hat{X}_t^{N_c} = (\hat{V}_t^{N_c}, \hat{m}_t^{N_c}, \hat{n}_t^{N_c})$. The continuous component $\hat{V}_t^{N_c}$ will evolve just as in the previous case, that is

$$\frac{d\hat{V}_t}{dt} = I - g_{\mathrm{L}}(\hat{V}_t - V_{\mathrm{L}}) - g_{\mathrm{Ca}} m_t (\hat{V}_t - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_t (\hat{V}_t - V_{\mathrm{K}}),$$

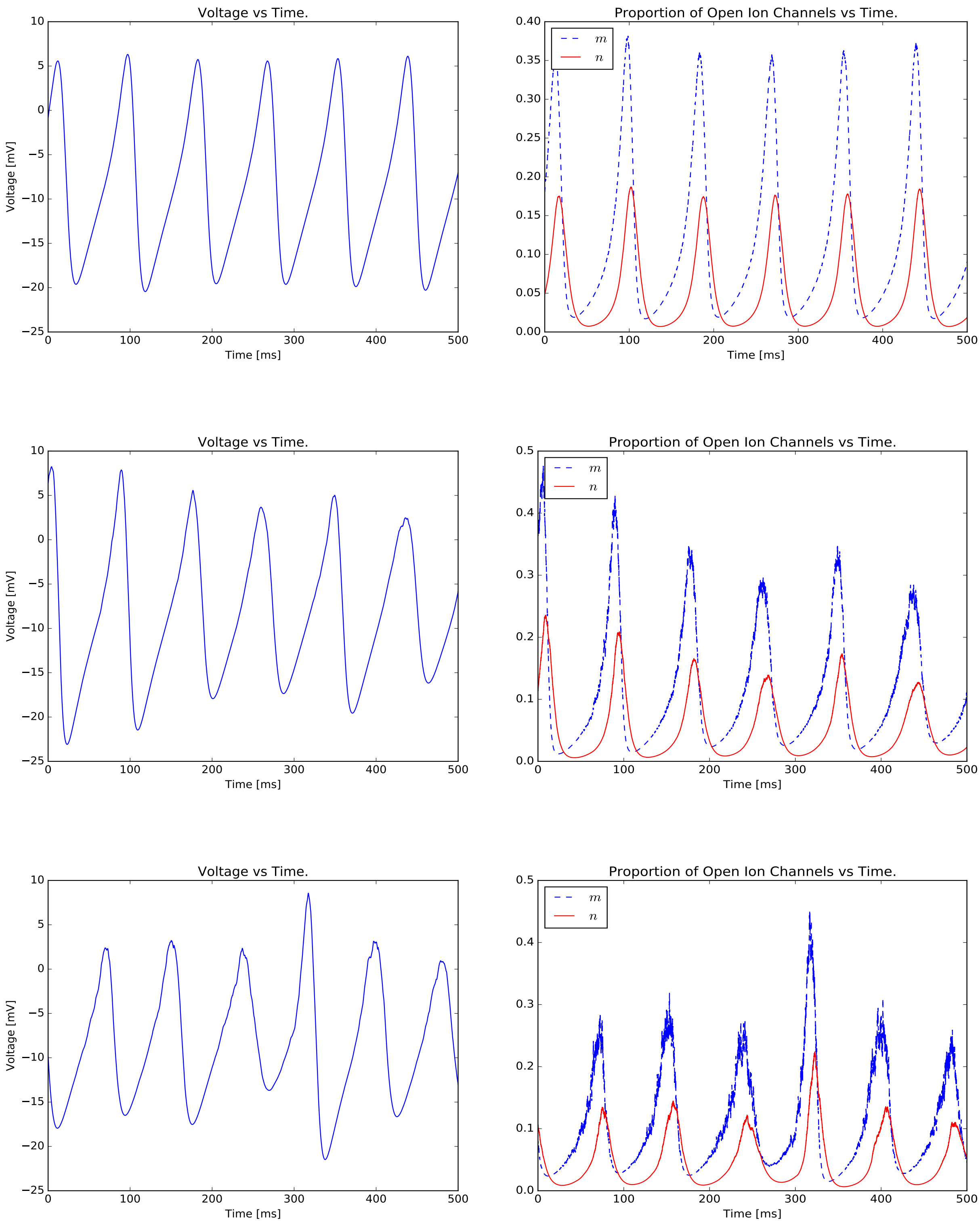


Figure 4.2.7: Stochastic system (4.2.10) for different values of σ. First row: $\sigma = 0.1$. Second row: $\sigma = 0.5$. Third row: $\sigma = 1$.

but the jump processes will be now characterized by the jump rates

$$\begin{aligned}\hat{\lambda}_0^x(V,u) &= N_c(1-u)\rho_x(V) + \frac{{N_c}^2}{2}\sigma_x^2(V,u)\chi^2(u),\\ \hat{\lambda}_1^x(V,u) &= N_c u\zeta_x(V) + \frac{{N_c}^2}{2}\sigma_x^2(V,u)\chi^2(u),\end{aligned} \tag{4.2.11}$$

where

$$\sigma_x(V,u) = \sigma\sqrt{|\rho_x(V)(1-u)+\zeta_x(V)u|},$$

and then, the jump rate of the process $x = m, n$, is given by

$$\hat{\lambda}^x(V,u) = \hat{\lambda}_0^x(V,u) + \hat{\lambda}_1^x(V,u).$$

Later we will check that such a process can be constructed in a unique globally way.

As before, the processes $\hat{m}$ and $\hat{n}$ are semimartingales with decomposition

$$\hat{x}_t^{N_c} = \hat{x}_0^{N_c} + \int_0^t (1-\hat{x}_s^{N_c})\rho_x(\hat{V}_s^{N_c}) - \hat{x}_s^{N_c}\zeta_x(\hat{V}_s^{N_c})ds + \hat{M}_t^x, \tag{4.2.12}$$

but in this case,

$$\begin{aligned}\left\langle \hat{M}^x \right\rangle_t &= \int_0^t\int_0^\infty \left(\frac{1}{N_c}\mathbb{1}_{\{\lambda\le\hat{\lambda}_0^x(\hat{V}_s^{N_c},\hat{x}_s^{N_c})\}}\right)^2 d\lambda ds + \int_0^t\int_0^\infty \left(\frac{-1}{N_c}\mathbb{1}_{\{\lambda\le\hat{\lambda}_1^x(\hat{V}_s^{N_c},\hat{x}_s^{N_c})\}}\right)^2 d\lambda ds\\ &= \frac{1}{{N_c}^2}\int_0^t \hat{\lambda}_0^x(\hat{V}_s^{N_c},\hat{x}_s^{N_c}) + \hat{\lambda}_1^x(\hat{V}_s^{N_c},\hat{x}_s^{N_c})ds\\ &= \frac{1}{N_c}\int_0^t (1-\hat{x}_s^{N_c})\rho_x(\hat{V}_s^{N_c}) + \hat{x}_s^{N_c}\zeta_x(\hat{V}_s^{N_c})ds + \int_0^t \sigma^2(\hat{V}_s^{N_c},\hat{x}_s^{N_c})\chi(\hat{x}_s^{N_c})ds.\end{aligned}$$

Proposition 4.2.6. *When the number of channels N_c goes to infinity, the perturbed hybrid process $(\hat{X}_t^{N_c})_{0\le t<\infty}$*

$$\begin{aligned}\hat{V}_t^{N_c} &= \hat{V}_0^{N_c} + \int_0^t I - g_L(\hat{V}_s^{N_c}-V_L) - g_{Ca}m_t^{N_c}(\hat{V}_s^{N_c}-V_{Ca}) - g_K\hat{n}_t^{N_c}(\hat{V}_s^{N_c}-V_K)ds\\ \hat{m}_t^{N_c} &= \hat{m}_0^{N_c} + \int_0^t (1-\hat{m}_s^{N_c})\rho_m(\hat{V}_s^{N_c}) - \hat{m}_s^{N_c}\zeta_m(\hat{V}_s^{N_c})ds + \hat{M}_t^m.\\ \hat{n}_t^{N_c} &= \hat{n}_0^{N_c} + \int_0^t (1-\hat{n}_s^{N_c})\rho_n(\hat{V}_s^{N_c}) - \hat{n}_s^{N_c}\zeta_n(\hat{V}_s^{N_c})ds + \hat{M}_t^n.\end{aligned} \tag{4.2.13}$$

converges in distribution to the solution of (4.2.10).

The proof is a straightforward application of the following theorem.

Theorem 4.2.7 (Theorem 4.1 in [28, p. 354])**.** *Let $a = (a_{ij})$ be a continous, symmetric, nonnegative definite, dxd matrix-valued function on $\mathbb{R}^d$ and let $b : \mathbb{R}^d \to \mathbb{R}^d$ be continuous. Let*

$$\mathcal{A} = \left\{\left(f, Gf := \frac{1}{2}\sum a_{ij}\partial_i\partial_j f + \sum b_i\partial_i f\right) : f \in \mathcal{C}_c^\infty(\mathbb{R}^d)\right\},$$

and suppose that the $C_{\mathbb{R}^d}$ martingale problem for A is well posed.

For $n = 1, 2, \ldots$, let X_n and B_n be processes with sample paths in $D_{\mathbb{R}^d}[0, \infty)$, and let $A_n = (A_n^{i,j})$ be a symmetric $d \times d$ matrix-valued process such that A_n^{ij} has sample paths in $D_{\mathbb{R}}[0, \infty)$ and $A_n(t) - A_n(s)$ is nonnegative definite for $t > s \geq 0$. Set $\mathcal{F}_t^n = \sigma(X_n(s), B_n(s), A_n(s) : s \leq t)$.

Let $\tau_n^r = \inf\{t : \|X_n(t)\| \geq r$ or $\|X_n(t-)\| \geq r\}$, and suppose that

$$M_n := X_n - B_n \tag{4.2.14}$$

and

$$M_n^i M_n^j - A_n^{ij}, \quad i, j = 1, 2, \ldots, d, \tag{4.2.15}$$

are $\mathcal{F}_t^n$-local martingales, and for each $r > 0$, $T > 0$, and $i, j = 1, \ldots, d$,

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t \leq T \wedge \tau_n^r} \|X_n(t) - X_n(t-)\|^2\right] = 0, \tag{4.2.16}$$

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t \leq T \wedge \tau_n^r} \|B_n(t) - B_n(t-)\|^2\right] = 0, \tag{4.2.17}$$

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t \leq T \wedge \tau_n^r} \|A_n^{ij}(t) - A_n^{ij}(t-)\|^2\right] = 0, \tag{4.2.18}$$

$$\sup_{t \leq T \wedge \tau_n^r} \left| B_n^{(i)}(t) - \int_0^t b_i(X_n(s))ds \right| \xrightarrow{\mathbb{P}} 0, \tag{4.2.19}$$

$$\sup_{t \leq T \wedge \tau_n^r} \left| A_n^{ij}(t) - \int_0^t a_{ij}(X_n(s))ds \right| \xrightarrow{\mathbb{P}} 0. \tag{4.2.20}$$

Suppose that $PX_n(0)^{-1} \Rightarrow \nu \in \mathcal{P}(\mathbb{R}^d)$. Then $\{X_n\}$ converges in distribution to the solution of the martingale problem for $(\mathcal{A}, \nu)$.

Proof of Proposition 4.2.6. In our case, is clear that hypotheses (4.2.14), (4.2.15), (4.2.17), (4.2.18), (4.2.19) hold. Since the well posedness of the limit equation is established in [15], all we need to check are hypotheses (4.2.16) and (4.2.20).

1. Let $\tau^r_{N_c} = \inf\{t \geq 0 : |\hat{X}^{N_c}_t| \geq r \vee |\hat{X}^{N_c}_{t-}| \geq r\}$.Thanks to the continuity of $\hat{V}^{N_c}_\cdot$

$$\begin{aligned}
\lim_{N_c\to\infty} \mathbb{E}\Big[\sup_{t\leq T\wedge\tau^r_{N_c}} \|\hat{X}^{N_c}_t - \hat{X}^{N_c}_{t-}\|^2\Big] \\
&= \lim_{N_c\to\infty} \mathbb{E}\left[\sup_{t\leq T\wedge\tau^r_N} |\hat{V}^{N_c}_t - \hat{V}^{N_c}_{t-}|^2 + |\hat{m}^{N_c}_t - \hat{m}^{N_c}_{t-}|^2 + |\hat{n}^{N_c}_t - \hat{n}^{N_c}_{t-}|^2\right] \\
&= \lim_{N_c\to\infty} \mathbb{E}\left[\sup_{t\leq T\wedge\tau^r_{N_c}} |\hat{m}^{N_c}_t - \hat{m}^{N_c}_{t-}|^2 + |\hat{n}^{N_c}_t - \hat{n}^{N_c}_{t-}|^2\right] \\
&\leq \lim_{N_c\to\infty} \frac{2}{{N_c}^2} = 0.
\end{aligned}$$

   Then (4.2.16) holds.

2. Thanks to the form of $A^{ij}_{N_c}$, we have

$$\left|\frac{1}{N_c}\int_0^t (1-\hat{x}^{N_c}_s)\rho_x(\hat{V}^{N_c}_s) + \hat{x}^{N_c}_s\zeta_x(\hat{V}^{N_c}_s)ds + \int_0^t \sigma^2(\hat{V}^{N_c}_s,\hat{x}^{N_c}_s)\chi(\hat{x}^{N_c}_s)ds - \int_0^t \sigma^2(\hat{V}^{N_c}_s,\hat{x}^{N_c}_s)\chi(\hat{x}^{N_c}_s)ds\right| \leq \frac{Ct}{N_c}.$$

   So, the left-hand side converges almost surely to 0, and hence in probability. Then we have (4.2.20).

Summing up, the hypotheses of the Theorem 4.2.7 are fulfilled, and our Proposition holds. □

Remark 4.2.8. There are several facts to notice about this result:

1. The convergence we obtain thanks to Theorem 4.2.7 is in distribution and without any information on the rate of convergence. To quantify the rate of convergence of a sequence of jump processes to their diffusive limit, when there is one, is an interesting mathematical problem without a general answer. In the context of Boltzmann equation Godinho in [36] find a coupling between a compensated Poisson Integral and a time-space white noise, but it is not clear that the ideas in [36] can be extended to the context of this thesis.

2. Although this result is a simple application of Theorem 4.2.7, it shows that the diffusive approximation proposed in [8] (4.2.10) is not the limit of the original hybrid model (4.2.8), and then should not surprise us that it has different properties.

3. In the original hybrid model, the jump rates where linear on N_c; whereas, in the perturbed hybrid model, we add in each jump rate the quadratic term

$$\frac{{N_c}^2}{2}\sigma^2(\hat{X}^{N_c}_{t-})\chi(\hat{x}^{N_c}_{t-}).$$

   This term could be interpreted as the increment on the jump rates of the channels due to the presence of the other channels.

4. Notice that in the original hybrid model the jump law of $x_t^{N_c}$ is given by

$$\mu_{N_c}(V_t^{N_c}, x, x') = \frac{(1-x)\rho_x(V_t^{N_c})}{(1-x)\rho_x(V_t^{N_c}) + x\zeta_x(V_t^{N_c})}\mathbb{1}_{\{x+\frac{1}{N_c}\}}(x') + \frac{u\zeta_x(V_t^{N_c})}{(1-x)\rho_x(V_t^{N_c}) + x\zeta_x(V_t^{N_c})}\mathbb{1}_{\{x-\frac{1}{N_c}\}}(x'),$$

and it is independent of the number of channels N_c. While in the contrary, in the perturbed model,

$$\hat{\mu}_{N_c}(\hat{V}_t^{N_c}, x, x') = \frac{(1-x)\rho_x(\hat{V}_t^{N_c}) + N_c\sigma^2(\hat{V}_t^{N_c}, x)\chi(x)/2}{(1-u)\rho_x(\hat{V}_t^{N_c}) + x\zeta_x(\hat{V}_t^{N_c}) + N_c\sigma^2(\hat{V}_t^{N_c}, x)\chi(x)}\mathbb{1}_{\{x+\frac{1}{N_c}\}}(x') + \frac{x\zeta_x(\hat{V}_t^{N_c}) + N_c\sigma^2(\hat{V}_t^{N_c}, x)\chi(x)/2}{(1-u)\rho_x(\hat{V}_t^{N_c}) + x\zeta_x(\hat{V}_t^{N_c}) + N_c\sigma^2(\hat{V}_t^{N_c}, u)\chi(x)}\mathbb{1}_{\{x-\frac{1}{N_c}\}}(x').$$

so when $N_c >> 1$ the jump law is

$$\hat{\mu}_{N_c}(\hat{V}_t^{N_c}, x, x') \approx \frac{1}{2}\mathbb{1}_{\{x+\frac{1}{N_c}\}}(x') + \frac{1}{2}\mathbb{1}_{\{x-\frac{1}{N_c}\}}(x').$$

Here we find another qualitative difference between the unperturbed hybrid model and the perturbed one, and consequently, between the former and the diffusive approximation.

Now that we have discussed the modeling issues arround the stochastic versions of the Morris-Lecar model, we start the study of network of neurons in this setting.

4.3 Interaction between neurons

In what follows, we will consider a network of N neurons interacting with each other in a fully connected net. Neurons are not homogeneous. In fact, in the human brain there are around $10,000$ different kind of neurons[1]. Nevertheless, we will assume that all the neurons are of the same kind, because the generalization to several populations it is not difficult.

An connection between neurons is called a synapse. Synapses could be chemical or electrical. We refer briefly to both types in the next paragraphs. For details see Chapter 8 in [27].

Chemical Synapses In a chemical synapses a neurotransmitter is released to the synaptic cleft (see Figure 1.3.2a) from the pre-synaptic neuron to the post-synaptic one. The release of this neurotransmitter will occur through some channels, which will open and close depending on the membrane voltage in a similar way as for the Calcium and Potassium channels. Let

[1] See [27, p.4]

us denote by y_t the proportion of open neurotransmitters channels open at time t, and by

$$\rho_y(V) = a_r \frac{C_T}{1+e^{-\Lambda_T(V-V_T)}}, \quad \zeta_y(V) = a_d,$$

where the constants a_r, a_d, C_T, Λ_T and V_T are determined experimentally. So, the jump rates for the process y will be given by

$$\begin{aligned} \lambda_0^y(V,y) &= N_c(1-y)\rho_y(V), \\ \lambda_1^y(V,y) &= N_c y \zeta_y(V), \end{aligned} \tag{4.3.1}$$

or, in the perturbed case, by

$$\begin{aligned} \lambda_0^y(V,y) &= N_c(1-y)\rho_y(V) + \frac{N_c^2}{2}\sigma_y^2(V,y)\chi^2(y) \\ \lambda_1^y(V,y) &= N_c a_d \zeta_y(V) + \frac{N_c^2}{2}\sigma_y^2(V,y)\chi^2(y), \end{aligned} \tag{4.3.2}$$

where as before

$$\sigma_y(V,y) = \sigma\sqrt{|(1-y)\rho_y(V) + y\zeta_y(V)|}. \tag{4.3.3}$$

Then, for the i-th neuron, the dynamic for the proportion of open neurotransmitter channels will be

$$y_t^{i,N_c} = y_0^{i,N_c} + \int_0^t (1-y_s^{i,N_c})\rho_y(V_s^{i,N_c}) - y_s^{i,N_c}\zeta_y(V_s^{i,N_c})ds + M_t^{y,i},$$

and due to the chemical synapses coming from neuron j, the contribution to the dynamic of the voltage of the neuron i will be

$$-J_{\text{Ch}} y_s^{j,N_c}(V_s^{i,N_c} - V_{\text{rev}}),$$

where J_{Ch} is the chemical conductance of the network, and V_{rev} is the reversal potential, which indicates if the synapses are inhibitory or excitatory. Some values from the literature for the constants $a_d, a_r, C_T, \Lambda_T, V_T$ appear in Table 4.3.1a. In Figure 4.3.1 we show the rate functions ρ_y and ζ_y.

Electrical Synapses Sometimes the interior of one neuron can be directly connected with the interior of another neuron through an intercellular channel called *gap junction*. These gap junctions allows the constant flow of current between neurons as a result of the difference of potential between them.

Electrical synapses are less frequent than the chemical ones, but they transmite information faster, and also have been observed to induce synchronization between neurons. See [43].

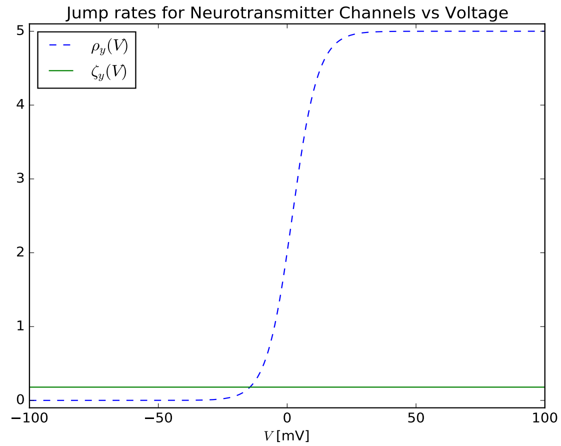
Since current flows from higher voltages to lower ones, the contribution to the dynamic of the voltage of the neuron i, due to the electrical synapse with neuron j, will be

$$-J_{\text{E}}(V_s^{i,N_c} - V_s^{j,N_c}),$$

where J_{E} is the electrical conductance and can be thought of as a measure of the connectivity of the network. For a further reading on electrical synapses see [43].

Parameter	Value
V_{rev}	1
C_T	1.0
V_T	2.0
Λ_T	0.2
a_r	5.0
a_d	0.18

(a) Example of values for the parameters of the dynamics for the neurotransmitter channel. The first four rows are taken from [8], others from [27].



(b) Rate functions for the neurotransmitter channel.

Figure 4.3.1: Values for the parameters and rate functions for the Neurotransmitter dynamics

Dynamic for the Voltage of a Neuron in a Network Finally, the dynamic for the potential of the membrane of the neuron in a network will be given by

$$\begin{aligned} V_t^{i,N_c} = V_0^{i,N_c} &+ \int_0^t I - g_{\text{Ca}} m_s^{i,N_c} t(V_s^{i,N_c} - V_{\text{Ca}}) - g_{\text{K}} n_s^{i,N_c}(V_s^{i,N_c} - V_{\text{K}}) ds \\ &- \int_0^t g_{\text{L}}(V_s^{i,N_c} - V_{\text{L}}) ds - \frac{1}{N}\sum_{j=1}^{N} \int_0^t J_{\text{Ch}} y_s^{j,N_c}(V_s^{i,N_c} - V_{\text{rev}}) ds \\ &- \frac{1}{N}\sum_{j=1}^{N} \int_0^t J_{\text{E}}(V_s^{i,N_c} - V_s^{j,N_c}) ds. \end{aligned} \tag{4.3.4}$$

4.3.1 The Hybrid Model for a Network of Interacting Neurons

In a fully connected network of N neurons, each neuron in the network will be described by a 4-tuple $(V^{i,N_c}, m^{i,N_c}, n^{i,N_c}, y^{i,N_c})$. As before, to describe the dynamics of the different channels, we will consider $\mathcal{N}_0^{n,i}$,$\mathcal{N}_1^{n,i}$, $\mathcal{N}_0^{m,i}$, $\mathcal{N}_1^{m,i}$, $\mathcal{N}_0^{y,i}$ and $\mathcal{N}_1^{y,i}$, independent Poisson Measures on $\mathbb{R}_+ \times \mathbb{R}_+$, each of them with intensity $d\lambda \otimes ds$.

For each neuron, the dynamic of its membrane voltage will be given by (4.3.4). On the other hand, for the proportion of open channels we will consider either the jump rates given by (4.2.5) and (4.3.1), or in the perturbed case, the jump rates given by (4.2.11) and (4.3.2). With all these notations, the hybrid model for the system will be given by

$$\begin{aligned} V_t^{(i)} &= V_0^{(i)} + \int_0^t I - g_{\text{Ca}} m_s^{(i)}(V_s^{(i)} - V_{\text{Ca}}) - g_{\text{K}} n_s^{(i)}(V_s^{(i)} - V_{\text{K}}) ds \\ &\quad - \int_0^t g_{\text{L}}(V_s^{(i)} - V_{\text{L}}) ds - \int_0^t J_{\text{Ch}} \bar{y}_s^N (V_s - V_{\text{rev}}) ds \\ &\quad - \int_0^t J_{\text{E}}(V_s^{i,N_c} - \bar{V}_s^N) ds, \qquad (4.3.5) \\ x_t^{(i)} &= x_0^{(i)} + \int_0^{t+} \int_0^\infty \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_0^x(V_{s-}^{(i)}, x_{s-}^{(i)})\}} \mathcal{N}_0^{x,i}(d\lambda, ds) \\ &\quad - \int_0^{t+} \int_0^\infty \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_1^x(V_{s-}^{(i)}, x_{s-}^{(i)})\}} \mathcal{N}_1^{x,i}(d\lambda, ds), \quad \text{for } x = m, n, y. \end{aligned}$$

where

$$\bar{V}_s^N = \frac{1}{N} \sum_{j=1}^N V_s^{j,N_c}), \quad \bar{y}_s^N = \frac{1}{N} \sum_{j=1}^N y_s^{j,N_c}.$$

In the following proposition we will prove the boundedness of the solutions of the hybrid models. This will be crucial to obtain the main results of this chapter.

Proposition 4.3.1. *If the initial condition $V_0 = (V_0^{1,N_c}, \ldots, V_0^{N_c,N})$ is bounded, (i.e.) for all $i = 1, \ldots, N$, $|V_0^{i,N_c}| \le V_0^{max}$, then there exist a constant $V_\infty^{max} = V_\infty^{max}(V_0^{max})$ independent of N_c and N, such that for all $t \ge 0$*

$$|V_t^{i,N_c}| \le V_\infty^{max}.$$

Before we prove last proposition, we recall a version of Gronwall's Lemma which appears in [4, p. 88].

Lemma 4.3.2. *Let $x : [0, +\infty) \to \mathbb{R}$ be a locally absolutely continuous function, let $a, b \in L^1_{loc}([0, +\infty))$ be given functions satisfying, for $\lambda \in \mathbb{R}$,*

$$\frac{d}{dt} x^2(t) + 2\lambda x^2(t) \le a(t) + 2b(t)x(t) \text{ for } \mathcal{L}^\infty - a.e.\ t > 0.$$

Then for every $T > 0$ we have

$$e^{\lambda T} |x(T)| \le \left(x^2(0) + \sup_{t \in [0,T]} \int_0^t e^{2\lambda s} a(s) ds \right)^{1/2} + 2 \int_0^T e^{\lambda t} |b(t)| dt.$$

Proof of Proposition 4.3.1. To simplify the notation we will remove the N_c superscript in this proof. Notice that calling

$$R_s^{(i)} = I + g_{\text{Ca}} V_{\text{Ca}} m_s^{(i)} + g_{\text{K}} V_{\text{K}} n_s^{(i)} + g_{\text{L}} V_{\text{L}} + J_{\text{Ch}} V_{\text{rev}} \bar{y}_s^N,$$

$$A_s^{(i)} = g_{\mathrm{Ca}} m_s^{(i)} + g_{\mathrm{K}} n_s^{(i)} + g_{\mathrm{L}} + J_{\mathrm{Ch}} \bar{y}_s^N,$$

the dynamic for the potential can be written as

$$V_t^{(i)} = V_0^{(i)} + \int_0^t R_s^{(i)} - A_s^{(i)} V_s^{(i)} - J_{\mathrm{E}} V_s^{(i)} + J_{\mathrm{E}} \bar{V}_s^N ds. \tag{4.3.6}$$

Then

$$\left(V_t^{(i)}\right)^2 = \left(V_0^{(i)}\right)^2 + 2\int_0^t R_s^{(i)} V_s^{(i)} - A_s^{(i)} \left(V_s^{(i)}\right)^2 - J_{\mathrm{E}} \left(V_s^{(i)}\right)^2 + J_{\mathrm{E}} \bar{V}_s^N V_s^{(i)} ds.$$

If we add for $i = 1, \ldots, N$ we obtain for $V = (V^{(1)}, \ldots, V^{(N)})$

$$\|V_t\|_2^2 = \|V_0\|_2^2 + 2\int_0^t \sum_{i=1}^N \left[R_s^{(i)} V_s^{(i)} - A_s^{(i)} \left(V_s^{(i)}\right)^2\right] - J_{\mathrm{E}} \|V_s\|_2^2 + N J_{\mathrm{E}} (\bar{V}_s^N)^2 ds.$$

Notice that

$$(\bar{V}_s^N)^2 = \frac{1}{N^2} \sum_{i,j=1}^N V_s^{(i)} V_s^{(j)} \leq \frac{1}{2N^2} \sum_{i,j=1}^N (V_s^{(i)})^2 + (V_s^{(j)})^2 = \frac{1}{N} \|V_s\|_2^2.$$

Then we have

$$\|V_t\|_2^2 \leq \|V_0\|_2^2 + 2\int_0^t \|R_s\|_2 \|V_s\|_2 - g_{\mathrm{L}} \|V_s\|_2^2 ds,$$

or in its differential form

$$\frac{d}{dt} \|V_t\|_2^2 + 2 g_{\mathrm{L}} \|V_t\|_2^2 \leq 2 \|R_t\|_2 \|V_t\|_2.$$

Applying Lemma 4.3.2 to this last inequality, we conclude

$$\|V_t\|_2 \leq \|V_0\|_2 e^{-g_{\mathrm{L}} t} + 2 e^{-g_{\mathrm{L}} t} \int_0^t e^{g_{\mathrm{L}} s} \|R_s\|_2 ds.$$

But we have

$$|R_s^{(i)}| \leq \max_{m,n,y \in [0,1]} |I + g_{\mathrm{Ca}} V_{\mathrm{Ca}} m + g_{\mathrm{K}} V_{\mathrm{K}} n + g_{\mathrm{L}} V_{\mathrm{L}} + J_{\mathrm{Ch}} V_{\mathrm{rev}} y| =: R_{\max},$$

so that

$$\|V_t\|_2 \leq \|V_0\|_2 e^{-g_{\mathrm{L}} t} + \frac{2\sqrt{N} R_{\max}}{g_{\mathrm{L}}} e^{-g_{\mathrm{L}} t} (e^{g_{\mathrm{L}} t} - 1) = \sqrt{N} \left(V_0^{\max} - \frac{2 R_{\max}}{g_{\mathrm{L}}}\right) e^{-g_{\mathrm{L}} t} + \frac{2\sqrt{N} R_{\max}}{g_{\mathrm{L}}},$$

and we deduce that

$$\sqrt{\frac{1}{N} \sum_{i=1}^N (V_t^{(i)})^2} \leq \left(V_0^{\max} - \frac{2 R_{\max}}{g_{\mathrm{L}}}\right) e^{-g_{\mathrm{L}} t} + \frac{2 R_{\max}}{g_{\mathrm{L}}}.$$

On the other hand, notice that

$$\left|\bar{V}_t^N\right| = \left|\frac{1}{N}\sum_{i=1}^N V_t^{(i)}\right| \leq \frac{1}{N}\sum_{i=1}^N |V_t^{(i)}| \leq \sqrt{\frac{1}{N}\sum_{i=1}^N (V_t^{(i)})^2},$$

from where

$$\left|\bar{V}_t^N\right| \leq \left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)e^{-g_{\mathrm{L}}t} + \frac{2R_{\max}}{g_{\mathrm{L}}}.$$

If we go back to (4.3.6), we observe that

$$(V_t^{(i)})^2 = (V_0^{(i)}) + 2\int_0^t (R_s^{(i)} + J_{\mathrm{E}}\bar{V}_s^N)V_s^{(i)} - (A_s^{(i)} + J_{\mathrm{E}})(V_s^{(i)})^2 ds,$$

from where it follows that

$$\begin{aligned}\frac{d}{dt}(V_t^{(i)})^2 + 2(g_{\mathrm{L}} + J_{\mathrm{E}})(V_t^{(i)})^2 &\leq 2|R_t^{(i)} + J_{\mathrm{E}}\bar{V}_t^N||V_t^{(i)}| \\ &\leq 2\left(R_{\max} + J_{\mathrm{E}}\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)e^{-g_{\mathrm{L}}t} + \frac{2J_{\mathrm{E}}R_{\max}}{g_{\mathrm{L}}}\right)|V_t^{(i)}|.\end{aligned}$$

Applying one more time Lemma 4.3.2, we obtain

$$\begin{aligned}\left|V_t^{(i)}\right| &\leq V_0^{\max}e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t} \\ &\quad + 2e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t}\int_0^t e^{(g_{\mathrm{L}}+J_{\mathrm{E}})s}\left(R_{\max}\left(\frac{g_{\mathrm{L}}+2J_{\mathrm{E}}}{g_{\mathrm{L}}}\right) + J_{\mathrm{E}}\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)e^{-g_{\mathrm{L}}s}\right)ds \\ &= V_0^{\max}e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t} + 2R_{\max}\left(\frac{g_{\mathrm{L}}+2J_{\mathrm{E}}}{g_{\mathrm{L}}(g_{\mathrm{L}}+J_{\mathrm{E}})}\right)(1 - e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t}) \\ &\quad + 2\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)\left(e^{-g_{\mathrm{L}}t} - e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t}\right) \\ &= \frac{2R_{\max}}{g_{\mathrm{L}}}\left(\frac{g_{\mathrm{L}}+2J_{\mathrm{E}}}{g_{\mathrm{L}}+J_{\mathrm{E}}}\right) + 2\left(V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\right)e^{-g_{\mathrm{L}}t} \\ &\quad + \left(\frac{4R_{\max}}{g_{\mathrm{L}}} - V_0^{\max} - \frac{2R_{\max}}{g_{\mathrm{L}}}\left(\frac{g_{\mathrm{L}}+2J_{\mathrm{E}}}{g_{\mathrm{L}}+J_{\mathrm{E}}}\right)\right)e^{-(g_{\mathrm{L}}+J_{\mathrm{E}})t}.\end{aligned}$$

If $V_\infty^{\max}$ denote the supremum over $t \geq 0$ of the right-hand side of last expresion, which is finite since the right side is bounded, we conclude that $\forall i = 1, \ldots, N$

$$\sup_{t\geq 0}\left|V_t^{(i)}\right| \leq V_\infty^{\max}.$$

□

Remark 4.3.3. Notice that in the last proof, we only used the boundedness of the processes m, n and y.

At this point we will show that the hybrid model (4.3.5) is well posed.

Lemma 4.3.4. *If the initial condition are almost surely bounded, then there exists a unique strong solution for the system* (4.3.5).

Remark 4.3.5. Notice that the rate functions λ_i^x are Lipchitz, since they are a combination of ρ_x, ζ_x and σ_x^2.

Remark 4.3.6. Notice that one of the usual hypotheses for an equation like (4.3.5) to be well-posed is that

$$\int_0^\infty |f(v_1,x_1,\lambda) - f(v_2,x_2,\lambda)|^2 d\lambda \leq K\left[|v_1-v_2|^2 + |x_1-x_2|^2\right],$$

which corresponds to a certain Lipschitz property of the coefficients of the equation.

Since, in our case, $f(v,x,\lambda) = \mathbb{1}_{\{\lambda \leq \lambda_i^x(v,x)\}}$, we have

$$|\mathbb{1}_{\{\lambda\leq\lambda_i^x(v_1,x_1)\}} - \mathbb{1}_{\{\lambda\leq\lambda_i^x(v_2,x_2)\}}|^2 = \mathbb{1}_{\{\lambda_i^x(v_1,x_1)\wedge\lambda_i^x(v_2,x_2)\leq\lambda\leq\lambda_i^x(v_1,x_1)\vee\lambda_i^x(v_2,x_2)\}},$$

and then

$$\begin{aligned}\int_0^\infty |\mathbb{1}_{\{\lambda\leq\lambda_i^x(v_1,x_1)\}} - \mathbb{1}_{\{\lambda\leq\lambda_i^x(v_2,x_2)\}}|^2 d\lambda &= \int_0^\infty \mathbb{1}_{\{\lambda_i^x(v_1,x_1)\wedge\lambda_i^x(v_2,x_2)\leq\lambda\leq\lambda_i^x(v_1,x_1)\vee\lambda_i^x(v_2,x_2)\}} d\lambda \\ &= |\lambda_i^x(x_1,v_1) - \lambda_i^x(x_2,v_2)|. \\ &\leq K_\lambda^{i,x}\left[|v_1-v_2| + |x_1-x_2|\right].\end{aligned}$$

Thus, in order to prove the well-posedness of equation (4.3.5), we can't rely on known results and we need to do the construction by hand.

Proof of Lemma 4.3.4. We will construct a solution and later we will prove that such solution is unique.

Existence: Since in the system (4.3.5) involves only integrals with respect to the Lebesgue measure or with respect to Poisson random measures, we can interpret this system in a pathwise sense, or in more informal words, ω by ω.

Just as in the previous proof we omit the N_c superscript to make the notation easier to read. In what follows we will call $V = (V^{(1)},\dots,V^{(N)})$, $m = (m^{(1)},\dots,m^{(N)})$, $n = (n^{(1)},\dots,n^{(N)})$ and $y = (y^{(1)},\dots,y^{(N)})$.

Let us fix $\omega \in \Omega$. If we replace m_s, n_s, y_s for some fix $m^*, n^*, y^* \in [0,1]^N$ the equation for V in (4.3.5) becomes a linear equation with constant coefficient, and then it has an explicit solution $\Phi_t(V_0(\omega), m^*, n^*, y^*)$ which, according to Proposition (4.3.6), is uniformly bounded in time.

On the other hand, since the functions ρ_x and ζ_x are bounded, we have, in the non perturbed case

$$\lambda_i^x(v,x) \leq CN_c|x|, \text{ for } x = m,n,y, \text{ and } i = 0,1,$$

whereas, in the perturbed case

$$\lambda_i^x(v,x) \le C(N_c + N_c^2)|x|, \text{ for } x = m, n, y, \text{ and } i = 0, 1.$$

So, in both cases we have that $\lambda_i^x(v,x) \le C_\lambda$ for some $C_\lambda > 0$ uniformly on (v,x).

Notice also the different Poisson measures appearing in (4.3.5) can be packed in one Poisson random measure $\mathcal{N}$ on $\mathbb{R}_+ \times \mathbb{R}_+ \times \{1,\dots,N\} \times \{m,n,y\} \times \{0,1\}$, which have almost surely a finite number of atoms in the set $R = [0,T] \times [0, C_\lambda] \times \{1,\dots,N\} \times \{m,n,y\} \times \{0,1\}$. Since the indicator functions appearing in the Poisson integrals for $m^{(i)}$, $n^{(i)}$ and $y^{(i)}$ in (4.3.5) are equal to 0 for $\lambda \ge C_\lambda$, the only atoms of the Poisson measure we care about, are the $\mathcal{N}_\omega(R)$ atoms in the set R. Let us call $\{J_n\}_{n=1}^{\mathcal{N}_\omega(R)}$ the sequence of jump times. For notational convenience $J_0 := 0$. We will construct the solution for (4.3.5) in each interval $[J_n, J_{n+1}]$ inductively.

In the interval $[0, J_1(\omega))$, $m_t(\omega) = m_0(\omega)$, $n_t(\omega) = n_0(\omega)$, $y_t(\omega) = y_0(\omega)$ so for all $t \in [0, J_1(\omega))$

$$V_t = \Phi_t(V_0(\omega), m_0(\omega), n_0(\omega), y_0(\omega)),$$

and then, the Poisson integrals for $m^{(i)}, n^{(i)}$ and $y^{(i)}$ are well defined in $[0, J_1(\omega))$. Now we extend the solution to the interval $[J_1(\omega), J_2(\omega))$.

Without loss of generality let us assume that the atom associated to $J_1(\omega)$ is equal to $(J_1(\omega), \lambda^*, i, m, 0)$. This means that a closed Calcium channel of the i-th neuron is candidate to open. If

$$\lambda_0^m\left(V^{(i)}_{J_1(\omega)-}, m^{(i)}_{J_1(\omega)-}\right) \ge \lambda^*,$$

then the Calcium channel will change its state to open and m will change to $\hat{m}$, with $\hat{m}^{(i)} = m^{(i)} + 1/N_c$. Then, for $t \in [J_1(\omega), J_2(\omega))$

$$V_t = \Phi_t(V_{J_1(\omega)-}, \hat{m}, n_0(\omega), y_0(\omega)).$$

If

$$\lambda_0^m\left(V^{(i)}_{J_1(\omega)-}, m^{(i)}_{J_1(\omega)-}\right) < \lambda^*,$$

then the Calcium channel will remain closed, and then for $t \in [J_1(\omega), J_2(\omega))$

$$V_t = \Phi_t(V_0(\omega), m_0(\omega), n_0(\omega), y_0(\omega)).$$

Iterating this procedure $\mathcal{N}_\omega(R)$ times, we obtain a solution on $[0,T]$ for each ω in a subset of probability one of the subjacent probability space.

Uniqueness: Let $(V_t, m_t, n_t, y_t)_{t\in[0,T]}$ and $(U_t, \alpha_t, \beta_t, \gamma_t)_{t\in[0,T]}$, be two solutions for (4.3.5) for the same initial condition. Then

$$\begin{aligned} |V_t^{(i)} - U_t^{(i)}| \le & \int_0^t g_{\mathrm{Ca}}|V_{\mathrm{Ca}} - V_s^{(i)}||m_s^{(i)} - \alpha_s^{(i)}| + g_{\mathrm{K}}|V_{\mathrm{K}} - V_s^{(i)}||n_s^{(i)} - \beta_s^{(i)}|ds \\ & + \int_0^t |g_{\mathrm{L}} + g_{\mathrm{Ca}}\alpha_s^{(i)} + g_{\mathrm{K}}\beta_s^{(i)} + J_{\mathrm{Ch}}\bar{\gamma}_s^N + J_{\mathrm{E}}||V_s^{(i)} - U_s^{(i)}|ds \\ & + \int_0^t |J_{\mathrm{Ch}}V_{\mathrm{rev}}||\bar{y}_s^N - \bar{\gamma}_s^N| + J_{\mathrm{E}}|\bar{V}_s^N - \bar{U}_s^N|ds. \end{aligned}$$

Thanks to the uniform bound on V, if we add this last bound for $i = 1, \dots, N$ we have

$$\|V_t - U_t\|_1 \leq C_V \int_0^t \|m_s - \alpha_s\|_1 + \|n_s - \beta_s\|_1 + \|y_s - \gamma_s\|_1 + \|V_s - U_s\|_1 ds.$$

On the other hand,

$$\begin{aligned} m_t^{(i)} - \alpha_t^{(i)} = & \int_0^{t+}\int_0^\infty \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_0^m(V_{s-}^{(i)}, m_{s-}^{(i)})\right\}} - \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_0^m(V_{s-}^{(i)}, \alpha_{s-}^{(i)})\right\}} \mathcal{N}_0^m(d\lambda ds) \\ & - \int_0^{t+}\int_0^\infty \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_1^m(V_{s-}^{(i)}, m_{s-}^{(i)})\right\}} - \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_1^m(V_{s-}^{(i)}, \alpha_{s-}^{(i)})\right\}} \mathcal{N}_1^m(d\lambda ds). \end{aligned}$$

Then, we obtain

$$\begin{aligned} \mathbb{E}\left|m_t^{(i)} - \alpha_t^{(i)}\right| \leq & \ \mathbb{E}\int_0^{t+}\int_0^\infty \left|\frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_0^m(V_{s-}^{(i)}, m_{s-}^{(i)})\right\}} - \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_0^m(V_{s-}^{(i)}, \alpha_{s-}^{(i)})\right\}}\right| \mathcal{N}_0^m(d\lambda ds) \\ & + \mathbb{E}\int_0^{t+}\int_0^\infty \left|\frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_1^m(V_{s-}^{(i)}, m_{s-}^{(i)})\right\}} - \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_1^m(V_{s-}^{(i)}, \alpha_{s-}^{(i)})\right\}}\right| \mathcal{N}_1^x(d\lambda ds) \\ = & \ \mathbb{E}\int_0^{t+}\int_0^\infty \left|\frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_0^m(V_{s}^{(i)}, m_{s}^{(i)})\right\}} - \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_0^m(V_{s}^{(i)}, \alpha_{s}^{(i)})\right\}}\right| d\lambda ds \\ & + \mathbb{E}\int_0^{t+}\int_0^\infty \left|\frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_1^m(V_{s}^{(i)}, m_{s}^{(i)})\right\}} - \frac{1}{N_c}\mathbb{1}_{\left\{\lambda \leq \lambda_1^m(V_{s}^{(i)}, \alpha_{s}^{(i)})\right\}}\right| d\lambda ds \\ = & \ \mathbb{E}\int_0^{t}\int_0^\infty \frac{1}{N_c}\mathbb{1}_{\left\{\lambda_0^m(V_s^{(i)}, m_s^{(i)}) \wedge \lambda_0^m(V_s^{(i)}, \alpha_s^{(i)}) \leq \lambda \leq \lambda_0^m(V_s^{(i)}, m_s^{(i)}) \vee \lambda_0^m(V_s^{(i)}, \alpha_s^{(i)})\right\}} d\lambda ds \\ & + \mathbb{E}\int_0^{t}\int_0^\infty \frac{1}{N_c}\mathbb{1}_{\left\{\lambda_1^m(V_s^{(i)}, m_s^{(i)}) \wedge \lambda_1^m(V_s^{(i)}, \alpha_s^{(i)}) \leq \lambda \leq \lambda_1^m(V_s^{(i)}, m_s^{(i)}) \vee \lambda_1^m(V_s^{(i)}, \alpha_s^{(i)})\right\}} d\lambda ds \\ = & \ \mathbb{E}\int_0^t \frac{1}{N_c}|\lambda_0^m(V_s^{(i)}, m_s^{(i)}) - \lambda_0^m(V_s^{(i)}, \alpha_s^{(i)})| ds \\ & + \mathbb{E}\int_0^t \frac{1}{N_c}|\lambda_1^m(V_s^{(i)}, m_s^{(i)}) - \lambda_1^m(V_s^{(i)}, \alpha_s^{(i)})| ds \\ \leq & \ K_m \int_0^t \left(\mathbb{E}\left|m_s^{(i)} - \alpha_s^{(i)}\right| + \mathbb{E}\left|V_s^{(i)} - V_s^{(i)}\right|\right) ds. \end{aligned}$$

Adding over $i = 1, \dots, N$ we conclude that

$$\mathbb{E}\left[\|m_t - \alpha_t\|_1\right] \leq K_m \int_0^t \mathbb{E}\left[\|m_s - \alpha_s\|_1\right] + \mathbb{E}\left[\|V_s - V_s\|_1\right] ds.$$

For the other coordinates the argument is similar. Then, we have

$$\begin{aligned} \mathbb{E}\big[\|V_t - U_t\|_1 + & \|m_t - \alpha_t\|_1 + \|n_t - \beta_t\|_1 + \|y_t - \gamma_t\|_1\big] \\ & \leq K \int_0^t \mathbb{E}\left[\|V_s - U_s\|_1 + \|m_s - \alpha_s\|_1 + \|n_s - \beta_s\|_1 + \|y_s - \gamma_s\|_1\right] ds, \end{aligned}$$

from where Gronwall's Lemma yields

$$\mathbb{E}\big[\|V_t - U_t\|_1 + \|m_t - \alpha_t\|_1 + \|n_t - \beta_t\|_1 + \|y_t - \gamma_t\|_1\big] = 0.$$

Since the trajectories of $(V_t, m_t, n_t, y_t)_{t\in[0,T]}$ and $(U_t, \alpha_t, \beta_t, \gamma_t)_{t\in[0,T]}$ are *càdlàg*, we conclude that both processes are indistinguishable on $[0, T]$.

Since the solution of (4.3.5) is uniformly bounded for $t \geq 0$, we can extend it to $[0, \infty)$. □

Remark 4.3.7. Notice that the construction procedure of the solution for (4.3.5) provide also an algorithm to simulate it.

Remark 4.3.8. This proof shows that the single neuron equation is also well posed.

4.3.2 Continuous Approximations for the hybrid models

As in the case of a single neuron, the hybrid interacting system can be approximated by a continuous one. Let us recall that after compensating the Poisson measures the system (4.3.5) is equivalent to

$$\begin{aligned} V_t^{i,N_c} &= V_0^{i,N_c} + \int_0^t I - g_{\mathrm{Ca}} m_s^{i,N_c}(V_s^{i,N_c} - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_s^{i,N_c}(V_s^{i,N_c} - V_{\mathrm{K}}) ds \\ &\quad - \int_0^t g_{\mathrm{L}}(V_s^{i,N_c} - V_{\mathrm{L}}) ds - \int_0^t J_{\mathrm{Ch}} \bar{y}_s^N (V_s^{i,N_c} - V_{\mathrm{rev}}) ds \\ &\quad - \int_0^t J_{\mathrm{E}}(V_s^{i,N_c} - \bar{V}_s^N) ds \\ x_t^{i,N_c} &= x_0^{i,N_c} + \int_0^t (1 - x_s^{i,N_c})\rho_x(V_s^{i,N_c}) - x_s^{i,N_c}\zeta_x(V_s^{i,N_c}) ds + M_t^{x,i}, \quad x = m, n, y, \end{aligned} \tag{4.3.7}$$

where recall $M^{x,i}$ is the difference of two stochastic integrals with respect to compensated Poisson measures.

Proposition 4.3.9. *Fix $T > 0$ and $N \geq 1$, let*

$$X_t^{N,N_c} = (V_t^{1,N_c}, n_t^{1,N_c}, m_t^{1,N_c}, y_t^{1,N_c}, \dots, V_t^{N,N_c}, n_t^{N,N_c}, m_t^{N,N_c}, y_t^{N,N_c}),$$

be the stochastic process defined by (4.3.7) when the jump rate functions for m and n are given by (4.2.5) and the jump rate function for y is given by (4.3.1). That is, the hybrid model is not perturbed. Let

$$X_t^N = (V_t^{(1)}, n_t^{(1)}, m_t^{(1)}, h_t^{(1)}, y_t^{(1)}, \dots, V_t^N, n_t^N, m_t^N, h_t^N, y_t^N),$$

be the solution of

$$\begin{aligned} V_t^{(i)} &= V_0^{i,} + \int_0^t I - g_{Ca} m_s^{(i)}(V_s^{(i)} - V_{Ca}) - g_K n_s^{(i)}(V_s^{(i)} - V_K) ds \\ &\quad - \int_0^t g_L(V_s^{(i)} - V_L) ds - \int_0^t J_{Ch} \bar{y}_s^N (V_s^{(i)} - V_{rev}) ds \\ &\quad - \int_0^t J_E(V_s^{(i)} - \bar{V}_s^N) ds \\ x_t^{(i)} &= x_0^{(i)} + \int_0^t (1 - x_s^{(i)})\rho_x(V_s^{(i)}) - x_s^{(i)}\zeta_x(V_s^{(i)}) ds, \quad x = m, n, y. \end{aligned} \tag{4.3.8}$$

Then

$$\mathbb{E}\left[\sup_{0\le t\le T}\|X_t^{N,N_c}-X_t^N\|_1\right]\le\frac{NCe^{CT}}{\sqrt{N_c}}.$$

In particular

$$\lim_{N_c\to\infty}\mathbb{E}\left[\sup_{0\le t\le T}\|X_t^{N,N_c}-X_t^N\|_1\right]=0.$$

The proof of this Proposition is essentially the same of Proposition 4.2.4, so we omit it. The only relevant thing to notice, is that the bound is linear in the number of neurons.

In the perturbed case, as before, we only have convergence in distribution.

Proposition 4.3.10. *Fix $T>0$ and $N\ge 1$, let*

$$X_t^{N,N_c}=(V_t^{1,N_c},n_t^{1,N_c},m_t^{1,N_c},y_t^{1,N_c},\dots,V_t^{N,N_c},n_t^{N,N_c},m_t^{N,N_c},y_t^{N,N_c}),$$

be the stochastic process defined by (4.3.7). If the jump rate functions for m and n are given by (4.2.11) and the jump rate function for y is given by (4.3.2), that is the model is perturbed, and N_c goes to infinity, then the process X_t^{N,N_c} converges in distribution to

$$X_t^N=(V_t^{(1)},n_t^{(1)},m_t^{(1)},h_t^{(1)},y_t^{(1)},\dots,V_t^N,n_t^N,m_t^N,h_t^N,y_t^N),$$

the solution of

$$\begin{aligned}
V_t^{(i)}&=V_0^{i,}+\int_0^t I-g_{Ca}m_s^{(i)}t(V_s^{(i)}-V_{Ca})-g_Kn_s^{(i)}(V_s^{(i)}-V_K)ds\\
&\quad-\int_0^t g_L(V_s^{(i)}-V_L)ds-\int_0^t J_{Ch}\bar{y}_s^N(V_s^{(i)}-V_{rev})ds\\
&\quad-\int_0^t J_E(V_s^{(i)}-\bar{V}_s^N)ds\\
x_t^{(i)}&=x_0^{(i)}+\int_0^t(1-x_s^{(i)})\rho_x(V_s^{(i)})-x_s^{(i)}\zeta_x(V_s^{(i)})ds\\
&\quad+\int_0^t\sigma_x(V_s^{(i)},x_s^{(i)})\chi(x_s^{(i)})dW_s^{i,x},\quad x=m,n,y.
\end{aligned}\tag{4.3.9}$$

Remark 4.3.11. If we go back to the proof of Proposition 4.3.1, we will see that boundedness is the only property of the variables m,n,y that we have used. Then, the same computation shows that the solutions of (4.3.8) and (4.3.9) are also bounded.

Proof of Proposition 4.3.10. As before, the proof of this proposition is a straightforward application of Theorem 4.2.7. Again, most of the hypotheses hold trivially, and we just have to check that (4.2.16) and (4.2.20) holds. We have:

1. Thanks to the continuity of V_t^{i,N_c}

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\leq T\wedge\tau^r_{N_c}} \|X_t^{N,N_c} - X_{t-}^{N,N_c}\|^2\right] &\leq \mathbb{E}\left[\sup_{t\leq T\wedge\tau^r_{N_c}} \sum_{i=1}^N |V_t^{i,N_c} - V_{t-}^{i,N_c}|^2\right] \\
&+ \mathbb{E}\left[\sup_{t\leq T\wedge\tau^r_{N_c}} \sum_{i=1}^N \left(\sum_{x\in\{n,m,y\}} |x_t^{i,N_c} - x_{t-}^{i,N_c}|^2\right)\right] \\
&= \mathbb{E}\left[\sup_{t\leq T\wedge\tau^r_{N_c}} \sum_{i=1}^N \left(\sum_{x\in\{n,m,h,y\}} |x_t^{i,N_c} - x_{t-}^{i,N_c}|^2\right)\right] \\
&\leq \frac{3N}{N_c^2}.
\end{aligned}
$$

Since N remains fix as N_c goes to infinity, (4.2.16) holds.

2. For $x = m, n$ we have that the difference between the quadratic variation of M^{x,i,N_c} and the diffusion term of the generator of X^N valued at X^{N,N_c} is given by

$$
\left|\frac{1}{N_c}\int_0^t (1 - x_s^{i,N_c})\rho_x(V_s^{i,N_c}) + x_s^{i,N_c}\zeta_x(V_s^{i,N_c})ds + \int_0^t \sigma_x^2(V_s^{i,N_c}, x_s^{i,N_c})\chi(x_s^{i,N_c})ds \right.
$$
$$
\left. - \int_0^t \sigma_x^2(V_s^{i,N_c}, x_s^{i,N_c})\chi(x_s^{i,N_c})ds\right| \leq \frac{Ct}{N_c},
$$

and then the left-hand side converges almost surely to 0, from where (4.2.20) holds. The argument for y is similar.

□

4.4 Propagation of Chaos

In the human brain there are from 78.92×10^9 to 95.40×10^9 neurons (See [6]). This huge amount of interacting objects resemble the mechanical statistic context, and nowadays is classical in mathematical neuroscience to study the mean field limits of this interacting network of neurons. There are several works in this direction: [8], [15], [33] and [62], just to mention a few. A review of the mean field approach in neuroscience can be found in [29]. To the best of our knowledge, mean field limits are not been studied for the hybrid model, and in that context our work has some novelty. On the other hand, the existence and convergence to the mean field limit for a different version of our diffusive model, was proposed in [8] and proved rigorously in [15]. In contrast with our models, the authors of [8] and [15] consider:

1. A context of several populations.

2. Random processes as conductance coefficients $\{J_{\mathrm{E}}(t)\}_{t\geq 0}$, $\{J_{\mathrm{Ch}}(t)\}_{t\geq 0}$.
3. A white noise appearing in the dynamic for the potential V.

The first point is not difficult to incorporate in our model, in fact it only makes the computations a little harder to write. The second point is a big difference, and in fact is one of the reasons why in [15] the authors do not obtain an optimal rate of convergence for the propagation of chaos. We could incorporate bounded processes for the conductances, but in a more general case we do not know for sure how much of our proof would have to be changed. The third point is interesting, because on one hand, our results relay heavily on the fact that almost surely the voltage component is uniformly bounded for $t \geq 0$, which is not true if we add a Brownian motion in the dynamic for V. On the other hand, since we are considering the hybrid model as starting point, it is very natural to consider noise only at the channels.

In this section, following [8] and [15], we will only consider chemical synapses. That is, we take $J_{\mathrm{E}} = 0$. We treat in Section 4.4.2 the Hybrid model, and in Section 4.4.3 the diffusive model.

4.4.1 Some preliminaries on Propagation of Chaos.

Let us consider a system of N particles $X^{(1)}, \dots, X^{(N)}$ interacting in a mean field fashion. That is, the aggregate effect of system over a single particle depends only on the empirical measure of the whole particle system. Let μ_t^N be the empirical law of the system at time t, then the dynamic of $X^{(i)}$ is given by

$$X_t^{(i)} = X_0^{(i)} + \int_0^t \sigma[X_s^{(i)}, \mu_s^N] dW_s^{(i)} + \int_0^t b[X_s^{(i)}, \mu_s^N] ds, \tag{4.4.1}$$

where σ and b are suitable functions and $W^{(i)}$ are independent Brownian motions.

Associated to the particle system given by (4.4.1) there exists a nonlinear process in the sense of McKean [59] given as the solution of

$$\overline{X}_t = \overline{X}_0 + \int_0^t \sigma[\overline{X}_s, \mu_s] dW_s^{(i)} + \int_0^t b[\overline{X}_s, \mu_s] ds, \quad \mu_s = \mathrm{law}(X_s). \tag{4.4.2}$$

Under regularity assumptions on b and σ, and exchangeability of the particles in (4.4.1) for i.i.d. initial conditions, it can be shown that for all $T > 0$, there exists $C(T)$ such that

$$N \mathbb{E}\left[\sup_{t \leq T} |X_t^{(i)} - \overline{X}_t^{(i)}|^2\right] \leq C(T). \tag{4.4.3}$$

From this property, it follows that for any $k \in \mathbb{N}$, the law of a set of k particles of the system (4.4.1) converges as the total number of particles goes to infinity to the joint law of k independent copies of the nonlinear process. In particular, each fixed finite subset of k particles is asymptotically independient. This property is called propagation of chaos. For further background see [60] and [74].

A posible interpretation for the propagation of chaos is the following: Given an interacting particle system with N large, the evolution of any particle can be approximated by an average particle given by the nonlinear process. The error of this approximation will be of order $C(T)/N$. Nevertheless, the function $C(T)$ in the right-hand side of (4.4.3) is usually increasing, thus if we consider larger time windows the approximation given by the nonlinear process degrades as T increase.

To have an approximation for the system uniform on time, the particle system has to *propagates chaos uniformly*, that is

$$N \sup_{t\geq 0} \mathbb{E}\left[|X_t^{(i)} - \overline{X}_t^{(i)}|^2 \right] \leq C. \tag{4.4.4}$$

Naturally, the uniform propagation of chaos is much more difficult to prove.

4.4.2 Propagation of Chaos for the Hybrid Model

We are interested in the behavior of the model (4.3.7) when the number of neurons goes to infinity. In what follows, we fix the number of ionic channels N_c, and for the sake of clarity, we drop the superscript N_c in the notation. Let us recall the hybrid model.

$$\begin{aligned}
V_t^{(i)} &= V_0^{(i)} + \int_0^t I - g_{\mathrm{Ca}} m_s^{(i)}(V_s^{(i)} - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_s^{(i)}(V_s^{(i)} - V_{\mathrm{K}}) ds \\
&\quad - \int_0^t g_{\mathrm{L}}(V_s^{(i)} - V_{\mathrm{L}}) ds - \int_0^t J_{\mathrm{Ch}} \bar{y}_s^N (V_s - V_{\mathrm{rev}}) ds, \\
x_t^{(i)} &= x_0^{(i)} + \int_0^{t+} \int_0^\infty \frac{1}{N_c} \mathbb{1}_{\{\lambda \leq \lambda_0^x(V_{s-}^{(i)}, x_{s-}^{(i)})\}} \mathcal{N}_0^{x,i}(d\lambda ds) \\
&\quad - \int_0^{t+} \int_0^\infty \frac{1}{N_c} \mathbb{1}_{\{\lambda \leq \lambda_1^x(V_{s-}^{(i)}, x_{s-}^{(i)})\}} \mathcal{N}_1^{x,i}(d\lambda ds), \quad \text{for } x = m, n, y.
\end{aligned}$$

The nonlinear process

Our goal is to show that when the number of neurons increases to infinity (i.e. $N \to \infty$), any finite set of neurons behave as independent copies of the following nonlinear process.

$$\begin{aligned}
\tilde{V}_t &= \tilde{V}_0 + \int_0^t I - g_{\mathrm{Ca}}\tilde{m}_s(\tilde{V}_s - V_{\mathrm{Ca}}) - g_{\mathrm{K}}\tilde{n}_s(\tilde{V}_s - V_{\mathrm{K}})ds \\
&\quad - \int_0^t g_{\mathrm{L}}(\tilde{V}_s - V_{\mathrm{L}})ds - \int_0^t J_{\mathrm{Ch}}\langle \tilde{\mu}_s^y\rangle_1(\tilde{V}_s^{(i)} - V_{\mathrm{rev}})ds, \\
\tilde{x}_t &= x_0 + \int_0^{t+}\int_0^{\infty} \frac{1}{N_c}\mathbb{1}_{\{\lambda \le \lambda_0^x(\tilde{V}_{s-},\tilde{x}_{s-})\}}\mathcal{N}_0^x(d\lambda ds) \\
&\quad - \int_0^{t+}\int_0^{\infty} \frac{1}{N_c}\mathbb{1}_{\{\lambda \le \lambda_0^x(\tilde{V}_{s-},\tilde{x}_{s-})\}}\mathcal{N}_1^x(d\lambda ds), \quad \text{for } x = m, n, y. \\
\langle \tilde{\mu}_s^y\rangle_1 &= \int_0^{(1)} q\tilde{\mu}_s^y(dq), \ \tilde{\mu}_s^y = \mathrm{law}(\tilde{y}_s),
\end{aligned} \tag{4.4.5}$$

where for simplicity for all $i = 1, \ldots, N$, law$\tilde{V}_0) = \mathrm{law}(V_0^{(i)})$. Consequently,

$$|\tilde{V}_0| \in L^{\infty}(\Omega). \tag{4.4.6}$$

Our program is the following:

1. To prove the well-posedness of equation (4.4.5). To do so, we introduce an auxiliary linear problem, and then we use a fixed point argument.
2. To prove the strong convergence for a particle of the system (4.3.5) to the solution of (4.4.5).

The auxiliary problem

To prove the existence of the nonlinear process, we introduce an auxiliary problem, and follow some ideas from [11]. In what comes next we will consider $T > 0$. Let us denote $\mathcal{R} = \mathbb{R} \times [0,1] \times [0,1] \times [0,1]$, and $\mathcal{P}$ the set of probability measures over $\mathcal{R}$, then

$$\mathcal{P}_1 = \left\{ Q : Q \in \mathcal{P}, \int_{\mathcal{R}} |v| + |u_1| + |u_2| + |u_3| Q(dv, du_1, du_2, du_3) < \infty \right\}.$$

$\mathcal{P}_1$ is a complete and separable metric space when it is endowed with the Wasserstein metric

$$\mathcal{W}(Q^1, Q^2) = \inf_{\pi \in \Pi(Q^1,Q^2)} \int_{\mathcal{R}\times\mathcal{R}} \|x - y\|_1 d\pi(x,y),$$

where $\Pi(Q^1, Q^2)$ is the set of all couplings between Q^1 and Q^2, that is the set of all probability measures π over $\mathcal{R} \times \mathcal{R}$, whose marginals are Q^1 and Q^2. For details on Wasserstein metrics see [70], [78, 79].

We also consider the set

$$\mathcal{C}\left([0,T],\mathcal{P}_1\right)=\left\{\nu:[0,T]\to\mathcal{P}_1,\,\nu\text{ continuous}\right\},$$

endowed with the metric

$$\mathbb{W}(\nu^1,\nu^2)=\sup_{0\le t\le T}\mathcal{W}(\nu^1_t,\nu^2_t).$$

The metric space $\left(\mathcal{C}\left([0,T],\mathcal{P}_1\right),\mathbb{W}\right)$ is complete.

For $\nu\in\mathcal{C}\left([0,T],\mathcal{P}_1\right)$, and $t\in[0,T]$, we will denote by ν^y_t the marginal law of ν_t with respect to its fourth coordinate, that is, for any $A\subset[0,1]$ measurable,

$$\nu^y_t(A)=\int_{\mathbb{R}}\int_0^1\int_0^1\int_0^1 \mathbb{1}_A(y)\nu_t(dv,dm,dn,dy),$$

and we will denote by $\langle\nu^y_t\rangle_1$ the first moment of ν^y_t, that is

$$\langle\nu^y_t\rangle_1=\int_0^1 y\nu^y_t(dy).$$

Notice that for any $\nu\in\mathcal{C}\left([0,T],\mathcal{P}_1\right)$ the function $g:[0,T]\to\mathbb{R}$ given by $g(t)=\langle\nu^y_t\rangle_1$ is continuous. Indeed, the continuity of g is a direct consequence of the bound

$$|\langle\nu^y_t\rangle_1-\langle\nu^y_s\rangle_1|\le\mathcal{W}(\nu_t,\nu_s),$$

since for any coupling between π_{st} of ν_t and ν_s we have that

$$|\langle\nu^y_t\rangle_1-\langle\nu^y_s\rangle_1|\le\int_{\mathcal{R}\times\mathcal{R}}\|x-y\|d\pi_{st}(x,y).$$

Let $\nu\in\mathcal{C}\left([0,T],\mathcal{P}_1\right)$, and consider the equation

$$\begin{aligned}\tilde{V}_t&=\tilde{V}_0+\int_0^t I-g_{\mathrm{Ca}}\tilde{m}_s(\tilde{V}_s-V_{\mathrm{Ca}})-g_{\mathrm{K}}\tilde{n}_s(\tilde{V}_s-V_{\mathrm{K}})ds\\&\quad-\int_0^t g_{\mathrm{L}}(\tilde{V}_s-V_{\mathrm{L}})ds-\int_0^t J_{\mathrm{Ch}}\langle\nu^y_s\rangle_1(\tilde{V}^{(i)}_s-V_{\mathrm{rev}})ds,\\\tilde{x}_t&=x_0+\int_0^{t+}\int_0^\infty\frac{1}{N_c}\mathbb{1}_{\{\lambda\le\lambda^x_0(\tilde{V}_{s-},\tilde{x}_{s-})\}}\mathcal{N}^x_0(d\lambda ds)\\&\quad-\int_0^{t+}\int_0^\infty\frac{1}{N_c}\mathbb{1}_{\{\lambda\le\lambda^x_1(\tilde{V}_{s-},\tilde{x}_{s-})\}}\mathcal{N}^x_1(d\lambda ds),\quad\text{for: }x=m,n,y.\end{aligned}\tag{4.4.7}$$

Proposition 4.4.1. *Assume equation (4.4.7) has a solution $(\tilde{V}_t,\tilde{m}_t,\tilde{n}_t,\tilde{y}_t)_{0\le t\le\infty}$. Then, there exists a constant V_+ depending on V_0 and the parameters of the system, such that*

$$\sup_{0\le t\le\infty}|\tilde{V}_t|\le V_+\ a.s.\tag{4.4.8}$$

Proof. The key point is to notice that the equation for $\tilde{V}$ can be written as

$$\tilde{V}_t = \tilde{V}_0 + \int_0^t R_s - A_s \tilde{V}_s ds,$$

where

$$R_s = I + g_{\mathrm{L}} V_{\mathrm{L}} + g_{\mathrm{Ca}} V_{\mathrm{Ca}} \tilde{m}_s + g_{\mathrm{K}} V_{\mathrm{K}} \tilde{n}_s + J_{\mathrm{Ch}} V_{\mathrm{rev}} \langle \nu_s^y \rangle_1,$$

and

$$A_s = g_{\mathrm{L}} + g_{\mathrm{Ca}} \tilde{m}_s + g_{\mathrm{K}} \tilde{n}_s + J_{\mathrm{Ch}} \langle \nu_s^y \rangle_1.$$

Then

$$\tilde{V}_t = V_0 \exp\left(-\int_0^t A_s ds\right) + \int_0^t R_s \exp\left(-\int_s^t A_u ds\right) ds. \tag{4.4.9}$$

Let

$$R_+ = \max_{m,n,\nu \in [0,1]} I + g_{\mathrm{L}} V_{\mathrm{L}} + g_{\mathrm{Ca}} V_{\mathrm{Ca}} m + g_{\mathrm{K}} V_{\mathrm{K}} n + J_{\mathrm{Ch}} V_{\mathrm{rev}} \nu,$$

then $R_s \le R_+$. Since $A_s \ge g_{\mathrm{L}}$,

$$\tilde{V}_t \le |V_0| e^{-g_{\mathrm{L}} t} + \int_0^t R_+ e^{-g_{\mathrm{L}}(t-s)} ds = \left(|V_0| - \frac{R_+}{g_{\mathrm{L}}}\right) e^{-g_{\mathrm{L}} t} + \frac{R_+}{g_{\mathrm{L}}}.$$

Then the bound (4.4.8) follows taking

$$V_+ = \max\left\{|V_0|, \frac{R_+}{g_{\mathrm{L}}}\right\}.$$

□

Lemma 4.4.2. *The linear auxiliary problem (4.4.7) has a unique strong solution.*

The proof of this Lemma is the very similar to the proof of Lemma 4.3.4, so we omit it.

Proposition 4.4.3. *Let $(\tilde{V}_t, \tilde{m}_t, \tilde{n}_t, \tilde{y}_t)_{t\in[0,T]}$ be the solution of (4.4.7). Then the function $t \to law(\tilde{V}_t, \tilde{m}_t, \tilde{n}_t, \tilde{y}_t)$ is a continuous function from $[0,T]$ to $\mathcal{P}_1$. That is*

$$law(\tilde{V}_\cdot, \tilde{m}_\cdot, \tilde{n}_\cdot, , \tilde{y}_\cdot) \in \mathcal{C}\left([0,T], \mathcal{P}_1\right).$$

Proof. Notice that

$$\mathbb{E}\Big(|\tilde{V}_t - \tilde{V}_u|\Big) \le$$
$$\mathbb{E}\int_u^t |I| + g_{\mathrm{Ca}}|\tilde{V}_s - V_{\mathrm{Ca}}| + g_{\mathrm{K}}|\tilde{V}_s - V_{\mathrm{K}}| + g_{\mathrm{L}}|\tilde{V}_s - V_{\mathrm{L}})ds| + J_{\mathrm{Ch}}|\tilde{V}_s^{(i)} - V_{\mathrm{rev}}|ds.$$

On the other hand, for $x = m, n, y$

$$\begin{aligned}\mathbb{E}\left(|\tilde{x}_t - \tilde{x}_u|\right) &\le \mathbb{E}\int_u^t \int_0^\infty \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_0^x(\tilde{V}_s, \tilde{x}_s)\}} d\lambda ds + \mathbb{E}\int_u^t \int_0^\infty \frac{1}{N_c} \mathbb{1}_{\{\lambda \le \lambda_1^x(\tilde{V}_s, \tilde{x}_s)\}} d\lambda ds \\ &= \frac{1}{N_c} \mathbb{E}\int_u^t \lambda_0^x(\tilde{V}_s, \tilde{x}_s) + \lambda_1^x(\tilde{V}_s, \tilde{x}_s) ds\end{aligned}$$

From this last inequalities and the boundedness of the integrands, is easy to see

$$\mathbb{E}\left(|\tilde{V}_t - \tilde{V}_u|\right) + \mathbb{E}\left(|\tilde{m}_t - \tilde{m}_u|\right) + \mathbb{E}\left(|\tilde{n}_t - \tilde{n}_u|\right) + \mathbb{E}\left(|\tilde{y}_t - \tilde{y}_u|\right) \leq C|t-u|,$$

from where

$$\mathcal{W}_1(\text{law}(\tilde{V}_t, \tilde{m}_t, \tilde{n}_t, \tilde{y}_t), \text{law}(\tilde{V}_u, \tilde{m}_u, \tilde{n}_u, \tilde{y}_u)) \leq C|t-u|.$$

Thus, the function $t \to \text{law}(\tilde{V}_t, \tilde{m}_t, , \tilde{n}_t, \tilde{y}_t)$ is not just a continuous function from $[0,T]$ to $\mathcal{P}_1$, it is indeed locally Lipschitz. □

Let us consider now the operator

$$\begin{aligned}\Phi : \mathcal{C}\left([0,T], \mathcal{P}_1\right) &\to \mathcal{C}\left([0,T], \mathcal{P}_1\right)\\ \nu &\to \text{law}(\tilde{V}_\cdot, \tilde{m}_\cdot, \tilde{n}_\cdot, \tilde{y}_\cdot),\end{aligned}$$

where $(\tilde{V}_t, \tilde{m}_t, \tilde{n}_t, \tilde{y}_t)_{t\in[0,T]}$ is the solution of (4.4.7). Notice that thanks to the last Lemma Φ is well defined.

Remark 4.4.4. The same computation done to prove (4.4.8), shows that for all $t \geq 0$, and for all $\nu \in \mathcal{C}\left([0,T], \mathcal{P}_1\right)$,

$$\text{supp}(\Phi(\nu)_t) \subset [-V_+, V_+] \times [0,1] \times [0,1] \times [0,1].$$

Lemma 4.4.5. *The operator Φ has a unique fixed point, or equivalently the nonlinear equation (4.4.5) has a unique solution.*

Proof. Let $\nu^1, \nu^2 \in \mathcal{C}\left([0,T], \mathcal{P}_1\right)$ and $\tilde{X}^{(i)}_t = (\tilde{V}^{(i)}, \tilde{m}^{(i)}, \tilde{n}^{(i)}, \tilde{y}^{(i)})$, $i = 1,2$, the solutions of (4.4.5) for ν^1 and ν^2 respectively. Then

$$\begin{aligned}\left|\tilde{V}^{(1)}_t - \tilde{V}^{(2)}_t\right| \leq &\int_0^t \left|g_{\text{Ca}}\tilde{m}^{(2)}_s + g_{\text{K}}\tilde{n}^{(2)}_s + g_{\text{L}} + J_{\text{Ch}}\langle\nu^{1,y}_s\rangle_1\right| \left|\tilde{V}^{(1)}_s - \tilde{V}^{(2)}_s\right| ds\\ &+ \int_0^t g_{\text{Ca}}\left|\tilde{V}^{(1)}_s - V_{\text{Ca}}\right| \left|\tilde{m}^{(1)}_s - \tilde{m}^{(2)}_s\right| + g_{\text{K}}\left|\tilde{V}^{(1)}_s - V_{\text{K}}\right| \left|\tilde{n}^{(1)}_s - \tilde{n}^{(2)}_s\right| ds\\ &+ \int_0^t J_{\text{Ch}}\left|\left(\tilde{V}^{(2)}_s - V_{\text{rev}}\right)\right| \left|\langle\nu^{2,y}_s\rangle_1 - \langle\nu^{2,y}_s\rangle_1\right| ds.\end{aligned}$$

If we call

$$K_V = \max\left\{g_{\text{Ca}}\left[|V_{\text{Ca}}| + V_+\right], g_{\text{K}}\left[|V_{\text{K}}| + V_+\right], \left[g_{\text{L}} + g_{\text{Ca}} + g_{\text{K}} + J_{\text{Ch}}\right]\right\},$$

and

$$\Delta_s(\nu^1, \nu^2) := \left|\langle\nu^{1,y}_s\rangle_1 - \langle\nu^{2,y}_s\rangle_1\right|,$$

we have

$$\left|\tilde{V}^{(1)}_t - \tilde{V}^{(2)}_t\right| \leq K_V \int_0^t \|\tilde{X}^{(1)}_s - \tilde{X}^{(2)}_s\|_1 ds + \int_0^t J_{\text{Ch}}\left(V_+ + |V_{\text{rev}}|\right)\Delta_s(\nu^1, \nu^2) ds.$$

Proceeding in the same way as before, we can prove that for $x = m, n, y$

$$\mathbb{E}\left|\tilde{x}_t^{(1)} - \tilde{x}_t^{(2)}\right| \leq K_x \int_0^t \left(\mathbb{E}\left|\tilde{x}_s^{(1)} - \tilde{x}_s^{(2)}\right| + \mathbb{E}\left|\tilde{V}_s^{(1)} - \tilde{V}_s^{(2)}\right|\right) ds.$$

This yields

$$\begin{aligned}\mathbb{E}\left(\|\tilde{X}_t^{(1)} - \tilde{X}_t^{(2)}\|_1\right) \leq (K_V + K_m + K_n + K_y) \int_0^t \mathbb{E}\left(\|\tilde{X}_s^{(1)} - \tilde{X}_s^{(2)}\|_1\right) ds \\ + \int_0^t J_{\mathrm{Ch}}(V_+ + |V_{\mathrm{rev}}|)\Delta_s(\nu^1, \nu^2) ds,\end{aligned}$$

and by Growall's Lemma it yields,

$$\mathbb{E}\left(\|\tilde{X}_t^{(1)} - \tilde{X}_t^{(2)}\|_1\right) \leq \exp(Kt) \int_0^t J_{\mathrm{Ch}}(V_+ + |V_{\mathrm{rev}}|)\Delta_s(\nu^1, \nu^2) ds, \tag{4.4.10}$$

where $K = K_V + K_m + K_n + K_y$. But if π_s is any coupling between ν_s^1 and ν_s^2, we have

$$\begin{aligned}\Delta_s(\nu^1, \nu^2) &= \left|\int_0^1 x_3\, \nu_s^{1,y}(dx_3) - \int_0^1 x_3' \nu_s^{2,y}(dx_3')\right| \\ &\leq \left|\int_{\mathbb{R}} v\, \nu_s^{1,V}(dv) - \int_{\mathbb{R}} v' \nu_s^{2,V}(dv')\right| + \left|\int_0^1 x_1\, \nu_s^{1,m}(dx_1) - \int_0^1 x_1' \nu_s^{2,m}(dx_1')\right| \\ &\quad + \left|\int_0^1 x_2\, \nu_s^{1,n}(dx_2) - \int_0^1 x_2' \nu_s^{2,n}(dx_2')\right| + \left|\int_0^1 x_3\, \nu_s^{1,y}(dx_3) - \int_0^1 x_3' \nu_s^{2,y}(dx_3')\right| \\ &= \left|\int_{\mathbb{R}^2} v - v'\, \pi_s^V(dv, dv')\right| + \left|\int_{[0,1]^2} x_1 - x_1'\, \pi_s^m(dx_1, dx_1')\right| \\ &\quad + \left|\int_{[0,1]^2} x_2 - x_2'\, \pi_s^n(dx_2, dx_2')\right| + \left|\int_{[0,1]^2} x_3 - x_3'\, \pi_s^y(dx_3, dx_3')\right| \\ &\leq \int_{(\mathbb{R}\times[0,1]^3)^2} \|X - X'\|_1 \pi_s(dv dx_1 dx_2 dx_3, dv' dx_1' dx_2' dx_3'),\end{aligned}$$

where $\|X - X'\|_1 = |v - v'| + |x_1 - x_1'| + |x_2 - x_2'| + |x_3 - x_3'|$. Since the last inequality holds for any coupling between v_s^1 and v_s^1, we have

$$\Delta_s(\nu^1, \nu^2) \leq \mathcal{W}(\nu_s^1, \nu_s^2).$$

On the other hand

$$\mathcal{W}(\Phi(\nu^1)_t, \Phi(\nu^2)_t) \leq \mathbb{E}\left(\|\tilde{X}_t^{(1)} - \tilde{X}_t^{(2)}\|_1\right).$$

Putting these two bounds in (4.4.10), we have

$$\mathcal{W}(\Phi(\nu^1)_t, \Phi(\nu^2)_t) \leq \exp(Kt) \int_0^t J_{\mathrm{Ch}}(V_+ + |V_{\mathrm{rev}}|)\mathcal{W}(\nu_s^1, \nu_s^2) ds,$$

so if we call $\kappa_T = J_{\mathrm{Ch}}(V_+ + |V_{\mathrm{rev}}|)\exp(KT)$, we deduce that

$$\sup_{s\leq t} \mathcal{W}(\Phi(\nu^1)_s, \Phi(\nu^2)_s) \leq \kappa_T \int_0^t \sup_{u\leq s} \mathcal{W}(\nu_u^1, \nu_u^2) ds. \tag{4.4.11}$$

Let us consider now the sequence of probability flows $\nu^{n+1} = \Phi(\nu^n)$, with $\nu^0 = \nu \in \mathcal{P}_1$. Iterating the last inequality we conclude that

$$\sup_{t\leq T} \mathcal{W}(\nu_t^{n+1}, \nu_t^n) \leq \frac{(\kappa_T T)^n}{n!} \sup_{t\leq T} \mathcal{W}(\nu_t^1, \nu_t^0).$$

Therefore, if $m \geq n$ we have

$$\sup_{t\leq T} \mathcal{W}(\nu_t^m, \nu_t^n) \leq \sup_{t\leq T} \mathcal{W}(\nu_t^1, \nu_t^0) \sum_{k=n}^{m-1} \frac{(\kappa_T T)^k}{k!},$$

but the right-hand side is the tail of a convergent serie, and so goes to zero when n goes to infinity. Hence $\{\nu^n\}_{n\geq 0}$ is a Cauchy sequence in $\mathcal{C}([0,T], \mathcal{P}_1)$, and since this is a complete metric space, we conclude that the sequence converges. Is easy to show that the limit of the sequence is the unique fixed point of Φ. □

Remark 4.4.6. Thanks to Lemma 4.4.5 we have that the nonlinear equation (4.4.5) is well posed in $[0,T]$ for every $T > 0$. Since the restriction of the solution in the interval $[0,T']$ to the interval $[0,T]$ with $T < T'$ will be a solution, because of the uniqueness of solutions in $[0,T]$, we have that the solution for (4.4.5) can be extended to $[0,\infty)$.

Propagation of Chaos

Lemma 4.4.7. *Let $T > 0$ and $\tilde{X}_t^{(i)} = (\tilde{V}_t^{(i)}, \tilde{m}_t^{(i)}, \tilde{n}_t^{(i)}, \tilde{y}_t^{(i)})$ a solution of the nonlinear equation* (4.4.5) *and $X_t^{(i)} = (V_t^{(i)}, m_t^{(i)}, n_t^{(i)}, y_t^{(i)})$ a solution for the N neurons system* (4.3.5)*, then*

$$\sqrt{N} \sup_{0\leq t\leq T} \mathbb{E}\left(\left\|\tilde{X}_t - X_t\right\|_1\right) \leq T J_{Ch}(V_+ + |V_{rev}|)e^{\tilde{K}T}.$$

Proof. Proceeding as before we can show for $x = m, n, y$

$$\mathbb{E}\left(|\tilde{x}_t^{(i)} - x_t^{(i)}|\right) \leq K_x \int_0^t \left[\mathbb{E}\left(|\tilde{x}_t^{(i)} - x_t^{(i)}|\right) + \mathbb{E}\left(|\tilde{V}_t^{(i)} - V_t^{(i)}|\right)\right] ds,$$

and for V

$$\begin{aligned} \left|\tilde{V}_t^{(1)} - \tilde{V}_t^{(2)}\right| &\leq K_V \int_0^t \|\tilde{X}_s^{(1)} - \tilde{X}_s^{(2)}\|_1 ds \\ &+ \int_0^t J_{\mathrm{Ch}} \mathbb{E}\left(\left|\left(\tilde{V}_s^{(i)} - V_{\mathrm{rev}}\right)\langle\tilde{\mu}_s^{i,y}\rangle_1 - \left(V_s^{(i)} - V_{\mathrm{rev}}\right)\bar{y}_s^N\right|\right) ds, \end{aligned} \tag{4.4.12}$$

where $\tilde{\mu}_s^{(i)} = \mathrm{law}(\tilde{y}_t^{(i)})$, $\langle\tilde{\mu}_s^{i,y}\rangle_1 = \int_0^{(1)} y\tilde{\mu}_s^{(i)}(dy)$, and $\bar{y}_s^N$ is the empirical mean of the variables $y_s^{(i)}$.

To bound the second term in (4.4.12) we introduce the empirical measure of the fourth component of the nonlinear process $\tilde{y}^N_s = \sum_{i=1}^N \tilde{y}^{(i)}_s/N$. Then

$$\begin{aligned}\mathbb{E}\left(\left|\left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\langle\tilde{\mu}^{i,y}_s\rangle_1 - \left(V^{(i)}_s - V_{\text{rev}}\right)\bar{y}^N_s\right|\right)& \\ \leq \mathbb{E}\left(\left|\left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\langle\tilde{\mu}^{i,y}_s\rangle_1 - \left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\tilde{y}^N_s\right|\right)& \\ + \mathbb{E}\left(\left|\left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\tilde{y}^N_s - \left(V^{(i)}_s - V_{\text{rev}}\right)\bar{y}^N_s\right|\right).&\end{aligned} \tag{4.4.13}$$

Since $\tilde{V}^{(i)}_t$ is uniformly bound on $[0,T]$ we have

$$\begin{aligned}\mathbb{E}\Big(\Big|&\left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\langle\tilde{\mu}^{i,y}_s\rangle_1 - \left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\tilde{y}^N_s\Big|\Big) \\ &\leq (V_+ + |V_{\text{rev}}|)\,\mathbb{E}\left(\left|\langle\tilde{\mu}^{j,y}_s\rangle_1 - \tilde{y}^N_s\right|\right) \\ &= (V_+ + |V_{\text{rev}}|)\,\mathbb{E}\left(\left|\frac{1}{N}\sum_{j=1}^N \langle\tilde{\mu}^{j,y}_s\rangle_1 - \tilde{y}^{(j)}_s\right|\right) \\ &\leq \frac{(V_+ + |V_{\text{rev}}|)}{N}\sqrt{\mathbb{E}\left(\left[\sum_{j=1}^N \langle\tilde{\mu}^{j,y}_s\rangle_1 - \tilde{y}^{(j)}_s\right]^2\right)} \\ &= \frac{(V_+ + |V_{\text{rev}}|)}{N}\sqrt{\mathbb{E}\left(\sum_{j=1}^N\sum_{k=1}^N\left(\langle\tilde{\mu}^{j,y}_s\rangle_1 - \tilde{y}^{(j)}_s\right)\left(\langle\tilde{\mu}^{k,y}_s\rangle_1 - \tilde{y}^{(k)}_s\right)\right)},\end{aligned}$$

but the variables $\tilde{y}^{(j)}$ are pairwise independent and identically distributed, so

$$\mathbb{E}\left(\sum_{j=1}^N\sum_{k=1}^N\left(\langle\tilde{\mu}^{j,y}_s\rangle_1 - \tilde{y}^{(j)}_s\right)\left(\langle\tilde{\mu}^{i,y}_s\rangle_1 - \tilde{y}^{(k)}_s\right)\right) = N\mathbb{E}\left(\left(\langle\tilde{\mu}^{1,y}_s\rangle_1 - \tilde{y}^{(1)}_s\right)^2\right) \leq N.$$

So,

$$\mathbb{E}\left(\left|\left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\langle\tilde{\mu}^{i,y}_s\rangle_1 - \left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\tilde{y}^N_s\right|\right) \leq \frac{V_+ + |V_{\text{rev}}|}{\sqrt{N}}.$$

For the second term in (4.4.13) we have

$$\begin{aligned}\mathbb{E}\Big(\Big|&\left(\tilde{V}^{(i)}_s - V_{\text{rev}}\right)\tilde{y}^N_s - \left(V^{(i)}_s - V_{\text{rev}}\right)\bar{y}^N_s\Big|\Big) \\ &= \mathbb{E}\left(\left|\left(\tilde{V}^{(i)}_s - V^{(i)}_s\right)\tilde{y}^N_s - \left(V^{(i)}_s - V_{\text{rev}}\right)\left(\tilde{y}^N_s - \bar{y}^N_s\right)\right|\right) \\ &\leq \mathbb{E}\left(\left|\left(\tilde{V}^{(i)}_s - V^{(i)}_s\right)\right|\right) + (V_+ + |V_{\text{rev}}|)\,\mathbb{E}\left(\left|\tilde{y}^N_s - \bar{y}^N_s\right|\right),\end{aligned}$$

because $V^{(i)}_s$ are uniformly bounded on $[0,T]$. By symmetry of the N neurons system and since the nonlinear processes are i.i.d, we have

$$\mathbb{E}\left(\left|\tilde{y}^N_s - \bar{y}^N_s\right|\right) \leq \frac{1}{N}\sum_{j=1}^N \mathbb{E}\left(\left|\tilde{y}^{(j)}_s - y^{(j)}_s\right|\right) = \mathbb{E}\left(\left|\tilde{y}^{(i)}_s - y^{(i)}_s\right|\right).$$

Putting these last computations in (4.4.13), we have

$$\mathbb{E}\Big(\Big|\left(\tilde{V}_s^{(i)}-V_{\text{rev}}\right)\langle\tilde{\mu}_s^{i,y}\rangle_1-\left(V_s^{(i)}-V_{\text{rev}}\right)\bar{y}_s^N\Big|\Big)$$
$$\leq \mathbb{E}\left(\left|\left(\tilde{V}_s^{(i)}-V_s^{(i)}\right)\right|\right)+(V_++|V_{\text{rev}}|)\,\mathbb{E}\left(\left|\tilde{y}_s^{(i)}-y_s^{(i)}\right|\right)+\frac{V_++|V_{\text{rev}}|}{\sqrt{N}}.$$

Putting these computations back on (4.4.12), we find

$$\left|\tilde{V}_t^{(1)}-\tilde{V}_t^{(2)}\right|\leq \tilde{K}_V\int_0^t\|\tilde{X}_s^{(1)}-\tilde{X}_s^{(2)}\|_1 ds+\frac{J_{\text{Ch}}(V_++|V_{\text{rev}}|)}{\sqrt{N}}.,$$

and then summarizing all the previous computations we have

$$\mathbb{E}\left(\left\|\tilde{X}_t^{(i)}-X_t^{(i)}\right\|_1\right)\leq \tilde{K}\int_0^t\mathbb{E}\left(\left\|\tilde{X}_s^{(i)}-X_s^{(i)}\right\|_1\right)ds+\frac{tJ_{\text{Ch}}(V_++|V_{\text{rev}}|)}{\sqrt{N}},$$

and then thanks to Gronwall's inequality we can conclude

$$\sqrt{N}\sup_{0\leq t\leq T}\mathbb{E}\left(\left\|\tilde{X}_t^{(i)}-X_t^{(i)}\right\|_1\right)\leq TJ_{\text{Ch}}(V_++|V_{\text{rev}}|)e^{\tilde{K}T},$$

in other words for a fixed number of channels N_c propagation of chaos holds. □

Remark 4.4.8. The result of propagation of chaos can be easily extended to a model with more than one population and with J_{Ch}, the maximal conductance coefficient, being a stochastic process uniformly bounded for $t\geq 0$.

Remark 4.4.9. The constant $\tilde{K}=\tilde{K}_V+K_m+K_n+K_y$, and the constant K_x depend on the functions ρ_x,ζ_x, and are independent of N_c in the non perturbed case, and linear on N_c in the perturbed one.

Remark 4.4.10. It is very natural to try to obtain the bound of the propagation of chaos with the supremum inside of the expectation, that is,

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left\|\tilde{X}_t^{(i)}-X_t^{(i)}\right\|_1\right]\leq\frac{C}{\sqrt{N}}.$$

But in this case, this does not seem easy to do. If we apply the Burkholder Davis Gundy inequality to the martingale part of the hybrid model, it will appear the square root of the quadratic variation which will not allows to conclude with Gronwall's inequality. Alternatively, someone could suggest to consider $\|\cdot\|_2$ instead of $\|\cdot\|_1$, but we already explain before that the indicators inside of the integrals with respect to the Poisson measures, does not get along with the $\|\cdot\|_2$.

4.4.3 Propagation of Chaos for the Diffusive Model

In this section we are going to prove that the diffusive model for the interacting particle system (4.3.9) converges, as the number of neurons goes to infinity, to the solution of the nonlinear equation

$$
\begin{aligned}
\tilde{V}_t &= \tilde{V}_0 + \int_0^t I - g_{\mathrm{Ca}}\tilde{m}_s(\tilde{V}_s - V_{\mathrm{Ca}}) - g_{\mathrm{K}}\tilde{n}_s(\tilde{V}_s - V_{\mathrm{K}})ds \\
&\quad - \int_0^t g_{\mathrm{L}}(\tilde{V}_s - V_{\mathrm{L}})ds - \int_0^t J_{\mathrm{Ch}}\langle \tilde{\mu}_s^y\rangle_1(\tilde{V}_s - V_{\mathrm{rev}})ds \\
\tilde{x}_t &= \tilde{x}_0 + \int_0^t (1-\tilde{x}_s)\rho_x(\tilde{V}_s) - \tilde{x}_s\zeta_x(\tilde{V}_s)ds \\
&\quad + \int_0^t \sigma_x(\tilde{V}_s, \tilde{x}_s)\chi(\tilde{x}_s)dW_s^x, \quad x = m, n, y \\
\tilde{\mu}_s^y &= \mathrm{law}(\tilde{y}_s).
\end{aligned}
\tag{4.4.14}
$$

This equation has been studied in [15], where have been proved the well posedness of the equation and a general propagation of chaos result, although with a non optimal rate of convergence. In our particular case, we obtain a propagation of chaos result with optimal rate of convergence $\sqrt{N}$.

Proposition 4.4.11. *Let $T > 0$ and $\tilde{X}_t^{(i)} = (\tilde{V}_t^{(i)}, \tilde{m}_t^{(i)}, \tilde{n}_t^{(i)}, \tilde{y}_t^{(i)})$ a solution of the nonlinear equation (4.4.14) and $X_t^{(i)} = (V_t^{(i)}, m_t^{(i)}, n_t^{(i)}, y_t^{(i)})$ a solution for the N neurons system (4.3.9). Assume that*

$$
\min\left\{\inf_{v\in\mathbb{R}} \rho_x(V),\ \inf_{v\in\mathbb{R}} \zeta_x(V)\right\} > 0. \tag{4.4.15}
$$

Then we have

$$
N\mathbb{E}\left(\sup_{0\le t\le T}\left\|\tilde{X}_t^{(i)} - X_t^{(i)}\right\|^2\right) \le K(1+T)e^{K_2 T}.
$$

Remark 4.4.12. Assumption (4.4.15) is also made in [15]. Notice that in our case, since ρ_x and ζ_x are continuous and the V components of the systems involve are confined in a compact interval K, we can always replace ρ_x and ζ_x by a some functions ρ'_x and ζ'_x, such that ρ_x and ρ'_x agree on K, ζ_x and ζ'_x too, but ρ'_x and ζ'_x satisfy (4.4.15).

Proof. Let $\Delta V_t^{(i)} = \tilde{V}_t^{(i)} - V_t^{(i)}$, and for $x = m, n, y$, $\Delta x_t^{(i)} = \tilde{x}_t^{(i)} - x_t^{(i)}$. Just as before, under the hypothesis of $\tilde{V}_0$ being bounded, we have $|\tilde{V}_t| \le V_+$ *a.s.* Then

$$
\begin{aligned}
(\Delta V_t^{(i)})^2 &= -2\int_0^t g_{\mathrm{Ca}}(V_{\mathrm{Ca}} - \tilde{V}_s^{(i)})\Delta m_s^{(i)}\Delta V_s^{(i)}ds - 2\int_0^t g_{\mathrm{K}}(V_{\mathrm{K}} - \tilde{V}_s^{(i)})\Delta n_s^{(i)}\Delta V_s^{(i)}ds \\
&\quad - 2\int_0^t J_{\mathrm{Ch}}(V_{\mathrm{rev}} - \tilde{V}_s^{(i)})(\langle\tilde{\mu}_s^y\rangle_1 - \bar{y}_s^N)\Delta V_s^{(i)}ds \\
&\quad - 2\int_0^t \left(g_{\mathrm{L}} + g_{\mathrm{Ca}}m_s^{(i)} + g_{\mathrm{K}}n_s^{(i)} + J_{\mathrm{Ch}}\bar{y}_s^N\right)\left(\Delta V_s^{(i)}\right)^2 ds \\
&\le K_1\int_0^t \sup_{u\le s}\{\left(\Delta V_u^{(i)}\right)^2 + \left(\Delta m_u^{(i)}\right)^2 + \left(\Delta n_u^{(i)}\right)^2\}ds \\
&\quad + \int_0^t (\langle\tilde{\mu}_s^y\rangle_1 - \bar{y}_s^N)^2 ds.
\end{aligned}
$$

Since the right side is increasing on t, we can take supremum and then expectation to get

$$\mathbb{E}\left[\sup_{s\le t}(\Delta V_t^{(i)})^2\right] \le K_1 \int_0^t \mathbb{E}\left[\sup_{u\le s}\{\left(\Delta V_u^{(i)}\right)^2 + \left(\Delta m_u^{(i)}\right)^2 + \left(\Delta n_u^{(i)}\right)^2\}\right] ds + \int_0^t \mathbb{E}\left[(\langle\tilde{\mu}_s^y\rangle_1 - \bar{y}_s^N)^2\right] ds.$$

The same computation we did in the discrete case, shows that

$$\mathbb{E}\left[(\langle\tilde{\mu}_s^y\rangle_1 - \bar{y}_s^N)^2\right] \le \mathbb{E}\left[(\Delta y_s^{(i)})^2\right] + \frac{K_2}{N},$$

from where

$$\mathbb{E}\left[\sup_{s\le t}(\Delta V_t^{(i)})^2\right] \le K_1 \int_0^t \mathbb{E}\left[\sup_{u\le s}\left\|\tilde{X}_s^{(i)} - X_s^{(i)}\right\|^2\right] ds + \frac{K_2 t}{N}.$$

On the other hand for $x = m, n, y$, we have

$$(\Delta x_t^{(i)})^2 = \Big(\int_0^t (1 - x_t^{(i)})(\rho_x(\tilde{V}_t^{(i)}) - \rho_x(V_t^{(i)}))ds + \int_0^t x_t^{(i)}(\zeta_x(\tilde{V}_t^{(i)}) - \zeta_x(V_t^{(i)}))ds - \int_0^t \left(\rho_x(\tilde{V}_s^{(j)}) + \zeta_x(\tilde{V}_s^{(j)})\right)\Delta x_s^{(i)} ds + \int_0^t \sigma_x(\tilde{V}_s^{(i)}, \tilde{x}_s^{(i)}) - \sigma_x(V_s^{(i)}, x_s^{(i)}) dW_s^{(i)}\Big)^2.$$

Since $(a+b+c+d)^2 \le 4(a^2+b^2+c^2+d^2)$, and thanks to Jensen's inequality

$$\left(\int_0^t f(s)ds\right)^2 \le t\int_0^t f(s)^2 ds,$$

we have

$$\begin{aligned}(\Delta x_t^{(i)})^2 \le{}& 4t\int_0^t (\rho_x(\tilde{V}_t^{(i)}) - \rho_x(V_t^{(i)}))^2 ds + 4t\int_0^t (\zeta_x(\tilde{V}_t^{(i)}) - \zeta_x(V_t^{(i)}))^2 ds \\ &+ 4t\int_0^t \left(\rho_x(\tilde{V}_s^{(j)}) + \zeta_x(\tilde{V}_s^{(j)})\right)^2 (\Delta x_s^{(i)})^2 ds \\ &+ 4\left(\int_0^t \sigma_x(\tilde{V}_s^{(i)}, \tilde{x}_s^{(i)}) - \sigma_x(V_s^{(i)}, x_s^{(i)}) dW_s^{(i)}\right)^2.\end{aligned}$$

The functions ρ_x and ζ_x are bounded and Lipschitz, then

$$\begin{aligned}(\Delta x_t^{(i)})^2 \le{}& K_3 t\int_0^t (\Delta V_s^{(i)})^2 + (\Delta x_s^{(i)})^2 ds \\ &+ 4\left(\int_0^t \sigma_x(\tilde{V}_s^{(i)}, \tilde{x}_s^{(i)}) - \sigma_x(V_s^{(i)}, x_s^{(i)}) dW_s^{(i)}\right)^2,\end{aligned}$$

from where

$$\begin{aligned}\mathbb{E}\left[\sup_{s\le t}(\Delta x_s^{(i)})^2\right] \le{}& K_3 T\int_0^t \mathbb{E}\left[\sup_{u\le s}\{(\Delta V_u^{(i)})^2 + (\Delta x_u^{(i)})^2\}\right] ds \\ &+ 4\mathbb{E}\left[\sup_{s\le t}\left(\int_0^s \left(\sigma_x(\tilde{V}_s^{(i)}, \tilde{x}_s^{(i)}) - \sigma_x(V_s^{(i)}, x_s^{(i)})\right) dW_s^{(i)}\right)^2\right].\end{aligned}$$

From the Burkholder-Davis-Gundy inequality, and the Lipschitz property of σ_x, it follows

$$\mathbb{E}\Big[\sup_{s\le t}\Big(\int_0^s\Big(\sigma_x(\tilde V_s^{(i)},\tilde x_s^{(i)})-\sigma_x(V_s^{(i)},x_s^{(i)})\Big)dW_s^{(i)}\Big)^2\Big]$$
$$\le K_4\mathbb{E}\left[\int_0^t\Big(\sigma_x(\tilde V_s^{(i)},\tilde x_s^{(i)})-\sigma_x(V_s^{(i)},x_s^{(i)})\Big)^2 ds\right].$$

Under assumption (4.4.15) it is straightforward to prove that the function σ_x is Lipschitz. Thus we have

$$\mathbb{E}\Big[\sup_{s\le t}\Big(\int_0^s\Big(\sigma_x(\tilde V_s^{(i)},\tilde x_s^{(i)})-\sigma_x(V_s^{(i)},x_s^{(i)})\Big)dW_s^{(i)}\Big)^2\Big]$$
$$\le K_4\int_0^t\mathbb{E}\left[\sup_{u\le s}\{(\Delta V_u^{(i)})^2+(\Delta x_u^{(i)})^2\}\right]ds.$$

Then we have

$$\mathbb{E}\left[\sup_{s\le t}(\Delta x_s^{(i)})^2\right]\le K_x(1+T)\int_0^t\mathbb{E}\left[\sup_{u\le s}\{(\Delta V_u^{(i)})^2+(\Delta x_u^{(i)})^2\}\right]ds.$$

Putting all the computations together we obtain

$$\mathbb{E}\left[\sup_{s\le t}\left\|\tilde X_s^{(i)}-X_s^{(i)}\right\|^2\right]\le K(1+T)\int_0^t\mathbb{E}\left[\sup_{u\le s}\left\|\tilde X_s^{(i)}-X_s^{(i)}\right\|^2\right]ds+\frac{K_2T}{N},$$

from where we conclude thanks to Gronwall's Lemma. □

4.5 Long Time Behavior

From the mathematical point of view, it is very natural to try to answer the following three set of questions:

1. Is there an invariant distribution for the particle system? In that case, Is there convergence of the law of the system to this distribution as $t\to\infty$? Is this convergence to equilibrium homogeneous in the size of the network N?
2. Is there an invariant distribution for the nonlinear process? In that case, Is there convergence of the law of the process to this distribution as $t\to\infty$?
3. Does uniform propagation of chaos holds? That is

$$\sqrt{N}\sup_{t\ge0}\mathbb{E}\left(\left\|\tilde X_t^{(i)}-X_t^{(i)}\right\|_1\right)\le C.$$

Notice that if all these questions have a positive answer, then the invariant law of the nonlinear process can be approximated by the law of the particle system for N and t large. For example, if f is nice enough,

$$\left|\mathbb{E}\left[f(\tilde{X}_\infty)\right] - \mathbb{E}\left[f(X_t^N)\right]\right| \leq \left|\mathbb{E}\left[f(\tilde{X}_\infty)\right] - \mathbb{E}\left[f(\tilde{X}_t)\right]\right| + \left|\mathbb{E}\left[f(\tilde{X}_t)\right] - \mathbb{E}\left[f(X_t^N)\right]\right|,$$

and the first term in the right side can be controlled by the convergence to the equilibrium of the nonlinear process, whereas the second one can be controlled by the uniform propagation of chaos.

Unfortunately, we have not been able yet to give a positive answer to these questions, which have proved to be very difficult. In the case of the diffusive model, the system for the neurons (4.3.9) is hypoelliptic, not Hamiltonian and with non constant diffusion coefficients. Although there are some works in the literature addressing models with some of these difficulties, for example in [12] Bolley et al. study a kinetic equation which is hypoelliptic and not Hamiltonian or, in [76], Tugaut studies a nonlinear process subject to a non-convex potencial, the techniques they proposed are very *ad-hoc* and we did not succeed in adapting them to our case. In the case of the hybrid models, the main obstacle we found was the incompatibility between the $\|\cdot\|_2$ norm and the dynamic of the process. The hybrid model is in some degree dissipative, but exploiting this property requieres to deal with estimations in $\|\cdot\|_2$, which leads to bounds we were not able to close. To avoid these difficulties, in [9], Benaïn et al. consider a metric for the probability laws of hybrid models consisting in the Wasserstain distance for the continuous components and total variation distance for the proportion processes. We will comment later more details of that work.

Nevertheless, in the course of trying to understand the long term behavior of the particle system, we perform numerical simulations and we observed very interesting phenomena of synchronization, that we discuss in Sections 4.5.2 and 4.5.2.

We continue this section by commenting some known results about the long term behavior of system of neurons and the nonlinear process.

4.5.1 Some known results from the literature

Although in the literature there are other works that investigate the long time behavior of neuron models (see for example [33], [42]), we would like to remark two works which we consider closer to our case.

In the discrete context, we would like to comment the work by Benaïm et al. [9]. Let us recall the main results of the article.

Theorem 4.5.1 (Theorem 1.15 in [9]). *Let E be a finite set, and for any $i \in E$, $F^i : \mathbb{R}^d \to \mathbb{R}^d$ and smooth field. Let $(Z_t)_{t\geq 0} = (X_t, I_t)_{t\geq 0}$ be a Markov Process taking values on $\mathbb{R}^d \times E$, defined by its extended generator*

$$\mathcal{L}f(x,i) = \langle F^i(x), \nabla_x f(x,i)\rangle + \sum_{j\in E} a(x,i,j)\left[f(x,j) - f(x,i)\right].$$

Consider the following asumptions:

1. *There exists $\underline{a} > 0$ and $K > 0$ such that, for any $x, \tilde{x} \in \mathbb{R}^d$ and $i, j \in E$,*

$$a(x,i,j) \geq \underline{a} \quad \text{and} \quad \sum_{j\in E} |a(x,i,j) - a(\tilde{x},i,j)| \leq |x - \tilde{x}|.$$

2. *There exists $\alpha > 0$ such that,*

$$\langle x - \tilde{x}, F^i(x) - F^i(\tilde{x})\rangle \leq -\alpha|x - \tilde{x}|^2, \; x, \tilde{x} \in \mathbb{R}^d, i \in E.$$

Let η_t (resp. $\tilde{\eta}_t$) be the distribution at time t of the process Z_t starting from η_0 (resp.$\tilde{\eta}_0$), then there exist positive constants r, c and γ such that

$$\mathcal{W}_1(\eta_t, \tilde{\eta}_t) \leq (1 + 2r)(1 + ct)\exp\left(-\frac{\alpha}{1 + \alpha/\gamma}t\right), \tag{4.5.1}$$

where

$$\mathcal{W}_1(\eta, \tilde{\eta}) = \inf\left\{\left(\mathbb{E}\left[|X - \tilde{X}|^p\right]\right)^{1/p} + \mathbb{P}(I \neq \tilde{I}) : (X, I) \sim \eta(\tilde{X}, \tilde{I}) \sim \tilde{\eta}\right\},$$

and γ depends on the coalescence time of two independent processes defined on E as the second coordinates of independent copies of Z.

Corollary 4.5.2 (Corollary 1.16 in [9]). *Under the hypothesis of the previous Theorem, the process Z admits a unique invariant measure η_∞ and*

$$\mathcal{W}_1(\eta_t, \eta_\infty) \leq (1 + 2r)(1 + ct)\exp\left(-\frac{\alpha}{1 + \alpha/\gamma}t\right).$$

As an example, the authors of [9] apply their main result to a single neuron with a stochastic dynamics given by (4.2.8). Notice that in this case, in the notation of [9], $d = 1$, $E = \{1/N_c, 2/N_c, \ldots, 1 - 1/N_c, 1\}^2$, $i = (m, n)$ and

$$F^{(m,n)}(V) = I - g_{\mathrm{L}}(V - V_{\mathrm{L}}) - g_{\mathrm{Ca}}m(V - V_{\mathrm{Ca}}) - g_{\mathrm{K}}n(V - V_{\mathrm{K}}).$$

This model does not satisfy the uniform lower bound for the jump rates appearing in Hypothesis 1. In fact,

$$a\left(x, (m,n), \left(m + \frac{2}{N_c}, n\right)\right) = 0.$$

However, reviewing the proof of Theorem 4.5.1 it can be noticed that the hypothesis of uniform lower bound for the jump rates is only used to ensure that the coalescent time of the discrete components of two independent copies of the process starting from different initial conditions, $T_c(x, i, x', i')$, is stochastically bounded by an exponential random variable, uniformly on x, x', i, i'. That is, there exists $\beta > 0$ such that

$$\forall\, x, x' \in \mathbb{R}^d,\ \forall i, i' \in E\ \ \mathbb{P}\left(T_c(x, i, x', i') > t\right) \leq e^{-\beta t}.$$

In the case of the hybrid model (4.2.8), this property holds, thanks to the boundedness of the continuous part of the hybrid model and the continuity of $\rho_x(\cdot)$ and $\zeta_x(\cdot)$. On the other hand, the fields driving the continuous dynamic satisfy the hypothesis of coercitivity. Indeed,

$$\begin{aligned}\langle V - \tilde{V}, F^{(m,n)}(V) - F^{(m,n)}(\tilde{V})\rangle &= -g_{\mathrm{L}}(V - \tilde{V})^2 - g_{\mathrm{Ca}}m(V - \tilde{V})^2 - g_{\mathrm{K}}n(V - \tilde{V})^2 \\ &\geq -g_{\mathrm{L}}(V - \tilde{V})^2.\end{aligned}$$

In conclusion, the authors of [9] prove that the hybrid model for a single neuron has a unique invariant measure, and that the law of the process converges to it as time goes to infinity.

Following the arguments of Benaïm et al. we can apply the same variant of their main Theorem and obtain the following corollaries.

Corollary 4.5.3. *There exists an unique invariant distribution for the perturbed hybrid model* (4.2.13)*, and as time goes to infinity the law of the solution of* (4.2.13) *converges to the invariant distribution.*

Proof. Is enough to notice that the jump rates for the perturbed hybrid model are bounded from below by the jump rates of the original hybrid model, whereas the dynamic for V is bounded as well. □

Corollary 4.5.4. *There exists an unique invariant distribution for the particle systems* (4.3.7) *interacting under pure chemical synapses. Additionally, as time goes to infinity, the law of the solution of* (4.3.7) *converges to the invariant distribution.*

Proof. In the notation of [9], in this case $d = N$, $E = \{1/N_c, \ldots, 1\}^{3N}$, and

$$\begin{aligned}\langle V - \tilde{V}, F^{(m,n,y)}(V) - F^{(m,n,y)}(\tilde{V})\rangle &= -\sum_{i=1}^{N}\left[g_{\mathrm{L}} + g_{\mathrm{Ca}}m_i + g_{\mathrm{K}}n_i + J_{\mathrm{Ch}}\sum_{j=1}^{N} y_j\right](V_i - \tilde{V}_i)^2 \\ &\geq -g_{\mathrm{L}}\sum_{i=1}^{N}(V_i - \tilde{V})_i^2 \\ &= -g_{\mathrm{L}}\|V - \tilde{V}\|^2,\end{aligned}$$

so the hypothesis of coercitivity holds. The hypothesis about the exponentially decaying coalescent time for the discrete component of two independent copies of the process holds just as before. □

Remark 4.5.5. From the proof of the main Theorem of [9], we observe that the exponential rate of convergence decreases as the coalescence time $T_c(x,i,x',i')$ increases. Also, it is intuitive that the coalescence time grows with the number of element in E (in the hybrid model for N neurons $|E| = N_c^N$). Thus, if we think of the models (4.2.10) or (4.3.9) as approximations of the discrete models when the number of channels goes to infinite, it is not clear from these results that the diffusive models will inherit the existence of an invariant measure.

In the diffusive context, we would like to mention the work by Mischler, Quiñinao and Toubul [62]. In this work, the authors consider a stochastic version of the FitzHugh-Nagumo (FHN) model [32, 64] for a network of neurons interacting under electrical synapses.

In the FHN model, each neuron is described by its membrane potential V and a recovery variable w. The dynamics of the network is given by

$$\begin{aligned} V_t^{(i)} &= V_0^{(i)} + \int_0^t I_s + V_s^{(i)}(V_s^{(i)} - \lambda)(1 - V_s^{(i)}) - w_s^{(i)} + \varepsilon(V_s^{(i)} - \bar{V}_s^N)ds + \sigma W_t^{(i)}, \\ w_t^{(i)} &= w_0^{(i)} + \int_0^t bV_s^{(i)} - aw_s^{(i)}ds, \end{aligned} \tag{4.5.2}$$

where a, b are positive constants representing the timescale and coupling intensity between the two variables, $\sigma > 0$ and $\{W^{(i)}\}$ is a family of independent Browian motions. Notice that in this model the electrical interaction has the opposite sign than in our models.

It is proved in [62], that the property of propagation of chaos holds for system (4.5.2), and that the density of the nonlinear process satisfies the nonlinear PDE

$$\begin{gathered} \partial_t f = \partial_w(Af) + \partial_V(B_\varepsilon(\mathscr{J}_f)f) + \partial_{VV}f \ \text{ on } (0,\infty)\times\mathbb{R}^2, \\ A = A(V,w) = aw - bV, \ B_\varepsilon(\mathscr{J}_f) = B(w,V;\varepsilon,\mathscr{J}_f), \\ \mathscr{J}_f = \mathscr{J}(f) = \int_{\mathbb{R}^2} Vf(w,V)dVdw, \\ B(w,V;\varepsilon,j) = V(V-\lambda)(V-1) + w - \varepsilon(V-j) + I. \end{gathered} \tag{4.5.3}$$

The main results of [62] are the well-posedness of (4.5.3), existence of stationary solutions for the PDE for $\varepsilon \geq 0$ and uniqueness of stationary solutions for small values of ε. Notice that the case $\varepsilon = 0$ corresponds to a single neuron, so the result of Mischler et al. answers the question about the existence and uniqueness of a stationary measure in that case.

Summarizing, there is evidence to presume that at least some of the question we present at the beginning of the section have a positive answer, even if we have not been able to prove it yet. In the next sections we show the result of some simulations, and we establish the synchronization of a network of neurons interacting through pure electrical synapses.

4.5.2 Synchronization

Synchronization is a recurrent phenomena in nature. The synchronized flash of hundreds of fireflies is an example of how a system of interacting particles can self organize. In the case of neurons, their synchronization is related to the generation of rhythms as the respiratory one or the heartbeat. At brain level, synchronization in neurons is related to memory formation [5], and with epileptic seizures [48].

In this section we present the result of some numerical simulations of a network of Morris-Lecar neurons under electrical or chemical interactions. The simulations we present where coded in Python3.0 and performed in the cluster of the National Laboratory for High Performance Computing[2].

In the case of electrical interaction, we have found conditions on the parameters of the system that ensure the synchronization of the network, in a sense to be made precise later.

Pure Electrical Interaction

Let us consider the system with pure electrical interaction

$$\begin{aligned} V_t^{(i)} &= V_0^{(i)} + \int_0^t I - g_{\mathrm{Ca}} m_s^{(i)}(V_s^{(i)} - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_s^{(i)}(V_s^{(i)} - V_{\mathrm{K}}) ds \\ &\quad - \int_0^t g_{\mathrm{L}}(V_s^{(i)} - V_{\mathrm{L}}) ds - \int_0^t J_{\mathrm{E}}(V_s^{(i)} - \bar{V}_s^N) ds \\ x_t^{(i)} &= x_0^{(i)} + \int_0^t \rho_x(V_s^{(i)})(1 - x_s^{(i)}) - \zeta_x(V_s^{(i)}) x_s^{(i)} ds \\ &\quad + \int_0^t \sigma_x(V_s^{(i)}, x_s^{(i)}) dW_s^{x,i}, \quad x = m, n. \end{aligned} \tag{4.5.4}$$

Recall that, thanks to Proposition 4.3.1 and Remark 4.3.11, the trajectories of the system are uniformly bounded on t.

Before passing to the main result of this section, we comment the results of some simulations that motivate us. In Figure 4.5.1 we display the results for the simulation of a network of 100 neurons (first row), 1000 neurons (second row) and 10000 neurons (third row). It is clear that the system synchronizes, and compared to a system with chemical synapses, the system synchronizes fast. This result motivate us proving the following result.

Proposition 4.5.6. *Let $i, j \in \{1, \dots, N\}$, and $X_t^{(k)} = (V_t^{(k)}, m_t^{(k)}, n_t^{(k)})$, $k = i, j$, two neurons of the interacting system (4.5.4) starting from initial conditions $X_0^{(i)} = (V_0^{(i)}, m_0^{(i)}, n_0^{(i)})$. Assume that*

$$\eta_x = \inf_{V \in \mathbb{R}} \{\rho_x(V) + \zeta_x(V)\} > 0, \tag{4.5.5}$$

[2] http://www.nlhpc.cl/en/

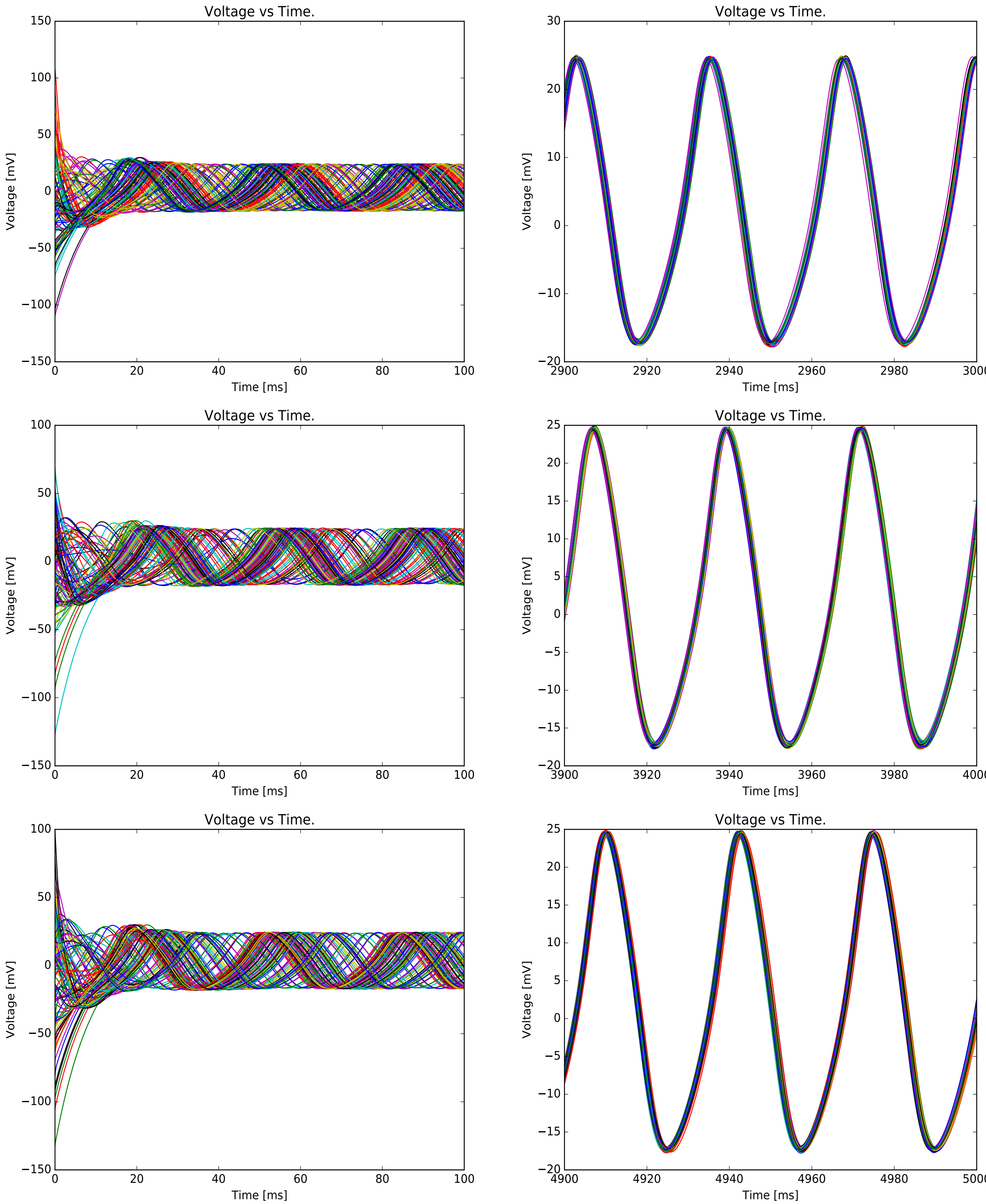


Figure 4.5.1: Trajectories of the Voltage at the beginning of the simulation (left), and when the system synchronizes (right), for $J_{\mathrm{E}} = 0.1$ and $\sigma = 0.1$. First row: Results for 100 neurons. Second row: Results for 1000 neurons. Third row: Results for 5000 neurons.

and moreover that J_E is big enough. Then there exists $\eta > 0$ not depending on N, such that

$$\mathbb{E}\left(\|X_t^{(i)} - X_t^{(j)}\|_2^2\right) \leq \|X_0^{(i)} - X_0^{(j)}\|_2^2 e^{-\eta t} + \sigma^2 C. \tag{4.5.6}$$

In particular

$$\limsup_{t\to\infty} \mathbb{E}\left(\|X_t^{(i)} - X_t^{(j)}\|_2^2\right) \leq \sigma^2 C.$$

Proof. Let $\Delta V_t = V_t^{(i)} - V_t^{(j)}$, $\Delta m_t = m_t^{(i)} - m_t^{(j)}$ and $\Delta n_t = n_t^{(i)} - n_t^{(j)}$. Then

$$\begin{aligned}
(\Delta V_t)^2 &= (\Delta V_0)^2 - 2\int_0^t g_{\mathrm{Ca}}(V_s^{(j)} - V_{\mathrm{Ca}})\Delta m_s \Delta V_s + g_{\mathrm{K}}(V_s^{(i)} - V_{\mathrm{K}})\Delta n_s \Delta V_s ds \\
&\quad - 2\int_0^t \left(g_{\mathrm{L}} + J_{\mathrm{E}} + g_{\mathrm{Ca}} m_s^{(i)} + g_{\mathrm{K}} n_s^{(i)}\right)(\Delta V_s)^2 ds \\
&\leq (\Delta V_0)^2 + \int_0^t \varepsilon(V,m)\,(\Delta m_s)^2 + \varepsilon(V,n)\,(\Delta n_s)^2 ds \\
&\quad - \int_0^t \left(2g_{\mathrm{L}} + 2J_{\mathrm{E}} - \frac{g_{\mathrm{Ca}}^2 (V_+ + |V_{\mathrm{Ca}}|)^2}{\varepsilon(V,m)} - \frac{g_{\mathrm{K}}^2 (V_+ + |V_{\mathrm{K}}|)^2}{\varepsilon(V,n)}\right)(\Delta V_s)^2 ds,
\end{aligned}$$

where $\varepsilon(V,x) \in (0,1)$ will be chosen later.

On the other hand, for $x = m, n$ we have

$$\begin{aligned}
\mathbb{E}\left[(\Delta x_t)^2\right] &= \mathbb{E}\left[(\Delta x_0)^2\right] + 2\int_0^t \mathbb{E}\left[(1 - x_t^{(i)})(\rho_x(V_t^{(j)}) - \rho_x(V_t^{(i)}))\Delta x_s\right] ds \\
&\quad + 2\int_0^t \mathbb{E}\left[x_t^{(i)}(\zeta_x(V_t^{(j)}) - \zeta_x(V_t^{(i)}))\Delta x_s\right] ds \\
&\quad - 2\int_0^t \mathbb{E}\left[\left(\rho_x(V_s^{(j)}) + \zeta_x(V_s^{(j)})\right)(\Delta x_s)^2\right] ds \\
&\quad + \int_0^t \mathbb{E}\left[\sigma_x^2(V_s^{(j)}, x_s^{(j)}) + \sigma_x^2(V_s^{(i)}, x_s^{(i)})\right] ds.
\end{aligned}$$

But, ρ_x, ζ_x are Lipschitz, and

$$\sigma_x^2(V,x) = \sigma^2|(1-x)\rho_x(V) + x\zeta_x(V)| \leq \sigma^2\left(\|\rho_x\|_\infty + \|\zeta_x\|_\infty\right),$$

So, we have

$$\begin{aligned}
\mathbb{E}\left[(\Delta x_t)^2\right] &\leq \mathbb{E}\left[(\Delta x_0)^2\right] + \int_0^t \mathbb{E}\left[\frac{L_{\rho_x}^2 + L_{\zeta_x}^2}{\varepsilon(V,x)}(\Delta V_s)^2\right] ds \\
&\quad - \int_0^t \mathbb{E}\left[(2\eta_x - 2\varepsilon(V,x))\,(\Delta x_s)^2\right] ds + \int_0^t \sigma^2\left(\|\rho_x\|_\infty + \|\zeta_x\|_\infty\right) ds.
\end{aligned}$$

Adding up,

$$\begin{aligned}\mathbb{E}\left(\|X_t^{(i)} - X_t^{(j)}\|_2^2\right) \leq \mathbb{E}\left[\|X_0^{(i)} - X_0^{(j)}\|_2^2\right] \\ - \int_0^t \mathbb{E}\Big[\Big(2g_{\mathrm{L}} + 2J_{\mathrm{E}} - \frac{g_{\mathrm{Ca}}^2(V_+ + |V_{\mathrm{Ca}}|)^2 + L_{\rho_m}^2 + L_{\zeta_m}^2}{\varepsilon(V,m)} \\ - \frac{g_{\mathrm{K}}^2(V_+ + |V_{\mathrm{K}}|)^2 L_{\rho_n}^2 + L_{\zeta_n}^2}{\varepsilon(V,n)} - \Big)(\Delta V_s)^2\Big] ds \\ - \int_0^t \mathbb{E}\left[(2\eta_m - 3\varepsilon(V,m))\,(\Delta m_s)^2\right] ds \\ - \int_0^t \mathbb{E}\left[(2\eta_n - 3\varepsilon(V,n))\,(\Delta n_s)^2\right] ds \\ + \int_0^t \sigma^2\left(\|\rho_m\|_\infty + \|\zeta_m\|_\infty\right) ds + \int_0^t \sigma^2\left(\|\rho_n\|_\infty + \|\zeta_n\|_\infty\right) ds\end{aligned}$$

For J_{E} big enough, we can optimize on the parameters $\varepsilon(V,x)$, in order that

$$\begin{aligned}\eta = \min\Big\{2g_{\mathrm{L}} + 2J_{\mathrm{E}} - \frac{g_{\mathrm{Ca}}^2(V_+ + |V_{\mathrm{Ca}}|)^2 + L_{\rho_m}^2 + L_{\zeta_m}^2}{\varepsilon(V,m)} - \frac{g_{\mathrm{K}}^2(V_+ + |V_{\mathrm{K}}|)^2 L_{\rho_n}^2 + L_{\zeta_n}^2}{\varepsilon(V,n)}, \\ 2\eta_m - 3\varepsilon(V,m),\, 2\eta_n - 3\varepsilon(V,n)\Big\} > 0.\end{aligned}$$

Then, we have for $C_{\zeta,\rho} = \|\rho_m\|_\infty + \|\zeta_m\|_\infty + \|\rho_n\|_\infty + \|\zeta_n\|_\infty$

$$\mathbb{E}\left(\|X_t^{(i)} - X_t^{(j)}\|_2^2\right) \leq \mathbb{E}\left[\|X_0^{(i)} - X_0^{(j)}\|_2^2\right] - \eta \int_0^t \mathbb{E}\left(\|X_s^{(i)} - X_s^{(j)}\|_2^2\right) + \int_0^t \sigma^2 C_{\zeta,\rho} ds.$$

If we apply to this inequality the Lemma 4.3.2, we obtain

$$\sqrt{\mathbb{E}\left(\|X_t^{(i)} - X_t^{(j)}\|_2^2\right)} \leq e^{-\frac{\eta t}{2}}\left(\mathbb{E}\left[\|X_0^{(i)} - X_0^{(j)}\|_2^2\right] + \int_0^t e^{\eta s}\sigma^2 C_{\zeta,\rho} ds\right)^{1/2},$$

from where the result follows. □

Remark 4.5.7. From the proof we can find explicit bounds for J_{E}, but this bounds seem to be sufficient but not necessary. In the simulations we observe synchronization for parameters that do not satisfy the bounds of the proof.

Remark 4.5.8. For $\sigma = 0$, the last proposition implies the full synchronization of the interacting system, which in this case corresponds to the deterministic interacting model (4.3.8).

Remark 4.5.9. In [62], the synchronization of a network of FitzHugh Nagumo neurons interacting through electrical synapses is observed in numerical simulations for J_{E} large in comparison with the noise of the system.

Pure Chemical Interaction

Let us start by recalling that in this case the evolution of the system is given by

$$\begin{aligned} V_t^{(i)} &= V_0^{i,} + \int_0^t I - g_{\mathrm{Ca}} m_s^{(i)} t (V_s^{(i)} - V_{\mathrm{Ca}}) - g_{\mathrm{K}} n_s^{(i)} (V_s^{(i)} - V_{\mathrm{K}}) ds \\ &\quad - \int_0^t g_{\mathrm{L}} (V_s^{(i)} - V_{\mathrm{L}}) ds - \int_0^t J_{\mathrm{Ch}} \bar{y}_s^N (V_s^{(i)} - V_{\mathrm{rev}}) ds \\ x_t^{(i)} &= x_0^{(i)} + \int_0^t (1 - x_s^{(i)}) \rho_x (V_s^{(i)}) - x_s^{(i)} \zeta_x (V_s^{(i)}) ds \\ &\quad + \int_0^t \sigma_x (V_s^{(i)}, x_s^{(i)}) \chi (x_s^{(i)}) dW_s^{i,x}, \quad x = m, n, y. \end{aligned} \tag{4.5.7}$$

In this case we have observed a different kind of synchronization than in the case of electrical synapses. Now we observe that the system self organizes in two clusters in anti-phase synchrony. This can be seen in Figure 4.5.2 where we show the simulation of a network of ten neurons following the dynamic (4.5.7). In Figure 4.5.3 we show the result of a simulation for 1000 neurons. To get better graphs, we only display the trajectories of 100 neurons. We observe that the anti-phase synchronization is still observed, but in this case the system needs more time to self organize.

For pure chemical interaction we have not been able to prove the anti-phase synchronization observed in the simulations. Looking at the proof of Proposition 4.5.6 one can notice that the same proof can be made for chemical interaction when the g_{L} parameter is big enough. However this case is not very interesting, since for g_{L} large in comparison with g_{Ca} and g_{K}, the solutions of the system (4.5.7) are not oscillatory, instead they goes very fast to an equilibrium point.

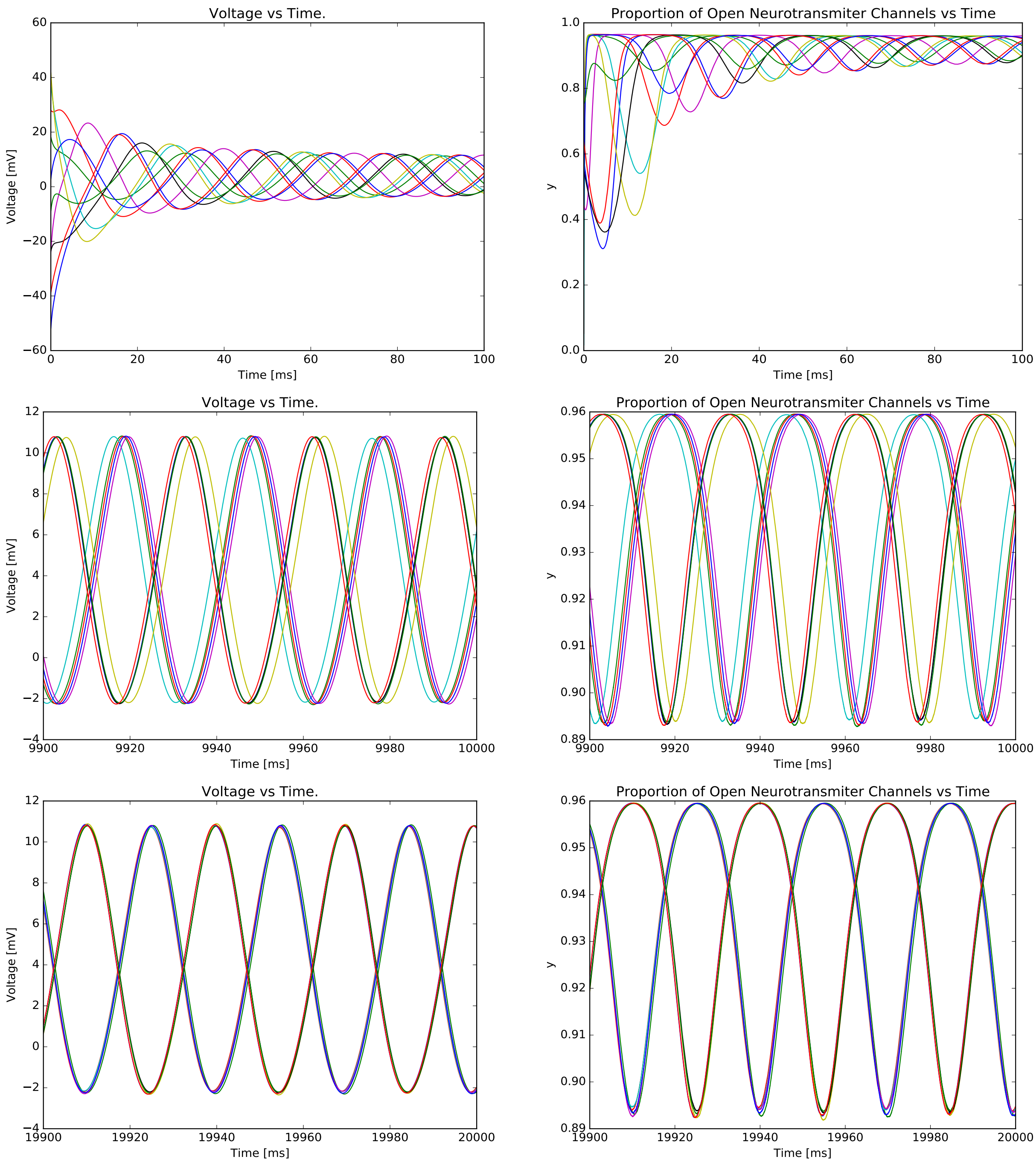


Figure 4.5.2: Trajectories of the Voltage $V^{(i)}$ (left) and proportion of open neurotransmitter channels $y^{(i)}$ (right) for 10 neurons, $J_{\text{Ch}} = 3$, $\sigma = 0.01$. In the first row we observe the beginning of the evolution, in the second row we observe that the neurons are almost synchronized. In the third row the neurons have reached the anti-phase synchronization.

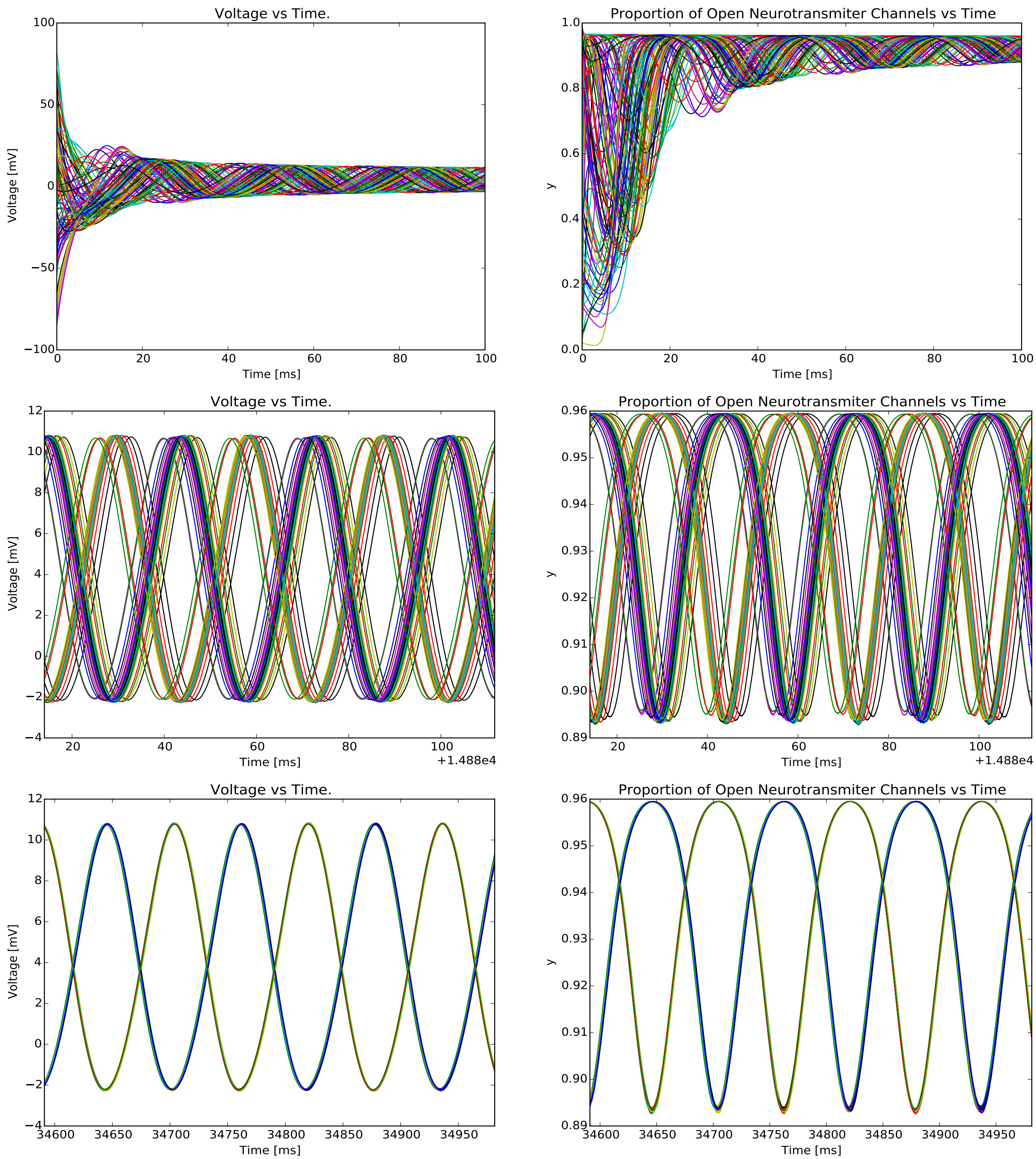


Figure 4.5.3: Trajectories of the Voltage $V^{(i)}$ (left) and proportion of open neurotransmitter channels $y^{(i)}$ (right) for 100 neurons evolving in a network of 1000 neurons. For $J_{\text{Ch}} = 3$ and $\sigma = 0.01$. In the first row we observe the beginning of the evolution, in the second row we observe that the neurons are starting to synchroniz. In the third row the neurons have reached the anti-phase synchronization.

4.6 Conclusions and Open Questions

.

In this part of the thesis we have discussed several aspect of the stochastic version of the Morris-Lecar model. To be more precise:

1. In Section 4.2 we have discussed different stochastic versions of the Morris-Lecar model, trying to understand the relationship between them, but alto their differences. In concrete we have introduced a perturbed hybrid model and we have proved how the "standard" diffusive approximation of the hybrid model is in fact the limit in distribution of this perturbed version.

2. In Section 4.3 we have incorporated interaction into the models discussed in the previous section and we have showed that the solutions of the different models are almost surely uniformly bounded for $t \geq 0$. This fact has resulted crucial for the proof of all our results.

3. In Section 4.4 we have studied the behavior of a network of neurons in a finite time window when the number of neurons goes to infinity. In 4.4.2 we have proved the propagation of chaos property for the hybrid model, which to the best of our knowledge was not known. On the other hand, in Section 4.4.3 we have recovered the propagation of chaos property for the diffusive model, which was already known, but we have been able to recover the optimal rate of convergence $\sqrt{N}$.

   Notice that the propagation of chaos for the diffusive model was already known, but since the hybrid model and its diffusive approximation can behave quite differently, it seems natural to ask if the known propagation of chaos result for the diffusive process is a characteristic of the model, or a characteristic of the phenomena which the model represents. Our result incline the balance to the second option.

4. In Section 4.5, we have proved the synchronization of the interacting particle system under electrical interaction when the time goes to infinite.

   This result is specially interesting in the light of the inital questions of this section. Notice that at first sight uniform propagation of chaos and synchronization seem to be incompatible properties. The first one tells us that for any time, when the number of neurons goes to infinite, the neurons become less and less correlated uniformly in time, whereas the second one tells us that for any number of neurons, as the time goes to infinite, the neurons synchronize i.e. they become highly correlated, but this incompatibility is only apparent. In fact, it could happen for example, that the network of neurons propagates chaos uniformly, and its nonlinear process after a long time start to follow a nearly deterministic trajectory, then by the propagation of chaos property for a large number of neurons each one of them will starts to behave like the nonlinear process, and after enough time the finite system will be synchronized. Notice that a stochastic process follows a deterministic trajectory if its law is a Dirac mass. It seems that we should study the density of the law of the nonlinear process with respect to the Lebesgue measure on $\mathbb{R}^4$. If we show some bounds uniform on time for the density of the nonlinear process, this could allow us to conclude that uniform propagation of chaos does not holds.

In regard of open questions, our initial questions about the long time behavior of the particle systems and the nonlinear processes remain unanswered, and the results of Benaïm et al. [9], and Mischler et al. [62] motivate us to follow the research in that direction.

Chapter 5

Final Remarks

We would like to start these final remarks by summarizing the main results of this thesis:

- In the first part of this thesis we have obtained a very concrete result, because we start with a very concrete question. For the equation

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma |X_s|^\alpha dW_s, \tag{5.0.1}$$

  we have proved, under suitable hypothesis over the coefficient, the strong convergence of the symmetrized Milstein scheme at rate Δt.

- In the second part of this thesis we have worked with three stochastic versions of the Morris-Lecar model: The hybrid model, the perturbed hybrid model and the diffusive model. We have obtained the following results:

  – We have obtain the convergence in L^1 of the hybrid model to the classical deterministic model when the number of channels goes to infinity. (The convergence in probability was already known).

  – We have characterized the diffusive model as the limit in distribution of a perturbed hybrid model. This is interesting because contribute to the discussion about the pertinence of the diffusive model as an approximation of the classical hybrid model.

  – We have obtained the property of propagation of chaos on finite time windows for the hybrid models, and for the diffusive model we have recovered the already known propagation of chaos property on finite time windows, but we have been able to obtain a better convergence rate in our specific case.

  – We have extended a result of existence of invariant measure and convergence towards it by Benaïm et al. [9] to the perturbed hybrid model, and to networks of neurons, although the convergence is not uniform neither in the number of channels or the number of neurons, so does not tell us anything about the diffusive model or the non linear process.

- We have proved the synchronization of the finite system under electrical interactions when the time goes to infinity.

We end this thesis by taking a look into the future.

A very natural continuation to the work of the first part of this thesis would be to study the weak convergence of the symmetrized Milstein scheme. In the weak analysis framework instead of studying

$$\mathbb{E}(|X_t - \overline{X}_t|),$$

we are interested in study the *weak error*

$$\left|\mathbb{E}[f(X_t)] - \mathbb{E}[f(\overline{X}_t)]\right|,$$

for f in a suitable set of functions. Of course, if f is Lipchitz, the later is bounded by the former, but a direct analysis of the weak error could led us to better hypothesis on the coefficients. In particular, could allows to relax the hypothesis over $b(0)$ and σ^2.

The importance of the weak error analysis comes from applications. In many occasion, particularly in finance, the object of interest is not the process itself but a functional of the process, so the natural error to control is the weak one.

Another posible continuation in this line of research would be try to apply a suitable version of the Symmetrized Milstein Scheme to the Heston model [38] for assets with stochastic volatility

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{\nu_t} S_t dW_t^S \\ d\nu_t &= \kappa(\theta - \nu_t)dt + \sigma\sqrt{\nu_t} dW_t^\nu, \end{aligned}$$

where (W_t^S, W_t^ν) is a two dimensional Brownian motion with correlated coordinates.

With respect to the second part of the thesis is more evident which are the questions that remain open. The existence and characterization of invariant measures, and the convergence toward them is a very interesting problem, specially in the light of the synchronization phenomena that is present in this models.

Synchronization by itself is another posible continuation for the present work. Here we have attacked the problem from a trajectorial perspective, mainly because existing techniques to analyze these type of models (e.g. from dynamical systems, see [53]) do not seem to be readily applicable or extendable to our context, in which the dimension is going to infinity and where the role of the randomness cannot be neglected. Nevertheless, it would be interesting to find a suitable approach to adapt those techniques to this setting, taking for instance into account the symmetries and exchangeability of the systems of neurons.

Uniform propagation of chaos is also a interesting question that is still open. The propagation of chaos property in finite time windows allows, for a fixed time interval, to approximate the interacting neuron system by an average neuron. But as the time grows, if the number of neurons remains fix, this approximation degrades quite fast.

In the case of the hybrid model, we have seen that the use of the $\|\cdot\|_1$ does not exploit the fact that the dynamics are dissipative. From the other side the norm $\|\cdot\|_2$ which is more appropriate to work with dissipative systems, does not work fine with the stochastic integral representation of the proportion processes. So, another extension for the work we have done is to find the right metric to work with the hybrid models.

Other open question is if the synchronization of the interacting system under electrical interactions can be exploit to simulate large networks more efficiently.

Another open question that this part of the thesis has left us is the convergence of the perturbed hybrid model to the diffusive model. Up to now, we know that this convergence is in distribution, but would be very interesting to obtain a stronger convergence.

We want to end this thesis with a thought that remind us that working in research is about the journey, and not about the destination.

"We have not succeeded in answering all our problems. The answers we have found only serve to raise a whole set of new questions. In some ways we feel we are as confused as ever, but we believe we are confused on a higher level and about more important things.[1]*"*

[1] This quote appears in [65, p. 4], but according to Oksendal, was posted outside the mathematics reading room in Tromsø University.

Bibliography

[1] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications*, 11:355–384, December 2005. 4, 24, 41

[2] A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Mathematics of Computation*, 79(269):209–237, 2010. 4

[3] A. Alfonsi. Strong order one convergence of a drift implicit Euler Scheme: Application to CIR process. *Statistic & Probability Letters*, 83(2):602–607, 2013. 4, 24, 41, 43, 51

[4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008. 92

[5] N. Axmacher, F. Mormann, G. Fernández, C. E. Elger, and J. Fell. Memory formation by neuronal synchronization. *Brain Research Reviews*, 52(1):170 – 182, 2006. 118

[6] F. Azevedo, L. Carvalho, L. Grinberg, J. Farfel, R. Ferretti, R. Leite, W. Filho, R. Lent, and S. Herculano-Houzel. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of Comparative Neurology*, 513(5):532–541, 2009. 9, 100

[7] M. Bachelier. Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, pages 21–86, 1900. iii, v

[8] J. Baladron, D. Fasoli, O. Faugeras, and J. Toubul. Mean field description of and propagation of chaos in networks of hodgkin-huxley and fitzhugh-nagumo neurons. *Journal of Mathematical Neuroscience*, 2(10), 2012. 9, 72, 84, 88, 91, 100, 101

[9] M. Benaïm, S. Le Borgne, F. Malrieu, and P.A. Zitt. Quantitative ergodicity for some switched dynamical systems. *Electronic Communications in Probability*, 17(56):1–14, 2012. 114, 115, 116, 117, 126, 127

[10] A. Berkaoui, M. Bossy, and A. Diop. Euler scheme for SDEs with non-Lipschitz diffusion coeffcient : strong convergence. *ESAIM Probability and Statistics*, 12:1–11, 2008. 4, 24, 28, 41, 42, 43

[11] F. Bolley, J. A. Cañizo, and J. A. Carrillo. Stochastic mean-field limit: Non-lipschitz forces and swarming. *Mathematical Models and Methods in Applied Sciences*, 21(11):2179–2210, 2011. 103

[12] F. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent vlasov-fokker-planck equation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44(5):867–884, 009 2010. 114

[13] A.N. Borodin and P. Salminen. *Handbook of Brownian Motion: Facts and Formulae*. Operator Theory, Advances and Applications. Birkhäuser, 2002. 54

[14] M. Bossy and A. Diop. Euler sheme for one dimensional SDEs with a diffusion coeffcient function of the form $|x|^a$, a in $[1/2, 1)$. *Preprint arXiv:1508.04573*, 2004. 4, 25, 59

[15] M. Bossy, O. Faugeras, and D. Talay. Clarification and complement to "mean-field description and propagation of chaos in networks of hodgkin–huxley and fitzhugh–nagumo neurons". *The Journal of Mathematical Neuroscience (JMN)*, 5(1):1–23, 2015. 9, 72, 84, 87, 100, 101, 111

[16] M. Bossy and H. Olivero. Strong convergence of the symmetrized milstein scheme for some cev-like sdes. *Preprint*, 2015. iii, v, 23

[17] F. Buchholtz, N. Schinor, and F. W. Schneider. Stochastic nonlinear dynamics:âĂŤ how many ion channels are in a single neuron? *The Journal of Physical Chemistry B*, 106(19):5086–5090, 2002. 78

[18] A. N. Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological Cybernetics*, 95(1):1–19, 2006. 7

[19] J.-F. Chassagneux, A. Jacquier, and I. Mihaylov. An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients. *Preprint*, 2015. 24, 39, 41, 42, 43, 45, 51

[20] J. Cox, J. Ingersoll, and S. Ross. A theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407, Mar. 1985. 2, 49

[21] C. E. Dangerfield, D. Kay, and K. Burrage. Modeling ion channel dynamics through reflected stochastic differential equations. *Physical Review E*, 85(5):051907, 2012. 84

[22] M. HA. Davis. Piecewise-deterministic markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 353–388, 1984. 77

[23] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of mckean–vlasov type. *The Annals of Applied Probability*, 25(4):2096–2133, 2015. 9

[24] Freddy Delbaen and Hiroshi Shirakawa. A note on option pricing for the constant elasticity of variance model. *Asia-Pacific Financial Markets*, 9(2):85–99, 2002. 23

[25] T. D Dreeben and S. B Pope. Wall-function treatment in pdf methods for turbulent flows. *Physics of Fluids (1994-present)*, 9(9):2692–2703, 1997. 4

[26] A. Einstein. The theory of the brownian movement. *Ann. der Physik*, 17:549, 1905. 2

[27] GB Ermentrout and D Terman. *Foundations of Mathematical Neuroscience.* Springer, Berlin, 2010. 5, 9, 73, 89, 91

[28] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley Series in Probability and Statistics, 1986. 86

[29] O. Faugeras, J. Touboul, and B. Cessac. A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Computational Neuroscience*, 3:1, 2009. 9, 100

[30] O. Faure. *Simulation du mouvement brownien et des diffusions.* PhD thesis, Ecole Nationale des Ponts et Chaussees, 1992. 14

[31] W. Feller. Two singular diffusion problems. *Annals of mathematics*, pages 173–182, 1951. 3

[32] R FitzHugh. Mathematical models of threshold phenomena in the nerve membrane. *Bull Math Biol*, 17(4):257–278, 1955. 71, 117

[33] N. Fournier and E. Löcherbach. On a toy model of interacting neurons. *https://arxiv.org/abs/1410.3263*, 2014. 9, 100, 114

[34] M. Giles. Improved Multilevel Monte Carlo Convergence using the Milstein Scheme. In Alexander Keller, Stefan Heinrich, and Harald Niederreiter, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pages 343–358. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. 49, 50

[35] M. B. Giles. Multilevel Monte Carlo Path Simulation. *Operations Research*, 56(3):607–617, 2008. 49

[36] D. Godinho. Asymptitic of grazing collisions for the spatially homogeneus boltzmann equation for soft and coloumb potentials. *Stochastic Processes and their application*, 2013. 88

[37] P. Greenwood and L. Ward. *Stochastic Neuron Models.* Springer International Publishing, 2016. 72

[38] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343, 1993. 128

[39] D. J Higham, X. Mao, and L. Szpruch. Convergence, non-negativity and stability of a new Milstein scheme with applications to finance. *Discrete Contin. Dyn. Syst. Ser. B*, 18(8):2083–2100, 2013. 4

[40] D.J. Higham. An Introduction to Multilevel Monte Carlo for Option Valuation. *Preprint*, 2015. 49

[41] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117, 1952. 8, 71

[42] R. Höpfner, E. Löcherbach, and M. Thieullen. Ergodicity for a stochastic hodgkin-huxley model driven by ornstein-uhlenbeck type input. *https://arxiv.org/abs/1311.3458*, 2013. 114

[43] S. G. Hormuzdi, M. A. Filippov, G. Mitropoulou, H. Monyer, and R. Bruzzone. Electrical synapses: a dynamic signaling system that shapes the activity of neuronal networks. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1662(1–2):113 – 137, 2004. The Connexins. 90

[44] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland, 1981. 1, 13, 82

[45] K. Itô. Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519–524, 1944. 2

[46] K. Itô. On a stochastic integral equation. *Proceedings of the Japan Academy*, 22(2):32–35, 1946. 2

[47] M. Jacobsen. *Point process theory and applications: marked point and piecewise deterministic processes*. Springer Science & Business Media, 2006. 77

[48] P. Jiruska, M. de Curtis, J. G. R. Jefferys, C. A. Schevon, S. J. Schiff, and K. Schindler. Synchronization and desynchronization in epilepsy: controversies and hypotheses. *The Journal of Physiology*, 591(4):787–797, 2013. 118

[49] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 1955, pages 171–197, 1954. 9

[50] C. Kahl and P. Jäckel. Fast strong approximation Monte Carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006. 4, 24, 42

[51] C. Kahl and H. Schurz. Balanced Milstein methods for ordinary SDEs. *Monte Carlo Methods and Applications*, 12(2):143–170, 2006. 4, 41, 42

[52] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer, 2nd edition, 1991. 1, 13, 17

[53] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer Berlin Heidelberg, 1984. 128

[54] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance, Second Edition*. Chapman & Hall/CRC Financial Mathematics Series. Taylor & Francis, 1996. 49

[55] P.-L. Lions and M. Musiela. Correlations and bounds for stochastic volatility models. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 24(1):1 – 16, 2007. 23

[56] C. Liu, X. Liu, and S. Liu. Bifurcation analysis of a morris–lecar neuron model. *Biological Cybernetics*, 108(1):75–84, 2014. 73

[57] G. Maruyama. Continuous markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo (1952 -)*, 4(1):48–90, 01-04 1955. 3, 14

[58] W. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943. 7

[59] H. P. McKean. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57, 1967. 101

[60] S. Méléard. Asymptotic behaviour of some interacting particle systems; mckean-vlasov and boltzmann models. In *Probabilistic models for nonlinear partial differential equations*, pages 42–95. Springer, 1996. 9, 101

[61] G. N. Milstein. Approximate Integration of Stochastic Differential Equations. *Theory of Probability & Its Applications*, 19:557–562, 1974. 14, 24

[62] S. Mischler, C. Quiñinao, and J. Touboul. On a kinetic fitzhugh–nagumo model of neuronal network. *Communications in Mathematical Physics*, 342(3):1001–1042, 2016. 9, 100, 117, 121, 126

[63] C. Morris and H. Lecar. Voltage oscillations in the barnacle gian muscle fiber. *Biophysical Journal.*, 31(1):193–213, 1981. 8, 71, 72, 73, 76

[64] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 1962. 71, 117

[65] B. Oksendal. *Stochastic Differential Equations.* Springer-Verlag Berlin Heidelberg, 6 th edition, 2003. 13, 129

[66] K. Pakdaman, M. Thieullen, and G. Wainrib. Diffusion approximation of birth-death processes: Comparison in term of large deviations and exit points. *Statistic and Probability Letters*, 2010. 84

[67] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with aplication to neuron models. *Arxiv*, 2010. 8, 72, 76, 83

[68] E. Pardoux. Backward Stochastic Differential Equations and Viscosity Solutions of Systems of Semilinear Parabolic and Elliptic PDEs of Second Order. In Laurent Decreusefond, Bernt Øksendal, Jon Gjerde, and Ali Süleyman Üstünel, editors, *Stochastic Analysis and Related Topics VI: Proceedings of the Sixth Oslo—Silivri Workshop Geilo 1996*, pages 79–127. Birkhäuser Boston, Boston, MA, 1998. 30

[69] P. E. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, 2nd edition, 2005. 1, 13

[70] S.T. Rachev and L. Rüschendorf. *Mass Transportation Problem*, volume I and II. Springer, 1998. 103

[71] B. Raesi, M. Hassanpour-Ezatti, and H. Zahmati. Classification of global phase portraits of morris-lecar model for spiking neuron. *Advanced Studies in Biology*, 4(4):175–194, 2012. 73

[72] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 1st edition, 1991. 58

[73] G.F. Simmons. *Differential equations: with applications and historical notes*. International series in pure and applied mathematics. McGraw-Hill, 1972. 30

[74] A.-S. Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer, 1991. 9, 101

[75] D. Talay. Probabilistic numerical methods for partial differential equations: elements of analysis. In D. Talay and L. Tubaro, editors, *Probabilistic Models for Nonlinear Partial Differential Equations*, volume 1627 of *Lecture Notes in Mathematics*, pages 148–196. Springer-Verlag, CIME Lectures, 1996. 14

[76] J. Tugaut. Self-stabilizing processes in multi-wells landscape in -convergence. *Stochastic Processes and their Applications*, 123(5):1780 – 1801, 2013. 114

[77] O. Vasicek. An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2):177–188, 1977. 2

[78] C. Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, 2003. 103

[79] C. Villani. *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften. Springer, 2009. 103

[80] N. Wiener. Differential-space. *Journal of Mathematics and Physics*, 2(1):131–174, 1923. 2

[81] N Wiener. Un problème de probabilité dénombrables. *Bulletin de la Société mathématique de France*, 52:569–578, 1924. 2