

2012

Sample Path Analysis of Integrate-and-Fire Neurons

Thibaud Taillefumier

Follow this and additional works at: http://digitalcommons.rockefeller.edu/student_theses_and_dissertations



Part of the [Life Sciences Commons](#)

Recommended Citation

Taillefumier, Thibaud, "Sample Path Analysis of Integrate-and-Fire Neurons" (2012). *Student Theses and Dissertations*. Paper 251.

This Thesis is brought to you for free and open access by Digital Commons @ RU. It has been accepted for inclusion in Student Theses and Dissertations by an authorized administrator of Digital Commons @ RU. For more information, please contact mcsweej@mail.rockefeller.edu.



SAMPLE PATH ANALYSIS OF INTEGRATE-AND-FIRE NEURONS

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
Thibaud Taillefumier
June 2012

SAMPLE PATH ANALYSIS OF INTEGRATE-AND-FIRE NEURONS

Thibaud Taillefumier, Ph.D.

The Rockefeller University, 2012

Computational neuroscience is concerned with answering two intertwined questions that are based on the assumption that spatio-temporal patterns of spikes form the universal language of the nervous system. First, what function does a specific neural circuitry perform in the elaboration of a behavior? Second, how do neural circuits process behaviorally-relevant information?

Non-linear system analysis has proven instrumental in understanding the coding strategies of early neural processing in various sensory modalities. Yet, at higher levels of integration, it fails to help in deciphering the response of assemblies of neurons to complex naturalistic stimuli. If neural activity can be assumed to be primarily driven by the stimulus at early stages of processing, the intrinsic activity of neural circuits interacts with their high-dimensional input to transform it in a stochastic non-linear fashion at the cortical level.

As a consequence, any attempt to fully understand the brain through a “system analysis” approach becomes illusory. However, it is increasingly advocated that neural noise plays a constructive role in neural processing, facilitating information transmission.

This prompts to gain insight into the neural code by studying the stochasticity of neuronal activity, which is viewed as biologically relevant. Such an endeavor requires the design of guiding theoretical principles to assess

the potential benefits of neural noise.

In this context, meeting the requirements of biological relevance and computational tractability, while providing a stochastic description of neural activity, prescribes the adoption of the integrate-and-fire model.

In this thesis, founding ourselves on the path-wise description of neuronal activity, we propose to further the stochastic analysis of the integrate-and-fire model through a combination of numerical and theoretical techniques.

To begin, we expand upon the path-wise construction of linear diffusions, which offers a natural setting to describe leaky integrate-and-fire neurons, as inhomogeneous Markov chains. Based on the theoretical analysis of the first-passage problem, we then explore the interplay between the internal neuronal noise and the statistics of injected perturbations at the single-unit level, and examine its implications on the neural coding. At the population level, we also develop an exact event-driven implementation of a Markov network of perfect integrate-and-fire neurons with both time-delayed instantaneous interactions and arbitrary topology.

We hope our approach will provide new paradigms to understand how sensory inputs perturb neural intrinsic activity and accomplish the goal of developing a new technique for identifying relevant patterns of population activity. From a perturbative perspective, our study shows how injecting frozen noise in different flavors can help characterize internal neuronal noise, which is presumably functionally relevant to information processing. From a simulation perspective, our event-driven framework is amenable to scrutinize the stochastic behavior of simple recurrent motifs as well as temporal dynamics of large scale networks under spike-timing-dependent plasticity.

A mon grand-père, André.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor, Prof. Marcelo Magnasco, for enlightening me with his constant creativity and for sharing his immense knowledge. His guidance helped shape my research while offering me the opportunity to freely explore my interests. This work would not have been possible without him. In addition, he was instrumental in deciding the pursuit of my career.

Besides my advisor, I would like to kindly thank my thesis committee members, Prof. George Reeke, Prof. Jonathan Victor, Prof. Bruce Knight and Prof. Eric Vanden-Eijnden, for their insightful comments, and astute questions.

I am particularly grateful to Prof. Maria Geffen who introduced me to the joy of electrophysiology and Prof. Jonathan Touboul who fostered my liking for theoretical neuroscience.

I am deeply indebted to my colleagues at the Physics Center: Dr. Daniel Andor, Dr. John Agapiou, Dr. Anton Zilman, Dr. Pradeep Kumar, Dr. Eleni Katifori, Philip Kidd, and Jacob Oppenheim for stimulating discussions and enriching interactions. Many thanks to Melanie Lee for being the organizing force behind everything.

It was a pleasure to have my graduate friends Jason, Yas, and Guillaume for roommates. They provided me with more entertainment than should be allowed during a thesis.

I also extend my thanks to the Dean's office and the staff of the Rockefeller University for their invaluable help.

I feel privileged to have entered the lives of my wife's parents, Al and Mary, who gave me a second home.

A special mention should be made of my brothers. Romain for his unwavering support, and Corentin for the promises of his early scientific curiosity. I would like to thank my parents, Patrice and Annie, and my grand mother, Yvette, for their dedication and sacrifices during my upbringing. Despite the geographical distance, my family was always there with me.

Finally, and most importantly, my deepest gratitude belongs to my wife, Jen, who has tolerated my sometimes unsavory temper, has always been inexplicably oblivious of my shortcomings, and at the same time, has shown remarkable abilities to celebrate any of my small victories.

Table of Contents

1	Introduction	1
1.1	Context	2
1.1.1	Classical Treatment in Early Sensory Pathways	2
1.1.2	Perturbative Approach suggested by New Experimental Possibilities	4
1.1.3	Including Noise at Every Stage of Neural Processing	5
1.1.4	Path-wise Approach and Integrate-and-Fire Model	8
1.1.5	Network of Integrate-and-Fire Neurons	11
1.2	General Results	12
1.2.1	Path-wise Description of Linear Diffusions	12
1.2.2	First-Passage Time Problem Analysis	13
1.2.3	Rigorous Markov Framework and Numerical Simulations	15
2	Muti-resolution Construction of Linear Diffusions	19
2.1	Rationale of the Construction	24
2.1.1	Multidimensional Linear Diffusions	24
2.1.2	Multi-Resolution Description of Linear Diffusions	30
2.2	System of Dual Schauder Bases	41
2.2.1	Underlying System of Orthonormal Functions	44
2.2.2	Generalized Dual Operators	49

2.2.3	Dual Basis of Generalized Functions	53
2.3	Inductive Construction of Linear Diffusions	56
2.3.1	The Space of Sample Paths	57
2.3.2	Finite-Dimensional Approximations	61
2.3.3	The Lévy-Cesielski Expansion	65
2.4	Derivation of the Bases for Some Classical Processes	69
2.4.1	One-Dimensional Case	69
2.4.2	Multidimensional Case	73
3	First-Passage Problem for Gauss-Markov Processes	83
3.1	Theoretical Approaches for the Wiener First-Passage Problem.	86
3.1.1	Direct Probabilistic Treatment	86
3.1.2	Partial Differential Equation Treatment	91
3.1.3	Integral Equation Formulation	97
3.2	Paradigm of a Piecewise Linear Barrier	102
3.2.1	Effect of the Slope Discontinuity of the Barrier	103
3.2.2	Effect of the Initial Probability Distribution	111
3.2.3	Heuristic Approach to Barrier's Reachability	115
3.3	Absolute Continuity of First-Passage Density and Barrier Regularity .	120
3.3.1	Probabilistic Treatment	121
3.3.2	Analytical Treatment	130
4	First-Passage Markov Framework	143
4.1	First-Passage Markov Chain	146
4.1.1	Cyclically Driven Leaky Integrate-and-Fire Neuron	146
4.1.2	Ergodicity of the Markov Chain	151
4.1.3	Numerical Simulation of the Markov Chain	158
4.2	First-Passage Time Monte-Carlo Algorithm	164

4.2.1	Probabilistic Dichotomic Search Algorithm	167
4.2.2	Analysis of the Algorithm	179
4.3	Simulation of a First-Passage Markov Chain	191
4.3.1	Transition between two Spiking Regimes	192
4.3.2	Spiking Precision and Spiking Reliability	196
4.3.3	Path-wise Perspective	200
4.3.4	Neural Code	202
4.4	First-Passage Neural Networks	209
4.4.1	Simulation of Integrate-and-Fire Networks	211
4.4.2	Stochastic Event-driven Strategy	213
4.4.3	New Practical Implementation	214
4.5	Exact Markovian Framework	216
4.5.1	Perfect Integrate-and-Fire Networks	216
4.5.2	Local Neuron Update Rules	219
4.5.3	Network Update Rule	223
4.6	Event-Driven Implementation	226
4.6.1	Pseudo-Code	226
4.6.2	Monte-Carlo Sampling	230
4.6.3	Complexity	231
4.7	Simulation of Networks	234
4.7.1	Inhibitory Interactions	235
4.7.2	Excitatory Interactions	236
4.7.3	Balanced Network	240
A	Optimality Criterion of the Multi-resolution Decomposition	247
A.1	Sample Paths Space as a Reproducing Hilbert Kernel Space	248
A.2	Finite-Dimensional Processes as Orthogonal Projections	253
A.3	Optimality Criterion of the Sample Paths	256

B	Random Generator for Monte-Carlo Simulation	261
B.1	First-Passage Time of Drifted Wiener Process	262
B.2	Inverse-Gaussian Bridge	264
B.3	First-Passage Brownian Bridge	266
	 Bibliography	 269

List of Figures

1.1	Spike Jittering and Barrier Regularity	10
2.1	Lévy-Ciesielski Construction of the Wiener Process	21
2.2	Recursive Construction and Markov Bridge	32
2.3	Lévy-Ciesielski Construction of the Ornstein-Uhlenbeck Process	72
2.4	Construction of the 2-dimensional rotating Ornstein-Uhlenbeck process.	75
2.5	Schauder Basis for the Doubly-Integrated Wiener Process.	80
3.1	Inverse Gaussian Distribution	90
3.2	Extension of the Density Kernel for an Increasing Affine Barrier . . .	96
3.3	Extension of the Density Kernel for a Decreasing Affine Barrier . . .	97
3.4	Density Kernel for a Concave Piecewise Linear Barrier	104
3.5	Density Kernel for a Convex Piecewise Linear Barrier	105
3.6	First-Passage Densities through a Concave Piecewise Linear Barrier .	110
3.7	First-Passage Densities through a Convex Piecewise Linear Barrier . .	111
3.8	First-Passage Sample Paths to a Concave Piecewise Barrier	117
3.9	First-Passage Sample Paths to a Convex Piecewise Barrier	118
4.1	Effective Markov Chain	149
4.2	Illustration of the Feller Property	154
4.3	Continuous Family of Hölder Continuous Barriers	163
4.4	Probabilistic Dichotomic Search of a First-Passage Time	180

4.5	Simulated Distribution of First-Passage Times	181
4.6	Computational Cost of the Algorithm.	187
4.7	Cost Comparison with the Traditional Euler Schema.	190
4.8	Peri-Stimulus Histogram, Low Reliability	193
4.9	Peri-Stimulus Histogram, High Reliability	193
4.10	Raster Plot over 1s	197
4.11	Raster Plot over 32ms	198
4.12	Raster Plot over 1ms	199
4.13	Empiric Rate Entropy	205
4.14	Kolomogorv-Smirnov Test	206
4.15	Recurrent Neural network	212
4.16	Non-interacting Perfect Integrate-and-fire Neurons	218
4.17	Local Neuron Update Rules	220
4.18	Principle of the Load Rule	225
4.19	Heap Structure of Spike Deliveries	229
4.20	Inhibitory Network of 4 Neurons in Asynchronous Time	235
4.21	Excitatory Network of 4 Neurons in Asynchronous Time	237
4.22	Excitatory Network of 4 Neurons with Anticipations	239
4.23	Balanced Network of 100 Neurons	241
4.24	Raster Plot of a Network of 200 Neurons	242
4.25	Computational time for a network of 200 neurons	244
B.1	Inverse-Gaussian Bridge Distribution	265
B.2	First-Passage Bridge Distribution	267

List of Algorithms

1	Recursive Sampling Algorithm	169
2	Naive Search Algorithm	170
3	Naive Search Algorithm	171
4	Probabilistic Dichotomic Search Algorithm	173
5	Basecase of the Recursive Search	176
6	Update Rule of the Network of IF Neurons	228
7	Random Generator for the Inverse-Gaussian Law	263
8	Random Generator for the Inverse-Gaussian Bridge Law	265
9	Random Generator for the Bessel Bridge Law	268

Chapter 1

Introduction

1.1 Context

The ultimate goal of sensory neuroscience is to understand how the brain parses its environment in order to elaborate behavioral responses. This aims at deciphering the encoding strategies of different sensory modalities, which transform an afferent physical stimulus into relevant patterns of neural activity, in the form of trains of spikes. Ideally, relevance should be understood in terms of conveying information about components of the stimuli that cause or affect a behavior.

1.1.1 Classical Treatment in Early Sensory Pathways

The stepping stone to tackle the encoding problem invariably consists of embedding the stimuli into a finite-dimensional feature space. The classical approach [138] then considers sensory processing as a stationary non-linear transformation of a temporal representation \mathbf{x}_s into an instantaneous rate function of spiking events r :

$$r(t) = \mathcal{F}(\{\mathbf{x}_s, 0 \leq s < t\}).$$

The coding strategy is thus implicitly assumed to be based on rate of events (rate coding) as opposed to temporal sequences of events (temporal coding). The dependence of the firing rate on the stimulus features is encoded by \mathcal{F} . This functional is characterized through standard Wiener-Itô chaos decomposition [95], by its Wiener kernels of successive orders \mathbf{k}_n , giving a theoretical ground to the concept of receptive fields [98]: more precisely, if $d\mathbf{W}_s$ is a multidimensional Gaussian white noise, the kernels \mathbf{k}_n are such that

$$\mathcal{F}(\{d\mathbf{W}_s, 0 \leq s < t\}) = \sum_{n=0}^{\infty} I_n(\mathbf{k}_n)$$

where $I_n(\mathbf{k}_n)$ denote the Itô iterated integrals

$$I_n(\mathbf{k}_n) = \int_0^t \int_0^{s_n} \dots \int_0^{s_1} \mathbf{k}_n(s_1, \dots, s_{n-1}, s_n) d\mathbf{W}_{s_1} \dots d\mathbf{W}_{s_{n-1}} d\mathbf{W}_{s_n} .$$

Wiener analysis provides a very fertile framework to develop analytic techniques that yield fine assessments of visual receptive fields [42], and with a lower degree of success, reliable spectro-temporal receptive fields in the auditory system [112].

A number of intrinsic limitations [154] prevent Wiener analysis from being the adequate setting to define the notion of receptive fields in a more general context. For instance as pointed in [9], Wiener analysis supposes that the encoding system has finite memory, which precludes to account for chaotic dynamics. More importantly, generalized Wiener analysis makes a very restraining assumption of rate coding and depends heavily on the use of unnatural stereotyped input signals [232].

Direct spike-triggered average methods (or reverse correlation, see e.g. [53]) approximate Wiener kernels when computed with appropriate input signals and allow the use of more naturalistic stimuli [192, 219]. If the presentation of an ensemble of stimuli elicits a spike train $t_1 < \dots < t_n$ in a neuron, the formula

$$\mathbf{k}(u) = \frac{1}{N} \int \Sigma^{-1}[\mathbf{x}](u, v) \left(\sum_{i=1}^N \mathbf{x}(v - t_i) \right) dv ,$$

gives an estimate of the linear response kernel, where $\Sigma^{-1}[\mathbf{x}]$ is the inverse covariance function of the stimuli. However, this paradigm invariably relies on the existence of a strict time-locking of spike events to the stimulus [222], which is a very stringent hypothesis at higher level of integration along neural pathways.

More generally, the notion of receptive fields is difficult to generalize to higher level neurons, whose activity presumably results from a cascade of non-linear filtering. Using spike-triggered averaging very quickly yields to the curse of dimensionality when estimating kernels of several dimensions and requires the use of ad-hoc,

non-linear, embedding techniques inspired from machine learning theory [196, 71]. The two previous complications dramatically increase when trying to explain the collective behavior of many neurons [234].

1.1.2 Perturbative Approach suggested by New Experimental Possibilities

An increasing body of data has been acquired from neural assembly recordings from behaviorally relevant complex stimuli, that shows that the very paradigm of receptive fields often fails to have explanatory powers (see e.g. [12]). Still, upon finding a complex stimulus that elicits a neural response, extremely reliable volleys of spikes are observed at cortical levels after repeated presentation of the same stimulus [3, 170, 13]. It has also been shown that such reliable patterns convey behaviorally relevant information [96, 99, 249].

This demonstrates the ability of the brain to encode information reliably after many integrative steps [171], in a form that conceptually does not require precise spatio-temporal patterns. Correlation coding [246], sparse coding [152], and rank coding [220] are examples of alternate candidate encoding strategies that can accommodate degrees of spatio-temporal plasticity. Consequently, a general program to understand neural coding can be cast in the terms a *perturbative approach*:

1. Identify a stimulus causing a neural response (i.e a burst of activity) within a set of naturally relevant stimuli whose representations are ideally embedded in a space of low dimensions (statistical representation such as sounds [143] or visual texture [167]).
2. Assess the reliability of the response in the sense of identifying some spatio-temporal patterns whose conservation through trials is statistically significant

when compared to spontaneous activity [139, 89, 162].

3. Perturb the identified stimulus in order to characterize the neural response sensitivity and retrospectively identify the relevant component of the stimulus (see e.g. [203]).

While the first step of the program is essentially the task of behavioral neuroscience and signal processing, the two other stages require the elaboration of a theoretical machinery in order to guide the analysis of integrated neuronal assemblies.

As argued in [30], most of the theory devoted to neural signal analysis was designed in view of analyzing a one-dimensional continuously-valued process, as opposed to the multi-dimensional point processes [49, 50] that best model spiking activity. In addition, recent experimental results [12] demonstrate that, as early as in the sensory cortical areas, neural activity cannot be assumed to be primarily driven by the stimulus. In cortex, the intrinsic activity of neural circuits is assumed to interact with their input to transform it in a non-trivial way [63].

The previous fact necessitates a paradigm shift. If, at early stages, spontaneous cortical firing can be conceived of as a nuisance to the transformation of the input, then, at higher levels of integration, the network’s intrinsic activity should become the central feature involved in the transformation and distribution of the input.

1.1.3 Including Noise at Every Stage of Neural Processing

Shifting away from the view that neural activity is primarily driven by sensory input requires two different forms of reasoning:

- First, as opposed to early sensory pathway stages, a neuron must always be considered to be embedded in a network of interacting units.
- Second, neuronal input is more likely to consist in highly preprocessed distributed spiking motifs, rather than in massive feedforward barrages.

Both these realizations depict neuronal networks as indirectly driven, weakly interacting systems, that are necessarily subjected to noise perturbations. This view directly echoes the long-claimed pervasive role of neural noise for single unit neural integration, as well as for cell assemblies' activity [65].

Single-Unit Level

Based on long-standing physiological observations, the sequence of events that leads to a spike generation is well-understood at the single-neuron level [93]. Schematically, the membrane voltage of a neuron fluctuates in response to synaptic inputs as well as to internal noise. When the voltage reaches the triggering threshold, nonlinear avalanche biological mechanisms cause the neuron to generate an action potential or spike.

In this respect, it is important to recognize the crucial role played by noise in shaping the propagation of the neural response [65]. Conceptually, the noise that accounts for the observed variability in neural responses comes into two flavors. The internal noise translates the inherent stochasticity of the molecular mechanisms that underpin the electrical activity of neurons and that agglomerate the effects of thermal noise and stochastic genetic expression. The external noise is the by-product of the spontaneous activity of surrounding neurons and summarized the highly variable bombardment that a neuron undergoes when embedded in an assembly.

At the single-unit level, the complex interplay between internal and external noise is best exemplified by considering the seminal experiment of Mainen and Sejnowski [133]. In their study, a steady current is injected to a neuron while recording the train of spikes the neuron emits in response. Upon repeating the exact same injection, each train of spikes tends to start with the same delay from the onset of the injection and stabilizes at the same given firing frequency. Over time, the spiking times gradually desynchronize, as if random perturbations were added independently

to each inter-spike interval. Thus, neuronal noise appears to be a nuisance resulting from the internal fluctuation of the cellular machinery, ultimately precluding spike timing to convey information.

However, if the cell is injected with the same steady input disturbed by a succession of many transient pulses, the neuron no longer fires regularly. Instead, it repeatedly produces precise patterns of firing that are closely time-locked to the time-varying input. In other words, what appears as transient random external fluctuations leads to a reliable precise neural response, allowing us to say that the neuron is driven by the noisy input. To emphasize that, despite being seemingly random, the previous perturbed input is actually a control input, it is referred as *frozen noise* [137]. The controlled injection of frozen noise reveals that, if a single-unit neuron displays a certain level of intrinsic noise, its spike generation mechanism is highly sensitive to transient features of the input and, incidentally, to the pervasive noisy fluctuations of afferent neural activity.

Neural Assemblies

The dynamics of spike generation within an assembly of neurons underlies neural coding, defined as the functional relation between the encoded information and the spatio-temporal spiking patterns [181, 53]. In this population context, neurons that are embedded into cortical cell assemblies receive about 10,000 connections from surrounding neurons [27]. As a result, neurons are constantly bombarded by spontaneous spike emissions of upstream neurons, ultimately giving rise to a background of external noise [66]. This noise can be thought of as an activity-dependent perturbation that superimposes itself onto the internal noise, and is thus a nuisance to the faithful transmission of temporal neural information [130]. At the same time, it has been argued that this inherent stochasticity is essential in moderating the collective activity of cell-assemblies which are otherwise prone to avalanche phenomenon [116].

However, the role of noise at the population level is not only that of a nuisance since stochastic variability can actually prove beneficial to neural transmission. For instance, it is well-documented that the presence of noise enhances the ability of threshold-like systems to detect weak signals, giving rise to the concept of stochastic resonance [189, 210]. From a more computational standpoint, the irregularity of seemingly random, yet reproducible, spiking patterns can only result from a high level of synchrony in the activity of population of neurons [36, 213]. Noise itself has been argued to cause synchrony, provided that the perturbations fed in different neurons are correlated [188, 75].

The role of noise in the central nervous system is increasingly seen as constructive, to such a point that its effect on neural functioning has recently been termed *stochastic facilitation* [144]. The experimental approach suggested by the framework of stochastic facilitation is essentially a perturbative method: assuming the noise is *relevant*, it proposes the injection a control noise in order to investigate a biological system by exploring its specific noise tuning. However, as pointed out in [166], such an inquiry requires a prior solid computational hypothesis, thus founding the need for a neural stochastic theory that is able to accommodate controlled perturbation by fluctuating input.

1.1.4 Path-wise Approach and Integrate-and-Fire Model

The adoption of a sample path description of stochastic neurons is the natural way to study the contribution of noise to their dynamics, while also preserving the essential mechanistic principle of neural integration. The central dogma of stochastic integrate-and-fire models consists in the assimilation of the fluctuating membrane potential to a continuous stochastic process X . In all generality, the dynamics of these processes

are described by a stochastic differential equation (or Langevin equation) of the type

$$dX_t = F(X_t, I(t), t) dt + \sigma(X_t, I(t), t) dW_t, \quad (1.1)$$

where F and σ are both dependent on the time-varying input I of the neuron. These models have been shown to reproduce neuronal firing better than models that are purely based on firing rate [174, 9, 77]. In the probability literature, equation (4.11) describes the evolution of a Markov diffusion process and a great deal is known about their properties [216]. However, these processes are described *in law* [100] and, despite recent advancement [131], path-wise properties are only available for certain simple processes.

Bearing this in mind, we propose to adopt the leaky integrate-and-fire model [180, 121, 179, 34, 35] as the fundamental encoding schema that transforms neural input into a train of spikes. For the sake of simplicity, we restrain our study to its simplest instance, the current-based model, where neuronal input takes the form of an injected current. The interesting aspect of this model lies primarily in the fact that it is admittedly an advantageous compromise between biological relevance and mathematical tractability [77]. Its simplicity primarily stems from its representation of the membrane potential as the only linear diffusion process with a stationary distribution: the Ornstein-Uhlenbeck process [230]. Incidentally, for being closely related to Brownian motion, the sample paths properties of such a process are known in depth.

The non-linearity of the model lies entirely in the membrane reset rule that is implemented after reaching the threshold and triggering a spiking event. Although it is simply stated, this spike-generation mechanism is more complex than it seems. Specifically, its definition as a first-passage time [173] means that each spike is an isolated discrete event that depends on the continuous past-history of the fluctuating

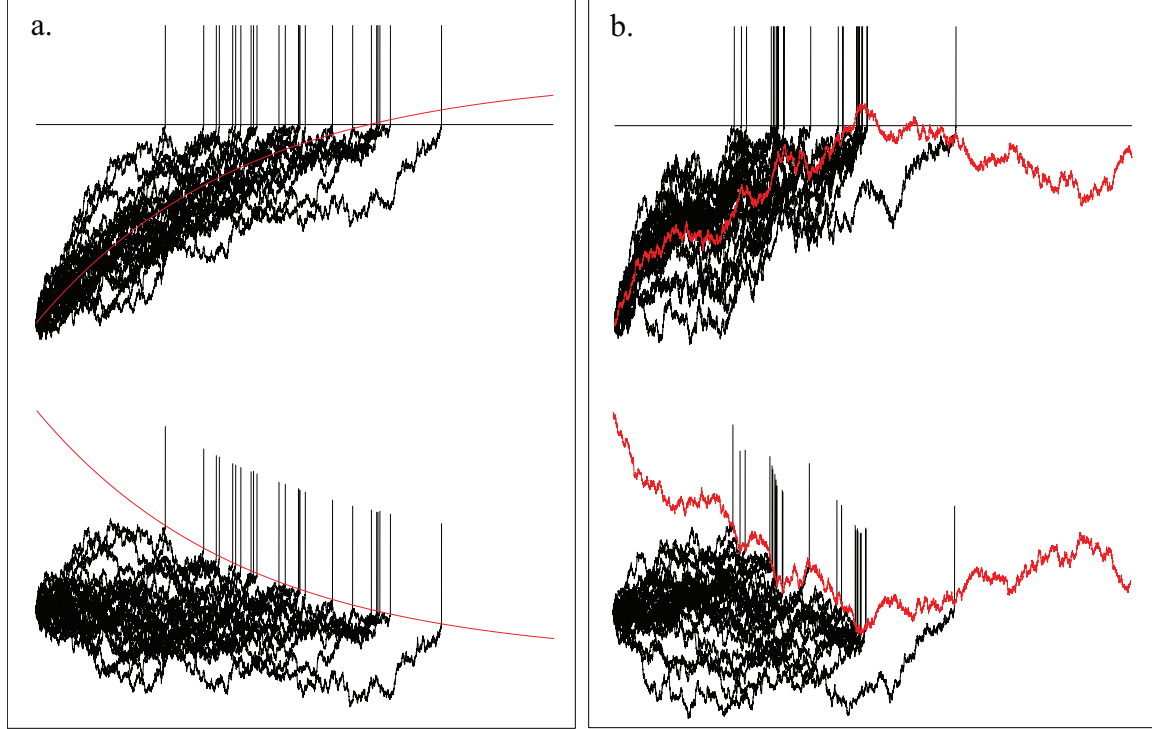


Figure 1.1: **a.** Leaky integrate-and-fire neuron injected with a steady current. Top: each black trace represents a membrane potential's trajectory, the red trace is the drift resulting from the integration of the input current. Bottom: the drift is subtracted from the voltage traces of **a.** to show the Ornstein-ulhenbeck sample paths. **b.** Leaky integrate-and-fire neuron injected with a frozen noise current. Top: the drift is a frozen Orstein-Ulhenbeck path added to the original exponential rise, spiking times tend to cluster in specific time regions. Bottom: subtracting the drift shows that frozen noise contribution can be encapsulated in terms of a fluctuating effective barrier.

input. This conceptual difficulty echoes the mixed nature of neural coding, which combines continuous (spike time) and discrete (spike count) aspects [236, 235]. To illustrate the cogency of a sample path based description in the context of a perturbative approach, we represent in Figure 1.1 a numerical experiment of a single neuron injected with a steady input and a frozen noise.

1.1.5 Network of Integrate-and-Fire Neurons

When analyzing the population dynamics of networks [153, 113], two issues quickly emerge [116]:

Transmission: What global architecture allows the network to display functional modular units whose activity can be reliably propagated between modules?

Transformation: What rules can a sub-population of neurons implement when carrying out information processing?

Purely random networks appear to be unable to sustain or convey steady neural activity. As such, a simple network organization, the synfire chain [2], is proposed. It consists of postulating the existence of individualized populations of preferentially recurrently connected neurons. At the same time, it prescribes the neurons within each population to connect to neurons in another population, in a feed-forward way. It is demonstrated that such organization ensures stable propagation of spiking activity while also promoting synchrony along the feed-forward circuit [111].

Accordingly, transformation can be naturally carried out conforming to two modalities: within module through potentially complex distributed recurrent circuitry, or through large scale integration at the distribution stage. Simulation of integrate-and-fire networks proves that local recurrent circuitry can implement basic logic operations [237], while networks with a detailed balance of excitation and inhibition allow for logical gating of whole sub-population of neurons [238].

In this approach noise is treated solely as a source of perturbations for the transmission and transformation of neural activity. Actually, in most simulations, the evolution of single-unit neurons is deterministic and the noise is applied externally and thought of as the chaotic behavior of a large ambient neural population. This renders most current simulation networks unsuited to developing a perturbative ap-

proach to the analysis of networks, which are assumed to have endogenous, potentially relevant, sources of noise. To remedy this situation, a promising Markov framework for the simulation of interacting integrate-and-fire neurons has recently been sketched, by analogy with telecommunication networks [226].

1.2 General Results

In this thesis, we exploit the path-wise properties of the integrate-and-fire encoding schema to contribute to the development of a stochastic perturbative theory, both at the single unit level and at the level of cell assemblies.

1.2.1 Path-wise Description of Linear Diffusions

Multidimensional linear diffusions form a direct generalization of the Ornstein-Uhlenbeck process that is routinely used in modeling the membrane potential of integrate-and-fire neurons. Recently, it has been suggested that integrated stochastic processes [224], which are Markov if considered as multidimensional processes, offer a convenient way to simulate smoothed neural input.

Here, we provide a general path-wise construction of multidimensional linear diffusions, which proceeds by refining asymptotically path-wise coarse-grained approximations of the process. The ideas underlying this work can be directly traced back to the original work of Lévy [126]. The construction relies on the existence of particular basis of functions, the Schauder basis, that are especially suited to expand the diffusion processes as a linear random expansion.

These functions exhibit properties that are of interest for numerical computations: *i*) all basis elements have compact support on an finite closed interval; *ii*) these intervals have a nested structure and are ever decreasing for larger indices of the basis element, and *iii*) for any interval endpoint, only a finite number of basis elements is

nonzero at that point. Thus, the expansion in our basis, when evaluated at an interval endpoint (e.g. dyadic rational), terminates in a finite number of steps. We later advantageously use these properties for the case of a one-dimensional linear diffusion with continuous barriers to elaborate a probabilistic dichotomic search algorithm for first passage time. From a more theoretical point of view, but with direct practical implications, we also extend some well-known results of interpolation theory in signal processing [55, 109, 110]. In particular, we associate with some canonical multidimensional stochastic processes their basis of decomposition as an extension of an already known family of functions: the natural splines of interpolation. This incidentally allows us to generalize a central result to support vector machine learning [197], the representer theorem [108], to matrix-valued kernels.

These results, exposed in Chapter 2, are fairly technical and are summed in the chapter's introduction. Only the result obtained for the one-dimensional case are used in the rest of the thesis.

There are a number of practical applications where applying the Schauder basis framework clearly provides an advantage compared to standard stochastic calculus methods. These include: first-hitting times of stochastic processes, pricing of multidimensional path-dependant options [16, 18, 17, 82], regularization techniques for machine learning [197], and more theoretical work on uncovering the differential geometry structure of the space of linear diffusions [201].

1.2.2 First-Passage Time Problem Analysis

The time at which the continuous sample path of a stochastic process first reaches a given boundary is a deceptively simple problem with many practical applications [106, 182, 231]. Its history can be traced back to Bachelier's study of speculation [14]. Since then, its mathematical study has expanded at the nexus of analysis and probability theory, involving both partial differential equations and stochastic

differential equations [114]. Aside from providing the basis for the spike generating process of integrate-and-fire neurons, the first attainment of the boundary can model the onset of a chemical reaction [119, 58], the triggering of a limit order for a commodity [247], or the lifetime of an infrastructure in structural analysis [120].

As apparent in Figure 1.1, spiking times of the linear integrate-and-fire neuron are first-passage times of an Ornstein-Uhlenbeck process with a fluctuating barrier [70, 207]. Here, we study in-depth the first-passage problem corresponding to a barrier generated through the integration of a frozen noise. In particular, we investigate the effect of the noise singularity that is encapsulated by the roughness of the barrier’s profile on the qualitative nature of its distribution. We focus on learning if the first-passage time admits a density function, a property that is readily assumed in most applications, but that can very well fail to hold true. In this regard, we show that the first-passage time admits a density as long as the barrier resulting from the current injection has a larger Hölder exponent H than the typical sample path of the Ornstein-Uhlenbeck process. The exponent H quantifies the strength of the frozen noise fluctuation and for an Ornstein-Uhlenbeck process, it is valued $H = 1/2$. These new results are made more precise and fully explained in Chapter 3. They are extensively used in Chapter 4 to shed light on the simulations of a single-unit neuron. Even if this work relies chiefly on the early results of Gevrey [78], later actualized in modern form [38], it makes full use of probability theory and analysis, and ultimately further their connection in the first-passage problem. In particular, we establish the validity of an integral equation that was previously only thought valid for differentiable barriers [33]. We also provide some heuristics about the nature of the first-passage distribution in the case of a frozen noise that is more singular than the internal noise, thus laying the groundwork to interpret the spiking activity of some later simulated integrate-and-fire neurons.

1.2.3 Rigorous Markov Framework and Numerical Simulations

Cyclically-Perturbed Leaky Integrate-and-Fire Neurons

Inspired from the seminal experiment of Mainen and Sejnowski [133], we identify the mathematical structure describing the firing activity of a leaky integrate-and-fire neuron that is cyclically driven by an input. Under mild hypothesis on the nature of the input, we show that spikes happen as generated through an ergodic Markov chain [90, 151, 214], whose Markov transition kernels are naturally deduced from the first-passage distribution for different reset times:

$$\tau_s = \inf \{t \mid X_t > V, X_s = v\}$$

where v is the reset value and V is the threshold. Within this rigorous framework, the distribution of spiking events converges independently of the initial state of the neuron to a unique stationary distribution μ , which is in turn identified with the instantaneous firing rate.

Next, we construct a family of frozen noise currents I_H , exhibiting varying degrees of fluctuation strength parametrized by H . This family of injected currents is actually constructed by gradually altering a frozen Gaussian white noise ($H = 1/2$) that has the same amplitude as the internal noise of the neuron. To circumvent the effect of varying firing rates over encoding capabilities of the neuron [228, 56, 47], we operate under the important constraint of constant mean firing activity. Thus, to the currents I_H , correspond a family of instantaneous spiking rates; the unique stationary distribution μ_H .

The coefficient H , controlling the strength of the frozen noise fluctuation, is identified to an Hölder singularity exponent or more convenient to the local Hölder continuity exponent of the drift it induces.

In dealing with highly irregular inputs, we resort to numerical Monte-Carlo simulations [146, 183] to infer the instantaneous firing rates by generating extremely long histories of spiking time, with exquisite precision. We demonstrate the existence of two regimes of spiking with constant mean firing rate. First, if the injected frozen noise is more singular than internal neuronal noise ($H < 1/2$), the driven neuron spikes with high temporal precision and high information rate (singular spiking mode). Second, if the frozen injected noise is less singular than internal neuronal noise ($H > 1/2$), the neuron has diminished temporal precision and low information rate (continuous spiking mode). Interestingly, the singular regime offers a simple solution to overcome the rate-specificity trade-off, which stipulates that neurons with high rate of information tend to exhibit low signal-specificity [215, 22, 181]. This observation appears to be generalizable to all current-based integrate-and-fire neurons and becomes the signature of the competition between drift and diffusive components of the dynamics, which shapes the voltage sample path in the vicinity of the threshold.

Network of Perfect Integrate-and-Fire Neurons

Embedding integrate-and-fire neurons in a network with interactions amounts to coupling the Markov chains that describe the spiking generation of non-interacting neuron. If these interactions are themselves Markov, the network dynamics *with coupling* can be described as a multidimensional Markov chain. However, the situation is much more entangled than for the case of a single unit [191, 29, 226]. For an individual neuron, a Markov chain proceeds from spiking time to spiking time, and there is no ambiguity about the nature of the next event. In the case of an interacting network, due to the intrinsically stochastic nature of its neurons' dynamics, there is no direct way to predict independently the nature of the next events for each neuron. Any time a neuron spikes, the neurons to which it is connected are bound to interact with the spiking neuron, thus invalidating any previous predictions about their spiking

times. Here, we elucidate the Markovian nature of the network’s evolution for a simple type of interaction but without any constraint on the topology of the network. The key element is carefully time-ordering the network events, together with their corresponding chain of interactions. This task is made possible by reasoning on the underlying sample paths of the neurons’ membrane potentials, which in turn yields a direct event-driven strategy to simulate the network [223, 187, 177, 142, 190].

In a current-based model, neuronal interactions necessarily operate through the delivery of spiking impulses. We adopt the simplest form of such interactions as sums of time-delayed instantaneous Dirac delta currents. More formally, if $n_i(t)$ denotes the number of times a neuron j has spiked before time t , with corresponding spiking history $t_1^j < \dots < t_n^j$, another neuron i receives a current I_i summing the contribution from the set of upstream neurons $U(j)$ that connects to it:

$$I_j(t) = \sum_{i \in U(j)} \sum_{n < n_i(t)} w_{i,j} \cdot \delta(t - (t_n^i - \delta_{i,j})).$$

In the previous expression, the weight $w_{i,j}$ is positive (negative) for excitatory (inhibitory) synapses and $\delta_{i,j}$ denotes the finite time of propagation from neuron i to neuron j ¹.

We stress that the inclusion of delays is absolutely compulsory for the well-posedness of the problem [103, 85]. Indeed, not only do interaction delays correspond to physical propagation periods along the axons, they also introduce an implicit ordering in the succession of individual neuron updates, that comprise a full network update. As explained later, in the case of instantaneous interaction, there is no unambiguous way to update a recurrent network (including feed-backs).

The special features of the linear stochastic integrate-and-fire model make it possible to formulate the evolution of the Markov network chain in a convenient way,

¹The weight and delay of interactions are actually allowed to be random variable as long as they only depend on the instantaneous state of the network when the spike is emitted.

through single-unit asynchronous update rules [226]. Schematically, the order with which these rules are called is determined as follows. For each instantaneous state of the network, suppose we estimate the next spiking event for every neuron that it comprises, as if the network was not interacting. Selecting the first of these spiking events is then equivalent to drawing the next network event (no other event happens before it!). Then, the local update rule corresponding to the neuron that spikes first is applied.

However, each time an update rule is implemented, it propagates interaction to other neurons. Maintaining the Markovian property demands that updated neurons bear the memory of the fact that, in the absence of interactions, they would spike at a known time in the future. While this book-keeping process appears cumbersome at first-sight, we manage to circumvent it in an efficient way, by extensively conditioning back and forth the neurons' membrane potential sample paths. As a result, the event-driven implementation that we found on the preceding principles outperforms computationally any other simulation method. In addition to being extremely efficient, this method presents the advantage of being mathematically exact, while proving particularly suited to the implementation of spike-timing dependent plasticity [53, 1, 39].

Chapter 2

Muti-resolution Construction of Linear Diffusions

Adopting a discrete representation of a process that allows the inference of sample paths properties from finite-dimensional approximating processes, alleviates the difficulty of working in multidimensional spaces. This is achieved through writing a process \mathbf{X} as an *almost-sure path-wise* convergent series of random functions

$$\mathbf{X}_t = \lim_{N \rightarrow \infty} \mathbf{X}_N \quad \text{with} \quad \mathbf{X}_N = \sum_{n=0}^N \mathbf{f}_n(t) \cdot \boldsymbol{\Xi}_n ,$$

where the \mathbf{f}_n are deterministic functions and the $\boldsymbol{\Xi}_n$ are independently identically distributed random variables.

The Lévy-Ciesielski construction [126] of the d -dimensional Brownian motion \mathbf{W} provides us with an example of discrete representation for a continuous stochastic process, whose principle is represented for the one-dimensional case in Figure 2.1. The simple form of the probability density of a Brownian bridge results in a representation based on completing sample paths by interpolation according to the conditional probabilities of the Wiener process. Specifically, the coefficients $\boldsymbol{\Xi}_n$ are Gaussian independent and the elements \mathbf{f}_n are called Schauder elements and denoted \mathbf{s}_n : $\mathbf{s}_{0,0}(t) = t \mathbf{I}_d$ and $\mathbf{s}_{n,k}(t) = s_{n,k}(t) \mathbf{I}_d$, where for all $n > 0$

$$s_{n,k}(t) = \begin{cases} 2^{\frac{n-1}{2}}(t - l_{n,k}) , & k2^{-n+1} \leq t \leq (2k+1)2^{-n} , \\ 2^{\frac{n-1}{2}}(r_{n,k} - t) , & (2k+1)2^{-n} \leq t \leq (k+1)2^{-n+1} , \\ 0 , & \text{otherwise} . \end{cases}$$

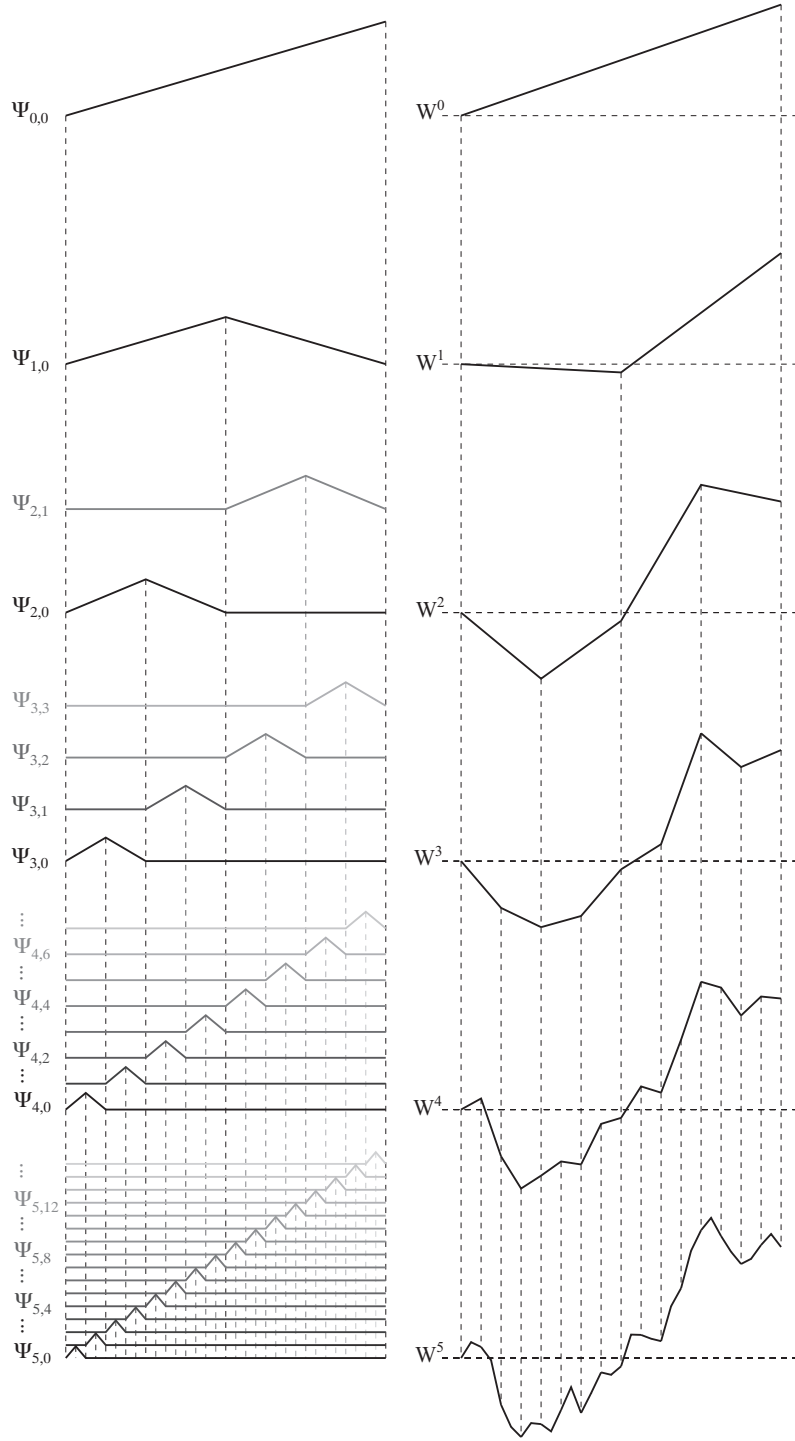


Figure 2.1: In the left column, the elements of the basis $s_{n,k}$ are represented for each rank n with $0 \leq n < 6$. In the right column, the partial sums $W^n(\omega)$ are shown for a given set of realizations ω . Note that each element $s_{n,k}$ has a compact support delimited by dyadic numbers in $D_n = \{k2^{-n} \mid 0 \leq k \leq 2^n\}$ and that all $s_{n',k}$ is zero on D_n for $n' > n$.

The Schauder elements are obtained by time-dependent integration of the Haar basis elements: $h_{0,0}(t) = 1$ and $h_{n,k}(t)$, where for all $n > 0$

$$h_{n,k}(t) = \begin{cases} 2^{\frac{n-1}{2}}, & k2^{-n+1} \leq t \leq (2k+1)2^{-n}, \\ -2^{\frac{n-1}{2}}, & (2k+1)2^{-n} \leq t \leq (k+1)2^{-n+1}, \\ 0, & \text{otherwise.} \end{cases}$$

The introduction of the Haar system is of relevance since it has several interesting properties. First, the functions $h_{n,k}$ form a complete orthonormal basis of $L^2([0, 1])$ for the scalar product $(f, g) = \int_0^1 f(t)g(t)dt$. Second, each element $h_{n,k}$ has a compact support

$$S_{n,k} = [k \cdot 2^{-n+1}, (k+1)2^{-n+1}]$$

and, for a given n , the collection of supports $S_{n,k}$ represents a partition of $[0, 1]$ up to the endpoints. Third, the functions $h_{n,k}$ build up a wavelet basis of $L^2([0, 1])$, since we have the scale-invariant construction rule

$$h_{n,k}(t) = 2^{\frac{n-1}{2}} \cdot h_{1,0}(2^{n-1}t - k). \quad (2.1)$$

For our purpose, it is well-known that the Hilbert property of this basis greatly simplifies the demonstration of the existence of the Wiener process [43]. It is also important for our purpose to realize that the Schauder elements \mathbf{s}_n have compact supports that exhibit a nested structure: this fact entails that the finite sums \mathbf{W}_N are processes that interpolate the limit process \mathbf{W} on the endpoints of the supports, i.e. on the dyadic points $k2^{-N}$, $0 \leq k \leq 2^N$. One of the specific goals of our construction is to maintain such a property in the construction of all multidimensional linear diffusions \mathbf{X} , being successively approximated by finite dimensional processes \mathbf{X}^N that interpolate \mathbf{X} at ever finer resolution. It is only in that sense that we refer to

our framework as a multi-resolution approach as opposed to wavelet multi-resolution theory [134]. Extensions of this method to the fractional Brownian motion were also developed [147], but applied to some very specific processes.

We first develop a heuristic approach for the construction of stochastic processes reminiscent of the midpoint displacement technique [126, 43], before rigorously deriving the multi-resolution basis. This set of functions is then studied as a multi-resolution Schauder basis of functions. It is important to recall that a Schauder basis f_n for a given space \mathcal{X} , is defined as a set of functions on which any element x of \mathcal{X} admits a unique decomposition. It is not to be confused with the Schauder elements, which only forms one exemplar of Schauder bases for the space of Wiener sample-paths. In particular, we derive explicitly from the multi-resolution basis an Haar-like Hilbert basis, which is the underlying structure explaining the dual relationship between basis elements and coefficients. Based on these results, we study the construction application and its inverse, the coefficient applications, that relate coefficients on the Schauder basis to sample paths. We pursue by proving the convergence of the process having independent standard normal coefficients on the Schauder basis to a linear diffusion. We also show that our decomposition is optimal in a sense that is strongly evocative of spline interpolation theory [54]: the construction yields successive interpolations of the process at the interval endpoints that minimizes the Dirichlet energy induced by the differential operator associated with the linear diffusion [72, 163]. We also provide a series of examples for which the proposed Schauder framework yields bases functions that have simple closed form formulae. In addition to the simple one-dimensional Markov processes, we explicit our framework for two classes of multidimensional processes: Gauss-Markov rotations and iteratively integrated Wiener processes (see e.g [145, 84, 124]).

2.1 Rationale of the Construction

In order to provide a discrete multi-resolution description of linear diffusions, we first establish basic results about the law of Gauss-Markov bridges in the multidimensional setting. We then use them to infer the candidate expressions for our desired bases of functions, while imposing its elements to be compactly supported on a nested sequence of segments.

2.1.1 Multidimensional Linear Diffusions

After recalling the definition of multidimensional linear diffusions in terms of stochastic integrals, we use the well known conditioning formula for Gaussian vectors to characterize the law of Gauss-Markov bridge processes.

Notations and Definitions

A d -dimensional continuous-time stochastic process \mathbf{X} is a collection of real random variables \mathbf{X}_t defined for a continuous index set of t on some abstract underlying measurable space (Ω, \mathcal{F}) , where Ω is the sample space and \mathcal{F} denotes its associated σ -field. Assuming the index set to be $[0, \infty)$, the process \mathbf{X} takes values in the space of functions $[0, \infty) \times \mathbb{R}^d$. For every realization ω in Ω , we call the outcome function $t \mapsto \mathbf{X}_t(\omega)$ a sample path or trajectory of the process \mathbf{X} . The natural filtration $\mathcal{F}_t \subset \mathcal{F}$ is the intersection of the σ -algebra $\sigma(\mathbf{X}_s)$ generated by \mathbf{X}_s for $0 \leq s < t$ and represents the *past history* of the process at time t .

In the following, we only deal with processes \mathbf{X} that are continuous, i.e. that it has continuous pathways $t \mapsto \mathbf{X}_t(\omega)$ for any ω in Ω . The state space of \mathbf{X} is the set of continuous functions on $[0, \infty)$ that are zero-valued at zero. Such a space is called the Wiener space $C_0[0, \infty)$ and is naturally provided with the σ -field $\mathcal{B}(C_0[0, \infty))$ generated by the cylinder sets $\mathcal{C}_t(A) = \{\mathbf{x} \in C_0[0, \infty) \mid \mathbf{x}(t) \in A\}$, where A is a real

Borelian in $\mathcal{B}(\mathbb{R}^d)$. Then, it is always possible to equip (Ω, \mathcal{F}) with a probability measure \mathbb{P} , so that \mathbf{X} induces a measure \mathbf{P} on $\mathcal{B}(C_0(0, \infty))$ defined on the generating cylinder sets $\mathcal{C}_t(A)$ by:

$$\mathbf{P}(\mathcal{C}_t(A)) = \mathbb{P}(\{\omega \mid X_t(\omega) \in A\}) \stackrel{def}{=} \mathbb{P}(X_t \in A),$$

Thus specified on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the process \mathbf{X} is entirely characterized by the measure \mathbb{P} and is referred as the law of the process.

We are now in a position to define the Gauss-Markov processes. Given some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let \mathbf{X} be a continuous stochastic process with natural filtration \mathcal{F}_t . Then \mathbf{X} is a Gauss-Markov process if it satisfies both Gaussian and Markov properties:

1. X is a Gaussian process if, for any integers k and positive reals $t_1 < t_2 < \dots < t_k$, the random matrix $(\mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \dots, \mathbf{X}_{t_k})$ has a joint Gaussian distribution.
2. \mathbf{X} is a Markov process if, for any $s, t \geq 0$ and $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\mathbb{P}(\mathbf{X}_{t+s} \in A \mid \mathcal{F}_s) = \mathbb{P}(\mathbf{X}_{t+s} \in A \mid \mathbf{X}_s),$$

which states that the conditional probability distribution of future states \mathbf{X}_{t+s} , given the present state and all past states \mathcal{F}_s , depends only upon the present state \mathbf{X}_s .

Linear Diffusion Processes

The canonical class of multidimensional Gauss-Markov processes is made of these processes that also are diffusions [216], meaning that their temporal dynamic is prescribed equivalently through a stochastic differential equation or a Fokker-Planck equation [169, 182]. When focusing on Gauss-Markov process, such a class of dif-

fusions is actually restrained to these stochastic processes whose time-evolution is linear. In the one-dimensional case, Gauss-Markov processes and diffusion processes can actually be identified [60]. In this chapter, we adopt the stochastic integral formalism (see e.g. [169]) since we find it is most amenable to develop a multidimensional construction in a rigorous way. Yet, since we emphasize on the path-wise description of processes, a point that stochastic integration disregard for the most part, we only use this formalism as a stepping stone to our study.

Let $(\mathbf{W}_t, \mathcal{F}_t, t \in [0, 1])$ be a m -dimensional Wiener process, consider the continuous functions $\boldsymbol{\alpha} : [0, 1] \rightarrow \mathbb{R}^{d \times d}$ $\sqrt{\boldsymbol{\Gamma}} : [0, 1] \rightarrow \mathbb{R}^{d \times m}$ as a time-dependent leak constant and define the positive bounded continuous function $\boldsymbol{\Gamma} = \sqrt{\boldsymbol{\Gamma}} \cdot \sqrt{\boldsymbol{\Gamma}}^T : [0, 1] \rightarrow \mathbb{R}^{d \times d}$ as an instantaneous diffusivity matrix. The linear diffusion associated process with these parameters is the d -dimensional Ornstein-Uhlenbeck process solution of the equation

$$d\mathbf{X}_t = \boldsymbol{\alpha}(t) \cdot \mathbf{X}_t + \sqrt{\boldsymbol{\Gamma}(t)} \cdot d\mathbf{W}_t, \quad (2.2)$$

and with initial condition \mathbf{X}_{t_0} in t_0 , it reads

$$\mathbf{X}_t = \mathbf{F}(t_0, t) \cdot \mathbf{X}_{t_0} + \mathbf{F}(t_0, t) \cdot \int_{t_0}^t \mathbf{F}(s, t_0) \cdot \sqrt{\boldsymbol{\Gamma}(s)} \cdot d\mathbf{W}_s, \quad (2.3)$$

where $\mathbf{F}(t_0, t)$ is the flow of the equation, namely the solution in $\mathbb{R}^{d \times d}$ of the linear equation:

$$\begin{cases} \frac{\partial \mathbf{F}(t_0, t)}{\partial t} &= \boldsymbol{\alpha}(t) \mathbf{F}(t) \\ \mathbf{F}(t_0, t_0) &= \mathbf{I}_d \end{cases}. \quad (2.4)$$

An Ornstein-Uhlenbeck process [230] is usually defined in the one-dimensional setting for constant parameters value α and Γ . However, we refer to \mathbf{X} since its expression as a stochastic integrale (2.3) are essentially the same as for a regular Ornstein-Uhlenbeck

process, with the restriction that, for being matrix-valued, all the intervening time-dependent functions are non-commuting. However, note that the flow $\Phi(t_0, t)$ enjoys the chain rule property:

$$\mathbf{F}(t_0, t) = \mathbf{F}(t_1, t) \cdot \mathbf{F}(t_0, t_1).$$

From multidimensional Gaussian calculus, the vectors \mathbf{X}_t and \mathbf{X}_s , $t_0 < s, t$, admit the covariance

$$\begin{aligned} \mathbf{C}_{t_0}(s, t) &= \mathbf{F}(t_0, t) \left(\int_{t_0}^{t \wedge s} \mathbf{F}(w, t_0) \mathbf{\Gamma}(w) \mathbf{F}(w, t_0)^T dw \right) \mathbf{F}(t_0, s)^T \\ &= \mathbf{F}(t_0, t) \mathbf{h}_{t_0}(s, t) \mathbf{F}(t_0, s)^T \end{aligned} \quad (2.5)$$

where we further defined $\mathbf{h}_u(s, t)$ the function

$$\mathbf{h}_u(s, t) = \int_s^t \mathbf{F}(w, u) \cdot \mathbf{\Gamma}(w) \cdot \mathbf{F}(w, u)^T dw$$

which will be of particular interest in the sequel. Note that because of the chain rule property of the flow, we have:

$$\mathbf{h}_v(s, t) = \mathbf{F}(v, u) \mathbf{h}_u(s, t) \mathbf{F}(v, u)^T \quad (2.6)$$

We suppose that the process \mathbf{X} is never degenerated, that is, for all $t_0 < u < v$, all the components of the vector \mathbf{X}_v knowing \mathbf{X}_u are non-deterministic random variables, which is equivalent to saying that the covariance matrix of \mathbf{X}_v knowing \mathbf{X}_u , denoted $\mathbf{C}_u(v, v)$ is symmetric positive definite for any $u \neq v$. Therefore, assuming the initial condition $\mathbf{X}_0 = \mathbf{0}$, the multidimensional centered process \mathbf{X} has a representation

(similar to Doob's representation for one-dimensional processes, see [106]) of form:

$$\mathbf{X}_t = \mathbf{g}(t) \int_0^t \mathbf{f}(s) \cdot d\mathbf{W}_s,$$

with $\mathbf{g}(t) = \mathbf{F}(0, t)$ and $\mathbf{f}(t) = \mathbf{F}(t, 0) \cdot \sqrt{\Gamma(t)}$.

Note that the processes at stake are defined on the time interval $[0, 1]$. However, because of the time-rescaling property of these processes, considering the processes on this time interval is equivalent to considering the process on any other bounded interval without loss of generality.

Conditional Law and Linear Diffusion Bridges

As stated previously, we aim at defining a multi-resolution description of linear diffusion processes. Such a description can be seen as an interpolation of the process over a series of regularly sampled times, with an inter-timing interval which gets increasingly finer. This principle, in addition to the Markov property, prescribes to characterize the law of \mathbf{X} at a time t , knowing the values \mathbf{x} and \mathbf{z} that \mathbf{X} takes at framing times $t_x < t < t_z$. This corresponds to the law of the corresponding linear diffusion bridge $\{\mathbf{X}_t \mid \mathbf{X}_{t_x} = \mathbf{x}, \mathbf{X}_{t_z} = \mathbf{z}\}$, that is the law of the Gauss-Markov process \mathbf{X} conditioned on its initial and final values. The bridge process of a Gauss process is still a Gauss process and, for a linear diffusion, its law can be computed as follows from Gaussian calculus:

Proposition 1. *Let $t_x \leq t_z$ be two times in the interval $[0, 1]$. For any $t \in [t_x, t_z]$, the random variable \mathbf{X}_t conditioned on $\mathbf{X}_{t_x} = \mathbf{x}$ and $\mathbf{X}_{t_z} = \mathbf{z}$ is a Gaussian variable with covariance matrix $\Sigma(t)$ and mean vector $\boldsymbol{\mu}(t)$ given by:*

$$\begin{aligned} \Sigma(t; t_x, t_z) &= \mathbf{h}_t(t_x, t) (\mathbf{h}_t(t_x, t_z))^{-1} \mathbf{h}_t(t, t_z), \\ \boldsymbol{\mu}(t) &= \boldsymbol{\mu}^l(t; t_x, t_z) \cdot \mathbf{x} + \boldsymbol{\mu}^r(t; t_x, t_z) \cdot \mathbf{z}, \end{aligned}$$

where the continuous matrix functions $\boldsymbol{\mu}^l(\cdot; t_x, t_z)$ and $\boldsymbol{\mu}^r(\cdot; t_x, t_z)$ of $\mathbb{R}^{d \times d}$ are given by:

$$\begin{cases} \boldsymbol{\mu}^l(t; t_x, t_z) &= \mathbf{F}(t_x, t) \mathbf{h}_{t_x}(t, t_z) (\mathbf{h}_{t_x}(t_x, t_z))^{-1} \\ \boldsymbol{\mu}^r(t; t_x, t_z) &= \mathbf{F}(t_z, t) \mathbf{h}_{t_z}(t_x, t) (\mathbf{h}_{t_z}(t_x, t_z))^{-1}. \end{cases}$$

Note that the functions μ^l and μ^r have the property that $\boldsymbol{\mu}^l(t_x; t_x, t_z) = \boldsymbol{\mu}^r(t_z; t_x, t_z) = \mathbf{I}_d$ and $\boldsymbol{\mu}^l(t_z; t_x, t_z) = \boldsymbol{\mu}^r(t_x; t_x, t_z) = \mathbf{0}$ ensuring that the process is indeed equal to \mathbf{x} at time t_x and \mathbf{z} at time t_z .

Proof. Let t_x, t_z be two times of the interval $[0, 1]$ such that $t_x < t_z$, and let $t \in [t_x, t_z]$. We consider the Gaussian random variable $\boldsymbol{\xi} = (\mathbf{X}_t, \mathbf{X}_{t_z})$ conditioned on the fact that $\mathbf{X}_{t_x} = \mathbf{x}$. Its mean can be easily computed from the expression (2.3) and reads:

$$(\mathbf{m}_t, \mathbf{m}_{t_z}) = (\mathbf{F}(t_x, t)\mathbf{x}, \mathbf{F}(t_x, t_z)\mathbf{x}) = (\mathbf{g}(t)\mathbf{g}^{-1}(t_x)\mathbf{x}, \mathbf{g}(t_z)\mathbf{g}^{-1}(t_x)\mathbf{x}),$$

and its covariance matrix, from equation (2.5), reads:

$$\begin{aligned} \begin{bmatrix} \mathbf{C}_{t,t} & \mathbf{C}_{t,t_z} \\ \mathbf{C}_{t_z,t} & \mathbf{C}_{t_z,t_z} \end{bmatrix} &= \begin{bmatrix} \mathbf{F}(t_x, t)\mathbf{h}_{t_x}(t_x, t)\mathbf{F}(t_x, t)^T & \mathbf{F}(t_x, t)\mathbf{h}_{t_x}(t_x, t)\mathbf{F}(t_x, t_z)^T \\ \mathbf{F}(t_x, t_z)\mathbf{h}_{t_x}(t_x, t)\mathbf{F}(t_x, t)^T & \mathbf{F}(t_x, t_z)\mathbf{h}_{t_x}(t_x, t_z)\mathbf{F}(t_x, t_z)^T \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{h}_t(t_x, t) & \mathbf{h}_t(t_x, t)\mathbf{F}(t, t_z)^T \\ \mathbf{F}(t, t_z)\mathbf{h}_t(t_x, t) & \mathbf{F}(t, t_z)\mathbf{h}_t(t_x, t_z)\mathbf{F}(t, t_z)^T \end{bmatrix}. \end{aligned}$$

From there, we apply the conditioning formula for Gaussian vectors (see e.g. [24]) to infer the law of \mathbf{X}_t conditioned on $\mathbf{X}_{t_x} = \mathbf{x}$ and $\mathbf{X}_{t_z} = \mathbf{z}$, that is the law $\mathcal{N}(\boldsymbol{\mu}(t), \boldsymbol{\Sigma}(t; t_x, t_z))$ of \mathbf{B}_t where \mathbf{B} denotes the bridge process obtained by pinning

\mathbf{X} in t_x and t_z . The covariance matrix is given by

$$\begin{aligned}
\Sigma(t; t_x, t_z) &= \mathbf{C}_{y,y} - \mathbf{C}_{y,z} \mathbf{C}_{z,z}^{-1} \mathbf{C}_{z,y}, \\
&= \mathbf{h}_t(t_x, t) - \mathbf{h}_t(t_x, t) (\mathbf{h}_t(t_x, t_z))^{-1} \mathbf{h}_t(t_x, t), \\
&= \mathbf{h}_t(t_x, t) (\mathbf{h}_t(t_x, t_z))^{-1} \mathbf{h}_t(t, t_z),
\end{aligned}$$

and the mean reads

$$\begin{aligned}
\boldsymbol{\mu}(t) &= \mathbf{m}_y + \mathbf{C}_{y,z} \mathbf{C}_{z,z}^{-1} (\mathbf{z} - \mathbf{m}_z) \\
&= \mathbf{F}(t_x, t) (\mathbf{I}_d - \mathbf{h}_{t_x}(t_x, t) (\mathbf{h}_{t_x}(t_x, t_z))^{-1}) \mathbf{x} \\
&\quad + \mathbf{F}(t_z, t) \mathbf{h}_{t_z}(t_x, t) (\mathbf{h}_{t_z}(t_x, t_z))^{-1} \mathbf{z}, \\
&= \underbrace{\mathbf{F}(t_x, t) \mathbf{h}_{t_x}(t, t_z) (\mathbf{h}_{t_x}(t_x, t_z))^{-1}}_{\boldsymbol{\mu}^l(t; t_x, t_z)} \cdot \mathbf{x} \\
&\quad + \underbrace{\mathbf{F}(t_z, t) \mathbf{h}_{t_z}(t_x, t) (\mathbf{h}_{t_z}(t_x, t_z))^{-1}}_{\boldsymbol{\mu}^r(t; t_x, t_z)} \cdot \mathbf{z},
\end{aligned}$$

where we have used the fact that $\mathbf{h}_{t_x}(t_x, t_z) = \mathbf{h}_{t_z}(t_x, t) + \mathbf{h}_{t_x}(t, t_z)$. The regularity of the thus-defined functions $\boldsymbol{\mu}_x$ and $\boldsymbol{\mu}_z$ directly stems from the regularity of the flow operator \mathbf{F} . Moreover, since for any $0 \leq t, u \leq 1$, we observe that $\mathbf{F}(t, t) = \mathbf{I}_d$ and $\mathbf{h}_u(t, t) = \mathbf{0}$, we clearly have $\boldsymbol{\mu}_x(t_x) = \boldsymbol{\mu}_y(t) = \mathbf{I}_d$ and $\boldsymbol{\mu}_x(t) = \boldsymbol{\mu}_y(t_x) = \mathbf{0}$. \square

2.1.2 Multi-Resolution Description of Linear Diffusions

Recognizing the Gauss property and the Markov property as the two crucial elements for a stochastic process to be expanded *à la* Lévy-Cesielsky, our approach first undertakes to exhibit bases of deterministic functions that would play the role of the Schauder bases for the Wiener process. In this regard, we first expect such functions to be continuous and compactly supported on increasingly finer supports (i.e. sub-intervals of the definition interval $[0, 1]$) in a similar nested binary tree structure.

Then, as in the Lévy-Ciesielsky construction, we envision that, at each resolution (i.e. on each support), the partially constructed process (up to the resolution of the support) has the same conditional expectation as the linear diffusion when conditioned on the endpoints of the supports. The partial sums obtained with independent Gaussian coefficients of law $\mathcal{N}(0, 1)$ will thus approximate the targeted linear diffusion in a multi-resolution fashion, in the sense that, at every resolution, considering these two processes on the intervals endpoints yields finite-dimensional Gaussian vectors which have the same law.

Nested Structure of the Sequence of Supports

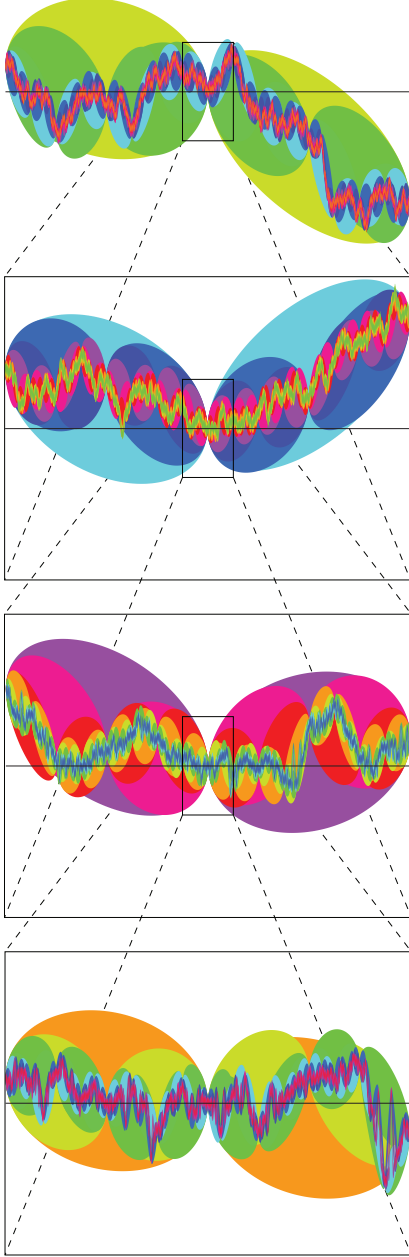
Here, we define the nested sequence of segments that constitute the supports of the multi-resolution basis. We construct such a sequence by recursively partitioning the interval $[0, 1]$.

More precisely, starting from $S_{1,0} = [l_{1,0}, r_{1,0}]$ with $l_{1,0} = 0$ and $r_{1,0} = 1$, we iteratively apply the following operation. Suppose that, at the n th step, the interval $[0, 1]$ is decomposed into 2^{n-1} intervals $S_{n,k} = [l_{n,k}, r_{n,k}]$, called supports, such that $l_{n,k+1} = r_{n,k}$ for $0 \leq k < 2^{n-1}$. Each of these intervals is then subdivided into two child intervals, a left child $S_{n+1,2k}$ and a right child $S_{n+1,2k+1}$, and the subdivision point $r_{n+1,2k} = l_{n+1,2k+1}$ is denoted $m_{n,k}$. Therefore, we have defined three sequences of real $l_{n,k}$, $m_{n,k}$, and $r_{n,k}$ for $n > 0$ and $0 \leq k < 2^{n-1}$ satisfying $l_{0,0} = 0 \leq l_{n,k} < m_{n,k} < r_{n,k} \leq r_{0,0} = 1$ and

$$l_{n+1,2k} = l_{n,k}, \quad m_{n,k} = r_{n+1,2k} = l_{n+1,2k+1}, \quad r_{n+1,2k+1} = r_{n,k}.$$

where we have posited $l_{0,0} = 0$ and $r_{0,0} = 1$ and $S_{0,0} = [0, 1]$. The resulting sequence of supports $\{S_{n,k}; n \geq 0, 0 \leq k < 2^{n-1}\}$ clearly has a binary tree structure.

a.



b.

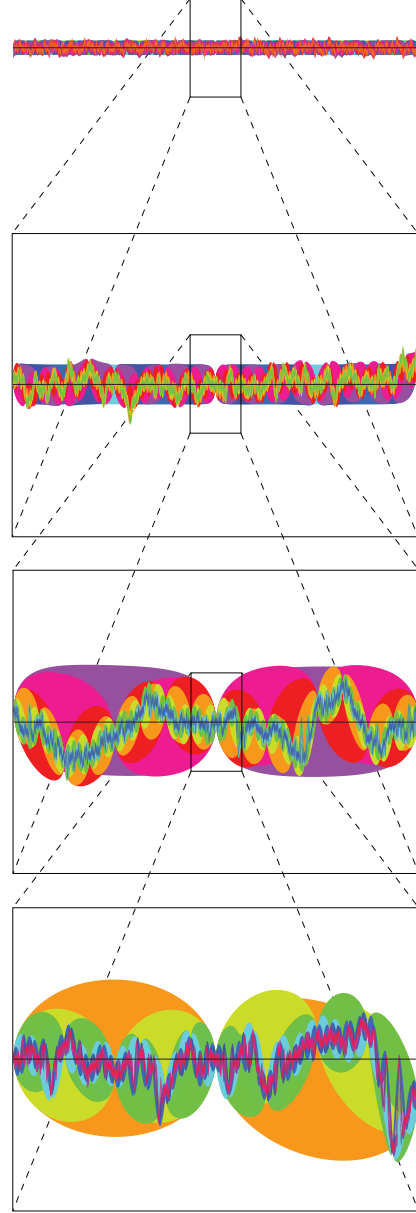


Figure 2.2: Recursive use of the Markov bridge to build a sample path of a Wiener (**a.**) and an Ornstein-Uhlenbeck (**b.**) process. At a given scale, knowing two consecutive sample points, we draw the next value according to bridge law of the process for an intermediary time: we represent graphically this law by coloring the region a standard deviation away from its mean value (colored blob). Iterating these drawings yields the apparent nested structure allowing us to sample the process accurately at any scale. At high resolution, an Ornstein-Uhlenbeck process approaches a Wiener process.

For the sake of compactness of notation, we define \mathcal{I} the set of indices

$$\mathcal{I} = \bigcup_{n < N} \mathcal{I}_n \quad \text{with} \quad \mathcal{I}_N = \{(n, k) \in \mathbb{N}^2 \mid 0 < n \leq N, 0 \leq k < 2^{n-1}\},$$

and for $N > 0$, we define $D_N = \{m_{n,k}, (n, k) \in \mathcal{I}_{N-1}\} \cup \{0, 1\}$, the set of endpoints of the intervals $S_{N,k}$. We additionally require that there exists $\rho \in (0, 1)$ such that for all $(n, k) \in \mathcal{I}$ $\max(r_{n,k} - m_{n,k}, m_{n,k} - l_{n,k}) < \rho(r_{n,k} - l_{n,k})$ which in particular implies that

$$\lim_{n \rightarrow \infty} \sup_k r_{n,k} - l_{n,k} = 0.$$

and ensures that the set of endpoints $\cup_{N \in \mathbb{N}} D_N$ is dense in $[0, 1]$. The simplest case of such partitions is the dyadic partition of $[0, 1]$, where the end points for $(n, k) \in \mathcal{I}$ read

$$l_{n,k} = k 2^{-n+1}, \quad m_{n,k} = (2k+1)2^{-n}, \quad r_{n,k} = (k+1)2^{-n+1}.$$

in which case the endpoints are simply the dyadic points $\cup_N D_N = \{k 2^{-N} \mid 0 \leq k \leq 2^N\}$. The nested structure of the supports, together with constraint of continuity of the bases elements, implies that only a finite number of coefficients are needed to construct the exact value of the process at a given endpoint, thus providing us with an exact schema to simulate sample values of the process on the endpoint up to an arbitrary resolution, as we will further explore.

Innovation processes for Linear Diffusions

For \mathbf{X}_t a multidimensional linear diffusion, we call multi-resolution description of a process the sequence of conditional expectations on the nested sets of endpoints D_n . In detail, if we denote \mathcal{F}_N the filtration generated by $\{\mathbf{X}_t; t \in D_N\}$ the values of the process at the endpoints D_N of the partition, we introduce the sequence of Gaussian

processes $(\mathbf{Z}_t^N)_{N \geq 1}$ defined by:

$$\mathbf{Z}_t^N = \mathbb{E} [\mathbf{X}_t \mid \mathcal{F}_N] = \mathbb{E}_N [\mathbf{X}_t] .$$

These processes \mathbf{Z}^N constitute a martingale taking values in the processes spaces that can be naturally viewed as an interpolation of the process \mathbf{X} sampled at the increasingly finer partitions times D_N , since for all $t \in D_N$ we have $\mathbf{Z}_t^N = \mathbf{X}_t^N$: indeed for all N, M with $N > M$

$$\mathbb{E} [\mathbf{Z}^N \mid \mathcal{F}_M] = \mathbb{E} [\mathbb{E} [\mathbf{X} \mid \mathcal{F}_N] \mid \mathcal{F}_M] = \mathbb{E} [\mathbf{X} \mid \mathcal{F}_M] = \mathbf{Z}^M .$$

The innovation process $(\boldsymbol{\delta}_t^N, \mathcal{F}_t, t \in [0, 1])$ is defined as the update transforming the process \mathbf{Z}_t^N into \mathbf{Z}_t^{N+1} , i.e.

$$\boldsymbol{\delta}_t^N = \mathbf{Z}_t^{N+1} - \mathbf{Z}_t^N . \tag{2.7}$$

It corresponds to the difference which the additional knowledge of the process at the points $m_{N,k}$ makes on the conditional expectation of the process. This process satisfies the following important properties found in our multi-resolution construction

Proposition 2. *The innovation process $\boldsymbol{\delta}_t^N$ is a centered Gaussian process independent of the processes \mathbf{Z}_t^n for any $n \leq N$. For $s \in S_{N,k}$ and $t \in S_{N,p}$ with $k, p \in \mathcal{I}_N$, the covariance of the innovation process reads:*

$$\mathbb{E}_N [\boldsymbol{\delta}_t^N \cdot (\boldsymbol{\delta}_s^N)^T] = \begin{cases} \boldsymbol{\mu}_{N,k}(t) \cdot \boldsymbol{\Sigma}_{N,k} \cdot \boldsymbol{\mu}_{N,k}(t)^T & \text{if } k = p \\ \mathbf{0} & \text{if } k \neq p, \end{cases} \tag{2.8}$$

where

$$\boldsymbol{\mu}_{N,k}(t) = \begin{cases} \boldsymbol{\mu}^r(t; l_{N,k}, m_{N,k}) & t \in [l_{N,k}, m_{N,k}] \\ \boldsymbol{\mu}^l(t; m_{N,k}, r_{N,k}) & t \in [m_{N,k}, r_{N,k}]. \end{cases}$$

with $\boldsymbol{\mu}^l$, $\boldsymbol{\mu}^r$ and $\Sigma_{N,k} = \Sigma(m_{N,k}; l_{N,k}, r_{N,k})$ as defined in Proposition 1.

Proof. Because of the Markovian property of the process \mathbf{X} , the law of the process \mathbf{Z}^N can be computed from the bridge formula derived in Proposition 1 and we have:

$$\mathbf{Z}_t^N = \boldsymbol{\mu}^l(t; l_{N,k}, r_{N,k}) \cdot \mathbf{X}_{l_{N,k}} + \boldsymbol{\mu}^r(t; l_{N,k}, r_{N,k}) \cdot \mathbf{X}_{r_{N,k}}.$$

and

$$\mathbf{Z}_t^{N+1} = \begin{cases} \boldsymbol{\mu}^l(t; l_{N,k}, m_{N,k}) \cdot \mathbf{X}_{l_{N,k}} + \boldsymbol{\mu}^r(t; l_{N,k}, m_{N,k}) \cdot \mathbf{X}_{m_{N,k}}, \\ \text{for } t \in [l_{N,k}, m_{N,k}], \\ \boldsymbol{\mu}^l(t; m_{N,k}, r_{N,k}) \cdot \mathbf{X}_{m_{N,k}} + \boldsymbol{\mu}^r(t; m_{N,k}, r_{N,k}) \cdot \mathbf{X}_{r_{N,k}}, \\ \text{for } t \in [m_{N,k}, r_{N,k}]. \end{cases}$$

Therefore, the innovation process can be written for $t \in S_{N,k}$ as

$$\boldsymbol{\delta}_t^N = \boldsymbol{\mu}_{N,k}^N(t) \cdot \mathbf{X}_{m_{N,k}} + \boldsymbol{\nu}^N(t) \cdot \mathbf{Q}_t^N$$

where \mathbf{Q}_t^N is a \mathcal{F}_N measurable process, $\boldsymbol{\nu}^N(t)$ a deterministic matrix function and

$$\boldsymbol{\mu}_{N,k}(t) = \begin{cases} \boldsymbol{\mu}^r(t; l_{N,k}, m_{N,k}) & t \in [l_{N,k}, m_{N,k}] \\ \boldsymbol{\mu}^l(t; m_{N,k}, r_{N,k}) & t \in [m_{N,k}, r_{N,k}]. \end{cases}$$

The expressions of $\boldsymbol{\nu}$ and \mathbf{Q} are quite complex, but are highly simplified when one

notes that

$$\begin{aligned}
\mathbb{E}[\boldsymbol{\delta}_t^N | \mathcal{F}_N] &= \mathbb{E}[\mathbf{Z}_t^{N+1} | \mathcal{F}_N] - \mathbf{Z}_t^N \\
&= \mathbb{E}\left[\mathbb{E}[\mathbf{Z}_t | \mathcal{F}_{N+1}] | \mathcal{F}_N\right] - \mathbf{Z}_t^N \\
&= \mathbf{0}
\end{aligned}$$

which directly implies that $\boldsymbol{\nu}(t) \cdot \mathbf{Q}_t^N = \boldsymbol{\mu}^N(t) \cdot \mathbf{Z}_{m_{N,k}}^N$ and which yields the remarkably compact expression:

$$\boldsymbol{\delta}_t^N = \boldsymbol{\mu}_{N,k}(t) \cdot (\mathbf{X}_{m_{N,k}} - \mathbf{Z}_{m_{N,k}}^N). \quad (2.9)$$

This process is a centered Gaussian process. Moreover, observing that it is \mathcal{F}_N -measurable, it can be written as:

$$\boldsymbol{\delta}_t^N = \boldsymbol{\mu}_{N,k}(t) \cdot (\{\mathbf{X}_{m_{N,k}} | \mathcal{F}_N\} - \mathbf{Z}_{m_{N,k}}^N).$$

and the process $\{\mathbf{X}_{m_{N,k}} | \mathcal{F}_N\}$ appears as the Gauss-Markov bridge conditioned at times $l_{N,k}$ and $r_{N,k}$, and whose covariance is given by Proposition 1 and that has the expression

$$\begin{aligned}
\Sigma_{N,k} &= \Sigma(m_{N,k}; l_{N,k}, r_{N,k}) \\
&= \mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k}) (\mathbf{h}_{m_{n,k}}(l_{n,k}, r_{n,k}))^{-1} \mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k}). \quad (2.10)
\end{aligned}$$

Let $(s, t) \in [0, 1]^2$, and assume that $s \in S_{N,k}$ and $t \in S_{N,p}$. If $k \neq p$, then because of the Markov property of the process \mathbf{X} , the two bridges are independent and therefore the covariance $\mathbb{E}_N [\boldsymbol{\delta}_t^N \cdot (\boldsymbol{\delta}_s^N)^T]$ is zero. If $k = p$, we have:

$$\mathbb{E}_N [\boldsymbol{\delta}_t^N \cdot (\boldsymbol{\delta}_s^N)^T] = \boldsymbol{\mu}_{N,k}(t) \cdot \Sigma_{N,k} \cdot \boldsymbol{\mu}_{N,k}(s)^T.$$

Eventually, the independence property stems from simple properties of the conditional

expectation. Indeed, let $n \leq N$. We have:

$$\begin{aligned}
\mathbb{E} \left[\mathbf{Z}_t^n \cdot (\boldsymbol{\delta}_s^N)^T \right] &= \mathbb{E} \left[\mathbf{Z}_t^n \cdot (\mathbf{Z}_s^{N+1} - \mathbf{Z}_s^N)^T \right] \\
&= \mathbb{E} \left[\mathbb{E} [\mathbf{X}_t | \mathcal{F}_n] \cdot (\mathbb{E} [\mathbf{X}_s^T | \mathcal{F}_{N+1}] - \mathbb{E} [\mathbf{X}_s^T | \mathcal{F}_N]) \right] \\
&= \mathbb{E} \left[\mathbb{E} [\mathbf{X}_t | \mathcal{F}_n] \cdot \mathbb{E} [\mathbf{X}_s^T | \mathcal{F}_{N+1}] \right] - \mathbb{E} \left[\mathbb{E} [\mathbf{X}_t | \mathcal{F}_n] \cdot \mathbb{E} [\mathbf{X}_s^T | \mathcal{F}_N] \right] \\
&= \mathbb{E} \left[\mathbf{Z}_t^n (\mathbf{Z}_s^n)^T \right] - \mathbb{E} \left[\mathbf{Z}_t^n (\mathbf{Z}_s^n)^T \right] \\
&= \mathbf{0}
\end{aligned}$$

and the fact that a zero covariance between two Gaussian processes implies the independence of these processes, which concludes the proof. \square

Derivation of the Candidate Multi-Resolution Bases of Functions

We deduce from the previous proposition the following fundamental theorem of this paper

Theorem 1. *For all $N \in \mathbb{N}$, there exists a collection of functions $\boldsymbol{\psi}_{N,k} : [0, 1] \mapsto \mathbb{R}^{d \times d}$ that are zero outside the sub-interval $S_{N,k}$ and such that in distribution we have:*

$$\boldsymbol{\delta}_t^N = \sum_{k \in \mathcal{I}_N} \boldsymbol{\psi}_{N,k}(t) \cdot \boldsymbol{\Xi}_{N,k}$$

where $\boldsymbol{\Xi}_{N,k}$ are independent d -dimensional standard normal random variables (i.e. of law $\mathcal{N}(0, \mathbf{I}_d)$). This basis of functions is unique up to an orthogonal transformation.

Proof. The two processes $\boldsymbol{\delta}_t^N$ and $\mathbf{d}_t^N \stackrel{\text{def}}{=} \sum_{k \in \mathcal{I}_N} \boldsymbol{\psi}_{N,k}(t) \cdot \boldsymbol{\Xi}_{N,k}$ are two Gaussian processes of mean zero. Therefore, we are searching for functions $\boldsymbol{\psi}_{N,k}$ vanishing outside $S_{N,k}$ and ensuring that the two processes have the same probability distribution. A necessary and sufficient condition for the two Gaussian processes to have the same probability distribution is to have the same covariance function (see e.g. [24]). We therefore need to show the existence of a collection of functions $\boldsymbol{\psi}_{N,k}(t)$ functions

that vanish outside the sub-interval $S_{N,k}$ and that ensure that the covariance of the process \mathbf{d}^N is equal to the covariance of $\boldsymbol{\delta}^N$. Let $(s, t) \in [0, 1]$ such that $s \in S_{N,k}$ and $t \in S_{N,p}$. If $k \neq p$, the assumption that the functions $\psi_{N,k}$ vanish outside $S_{N,k}$ implies that

$$\mathbb{E} [\mathbf{d}_t^N \cdot (\mathbf{d}_s^N)^T] = \mathbf{0}.$$

If $k = p$, the covariance reads:

$$\begin{aligned} \mathbb{E} [\mathbf{d}_t^N \cdot (\mathbf{d}_s^N)^T] &= \mathbb{E} [\boldsymbol{\psi}_{N,k}(t) \cdot \boldsymbol{\Xi}_{N,k} \cdot \boldsymbol{\Xi}_{N,k}^T \cdot (\boldsymbol{\psi}_{N,k}(s))^T] \\ &= \boldsymbol{\psi}_{N,k}(t) \cdot (\boldsymbol{\psi}_{N,k}(s))^T \end{aligned}$$

which needs to be equal to the covariance of $\boldsymbol{\delta}^N$, namely:

$$\boldsymbol{\psi}_{N,k}(t) \cdot (\boldsymbol{\psi}_{N,k}(s))^T = \boldsymbol{\mu}_{N,k}(t) \cdot \boldsymbol{\Sigma}_{N,k} \cdot (\boldsymbol{\mu}_{N,k}(s))^T. \quad (2.11)$$

Therefore, since $\boldsymbol{\mu}_{N,k}(m_{N,k}) = \mathbf{I}_d$, we have:

$$\boldsymbol{\psi}_{N,k}(m_{N,k}) \cdot (\boldsymbol{\psi}_{N,k}(m_{N,k}))^T = \boldsymbol{\Sigma}_{N,k}$$

meaning that $\boldsymbol{\sigma}_{N,k} \stackrel{def}{=} \boldsymbol{\psi}_{N,k}(m_{N,k})$ is a square root of the symmetric positive matrix $\boldsymbol{\Sigma}_{N,k}$. Moreover, by fixing $s = m_{N,k}$ in equation (2.11), we get:

$$\boldsymbol{\psi}_{N,k}(t) \cdot \boldsymbol{\sigma}_{N,k}^T = \boldsymbol{\mu}(t) \cdot \boldsymbol{\sigma}_{N,k} \cdot \boldsymbol{\sigma}_{N,k}^T$$

Eventually, since by assumption we have $\boldsymbol{\Sigma}_{N,k}$ invertible, so is $\boldsymbol{\sigma}_{N,k}$, and the functions $\boldsymbol{\psi}_{N,k}$ can be written as:

$$\boldsymbol{\psi}_{N,k}(t) = \boldsymbol{\mu}_{N,k}(t) \cdot \boldsymbol{\sigma}_{N,k} \quad (2.12)$$

with $\boldsymbol{\sigma}_{N,k}$ a square root of $\boldsymbol{\Sigma}_{N,k}$. Square roots of positive symmetric matrices are

uniquely defined up to an orthogonal transformation. Therefore, all square roots of $\Sigma_{N,k}$ are related by orthogonal transformations $\sigma'_{N,k} = \sigma_{N,k} \cdot \mathbf{O}_{N,k}$ where $\mathbf{O}_{N,k} \cdot \mathbf{O}_{N,k}^T = \mathbf{I}_d$. This property immediately extends to the functions $\psi_{N,k}$ we are studying: two different functions $\psi_{N,k}$ and $\psi'_{N,k}$ satisfying the theorem differ by an orthogonal transformation $\mathbf{O}_{N,k}$. We proved that, for $\psi_{N,k}(t) \cdot \Xi_{N,k}$ to have the same law as $\delta^N(t)$ in the interval $S_{N,k}$, the function $\psi_{N,k}$ with support in $S_{N,k}$ are necessarily of the form $\mu_{N,k}(t) \cdot \sigma_{N,k}$. It is straightforward to show the sufficient condition that provided such a set of functions, the processes δ_t^N and d_t^N are equal in law, which ends the proof of the theorem. \square

Using the expressions obtained in Proposition 1, we can make completely explicit the form of the basis in terms of the functions \mathbf{f} , \mathbf{g} and \mathbf{h} .

$$\psi_{n,k}(t) = \begin{cases} \mathbf{g}(t) \mathbf{g}^{-1}(m_{n,k}) \mathbf{h}_{m_{n,k}}(l_{n,k}, t) (\mathbf{h}_{m_{n,k}}((l_{n,k}, m_{n,k})))^{-1} \sigma_{n,k}, \\ \text{for } l_{n,k} \leq t \leq m_{n,k}, \\ \\ \mathbf{g}(t) \mathbf{g}^{-1}(m_{n,k}) \mathbf{h}_{m_{n,k}}(t, r_{n,k}) (\mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k}))^{-1} \sigma_{n,k}, \\ \text{for } m_{n,k} \leq t \leq r_{n,k}, \end{cases} \quad (2.13)$$

and $\sigma_{n,k}$ satisfies

$$\sigma_{n,k} \cdot \sigma_{n,k}^T = \mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k}) (\mathbf{h}_{m_{n,k}}(l_{n,k}, r_{n,k}))^{-1} \mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k}).$$

Note that $\sigma_{n,k}$ can be defined uniquely as the symmetric positive square root, or as the lower triangular matrix resulting from the Cholesky decomposition of $\Sigma_{n,k}$.

Let us now define the function $\psi_{0,0} : [0, 1] \mapsto \mathbb{R}^{d \times d}$ such that the process $\psi_{0,0}(t) \cdot \Xi_{0,0}$ has the same covariance as \mathbf{Z}_t^0 , which is computed using exactly the same technique

as developed in the proof of Theorem 1 and that has the expression

$$\boldsymbol{\psi}_{0,0}(t) = \boldsymbol{g}(t) \boldsymbol{h}_0(l_{0,0}, t) (\boldsymbol{h}_0(l_{0,0}, r_{0,0}))^{-1} \boldsymbol{g}^{-1}(r_{0,0}) \boldsymbol{\sigma}_{0,0},$$

for $\boldsymbol{\sigma}_{0,0}$ a square root of $\boldsymbol{C}_{r_{0,0}}$ the covariance matrix of $\boldsymbol{X}_{r_{0,0}}$ which from equation (2.5) reads:

$$\boldsymbol{F}(0, 1) \boldsymbol{h}_0(1, 1) \boldsymbol{F}(0, 1)^T = \boldsymbol{g}(1) \boldsymbol{h}_0(1, 1) (\boldsymbol{g}(1))^T.$$

We are now in position to show the following corollary of Theorem 1

Corollary 1. *The Gauss-Markov process \boldsymbol{Z}_t^N is equal in law to the process*

$$\boldsymbol{X}_t^N = \sum_{n=0}^{N-1} \sum_{k \in \mathcal{I}_n} \boldsymbol{\psi}_{n,k}(t) \cdot \boldsymbol{\Xi}_{n,k}$$

where $\boldsymbol{\Xi}_{n,k}$ are independent standard normal random variables $\mathcal{N}(0, \boldsymbol{I}_d)$.

Proof. We have:

$$\begin{aligned} \boldsymbol{Z}_t^N &= (\boldsymbol{Z}_t^N - \boldsymbol{Z}_t^{N-1}) + (\boldsymbol{Z}_t^{N-1} - \boldsymbol{Z}_t^{N-2}) + \dots + (\boldsymbol{Z}_t^2 - \boldsymbol{Z}_t^1) + \boldsymbol{Z}_t^1 \\ &= \sum_{n=1}^{N-1} \boldsymbol{\delta}_t^n + \boldsymbol{Z}_t^1 \\ &= \sum_{n=1}^{N-1} \sum_{k \in \mathcal{I}_n} \boldsymbol{\psi}_{n,k}(t) \cdot \boldsymbol{\Xi}_{n,k} + \boldsymbol{\psi}_{0,0}(t) \cdot \boldsymbol{\Xi}_{0,0} \\ &= \sum_{n=0}^{N-1} \sum_{k \in \mathcal{I}_n} \boldsymbol{\psi}_{n,k}(t) \cdot \boldsymbol{\Xi}_{n,k} \end{aligned}$$

□

We therefore identified a collection of functions $\{\boldsymbol{\psi}_{n,k}\}_{(n,k) \in \mathcal{I}}$ that allows a simple construction of the linear diffusions iteratively conditioned on increasingly finer partitions of the interval $[0, 1]$. We will show that this sequence \boldsymbol{Z}_t^N converges almost surely towards the linear diffusion \boldsymbol{X}_t used to construct the basis, proving that these

finite-dimensional continuous functions \mathbf{Z}_t^N form an asymptotically accurate description of the initial process. Beforehand, we rigorously study the Hilbertian properties of the collection of functions we just defined.

2.2 System of Dual Schauder Bases

The above analysis motivates the introduction of a set of functions $\{\psi_{n,k}\}_{(n,k) \in \mathcal{I}}$ we now study in details. In particular, we enlighten the structure of the collection of functions $\psi_{n,k}$ as a Schauder basis in a certain space \mathcal{X} of continuous functions from $[0, 1]$ to \mathbb{R}^d . The Schauder structure was defined in [193, 194], and its essential characterization is the unique decomposition property: namely that every element x in \mathcal{X} can be written as a well-formed linear combination

$$x = \sum_{(n,k) \in \mathcal{I}} \psi_{n,k} \cdot \xi_{n,k} ,$$

and that the coefficients satisfying this relation are unique.

To complete this program, we need to introduce some quantities that will play a crucial role in expressing the family $\psi_{n,k}$ as a Schauder basis for some given space. In equation (2.13), two constant matrices $\mathbb{R}^{d \times d}$ appear, that will have a particular importance in the sequel for (n, k) in \mathcal{I} with $n \neq 0$:

$$\begin{aligned} \mathbf{L}_{n,k} &= \mathbf{g}^T(m_{n,k}) (\mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k}))^{-1} \boldsymbol{\sigma}_{n,k} \\ &= (\mathbf{h}(l_{n,k}, m_{n,k}))^{-1} \mathbf{g}^{-1}(m_{n,k}) \boldsymbol{\sigma}_{n,k} , \\ \mathbf{R}_{n,k} &= \mathbf{g}^T(m_{n,k}) (\mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k}))^{-1} \boldsymbol{\sigma}_{n,k} \\ &= (\mathbf{h}(m_{n,k}, r_{n,k}))^{-1} \mathbf{g}^{-1}(m_{n,k}) \boldsymbol{\sigma}_{n,k} , \end{aligned}$$

where \mathbf{h} stands for \mathbf{h}_0 . We further define the matrix

$$\mathbf{M}_{n,k} = \mathbf{g}^T(m_{n,k}) \boldsymbol{\sigma}_{n,k}^{-1T}$$

and we recall that $\boldsymbol{\sigma}_{n,k}$ is a square root of $\boldsymbol{\Sigma}_{n,k}$, the covariance matrix of $\mathbf{X}_{m_{n,k}}$ knowing $\mathbf{X}_{l_{n,k}}$ and $\mathbf{X}_{r_{n,k}}$ given in equation (2.10). We stress that the matrices $\mathbf{L}_{n,k}$, $\mathbf{R}_{n,k}$, $\mathbf{M}_{n,k}$ and $\boldsymbol{\Sigma}_{n,k}$ are all invertible and satisfy the important following properties:

Proposition 3. *For all (n, k) in \mathcal{I} , $n \neq 0$, we have:*

$$i. \quad \mathbf{M}_{n,k} = \mathbf{L}_{n,k} + \mathbf{R}_{n,k} \text{ and}$$

$$ii. \quad \boldsymbol{\Sigma}_{n,k}^{-1} = (\mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k}))^{-1} + (\mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k}))^{-1}.$$

To prove this proposition, we first establish the following simple lemma of linear algebra:

Lemma 1. *Given two invertible matrices A and B in $GL_n(\mathbb{R})$ such that $C = A + B$ is also invertible, if we posit $D = AC^{-1}B$, we have the following properties:*

$$i. \quad D = AC^{-1}B = BC^{-1}A \text{ and}$$

$$ii. \quad D^{-1} = A^{-1} + B^{-1}$$

Proof. i) $D = AC^{-1}B = (C - B)C^{-1}B = B - BC^{-1}B = B(I - C^{-1}B) = BC^{-1}(C - B) = BC^{-1}A$.

ii) $(A^{-1} + B^{-1})D = A^{-1}D + B^{-1}D = A^{-1}AC^{-1}B + B^{-1}BC^{-1}A = C^{-1}(B + A) = C^{-1}C = I$. \square

Proof of Proposition 3.

(ii) directly stems from Lemma 1, item (ii) by posing $A = \mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k})$, $B =$

$\mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k})$ and $C = A + B = \mathbf{h}_{m_{n,k}}(l_{n,k}, r_{n,k})$. Indeed, the lemma implies that

$$\begin{aligned} D^{-1} &= A^{-1}CB^{-1} \\ &= \mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k})^{-1} \mathbf{h}_{m_{n,k}}(l_{n,k}, r_{n,k}) \mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k})^{-1} \\ &= \Sigma_{n,k}^{-1} \end{aligned}$$

(i) We have:

$$\begin{aligned} \mathbf{L}_{n,k} + \mathbf{R}_{n,k} &= \mathbf{g}(m_{n,k})^T (\mathbf{h}(l_{n,k}, m_{n,k})^{-1} + \mathbf{h}(m_{n,k}, r_{n,k})^{-1}) \boldsymbol{\sigma}_{n,k} \\ &= \mathbf{g}(m_{n,k})^T \Sigma_{n,k}^{-1} \boldsymbol{\sigma}^{n,k} \\ &= \mathbf{g}(m_{n,k})^T (\boldsymbol{\sigma}_{n,k}^{-1})^T \end{aligned}$$

which ends the demonstration of the proposition. \square

Let us define $\mathbf{L}_{0,0} = (\mathbf{h}(l_{0,0}, r_{0,0}))^{-1} \mathbf{g}^{-1}(r_{0,0}) \boldsymbol{\sigma}_{0,0}$. With this notations we define the functions in a compact form as:

Definition 1. For every (n, k) in \mathcal{I} with $n \neq 0$, the continuous functions $\boldsymbol{\psi}_{n,k}$ are defined on their support $S_{n,k}$ as

$$\boldsymbol{\psi}_{n,k}(t) = \begin{cases} \mathbf{g}(t) \mathbf{h}(l_{n,k}, t) \cdot \mathbf{L}_{n,k}, & l_{n,k} \leq t \leq m_{n,k}, \\ \mathbf{g}(t) \mathbf{h}(t, r_{n,k}) \cdot \mathbf{R}_{n,k}, & m_{n,k} \leq t \leq r_{n,k}, \end{cases} \quad (2.14)$$

and the basis element $\boldsymbol{\psi}_{0,0}$ is given on $[0, 1]$ by

$$\boldsymbol{\psi}_{0,0}(t) = \mathbf{g}(t) \mathbf{h}(l_{0,0}, t) \cdot \mathbf{L}_{0,0}.$$

The definition implies that the $\boldsymbol{\psi}_{n,k}$ are continuous functions in the space of piecewise derivable functions with piecewise continuous derivative which takes value zero at zero. We denote such a space $C_0^1([0, 1], \mathbb{R}^{d \times d})$.

Before studying the property of the functions $\psi_{n,k}$, it is worth remembering that their definitions include the choice of a square root $\sigma_{n,k}$ of $\Sigma_{n,k}$. Properly speaking, there is thus a class of bases $\psi_{n,k}$ and all the points we develop in the sequel are valid for this class. However, for the sake of simplicity, we consider from now on that the basis under scrutiny results from choosing the unique square root $\sigma_{n,k}$ that is lower triangular with positive diagonal entries (Cholesky decomposition).

2.2.1 Underlying System of Orthonormal Functions

We first introduce a family of functions $\phi_{n,k}$ and show that it constitutes an orthogonal basis on a certain Hilbert space. The choice of this basis can seem arbitrary at first sight, but the definition of these function will appear natural for its relationship with the functions $\psi_{n,k}$ and $\Phi_{n,k}$ that is made explicit in the sequel, and the mathematical rigor of the argument leads us to choose this apparently artificial introduction.

Definition 2. For every (n, k) in \mathcal{I} with $n \neq 0$, we define a continuous function $\phi_{n,k} : [0, 1] \rightarrow \mathbb{R}^{m \times d}$ which is zero outside its support $S_{n,k}$ and has the expressions:

$$\phi_{n,k}(t) = \begin{cases} \mathbf{f}(t)^T \cdot \mathbf{L}_{n,k} , & \text{if } l_{n,k} \leq t < m_{n,k} , \\ \mathbf{f}(t)^T \cdot \mathbf{R}_{n,k} , & \text{if } m_{n,k} \leq t < r_{n,k} . \end{cases} \quad (2.15)$$

The basis element $\phi_{0,0}$ is defined on $[0, 1]$ by

$$\phi_{0,0}(t) = \mathbf{f}(t)^T \cdot \mathbf{L}_{0,0} . \quad (2.16)$$

Remark that the definitions make apparent the fact that these two families of functions are linked for all (n, k) in \mathcal{I} through the simple relation

$$\psi'_{n,k} = \alpha \cdot \psi_{n,k} + \sqrt{\Gamma} \cdot \phi_{n,k} . \quad (2.17)$$

Moreover, this collection of functions $\phi_{n,k}$ constitutes an orthogonal basis of functions, in the following sense:

Proposition 4. *Let $L_{\mathbf{f}}^2$ be the closure of*

$$\{\mathbf{u} : [0, 1] \rightarrow \mathbb{R}^m \mid \exists \mathbf{v} \in L^2([0, 1], \mathbb{R}^d), \mathbf{u} = \mathbf{f}^T \cdot \mathbf{v}\},$$

equipped with the natural norm of $L^2([0, 1], \mathbb{R}^m)$. It is a Hilbert space, and moreover for all $0 \leq j < d$, the family of functions $c_j(\phi_{n,k})$ defined as the columns of $\phi_{n,k}$, namely

$$c_j(\phi_{n,k}) = [(\phi_{n,k})_{i,j}]_{0 \leq i < m},$$

forms a complete orthonormal basis of $L_{\mathbf{f}}^2$.

Proof. The space $L_{\mathbf{f}}^2$ is clearly a Hilbert space as a closed subspace of the larger Hilbert space $L^2([0, 1], \mathbb{R}^m)$ equipped with the standard scalar product:

$$\forall \mathbf{u}, \mathbf{v} \in L^2([0, 1], \mathbb{R}^d), \quad (\mathbf{u}, \mathbf{v}) = \int_0^1 \mathbf{u}(t)^T \cdot \mathbf{v}(t) dt.$$

We now proceed to demonstrate that the columns of $\phi_{n,k}$ form an orthonormal family which generates a dense subspace of $L_{\mathbf{f}}^2$. To this end, we define $M([0, 1], \mathbb{R}^{m \times d})$ the space of functions

$$\{\mathbf{A} : [0, 1] \rightarrow \mathbb{R}^{m \times d} \mid \forall j : 0 \leq j < d, t \mapsto [\mathbf{A}_{i,j}(t)]_{0 \leq i < m} \in L^2([0, 1], \mathbb{R}^m)\},$$

that is, the space of functions which take values in the set of $m \times d$ -matrices whose columns are in $L^2([0, 1], \mathbb{R}^m)$. This definition allows us to define the bilinear function $\mathcal{P} : M([0, 1], \mathbb{R}^{m \times d}) \times M([0, 1], \mathbb{R}^{m \times d}) \rightarrow \mathbb{R}^{d \times d}$ as

$$\mathcal{P}(\mathbf{A}, \mathbf{B}) = \int_0^1 \mathbf{A}(t)^T \cdot \mathbf{B}(t) dt \quad \text{satisfying} \quad \mathcal{P}(\mathbf{B}, \mathbf{A}) = \mathcal{P}(\mathbf{A}, \mathbf{B})^T,$$

and we observe that the columns of $\phi_{n,k}$ form an orthonormal system if and only if

$$\forall ((p,q), (n,k)) \in \mathcal{I} \times \mathcal{I}, \quad \mathcal{P}(\phi_{n,k}, \phi_{p,q}) = \int_0^1 \phi_{n,k}(t)^T \cdot \phi_{p,q}(t) dt = \delta_{p,q}^{n,k} \mathbf{I}_d,$$

where $\delta_{p,q}^{n,k}$ is the Kronecker delta function, whose value is 1 if $n = p$ and $k = q$, and 0 otherwise.

First of all, since the functions $\phi_{n,k}$ are zero outside the interval $S_{n,k}$, the matrix $\mathcal{P}(\phi_{n,k}, \phi_{p,q})$ are non-zero only if $S_{n,k} \cap S_{p,q} \neq \emptyset$. In such cases, assuming that $n \neq p$ and for example that $n < p$, we necessarily have $S_{n,k}$ strictly included in $S_{p,q}$: more precisely, $S_{n,k}$ is either included in the left-child support $S_{p+1,2q}$ or in the right-child support $S_{p+1,2q+1}$ of $S_{p,q}$. In both cases, writing the matrix $\mathcal{P}(\phi_{n,k}(t), \phi_{p,q})$ shows that it is expressed as a matrix product whose factors include $\mathcal{P}(\phi_{n,k}, \mathbf{f}^T)$. We then show that:

$$\begin{aligned} \mathcal{P}(\phi_{n,k}, \mathbf{f}^T) &= \int_0^1 \phi_{n,k}(t)^T \cdot \mathbf{f}(t)^T \\ &= \mathbf{L}_{n,k}^T \cdot \int_{l_{n,k}}^{m_{n,k}} \mathbf{f}(u) \cdot \mathbf{f}^T(u) du - \mathbf{R}_{n,k}^T \cdot \int_{m_{n,k}}^{r_{n,k}} \mathbf{f}(u) \cdot \mathbf{f}^T(u) du, \\ &= \mathbf{L}_{n,k}^T \cdot \mathbf{h}(l_{n,k}, m_{n,k}) - \mathbf{R}_{n,k}^T \cdot \mathbf{h}(m_{n,k}, r_{n,k}) \\ &= \boldsymbol{\sigma}_{n,k}^T \mathbf{g}^{-1}(m_{n,k})^T - \boldsymbol{\sigma}_{n,k}^T \mathbf{g}^{-1}(m_{n,k})^T, \end{aligned}$$

which entails that $\mathcal{P}(\phi_{n,k}, \mathbf{f}^T) = 0$ if $n < p$. If $n > p$, we use the fact that $\mathcal{P}(\phi_{n,k}, \phi_{p,q}) = \mathcal{P}(\phi_{p,q}, \phi_{n,k})^T$, and we conclude that $\mathcal{P}(\phi_{n,k}, \phi_{p,q}) = \mathbf{0}$ from the preceding case. For $n = p$, we directly compute for $n > 0$ the only non-zero term

$$\begin{aligned} \mathcal{P}(\phi_{n,k}, \phi_{n,k}) &= \mathbf{L}_{n,k}^T \cdot \int_{l_{n,k}}^{m_{n,k}} \mathbf{f}(u) \cdot \mathbf{f}^T(u) du \cdot \mathbf{L}_{n,k} \\ &\quad + \mathbf{R}_{n,k}^T \cdot \int_{m_{n,k}}^{r_{n,k}} \mathbf{f}(u) \cdot \mathbf{f}^T(u) du \cdot \mathbf{R}_{n,k}, \\ &= \boldsymbol{\sigma}_{n,k}^T \mathbf{g}^{-1}(m_{n,k})^T (\mathbf{h}(l_{n,k}, m_{n,k}))^{-1} \mathbf{g}^{-1}(m_{n,k}) \boldsymbol{\sigma}_{n,k} \\ &\quad + \boldsymbol{\sigma}_{n,k}^T \mathbf{g}^{-1}(m_{n,k})^T (\mathbf{h}(m_{n,k}, r_{n,k}))^{-1} \mathbf{g}^{-1}(m_{n,k}) \boldsymbol{\sigma}_{n,k}. \end{aligned}$$

Using the passage relationship between the symmetric functions \mathbf{h} and $\mathbf{h}_{m_{n,k}}$ given in equation (2.6), we can then write

$$\begin{aligned} \mathcal{P}(\phi_{n,k}, \phi_{n,k}) &= \boldsymbol{\sigma}_{n,k}^T (\mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k}))^{-1} \boldsymbol{\sigma}_{n,k} \\ &\quad + \boldsymbol{\sigma}_{n,k}^T (\mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k}))^{-1} \boldsymbol{\sigma}_{n,k}. \end{aligned}$$

Proposition 3 implies that $\mathbf{h}_{m_{n,k}}(l_{n,k}, m_{n,k})^{-1} + \mathbf{h}_{m_{n,k}}(m_{n,k}, r_{n,k})^{-1} = \boldsymbol{\Sigma}_{n,k}^{-1} = (\boldsymbol{\sigma}_{n,k}^{-1})^T \boldsymbol{\sigma}_{n,k}^{-1}$ which directly implies that $\mathcal{P}(\phi_{n,k}, \phi_{n,k}^T) = \mathbf{I}_d$. For $n = 0$, a computation of the exact same flavor yields that $\mathcal{P}(\phi_{0,0}, \phi_{0,0}) = \mathbf{I}_d$. Hence, we have proved that the collection of columns of $\phi_{n,k}$ forms an orthonormal family of functions in $L_{\mathbf{f}}^2$ (the definition of $\phi_{n,k}$ clearly states that its columns can be written in the form of elements of $L_{\mathbf{f}}^2$).

The proof now amounts to showing the density of the family of functions which we consider. Before showing this density property, we introduce for all (n, k) in \mathcal{I} the functions $\mathbf{P}_{n,k} : [0, 1] \rightarrow \mathbb{R}^{d \times d}$ with support on $S_{n,k}$ defined by:

$$\mathbf{P}_{n,k}(t) = \begin{cases} \mathbf{L}_{n,k} & \text{if } l_{n,k} \leq t < m_{n,k} \\ -\mathbf{R}_{n,k} & \text{if } m_{n,k} \leq t < r_{n,k} \end{cases} \quad n \neq 0 \quad \text{and} \quad \mathbf{P}_{0,0}(t) = \mathbf{L}_{0,0}$$

To show that the family of columns of $\phi_{n,k}$ is dense in $L_{\mathbf{f}}^2$ is equivalent to show that the column vectors of the matrices $\mathbf{P}_{n,k}$ seen as a function of t , are dense in $L^2([0, 1], \mathbb{R}^d)$. It is enough to show that the span of such functions contains the family of piecewise continuous \mathbb{R}^d -valued functions that are constant on $S_{n,k}$, (n, k) in \mathcal{I} . The density of the endpoints of the partition $\cup_{N \in \mathbb{N}} D_N$ entails that the latter family generates $L^2([0, 1], \mathbb{R}^d)$.

In fact, we show that the span of functions

$$V_N = \text{span} \left\{ t \mapsto c_j(\mathbf{P}_{n,k})(t) \mid 0 \leq j < d, (n, k) \in \mathcal{I}_N \right\}$$

is exactly equal to the space K_N of piecewise continuous functions from $[0, 1]$ to \mathbb{R}^d that are constant on the supports $S_{N+1,k}$, for any $(N+1, k)$ in \mathcal{I} . The fact that V_N is included in K_N is clear from the fact that the matrix-valued functions $\mathbf{P}_{N,k}$ are defined constant on the support $S_{N+1,k}$, for (N, k) in I .

We prove that K_N is included in V_N by induction on $N \leq 0$. The property is clearly true at rank $N = 0$ since $\mathbf{P}_{0,0}$ is then equal to the constant invertible matrix $\mathbf{L}_{0,0}$. Assuming the proposition true at rank $N-1$ for a given $N > 0$, let us consider a piecewise continuous function $\mathbf{c} : [0, 1] \rightarrow \mathbb{R}^d$ in K_{N-1} . Remark that for every (N, k) in \mathcal{I} , the function \mathbf{c} can only take two values on $S_{N,k}$ and can have discontinuity jumps in $m_{N,k}$: let us denote these jumps as

$$\mathbf{d}_{N,k} = \mathbf{c}(m_{N,k}^+) - \mathbf{c}(m_{N,k}^-).$$

Now, remark that for every (N, k) in \mathcal{I} , the matrix-valued functions $\mathbf{P}_{N,k}$ takes only two matrix values on $S_{N,k}$, namely $\mathbf{L}_{N,k}$ and $-\mathbf{R}_{N,k}$. From Proposition 3, we know that $\mathbf{L}_{N,k} + \mathbf{R}_{N,k} = \mathbf{M}_{N,k}$ is invertible. This fact implies that there exists vectors $\mathbf{a}_{N,k}$, for any (N, k) in \mathcal{I} , such that $\mathbf{d}_{N,k} = (\mathbf{L}_{N,k} + \mathbf{R}_{N,k})(-\mathbf{a}_{N,k})$. We then necessarily have that the function $\mathbf{c}' = \mathbf{c} + \mathbf{P}_{n,k} \cdot \mathbf{a}_{n,k}$ is piecewise constant on the supports $S_{N,k}$, (N, k) in \mathcal{I} . By recurrence hypothesis, \mathbf{c}' belongs to V_{N-1} , so that \mathbf{c} belongs to V_N , and we have proved that $K_N \subset V_N$.

Therefore, the space generated by the column vectors $P_{n,k}$ is dense in $L^2[0, 1]$, which completes the proof that the functions $t \mapsto [(\phi_{n,k}(t))_{i,j}]_{0 \leq i < m}$ form a complete orthonormal family of $L^2[0, 1]$. \square

The fact that the column functions of $\phi_{n,k}$ form a complete orthonormal system of $L^2_{\mathbf{f}}$ implies the following decomposition of the identity on $L^2_{\mathbf{f}}$:

Corollary 2. *If δ is the real Dirac delta function, we have*

$$\sum_{(n,k) \in \mathcal{I}} \phi_{n,k}(t) \cdot \phi_{n,k}^T(s) = \delta(t-s) Id_{L_{\mathbf{f}}^2}, \quad (2.18)$$

Proof. Indeed it is easy to verify that for all \mathbf{v} in $L_{\mathbf{f}}^2$, we have for all $N > 0$

$$\begin{aligned} \int_U \sum_{(n,k) \in \mathcal{I}_N} (\phi_{n,k}(t) \cdot \phi_{n,k}^T(s)) \mathbf{v}(s) ds &= \sum_{(n,k) \in \mathcal{I}_N} \phi_{n,k}(t) \cdot \mathcal{P}(\phi_{n,k}, \mathbf{v}), \\ &= \sum_{(n,k) \in \mathcal{I}_N} \sum_{p=0}^{d-1} c_p(\phi_{n,k})(c_p(\phi_{n,k}), \mathbf{v}), \end{aligned}$$

where $(c_p(\phi_{n,k}), \mathbf{v})$ denotes the inner product in $L_{\mathbf{f}}^2$ between \mathbf{v} and the p -column of $\psi_{n,k}$. Therefore, by the Parseval identity, we have in the $L_{\mathbf{f}}^2$ sense

$$\int_U \sum_{(n,k) \in \mathcal{I}} (\phi_{n,k}(t) \cdot \phi_{n,k}^T(s)) \mathbf{v}(s) ds = \mathbf{v}(t).$$

□

From now on, abusing language, we will say that the family of $\mathbb{R}^{m \times d}$ -valued functions $\phi_{n,k}$ is an orthonormal family of functions to refer to the fact that the columns of such matrices form an orthonormal set of $L_{\mathbf{f}}^2$. We now make explicit the relationship between this orthonormal basis and our functions $(\psi_{n,k})$ derived in our analysis of multi-dimensional linear diffusions.

2.2.2 Generalized Dual Operators

The integral operator \mathcal{K} The basis $\phi_{n,k}$ is of great interest in this article for its relationship to the functions $\psi_{n,k}$ that naturally arise in the decomposition of linear diffusions. Indeed, the collection $\psi_{n,k}$ can be generated from the orthonormal basis $\phi_{n,k}$ through the action of the integral operator \mathcal{K} defined on $L^2([0, 1], \mathbb{R}^m)$ into

$L^2([0, 1], \mathbb{R}^d)$ by:

$$\mathbf{u} \mapsto \mathcal{K}[\mathbf{u}] = \left\{ t \mapsto \mathbf{g}(t) \cdot \int_U \mathbb{1}_{[0,t]}(s) \mathbf{f}(s) \mathbf{u}(s) ds \right\} \quad (2.19)$$

where $U \supset [0, 1]$ is an open set and where, for any set $E \subset U$, $\mathbb{1}_E(\cdot)$ denotes the indicator function of E . Indeed, realizing that \mathcal{K} acts on $M([0, 1], \mathbb{R}^{m \times d})$ into $M([0, 1], \mathbb{R}^{d \times d})$ through

$$\forall \mathbf{A} \in M([0, 1], \mathbb{R}^{m \times d}), \quad \mathcal{K}[\mathbf{A}] = \left[\mathcal{K}[c_0(\mathbf{A})], \dots, \mathcal{K}[c_{d-1}(\mathbf{A})] \right],$$

where $c_j(\mathbf{A})$ denote the j -th \mathbb{R}^m -valued column function of \mathbf{A} , we easily see that for all (n, k) in \mathcal{I} , $0 \leq t \leq 1$:

$$\boldsymbol{\psi}_{n,k}(t) = \mathbf{g}(t) \cdot \int_0^t \mathbf{f}(s) \cdot \boldsymbol{\phi}_{n,k}(s) ds = \mathcal{K}[\boldsymbol{\phi}_{n,k}](t). \quad (2.20)$$

It is worth noticing that the introduction of the operator \mathcal{K} can be considered natural since it characterizes the centered linear diffusions \mathbf{X} through formally writing $\mathbf{X} = \mathcal{K}[\dot{\mathbf{W}}]$.

In order to exhibit a dual family of functions to the basis $\boldsymbol{\psi}_{n,k}$, we further investigate the property of the integral operator \mathcal{K} . In particular, we study the existence of an inverse operator \mathcal{D} , whose action on the orthonormal basis $\boldsymbol{\phi}_{n,k}$ will conveniently provide us with a dual basis to $\boldsymbol{\psi}_{n,k}$. Such an operator does not always exist, nevertheless, under special assumptions, it can be straightforwardly expressed as a generalized differential operator.

The differential operator \mathcal{D} Here, we make the assumptions that when $m = d$, for all t , $\mathbf{f}(t)$ is invertible in $\mathbb{R}^{d \times d}$ and that \mathbf{f} and \mathbf{f}^{-1} have continuous derivatives, which especially implies that $L_{\mathbf{f}}^2 = L^2(\mathbb{R}^d)$. In this setting, we define the space

$D_0(U, \mathbb{R}^d)$ of functions in $C_0^\infty(U, \mathbb{R}^d)$ that are zero at zero, and denote by $D'_0(U, \mathbb{R}^d)$ its dual in the space of distributions (or generalized functions). Under the assumptions just made, the operator $\mathcal{K} : D_0(U, \mathbb{R}^d) \mapsto D_0(U, \mathbb{R}^d)$ admits the differential operator $\mathcal{D} : D_0(U, \mathbb{R}^d) \mapsto D_0(U, \mathbb{R}^d)$ defined by

$$\mathbf{u} \in D_0(U, \mathbb{R}^d) \mapsto \mathcal{D}[\mathbf{u}] = \left\{ t \mapsto \mathbf{f}^{-1}(t) \frac{d}{dt} \left(\mathbf{g}^{-1}(t) \mathbf{u}(t) \right) \right\}.$$

as its inverse, that is, when restricted to $D_0(U, \mathbb{R}^d)$, we have $\mathcal{D} \circ \mathcal{K} = \mathcal{K} \circ \mathcal{D} = Id$ on $D_0(U, \mathbb{R}^d)$. The dual operators of \mathcal{K} and \mathcal{D} are expressed, for any \mathbf{u} in $D_0(U, \mathbb{R}^d)$, as

$$\begin{aligned} \mathcal{D}^*[\mathbf{u}] &= \left\{ t \mapsto -(\mathbf{g}^{-1}(t))^T \frac{d}{dt} \left((\mathbf{f}^{-1}(t))^T \mathbf{u}(t) \right) \right\}, \\ \mathcal{K}^*[\mathbf{u}] &= \left\{ t \mapsto -\mathbf{f}(t)^T \int_U \mathbb{1}_{[0,t]}(s) \mathbf{g}^T(s) \mathbf{u}(s) ds \right\}. \end{aligned}$$

They satisfy (from the properties of \mathcal{K} and \mathcal{D}) $\mathcal{D}^* \circ \mathcal{K}^* = \mathcal{K}^* \circ \mathcal{D}^* = Id$ on $D_0(U, \mathbb{R}^d)$. By dual pairing, we extend the definition of the operators \mathcal{K} , \mathcal{D} as well as their dual operators, to the space of generalized function $D'_0(U, \mathbb{R}^d)$. In details, for any distribution T in $D'_0(U, \mathbb{R}^d)$ and test function \mathbf{u} in $D_0(U, \mathbb{R}^d)$, define \mathcal{K} and \mathcal{K}^* by

$$(\mathcal{D}[T], \mathbf{u}) = (T, \mathcal{D}^*[\mathbf{u}]) \quad \text{and} \quad (\mathcal{K}[T], \mathbf{u}) = (T, \mathcal{K}^*[\mathbf{u}]),$$

and reciprocally for the dual operators \mathcal{D}^* and \mathcal{K}^* .

Candidate Dual Basis We are now in a position to use the orthonormality of the $\phi_{n,k}$ to infer a dual family to the basis of elements $\psi_{n,k}$. For any function \mathbf{u} in $L^2(U, \mathbb{R}^d)$, the generalized function $\mathcal{K}[\mathbf{u}]$ belongs to $C_0(U, \mathbb{R}^d)$, the space of continuous function which are zero at zero. We equip this space with the uniform norm and denote its topological dual $R_0(U, \mathbb{R}^d)$, the set of d -dimensional Radon measures with

$R_0(U, \mathbb{R}^d) \subset D'_0(U, \mathbb{R}^d)$. Consequently, operating in the Gelfand triple

$$C_0(U, \mathbb{R}^d) \subset L^2(U, \mathbb{R}^d) \subset R_0(U, \mathbb{R}^d), \quad (2.21)$$

we can write, for any function \mathbf{u}, \mathbf{v} in $L^2(U, \mathbb{R}^d) \subset R_0(U, \mathbb{R}^d)$

$$(\mathbf{u}, \mathbf{v}) = \left((\mathcal{D} \circ \mathcal{K})[\mathbf{u}], \mathbf{v} \right) = (\mathcal{K}[\mathbf{u}], \mathcal{D}^*[\mathbf{v}])$$

The first equality stems from the fact that, when \mathcal{K} and \mathcal{D} are seen as generalized functions, they are still inverse of each other, so that in particular $\mathcal{D} \circ \mathcal{K} = Id$ on $L^2(U, \mathbb{R}^d)$. The dual pairing associated with the Gelfand triple (2.21) entails the second equality where \mathcal{D}^* is the generalized operator defined on $D'_0(U, \mathbb{R}^d)$ and where $\mathcal{D}^*[\mathbf{v}]$ is in $R_0(U, \mathbb{R}^d)$.

As a consequence, defining the functions $\boldsymbol{\delta}_{n,k}$ in $R_0(U, \mathbb{R}^{d \times d})$, the $d \times d$ -dimensional space of Radon measures, by

$$\boldsymbol{\delta}_{n,k} = \mathcal{D}^*(\phi_{n,k}) = [\mathcal{D}^*[c_1(\phi_{n,k})], \dots, \mathcal{D}^*[c_d((\phi_{n,k}))]]$$

provides us with a family of $d \times d$ -generalized functions which are dual to the family $\psi_{n,k}$ in the sense that, for all $((n, k), (p, q))$ in $\mathcal{I} \times \mathcal{I}$, we have

$$\mathcal{P}(\boldsymbol{\delta}_{n,k}, \psi_{p,q}) = \delta_{p,q}^{n,k} \mathbf{I}_d,$$

where the definition of \mathcal{P} has been extended through dual pairing: given any \mathbf{A} in $R_0(U, \mathbb{R}^{m \times d})$ and any \mathbf{B} in $C_0(U, \mathbb{R}^{m \times d})$, we have

$$\mathcal{P}(\mathbf{A}, \mathbf{B}) = \left[(c_i(\mathbf{A}), c_j(\mathbf{B})) \right]_{0 \leq i, j < d}$$

with $(c_i(\mathbf{A}), c_j(\mathbf{B}))$ denoting the dual pairing between the i -th column of \mathbf{A} taking value in $R_0(U, \mathbb{R}^d)$ and the j -th column of \mathbf{B} taking value in $C_0(U, \mathbb{R}^d)$. Under the favorable hypothesis of this section, the $d \times d$ -generalized functions $\delta_{n,k}$ can actually be easily computed, since considering the definition of $\phi_{n,k}$ shows that the functions $(\mathbf{f}^{-1})^T \cdot \phi_{n,k}$ have support $S_{n,k}$ and are constant on $S_{n+1,2k}$ and $S_{n+1,2k+1}$ in $\mathbb{R}^{d \times d}$. Only the discontinuous jumps in $l_{n,k}$, $m_{n,k}$ and $r_{n,k}$ intervene, leading to express for (n, k) in \mathcal{I} , $n \neq 0$

$$\delta_{n,k}(t) = (\mathbf{g}(t)^{-1})^T \cdot (\mathbf{M}_{n,k} \delta(t - m_{n,k}) - (\mathbf{L}_{n,k} \delta(t - l_{n,k}) + \mathbf{R}_{n,k} \delta(t - r_{n,k})))$$

and $\delta_{0,0}(t) = (\mathbf{g}(t)^{-1})^T \cdot \mathbf{L}_{0,0}$, where $\delta(\cdot)$ denotes the standard Dirac delta function (centered in 0). These functions can be extended to the general setting of the article since its expressions do not involve the assumptions made on the invertibility and smoothness of $\mathbf{f}(t)$. We now show that these functions, when defined in the general setting still provide a dual basis of the functions $\psi_{n,k}$.

2.2.3 Dual Basis of Generalized Functions

The expression of the basis $\delta_{n,k}$ that has been found under favorable assumptions makes no explicit reference to these assumptions. It suggest to define functions $\delta_{n,k}$ formally as linear combinations of Dirac delta functions acting by duality on $C_0(U, \mathbb{R}^{d \times d})$:

Definition 3. For (n, k) in \mathcal{I} , the family of generalized functions $\delta_{n,k}$ in $R_0(U, \mathbb{R}^{d \times d})$ is given $n \neq 0$ by

$$\delta_{n,k}(t) = (\mathbf{g}(t)^{-1})^T \cdot (\mathbf{M}_{n,k} \delta(t - m_{n,k}) - (\mathbf{L}_{n,k} \delta(t - l_{n,k}) + \mathbf{R}_{n,k} \delta(t - r_{n,k}))) , ,$$

and $\delta_{0,0}(t) = (\mathbf{g}(t)^{-1})^T \cdot \mathbf{L}_{0,0}$, where δ is the standard Dirac distribution.

Notice that the basis $\delta_{n,k}$ is defined for the open set U . For the sake of consistency, we extend the definition of the families $\psi_{n,k}$ and $\phi_{n,k}$ on U by setting them to zero on $U \setminus [0, 1]$, except for $\psi_{0,0}$ which is continued for $t > 1$ by a continuous function \mathbf{c} that is compactly supported in $[1, a)$ for a given a in U , $a > 1$ and satisfies $\mathbf{c}(1) = \psi_{0,0}(1)$. We can now formulate:

Proposition 5. *Given the dual pairing in $C_0(U) \subset L^2(U) \subset R(U)$ where U is a bounded open set of \mathbb{R} containing $[0, 1]$, the family of continuous functions $\psi_{n,k}$ in $C_0(U)$ admits as a dual family in $R(U)$, the set of distributions $\delta_{n,k}$.*

Proof. We have to demonstrate that, for all $((n, k), (p, q))$ in $\mathcal{I} \times \mathcal{I}$,

$$\mathcal{P}(\delta_{p,q}, \psi_{n,k}) = \delta_{p,q}^{n,k} \mathbf{I}_d.$$

Suppose first, $n, p > 0$. If $p < n$, $\mathcal{P}(\delta_{n,k}, \psi_{p,q})$ can only be non-zero if the support $S_{p,q}$ is strictly included in $S_{n,k}$. We then have

$$\begin{aligned} \mathcal{P}(\delta_{p,q}, \psi_{n,k}) &= \mathbf{M}_{p,q}^T \mathbf{g}^{-1}(m_{p,q}) \psi_{n,k}(m_{p,q}) \\ &\quad - (\mathbf{L}_{p,q}^T \mathbf{g}^{-1}(l_{p,q}) \psi_{n,k}(l_{p,q}) + \mathbf{R}_{p,q}^T \mathbf{g}^{-1}(r_{p,q}) \psi_{n,k}(r_{p,q})) . \end{aligned}$$

Assume $S_{p,q}$ is to the left of $m_{n,k}$, that is, $S_{p,q}$ is a left-child of $S_{n,k}$ in the nested binary tree of supports and write

$$\mathcal{P}(\delta_{p,q}, \psi_{n,k}) = \left(\mathbf{M}_{p,q}^T \mathbf{h}(l_{n,k}, m_{p,q}) - \mathbf{L}_{p,q}^T \mathbf{h}(l_{n,k}, l_{p,q}) - \mathbf{R}_{p,q}^T \mathbf{h}(l_{n,k}, r_{p,q}) \right) \mathbf{L}_{n,k}.$$

Using the fact that $\mathbf{M}_{p,q} = \mathbf{L}_{p,q} + \mathbf{R}_{p,q}$ and that the function $h(x, y)$, as any integral between x and y , satisfies the chain rule $\mathbf{h}(x, y) = \mathbf{h}(x, z) + \mathbf{h}(z, y)$ for all (x, y, z) ,

we obtain:

$$\begin{aligned}
\mathcal{P}(\boldsymbol{\delta}_{p,q}, \boldsymbol{\psi}_{n,k}) &= \left(-\mathbf{L}_{p,q}^T (\mathbf{h}(l_{n,k}, m_{p,q}) - \mathbf{h}(l_{n,k}, l_{p,q})) \right. \\
&\quad \left. + \mathbf{R}_{p,q}^T (h(l_{n,k}, m_{p,q}) - h(l_{n,k}, r_{p,q})) \right) \mathbf{L}_{n,k} \\
&= \left(-\mathbf{L}_{p,q}^T \mathbf{h}(l_{p,q}, m_{p,q}) + \mathbf{R}_{p,q}^T h(r_{p,q}, m_{p,q}) \right) \mathbf{L}_{n,k} \\
&= \left(-\boldsymbol{\sigma}_{p,q}^T (\mathbf{g}^{-1}(m_{p,q}))^T (\mathbf{h}(l_{p,q}, m_{p,q})^{-1})^T \cdot \mathbf{h}(l_{p,q}, m_{p,q}) \right. \\
&\quad \left. + \boldsymbol{\sigma}_{p,q}^T (\mathbf{g}^{-1}(m_{p,q}))^T (\mathbf{h}(m_{p,q}, r_{p,q})^{-1})^T \cdot \mathbf{h}(m_{p,q}, r_{p,q}) \right) \cdot \mathbf{L}_{n,k} \\
&= \mathbf{0}
\end{aligned}$$

The same result is true if $S_{p,q}$ is a right-child of $S_{n,k}$ in the nested binary tree of supports. If $p = n$, necessarily the only non-zero term is for $q = p$, i.e.

$$\begin{aligned}
\mathcal{P}(\boldsymbol{\delta}_{p,q}, \boldsymbol{\psi}_{n,k}) &= \mathbf{M}_{n,k}^T \mathbf{g}^{-1}(m_{n,k}) \boldsymbol{\psi}(m_{n,k}) \\
&= \mathbf{M}_{n,k}^T \mathbf{h}(l_{n,k}, m_{n,k}) \mathbf{L}_{n,k} \\
&= \boldsymbol{\sigma}_{p,q}^{-1} \mathbf{g}(m_{n,k}) \mathbf{h}(l_{n,k}, m_{n,k}) \mathbf{h}(l_{n,k}, m_{n,k})^{-1} \mathbf{g}^{-1}(m_{n,k}) \boldsymbol{\sigma}_{p,q} \\
&= \mathbf{I}_d.
\end{aligned}$$

If $p > n$, $\mathcal{P}(\boldsymbol{\delta}_{n,k}, \boldsymbol{\psi}_{p,q})$ can only be non-zero if the support $S_{n,k}$ is included in $S_{p,q}$, but then $\boldsymbol{\psi}_{n,k}$ is zero in $l_{p,q}, m_{p,q}, r_{p,q}$ so that $\mathcal{P}(\boldsymbol{\delta}_{n,k}, \boldsymbol{\psi}_{p,q}) = \mathbf{0}$.

Otherwise, if $n = 0$ and $p > 0$, we directly have

$$\begin{aligned}
\mathcal{P}(\boldsymbol{\delta}_{p,q}, \boldsymbol{\psi}_{0,0}) &= \mathbf{M}_{p,q}^T \mathbf{g}^{-1}(m_{p,q}) \boldsymbol{\psi}_{0,0}(m_{p,q}) \\
&\quad - (\mathbf{L}_{p,q}^T \mathbf{g}^{-1}(l_{p,q}) \boldsymbol{\psi}_{0,0}(l_{p,q}) + \mathbf{R}_{p,q}^T \mathbf{g}^{-1}(r_{p,q}) \boldsymbol{\psi}_{0,0}(r_{p,q})) , \\
&= \left(-\mathbf{L}_{p,q}^T \mathbf{h}(l_{p,q}, m_{p,q}) + \mathbf{R}_{p,q}^T \mathbf{h}(m_{p,q}, r_{p,q}) \right) \mathbf{L}_{0,0} , \\
&= \left(-\boldsymbol{\sigma}_{p,q}^T \mathbf{g}^{-1}(m_{p,q})^T + \boldsymbol{\sigma}_{p,q}^T \mathbf{g}^{-1}(m_{p,q})^T \right) \mathbf{L}_{n,k} , \\
&= \mathbf{0} .
\end{aligned}$$

Finally, if $p = 0$, given the simple form of $\delta_{0,0}$ with a single Dirac delta function centered in $r_{0,0}$, we clearly have $\mathcal{P}(\delta_{0,0}, \psi_{n,k}) = 0$, and if $n > 0$

$$\begin{aligned}
\mathcal{P}(\delta_{0,0}, \psi_{0,0}) &= \mathbf{L}_{0,0}^T \mathbf{h}(l_{0,0}, r_{0,0}) \mathbf{L}_{0,0}, \\
&= \boldsymbol{\sigma}_{0,0}^T (\mathbf{g}^{-1}(r_{0,0}))^T \mathbf{L}_{0,0}, \\
&= \boldsymbol{\sigma}_{0,0}^T (\mathbf{g}^{-1}(r_{0,0}))^T (\mathbf{h}(l_{0,0}, r_{0,0}) \mathbf{L}_{0,0})^{-1} \mathbf{g}^{-1}(r_{0,0}) \boldsymbol{\sigma}_{0,0}, \\
&= \boldsymbol{\sigma}_{0,0}^T \mathbf{C}_{r_{0,0}}^{-1} \boldsymbol{\sigma}_{0,0},
\end{aligned}$$

and using the fact that (by definition of $\boldsymbol{\sigma}_{0,0}$) we have $\boldsymbol{\sigma}_{0,0} \cdot \boldsymbol{\sigma}_{0,0}^T = \mathbf{C}_{r_{0,0}}$, this last expression is equal to:

$$\mathcal{P}(\delta_{0,0}, \psi_{0,0}) = \boldsymbol{\sigma}_{0,0}^T (\boldsymbol{\sigma}_{0,0}^T)^{-1} \cdot (\boldsymbol{\sigma}_{0,0})^{-1} \boldsymbol{\sigma}_{0,0} = I_d,$$

which completes the proof. □

This proposition directly implies the main result of the section:

Theorem 2. *The collection of functions $(\psi_{n,k} ; (n,k) \in \mathcal{I})$ constitute a Schauder basis of functions on $C_0(U, \mathbb{R}^d)$.*

This theorem provides us with a complementary view of stochastic processes: in addition to the standard sample paths view, the dual structure allows us to see linear diffusions as coefficients on the computed basis. This duality is developed in the sequel.

2.3 Inductive Construction of Linear Diffusions

Up to this point, we have rigorously defined the dual spaces of sample paths ${}_x\Omega'$ and coefficients ${}_x\Omega'$. Through the use of the Schauder basis $\psi_{n,k}$ and its dual family of generalized functions $\delta_{n,k}$, we have defined inverse measurable bijections Ψ and

Δ which transform one space into the other. In doing so, we have unraveled the fundamental role played by the underlying orthonormal basis $\phi_{n,k}$. We now turn to use this framework to formulate a path-wise construction of the linear diffusions in the exact same flavor as the Levy-Ciesielski construction of the Wiener process.

2.3.1 The Space of Sample Paths

The Construction Application The Schauder basis of functions with compact supports constructed above allows us to define functions by considering the coefficients on this basis, which constitute sequences of real numbers in the space:

$${}_{\xi}\Omega = \left\{ \xi = \{\xi_{n,k}\}_{\mathcal{I}}; \forall (n,k) \in \mathcal{I}, \xi_{n,k} \in \mathbb{R}^d \right\} = (\mathbb{R}^d)^{\mathcal{I}}.$$

We equip ${}_{\xi}\Omega$ with the uniform norm $\|\xi\|_{\infty} = \sup_{(n,k) \in \mathcal{I}} |\xi_{n,k}|$, where we write $|\xi_{n,k}| = \sup_{0 \leq i < d} |(\xi_{n,k})_i|$. We denote by $\mathcal{B}({}_{\xi}\Omega)$ the Borelian sets of the topology induced by the uniform norm and we recall that $C({}_{\xi}\Omega)$, the cylinder sets of ${}_{\xi}\Omega$, form a generative family of Borelian sets. Note that not every sequence of coefficients provides a continuous function, and one needs to assume a certain decrease in the coefficients to get convergence. A sufficient condition to obtain convergent sequences is to consider coefficients in the space:

$${}_{\xi}\Omega' = \left\{ \xi \in {}_{\xi}\Omega \mid \exists \delta \in (0, 1), \exists N \in \mathbb{N}, \forall (n,k) \in \mathcal{I} \setminus \mathcal{I}_N, |\xi_{n,k}| < 2^{\frac{n\delta}{2}} \right\}$$

This set is clearly a Borelian set of ${}_{\xi}\Omega$, since it can be written as a countable intersection and union of cylinders, namely, by denoting \mathcal{J} the set of finite subset of \mathbb{N} and $\delta_p = 1 - 1/p$, $p > 1$,

$${}_{\xi}\Omega' = \bigcup_{p>1} \bigcup_{J \in \mathcal{J}} \bigcap_{n \in \mathbb{N} \setminus J} \left\{ \xi \in {}_{\xi}\Omega \mid \max_{0 \leq k < 2^{n-1}} |\xi_{n,k}| < 2^{\frac{n\delta_p}{2}} \right\}.$$

It is also easy to verify that ${}_{\xi}\Omega'$ forms a vectorial subspace of ${}_{\xi}\Omega$.

After these definitions, we are in position to introduce the following useful function:

Definition 4. We denote by Ψ^N the *partial construction application*:

$$\Psi^N = \begin{cases} {}_{\xi}\Omega & \longrightarrow C_0([0, 1], \mathbb{R}^d) \\ \xi & \longmapsto \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k}(t) \cdot \xi_{n,k} . \end{cases}$$

where the $C_0([0, 1], \mathbb{R}^d)$ is the d -dimensional Wiener space, which is complete under the uniform norm $\|\mathbf{x}\|_{\infty} = \sup_{0 \leq t \leq 1} |\mathbf{x}(t)|$.

This sequence of partial construction applications is shown to converge to the construction application in the following:

Proposition 6. *For every ξ in ${}_{\xi}\Omega'$, $\Psi^N(\xi)$ converges uniformly toward a continuous function in $C_0([0, 1], \mathbb{R}^d)$. We will denote this function $\Psi(\xi)$, defined as:*

$$\Psi : \begin{cases} {}_{\xi}\Omega' & \longrightarrow C_0([0, 1], \mathbb{R}^d) \\ \xi & \longmapsto \sum_{(n,k) \in \mathcal{I}} \psi_{n,k}(t) \cdot \xi_{n,k} \end{cases}$$

and this application will be referred to as the construction application.

The image of this function constitutes a subset of the Wiener space of continuous functions $C_0([0, 1], \mathbb{R}^d)$. Let us now define the vectorial subspace ${}_x\Omega' = \Psi({}_{\xi}\Omega')$ of $C_0([0, 1], \mathbb{R}^d)$ so that Ψ appears as a bijection.

It is important to realize that, in the multidimensional case, the space ${}_x\Omega'$ depends on $\mathbf{\Gamma}$ and $\boldsymbol{\alpha}$ in a non-trivial way. For instance, assuming $\boldsymbol{\alpha} = 0$, the space ${}_x\Omega'$ depends obviously crucially on the rank of $\mathbf{\Gamma}$. To fix this idea, for a given constant $\sqrt{\mathbf{\Gamma}(\mathbf{t})} = [0, 0 \dots 1]^T$ in $\mathbb{R}^{d \times 1}$, we expect the space ${}_x\Omega'$ to only include sample paths of $C_0([0, 1], \mathbb{R}^d)$ for which the $n-1$ first components are constant. Obviously, a process with such sample paths is degenerate in the sense that its covariance matrix is not

invertible.

Yet, if we additionally relax the hypothesis that $\alpha \neq 0$, the space ${}_x\Omega'$ can be dramatically altered: if we take

$$\alpha(t) = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & 1 & \\ & & & & 0 \end{bmatrix}$$

the space ${}_x\Omega'$ will represent the sample space of the $d-1$ -integrated Wiener process, a non-degenerate d -dimensional process we fully develop in the example section.

However, the situation is much simpler in the one-dimensional case: because the uniform convergence of the sample paths is preserved as long as α is continuous and Γ is non-zero, the definition ${}_x\Omega'$ does not depend on α or Γ . Moreover, in this case, the space ${}_x\Omega'$ is large enough to contain reasonably regular functions.

Proposition 7. *In the one-dimensional case, the space ${}_x\Omega'$ contains the space of uniformly Hölder continuous functions H defined as*

$$H = \left\{ x \in C[0, 1] \mid \exists \delta > 0, \sup_{0 \leq s, t \leq 1} \frac{|x(t) - x(s)|}{|t - s|^\delta} < +\infty \right\}.$$

Remark 1. This point can be seen as a direct consequence of the characterization of the local Hölder exponent of a continuous real function in term of the asymptotic behavior of its coefficients in the decomposition on the Schauder basis [51].

Proof. To underline that we place ourselves in the one-dimensional case, we drop the bold notations that indicates multidimensional quantities. Suppose x is uniformly Hölder continuous for a given $\delta > 0$, there always exist ξ such that $\Psi^N(\xi)$ coincides with x on D_N : it is enough to take ξ such that for all (n, k) in \mathcal{I}_N , $\xi_{n,k} = (\delta_{n,k}, x)$.

We can further write for $n > 0$

$$\begin{aligned} (x, \delta_{n,k}) &= M_{n,k} \frac{x(m_{n,k})}{g(m_{n,k})} - \left(L_{n,k} \frac{x(l_{n,k})}{g_\alpha(l_{n,k})} + R_{n,k} \frac{x(r_{n,k})}{g(r_{n,k})} \right), \\ &= L_{n,k} \left(\frac{x(m_{n,k})}{g(m_{n,k})} - \frac{x(l_{n,k})}{g(l_{n,k})} \right) + R_{n,k} \left(\frac{x(m_{n,k})}{g(m_{n,k})} - \frac{x(r_{n,k})}{g(r_{n,k})} \right). \end{aligned}$$

For a given function α , posing $N_\alpha = \frac{\sup_{0 \leq t \leq 1} f_\alpha(t)}{\inf_{0 \leq t \leq 1} f_\alpha^2(t)}$, we have

$${}_\alpha M_{n,k} \leq N_\alpha 2^{\frac{n+1}{2}}, \quad {}_\alpha L_{n,k} \leq N_\alpha 2^{\frac{n-1}{2}}, \quad {}_\alpha R_{n,k} \leq N_\alpha 2^{\frac{n-1}{2}}.$$

Moreover, if α is in H , it is straightforward to see that g_α has a continuous derivative. Then, since x is δ -Hölder, for any $\epsilon \geq 0$, there exists $C > 0$ such that $|t - s| \leq \epsilon$ entails

$$\left| \frac{x(t)}{g(t)} - \frac{x(s)}{g(s)} \right| \leq C \epsilon^\delta,$$

from which we directly deduce

$$|\xi_{n,k}| \leq \frac{N_\alpha C}{\sqrt{2}} 2^{n(\frac{1}{2}-2\delta)}.$$

This demonstrates that $\{\xi_{n,k}\}$ belongs to ${}_\xi \Omega'$ and ends the proof. \square

We equip the space ${}_x \Omega'$ with the topology induced by the uniform norm on $C_0([0, 1], \mathbb{R}^d)$. As usual, we denote $\mathcal{B}({}_x \Omega')$ the corresponding Borelian sets.

Proposition 8. *The function $\Psi : ({}_\xi \Omega', \mathcal{B}({}_\xi \Omega')) \rightarrow ({}_x \Omega', \mathcal{B}({}_x \Omega'))$ is a bounded continuous bijection.*

We thus have a continuous bijection mapping the coefficients onto the sample paths. We now turn to study its inverse, the coefficient application, mapping sample paths on coefficients over the Schauder basis.

The Coefficient Application Reciprocally, we introduce and study the properties of the following function:

Definition 5. We call *coefficient application* and denote by Ξ the function defined by:

$$\Xi : \begin{cases} C_0([0, 1], \mathbb{R}^d) & \longrightarrow \xi\Omega = (\mathbb{R}^d)^{\mathcal{I}} \\ \mathbf{x} & \longmapsto \Delta(\mathbf{x}) = \{\Delta(\mathbf{x})\}_{(n,k) \in \mathcal{I}}, \quad \{\Delta(\mathbf{x})\}_{n,k} = \mathcal{P}(\delta_{n,k}, \mathbf{x}) . \end{cases} \quad (2.22)$$

Should a function x admit a uniformly convergent decomposition in terms on the basis of elements $\psi_{n,k}$, the function Δ gives its coefficients in such a representation. More precisely, we have:

Theorem 3. *The function $\Delta : ({}_x\Omega', \mathcal{B}({}_x\Omega')) \rightarrow (\xi\Omega', \mathcal{B}(\xi\Omega'))$ is a measurable linear bijection whose inverse is $\Psi = \Delta^{-1}$.*

2.3.2 Finite-Dimensional Approximations

Considering the infinite dimensional subspace ${}_x\Omega'$ of $C_0([0, 1], \mathbb{R}^d)$, let us introduce the equivalence relation \sim_N as

$$\mathbf{x} \sim_N \mathbf{y} \iff \forall t \in D_N, \quad \mathbf{x}(t) = \mathbf{y}(t) .$$

We can use the functions Ψ to carry the structure of \sim_N on the infinite-dimensional space of coefficients $\xi\Omega'$:

$$\xi \sim_N \eta \iff \Psi(\xi) \sim_N \Psi(\eta) \iff \forall (n, k) \in \mathcal{I}_N, \quad \xi_{n,k} = \eta_{n,k} ,$$

which clearly entails that $\mathbf{x} \sim_N \mathbf{y}$ if and only if $\Delta(\mathbf{x}) \sim_N \Delta(\mathbf{y})$. We denote the sets of equivalence classes of ${}_x\Omega' / \sim_N$ and $\xi\Omega' / \sim_N$ as ${}_x\Omega_N$ and $\xi\Omega_N$ respectively. Both

these sets are isomorphic since ${}_x\Omega_N = (\mathbb{R}^d)^{\mathcal{I}} = {}_\xi\Omega_N$. For every $N > 0$, we define the finite-dimensional operators $\Psi_N = {}_x\mathbf{i}_N \circ \Psi \circ {}_\xi\mathbf{p}_N$ and $\Delta_N = {}_\xi\mathbf{i}_N \circ \Delta \circ {}_x\mathbf{p}_N$, with the help of the canonical projections ${}_\xi\mathbf{p}_N : {}_\xi\Omega' \rightarrow {}_\xi\Omega_N$, ${}_x\mathbf{p}_N : {}_x\Omega' \rightarrow {}_x\Omega_N$ and the inclusion map ${}_\xi\mathbf{i}_N : {}_\xi\Omega_N \rightarrow {}_\xi\Omega'$, ${}_x\mathbf{i}_N : {}_x\Omega_N \rightarrow {}_x\Omega'$.

The results of the preceding sections straightforwardly extend on the equivalence classes, and in particular we see that the function $\Psi_N : {}_\xi\Omega_N \rightarrow {}_x\Omega_N$ and $\Delta_N : {}_x\Omega_N \rightarrow {}_\xi\Omega_N$ are linear finite-dimensional bijections satisfying $\Psi_N = \Delta_N^{-1}$. We write $\mathbf{e} = \{\mathbf{e}_{p,q}\}_{(p,q) \in \mathcal{I}}$ (resp. $\mathbf{f} = \{\mathbf{f}_{p,q}\}_{(p,q) \in \mathcal{I}}$) the canonical basis of ${}_\xi\Omega_N$ (resp. ${}_x\Omega_N$) when listed in the recursive dyadic order. In these bases, the matrices Ψ_N and Δ_N are lower block-triangular. Indeed, denoting Ψ_N in the natural basis $\mathbf{e} = \{\mathbf{e}_{p,q}\}_{(p,q) \in \mathcal{I}}$ and $\mathbf{f} = \{\mathbf{f}_{p,q}\}_{(p,q) \in \mathcal{I}}$ by

$$\Psi_N = [\psi_{n,k}(m_{i,j})] = [\Psi_{n,k}^{i,j}],$$

where $\Psi_{n,k}^{i,j}$ is a $d \times d$ matrix, the structure of the nested-support $S_{n,k}$ entails the block-triangular structure (where only possibly non-zero coefficients are written):

$${}_\alpha\Psi_N = \begin{bmatrix} \psi_{0,0}^{0,0} & & & & \\ \psi_{0,0}^{1,0} & \psi_{1,0}^{1,0} & & & \\ \psi_{0,0}^{2,0} & \psi_{1,0}^{2,0} & \psi_{2,0}^{2,0} & & \\ \psi_{0,0}^{2,1} & \psi_{1,0}^{2,1} & & \psi_{2,1}^{2,1} & \\ \psi_{0,0}^{3,0} & \psi_{1,0}^{3,0} & \psi_{2,0}^{3,0} & \psi_{3,0}^{3,0} & \\ \psi_{0,0}^{3,1} & \psi_{1,0}^{3,1} & \psi_{2,0}^{3,1} & & \psi_{3,1}^{3,1} \\ \psi_{0,0}^{3,2} & \psi_{1,0}^{3,2} & & \psi_{2,1}^{3,2} & \psi_{3,2}^{3,2} \\ \psi_{0,0}^{3,3} & \psi_{1,0}^{3,3} & & \psi_{2,1}^{3,3} & \psi_{3,3}^{3,3} \\ \vdots & & & & \ddots \end{bmatrix}.$$

Similarly, for the matrix representation of Δ_N in the natural basis $e_{n,k}$ and $f_{i,j}$

$$\Delta_N = \begin{bmatrix} \Delta_{i,j}^{n,k} \end{bmatrix}$$

proves to have the following triangular form:

$$\Delta_N = \begin{bmatrix} \mathbf{g}^{-1}(t_{0,0})^T \mathbf{M}_{0,0} & & & & \\ -\mathbf{g}^{-1}(t_{0,0})^T \mathbf{R}_{1,0} & \mathbf{g}^{-1}(t_{1,0})^T \mathbf{M}_{1,0} & & & \\ & -\mathbf{g}^{-1}(t_{0,0})^T \mathbf{R}_{2,0} & \mathbf{g}^{-1}(t_{1,0})^T \mathbf{M}_{2,0} & & \\ -\mathbf{g}^{-1}(t_{0,0})^T \mathbf{R}_{2,1} & -\mathbf{g}^{-1}(t_{1,0})^T \mathbf{L}_{2,1} & & \mathbf{g}^{-1}(t_{2,1})^T \mathbf{M}_{2,1} & \\ \vdots & & & & \ddots \end{bmatrix}.$$

By Proposition 5, the duality property simply reads for all $0 \leq n < N$ and $0 \leq k < 2^{n-1}$, $0 \leq p < N$ and $0 \leq p < 2^{q-1}$

$$\mathcal{P}(\delta_{p,q}, \psi_{n,k}) = \sum_{(n,k) \in \mathcal{I}_N} \Delta_{i,j}^{p,q} \cdot \Psi_{n,k}^{i,j} = \delta_{n,k}^{p,q} \mathbf{I}_d.$$

that is, $\Delta_N \cdot \Psi_N = Id_{\xi_{\Omega_N}}$. But because, we are now in a finite-dimensional setting, we also have $\Psi_N \cdot \Delta_N = Id_{x_{\Omega_N}}$:

$$\delta_{k,l}^{i,j} \mathbf{I}_d = \sum_{(p,q) \in \mathcal{I}_N} \Psi_{p,q}^{i,j} \cdot \Delta_{k,l}^{p,q}.$$

Realizing that $\delta_{k,l}^{i,j} \mathbf{I}_d$ represents the class of functions \mathbf{x} in ${}_x\Omega'$ whose value are zero on every dyadic points of D_N except for $\mathbf{x}(l2^k) = \mathbf{I}_d$, $\{\Delta_{k,l}^{p,q}\}_{(p,q) \in \mathcal{I}_N}$ appear as the coefficients of the decomposition of such functions in the bases $\psi_{p,q}$ for (p,q) in \mathcal{I}_N .

Denoting $\Xi = \{\Xi_{n,k}\}_{(n,k) \in I}$, a set of independent Gaussian variables of law $\mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ on $(\Omega, \mathcal{F}, \mathbf{P})$, and for all $N > 0$, we form the finite dimensional Gauss-

Markov vector $[\mathbf{X}_{i,j}^N]_{(i,j) \in \mathcal{I}_N}$ as

$$\mathbf{X}_{i,j}^N = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k}(m_{i,j}) \cdot \Xi_{n,k},$$

which, from Corollary 1, has the same law as $[\mathbf{X}_t]_{t \in D_N}$, the finite-dimensional random vector obtained from sampling \mathbf{X} on D_N (modulo a permutation on the indices). We then prove the following lemma that sheds light on the meaning of the construction:

Lemma 2. *The Cholesky decomposition of the finite-dimensional covariance block matrix Σ_N is given by $\Sigma_N = \Psi_N \cdot \Psi_N^T$.*

Proof. For every $0 \leq t, s \leq 1$, we compute the covariance of the finite-dimensional process \mathbf{X}^N as

$$\mathbf{C}_N(t, s) = \mathbb{E} [\mathbf{X}_t^N \cdot (\mathbf{X}_s^N)^T] = \sum_{n=0}^N \sum_{0 \leq k < 2^{n-1}} \psi_{n,k}(t) \cdot (\psi_{n,k}(s))^T,$$

From there, we write the finite-dimensional covariance block matrix Σ_N in the recursively ordered basis $\mathbf{f}_{i,j}$ for $0 \leq i \leq N$, $0 \leq j < 2^{i-1}$, as

$$[\Sigma_N]_{k,l}^{i,j} = \mathbf{C}_N(m_{i,j}, m_{k,l}) = \sum_{n=0}^N \sum_{0 \leq k < 2^{n-1}} \Psi_{n,k}^{i,j} \cdot \Psi_{n,k}^{k,l}.$$

We already established that the matrix Ψ_N was triangular with positive diagonal coefficient, which assures that the preceding equality provides us with the Cholesky decomposition of Σ . \square

In the finite-dimensional case, the inverse covariance or potential matrix is a well-defined quantity and we straightforwardly have the following corollary:

Corollary 3. *The Cholesky decomposition of the finite-dimensional inverse covariance matrix Σ_N^{-1} is given by $\Sigma_N^{-1} = \Delta_N^T \cdot \Delta_N$.*

Proof. The result stems from the equalities: $\Sigma_N^{-1} = (\Psi_N \cdot \Psi_N^T)^{-1} = (\Psi_N^{-1})^T \cdot \Psi_N^{-1} = \Delta_N^T \cdot \Delta_N$. \square

2.3.3 The Lévy-Ciesielski Expansion

We now show that asymptotically, the basis $\psi_{n,k}$ allows us to faithfully build the linear diffusion from which we have derived its expression. In this perspective we consider $\Xi = \{\Xi_{n,k}\}_{(n,k) \in \mathcal{I}}$, a set of independent Gaussian variables of law $\mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ on $(\Omega, \mathcal{F}, \mathbf{P})$, and for all $N > 0$, we form the finite dimensional continuous Gaussian process \mathbf{Z}^N , defined for $0 \leq t \leq 1$ by

$$\mathbf{X}_t^N = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k}(t) \cdot \Xi_{n,k},$$

which, from the result of Theorem 1, has the same law $\mathbf{Z}_t^N = \mathbb{E}[X_t | \mathcal{F}_N]$. We prove the following lemma:

Lemma 3. *The sequence of processes \mathbf{X}^N almost surely converges towards a continuous Gaussian process denoted \mathbf{X}^∞ .*

Proof. For all fixed $N > 0$ and for any ω in Ω , we know that $t \mapsto \mathbf{X}_t^N(\omega)$ is continuous. Moreover, we have established, that for every ξ in ${}_\xi\Omega'$, $\mathbf{X}^N(\xi)$ converges uniformly in t toward a continuous limit denoted $\mathbf{X}^N(\xi)$. Therefore, in order to prove that $\lim_{N \rightarrow \infty} \mathbf{X}^N$ defines almost surely a process \mathbf{X} with continuous paths, it is sufficient to show that $\mathbf{P}_\xi({}_\xi\Omega') = 1$, where $\mathbf{P}_\xi = \mathbf{P}_{\Xi^{-1}}$ is the Ξ -induced measure on ${}_\xi\Omega$, which stems from a classical Borel-Cantelli argument. For ξ a random variable of normal law $\mathcal{N}(0, 1)$, and $a > 0$, we have

$$\mathbf{P}(|\xi| > a) = \sqrt{\frac{2}{\pi}} \int_a^\infty e^{-u^2/2} du \leq \sqrt{\frac{2}{\pi}} \int_a^\infty \frac{u}{a} e^{-u^2/2} du = \sqrt{\frac{2}{\pi}} \frac{e^{-a^2/2}}{a}.$$

Then, for any $\delta > 0$

$$\mathbf{P}_\xi \left(\max_{0 \leq k < 2^{n-1}} |\boldsymbol{\xi}_{n,k}|_\infty > 2^{\frac{n\delta}{2}} \right) \leq d2^n \mathbf{P}(|\xi| > d2^{\frac{n\delta}{2}}) = \sqrt{\frac{2}{\pi}} 2^{(1-\delta/2)n} \exp(-2^{n\delta-1}).$$

Since the series

$$\sum_{n=0}^{\infty} \sqrt{\frac{2}{\pi}} 2^{(1-\delta/2)n} \exp(-2^{n\delta-1}) \quad (2.23)$$

is convergent, the Borel-Cantelli argument implies that $\mathbf{P}_\xi(\xi\Omega') = 1$. Eventually, the continuous almost-sure limit process \mathbf{X}_t^∞ is Gaussian as a countable sum of Gaussian processes. \square

Now that these preliminary remarks have been made, we can evaluate, for any t and s in $[0, 1]$, the covariance of \mathbf{X} as the limit of the covariance of \mathbf{X}^N .

Lemma 4. *For any $0 \leq t, s \leq 1$, the covariance of $\mathbf{X}^\infty = \{\mathbf{X}_t^\infty = \boldsymbol{\Psi}_t \circ \boldsymbol{\Xi}; 0 \leq t \leq 1\}$ is*

$$\mathbf{C}(t, s) = \mathbb{E} \left[\mathbf{X}_t^\infty \cdot (\mathbf{X}_s^\infty)^T \right] = \mathbf{g}(t) \mathbf{h}(t \wedge s) \mathbf{g}(s)^T. \quad (2.24)$$

Proof. As $\boldsymbol{\Xi}_{n,k}$ are independent Gaussian random variables of normal law $\mathcal{N}(\mathbf{0}, \mathbf{I}_d)$, we see that the covariance of \mathbf{X}^N is given by

$$\mathbf{C}^N(t, s) = \mathbb{E} \left[\mathbf{X}_t^N \cdot (\mathbf{X}_s^N)^T \right] = \sum_{(n,k) \in \mathcal{I}_N} \boldsymbol{\psi}_{n,k}(t) \cdot (\boldsymbol{\psi}_{n,k}(s))^T.$$

To compute the limit of the right-hand side, we recall that the element of the basis $\boldsymbol{\psi}_{n,k}$ and the functions $\phi_{n,k}$ are linked by the following relation

$$\boldsymbol{\psi}_{n,k}(t) = \mathcal{K}[\phi_{n,k}] = \mathbf{g}(t) \int_U \mathbb{1}_{[0,t]}(s) \mathbf{f}(s) \phi_{n,k}(s) ds$$

from which we deduce

$$\begin{aligned} \mathbf{C}^N(t, s) &= \mathbf{g}(t) \left(\sum_{(n,k) \in \mathcal{I}_N} \left(\int_U \mathbb{1}_{[0,t]}(u) \mathbf{f}(u) \phi_{n,k}(u) du \right) \right. \\ &\quad \left. \left(\int_U \mathbb{1}_{[0,s]}(v) \mathbf{f}(v) \phi_{n,k}(v) dv \right)^T \right) \mathbf{g}(s)^T. \end{aligned}$$

Defining the auxiliary $\mathbb{R}^{d \times d}$ -valued function

$$\boldsymbol{\kappa}_{n,k}(t) = \int_U \mathbb{1}_{[0,t]}(u) \mathbf{f}(u) \phi_{n,k}(u) du,$$

we observe that (i, j) -coefficient function reads

$$\begin{aligned} (\boldsymbol{\kappa}_{n,k})_{i,j}(t) &= \int_U \mathbb{1}_{[0,t]}(u) \left(l_i(\mathbf{f}(u))^T \cdot c_j(\phi_{n,k}(u)) \right) du, \\ &= \int_U \left(\mathbb{1}_{[0,t]}(u) c_i(\mathbf{f}^T(u))^T \cdot c_j(\phi_{n,k}(u)) \right) du, \end{aligned}$$

where $\mathbb{1}_{[0,t]}$ is the real function that is one if $0 \leq u \leq t$ and zero otherwise. As we can write

$$\mathbb{1}_{[0,t]}(u) c_i(\mathbf{f}^T(u)) = \mathbf{f}^T(u) \cdot \begin{bmatrix} 0 \\ \vdots \\ \mathbb{1}_{[0,t]}(u) \\ \vdots \\ 0 \end{bmatrix} \leftarrow i,$$

we see that the function $\mathbf{f}_{i,t} = \mathbb{1}_{[0,t]} c_i(\mathbf{f}^T)$ belongs to $L_{\mathbf{f}}^2$, so that we can write $(\boldsymbol{\kappa}_{n,k})_{i,j}(t)$ as a scalar product in the Hilbert space $L_{\mathbf{f}}^2$:

$$(\boldsymbol{\kappa}_{n,k})_{i,j}(t) = \int_U \mathbf{f}_{i,t}^T(u) \cdot c_j(\phi_{n,k}(u)) du = \left(\mathbf{f}_{i,t}, c_j(\phi_{n,k}) \right).$$

We then specify the (i, j) -coefficient of $\mathbf{g}^{-1}(t) \mathbf{C}^N(t, s) (\mathbf{g}^{-1}(s))^T$ writing

$$\sum_{(n,k) \in \mathcal{I}_N} \left(\boldsymbol{\kappa}(t) \cdot \boldsymbol{\kappa}(s)^T \right)_{i,j} = \sum_{(n,k) \in \mathcal{I}_N} \sum_{p=0}^{d-1} \left(\mathbf{f}_{i,t}, c_j(\boldsymbol{\phi}_{n,k}) \right) \left(\mathbf{f}_{j,s}, c_j(\boldsymbol{\phi}_{n,k}) \right)$$

and, remembering that the family of functions $c_j(\boldsymbol{\phi}_{n,k})$ forms a complete orthonormal system of $L^2_{\mathbf{f}}$, we can use the Parseval identity, which reads

$$\begin{aligned} \sum_{(n,k) \in \mathcal{I}} \left(\boldsymbol{\kappa}(t) \cdot \boldsymbol{\kappa}(s)^T \right)_{i,j} &= \left(\mathbf{f}_{i,t}, \mathbf{f}_{j,s} \right) \\ &= \int_U \mathbb{1}_{[0,t]}(u) c_i(\mathbf{f}^T(u))^T \cdot \mathbb{1}_{[0,s]}(u) c_j(\mathbf{f}^T(u)) du, \\ &= \int_0^{t \wedge s} (\mathbf{f} \cdot \mathbf{f}^T)_{i,j}(u) du. \end{aligned}$$

Thanks to this relation, we can conclude the evaluation of the covariance since

$$\lim_{N \rightarrow \infty} \mathbf{C}^N(t, s) = \mathbf{g}(t) \left(\int_0^{t \wedge s} (\mathbf{f} \cdot \mathbf{f}^T)(u) du \right) \mathbf{g}(s)^T = \mathbf{g}(t) \mathbf{h}(t \wedge s) \mathbf{g}(s)^T,$$

□

We stress the fact that the relation

$$\mathbf{C}(t, s) = \sum_{(n,k) \in \mathcal{I}} \boldsymbol{\psi}_{n,k}(t) \cdot (\boldsymbol{\psi}_{n,k}(s))^T = \boldsymbol{\Psi}(t) \circ \boldsymbol{\Psi}^T(s).$$

provides us with a continuous version of the Cholesky decomposition of the covariance kernel \mathbf{C} . Indeed, if we chose $\boldsymbol{\sigma}_{n,k}$ as the Cholesky square root of $\boldsymbol{\Sigma}_{n,k}$, we remark that the operators $\boldsymbol{\Psi}$ are triangular in the following sense: consider the chain of nested vectorial spaces $\{F_{n,k}\}_{(n,k) \in \mathcal{I}}$

$$F_{0,0} \subset F_{1,0} \subset (F_{2,0} \subset F_{2,1}) \dots \subset (F_{n,0} \subset \dots \subset F_{n,2^n-1}) \dots \subset \xi \Omega'$$

with $F_{n,k} = \text{span} \{ \mathbf{f}_{i,j} \mid 0 \leq i \leq n, 0 \leq j \leq k \}$, then for every (n, k) in \mathcal{I} , the operator Ψ transforms the chain $\{F_{n,k}\}_{(n,k) \in \mathcal{I}}$ into the chain

$$\mathcal{E}_{0,0} \subset \mathcal{E}_{1,0} \subset (\mathcal{E}_{2,0} \subset \mathcal{E}_{2,1}) \dots \subset (\mathcal{E}_{n,0} \subset \dots \subset \mathcal{E}_{n,2^n-1}) \dots \subset {}_x\Omega'$$

with $\mathcal{E}_{n,k} = \text{span} \{ \Psi_{i,j} \mid 0 \leq i \leq n, 0 \leq j \leq k \}$.

The fact that this covariance is equal to the covariance of the process \mathbf{X} , the solution of equation (2.2) implies that we have the following fundamental result:

Theorem 4. *The process \mathbf{X}^∞ is equal in law to the initial linear diffusion \mathbf{X} used to construct the basis of functions.*

2.4 Derivation of the Bases for Some Classical Processes

2.4.1 One-Dimensional Case

In the one-dimensional case, the construction of the Gauss-Markov processes, which coincide with the linear diffusions, is considerably simplified since we do not have to consider the potential degeneracy of matrix-valued functions. Indeed, in this situation, the centered Gauss-Markov process X is the solution of the one-dimensional stochastic equation

$$dX_t = \alpha(t) X_t dt + \sqrt{\Gamma(t)} dW_t,$$

with α homogeneously Hölder continuous and positive continuous function Γ . We then have the Doob representation

$$X_t = g(t) \int_0^t f(s) dW_s, \quad \text{with} \quad g(t) = e^{\int_0^t \alpha(v) dv}, \quad f(t) = \sqrt{\Gamma(t)} e^{-\int_0^t \alpha(v) dv}.$$

Define the function h as

$$h(t) = \int_0^t f^2(s) ds,$$

then the covariance of the process reads for any $0 \leq t, s \leq 0$ as

$$C(t, s) = g(t)g(s)h(t \wedge s).$$

The variance of the Gauss-Markov bridge B_t pinned in t_x and t_z yields

$$(\sigma_{t_x, t_z}(t))^2 = g(t)^2 \frac{(h(t) - h(t_x))(h(t_z) - h(t))}{h(t_z) - h(t_x)}.$$

These simple relations insures that the functions $\psi_{n,k}$ are defined on their supports $S_{n,k}$ by $\psi_{n,k}(t)^2 = \mathbb{E}[(\delta^n(t))^2]$ with

$$\begin{aligned} \mathbb{E}[(\delta^n(t))^2] &= (\sigma_{l_{n,k}, r_{n,k}}(t))^2 \\ &\quad - \left(\mathbb{1}_{S_{n+1, 2k}}(t) (\sigma_{l_{n,k}, m_{n,k}}(t))^2 + \mathbb{1}_{S_{n+1, 2k+1}}(t) (\sigma_{l_{n,k}, m_{n,k}}(t))^2 \right). \end{aligned}$$

This reads on $S_{n+1, 2k}$ as

$$\begin{aligned} \psi_{n,k}(t)^2 &= g(t)^2 \left[\frac{(h(t) - h(l_{n,k}))(h(r_{n,k}) - h(t))}{h(r_{n,k}) - h(l_{n,k})} \right. \\ &\quad \left. - \frac{(h(t) - h(l_{n,k}))(h(m_{n,k}) - h(t))}{h(m_{n,k}) - h(l_{n,k})} \right], \end{aligned}$$

and on $S_{n+1, 2k+1}$ as

$$\begin{aligned} \psi_{n,k}(t)^2 &= g(t)^2 \left[\frac{(h(t) - h(l_{n,k}))(h(r_{n,k}) - h(t))}{h(r_{n,k}) - h(l_{n,k})} \right. \\ &\quad \left. - \frac{(h(t) - h(m_{n,k}))(h(r_{n,k}) - h(t))}{h(r_{n,k}) - h(m_{n,k})} \right], \end{aligned}$$

and therefore we have:

$$\psi_{n,k}(t) = \begin{cases} \frac{\sigma_{n,k} g(t)(h(t) - h(l_{n,k}))}{g(m_{n,k})(h(m_{n,k}) - h(l_{n,k}))}, & l_{n,k} \leq t \leq m_{n,k}, \\ \frac{\sigma_{n,k} g(t)(h(r_{n,k}) - h(t))}{g(m_{n,k})(h(r_{n,k}) - h(m_{n,k}))}, & m_{n,k} \leq t \leq r_{n,k}, \end{cases}$$

with

$$\sigma_{n,k} = \sqrt{\frac{(h(r_{n,k}) - h(m_{n,k}))(h(m_{n,k}) - h(l_{n,k}))}{h(r_{n,k}) - h(l_{n,k})}}.$$

As for the first element, it simply results from the conditional expectation of the one-dimensional bridge pinned in $l_{0,0} = 0$ and $r_{0,0} = 1$:

$$\psi_{0,0}(t) = \frac{g(t)(h(t) - h(l_{0,0}))}{\sqrt{h(r_{0,0}) - h(l_{0,0})}}.$$

In this class of processes, two paradigmatic processes are the Wiener process and the Ornstein-Uhlenbeck processes with constant coefficients. In the case of the Wiener process, $h(t) = t$ and $g(t) = 1$, which yields the classical triangular-shaped Schauder functions used by Lévy [126]. As for the Ornstein-Uhlenbeck process with constant coefficients α and $\sqrt{\Gamma}$, we have $g(t) = \exp(\alpha t)$, $f(t) = \sqrt{\Gamma} \exp(-\alpha t)$ and $h(t) = \frac{\Gamma}{2\alpha}(1 - e^{-2\alpha t})$, yielding for the construction basis the expression:

$$\psi_{n,k}(t) = \begin{cases} \sqrt{\frac{\Gamma}{\alpha}} \frac{\sinh(\alpha(t - l_{n,k}))}{\sqrt{\sinh(\alpha(m_{n,k} - l_{n,k}))}}, & l_{n,k} \leq t \leq m_{n,k}, \\ \sqrt{\frac{\Gamma}{\alpha}} \frac{\sinh(\alpha(r_{n,k} - t))}{\sqrt{\sinh(\alpha(m_{n,k} - l_{n,k}))}}, & m_{n,k} \leq t \leq r_{n,k}, \end{cases}$$

and

$$\psi_{0,0}(t) = \sqrt{\frac{\Gamma}{\alpha}} \frac{e^{-\alpha/2} \sinh(\alpha t)}{\sqrt{\sinh(\alpha)}}$$

which were already evidenced in [217].

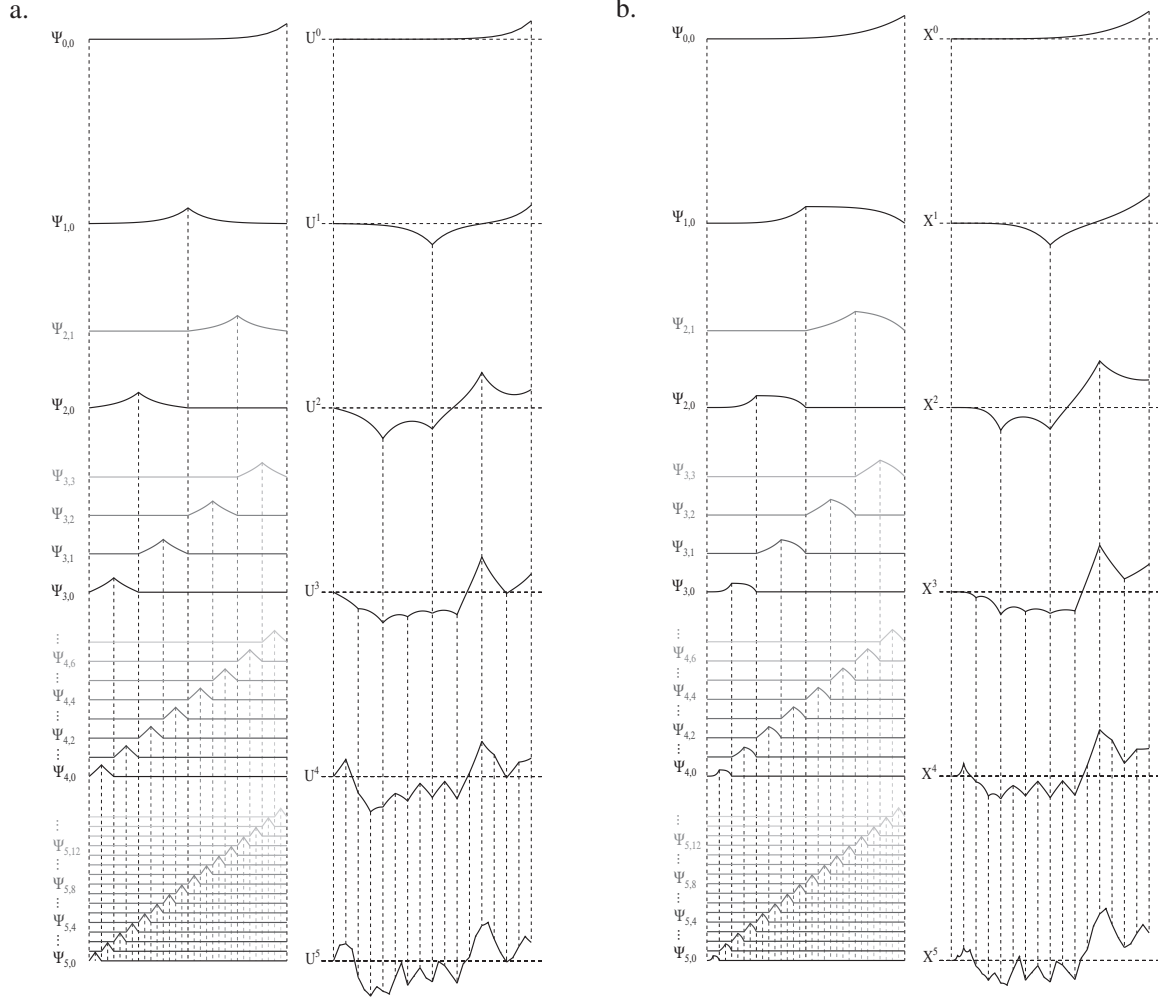


Figure 2.3: For each panel, in the left column, the elements of the basis $\Psi_{n,k}$ are represented for each rank n with $0 \leq n < 6$. In the right column, the partial sums $U^n(\omega)$ or $X^n(\omega)$ are shown for a given set of realizations ω . U refer to a canonical Ornstein-Uhlenbeck process, whereas X is a linear diffusion resulting with time-dependent drift $\alpha(t) = -1/(1+t)$. **a.** For large scale, the non-zero leak term cause the basis elements to exponentially attenuated, whereas at small scale we asymptotically recover the Schauder elements of the Wiener process. **b.** The asymptotical behavior toward the Wiener process is observable as expected. However, notice that we recover the Schauder elements sooner for large time since α vanishes accordingly, and also notice that the time-dependence breaks the time-reversal symmetry.

2.4.2 Multidimensional Case

In the multidimensional case, the explicit expressions for the basis functions $\psi_{n,k}$ make fundamental use of the flow \mathbf{F} of the underlying linear equation (2.4) for a given function α . For commutative forms of α (i.e such that $\alpha(t) \cdot \alpha(s) = \alpha(s) \cdot \alpha(t)$ for all t, s), the flow can be formally expressed as an exponential operator. It is however a notoriously difficult problem to find a tractable expression for general α . As a consequence, it is only possible to provide closed-form formulae for our basis functions in very specific cases.

Multi-Dimensional Gauss-Markov Rotations

We consider in this section α antisymmetric and constant and $\sqrt{\Gamma} \in \mathbb{R}^{d \times m}$ such that $\Gamma = \sigma^2 \mathbf{I}_d$. Since $\alpha^T(t) = -\alpha(t)$, we have:

$$\mathbf{F}(s, t)^T = \mathbf{F}(s, t)^{-1},$$

i.e. the flow is unitary. This property implies that

$$\mathbf{h}_u(s, t) = \sigma^2 \int_s^t \mathbf{F}(w, u) \mathbf{F}(w, u)^T dw = \sigma^2(t - s) \mathbf{I}_d,$$

which yields by definition of $\sigma_{n,k}$

$$\sigma_{n,k} \cdot \sigma_{n,k}^T = \sigma^2 \frac{(m_{n,k} - l_{n,k})(r_{n,k} - m_{n,k})}{r_{n,k} - l_{n,k}} \mathbf{I}_d.$$

The square root $\sigma_{n,k}$ is then uniquely defined (both by choosing Cholesky and symmetrical square root) by

$$\sigma_{n,k} = \sigma \sqrt{\frac{(m_{n,k} - l_{n,k})(r_{n,k} - m_{n,k})}{(r_{n,k} - l_{n,k})}} \mathbf{I}_d,$$

and $\psi_{n,k}(t)$ reads

$$\psi_{n,k}(t) = \begin{cases} \sigma \sqrt{\frac{r_{n,k} - m_{n,k}}{(m_{n,k} - l_{n,k})(r_{n,k} - l_{n,k})}} (t - l_{n,k}) \mathbf{F}(m_{n,k}, t), & l_{n,k} \leq t \leq m_{n,k}, \\ \sigma \sqrt{\frac{m_{n,k} - l_{n,k}}{(r_{n,k} - m_{n,k})(r_{n,k} - l_{n,k})}} (r_{n,k} - t) \mathbf{F}(m_{n,k}, t), & l_{n,k} \leq t \leq m_{n,k}. \end{cases}$$

Recognizing the (n, k) element of the Schauder basis for the construction of the one-dimensional Wiener process

$$s_{n,k}(t) = \begin{cases} \sqrt{\frac{r_{n,k} - m_{n,k}}{(r_{n,k} - l_{n,k})(m_{n,k} - l_{n,k})}} (t - l_{n,k}), & l_{n,k} \leq t \leq m_{n,k}, \\ \sqrt{\frac{m_{n,k} - l_{n,k}}{(r_{n,k} - l_{n,k})(r_{n,k} - m_{n,k})}} (r_{n,k} - t), & l_{n,k} \leq t \leq m_{n,k}, \end{cases}$$

we obtain the following formula:

$$\psi_{n,k}(t) = \sigma s_{n,k}(t) \mathbf{F}(t - m_{n,k}).$$

This form shows that the Schauder basis for multidimensional rotations results from the multiplication of the triangular-shaped elementary function used for the Lévy-Ciesielski construction of the Wiener process with the flow of the equation, i.e. the elementary rotation.

The simplest example of this kind is the stochastic the 2-dimensional rotating process corresponding to:

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \sqrt{\Gamma} = \sigma^2 \mathbf{I}_2. \quad (2.25)$$

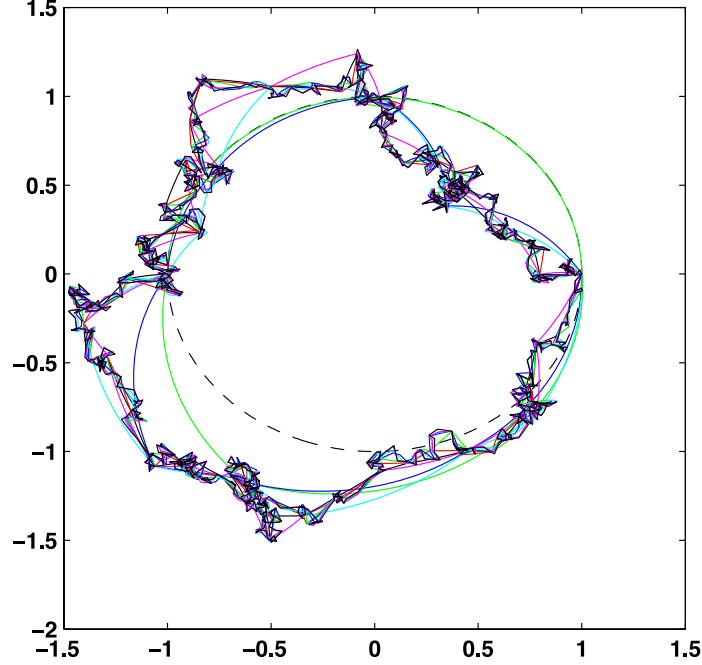


Figure 2.4: Multi-resolution construction of the 2-dimensional rotating Ornstein-Uhlenbeck process.

In that case, $\psi_{n,k}$ has the expression

$$\psi_{n,k}(t) = s_{n,k}(t) \begin{pmatrix} \cos(t - m_{n,k}) & -\sin(t - m_{n,k}) \\ \sin(t - m_{n,k}) & \cos(t - m_{n,k}) \end{pmatrix}$$

Interestingly, the different basis functions have the structure of the solutions of the non-stochastic oscillator equation. One of the equations perturbs the trajectory in the radial component of the deterministic solution, and the other one in the tangential direction. We represent such a construction in Figure 2.4 with the additional conditioning that $\mathbf{X}_1 = \mathbf{X}_0$, i.e. imposing that the trajectory forms a loop between time 0 and 1.

The Successive Primitives of the Wiener Process

In applications, it often occurs that people use smooth stochastic processes to model the integration of noisy signals. This is for instance the case of a particle subject to

a Brownian forcing or of the synaptic integration of noisy inputs [225]. Such smooth processes involve in general integrated martingales, and the simplest example of such processes are the successive primitives of a standard Wiener process.

Let $d > 2$ and denote by X_t^d the $d - 1$ th order primitive of the Wiener process. This process can be defined via the lower order primitives X_t^k for $k < d$ via the relations:

$$\begin{cases} dX_t^{k+1} &= X_t^k dt & k < d \\ dX_t^1 &= dW_t \end{cases}$$

where W_t is a standard real Wiener process. These equations can be written in our formalism

$$d\mathbf{X}_t = \boldsymbol{\alpha}(t) \cdot \mathbf{X}_t + \sqrt{\boldsymbol{\Gamma}(t)} \cdot dW_t,$$

with

$$\boldsymbol{\alpha}(t) = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & 1 & \\ & & & & 0 \end{bmatrix}, \quad \sqrt{\boldsymbol{\Gamma}(t)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

In particular, though none of the integrated processes X^k for $K > 1$ is Markov by itself, the d -tuple $\mathbf{X} = (X^d, \dots, X^1)$ is a Gauss-Markov process.

Furthermore because of the simplicity and the sparsity of the matrices involved, we can identify in a compact form all the variables used in the computation of the construction basis for these processes. In particular, the flow \mathbf{F} of the equation is the exponential of the matrix $\boldsymbol{\alpha}$, and since α is nilpotent, it is easy to show that \mathbf{F} has

the expression:

$$\mathbf{F}(s, t) = \begin{bmatrix} 1 & (t-s) & \frac{(t-s)^2}{2} & \cdots & \frac{(t-s)^{d-1}}{(d-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{(t-s)^2}{2} \\ & & & \ddots & (t-s) \\ & & & & 1 \end{bmatrix}$$

and the only non-zero entry of the $d \times d$ matrix $\mathbf{\Gamma}$ is one at position $(d-1, d-1)$.

Using this expression and the highly simple expression of $\mathbf{\Gamma}$, we can compute the general element of the matrix $\mathbf{h}_u(t, s)$, which reads:

$$(\mathbf{h}_u(s, t))_{i,j} = (-1)^{i+j} \frac{(t-u)^{2d-1-(i+j)} - (s-u)^{2d-1-(i+j)}}{(2d-1-(i+j))(d-1-i)!(d-1-j)!}.$$

Eventually, we observe that the functions $\psi_{n,k}$ yielding the multi-resolution description of the integrated Wiener processes, are directly deduced from the matrix-valued function

$$(\mathbf{c}_{n,k}(t))_{i,j} = \begin{cases} \psi_{n,k} \cdot \mathbf{L}_{n,k}^{-1} = \mathbf{g}(t) \mathbf{h}(l_{n,k}, t), & l_{n,k} \leq t \leq m_{n,k}, \\ \psi_{n,k} \cdot \mathbf{R}_{n,k}^{-1} = \mathbf{g}(t) \mathbf{h}(t, r_{n,k}), & m_{n,k} \leq t \leq r_{n,k}, \end{cases}$$

whose components are further expressed as

$$(\mathbf{c}_{n,k}(t))_{i,j} = \sum_{p=i}^{d-1} (-1)^{p+j} \frac{t^{i-p}}{(i-p)!} \frac{t^{2d-1-(p+j)} - l_{n,k}^{2d-1-(p+j)}}{(2d-1-(p+j))(d-1-p)!(d-1-j)!},$$

for $l_{n,k} \leq t \leq m_{n,k}$ and as

$$(\mathbf{c}_{n,k}(t))_{i,j} = \sum_{p=i}^{d-1} (-1)^{p+j} \frac{t^{i-p}}{(i-p)!} \frac{m_{n,k}^{2d-1-(p+j)} - t^{2d-1-(p+j)}}{(2d-1-(p+j))(d-1-p)!(d-1-j)!},$$

for $m_{n,k} \leq t \leq r_{n,k}$. The final computation of the $\psi_{n,k}$ involves the computation of $\mathbf{L}_{n,k}$ and $\mathbf{R}_{n,k}$, which in the general case can become very complex. However, this expression is highly simplified if one assumes that $m_{n,k}$ is the middle of the interval $[l_{n,k}, r_{n,k}]$. Indeed, in that case, we observe that for any (i, j) such that $i + j$ is odd, $(\mathbf{h}_m(l, r))_{i,j} = \mathbf{0}$ which induces the same property on the covariance matrix $\Sigma_{n,k}$ and on the polynomials $(\mathbf{c}_{n,k}(t))_{i,j}$. This property gives therefore a preference to the dyadic partition that provides simple expressions for the basis elements in any dimension, and allows simple computation of the basis.

Remark 2. Observe that for all $0 \leq i < d - 1$, we have

$$\begin{aligned} (\mathbf{c}_{n,k}(t))'_{i,j} &= (\mathbf{c}_{n,k}(t))_{i+1,j} \pm \sum_{p=i}^{d-1} (-1)^{p+j} \frac{t^{i-p}}{(i-p)!} \frac{t^{2d-1-(p+j)} - l_{n,k}^{2d-2-(p+j)}}{(d-1-p)!(d-1-j)!}, \\ &= (\mathbf{c}_{n,k}(t))_{i+1,j} \pm \frac{t^{d-j-1}}{(d-j-1)!} \underbrace{\sum_{q=0}^{d-1-i} \frac{(-t)^q t^{(d-1-i)-p}}{p!((d-1-i)-p)!}}_0. \end{aligned}$$

As $\mathbf{L}_{n,k}$ and $\mathbf{R}_{n,k}$ are constant, we immediately deduce the important relation that for all $0 \leq i \leq d-1$, $(\psi_{n,k}(t))_{0,j}^{(i)} = (\psi_{n,k}(t))_{i,j}$. This indicates that each finite-dimensional sample path of our construction has components that satisfy the non-deterministic equation associated with the iteratively integrated Wiener process. Actually, this fact is better stated remembering that the Schauder basis $\psi_{n,k}$ and the corresponding orthonormal basis $\phi_{n,k} : [0, 1] \rightarrow \mathbb{R}^{1 \times d}$ are linked through the equation (2.17), which

reads

$$\begin{aligned}
\begin{bmatrix} (\boldsymbol{\psi}_{n,k})'_{0,0} & \cdots & (\boldsymbol{\psi}_{n,k})'_{0,d-1} \\ \vdots & & \vdots \\ (\boldsymbol{\psi}_{n,k})'_{d-2,0} & \cdots & (\boldsymbol{\psi}_{n,k})'_{d-2,d-1} \\ (\boldsymbol{\psi}_{n,k})'_{d-1,0} & \cdots & (\boldsymbol{\psi}_{n,k})'_{d-1,d-1} \end{bmatrix} &= \begin{bmatrix} (\boldsymbol{\psi}_{n,k})_{1,0} & \cdots & (\boldsymbol{\psi}_{n,k})_{1,d-1} \\ \vdots & & \vdots \\ (\boldsymbol{\psi}_{n,k})_{d-1,0} & \cdots & (\boldsymbol{\psi}_{n,k})_{d-1,d-1} \\ 0 & \cdots & 0 \end{bmatrix} \\
&+ \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ (\boldsymbol{\phi}_{n,k})_{0,0} & \cdots & (\boldsymbol{\phi}_{n,k})_{0,d-1} \end{bmatrix}.
\end{aligned}$$

Additionally, we realize that the orthonormal basis is entirely determined by the one-dimensional families $(\boldsymbol{\phi}_{n,k})_{0,j}$, which are mutually orthogonal functions satisfying $(\boldsymbol{\phi}_{n,k})_{0,j} = (\boldsymbol{\psi}_{n,k})_{0,j}^{(d)}$.

We study in more details the case of the integrated and doubly-integrated Wiener processes for $d = 2$ and $d = 3$. As expected from the optimality characterization of Appendix A, the first row of the basis functions for the integrated Wiener process turns out to be the well-known cubic Hermite splines [48]. These functions have been widely used in numerical analysis and actually constitute the basis of lowest degree in a wider family of bases known as the natural basis of polynomial splines of interpolation [110]. Such bases are used to interpolate data points with constraint of smoothness of different degree (for instance the cubic Hermite splines ensure that the resulting interpolation is in $C^1[0, 1]$). The next family of splines of interpolation (corresponding to the C^2 constraint) is naturally retrieved by considering the construction of the doubly-integrated Wiener process: we obtain a family of three 3-dimensional functions, that constitutes the columns of a 3×3 matrix that we denote $\boldsymbol{\psi}$ and that we represent in Figure 2.5. The top row is made of polynomials of degree five, which have again simple expressions when $m_{n,k}$ is the middle of the interval $[l_{n,k}, r_{n,k}]$.

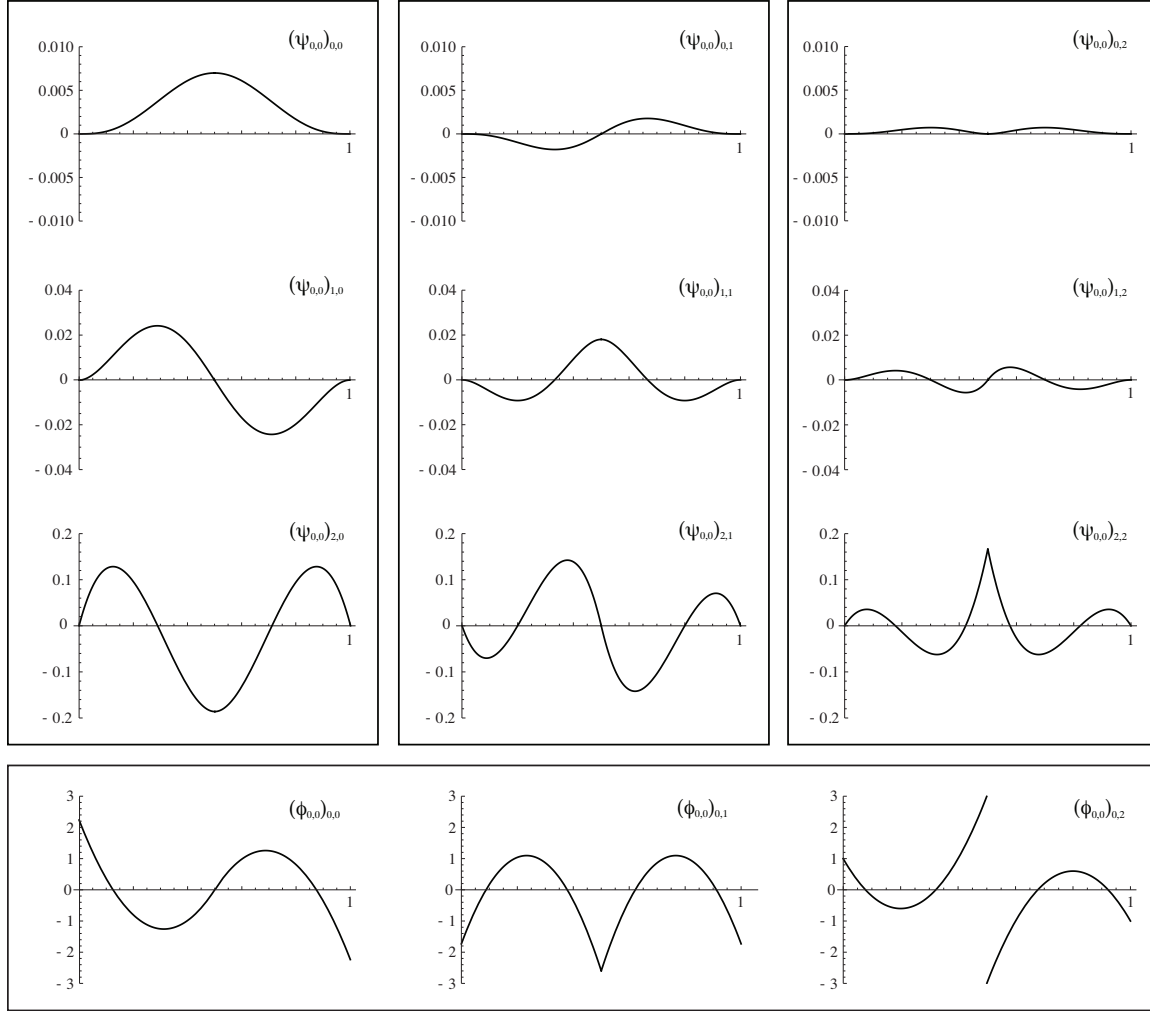


Figure 2.5: Basis element for the construction of the doubly-integrated Wiener process ($d = 3$). The three columns form the square matrix function $\psi_{0,0}$, notice that in each column $(\psi_{0,0})_{i+1,j} = (\psi_{0,0})'_{i,j}$, for $0 \leq j < 3$ and $0 \leq i < 2$. These functions are known as the polynomial splines of interpolation of degree 5. The bottom line of functions $(\phi_{0,0})_{0,j}$, $0 \leq j < 3$, form the underlying orthonormal system.

The ideas underlying this work can be directly traced back to the original work of Lévy. Here, we intend to develop a self-contained Schauder dual framework to further the description of multidimensional linear diffusions, and in doing so, we extend some well-known results of interpolation theory in signal processing [55, 109, 110]. There are a number of practical applications where applying the Schauder basis framework clearly provides an advantage compared to standard stochastic calculus methods, among which first-hitting times of stochastic processes, pricing of multidimensional path-dependant options [16, 18, 17, 82], regularization techniques for support vector machine learning [197] and more theoretical work on uncovering the differential geometry structure of the space of linear diffusions [201].

Chapter 3

First-Passage Problem for Gauss-Markov Processes

The leaky integrate-and-fire model [180, 121, 179, 34, 35] represents an idealized mechanism for stochastic non-linear spike generation. In the current-based version of this model, the neuronal state is encoded by a time-dependent internal variable X_t representing the membrane potential. This potential fluctuates according to a stochastic diffusion equation that integrates internal noise and an external, potentially noisy, current input. Whenever the membrane's voltage reaches a triggering threshold V , a spiking event is generated.

As a simplistic approximation, the stochastic sub-threshold dynamic of X_t is entirely determined by a leak constant α and an internal noise intensity Γ through the linear diffusion equation:

$$dX_t = -\alpha(t)X_t dt + I(t) dt + \sqrt{\Gamma}(t)dW_t,$$

where I denotes the instantaneous input and dW_t is the mathematical representation of a Gaussian white noise. The resulting trajectories of X_t are sample paths of one-dimensional Gauss-Markov processes, that are instantaneously reset to a resting potential v upon reaching the spiking threshold V .

We are thus led to consider spiking time as the first-passage time τ of the stochastic process X_t (which includes a deterministic drift resulting from the integration of I) with the threshold V , given that X starts at v for some initial time t_0 :

$$\tau = \inf\{t > t_0 \mid X_t > V, X_{t_0} = v\}. \quad (3.1)$$

We can formally write X as the sum of its purely stochastic part U

$$U_t = ve^{-\int_{t_0}^t \alpha(s) ds} + \int_{t_0}^t e^{-\int_{t_0}^s \alpha(u) du} dW_s,$$

and its deterministic drift l given as

$$l(t) = \int_{t_0}^t e^{-\int_0^s \alpha(u) du} I(s) ds,$$

thus defining τ as the first-passage of a Gauss-Markov process with a fluctuating effective barrier [61, 62, 70]:

$$\tau = \inf\{t > t_0 \mid U_t > L(t) = V - l(t), U_{t_0} = v\}. \quad (3.2)$$

Moreover, it is known from Doob's representation [60] that the one-dimensional Gauss-Markov process U_t has the same law as the stochastic process $g(t)W_{h(t)}$ with

$$g(t) = e^{\int_{t_0}^t \alpha(s) ds}, \quad h(t) = \int_{t_0}^t \Gamma(s) e^{-2 \int_{t_0}^s \alpha(s) ds},$$

and where W is a Wiener process satisfying $W_{t_0} = U_{t_0}$. As a consequence, the law of the first-passage time τ is deduced from the first-passage time σ

$$\sigma = \inf\{s > s_0 \mid W_s > g^{-1}(s)L(h^{-1}(s)), W_{t_0} = v\}. \quad (3.3)$$

through the law of the transformed random variable $h^{-1}(\sigma)$. Assuming smooth time-dependance for the leak constant and noise intensity, the transformation h is itself smooth. As such, it is adequate to study the first-passage time of the Wiener process with a fluctuating barrier to characterize the properties of the distribution of spiking events [61, 62]. For the sake of simplicity, we consider that α and Γ are constant if not stipulated otherwise, in which case the Gauss-Markov process is the canonical Ornstein-Uhlenbeck process [179].

In this chapter, we adopt the point of view of an effective fluctuating boundary to

study theoretically the relationship linking the regularity of a fluctuating barrier with the law of the time at which a Wiener process first crosses this barrier. In the first section, we introduce various methods used to study first-passage distributions based on the elementary but seminal example of an affine barrier. In the second section, we exhibit the essential features of a fluctuating barrier that can qualitatively alter the distribution of first-passage events by focusing on piecewise linear barriers. Finally, in the third section, we give a formal account of the results inferred from intuitions linking regularity of the barrier and the distribution of first-passage occurrences.

3.1 Theoretical Approaches for the Wiener First-Passage Problem.

Completing our program requires familiarity with different mathematical approaches, so we provide a short survey of the theory of the first-passage while focusing on an elementary example. The situation of a Wiener process hitting an affine barrier L is the simplest first-passage problem that one can conceive. Its study dates back the original work of Schrödinger [199] and Smoluchowski [240] who first exhibited the density of the first-passage time to a constant threshold in closed form. Incidentally, thanks to the availability of simple analytical formulas, the case of an affine barrier is especially amenable to the introduction of the different intertwined theoretical approaches in elucidating the first-passage problem.

3.1.1 Direct Probabilistic Treatment

The many properties of the Wiener process allow us to establish the density of first-passage time to a simple barrier through direct probabilistic arguments. For being a classical result, we only derive the result heuristically following a standard symmetry argument [106] called the *reflection principle* [126]. Given a Wiener process W start-

ing in x at time zero, let us consider the first-passage time τ to a constant boundary at level $l > 0$. The key point to establish the law of τ is to consider the continuation of the sample path $t \mapsto W_t(\omega)$ after it reaches the threshold. At a time $t > \tau$, we are presented with two mutually exclusive cases, either the Wiener process lies strictly above l or below l meaning:

$$\mathbb{P}(\tau < t) = \mathbb{P}(\tau < t \mid W_t > l)\mathbb{P}(W_t > l) + \mathbb{P}(\tau < t \mid W_t \leq l)\mathbb{P}(W_t \leq l).$$

The probability law of W' , the continuation of the Wiener process after the first-passage time $W'_{t-\tau} = \{W_t - W_\tau\}_{t \geq \tau}$, is the same as a Wiener process conditioned to start in zero at time zero. In other words, after the first time W reaches the threshold l , the following trajectory $\{W(\omega)\}_{t \geq \tau}$ unfolds irrespective of its past history. We stress that this property is more stringent than the Markov property, since the process is conditioned at the random time τ . It is referred to in the mathematical literature as the strong Markov property [106].

To each sample path $\{W_t(\omega)\}_{t \geq \tau}$, we can associate a “shadow” path by reflection at the level of the threshold l : $\{B - W_t(\omega)\}_{t \geq \tau}$. At any given time $t > \tau$, the reflection transforms the set of sample paths for which $W_t > l$ into the set for which $W_t < l$ and, reciprocally, so that we can convince ourselves that, by symmetry, $\mathbb{P}(\tau < t \mid W_t > l) = \mathbb{P}(\tau < t \mid W_t \leq l)$. We then observe that, knowing $W_t > l$, the previous point entails that the first-passage happens for $\tau < t$ and that we have

$$\mathbb{P}(\tau < t) = 2\mathbb{P}(\tau < t \mid W_t > l) = 2 \left(1 - \Phi \left(\frac{l}{\sqrt{t}} \right) \right) = \sqrt{\frac{2}{\pi}} \int_{-l/\sqrt{t}}^{\infty} e^{-\frac{x^2}{2t}} dx. \quad (3.4)$$

where Φ is the normal cumulative distribution. By differentiation on t , we recover the density of τ as the well-known Wald distribution

$$q(t) = \frac{l}{\sqrt{2\pi t^3}} \cdot \exp \left(-\frac{l^2}{2t} \right).$$

This expression is marginally altered when the barrier L has a linear time-dependence $L(t) = l + \alpha t$. To establish this, we must consider the drifted Wiener process $W' = W - \alpha t$ since finding the first-passage time τ of W with L is equivalent to finding the first-passage time τ' of W' with a constant threshold l . Thus, denoting \mathbf{P} as the probability law of W , and \mathbf{P}' as the probability law of W' , we have $\mathbf{P}(\tau < t) = \mathbf{P}'(\tau' < t)$. When the drift rate α is non-zero, W' has a law that clearly differs from the law of W but is still one of a Gauss-Markov process. Actually, since the addition of a deterministic drift corresponds to a simple translation of the sample paths, the probability measures \mathbf{P} and \mathbf{P}' are closely related.

To illustrate this, let us consider a random Gaussian vector $X = (X_1, \dots, X_n)$, whose components are independent and normally distributed according to $\mathcal{N}(0, \sigma_i)$, $1 \leq i \leq n$ and the translated vector $X' = (X_1 - \mu_1, \dots, X_n - \mu_n)$. Clearly, X' and X take values in the same space \mathbb{R}^n and the probability density p' of X' reads

$$\begin{aligned} p'(x'_1, \dots, x'_n) &= \frac{1}{(2\pi)^{n/2}} \exp \left(- \sum_{i=1}^n \frac{(x'_i - \mu_i)^2}{2\sigma_i^2} \right) \\ &= p(x'_1, \dots, x'_n) \exp \left(\sum_{i=1}^n \frac{2x'_i \mu_i - \mu_i^2}{2\sigma_i^2} \right) \end{aligned}$$

where p denotes the probability density of the original vector X . Simply stated, there exists a simple function

$$f(x_1, \dots, x_n) = \exp \left(\sum_{i=1}^n \frac{2x_i \mu_i - x_i^2}{2\sigma_i^2} \right),$$

equating the ratio of the probability density of \mathbf{P}' with the density of \mathbf{P} , when evaluated at the same sample point. We recognize in such a function the Radon-Nikodym derivative of the probability measure of X' with respect to X since we can write for

any Lebesgue measurable set B of \mathbb{R}^n

$$\mathbf{P}'(B) = \int_B d\mathbf{P}'(x) = \int_B p'(x) dx = \int_B f(x)q(x) dx = \int_B f(x) d\mathbf{P}(x).$$

The Cameron-Martin theorem [37] and its extension as the Girsanov theorem [81] generalize this result to the infinite-dimensional setting of continuous diffusion processes. Namely, for the simple case of our translated processes W' and W of respective measure \mathbf{P}' and \mathbf{P} , it states that \mathbf{P}' admits a Radon-Nikodym derivative with respect to \mathbf{P} under the form

$$\frac{d\mathbf{P}'}{d\mathbf{P}}(\omega) = \exp\left(\alpha W_t(\omega) - \frac{\alpha^2 t}{2}\right),$$

where the sample space of W and W' is identified with the Wiener space $C_0(0, t)$. Equipped with this result, it is simple to transform our first-passage formulation for an affine barrier into one for a constant threshold:

$$\mathbf{P}'(\tau' < t) = \int_{\{\tau' < t\}} d\mathbf{P}'(\omega) = \int_{\{\tau' < t\}} \frac{d\mathbf{P}'}{d\mathbf{P}}(\omega) d\mathbf{P}(\omega) = \exp\left(\alpha l - \frac{\alpha^2 t}{2}\right) \mathbf{P}(\tau' < t).$$

Since we recall that the event τ' denotes the first-passage time to a constant threshold at l , the measure of the event $\{\tau' < t\}$ under the law of a canonical Wiener process \mathbf{P} follows from (3.4), and we finally get:

$$q(t) = \frac{l}{\sqrt{2\pi t^3}} \cdot \exp\left(-\frac{(l - \alpha t)^2}{2t}\right).$$

When α is negative, that is when the drift pushes away from the barrier, q does not integrate to one over the real half-line. This indicates that, in such a situation, the probability that a first-passage occurs with finite time is not sure: $\mathbb{P}(\tau < \infty) < 1$. Otherwise, when parametrized by l and α , the family of density $q_{\alpha, l}$ corresponds to a family of distributions known as the *inverse Gaussian distributions* represented in Figure 3.1. These distributions play an important role in mathematical and physical

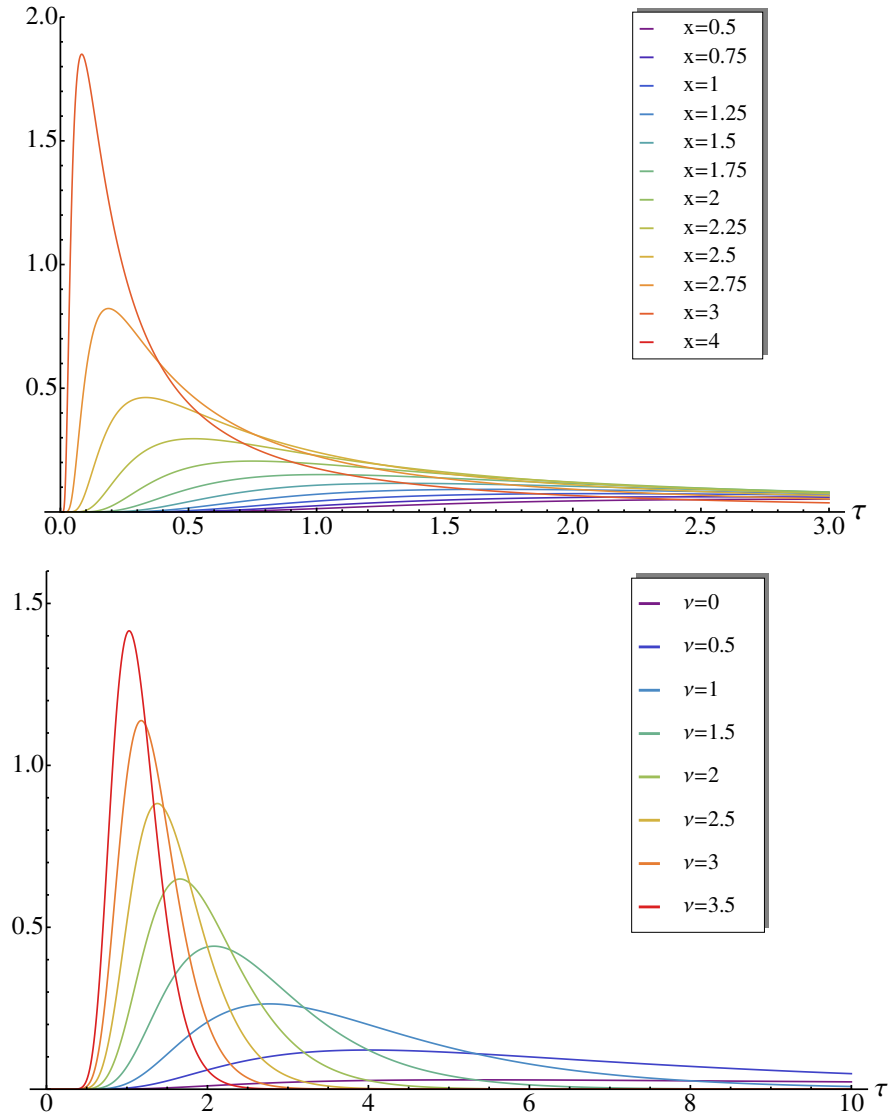


Figure 3.1: Density function of the first-passage time τ of a Wiener process with constant barriers of different height x on top, with affine barriers of the form $t \mapsto 4 - \nu t$ on the bottom.

probability, allowing the construction and simulation of non-Markovian anomalous diffusion processes by subordination [132, 86].

From a mathematical point of view, a more powerful approach consists in focusing on the running maximum of the Wiener process $M_t = \max_{0 \leq s \leq t} W_s$ and noticing that the event $\{M_t > l\}$ coincides with $\{\tau < l\}$. By slightly altering the precedent reflection principle heuristics [106], the joint law of M_t and W_t can be shown to be

$$\mathbb{P}(W_t \in dx, M_t \in dl) = \frac{2(2l - x)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(2l - x)^2}{2t}\right). \quad (3.5)$$

From there, it becomes possible to study the first-passage time when the process W is conditioned to a given value at a later time t : this is the first-passage problem for the Brownian bridge and an affine barrier. In particular, after integrating over $\{M_t > l\}$ the ratio of expression (3.5) with $\mathbb{P}(W_t \in dx) = k(0, 0; t, x)$, we find the probability of occurrence of a first-passage for a Wiener process satisfying $W_t = x < l$:

$$\mathbb{P}(\tau < t | W_t = x) = \mathbb{P}(M_t > l | W_t = x) = \exp\left(-\frac{2l(l + \alpha t - x)}{t}\right).$$

As expected, this probability is strictly less than one when x lies below the barrier. If x is above the barrier, the occurrence of a crossing is certain and the previous expression bears no probabilistic meaning.

3.1.2 Partial Differential Equation Treatment

The many symmetries displayed by the Wiener process make it possible to analytically express the density of first-passage times to an affine barrier directly. However, once the time-dependence becomes more elaborate, the previous argument fails to deliver a closed form answer due to the impossibility of computing conditional expectations of intervening stochastic integrals. Fortunately, there is a deep connection between the probability theory of diffusion processes and the theory of partial differential equa-

tions [114]. In particular, linear diffusion processes can be characterized through their dynamics as prescribed by a stochastic linear equation [216], as well as through their transition kernel seen as the solution of a linear Fokker-Planck equation [182, 169]. When dealing with the Wiener process, this parabolic partial differential equation reduces to the one-dimensional *heat equation* [38]

$$\frac{\partial}{\partial t}p(t, y) = \frac{1}{2} \frac{\partial^2}{\partial y^2}p(t, y) \quad (3.6)$$

For any initial probability distribution μ at time s , the usual one-dimensional heat equation admits a unique solution $(t, y) \mapsto p_t(y)$ under the constraint that this solution is positive at all future times [248]. The solution of this inhomogeneous problem is naturally expressed in terms of the *Green's function* k (or heat kernel) under the integral form:

$$p_t(y) = \int_{\mathbb{R}} k(s, x, t, y) \mu(dx). \quad (3.7)$$

Recall that k is simply the transition kernel of W , i.e. for $s < t$, $\mathbb{P}(W_t \in dy \mid W_s = x) = k(s, x; t, y)$. Moreover, because of the analyticity of the heat kernel for all $t > 0$, the solution $(t, y) \mapsto p_t(y)$ is also analytic for $t > 0$.

We are specifically interested in studying the first-passage problem of a Wiener process W with a barrier $L(t) = l + \alpha t$. This necessitates looking for “new” Green functions to (3.6) under the putative form of a transition kernel for the corresponding *killed* Wiener process. Anytime a Wiener sample path hits the barrier, we can disregard its continuation at later times in the evaluation of a future first-passage. As such, we are able to consider that it vanishes when reaching the barrier.

Such a killed process is still a Markov process and its probability measure is characterized by its transition kernel κ . Because a sample path is killed at the boundary, the transition kernel is zero for $y \geq L(t)$, which entails that κ is non-gaussian. Moreover, since there is a non-zero probability that a sample path dies in any finite time

interval, the transition kernel is degenerate and does not conserve probability:

$$\int_{-\infty}^{L(t)} \kappa(s, x; t, y) dy < 1 \quad \text{for all } t > s.$$

The kernel κ can always be expressed as the product of the Wiener heat kernel $k(s, x; t, y)$ which measures the probability that a sample path starting at (s, x) finishes in the vicinity of y at t , times the conditional probability $s(x, s; t, y)$ for such sample paths to remain strictly below the barrier between s and t :

$$\kappa(s, x; t, y) = k(s, x, t, y) s(x, s; t, y), \quad s(x, s; t, y) = \mathbb{P}(\tau < t \mid W_s = x, W_t = y).$$

In the case of an affine barrier $L(t) = l + \alpha(t - s)$, the survival probability s reads

$$s(s, x; t, z) = \left(1 - \exp \left(- \frac{2(l - x)(l + \alpha(t - s) - y)}{(t - s)} \right) \right).$$

When dealing with the killed process, we naturally relax the constraint of positivity, only enforcing the solution to be positive in the lower half-region $\{(t, x) \mid x < L(t)\}$. Under this new constraint, it is straightforward to verify that the kernel κ satisfies the heat equation (3.6) with initial condition $\kappa(s, x; t, y) = \delta_x(y)$ and with the constraint that $\kappa(s, x; t, L(t)) = 0$ for all $t \geq s$. Notice that the kernel $(t, z) \mapsto \kappa(s, x; t, z)$ is also analytic as soon as $t > s$, so that it has the same regularizing properties as the heat kernel.

In terms of partial differential equations, the kernel κ corresponds to the solution of (3.6) with *absorbing boundary conditions* at the barrier, i.e. $p(t, L(t)) = 0$. The existence and unicity of a solution to the Heat equation for given boundary conditions is a vast topic [38], which we discuss in Section 3.3.2.

Once the expression of κ is established, recovering the first-passage time distribution is a simple matter. As $\kappa(s, x; t, y)$ gives the probability density to find a Wiener

sample path starting at (s, x) in the vicinity of y at t without reaching L , the survival probability of the killed process W is defined as

$$S_t = \int_{-\infty}^{L(t)} \kappa(s, x; t, y) dy. \quad (3.8)$$

The function $t \mapsto S_t$ is a decreasing function and if S admits a density with respect to the Lebesgue measure on the real half-line $[t_x, +\infty)$, this density is naturally the opposite of the probability density q of the first-passage time τ . Differentiating the following relation with respect to t leads to

$$q(t) = -\frac{\partial S_t}{\partial t} = -\frac{1}{2} \left[\frac{\partial \kappa(s, x; t, y)}{\partial y} \right]_{y=L(t)} \quad (3.9)$$

showing that the killing rate increases in keeping with the slope of the density of the killed process in the vicinity of the barrier. Rewriting the heat equation (3.6) as a local law of conservation of probability

$$\frac{\partial}{\partial t} p(t, y) + \frac{\partial}{\partial y} J(t, y) = 0,$$

with

$$J(t, y) = -\frac{1}{2} \partial_y p(t, y),$$

expression (3.9) shows that the probability density of first-passage times $q(t)$ is the probability current $J(t, L(t))$ at the barrier.

To understand intuitively the effect of the absorbing boundary condition, we extend the domain of definition of κ to the entire real line, as opposed to only considering the lower half-plane $\{(t, z) \mid z < L(t)\}$. Following the *method of images*, we then reformulate the original problem into a free-boundary partial differential problem by judiciously choosing the initial conditions.

Consider the case of an horizontal boundary at l . If the starting point is x at time

s , we define δ_{2l-x} as the mirror image of δ_x with respect to the boundary. Forming the new initial distribution $\mu = \delta_x - \delta_{2l-x}$, by linearity, the solution $(t, x) \mapsto p_t(x)$, to the free-boundary heat equation is the sum of the two solutions corresponding to the initial conditions δ_x and $-\delta_{2l-x}$. The symmetry of the problem entails the absorbing conditions, i.e. $p_t(l) = 0$ for all $t > s$ and $(t, y) \mapsto p_t(y)$, and is the solution to the absorbing boundary problem.

Having informally set the framework of the method of images, we can turn to the first-passage problem with an affine barrier $L(t) = l + \alpha(t - s)$. Extending analytically the Green function κ over the entire real line, we infer the initial distribution of the corresponding free-boundary problem by investigating for a fixed (s, x) , the quantity

$$\lim_{t \rightarrow s^+} \kappa(s, x; t, y), \quad (3.10)$$

defined in the sense of the limit in distribution. This yields the expression

$$\mu = \delta_x - e^{-2\alpha(x-l)} \delta_{2l-x} \quad (3.11)$$

for the initial distribution, which, once convolved with the heat kernel k , provides us with the correct solution to the problem with absorbing boundary conditions. Furthermore, this approach provides a sufficient condition about the initial distribution μ for the function $(t, y) \mapsto p_t(y)$ to be solution of a first-passage problem with a given affine boundary. For any initial arbitrary probability density μ^+ defined on $(-\infty, l]$, setting $\mu = \mu^+ + \mu^-$ with

$$\mu^-(dx) = -e^{-\frac{2\alpha}{\Gamma}(x-l)} \mu^+(2l - dx).$$

ensures that the fundamental solution to the free-boundary heat equation has the property $p_t(L(t)) = 0$.

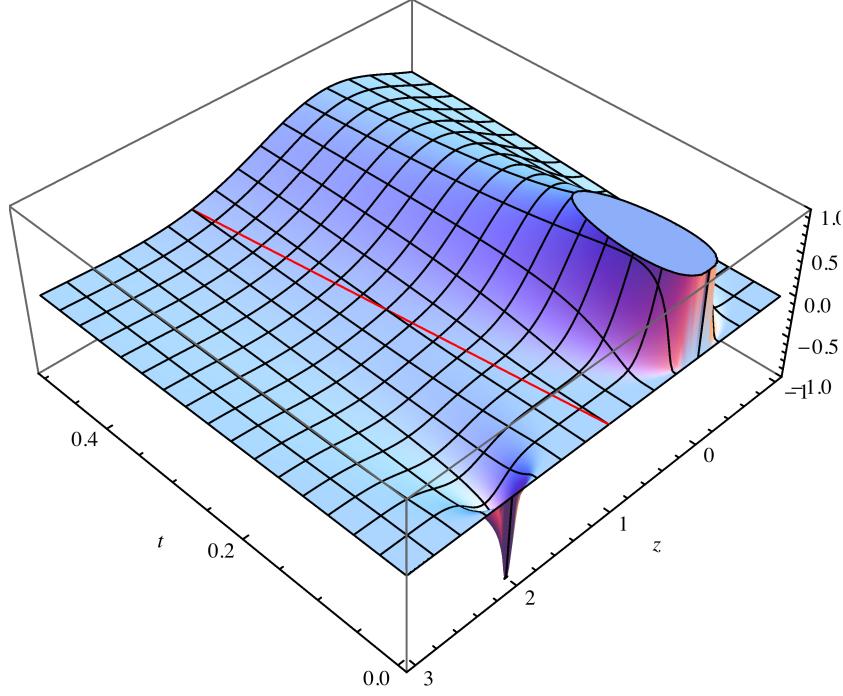


Figure 3.2: Density kernel κ solution of the heat equation with a two-piece linear absorbing boundary given by $B_t = 1 + t$.

To summarize, the theory of partial differential equations provides us with a flexible setting to study the distribution of first-passage times from the solutions of the heat equation with absorbing boundary conditions on the barrier and positive constraint on the lower half-plane. Central to the successful use of this framework is checking for the admissibility of a unique solution to the previous problem for different barrier profiles. In the positive alternative, if the solution is semi-derivable under the barrier, we can equate the opposite of the probability flux through the barrier with the first-passage density. However, proving the existence and unicity of such solutions often demands a reformulation of the partial differential problem into a more amenable integral equation.

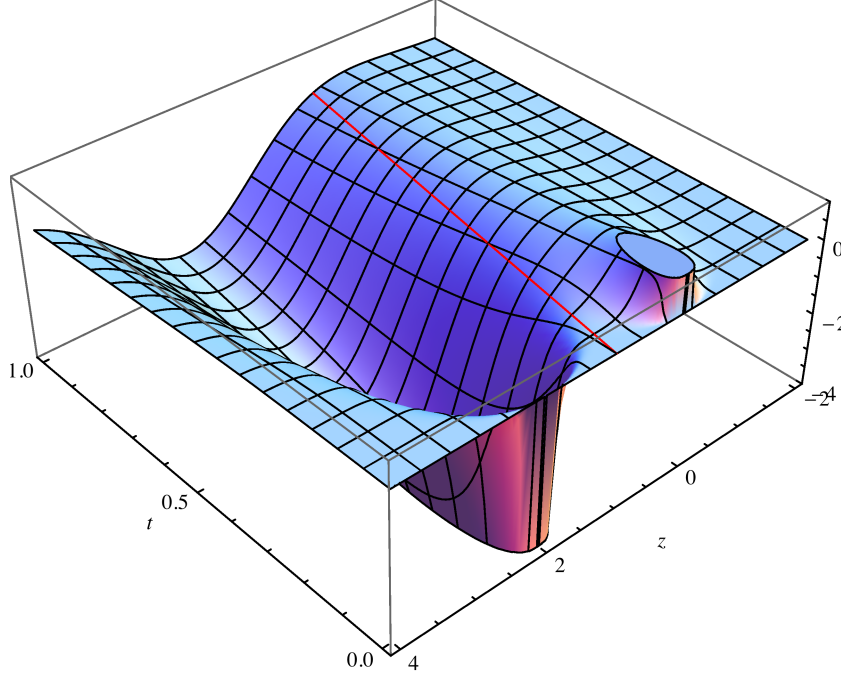


Figure 3.3: Density kernel κ solution of the heat equation with an affine absorbing boundary given by $B_t = 1 - t$.

3.1.3 Integral Equation Formulation

We finish the introduction to our topic by discussing the theory of integral equations in connection to the first-passage problem. The idea of establishing integral equations to compute first-passage time properties (such as integral representations of its moment) dates back the work of Siegert [206, 52] and is still a very active field motivated by financial applications [6, 88, 161]. In addition to naturally arising in the context of the heat equation problem, this theory bears its own interest for its numerical merit. Indeed, as solving the first-passage problem quickly becomes intractable analytically, it is desirable to develop a theoretical framework suited to yield efficient numerical methods. The theory of integral equation provides us with such a framework [200]. The occurrence of integral equations for the cumulative distribution of the first-passage time of a Wiener process naturally stems from probabilistic arguments that we recount in the following. Consider the event $\{W_t > x\}$ for an affine barrier L

satisfying $x > L(t)$. Then, the first-passage time τ with L occurs certainly before t and we can condition this event with respect to τ , which yields

$$\mathbb{P}(W_t > x) = \mathbb{E} [\mathbb{P}(W_t > x \mid \tau)] = \int_0^t \mathbb{P}(W_t > x \mid \tau = s) q(s) ds$$

where q denotes the first-passage time probability. Using the strong Markov property, on $\{\tau = s\}$, we can disregard the past-trajectory of W and equate the probabilities $\mathbb{P}(W_t > x \mid \tau = s)$ and $\mathbb{P}(W_{t-s} > x - L(s))$. Writing the normal cumulative distribution Φ , we end up with:

$$\Phi\left(\frac{x}{\sqrt{t}}\right) = \int_0^t \Phi\left(\frac{x - L(s)}{\sqrt{t-s}}\right) q(s) ds.$$

Since Φ and L are smooth, we can make the arbitrary value x tend toward the barrier $L(t)$ by superior value and, through the dominated convergence theorem, we get the following integral equation [158, 159]:

$$\Phi\left(\frac{L(t)}{\sqrt{t}}\right) = \int_0^t \Phi\left(\frac{L(t) - L(s)}{\sqrt{t-s}}\right) q(s) ds. \quad (3.12)$$

Notice that we recover the reflection principle for a constant threshold at level l .

$$\Phi\left(\frac{l}{\sqrt{t}}\right) = 2\mathbb{P}(\tau < t).$$

Moreover, if the barrier has a non-zero slope, the integral becomes of the convolution type and can be solved by the Laplace transform method.

When the region of integration is variable, an integral equation is said to be a *Volterra equation* [129]. In our case, equation (3.12) pertains to the class of Volterra equations of the *first-kind* since the unknown function q only appears under the integral sign. Studying such an equation is generally cumbersome and it is usually more convenient

to transform it into a Volterra equation of the *second-kind*, which is of the general form

$$f(t) = q(t) + \int_0^t K(t, s, q(s)) ds$$

where f and K are given functions and q the unknown. For instance, when concerned with (3.12), we get an equation of the second-type for q by differentiating with respect to t leading to

$$\frac{d}{dt} \left[\Phi \left(\frac{L(t)}{\sqrt{t}} \right) \right] = \frac{q(t)}{2} + \int_0^t \frac{d}{dt} \left[\Phi \left(\frac{L(t) - L(s)}{\sqrt{t-s}} \right) \right] q(s) ds.$$

Integral equations of the second-kind admit a unique solution under rather mild conditions of regularity on the kernel K . In particular, it can be shown that equation (3.12) admits a unique continuous solution despite the fact that the kernel K exhibits a singularity for $s \rightarrow t^-$. It is crucial to bear in mind the importance of the regularity of the kernel since the Volterra kernels intervening in the first-passage integral always implicitly depends on the barrier, and thus inherits its regularity properties.

Depending on the intent, it is often possible to further transform integral equations into new ones, potentially more amenable to numerical simulation or analysis. For instance, the system of integral equation arising from iterated differentiation and integration under the sign integral in (3.12) has been studied in depth [160].

More recently, a general probabilistic method has been developed to elaborate integral equations for first-passage time density [104], from the point of view of the martingale theory [184]. The starting point is to realize that, if we define the random process

$$X_s = \Phi \left(\frac{x - W_s}{\sqrt{t-s}} \right) \tag{3.13}$$

equation (3.12) can be written as an expectation

$$\mathbb{E} \left(X_\tau \mathbb{1}_{\{\tau \leq t\}} \right) = X_0, \quad (3.14)$$

where X_s is actually a martingale process, that is, for all u, v such that $0 < u < v < t$, we have $\mathbb{E}(X_v | X_u) = X_u$. This martingale property is a direct consequence of the fact that the function $(x, t) \mapsto \Phi(x/\sqrt{t})$ satisfies the heat equation.

When dealing with martingales, we are provided with the powerful optional sampling theorem which, in our situation, states that the martingale relation still holds when sampled at the random time τ

$$\mathbb{E}(X_\tau) = X_0, \quad (3.15)$$

a statement that is very close in writing to (3.14). Then, the next question is the following. Is it possible to construct martingales from W so that we can deduce an integral equation from the optional sampling theorem? We always have

$$\mathbb{E}(X_\tau) = \mathbb{E}(X_\tau \mathbb{1}_{\{\tau \leq t\}}) + \mathbb{E}(X_\tau \mathbb{1}_{\{\tau > t\}})$$

so that the martingale relation entails a Volterra equation for the first-passage time density as soon as $\mathbb{E}(X_\tau \mathbb{1}_{\{\tau > t\}}) = 0$. If X is positive, this is equivalent to $X_\tau \mathbb{1}_{\{\tau > t\}} = 0$ holding almost surely and it is necessarily true for X_s defined as in (3.13). Indeed, reasoning on the event $\{\tau > t\}$, we see that since $W_t < x$ we have

$$\lim_{s \rightarrow t} \frac{W_t - x}{\sqrt{t - s}} = -\infty,$$

which directly entails the desired result since $\lim_{x \rightarrow -\infty} \Phi = 0$. Accordingly, there are as many integral equations of the type (3.15) as there are ways to construct martingales X_s as a parametric function of t, s and W_s that satisfies the criterion $X_\tau \mathbb{1}_{\{\tau > t\}} = 0$ with probability one. It turns out that such martingales can be char-

acterized and admit a known representation, leading to the generation of a class of Volterra equation for the first-passage time of a Wiener process that includes all previously known ones.

3.2 Paradigm of a Piecewise Linear Barrier

If the first-passage problem is simply formulated, considering its most basic exemplar, a Wiener process and an affine barrier, already requires sophisticated mathematical concepts. Once equipped with an analytical solution, we can reformulate such a problem in the field of analysis, whether taking the point of view of partial differential equation or integral equation. In this regard, it is important to realize that the analytical points made in the previous section do not depend on the form of the barrier. They are thus readily generalizable to arbitrary barriers, as long as these barriers satisfy minimal regularity conditions as discussed later.

Our interest lies primarily in investigating how the regularity of the barrier affects the property of the law of first-passage time. By regularity of the barrier, we mean more than merely piecewise derivability since we are including barriers that result from the integration of noisy input, and as such, are nowhere differentiable. We quantify the regularity of the barrier L , assuming this property to be homogeneous, through the Hölder exponent $H > 0$ defined as:

$$H = \inf \left\{ h > 0 \mid \exists \delta > 0, \sup_{0 \leq s, t \leq 1} \frac{|L(t) - L(s)|}{|t - s|^h} < +\infty \right\}.$$

We say that a function admitting a Hölder exponent $H > 0$ is H -continuous. Intuitively, the Hölder exponent indicates how strongly the profile of a continuous curve deviates from smoothness: the lower the exponent is, the more locally singular is the curve. Recall that if $H > 1$, the barrier is differentiable and it fails to be so for $H < 1$. In the limit $H \rightarrow 0$, it actually becomes discontinuous. A central example is the Wiener sample path representing the Brownian motion trajectory. It is known to have an exponent $H = 1/2$, which is also the Hurst roughness exponent characterizing its local scale invariance.

We can approximate any H -continuous barrier by a piecewise linear approximation,

interpolating the original function at successive times. For a rough barrier with a Hölder exponent H such that $0 < H < 1$, the slope of these approximations displays discontinuities at interpolating points that tend to diverge as the resolution of the interpolation becomes finer. The speed at which the slope's jumps increase while refining the interpolation, characterizes the underlying curve's Hölder regularity.

The distribution of the first-passage time to these piecewise linear approximations provides us with a rigorous estimate of the distribution of the original first-passage. Intending to capture the effect of the Hölder regularity of the barrier onto the first-passage time, we propose to first investigate the effect of slope discontinuity on its distribution. Fortunately, we can carry out this program analytically for the simplest case of a two-piece linear barrier.

3.2.1 Effect of the Slope Discontinuity of the Barrier

As in the case of an affine barrier, we can define the transition kernel of the Wiener process killed on a two-piece linear barrier. Specifically, consider the barrier L satisfying

$$L(t) = \begin{cases} L_x \cdot \frac{t_y - t}{t_y - t_x} + L_y \cdot \frac{t - t_x}{t_y - t_x} & \text{if } t_x \leq t \leq t_y \\ L_y \cdot \frac{t_z - t}{t_z - t_y} + L_z \cdot \frac{t - t_y}{t_z - t_y} & \text{if } t_y \leq t \leq t_z \end{cases}.$$

and denote the two different values taken by the slope as:

$$\alpha = L_y - L_x/t_y - t_x, \quad \beta = \frac{L_z - L_y}{t_z - t_y}.$$

For this simple barrier, the transition kernel of the killed process $\rho_{\alpha,\beta}$ can be evaluated formally by direct integration. For time $t < t_y$, $\rho_{\alpha,\beta}$ simply amounts to the kernel of

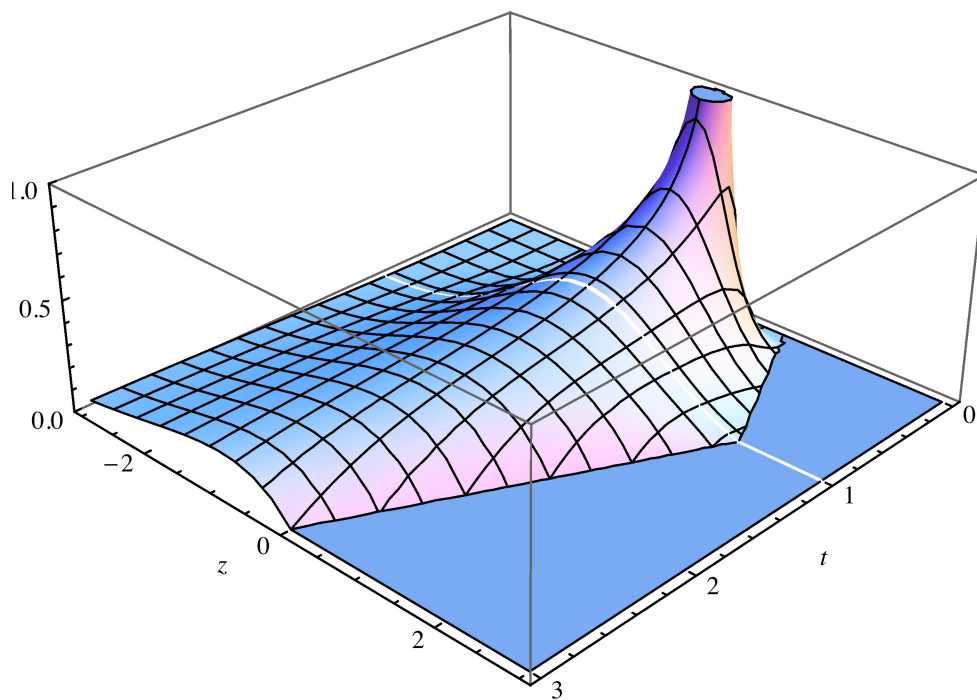


Figure 3.4: Density kernel $\rho_{\alpha,\beta}$ solution of the heat equation with a two-piece linear absorbing boundary given by $L_t = 1 + t$ for $t \leq 1$ and $L_t = 3 - t$ for $t > 1$.

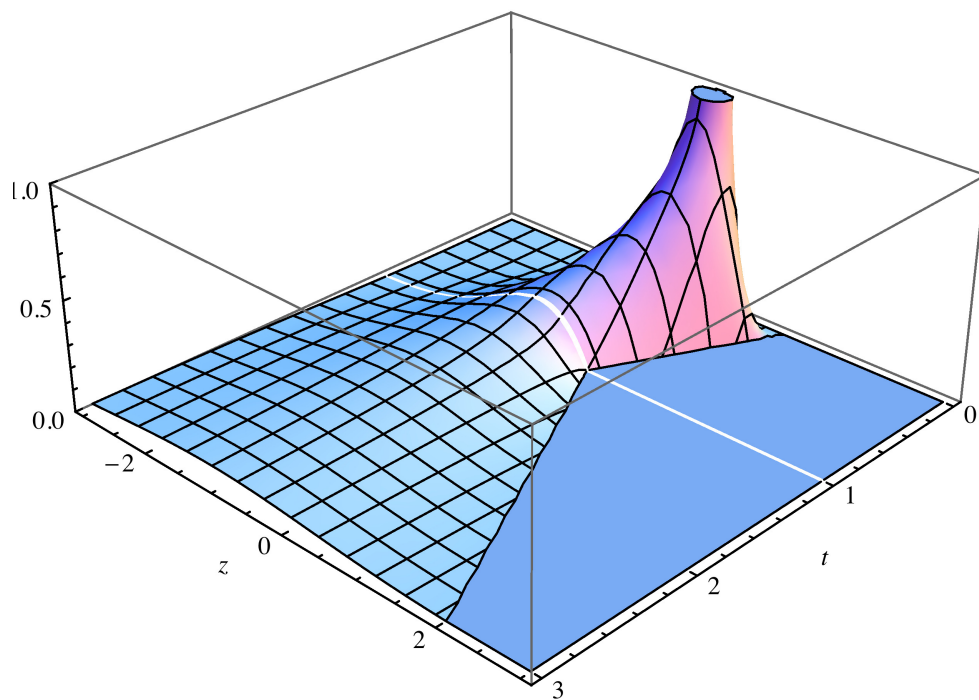


Figure 3.5: Density kernel $\rho_{\alpha,\beta}$ solution of the heat equation with a two-piece linear absorbing boundary given by $L_t = 1 - t$ for $t \leq 1$ and $L_t = t - 1$ for $t > 1$.

a killed process with a perfectly linear barrier, whereas for later time, $\rho_{\alpha,\beta}$ is

$$\rho_{\alpha,\beta}(t_x, x; t_z, z) = \int_{-\infty}^{L_y} \kappa_{\alpha}(t_x, x; t_y, y) \kappa_{\beta}(t_y, y; t_z, z) dy ,$$

where κ_{α} is the kernel corresponding to the affine barrier $t \mapsto L_x + \alpha t$ for $t < t_y$, and κ_{β} is the kernel corresponding to the affine barrier $t \mapsto L_y + \alpha(t - t_y)$ for $t > t_y$. After direct evaluation, the kernel $\rho_{\alpha,\beta}$ can be written under the form

$$\rho_{\alpha,\beta}(t_x, x; t_z, z) = k(t_x, x; t_z, z) \left(\sum_{k=1}^{k=4} \iota_k(t_x, x; t_z, z) \right) , \quad (3.16)$$

where the sum of terms ι is the survival probability of the killed process W conditioning to $W_{t_x} = x$ and $W_{t_z} = z$. The term ι_1 gives the probability that a sample path satisfies $W_{t_y} \leq L_y$ and reads

$$\iota_1(t_x, x; t_z, z) = \Phi \left(\frac{L_y - \mu_y}{\sigma_y} \right) ,$$

with Φ denoting the cumulative function of the normal probability and with

$$\sigma_y = \sqrt{\frac{(t_y - t_x)(t_z - t_y)}{t_z - t_x}} \quad \text{and} \quad \mu_y = \frac{t_z - t_y}{t_z - t_x} x + \frac{t_y - t_x}{t_z - t_x} z . \quad (3.17)$$

The term ι_2 and ι_3 gives the probability that a sample path satisfying $W_{t_y} \leq L_y$ crosses the boundary respectively before t_y and after t_y

$$\begin{aligned} \iota_2(t_x, x; t_z, z) = & -\Phi \left(\frac{L_y - \mu_y}{\sigma_y} - 2a_x \frac{L_x - x}{\sigma_y} \right) \\ & \exp \left(-2 \frac{(L_x - x)(L_y - (a_x L_x + a_z z))}{(t_y - t_x)} \right) , \end{aligned}$$

$$\begin{aligned}\iota_3(t_x, x; t_z, z) = & -\Phi\left(\frac{L_y - \mu_y}{\sigma_y} - 2a_z \frac{L_z - z}{\sigma_y}\right) \\ & \exp\left(-2 \frac{(L_z - z)(L_y - (a_x x + a_z L_z))}{(t_z - t_y)}\right),\end{aligned}$$

whereas the last term indicates the probability that crossings happens both before and after t_y with $W_{t_y} \leq L_y$

$$\begin{aligned}\iota_4(t_x, x; t_z, z) = & \Phi\left(\frac{L_y - \mu_y}{\sigma_y} - 2a_x \frac{L_x - x}{\sigma_y} - 2a_z \frac{L_z - z}{\sigma_y}\right) \\ & \exp\left(-2(L_y - a_x L_x - a_z L_z) \left(\frac{L_x - x}{(t_y - t_x)} + \frac{L_z - z}{(t_z - t_y)}\right)\right).\end{aligned}$$

with the coefficients a_x, a_z defined as follows

$$a_x = \frac{t_z - t_y}{t_z - t_x} \quad a_z = \frac{t_y - t_x}{t_z - t_x}$$

We can easily verify that if $z = L_z$, the terms ι_1 and ι_3 are equal to each other as well as the terms ι_2 and ι_4 , which is consistent with the absorbing barrier constraint, i.e. $\rho(t_x, x; t_z, L_z) = 0$. Moreover, placing ourselves in the affine case by setting $L_y = a_x L_x + a_z L_z$, we retrieve the usual conditional survival probability:

$$1 - \exp\left(\frac{-2(L_z - z)(L_x - x)}{t_z - t_x}\right).$$

Figure 3.4 and Figure 3.5 represents the kernel $\rho_{\alpha, \beta}$ for values $\alpha = 1, \beta = -1$ and $\alpha = -1, \beta = 1$ respectively.

We now study the first-passage density of the Wiener process killed onto the piecewise linear boundary, with a special interest for the effect of the slope discontinuity. By standard manipulation of the heat equation with absorbing boundary condition, we see that the distribution of first-passage time of a Wiener process starting in (t_x, x)

admits the probability density

$$q_{\alpha,\beta}(t_x, x; t_z) = -\frac{1}{2} \left[\frac{\partial}{\partial z} \rho_{\alpha,\beta}(t_x, x; t_z, z) \right]_{z=L(t_z), t_z > t_y}, \quad (3.18)$$

where $t_z > t_x$ is the time of absorption. Applying relation (3.18) to the analytical probability distribution $\rho_{\alpha,\beta}$, we obtain

$$\begin{aligned} q_{\alpha,\beta}(t_x, x; t_z) = & \quad (3.19) \\ & e^{-\frac{(L_x - x + \alpha(t_y - t_x) + \beta(t_z - t_y))^2}{2(t_z - t_x)}} \left[\frac{(L_x - x + (\alpha - \beta)(t_y - t_x)) \Phi\left(\frac{L_x - x + (\alpha - \beta)(t_y - t_x)}{\varsigma(t_z)}\right)}{\sqrt{2\pi(t_z - t_x)^3}} \right. \\ & \left. + e^{-\frac{2(L_x - x)(\alpha - \beta)(t_z - t_y)}{(t_z - t_x)}} (L_x - x + (\beta - \alpha)(t_y - t_x)) \Phi\left(\frac{x - L_x + (\alpha - \beta)(t_y - t_x)}{\varsigma(t_z)}\right) \right] \end{aligned}$$

with

$$\varsigma(t_z) = \sqrt{\frac{(t_z - t_x)(t_y - t_x)}{t_z - t_y}}.$$

This density is continuous over the whole time-line and is smooth away from the slope discontinuity. We have represented such curves in Figure 3.6 for varying slope α and in Figure 3.7 for varying slope β .

As expected, since a Wiener process is Markov, the density of first-passage times is not affected for time preceding the slope's discontinuity but is greatly affected by the change of slope for later time. Obviously if the barrier's slope is steeper on the second segment than on the first segment, that is if $\beta < \alpha$ the probability density is expected to locally increased in the right half-vicinity of t_y . A Taylor expansion of expression (3.19) at t_y for $t_z = t_y + \Delta t$ with $\Delta t > 0$ yields

$$\begin{aligned} q_{\alpha,\beta}(t_x, x; t_y + \Delta t) = & k(t_x, x; t_y, L_y) \left(\frac{L_x - x}{t_y - t_x} + \frac{2\sqrt{2}(L_x - x)(\alpha - \beta)}{\sqrt{\pi}(t_y - t_x)} \sqrt{\Delta t} \right. \\ & \left. - \left[\frac{(L_x - x)^3}{(t_y - t_x)^3} + \frac{(L_x - x)(3(\alpha - \beta)^2 - \beta^2 - 3)}{(t_y - t_x)^2} \right] \Delta t \right) + O(\Delta t^{3/2}), \end{aligned} \quad (3.20)$$

where k is the heat kernel. We notice that the probability density is not derivable on the right side of the discontinuity. It exhibits a singularity in $\sqrt{\Delta t}$, i.e. of power $1/2$, with a pre-factor of the same sign as $\alpha - \beta$, which is consistent with our previous intuitive observation. Moreover, we check that if $\alpha = \beta$, we recover the first-terms of the Taylor expansion at t_y for $t_z > t_y$ of the probability density of first-passage times for a linear barrier with slope α

$$q_\alpha(t_x x; t_z) = \frac{L_x - x}{\sqrt{2\pi(t_z - t_x)^3}} \exp\left(-\frac{(L_x - x + \alpha(t_z - t_x))^2}{2(t_z - t_x)}\right).$$

Actually, if the Taylor expansion (3.20) is continued for higher order, we see that the series exhibits terms at every integer and half-integer order, with the property that every half-integer term is zero when $\alpha = \beta$ (we recover the Taylor expansion of $t_z \mapsto q_\alpha(t_x, x; t_z)$, which is infinitely derivable at $t_y > t_x$). Thus, it appears that the presence of a slope discontinuity perturbs locally the otherwise smooth density by the addition of singular terms of half-integers orders.

This observation already allows us to develop some intuition about how the first-passage density is affected by the local Hölder regularity of the barrier. Suppose that the two-piece barrier results from the interpolation of the function

$$t \mapsto L_h^\pm(t) \begin{cases} 1 & \text{if } t < t_y \\ 1 \pm (t - t_y)^h & \text{otherwise} \end{cases},$$

that exhibits a local Hölder singularity of strength h in t_y . Assume also that the sample time t_z is such that $t_z - t_y = \Delta t$ tends to zero. Then, the slope β scales as Δt^{h-1} and diverges when refining the interpolation.

It is natural to investigate the condition under which the Taylor expansion of the density $q_{\alpha,\beta}$ remains finite through this scaling process. This ensures that, despite

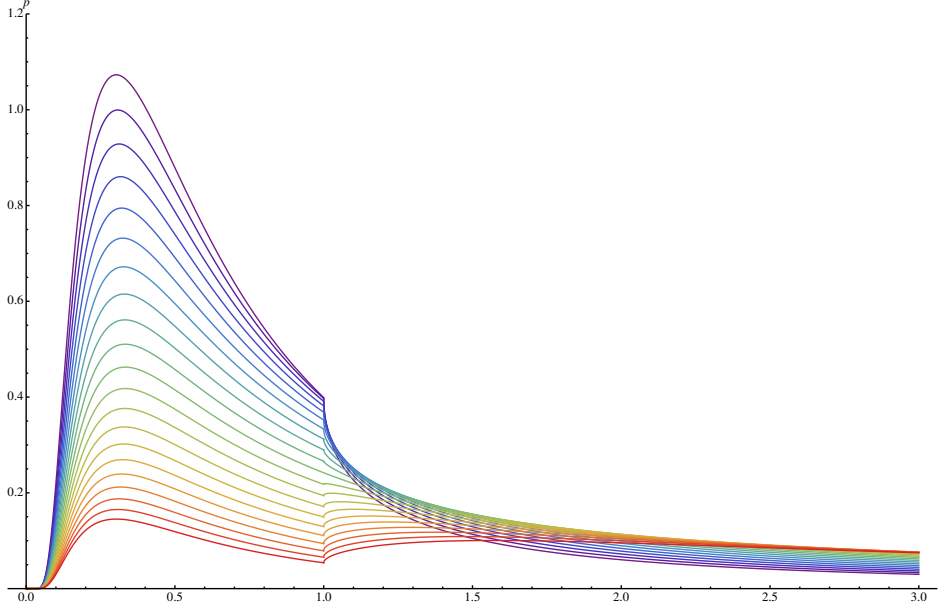


Figure 3.6: Probability current through the two-piece linear absorbing boundary given by $L_t = 1 + \alpha t$ for $t \leq 1$ and $L_t = 2 + \alpha - t$ for $t > 1$, with $\alpha = -10, -9 \dots 10$.

the perturbation entailed by the change of slope, the first-passage time still admits a finite non-zero density after the discontinuity. Since expression β only intervenes in the Taylor expansion in powers of $\beta\sqrt{\Delta t}$, which scales as $\Delta t^{h-1/2}$, the series converges for Hölder singularity $h > 1/2$ but diverges for $h < 1/2$. As one expects, the density appears to be well-defined if its singularity is weaker than the typical roughness of a Brownian trajectory. In the case of a stronger singularity $h < 1/2$, the density seems to exhibit a singularity itself: it should diverge in t_y^+ for L_h^- and becomes zero in t_y^+ for L_h^+ (essential singularity).

However, this situation is more complex than it looks. For instance, we have completely disregarded the fact that the slope discontinuity $\beta - \alpha$ appears in expression (3.20), thus showing that the behavior after the discontinuity is affected by the profile of the curve before the discontinuity.

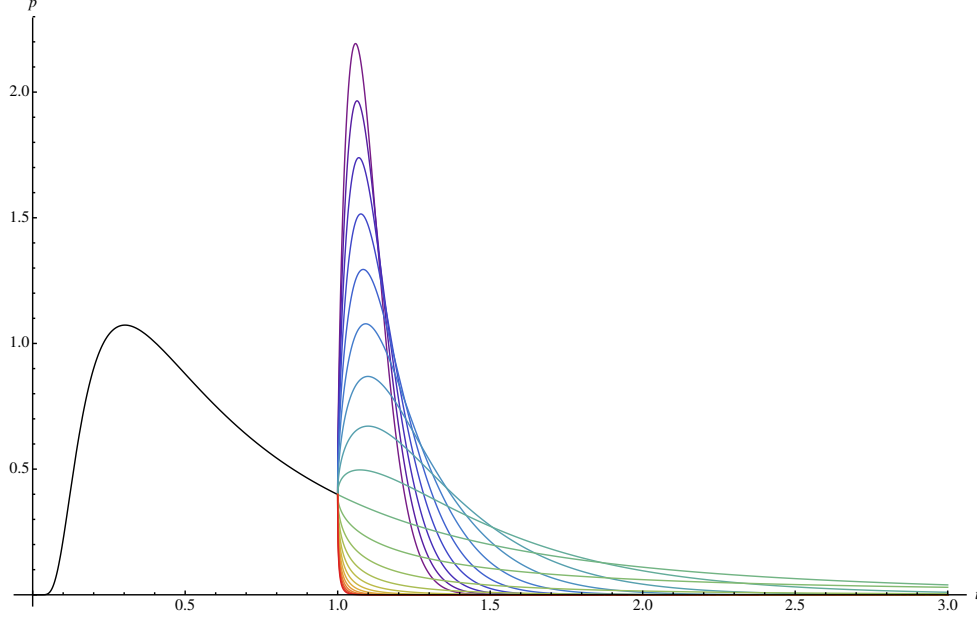


Figure 3.7: Probability current through the two-piece linear absorbing boundary given by $L_t = 1 - t$ for $t \leq 1$ and $L_t = \beta(t - 1)$ for $t > 1$, with $\beta = -10, -9 \dots 10$.

3.2.2 Effect of the Initial Probability Distribution

To further understand the effect of a slope discontinuity, we observe that, since the process cannot anticipate a jump in the slope, its probability density at the time of the discontinuity is the same as if the process was being absorbed on a straight barrier. This probability satisfies the heat equation with corresponding absorbing boundary conditions, thus constraining its shape. In turn, the sudden change of boundary conditions causes the appearance of a local singularity in the first-passage density, seen as the flux of probability at the barrier.

Here, instead of studying directly this effect, we take a more general reverse approach, considering an arbitrary analytical distribution $y \mapsto p_{t_y}(y)$ at a given time t_y (corresponding to the time of the discontinuity) and a barrier $t \mapsto L_y + \beta(t - t_y)$ for the ensuing propagation. The density probability of first-passages for a future time

$t_z > t_y$ is given by the integral

$$q(t_z) = \int_{-\infty}^{L_y} q_\beta(t_y, t_z, x) p_{t_y}(y) dy,$$

which is well-defined everywhere but is not analytical for $t_z \rightarrow t_y^+$.

However, for small time intervals $\Delta t = t_z - t_y > 0$, we can safely assume that only the sample paths that are in the vicinity of the barrier at t_y are likely to cross the barrier before t_z . Accordingly, the behavior of q in the right half-neighborhood of t_y is dictated by the local behavior of p_{t_y} in the vicinity of the barrier at L_y . To verify this point, we use the analytical expansion of p_{t_y} in L_y (p_{t_y} is analytic for piecewise linear barriers) and write

$$p_{t_y}(y) - \sum_{k=1}^{n-1} p_{t_y}^{(k)}(L_y) \frac{y^k}{k!} = y^n h(y), \quad M_n^h = \sup_{y < 0} h(y) < \infty.$$

A change of variable shows that

$$\int_{-\infty}^0 \frac{|x| e^{-\frac{x^2}{2}}}{\sqrt{2\pi t^3}} |x|^n h(x) dx = t^{\frac{n-1}{2}} \int_{-\infty}^0 |u|^{n+1} e^{-\frac{u^2}{2}} h(\sqrt{t}u) dy \leq M_n^h (\sqrt{t})^{n-1},$$

because the integral on the right-hand side is convergent. Incidentally, we have that

$$q_\beta(t_z) = \sum_{k=1}^3 p_{t_y}^{(k)}(L_y) \int_{-\infty}^{L_y} q_\beta(t_y, t_z, y) \frac{y^k}{k!} dy + O((t_z - t_y)^{3/2}),$$

and now remains to express the integral coefficients of the expansion to study the behavior of the density q . It is possible to give a general formula for these coefficients through a recurrence argument. However, this is beyond the scope of the present

discussion, and direct integration yields

$$q(t_z) = -\frac{1}{2}p_{t_y}^{(1)}(L_y) - \sqrt{\frac{2}{\pi}} \left(\beta p_{t_y}^{(1)}(L_y) + \frac{1}{2}p_{t_y}^{(2)}(L_y) \right) \sqrt{\Delta t} + \quad (3.21)$$

$$-\frac{1}{2} \left(\beta^2 p_{t_y}^{(1)}(L_y) + \frac{1}{2} \left(3\beta p_{t_y}^{(2)}(L_y) p_{t_y}^{(3)}(L_y) \right) \right) \Delta t + O(\Delta t^{3/2}),$$

where $t_z - t_y = \Delta t > 0$ and $p^k(L_y)$, $k \geq 1$ denotes the derivative of p_{t_y} of order k . It is apparent that for an arbitrary analytical probability density at time t_y , terms of half-integer order in Δt naturally arise in the Taylor expansion of the density of first-passage times.

We now intend to figure out the conditions under which these quantities are non-zero. Suppose that the barrier is perfectly linear of slope β for every $t_y > t_x$. By the heat equation, we have that

$$\beta p_{t_y}^{(1)}(L_y) + \frac{1}{2}p_{t_y}^{(2)}(L_y) = \left[\beta \frac{\partial}{\partial y} p_{t_y}(y) + \frac{\partial}{\partial t_y} p_{t_y}(y) \right]_{L(t_y)}. \quad (3.22)$$

The right-hand term of the previous expression appears to be zero since differentiating the absorption constraint $p_{t_y}(L(t_y)) = 0$ with respect to t and with $L(t) = L_y + \beta t_y$ leads to

$$[\nabla_{t,y} p_t(y)]_{(t,L(t))} \perp (1, \beta).$$

Actually, in the Taylor expansion (3.21), the quantity (3.22) factorizes in the coefficient of every half integer term. Therefore, away from the slope's discontinuity, when the absorption constraint is satisfied, the probability density is smooth, no matter what the profile of the initial distribution was at the discontinuity.

Let us then focus on what happens after the discontinuity. For being a solution of the heat equation $(t, z) \mapsto p_{t_z}(z)$ necessarily admits a continuous derivative in z . Moreover, on the open intervals (t_x, t_y) and (t_y, ∞) for which the barrier has respectively a slope of α and β , the corresponding absorbing conditions of type (3.22) are satisfied.

As a result, the time derivative $t_y \mapsto \partial_{t_y} p_{t_y}(y)$ goes through a discontinuity and so does $t_y \mapsto p_{t_y}^{(2)}(y)$, as imposed by the heat equation.

We can then understand the Taylor expansion of the first-passage time density after the discontinuity, by plugging the jump conditions of $p_{t_y}^{(k)}(L_y)$ in expression (3.21). We express the values of $p_{t_y}^{(k)}(L_y)$, $k \geq 1$ by computing the iterated derivatives in y of the transition kernel κ_α for a linear barrier of slope α . Once we have these expressions that are naturally given in terms of α , we inject them in equation (3.21) and we verify that we effectively recover the Taylor expansion of the probability density of first-passages right after the slope discontinuity. For example, we have

$$\left[\frac{\partial}{\partial y} \kappa_\alpha(t_x, x; t_y, y) \right]_{y=L(t_y)} = - \underbrace{\frac{2(L_x - x)}{\sqrt{2\pi^3(t_y - t_x)^3}} \exp\left(-\frac{(L_x - x - \alpha(t_y - t_x))^2}{2(t_y - t_x)}\right)}_{2q_\alpha(t_x, x; t_y)}$$

and we see that the two first-order terms in (3.21) are recovered from

$$\begin{aligned} p^{(1)}(L_y) &= \left[\frac{\partial}{\partial y} p_{t_y}(y) \right]_{L(t_y)} = -2 \int_{-\infty}^{L_x} q_\alpha(t_x, x; t_y) p_{t_x}(x) dx = -2q(t_y), \\ p^{(2)}(L_y) &= -2\alpha p^{(1)}(L_y). \end{aligned}$$

As in the previous section, suppose that the two-piece barrier interpolates some function in t_x , t_y and t_z and that this function has one single Hölder singularity point in t_y of strength $0 < h < 1$. Notice that such a function can be left-Hölder singular as the function

$$t \mapsto L(t) \begin{cases} (t_y - t)^h & \text{if } t < t_y \\ 0 & \text{otherwise} \end{cases}.$$

right-Hölder singular as the function

$$t \mapsto R(t) \begin{cases} 0 & \text{if } t < t_y \\ (t - t_y)^h & \text{otherwise} \end{cases}.$$

or both as the function $t \mapsto |t - t_y|^h$ and $t \mapsto \text{sgn}(t - t_y) |t - t_y|^h$.

The above analysis demonstrates how, at any given time, the transition kernel of a killed Wiener process is affected by the past profile of the barrier. Specifically, in the case of an affine barrier, the absorbing boundary condition enforces through the heat equation a relation between the first and second space derivative of the transition kernel. This relation effectively constrains locally the probability of a sample path to be in the vicinity of the barrier. At the same time, the slope of the barrier can abruptly change, and the diffusion of the Wiener process is constrained by a different relation in the vicinity of the barrier. The local amount of probability that accounts for the discrepancy is responsible for the singular behavior of the first-passage density. The probability to find a sample path near the slope discontinuity is too high (low), which is later eliminated through a local increase (decrease) of the first-passage density in $\sqrt{\Delta t}$.

3.2.3 Heuristic Approach to Barrier's Reachability

In order to develop some intuition about the problem, we are going to break it up into two parts, a geometrical optics part, in which most first passages can be accounted for by simple visibility considerations, and then a diffractive correction in which we take into account that random walkers can turn around corners. The geometrical part is simple: most first passages are generated by the walker running into a hard-to-avoid obstacle, as shown in Figure 3.9. The intuition is that the walkers are moving left to right, rising onto a ceiling from which features are hanging, and as the walkers rise they collide with some feature. The problem is thus twice symmetry-broken:

what matters are local minima of the boundary, not the maxima, which are hard to get into; and the walkers only spontaneously run onto the left flank of a local minimum. Therefore, a good first order approximation follows from observing that most of the first passages occur on the left flanks of local minima, and deeper local minima cast shadows on subsequent less shallow minima close by. However, there is a finite probability that a walker may narrowly avoid a local minimum and pass just under it, only to rapidly rise afterwards and hit the right rising flank of the barrier, as shown in Figure 3.8. This is, effectively, a race between the boundary and the walker: if the walker can rise faster than the boundary, then there is some probability of passage right off the minimum. But if the boundary rises faster than a walker can catch up with it, then the probability of passage right of the minimum can be very small (see 3.8).

Finally, we explain the consequences of our intuitive reasoning on simple examples. Let us consider a simple family of functions that admits one single Hölder singular point of strength $h > 1/2$ at time s . Notice that such a function can be left-Hölder singular as the function

$$t \mapsto L_h(t) \begin{cases} (s-t)^h & \text{if } t < s \\ 0 & \text{otherwise} \end{cases}.$$

right-Hölder singular as the function

$$t \mapsto R_h(t) \begin{cases} 0 & \text{if } t < s \\ (t-s)^h & \text{otherwise} \end{cases}.$$

Based on the previous heuristics, the functions $-L_h$ and R_h rise faster than a typical Brownian sample path is able to, so that the first-passage density vanishes in s^- for $-L_h$ and in t^+ for R_h . Reciprocally, the functions L_h and $-R_h$ locally represent an obstacle that a Brownian sample path cannot escape if it wanders in its vicinity, so

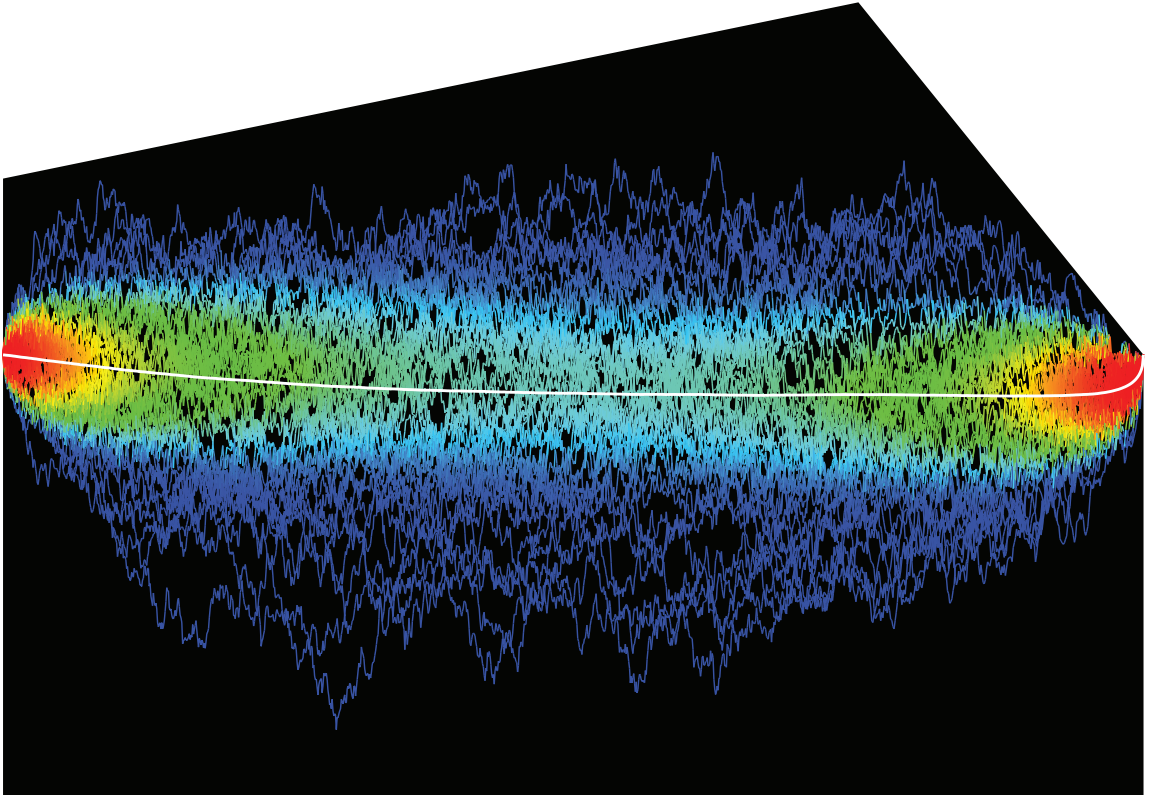


Figure 3.8: Fifty sample paths conditioned to first-hit the boundary at the same time. The color scale codes for the conditional probability of the underlying first-passage bridge process, i.e. the Wiener process conditioned to first hit the barrier at a given time. The white trace is the mean deterministic trajectory, whose slope diverges at the time of first-passage.

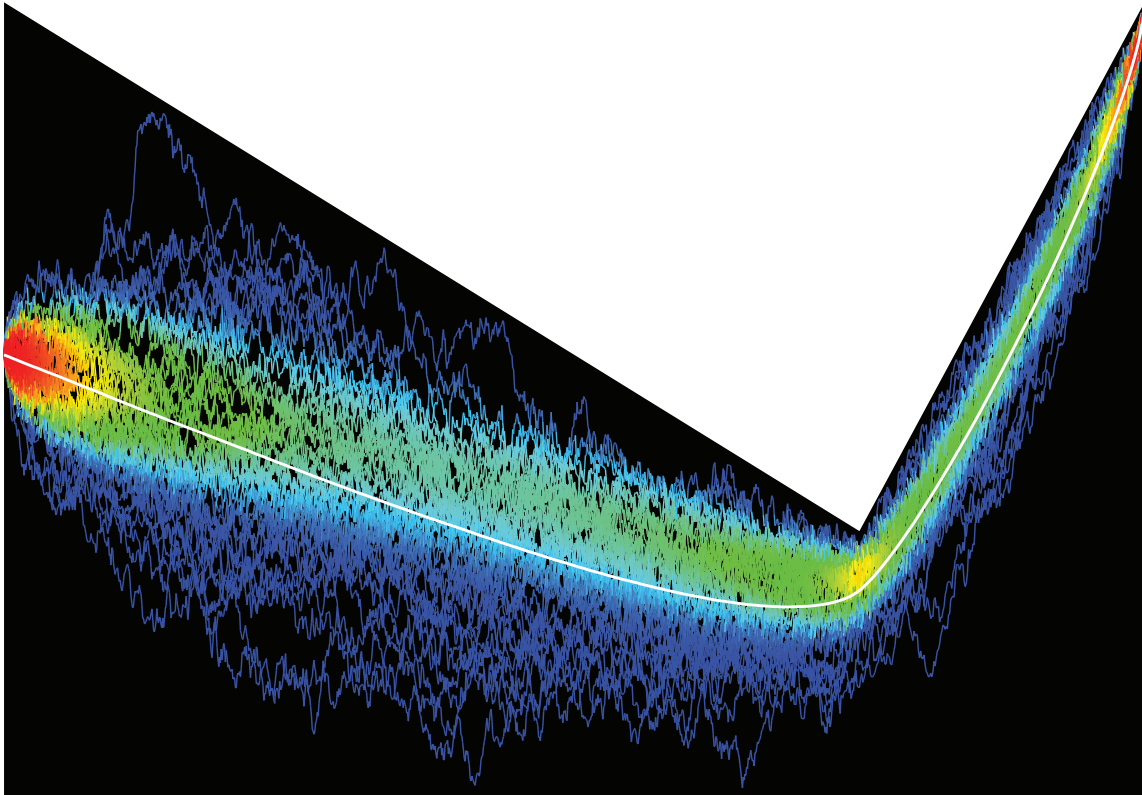


Figure 3.9: Fifty sample paths conditioned to first-hit the boundary at the same time. The color scale codes for the conditional probability of the underlying first-passage bridge process, i.e. the Wiener process conditioned to first hit the barrier at a given time. The white trace is the mean deterministic trajectory, whose slope diverges at the time of first-passage.

that the first-passage density diverges in s^- for L_h and in t^+ for $-R_h$.

The situation is more intricate for barriers that have more general Hölder singularities like the barriers $L_h + R_k$, $L_h - R_k$, $-L_h + R_k$, and $-L_h - R_k$. In this situation, the dissymmetry between left-flank and right flank needs to be taken into account. By the Markov property of the Wiener process, the behavior the first-passage density on the left flank does not depend on the later profile of the barrier. However, the behavior on the right flank can be shadowed or overexposed by the left flank, thus potentially dictating the accessibility to the right flank of the barrier.

In the following, the general issue of investigating the interaction between the strength of singularity on the left side and the strength of the right side is not addressed for $h < 1/2$.

3.3 Absolute Continuity of First-Passage Density and Barrier Regularity

In this section, we provide a mathematical ground to the propositions we have explored intuitively in the case of piecewise linear barriers. Precisely, we want to verify the existence of a density function for the first-passage time of a Wiener process absorbed on a H -continuous barrier with $H > 1/2$, a property that is formally referred to as the absolute continuity of the first-passage time distribution with the Lebesgue measure on the real half-line. In doing so, we are going to alternate between the view of a first-passage problem holding for a canonical Wiener process with a fluctuating barrier and a corresponding first-passage problem holding for a drifted Wiener process and a constant threshold.

These two points of view offer different benefits. In defining Wiener drifted processes, we are led to study killed processes that are defined on the same sample space (the set of continuous function starting at zero that dies the first-time they hit a threshold $l > 0$), and it will prove possible to define a Radon-Nykodim derivative between the measures of the two drifted processes. In turn, the consideration of the fluctuating barrier allows us to use the powerful machinery of the heat equation theory with boundary conditions in conjunction with the theory of Volterra integral equations.

In the following section, we first take the probabilistic approach to the drifted processes for a class of regular drift functions which are connected to analytical tools developed to study killed diffusions [141]. We show that the density of the first-passage to a constant threshold has an expression where a conditional expectation intervenes. As long as the drift is such that the conditional expectation is well-defined, the first-passage always admits a density. These drift functions include curves that have a finite number of singularity points of Hölder type $h > 1/2$; that is, functions whose profiles at singularity times t look like $s \mapsto |s - t|^h$ with $h > 1/2$. In addition to

providing insight into the relation between the local Hölder regularity of the barrier and the existence of a density for first-passages, the analysis formally establishes the result we heuristically derived in the preceding section.

Barriers that are H -homogeneous have an uncountable number of singularity points and we must resort to indirect methods to elucidate this case. Adopting the point of view of a fluctuating barrier, we use the many results known about the heat equation to show that, even in the case of homogeneous Hölder H -functions, a first-passage density is well-defined for $H > 1/2$. These results, that improve on the current literature [125], are rooted in the ground-breaking work of Gevrey [78] for the study of parabolic differential equations, later actualized in a modern form by Rozier [38].

3.3.1 Probabilistic Treatment

Following usual notation, $W = (W_t, \mathcal{F}_t; t \in \mathbb{R}^+)$ denotes a standard one-dimensional real Wiener process with associated natural filtration \mathcal{F}_t . In this section, we identify the probability space on which W is defined with the space of continuous functions on \mathbb{R}^+ starting in zero at time zero and denoted $C_0(\mathbb{R}^+)$, endowed with the σ -algebra \mathcal{B} generated by the canonical cylinder sets:

$$C_{t_1, \dots, t_n}(B_1, \dots, B_n) = \{\omega \in \Omega \mid \omega(t_1) \in B_1, \dots, \omega(t_n) \in B_n\}.$$

and equipped with the W -induced Wiener measure \mathbf{P} . We write $W_t(\omega) = \omega(t)$ and when there is no ambiguity, $\omega(t_i) = \omega_i$. Given a continuous barrier L , L_N is a piecewise linear barrier interpolating L on the dyadic points $D_N = \{k2^{-n}, k \in \mathbb{N}\}$.

Convergent Sequence of First-Passage Time

We intend to investigate the density (if it exists) of the first-passage time $\tau_N = \inf\{t > 0 \mid W_t > L_N(t)\}$ knowing $W_0 = 0 < L(0)$: if L is homogeneously Hölder continuous,

we know that L_N converges uniformly toward L on every finite interval of \mathbb{R}^+ , which implies that τ_N converges in distribution toward $\tau = \inf\{t > 0 \mid W_t > L(t)\}$ as we show in the following section.

Proposition 9. *Given a real continuous process X and a real x_0 , if a sequence of bounded continuous functions $\{L_n\}$ satisfies $L_n(0) > x_0$ and converges uniformly to L on \mathbb{R}^+ , then $\tau_n = \inf\{t > 0 \mid X_t > L_n(t)\}$ converges in law to $\tau = \inf\{t > 0 \mid X_t > L(t)\}$.*

Proof. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $X : \Omega \rightarrow C_{x_0}(\mathbb{R}^+)$ be a real continuous process with natural filtration \mathcal{F}_t taking values in the set of continuous functions satisfying $x(0) = x_0$. Note that for any bounded continuous function L in $C^b(\mathbb{R}^+)$ and for every ω in Ω , the first-passage time $\tau(\omega)$ with the function L can be seen as a real function defined on $C_b(\mathbb{R}^+)$ by $L \mapsto \tau_L(\omega)$. Moreover, in the complete separable metric space $(C_b(\mathbb{R}^+), |\cdot|_\infty)$, it is a continuous function of its argument. Indeed, given a continuous barrier L and a positive real $\epsilon > 0$, by definition of τ , for any ω we have $X_t(\omega) < L(t)$ if $t < \tau_L(\omega) - \epsilon$, so that, by continuity of the sample paths, we can define

$$\sup_{0 < t < \tau_L(\omega) - \epsilon} (L(t) - X_t(\omega)) = \delta^- > 0.$$

In the same way, we can define

$$\sup_{\tau_L(\omega) < t < \tau_L(\omega) + \epsilon} (X_t(\omega) - L(t)) = \delta^+ > 0.$$

Setting $\delta = \min(\delta^+, \delta^-)$, for every L' such that $|L' - L|_\infty < \delta$, we have that $|\tau_L(\omega) - \tau_{L'}(\omega)| < \epsilon$.

As a consequence, since L_n converges uniformly toward L when n tends toward infinity, the previous continuity result implies the point-wise sure convergence

$$\lim_n \tau_n(\omega) = \tau(\omega),$$

which entails the convergence in law. □

Therefore, if there exist densities q_N for τ_N and they admit a limit density $q = \lim_N q_N$, then q is the density of τ by Scheffé's theorem. Accordingly, we restrain our discussion to barriers L that are bounded on \mathbb{R}^+ , which ensures that $\mathbb{P}(\tau_N < \infty) = 1$ so that the distributions of τ_N and τ are proper probabilities.

Probability Space of Killed Processes

Before treating the case of the barrier L_N , consider a Wiener process starting in $W_s = x < L(s)$. Recall the probability that W_t is in dy with $y < L(t)$ without hitting the barrier is

$$\kappa(s, x; t, y) dy = k(s, x; t, y) \left(1 - \exp \left(-\frac{2(L(s) - x)(L(t) - y)}{t - s} \right) \right) dy.$$

where κ is referred to as the killed density transition kernel and where k is the heat kernel. This result is instrumental for the treatment of piecewise linear barriers because it readily shows that τ_N admits a density q_N .

There is no loss of generality in considering that the points at which the piecewise linear barrier exhibits a discontinuity are in the set of dyadic points. Moreover, we assume that this set of points defines a quasi-finite partition \mathcal{P} of \mathbb{R}^+ , which means that, when restrained to a compact interval, \mathcal{P} is finite. Writing the partition \mathcal{P} as an ordered sequence of positive reals $\{s_i\}_{i \in \mathbb{N}}$, the density of first-passage q_N is expressed at time t as

$$q_N(t) = -\frac{1}{2} \left[\frac{\partial}{\partial x} \int_{-\infty}^{L(s_1)} \dots \int_{-\infty}^{L(s_p)} \kappa_N(t_0, \omega_0; s_1, w_1) \dots \kappa_N(s_p, \omega_p; t, x) dw_1 \dots dw_p \right]_{x=L(t)} \quad (3.23)$$

where $p = p(\mathcal{P}, t) = \inf\{i \in \mathbb{N} \mid s_i < t\}$ and where the index N in κ_N is a reminder that the kernel is associated to the barrier L_N . For any $t > 0$, we posit

$$\kappa_N(\mathcal{P}; t, \omega) = \kappa_N(s_0, \omega_0; s_1, w_1) \dots \kappa_N(s_p, \omega_p; t, x)$$

and equation (3.23) can be rewritten

$$\begin{aligned} q_N(t) &= -\frac{1}{2} \left[\frac{\partial}{\partial x} \int_{-\infty}^{L(s_1)} \dots \int_{-\infty}^{L(s_p)} \kappa_N(\mathcal{P}; t, \omega) dw_1 \dots dw_p \right]_{x=L(t)} \\ &= -\frac{1}{2} \left[\frac{\partial}{\partial x} \int_{-\infty}^0 \dots \int_{-\infty}^0 \tilde{\kappa}_N(\mathcal{P}; t, \omega) dw_1 \dots dw_p \right]_{x=L(t)} \end{aligned}$$

where $\tilde{\kappa}$ denotes the transformed kernel obtained by the translated process $\tilde{W}_N = W - L_N$ by a change of variable:

$$\tilde{\kappa}_N(\mathcal{P}; t, \omega) = \kappa_N(s_0, \omega_0 - L(s_0); s_1, w_1 - L(s_1)) \dots \kappa_N(s_p, \omega_p - L(s_p); t, x - L(t)).$$

The auxiliary process \tilde{W}_N is introduced so that, instead of studying killed processes defined on different state spaces corresponding to different barriers L_N , we can consider translated versions of the Wiener process W that are all similarly killed when reaching zero.

To rigorously define the measure of the processes \tilde{W}_N killed at zero, we introduce $\tau_0(\omega) = \inf\{t > 0 \mid w_t > 0, w_0 = -L_N(0) = -L(0)\}$. Since we disregard the trajectory of ω after it dies at $\tau_0(\omega)$, we identify all sample-paths that are equal up to their first-passage. The equivalence relationship \mathcal{R} corresponding to this identification defines a quotient space

$$\Omega_0 = \Omega / \mathcal{R} \sim \cup_{t \in \mathbb{R}^+} \{ \{ \omega_s; 0 \leq s \leq t \} \mid \omega \in \Omega, \tau_0(\omega) = t \}.$$

The probability law \mathbf{P}_N of \tilde{W}_N induces a measure P_N on Ω_0 equipped with the Borelian sets \mathcal{B}_0 defined by the quotient topology: denoting the quotient projection map $j : \Omega \rightarrow \Omega_0$, we posit

$$\forall B \in \mathcal{B}_0, \quad P_N(B) = \mathbf{P}_N(j^{-1}(B)).$$

Because it is defined through the continuous surjection j , \mathcal{B}_0 appears as the σ -algebra generated by the killed cylinder sets $C_{t_1, \dots, t_n; \tau}(B_1, \dots, B_n; B_\tau)$ of the form

$$\left\{ \{\omega_s; 0 \leq s \leq \tau_0(\omega)\} \mid \omega \in \Omega, \omega(t_1) \in B_1, \dots, \omega(t_n) \in B_n; \tau_0 \in t_n + B_\tau \right\},$$

where B_1, \dots, B_n are in $\mathcal{B}(\mathbb{R}^-)$ and B_τ is in $\mathcal{B}(\mathbb{R}^+)$. Under the probability law P_N , any such killed cylinder set has measure

$$P_N(C_{t_1, \dots, t_n; \tau}(B_1, \dots, B_n; B_\tau)) = \int_{t_n + B_\tau} \left(\int_{B_1} \dots \int_{B_n} \partial_x \tilde{\kappa}_N(\mathcal{P}; \tau, \omega) dw_1 \dots dw_p \right) d\tau, \quad (3.24)$$

where $\mathcal{P}(\tau) = \{t \in \mathcal{P} \mid t < \tau\}$. Let us define the cross-section B_t of a measurable set B in \mathcal{B}_0 as the measurable limit set

$$B_t = B \cap j(\{\tau = t\}) = \bigcap_{n \leq 0} j(B \cap \{\tau \in [t, t + 2^{-n})\}).$$

on which the finite-dimensional kernels $\tilde{\kappa}_N(\mathcal{P}(t); t, \cdot)$ are naturally defined for finite partitions $\mathcal{P}(t)$ comprising times inferior to t . Reasoning on $\Omega_t = j(\{\tau = t\})$, we use the Kolmogorov-Čentsov extension theorem to construct the conditional measures $P_{N,t}$ from the consistent finite-dimensional distributions $P_{N,t}^{\mathcal{P}}$ associated with the kernels $\tilde{\kappa}_N(\mathcal{P}(t); t, \cdot)$. Then, the formulation (3.24) of the measure P_N on cylinder sets makes apparent that the cross-section measure Σ_N of measurable sets B_t for B in \mathcal{B}_0

is naturally defined in Ω_0 as

$$\Sigma_N(B_t) = \int_{B_t} P_{N,t}(d\omega).$$

This provides us with a natural decomposition of the measure P_N in a family of conditional measures $P_{N,t}$ with respect to τ_0 . This satisfies $dP(t, \omega) = P_{N,t}(d\omega) dt$. Notice that the cross-section measure of Ω_t appears as the density of τ_0 under \mathbf{P}_N (or equivalently τ_N under \mathbf{P}), i.e. $\Sigma_N(\Omega_t) = p_N(t)$.

We refer to $(\Omega_0, \mathcal{B}_0, P_N)$ as the conditional probability space of the killed process \tilde{W}_N .

Absolute Continuity through Cameron-Martin Theorem

Reasoning on the space Ω_0 , we want to study the properties of absolute continuity of the distribution P_N with respect to the distribution P_0 , which corresponds to the case of a constant barrier $L = L(0)$. In the case of a constant barrier, all the properties of p_0 , ρ_0 and P_0 are known and, in particular, p_0 is smooth. As such, the underlying idea is simple. If the conditioned measures $P_{N,t}$ and $P_{0,t}$ are mutually absolutely continuous (equivalent), then there is a random process X_t^N such that $dP_{N,t}(\omega) = X_t^N(\omega) dP_{0,t}(\omega)$. Since we want to find a function f_N such that $p_N = f_N \cdot p_0$, we necessarily have for all $t > 0$

$$\int_{\Omega_t} dP_{N,t}(\omega) - f_N(t) dP_{0,t}(\omega) = \int_{\Omega_t} (X_t^N(\omega) - f_N(t)) dP_{0,t}(\omega) = 0$$

which imposes that f_N takes the form

$$f_N(t) = \frac{1}{p_0(t)} \int_{\Omega_t} X_t^N(\omega) dP_{0,t}(\omega)$$

We hope that such an expression allows one to study the effect of the barrier's regularity on the local asymptotic behavior of f_N .

As a starting point, we show the following property for a piecewise linear barrier L_N :

Property 1. *The Radon-Nykodym derivative of p_N with respect to p_0 is expressed as the conditional expectation*

$$\frac{p_N(t)}{p_0(t)} = \mathbb{E}_{P_0} [\Xi_N | \tau_0 = t] ,$$

where the random variable Ξ_N is defined on $(\Omega_0, \mathcal{B}_0)$ as

$$\Xi_N(\omega) = \exp \left((\omega, l_N)_{\tau_0(\omega)} - \frac{1}{2} \|l_N\|_{\tau_0(\omega)}^2 \right) , \quad \mathbb{E}_{P_0} [\Xi_N] = 1 ,$$

with $(\cdot, \cdot)_t$ denoting the Wiener-Payley integral on $[0, t]$ and $\|\cdot\|_t$ the norm on $L^2(0, t)$.

The Cameron-Martin theorem plays a central role in proving this property and we recall its simplest formulation:

Theorem 5 (adapted from [37]). *On $\Omega \sim C_0(0, t)$, the probability law \mathbf{P} of W and \mathbf{Q} of $W - L$ are equivalent if and only if L belongs to the Dirichlet space*

$$\mathcal{D} = \left\{ F \in C_0(0, t) \mid \exists f \in L^2[0, t], L(t) = \int_0^t l(s) ds \right\} ,$$

in which case, their Radon-Nykodim derivative is almost surely

$$\frac{d\mathbf{Q}}{d\mathbf{P}}(t, \omega) = \exp \left((\omega, l)_t - \frac{1}{2} \|l\|_t^2 \right)$$

Moreover, if L does not belong to \mathcal{D} , the measures \mathbf{P} and \mathbf{Q} are orthogonal, i.e. there exists a measurable set B such that $\mathbf{P}(B) = 0$ and $\mathbf{Q}(\Omega \setminus B) = 0$

We can now proceed to demonstrate Property 1 expressing the density p_N of first-passage as a conditional expectation:

Proof. For any partition \mathcal{P} containing D_N , the finite-dimensional distribution $P_{N,t}^{\mathcal{P}}$ and $P_{0,t}^{\mathcal{P}}$ are equivalent on Ω_t and their Radon-Nykodim derivative is given by the

ratio of their densities

$$\begin{aligned} \frac{dP_{N,t}^{\mathcal{P}}}{dP_{0,t}^{\mathcal{P}}}(\omega) &= \frac{\tilde{\kappa}_N(\mathcal{P}; t, \omega)}{\tilde{\kappa}_0(\mathcal{P}; t, \omega)} \\ &= \exp \left(\sum_{i=1}^{p(\mathcal{P}, t)} \frac{1}{\Delta t_i} \left(\omega_i (L(t_i) - L(t_{i-1})) - \frac{1}{2} (L(t_i) - L(t_{i-1}))^2 \right) \right). \end{aligned} \quad (3.25)$$

Denoting \mathbf{P}_0 and \mathbf{P}_N the probabilities law of $W - L(0)$ and $\tilde{W}_N = W - L_N$ on $\Omega \sim C_{-L(0)}(0, t)$, expression (3.25) appears as a finite-dimensional Radon-Nykodim derivative of \mathbf{P}_N with respect to \mathbf{P}_0 (with the additional condition that $\tau_0(\omega) = t$). Since the piecewise linear barrier L_N is in \mathcal{D} , we deduce from the Cameron-Martin theorem that there exists a random process X^N defined almost surely on $(\Omega, \mathcal{B}, \mathbf{P}_0)$ as

$$X_t^N(\omega) = \exp \left((\omega, l_N)_t - \frac{1}{2} \|l_N\|_t^2 \right)$$

such that $d\mathbf{P}_N(t, \omega) = X_t^N(\omega) d\mathbf{P}_0(t, \omega)$. In particular the process X^N is a martingale on Ω . We now write the Radon-Nykodym derivative of P_N with respect to P_0 on Ω_0 as

$$\frac{dP_{N,t}}{dP_{0,t}}(\omega) = \mathbb{1}_{\Omega_t}(\omega) \frac{d\mathbf{P}_N}{d\mathbf{P}_0}(t, \omega) = \mathbb{1}_{\{\tau_0=t\}}(\omega) X_t^N(\omega) = X_{\tau_0}^N(\omega)$$

By the optional sampling theorem, the random variable Ξ_N

$$\Xi_N(\omega) = X_{\tau(\omega)}^N(\omega)$$

satisfies $\mathbb{E}_{\mathbf{P}_0} [\Xi_N] = 1$. At this point, it is crucial to realize that the definition of Ξ_N when restricted on Ω_0 still yields $\mathbb{E}_{P_0} [\Xi_N] = 1$.

The density of the first-passage with L_N at time t is

$$p_N(t) = \int_{\Omega_t} \frac{dP_N}{dP_0}(t, \omega) dP_0(t, \omega) = \int_{\Omega_t} \Xi_N(\omega) dP_0(\omega) = \mathbb{E}_{P_0} [\Xi_N | \tau_0 = t] p_0(t),$$

showing that the Radon-Nykodym derivative of p_N with respect to p_0 is naturally expressed as the conditional expectation of the random variable Ξ_N defined on Ω_0 . \square

It is then direct to give a criterion ensuring that τ admits a density function. Defining the martingale X_t^N with respect to the law \mathbf{P}_0 of $W - L(0)$ on (Ω, \mathcal{B}) as

$$X_t^N(\omega) = \exp \left((\omega, l_N)_t - \frac{1}{2} \|l_N\|_t^2 \right),$$

we notice that if L_N converges toward L in \mathcal{D} , the Cameron-Martin theorem yields

$$\lim_{N \rightarrow \infty} X_t^N(\omega) = \exp \left((\omega, l_N)_t - \frac{1}{2} \|l_N\|_t^2 \right) = X_t(\omega)$$

where the convergence is almost sure. As a consequence, we directly deduce the pointwise convergence on \mathbb{R}^+

$$\lim_N p_N(t) = \mathbb{E}_{P_0} [\Xi | \tau = t] p_0(t) = p(t), \quad \Xi = X_{\tau_0}.$$

We can then proceed just as in the proof of Property 1 to show that p thus-defined is a probability density. Indeed, for any $t > 0$, the random variable X_t is a martingale with respect to \mathbf{P}_0 , so that the optional sampling theorem entails that $\Xi = X_{\tau_0}$ is itself a martingale with respect to \mathbf{P}_0 (since τ_0 is a stopping time). Moreover, Ξ is equally well-defined on $(\Omega_0, \mathcal{B}_0)$ and is a martingale with respect to P_0 and $\mathbb{E}_{P_0} [\Xi] = 1$. Therefore p appears as a one-dimensional probability on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, and by Scheffé's theorem, we immediately conclude that τ admits p as a density.

3.3.2 Analytical Treatment

Thanks to direct probabilistic methods, we have elucidated the link between the local Hölder regularity of a barrier and the local admissibility of a density for the first-passage time. However, for being sample path based approaches, such methods are of no use when the barrier becomes homogeneously Hölder continuous. In this situation, the measure of the canonical Wiener process and its translated versions become singular with one another, such that they do not admit any Radon-Nykodim derivatives with respect to one another. Thus, to explore whether the first-passage still admits a density when the barrier is H -continuous with $H > 1/2$, we use analysis and disregard the actual path-wise formulation of the problem.

Partial Differential Formulation

In order to gain insight into the analytic methods to be developped, we first draw a link between the previous probabilistic treatment and the partial differential treatment of a killed Wiener process. Suppose the barrier L at stake is in the Dirichlet space \mathcal{D} , there exists a function l in $L^2(0, t)$ such that $L(t) = \int_0^t l(s) ds$. If W is a Wiener process starting at $W_s = x < 0$ at time zero, the expression

$$X_{u,v} = \exp \left(\int_u^v l(s) dW_s - \frac{1}{2} \int_u^v l^2(s) ds \right)$$

defines a well-defined random variable on Ω in the sense that it is almost surely finite.

Now, by the discussion above, we have the relation

$$\mathbb{E}(X_{s,t} | \tau = t) = \frac{q(s, x; t)}{q_0(s, x; t)}, \quad (3.26)$$

where q_0 is the density of first-passage to level zero knowing $W_s = x$ and $t \mapsto q(s, x; t)$ is the density for a first-passage to L for the same conditioning. By the same reasoning

as in the previous section, we can justify [141] that

$$\mathbb{E}(X_{s,t} | \tau > t, W_t = y) = \frac{\tilde{\kappa}(s, x; t, y)}{\tilde{\kappa}_0(s, x; t, y)}, \quad (3.27)$$

where respectively, $\tilde{\kappa}_0$ and $\tilde{\kappa}$ are the transition kernel of the canonical, and for the drifted Wiener process killed at level zero. For instance, we can show that the Kolmogorov equation for the killed drifted Wiener process is equivalent to the conditional time-evolution law of $X_{r,t}$

$$\begin{aligned} \mathbb{E}(X_{r,t} | \tau > t, W_t = z) &= \mathbb{E}(\mathbb{E}(X_{r,t} | \tau > t, W_t = z) | W_s), \\ &= \int_{-\infty}^0 \mathbb{E}(X_{r,s} | \tau > t, W_t = z, W_s = y) \mathbb{E}(X_{s,t} | \tau > t, W_s = y) \\ &\quad \mathbb{P}(W_s \in dy | \tau > t, W_r = x, W_t = z). \end{aligned} \quad (3.28)$$

Indeed, by the Markov property of the Wiener process killed at level zero, we know that the law of the Markov bridge reads

$$\mathbb{P}(W_s \in dy | \tau > t, W_r = x, W_t = z) = \frac{\tilde{\kappa}_0(r, z; s, y) \tilde{\kappa}_0(s, y; t, z)}{\tilde{\kappa}_0(r, x; t, z)}$$

so that, after substituting in (3.28) the expectation expressions of (3.27), we recover the Kolmogorov law

$$\tilde{\kappa}(r, x; t, z) = \int_{-\infty}^0 \tilde{\kappa}(r, z; s, y) \tilde{\kappa}(s, y; t, z) dy$$

for the killed drifted Wiener process. Incidentally, the heat equation for the drifted process killed at zero yields

$$q(s, x; t) = -\frac{1}{2} \left[\frac{\partial \tilde{\kappa}(s, x; t, y)}{\partial y} \right]_{y=0},$$

which is equivalent to the intuitive fact that $\lim_{y \rightarrow 0} \mathbb{E}(X_{s,t} | \tau > t, W_t = y) = \mathbb{E}(X_{s,t} | \tau = t)$, meaning that the expectation is continuous on $\{\tau = t\}$ when such an event is seen as the limit conditioning $\{\tau > t, W_t = y_n\}$, with $\lim_{n \rightarrow \infty} y_n = 0$. Indeed, the absorbing boundary condition states that $\lim_{y \rightarrow 0} \tilde{\kappa}(s, x; t, y) = \tilde{\kappa}_0(s, x; t, y) = 0$, while $\lim_{y \rightarrow 0} \partial_y \tilde{\kappa}(s, x; t, y) = -2q(s, x; t)$ and $\lim_{y \rightarrow 0} \partial_y \tilde{\kappa}_0(s, x; t, y) = -2q_0(s, x; t)$ are well-defined, so that the L'Hôpital rule reads

$$\lim_{y \rightarrow 0} \mathbb{E}(X_{s,t} | \tau > t, W_t = y) = \lim_{y \rightarrow 0} \frac{\tilde{\kappa}(s, x; t, y)}{\tilde{\kappa}_0(s, x; t, y)} = \frac{q(s, x; t)}{q_0(s, x; t)} = \mathbb{E}(X_{s,t} | \tau = t) .$$

The previous discussion reveals that this continuity holds when the barrier admits a point of local singularity at t that has a Hölder exponent $h > 1/2$. However, if L admits an infinity of such singularities, the random variable $X_{s,t}$ vanishes or diverges almost surely and its expectation cannot be properly defined.

The direct probabilistic method then falls short of elucidating the case of an H -continuous function for $H > 1/2$ and we must resort to using indirect arguments from analysis. As the densities of both the killed process and the first-passage are expressed as conditional expectations, it is advantageous to distance ourselves from a purely path-wise description. Though the path-wise defined random variable $X_{s,t}$ has no meaning when L does not belong to the space \mathcal{D} , nevertheless the expectation (3.26) and (3.27) can remain valid as limit objects

$$\begin{aligned} \frac{q(s, x; t)}{q_0(s, x; t)} &= \lim_{N \rightarrow \infty} \mathbb{E}(X_{s,t}^N | \tau = t) \\ \frac{\tilde{\kappa}(s, x; t, y)}{\tilde{\kappa}_0(s, x; t, y)} &= \lim_{N \rightarrow \infty} \mathbb{E}(X_{s,t}^N | \tau > t, W_t = y) \end{aligned}$$

where the $X_{s,t}^N$ corresponds to the piecewise approximation L_N . As the expectations of $X_{s,t}^N$ are well-defined, the question becomes whether the previous limit is well-behaved for the case of an H -continuous function L with $H > 1/2$.

Before answering this question positively, we draw attention to the fact that such a result would not be surprising. Consider the case of a drifted Wiener process *without killing* with an arbitrary H -continuous drift L , $0 < H < 1$, and its usual piecewise linear interpolations L_N . There always exist generalized functions l_N which satisfies $L_N(t) = \int_0^t l_N(s) ds$ and these “derivatives” belong to $L^2(0, t)$ for all $t > 0$. Because L does not belong to $L^2(0, t)$ for any $t > 0$, the random X^N does not converge in distribution to any random variable. However, the conditional expectations of X^N can be convergent: for instance, we have:

$$\lim_{N \rightarrow \infty} \mathbb{E}(X_{s,t}^N | W_t = y) = \lim_{N \rightarrow \infty} \frac{\mathbb{P}(W_t + L_N(t) \in dy)}{\mathbb{P}(W_t \in dy)} = \exp\left(-\frac{1}{2t}(L(t)(L(t) - 2y))\right).$$

Integral Representation and Single-Layer Potential

We want to establish the existence of a solution to the heat equation with absorbing boundary specified on H -continuous function with $H > 1/2$. Benefiting from the insight of our probabilistic treatment, we first infer an integral representation for the solution where the potential first-passage density intervenes. We proceed heuristically by observing that the transition kernel κ of the Wiener process killed on L can always be written

$$\kappa(s, x; t, y) = k(s, x; t, y)s(s, x; t, y)$$

where the conditional survival function $s(s, x; t, y)$ gives the probability that W does not hit the barrier knowing that $W_s = x$ and $W_t = y$. If the first-passage time τ to L admits a density function $t \mapsto q(s, x; t)$, the strong Markov property entails that for $y < L(t)$, the survival probability s satisfies

$$1 - s(s, x; t, y) = \frac{1}{k(s, x; t, y)} \int_s^t q(s, x; \tau) k(\tau, L(\tau); t, y) d\tau.$$

This suggests that one looks for solutions to the heat equation with absorbing boundary condition on L under the following integral representation

$$p(x, s; t, y) = k(s, x; t, y) - \underbrace{\int_s^t k(\tau, L(\tau); t, y) q(s, x; \tau) d\tau}_{w(s, x; t, y)}, \quad (3.29)$$

where q is the unknown function that parametrized the problem. It is straightforward to verify that p satisfies the heat equation on $\{(t, y) \in \mathbb{R}^+ \times \mathbb{R} \mid y \neq L(t)\}$ for general forms of functions q . Indeed for any cumulative probability distribution Q on \mathbb{R}^+ , the integral expression

$$\int_0^t k(\tau, L(\tau); t, y) dQ(\tau),$$

is a well-formed solution, referred to as a *single-layer potential*, and is analytical for all (t, y) such that $t > 0$ and $y \neq L(t)$.

Obviously, we say that the function p solves our problem if it satisfies the appropriate boundary conditions for our first-passage problem. In this regard, we can formulate two types of boundary conditions:

Explicitly as Dirichlet conditions: by stipulating the absorbing nature of the barrier

$$p(x, s; t, L(t)) = 0.$$

Implicitly as Neumann conditions: by imposing the relation between the probability current at the barrier and the first-passage density

$$2q(s, x; t) = -\partial_y p(x, s; t, L(t)).$$

Both boundary conditions lead to Volterra equations, ultimately expressed as integrals of the second-type. To ensure the existence and unicity of a solution to these equations, we need the following powerful result:

Theorem 6 (adapted from [45, 229]). *The linear volterra eqution of the second-kind*

$$g(t) = f(t) + \int_0^t K(t, s)f(s) ds ,$$

where g is a piecewise continuous function has a unique piecewise continuous solution f for all $t > 0$ if K is bounded on $0 < s < t$ and if there exist a monotone increasing function α with $\lim_{t \rightarrow 0} \alpha(t) = 0$, such that for all $0 < s < t$

$$\int_s^t |K(t, \tau)| d\tau \leq \alpha(t - s) .$$

We now proceed to analyze the boundary conditions.

Volterra Equation from Dirichlet Boundary Condition

As a first step toward formulating the Dirichlet boundary condition, we must investigate the limit behavior of p when its argument approaches L . In this respect, it is important to observe that as long as L is H -continuous with $H > 1/2$, we have

$$\lim_{\tau \rightarrow t^-} \frac{L(t) - L(\tau)}{\sqrt{t - \tau}} = 0 ,$$

and we can apply the dominated convergence theorem to the integrable family of function $\tau \mapsto k(\tau, L(\tau); t, y)q(s, x; \tau)$ when y tends toward $L(t)$, so that we verify that p is actually continuous on the whole half-plan $\mathbb{R}^+ \times \mathbb{R}$. Then, letting y tend toward L in (3.29), we can rewrite the absorbing boundary conditions $p(s, x; t, L(t)) = 0$ as the integral equation

$$k(s, x; t, L(t)) = \int_s^t k(\tau, L(\tau); t, L(t)) dQ(\tau) . \quad (3.30)$$

This equation can also be derived through probabilistic means such as differentiating equation (3.12) with respect to x and then letting x tend toward $L(t)$ or by simply

realizing that $k(s, x; W_s, t)$ indexed by s is a martingale.

Unfortunately, equation (3.30) is a Volterra equation of the first-kind and as such cannot be dealt with directly. However for barriers L that are H -continuous, it can be recognized as a linear generalized Abel integral equation, that is an equation of the type

$$g(t) = \int_s^t \frac{K(t, \tau)f(\tau)}{(t - \tau)^h} d\tau$$

where f is the unknown, g is a continuous function, and K is a continuous kernel for $s \leq t$ and $0 < h < 1$.

Abel integral equations are frequently encountered in physics and there exist methods to prove the existence and unicity of a solution by transforming the original equation into a Volterra equation of the second-kind. In our case, it proceeds through the use of the Abel integral transform, which is designed to solve the canonical Abel equation

$$g(t) = \int_s^t \frac{f(\tau)}{\sqrt{t - \tau}} d\tau .$$

It is well-known that the unique solution is given as

$$f(t) = \mathcal{A}[g](t) = \frac{1}{\pi} \frac{d}{dt} \left[\int_s^t \frac{g(\tau)}{\sqrt{t - \tau}} d\tau \right]$$

where \mathcal{A} is the Abel inverse operator. Before showing how applying \mathcal{A} to equation (3.30) reduces the problem to a Volterra equation of the second-kind, we state the result:

Proposition 10 (adapted from [38]). *If L is H -continuous with $H > 1/2$, through the application of the Abel operator, the Volterra equation of the first-kind (3.30) is*

equivalent to the new Volterra equation of the second-kind

$$\sqrt{2\pi}\mathcal{A}[g](t) = q(t) + \frac{1}{\pi} \int_s^t K(t, \tau) q(\tau) d\tau,$$

with the kernel K being defined as

$$K(t, \tau) = \frac{\partial}{\partial t} \left\{ \int_{\tau}^t (t - \sigma)^{-1/2} (\sigma - \tau)^{-1/2} \exp \left(-\frac{(L(\sigma) - L(\tau))^2}{2(\sigma - \tau)} \right) d\sigma \right\} d\tau,$$

and g denotes the continuous function $g(t) = k(s, x; t, L(t))$.

Proof. To simplify notations, we suppose $x = s = 0$ and we write

$$w(t) = \int_0^t k(\tau, L(\tau); t, L(t)) q(\tau) d\tau.$$

By Fubini's theorem, we have

$$\begin{aligned} \mathcal{A}[w](t) &= \frac{1}{\pi} \frac{d}{dt} \left[\int_0^t (t - \tau)^{-1/2} \left(\int_0^{\tau} k(\sigma, L(\sigma); \tau, L(\tau)) q(\sigma) d\sigma \right) d\tau \right], \\ &= \frac{1}{\sqrt{2\pi^3}} \frac{d}{dt} \left[\int_0^t q(\tau) \underbrace{\left(\int_{\tau}^t (t - \sigma)^{-1/2} k(\tau, L(\tau); \sigma, L(\sigma)) d\sigma \right)}_{Q(t, \tau)} d\tau \right]. \end{aligned}$$

We want to apply Leibniz's rule to express the derivative of the time dependent interval. In order to do so, we must study the limit behavior of $Q(t, \tau)$ when $\tau \rightarrow t^-$.

As usual, when dealing with the Abel transform we use the fact that

$$\int_{\tau}^t \frac{1}{\sqrt{(t-s)(s-\tau)}} ds = \pi,$$

to write the expression

$$\begin{aligned}
\lim_{\tau \rightarrow t^-} Q(t, \tau) &= \lim_{\tau \rightarrow t^-} \pi \exp \left(-\frac{(L(t) - L(\tau))^2}{2(t - \tau)} \right), \\
&+ \left\{ Q(t, \tau) - \pi \exp \left(-\frac{(L(t) - L(\tau))^2}{2(t - \tau)} \right) \right\}, \\
&= \pi + \lim_{\tau \rightarrow t^-} \left\{ \int_{\tau}^t \frac{\exp \left(-\frac{(L(\sigma) - L(\tau))^2}{2(\sigma - \tau)} \right) - \exp \left(-\frac{(L(t) - L(\tau))^2}{2(t - \tau)} \right)}{\sqrt{(t - \tau)(\sigma - \tau)}} d\sigma \right\},
\end{aligned}$$

where we use the fact that L is H -continuous with $H > 1/2$:

$$\lim_{\tau \rightarrow t^-} \exp \left(-\frac{(L(t) - L(\tau))^2}{2(t - \tau)} \right) = 1.$$

Finally, applying the mean-value theorem to the expression between braces shows that it is dominated by $O(t - \tau)^{2H-1}$, and thus since $H > 1/2$, we have $\lim_{\tau \rightarrow t^-} Q(t, \tau) = \pi$.

Consequently, we can apply Leibniz's rule, which yields:

$$\mathcal{A}[w](t) = \frac{1}{\sqrt{2\pi}} \left(q(t) + \frac{1}{\pi} \int_s^t K(t, \tau) q(\tau) d\tau \right),$$

where K is defined as in Proposition 10. □

A careful study of the kernel K shows that the integral equation of Proposition 10 satisfies the conditions of Theorem 6 [38]. It demonstrates that the integral equation (3.31) obtained through the Abel transform admits a unique continuous solution. However, the form of the kernel K is complicated. In the following section, we show that once the existence of the continuous solution is admitted, this solution satisfies another well-posed Volterra equation of the second-kind.

Volterra Equation from Neumann Boundary Condition

For physical applications, the heat equation is studied in a three-dimensional setting and is likened to a Poisson equation for a static electric potential. In this context, if

the charge density σ is prescribed on the surface Σ of a space domain D , the solution to the Poisson equation is given by the single-layer potential

$$p(\mathbf{x}) = \int_{\Sigma} k(\mathbf{x}, \mathbf{u}) \sigma(\mathbf{u}) d\Sigma(\mathbf{u}),$$

where k is the Green function of the corresponding Laplace equation. For such potentials, the electric field is seen as the gradient $\nabla_{\mathbf{x}} p$ of the potential p , is well-defined on both sides of the surface, and obeys a *jump condition* when passing through the surface. This jump condition consists of a discontinuity within the component of the gradient that is normal to the surface and is directly related to the local charge density according to

$$[\nabla_{\mathbf{x}} p(\mathbf{u}) \cdot \mathbf{n}_{\mathbf{u}}]_{\mathbf{u}-}^{\mathbf{u}+} = \sigma(\mathbf{x}), \quad \mathbf{n}_{\mathbf{u}} \perp d\Sigma(\mathbf{u})$$

where \mathbf{u}^+ and \mathbf{u}^- mean that the gradient on the surface is computed by approaching the surface from the interior and exterior, respectively.

The time-dependence at play in the heat equation does not fundamentally change this result. In our one-dimensional case, the instantaneous space domain is $(-\infty, L(t))$ and the charge is a Dirac delta function $q(t)\delta_{L(t)}$. In turn, we identify the gradient with its space derivative, whose upper and lower limit on the barrier L can be shown to satisfy the following jump condition:

Proposition 11 (adapted from [38]). *If L is an H -continuous function with $H > 1/2$ and if q is a bounded function, then the space derivative $\partial_y w$ is a continuous function on $\{(t, y) \mid y < L(t)\}$ and $\{(t, y) \mid y > L(t)\}$, and admits an upper and lower limit*

when approaching the barrier L , and these are given as

$$\begin{aligned}\lim_{y \rightarrow L(t)^-} \partial_y w(s, x; t, y) &= q(s, x; t) + \int_s^t \partial_y k(\tau, L(\tau); t, L(t)) q(s, x; \tau) d\tau, \\ \lim_{y \rightarrow L(t)^+} \partial_y w(s, x; t, y) &= q(s, x; t) + \int_s^t \partial_y k(\tau, L(\tau); t, L(t)) q(s, x; \tau) d\tau.\end{aligned}$$

and $\lim_{y \rightarrow L(t)^-} \partial_y w(s, x; t, y)$ as a function of t is itself H -continuous

The proof of this point is essentially technical and we refer to [38] for details. It is important to note that if L is H -continuous with $H < 1/2$, the integral expression in Proposition 11 ceases to be well-posed.

At this point, we know that if we constrain q to be homogeneously Hölder continuous with $H > 1/2$, the integral solution p to the heat equation admits a continuous spatial derivative on $\{(t, y) \mid y < L(t)\}$. Instead of trying to solve the problem for the Dirichlet boundary condition, that is $p(s, x; t, L(t)) = 0$, we now consider the Neumann condition. This stipulates that the probability current at the barrier (i.e. the spatial derivative) yields the first-passage distribution q as

$$\begin{aligned}2q(s, x; t) &= - \left[\frac{\partial}{\partial y} p(s, x; t, y) \right]_{L(t)^-} \\ &= -\partial_y k(s, x; t, L(t)) + q(s, x; t, L(t)) \\ &\quad + \int_s^t \partial_y k(\tau, L(\tau); t, L(t)) q(s, x; \tau) d\tau.\end{aligned}$$

This self-consistent equation takes the form of a Volterra equation of the second-kind whose validity has been previously established only for differentiable barriers [178]. We recapitulate the present discussion in the following proposition.

Proposition 12. *Consider L is H -continuous with $H > 1/2$ such that $L(s) > W_s$, the first-passage $\tau = \{t > s \mid W_t > L(t)\}$ admits a bounded density $t \mapsto q(s, x; t)$, that*

is the unique solution of the Volterra equation of the second type

$$\frac{L(t)}{t^{3/2}}\phi\left(\frac{L(t)}{\sqrt{t}}\right) = q(s, x; t) + \int_s^t \frac{L(t) - L(\tau)}{(t - \tau)^{3/2}}\phi\left(\frac{L(t) - L(\tau)}{\sqrt{t - \tau}}\right) q(s, x; \tau) d\tau, \quad (3.31)$$

where $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$.

Chapter 4

First-Passage Markov Framework

In their seminal study, Mainen and Sejnowski [133] draw attention to the concept of spike timing reliability by presenting the following electrophysiological results. When a neocortical neuron is injected repeatedly with a steady current, the first spikes tend to occur at the same time after the injection of the current, but later on, the successive spikes progressively desynchronized, as if each inter-spike interval was independently perturbed. Contrarily, if an highly fluctuated signal is superimposed, the elicited train of spikes tends to line up reliably, while exhibiting little spike temporal jittering when spikes are recognized as the same across trials. Thus, a neuron’s firing pattern differs profoundly according to the nature of its input.

From there, the concept of reliability is understood as reproducibility of the firing pattern in response to cyclical stimulation. This idea connects closely to spiking precision of the neural response, that is, the observed weak temporal variability of spiking events thought as being the instances of the same encoding unit [15, 20]. However, the two notions of precision and reliability are far from overlapping. For instance, if precision appears to be generally message-dependent, spike reliability can very well be independent of it at the same time [175].

From a theoretical perspective, this complex message dependence is explained by showing that the spiking precision is a function of the input for virtually all spike-generating mechanisms [41]. For instance, in the case of the leaky integrate-and-fire model, the repeated occurrence of spikes stems from the fact that the effective barrier exhibits a downward peak that lies lower than the equilibrium value of the voltage, while the spike precision is related to the slope of the barrier on the way to the peak’s minimum. This view demonstrates that spike precision and spike reliability can actually be independent of each other. It has also been noted that these general operative definitions appear to hold for all spiking models [28].

Because reliability and precision are the two concepts used experimentally to pool spikes in the same meaningful event, it is fundamental to understand the interplay

of these two notions in hopes of elucidating the neural encoding. This issue has long been identified in addressing the question of whether the neural encoding exhibit a dichotomy between temporal or rate coding [117]. We propose to revisit the integrate-and-fire model to further the study of spike reliability and spike precision by focusing on the effect entailed by the stochasticity of the model. Previous studies that have undertaken a similar task have narrowed down the concept of time-scale matching [204]; spike reliability and spiking interval precision are maximized if the typical fluctuation of the barrier match the mean spike interval in the absence of fluctuations [165]. In the same spirit of quasi-periodic forcing, the integrate-and-fire model has been considered from the perspective of stochastic bifurcation theory [10]. In this context, firing patterns are likened to attractors for train of spikes, and reliability is understood as stability of the attractor in the face of noise perturbation [221].

Here, we clarify the crucial role played by the neuronal noise (intrinsic or extrinsic) in explaining what messages are prone to elicit precise spiking and how it relates to spike reliability. We first formulate the most general Markovian framework amenable to simulate the situation of the Mainen-Sejnowski experiment. As in previous work [205], we found ourselves on the study of the first-passage kernel to *rigorously* define the succession of spikes as the outcome of an inhomogeneous Markov chain. Then, identifying the Hölder exponent of the effective barrier as the controlling parameter of the overall time-precision of spike timing, we simulate our Markov chain for injected current giving rise to various H -continuous effective barriers. Based on our numerical experiments, we conclude by discussing the message dependence of the leaky integrate-and-fire encoding.

4.1 First-Passage Markov Chain

Consider the stochastic leaky integrate-and-fire model for a spike triggering membrane threshold V and a post-spiking reset value $v < V$. Suppose a spike is emitted at time $t_i > 0$. With initial condition $X_{t_i}^+ = v$, the inhomogeneous linear stochastic differential system

$$dX_t = \alpha X_t dt + \sigma dW_t + dC(t), \quad t > t_i, \quad (4.1)$$

describes the ensuing sub-threshold noisy dynamic of the potential when driven by the input current $dC(t)$. Here, $dC(t)$ shall be considered as the infinitesimal increment of a time-varying load function $C(t)$ that is homogeneously Hölder continuous for a given exponent $H > 0$, i.e. for every $T > 0$, there exists a constant $c_T > 0$ such that for all $0 < t, s < T$

$$\lim_{\delta \rightarrow 0+} \sup_{|t-s| \leq \delta} \frac{|C(t) - C(s)|}{|t - s|^H} \leq c_T.$$

Notice that, at the cost of rescaling X and I by σ , we can restrain ourselves to the study of the case $\sigma = 1$.

4.1.1 Cyclically Driven Leaky Integrate-and-Fire Neuron

The nonlinearity of the leaky integrate-and-fire model lies entirely in the spike generation and subsequent reset, so that we can separately integrate input and noise between spikes. Thus, our first-passage problem for constant threshold V and varying forcing dC becomes a first-passage problem without driving forces to a fluctuating effective barrier. Precisely, we solve (4.1) writing $X = U^i + l^i$, where we separate the stochastic part U^i (the Ornstein-Uhlenbeck process obtained for $dC = 0$) and the

deterministic part l^i arising from the integration of the input $dC(t)$:

$$\begin{aligned} U_t^i &= v e^{\alpha(t-t_i)} + \int_{t_i}^t e^{\alpha(t-s)} dW_s, \\ l^i(t) &= \int_{t_i}^t e^{\alpha(t-s)} dC(s). \end{aligned}$$

Determining the next spiking time t_{i+1} can be cast in terms of a first-passage problem for the process U^i with the effective barrier $t \mapsto L^i(t) = V - l^i(t)$:

$$t_{i+1} = \inf\{t > t_i \mid U_t^i > L^i(t)\}. \quad (4.2)$$

Therefore, a train of spikes $t_0 < t_1 < \dots < t_n$ is determined by solving consecutively the first-passage problems (4.2). Note that, due to the reset rule, the effective barriers do not agree at spiking times $L^{i-1}(t_i^-) \neq L^i(t_i^+) = V$. However, for all $i > 0$, we have for $t > t_i$:

$$\left\{U_t^i < V - l^i(t)\right\} = \left\{U_t^i - \int_{t_{i-1}}^{t_i} e^{\alpha(t-s)} dC(s) < V - l^{i-1}(t)\right\}$$

Making the left-hand term U_t^i of the second inequality explicit, we have

$$U_t^i = e^{\alpha(t-t_i)} \left(v - \int_{t_{i-1}}^{t_i} e^{\alpha(t_i-s)} dC(s) \right) + \int_{t_i}^t e^{\alpha(t-s)} dW_s,$$

and we recognize U_t^i as the solution of (4.1) for $dC = 0$, with the new initial condition:

$$U_{t_i^+}' = v - \int_{t_{i-1}}^{t_i} e^{\alpha(t_i-s)} dC(s) = L^{i-1}(t) - (V - v).$$

Therefore, by a straightforward recurrence argument, the train of spikes $t_0 < t_1 < \dots < t_n$ is determined by the sequence of first-passage problem:

$$t_{i+1} = \inf\{t > t_i \mid U_t^i > L^0(t)\}, \quad (4.3)$$

where U^i is the standard Ornstein-Uhlenbeck process with initial condition $U_{t_i+}^i = L^0(t) - (T - v)$. In other words, by altering the reset rule, the linearity of the stochastic dynamics allows us to recast the successive first-passage problems (4.2) in terms of a sequence of first-passage problems for one single continuous barrier (4.3).

In a typical experiment, the spiking history of a neuron is recorded in response to repeated presentations of the same stimulus. We idealize this situation by studying the distribution of spiking events when an input cyclically forces a leaky-integrate-and-fire neuron. To avoid discontinuity effects, we choose a barrier satisfying $L(T) = L(0)$ for some $T > 0$ and then extend the definition of L on the whole time-line by periodization $L(t) = L(t \bmod T)$. Then, the sequence of random times $\mathcal{T}_n = (\tau_n \bmod T)$, where τ_n denotes successive first-passage times to L , defines a discrete-time Markov chain \mathcal{T} over the finite time period $[0, T)$, seen as an oriented circle¹.

To make it more formal, assume we can choose a load function satisfying for some $T > 0$

$$\int_0^T e^{\alpha(T-s)} dC(s) = 0, \quad (4.4)$$

which amounts to having a periodic effective barrier by setting $L(t) = L(t \bmod T)$. For any time s in $[0, T)$, consider the first passage time τ_s for an Ornstein-Uhlenbeck process starting at $U_s = L(t) - (T - v)$ and the barrier L . Because $L(t)$ is a continuous function, it is known that the random variable τ_s admits a continuous non-decreasing

¹ The passage of time orients the circle and we identify the future time T with the past time 0

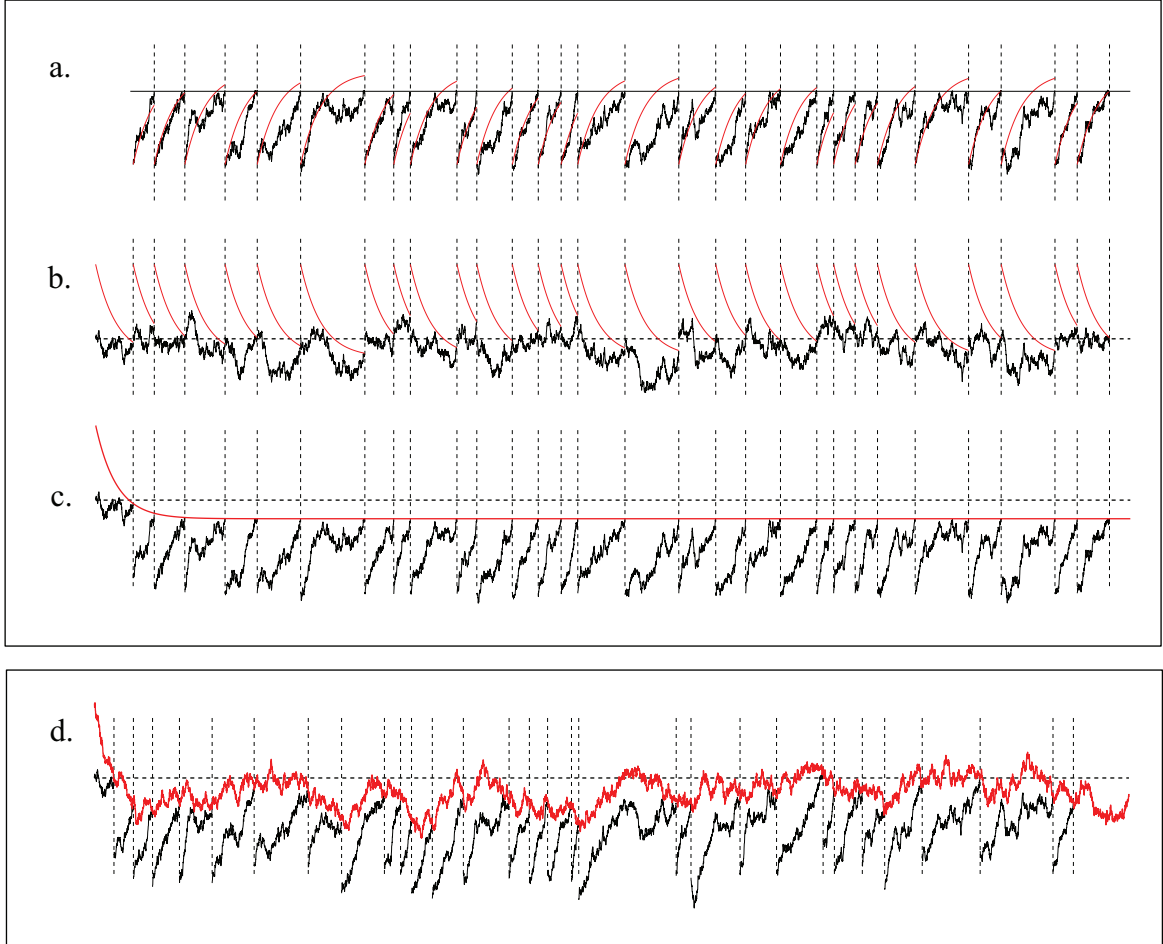


Figure 4.1: **a,b,c.** Equivalent representations of one cycle of a first-passage Markov chain for an Ornstein-Uhlenbeck process with a constant threshold and a constant positive drift: **a.** The drift is added to the sample path causing them to increase up to the threshold, **b.** The drift is subtracted from the threshold after each reset, **c.** The sample path is similarly obtained for one single effective boundary. **d.** One cycle of the first-passage Markov chain generated for the same Ornstein-Uhlenbeck process for constant positive drift with injected Gaussian white noise: the effective barrier is an Ornstein-Uhlenbeck bridge.

cumulative distribution function $F_s : [s, \infty) \rightarrow [0, 1]$. We then define the measure k_s on the Borel sets of $[s, \infty)$ by setting for every open set $O_{a,b} = (a, b) \subset [s, \infty)$, $s < a < b$:

$$k_s(O_{a,b}) = F_s(b) - F_s(a).$$

Moving forward, we identify $[0, T)$ with the circle $\mathbb{C} = \mathbb{R}/T\mathbb{Z}$, which is compact and for which the open arc circles $O_{(a,b)}$, are oriented counter-clockwise from a to b , and generate the collection of Borel sets $\mathcal{B}(\mathbb{C})$. For every s , the surjection $\pi_s : [s, \infty) \rightarrow [s, \infty)/T\mathbb{N} \sim \mathbb{C}$ is continuous for the quotient topology. We can then form on the quotient space \mathbb{C} , the measurable functions k_s^T by setting for all open $O_{(a,b)}$

$$k_s^T(O_{(a,b)}) = k_s(\pi^{-1}(O_{(a,b)})).$$

The functions k_s^T are transition kernels on the compact measurable state space $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$. Given an initial probability measure μ_0 on \mathbb{C} , they define a continuous state, discrete time Markov chain [90, 151, 214] $\mathcal{T} = (\mathcal{T}, \mathcal{P})$ on $(\Omega, \mathcal{M}) = (\mathbb{C}, \mathcal{B}(\mathbb{C}))^{\mathbb{N}}$, whose probability \mathcal{P} satisfies:

$$\forall n \in \mathbb{N}, \quad \mathcal{P}(d\tau_n, \dots, \mathcal{T}_0 \in d\tau_0) = \mathbb{P}(\mathcal{T}_n \in d\tau_n, \dots, \mathcal{T}_0 \in d\tau_0) = k_{\tau_{n-1}}(d\tau_n) \dots k_{\tau_0}(d\tau_1) \mu_0(d\tau_0).$$

In particular, for all u, v in \mathbb{C} , $v \mapsto k_s^T(O_{(u,v)})$ is continuous in v with $k_s^T(\mathbb{C}) = 1$.

We shall see k_s^T as the cumulative distribution of τ_n when the underlying process U^n starts at $U_s^n = L(s) - (V - v)$, i.e. the distribution of a spiking event knowing that the previous spike occurs at t . As such, the kernels k_s^T need not admit a density κ satisfying $k_s^T(dt) = \kappa(s, t) dt$, similarly to the ‘‘Devil’s staircase’’ resulting from the integration the uniform measure over the triadic Cantor set [135]. However, if we do not suppose the existence of a density, we stress the fact that, by continuity of the F_s with $\lim_{t \rightarrow s^+} F_s([s, t]) = 0$, every open set of the form $O(a, b)$ is a τ_s -continuity set

$(k_s^T(\partial O) = 0)$, and so are the Borelians $\mathcal{B}(\mathbb{C})$.

4.1.2 Ergodicity of the Markov Chain

We are interested in using this Markov framework to elucidate the distribution of spiking events when a neuron is driven cyclically by an input defined (4.4). In view of this, we first justify the crucial assumption that, under periodic forcing, the average spiking activity sums up the neurons encoding capabilities. To ensure that the instantaneous firing rate and the probability of spiking coincide, we show that the Markov Chain $(\mathcal{T}, \mathcal{P})$ is *ergodic*, a notion we define in the following.

An distribution μ is invariant by $(\mathcal{T}, \mathcal{P})$ if it satisfies

$$\mu(dt) = \int_0^T k_s^T(dt) \mu(ds),$$

so that if \mathcal{T}_n is distributed according to μ , so is \mathcal{T}_{n+1} . When there exists a unique such measure μ , for any initial distribution μ_0 and any measurable set B on the circle \mathbb{C}

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{1}_B(\mathcal{T}_n) = \mu(B), \quad \mathbb{1}_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \notin B \end{cases},$$

and the Markov chain is said to be ergodic. Simply stated, the mean sojourn-time of the Markov chain in B tends toward the measure of B under μ . Practically, this property justifies considering the normalized instantaneous mean firing rate as a probability of spiking.

We can show that the Markov chain $(\mathcal{T}, \mathcal{P})$ is indeed ergodic for a very general class of barriers L : the homogeneously Hölder continuous functions. Since the state space \mathbb{C} of $(\mathcal{T}, \mathcal{P})$ is compact, it is enough to show that it has the strong-Feller property [91] to prove the existence of invariant measures.

Property 2. *If the barrier L is homogeneously Hölder continuous, the Markov chain*

$(\mathcal{T}, \mathcal{P})$ is strong-Feller, i.e.

$$\forall B \in \mathcal{B}(\mathbb{C}), \quad s_n \rightarrow s \in \mathbb{C}, \quad \Rightarrow \quad k_{s_n}(B) \rightarrow k_s(B)$$

To establish the unicity of the invariant measure μ , it is enough to show that the Markov chain $(\mathcal{T}, \mathcal{P})$ has the irreducible property [91]:

Property 3. *The Markov chain $(\mathcal{T}, \mathcal{P})$ is irreducible, i.e.*

$$\forall B \in \mathcal{B}(\mathbb{C}), \quad \forall s \in \mathbb{C}, \quad k_s(B) > 0.$$

The two previous properties are deduced from the general properties of the first-passage times.

The Feller property specifies that, if two identical leaky integrate-and-fire neurons spike respectively at times s and t , then, when s asymptotically approaches t , the probability that the first neuron later spikes in a given time interval becomes the same as for the other neuron. In other words, close initial conditions entail similar probability laws for the occurrence of the next spiking events (in the sense of the Kolmogorov test). Intuitively, this property holds for our first-passage Markov chain for two reasons. First, the continuity of the barrier which ensures the continuity of the cumulative distributions of the transition kernels. Second, the non-zero reset rules which constrain the membrane potential to be reset away from the barrier, thus avoiding pathological situations such as immediate absorption. Anticipating future results, we illustrate the Feller property for a simulated first-passage Markov chain in Figure 4.1.2.

The irreducible property, which states that if one spiking time is achievable for a given starting condition (previous reset time), it is attainable for any starting time, similarly stems from these two intuitive observations. Once the membrane potential has reset, its underlying trajectory can explore the region of the phase space lying

under the barrier. If one trajectory starting at t has a non-zero probability to hit this barrier in a certain time-region, we can easily convince ourselves that another trajectory starting at any s has a non-zero probability to be close to the reset value in t , and from there, unfold as a trajectory that has been reset in t .

In the following, we prove rigorously the two properties. We first proceed to prove the Feller property 2:

Proof. Since every open set is a set of continuity, it is enough to show the property for the open sets of the form $O_{(a,b)}$. We proceed in three steps:

i) Uniform tightness of the first-passage distributions: Given the continuous surjection $\pi_s : [s, +\infty) : \mathbb{R}/T\mathbb{Z}$, posit $A_{a,b} = \pi_s^{-1}(O_{(a,b)})$. Since the barrier L can be seen as a periodic continuous function on $[s, +\infty)$, there exist $M > \sup_{t \in \mathbb{R}^+} L(t)$. But for all $x < M$, we have $\tau_s \leq \tau_s^M$ where $\tau_s^M = \inf\{t > s \mid X_t > M\}$, then $\{\tau_s^M < +\infty\} \subset \{\tau_s < +\infty\}$ entails $\mathbb{P}(\tau_s < +\infty) \geq \mathbb{P}(\tau_s^M < +\infty) = 1$. As a consequence, $\mathbb{P}(\tau_s < N)$ uniformly vanishes in s when N tends to infinity: for any $\epsilon > 0$, there exists $N_\epsilon > T$, such that for all s in $[0, T)$, $k_s^T(A_{a,b} \cap [N_\epsilon, \infty)) < \epsilon$.

ii) Convergence in law of the first-passage times: Considering a sequence $s_n \rightarrow s$, define the first-passage time $\tau_n = \inf\{t > s_n \mid X_t > L(t)\}$ with $X_{s_n^+} = L(s_n) - (V - v)$. The random variable τ_n is equivalently defined as $\tau_n = \tau'_n + (s_n - s)$, where we denote the translated first-passage time $\tau'_n = \inf\{t > s \mid X_t > L_n(t)\}$ with $X_{s^+} = L(s) - (V - v)$ and

$$L_n(t) = L(t - (s - s_n)) + (L(s) - L(s_n))e^{\alpha(t-s_n)}.$$

Since L is homogeneously Hölder continuous for a given exponent $H > 0$, for any

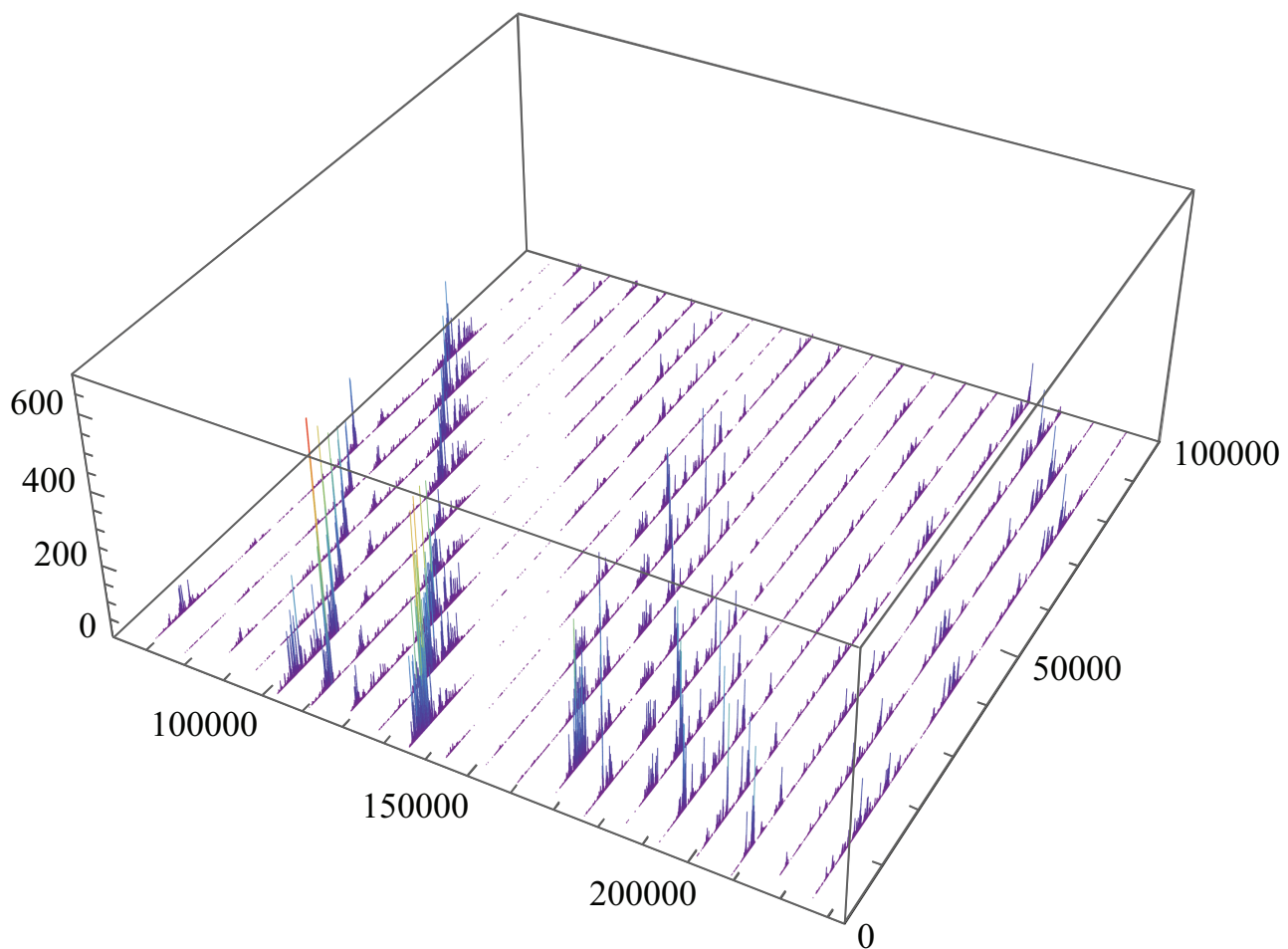


Figure 4.2: Monte-Carlo simulation of transition kernels k_s for a first-passage Markov chain with a Hölder continuous barrier ($H = 1/2$) and for different starting conditions s . Notice that, despite the apparent irregularity of the histograms of first-passage times, the overall probability that lies in a given time region varies continuously.

$N > 0$, there exists c_N such that

$$\forall t, s \in [0, N], |t - s| \leq \delta \quad \Rightarrow \quad |L(t) - L(s)| \leq c_N \delta^h,$$

so that we have, writing $s^- = \inf s_n > 0$

$$|L(t) - L_n(t)| \leq \left(1 + e^{-\alpha s^-}\right) c_N |s - s_n|^h.$$

Since L_n uniformly converges to L on every compact of \mathbb{R}^+ , τ'_n converges in law toward τ by Proposition 9. Now, since the deterministic quantity $|\tau'_n - \tau_n| = |s - s_n| \rightarrow 0$ when n tends to infinity, the sequence τ_n converges in law toward τ from [23].

iii) Feller property: Define the open sets $A'_{a,b} = \pi_s^{-1}(O_{(a,b)}) \cap [0, N_{\epsilon/4})$, where $N_{\epsilon/4}$ is defined as in *i*). Since for all s in $\mathbb{R}/T\mathbb{Z}$, any open set of $[0, N_{\epsilon/4}]$ is a τ_s -continuity set, convergence in law is equivalent to convergent in distribution: there exists $n_{(a,b),\epsilon} > 0$ such that:

$$\forall n > n_{(a,b),\epsilon}, \quad \left| \mathbb{P}(\tau_n \in A'_{a,b}) - \mathbb{P}(\tau \in A'_{a,b}) \right| \leq \epsilon/2.$$

Recapitulating, we have for all $n > n_{(a,b),\epsilon}$:

$$\begin{aligned} \left| k_{s_n}^T(O_{(a,b)}) - k_s^T(O_{(a,b)}) \right| &\leq \left| k_{s_n}(\pi_s^{-1}(O_{(a,b)})) - k_s(\pi_s^{-1}(O_{(a,b)})) \right|, \\ &\leq \left| k_{s_n}(A'_{a,b}) - k_s(A'_{a,b}) \right| \\ &\quad + \left| k_{s_n}(A_{a,b} \setminus A'_{a,b}) \right| + \left| k_s(A_{a,b} \setminus A'_{a,b}) \right|, \\ &\leq \epsilon/2 + \epsilon/4 + \epsilon/4 = \epsilon, \end{aligned}$$

proving the Feller continuity property. □

We now demonstrate the property of irreducibility 3:

Proof. For all s , consider $t \mapsto L_s(t)$ the periodic continuation of L for $t > s$. It is known [243] that the first-passage problems τ_s for an Ornstein-Uhlenbeck process U

satisfying

$$dU_t = \alpha U_t dt + dW_t, \quad \alpha < 0, \quad U_{s+} = L - (V - v),$$

with L_s can be mapped to first-passage problems σ_s for the standard Wiener process starting at $L - (V - v)$ and the barriers $B_s = \mathcal{D}_s[L]$ modified through the inverse Doob's transformation \mathcal{D} :

$$\mathcal{D}_s : L \mapsto \{B_s : t \geq s \mapsto L(s + \lambda_s(t)) \sqrt{1 - 2\alpha(t - s)}\},$$

where λ_s is the strictly increasing function

$$\lambda_s(t) = -\ln(1 + 2\alpha(t - s))/(2\alpha).$$

Notice that for all open $O_{a,b} = (a, b)$, $s < a < b$, we have

$$k_s(O_{a,b}) = \mathbb{P}(\tau_s \in O_{a,b}) = \mathbb{P}(\sigma_s \in O_{s+\lambda_s(a), s+\lambda_s(b)})$$

It is then enough to show the irreducible property reasoning on first-passage problem for the Wiener process with the $B_s = \mathcal{D}_s[L]$.

Consider a mollifier ψ and denote $\psi_\epsilon = \epsilon\psi(\cdot/\epsilon)$. Since B_s is continuous, we form a sequence of C^∞ -functions $B_s * \psi_{2^{-n}}$ that converges uniformly to B on every compact when n tends to infinity. Posit $\delta_n = \sup |B - B * \psi_{2^{-n}}| + 1/2^{-n}$, then $B_s^n = B_s * \psi_{2^{-n}} - \delta_n$ converges uniformly to B with $B_s^n < B_s$. Now, for all $s < a < b$, consider the discontinuous barrier D_s^n defined as

$$D_s^n(t) = \begin{cases} B_s^n(t) & \text{if } s \geq t \geq a \\ \sup_{a \leq s \leq b} L(s) & \text{if } t > a \end{cases}$$

Choose n such that $D_s^n(0) > L - (V - v)$ and denote σ_s^n the first-passage to D_s^n

with standard initial condition. Since $D_s^n(t) < B_s(t)$ for $t < a$ and $D_s^n(t) \geq B_s(t)$ otherwise, it is easy to see that

$$\mathbb{P}(\sigma_s \in O_{a,b}) > \mathbb{P}(\sigma_s^n \in O_{a,b}) .$$

We will finally show that $\mathbb{P}(\sigma_s^n \in O_{a,b}) > 0$. For all $t < a$, the probability density p of the Wiener process killed on D_s^n is the unique continuous solution of the Heat equation

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} \quad \text{with} \quad p_s = \delta_{L(s)-(V-v)} .$$

with absorbing boundary conditions $p_t(D_s^n(t)) = 0$ [38]. By the strong maximum principle, we know that $x \mapsto p_{a-}(x) > 0$ for $x < D_s^n(a)$, and consequently

$$\mathbb{P}(\sigma_s^n \in O_{a,b}) = \int_a^b \int_{-\infty}^{D_s^n(a^-)} q_s(D_s^n(a^+) - x) p_{a-}(x) dx ds > 0 ,$$

with q_s being the first-passage density of a Wiener process to a constant barrier at height x :

$$q_s(x) = \frac{1}{\sqrt{2\pi s^3}} e^{-\frac{x^2}{2s}} > 0 \quad \text{if} \quad s > 0 .$$

□

To summarize, when a leaky integrate-and-fire neuron is cyclically driven by an input, it produces trains of spikes which come under a statistical flavor. As a first-passage model, we are led to view the successive spiking times as the sample path of a discrete inhomogeneous Markov chain $(\mathcal{T}, \mathcal{P})$ with continuous space. The specifics of the model ensure that the resulting Markov chain $(\mathcal{T}, \mathcal{P})$ is ergodic: irrespective of the starting time, the mean sojourn-time in any time interval dt tends to the probability of dt under the unique invariant measure μ . Incidentally, we establish that, for each deterministic input, there exists a unique corresponding invariant spiking probability measure. In the next section, we use this well-posed framework to set up a numerical

method to investigate the effect of the barrier regularity on the measure it gives rise to.

4.1.3 Numerical Simulation of the Markov Chain

If the first-passage Markov chain $(\mathcal{T}, \mathcal{P})$ is ergodic, due to the possible irregularity of the barrier, numerical simulation of its invariant measure demands we resort to approximation schema. To justify this approach, we adapt a general result from [107], setting out sufficient conditions for a sequence of Markov chains $(\mathcal{T}^N, \mathcal{P}^N)$ to converge toward a limit chain $(\mathcal{T}, \mathcal{P})$ in the sense that the law \mathcal{P}^N converges weakly toward \mathcal{P} when N tends to infinity.

Theorem 7 (adapted from [107]). *If a sequence of strongly-Feller Markov chains $(\mathcal{X}^N, \mathcal{Q}^N)$ defined on a compact state space \mathbb{S} is such that:*

- *for any s in \mathbb{S} , the kernel probability measures q_s^N converge in law toward a limit probability measure q_s ,*

then,

- *any limit in law of a sequence ν_n of invariant measures of $(\mathcal{X}^N, \mathcal{Q}^N)$, is an invariant measure of the Markov chain $(\mathcal{X}, \mathcal{Q})$ corresponding to the limit kernel q .*

In particular, if all $(\mathcal{X}^N, \mathcal{Q}^N)$ and $(\mathcal{X}, \mathcal{Q})$ are ergodic, the sequence ν_n is uniquely defined and so is its limit distribution ν , which is the stationary measure of $(\mathcal{X}, \mathcal{Q})$.

For our purpose, an efficient approximation strategy of μ consists in exhibiting a sequence of ergodic strongly-Feller Markov chains $(\mathcal{T}^N, \mathcal{P}^N)$ whose kernels k_s^N converge to k_s^T in law. This is accomplished by considering a sequence of first-passage Markov chains $(\mathcal{T}^N, \mathcal{P}^N)$ defined for the piecewise continuous periodic barriers L_N

that interpolates L on the dyadic points $D_N = \{k2^{-N}T \mid 0 \leq k < 2^N\}$:

$$L_N : t \in \mathbb{C} \mapsto \mathbb{E} [U_t \mid U_{k2^{-N}T} = L(k2^{-N}T), 0 \leq k < 2^N]$$

where \mathbb{E} denotes the expectation with respect to the law of U (see [217]). Such Markov chains are ergodic by the same argument as for $(\mathcal{T}, \mathcal{P})$. Moreover, since we restrain ourselves to barriers L that are homogeneously Hölder continuous, the sequence L_N converges uniformly toward L (actually there exist $c_T > 0$ such that $|L - L_N|_\infty < c_T 2^{-Nh}$ where H is the Hölder exponent). From Proposition 9 and the same argument as in the proof of Property 2, the uniform convergence of L_N to L entails the convergence in law (and in distribution) of k_s^N toward k_s^T , showing the cogency of approximating L by L^N .

In addition to providing a valid numerical method, the previous approach provides an easy description of the input dC that gives rise to L . The central results is adapted from [218]:

Theorem 8. *There exists a Schauder basis of continuous functions $\psi_{n,k}$ compactly supported on $S_{n,k} = [k2^{-n+1}T, (k+1)2^{-n+1}T]$ such that, for all $N > 0$,*

$$\mathbb{E} [U_t \mid U_{k2^{-N}T}, 0 \leq k < 2^N] = \sum_{0 \leq n < N} \sum_{0 \leq k < 2^{n-1}} \psi_{n,k}(t) \cdot \Xi_{n,k}$$

where the $\xi_{n,k}$ are the independent standard Gaussian variables

$$\Xi_{n,k} = \int_0^T \phi_{n,k}(t) dW_t, \quad \phi_{n,k} = \psi'_{n,k} - \alpha \psi_{n,k}.$$

and the thus-defined functions $\phi_{n,k}$ form an orthonormal system of $L^2[0, T]$

Equipped with this result, it is easy to see that writing the input dC as a “Gaussian

white noise”

$$dC_t = \sum_{0 \leq n} \sum_{0 \leq k < 2^{n-1}} \phi_{n,k}(t) \cdot \Xi_{n,k}, \quad \Xi_{n,k} \text{ i.i.d} \sim \mathcal{N}(0, 1),$$

the statistics of the resulting random barrier

$$L_t = V - \int_0^t e^{\alpha(t-s)} dC(s) = V - \sum_{0 \leq n} \sum_{0 \leq k < 2^{n-1}} \psi_{n,k}(t) \cdot \Xi_{n,k},$$

is the same as for an Ornstein-Uhlenbeck process centered around zero and translated upward by V . Moreover, setting $\xi_{0,0} = 0$, we naturally enforce the periodic condition $L(t) = L(T) = V$.

However, we are interested in studying the distribution of spiking events of a neuron cyclically driven by a deterministic input. Accordingly, suppose now $dC(t) = dC_t(\omega)$ is a realization of our “Gaussian white noise”, i.e. a frozen noise. Then, $L(t) = L_t(\omega)$ is the sample path of an Ornstein-Uhlenbeck bridge translated upward, which we know is almost-surely a homogeneously Hölder continuous function of exponent $1/2^+$ (with probability one, L is H -Hölder continuous if and only if $H > 1/2$). For this reason, we denote such an input $dC^{1/2}$, the associated barrier $L^{1/2}$ and the coefficients $\xi_{n,k}^{1/2}$.

From there, let us consider Ω_ξ the set of coefficients $\xi_{n,k}$ for which the continuous barriers of the form

$$L_N(t) = V - \sum_{0 \leq n < N} \sum_{0 \leq k < 2^{n-1}} \psi_{n,k}(t) \cdot \xi_{n,k},$$

converge uniformly on \mathbb{C} . It can be shown [218] that Ω_ξ contains the set

$$\Omega'_\xi = \{ \xi_{n,k} \in \mathbb{R}^\mathbb{N} \mid \exists \delta < 1, \exists N > 0, \forall n > N, \max_k |\xi_{n,k}| \leq 2^{n\delta/2} \}.$$

From this, we immediately deduce that given $L^{1/2}$, for any real H such that $0 < h < 1$, the barrier L^H

$$L^H(t) = V - \sum_{0 \leq n} \sum_{0 \leq k < 2^{n-1}} \psi_{n,k}(t) \cdot \xi_{n,k}^H, \quad \xi_{n,k}^H = 2^{n(H-1/2)} \xi_{n,k},$$

is well-defined as a continuous function of \mathbb{C} . Keeping this in mind, we have at our disposal a well-known result [51] relating the local Hölder exponent of a function to the asymptotic behavior of the coefficients of its decomposition in the Schauder basis, which we have proven in Chapter 2. Adjusting to our situation, it directly entails that for all H , $0 < H < 1$, the barriers L^H are almost-surely homogeneously H^+ -Hölder continuous. Further consideration actually reveals that we have the following:

Property 4. *Given the countable sequence of reals $\xi_{n,k} = \Xi_{n,k}(\omega)$ where $\Xi_{n,k}$ are i.i.d. $\mathcal{N}(0,1)$, the application $H \mapsto L^H(\xi)$, defined from $(0,1)$ to the set of periodic continuous functions $C(\mathbb{C})$ equipped with the topology of the L^∞ -norm, is a continuous mapping and for all H , L^H is H^+ -Hölder continuous.*

Therefore, we can continuously control the asymptotic Hölder continuity of the effective barrier driving the activity of a leaky integrate-and-fire neuron by smoothly changing the coefficient $\xi_{n,k}^H$ used to construct piecewise approximations L_N^H . In Section 4.3, we use such a construction to simulate numerically the instantaneous mean firing rate of an integrate-and-fire neuron that is cyclically driven by a steady input current and periodically perturbed by a “frozen” Gaussian white noise. In order to emphasize the effect of the varying Hölder regularity, we adopt a slightly modified version of our barriers L^H , by weighting them with a continuous function $H \mapsto c(H)$ under the form $L^H = c(H)(L^H - L^H(0)) + L^H(0)$. The function c is chosen so that the newly formed barriers cause the neuron to fire with an overall mean firing rate (as opposed to the instantaneous mean firing rate which is time-dependent) remains constant when changing H . Formally, this constraint is equivalent to holding a constant

mean inter-spike time

$$\int_0^T \left(\int_s^\infty (t-s) \kappa_s^H(dt) \right) \mu_H(ds)$$

while varying H . Notice that for the sake of well-posedness, the kernels that intervene in the formulation of the mean inter-spike time are computed for a periodic barrier L^H but defined on $[0, \infty)$ instead of being wrapped on $[0, T)$. We represent in Figure 4.1.3 a continuous family of barriers that are generated in such a fashion.

To complete our numerical program aiming at simulating our first-passage Markov chain, we have yet to propose a computational method to evaluate the first-passage transition kernels k_s^T for any H -continuous barrier. Since we believe the algorithm we produce can be generalized to the multidimensional setting, we devote the entire following section to the self-contained explanation of its principle.

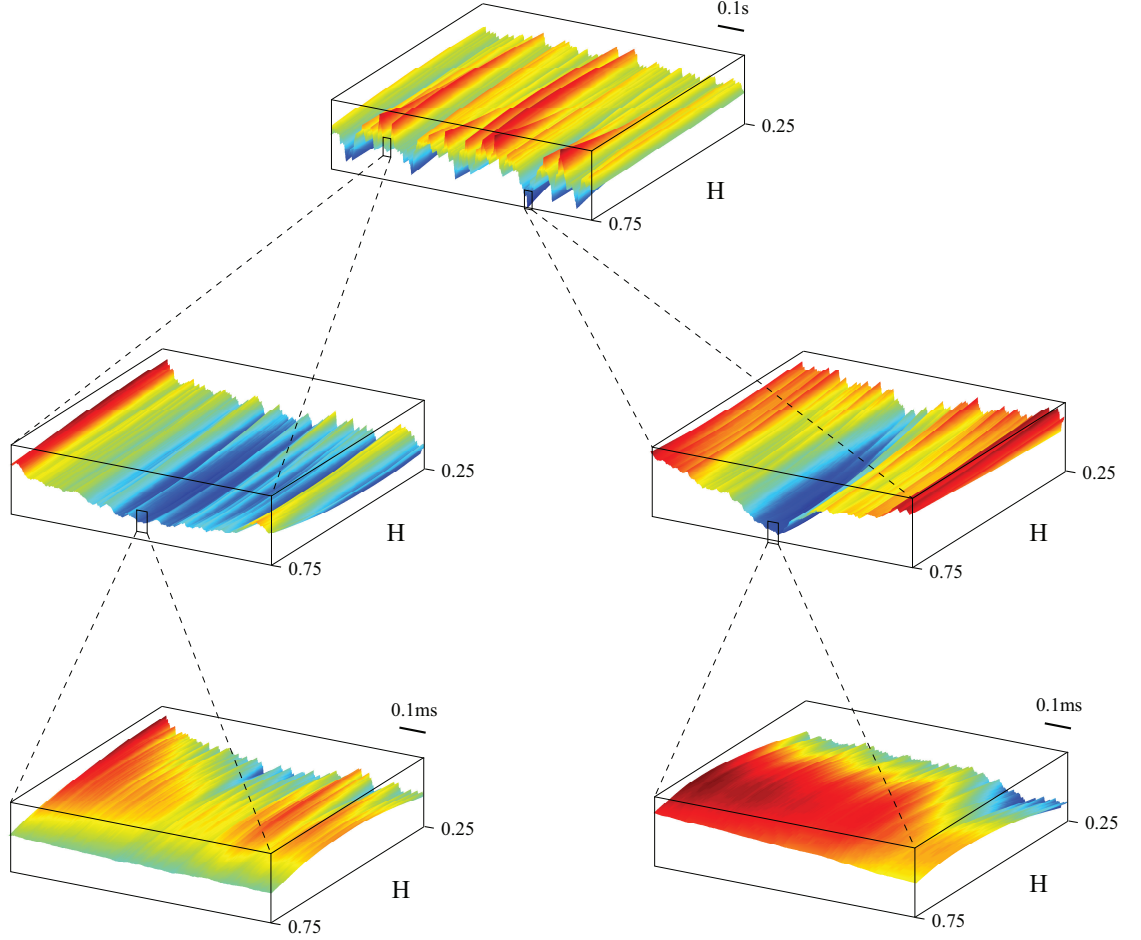


Figure 4.3: Continuous family of Hölder continuous barriers for Hölder exponent H varying from 0.25 to 0.75. The color encodes relative heights of the barrier in each graph independently. For each exponent H , the barrier is made using the construction discussed in the main text for an Ornstein-Uhlenbeck bridge of leak $\alpha = -100\text{s}^{-1}$. Each barrier is weighted with a continuous function c that is chosen so that the first-passage Markov chain corresponding to a leaky integrate-and-fire neuron with same α , has the same overall mean firing. The barriers get rougher with decreasing Hölder exponents, as shown when zooming-in the trajectory below the persistence length $|\alpha|^{-1} = 10\text{ms}$.

4.2 Algorithm for the First-Passage Times with Hölder Continuous Boundaries

The Markov chain framework proves amenable to the numerical simulation of the mean firing activity of cyclically driven integrate-and-fire neurons. The mean firing function can be viewed as the unique invariant measure of an ergodic first-passage Markov chain. Thus, the numerical estimation of the firing rate requires applying the theory of Markov chains [90, 151, 214], which lies at the heart of the numerical methodologies used for Monte-Carlo simulations. However, this supposes that we can generate stochastic spiking events according to the law of a first-passage time to a fluctuating barrier with varying initial conditions in an efficient way. The need for simulating a continuum of transition kernels², precludes any attempt to tabulate their probability laws.

As exemplified by Chapter 4, studying the first-passage time problem remains among the “tough” problems in the theory of stochastic processes despite considerable effort. Probability distributions of passage times are known analytically only for the most trivial situations, such as a Wiener process first crossing affine boundaries. For our practical problem, first-passage times must be computed by numerical integration. When concerned with the leaky integrate-and-fire model, numerical integration through the discretization of Volterra equation is often advocated to directly compute the density of first-passages [156, 5]. Nevertheless, this otherwise very-fast method fails to produce a convergent schema for non-differentiable barrier functions and is not adapted to situations when the distribution of passage times becomes singular.

Focusing on irregular barriers, our situation demands that we estimate the distribution of spiking events by Monte-Carlo methods, thus simulating many sample paths of

²Each kernel corresponds to a particular initial condition, prescribed by the timing of the preceding spike.

the process until it first crosses the boundary. Such an approach traditionally carries both practical and theoretical difficulties whether focusing on the computational cost of the method or on the accuracy of the returned times [164, 94, 26].

In this section, we improve the numerical approach by designing a stochastic dichotomic-search algorithm for first-passage time, based on the construction of linear diffusion developed in Chapter 2 and the first-passage analysis of the Wiener process in Chapter 3.

Since the leaky integrate-and-fire neuron is set in a one-dimensional setting, we only consider the class of real Gauss-Markov processes [105, 92]. For this class which comprises the Ornstein-Uhlenbeck process [230], numerical integration can be performed in a completely different way than it is in naive time-discretization computational scheme. As closed-form knowledge of the conditioning formula permits error-free sample path constructions at large time steps [79, 40, 19], the only issue is estimating if the process has exceeded the boundary within a given timestep. Thus, the process may be path-wise simulated at high resolution only on regions where such a resolution is warranted [73, 186]. In particular, numerical effort can be devoted to accurately reconstruct the process only when close to the boundary (for example through a dichotomic refining procedure), but not when far from the boundary [80]. Such an approach has already been advantageously used to evaluate expectation [82, 83].

We study computationally the first-passage problem of generic Gauss-Markov processes, viewed as time-dependent Ornstein-Uhlenbeck processes, to address the biologically relevant situation of a time-dependent leak [212, 76]. Such processes are solutions of the following stochastic equation

$$dX_t = \alpha(t) \cdot X_t dt + \sigma(t) \cdot dW_t, \quad (4.5)$$

where α is the leak constant and $\sigma > 0$ is the diffusion coefficient. As noticed

when discussing the first-passage time of a Wiener process with a fluctuating barrier, we already know that two factors are important: the regularity of the barrier, and the regularity of the coefficients in equation (4.5). In addition to determining the existence and regularity of a continuous density function for first-passage time, it prescribes the speed of convergence of first-passage times' computation [244, 242, 168]. We therefore make the two following assumptions:

Coefficient Regularity Assumption: The Gauss-Markov processes are solution of a linear stochastic equation with time-dependent *non-positive, bounded* function α and with time-dependent positive, homogeneously Hölder continuous function σ .

Barrier Regularity Assumption: The barrier function is assumed to be *homogeneously Hölder continuous* and *non-negative*.

Since our algorithm is probabilistic in nature, it can return erroneous approximate time values: in such cases, it always produces approximate crossing times that are not first-passage times. However, the probability of occurrences of such errors can be tightly controlled. Our algorithm is designed not to search for first-passages in time intervals where the probability of such crossings is known to be less than a parameter value $\epsilon > 0$. As a consequence, our algorithm has the following essential property:

Error Bound: Given a real $\eta > 0$ and a recursive depth N , choosing the parameter $\ln \epsilon$ to satisfy the criterion

$$\ln \epsilon < \ln \eta - (N - 1) \ln 2,$$

ensures that our algorithm returns an approximate first-passage time τ_N with a resolution of 2^{-N} , and with an error tolerance

$$\mathcal{E}(N, \epsilon) = \mathbb{P}(\tau^N - \tau > 2^{-N-1}) \leq \eta,$$

where τ is the true first-passage.

4.2.1 Probabilistic Dichotomic Search Algorithm

Our algorithm consists in implementing recursively a dichotomic search for the first crossing of simulated sample paths with the boundary, assuming the following facts:

Assumption 1. The Gauss-Markov process X under scrutiny are solutions of (4.5) with σ in the set of homogeneously Hölder continuous functions and with α being a non-positive function, bounded on every compact support $[0, T]$, $T > 0$.

Assumption 2. The barrier L is a non-negative homogeneously Hölder continuous function on every compact support $[0, T]$, $T > 0$.

Bearing in mind these restrictions under which we operate, we proceed to explain the algorithm in several stages.

First, we explicit the recursive schema for simulating Gauss-Markov sample paths. Second, we give the plain dichotomic search algorithm for first-passages. Third, we elaborate the probabilistic version of the dichotomic search by adding a probabilistic screening at each recursive step: the recursive construction of sample paths is only further if the probability of a crossing within a given time interval, is estimated greater than some small parameter ϵ . Fourth, in order to check the previous probabilistic screening, we establish an upper bound to the conditional probability that a crossing happens in a given segment, knowing the value of the sample paths at the endpoints of that segment. Fifth, we detail the base case responsible for the termination of the algorithm at the resolution limit: we simulate a first-passage in an end segment with the exact same probability as the probability that the Gauss-Markov process crosses a particular continuous function that interpolates the barrier at the endpoints. Sixth, we summarize formally the whole algorithm by giving its condensed mathematical

formulation. Finally, we illustrate the algorithm on two simple examples, one of which is analytically solvable.

Recursive Schema for the Sample Paths

Here, we are just translating the Lévy-Ciesielski construction for Gauss-Mrkov processes into a simple recursive algorithm. But beforehand, let us recall some short notations to simplify the writing of forthcoming expressions

$$l_{n,k} = (2k) 2^{-n}, m_{n,k} = (2k+1) 2^{-n}, r_{n,k} = 2(k+1) 2^{-n}.$$

As usual, we write $S_{n,k} = [l_{n,k}, r_{n,k}]$ where $S_{n,k}$ is the support of an elementary function used for the Lévy-Ciesielski construction of the Gauss-Markov process. With these notations, remark that $D_{N-1} = \{l_{N,k}, r_{N,k} \mid 0 \leq k < 2^{N-1}\}$ and that $D_N \setminus D_{N-1} = \{m_{N,k} \mid 0 \leq k < 2^{N-1}\}$. Also notice that the support $S_{n,k}$ constitute a binary tree of nested compact supports: up to the point $m_{n,k}$, we have the partition $S_{n,k} = S_{n+1,2k} \cup S_{n+1,2k+1}$ and we can identify $m_{n,k}$ with $r_{n+1,2k}$ and $l_{n+1,2k+1}$. We say that a sample path is simulated up to depth N or with resolution 2^{-N} if the value of $t \mapsto X_t(\omega)$ for a given ω have been simulated for every time t in the dyadic set D_N . Now, assume we can generate a collection of random variables $\Xi_{n,k}$ simulating independent normal laws $\mathcal{N}(0, 1)$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, remember that the law of a generic Gauss-Markov bridge $\{X_t \mid X_l = x, X_r = z\}$ follows the distribution of $\mathcal{N}(\mu(t), \sigma^2(t))$ where $\mu(t)$ is the time-dependent mean

$$\mu(t) = \frac{g(t)}{g(l)} \cdot \frac{h(r) - h(t)}{h(r) - h(l)} \cdot x + \frac{g(t)}{g(r)} \cdot \frac{h(t) - h(l)}{h(r) - h(l)} \cdot z, \quad (4.6)$$

and $\sigma(t_y)$ is the time-dependent standard deviation defined by

$$\sigma(t)^2 = g^2(t) \cdot \frac{(h(t) - h(l))(h(r) - h(t))}{h(r) - h(l)}. \quad (4.7)$$

with the conventions:

$$g(t) = e^{\int_0^t \alpha(v) dv}, f(t) = \sigma(t) e^{-\int_0^t \alpha(v) dv}, h(t) = \int_0^t f^2(s) ds,$$

Expressions (4.6) and (4.7) provide us with an exact iterative schema to simulate a sample path of X appropriately defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and up to a depth N . Indeed, supposing that we have simulated the values of X_t for two consecutive dyadic points $l_{n,k}$ and $r_{n,k}$ in D_{N-1} , we can simulate the outcome of X_t at the midpoint $m_{n,k}$ in $D_N \setminus D_{N-1}$ by drawing a random variable

$$X_{m_{n,k}} = \sigma(m_{n,k}) \cdot \Xi_{n,k} + \mu(m_{n,k}),$$

where $\sigma(m_{n,k})$ and $\mu(m_{n,k})$ are given by the conditioning formula at $m_{n,k}$. The basis step of the recurrence, i.e. the simulation of a value $z_{0,0}$ at $r_{0,0}$ knowing $x_{0,0}$ at $l_{0,0}$, is immediate by direct application of the forward conditioning formula:

$$p(x, t | x_0, t_0) = \frac{1}{g(t) \sqrt{2\pi(h(t) - h(t_0))}} \cdot \exp \left(-\frac{\left(\frac{x}{g(t)} - \frac{x_0}{g(t_0)} \right)^2}{2(h(t) - h(t_0))} \right).$$

From there, the refinement of a simulated sample path $t \mapsto X_t(\omega)$ for a given ω satisfying $X_{l_{n,k}}(\omega) = x_{n,k}$ and $X_{r_{n,k}}(\omega) = x_{n,k}$ and up to depth $N \geq n$, is implemented recursively as follows:

The function *basecase* in Procedure 1 merely implements the termination of the

Procedure 1 *subdivide*($l_{n,k}, x_{n,k}, r_{n,k}, z_{n,k}$)

if $r_{n,k} - l_{n,k} = 2^{-N}$ **then**
 basecase($l_{n,k}, x_{n,k}, r_{n,k}, z_{n,k}$)
else
 simulate the value $y_{n,k}$ of the sample path at $m_{n,k}$
 subdivide($l_{n+1,2k}, x_{n,k}, r_{n+1,2k}, y_{n,k}$)
 subdivide($l_{n+1,2k+1}, y_{n,k}, r_{n+1,2k+1}, z_{n,k}$)
end if

recursion. The interest of that schema lies in the fact that it is constructed on a binary tree of nested supports, allowing us to refine the simulation of a sample path on any given $S_{n,k}$ independently of others disjoint supports $S_{n',k'}$.

Recursive Dichotomic Search of First-Passage Times

In order to compute first-passage times in a support $S_{n,k}$, a straightforward algorithm consists in constructing sample paths with depth N until a drawn value occurs above the barrier L , or until the constructed path reaches the endpoint $r_{n,k}$ of $S_{n,k}$ staying below L .

As observable in Procedure 9, we just have to alter the recursive schema in Procedure 1 to integrate the following idea: if there is a time s in $S_{n,k}$ for which $X_s(\omega) \geq L(s)$, by continuity of the sample paths, we know that a crossing has occurred before s ; therefore, we have to disregard continuing the simulation of the sample path $t \mapsto X_t(\omega)$ for time t following s .

Procedure 9 will form the backbone of our algorithm and we refer to it as a dichotomic

Procedure 2 $passage(l_{n,k}, x_{n,k}, r_{n,k}, z_{n,k})$

```

if  $r_{n,k} - l_{n,k} = 2^{-N}$  then
  return  $basecase(l_{n,k}, x_{n,k}, r_{n,k}, z_{n,k})$ 
else
  simulate the value  $y_{n,k}$  of the sample path at  $m_{n,k}$ 
  if  $y_{n,k} \geq L(m_{n,k})$  then
    return  $passage(l_{n,k}, x_{n,k}, m_{n,k}, y_{n,k})$ 
  else
    if  $time = passage(l_{n,k}, x_{n,k}, m_{n,k}, y_{n,k}) > 0$  then
      return  $time$ ;
    else
      return  $passage(m_{n,k}, y_{n,k}, r_{n,k}, z_{n,k})$ 
    end if
  end if
end if

```

search algorithm. For any simulated sample path $t \mapsto X_t(\omega)$, the method relies on the recursive exploration of the binary tree of dyadic segments $S_{n,k}$, effectively

investigating in the prefix order every segment for which $x_{n,k}$ is below the barrier L . However, for a given sample path, it is obviously possible that no first-passage occurs within $S_{0,0}$, the root segment of the binary tree of supports. We then need to simulate such sample paths $t \mapsto X_t(\omega)$ for time t later than $r_{0,0}$. In other words, we have to simulate values of the sample path for every time step $d_0 = r_{0,0} - l_{0,0}$ by successive application of the forward conditioning formula (4.8), and then initiate a recursive search during any of these time steps.

Procedure 3 *searchfirstpassage*

```

 $r_{0,0} = 0$ 
 $z_{0,0} = 0$ 
time = 0
while time = 0 do
   $l_{0,0} \leftarrow r_{0,0}$ 
   $x_{0,0} \leftarrow z_{0,0}$ 
   $r_{0,0} \leftarrow r_{0,0} + d_0$ 
  simulate the value  $z_{0,0}$  of the sample path at  $r_{0,0}$ 
  time = passage( $l_{0,0}, x_{0,0}, r_{0,0}, z_{0,0}$ )
end while
return time

```

Remark 3. It is possible that the expected first-passage time diverges (for instance, in the case of a Wiener process with a constant barrier). To circumvent this predicament, we have to limit the scope of the search to a given compact segment $[0, T]$.

Probabilistic Screening of First-Passage Times

Denote $P_{n,k}$ the probability for the process X to cross the boundary L between $l_{n,k} = k2^{-n+1}$ and $r_{n,k} = (k+1)2^{-n+1}$ knowing $X_{l_{n,k}}$ and $X_{r_{n,k}}$, the value of X on two successive points in D_{N-1} . Thus defined, $P_{n,k}$ is a random variable on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. More formally, we write $P_{n,k}$ as the conditional probability

$$P_{n,k} = \mathbb{P} \left(\tau_{l_{n,k}}^L \in S_{n,k} = [l_{n,k}, r_{n,k}] \mid X_{l_{n,k}}, X_{r_{n,k}} \right).$$

Still, as X is a Markov process, $P_{n,k}$ knowing $X_{l_{n,k}} = x_{n,k}$ and $X_{r_{n,k}} = y_{n,k}$, becomes a deterministic function of the times $l_{n,k}$ and $r_{n,k}$ and the corresponding values $x_{n,k}$ and $y_{n,k}$. We will simply write the outcome of $P_{n,k}$ as $P_{n,k}(\omega) = P_{n,k}(x_{n,k}, z_{n,k})$ making the dependence on $l_{n,k}$ and $r_{n,k}$ implicit.

Assuming that we are provided with an estimation of $P_{n,k}(x_{n,k}, z_{n,k})$, we can choose to refine the simulation of $t \mapsto X_t(\omega)$ only if the probability for a crossing to happen is larger than what is admissible: typically a small positive real ϵ chosen according to the total number of simulated paths and the desired level of accuracy. Formally stated, at a given depth n , we only investigate a sample path $t \mapsto X_t(\omega)$ between two consecutive dyadic times $l_{n,k}$ and $r_{n,k}$ if the outcome of $P_{n,k}$ for the particular occurrence ω is larger than ϵ .

The idea behind this probabilistic screening is to search for first-passage times by only simulating the process on the dyadic points where the outcomes happen close enough to the boundary L . Since exact computation of $P_{n,k}(x_{n,k}, z_{n,k})$ is impossible, we need a simple analytical upper bound $B_{n,k}$ to $P_{n,k}$: we can always discard recursive searches in supports $S_{n,k}$ for sample path satisfying $P_{n,k}(\omega) \leq B_{n,k}(\omega) \leq \epsilon$.

We refer to the underlying algorithm of Procedure 4 as a probabilistic dichotomic search algorithm. Two rules of exploration of the binary tree $S_{n,k}$ are implemented: we discard the branches of the tree issued from a root segment for which $B_{n,k}(\omega) \leq \epsilon$; we only explore the branches occurring before any dyadic point t for which $X_t(\omega)$ exceeds $L(t)$. We detail in the following section how to compute an upper bound to the probability $P_{n,k}$.

Procedure 4 $passage(l_{n,k}, x_{n,k}, r_{n,k}, z_{n,k})$

```

if  $r_{n,k} - l_{n,k} = 2^{-N}$  then
  return  $basecase(l_{n,k}, x_{n,k}, r_{n,k}, z_{n,k})$ 
else
  compute the upper bound  $B_{n,k}(x_{n,k}, z_{n,k})$ 
  if  $B_{n,k}(x_{n,k}, z_{n,k}) > \epsilon$  then
    simulate the value  $y_{n,k}$  of the sample path at  $m_{n,k}$ 
    if  $y_{n,k} \geq L(m_{n,k})$  then
      return  $passage(l_{n,k}, x_{n,k}, m_{n,k}, y_{n,k})$ 
    else
      if  $time = passage(l_{n,k}, x_{n,k}, m_{n,k}, y_{n,k}) > 0$  then
        return  $time$ ;
      else
        return  $passage(m_{n,k}, y_{n,k}, r_{n,k}, z_{n,k})$ 
      end if
    end if
  end if
end if

```

Upper Bound to the Probability of First-Passage Time

For an homogeneously Hölder continuous non-negative boundary, we introduce the binary tree of minima of L on the compact supports $S_{n,k}$ as

$$\underline{L}_{n,k} = \inf_{t \in S_{n,k}} L(t).$$

We stress that this structure needs to be computed for the implementation of the algorithm: if the threshold function L is of simple analytical expression, the values $\underline{L}_{n,k}$ can be evaluated dynamically; otherwise the tree of minima should be evaluated numerically once and for all.

It is possible to show [217] that, as long as the function α remains non-positive on $S_{n,k}$ and that $\underline{L}_{n,k} \geq \max(x, 0)$, we have

$$\mathbb{P}(\tau_{l_{n,k}}^L < r_{n,k} \mid X_{l_{n,k}} = x, X_{r_{n,k}} = z) \leq \mathbb{P}\left(\tau_{l_{n,k}}^{\underline{L}_{n,k}} < r_{n,k} \mid Y_{l_{n,k}} = x, Y_{r_{n,k}} = \frac{g(l_{n,k})}{g(r_{n,k})}z\right),$$

with $\tau'_{l_{n,k}}^{\underline{L}_{n,k}} = \inf\{t > l_{n,k} \mid Y_t \geq \underline{L}_{n,k}\}$, where Y is a scaled time-changed Wiener process defined by

$$Y_t = \frac{g(t)}{g(t_x)} W_{h_{t_x}(t)}, \quad h_{t_x}(t) = g^2(t_x)(h(t) - h(t_x)).$$

If we denote $B_{n,k}(x, z)$ the probability of Y to reach a constant threshold $\underline{L}_{n,k}$ knowing that $Y_{l_{n,k}} = x$ and $Y_{r_{n,k}} = g(l_{n,k})z/g(r_{n,k})$, expression (??) allows us to define a random variable $B_{n,k}$ on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$

$$B_{n,k} = \exp \left(-2 \cdot \frac{(\underline{L}_{n,k} - X_{l_{n,k}}) \left(\frac{g(r_{n,k})}{g(l_{n,k})} \underline{L}_{n,k} - X_{r_{n,k}} \right)}{g(r_{n,k})g(l_{n,k})(h(r_{n,k}) - h(l_{n,k}))} \right), \quad (4.8)$$

satisfying the desired upper bound condition $P_{n,k} \leq B_{n,k}$.

We emphasize that the need to compute a binary tree of minimum $\underline{L}_{n,k}$ is a crucial step to implement our algorithm. Indeed, we can be in situations where the starting time t_0 of the process is not known in advance, such as for the simulation of a train of spikes for a noisy “leaky integrate and fire” neuron. In this model, after a first-passage time (or spiking event) with a given barrier (or firing threshold), the neuron is reset to its initial condition (or resting potential), its state (or electric potential) then evolves as an Ornstein-Uhlenbeck process until it reaches again the firing threshold. It is easy to show that the evaluation of consecutive first-passage events is actually equivalent to consider a first-passage problem with a class of time-shifted barrier for varying initial condition [179]. In such cases, our construction schema uses a *floating* binary tree of supports $S'_{n,k}$ which depends on the time shift $t_0 = l_{0,0}$. We cannot assume it will superimpose the binary tree of minimum $\underline{L}_{n,k}$, since for computational cost we compute it once and for all on the partition $S_{n,k}$ for a given shift, say $t_0 = 0$.

This complication is overcome by defining

$$B_{n,k} = \exp \left(-2 \cdot \frac{(\underline{L}'_{n,k} - X_{l_{n,k}}) \left(\frac{g(r_{n,k})}{g(l_{n,k})} \underline{L}'_{n,k} - X_{r_{n,k}} \right)}{g(r_{n,k})g(l_{n,k})(h(r_{n,k}) - h(l_{n,k}))} \right)$$

with $\underline{L}'_{n,k} = \min(\underline{L}_{n,k_1}, \underline{L}_{n,k_2})$, where k_1, k_2 are computed as the only two admissible integers satisfying $S'_{n,k} \cap S_{n,k_1} \neq \emptyset$ and $S'_{n,k} \cap S_{n,k_2} \neq \emptyset$ up to the endpoints. Indeed the minimum of the barrier L on given *floating* supports $S'_{n,k}$ is always less than the minimum of L on $S_{n,k_1} \cup S_{n,k_2} \supset S'_{n,k}$.

Treatment of the Base Case

If N denotes the maximum depth of exploration, a naive idea is to consider that no crossing has occurred in a limit segment $S_{N+1,k}$ of length 2^{-N} , unless the last simulated value $z_{N+1,k}$ is larger than $L(z_{N+1,k})$. As the algorithm explores the tree of nested support $S_{n,k}$ in the prefix order³, any such crossing in a limit segment $S_{N+1,k}$ has to be a first-passage, whose timing can be arbitrarily set to $m_{N+1,k}$. However, this approach is very unsatisfactory because it neglects the potential occurrence of a crossing between two consecutive points in D_N which are below the barrier L . This eventuality cannot be discarded and is a major source of error in the simulation of first-passage times. To circumvent this issue, it is much more preferable to assess $P_{N+1,k}(x_{N+1,k}, z_{N+1,k})$, the probability of occurrence of a crossing in a limit segment $S_{N+1,k}$.

Since the function $t \mapsto \underline{L}_{N+1,k}(t)$ is a piecewise constant function on $S_{0,0}$, the first-passage problem for this presumably discontinuous barrier is not well posed and the upper bound $B_{N+1,k}(x_{N+1,k}, z_{N+1,k})$ yields an inconvenient estimate of $P_{N+1,k}$. For some x and z such that $x \leq L(l_{N,k})$ and $z \leq L(r_{N+1,k})$, it is actually preferable to

³Remember that for any explored segment $S_{n,k}$, we necessarily have $x_{n,k} < L(l_{n,k})$.

approximate $P_{N+1,k}(x, z)$ by

$$Q_{N+1,k}(x, z) = \exp \left(-\frac{2 \left(L(l_{N+1,k}) - x \right) \left(L(r_{N+1,k}) - z \right)}{g(r_{N+1,k})g(l_{N+1,k}) \left(h(r_{N+1,k}) - h(l_{N+1,k}) \right)} \right). \quad (4.9)$$

The quantity $Q_{N+1,k}(x, z)$ just gives the probability that the process X conditioned by $X_{l_{N+1,k}} = x$ and $X_{r_{N+1,k}} = z$, crosses the barrier μ^N , which is defined such that for any t in $S_{N+1,k}$, $\mu^N(t)$ is the expected value of X_t knowing $X_{l_{N+1,k}} = L(l_{N+1,k})$ and $X_{r_{N+1,k}} = L(r_{N+1,k})$. Thus, the algorithm approximates the barrier L by a piecewise continuous function μ^N interpolating L on the dyadic numbers D_N . With such a piecewise barrier, the first-passage problem for the Wiener process is well-posed and yields consistent estimates of $P_{N+1,k}$ at the limit depth when N tends to infinity. Moreover, notice that in the limit of large N , the interpolating approximation of the barrier rapidly converges to the linear piecewise interpolation.

Procedure 5 *basecase*($l_{n,k}, x_{n,k}, r_{n,k}, z_{n,k}$)

```

if  $z_{n,k} \geq L(r_{n,k})$  then
  return  $m_{n,k}$ 
else
  compute the estimate  $Q_{n,k}(x_{n,k}, z_{n,k})$ 
  draw  $\Xi$  a uniformly between zero and one
  if  $\Xi \leq Q_{n,k}(x_{n,k}, z_{n,k})$  then
    return  $m_{n,k}$ 
  else
    return 0
  end if
end if

```

Formal Definition of the Algorithm

We now recapitulate formally the recursive procedure for the dichotomic search of a first-passage within the segment $[0, 1]$. We aim at computing approximate occurrences

τ^N of the true first-passage time τ , with an resolution of 2^{-N} . In that perspective, assume we are provided with two families of independent identically distributed random variables $\{\Xi_{n,k}\}$, $0 \leq n < N$, $0 \leq k < 2^{n-1}$ of normal law $\mathcal{N}(0, 1)$, and $\{v_k\}$, $0 \leq k < 2^N$, of uniform law $U(0, 1)$.

Definition 6. The definition of the approximate sample path τ^N of our algorithm proceeds as follows:

1. The values of a sample path X at the endpoints of $[0, 1]$ are given by

$$\mathcal{I}_0 = \{0, 1\}, \quad X_0^0 = 0, \quad X_1^0 = g(1)\sqrt{h(1)} \cdot \Xi_{0,0}.$$

By induction, for $1 \leq n \leq N$, we define:

- (a) The set of indices

$$\mathcal{K}_n = \{0 \leq k < 2^{n-1} \mid B_{n,k}(X_{l_{n,k}}^{n-1}, X_{r_{n,k}}^{n-1}) > \epsilon\},$$

indicating the segments in which a crossing is to be investigated. The supports $S_{n,k}$, with k in \mathcal{K}_n , is thus inductively defined as the set of supports with dyadic endpoints in D_{n-1} , for which the corresponding upper bound $B_{n,k}$ is not small enough for the probability of a crossing to be neglected.

- (b) Corresponding to the supports $S_{n,k}$, for k in \mathcal{K} , we define the set of dyadic times

$$\mathcal{L}_n = \{l_{n,k} \mid k \in \mathcal{K}_n\}, \quad \mathcal{M}_n = \{m_{n,k} \mid k \in \mathcal{K}_n\}, \quad \mathcal{R}_n = \{r_{n,k} \mid k \in \mathcal{K}_n\},$$

which consists of the left points, midpoints and right points.

(c) The values of X_t^n , for dyadic times in $\mathcal{L}_n \cup \mathcal{M}_n \cup \mathcal{R}_n$ by

$$\begin{aligned} \forall t \in \mathcal{L}_n \cup \mathcal{R}_n, \quad X_t^n &= X_t^{n-1}, \\ \forall t \in \mathcal{M}_n, \quad X_t^n &= \sigma(t) \cdot \Xi_{n,2^N t} + \mu(t). \end{aligned}$$

from the values of $\{X_t^{n-1}\}_{t \in D_{n-1}}$ and the random drawings of $\Xi_{n,k}$.

2. For each $1 \leq n \leq N$, the algorithm disregard any time following the occurrence of a value above the barrier L , that is, defining

$$t^n = \inf \{t \in \mathcal{L}_n \cup \mathcal{M}_n \cup \mathcal{R}_n \mid X_t^n \geq L(t)\},$$

the algorithm investigates for a first-passage time in dyadic segments $S_{n,k}$ delimited by endpoints in the set

$$\mathcal{I}_n = \{t \in \mathcal{L}_n \cup \mathcal{M}_n \cup \mathcal{R}_n \mid t \leq t^n\}.$$

In the previous definition of t^n , we observe the convention that $\inf \emptyset = 1$ to account for the possibility of a first-passage, even if all the simulated value of $\{X_t^n\}_{t \in D_n}$, are below the barrier.

3. Finally, we define formally the approximate first-passage time τ^N as

$$\tau_N = \inf \{t \in \mathcal{I}_N \mid Q_{N+1,2^N t}(X_t^N, X_{t+2^N t}^N) \geq v_{2^N t}\} + 2^{-N-1}.$$

The time τ_N is then the midpoint of the first support $S_{N,k}$, k in \mathcal{I}_N on which a simulated sample path of X interpolating $\{X_t^N\}_{t \in \mathcal{I}_N}$ would cross $\mu^N(t)$, the piecewise continuous approximation of L interpolating its value on the dyadic points D_N .

Examples

We illustrate the use of the algorithm in the simple case of an Ornstein-Uhlenbeck process U with an elastic coefficient $\alpha = -1$ and a diffusion coefficient $\sigma = 1$.

Considering the case of a constant barrier at $\Lambda = 1$, we represent in Figure 4.4 the simulation of a sample path up to a first-passage time. Remark that the sample path is simulated with increasing precision close to the barrier. The zooming operations in time regions where the path is about to cross the threshold underscore this fact. Figure 4.4 exemplifies a rather unfavorable situation: the first-passage occurs relatively late in time, about 4 times later than its expected value, and the sample path wanders three times in the close vicinity of the barrier. Notice that, despite this, the algorithm only needs computing 683 sample points, when the simulation of a sample path at full resolution would have required to compute more than $8 \cdot 10^6$ sample points.

In the particular case of a constant barrier set to $\Lambda = 0$, the probability density of the first-passage time is known analytically. If we assume the initial condition $U_0 = -1$, we actually have [6]

$$P({}_U\tau_{-1}^0 \in dt) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sinh(t)} \right)^{\frac{3}{2}} \exp \left(-\frac{e^{-t}}{2\sinh t} + \frac{t}{2} \right) dt.$$

In Figure 4.5, we compare the inferred distribution function obtained from binning the first-passages of our algorithm with this true probability density function. As apparent in logarithmic coordinates, the agreement is excellent.

4.2.2 Analysis of the Algorithm

In the present part, we analyze the behavior of the algorithm. First, due to its analytical tractability, we review the properties of the probabilistic screening in the seminal case of a Wiener process with a constant threshold. Second, we study in

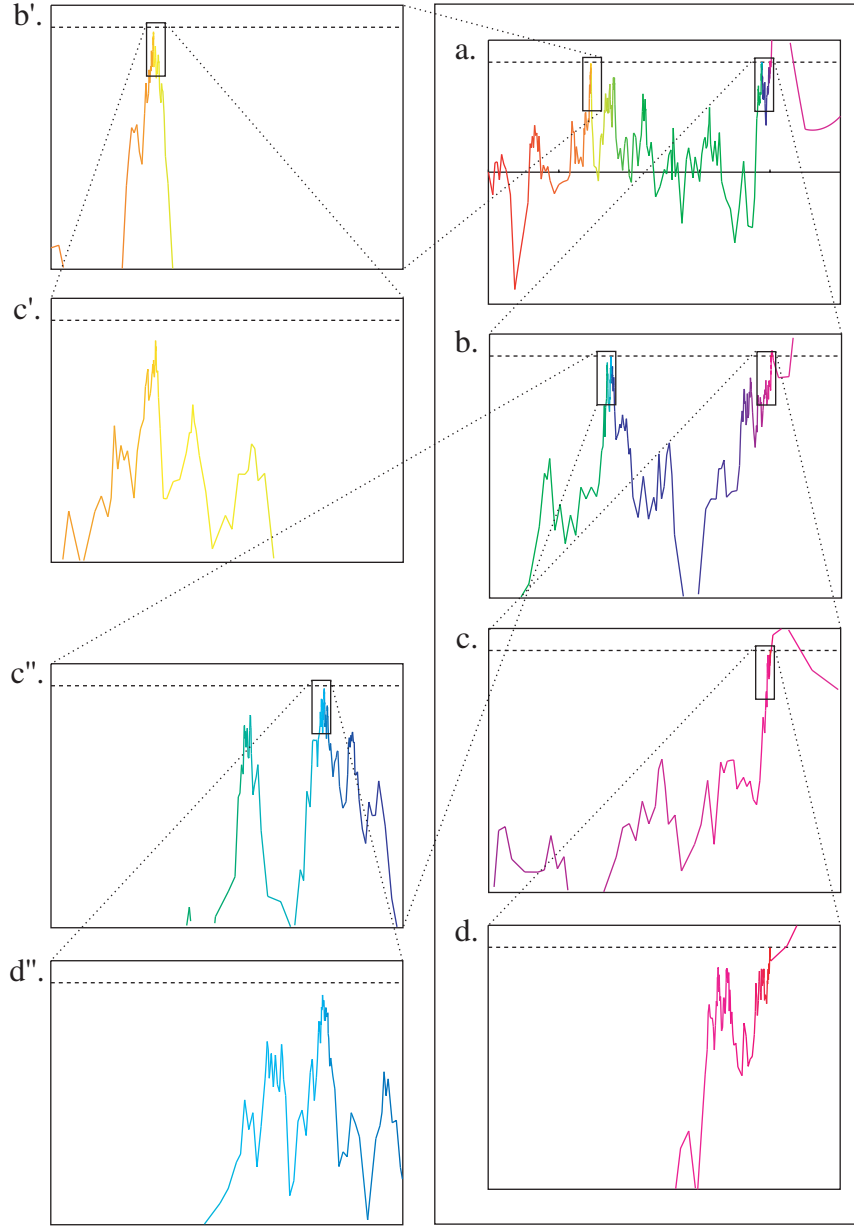


Figure 4.4: We consider the first-passage problem for an Ornstein-Uhlenbeck process U with an elastic coefficient $\alpha = -1$ and a diffusion coefficient $\sigma = 1$, the barrier is constant $\Lambda = 1$ and the initial condition is $U_0 = 0$. We represent a realization $U_t(\omega)$ for which the algorithm returns a first-time passage $\tau(\omega) = 8.00469684$ with a resolution $\delta t = 2^{-21} = 5 \cdot 10^{-7}$. The whole sample path is represented in **a.** and a series of zooms is carried out around $\tau(\omega)$ in **b.,c.,d.**. The series **b'.c'.** and **c''.d''.** zooms on a region where the sample path gets close to the barrier. The simulation of the sample path has required 683 subdivisions and illustrates an unfavorable situation since the expected number of divisions is approximatively 284 for this particular setting.

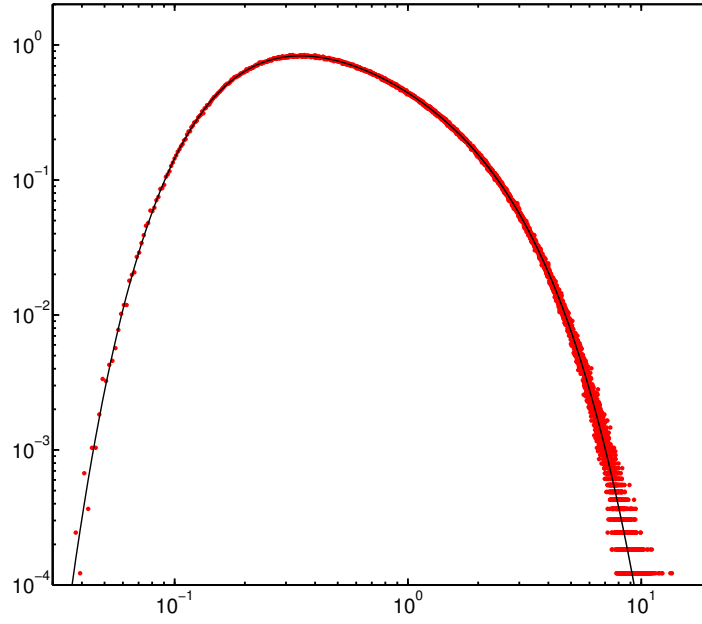


Figure 4.5: We consider the first-passage problem for an Ornstein-Uhlenbeck process U with an elastic coefficient $\alpha = -1$ and a diffusion coefficient $\sigma = 1$, the barrier is constant $\Lambda = 0$ and the initial condition is $U_0 = -1$. The black thin line represents the exact distribution of first-passage times, which is known analytically in the specific case of a barrier set to zero. The red dots are obtained by populating an histogram of 10^3 bins with 10^6 first-passage times simulations.

detail the computational efficiency and accuracy of our algorithm in the general case set by Assumption 1 and 2

We show that the essential features of the probabilistic screening similarly hold in the general case of a Gauss-Markov process X with a continuous barrier L satisfying Assumption 1 and 2. In particular, we verify the fast decay of the upper bounds $B_{n,k}$ for finer scale. Such a property results in two forms of advantage for computational accuracy and efficiency of our algorithm:

Computational Efficiency: while exploring sample paths at finer scale, the condition $B_{n,k} \leq \epsilon$ is satisfied after a few steps of the recursive exploration, saving us the computational cost of simulating the path at finer resolution.

Computational Accuracy: The limit probability ϵ of $B_{n,k}$ for which we neglect to continue the search for a passage, can be set all the smaller as $B_{n,k}$ vanishes fast for increasing n , thus lowering the overall probability to overlook a first-passage.

The present section is organized as follows. First, we establish an upper bound to the probability of returning an erroneous approximate first-passage time. Second, we justify the fast decay of the upper bounds $B_{n,k}$ used in the algorithm and we give a criterion to measure the algorithm efficacy. Finally, we explain the strategy to set the values of the algorithm parameters.

Algorithmic Accuracy

In the erroneous case, the algorithm always returns a crossing time that is not the true first-passage time. This happens when, while exploring sample paths, the algorithm dismisses regions where a first-passage actually occurs against the odds. As a result, the algorithm delays the first-passage time.

It is possible to naively estimate an upper bound to the probability of occurrence of such errors when the algorithm returns a putative first-passage time.

Property 13. Given a parameter value $\epsilon > 0$ and a time resolution of 2^{-N} , the probability of error

$$\mathcal{E}(N, \epsilon) = \mathbb{P}(\tau^N > \tau + 2^{-N-1}),$$

i.e the probability that a simulated first-passage in $[0, 1]$ does not approximate a true first-passage, satisfies

$$\mathcal{E}(N, \epsilon) \leq \epsilon 2^{N-1}.$$

Proof. False timings possibly occur when the algorithm halts the search for a first-passage in a support $S_{n,k}$ because the simulated values $x_{n,k}$ and $z_{n,k}$ are such that $B_{n,k}(x_{n,k}, z_{n,k}) \leq \epsilon$. For every such halting situations, the algorithm possibly neglects the occurrence of a crossing, the probability of such an error is then

$$\begin{aligned} \mathbb{P}(\exists t \in S_{N,k}, X_t \geq L(t) \mid B_{n,k} \leq \epsilon) = \\ \mathbb{E} [P_{n,k}(X_{l_{n,k}}, X_{r_{n,k}}) \mid B_{n,k}(X_{l_{n,k}}, X_{r_{n,k}}) \leq \epsilon] . \end{aligned}$$

By definition of $B_{n,k}$ as an upper bound to the probability of a crossing $P_{n,k}$ within $S_{n,k}$, the probability that the algorithm disregards a crossing is clearly dominated above by ϵ , i.e.

$$\mathbb{P}(\exists t \in S_{N,k}, X_t \geq L(t) \mid B_{n,k} \leq \epsilon) \leq \epsilon.$$

If we set the limit depth to be N , there are at most 2^{N-1} halts during each call of the recursive search procedure. Indeed the algorithm explores the binary tree of supports $S_{n,k}$ for $n \geq 1$, such that a given segment $S_{n,k}$ admits children supports if $B_{n,k} \leq \epsilon$. Such a tree has a depth of at most N levels and a support $S_{n,k}$ is a leaf if $n = N + 1$

or $B_{n,k} < \epsilon$. There are at most 2^{N-1} leaves for such a binary tree structure.

This corresponds to the worst-case scenario for which the algorithm explores the sample path up to depth N on the whole segments and find $B_{N,k} \leq \epsilon$ for every k .

Then we have

$$\begin{aligned}
\mathcal{E}(N, \epsilon) &= \mathbb{P}(\exists k, 0 \leq k < 2^{N-1} \mid \exists t \in S_{N,k}, X_t \geq L(t), B_{N,k} \leq \epsilon) \\
&\leq \sum_{k=0}^{2^{N-1}-1} \mathbb{P}(\exists t \in S_{N,k}, X_t \geq L(t), B_{N,k} \leq \epsilon) \\
&\leq 2^{N-1} \sup_{0 \leq k < 2^{N-1}} \mathbb{P}(\exists t \in S_{N,k}, X_t \geq L(t), B_{N,k} \leq \epsilon) \\
&\leq \epsilon 2^{N-1}
\end{aligned}$$

In other words, if we look for first-passage times with a resolution of 2^{-N} , the probability of an error $\mathcal{E}(N, \epsilon)$ per recursive call is inferior to $\epsilon 2^{N-1}$. \square

Algorithmic Efficiency

Under Assumption 1 and 2, we ideally wish to establish an upper bound to the complexity of the algorithm. Since the complexity increases linearly with the number of dichotomy operations, we see that the worst case scenario consists of the situation when every sample point needs to be simulated, which yields the complexity of the classic Runge-Kutta method. Unfortunately, there appears to be no direct way to establish analytically the complexity of the algorithm. Nevertheless, because of its obvious dichotomic structure, our algorithm substantially outperforms Euler and Runge-Kutta.

We then prefer to measure the computational efficiency of the algorithm indirectly. The exploration of a branch of the tree of supports $S_{n,k}$ stops as soon as $B_{n,k} \leq \epsilon$. The value of $n \leq N$ at such a halt defines the local depth of exploration for times in the support $S_{n,k}$.

Definition 7. For all s in $[0, 1]$, the local depth of exploration d_s is the discrete

random variable

$$d_s = \sup\{n \leq N \mid s \in S_{n,k}, B_{n,k}(X_{l_{n,k}}, X_{r_{n,k}}) > \epsilon\}.$$

The smaller the typical depth of exploration, the fewer sample points are simulated in regions where a sample path occurs strictly below the barrier, which is desirable for computational efficiency.

We thus want to estimate the local depth of exploration d_s around s to measure the algorithmic efficiency. This is determined by the speed at which $B_{n,k}$ concentrates its distribution on vanishing values for finer resolution. In that respect, bearing in mind the case of a standard Wiener process and a constant threshold, we now show that the doubly exponential speed of decay still holds in the general case.

To be more precise, assume that for some s , the sample path occurs strictly below the continuous barrier $X_s = y < L(s)$, so that we can find a neighboring of s for which the sample points are strictly below L . There exists a unique sequence of indices k_n such that $0 \leq k_n \leq 2^{n-1}$ and s satisfy $k_n 2^{-n+1} \leq s < (k_n + 1) 2^{-n+1}$. It defines a sequence of nested supports S_{n,k_n} such that we have $\cap_n S_{n,k_n} = \{s\}$ and by continuity of L , we have $\lim_{n \rightarrow \infty} \underline{L}_{n,k_n} = L(s)$. In view of this, if the increasing sequence \underline{L}_{n,k_n} is strictly above y and if the diffusion coefficient function σ is continuous, it is possible to show [217] that:

Property 14. If L and σ are respectively homogeneously H_L - and H_σ -Hölder continuous, then conditionally to $X_s = y < L(s)$, for every ω in Ω , we have the following asymptotic equivalent when n tends to infinity: if $\min(H_L, H_\sigma) < 1/2$, we have

$$\ln B_{n,k_n}(\omega) = -2^{n+1} \frac{(L(s) - y)^2}{\sigma^2(s)} + O(2^{n(1-\min(H_L, H_\sigma))}) ,$$

and otherwise for every $0 < H < 1/2$, we have

$$\ln B_{n,k_n}(\omega) = -2^{n+1} \frac{(L(s) - y)^2}{\sigma^2(s)} + O(2^{n(1-H)}) .$$

Moreover, if P_{n,k_n} denotes the conditional probability of a crossing, we have $\ln B_{n,k_n} \sim \ln P_{n,k_n}$ when n tends to infinity.

We stress that we only account for the certainty to reach an asymptotic regime without information about the time when this regime is reached. The speed at which such a regime is attained is set by the constant bounds implicitly present in the Landau notations of Proposition 14. Introducing the (ϵ, H) -modulus $C_f(\epsilon, H)$ of a H -Hölder function f as

$$C_f(\epsilon, H) = \sup_{|t-s| < \epsilon} \frac{|f(t) - f(s)|}{|t - s|^H} ,$$

It is possible to show [217] that the previous constants are directly set by the value of the moduli $C_L(2^{-N}, H_L)$, $C_\sigma(2^{-N}, H_\sigma)$ and $C_{X(\omega)}(2^{-N}, H')$, $H' > 1/2$, for every ω in Ω . Large values of these moduli delay the onset of the decaying behavior. The modulus $C_X(2^{-N}, H')$, $H' > 1/2$ plays a specific part here since it clearly appears as random variable. However, the Lévy's modulus of continuity theorem ensures that $\lim_{N \rightarrow \infty} C_X(2^{-N}, H')$ is bounded by one.

Keeping in mind the previous limitations, the asymptotic behavior in Property 14 is similar to the case of the Wiener process and the upper bound approximation becomes exact at vanishing scale. It suggests to approximate the local depth of exploration d_s by the following quantity:

Definition 8. For all s in $[0, 1]$, we define

$$N(\epsilon, y) = \log_2(-\ln \epsilon) + \log_2 \left(\frac{\sigma^2(s)}{(L(s) - y)^2} \right) .$$

as an approximation of the local depth of exploration d_s .

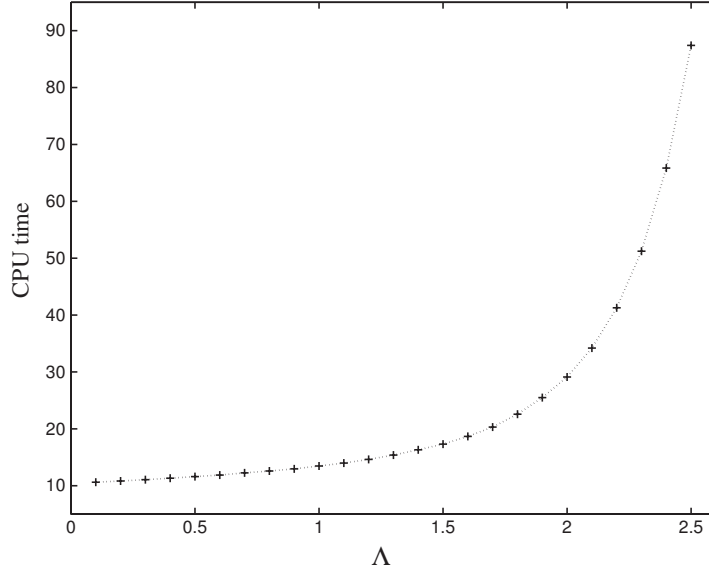


Figure 4.6: Computational cost of our algorithm for the first-passages of an Ornstein-Uhlenbeck process U (with elastic coefficient $\alpha = -1$ and diffusion coefficient $\sigma = 1$) through a constant boundary $\Lambda = 0.1, 0.2 \dots 2.5$ and initial condition $U_0 = 0$. The termination of the recursion is set by the limit resolution $\delta t = 2^{-21} = 5 \cdot 10^{-7}$, and by the tolerance for a false first-passage $\mathcal{E} \leq 10^{-10}$. The computational cost blows up when the threshold exceeds the persistence length of the Ornstein-Uhlenbeck process $2/|\alpha| = 2$. CPU time is quoted in microseconds per first-passage computed for a single core in a 2.66GHz Xeon processor.

This quantity results from the sum of two contributions: one from a statistical term expressing the stringency of the statistical screening and a geometrical term stemming from the interplay of the distance to the barrier and the diffusion coefficient. Notice that the dependence on ϵ through an iterated logarithm indicates that the statistics of d_s should vary very weakly with the parameter ϵ which can be set very small. A straightforward criterion to measure the algorithm computational advantage, is then to compare the typical value $N(\epsilon) = N(\epsilon, 0)$ with the total number of recursions allowed by the resolution. We will assess such a criterion in the following section, after describing how to choose the parameters values. Beforehand, we

illustrate the computational cost of our algorithm numerically in Figure 4.6.

Choice of the Parameters

Here, we explain the strategy to set the value of the various parameters intervening in the implementation in a segment $[0, T]$, where T is a positive integer. Focusing on reliability, we request that the probability of an erroneous result \mathcal{E}_T is inferior to a fixed parameter $\eta > 0$. Since we simulate first-passage times to describe a full-statistics, we require η to be very small, typically of order 10^{-10} . It is straightforward to see that \mathcal{E}_T as a function of the parameter ϵ and the limit depth N , is inferior to $T\mathcal{E}(N, \epsilon)$. It is then clearly enough to chose ϵ satisfying

$$\ln \epsilon \leq \ln \eta - (N - 1) \ln 2 - \ln T ,$$

and we consequently set $\epsilon(N, \eta, T) = \ln \eta - (N - 1) \ln 2 - \ln T$. With such $\epsilon(N, \eta, T)$, the algorithm necessarily returns an erroneous first-passage with probability less than $\mathcal{E}_T \leq \eta$. Concretely, if $\eta = 10^{-10}$, let us set the limit depth to $N = 21$ so that the accuracy is $2^{-21} \simeq 5 \times 10^{-7}$. Taking a time window of total length $T = 100$, we need to chose $-\ln \epsilon(N, \eta)$ close to 40 to satisfy $\mathcal{E}_T \leq \eta$. Thus setting the parameters ensures that each returned time is a valid first-passage with probability superior to $1 - \eta$.

The question is then to inquire wether such values of ϵ actually saves us a considerable computational cost, using our previously stated criterion. The answer to that question depends on the particular nature of the first-passage problem, but a prototypical answer can be given in the situation of constant α and σ for an Ornstein-Uhlenbeck process. Under the constraint that the limit resolution satisfies $2^{-N} \ll 1/|\alpha|$, if the local height of the barrier and the diffusion coefficient are of same order, the main contribution to the local depth of exploration d_s in (8) is determined by the screening

term $\log_2(-\ln \epsilon(N, \eta))$. As an example, if $\eta = 10^{-10}$ and $\alpha = -1$ for a resolution of 2^{-21} , the typical depth of exploration is $d = \log_2(-\ln \epsilon(N, \eta)) \simeq 6$. This has to be compared with the depth of the recursion necessary to simulate exhaustively the sample path up to the limit scale, i.e $N = 21$. Clearly, the typical recursive exploration of a time region halts notably before the limit depth and, because of the iterated logarithm, this behavior is very slowly varying with the parameters ϵ or η .

We illustrate this approach for the first-passage time problem of the Ornstein-Uhlenbeck U process with $\alpha = -1$ and $\sigma = 1$ for a constant boundary of value $\Lambda > 0$ and an initial condition $U_0 = 0$. The computational efficiency of the algorithm is essentially determined by the number of calls to the Gaussian random generator per simulated first-passage. For the Euler method, the number of Gaussian draws is just the number of time steps necessary to reach a first-passage and is therefore linear with the typical first-passage time. Equivalently stated, this number is linearly increasing with the inverse of the limit time step $1/\delta t = 2^N$ as opposed to our method which is strongly sub-linear with the inverse of the time resolution 2^N . Given the parameters of the simulation, Figure 4.7 demonstrates that the performance of the algorithm is four to five orders of magnitude faster than an Euler implementation with a comparable temporal resolution, while keeping a strong guarantee that the passage determined is indeed the first one.

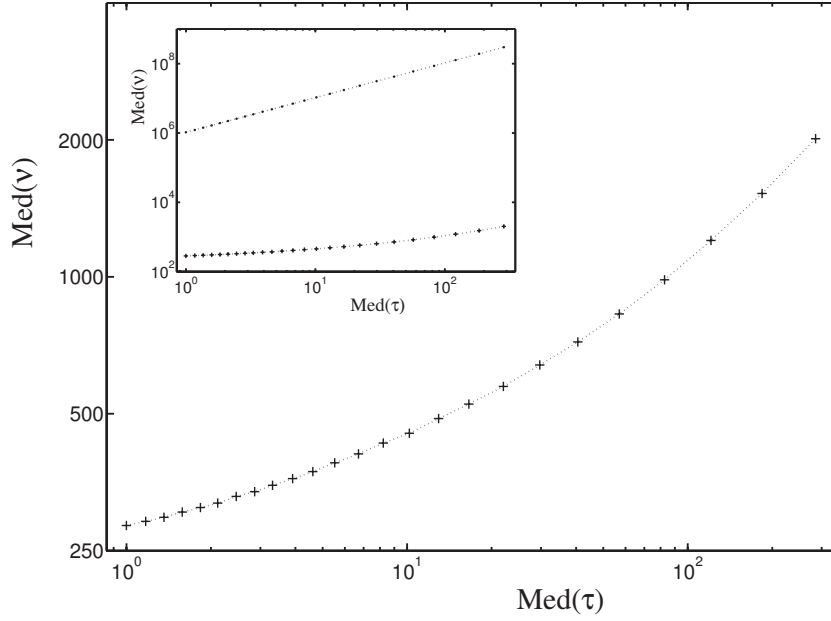


Figure 4.7: Parametric plot of the median number of Gaussian draws and the median of the first-passage time for the same conditions as in Figure 4.6 with $\Lambda = 0.1, 0.2 \dots 2.5$. Notice the logarithmic scale of both axes. As the boundary is moved away from zero, the mean first-passage time increases rapidly. **Inset:** We represent the number of calls to the gaussian random generators for our method (+) and the classical Euler method (•) on the same log-log plot. In an Euler scheme, the number of random numbers drawn is equal to the first passage time divided by the time resolution δt . Given our time accuracy, the number of draws is many orders of magnitude smaller with our algorithm. Moreover, it is strongly sublinear in the length elapsed when increasing the time accuracy.

4.3 Simulation of a First-Passage Markov Chain

In Section 4.1, we developed the simulation framework of a cyclically driven leaky integrate-and-fire neuron as an inhomogeneous Markov chain, while in Section 4.2 we described a computational method to simulate the first-passage kernel of the same Markov chain. Equipped with these results, in this section we intend to investigate the effect of the Hölder regularity of the effective barrier on spike coding. As usual, the leaky integrate-and-fire model is characterized by the following parameters: the leak constant α (with persistent time $|\alpha^{-1}|$), the diffusion coefficient σ and the threshold value V (the reset is set to $v = 0$). Here, we set these parameters to their most generally accepted value [76], namely $\alpha = -10\text{ms}$, $\sigma = 5\text{mV}$ and $l = 50\text{mV}$. These values are deduced from the physiological measurements of neuronal membrane potentials and are supported by simple models of neuronal noise, where the noise is modeled in the diffusive limit as the result of the background bombardment from weakly connected surrounding neurons [34].

We simulate the stationary measures μ_H corresponding to a certain family of H -continuous barriers L^H represented in Figure 4.1.3 and corresponding to the cyclical injection of a frozen current dC_H over a period $T = 1\text{s}$. The current includes a steady component that drives the barrier close to the zero mean value of the Ornstein-Uhlenbeck process simulating a neuron's voltage. In addition, we superimpose a frozen noise to form a current dC_H (taken with the same amplitude as the intrinsic noise for $H = 1/2$) that, once integrated through the equation of the leaky integrate-and-fire model, gives rise to an H -continuous effective barrier (see Section 4.1.3). Although our first-passage algorithm is rigorously valid for strictly non-negative barriers, it remains very accurate when this constraint is relaxed, if the limit resolution is well below the persistence scale. Since this is always the case in our simulation, we can impose a steady component that causes the barrier to cross the zero. However, in the results we present here, we only deal with positive barriers.

Finally, we constrain the family of noise currents dC_H to be such that the average firing rate of the neuron is held constant at 5Hz. Thus, the stationary measures μ_H are sampled temporally with the same statistical power. As previously stated, this constraint is satisfied by modulating the amplitude of the injected noise with a continuous function $H \mapsto c(H)$. This treatment increases the amplitude of the noise with larger H , i.e for smoother barriers, so that over the time course of one period T , the barriers L^H can be seen as sample paths of stationary processes, whose stationary measures have similar variance. Heuristically, this approach is only numerically correct as long as the mean spiking interval (here 0.2s) is larger than the typical fluctuations of the barriers between extreme values for all H , if we operate below the time-scale matching regime [204].

Now that we have accounted for the details of the simulation, we discuss the results in light of the the mathematical study of the first-passage time of Chapter 3.

4.3.1 Transition between two Spiking Regimes

In Figure 4.8 and Figure 4.9, we depict the peri-stimulus time histograms of the numerical experiments corresponding to Hölder exponents H varying between 0.25 and 0.75 . We form the histograms by binning $N = 100,000,000$ spiking times that have been computed with an end resolution of $1\mu s$ for the values $H = 0.25, 0.26, \dots, 0.75$. The representation we adopt is non-customary in the sense that, for every H -continuous barrier, we color-code in logarithmic scale the number of spikes occurring in a time bin. On both figures, the top bars of the graphs comprise side-to-side the histograms computed for different H during the whole cycle of repetition with a resolution of 1ms. The middle and bottom bars of the figures result from two successive magnifications of the top bars with respective resolutions of 0.03ms and $1\mu s$. On Figure 4.8, we zoom into a region of relatively low spiking probability, which implies weak reliability, whereas on Figure 4.9, we zoom into a region of high spiking proba-

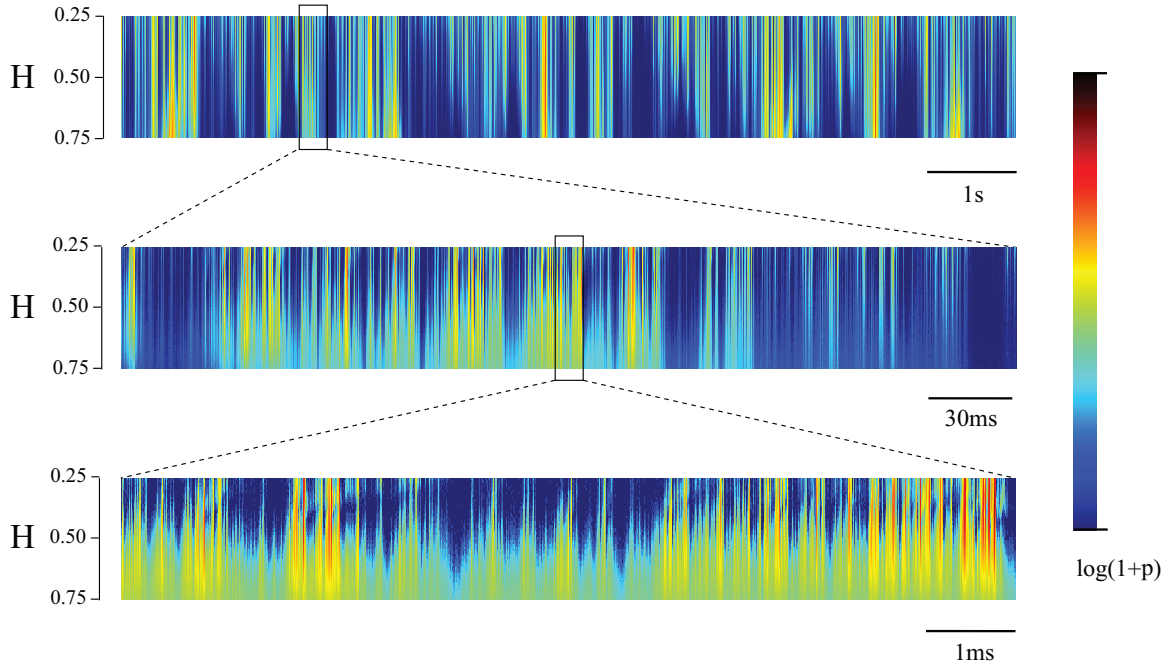


Figure 4.8: Histograms elicited by the family of H -continuous barrier L^H with progressive zooming onto a region of spiking of low reliability.

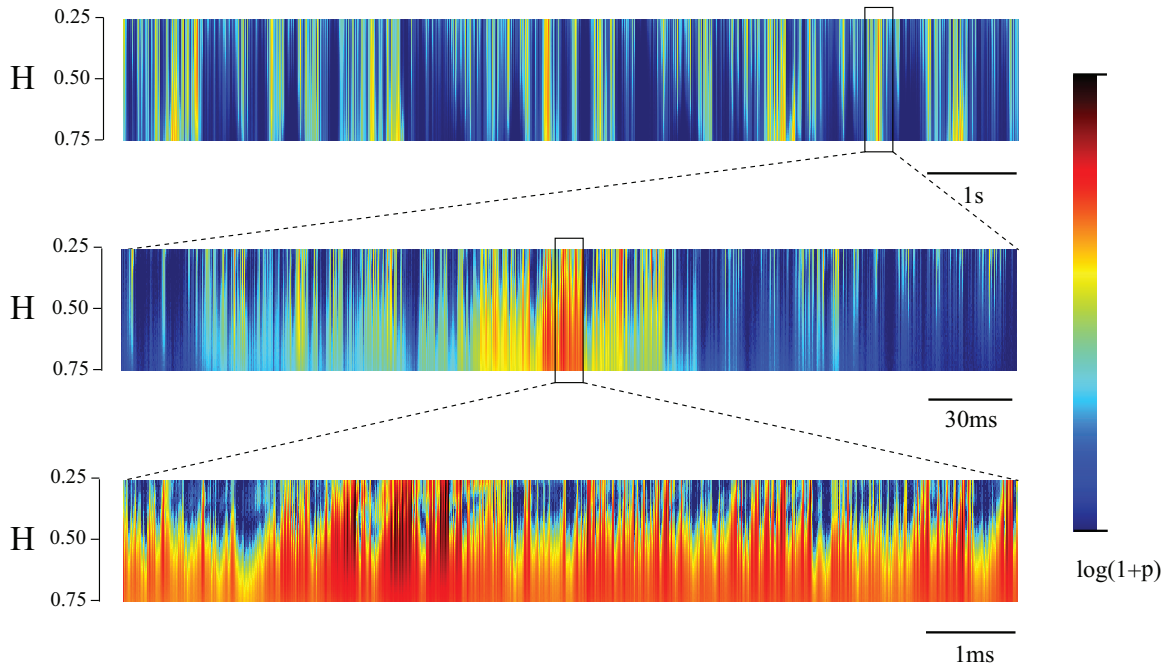


Figure 4.9: Histograms elicited by the family of H -continuous barrier L^H with progressive zooming into a region of spiking of high reliability.

bility, which implies strong reliability.

For a given H , the occupancy of the bins represents the amplitude of the stationary distributions μ_H , which is indicative of the instantaneous firing rate of the cyclically driven leaky integrate-and-fire neurons. Thus, aligning the empirical histograms alongside elucidates how the mean firing rate is altered through gradually varying H . At a low resolution, the overall profile of the histograms is only mildly affected by the change of Hölder continuity, which is related to the fact that the persistence time of the neuron $|\alpha^{-1}| = 10\text{ms}$ is close to the binning resolution 1ms . In first approximation, the spiking events happen similarly as if they were simulated independently in each bin i , with probability $p_i = \nu(-\infty, L^{H,i})$, where ν is the stationary distribution of the Ornstein-Uhlenbeck process and $L^{H,i}$ is the mean value of the barrier in the bin i .

However, when the histograms are magnified to resolutions well-under the persistence scale, we observe that there are two different spiking regimes. As expected for $H > 1/2$, representing the histograms at high resolutions reveals that they are almost everywhere non-zero, and that while zooming-in, their profiles become smoother. Contrarily, for $H < 1/2$, the histograms show widening silent gaps and an increasing number of isolated hot regions, signaling very high rates of instantaneous firing. As inferred from theoretical consideration, this strongly suggests that the distribution μ_H becomes singular when the barrier becomes rougher than the typical voltage sample path for $H = 1/2$.

This observation demonstrates the existence of a qualitative transition for neural firing at $H = 1/2$. Beyond this limit, when the neuron is driven by a frozen current that is more singular than Gaussian white noise, the temporal precision of its firing pattern becomes asymptotically infinite. To grasp how the time precision diverges, suppose we assess the normalized local firing rate $r_M(t)$ for a bin centered on t of size

T/M , $M > 0$ as

$$r_M(t) = \frac{n_M(t) \times M}{N \times T},$$

where $n_M(t)$ is the number of spikes falling in the bin, N is the total number of recorded spikes. In the vanishing scale limit, that is when M tends towards infinity while holding M/N constant, the mean firing rate tends towards zero for almost all time t , except for an ever-decreasing set of times where, in turn, the firing rate ever-increases toward infinity. For $H < 1/2$, at infinite resolution, the neuron would appear silent almost always, but for a negligible set of time at which the firing rate is infinite. On the contrary, we have establish that μ_H admits a continuous density for $H > 1/2$ and the estimate $r_M(t)$ is everywhere convergent to a finite value in the limit of vanishing scale.

It is tempting to speculate on the mathematical nature of such a transition phenomenon. More formally stated, the preceding discussion means that the support of μ_H has Lebesgue measure zero for $H < 1/2$ and is concentrated on singular spiking times. We know from mathematical considerations, that the cumulative distribution of μ_H is continuous, thus constraining the set of spiking times to be continuously singular in the sense that, as opposed to singularities of the Dirac delta type, they do not cause discontinuities once integrated. This situation is reminiscent of the best-known continuous singular measure, the Devil's staircase function [135]. However, the cumulative functions of μ_H are much more complex due to the distribution of its necessarily infinite number of singular spiking times. Indeed, the local minima of the rough boundaries are evident candidates for being spiking times, which suggests that the set of singular spiking times is dense in the time line $[0, T]^4$.

As a result, direct characterization of μ_H as a singular function is uneasy. For in-

⁴On the contrary, the closure of the singularity points of Devil's staircase function is of Lebesgue measure zero.

stance, direct use of the multi-fractal formalism provides ambiguous outcomes. We hypothesize that the distribution of μ_H is similar in nature to the distributions called stable subordinators, which are fractal measures that are known to be related to the first-passage time of certain stochastic processes [106].

4.3.2 Spiking Precision and Spiking Reliability

In order to better connect our simulations with the concepts of spike reliability and spike precision, we provide another representation of our results as raster plots. In Figure 4.10, we draw the raster plots made of 5,000 spikes for 6 different H -continuous barriers, on as many panels (**a.** through **f.**). In a given raster plot, we color each spike in keeping with the instantaneous firing rate at its time of occurrence, as recapitulated by the empiric spiking distribution. On top of each panel, we picture the barrier that has elicited the spiking activity and we color-code its profile according to the same principle as for the raster plot. It is interesting to observe that if the barriers all exhibit similar amplitudes of fluctuation, the smoothing of the barriers with larger values of H is accompanied by the formation of a few isolated downward peaks. Accordingly, the stronger the barriers' Hölder regularity, the higher the spike reliability at large scale, as shown by the well-individualized band of firing on the raster plots for $H > 1/2$.

In Figure 4.11 and 4.12, we represent raster plots in the exact same fashion, except that their panels depict the magnified regions of low spike-reliability, whose histograms are explored in Figure 4.8. In these two figures, notice that the barrier is scaled with the same aspect ratio on each panel according to the scaling of an Ornstein-Uhlenbeck process ($H = 1/2$). We deliberately choose this type of scaling to emphasize the “point of view” of a voltage trace exploring the vicinity of the barrier. Zooming-in on H -continuous barriers with $H > 1/2$, gradually smoothens their profiles, while the same operation for $H < 1/2$ yields seemingly rougher traces. In

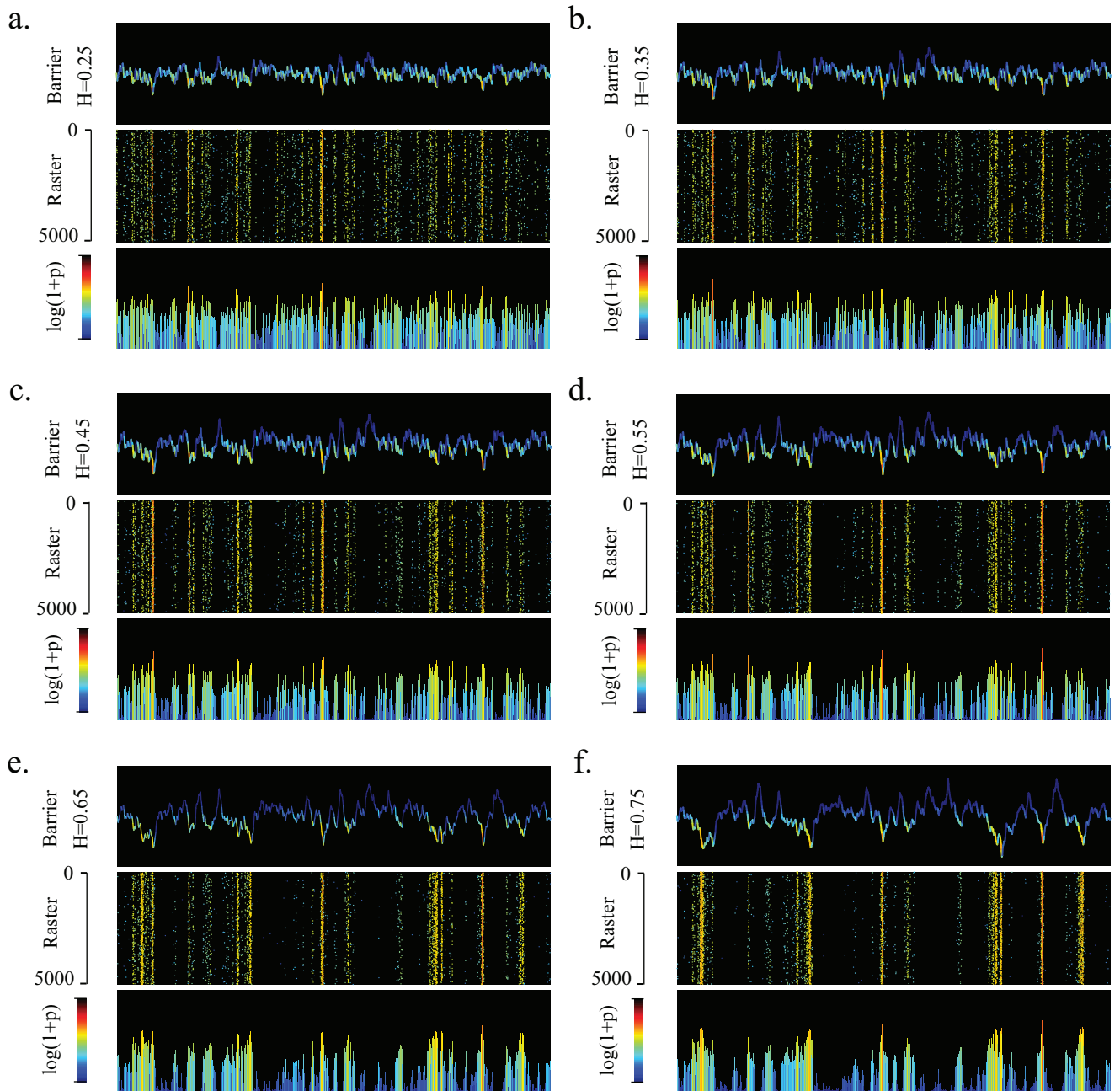


Figure 4.10: Raster plot of a cyclically driven leaky integrate-and-fire neuron over 1s for varying Hölder exponent H .

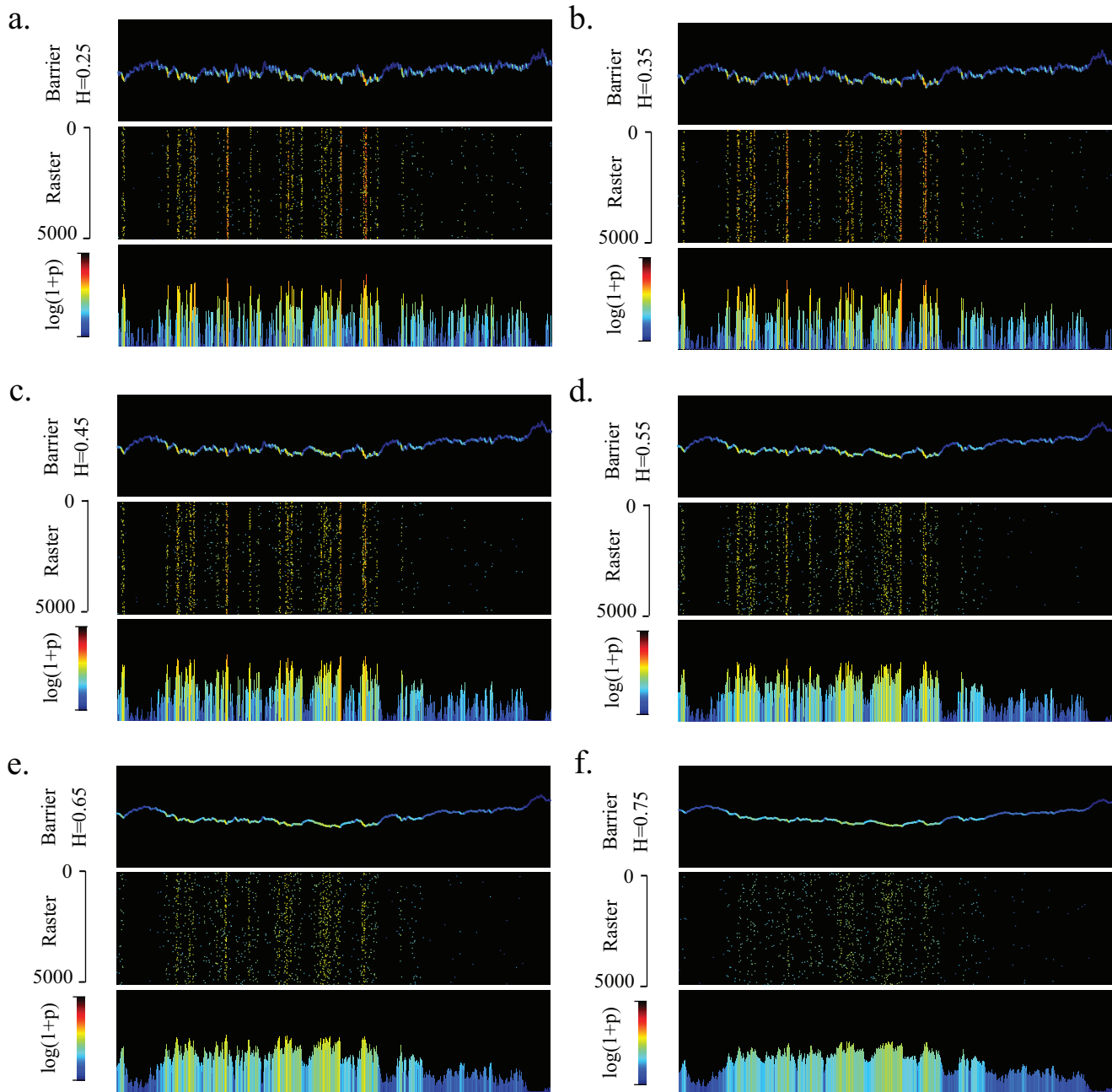


Figure 4.11: Raster plot of a cyclically driven leaky integrate-and-fire neuron over 32ms for varying Hölder exponent H

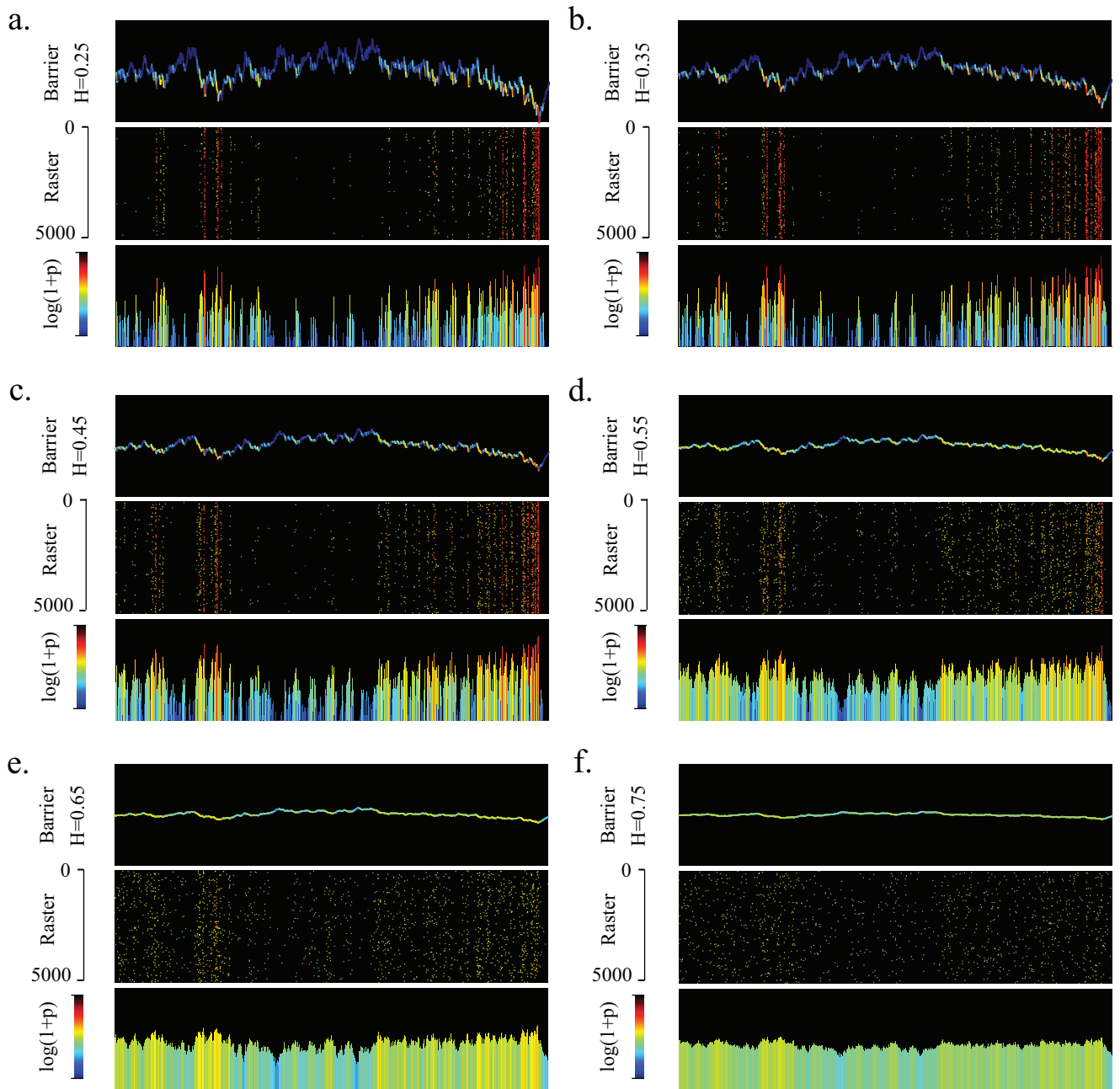


Figure 4.12: Raster plot of a cyclically driven leaky integrate-and-fire neuron over 1ms for varying Hölder exponent H

agreement with this relative smoothing or roughening of the barrier, it is observed that the spike precision, as indicated by perfect lining up of spiking events on the raster plot, changes abruptly when the barriers have the same roughness as the typical voltage trace ($H = 1/2$).

To summarize, under the constraint of constant average activity, varying the degree of singularity of the injected current causes a leaky integrate-and-fire neuron to transition between two regimes: a regime of high spiking reliability but poor temporal precision for weakly singular input, and a regime of poor reliability and exquisite temporal precision for highly singular input. From the mathematical point of view, this transition happens when the effective barrier becomes as rough as the underlying typical voltage trace.

We emphasize that these results differ from the already recognized fact that if the amplitude of the barrier's fluctuations is large enough to overcome the intrinsic noise, the frozen noisy input can drive the neuron to spike reliably with precision [74]. Actually, while being chosen to enforce a steady overall firing rate, the weight function $H \mapsto c(H)$ is such that the resulting barrier's fluctuations are of the same amplitude as the fluctuations caused by the neuron's intrinsic noise.

4.3.3 Path-wise Perspective

The transition between the two encoding regimes can be qualitatively explained by path-wise considerations. As a preamble, we underline the fact that the effective barrier determines a neuron's spiking pattern through the strength of its fluctuation, and also through its Hölder regularity. In actuality, if the strength of fluctuations is the dominant effect at large time-scale, the Hölder continuity is always the decisive property that dictates the distribution of spike timing. In first approximation, a voltage trace that is tasked with hitting a barrier reliably runs into the accessible downward excursions of the low resolution depiction of the barrier. Due to the leak

term, this process of traveling from one major obstacle to another is essentially memoryless. Yet, at small time scale, the continuous voltage trace bears the memory of the position it just visited. This memory effect plays a crucial part when the trace lies in the vicinity of the barrier

To see this, we must recall that the Hölder exponent can be likened to the exponent of the scaling operation that leaves the profile of a curve statistically invariant. This explains why zooming on H -continuous barriers with the scale invariance of the Ornstein-Uhlenbeck process ($H = 1/2$), causes them to smoothen for $H > 1/2$ or roughen for $H < 1/2$. As a result, if on its way to hit a sheltered part of the barrier, the voltage trace always runs into an easy-to-access downward fluctuation for $H < 1/2$. Indeed, such obstacles keep presenting themselves at every scale and become increasingly harder to avoid, since they scale with a lower exponent (higher roughness). Therefore, while on its way to hit a sheltered part of the barrier, the voltage trace *has* to die on the left-flank of one of these obstacle first. For barriers that are everywhere rougher than the typical voltage trace, the barrier's profile is almost surely entirely sheltered, causing the distribution of accessible time to be singular. For $H > 1/2$, the obstacles become easier to avoid while going down in scale. Therefore, as the voltage trace gets closer to hit the barrier, it increasingly ignores the obstacles and can reach any part of the curve indiscriminately, giving rise to a continuous density of spiking time.

Thus, understood as a path-wise phenomenon, this principle is readily generalizable to other types of integrate-and-fire neurons. For instance, if the voltage traces are sample paths of a n -times integrated Gauss-Markov process (see Section 2.4.2), the transition to singularity happens for barriers that are n -derivable with H -continuous n -derivative and $H < 1/2$. Similarly, if the noise is a fractional Gaussian noise with Hurst coefficient H , $0 < H < 1$ [136], the transition of the leaky integrate-and-fire neurons happens for barriers that are rougher than the H -continuous voltage paths.

Being very general in nature, we hypothesize that such a behavior is the signature of a stochastic integrate-and-fire encoding schema.

4.3.4 Neural Code

In the perspective of neural coding, the injected noisy current is often viewed as a signal that conveys information. If we consider a neuron as a black box that responds to the current injection, each neuron gives rise to trains of spikes, whose stochastic patterns presumably encode some features of the signal. Here, we do not consider the neuron as an information channel transforming probabilistically an ensemble of stimuli into a set of neural responses [46]. Instead, we identify the cyclically driven neuron as a source of signs (spiking times) parametrized by the current signal. For the leaky integrate-and-fire model, the neuronal source generates stochastic spikes through a Markov chain, whose statistics is a function of the injected current. If the signal is equated with the injected current over a time period, it is possible to endow the signal with a notion of computational complexity, such as the minimum number of instructions required to give an accurate description of it (e.g. the minimum description length principle [87]). In our numerical model, all currents parametrized by H exhibit the same complexity for being finite-dimensional Gaussian noise constructed as

$$dC_H(t) = c(H) \cdot \sum_{0 \leq n} \sum_{0 \leq k < 2^{n-1}} \phi_{n,k}(t) \cdot \xi_{n,k}^H, \quad \text{with} \quad \xi_{n,k}^H = 2^{n(H-1/2)} \xi_{n,k},$$

with $\xi_{n,k}$ independent realization of the standard law $\mathcal{N}(0,1)$. In this context, it is important to investigate how much information this source produces, ultimately giving us an idea of the range of encoding patterns that a signal can elicit. An even more crucial issue is assessing if the information produced by the source is actually about the signal [211]. This later point directly leads to the *inverse first-passage problem* aimed at inferring the shape of the barrier from the distribution of first-passage

time [4], that can be solved for smooth barriers in the bayesian framework [157].

Rate of Information

Considering the cyclically driven neuron as a source, the amount of information produced without necessarily being signal-specific is quantified by the entropy of the spike-generating process. When dealing with an ergodic inhomogeneous Markov chain, the quantity of information produced per spike is the rate of entropy production of the chain denoted ρ_H [140]. Intuitively, the rate of entropy is an average measure of the incertitude with which each spike is generated. Therefore, it is closely related to the spiking reliability [22]. If for a given initial reset time, the next spike is reliable, happening consistently in the same narrow time region, the incertitude about the spiking event is limited. Contrarily, if the reset time is poorly indicative of the position of the next spiking time, the incertitude per event is important, leading to a high rate of entropy.

In considering the rate of information, we benefit from the availability of a simple formula for the case of an ergodic inhomogeneous Markov chain:

$$\rho_H = \int_0^T \left(\int_0^T \kappa^H(t, s) \ln(\kappa^H(t, s)) ds \right) \mu_H(dt). \quad (4.10)$$

The previous relation, which uses the notation of Section 4.1, only holds if the the Markov kernels are absolutely continuous with density κ^H . Notice that expression (4.10) explicitly uses the stationary measure μ_H as the weighting function of the entropy of each conditional Markov transition.

Measuring information quantities is a vast a topic [155], with a special interest in developing non-parametric methods [233]. For our purposes, a number of sophisticated estimators have been developed to infer ρ_H , based for instance on clustering algorithms [127] and on complexity analysis [8]. Operating with a huge sample size,

we assess the continuous rate of entropy resorting to the naive discrete estimator. Expression (4.10) becomes ill-defined when the kernels are not absolutely continuous with respect to the Lebesgue measure. However, we can still naively estimate it after the reckoning that our simulation introduces a cut-off time-length, beyond which the distribution of spiking times is continuous.

We expect the rate of information to be low for values of $H > 1/2$, cases for which the distributions are smoothly peaked around regions of high reliability, so that most of the Markov transitions occur from one peak to another. On the contrary, when the attainable positions become singularly distributed, the rate of entropy estimator experiences a flatten distribution of spiking times, since the times are then densely sampled in a more homogeneous fashion.

We sketch the rate of information ρ_H computed for each H -continuous barrier in Figure 4.13. The data points lie on a smooth curve, thus suggesting a smooth transition to singularity as opposed to a finite-order phase transition phenomenon. Except for the portion of the trajectory corresponding to very rough boundaries, the curve presents an overall downward profile that is consistent with our expectations. At a fixed overall firing rate, low values of a Holder exponent enforces time precision at the cost of temporal reliability, which results in increasing the rate of information. As apparent on Figure 4.13, the speed of this change is maximal at the transition for $H = 1/2$, where the slope of the curve is extremal. Contrary to other studies [25], we stress the fact that the information rate of the leaky integrate-and-fire neuron is computed at constant firing rate.

Signal Specificity

The leaky integrate-and-fire neuron, seen as a Markov source, generates more information when driven by a singular input for which $H < 1/2$. In the context of

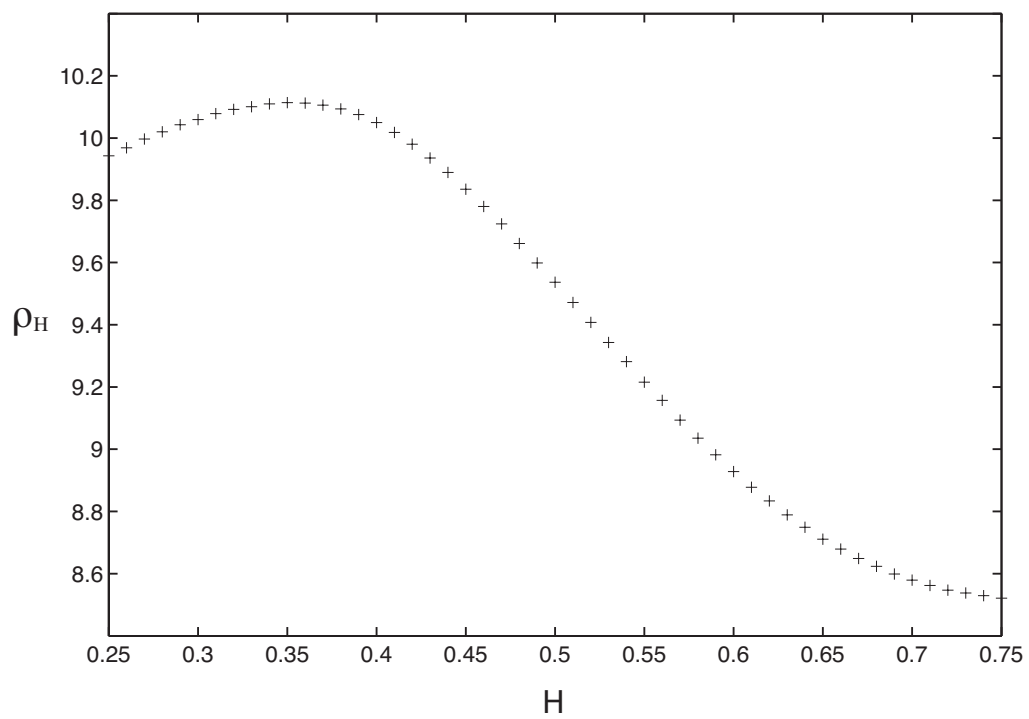


Figure 4.13: Dependence of the rate of entropy of the first-passage Markov chain on the Hölder exponent H .

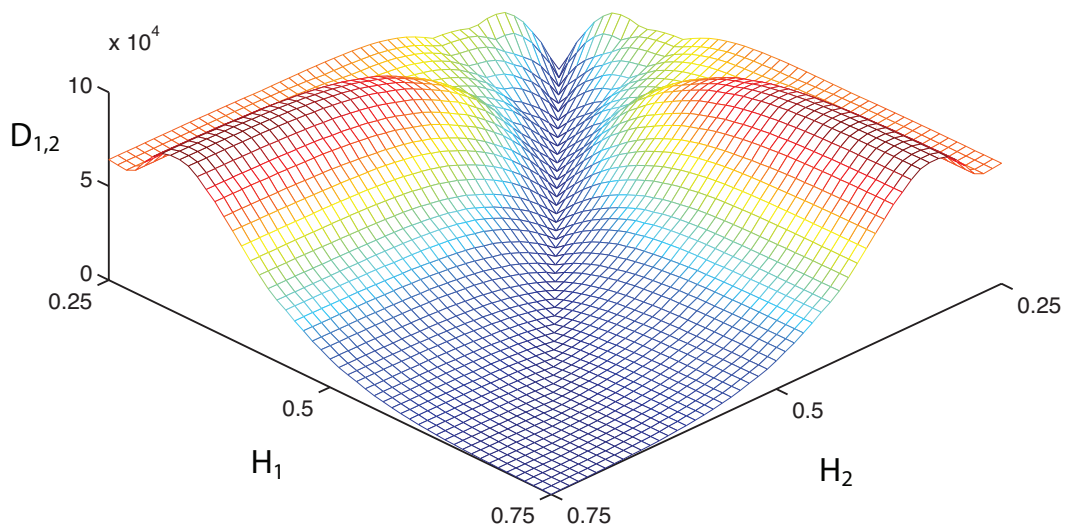


Figure 4.14: Supremum of the amplitude of the difference of the empiric distribution used in the Kolmogorov-Smirnov test.

information theory, this entropy measure does not say whether this flux of information encodes any specific features of the signal. This is especially apparent once one remembers that the entropy of a finite distribution is independent by permutation. To investigate information content that is relevant about a stimulus, information theoretical approaches, such as the *information bottleneck* method, investigate the mutual information between the response and some statistical stimulus feature [208].

However, the signals at stake here are not drawn from a distribution; they are frozen with a parametric dependance on H . This suggests the problem of information relevance can be tackled at a lower level, from the point of view of a discrimination task. We measure the signal-specificity of the neuronal firing rate by subjecting the stationary distributions μ_H to a series of two-sample Kolmogorov-Smirnov tests [115, 209]. Since all the empirical distributions μ_H are evaluated for the same sample size, we can directly compare the statistics of the test $D(H_1, H_2)$, defined as the maximum absolute difference between the two empirical cumulative histograms that are evaluated for barriers L^{H_1} and L^{H_2} . We represent these statistics in Figure , which shows that, based on the observation of the mean instantaneous firing rate, the power of discrimination between two barriers is much higher for barriers satisfying $(H_1, H_2) \in \mathcal{H}_1$.

This observation is not surprising in the sense that these distributions become singular in the infinite resolution limit, so that the statistics should actually diverge. Yet, as observable on Figure 4.3.4, the set of values of D lies on a seemingly smooth sheet in our numerical experiments. This is due to the fact that we are dealing with a huge sample size while assuming that the distributions smoothed beyond a cut-off time scale, which incidentally confirms that the transition to singularity occurs in a smooth fashion.

The multiple Kolmogorov-Smirnov tests unambiguously demonstrate that the leaky integrate-and-fire neurons have a superior discriminative power when being driven by highly singular currents, while achieving higher rates of information. This result

is more original than it appears; it is at odds with the vastly recognized existence of a trade-off between rate of information and signal-specificity. For instance, at early stages of auditory processing where temporal coding is at play, neurons that exhibit the highest input selectivity are also the ones that exhibit the lowest rate of information [64, 57].

4.4 First-Passage Neural Networks

In cortical cell assemblies, a neuron forms synapses onto about 10000 target neurons [27]. The resulting neural circuits are weakly driven by behaviorally relevant, highly preprocessed spiking patterns [69, 31, 123]. In turn, the neural processing performed by these circuits is crucially affected by noise, which results from background spontaneous activity or from inherently stochastic mechanisms of spike generation [185, 65, 66]. However, when studied *in vivo*, cortical neural circuits can exhibit precise and reliable neural response [3, 20, 170, 176, 192, 13]: upon repeated excitation by a stimulus, neural circuits consistently produces patterns of activity where spike timing are precisely time-locked to the features of the stimulus. This suggests that neural circuit process information through temporal coding [220], whereby the actual spiking times convey meaningful information [99, 249].

Understanding how indirectly-driven, noisy, interacting neurons can produce reliable and temporally precise neural response is central to the problem of neural transmission [116, 117]. In that respect, noise plays a crucial and ambiguous part: it alternatively appears as a nuisance constraining neural circuits to operate through rate coding [130] or as a benefit to precise temporal coding through stochastic facilitation [144]. Computational neuroscience addresses the problem of neural transmission by studying *in silico* models of neural networks, which model the internal dynamics of isolated single-units as well as their interactions, while explicitly including noise. When concerned with the temporal coding of a neural population, an adequate model of spiking neuron is the *integrate-and-fire* neuron [122, 113].

In this model, at any time t , the internal state of a neuron is given by its membrane potential X_t , which evolves according to a stochastic differential equation (or Langevin

equation) of the diffusion type [106, 216]

$$dX_t = F(X_t, I(t), t) dt + \sigma(X_t, I(t), t) dW_t, \quad (4.11)$$

where F is a drift function, σ the diffusion coefficient, I is the neuronal input and W is a Wiener process [76, 77, 102, 101]. In all generality, notice that F and σ are both dependent on the time-varying input I of the neuron. Whenever the potential X_t reaches a given threshold l , the neuron spikes and the membrane potential instantaneously resets to its base level. Integrate-and-fire models are advantageous because they articulate the discrete nature of spike counts and the continuous nature of spike timing [235], by prescribing the subthreshold dynamics of the membrane potential. It thus offers the possibility to readily include noise contribution in the model and to model temporally localized neuronal interactions as recapitulated by I , seen as the integrated synaptic input.

In the context of a population, an integrate-and-fire neuron initiates a spike in response to the temporal spiking patterns of upstream neurons, when the corresponding synaptic deliveries reliably drive the neuron’s membrane potential to its threshold. The transient nature of synaptic transmission and the existence of finite propagation delays conspire to render the integration of afferent activity extremely sensitive to the perturbations of incoming spikes’ timing [103, 85]. Thus, temporal precision is absolutely crucial to compute the activity of a neural population in response to both past and external spiking patterns.

Moreover, as the result of an intrinsically noisy dynamics, the spiking patterns of a population of stochastic integrate-and-fire neurons are realizations of a highly dimensional stochastic process. A *stochastic* simulation method is then temporally exact if its algorithm produces exact samples of the intricate law of the network’s evolution.

Practically, the implementation of exact stochastic algorithms yields spiking patterns with a temporal precision that is only limited by the exactness of the random generators and the numerical precision of the computer, thus ensuring exquisite temporal precision.

Despite substantial work [142, 177, 187], the many difficulties inherent to the study of interacting stochastic processes has hindered the development of temporally exact algorithms with noisy dynamics [29]. It is the purpose of this paper to remedy this situation for the simplest network of stochastic integrate-and-fire neurons.

4.4.1 Simulation of Integrate-and-Fire Networks

The simulation of networks of integrate-and-fire neurons invariably consists of the implementation of propagation periods between two consecutive events (spike emission or reception) and network update rules following each event (modifying neuronal internal variables if necessary). Such algorithms generally fall into two categories [29]:

Clock-driven Algorithm: The integrative schema corresponding to the network evolution is discretized with a finite time-step (typically using Runge-Kutta methods) and is paced by a central clock [150].

Event-driven Algorithm: The network evolves according to simple equations and, given the network state at a particular spiking event, there exist analytical formulae predicting the next spiking event and network state [142].

Due to their flexibility, clock-driven algorithms have received the most investigation. Nevertheless, when concerned with precise temporal dynamics, these algorithms present two major flaws. First, their computational complexity is set by the duration of the time step and most of the computational time is devoted to simulate the unobserved subthreshold dynamics of neurons that, in addition, can remain silent most of the time. Second, its discrete flavor introduces errors: noisy membrane trajectories

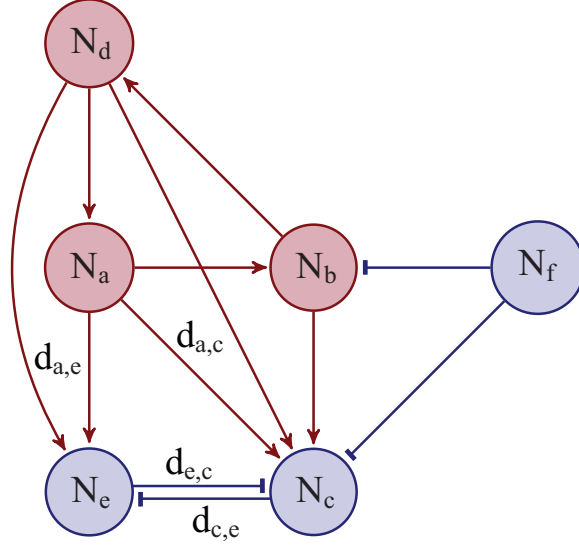


Figure 4.15: Recurrent neural network with time delays.

can very-well cross the spiking threshold in between two subthreshold sample-points [217].

Another key issue is the existence of finite time for spike propagation, which affects the core functioning of a neuronal network. For instance, reasoning on the recurrent network depicted in Figure 4.15, suppose we need to update the network after neuron N_a fires and that receiving one spike is enough to elicit a response. Then if the delays are such that $d_{a,c} + d_{c,e} < d_{a,e}$, neuron N_c is activated and N_e is inhibited, whereas the situation is reversed if $d_{a,e} + d_{e,c} < d_{a,c}$. Actually in the absence of delays, it is impossible to order the sequence of local updates consistently and the problem is ill-posed.

As opposed to stochastic clock-driven algorithms, stochastic event-driven algorithms devote most of their computational power to the implementation of neuronal interactions, naturally include finite propagation times in their event scheduling and are exact in principle. However, since the existence of analytical transition formulae is still required [177, 187], event-driven algorithms remain restrained to a few deter-

ministic network models [190, 223]. A natural candidate spiking model to develop an event-driven strategy suited to noisy dynamics, is the *linear integrate-and-fire* neuron [121, 179, 32, 34, 35] with

$$F(X_t^a, I(t), t) = -\alpha X_t^a + I_t^a, \quad \text{and} \quad \sigma(X_t^a, I(t), t) = \sigma_a,$$

and Dirac-like time-delayed interactions as a weighted-sum of Dirac functions

$$I^a(t) = \sum_b \sum_n w_{b,a} \cdot \delta(t - (t_b^n + d_{b,a})). \quad (4.12)$$

$w_{b,a}$ and $d_{b,a}$ denotes the weight and delay of the synapse from neuron b to neuron a and t_b^n refers to the timing of the n -th spiking event of neuron b ⁵.

4.4.2 Stochastic Event-driven Strategy

Recently, the Markovian structure of networks of linear integrate-and-fire neurons was identified and translated conceptually in terms of a stochastic event-driven strategy [226]. The special features of the linear integrate-and-fire model make it possible to design local update rules which stipulates how a neuron behaves when receiving or emitting spikes. Incidentally, the network evolution of such neurons appears as a Markovian chain of single-unit asynchronous updates. The time sequence with which these rules are called is determined as follows. For each instantaneous state of the network, the timing of the next event is estimated independently for every neuron, as if the neurons were not interacting. Selecting the first of these events is tantamount to drawing the next network event (since no other event happens before it) and the corresponding local update rule can be safely applied.

⁵The weights and delays can be taken as independent random variables without changing the analysis.

However, each time a local update rule is implemented, it propagates interactions to other neurons. Maintaining the Markovian property demands that updated neurons bear the memory of the fact that, in the absence of interactions, they would spike at a known time in the future. From a mathematical standpoint, the book-keeping of intermediary events means that the membrane potentials are sampled from random processes that are conditioned in the future. Stochastic event-driven algorithms implement consistently these conditional updates at the network level. They proceed by using information about the current state of the network to refine provisional estimates of future events, until each of these refined estimates sequentially becomes the next network event.

Such algorithms are efficient when they minimize the time instances at which the underlying membrane potential needs to be sampled. This sampling number is determined by how often a neuron’s membrane is sampled when receiving inputs and how many estimates are required for a provisional neuronal event to become the next network event. Figuratively, the task of a stochastic event-driven algorithm is to simulate future potential scenarii at the cost of later discarding parts of it, so that the individual neurons are only simulated for those times when their update is meaningful at the network level.

4.4.3 New Practical Implementation

Here, elaborating on results in [226], we rigorously investigate the Markovian nature of networks of *perfect stochastic integrate-and-fire* neurons with time-delayed Dirac-like interactions and produce an *efficient* and *exact* event-driven implementation of their dynamics. A linear Integrate-and-fire neuron is deemed “perfect” when there is no leak term ($\alpha = 0$) in the equation of its subthreshold membrane dynamics.

Such a model is general in the sense that the Markovian setting of its networks can be extended to any linear integrate-and-fire neuron. However, networks of perfect stochastic integrate-and-fire neurons are of special interest because they are simple enough to be practically implemented in an event-driven fashion.

Indeed, we show that the conditioned processes that intervene in the Markovian description of the network are well-known stochastic processes and are closely related to the canonical Wiener process [106]. From there, we implement *explicitly* our event-driven strategy as a Monte-Carlo algorithm, whose efficiency is ensured by the existence of exact efficient random generators for our conditioned processes. An exact random generator is an algorithmic procedure that, given exact samples from a simple probability law (typically the uniform law over $[0, 1]$), produces a number whose statistics follows exactly the law of the desired random variable. The efficiency of the generator is measured by the expected number of arithmetic operations that are required to simulate a sample. Moreover, we organize the scheduling of events so that the effective number of updates of a given neuron is *less* than the number of spikes it receives, while discarding little information in the refinement of provisional estimates. Incidentally, despite being a *fully stochastic exact* scheme, our algorithm operates with the same complexity as a *deterministic* event-driven method.

As a result, we argue that our event-driven implementation computationally outperforms any other stochastic simulation methods and is especially suited to the study of temporal coding. This is of particular interest for recurrent networks evolving under spike-timing-dependent plasticity [53, 1, 39]. From another perspective, networks of integrate-and-fire neurons have recently inspired bayesian inference methods to analyze experimental neural data [44, 149]. Our algorithm provides a natural candidate to benchmark these methods on synthetic data.

The manuscript is organized as follows. We first describe the Markovian evolution of a network of perfect stochastic integrate-and-fire neurons and explicit the corresponding local and network rules that allow us to design an event-driven strategy. We then devise its practical implementation and bound its complexity. Finally, we illustrate its behavior by successively considering the case of an inhibitory network, an excitatory network and a network with both inhibition and excitation.

4.5 Exact Markovian Framework

In this section, we first define the evolution of a single-unit embedded in a network of neurons as a first-passage Markov chain, where the interactions are recapitulated in terms of a fluctuating barrier. We then exhibit the update rules for an interacting single-unit and establish their schedule at the network level, in order to faithfully simulate the network's evolution.

4.5.1 Perfect Integrate-and-Fire Networks

The dynamics of the membrane potential X^a of a neuron a integrates a time-dependent input current I^a through the stochastic differential equation of a drifted Brownian motion

$$dX_t^a = \left(\nu_a + I_a(t) \right) dt + \sigma_a dW_t^a ,$$

where $\{\nu_a\}$ are constant positive drifts, $\{\sigma_a\}$ are diffusion coefficients and $\{W^a\}$ denote independent standard Wiener processes [34, 35]. When the membrane potential X^a reaches a threshold l^a , that is at the first-passage time $t^a = \inf\{t > 0 \mid X_t^a \geq l_a\}$ assuming $X_0^a < l^a$, a spiking event occurs and the membrane potential resets $X_{t^a-}^a = 0$. A spiking event entails the delivery of spikes to the set of downstream neurons $D(a)$

that are connected to a . The past spiking history $h^a(t)(\omega)$ of a neuron a up to time t is the finite time-ordered sequence $\{t_i^a(\omega)\}_{0 \leq i < C_a(t)}$ that satisfies

$$t_{i+1}^a(\omega) = \inf\{s \mid X_s^a(\omega) \geq l_a \text{ and } t_i^a < s < t\}$$

and the instantaneous spike count $C_a(t)$ is the number of events in $h^a(t)$.

Here, we posit that the currents I_a only recapitulate the interactions between connected neurons and consist of Dirac impulses, modeling time-delayed deliveries of spikes from the set of upstream neurons $U(a)$ that are connected to a . Given the spiking histories $\{h^b = h^b(\infty)\}$ of neurons b in $U(a)$, the neuron a receives the input current

$$I_a(t) = \sum_{b \in U(a)} \sum_{i_b \in \mathbb{N}} w_{b,a} \cdot \delta(t - (t_{i_b}^b + d_{b,a})),$$

where $w_{b,a}$ is the weight of the synapse from b to a , and $d_{b,a}$ is the propagation delay from neuron b to a . No neuron self-connects and we impose that the matrices of weights and delays satisfy $w_{a,a} = 0$ and $d_{a,a} = 0$.

The statistics of spiking events in the absence of interactions is entirely specified by the effective height of the threshold l_a/σ_a and the effective drift ν_a/σ_a through the law of the first-passage time of X^a at level l_a

$$\tau^a = \inf\{t > 0 \mid X_t^a \geq l_a\} = \inf\{t > 0 \mid W_t^a \geq L_a(t) = l_a - \nu_a t\}.$$

From now on, we posit that $\sigma_a = 1$.

In the presence of interactions, if a neuron b spikes at time t^b , the membrane potential X^a of a downstream neuron a in $D(b)$ is instantaneously updated at time $t^b + d_{b,a}$ according to

$$X_{(t^b+d_{b,a})+}^a = X_{(t^b+d_{b,a})-}^a + w_{b,a}.$$

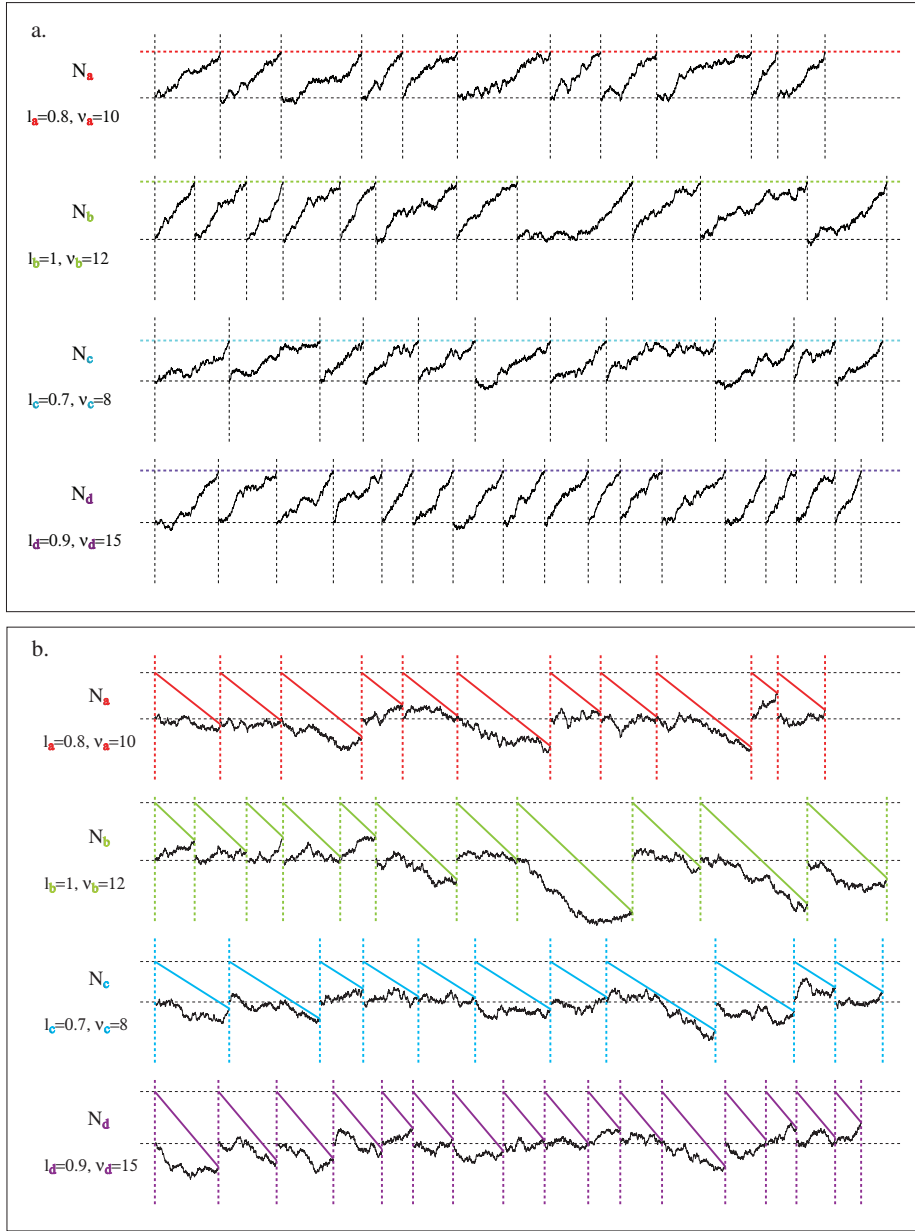


Figure 4.16: Sample paths of non-interacting stochastic perfect integrate-and-fire neurons. In the direct representation **a.**, each neuron spikes when its voltage, modeled as an upward drifted Wiener process, hits the firing threshold. In the effective representation **b.**, the drift is subtracted from the same voltage traces so that the sketched traces are Brownian trajectories. A stochastic event-driven simulation intends to simulate the spiking time while sampling the voltage trajectories on a minimal number of points.

Therefore, the spiking history h^a is made of first-passage times of the Wiener process W^a with the piecewise continuous effective barrier

$$\Lambda_a(t) = l_a - \nu_a t - \sum_{b \in U(a)} w_{b,a} \cdot C_b(t - d_{b,a}),$$

with the convention that a passage happens at a time of discontinuity t if $\Lambda_a(t^-) > X_t^a > \Lambda_a(t^+)$, whereas a passage cannot happen in t if $\Lambda_a(t^-) < \Lambda_a(t^+)$. In the formulation of interactions in terms of effective barriers, $\Lambda_a(t) + \nu_a t$ is a non-anticipating jump process that only depends on the past spiking history of upstream neurons $h^b(t)$, b in $U(a)$.

The preceding model is naturally endowed within a Markovian framework. To be more precise, at any given time t , suppose that we observe the last events of the spiking histories $\{t^a = \sup h^a(t)\}$ and that, for each neuron a and b , we maintain the schedule of spike deliveries $\{u_{a,b}\}$ that a neuron b is to receive as a result of the past activity of a . Then, we can consistently construct the ensuing dynamics of the network by simulating the sample-paths of the processes $\{X^a\}$. Event-driven strategies aim at simulating exactly the evolution of the network as a Markov chain, while simulating as few sample points as possible.

4.5.2 Local Neuron Update Rules

Before explaining our stochastic event-driven implementation, we describe the local updates that a neuron undergoes. If $U(a)$ is empty, neuron a evolves independently and each of its inter-spike time interval is an independent realization of the first-passage τ^a for a Wiener process and a straight barrier. The law of τ^a is known analytically and belongs to the class of inverse Gaussian distributions [202]. We denote this law as $IG(l_a, \nu_a)$ to underline its parametric dependence on the

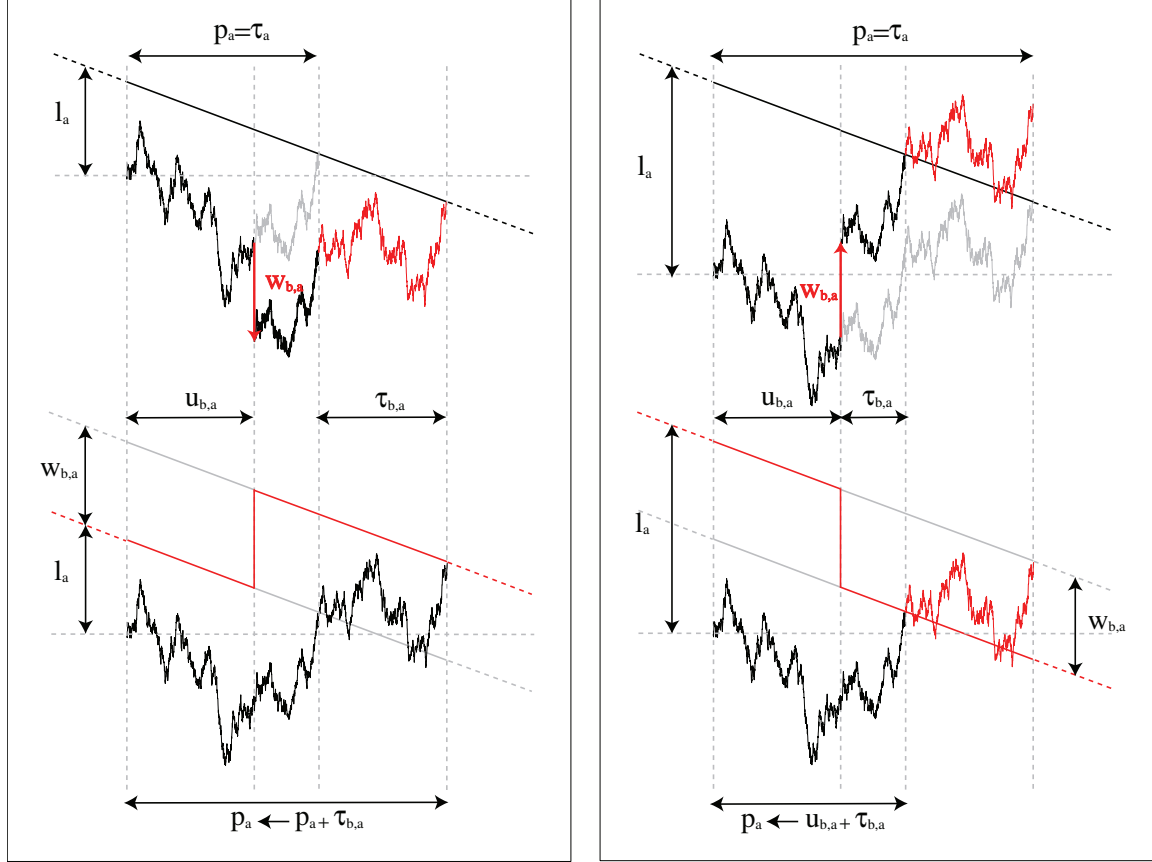


Figure 4.17: In both panels, the top figures represent the voltage trace in direct representation, the bottom figures are formulated in terms of an effective barriers. In the absence of interactions, the neuron spikes at the first-passage time τ_a . Upon arrival of a spike at time $u_{b,a}$, the trajectory is instantaneously updated: **Left panel: Inhibitory input.** The voltage trace is shifted downward and the estimate for the next spiking event p_a needs postponing, equivalent to the addition of a first-passage time $\tau_{a,b}$ to a barrier shifted by $w_{b,a}$. **Right Panel: Excitatory input.** The voltage trace is shifted upward. If the sample path does not instantaneously cross the barrier, the estimate for the next spiking event p_a needs to be advanced, amounting to substituting p_a with a new random variable with law given by equation (4.15).

threshold and drift values (see Appendix B). If one can design a random generator for $IG(l_a, \nu_a)$, it is thus possible to simulate the activity of neuron a without simulating the underlying sample-paths of X^a .

Now, assume we know that neuron a spikes at time $t_i^a(\omega)$ and denote the set of times $\{u_{b,a}\}$ of scheduled spike deliveries from neuron b in $U(a)$. In the absence of spike deliveries, $\{u_{b,a}\}$ is empty, the next spiking event t_{i+1}^a follows the law of $t_i^a + \tau^a$. If $\inf\{u_{b,a}\} < p_a = t_i^a(\omega) + \tau^a(\omega)$, a spike is received before neuron a would spike in the absence of interactions, and the estimate timing p_a for the next spike becomes incorrect. Indeed, if neuron a receives a spike from b at time $u_{b,a}$ with $0 < u_{b,a} - t_i^a(\omega) < \tau^a(\omega)$, the original continuous sample-path $t \mapsto X_t^a(\omega)$ instantaneously updates at $u_{b,a}$ upon receiving the spike:

$$X^a(\omega) \longleftarrow \bar{X}^a(\omega) : \begin{cases} X_t^a(\omega) & \text{if } t < u_{b,a} \\ X_t^a(\omega) + w_{b,a} & \text{if } t \geq u_{b,a} \end{cases}.$$

If the synapse is inhibitory, the update shifts the sample-path downward and delays the time neuron a would spike in the absence of later interactions. If the synapse is excitatory, the update shifts the sample-path upward. The shift may be strong enough to satisfy $X^a(\omega) + w_{b,a} > l_a$ and to cause neuron a to spike instantaneously. If $X^a(\omega) + w_{b,a} < l_a$, neuron a does not spike but the update hastens the time neuron a would spike in the absence of later interactions. Incidentally, the original estimate p_a appears as a provisional estimate of the next spiking time that inhibitory input postpones and excitatory input accelerates.

It is further possible to formulate these update rules rigorously from a probabilistic point of view. This requires some familiarity with conditioned stochastic processes that are related to the Wiener processes, namely the inverse Gaussian bridge process

and the first-passage bridge process [21]. Due the technicality of these results, we only give a brief account of their algorithmic formulation in the following, and justify their cogency in Figure 4.17. Details about the different probability laws of the conditioned processes are given in Appendix B.

The cases of inhibition and excitation need to be clearly distinguished. For an inhibitory input at time $u_{b,a}$, the provisional estimate p_a is updated by the addition of a delay whose distribution is given by $IG(w_{b,a}, \nu_a)$ according to

$$p_a \longleftarrow p_a + IG(w_{b,a}, \nu_a), \quad (4.13)$$

as explained in Figure 4.17. For an excitatory input at time $u_{b,a}$ and before the update, the membrane potential X^a follows the distribution of a drifted Wiener process conditioned to first reach the threshold l^a at time $t_i^a + p_a$, while starting in zero at t_i^a . Such a conditioned process X^a is referred as a first-passage bridge process (see Appendix B). It is known that $l_a - X^a$ follows the distribution of a 3-dimensional Bessel bridge $BES^3(l_a, 0)$ [21], making it possible to simulate the value of the conditioned sample-path X at time $u_{b,a}$. Updating the sample-path as

$$X_{u_{b,a}} \longleftarrow X_{u_{b,a}} + w_{b,a},$$

allows us to decide whether it causes the neuron to spike. In the event of a spike, p_a is later updated according to the reset rule

$$p_a \longleftarrow u_{b,a} + IG(l_a, \nu_a), \quad (4.14)$$

as explained in Figure 4.17. If the sample path remains below the spiking threshold, we still have to design an update rule for the provisional estimate p_a , that needs to be refined upon receiving excitatory input. As shown on Figure 4.17, the law of

the new provisional estimate is the first-passage time of a drifted Wiener process X^a at level $l_a - w_{b,a}$, knowing that X^a first-hit l^a at time $t_i^a + p_a$, while starting in zero at t_i^a . Such a conditional law follows the inverse Gaussian bridge distribution $IGB(w_{b,a}, l_a - w_{b,a}, u_{b,a}, p_a)$ (see Appendix B) and the update rule reads

$$p_a \leftarrow u_{b,a} + IGB(w_{b,a}, l_a - w_{b,a}, u_{b,a}, p_a). \quad (4.15)$$

Notice that we only need to simulate the value of X^a for those instances when neuron a receives an excitatory input.

4.5.3 Network Update Rule

Equipped with the previous local neuronal update rule, we can lay out an event-driven strategy that simulates exactly the spontaneous activity of an arbitrary network of stochastic perfect integrate-and-fire neurons. The crucial point of such an implementation is to carefully schedule the individual asynchronous neuron updates so that the Markovian nature of the dynamics is preserved. The general approach consists in maintaining

- a fixed size priority queue Q_n of neurons a that are time-stamped by their provisional estimate of spiking in the absence of interactions p_a ,
- a varying-size priority queue Q_d of spike deliveries, where receptions from neuron b to c are time-stamped by their time of delivery $\{u_{b,c}\}$.

At each iteration step, we select the next update event \mathcal{E} as the element of the two queues with highest priority, that is with minimum time stamp. If we select neuron a from Q_n , a spontaneous spiking of a occurs at time p_a , since a does not receive any input prior to its provisional spiking estimate p_a . It then entails the insertion of spike deliveries $\{u_{a,b} = p_a + d_{a,b}\}_b$ to downstream neurons $D(a)$ in Q_d . Contrarily,

if we select a time of delivery $u_{b,c}$ from Q_d , neuron c is updated according to the rules of the previous section. It can either stay silent or spike, in which case spike deliveries $\{u_{c,e} = u_{b,c} + d_{c,e}\}_e$ are scheduled.

From the mathematical standpoint, it is crucial to ensure that the order of events is proper, that is, the probability that two events happen at the same time is zero. This requirement entails the satisfactions of simple, albeit technical, constraints on the initial spiking times, on the initial deliveries and on the propagation delays. The initial spiking time of a neuron is defined as the last spike timing immediately preceding the zero time of a simulation, while initial deliveries consist of the schedule of spikes that are in transit at time zero. It is possible to show that, if the initial spiking times, the initial deliveries and the propagation delays are chosen from smooth distributions⁶, the probability of coincidental events is zero, which corresponds to the fact that the constraints of coincidence are met with negligible probability. Under such mild assumptions, the proper order of events guarantees the well-posedness of our algorithm.

To implement the strategy efficiently and consistently, we must make the following addition. If a neuron a receives a series of inputs between s and t from neurons in $U(a)$, the net update effect in t is to shift the sample-path by a value

$$\lambda_a(s, t) = \sum_{b \in U(a)} w_{b,a} (C_b(t) - C_b(s))$$

which we call the load of a between s and t . As long as $\lambda_a(s, t) < 0$, the input history of a between times s and t can only delay the occurrence of a spike. Instead of updating a at each spike delivery, we only update it if an excitatory interaction causes the load λ_a to become positive, and we reset it to zero after the update. This

⁶Actually, it is enough to choose the initial times and the propagation delays from distributions that are absolutely continuous with respect to the Lebesgue measure on their natural space of definition.

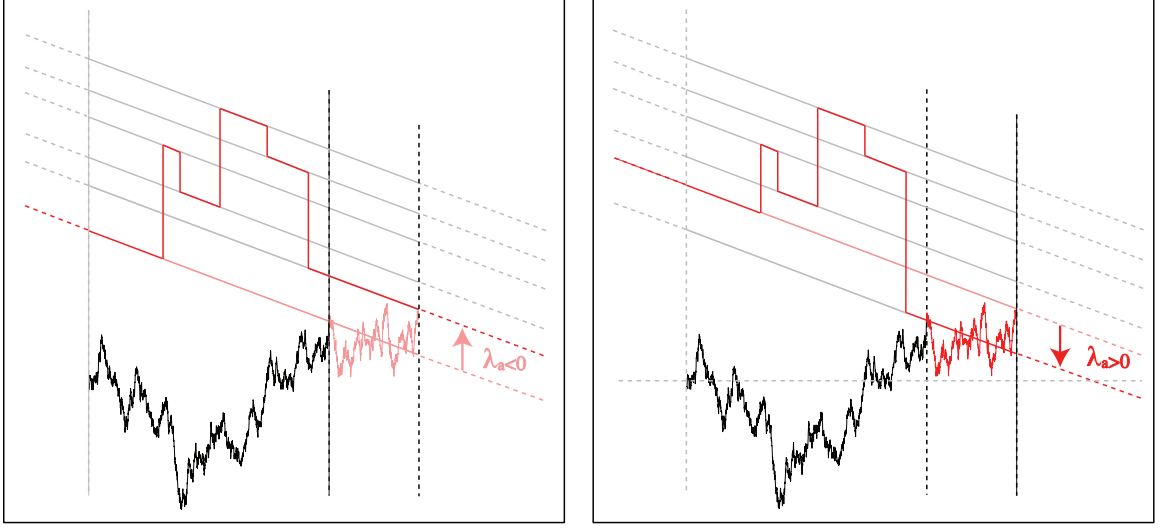


Figure 4.18: **Left panel: Net inhibitory input. Right Panel: Net excitatory input.** A negative load λ_a can only delay the spiking occurrences of a neuron a (lighter trace), as opposed to when the load becomes positive (darker trace). We only update a neuron a when a spike delivery causes λ_a to become positive, at the cost of treating separately the case when the next event \mathcal{E} is a “fake” spiking event.

implicitly defines the load of a neuron λ_a as the stochastic process obeying the update rules

$$\lambda_a \leftarrow \begin{cases} (\lambda_a + w_{b,a})^- & \text{if } \mathcal{E} \in Q_d \\ 0 & \text{if } \mathcal{E} \in Q_n \end{cases}.$$

Incidentally, it becomes possible to select for the next event a neuron a from Q_n with $\lambda_a < 0$.

In other words, the cost of considering λ_a is to introduce “fake” spontaneous spiking events. Suppose the next event is selected from Q_n by identifying neuron a as the neuron with highest priority. If $\lambda_a < 0$, the neuron has received some past inhibitory input, whose effect has not been realized yet. Neuron a should then remain silent and the event is deemed a “fake” spiking event. In such a case, we update p_a

according to

$$p_a \longleftarrow p_a + IG(-\lambda_a, \nu_a), \quad (4.16)$$

which corresponds to the inhibitory rule (4.13) where $w_{b,a}$ is substituted for $-\lambda$. Overall, p_a is solely updated forward in time when an update of neuron a is scheduled, with rule (4.16) for fake spiking events ($\lambda_a < 0$), or with rule (4.14) for real spiking events ($\lambda_a \geq 0$). In the meantime, p_a is only updated backward in time if the load of a becomes excitatory $\lambda_a > 0$ and according to the rule

$$p_a \longleftarrow u_{b,a} + IGB(\lambda_a, l_a - \lambda_a, u_{b,a}, p_a),$$

which corresponds to the excitatory rule (4.15) where λ_a replaces $w_{b,a}$. As a consequence, we considerably diminished the number of neurons' updates while ensuring that the successive estimates p_a are consistently time-ordered, as explained in Figure 4.18.

4.6 Event-Driven Implementation

In this section, we describe the algorithm of the stochastic event-driven strategy that has just been presented. We then discuss its implementation and analyze its complexity.

4.6.1 Pseudo-Code

As already mentioned, there are two sides to structure a program simulating a network of stochastic perfect integrate-and-fire neurons:

Updating: consists of locally updating the state of an individual neuron and issuing its time stamp as the time it would spike if solely driven by its internal dynamics.

There exist two types of updates:

- following the reception of an input, potentially triggering a spiking event,
- following a spiking event, scheduling spike deliveries to downstream neurons.

Sorting: manages the set of future events, which are prioritized according to provisional spike timing and selected based on highest priority

We summarize these principles by providing the pseudo-code that implements an iterative step of the algorithm in Procedure 6. We adopt the notations of the previous section:

Neuron parameters: The neurons' internal dynamics are entirely prescribed by their threshold value l_a , negative drift ν_a and refractory period ρ_a .

Network parameters: The topology of the network is described by the weight $w_{a,b}$ of the synapses, whereas its temporal properties are summed up in the delivery delays $d_{a,b}$.

Priority Stamps: Events are sorted according to the time separating their occurrence from the network's current time, coinciding with the time of the last update. The time to the next internal spiking event of neuron a is p_a , and $u_{a,b}$ is the time to the next spike delivery from neuron a to b .

History load: The spiking or update of a neuron a is effective if the past history of input of a gives rise to a non-negative load $\lambda_a \geq 0$.

The iteration of updates is managed through the first step of our algorithm, which consists in selecting the next update event: either an internal event from the fixed-size priority queue \mathcal{Q}_n or an interaction event from the variable-size priority queue \mathcal{Q}_d . To implement these two priority queues \mathcal{Q}_n and \mathcal{Q}_d , we use binary heaps,

Procedure 6 Update Network

Select $p_a = \inf p_\alpha$ and $u_{c,b} = \inf u_{\gamma,\beta}$
if $p_a < u_{c,b}$ **then**
 Update neuron a according to:
 if $\lambda_a \geq 0$ **then**,
 Schedule deliveries $\{p_a + d_{a,b}\}_b$
 Update $p_a \leftarrow p_a + IG(l_a, \nu_a)$ and $\lambda_a \leftarrow 0$
 else
 Update $p_a \leftarrow p_a + IG(-\lambda_a, \nu_a)$ and $\lambda_a \leftarrow 0$
 end if
else
 $\lambda_b \leftarrow \lambda_a + w_{c,b}$
 if $\lambda_b \geq 0$ **then**
 Update neuron b according to:
 if $\lambda_b + X^b \geq l_b$ **then**,
 Schedule deliveries $\{u_{c,b} + d_{b,e}\}_e$
 Update $p_b \leftarrow u_{c,b} + IG(l_b, \nu_b)$
 else
 Update $p_b \leftarrow u_{c,b} + IGB(\lambda_b, l_b - \lambda_b, u_{c,b}, p_b)$
 end if
 end if
end if

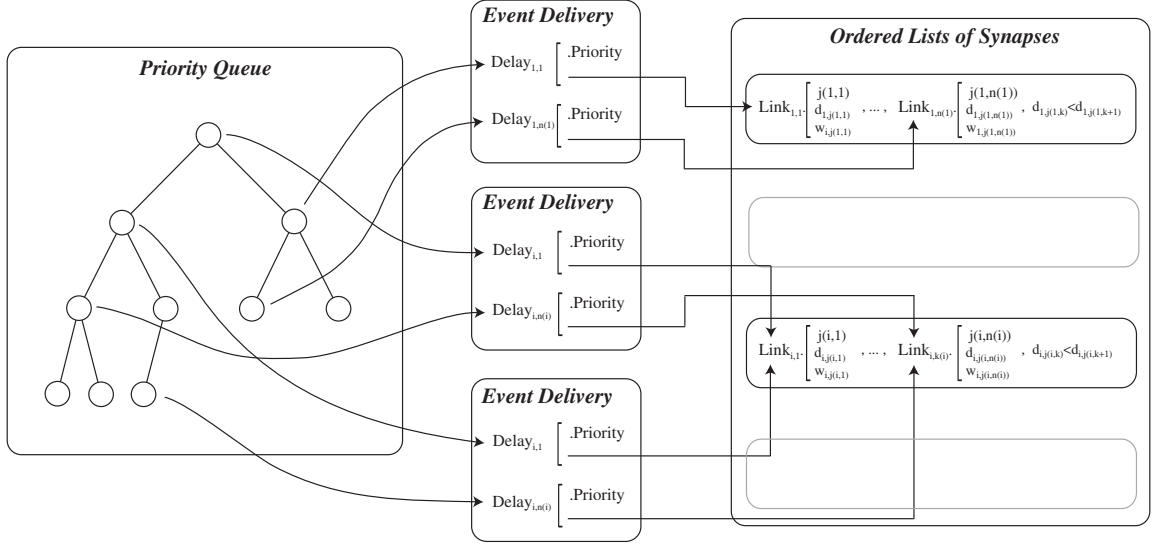


Figure 4.19: Heap Structure of Spike Deliveries. The synapses are organized as an array of lists of synapses ordered according to increasing delay. Each time a neuron i spikes, a delivery event is created for every neuron j to which i connects, with a priority stamp initialized to the delay $d_{i,j}$. The events are inserted into the dynamic heap structure that keeps track of the delivery schedule of spikes in the network.

a well-known data structure for which the cost of inserting an element (scheduling a future event) or deleting the root (selecting the next event) is $O(\log S)$, where S is the number of elements in the heap. As will be discussed later, the management of the priority queue \mathcal{Q}_d is essential to the computational efficiency of the algorithm and its principle is represented in Figure 4.19.

Including the refractory periods within our framework is tantamount to freezing the internal dynamics of a neuron after the last spiking event $t^a = \sup h^a(t)$ for a duration ρ_a . Refractory periods are implemented by updating the state of a neuron a for time t such that $t - t^a > \rho_a$, while altering the reset rule as $X_{t^a + \rho_a}^a = 0$. Since it is relatively straightforward to do so, we do not include the implementation of refractory periods in the pseudocode of Procedure 6.

4.6.2 Monte-Carlo Sampling

Before analyzing the complexity of the proposed algorithm, we must underline the fact that our implementation pertains to the class of Monte-Carlo methods [146, 183], and as such, its cogency relies crucially on the existence of exact and efficient random generators.

The evolution of the network of integrate-and-fire neurons is fundamentally stochastic and requires extensive sampling from the parametric laws IG , BES^3 , IGB . In simulating the network for a vast number of neurons, and incidentally for an even vaster number of connections, all interaction events need to be precisely time-ordered, thus making it necessary to employ exact random generators. Even a small bias can consistently introduce spurious ordering and cause the sustained propagation of “wrong” frustrated patterns of firing. From a more computational perspective, since a stochastic update is required each time an event occurs, the sampling process must demand as few arithmetic operations as possible, and specifically, we want the number of required operations to be bounded deterministically. This last point precludes the use of the most common exact sampling method, the *rejection algorithm* [239], that consists of sampling from a known distribution and then subjecting the sample to a test, to determine whether it is a true sample or a new a sequence of drawing and testing is needed.

When concerned with the parametric laws IG , BES^3 , IGB , it is possible to avoid the rejection algorithm thanks to a powerful statistical result of Michael, Schucany and Haas [148]. Its scope covers real random variables X for which there exists a polynomial P such that the transformed variable $U = P(X)$ follows a distribution for which an exact and efficient generator is already known. If P is invertible with known inverse function P^{-1} , simulating the original variable U

consists of first sampling from U and then returning $P^{-1}(U)$. A potential issue stems from the fact that P is often non-invertible since the equation $P(x) = u$ generally admits many roots $x_1(u), \dots, x_n(u)$. The essential merit of the Michael-Schucany-Haas method is that it provides a simple rule for deciding in which proportion to choose from the n possible roots. As opposed to the rejection algorithm, it always provides an exact sample of the law of X without ever rejecting a sample u from U , and thus in a finite deterministic number of steps. Fortunately, the laws of IG , BES^3 , IGB are closely connected to the χ^2 law, which is simulated as $1/\Xi^2$ where Ξ is a standard Gaussian variable of law $\mathcal{N}(0, 1)$. We give the implementations of the generators the laws IG , BES^3 , IGB in Supplementary Material.

4.6.3 Complexity

Equipped with such random generators, we are now in a position to discuss the complexity of the algorithm. As an event-driven algorithm, it is natural to investigate the computational cost of the algorithm per event of interest. In our case, that is the cost of generating a spike.

To establish an upper-bound to the cost of generating a spike, consider a network of N neurons for which each neuron connects, on average, to aN other neurons, and where $a < 1$ is the sparsity of the network. For the sake of simplicity, suppose that the network operates in steady regime and homogeneously, implying that each neuron has the same constant firing rate r . During each time period $1/r$, the homogeneity hypothesis entails that each neuron spikes on average once, scheduling the delivery of aN^2 spikes. Since we operate in steady regime, in the meantime, the network has received the delivery of as many spikes, potentially leading to an update if the load of the neuron becomes excitatory. From these observations, we can deduce the overall

computational cost per spike of our algorithm.

Update Cost: Since the random generators operate in a finite number of arithmetic steps, the cost of a local Monte-Carlo update is always finite. Its upper bound is denoted by c . In realizing that our implementation is such that the number of local update is always less than the number of spike deliveries (aN^2 for N spikes within the network), we can readily estimate the cost of the local update per spike as aNc .

Sorting Cost: Since we use binary heaps structures, the scheduling and deletion of aN^2 delivery events per duration $1/r$ incur a cost in $2aN^2 \log(S)$, where S is the size of the priority queue \mathcal{Q}_d in steady regime. The priority queue \mathcal{Q}_d acts as a buffer and its size S is the number of deliveries that have been scheduled and are still in transit, specifically $D \times r \times aN^2$ if D denotes the average delay. Similarly, the local updates (that occur less frequently than aN^2) may cause the reshaping of the priority queue of internal events \mathcal{Q}_n , which comprises N neurons, yielding $aN^2 \log(N)$ operations.

To summarize, we encapsulate the previous discussion in the following result:

Property 5. *For a network of N stochastic perfect integrate-and-fire neurons connected with sparsity $a < 1$ in steady state and homogeneous regime with firing rate r and time delay D , the complexity of the stochastic event-driven algorithm as the cost of generating a spike is bounded by*

$$\mathcal{C}(N, a, r, D, c) = aN(5 \log(N) + 2 \log(rDa) + c),$$

where c denotes the finite unit cost of a local Monte-Carlo update.

The computational cost of the algorithm lies primarily in the management of the priority queues \mathcal{Q}_n and \mathcal{Q}_d . For the case of the spike delivery events, it is essential to

optimize the implementation of insertion in the heap, by treating the volleys of spike deliveries that each spike elicits in a pre-ordered fashion (see Figure 4.19). This way, the reshaping of the binary tree takes as few steps as possible. More importantly, the size of the priority heap \mathcal{Q}_d grows enormously with the number of connections as $rDaN^2$, so that memory management and memory access become the limiting factor to the speed of execution. However, by relying exclusively on local update rules organized through the sorting of prioritized events, this algorithm is readily parallelizable, thus offering a simple way to alleviate these constraints.

Interestingly, decreasing the average time delay of spike propagation D , shortens the interaction memory of the network, offering another way to diminish the size of \mathcal{Q}_d . In the mammalian brain, the maximum spike delays are of the order of a few milliseconds [68] and the maximum firing rates is typically of the order of 100Hz [198], suggesting that low values of rD are the norm, thus limiting the size of \mathcal{Q}_d .

These results contrast with stochastic clock-based algorithms for which neurons are only allowed to emit and receive spikes at those times that are multiples of the simulation time-step Δt . In this context, it is possible to include finite-propagation delays that, for also being multiples of Δt , can be managed without being sorted [103]. However, at each time step and irrespective of the existence of an interaction event, the state of each neuron needs to be evaluated. Thus, the cost-per-spike for a network in homogeneous steady regime is $aN \times u + (r\Delta t)^{-1} \times c$, where u denotes the cost of an interaction update and c the cost of sampling the membrane potential. If the dependence of the complexity is only linear in N , invariably updating neurons at each time-step takes a fixed computational toll that becomes prohibitive even at reasonable precision.

Indeed, to avoid ambiguities about the order of networks updates, the precision should be high enough to ensure that all interaction events happens at separate times with

good confidence. It is easy to show that for homogeneous networks in steady regime, that requirement leads to chose $\Delta t = a^2 N^2 (q \times r)^{-1}$, where q represents the confidence of event separation

$$q \sim -\log \mathbb{P} [\exists t \neq s \in \mathcal{T} \mid \lfloor t/\Delta t \rfloor = \lfloor s/\Delta t \rfloor] ,$$

where \mathcal{T} contains the times of incoming spikes that a neuron receives during an inter-spike period⁷. The resulting cost-per-spike now becomes quadratic in the size of the network, which is worse than our stochastic event-driven strategy in $O(N \log(N))$. Moreover, we argue that for practical use, the situation is actually much more unpropitious to stochastic clock-driven methods. Indeed, neural networks generally exhibits highly fluctuating temporal spiking activity, that varies from one neuron to another. Clock-based implementations oversample neurons that spikes rarely and undersample neurons involved in firing cascades. By contrast, stochastic event-driven strategies are impervious to the previous sampling defect since their computations only affect interactions. In that regard, our implementation is especially amenable to simulate networks with strong inhibition. The introduction of inhibitory interactions slow down the overall rate of events, in turn downsizing the heap of spikes in transit \mathcal{Q}_d , whose sorting accounts for most of the computational cost.

4.7 Simulation of Networks

In this section, we illustrate concretely our stochastic event-driven algorithm exploiting the Markov nature of the network, seen as a collection of interacting homogeneous Markov chains. For the sake of clarity, we first distinguish between purely inhibitory and excitatory networks, and then proceed to networks including excitatory and inhibitory interactions. All figures have been elaborated from the outcomes of our

⁷Notice that p tends to zero for increasing confidence of separation, so that Δt diverges.

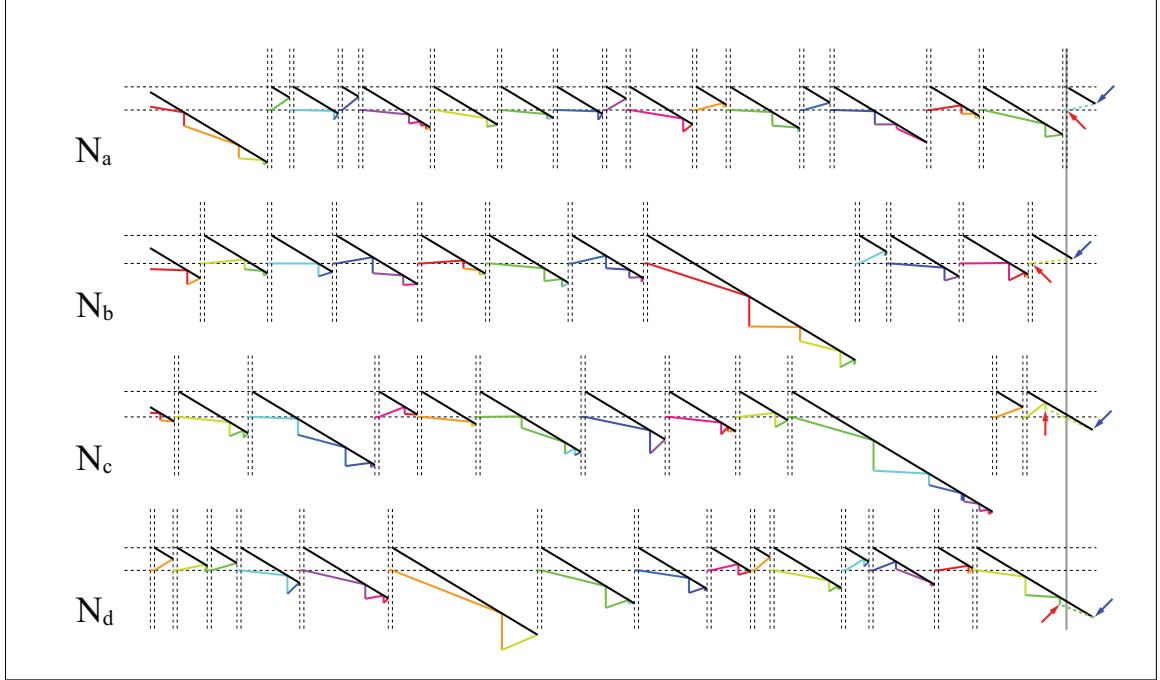


Figure 4.20: Sample paths of 4 neurons embedded in a fully-connected inhibitory network. The parameters of the simulation have been chosen for the sake of a clear illustration and should not be considered relevant.

machine-independent C-code implementation.

4.7.1 Inhibitory Interactions

In Figure 4.7.1, we represent the sample points of a neuron's voltage trajectory, as simulated by our stochastic event-driven strategy for an inhibitory network. The network is made of 4 fully interconnected neurons N_a, N_b, N_c and N_d .

For each neuron, the black downward traces represent the effective barriers. Every time a neuron spikes, the effective barrier is reset to start from the threshold value, after a refractory time materialized by silent periods between two vertical dashed lines. The succession of sample voltage values is sketched as follows. For a given neuron, the drawing of a provisional spiking time is shown as a point on the

effective threshold, since the voltage trajectory hits the barrier at that time. If the neuron receives some interaction input before this provisional time, the estimate of the next spiking time needs to be updated. Since the network is inhibitory, the occurrence of the spike is delayed according to the local update rule. This procedure is represented by shifting downward the neuron’s voltage by the amount of inhibition and drawing the next provisional estimate as for a non-interacting neuron. There are two types of internal events: the ones for which the neurons’ internal dynamics cause a spiking event (“true” event) or the ones that turn out not to be spiking events (“fake” event), for having ignored past inhibition.

Notice that the simulation proceeds asynchronously: each neuron’s current state is comprised of a time of last update t_n (light arrows) and an estimate of the next spiking time τ_n (dark arrows). The network’s implementation moves forward in time so that the network time stamp t , understood as the time of the last update (grey vertical line), satisfies $t_n \leq t < \tau_n$ for all neurons. For instance, in Figure 4.7.1, the network is stalled at the last time neuron N_a spikes, and thus, the last update state of N_a is the reset state. At the same time, neuron N_b has not interacted since its last spiking time and its last update state is also its reset state. Contrarily, neuron N_c and N_d have been updated once and twice, respectively. For a purely inhibitory network, the provisional estimates only move forward in time since they can only be postponed by interactions. Thus, the history of provisional spiking time translates exactly the history of the network interactions and no computational power is lost to simulate unnecessary sample points of the voltage trajectories.

4.7.2 Excitatory Interactions

In Figure 4.21 we sketch out the sample points of a neuron’s voltage trajectory, as simulated by our stochastic event-driven strategy for an excitatory network. Again,

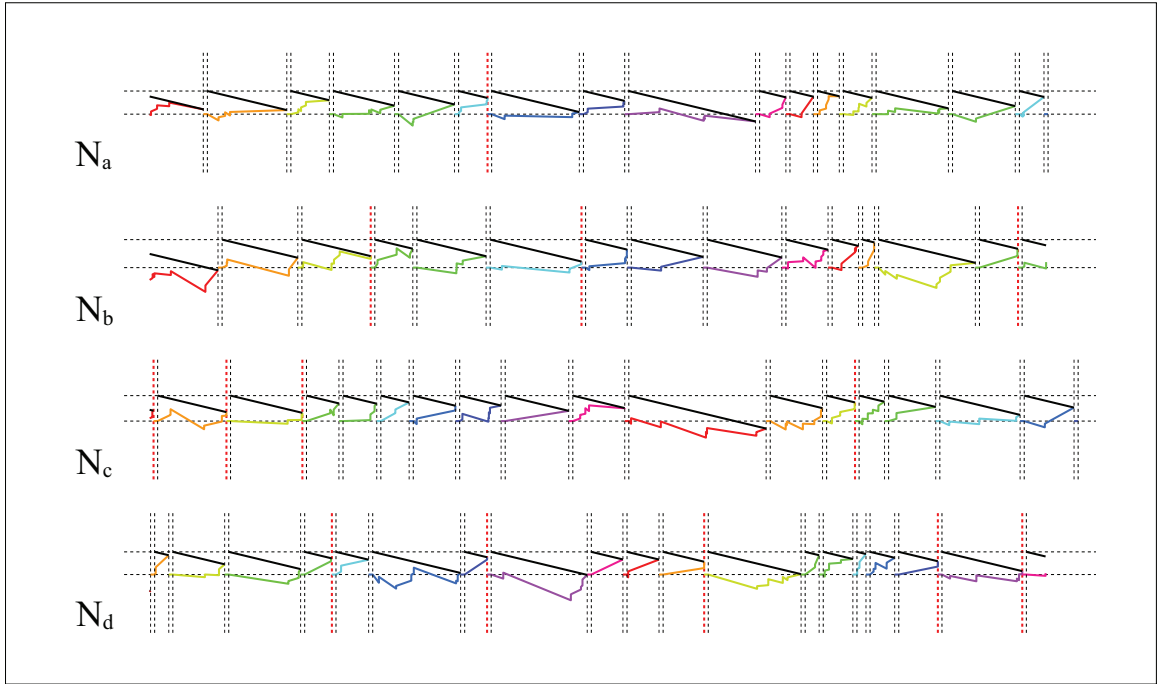


Figure 4.21: Sample paths of 4 neurons embedded in a fully-connected excitatory network in asynchronous time. The parameters of the simulation have been chosen for the sake of a clear illustration and should not be considered relevant.

the network is made of 4 neurons N_a, N_b, N_c and N_d that fully connect to one another.

Contrary to the inhibitory case, for each neuron, the voltage trajectory is updated by being shifted upward by the amount of excitatory input it receives when interacting with another neuron. Since interactions cannot delay a provisional estimate, the only possible internal events are true spiking events. Accordingly, the color of the voltage trace changes each time a spike happens. The spiking events can be distinguished in two classes: the spikes that result from a neuron's own internal dynamics driving the trajectory to hit the barrier (usual black dashed lines) and the spikes that directly follow the receiving of excitatory input (solid black lines). Notice that the network is stalled at the last time neuron N_c fires and all the other neurons are shown in the state of their last update, which is in keeping with their asynchronous timing.

To capture how the updating of the provisional estimate proceeds, we show the same network sample path with anticipations in Figure 4.22. By anticipations we mean that we represent the sample points of the provisional estimates, which are progressively hastened every time a neuron receives an excitatory input. Concretely, the history of successive provisional estimates lies on the barrier (since we are concerned with potential spiking time) and moves backward in time on this barrier. In order to clearly depict this process, we do not represent the effective barrier on Figure 4.22. The extensions of the voltage trajectories after an actual spiking event reveal how far in the future the original estimates of the spiking time are. In fact, the algorithm simulates admissible voltage traces from a future time back to the instant they cause the next network event, at which point the portion of the trajectory in the future can be considered irrelevant without altering the Markovian nature of the network evolution. Incidentally, if simulating the trajectory in the future ensures the consistency

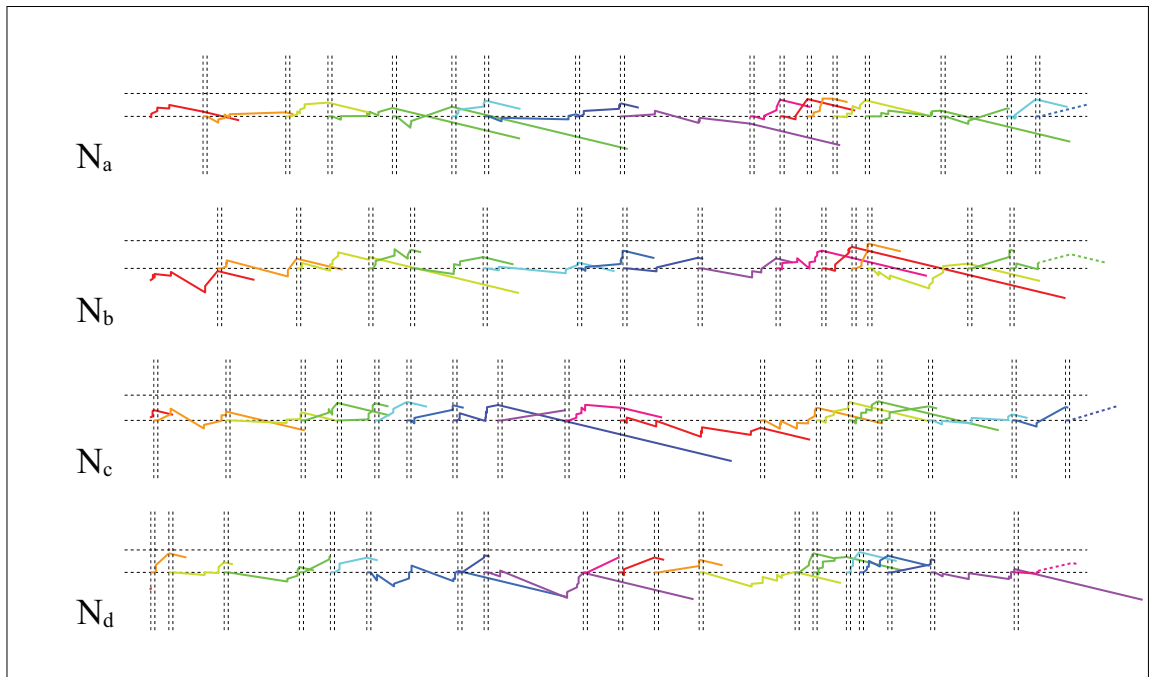


Figure 4.22: Sample paths of 4 neurons embedded in a fully-connected excitatory network with anticipations. The parameters of the simulation have been chosen for the sake of a clear illustration and should not be considered relevant.

of the network update, it devotes computational power to sampling trajectories that are not appearing in the sampled end results.

4.7.3 Balanced Network

Finally, we represent in Figure 4.23 the sample paths of 4 neurons embedded in a network with excitation and inhibition. The specific network under consideration is made of 100 neurons that fully connect to one another. We adopt the same representation as for the case study of excitatory networks, sketching the network's sample path in asynchronous time (**a.**) and with anticipations (**b.**).

Since the neurons N_a , N_b , N_c and N_d are the targets of both inhibitory and excitatory input, each of their sample paths presents the characteristics of excitatory and inhibitory updates. The asynchronous representation **a.** clearly shows that provisional estimates are regularly delayed by inhibition, while the representation with anticipations **b.** reveals that provisional estimates are also hastened by excitation. Notice that the update process is asymmetric: the downward shifts of the trajectories have much larger amplitude than the upward shifts, but are also less frequent. As long as a neuron's past history of interaction remains inhibitory, the implementation postpones local updates until the occurrence of an internal event, at which point the accumulated inhibition is realized. As a consequence, the number of times a trajectory needs to be sampled is drastically reduced as exemplified by Figure 4.23. For the given network at stake, if each spiking event causes a mean number of 100 spike deliveries, a trajectory leading up to a spiking time counts only 7 sample points on average.

To conclude on the efficiency of our stochastic event-driven algorithm, we simulate 100,000 spikes for a fully-connected network of 200 neurons (50 inhibitory neurons). This represents the management of 40,000 connections and 20,000,000 spike deliver-

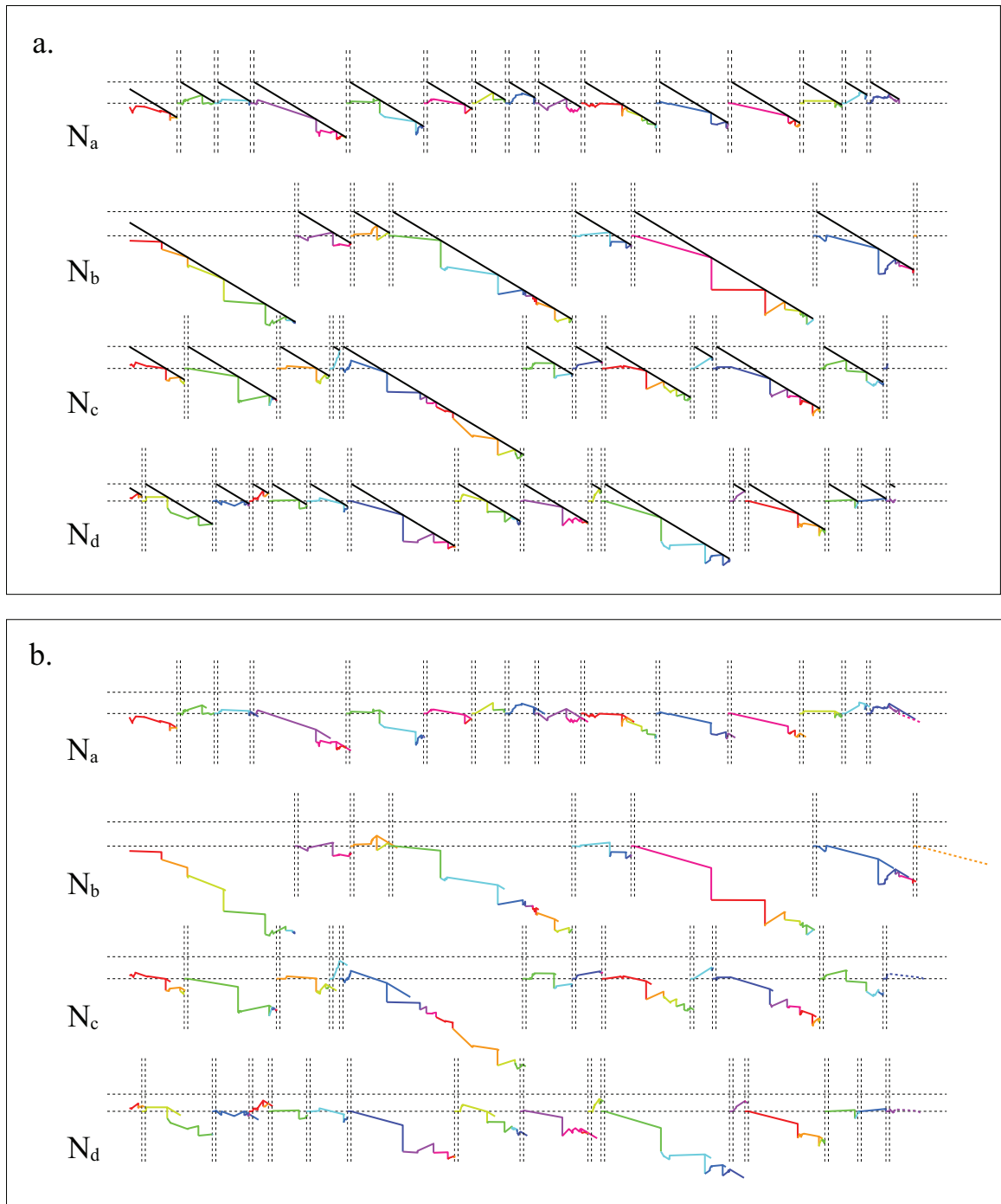


Figure 4.23: Sample paths of 100 neurons embedded in a fully-connected excitatory and inhibitory network. The parameters of the simulation have been chosen for the sake of a clear illustration and should not be considered relevant.

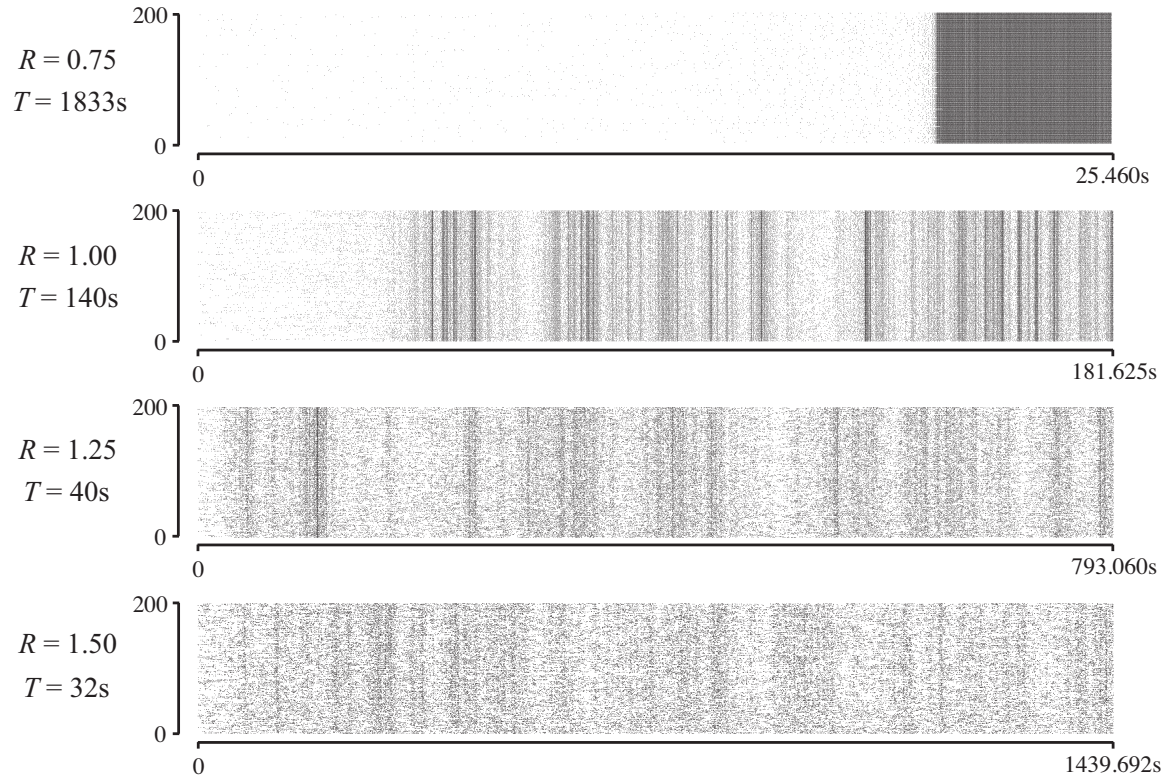


Figure 4.24: Raster plot of the spontaneous activity of a fully connected network of 200 neurons with a subpopulation of 50 inhibitory neurons. Each raster is made of 100,000,000 spiking events for different conditions of inhibition: R is the ratio of the weight of inhibitory synapses over the weight of excitatory synapses, T denotes the computational time required to run the simulation. The parameters of the simulation have been chosen to be relevant within the limits of a population of 200 neurons (see main text).

ies, which takes of the order of the minute on a standard computer for relevant range of parameters. In Figure 4.24, we show the raster plots of the population for varying degree of inhibition. The reduced parameters of every neuron are set to $l = 1V$ for the threshold and $\nu = 0.1V.s^{-1}$ for the drift. The network is embedded in a three-dimensional space where neurons are points that are uniformly distributed on a sphere. The interneuronal delays are taken proportional to the great circle distance on the sphere, so that the temporal radius is $1/\pi$ ms. Notice that the delays are thus drawn from a smooth probability distribution. The interactions of the network are prescribed by choosing at random to assign three fourths of the neurons to be excitatory with synaptic weight $w^+ = 0.01V$, and the other fourth to be inhibitory with synaptic weight at $w^- = Rw^+$, for varying values of R in the range $[0.75, 2]$. For each values of R , the network's evolution starts from the same initial condition, with initial times drawn from the stationary distribution of the corresponding non-interacting network ($w = 0$) and empty delivery schedules. Suppressing inhibition by lowering R facilitates the appearance of cascades of firing, until a point when the network becomes excitatory enough to saturate at extremely high firing. In Figure 4.25, we depict the fact that this transition is paralleled with a drastic increase in the computational cost of generating a spike.

Conclusion

We have designed an efficient algorithm for the simulation of stochastic perfect integrate-and-fire neurons. Our method, based on a Markovian framework [226], supposes the identification of closed-form expressions for the update rules of the network, which makes the exact simulation possible. The principles behind our implementation readily apply to general linear integrate-and-fire neurons. This encompasses the leaky integrate-and-fire model, for which the subthreshold dynamics follows an

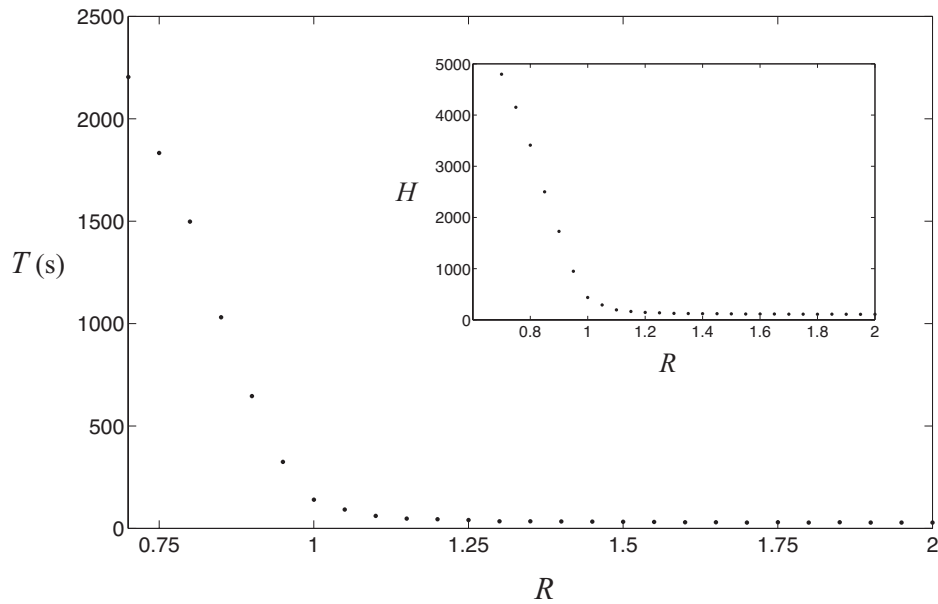


Figure 4.25: Computational time needed to simulate the network of 200 neurons of Figure 4.24. For inhibitory and excitatory synapses of the same weight $R = 1$, the computation takes 2 minutes. Notice the sharp increase of simulation time when alleviated inhibition causes the network to saturate. **Inset:** H is the mean size of the heap of spike deliveries, that is the average number of spikes in transit at each network update.

Ornstein-Uhlenbeck process and possesses a stationary distribution [230]. However, in the absence of closed-form expression, little is known about the simulation of the processes related to the first-passage times of such processes [217].

In all generality, neural activity should be described in terms of excitable systems [128], whose dynamics are inherently nonlinear. Unfortunately, the development of an event-based approach for nonlinear integrate-and-fire neurons is still elusive, with the notable exception of the deterministic quadratic integrate-and-fire neurons [223]. This is essentially due to the fact that a nonlinear subthreshold dynamics precludes the formulation of the first-passage problem with drift in terms of an effective barrier. As a result it is impossible to decouple the current-mediated interactions and the intrinsic dynamics of the neurons, and the formulation of simple update rules appears out of reach.

Despite only being valid for stochastic perfect integrate-and-fire neurons, we argue that our numerical method is of primary interest for its ability to simulate large-scale neuronal networks with little computational cost and exquisite temporal precision. Indeed, the spiking activity of a network is sensitive to perturbations of spiking times. This fact is best illustrated through the concept of polychronization [103], whereby only certain temporal spiking patterns of upstream neurons elicit the generation of specific cascades of spikes in downstream neurons. More importantly, networks are known to evolve according to spike-timing-dependent-plasticity rules [53, 1, 39], where the precise ordering of spikes' emissions and receptions is absolutely crucial. In the context of noisy neural assemblies, the exactness of our algorithm makes it especially suited to the simulation of such evolvable networks, while its efficiency shall facilitate the identification of polychronized motifs and elucidate the event structure of spike trains [67, 227].

Appendix A

Optimality Criterion of the Multi-resolution Decomposition

In this Appendix, we draw from the theory interpolating spline to further characterize the nature of our proposed basis for the construction of Gauss-Markov processes. We show, following [59, 108], that the finite-dimensional sample paths of our construction induce a nested sequence \mathcal{E}_N of reproducing Hilbert kernel space (RKHS). In turn, the finite-dimensional process \mathbf{X}^N naturally appears as the orthogonal projection of the infinite-dimensional process \mathbf{X} onto \mathcal{E}_N . We then show that such a RKHS structure allows us to define a unicity criterion for the finite-dimensional sample path as the only functions of \mathcal{E} that minimize a functional, called Dirichlet energy, under constraint of interpolation on D_N (equivalent to conditioning on the times D_N), thus extending well-known results to multidimensional kernel [195]. In this respect, we point out that the close relation between Markov processes and Dirichlet forms is the subject of a vast literature, largely beyond the scope of the present paper (see e.g. [72]).

A.1 Sample Paths Space as a Reproducing Hilbert Kernel Space

In order to define the finite-dimensional sample paths as a nested sequence of RKHS, let us first define the infinite-dimensional operator

$$\Phi : \begin{cases} l^2(\xi; \Omega) & \mapsto L_{\mathbf{f}}^2 \\ \xi & \mapsto \Phi[\xi] = \left\{ t \mapsto \sum_{(n,k) \in \mathcal{I}} \phi_{n,k}(t) \cdot \xi_{n,k} \right\} \end{cases}.$$

Since we know that the column functions of $\phi_{n,k}$ form a complete orthonormal system of $L_{\mathbf{f}}^2$, the operator Φ is an isometry and its inverse satisfies $\Phi^{-1} = \Phi^T$, which reads

for all \mathbf{v} in $L_{\mathbf{f}}^2$

$$[\Phi^{-1}[\mathbf{v}]]_{n,k} = \int_U \phi_{n,k}^T(t) \cdot \mathbf{v}(t) dt = \mathcal{P}(\phi_{n,k}, \mathbf{v}).$$

Equipped with this infinite-dimensional isometry, we then consider the linear operator $\mathcal{L} = \Phi \circ \Delta$ suitably defined on the set

$$\mathcal{E} = \{\mathbf{u} \in C_0(U, \mathbb{R}^d) \mid \mathcal{L}[\mathbf{u}] \in L_{\mathbf{f}}^2\} = \{\mathbf{u} \in C_0(U, \mathbb{R}^d) \mid \Delta[\mathbf{u}] \in l^2(\xi\Omega)\}.$$

with $\|\xi\|_2^2 = \sum_{n,k \in \mathcal{I}} |\xi_{n,k}|_2^2$, the l^2 norm of $\xi\Omega$. The set \mathcal{E} form an infinite-dimensional vectorial space that is naturally equipped with the inner product

$$\forall (\mathbf{u}, \mathbf{v}) \in \mathcal{E}^2, \quad \langle \mathbf{u}, \mathbf{v} \rangle = \int_U \mathcal{L}[\mathbf{u}](t)^T \cdot \mathcal{L}[\mathbf{v}](t) dt = (\mathcal{L}[\mathbf{u}], \mathcal{L}[\mathbf{v}]),$$

Moreover since $\mathbf{u}(0) = \mathbf{v}(0) = \mathbf{0}$, such an inner product is definite positive and consequently, \mathcal{E} forms an Hilbert space.

Remark 4. Two straightforward remarks are worth making. First, the space \mathcal{E} is strictly included in the infinite-dimensional sample paths space ${}_x\Omega'$. Second notice that, in the favorable case $m = d$, if \mathbf{f} is everywhere invertible with continuously differentiable inverse, we have $\mathcal{L} = \mathcal{D} = \mathcal{K}^{-1}$. More relevantly, the operator \mathcal{L} can actually considered a first-order differential operator from \mathcal{E} to $L_{\mathbf{f}}^2$ as a general left-inverse of the integral operator \mathcal{K} . Indeed, realizing that on $L_{\mathbf{f}}^2$, \mathcal{K} can be expressed as $\mathcal{K} = \Psi \circ \Phi^{-1}$, we clearly have

$$\mathcal{L} \circ \mathcal{K} = \Phi \circ \Delta \circ \Psi \circ \Phi^{-1} = Id_{L_{\mathbf{f}}^2}.$$

We now motivate the introduction of the Hilbert space \mathcal{E} by the following claim:

Proposition 15. *The Hilbert space $(\mathcal{E}, \langle, \rangle)$ is a reproducing kernel Hilbert space*

(RKHS) with $\mathbb{R}^{d \times d}$ -valued reproducing kernel \mathbf{C} , the covariance function of the process \mathbf{X} .

Proof. Consider the problem of finding all elements \mathbf{u} of \mathcal{E} solution of the equation $\mathcal{L}[\mathbf{u}] = \mathbf{v}$ for \mathbf{v} in $L^2_{\mathcal{F}}$. The operator \mathcal{K} provides us with a continuous $\mathbb{R}^{d \times m}$ -valued kernel function \mathbf{k}

$$\forall (t, s) \in U^2, \quad \mathbf{k}(t, s) = \mathbb{1}_{[0, t]}(s) \mathbf{g}(t) \cdot \mathbf{f}(s),$$

which is clearly the Green function for our differential equation. This entails that the following equality holds for every \mathbf{u} in \mathcal{E}

$$\mathbf{u}(t) = \int_U \mathbf{k}(t, s) \mathcal{L}[\mathbf{u}](s) ds.$$

Moreover, we can decompose the kernel \mathbf{k} in the $L^2_{\mathcal{F}}$ sense as

$$\mathbf{k}(t, s) = \sum_{(n, k) \in \mathcal{I}} \psi_{n, k}(t) \cdot \phi_{n, k}^T(s)$$

since we have

$$\begin{aligned} \mathbf{k}(t, s) &= \mathcal{K} \left[\delta_s Id_{L^2_{\mathcal{F}}} \right] (t) \\ &= \mathcal{K} \left[\sum_{(n, k) \in \mathcal{I}} \phi_{n, k} \cdot \phi_{n, k}^T(s) \right] \\ &= \sum_{(n, k) \in \mathcal{I}} \mathcal{K} [\phi_{n, k}] (t) \cdot \phi_{n, k}^T(s), \end{aligned}$$

with $\delta_s = \delta(\cdot - s)$. Then, we clearly have

$$\mathbf{C}(t, s) = \int_U \mathbf{k}(t, u) \cdot \mathbf{k}(s, u)^T du = \sum_{(n, k) \in \mathcal{I}} \psi_{n, k}(t) \cdot \psi_{n, k}^T(s)$$

where we recognize the covariance function of \mathbf{X} , which implies

$$\mathbf{k}(t, s) = \sum_{(n,k) \in \mathcal{I}} \psi_{n,k}(t) \cdot \mathcal{L}[\phi_{n,k}]^T(s) = \mathcal{L}[\mathbf{C}(t, \cdot)] .$$

Eventually, for all \mathbf{u} in $L_{\mathbf{f}}^2$, we have:

$$\mathbf{u}(t) = \int_U \mathcal{L}[\mathbf{C}(t, \cdot)](s) \cdot \mathcal{L}[\mathbf{u}](s) ds = \mathcal{P}\langle \mathbf{C}(t, \cdot), \mathbf{u} \rangle ,$$

where we have introduced the \mathcal{P} -operator associated with the inner product \langle, \rangle : for all $\mathbb{R}^{d \times d}$ -valued functions \mathbf{A} and \mathbf{B} defined on U such that the columns $c_i(\mathbf{A})$ and $c_i(\mathbf{B})$, $0 \leq i < d$, are in \mathcal{E} , we define the matrix $\mathcal{P}\langle \mathbf{A}, \mathbf{B} \rangle$ in $\mathbb{R}^{d \times d}$ by

$$\forall 0 \leq i, j < d, \quad \mathcal{P}\langle \mathbf{A}, \mathbf{B} \rangle_{i,j} = \langle c_i(\mathbf{A}), c_j(\mathbf{B}) \rangle .$$

By the Moore-Aronszajn theorem [11], we deduce that there is a unique reproducing kernel Hilbert space associated with a given covariance kernel. Thus, \mathcal{E} is the reproducing subspace of $C_0(U, \mathbb{R}^d)$ corresponding to the kernel \mathbf{C} , with respect to the inner product \langle, \rangle . □

Remark 5. From a more abstract point of view, it is well-know that the covariance operator of a Gaussian measure defines an associated Hilbert structure [118].

In the sequel, we will use the space \mathcal{E} as the ambient Hilbert space to define the finite-dimensional sample-paths spaces as a nested sequence of RKHS. More precisely, let us write for \mathcal{E}_N the finite-dimensional subspace of \mathcal{E}

$$\mathcal{E}_N = \{ \mathbf{u} \in C_0(U, \mathbb{R}^d) \mid \mathcal{L}[\mathbf{u}] \in L_{\mathbf{f},N}^2 \} ,$$

with the space $L_{\mathbf{f},N}^2$ being defined as

$$L_{\mathbf{f},N}^2 = \text{span} \left[\{c_i(\boldsymbol{\phi}_{n,k})\}_{n,k \in \mathcal{I}_N, 0 \leq i < d} \right] .$$

We refer to such spaces as finite-dimensional approximation spaces, since we remark that

$$\mathcal{E}_N = \text{span} \left[\{c_i(\boldsymbol{\psi}_{n,k})\}_{n,k \in \mathcal{I}_N, 0 \leq i < d} \right] = \boldsymbol{\Psi}_N [\xi \Omega_N] ,$$

which means the space \mathcal{E}_N is made of the sample space of the finite dimensional process \mathbf{X}_N . The previous definition makes obvious the nested structure $\mathcal{E}_0 \subset \mathcal{E}_1 \subset \dots \subset \mathcal{E}$, and it is easy to characterize each space \mathcal{E}_N as reproducing Hilbert kernel space:

Proposition 16. *The Hilbert spaces $(\mathcal{E}_N, \langle, \rangle)$ are reproducing kernel Hilbert space (RKHS) with $\mathbb{R}^{d \times d}$ -valued reproducing kernel \mathbf{C}_N , the covariance function of the process \mathbf{X}_N .*

Proof. The proof this proposition follows the exact same argument as in the case of \mathcal{E} , but with the introduction of finite-dimensional kernels \mathbf{k}_N

$$\forall (t, s) \in U^2, \quad \mathbf{k}_N(t, s) = \sum_{(n,k) \in \mathcal{I}_N} \boldsymbol{\psi}_{n,k}(t) \cdot \boldsymbol{\phi}_{n,k}^T(s) ,$$

and the corresponding covariance function

$$\forall (t, s) \in [0, 1]^2, \quad \mathbf{C}_N(t, s) = \sum_{(n,k) \in \mathcal{I}_N} \boldsymbol{\psi}_{n,k}(t) \cdot \boldsymbol{\psi}_{n,k}^T(s) .$$

□

A.2 Finite-Dimensional Processes as Orthogonal Projections

The framework set in the previous section offers a new interpretation of our construction. Indeed, for all $N > 0$, the column of $\{\psi_{n,k}\}_{(n,k) \in \mathcal{I}_N}$ form an orthonormal basis of the space \mathcal{E}_N :

$$\mathcal{P}\langle \psi_{n,k}, \psi_{p,q} \rangle = \mathcal{P}(\mathcal{L}[\psi_{n,k}], \mathcal{L}[\psi_{p,q}]) = \mathcal{P}(\phi_{n,k}, \phi_{p,q}) = \delta_{p,q}^{n,k}.$$

This leads to define the finite-dimensional approximation \mathbf{x}_N of an sample path \mathbf{x} of \mathcal{E} as the orthogonal projection of \mathbf{x} on \mathcal{E}_N with respect to the inner product \langle, \rangle . At this point, it is worth remembering that the space \mathcal{E} is strictly contained in ${}_x\Omega'$ and does not coincide with ${}_x\Omega'$: actually one can easily show that $\mathbf{P}(\mathcal{E}) = \mathbf{0}$. We devote the rest of this section to define the finite-dimensional processes $\mathbf{Z}^N = \mathbb{E}_N[X]$ resulting from the conditioning on D_N , as path-wise orthogonal projection of the original process \mathbf{X} on the sample space \mathcal{E}_N .

Proposition 17. *For any $N > 0$, the conditioned processes $\mathbb{E}_N[\mathbf{X}]$ can be written as the orthogonal projection of \mathbf{X} on \mathcal{E}_N with respect to \langle, \rangle*

$$\mathbb{E}_N[\mathbf{X}] = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k} \cdot \mathcal{P}\langle \psi_{n,k}, \mathbf{X} \rangle.$$

The only hurdle to prove Proposition 17 is purely technical in the sense that the process \mathbf{X} lives in a larger space than \mathcal{E} : we need to find a way to extend the definition of \langle, \rangle so that the expression bears a meaning. Before answering this point quite straightforwardly, we need to establish the following lemma:

Lemma 5. *Writing the Gauss-Markov process $\mathbf{X}_t = \int_0^1 \mathbf{k}(t, s) d\mathbf{W}_s$, for all $N > 0$,*

the conditioned process $\mathbf{Z}^N = \mathbb{E}_N[\mathbf{X}]$ is expressed as the stochastic integral

$$\mathbf{Z}^N = \int_0^1 \mathbf{k}_N(t, s) d\mathbf{W}_s \quad \text{with} \quad \mathbf{k}_N(t, s) = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k}(t) \cdot \phi_{n,k}^T(s).$$

Proof. In the previous section, we have noticed that the kernel \mathbf{k}_N converges toward the kernel \mathbf{k} (Green function) in the L_f^2 sense:

$$\begin{aligned} \mathbf{k}(t, s) &= \sum_{(n,k) \in \mathcal{I}} \psi_{n,k}(t) \cdot \phi_{n,k}^T(s) \\ &= \lim_{N \rightarrow \infty} \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k}(t) \cdot \phi_{n,k}^T(s) \mathbf{k}_N \\ &= \lim_{N \rightarrow \infty} \mathbf{k}_N(t, s). \end{aligned}$$

This implies that the process \mathbf{X} as the stochastic integral, can also be written as

$$\mathbf{X}_t = \mathbf{g}(t) \int_U \mathbb{1}_{[0,t]}(s) \mathbf{f}(s) d\mathbf{W}_s = \int_0^1 \mathbf{k}(t, s) d\mathbf{W}_s = \lim_{N \rightarrow \infty} \int_0^1 \mathbf{k}_N(t, s) d\mathbf{W}_s.$$

Specifying the decomposition of \mathbf{k}_N , we can then naturally express \mathbf{X} as the convergent sum

$$\mathbf{X}_t = \sum_{(n,k) \in \mathcal{I}} \psi_{n,k} \cdot \Xi_{n,k} \quad \text{with} \quad \Xi_{n,k} = \int_0^1 \phi_{n,k}^T(s) d\mathbf{W}_s,$$

where the orthonormality property of the $\phi_{n,k}$ with respect to (\cdot, \cdot) , makes the vectors $\Xi_{n,k}$ appears as independent d -dimensional Gaussian variables of law $\mathcal{N}(0, \mathbf{I}_d)$. It is then easy to see that by definition of the elements $\psi_{n,k}$, for almost every ω in Ω , we then have

$$\forall N > 0, 0 \leq t \leq 1, \quad \mathbf{Z}^N(\omega) = \mathbb{E}_N[\mathbf{X}](\omega) = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k} \cdot \Xi_{n,k}(\omega),$$

and we finally recognize in the previous expression that for all $0 \leq t \leq 1$

$$\mathbf{Z}_t^N = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k} \cdot \Xi_{n,k} = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k}(t) \cdot \int_U \phi_{n,k}^T(s) d\mathbf{W}_s = \int_0^1 \mathbf{k}_N(t, s) d\mathbf{W}_s.$$

□

We can now proceed to justify the main result of Proposition 17:

Proof. The finite-dimensional processes \mathbf{Z}^N defined through Lemma 5 have sample-paths $t \mapsto \mathbf{Z}_t^N(\omega)$ belonging to \mathcal{E}_N . Moreover, for almost every ω in Ω , and for all n, k in \mathcal{I}_N ,

$$\begin{aligned} \mathcal{P}\langle \psi_{n,k}, \mathbf{Z}^N(\omega) \rangle &= \mathcal{P}\left\langle \psi_{n,k}, \int_0^1 \mathbf{k}_N(t, s) d\mathbf{W}_s(\omega) \right\rangle, \\ &= \mathcal{P}\left\langle \psi_{n,k}, \sum_{(p,q) \in \mathcal{I}_N} \psi_{p,q}(\omega) \cdot \int_0^1 \phi_{p,q}^T(s) d\mathbf{W}_s(\omega) \right\rangle, \\ &= \int_0^1 \phi_{n,k}^T(s) d\mathbf{W}_s(\omega), \end{aligned}$$

because of the orthonormality property of $\psi_{n,k}$ with respect to $\langle \cdot, \cdot \rangle$. As the previous equalities holds for every $N > 0$, the applications $\mathbf{x} \mapsto \mathcal{P}\langle \psi_{n,k}, \mathbf{x} \rangle$ can naturally be extended on ${}_x\Omega'$ by continuity. Therefore, it makes sense to write for all (n, k) in \mathcal{I}_N , $\mathcal{P}\langle \psi_{n,k}, \mathbf{Z}^N \rangle = \lim_{N \rightarrow \infty} \mathcal{P}\langle \psi_{n,k}, \mathbf{Z}^N \rangle \stackrel{\text{def}}{=} \mathcal{P}\langle \psi_{n,k}, \mathbf{X} \rangle$ even if the \mathbf{X} is defined into a larger sample space than \mathcal{E} . In other words, we have

$$\mathcal{P}\langle \psi_{n,k}, \mathbf{X} \rangle = \int_0^1 \phi_{n,k}^T(s) d\mathbf{W}_s = \Xi_{n,k},$$

and we can thus express the conditioned process $\mathbf{Z}^N = \mathbb{E}_N[\mathbf{X}]$ as the orthogonal projection of \mathbf{X} onto the finite sample-path \mathcal{E}_N by writing

$$\mathbf{Z}^N = \sum_{(n,k) \in \mathcal{I}_N} \psi_{n,k} \cdot \mathcal{P}\langle \psi_{n,k}, \mathbf{X} \rangle.$$

□

A.3 Optimality Criterion of the Sample Paths

Proposition 17 elucidates the structure of the conditioned processes \mathbf{Z}_N as path-wise orthogonal projections of \mathbf{X} on the finite-dimensional RKHS \mathcal{E}_N . It allows us to cast the finite sample-paths in a geometric setting and incidentally, to give a characterization of them as the minimizer of some functionals. In doing so, we shed a new light on well-known results of interpolation theory [241, 172, 97] and extend them to the multidimensional case.

The central point of this section reads as follows:

Proposition 18. *Given a function \mathbf{x} in \mathcal{E} , the function $\mathbf{x}_N = (\Psi \circ \Delta_N)[\mathbf{x}]$ belongs to \mathcal{E}_N and is defined by the following optimal criterion: \mathbf{x}_N is the only function in \mathcal{E} interpolating \mathbf{x} on D_N such that the functional*

$$\langle \mathbf{y}, \mathbf{y} \rangle = \|\mathcal{L}[\mathbf{y}](t)\|_2^2 = \int_0^1 |\mathcal{L}[\mathbf{y}](t)|_2^2 dt, \quad (\text{A.1})$$

takes its unique minimal value over \mathcal{E} in \mathbf{x}_N .

Proof. The space \mathcal{E}_N has been defined as $\mathcal{E}_N = \Psi_N[\xi\Omega_N] = \Psi \circ \Delta_N[\mathcal{E}]$, so that for all \mathbf{x} in \mathcal{E} , \mathbf{x}_N clearly belongs to \mathcal{E}_N . Moreover, \mathbf{x}_N interpolates \mathbf{x} on D_N : indeed, we know that the finite-dimensional operator Δ_N and Ψ_N^{-1} are inverse of each other $\Delta_N = \Psi_N^{-1}$, which entails that for all t in D_N

$$\mathbf{x}_N(t) = (\Psi \circ \Delta_N)[\mathbf{x}](t) = (\Psi_N \circ \Delta_N)[\mathbf{x}](t) = \mathbf{x}(t),$$

where we use the fact that for any ξ in ${}_\xi\Omega'$, and for all t in D_N , $\Psi_N[\xi](t) = \Psi[\xi](t)$ (recall that $\psi_{n,k}(t) = \mathbf{0}$ if $n > N$ and t belongs to D_N).

Let us now show that \mathbf{x}_N is determined in \mathcal{E} by the announced optimal criterion.

Suppose \mathbf{y} belongs to \mathcal{E} and interpolates \mathbf{x} on D_N and remark that we can write

$$\langle \mathbf{y}, \mathbf{y} \rangle = \|\mathcal{L}[\mathbf{y}]\|_2^2 = \|(\Phi \circ \Delta)[\mathbf{y}](t)\|_2^2 = \|\Delta[\mathbf{y}]\|_2^2,$$

since Φ is an isometry. Then, consider $\Delta[\mathbf{y}]$ in $l^2(\xi\Omega)$ and remark that, since for all (n, k) in \mathcal{I}_N , $\delta_{n,k}$ are Dirac measures supported by D_N , we have

$$\forall (n, k) \in \mathcal{I}_N, \quad \Delta_{n,k}[\mathbf{y}] = \mathcal{P}(\delta_{n,k}, \mathbf{y}) = \mathcal{P}(\delta_{n,k}, \mathbf{x}) = \Delta_{n,k}[\mathbf{x}] = \Delta_{n,k}[\mathbf{x}_N].$$

This entails

$$\|\Delta[\mathbf{y}]\|_2^2 dt = \sum_{(n,k) \in \mathcal{I}} |\Delta_{n,k}[\mathbf{y}]|_2^2 \geq \sum_{(n,k) \in \mathcal{I}_N} |\Delta_{n,k}[\mathbf{y}]|_2^2 = \|\Delta[\mathbf{x}_N]\|_2^2 dt.$$

since by definition of \mathbf{x}_N , $\delta_{n,k}[\mathbf{x}_N] = \mathbf{0}$ if $n > N$. Moreover, the minimum $\langle \mathbf{x}_N, \mathbf{x}_N \rangle$ is only attained for \mathbf{y} such that $\delta_{n,k}[\mathbf{y}] = \mathbf{0}$ if $n > N$ and $\delta_{n,k}[\mathbf{y}] = \delta_{n,k}[\mathbf{x}]$ if $n \leq N$, which defines univocally \mathbf{x}_N . This establishes that for all \mathbf{y} in \mathcal{E} such that for all t in D_N , $\mathbf{y}(t) = \mathbf{x}(t)$, we have $\langle \mathbf{x}_N, \mathbf{x}_N \rangle \leq \langle \mathbf{y}, \mathbf{y} \rangle$ and the equality case holds if and only if $\mathbf{y} = \mathbf{x}_N$. \square

Remark 6. When \mathcal{L} represents a regular differential operator of order d , $\sum_{i=1}^d a_i(t) D^i$ where $D = \frac{d}{dt}$, that is for

$$d\mathbf{X}_t = \boldsymbol{\alpha}(t) \cdot \mathbf{X}_t + \sqrt{\Gamma(t)} \cdot dW_t,$$

with

$$\boldsymbol{\alpha}(t) = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ a_d & a_{d-1} & \dots & a_1 \end{bmatrix}, \quad \sqrt{\Gamma(t)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

the finite-dimensional sample paths coincide exactly the spline interpolation of order $2d + 1$, which are well-known to satisfy the previous criterion [108]. This example will be further explored in the example section.

The Dirichlet energy simply appears as the squared norm induced on \mathcal{E} by the inner product \langle, \rangle , which in turn can be characterized as a Dirichlet quadratic form on \mathcal{E} . Actually, such a Dirichlet form can be used to define the Gauss-Markov process, extending the Gauss-Markov property to processes indexed on multidimensional spaces parameter [163]. In particular, for a n -dimensional parameter space, we can condition such Gauss-Markov processes on a smooth $n - 1$ -dimensional boundary. Within the boundary, the sample paths of the resulting conditioned process (the solution to the prediction problem in [163]) are the solutions to the corresponding Dirichlet problems for the elliptic operator associated with the Dirichlet form.

The characterization of the basis as the minimizer of such a Dirichlet energy (A.1) gives rise to an alternative method to compute the basis as the solution of a Dirichlet boundary value problem for an elliptic differential operator:

Proposition 19. *Let us assume that α and $\sqrt{\Gamma}$ are continuously differentiable and that $\sqrt{\Gamma}$ is invertible. Then the functions $\mu_{n,k}$ are defined as:*

$$\mu_{n,k}(t) = \begin{cases} \mu^l(t) & t \in [l_{n,k}, m_{n,k}] \\ \mu^r(t) & t \in [m_{n,k}, r_{n,k}] \\ 0 & \text{else,} \end{cases}$$

where μ^l and μ^r are the unique solutions of the second order d -dimensional linear differential equation

$$\mathbf{u}'' + (\Gamma^{-1} (\alpha^T \Gamma - \Gamma') - \alpha) \mathbf{u}' - (\Gamma^{-1} (\alpha^T \Gamma - \Gamma') \alpha + \alpha') \mathbf{u} = 0 \quad (\text{A.2})$$

with the following boundary value conditions:

$$\begin{cases} \boldsymbol{\mu}^l(l_{n,k}) = \mathbf{0} \\ \boldsymbol{\mu}^l(m_{n,k}) = \mathbf{I}_d \end{cases} \quad \begin{cases} \boldsymbol{\mu}^r(m_{n,k}) = \mathbf{I}_d \\ \boldsymbol{\mu}^r(r_{n,k}) = \mathbf{0} \end{cases}$$

Proof. By Proposition 18, we know that $\boldsymbol{\mu}_{n,k}(t)$ minimizes the convex functional

$$\int_0^1 |\mathcal{L}[\mathbf{u}](s)|_2^2 ds$$

over \mathcal{E} , being equal to zero outside the interval $[l_{n,k}, r_{n,k}]$ and equal to one at the point $t = m_{n,k}$. Because of the hypotheses on $\boldsymbol{\alpha}$ and $\sqrt{\Gamma}$, we have $\mathcal{L} = \mathcal{D}$ and we can additionally restrain our search to functions that are twice continuously differentiable. Incidentally, we only need to minimize separately the contributions on the interval $[l_{n,k}, m_{n,k}]$ and $[m_{n,k}, r_{n,k}]$. On both intervals, this problem is a classical Euler-Lagrange problem (see e.g. [7]) and is solved using basic principles of calculus of variations. We easily identify the Lagrangian of our problem as

$$\begin{aligned} L(t, \mathbf{u}, \mathbf{u}') &= \left| \left(\mathbf{u}' - \boldsymbol{\alpha}(t) \mathbf{u}(t) \right) \left(\sqrt{\Gamma(t)} \right)^{-1} \right|_2^2 \\ &= \left(\mathbf{u}'(t) - \boldsymbol{\alpha}(t) \mathbf{u}(t) \right)^T \left(\Gamma(t) \right)^{-1} \left(\mathbf{u}'(t) - \boldsymbol{\alpha}(t) \mathbf{u}(t) \right). \end{aligned}$$

From there, after some simple matrix calculations, the Euler-Lagrange equations

$$\frac{\partial L(t, \mathbf{u}, \mathbf{u}')}{\partial u_i} - \frac{d}{dt} \left(\frac{\partial L(t, \mathbf{u}, \mathbf{u}')}{\partial u'_i} \right) = 0, \quad i = 1, \dots, d.$$

can be expressed under the form:

$$\mathbf{u}'' + \left(\Gamma^{-1} (\boldsymbol{\alpha}^T \Gamma - \Gamma') - \boldsymbol{\alpha} \right) \mathbf{u}' - \left(\Gamma^{-1} (\boldsymbol{\alpha}^T \Gamma - \Gamma') \boldsymbol{\alpha} + \boldsymbol{\alpha}' \right) \mathbf{u} = \mathbf{0}$$

which ends the proof. □

Remark 7. It is a simple matter of calculus to check that the expression of $\boldsymbol{\mu}$ given in Proposition 1 satisfies equation (A.2). Notice also that in the case $\boldsymbol{\Gamma} = \boldsymbol{I}_d$, the differential equation becomes

$$\boldsymbol{u}'' + (\boldsymbol{\alpha}^T - \boldsymbol{\alpha}) \boldsymbol{u}' - (\boldsymbol{\alpha}^T \boldsymbol{\alpha} + \boldsymbol{\alpha}') \boldsymbol{u} = \mathbf{0},$$

which is further simplified for constant or symmetrical $\boldsymbol{\alpha}$.

Under the hypotheses of Proposition 19, we can thus define $\boldsymbol{\mu}_{n,k}$ as the unique solution to the second-order linear differential equation (A.2) with the appropriate boundary values conditions. From this definition, it is then easy to derive the basis $\boldsymbol{\psi}_{n,k}$ by completing the following program:

1. Compute the $t \mapsto \boldsymbol{\mu}_{n,k}(t)$ by solving the linear ordinary differential problem.
2. Apply the differential operator \mathcal{D} to get the functions $\mathcal{D}[\boldsymbol{\mu}_{n,k}]$.
3. Orthonormalize the column functions $t \mapsto c_j(\mathcal{D}[\boldsymbol{\mu}_{n,k}(t)])$ by Gram-Schmidt process.
4. Apply the integral operator \mathcal{K} to get the desired functions $\boldsymbol{\psi}_{n,k}$ (or equivalently multiply the original function $t \mapsto \boldsymbol{\mu}_{n,k}(t)$ by the corresponding Gram-Schmidt triangular matrix).

Notice finally that each of these points are easily implemented numerically.

Appendix B

Random Generator for Monte-Carlo Simulation

In this appendix, we introduce various distributions related to the first-passage time of the standard Wiener process with an affine barrier. We also explicit the random generators used to draw exactly and efficiently from these distributions. Despite being technical, these points are absolutely crucial to the implementation of the algorithm.

B.1 First-Passage Time of Drifted Wiener Process

For a Wiener process W with $W_0 = 0$ and x a strictly positive real, the first-hitting time of W to the constant barrier of height x is defined as

$$\tau_x = \inf\{t > 0 \mid W_t \geq x\},$$

and admits the following continuous probability density

$$q_x(\tau) = \frac{x}{\sqrt{2\pi\tau^3}} \exp\left(-\frac{x^2}{2\tau}\right).$$

Such a density is proper since it is normalized to one. However, we have $\mathbb{E}[\tau] = \infty$ and none of its moments are defined.

The distribution of first-passage times of W with a barrier of negative slope $L(t) = x - \nu t$ is the same as for a Wiener process with positive drift $W' = \{W'_t = W_t + \nu t\}$ and a constant barrier at x

$$\tau_{x,\nu} = \inf\{t > 0 \mid W_t \geq x - \nu t\} = \inf\{t > 0 \mid W'_t \geq x\}.$$

Through standard argument of stochastic calculus (see e.g. [106]), it can be shown that such a distribution admits the continuous density

$$\begin{aligned} q_{x,\nu}(\tau) &= p_x(\tau) \exp\left(\nu x - \frac{\nu^2 \tau}{2}\right), \\ &= \frac{x}{\sqrt{2\pi\tau^3}} \exp\left(-\frac{(x - \nu\tau)^2}{2\tau}\right). \end{aligned}$$

This distribution is well defined as long as $\nu > 0$ and so are all its moments. For example, we have:

$$\mathbb{E}[\tau] = \frac{x}{\nu} \quad \text{and} \quad \mathbb{V}[\tau] = \frac{x}{\nu^3}.$$

When the drift ν is zero, the distribution $q_x = q_{x,0}$ is known as the Wald distribution, whereas the family of functions $q_{x,\nu}$ are collectively referred as the inverse-gaussian distributions and we denote them $IG(x, \nu)$. We represent the shape of this family of distributions in Figure 3.1.

We are interested in the case for which the slope of the barrier is strictly positive ($\nu > 0$). In this situation, there exist a very efficient method to simulate numerically samples from the the distribution $IG(x, \nu)$. The method generates exactly random deviates by transformation of standard normal variable [148] and its implementation reads:

Procedure 7 *Generate a sample of $IG(x, \nu)$*

```

 $q \sim \mathcal{N}(0, 1), u \sim \mathcal{U}(0, 1)$ 
 $q \leftarrow q \cdot q/\nu, \mu \leftarrow x/\nu$ 
 $t \leftarrow \nu^{-1} \left( h + q/2 - \sqrt{q(x + q/4)} \right)$ 
if  $u \leq \mu/(\mu + t)$  then
  return  $t$ 
else
  return  $\mu \cdot \mu/t$ 
end if

```

B.2 Inverse-Gaussian Bridge

Given two strictly positive constants x and y , we consider the two first-passage times $\tau_{x,\nu}$ and $\tau_{x+y,\nu}$. We obviously have $\tau_{x,\nu} < \tau_{x+y,\nu}$ and these two random variables are clearly correlated. However, the random variables $\tau_{x+y,\nu} - \tau_{x,\nu}$ and $\tau_{x,\nu}$ are independent, a fact that is formulated saying that the process $\tau_\nu = \{\tau_{x,\nu}\}_{x>0}$ has independent increments. In particular, τ_ν has the Markov property, which allows to define the probability densities of the conditioned random variables $\{\tau_{x,\nu} \mid \tau_{x+y,\nu}\}$ as

$$\begin{aligned} q_{x,y,\nu}(\tau \mid \tau_{x+y,\nu} = t) &= \frac{q_{x,\nu}(\tau)q_{y,\nu}(t-\tau)}{q_{x+y,\nu}(t)}, \\ &= \frac{1}{\sqrt{2\pi}} \frac{xy}{x+y} \left(\frac{t}{\tau(t-\tau)} \right)^{\frac{3}{2}}, \\ &\quad \exp \left(-\frac{x^2}{2\tau} - \frac{y^2}{2(t-\tau)} + \frac{(x+y)^2}{2t} \right), \end{aligned}$$

These functions form a parametric family of functions known as the inverse-gaussian bridge distributions and denoted $IGB(x, y, t)$. Actually, assuming that the underlying Wiener process hits a barrier $s \mapsto l - \nu s$, $l > 0$ at a time t , the conditional first-passage process $(\tau_\nu \mid \tau_{\nu,l} = t) = \{\tau_{\nu,x} \mid \tau_{\nu,l} = t\}_{0 \leq x \leq l}$ can be defined formally and is called the inverse bridge process [245]. Because of the conditioning, its law does not depend on ν and, consistently, the density $q_{x,y,\nu}(\cdot \mid \tau_{x+y,\nu} = t)$ can be written $q_{x,y}(\cdot \mid t)$, dropping the index ν .

The family of distributions $IGB(x, y, t)$ exhibits an interesting behavior when its parameter are varied since it can be either unimodal or bimodal. We account for this behavior in Figure B.1. In the very same fashion as for the distributions $IG(x, \nu)$, there exists an exact numerical method to simulate random variates from the law of $IGB(x, y, t)$ [245]. It again proceeds by transformation of a standard normal variable:

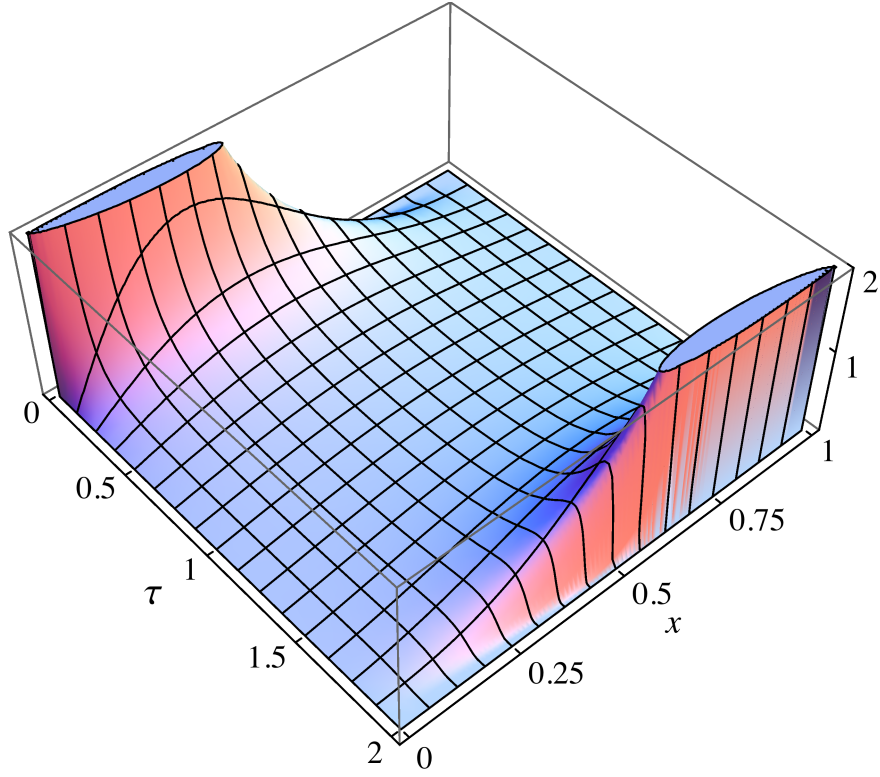


Figure B.1: Density function of the inverse Gaussian Bridge for $x + y = 1$ and $t = 2$.

Procedure 8 *Generate a sample of $IGB(x, y, s)$*

```

 $q \sim \mathcal{N}(0, 1), u \sim \mathcal{U}(0, 1)$ 
 $r \leftarrow y/x, p \leftarrow x \cdot y, q \leftarrow q \cdot q \cdot s$ 
 $t \leftarrow (x \cdot x)^{-1} \left( p + q/2 - \sqrt{q(p + q/4)} \right)$ 
if  $u \leq r(1 + t)((1 + r)(r + t))^{-1}$  then
    return  $s/(1 + t)$ 
else
     $t \leftarrow r \cdot r/t$ 
    return  $s/(1 + t)$ 
end if

```

B.3 First-Passage Brownian Bridge

Conditioning on $W_t = x < L(t)$, the probability that a Wiener process W_t hits the barrier $s \mapsto L(s) = l - \nu s$ before t , is given by

$$S_{t,\nu}(x) = \exp\left(-\frac{2l(l - \nu t - x)}{t}\right).$$

and $1 - S_{t,\nu}(x)$ is the probability that W , the Wiener process killed on the barrier, survives up to (x, t) . Accordingly, we can verify that $\kappa_t(x) = p_t(x)(1 - S_{t,\nu}(x))$ satisfies the Heat equation with absorbing boundary condition on the barrier L and initial condition $\kappa_0 = \delta$, which demonstrates that $\kappa_t(x)$ is the transition kernel of the killed Wiener process.

Even though the kernels $x \mapsto \kappa_t(x)$, $x < L(t)$ are non-gaussian kernels and define improper probability transitions on $(-\infty, L(t)]$, they satisfy the Markov property. Therefore, we can again exhibit the probability densities of the conditioned random variables $\{W_s \mid W_{t+s} = x, \tau_{\nu,l} > t + s\}$ as

$$\kappa_{s,t}(y \mid W_{s+t} = x, \tau_{\nu,l} > t + s) = \frac{\kappa_s(y)\kappa_t(x - y)}{\kappa_{s+t}(x)},$$

which is a proper density. Now letting x tend to $L(s+t)$, that is assuming the process W dies in $s+t$, the function $\kappa_{s,t}(\cdot \mid W_{t+s} = x, \tau_{\nu,l} > t + s)$ converges to a limit density

$$p_{s,t}(y \mid W_{s+t} = L(s+t)) \frac{(s+t)(l - \nu s - y)}{tl} \left(1 - \exp\left(-\frac{2l(l - \nu s - y)}{s}\right)\right),$$

where the function $p_{s,t}(\cdot \mid W_{s+t} = L(s+t))$ denotes the density of the Brownian bridge $(W_s \mid W_{s+t})$. The preceding limit function represents the density of the probability that the Wiener process belongs to $[y, y + dy)$ knowing that it first hits L in $t + s$. We denote such a density $\kappa_{s,t}(y \mid \tau_{\nu,l} = s + t)$ and the corresponding distributions as $FPB(s, t, \nu, l)$ (see Figure B.2).

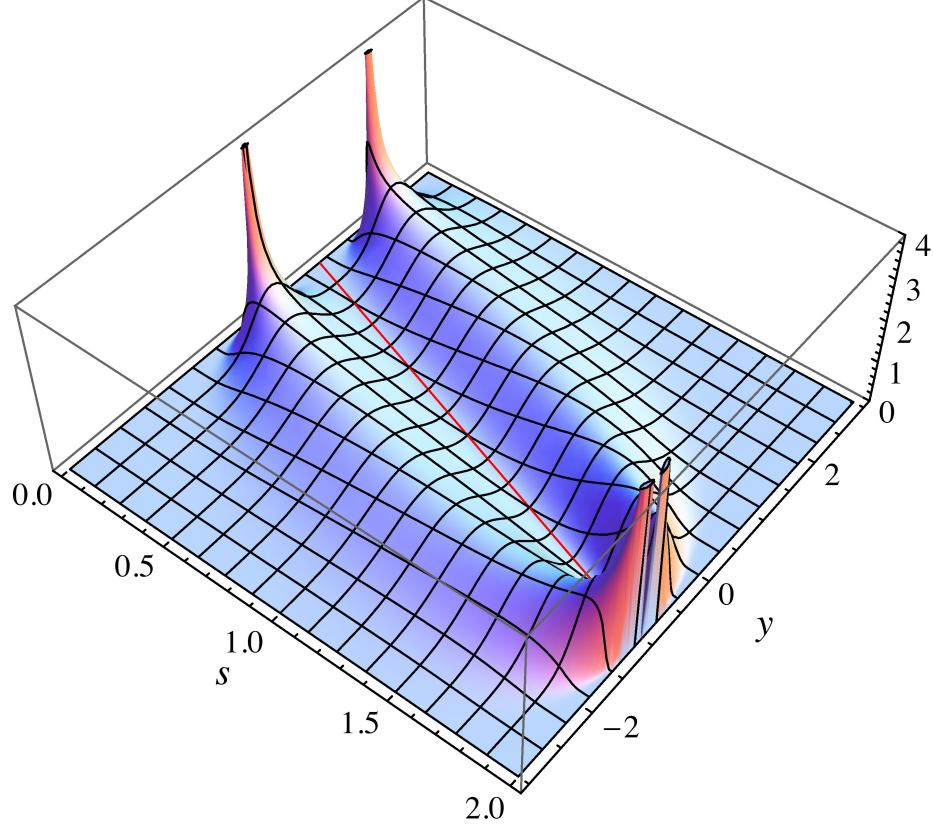


Figure B.2: Density function of the first-passage Brownian motion for a barrier $t \mapsto 1 - t$ and conditioning to $\tau = 2$.

The process $(W \mid \tau_{\nu,l} = t) = \{W_s \mid \tau_{\nu,l} = t\}_{0 \leq s \leq t}$ can be defined formally and is called the first-passage Brownian bridge [21]. The rigorous treatment of such a process is however beyond the scope of the present Appendix. We just mention that the law of the first-passage bridge $(W \mid \tau_{\nu,l} = t)$ has the same law as $(L(s) - B_s \mid B_0 = l, B_t = 0)$, where B denotes the 3-dimensional Bessel process. Since a 3-dimensional Bessel process B is defined as $B = \|\mathbf{W}\|_2$, the Euclidean norm of a 3-dimensional Wiener process \mathbf{W} , it is straightforward to simulate random deviates from $FPB(s, t, \nu, l)$ when equipped with a method to simulate the law $BES^3(s, t, x)$ of the 3-dimensional Bessel bridge $(B_s \mid B_0 = x, B_t = 0)$:

Procedure 9 Generate a sample of $BES^3(s, t, x)$

$q_1 \sim \mathcal{N}(0, 1), q_2 \sim \mathcal{N}(0, 1), q_3 \sim \mathcal{N}(0, 1)$
 $r \leftarrow \sqrt{s \cdot t / (s + t)}$
 $q_1 \leftarrow r \cdot q_1 - t \cdot x / (s + t), q_2 \leftarrow r \cdot q_2, q_3 \leftarrow r \cdot q_3$
return $\sqrt{q_1 \cdot q_1 + q_2 \cdot q_2 + q_3 \cdot q_3}$

Bibliography

- [1] L. F. Abbott and Sacha B. Nelson. Synaptic plasticity: taming the beast. *Nature Neurosciense*, 3:1178 – 1183, 2000.
- [2] M Abeles. *Corticonics: Neural Circuits of the Cerebral Cortex*. Cambridge University Press, 1991.
- [3] M. Abeles, H. Bergman, E. Margalit, and E. Vaadia. Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, 70(4):1629–1638, 1993.
- [4] Mario Abundo. Limit at zero of the first-passage time density and the inverse problem for one-dimensional diffusions. *Stochastic Analysis and Applications*, 24(6):1119–1145, 2011/08/31 2006.
- [5] Yashar Ahmadian, Jonathan W. Pillow, and Liam Paninski. Efficient markov chain monte carlo methods for decoding neural spike trains. *Neural Computation*, 23(1):46–96, 2011/08/17 2010.
- [6] L. Alili, P. Patie, and J. L. Pedersen. Representations of the first hitting time density of an Ornstein-Uhlenbeck process. *Stoch. Models*, 21(4):967–980, 2005.
- [7] Grégoire Allaire. *Numerical analysis and optimization*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2007. An intro-

duction to mathematical modelling and numerical simulation, Translated from the French by Alan Craig.

- [8] JoséM. Amigó, Janusz Szczepański, Elek Wajnryb, and Maria V. Sanchez-Vives. Estimating the entropy rate of spike trains via lempel-ziv complexity. *Neural Computation*, 16(4):717–736, 2004.
- [9] Blaise Agüera y Arcas and Adrienne L. Fairhall. What causes a neuron to spike? *Neural Computation*, 15(8):1789–1807, 2003.
- [10] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.
- [11] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.
- [12] Craig A. Atencio, Tatyana O. Sharpee, and Christoph E. Schreiner. Hierarchical computation in the canonical auditory cortical circuit. *Proceedings of the National Academy of Sciences*, 106(51):21894–21899, 2009.
- [13] Inbal Ayzenshtat, Elhanan Meirovithz, Hadar Edelman, Uri Werner-Reiss, Elie Bienenstock, Moshe Abeles, and Hamutal Slovin. Precise spatiotemporal patterns among visual cortical areas and their relation to visual stimulus processing. *Journal of Neuroscience*, 30(33):11232–11245, 2010.
- [14] L Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure, Sér. 3*, 3:21–86, 1900.
- [15] Wyeth Bair and Christof Koch. Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Computation*, 8(6):1185–1202, 1996.

- [16] Paolo Baldi. Exact asymptotics for the probability of exit from a domain and applications to simulation. *Ann. Probab.*, 23(4):1644–1670, 1995.
- [17] Paolo Baldi and Lucia Caramellino. Asymptotics of hitting probabilities for general one-dimensional pinned diffusions. *Ann. Appl. Probab.*, 12(3):1071–1095, 2002.
- [18] Paolo Baldi, Lucia Caramellino, and Maria Gabriella Iovino. Pricing complex barrier options with general features using sharp large deviation estimates. In *Monte Carlo and quasi-Monte Carlo methods 1998 (Claremont, CA)*, pages 149–162. Springer, Berlin, 2000.
- [19] Jingdong Bao, Yasuhisa Abe, and Yizhong Zhuo. An integral algorithm for numerical integration of one-dimensional additive colored noise problems. *Journal of Statistical Physics*, 90(3):1037–1045, 1998.
- [20] Michael J. Berry, David K. Warland, and Markus Meister. The structure and precision of retinal spike trains. *Proceedings of the National Academy of Sciences*, 94(10):5411–5416, 1997.
- [21] Jean Bertoin, Loïc Chaumont, and Jim Pitman. Path transformations of first passage bridges. *Electron. Comm. Probab.*, 8:155–166 (electronic), 2003.
- [22] W. Bialek and F. Rieke. Reliability and information transmission in spiking neurons. *Trends in Neurosciences*, 15(11):428–434, November 1992.
- [23] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [24] Vladimir I. Bogachev. *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.

- [25] Alexander Borst and Juergen Haag. Effects of mean firing on neural information rate. *Journal of Computational Neuroscience*, 10(2):213–221, March 2001.
- [26] Nicolas Bouleau and Dominique Lépingle. *Numerical methods for stochastic processes*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York, 1994. A Wiley-Interscience Publication.
- [27] V. Braitenberg and A. Schüz. *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer, Berlin, Germany, 1998. ISBN: 3-540-63816-4.
- [28] Romain Brette and Emmanuel Guigon. Reliability of spike timing is a general property of spiking model neurons. *Neural computation*, 15(2):279–308, February 2003.
- [29] Romain Brette, Michelle Rudolph, Ted Carnevale, Michael Hines, David Beman, James Bower, Markus Diesmann, Abigail Morrison, Philip Goodman, Frederick Harris, Milind Zirpe, Thomas Natschläger, Dejan Pecevski, Bard Ermentrout, Mikael Djurfeldt, Anders Lansner, Olivier Rochel, Thierry Vieville, Eilif Muller, Andrew Davison, Sami El Boustani, and Alain Destexhe. Simulation of networks of spiking neurons: A review of tools and strategies. *Journal of Computational Neuroscience*, 23(3):349–398, 2007.
- [30] Emery N Brown, Robert E Kass, and Partha P Mitra. Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nat Neurosci*, 7(5):456–461, 2004.
- [31] Randy M. Bruno and Bert Sakmann. Cortex is driven by weak but synchronously active thalamocortical synapses. *Science*, 312(5780):1622–1627, 2006.

- [32] A. Buonocore, L. Caputo, E. Pirozzi, and L. M. Ricciardi. On a stochastic leaky integrate-and-fire neuronal model. *Neural Computation*, 22(10):2558–2585, 2010.
- [33] A. Buonocore, A. G. Nobile, and L. M. Ricciardi. A new integral equation for the evaluation of first-passage-time probability densities. *Adv. in Appl. Probab.*, 19(4):784–800, 1987.
- [34] A. Burkitt. A review of the integrate-and-fire neuron model: 1. homogeneous synaptic input. *Biological Cybernetics*, 95(1):1–19, July 2006.
- [35] A. Burkitt. A review of the integrate-and-fire neuron model: 2. inhomogeneous synaptic input and network properties. *Biological Cybernetics*, 95(2):97–112, 2006.
- [36] Daniel A. Butts, Chong Weng, Jianzhong Jin, Chun-I Yeh, Nicholas A. Lesica, Jose-Manuel Alonso, and Garrett B. Stanley. Temporal precision in the neural code and the timescales of natural vision. *Nature*, 449(7158):92–95, 09 2007.
- [37] R. H. Cameron and W. T. Martin. Transformations of Wiener integrals under translations. *Ann. of Math. (2)*, 45:386–396, 1944.
- [38] John Rozier Cannon. *The one-dimensional heat equation*, volume 23 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1984. With a foreword by Felix E. Browder.
- [39] Natalia Caporale and Yang Dan. Spike timing-dependent plasticity: A hebbian learning rule. *Annual Review of Neuroscience*, 31(1):25–46, 2011/09/06 2008.
- [40] Guillermo A. Cecchi and Marcelo O. Magnasco. Negative resistance and rectification in Brownian transport. *Phys. Rev. Lett.*, 76(11):1968–1971, Mar 1996.

- [41] Guillermo A. Cecchi, Mariano Sigman, José-Manuel Alonso, Luis Martínez, Dante R. Chialvo, and Marcelo O. Magnasco. Noise in neurons is message dependent. *Proceedings of the National Academy of Sciences*, 97(10):5557–5561, 05 2000.
- [42] E.J. Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, 12(2):199–213, 2001.
- [43] Z. Ciesielski. Hölder conditions for realizations of Gaussian processes. *Trans. Amer. Math. Soc.*, 99:403–413, 1961.
- [44] Simona Cocco, Stanislas Leibler, and Rémi Monasson. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. *Proceedings of the National Academy of Sciences*, 106(33):14058–14062, 2009.
- [45] R. Courant and D. Hilbert. *Methods of mathematical physics. Vol. II: Partial differential equations.* (Vol. II by R. Courant.). Interscience Publishers (a division of John Wiley & Sons), New York-London, 1962.
- [46] Thomas M. Cover and Joy A. Thomas. *Elements of information theory.* Wiley-Interscience, New York, NY, USA, 1991.
- [47] Marshall Crumiller, Bruce Knight, Yunguo Yu, and Ehud Kaplan. Estimating the amount of information conveyed by a population of neurons. *Frontiers in Neuroscience*, 5(0), 2011.
- [48] W. Dahmen, B. Han, R.-Q. Jia, and A. Kunoth. Biorthogonal multiwavelets on the interval: cubic Hermite splines. *Constr. Approx.*, 16(2):221–259, 2000.
- [49] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes.*

- Vol. I. Probability and its Applications* (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.
- [50] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II. Probability and its Applications* (New York). Springer, New York, second edition, 2008. General theory and structure.
- [51] K. Daoudi, J. Lévy Véhel, and Y. Meyer. Construction of continuous functions with prescribed local regularity. *Constr. Approx.*, 14(3):349–385, 1998.
- [52] D. A. Darling and A. J. F. Siebert. The first passage problem for a continuous Markov process. *Ann. Math. Statistics*, 24:624–639, 1953.
- [53] Peter Dayan and LF Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, 2001.
- [54] Carl de Boor. *A practical guide to splines*, volume 27 of *Applied Mathematical Sciences*. Springer-Verlag, New York, revised edition, 2001.
- [55] Carl de Boor, Christian Gout, Angela Kunoth, and Christophe Rabut. Multivariate approximation: theory and applications. An overview. *Numer. Algorithms*, 48(1-3):1–9, 2008.
- [56] Jaime de la Rocha, Brent Doiron, Eric Shea-Brown, Kresimir Josic, and Alex Reyes. Correlation between neural spike trains increases with firing rate. *Nature*, 448(7155):802–806, 08 2007.
- [57] Michael R. DeWeese, Michael Wehr, and Anthony M. Zador. Binary spiking in auditory cortex. *The Journal of Neuroscience*, 23(21):7940–7949, 2003.
- [58] Charles R Doering, Khachik V Sargsyan, Leonard M Sander, and Eric Vandenberg. Asymptotics of rare events in birth–death processes bypassing the exact solutions. *Journal of Physics: Condensed Matter*, 19(6):065145, 2007.

- [59] C. L. Dolph and M. A. Woodbury. On the relation between Green's functions and covariances of certain stochastic processes and its application to unbiased linear prediction. *Trans. Amer. Math. Soc.*, 72:519–550, 1952.
- [60] J. L. Doob. Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statistics*, 20:393–403, 1949.
- [61] J. Durbin. The first-passage density of a continuous Gaussian process to a general boundary. *J. Appl. Probab.*, 22(1):99–122, 1985.
- [62] J. Durbin. The first-passage density of the Brownian motion process to a curved boundary. *J. Appl. Probab.*, 29(2):291–304, 1992. With an appendix by D. Williams.
- [63] Günter Ehret. New perspectives of information transformation through the auditory cortical layers. *Proceedings of the National Academy of Sciences*, 106(51):21463–21464, 2009.
- [64] Monty A. Escabí, Reza Nassiri, Lee M. Miller, Christoph E. Schreiner, and Heather L. Read. The contribution of spike threshold to acoustic feature selectivity, spike information content, and information throughput. *The Journal of Neuroscience*, 25(41):9524–9534, 10 2005.
- [65] A. Aldo Faisal, Luc P. J. Selen, and Daniel M. Wolpert. Noise in the nervous system. *Nat Rev Neurosci*, 9(4):292–303, 04 2008.
- [66] J. M. Fellous, M. Rudolph, A. Destexhe, and T. J. Sejnowski. Synaptic background noise controls the input/output characteristics of single cells in an in vitro model of in vivo activity. *Neuroscience*, 122(3):811–829, 2003.
- [67] Jean-Marc Fellous, Paul H. E. Tiesinga, Peter J. Thomas, and Terrence J.

- Sejnowski. Discovering spike patterns in neuronal responses. *The Journal of Neuroscience*, 24(12):2989–3001, 2004.
- [68] Stefano Ferraina, Martin Paré, and Robert H. Wurtz. Comparison of cortico-cortical and cortico-collicular signals for the generation of saccadic eye movements. *Journal of Neurophysiology*, 87(2):845–858, 02 2002.
- [69] Jozsef Fiser, Chiayu Chiu, and Michael Weliky. Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature*, 431(7008):573–578, 09 2004.
- [70] Nicolas Fourcaud-Trocmé, David Hansel, Carl van Vreeswijk, and Nicolas Brunel. How spike generation mechanisms determine the neuronal response to fluctuating inputs. *The Journal of Neuroscience*, 23(37):11628–11640, 12 2003.
- [71] Matthias O. Franz and Bernhard Schölkopf. A unifying view of Wiener and Volterra theory and polynomial kernel regression. *Neural Computation*, 18(12):3097–3118, 2006.
- [72] Masatoshi Fukushima, Yōichi Ōshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1994.
- [73] J. G. Gaines and T. J. Lyons. Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.*, 57(5):1455–1484, 1997.
- [74] Roberto F. Galán, G. Bard Ermentrout, and Nathaniel N. Urban. Optimal time scale for spike-time reliability: Theory, simulations, and experiments. *Journal of Neurophysiology*, 99(1):277–283, 01 2008.

- [75] Roberto F. Galán, Nicolas Fourcaud-Trocmé, G. Bard Ermentrout, and Nathaniel N. Urban. Correlation-induced synchronization of oscillations in olfactory bulb neurons. *The Journal of Neuroscience*, 26(14):3646–3655, 04 2006.
- [76] Wulfram Gerstner and Werner Kistler. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, 08 2002.
- [77] Wulfram Gerstner and Richard Naud. How good are neuron models? *Science*, 326(5951):379–380, 10 2009.
- [78] Maurice. Gevrey. *Sur les équations aux dérivées partielles du type parabolique*. Gauthier-Villars, Paris, 1913.
- [79] Daniel T. Gillespie. Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Phys. Rev. E*, 54(2):2084–2091, Aug 1996.
- [80] Maria Teresa Giraudo, Laura Sacerdote, and Cristina Zucca. A Monte Carlo method for the simulation of first passage times of diffusion processes. *Methodol. Comput. Appl. Probab.*, 3(2):215–231, 2001.
- [81] I. V. Girsanov. On transforming a class of stochastic processes by absolutely continuous substitution of measures. *Teor. Veroyatnost. i Primenen.*, 5:314–330, 1960.
- [82] Emmanuel Gobet. Weak approximation of killed diffusion using euler schemes. *Stochastic Processes and their Applications*, 87(2):167 – 197, 2000.
- [83] Emmanuel Gobet and Stéphane Menozzi. Exact approximation rate of killed hypoelliptic diffusions using the discrete euler scheme. *Stochastic Processes and their Applications*, 112(2):201 – 223, 2004.
- [84] Malcolm Goldman. On the first passage of the integrated Wiener process. *Ann. Mat. Statist.*, 42:2150–2155, 1971.

- [85] Dan F. M. Goodman and Romain Brette. Spike-timing-based computation in sound localization. *PLoS Comput Biol*, 6(11):e1000993–, 11 2010.
- [86] Rudolf Gorenflo, Francesco Mainardi, and Alessandro Vivoli. Continuous-time random walk and parametric subordination in fractional diffusion. *Chaos, Solitons & Fractals*, 34(1):87 – 103, 2007. In Search of a Theory of Complexity.
- [87] Peter D. Grünwald. *The Minimum Description Length Principle (Adaptive Computation and Machine Learning)*. The MIT Press, 2007.
- [88] R. Gutiérrez, L. M. Ricciardi, P. Román, and F. Torres. First-passage-time densities for time-non-homogeneous diffusion processes. *J. Appl. Probab.*, 34(3):623–631, 1997.
- [89] Robert Gütig, Ad Aertsen, and Stefan Rotter. Statistical significance of coincident spikes: Count-based versus rate-based statistics. *Neural Computation*, 14(1):121–153, 2002.
- [90] Olle Häggström. *Finite Markov chains and algorithmic applications*, volume 52 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2002.
- [91] Onésimo Hernández-Lerma and Jean Bernard Lasserre. *Markov chains and invariant probabilities*, volume 211. Birkhäuser Verlag, Basel, 2003.
- [92] Takeyuki Hida. Canonical representations of Gaussian processes and their applications. *Mem. Coll. Sci. Univ. Kyoto. Ser. A. Math.*, 33:109–155, 1960/1961.
- [93] Bertil Hille. *Ion Channels of Excitable Membranes*. Sinauer Associates, 07 2001.
- [94] Rebecca L. Honeycutt. Stochastic runge-kutta algorithms. i. white noise. *Phys. Rev. A*, 45(2):600–603, Jan 1992.

- [95] Christian Houdre and Victor Perez-Abreu. *Chaos Expansions, Multiple Wiener-Ito Integrals, and Their Applications*. CRC Press, April 5, 1994.
- [96] Toms Hromdka and Anthony M Zador. Representations in auditory cortex. *Current Opinion in Neurobiology*, 19(4):430–433, 2009.
- [97] Tailen Hsing and Haobo Ren. An RKHS formulation of the inverse regression dimension-reduction problem. *Ann. Statist.*, 37(2):726–755, 2009.
- [98] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.
- [99] Daniel Huber, Leopoldo Petreanu, Nima Ghitani, Sachin Ranade, Tomas Hromadka, Zach Mainen, and Karel Svoboda. Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. *Nature*, 451(7174):61–64, 2008.
- [100] Kiyosi Itô and Henry P. McKean, Jr. *Diffusion processes and their sample paths*. Springer-Verlag, Berlin, 1974. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125.
- [101] E. M. Izhikevich. Which model to use for cortical spiking neurons? *Neural Networks, IEEE Transactions on*, 15(5):1063–1070, September 2004.
- [102] Eugene M. Izhikevich. Simple model of spiking neurons. *IEEE Trans. Neural Networks*, pages 1569–1572, 2003.
- [103] Eugene M. Izhikevich. Polychronization: Computation with spikes. *Neural Computation*, 18(2):245–282, 2006.
- [104] S. Jaimungal, A. Kreinin, and A. Valov. Integral equations and the first passage time of Brownian motions. *ArXiv e-prints*, February 2009.

- [105] M. Kac and A. J. F. Siegert. An explicit representation of a stationary Gaussian process. *Ann. Math. Statistics*, 18:438–442, 1947.
- [106] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [107] Alan F. Karr. Weak convergence of a sequence of markov chains. *Probability Theory and Related Fields*, 33:41–48, 1975. 10.1007/BF00539859.
- [108] George Kimeldorf and Grace Wahba. Some results on Tchebycheffian spline functions. *J. Math. Anal. Appl.*, 33:82–95, 1971.
- [109] George S. Kimeldorf and Grace Wahba. A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. *Ann. Math. Statist.*, 41:495–502, 1970.
- [110] George S. Kimeldorf and Grace Wahba. Spline functions and stochastic processes. *Sankhyā Ser. A*, 32:173–180, 1970.
- [111] Werner M. Kistler and Wulfram Gerstner. Stable propagation of activity pulses in populations of spiking neurons. *Neural Computation*, 14(5):987–997, 2002.
- [112] D. J. Klein, D. A. Depireux, J. Z. Simon, and S. A. Shamma. Robust spectrotemporal reverse correlation for the auditory system: Optimizing stimulus design. *Journal of Computational Neuroscience*, 9(1):85–111, 2000.
- [113] BW Knight. Dynamics of encoding in a population of neurons. *The Journal of general physiology*, 59(6):734–766, 06 1972.
- [114] A. Kolmogoroff. Über die analytischen methoden in der wahrscheinlichkeit-rechnung. *Mathematische Annalen*, 104:415–458, 1931. 10.1007/BF01457949.

- [115] AN Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4:83–91, 1933.
- [116] Arvind Kumar, Stefan Rotter, and Ad Aertsen. Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *Journal of Neuroscience*, 28(20):5268–5280, 2008.
- [117] Arvind Kumar, Stefan Rotter, and Ad Aertsen. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nat Rev Neurosci*, 11(9):615–627, 2010.
- [118] Hui Hsiung Kuo. *Gaussian measures in Banach spaces*. Lecture Notes in Mathematics, Vol. 463. Springer-Verlag, Berlin, 1975.
- [119] M. Kurzyński and P. Chelminiak. Mean first-passage time in the stochastic theory of biochemical processes. application to actomyosin molecular motor. *Journal of Statistical Physics*, 110(1):137–181, 2003-01-21.
- [120] M. Labou. Solution of the first-passage problem by advanced monte carlo simulation technique. *Strength of Materials*, 35(6):588–593, 2003-11-01.
- [121] Petr Lansky and Susanne Ditlevsen. A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models. *Biol. Cybern.*, 99(4-5):253–262, 2008.
- [122] L. Lapicque. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarization. *Journal de physiologie et de pathologie générale*, 9:620–635, 1907.
- [123] Eric Larson, Ben P. Perrone, Kamal Sen, and Cyrus P. Billimoria. A robust and biologically plausible spike pattern recognition network. *The Journal of Neuroscience*, 30(46):15566–15572, 2010.

- [124] Mario Lefebvre and Éric Leonard. On the first hitting place of the integrated Wiener process. *Adv. in Appl. Probab.*, 21(4):945–948, 1989.
- [125] Axel Lehmann. Smoothness of first passage time distributions and a new integral equation for the first passage time density of continuous Markov processes. *Adv. in Appl. Probab.*, 34(4):869–887, 2002.
- [126] Paul Lévy. *Processus Stochastiques et Mouvement Brownien. Suivi d’une note de M. Loève*. Gauthier-Villars, Paris, 1948.
- [127] Tiger W. Lin and George N. Reeke. A continuous entropy rate estimator for spike trains using a k-means-based context tree. *Neural Computation*, 22(4):998–1024, 2011/08/29 2009.
- [128] B. Lindner, J. Garca-Ojalvo, A. Neiman, and L. Schimansky-Geier. Effects of noise in excitable systems. *Physics Reports*, 392(6):321–424, 3 2004.
- [129] Peter Linz. *Analytical and numerical methods for Volterra equations*, volume 7 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1985.
- [130] Michael London, Arnd Roth, Lisa Beeren, Michael Hausser, and Peter E. Latham. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature*, 466(7302):123–127, 07 2010.
- [131] Terry Lyons. Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.*, 1(4):451–464, 1994.
- [132] Marcin Magdziarz, Aleksander Weron, and Karina Weron. Fractional fokker-planck dynamics: Stochastic representation and computer simulation. *Phys. Rev. E*, 75(1):016708, Jan 2007.

- [133] ZF Mainen and TJ Sejnowski. Reliability of spike timing in neocortical neurons. *Science*, 268(5216):1503–1506, 06 1995.
- [134] Stephane G. Mallat. Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbf{R})$. *Trans. Amer. Math. Soc.*, 315(1):69–87, 1989.
- [135] Benoit Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman, 1982.
- [136] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.
- [137] Stephan L. Marguet and Kenneth D. Harris. State-dependent representation of amplitude-modulated noise stimuli in rat auditory cortex. *The Journal of Neuroscience*, 31(17):6414–6420, 04 2011.
- [138] Panos Z. Marmarelis and Ken-Ichi Naka. White-noise analysis of a neuron chain: An application of the Wiener theory. *Science*, 175(4027):1276–1278, 1972.
- [139] Laura Martignon, Gustavo Deco, Kathryn Laskey, Mathew Diamond, Winrich Freiwald, and Eilon Vaadia. Neural coding: Higher-order temporal patterns in the neurostatistics of cell assemblies. *Neural Computation*, 12(11):2621–2653, 2000.
- [140] Nathaniel F. G. Martin and James W. England. *Mathematical theory of entropy*, volume 12 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1981. With a foreword by James K. Brooks.
- [141] Servet Martínez and Jaime San Martín. Classification of killed one-dimensional diffusions. *Ann. Probab.*, 32(1A):530–552, 2004.
- [142] Maurizio Mattia and Paolo Del Giudice. Efficient event-driven simulation of large networks of spiking neurons and dynamical synapses. *Neural Computation*, 12(10):2305–2329, 2000.

- [143] Simoncelli E McDermott JH, Oxenham AJ. Sound texture synthesis via filter statistics. *Applications of Signal Processing to Audio and Acoustics*, pages 297 – 300, 2009.
- [144] Mark D. McDonnell and Lawrence M. Ward. The benefits of noise in neural systems: bridging theory and experiment. *Nat Rev Neurosci*, 12(7):415–426, 07 2011.
- [145] H. P. McKean, Jr. A winding problem for a resonator driven by a white noise. *J. Math. Kyoto Univ.*, 2:227–235, 1963.
- [146] Nicholas Metropolis and S. Ulam. The Monte Carlo method. *J. Amer. Statist. Assoc.*, 44:335–341, 1949.
- [147] Yves Meyer, Fabrice Sellan, and Murad S. Taqqu. Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. *J. Fourier Anal. Appl.*, 5(5):465–494, 1999.
- [148] John R. Michael, William R. Schucany, and Roy W. Haas. Generating random variates using transformations with multiple roots. *The American Statistician*, 30(2):88–90, 1976.
- [149] Remi Monasson and Simona Cocco. Fast inference of interactions in assemblies of stochastic integrate-and-fire neurons from spike recordings. *Journal of Computational Neuroscience*, 31:199–227, 2011. 10.1007/s10827-010-0306-8.
- [150] Abigail Morrison, Sirko Straube, Hans Ekkehard Plesser, and Markus Diesmann. Exact subthreshold integration with continuous spike times in discrete-time neural network simulations. *Neural Computation*, 19(1):47–79, 2006.
- [151] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical*

- and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.
Reprint of 1997 original.
- [152] Bruno A Olshausen and David J Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487, 2004.
 - [153] A. Omurtag, B. W. Knight, and L. Sirovich. On the simulation of large populations of neurons. *Journal of Computational Neuroscience*, 8(1):51–63, 2000-01-01.
 - [154] G. Palm and T. Poggio. The volterra representation and the Wiener expansion: Validity and pitfalls. *SIAM Journal on Applied Mathematics*, 33(2):195–216, 1977.
 - [155] Liam Paninski. Estimation of entropy and mutual information. *Neural Computation*, 15(6):1191–1253, 2011/08/31 2003.
 - [156] Liam Paninski, Adrian Haith, and Gabor Szirtes. Integral equation methods for computing likelihoods and their derivatives in the stochastic integrate-and-fire model. *Journal of Computational Neuroscience*, 24:69–79, 2008. 10.1007/s10827-007-0042-x.
 - [157] Liam Paninski, Jonathan Pillow, and Eero Simoncelli. Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. *Neural Computation*, 16(12):2533–2561, 12 2004.
 - [158] C. Park and S. R. Paranjape. Probabilities of Wiener paths crossing differentiable curves. *Pacific J. Math.*, 53:579–583, 1974.
 - [159] C. Park and F. J. Schuurmann. Evaluations of barrier-crossing probabilities of Wiener paths. *J. Appl. Probability*, 13(2):267–275, 1976.

- [160] Goran Peskir. On integral equations arising in the first-passage problem for Brownian motion. *J. Integral Equations Appl.*, 14(4):397–423, 2002.
- [161] Goran Peskir. On the American option problem. *Math. Finance*, 15(1):169–181, 2005.
- [162] Gordon Pipa and Sonja Grn. Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing*, 52-54:31–37, 2003.
- [163] Loren D. Pitt. A Markov property for Gaussian processes with a multidimensional parameter. *Archive for Rational Mechanics and Analysis*, 43(5):367–391, 1971.
- [164] Eckhard Platen. An introduction to numerical methods for stochastic differential equations. Research Paper Series 6, Quantitative Finance Research Centre, University of Technology, Sydney, April 1999.
- [165] Hans E. Plesser and Theo Geisel. Markov analysis of stochastic resonance in a periodically driven integrate-and-fire neuron. *Phys. Rev. E*, 59(6):7008–7017, Jun 1999.
- [166] Tomaso Poggio and David Marr. From understanding computation to understanding neural circuitry. *Neurosciences Res. Prog. Bull.*, 15:470–488, 1977.
- [167] Javier Portilla and Eero P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1):49–70, 2000.
- [168] Klaus Pötzelberger and Liqun Wang. Boundary crossing probability for Brownian motion. *J. Appl. Probab.*, 38(1):152–164, 2001.
- [169] Philip E. Protter. *Stochastic integration and differential equations*, volume 21

- of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.
- [170] Yifat Prut, Eilon Vaadia, Hagai Bergman, Iris Haalman, Hamutal Slovin, and Moshe Abeles. Spatiotemporal structure of cortical activity: Properties and behavioral relevance. *Journal of Neurophysiology*, 79(6):2857–2874, 1998.
 - [171] Rodrigo Quian Quiroga and Stefano Panzeri. Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci*, 10(3):173–185, 2009.
 - [172] Alain Rakotomamonjy and Stéphane Canu. Frames, reproducing kernels, regularization and learning. *J. Mach. Learn. Res.*, 6:1485–1515, 2005.
 - [173] Sidney Redner. *A Guide to First-passage Processes*. Cambridge University Press, 05 2007.
 - [174] Daniel S. Reich, Jonathan D. Victor, and Bruce W. Knight. The power ratio and the interval map: Spiking models and extracellular recordings. *The Journal of Neuroscience*, 18(23):10090–10104, 12 1998.
 - [175] Daniel S. Reich, Jonathan D. Victor, Bruce W. Knight, Tsuyoshi Ozaki, and Ehud Kaplan. Response variability and timing precision of neuronal spike trains in vivo. *Journal of Neurophysiology*, 77(5):2836–2841, 1997.
 - [176] Pamela Reinagel and R. Clay Reid. Temporal coding of visual information in the thalamus. *The Journal of Neuroscience*, 20(14):5392–5400, 2000.
 - [177] Jan Reutimann, Michele Giugliano, and Stefano Fusi. Event-driven simulation of spiking neurons with stochastic dynamics. *Neural Computation*, 15(4):811–830, 2003.

- [178] L. M. Ricciardi, L. Sacerdote, and S. Sato. On an integral equation for first-passage-time probability densities. *J. Appl. Probab.*, 21(2):302–314, 1984.
- [179] Luigi M. Ricciardi and Laura Sacerdote. The Ornstein-Uhlenbeck process as a model for neuronal activity. *Biological Cybernetics*, 35(1):1–9, March 1979.
- [180] Luigi M. Ricciardi and Shunsuke Sato. First-passage-time density and moments of the Ornstein-Uhlenbeck process. *J. Appl. Probab.*, 25(1):43–57, 1988.
- [181] Fred Rieke, David Warland, Rob De_ruyter_van_steveninck, and William Bialek. *Spikes: Exploring the Neural Code (Computational Neuroscience)*. The MIT Press, 06 1999.
- [182] H. Risken. *The Fokker-Planck equation*, volume 18 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, second edition, 1989. Methods of solution and applications.
- [183] Christian P. Robert and George Casella. *Monte Carlo statistical methods*. Springer Texts in Statistics. Springer-Verlag, New York, second edition, 2004.
- [184] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Foundations, Reprint of the second (1994) edition.
- [185] Edmund T. Rolls and Gustavo Deco. *The Noisy Brain : Stochastic Dynamics as a Principle of Brain Function*. OUP, January 2010.
- [186] Werner Römisch and Renate Winkler. Stepsize control for mean-square numerical methods for stochastic differential equations with small noise. *SIAM J. Sci. Comput.*, 28(2):604–625, 2006.
- [187] Eduardo Ros, Richard Carrillo, Eva M. Ortigosa, Boris Barbour, and Rodrigo Agís. Event-driven simulation scheme for spiking neural networks using lookup

- tables to characterize neuronal dynamics. *Neural Computation*, 18(12):2959–2993, 2006.
- [188] Michael Rosenblum, Arkady Pikovsky, and Jürgen Kurths. Phase synchronization of chaotic oscillators. *Physical Review Letters*, 76(11):1804–1807, Mar 1996.
- [189] Michael Rudolph and Alain Destexhe. Do neocortical pyramidal neurons display stochastic resonance? *Journal of Computational Neuroscience*, 11(1):19–42, 2001-07-01.
- [190] Michelle Rudolph and Alain Destexhe. Analytical integrate-and-fire neuron models with conductance-based dynamics for event-driven simulation strategies. *Neural Computation*, 18(9):2146–2210, 2006.
- [191] Michelle Rudolph and Alain Destexhe. How much can we trust neural simulation strategies? *Neurocomputing*, 70(10-12):1966–1969, 06 2007.
- [192] Nicole C. Rust, Odelia Schwartz, J. Anthony Movshon, and Eero P. Simoncelli. Spatiotemporal elements of macaque v1 receptive fields. *Neuron*, 46(6):945–956, 2005.
- [193] Julius Schauder. Bemerkungen zu meiner Arbeit “Zur Theorie stetiger Abbildungen in Funktionalräumen”. *Math. Z.*, 26(1):417–431, 1927.
- [194] Julius Schauder. Eine Eigenschaft des Haarschen Orthogonalsystems. *Math. Z.*, 28(1):317–320, 1928.
- [195] Bernhard Schölkopf, Ralf Herbrich, and Alex J. Smola. A generalized representer theorem. In *Computational learning theory (Amsterdam, 2001)*, volume 2111 of *Lecture Notes in Comput. Sci.*, pages 416–426. Springer, Berlin, 2001.

- [196] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5):1299–1319, 1998.
- [197] Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*. The MIT Press, 1st edition, December 2001.
- [198] C. E. Schreiner and M. W. Raggio. Neuronal responses in cat primary auditory cortex to electrical cochlear stimulation. ii. repetition rate coding. *Journal of Neurophysiology*, 75(3):1283–1300, 03 1996.
- [199] Erwin Schrödinger. Zur theorie der fall- und steigversuche an teilchen mit brownscher bewegung. *Physikalische Zeitschrift*, (16):289–295, 1915.
- [200] Štefan Schwabik, Milan Tvrdý, and Otto Vejvoda. *Differential and integral equations*. D. Reidel Publishing Co., Dordrecht, 1979. Boundary value problems and adjoints.
- [201] Jun Sekine. Information geometry for Symmetric Diffusions. *Potential Analysis*, 14(1):1–30, 2001.
- [202] V. Seshadri. *The inverse Gaussian distribution*, volume 137 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1999. Statistical theory and applications.
- [203] Tatyana Sharpee, Nicole C. Rust, and William Bialek. Analyzing neural responses to natural signals: Maximally informative dimensions. *Neural Computation*, 16(2):223–250, 2004.
- [204] T. Shimokawa, K. Pakdaman, and S. Sato. Time-scale matching in the response of a leaky integrate-and-fire neuron model to periodic stimulus with additive noise. *Phys. Rev. E*, 59(3):3427–3443, Mar 1999.

- [205] T. Shimokawa, K. Pakdaman, T. Takahata, S. Tanabe, and S. Sato. A first-passage-time analysis of the periodically forced noisy leaky integrate-and-fire model. *Biological Cybernetics*, 83:327–340, 2000. 10.1007/s004220000156.
- [206] Arnold J. F. Siegert. On the first passage time probability problem. *Physical Rev. (2)*, 81:617–623, 1951.
- [207] Lawrence Sirovich and Bruce Knight. Spiking neurons and the first passage problem. *Neural Computation*, 23(7):1675–1703, 2011/09/06 2011.
- [208] N. Slonim. *The Information Bottleneck: Theory And Applications*. PhD thesis, The Hebrew University, 2003.
- [209] N. Smirnov. Sur les écarts de la courbe de distribution empirique. *Rec. Math. N.S. [Mat. Sbornik]*, 6(48):3–26, 1939.
- [210] William C. Stacey and Dominique M. Durand. Synaptic noise improves detection of subthreshold signals in hippocampal ca1 neurons. *Journal of Neurophysiology*, 86(3):1104–1112, 09 2001.
- [211] Richard B. Stein, E. Roderich Gossen, and Kelvin E. Jones. Neuronal variability: noise or part of the signal? *Nat Rev Neurosci*, 6(5):389–397, 05 2005.
- [212] Charles Stevens and Anthony M. Zador. Novel integrate-and-fire-like model of repetitive firing in cortical neurons. In *In Proceedings of the 5th Joint Symposium on Neural Computation, UCSD*, 1998.
- [213] Charles F. Stevens and Anthony M. Zador. Input synchrony and the irregular firing of cortical neurons. *Nat Neurosci*, 1(3):210–217, 07 1998.
- [214] William J. Stewart. *Probability, Markov chains, queues, and simulation*. Princeton University Press, Princeton, NJ, 2009. The mathematical basis of performance modeling.

- [215] S. P. Strong, Roland Koberle, Rob R. de Ruyter van Steveninck, and William Bialek. Entropy and information in neural spike trains. *Physical Review Letters*, 80(1), 01 1998.
- [216] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
- [217] Thibaud Taillefumier and Marcelo Magnasco. A fast algorithm for the first-passage times of Gauss-Markov processes with Hölder continuous boundaries. *Journal of Statistical Physics*, 140(6):1–27, 2010.
- [218] Thibaud Taillefumier and Jonathan Touboul. Multi-resolution Schauder approach to multidimensional Gauss-Markov processes. *arXiv preprint: <http://arxiv.org/abs/1008.1954>, To Appear In International Journal of Stochastic Analysis*.
- [219] Frederic E. Theunissen, Kamal Sen, and Allison J. Doupe. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *Journal of Neuroscience*, 20(6):2315–2331, 2000.
- [220] Simon Thorpe, Arnaud Delorme, and Rufin Van Rullen. Spike-based strategies for rapid processing. *Neural Networks*, 14(6-7):715–725, 2001.
- [221] P. H. E. Tiesinga. Precision and reliability of periodically and quasiperiodically driven integrate-and-fire neurons. *Phys. Rev. E*, 65:041913, Apr 2002.
- [222] Gasper Tkacik and Marcelo O. Magnasco. Decoding spike timing: The differential reverse-correlation method. *Biosystems*, 93(1-2):90–100, 2007.
- [223] Arnaud Tonnelier, Hana Belmabrouk, and Dominique Martinez. Event-driven

- p>simulations of nonlinear integrate-and-fire neurons.
- Neural Computation*
- , 19(12):3226–3238, 2007.
- [224] Jonathan Touboul and Olivier Faugeras. The spikes train probability distributions: A stochastic calculus approach. *Journal of Physiology-Paris*, 101(1-3):78–98, 2007/5// 2007.
- [225] Jonathan Touboul and Olivier Faugeras. A characterization of the first hitting time of double integral processes to curved boundaries. *Adv. in Appl. Probab.*, 40(2):501–528, 2008.
- [226] Jonathan Touboul and Olivier Faugeras. A markovian event-based framework for stochastic spiking neural networks. *Journal of Computational Neuroscience*, pages 1–23, 2011. 10.1007/s10827-011-0327-y.
- [227] J. Vincent Toups, Jean-Marc Fellous, Peter J. Thomas, Terrence J. Sejnowski, and Paul H. Tiesinga. Finding the event structure of neuronal spike trains. *Neural Computation*, 23(9):2169–2208, 2012/02/13 2011.
- [228] Alessandro Treves, Stefano Panzeri, Edmund T. Rolls, Michael Booth, and Edward A. Wakeman. Firing rate distributions and efficiency of information transmission of inferior temporal cortex neurons to natural visual stimuli. *Neural Computation*, 11(3):601–631, 2011/09/06 1999.
- [229] F. G. Tricomi. *Integral equations*. Dover Publications Inc., New York, 1985. Reprint of the 1957 original.
- [230] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the Brownian motion. *Phys. Rev.*, 36(5):823–841, Sep 1930.
- [231] N. G. van Kampen. *Stochastic processes in physics and chemistry*, volume 888

- of *Lecture Notes in Mathematics*. North-Holland Publishing Co., Amsterdam, 1981.
- [232] Jonathan Victor. Asymptotic approach of generalized orthogonal functional expansions to Wiener kernels. *Annals of Biomedical Engineering*, 19(4):383–399, 1991.
 - [233] Jonathan D. Victor. Binless strategies for estimation of information from neural data. *Phys. Rev. E*, 66:051903, Nov 2002.
 - [234] Jonathan D Victor. Analyzing receptive fields, classification images and functional images: challenges with opportunities for synergy. *Nat Neurosci*, 8(12):1651–1656, 2005.
 - [235] Jonathan D Victor. Spike train metrics. *Current Opinion in Neurobiology*, 15(5):585–592, 10 2005.
 - [236] Jonathan D Victor and Keith P Purpura. Metric-space analysis of spike trains: theory, algorithms and application. *Network: Computation in Neural Systems*, 8(2):127–164, 2011/09/02 1997.
 - [237] Tim P. Vogels and L. F. Abbott. Signal propagation and logic gating in networks of integrate-and-fire neurons. *Journal of Neuroscience*, 25(46):10786–10795, 2005.
 - [238] Tim P Vogels and L F Abbott. Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. *Nat Neurosci*, 12(4):483–491, 2009.
 - [239] John von Neumann. Various techniques used in connection with random digits. *National Bureau of Standards, Applied Math Series*, 11:36–38, 1951.

- [240] M von Smoluchowski. Drei vorträge über diffusion, brownsche molekularbewegung und koagulation von kolloidteilchen. *Physikalische Zeitschrift*, 17:557–571, 585–599, 1915.
- [241] Grace Wahba. *Spline models for observational data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.
- [242] Liqun Wang and Klaus Pötzelberger. Boundary crossing probability for Brownian motion and general boundaries. *J. Appl. Probab.*, 34(1):54–65, 1997.
- [243] Liqun Wang and Klaus Pötzelberger. Crossing probabilities for diffusion processes with piecewise continuous boundaries. *Methodol. Comput. Appl. Probab.*, 9(1):1387–5841, 2007.
- [244] Liqun Wang and Klaus Pötzelberger. Crossing probabilities for diffusion processes with piecewise continuous boundaries. *Methodol. Comput. Appl. Probab.*, 9(1):21–40, 2007.
- [245] Nick Webber and Claudia Ribeiro. A Monte Carlo method for the normal inverse Gaussian option valuation model using an inverse Gaussian bridge. Technical Report 5, 2003.
- [246] Michael Wehr and Gilles Laurent. Relationship between afferent and central temporal patterns in the locust olfactory system. *Journal of Neuroscience*, 19(1):381–390, 1999.
- [247] G. A. Whitmore. First-passage-time models for duration data: Regression structures and competing risks. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 35(2):207–219, 01 1986.

- [248] D. V. Widder. Positive temperatures on an infinite rod. *Trans. Amer. Math. Soc.*, 55:85–95, 1944.
- [249] Yang Yang, Michael R DeWeese, Gonzalo H Otazu, and Anthony M Zador. Millisecond-scale differences in neural activity in auditory cortex can drive decisions. *Nat Neurosci*, 11(11):1262–1263, 2008.