Applied Mathematical Sciences

Zhongqiang Zhang
George Em Karniadakis

Numerical Methods for Stochastic Partial Differential Equations with White Noise

Springer

Applied Mathematical Sciences

Volume 196

More information about this series at http://www.springer.com/series/34

Zhongqiang Zhang • George Em Karniadakis

Numerical Methods for Stochastic Partial Differential Equations with White Noise

Springer

Zhongqiang Zhang
Department of Mathematical Sciences
Worcester Polytechnic Institute
Worcester, Massachusetts, USA

George Em Karniadakis
Division of Applied Mathematics
Brown University
Providence, Rhode Island, USA

ISSN 0066-5452 ISSN 2196-968X (electronic)
Applied Mathematical Sciences
ISBN 978-3-319-57510-0 ISBN 978-3-319-57511-7 (eBook)
DOI 10.1007/978-3-319-57511-7

Library of Congress Control Number: 2017941192

Mathematics Subject Classification (2010): 65Cxx, 65Qxx, 65Mxx

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

In his forward-looking paper [374] at the conference "Mathematics Towards the Third Millennium," our esteemed colleague at Brown University Prof. David Mumford argued that "...stochastic models and statistical reasoning are more relevant to i) the world, ii) to science and many parts of mathematics and iii) particularly to understanding the computations in our mind, than exact models and logical reasoning." Deterministic modeling and corresponding simulations are computationally much more manageable than stochastic simulations, but they also offer much less, i.e., a single point in the "design space" instead of a "sphere" of possible solutions that reflect the various random (or not) perturbations of the physical or biological problem we study.

In the last twenty years, three-dimensional simulations of physical and biological phenomena have gone from an exception to the rule, and they have been widely adopted in Computational Science and Engineering (CSE). This more realistic approach to simulating physical phenomena together with the continuing fast increase of computer speeds has also led to the desire in performing more ambitious and even more realistic simulations with multiscale physics. However, not all scales can be modeled directly, and stochastic modeling has been used to account for the un-modeled physics. In addition, there is a fundamental need to quantify the uncertainty of large-scale simulations, and this has led to a new emerging field in CSE, namely that of Uncertainty Quantification or UQ. Hence, computational scientists and engineers are interested in endowing their simulations with "error bars" that reflect not only numerical discretization errors but also uncertainties associated with unknown precise values of material properties, ill-defined or even unknown boundary conditions, or uncertain constitutive laws in the governing equations. However, performing UQ taxes the computational resources greatly, and hence the selection of the proper numerical method for stochastic simulations is of paramount importance, in some sense much more important than selecting a numerical method for deterministic modeling.

The most popular simulation method for stochastic modeling is the Monte Carlo method and its various extensions, but it requires a lot of computational effort to compute thousands and often millions of sample paths required to obtain certain statistics of the quantity of interest. Specifically, Monte Carlo methods are quite general, but they suffer from slow convergence so they are usually employed in conjunction with some variance reduction techniques to produce satisfactory accuracy in practice. More recently and for applications that employ stochastic models with *color* noise, deterministic integration methods in random space have been used with great success as they lead to high accuracy, especially for a modest number of uncertain parameters. However, they are not directly applicable to stochastic partial differential equations (SPDEs) with temporal white noise, since their solutions are usually non-smooth and would require a very large number of random variables to obtain acceptable accuracy.

Methodology. For linear SPDEs, we can still apply deterministic integration methods by exploiting the linear property of these equations for a long-time numerical integration. This observation has been made in [315] for Zakai-type equations, where a recursive Wiener chaos method was developed. In this book, we adopt this idea and we further formulate a recursive strategy to solve linear SPDEs of general types using Wiener chaos methods and stochastic collocation methods in conjunction with sparse grids for efficiency. In order to apply these deterministic integration methods, we first truncate the stochastic processes (e.g., Brownian motion) represented by orthogonal expansions. In this book, we show that the orthogonal expansions can lead to higher-order schemes with proper time discretization when Wiener chaos expansion methods are employed. However, we usually keep only a small number of truncation terms in the orthogonal expansion of Brownian motion to efficiently use deterministic integration methods for *temporal* noise. For *spatial* noise, the orthogonal expansion of Brownian motions leads to higher-order methods in both random space and physical space when the solutions to the underlying SPDEs are smooth.

The framework we use is the Wong-Zakai approximation and Wiener chaos expansion. Wiener chaos expansion is associated with the Ito-Wick product which was used intensively in [223]. The methodology and the proofs are introduced in [315] and in some subsequent papers by Rozovsky and his collaborators. In this framework, we are led to systems of deterministic partial differential equations with unknowns being the Wiener chaos expansion coefficients, and it is important to understand the special structure of these linear systems. Following another framework, such as viewing the SPDEs as infinite-dimensional stochastic differential equations (SDEs), admits a ready application of numerical methods for SDEs. We emphasize that there are multiple points of view for treating SPDEs and accordingly there are many views on what are proper numerical methods for SPDEs. However, irrespective of such views, the common difficulty remains: the solutions are typically

very rough and do not have first-order derivatives either in time or in space. Hence, no high-order (higher than first-order) methods are known except in very special cases.

Since the well-known monographs on numerical SDEs by Kloeden & Platen (1992) [259], Milstein (1995) [354], and Milstein & Tretyakov (2004) [358], numerical SPDEs with white noise have gained popularity, and there have been some new books on numerical SPDEs available, specifically:

- The book by Jentzen & Kloeden (2011) [251] on the development of stochastic Taylor's expansion for mild solutions to stochastic parabolic equations and their application to numerical methods.
- The book by Grigoriu (2012) [174] on the application of stochastic Galerkin and collocation methods as well as Monte Carlo methods to partial differential equations with random data, especially elliptic equations. Numerical methods for SODEs with random coefficients are discussed as well.
- The book by Kruse (2014) [277] on numerical methods in space and time for semi-linear parabolic equations driven by space-time noise addressing strong (L^p or mean-square) and weak (moments) sense of convergence.
- The book by Lord, Powell, & Shardlow (2014) [308] on the introduction of numerical methods for stochastic elliptic equations with color noise and stochastic semi-linear equations with space-time noise.

For numerical methods of stochastic differential equations with *color* noise, we refer the readers to [294, 485]. On the theory of SPDEs, there are also some new developments, and we refer the interested readers to the book [36] covering amplitude equations for nonlinear SPDEs and to the book [119] on homogenization techniques for effective dynamics of SPDEs.

How to use this book. This book can serve as a reference/textbook for graduate students or other researchers who would like to understand the state-of-the-art of numerical methods for SPDEs with white noise.

Reading this book requires some basic knowledge of probability theory and stochastic calculus, which are presented in Chapter 2 and Appendix A. Readers are also required to be familiar with numerical methods for partial differential equations and SDEs in Chapter 3 before further reading. The reader can also refer to Chapter 3 for MATLAB implementations of test problems. More MATLAB codes for examples in this book are available upon request. For those who want to take a glance of numerical methods for stochastic partial differential equation, they are encouraged to read a review of these methods presented in Chapter 3. Exercises with hints are provided in most chapters to nurture the reader's understanding of the presented materials.

Part I. Numerical stochastic ordinary differential equations. We start with numerical methods for SDEs with delay using the Wong-Zakai approximation

and finite difference in time in Chapter 4. The framework of Wong-Zakai approximation is used throughout the book. If the delay time is zero, we then recover the standard SDEs. We then discuss how to deal with strong nonlinearity and stiffness in SDEs in Chapter 5.

Part II. Temporal white noise. In Chapters 6–8, we consider SPDEs as PDEs driven by white noise, where discretization of white noise (Brownian motion) leads to PDEs with smooth noise, which can then be treated by numerical methods for PDEs. In this part, recursive algorithms based on Wiener chaos expansion and stochastic collocation methods are presented for linear stochastic advection-diffusion-reaction equations. Stochastic Euler equations in Chapter 9 are exploited as an application of stochastic collocation methods, where a numerical comparison with other integration methods in random space is made.

Part III. Spatial white noise. We discuss in Chapter 10 numerical methods for nonlinear elliptic equations as well as other equations with additive noise. Numerical methods for SPDEs with multiplicative noise are discussed using the Wiener chaos expansion method in Chapter 11. Some SPDEs driven by non-Gaussian white noise are discussed, where some model reduction methods are presented for generalized polynomial chaos expansion methods.

We have attempted to make the book self-contained. Necessary background knowledge is presented in the appendices. Basic knowledge of probability theory and stochastic calculus is presented in Appendix A. In Appendix B, we present some semi-analytical methods for SPDEs. In Appendix C, we provide a brief introduction to Gauss quadrature. In Appendix D, we list all the conclusions we need for proofs. In Appendix E, we present a method to compute convergence rate empirically.

MATLAB codes accompanying this book are available at the following website:

https://github.com/springer-math/Numerical-Methods-for-Stochastic-Partial-Differential-Equations-with-White-Noise

Acknowledgments. This book is dedicated to Dr. Fariba Fahroo who with her quiet and steady leadership at AFOSR and DARPA and with her generous advocacy and support has established the area of uncertainty quantification and stochastic modeling as an independent field, indispensable in all areas of computational simulation of physical and biological problems.

This book is based on our research through collaboration with Professor Boris L. Rozovsky at Brown University and Professor Michael V. Tretyakov at the University of Nottingham. We are indebted to them for their valuable advice. Specifically, Chapters 6, 7, and 8 are based on collaborative research with them. Chapter 5 is also from a collaboration with Professor Michael V. Tretyakov. We would also like to thank Professor Wanrong Cao at Southeast University for providing us numerical results in Chapter 4 and Professor Guang Lin at Purdue University and Dr. Xiu Yang at Pacific Northwest National Laboratory for providing their code for Chapter 9.

The research of this book was supported by OSD/MURI grant FA9550-09-1-0613, NSF/DMS grant DMS-0915077, NSF/DMS grant DMS-1216437, and ARO and NIH grants and also by the Collaboratory on Mathematics for Mesoscopic Modeling of Materials (CM4) which is sponsored by DOE. The first author was also supported by a start-up funding at Worcester Polytechnic Institute during the writing of the book.

Worcester, MA, USA — Zhongqiang Zhang
Providence, RI, USA — George Em Karniadakis

Contents

1

Prologue

Stochastic mathematical models have received increasing attention for their ability of representing intrinsic uncertainty in complex systems, e.g., representing various scales in particle simulations at molecular and mesoscopic scales, as well as extrinsic uncertainty, e.g., stochastic external forces, stochastic initial conditions, or stochastic boundary conditions. One important class of stochastic mathematical models is stochastic partial differential equations (SPDEs), which can be seen as deterministic partial differential equations (PDEs) with finite or infinite dimensional stochastic processes – either with color noise or white noise. Though white noise is a purely mathematical construction, it can be a good model for rapid random fluctuations, and it is also the limit of color noise when the correlation length goes to zero.

In the following text, we first discuss why random variables/processes are used using a simple model and introduce the color of noise. Then we present some typical models using PDEs with random (stochastic) processes. At last, we preview the topics covered in this book and the methodology we use.

1.1 Why random and Brownian motion (white noise)?

Consider the following simple population growth model

$$dy(t) = k(t)y(t)\,dt, \quad y(0) = y_0. \tag{1.1.1}$$

Here $y(t)$ is the size of population and $k(t)$ is the relative growth rate. In practice, $k(t)$ is not completely known and is disturbed around some known quantity $\bar{k}(t)$:

$$k(t) = \bar{k}(t) + \text{`perturbation (noise)'}.$$

Here $\bar{k}(t)$ is deterministic and is usually known while no exact behavior exists of the perturbation (noise) term. The uncertainty (lack of information) about

Z. Zhang, G.E. Karniadakis, *Numerical Methods for Stochastic Partial Differential Equations with White Noise*,
Applied Mathematical Sciences 196, DOI 10.1007/978-3-319-57511-7_1

$k(t)$ (the perturbation term) is naturally represented as a stochastic quantity, denoted as $v(t,\omega)$. Here ω represents the randomness.

To address the dependence on ω, we then write the ordinary differential equation as

$$dy(t,\omega) = k(t,\omega)y(t,\omega)\,dt,^{1} \quad y(0) = y_0. \tag{1.1.2}$$

Here y_0 can be a random variable but we take $y_0 = 1$ (deterministic) for simplicity.

What color is the noise?

In many situations, it is assumed that a stochastic process $v(t)$ satisfies the following properties

1. The expectation of $v(t)$ is zero for all t, i.e., $\mathbb{E}[v(t)] = 0$.
2. The covariance (two-point correlation) function of $v(t)$ is more or less known. That is,

$$\mathbb{C}\mathrm{ov}[(v(t), v(s))] = \mathbb{E}[(v(t) - \mathbb{E}[v(t)])(v(s) - \mathbb{E}[v(s)])].$$

When the covariance function is proportional to the Dirac function $\delta(t-s)$, the process $v(t)$ is called uncorrelated, which is usually referred to as *white noise*. Otherwise, it is correlated and is referred to as *color noise*. The white noise can be intuitively described as a stochastic process, which has independent values at each time instance and has an infinite variance.

Due to the simplicity of Gaussian processes, $v(t,\omega)$ is modeled with a Gaussian process. One important Gaussian process is Brownian motion in physics, which describes a random motion of particles in a fluid with constantly varying fluctuations. It is fully characterized by its expectation (usually taken as zero) and its covariance function, see Chapter 2.2.

Another way to represent the Gaussian noise is through Fourier series. A real-valued Gaussian process can be represented as

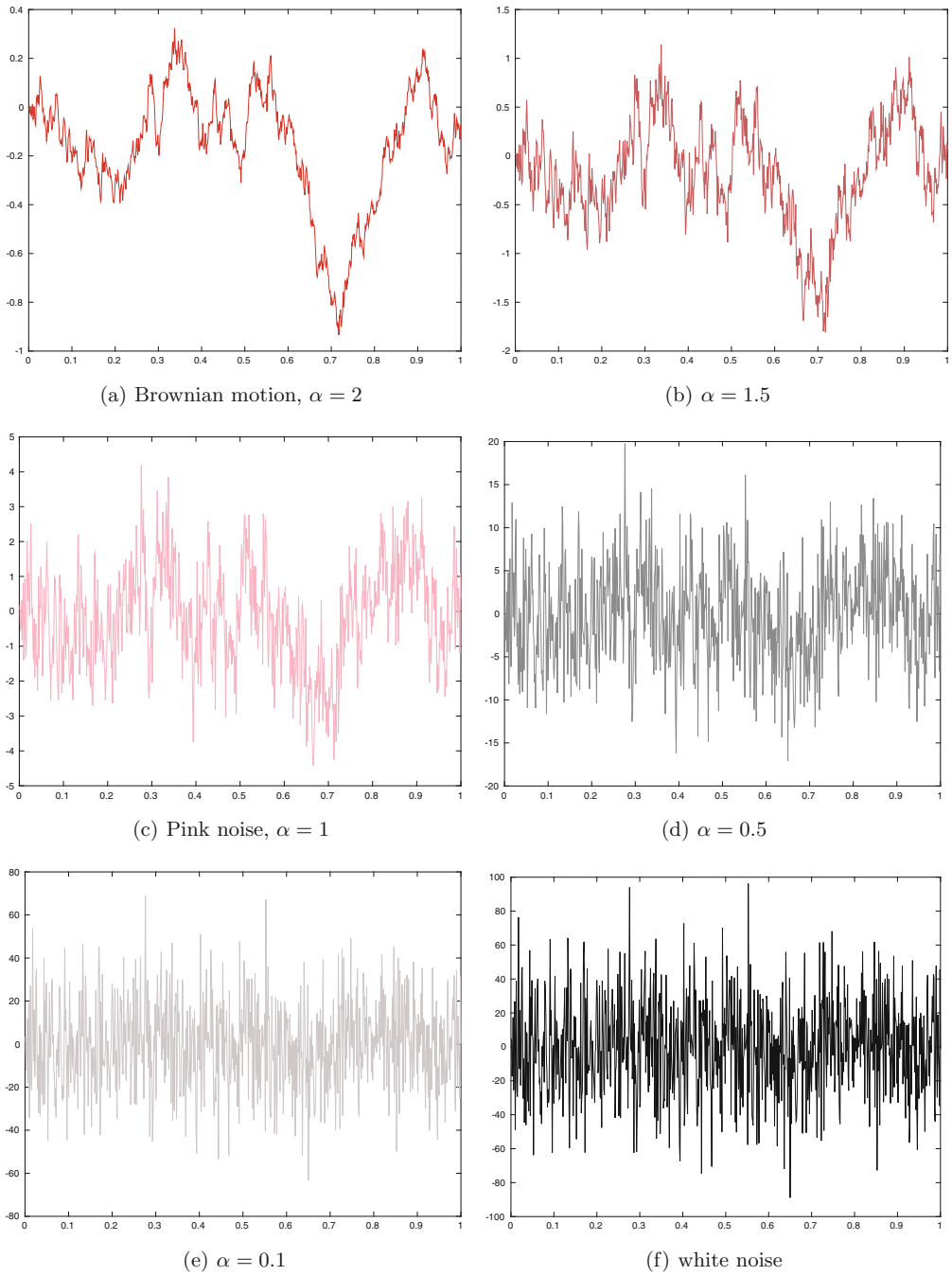
$$v(t) = \sum_{k=-\infty}^{\infty} e^{-ikt} a_k \xi_k,$$

where ξ_k's are i.i.d. standard Gaussian random variables (i.e., mean zero and variance 1) and i is the imaginary unit ($i^2 = -1$). When a_k's are the same constant, the process is called white noise. When $|a_k|^2$ is proportional to $1/k^\alpha$, it is called $1/f^\alpha$ noise. When $\alpha = 2$, the process is called red noise (Brownian noise). When $\alpha = 1$, the process is called pink noise. It is called blue noise when $\alpha = -1$. However, white noise ($\alpha = 0$) is not observed in experiments and nature, but $1/f^\alpha$ noise, $0 < \alpha \leq 2$ was first observed back in

[1] We assume that $k(t,\omega)$ is sufficiently smooth in t. It will be clear why some smoothness is needed when we introduce the stochastic product and calculus in Chapter 2.3.

1910s, e.g., $\alpha = 1$ [144]. When α is closer to 0, the process $v(t)$ becomes less correlated and when $\alpha = 0$, white noise can be treated as a limit of processes with extremely small correlation. This is illustrated in Figure 1.1, which is generated by the Matlab code in [429] using Mersenne Twister pseudorandom number generator with seed 100. The smaller α is, the closer is the noise to white noise.

Fig. 1.1. Sample paths of different Gaussian $1/f^{\alpha}$ processes.



(a) Brownian motion, $\alpha = 2$

(b) $\alpha = 1.5$

(c) Pink noise, $\alpha = 1$

(d) $\alpha = 0.5$

(e) $\alpha = 0.1$

(f) white noise

Solutions to (1.1.2)

When $\bar{k}(t) = 0$, $k(t,\omega) = v(t,\omega)$ may take the following form:

- $v(t,\omega) = \xi \sim \mathcal{N}(0,1)$ is a standard Gaussian random variable. The covariance (in this case variance) of $v(t,\omega)$ is 1.
- $v(t,\omega)$ where the covariance function of $V(t)$ is $\exp(-\frac{|t-s|}{A})$, with A the correlation time.
- $v(t,\omega) = W(t,\omega)$ is a standard Brownian motion where the covariance function is $\min(t,s)$.
- $v(t,\omega) = \dot{W}(t,\omega)$ is the white noise and the covariance function is $\delta(t-s)$.

When $k(t,\omega) =: \dot{W}(t,\omega)$, Equation (1.1.2) is understood in the Stratonovich sense

$$dy = y \circ dW(t), \quad y(0) = y_0 \tag{1.1.3}$$

so that one can apply the classical chain rule. We discuss what the circle means in Chapter 2.3, but roughly speaking the circle denotes Stratonovich product, which can be understood in some limit of Riemann sum using the midpoint rule.

The exact solution to Equation (1.1.2) is $y = y_0 \exp(K(t))$, where $K(t) = \int_0^t k(s)\, ds$ is again Gaussian with mean zero. It can be readily checked by the definition of moments that

- $K(t) \sim \mathcal{N}(0, t^2)$ when $k(t,\omega) =: \xi \sim \mathcal{N}(0,1)$;
- $K(t) \sim \mathcal{N}\big(0, 2At + 2A^2(\exp(-\frac{t}{A}) - 1)\big)$ when $k(t,\omega)$ has the two-point correlation function of $\exp(-\frac{|t_1-t_2|}{A})$.
- $K(t) \sim \mathcal{N}(0, \frac{t^3}{3})$ when $k(t,\omega) = W(t,\omega)$ is a standard Brownian motion.
- $K(t) \sim \mathcal{N}(0,t)$ is the Brownian motion when $k(t,\omega) = \dot{W}(t,\omega)$ is the white noise.

Then the moments of the solution y are, for $m = 1, 2, \cdots$,

$$\mathbb{E}[y^m(t)] = y_0^m \exp(\frac{m^2}{2}\sigma^2), \tag{1.1.4}$$

where $\sigma^2 = t^2,\ 2At + 2A^2(\exp(-\frac{t}{A}) - 1),\ \frac{t^3}{3},\ t$ for the listed processes, respectively. These results are useful and can be used in checking the accuracy of different numerical schemes applied to Equation (1.1.2).

1.2 Modeling with SPDEs

SPDEs with white noise have been formulated for various applications, such as nonlinear filtering (see, e.g., [499]), turbulent flows (see, e.g., [43, 348]), fluid flows in random media (see, e.g., [223]), particle systems (see, e.g., [268]), population biology (see, e.g., [98]), neuroscience (see, e.g., [463]), etc. Since

analytic solutions to SPDEs can rarely be obtained, numerical methods have to be developed to solve SPDEs. One of the motivations for numerical SPDEs in the early literature was to solve the Zakai equation of nonlinear filtering, see, e.g., [31, 78, 130, 150–152]. In the next section, we review some numerical methods for a semilinear equation (3.4.1), the advection-diffusion-reaction equation of nonlinear filtering (3.4.13), the stochastic Burgers equation (3.4.20), and the stochastic Navier-Stokes equation (1.2.4).

Let $W(t)$ be a r-dimensional Brownian motion, i.e., $W(t) = (W_1(t), \dots, W_r(t))^\top$ where $W_i(t)$'s are mutually independent Brownian motions – Gaussian processes with the covariance function $\min(t, s)$. For a rigorous definition of Brownian motion, see Chapter 2.2.

Example 1.2.1 (Zakai equation of nonlinear filtering) *Let $y(t)$ be a r-dimensional observation of a signal $x(t)$ and $y(t)$ given by*

$$y(t) = y_0 + \int_0^t h(x(s))\, ds + W(t),$$

where $h = (h_1, h_2, \dots, h_r)^\top$ is a $\mathbb{R}^r$-vector-valued function defined on $\mathbb{R}^d$ and the observational signal $x(t)$ satisfies the following stochastic differential equation

$$dx(t) = b(x(t))\, dt + \sum_{k=1}^{q} \sigma_k(x(t))\, dB_k, \quad x(0) = x_0,$$

where b and σ_k's are d-dimensional vector functions on $\mathbb{R}^d$, and $B(t) = (B_1(t), B_2(t), \dots, B_q(t))^\top$ is a q-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ and is independent of $W(t)$.

The conditional probability density function (un-normalized) of the signal $x(t)$ given $y(t)$ satisfies, see, e.g., [499],

$$\begin{aligned} du(t,x) = {} & \frac{1}{2} \sum_{i,j=1}^{d} D_i D_j [(\sigma\sigma^\top)_{ij} u(t,x)]\, dt - \sum_{i}^{d} D_i (b_i u(t,x))\, dt \\ & + \sum_{l=1}^{r} h_l u(t,x)\, dy_l(t), \quad x \in \mathbb{R}^d. \end{aligned} \tag{1.2.1}$$

Here $D_i := \partial_{x_i}$ is the partial derivative in the x_i-th direction and $\sigma^\top$ is the transpose of σ.

Equation (1.2.1) provides an analytical solution to the aforementioned filtering problem. Equation (1.2.1) and its variants have been a major motivation for the development of theory for SPDEs (see, e.g., [280, 408]) and corresponding numerical methods; see, e.g., [31, 78, 130, 150–152] and for a comprehensive review on numerical methods see [52].

Example 1.2.2 (Pressure equation) *The following model was introduced as an example of the pressure of a fluid in porous and heterogeneous (but isotropic) media at the point x over the physical domain $\mathcal{D}$ in $\mathbb{R}^d$ ($d \leq 3$):*

$$-\operatorname{div}(K(x)\nabla p) = f(x), \quad p|_{\partial\mathcal{D}} = 0, \tag{1.2.2}$$

where $K(x) > 0$ is the permeability of media at the point x and $f(x)$ is a mass source. In a typical porous media, $K(x)$ is fluctuating in an unpredictable and irregular way and can be modeled with a stochastic process.

In [140, 141], $K(x) = \exp(\int_{\mathcal{D}} \phi(x-y)\, dW(y) - \frac{1}{2}\|\phi\|^2)$ is represented as a lognormal process, see also [223] for a formulation of the lognormal process using the Ito-Wick product (see its definition in Chapter 2.3). Here ϕ is a continuous function in $\mathcal{D}$ and $\|\phi\| = (\int_{\mathcal{D}} \phi^2\, dx)^{1/2} < \infty$.

We define infinite dimensional Gaussian processes, see, e.g., [94, 408], as follows

$$W^Q(t,x) = \sum_{i\in\mathbb{N}^d} \sqrt{q_i} e_i(x) W_i(t), \tag{1.2.3}$$

where $W_i(t)$'s are mutually independent Brownian motions. Here $q_i \geq 0$, $i \in \mathbb{N}^d$ and $\{e_i(x)\}$ is an orthonormal basis in $L^2(\mathcal{D})$. The following noise is usually considered in literature: $\dot{W}^Q(t,x) = \sum_{i\in\mathbb{N}^d} \sqrt{q_i} e_i(x) \dot{W}_i(t)$. When $q_i = 1$ for all i's, we have a space-time *white noise*. When $\sum_{i=1}^{\infty} q_i < \infty$, we call it the space-time **color noise**. Sometimes it is called a Q-Wiener process. We call the noise finite-dimensional when $q_i = 0$ for all sufficient large i.

Example 1.2.3 (Turbulence model) *The stochastic incompressible Navier-Stokes equation reads, see, e.g., [108, 348]*

$$\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p = \sigma(\mathbf{u})\dot{W}^Q, \quad \operatorname{div}\mathbf{u} = 0, \tag{1.2.4}$$

where σ is Lipschitz continuous over the physical domain in $\mathbb{R}^d$ ($d = 2, 3$). Here $\mathbb{E}[W^Q(x,t)W^Q(y,s)] = q(x,y)\min(s,t)$ and $q(x,x)$ is square-integrable over the physical domain.

Example 1.2.4 (Reaction-diffusion equation)

$$du = \nu\Delta u + f(u) + \sigma(u)\dot{W}^Q(t,x).$$

This may represent a wide class of SPDEs:

- *In materials, the stochastic Allen-Cahn equation, where $f(u) = u(1-u)(1+u)$, $\sigma(u)$ is a constant and $\dot{W}^Q(t,x)$ is space-time white noise, see, e.g., in [204, 302].*
- *In population genetics, this equation has been used to model changes in the structure of population in time and space, where $\dot{W}^Q(t,x)$ is a Gaussian process white in time but color in space. For example, $f(u) = 0$, $\sigma(u) = \gamma\sqrt{\max(u,0)}$, where γ is a constant and u is the mass distribution of the population, see, e.g., [97]. Also, $f(u) = \alpha u - \beta$, $\sigma(u) = \gamma\sqrt{\max(u(1-u),0)}$, where α, β, γ are constants, see, e.g., [129].*

- *When $\sigma(u)$ is a small constant, the equation can be treated as random perturbation of deterministic equations ($\sigma(u) = 0$), see, e.g., [136].*

In the first and third cases, we say that the equation has an additive noise *as the coefficients of noise are independent of the solution. In the second case, we say that the equation has a* multiplicative noise *since the coefficient of noise depends on the solution itself.*

For more details on modeling with SPDEs, we refer to the books [223, 268]. For well-posedness (existence, uniqueness, and stability) of SPDEs, we refer to the books [94, 145, 170, 223, 399, 408]. For studies of the dynamics of SPDEs, we refer to [145] for asymptotic behavior of solutions to SPDEs; also to [92] for Kolmogrov equations for SPDEs, to [36] for amplitude equations of nonlinear SPDEs, and to [119] for homogenization of multiscale SPDEs. In this book, we present numerical methods for SPDEs. Specifically, we focus on *forward problems*, i.e., predicting the quantities of interest from known SPDEs with known inputs, especially those SPDEs driven by white noise.

1.3 Specific topics of this book

In this book, we focus on two issues related to numerical methods for SPDEs with white noise: one is deterministic integration methods in random space while the other one is the effect of truncation of the Brownian motion using the spectral approximation.[2] In Figure 1.2, we provide an overview on how we organize this book. In addition, we present some necessary details in the appendices to make the book self-contained.

In Chapter 2, we review some numerical methods for SPDEs with white noise addressing primarily their discretization in random space. We then explore the effect of truncation of Brownian motion using its spectral expansion (Wong-Zakai approximation) for stochastic differential equations with time delay in Chapter 4 and show that the Wong-Zakai approximation can facilitate the derivation of various numerical schemes.

For deterministic integration methods of SPDEs in random space, we aim at performing accurate long-time numerical integration of time-dependent equations, especially of linear stochastic advection-reaction-diffusion equations. We study Wiener chaos expansion methods (WCE) and stochastic collocation methods (SCM), compare their performance and prove their convergence order. To achieve longer time integration, we adopt the recursive WCE proposed in [315] for the Zakai equation for nonlinear filtering and develop algorithms for the first two moments of solutions. Numerical results show that when high accuracy is required, WCE is superior to

[2] There are different numerical methods for SPDEs using different approximation of Brownian motion and integration methods in random space, see Chapter 2 for a review, where discretization in time and space is also presented.

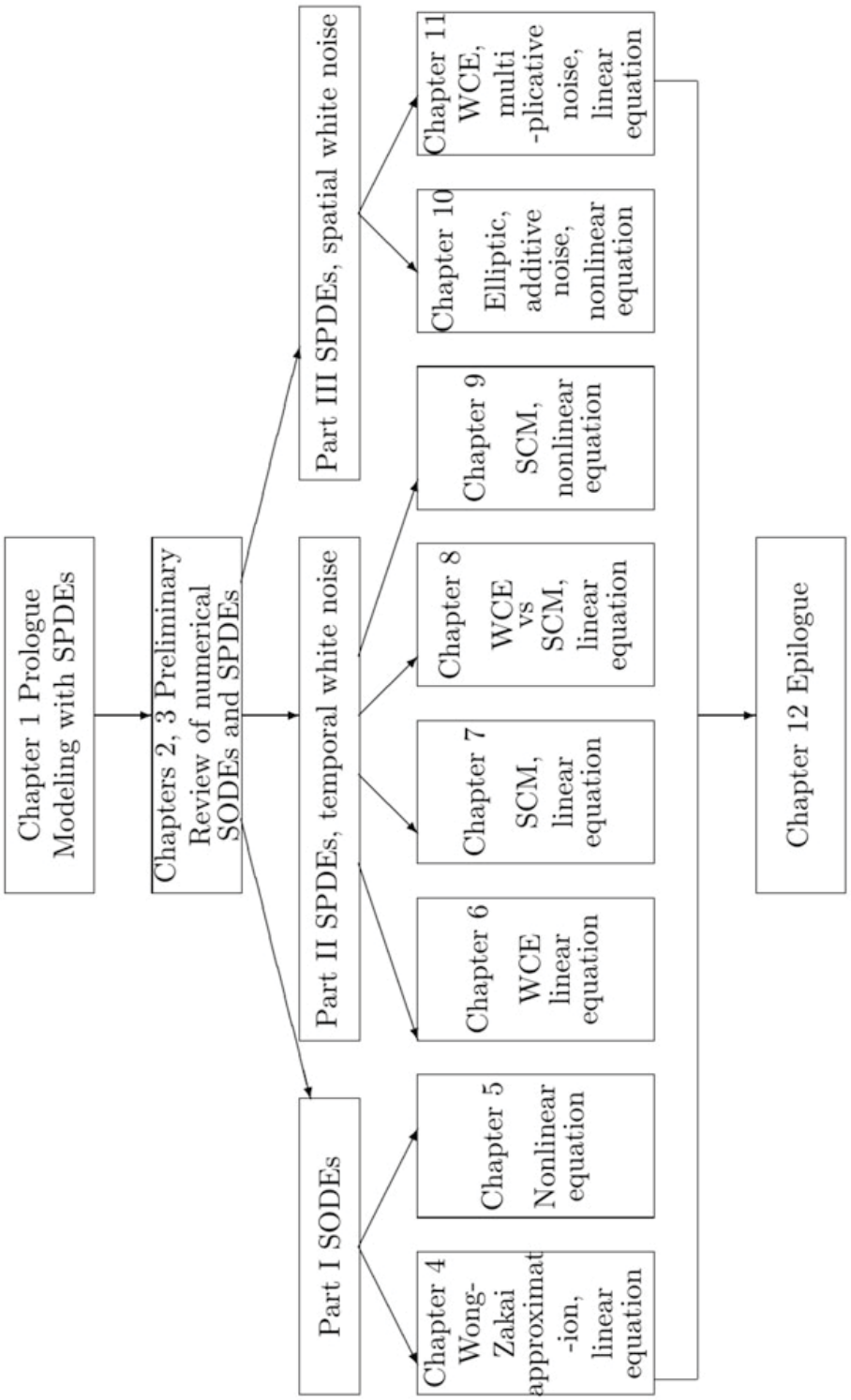
Monte Carlo methods while WCE is not as efficient if only low accuracy is sufficient, see Chapter 6. We show that the recursive approach for SCM for linear advection-reaction-diffusion equations is efficient for long-time integration in Chapter 7. We first analyze the error of SCM (sparse grid collocation of Smolyak type) with an Euler scheme in time for linear SODEs, and show that the error is small only when the noise magnitude is small and/or the integration time is relatively short. We compare WCE and SCM using the recursive procedure in Chapter 8, where we derive error estimates of WCE and SCM for linear advection-reaction-diffusion equations and show that WCE and SCM are competitive in practice by presenting careful numerical comparisons, even though WCE can be of one order higher than SCM.

Among almost all approximations for WCE and SCM, we use the Wong-Zakai approximation with spectral approximation of Brownian motion. The convergence order with respect to the number of truncation modes is half order, see Chapter 8. However, WCE can be of higher convergence order since it can preserve the orthogonality over the Wiener space (infinite dimensional) while SCM cannot as the orthogonality is only valid on discrete spaces (finite dimensional), see Chapter 8. In Chapter 9, we test the Wong-Zakai approximation in conjunction with the stochastic collocation method for the stochastic Euler equations modeling a stochastic piston problem and show the effectiveness of this approximation in a practical situation. To further investigate the effect of truncation of Brownian motions, we study the elliptic equation with additive white noise in Chapter 10. We show that the convergence of numerical solutions with truncation of Brownian motion depends on the smoothing effects of the resolvent of the elliptic operator. We also show similar convergence when finite element methods are used.

As shown in Figure 1.2, we focus on deterministic integration methods in random space, i.e., Wiener chaos and stochastic collocation, in Chapters 6–9 and compare their performance with Monte Carlo methods and/or quasi-Monte Carlo methods. In Chapter 6, we compare WCE and Monte Carlo methods and show that WCE is superior to Monte Carlo methods if high accuracy is needed. In Chapters 7 and 9, we show theoretically and numerically the efficiency of SCM for short time integration and for small magnitudes of noises. In Chapter 8, we compare WCE and SCM in conjunction with a recursive multistage procedure and show that they are comparable in performance. In Chapter 11, we consider WCE for elliptic equations with multiplicative noise. We use the Wick product for the interaction of two random variables as well as the Wick-Malliavin approximation to reduce the computational cost. We use Monte Carlo methods in Chapters 4 and 10 as the dimensionality in random space is beyond deterministic integration methods.

In all chapters except Chapters 5 and 7, we apply the Wong-Zakai approximation with the Brownian motion approximated by its spectral truncation. It is shown that the convergence of numerical schemes based on the Wong-Zakai approximation is determined by further discretization in space (Chapter 10) or in time (Chapter 4).

Fig. 1.2. Conceptual and chapter overview of this book.



2

Brownian motion and stochastic calculus

In this chapter, we review some basic concepts for stochastic processes and stochastic calculus as well as numerical integration methods in random space for obtaining statistics of stochastic processes.

We start from Gaussian processes and their representations in Chapter 2.1 and then introduce Brownian motion and its properties and approximations in Chapter 2.2. We discuss basic concepts in stochastic calculus: Ito integral in Chapter 2.3 and Ito formula in Chapter 2.4. We then focus on numerical integration methods in random space such as Monte Carlo methods, quasi-Monte Carlo methods, Wiener chaos method, and stochastic collocation method (sparse grid collocation method) in Chapter 2.5. Examples of applying these methods to a simple equation are provided in Chapter 2.5.5 with Matlab code. In Chapter 2.6 of bibliographic notes, we present a review on different types of approximation of Brownian motion and a brief review on *pros* and *cons* of numerical integration methods in random space. Various exercises are provided for readers to familiarize themselves with basic concepts presented in this chapter.

2.1 Gaussian processes and their representations

On a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ $(\Omega = \mathbb{R})$, if a cumulative distribution function of a random variable X is normal, i.e.,

$$\mathbb{P}(X < x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}\, dy, \quad \sigma > 0.$$

Z. Zhang, G.E. Karniadakis, *Numerical Methods for Stochastic Partial Differential Equations with White Noise*,
Applied Mathematical Sciences 196, DOI 10.1007/978-3-319-57511-7_2

then the random variable X is called a Gaussian (normal) random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Here X is completely characterized by its mean μ and its standard deviation σ. We denote $X \sim \mathcal{N}(\mu, \sigma^2)$. The probability density function of X is

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

When $\mu = 0$ and $\sigma = 1$, we call X a standard Gaussian (normal) random variable. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$, i.e., Z is a standard Gaussian (normal) random variable.

Example 2.1.1 *If X_i are mutually independent Gaussian random variables, then $\sum_{j=1}^{N} a_i X_i$ is a Gaussian random variable for any $a_i \in \mathbb{R}$. In particular, if $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are independent, then*

$$\alpha X_1 + \beta X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \alpha^2\sigma_1^2 + \beta^2\sigma_2^2).$$

Definition 2.1.2 (Gaussian random vector) *A $\mathbb{R}^n$-valued random vector $X = (X_1, X_2, \ldots, X_n)^\top$ has an n-variate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ if $X = \boldsymbol{\mu} + AZ$ where the matrix A is of size $n \times n$, $\Sigma = AA^\top$, and $Z = (Z_1, Z_2, \ldots, Z_n)^\top$ is a vector with independent standard Gaussian (normal) components.*

When $n = 1$, X is a (univariate) Gaussian random variable. The probability density of X is

$$p(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n}\,|\Sigma|^{1/2}} e^{-\frac{(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{2}}.$$

Example 2.1.3 *A set of random variables $\{X_i\}_{i=1}^n$ are called* jointly Gaussian *if $\sum_{i=1}^n a_i X_i$ is a Gaussian random variable for any $a_i \in \mathbb{R}$. Then $X = (X_1, X_2, \ldots, X_n)^\top$ is a Gaussian random vector.*

The correlation of two random variables (vectors) is a normalized version of the covariance, with values ranging from -1 to 1:

$$\mathrm{Corr}(X, Y) = \frac{\mathbb{C}\mathrm{ov}[(X, Y)]}{\sqrt{\mathbb{V}\mathrm{ar}[X]\mathbb{V}\mathrm{ar}[Y]}}, \quad \mathbb{C}\mathrm{ov}[(X, Y)] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^\top].$$

When $\mathrm{Corr}(X, Y) = 0$, we say X and Y are uncorrelated.

Definition 2.1.4 (Gaussian process) *A collection of random variables is called a Gaussian process, if the joint distribution of any finite number of its members is Gaussian. In other words, a Gaussian process is a $\mathbb{R}^d$-valued stochastic process with continuous time (or with index) t such that $(X(t_0), X(t_1), \cdots, X(t_n))^\top$ is a $n+1$-dimensional Gaussian random vector for any $0 \le t_0 < t_1 < \cdots < t_n$.*

The Gaussian process is denoted as $X = \{X(t)\}_{t\in T}$ where T is a set of indexes. Here $T = [0, \infty)$.

The consistency theorem of Kolmogorov [255, Theorem 2.2] implies that the finite dimensional distribution of a Gaussian stochastic process $X(t)$ is uniquely characterized by two functions: the mean function $\mu_t = \mathbb{E}[X(t)]$ and the covariance function $C(t, s) = \mathbb{C}\text{ov}[X(t), X(s)]$.

A Gaussian process $\{X(t)\}_{t\in T}$ is called a centered Gaussian process if the mean function $\mu(t) = \mathbb{E}[X(t)] = 0$ for all $t \in T$.

Given a function $\mu(t) : T \to \mathbb{R}$ and a nonnegative definite function $C(t, s) : T\times T \to \mathbb{R}$, there exists a Gaussian process $\{X(t)\}_{t\in T}$ with the mean function $\mu(t)$ and the covariance function $C(t, s)$.

To find such a Gaussian process, we can use the following expansion.

Theorem 2.1.5 (Karhunen-Loève expansion) *Let $X(t)$ be a Gaussian stochastic process defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $t \in [a, b]$, $-\infty < a < b < \infty$. Suppose that $X(t)$ has a continuous covariance function $C(t, s) = \mathbb{C}ov[(X(t), X(s))] = \mathbb{E}[(X(t) - \mathbb{E}[X(t)])(X(s) - \mathbb{E}[X(s)])]$. Then $X(t)$ admits the following representation*

$$X(t) = \mathbb{E}[X(t)] + \sum_{k=1}^{\infty} Z_k e_k(t),$$

where the convergence is in L^2, uniform in t (i.e., $\lim_{n\to\infty} \max_{t\in[a,b]} \mathbb{E}[(X(t) - \mathbb{E}[X(t)] - \sum_{k=1}^{n} Z_k e_k(t))^2] = 0$) and

$$Z_k = \int_a^b \big(X(t) - \mathbb{E}[X(t)]\big) e_k(t)\, dt.$$

Here the eigenfunctions e_k's of C_X with respective eigenvalues λ_k's form an orthonormal basis of $L^2([a, b])$ and

$$\int_a^b C(t, s) e_k(t)\, dt = \lambda_k e_k(s), \quad k \geq 1.$$

Furthermore, the random variables Z_k's have zero-mean, are uncorrelated, and have variance λ_k

$$\mathbb{E}[Z_k] = 0, \quad \textit{for all } k \geq 1 \qquad \textit{and} \qquad \mathbb{E}[Z_i Z_j] = \delta_{ij}\lambda_j, \quad \textit{for all } i, j \geq 1.$$

This is a direct application of Mercer's theorem [497] on a representation of a symmetric positive-definite function as a sum of a convergent sequence of product functions. The stochastic process $X(t)$ can be non-Gaussian.

The covariance function $C(t, s)$ can be represented as $C(t, s) = \sum_{k=1}^{\infty} \lambda_k e_k(t) e_k(s)$. The variance of $X(t)$ is the sum of the variances of the individual components of the sum:

$$\mathbb{V}\mathrm{ar}[X(t)] = \mathbb{E}[(X(t) - \mathbb{E}[X(t)])^2] = \sum_{k=0}^{\infty} e_k^2(t)\mathbb{V}\mathrm{ar}[Z_k] = \sum_{k=1}^{\infty} \lambda_k e_k^2(t).$$

Here Z_k are uncorrelated random variables.

The domain where the process is defined can be extended to domains in $\mathbb{R}^d$. In Table 2.1 we present a list of covariance functions commonly used in practice. Here the constant l is called correlation length, K_ν is the modified

Table 2.1. A list of covariance functions.

Wiener process	$\min(x, y),\ x, y \geq 0$
White Noise	$\sigma^2\delta(x-y), \quad x, y \in \mathbb{R}^d$
Gaussian	$\exp\left(-\frac{\lvert x-y\rvert^2}{2l^2}\right), \quad x, y \in \mathbb{R}^d$
Exponential	$\exp\left(-\frac{\lvert x-y\rvert}{l}\right), \quad x, y \in \mathbb{R}^d$
Matérn kernel	$\frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}\,\lvert x-y\rvert}{l}\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}\,\lvert x-y\rvert}{l}\right), \quad x, y \in \mathbb{R}^d$
Rational quadratic	$(1+\lvert x-y\rvert^2)^{-\alpha}, \quad x, y \in \mathbb{R}^d, \quad \alpha \geq 0$

Bessel function of order ν, and $\Gamma(\cdot)$ is the gamma function:

$$\Gamma(t) = \int_0^{\infty} x^{t-1}e^{-x}\,dx,\ t > 0.$$

Here are some examples of Karhunen-Loève expansion for Gaussian processes.

Example 2.1.6 (Brownian motion) *When $C(t, s) = \min(t, s)$, $t \in [0, 1]$, then the Gaussian process $X(t)$ can be written as*

$$X(t) = \sqrt{2}\sum_{k=1}^{\infty} \xi_k \frac{\sin\left(\left(k-\frac{1}{2}\right)\pi t\right)}{\left(k-\frac{1}{2}\right)\pi}.$$

Here ξ_k's are mutually independent standard Gaussian random variables. One can show that for $t, s \in [0, 1]$, the eigenvectors of the covariance function $\min(t, s)$ *are*

$$e_k(t) = \sqrt{2}\sin\left(\left(k-\tfrac{1}{2}\right)\pi t\right),$$

and the corresponding eigenvalues are

$$\lambda_k = \frac{1}{(k-\frac{1}{2})^2\pi^2}.$$

In the next section, we know that the process in Example 2.1.6 is actually a *Brownian motion*.

Example 2.1.7 (Brownian Bridge) *Let $X(t)$, $0 \le t \le 1$, be the Gaussian process in Example 2.1.6. Then $Y(t) = X(t) - tX(1)$, $0 \le t \le 1$, is also a Gaussian process and admits the following Karhunen-Loève expansion:*

$$Y(t) = \sum_{k=1}^{\infty} \eta_k \frac{\sqrt{2}\sin(k\pi t)}{k\pi}.$$

Here η_k's are mutually independent standard Gaussian random variables.

Example 2.1.8 (Ornstein-Uhlenbeck process) *Consider a centered one-dimensional Gaussian process with an exponential covariance function $\exp(-\frac{|t-s|}{l})$. The Karhunen-Loève expansion of such a Gaussian process over $[-a, a]$ is*

$$O(t) = \sum_{k=1}^{\infty} \xi_k \sqrt{\lambda_k} e_k(t),$$

where $\lambda_k = \frac{2l}{l^2\theta_k^2+1}$ and the corresponding eigenvalues are

$$e_{2i}(t) = \frac{\cos(\theta_{2i}t)}{\sqrt{2 + \frac{\sin(2\theta_{2i}a)}{2\theta_{2i}}}}, e_{2i-1}(t) = \frac{\sin(\theta_{2i-1}t)}{\sqrt{2 - \frac{\sin(2\theta_{2i-1}a)}{2\theta_{2i-1}}}}, \text{for all } i \ge 1,\ t \in [-a, a].$$

The θ_k's are solutions to the following transcendental equation

$$1 - l\theta\tan(a\theta) = 0 = l\theta + \tan(a\theta).$$

See p. 23, Section 2.3 of [155] or [245] for a derivation of such an expansion.

For more general forms of covariance functions $C(t, s)$, it may not be possible to find explicitly the eigenvectors and eigenvalues. The Karhunen-Loève expansion can be found numerically, and in practice only a finite number of terms in the expansion are required. Specifically, we usually perform a principal component analysis by truncating the sum at some N such that

$$\frac{\sum_{i=1}^{N} \lambda_i}{\sum_{i=1}^{\infty} \lambda_i} = \frac{\sum_{i=1}^{N} \lambda_i}{\int_a^b \mathbb{V}\mathrm{ar}[X(t)]\, dt} \ge \alpha.$$

Here α is typically taken as 0.9, 0.95, and 0.99. The eigenvalues and eigenfunctions are found by solving numerically the following eigenproblem (integral equation):

$$\int_a^b C(t, s)e_k(t)\, ds = \lambda_k e_k(s), \quad s \in [a, b] \text{ and } k = 1, 2, \ldots, N.$$

See, e.g., Section 2.3 of [155] or [419] for a Galerkin method for this problem. We can also apply the Nyström method or the quadrature method, where the integral is replaced with a representative weighted sum. Several numerical methods for representing a stochastic process with a given covariance kernel are presented in [308, Chapter 7], where numerical methods are not based on Karhunen-Loève expansion, but based on Fourier analysis and other methods.

The decay of eigenvalues in the Karhunen-Loève expansion depends on the smoothness of covariance functions.

Definition 2.1.9 ([419]) *A covariance function $C : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ is said to be piecewise analytic/smooth/$H^{p,q}$ on $\mathcal{D} \times \mathcal{D}$, $0 \le p, q \le \infty$ if there exists a partition $\mathfrak{D} = \{D_j\}_{j=1}^J$ into a finite sequence of simplexes D_j and a finite family $\mathfrak{G} = \{G_j\}_{j=1}^J$ of open sets in $\mathbb{R}^d$ such that*

$$\bar{D} = \cup_{j=1}^J \bar{D}_j,\ \bar{D}_j \subseteq G_j, \quad 1 \le j \le J$$

and such that $C|_{D_j \times D_j'}$ has an extension to $G_j \times G_{j'}$ which is analytic in $D_j \times D_{j'}$/is smooth in $D_j \times D_{j'}$/is in $H^p(D_j) \otimes H^q(D_{j'})$ for any pair (j, j').

The following conclusions on the eigenvalues in the Karhunen-Loève expansion are from [419].

Theorem 2.1.10 *Assume that $C \in L^2(\mathcal{D} \times \mathcal{D})$ be a symmetric covariance function which leads to a compact and nonnegative operator from $L^2(\mathcal{D})$ defined by $\mathfrak{C}u(x) = \int_{\mathcal{D}} C(x, y)u(y)\, dy$. If C is piecewise analytic on $\mathcal{D} \times \mathcal{D}$, then the eigenvalues λ_k in the Karhunen-Loève expansion satisfy that*

$$0 \le \lambda_k \le K_1 e^{-K_2 k^{1/d}}, \quad k \ge 1.$$

The constants K_1 and K_2 depend only on the covariance function C and the domain $\mathcal{D}$. If C is piecewise $H^p(\mathcal{D}) \otimes L^2(\mathcal{D})$ with $p \ge 1$, then the eigenvalues λ_k in the Karhunen-Loève expansion decay algebraically fast

$$0 \le \lambda_k \le K_3 k^{-p/d}, \quad k \ge 1.$$

For the Gaussian covariance function, $C(x, y) = \sigma^2 \exp(-(|x - y|^2)/\gamma^2/\mathrm{diam}(\mathcal{D}))$. Then the eigenvalues λ_k in the Karhunen-Loève expansion decay exponentially fast:

$$0 \le \lambda_k \le K_4 \gamma^{-k^{1/d}-2}/\Gamma(0.5k^{1/d}), \quad k \ge 1.$$

An different approach to show the decay of the eigenvalues is presented in [308, Chapter 7] using Fourier analysis for isotropic covariance kernels (the two-point covariance kernel depends only on distances of two points).

Theorem 2.1.11 ([419]) *Assume that the process $a(x, \omega)$ has a covariance function C, which is piecewise analytic/in $H^{p,q}$ on $\mathcal{D} \times \mathcal{D}$. Then the eigenfunctions are analytic/in H^p in each $\bar{D}_j \in \mathfrak{D}$.*

With further conditions on the domain D_j in $\mathfrak{D}$, it can be shown that the derivatives of eigenfunctions $e_k(x)$'s decay at the speed of $|\lambda_k|^{-s}$ when C is piecewise smooth where $s > 0$ is an arbitrary number.

2.2 Brownian motion and white noise

Definition 2.2.1 (One-dimensional Brownian motion) *A one-dimensional continuous time stochastic process $W(t)$ is called a standard Brownian motion if*

- *$W(t)$ is almost surely continuous in t,*
- *$W(t)$ has independent increments,*
- *$W(t) - W(s)$ obeys the normal distribution with mean zero and variance $t - s$.*
- *$W(0) = 0$.*

It can be readily shown that $W(t)$ is Gaussian process. We then call $\dot{W}(t) = \frac{d}{dt}W$, formally the first-order derivative of $W(t)$ in time, *white noise*.

By Example 2.1.6 and Exercise 2.7.7, then the Brownian motion $W(t)$, $t \in [0,1]$ can be represented by

$$W(t) = \sqrt{2}\sum_{i=1}^{\infty} \xi_i \frac{\sin\left(\left(i-\frac{1}{2}\right)\pi t\right)}{\left(i-\frac{1}{2}\right)\pi}, \quad t \in [0,1],$$

where ξ_i's are mutually independent standard Gaussian random variables. The Brownian motion and white noise can also be defined in terms of orthogonal expansions. Suppose that $\{m_k(t)\}_{k\geq 1}$ is a complete orthonormal system (CONS) in $L^2([0,T])$. The Brownian motion $W(t)$, $t \in [0,T]$ can be defined by (see, e.g., [315])

$$W(t) = \sum_{i=1}^{\infty} \xi_i \int_0^t m_i(s)\, ds, \ t \in [0,T], \tag{2.2.1}$$

where ξ_i's are mutually independent standard Gaussian random variables. It can be checked that the Gaussian process defined by (2.2.1) is indeed a standard Brownian motion by Definition 2.2.1. Correspondingly, the white noise is defined by

$$\dot{W}(t) = \sum_{i=1}^{\infty} \xi_i m_i(t), \ t \in [0,T]. \tag{2.2.2}$$

When $m_i(t) = \sqrt{2/T}\cos((i-1/2)\pi t/T)$, $i \geq 1$, then the representation (2.2.1) coincides with the *Karhunen-Loève expansion* of Brownian motion in Example 2.1.6 when $T = 1$.

Definition 2.2.2 (Multidimensional Brownian motion) *A continuous stochastic process $W_t = (W_1(t), \ldots, W_m(t))^\top$ is called an m-dimensional Brownian motion on $\mathbb{R}^m$ when $W_i(t)$ are mutually independent standard Brownian motions on $\mathbb{R}$.*

Definition 2.2.3 (Multidimensional Brownian motion, alternative definitions) *An $\mathbb{R}^d$-valued continuous Gaussian process $X(t)$ with mean function $\mu(t) = \mathbb{E}[X(t)]$ and the covariance function $C(t,s) = \mathbb{C}ov[(X(t), X(s))] = \mathbb{E}[(X(s) - \mu(s))(X(t) - \mu(t))^\top]$ is called a d-dimensional Brownian motion if for any $0 \leq t_0 < t_1 < \cdots < t_n$,*

- *$X(t_i)$ and $X(t_{i+1}) - X(t_i)$ are independent;*
- *the covariance function (a matrix) is a diagonal matrix with entries $\min(t_i, t_j)$, $0 \leq i, j \leq n$.*

When $\mu(t) = 0$ for all t and $C(t,s) = \min(t,s)$, the Gaussian process is called a standard Brownian motion.

2.2.1 Some properties of Brownian motion

Theorem 2.2.4 *The covariance $\mathbb{C}ov[(W(t), W(s))] = \mathbb{E}[W(t)W(s)] = \min(t,s)$.*

- *Time-homogeneity: For any $s > 0$, $\tilde{W}(t) = W(t+s) - W(s)$ is a Brownian motion, independent of $\sigma(W(u),\, u \leq s)$.*
- *Brownian scaling: For every $c > 0$, $cW(t/c^2)$ is a Brownian motion.*
- *Time inversion: Let $\tilde{W}(0) = 0$ and $\tilde{W}(t) = tW(1/t)$, $t > 0$. Then $\tilde{W}(t)$ is a Brownian motion.*

Corollary 2.2.5 (Strong law of large numbers for Brownian motion) *If $W(t)$ is a Brownian motion, then it holds almost surely that*

$$\lim_{t\to\infty} \frac{W(t)}{t} = 0.$$

Theorem 2.2.6 (Law of the iterated logarithm) *Let W_t be a standard Brownian motion. Then*

$$\mathbb{P}(\limsup_{t\to 0} \frac{W_t}{\sqrt{2t\,|\log\log(t)|}} = 1) = 1, \quad \mathbb{P}(\liminf_{t\to 0} \frac{W_t}{\sqrt{2t\,|\log\log(t)|}} = -1) = 1.$$

$$\mathbb{P}(\limsup_{t\to \infty} \frac{W_t}{\sqrt{2t\log\log(t)}} = 1) = 1, \quad \mathbb{P}(\liminf_{t\to \infty} \frac{W_t}{\sqrt{2t\log\log(t)}} = -1) = 1.$$

Example 2.2.7 (Ornstein-Uhlenbeck process) *Consider a centered one-dimensional Gaussian process with exponential covariance function $\exp(-\frac{|t-s|}{\sigma})$. The Gaussian process is usually called a Ornstein-Uhlenbeck process. Suppose that $W(t)$ is a standard Brownian motion. For $t \geq 0$, the Ornstein-Uhlenbeck process can be written as*

$$O(t) = e^{-\frac{t}{\sigma}} W(e^{\frac{2t}{\sigma}}).$$

Example 2.2.8 *The Brownian bridge $X(t)$ is a one-dimensional Gaussian process with time $t \in [0,1]$ and covariance $\mathbb{C}ov[(X(t), X(s))] = \min(t,s) - ts = \begin{cases} s(1-t),\ 0 \le s \le t \le 1. \\ t(1-s),\ 0 \le t \le s \le 1. \end{cases}$*

Suppose that $W(t)$ is a standard Brownian motion. Then $X(t)$ can be represented by

$$X(t) = W(t) - tW(1) = t\big(W(t) - W(1)\big) + (1-t)\big(W(t) - W(0)\big), \quad 0 \le t \le 1.$$

The process $X(t)$ bridges $W(t) - W(1)$ and $W(t) - W(0)$. It can be readily verified that $\mathbb{C}\text{ov}[(X(t),\, X(s))] = \min(t,s) - ts$ and $X(t)$ is continuous and starting from 0. Moreover,

$$W(t) = (t+1)X(\frac{t}{t+1}), \quad X(t) = (1-t)W(\frac{t}{1-t}).$$

Regularity of Brownian motion

For deterministic functions, $f(x)$, $x \in \mathbb{R}$ is Hölder continuous of order α if and only if there exists a constant C such that

$$|f(x+h) - f(x)| \le Ch^{\alpha}, \quad \text{for all } h > 0 \text{ and all } x.$$

When $\alpha = 1$, we call it Lipschitz continuous. When C depends on x, then we call it locally Hölder continuous of order α

$$|f(x+h) - f(x)| \le C(x)h^{\alpha}, \quad \text{for all small enough } h > 0.$$

Definition 2.2.9 *Consider two stochastic processes, $X(t)$ and $Y(t)$, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We call $Y(t)$ a modification (or version) of $X(t)$ if for every $t \ge 0$, we have*

$$\mathbb{P}(X(t) = Y(t)) = 1.$$

Theorem 2.2.10 (Kolmogorov and Centsov continuity theorem, [255, Section 2.2.B]) *Given a stochastic process $X(t)$ with $t \in [a,b]$, if there exist constants $p > r$, $K > 0$ such that*

$$\mathbb{E}[|X(t) - X(s)|^p] \le K\,|t-s|^{1+r}\,, \ \textit{for } t, s \in [a,b],$$

then $X(t)$ has a modification $Y(t)$ which is almost everywhere (in ω) continuous: for all $t, s \in [a,b]$,

$$|Y(t,\omega) - Y(s,\omega)| \le C(\omega)\,|t-s|^{\alpha}\,,\ 0 < \alpha < \frac{r}{p}.$$

For $X(\boldsymbol{t})$, $\boldsymbol{t} \in T \subseteq \mathbb{R}^d$, if there exist constants $p > r$, K such that

$$\mathbb{E}[|X(\boldsymbol{t}) - X(\boldsymbol{s})|^p] \leq K\,|\boldsymbol{t} - \boldsymbol{s}|^{d+r}\,, \text{ for } \boldsymbol{t}, \boldsymbol{s} \in T,$$

then $X(\boldsymbol{t})$ has a modification $Y(\boldsymbol{t})$ which is almost everywhere in ω continuous: for all $\boldsymbol{t}$, $\boldsymbol{s} \in T$,

$$\mathbb{E}[\big(\sup_{\boldsymbol{s}\neq\boldsymbol{t}} \frac{|Y(\boldsymbol{t},\omega) - Y(\boldsymbol{s},\omega)|^{\alpha}}{|\boldsymbol{t} - \boldsymbol{s}|}\big)^p] < \infty,\ 0 < \alpha < \frac{r}{p}.$$

Theorem 2.2.11 *For $\alpha < \frac{1}{2}$, the Brownian motion has a modification which is of locally Hölder continuous of order α.*

Proof. For integer $n \geq 1$, by Kolmogorov and Centsov continuity theorem, it only requires to show that

$$\mathbb{E}[|W(t) - W(s)|^{2n}] \leq C_n\,|t - s|^n\,.$$

Then recalling the conclusion from Exercise 2.7.1 leads to the conclusion.

Theorem 2.2.12 ([255, Section 2.9.D]) *The Brownian motion is nowhere differentiable: for almost all ω, the sample path (realization, trajectory) $W(t,\omega)$ is nowhere differentiable as function of t. Moreover, for almost all ω, the path (realization, trajectory) $W(t,\omega)$ is nowhere Hölder continuous with exponent $\alpha > \frac{1}{2}$.*

Definition 2.2.13 (p-variation) *The p-variation of a real-valued function f, defined on an interval $[a,b] \subset \mathbb{R}$, is the quantity*

$$|f|_{p,\mathrm{TV}} = \sup_{\Pi_n,\,|\pi_n|\to 0} \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)|^p,$$

where the supremum runs over the set of all partitions Π_n of the given interval.

Theorem 2.2.14 (Unbounded total variation of Brownian motion) *The paths (realizations, trajectories) of Brownian motion are of infinite total variation almost surely (a.s., with probability one).*

Proof. Without loss of generality, let's consider the interval $[0,1]$.

$$|W|_{1,\mathrm{TV}} = \sup_{\Pi_n} \sum_{i=0}^{n-1} |W(t_{i+1}) - W(t_i)| \geq \sum_{i=0}^{n-1} |W(\frac{i+1}{n}) - W(\frac{i}{n})| =: V_n.$$

Denote by $W(\frac{i+1}{n}) - W(\frac{i}{n}) = \frac{\xi_i}{\sqrt{n}}$. Then ξ_i's are i.i.d. $\mathcal{N}(0,1)$ random variables. Observe that $\mathbb{E}[V_n] = \sqrt{n}\mathbb{E}[|\xi_1|]$ and $\mathbb{V}\mathrm{ar}[V_n] = 1 - (\mathbb{E}[|\xi_1|])^2$.

Then it follows from the Chebyshev inequality (see Appendix D), we have

$$\begin{aligned}\mathbb{P}(V_n \geq \frac{1}{2}\mathbb{E}[|\xi_1|]\sqrt{n}) &= \mathbb{P}(V_n - \mathbb{E}[|\xi_1|]\sqrt{n} \geq -\frac{1}{2}\mathbb{E}[|\xi_1|]\sqrt{n}) \\ &\geq 1 - \mathbb{P}(\big|V_n - \mathbb{E}[|\xi_1|]\sqrt{n}\big| \geq \frac{1}{2}\mathbb{E}[|\xi_1|]\sqrt{n}) \\ &\geq 1 - \frac{\mathbb{V}\mathrm{ar}[V_n]}{\big(\frac{1}{2}\mathbb{E}[|\xi_1|]\sqrt{n}\big)^2} = 1 - 4\frac{1-(\mathbb{E}[|\xi_1|])^2}{n(\mathbb{E}[|\xi_1|])^2}.\end{aligned}$$

Thus we have

$$\mathbb{P}(|W|_{1,\mathrm{TV}} \geq \frac{\mathbb{E}[|\xi_1|]}{2}\sqrt{n}) \geq \mathbb{P}(V_n \geq \frac{\mathbb{E}[|\xi_1|]}{2}\sqrt{n}) = 1 - 4\frac{1-(\mathbb{E}[|\xi_1|])^2}{n(\mathbb{E}[|\xi_1|])^2}.$$

Letting $n \to \infty$, we obtain

$$\mathbb{P}(|W|_{1,\mathrm{TV}} = \infty) = 1.$$

2.2.2 Approximation of Brownian motion

According to the representation in Chapter 2.2, we have at least three approximate representations for Brownian motion by a finite number of random variables.

By Definition 2.2.1, the Brownian motion at time $\mathtt{t}_{n+1}$ can be approximated by

$$\sum_{i=0}^{n} \Delta W_i = \sum_{i=0}^{n} \sqrt{\Delta \mathtt{t}_i}\xi_i, \text{ where } \Delta W_i = W(\mathtt{t}_{i+1}) - W(\mathtt{t}_i), \text{ and } \Delta \mathtt{t}_i = \mathtt{t}_{i+1} - \mathtt{t}_i, \tag{2.2.3}$$

where ξ_i's are i.i.d. standard Gaussian random variables. A sample path (realization, trajectory) of Brownian motion is illustrated in Figure 2.1. Here is Matlab code for generating Figure 2.1.

Code 2.1. A sample path of Brownian motion.

```
% One realization of W(t) at time grids k*dt
clc, clear all
t =2.5;
n= 1000;
dt = t / n;
%  Increments of Brownian motion
Winc  = zeros ( n + 1,1);
% Declare the status of random number generator -- Mersenne
% Twister
rng(100, 'twister');
Winc(1:n) = sqrt ( dt ) * randn ( n, 1);
%  Brownian motion - cumulative sum of all previous
%  increments
```

```
W(2:n+1,1) = cumsum ( Winc(1:n) );
figure(10)
plot((1:n+1).'*dt,W,'b-','Linewidth',2);
xlabel('t')
ylabel('W(t)')
axis tight
```

One popular approximation of Brownian motion in continuous time is piecewise linear approximation (also known as polygonal approximation, see, e.g., [457, 481, 482] or [241, p. 396]), i.e.,

$$W^{(n)}(t) = W(\mathsf{t}_i) + (W(\mathsf{t}_{i+1}) - W(\mathsf{t}_i))\frac{t - \mathsf{t}_i}{\mathsf{t}_{i+1} - \mathsf{t}_i}, \quad t \in [\mathsf{t}_i, \mathsf{t}_{i+1}). \tag{2.2.4}$$

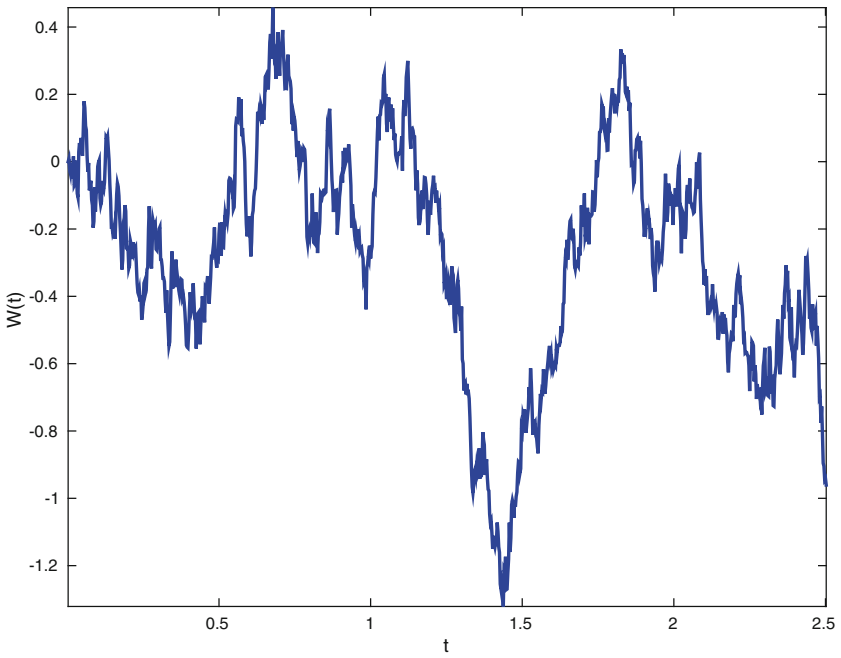
Another way to approximate Brownian motion is by a truncated *orthogonal expansion*:

$$W^{(\mathsf{n})}(t) = \sum_{i=1}^{\mathsf{n}} \xi_i \int_0^t m_i(s)\, ds, \quad \xi_j = \int_0^T m_j(t)\, dW, \quad t \in [0, T], \tag{2.2.5}$$

where $\{m_i(t)\}$ is a CONS in $L^2([0,T])$ and ξ_j are mutually independent standard Gaussian random variables.

In this book, we mostly use the cosine basis $\{m_l(s)\}_{l\geq 1}$ given by

Fig. 2.1. An illustration of a sample path of Brownian motion using cumulative summation of increments.



$$m_1(s) = \frac{1}{\sqrt{T}}, \quad m_l(s) = \sqrt{\frac{2}{T}} \cos(\frac{\pi(l-1)s}{T}), \quad l \geq 2, \quad 0 \leq s \leq T, \quad (2.2.6)$$

or a piecewise version of spectral expansion (2.2.5) by taking a partition of $[0, T]$, e.g., $0 = \mathsf{t}_0 < \mathsf{t}_1 < \cdots \mathsf{t}_{\mathsf{K}-1} < \mathsf{t}_{\mathsf{K}} = T$. We then have

$$W^{(\mathsf{n},\mathsf{K})}(t) = \sum_{k=1}^{\mathsf{K}} \sum_{l=1}^{\mathsf{n}} M_{k,l}(t)\xi_{k,l}, \quad \xi_{k,l} = \int_{I_k} m_{k,l}(s)\, dW(s) \qquad (2.2.7)$$

where $M_{k,l}(t) = \int_{\mathsf{t}_{k-1}}^{\mathsf{t}_k \wedge t} m_{k,l}(s)\, ds$ $(\mathsf{t}_k \wedge t = \min(\mathsf{t}_k, t))$, $\{m_{k,l}\}_{l=1}^{\infty}$ is a CONS in $L^2(I_k)$ and $I_k = [\mathsf{t}_{k-1}, \mathsf{t}_k)$. The random variables $\xi_{k,l}$ are i.i.d. standard Gaussian random variables. Sometimes (2.2.7) is written as

$$W^{(\mathsf{n},\mathsf{K})}(t) = \sum_{k=1}^{\mathsf{K}} \sum_{l=1}^{\mathsf{n}} \int_0^t \mathbf{1}_{I_k}(s) m_{k,l}(s)\, ds \xi_{k,l}, \qquad (2.2.8)$$

where $\mathbf{1}_{\cdot}$ is the indicator function.

Here different choices of CONS lead to different representations. The orthonormal piecewise constant basis over time interval $I_k = [\mathbf{t}_{k-1}, \mathbf{t}_k)$, with $h_k = (\mathbf{t}_k - \mathbf{t}_{k-1})/\sqrt{\mathsf{n}}$, is

$$m_{k,l}(t) = \frac{1}{\sqrt{h}} \chi_{[\mathbf{t}_{k-1}+(l-1)h_k, \mathbf{t}_{k-1}+lh_k)}, \quad l = 1, 2, \cdots, \mathsf{n}. \qquad (2.2.9)$$

When $\mathsf{n} = 1$, this basis gives the classical piecewise linear interpolation (2.2.4). The orthonormal Fourier basis gives Wiener's representation (see, e.g., [259, 358, 391]):

$$m_{k,1} = \frac{1}{\sqrt{\mathbf{t}_k - \mathbf{t}_{k-1}}}, \quad m_{k,2l} = \sqrt{\frac{2}{\mathbf{t}_k - \mathbf{t}_{k-1}}} \sin\big(2l\pi \frac{t - \mathbf{t}_{k-1}}{\mathbf{t}_k - \mathbf{t}_{k-1}}\big),$$
$$m_{k,2l+1}(t) = \sqrt{\frac{2}{\mathbf{t}_k - \mathbf{t}_{k-1}}} \cos\big(2l\pi \frac{t - \mathbf{t}_{k-1}}{\mathbf{t}_k - \mathbf{t}_{k-1}}\big), \; t \in [\mathbf{t}_{k-1}, \mathbf{t}_k). \quad (2.2.10)$$

Note that taking only $m_{k,1}$ $(\mathsf{n} = 1)$ in (2.2.10) again leads to the piecewise linear interpolation (2.2.4). Besides, we can also use the Haar wavelet basis, which gives the Levy-Ciesielsky representation [255].

Remark 2.2.15 *Once we have a formal representation (approximation) of Brownian motion, we then can readily obtain a formal representation (approximation) of white noise and thus we skip the formulas for white noise.*

Lemma 2.2.16 *Consider a uniform partition of $[0, T]$, i.e., $\mathsf{t}_k = k\Delta$, $\mathsf{K}\Delta = T$. For $t \in [0, T]$, there exists a constant $C > 0$ such that*

$$\mathbb{E}[\big(W(t) - W^{(\mathsf{n},\mathsf{K})}(t)\big)^2] \leq C\frac{\Delta}{n},$$

and for sufficient small $\epsilon > 0$

$$\left|W(t) - W^{(\mathsf{n},\mathsf{K})}(t)\right| \le \mathcal{O}\left(\left(\frac{\Delta}{n}\right)^{1/2-\epsilon}\right).^{1} \tag{2.2.11}$$

For $t = \mathsf{t}_k$*, we have*

$$W(\mathsf{t}_k) - W^{(\mathsf{n},\mathsf{K})}(\mathsf{t}_k) = 0, \tag{2.2.12}$$

if the CONS $\{m_{k,l}\}_{l=1}^{\infty}$ *contains* $\frac{1}{\sqrt{\mathsf{t}_k - \mathsf{t}_{k-1}}}$ *as its elements, i.e.,* $\displaystyle\int_{\mathsf{t}_{k-1}}^{\mathsf{t}_k} m_{k,l}(s) \frac{1}{\sqrt{\mathsf{t}_k - \mathsf{t}_{k-1}}}\, ds = \delta_{l,1}$.

Proof. By the spectral approximation of $W(t)$ (2.2.7) and the fact that $\xi_{k,l,i}$ are i.i.d., we have

$$\begin{aligned}
\mathbb{E}[(W(t) - W^{(\mathsf{n},\mathsf{K})}(t))^2] &= \sum_{k=1}^{\mathsf{K}} \sum_{l=\mathsf{n}+1}^{\infty} \left(\int_{\mathsf{t}_{k-1}}^{t\wedge\mathsf{t}_k} m_{k,l}(s)\, ds\right)^2 \\
&= \sum_{k=1}^{\mathsf{K}} \sum_{l=\mathsf{n}+1}^{\infty} \left(\int_{\mathsf{t}_{k-1}}^{\mathsf{t}_k} \chi_{[0,t]}(s) m_{k,l}(s)\, ds\right)^2 \\
&\le C\frac{\Delta}{n},
\end{aligned}$$

where $\mathsf{t}_k \wedge t = \min(\mathsf{t}_k, t)$ and we have applied the standard estimate for L^2-projection using piecewise orthonormal basis $m_{k,l}(s)$, see, e.g., [417] and have used the fact that the indicator function $\chi_{[0,t]}(s)$ belongs to the Sobolev space $H^{1/2}((0,T))$ for any $T > t$.

Once we have the L^2-estimate, we can apply the Borel-Cantelli lemma (see Appendix D) to obtain the almost sure (a.s.) convergence (2.2.11).

If $t = \mathsf{t}_k$, we have $\int_{\mathsf{t}_{k-1}}^{\mathsf{t}_k} m_{k,l}(s)\, ds = 0$ for any $l \ge 2$ if $m_{k,1} = \frac{1}{\sqrt{\mathsf{t}_k - \mathsf{t}_{k-1}}}$. and thus (2.2.12) holds.

Though any CONS in $L^2([0,T])$ can be used in the spectral approximation (2.2.5), we use a CONS containing a constant in the basis. Consequently, we have the following relation

$$\int_{t_n}^{t_{n+1}} dW^{(\mathsf{n})}(t) = \Delta W_n, \quad \Delta W_n = W(t_{n+1}) - W(t_n). \tag{2.2.13}$$

We will use these approximations in most of the chapters in the book for Wong-Zakai approximation.

[1] The big "$\mathcal{O}$" implies that the error is bounded by a positive constant times the term in the parenthesis.