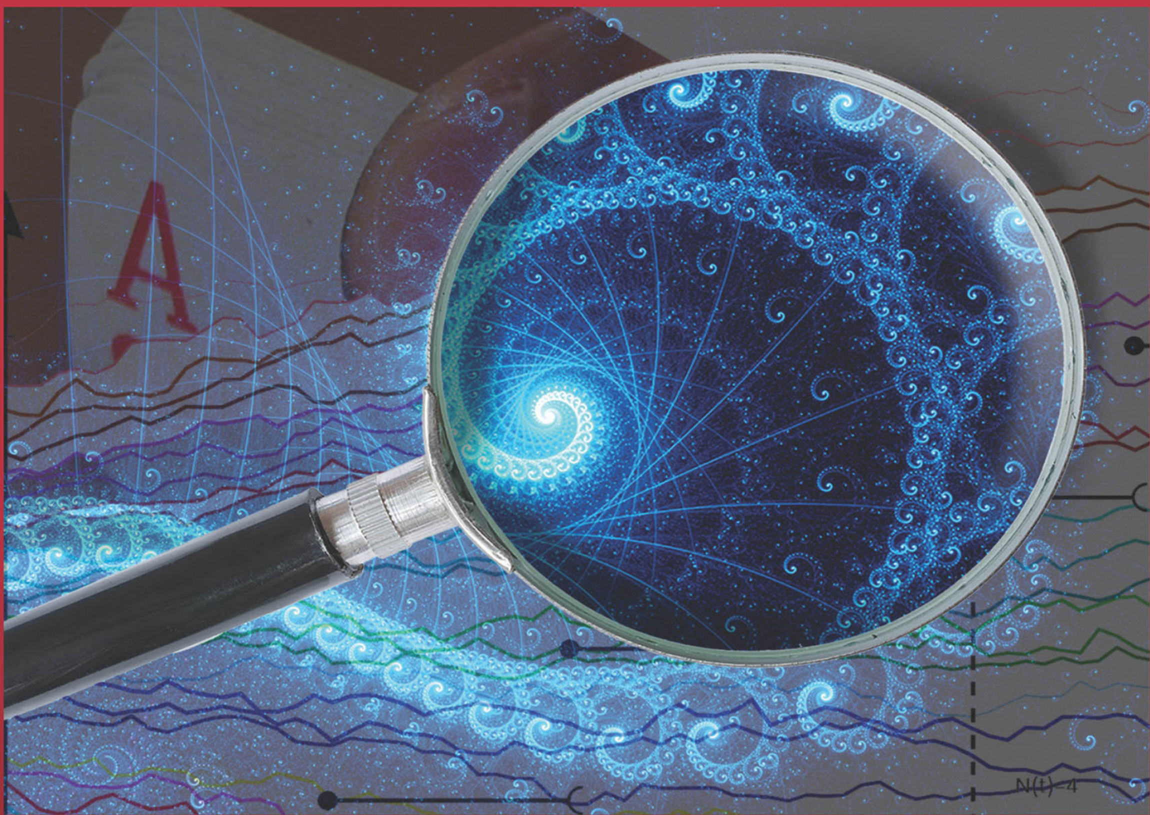
Texts in Statistical Science

Stochastic Processes

From Applications to Theory



Pierre Del Moral • Spiridon Penev

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

WITH VITALSOURCE® EBOOK

Stochastic Processes

From Applications to Theory

CHAPMAN&HA LL/CRC
TextsinStatis ticalSc ienceSe ries

SeriesEditors
FrancescaD ominici, *HarvardSc hoolofPublic H ealth,U SA*
JulianJ .F araway, *UniversityofBath,U K*
Martin Tanner,*N orthwesternU niversity,U SA*
JimZide k,*U niversityofBr itishC olumbia,C anada*

Statistical Theory: A Concise Introduction
F. Abramovich and Y. Ritov

Practical Multivariate Analysis, Fifth Edition
A. Afifi, S. May, and V.A. Clark

Practical Statistics for Medical Research
D.G. Altman

Interpreting Data: A First Course in Statistics
A.J.B. Anderson

Introduction to Probability with R
K. Baclawski

Linear Algebra and Matrix Analysis for Statistics
S. Banerjee and A. Roy

Mathematical Statistics: Basic Ideas and Selected Topics, Volume I, Second Edition
P. J. Bickel and K. A. Doksum

Mathematical Statistics: Basic Ideas and Selected Topics, Volume II
P. J. Bickel and K. A. Doksum

Analysis of Categorical Data with R
C. R. Bilder and T. M. Loughin

Statistical Methods for SPC and TQM
D. Bissell

Introduction to Probability
J. K. Blitzstein and J. Hwang

Bayesian Methods for Data Analysis, Third Edition
B.P. Carlin and T.A. Louis

Second Edition
R. Caulcutt

The Analysis of Time Series: An Introduction, Sixth Edition
C. Chatfield

Introduction to Multivariate Analysis
C. Chatfield and A.J. Collins

Problem Solving: A Statistician's Guide, Second Edition
C. Chatfield

Statistics for Technology: A Course in Applied Statistics, Third Edition
C. Chatfield

Analysis of Variance, Design, and Regression : Linear Modeling for Unbalanced Data, Second Edition
R. Christensen

Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians
R. Christensen, W. Johnson, A. Branscum, and T.E. Hanson

Modelling Binary Data, Second Edition
D. Collett

Modelling Survival Data in Medical Research, Third Edition
D. Collett

Introduction to Statistical Methods for Clinical Trials
T.D. Cook and D.L. DeMets

Applied Statistics: Principles and Examples
D.R. Cox and E.J. Snell

Multivariate Survival Analysis and Competing Risks
M. Crowder

Statistical Analysis of Reliability Data
M.J. Crowder, A.C. Kimber, T.J. Sweeting, and R.L. Smith

An Introduction to Generalized Linear Models, Third Edition
A.J. Dobson and A.G. Barnett

Nonlinear Time Series: Theory, Methods, and Applications with R Examples
R. Douc, E. Moulines, and D.S. Stoffer

Introduction to Optimization Methods and Their Applications in Statistics
B.S. Everitt

Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models, Second Edition
J.J. Faraway

Linear Models with R, Second Edition
J.J. Faraway

A Course in Large Sample Theory
T.S. Ferguson

Multivariate Statistics: A Practical Approach
B. Flury and H. Riedwyl

Readings in Decision Analysis
S. French

Discrete Data Analysis with R: Visualization and Modeling Techniques for Categorical and Count Data
M. Friendly and D. Meyer

Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Second Edition
D. Gamerman and H.F. Lopes

Bayesian Data Analysis, Third Edition
A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin

Multivariate Analysis of Variance and Repeated Measures: A Practical Approach for Behavioural Scientists
D.J. Hand and C.C. Taylor

Practical Longitudinal Data Analysis
D.J. Hand and M. Crowder

Logistic Regression Models
J.M. Hilbe

Richly Parameterized Linear Models: Additive, Time Series, and Spatial Models Using Random Effects
J.S. Hodges

Statistics for Epidemiology
N.P. Jewell

Stochastic Processes: An Introduction, Second Edition
P.W. Jones and P. Smith

The Theory of Linear Models
B. Jørgensen

Pragmatics of Uncertainty
J.B. Kadane

Principles of Uncertainty
J.B. Kadane

Graphics for Statistics and Data Analysis with R
K.J. Keen

Mathematical Statistics
K. Knight

Introduction to Multivariate Analysis: Linear and Nonlinear Modeling
S. Konishi

Nonparametric Methods in Statistics with SAS Applications
O. Korosteleva

Modeling and Analysis of Stochastic Systems, Second Edition
V.G. Kulkarni

Exercises and Solutions in Biostatistical Theory
L.L. Kupper, B.H. Neelon, and S.M. O'Brien

Exercises and Solutions in Statistical Theory
L.L. Kupper, B.H. Neelon, and S.M. O'Brien

Design and Analysis of Experiments with R
J. Lawson

Design and Analysis of Experiments with SAS
J. Lawson

A Course in Categorical Data Analysis
T. Leonard

Statistics for Accountants
S. Letchford

Introduction to the Theory of Statistical Inference
H. Liero and S. Zwanzig

Statistical Theory, Fourth Edition
B.W. Lindgren

Stationary Stochastic Processes: Theory and Applications
G. Lindgren

Statistics for Finance
E. Lindström, H. Madsen, and J. N. Nielsen

The BUGS Book: A Practical Introduction to Bayesian Analysis
D. Lunn, C. Jackson, N. Best, A. Thomas, and D. Spiegelhalter

Introduction to General and Generalized Linear Models
H. Madsen and P. Thyregod

Time Series Analysis
H. Madsen

Pólya Urn Models
H. Mahmoud

Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition
B.F.J. Manly

Introduction to Randomized Controlled Clinical Trials, Second Edition
J.N.S. Matthews

Statistical Rethinking: A Bayesian Course with Examples in R and Stan
R. McElreath

Statistical Methods in Agriculture and Experimental Biology, Second Edition
R. Mead, R.N. Curnow, and A.M. Hasted

Statistics in Engineering: A Practical Approach
A.V. Metcalfe

Statistical Inference: An Integrated Approach, Second Edition
H. S. Migon, D. Gamerman, and F. Louzada

Beyond ANOVA: Basics of Applied Statistics
R.G. Miller, Jr.

A Primer on Linear Models
J.F. Monahan

Stochastic Processes: From Applications to Theory
P.D Moral and S. Penev

Applied Stochastic Modelling, Second Edition
B.J.T. Morgan

Elements of Simulation
B.J.T. Morgan

Probability: Methods and Measurement
A. O'Hagan

Introduction to Statistical Limit Theory
A.M. Polansky

Applied Bayesian Forecasting and Time Series Analysis
A. Pole, M. West, and J. Harrison

Statistics in Research and Development, Time Series: Modeling, Computation, and Inference
R. Prado and M. West

Essentials of Probability Theory for Statisticians
M.A. Proschan and P.A. Shaw

Introduction to Statistical Process Control
P. Qiu

Sampling Methodologies with Applications
P.S.R.S. Rao

A First Course in Linear Model Theory
N. Ravishanker and D.K. Dey

Essential Statistics, Fourth Edition
D.A.G. Rees

Stochastic Modeling and Mathematical Statistics: A Text for Statisticians and Quantitative Scientists
F.J. Samaniego

Statistical Methods for Spatial Data Analysis
O. Schabenberger and C.A. Gotway

Bayesian Networks: With Examples in R
M. Scutari and J.-B. Denis

Large Sample Methods in Statistics
P.K. Sen and J. da Motta Singer

Spatio-Temporal Methods in Environmental Epidemiology
G. Shaddick and J.V. Zidek

Decision Analysis: A Bayesian Approach
J.Q. Smith

Analysis of Failure and Survival Data
P. J. Smith

Applied Statistics: Handbook of GENSTAT Analyses
E.J. Snell and H. Simpson

Applied Nonparametric Statistical Methods, Fourth Edition
P. Sprent and N.C. Smeeton

Data Driven Statistical Methods
P. Sprent

Generalized Linear Mixed Models: Modern Concepts, Methods and Applications
W. W. Stroup

Survival Analysis Using S: Analysis of Time-to-Event Data
M. Tableman and J.S. Kim

Applied Categorical and Count Data Analysis
W. Tang, H. He, and X.M. Tu

Elementary Applications of Probability Theory, Second Edition
H.C. Tuckwell

Introduction to Statistical Inference and Its Applications with R
M.W. Trosset

Understanding Advanced Statistical Methods
P.H. Westfall and K.S.S. Henning

Statistical Process Control: Theory and Practice, Third Edition
G.B. Wetherill and D.W. Brown

Generalized Additive Models: An Introduction with R
S. Wood

Epidemiology: Study Design and Data Analysis, Third Edition
M. Woodward

Practical Data Analysis for Designed Experiments
B.S. Yandell

Texts in Statistical Science

Stochastic Processes

From Applications to Theory

Pierre Del Moral

University of New South Wales
Sydney, Australia
and
INRIA Sud Ouest Research Center
Bordeaux, France

Spiridon Penev

University of New South Wales
Sydney, Australia

With illustrations by Timothée Del Moral

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2014 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper
Version Date: 20161005

International Standard Book Number-13: 978-1-4987-0183-9 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

To Laurence, Tiffany and Timothée;
to Tatiana, Iva and Alexander.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Contents

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

0

Introduction

A brief discussion on stochastic processes

These lectures deal with the foundations and the applications of stochastic processes also called random processes.

The term "stochastic" comes from the Greek word "stokhastikos" which means a skillful person capable of guessing and predicting. The first use of this term in probability theory can be traced back to the Russian economist and statistician Ladislaus Bortkiewicz (1868-1931). In his paper *Die Iterationen* published in 1917, he defines the term "stochastik" as follows: "The investigation of empirical varieties, which is based on probability theory, and, therefore, on the law of the large numbers, may be denoted as stochastic. But stochastic is not simply probability theory, but above all probability theory and its applications".

As their name indicates, stochastic processes are dynamic evolution models with random ingredients. Leaving aside the old well-known dilemma of determinism and freedom [26, 225] (solved by some old fashion scientific reasoning which cannot accommodate any piece of random "un-caused causations"), the complex structure of any "concrete" and sophisticated real life model is always better represented by stochastic mathematical models as a first level of approximation.

The sources of randomness reflect different sources of model uncertainties, including unknown initial conditions, model uncertainties such as misspecified kinetic parameters, as well as the external random effects on the system.

Stochastic modelling techniques are of great importance in many scientific areas, to name a few:

- Computer and engineering sciences: signal and image processing, filtering and inverse problems, stochastic control, game theory, mathematical finance, risk analysis and rare event simulation, operation research and global optimization, artificial intelligence and evolutionary computing, queueing and communication networks.

- Statistical machine learning: hidden Markov chain models, frequentist and Bayesian statistical inference.

- Biology and environmental sciences: branching processes, dynamic population models, genetic and genealogical tree-based evolution.

- Physics and chemistry: turbulent fluid models, disordered and quantum models, statistical physics and magnetic models, polymers in solvents, molecular dynamics and Schrödinger equations.

A fairly large body of the literature on stochastic processes is concerned with the probabilistic modelling and the convergence analysis of random style dynamical systems. As expected these developments are closely related to the theory of dynamical systems, to partial differential and integro-differential equations, but also to ergodic and chaos theory, differential geometry, as well as the classical linear algebra theory, combinatorics, topology, operator theory, and spectral analysis.

The theory of stochastic processes has also developed its own sophisticated probabilistic tools, including diffusion and jump type stochastic differential equations, Doeblin-Itō calculus, martingale theory, coupling techniques, and advanced stochastic simulation techniques.

One of the objectives of stochastic process theory is to derive explicit analytic type results for a variety of simplified stochastic processes, including birth and death models, simple random walks, spatially homogeneous branching processes, and many others. Nevertheless, most of the more realistic processes of interest in the real world of physics, biology, and engineering sciences are "unfortunately" highly nonlinear systems evolving in very high dimensions. As a result, it is generally impossible to come up with any type of closed form solutions.

In this connection, the theory of stochastic processes is also concerned with the modelling and with the convergence analysis of a variety of sophisticated stochastic algorithms. These stochastic processes are designed to solve numerically complex integration problems that arise in a variety of application areas. Their common feature is to mimic and to use repeated random samples of a given stochastic process to estimate some *averaging* type property using empirical approximations.

We emphasize that all of these stochastic methods are expressed in terms of a particular stochastic process, or a collection of stochastic processes, depending on the precision parameter being the time horizon, or the number of samples. The central idea is to approximate the expectation of a given random variable using the empirical averages (in space or in time) based on a sequence of random samples. In this context, the rigorous analysis of these complex numerical schemes also relies on advanced stochastic analysis techniques.

The interpretation of the random variables of interest depends on the application. In some instances, these variables represent the random states of a complex stochastic process, or some unknown kinetic or statistical parameters. In other situations, the random variables of interest are specified by a complex target probability measure. Stochastic simulation and Monte Carlo methods are used to predict the evolution of given random phenomena, as well as to solve several estimation problems using random searches. To name but a few: computing complex measures and multidimensional integrals, counting, ranking, and rating problems, spectral analysis, computation of eigenvalues and eigenvectors, as well as solving linear and nonlinear integro-differential equations.

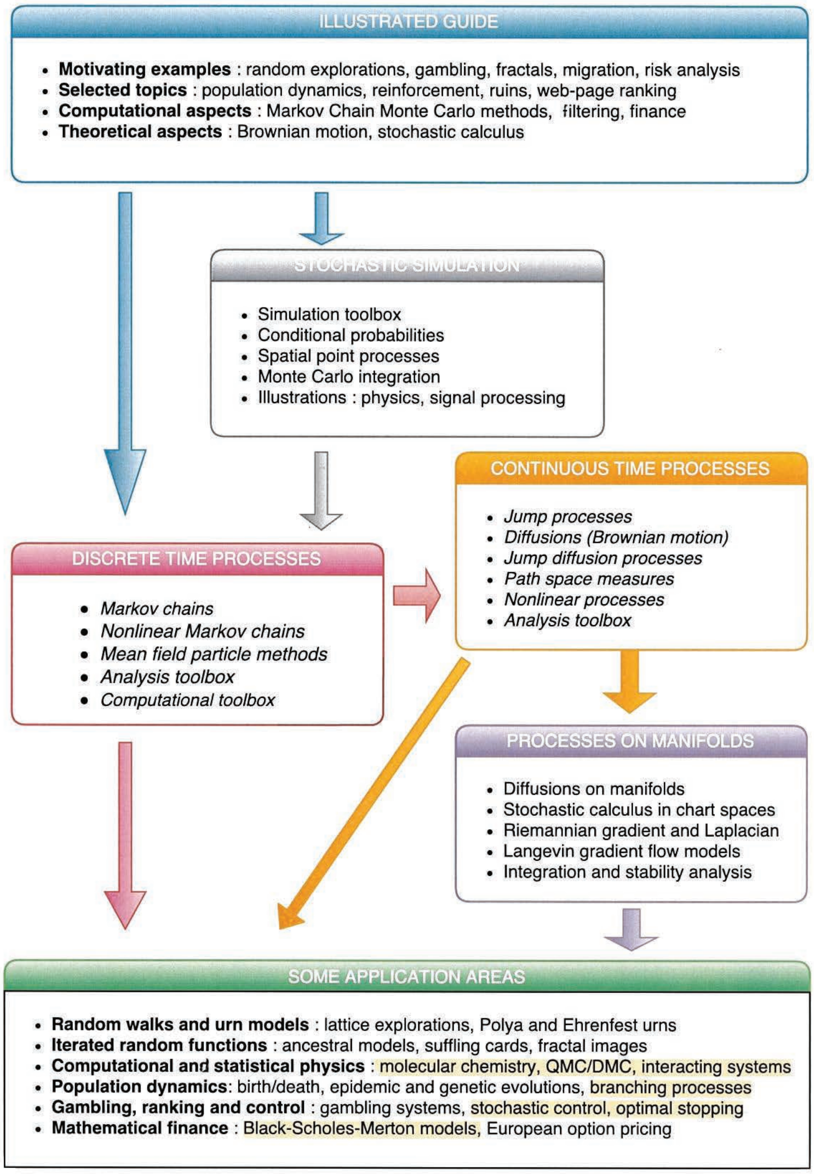
This book is almost self-contained. There are no strict prerequisites but it is envisaged that students would have taken a course in elementary probability and that they have some knowledge of elementary linear algebra, functional analysis and geometry. It is not possible to define rigorously stochastic processes without some basic knowledge on measure theory and differential calculus. Readers who lack such background should instead consult some introductory textbook on probability and integration, and elementary differential calculus.

The book also contains around 500 exercises with detailed solutions on a variety of topics, with many explicit computations. Each chapter ends with a section containing a series of exercises ranging from simple calculations to more advanced technical questions. This book can serve as a reference book, as well as a textbook. The next section provides a brief description of the organization of the book.

On page xxxii we propose a series of lectures and research projects which can be developed using the material presented in this book. The descriptions of these course projects also provide detailed discussions on the connections between the different topics treated in this book. Thus, they also provide a reading guide to enter into a given specialized topic.

Organization of the book

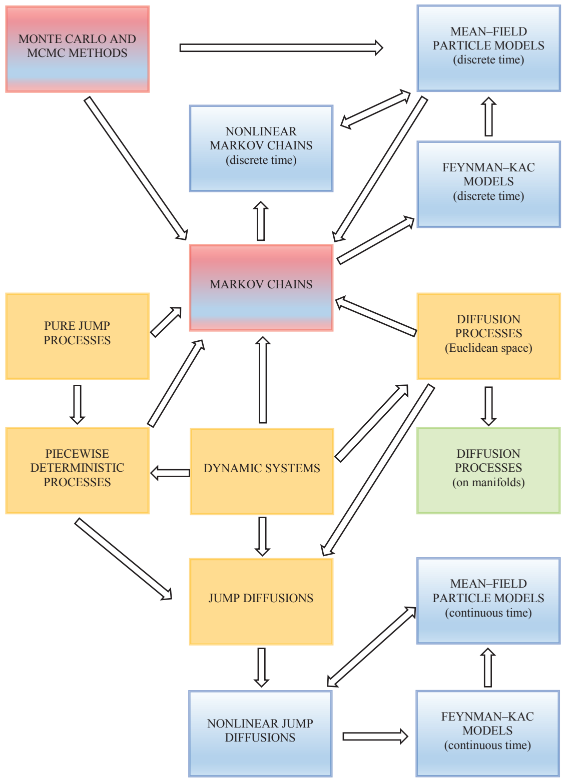
The book is organized in six parts. The synthetic diagram below provides the connections betweens the parts and indicates the different possible ways of reading the book.



Part I provides an illustrative guide with a series of motivating examples. Each of them is related to a deep mathematical result on stochastic processes. The examples include the recurrence and the transience properties of simple random walks, stochastic coupling techniques and mixing properties of Markov chains, random iterated functional models, Poisson processes, dynamic population models, Markov chain Monte Carlo methods, and the Doeblin-Itō differential calculus. In each situation, we provide a brief discussion on the mathematical analysis of these models. We also provide precise pointers to the chapters and sections where these results are discussed in more details.

Part II is concerned with the notion of randomness, and with some techniques to "simulate perfectly" some traditional random variables (*abbreviated r.v.*). For a more thorough discussion on this subject, we refer to the encyclopedic and seminal reference book of Luc Devroye [92], dedicated to the simulation of nonuniform random variables. Several concrete applications are provided to illustrate many of the key ideas, as well as the usefulness of stochastic simulation methods in some scientific disciplines.

The next three parts are dedicated to discrete and continuous time stochastic processes. The synthetic diagram below provides some inclusion type links between the different classes of stochastic processes discussed in this book.



Part III is concerned with discrete time stochastic processes, and more particularly with Markov chain models.

Firstly, we provide a detailed discussion on the different descriptions and interpretations of these stochastic models. We also discuss nonlinear Markov chain models, including self-interacting Markov chains and mean field type particle models.

Then we present a panorama of analytical tools to study the convergence of Markov chain models when the time parameter tends to ∞. The first class of mathematical techniques includes linear algebraic tools, spectral and functional analysis. We also present some more probabilistic-type tools such as coupling techniques, strong stationary times, and martingale limit theorems.

Finally, we review the traditional and the more recent computation techniques of Markov chain models. These techniques include Markov chain Monte Carlo methodologies, perfect sampling algorithms, and time inhomogeneous models. We also provide a discussion of the more recent Feynman-Kac particle methodologies, with a series of application domains.

Part IV of the book is concerned with continuous time stochastic processes and stochastic differential calculus, a pretty vast subject at the core of applied mathematics. Nevertheless, most of the literature on these processes in probability textbooks is dauntingly mathematical. Here we provide a pedagogical introduction, sacrificing from time to time some technical mathematical aspects.

Continuous time stochastic processes can always be thought as a limit of a discrete generation process defined on some time mesh, as the time step tends to 0. This viewpoint which is at the foundation of of stochastic calculus and integration is often hidden in purely theoretical textbooks on stochastic processes. In the reverse angle, more applied presentations found in textbooks in engineering, economy, statistics and physics are often too far from recent advances in applied probability so that students are not really prepared to enter into deeper analysis nor pursue any research level type projects.

One of the main new points we have adopted here is to make a bridge between applications and advanced probability analysis. Continuous time stochastic processes, including nonlinear jump-diffusion processes, are introduced as a limit of a discrete generation process which can be easily simulated on a computer. In the same vein, any sophisticated formula in stochastic analysis and differential calculus arises as the limit of a discrete time mathematical model. It is of course not within the scope of this book to quantify precisely and in a systematic way all of these approximations. Precise reference pointers to articles and books dealing with these questions are provided in the text.

All approximations treated in this book are developed through a single and systematic mathematical basis transforming the elementary Markov transitions of practical discrete generation models into infinitesimal generators of continuous time processes. The stochastic version of the Taylor expansion, also called the Doeblin-Itō formula or simply the Itō formula, is described in terms of the generator of a Markov process and its corresponding carré du champ operator.

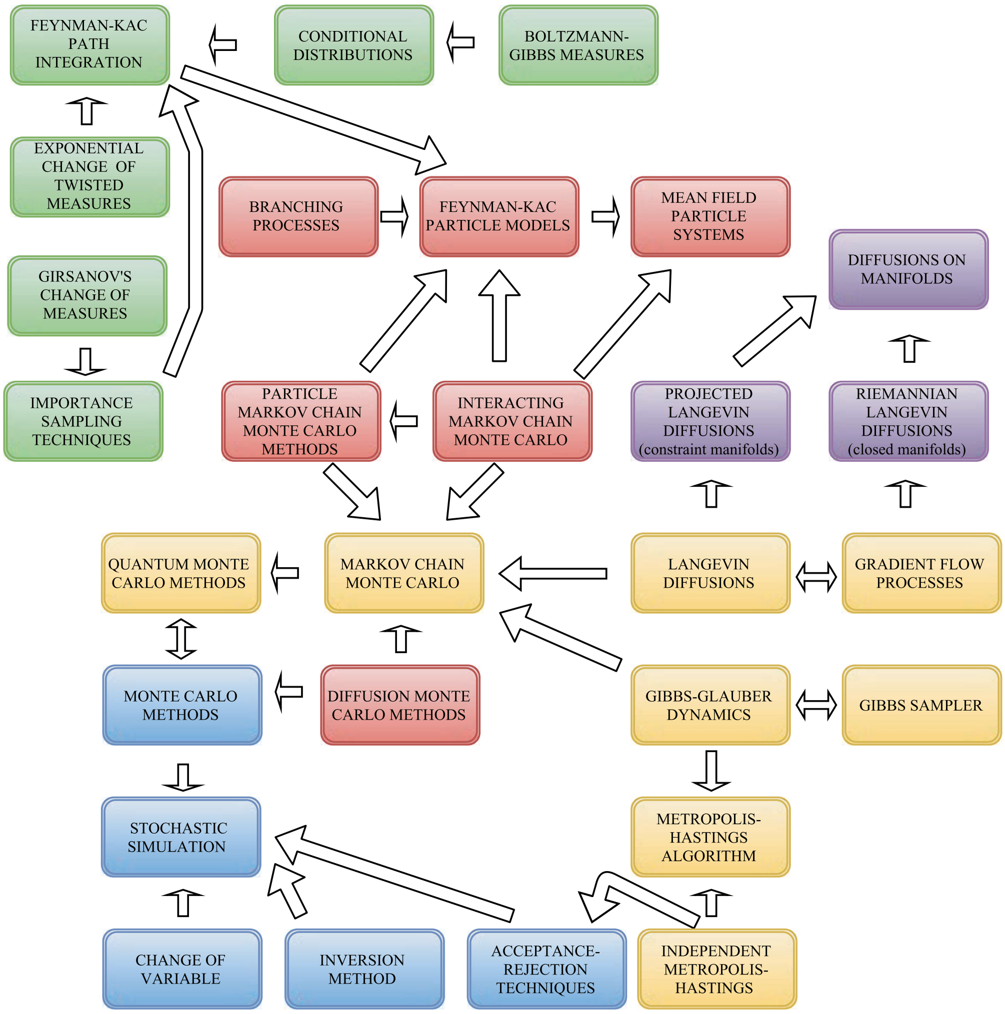
In this framework, the evolution of the law of the random states of a continuous time process resumes to a natural weak formulation of linear and nonlinear Fokker-Planck type equations (a.k.a. Kolmogorov equations). In this interpretation, the random paths of a given continuous time process provide a natural interpolation and coupling between the solutions of these integro-differential equations at different times. These path-space probability measures provide natural probabilistic interpretations of important Cauchy-Dirichlet-Poisson problems arising in analysis and integro-partial differential equation problems. The limiting behavior of these functional equations is also directly related to the long time behavior of a stochastic process.

All of these subjects are developed from chapter 10 to chapter 16. The presentation starts with elementary Poisson and jump type processes, including piecewise deterministic models, and develops forward to more advanced jump-diffusion processes and nonlinear Markov processes with their mean field particle interpretations. An abstract and universal class of continuous time stochastic process is discussed in section 15.5.

The last two chapters, chapter 17 and chapter 18 present respectively a panorama of analytical tools to study the long time behavior of continuous time stochastic processes, and path-space probability measures with some application domains.

Part V is dedicated to diffusion processes on manifolds. These stochastic processes arise when we need to explore state spaces associated with some constraints, such as the sphere, the torus, or the simplex. The literature on this subject is often geared towards purely mathematical aspects, with highly complex stochastic differential geometry methodologies. This book provides a self-contained and pedagogical treatment with a variety of examples and application domains. Special emphasis is given to the modelling, the stability analysis, and the numerical simulation of these processes on a computer.

The synthetic diagram below provides some inclusion type links between the different classes of stochastic computational techniques discussed in this book.



Part VI is dedicated to some application domains of stochastic processes, including random walks, iterated random functions, computational physics, dynamic population models, gambling and ranking, and mathematical finance.

Chapter 25 is mainly concerned with simple random walks. It starts with the recurrence or transience properties of these processes on integer lattices depending on their dimensions. We also discuss random walks on graphs, the exclusion process, as well as random walks on the circle and hypercubes and their spectral properties. The last part of the chapter is dedicated to the applications of Markov chains to analyze the behavior of urn type models such as the Polya and the Ehrenfest processes. We end this chapter with a series of exercises on these random walk models including diffusion approximation techniques.

Chapter 26 is dedicated to stochastic processes expressed in terms of iterated random functions. We examine three classes of models. The first one is related to branching processes and ancestral tree evolutions. These stochastic processes are expressed in terms of compositions of one to one mappings between finite sets of indices. The second one is concerned with shuffling cards. In this context, the processes are expressed in terms of random compositions of permutations. The last class of models are based on random compositions of linear transformations on the plane and their limiting fractal image compositions. We illustrate these models with the random construction of the Cantor's discontinuum, and provide some examples of fractal images such as a fractal leaf, a fractal tree, the Sierpinski carpet, and the Highways dragons. We analyze the stability and the limiting measures associated with each class of processes. The chapter ends with a series of exercises related to the stability properties of these models, including some techniques used to obtain quantitative estimates.

Chapter 27 explores some selected applications of stochastic processes in computational and statistical physics.

The first part provides a short introduction to molecular dynamics models and to their numerical approximations. We connect these deterministic models with Langevin diffusion processes and their reversible Boltzmann-Gibbs distributions.

The second part of the chapter is dedicated to the spectral analysis of the Schrödinger wave equation and its description in terms of Feynman-Kac semigroups. We provide an equivalent description of these models in terms of the Bra-kets and path integral formalisms commonly used in physics. We present a complete description of the spectral decomposition of the Schrödinger-Feynman-Kac generator, the classical harmonic oscillator in terms of the Hermite polynomials introduced in chapter 17. Last but not least we introduce the reader to the Monte Carlo approximation of these models based on the mean field particle algorithms presented in the third part of the book.

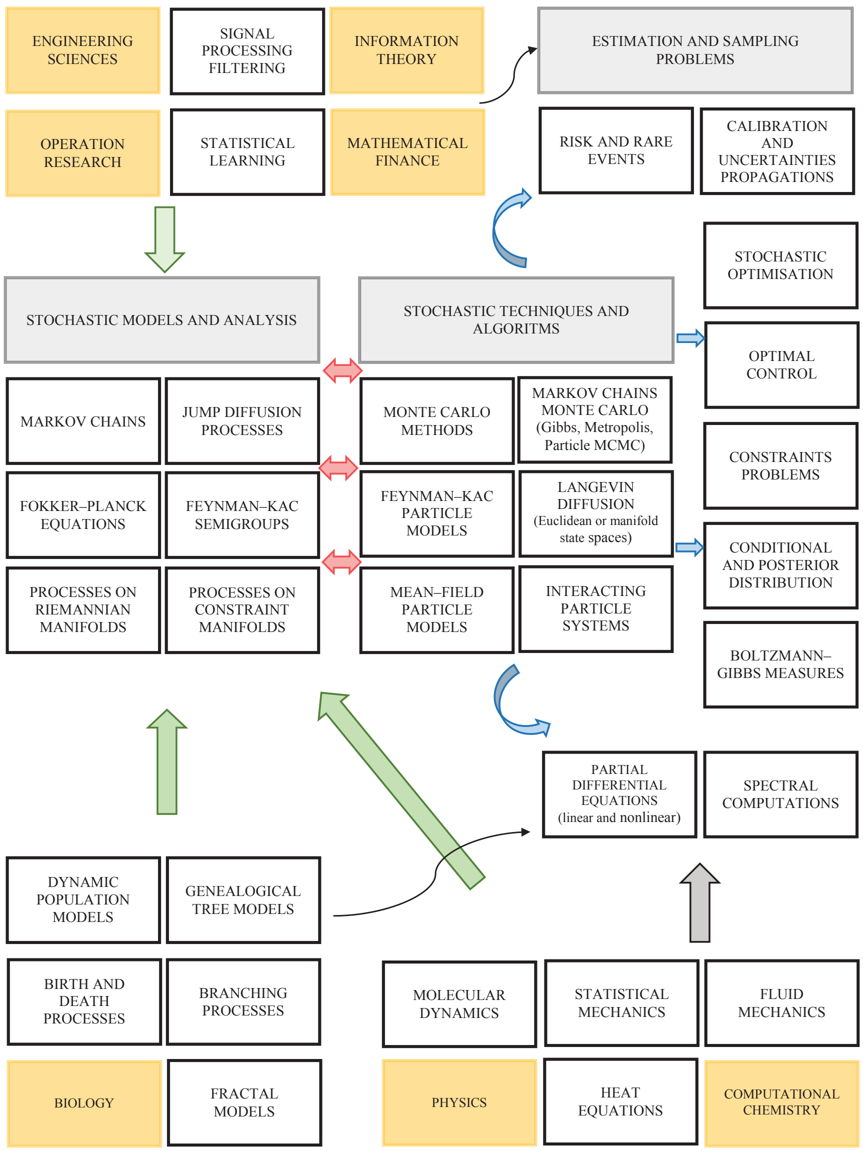
The third part of the chapter discusses different applications of interacting particle systems in physics, including the contact model, the voter and the exclusion process. The chapter ends with a series of exercises on the ground state of Schrödinger operators, quasi-invariant measures, twisted guiding waves, and variational principles.

Chapter 28 is dedicated to applications of stochastic processes in biology and epidemiology. We present several classes of deterministic and stochastic dynamic population models. We analyse logistic type and Lotka-Volterra processes, as well as branching and genetic processes. This chapter ends with a series of exercises related to logistic diffusions, bimodal growth models, facultative mutualism processes, infection and branching processes.

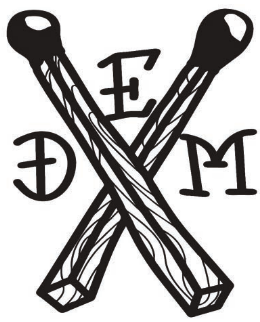
Chapter 29 is concerned with applications of stochastic processes in gambling and ranking. We start with the celebrated Google page rank algorithm. The second part of the chapter is dedicated to gambling betting systems. We review and analyze some famous martingales such as the St Petersburg martingale, the grand martingale, the D'Alembert and the Whittaker martingales. The third part of the chapter is mainly concerned with stochastic control and optimal stopping strategies. The chapter ends with a series of exercises on the Monty Hall game show, the Parrondo's game, the bold play strategy, the ballot and the secretary problems.

The last chapter, chapter 30, is dedicated to applications in mathematical finance. We discuss the discrete time Cox-Ross-Rubinstein models and the continuous Black-Scholes-Merton models. The second part of the chapter is concerned with European pricing options. We provide a discussion on the modelling of self-financing portfolios in terms of controlled martingales. We also describe a series of pricing and hedging techniques with some numerical illustrations. The chapter ends with a series of exercises related to neutralization techniques of financial markets, replicating portfolios, Wilkie inflation models, and life function martingales.

The synthetic diagram below provides some inclusion type links between the stochastic processes and the different application domains discussed in this book. The black arrows indicate the estimations and sampling problems arising in different application areas. The blue arrows indicate the stochastic tools that can be used to solve these problems. The red arrows emphasize the modelling, the design and the convergence analysis of stochastic methods. The green arrows emphasize the class of stochastic models arising in several scientific disciplines.



The book illustrations were hand-picked by the French artist Timothée Del Moral. He has been studying in the Fine Art School in Bordeaux and later joined several art studios in Toulouse, Nice and Bordeaux. Among his art works, he designed `the logo of the Biips INRIA software`, and posters for `the French Society of Mathematics and Industry`. He is now specialized in tattooing, developing `new styles and emulating old school-type illustrations and tattoos`.



Preview and objectives

This book provides an introduction to the probabilistic modelling of discrete and continuous time stochastic processes. The emphasis is put on the understanding of the main mathematical concepts and numerical tools across a range of illustrations presented through models with increasing complexity.

> *If you can't explain it simply, you don't understand it well enough.*
> Albert Einstein (1879-1955).

While exploring the analysis of stochastic processes, the reader will also encounter a variety of mathematical techniques related to matrix theory, spectral analysis, dynamical systems theory and differential geometry.

We discuss a large class of models ranging from finite space valued Markov chains to jump diffusion processes, nonlinear Markov processes, self-interacting processes, mean field particle models, branching and interacting particle systems, as well as diffusions on constraint and Riemannian manifolds.

Each of these stochastic processes is illustrated in a variety of applications. Of course, the detailed description of real-world models requires a deep understanding of the physical or the biological principles behind the problem at hand. These discussions are out of the scope of the book. We only discuss academic-type related situations, providing precise reference pointers for a more detailed discussion. The illustrations we have chosen are very often at the crossroads of several seemingly disconnected scientific disciplines, including biology, mathematical finance, gambling, physics, engineering sciences, operations research, as well as probability, and statistical inference.

The book also provides an introduction to stochastic analysis and stochastic differential calculus, including the analysis of probability measures on path spaces and the analysis of Feynman-Kac semigroups. We present a series of powerful tools to analyze the long time behavior of discrete and continuous time stochastic processes, including coupling, contraction inequalities, spectral decompositions, Lyapunov techniques, and martingale theory.

The book also provides a series of probabilistic interpretations of integro-partial differential equations, including stochastic partial differential equations (such as the nonlinear filtering equation), Fokker-Planck equations and Hamilton-Jacobi-Bellman equations, as well as Cauchy-Dirichlet-Poisson equations with boundary conditions.

Last but not least, the book also provides a rather detailed description of some traditional and more advanced computational techniques based on stochastic processes. These techniques include Markov chain Monte Carlo methods, coupling from the past techniques, as well as more advanced particle methods. Here again, each of these numerical techniques is illustrated in a variety of applications.

Having said that, let us explain some other important subjects which are not treated in this book. We provide some precise reference pointers to supplement this book on some of these subjects.

As any lecture on deterministic dynamical systems does not really need to start with a fully developed Lebesgue, or Riemann-Stieltjes integration theory, nor with a specialized course on differential calculus, we believe that a full and detailed description of stochastic integrals is not really needed to start a lecture on stochastic processes.

For instance, Markov chains are simply defined as a sequence of random variables indexed by integers. Thus, the mathematical foundations of Markov chain theory only rely on rather elementary algebra or Lebesgue integration theory.

It is always assumed implicitly that these (infinite) sequences of random variables are defined on a common probability space; otherwise it would be impossible to quantify, nor to define properly limiting events and objects. From a pure mathematical point of view, this innocent technical question is not so obvious even for sequences of independent random variables. The construction of these probability spaces relies on abstract projective limits, the well known `Carathéodory extension theorem` and the `Kolmogorov extension theorem`. The Ionescu-Tulcea theorem also provides a natural probabilistic solution on general measurable spaces. It is clearly not within the scope of this book to enter into these rather well known and sophisticated constructions.

As dynamical systems, continuous time stochastic processes can be defined as the limit of a discrete generation Markov process on some time mesh when the time step tends to 0. For the same reasons as above, these limiting objects are defined on a probability space constructed using projective limit techniques and extension type theorems to define in a unique way infinite dimensional distributions from finite dimensional ones defined on cylindrical type sets.

In addition, the weak convergence of any discrete time approximation process to the desired continuous time limiting process requires us to analyze the convergence of the finite dimensional distributions on any finite sequence of times. We will discuss these finite dimensional approximations in some details for jump as well as for diffusion processes.

Nevertheless, the weak convergence at the level of the random trajectories requires us to ensure that the laws of the random paths of the approximating processes are `relatively compact` in the sense that any sequence admits a convergent subsequence. For complete separable metric spaces, this condition is equivalent to a `tightness condition` that ensures that these sequences of probability measures are almost concentrated on a compact subset in the whole set of trajectories.

Therefore to check this compactness condition we first need to characterize the compact subsets of the set of continuous trajectories, or of the set of right continuous trajectories with left hand limits. By `the Arzelà-Ascoli theorem`, these compact sets are described in terms of equicontinuous trajectories. Besides the fact that the most of stochastic processes encountered in practice satisfy this tightness property, the proof of this condition relies on sophisticated probabilistic tools. A very useful and commonly used to check this tightness property is `the Aldous criterion` introduced by D. Aldous in his PhD dissertation and published in the *Annals of Probability* in 1978.

Here again, entering in some details into these rather well known and well developed subjects would cause digression. More details on these subjects can be found in the books

- R. Bass. *Stochastic Processes*. CUP (2011).
- S. Ethier,T. Kurtz.*Markov Processes: Characterization and Convergence.* Wiley (1986).
- J. Jacod, A. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer (1987).

The seminal book by J. Jacod and A. N. Shiryaev is highly technical and presents difficult material but it contains almost all the limiting convergence theorems encountered in the theory of stochastic processes. This reference is dedicated to researchers in pure and applied probability. The book by S.N. Ethier and T. G. Kurtz is in spirit closer to our approach based on the description of general stochastic processes in terms of their generators. This book is recommended to anyone interested into the precise mathematical definition of generators and their regularity properties. The book of R. Bass is more accessible and presents a detailed mathematical construction of stochastic integrals and stochastic calculus.

The integral representations of these limiting stochastic differential equations (abbreviated SDE) are expressed in terms of `stochastic integrals`. These limiting mathematical models coincide with the definition of the stochastic integral. The convergence can be made precise in a variety of well known senses, such as in probability, almost sure, or w.r.t. some $\mathbb{L}_p$-norms.

Recall that integral description of any deterministic dynamical system is based on the so-called `fundamental theorem of calculus`. In the same vein, the integral description of stochastic processes is now based on `the fundamental Doeblin-Itō differential calculus`. This stochastic differential calculus can be interpreted as the natural extension of `Taylor differential calculus` to random dynamical systems, such as diffusions and random jump type processes. Here again `the Doeblin-Itō differential calculus` can also be expressed in terms of natural Taylor type expansions associated with a discrete time approximation of a given continuous type stochastic process.

In the further development of this book, we have chosen to describe in full details these stochastic expansions for the most fundamental classes of stochastic processes including jump diffusions and more abstract stochastic models in general state spaces, as well as stochastic processes on manifolds.

Precise reference pointers to textbooks dedicated to deeper mathematical foundations of `stochastic integrals` and `the Doeblin-Itō differential calculus` are given in the text.

This choice of presentation has many advantages.

Firstly, the time discretization of continuous time processes provides a natural way to simulate these processes on a computer.

In addition, any of these simulation techniques also improves the physical and the probabilistic understanding of the nature and the long time behavior of the limiting stochastic process.

Furthermore, this pedagogical presentation does not really require a strong background in sophisticated stochastic analysis techniques, as it is based only on discrete time approximation schemes. As `stochastic integrals` and `the Doeblin-Itō differential calculus` are themselves defined in terms of limiting formulae associated with a discrete generation process, our approach provides a way to introduce these somehow sophisticated mathematical objects in a simple form.

In this connection, most of the limiting objects as well as quantitative estimates are generally obtained by using standard technical tools such as `the monotone` or the `dominated convergence` theorems, `Cauchy-Schwartz`, `Hölder` or `Minskowski` inequalities. Some of these results are provided in the text and in the exercises. Further mathematical details can also be found in the textbooks cited in the text.

Last, but not least, this choice of presentation also lays a great foundation for research, as most of the research questions in probability are asymptotic in nature. In order to understand these questions, the basic non-asymptotic manipulations need to be clearly understood.

Rather than pepper the text with repeated citations, we mention here (in alphabetical order) some other classical and supplemental texts on the full mathematical construction of stochastic integrals, the characterization and the convergence of continuous time stochastic processes:

- R. Durrett, *Stochastic Calculus*. CRC Press (1996).
- I. Gikhman, A. Skorokhod. *Stochastic Differential Equations.* Springer (1972).
- F.C Klebaner. *Introduction to Stochastic Calculus with Applications.* World Scientific Publishing (2012).
- P. Protter. *Stochastic Integration and Differential Equations.* Springer (2004).

More specialized books on Brownian motion, martingales with continuous paths, diffusions processes and their discrete time approximations are:

- N. Ikeda, S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland Publishing Co. (1989).
- I. Karatzas, S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer (2004).
- P. E. Kloeden, E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer (2011).
- B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer (2005).
- L. C. G. Rogers, D. Williams. *Diffusions, Markov Processes, and Martingales.* CUP (2000).

The seminal book by P. E. Kloeden and E. Platen provides an extensive discussion on the convergence properties of discrete time approximations of stochastic differential equations.

Other specialized mathematical textbooks and lecture notes which can be useful for supplemental reading on stochastic processes and analysis on manifolds are:

- K.D. Elworthy. *Stochastic Differential Equations on Manifolds*. CUP(1982).
- M. Emery. *Stochastic Calculus on Manifolds*. Springer (1989).
- E.P. Hsu. *Stochastic analysis on manifolds.* Providence AMS (2002).

Some lecture projects

The material in this book can serve as a basis for different types of advanced undergraduate and graduate level courses as well as master-level and post-doctoral level courses on stochastic processes and their application domains. As we mentioned above, the book also contains around 500 exercises with detailed solutions. Each chapter ends with a section containing a series of exercises ranging from simple calculations to more advanced technical questions. These exercises can serve for training the students through tutorials or homework.

> *Running overtime is the one unforgivable error a lecturer can make.*
> *After fifty minutes (one microcentury as von Neumann used to say)*
> *everybody's attention will turn elsewhere.* Gian Carlo Rota (1932-1999).

Introduction to stochastic processes (with applications)

This first type of a lecture course would be a pedagogical introduction to elementary stochastic processes based on the detailed illustrations provided in the first three chapters, chapter 1, chapter 2, and chapter 3. The lectures can be completed with the simulation techniques presented in chapter 4, the Monte Carlo methods discussed in chapter 5, and the illustrations provided in chapter 6.

The lecture presents discrete as well as continuous time stochastic processes through a variety of motivating illustrations including dynamic population models, fractals, card shuffling, signal processing, Bayesian statistical inference, finance, gambling, ranking, and stochastic search, reinforcement learning and sampling techniques.

Through a variety of applications, the lecture provides a brief introduction to Markov chains, including the recurrence properties of simple random walks, Markov chains on permutation groups, and iterated random functions. This course also covers Poisson processes, and provides an introduction to Brownian motion, as well as piecewise deterministic Markov processes, diffusions and jump processes.

This lecture course also covers computational aspects with an introduction of Markov chain Monte Carlo methodologies and the celebrated Metropolis-Hastings algorithm. The text also provides a brief introduction to more advanced and sophisticated nonlinear stochastic techniques such as self-interacting processes particle filters, and mean field particle sampling methodologies.

The theoretical topics also include an accessible introduction to stochastic differential calculus. The lecture provides a natural and unified treatment of stochastic processes based on infinitesimal generators and their carré du champ operators.

This book is also designed to stimulate the development of research projects that offer the opportunity to apply the theory of stochastic processes discussed during the lectures to an application domain selected among the ones discussed in this introductory course: *card shuffling, fractal images, ancestral evolutions, molecular dynamics, branching and interacting processes, genetic models, gambling betting systems, financial option pricing, advanced signal processing, stochastic optimization, Bayesian statistical inference, and many others.*

These research projects provide opportunity to the lecturer to immerse the student in a favorite application area. They also require the student to do a background study and perform personal research by exploring one of the chapters in the last part of the book dedicated to application domains. The diagrams provided on page xxiii and on page xxviii guide the students on the reading order to enter into a specific application domain and explore the different links and inclusions between the stochastic models discussed during the lectures.

More theoretical research projects can also be covered. For instance, the recurrence questions of random walks discussed in the first section of chapter 1 can be further developed by using the topological aspects of Markov chains presented in section 8.5. Another research project based on the martingale theory developed in section 8.4 will complement the discussion on gambling and ruin processes discussed in chapter 2. The discussion on signal processing and particle filters presented in chapter 3 can be complemented by research projects on Kalman filters (section 9.9.6), mean field particle methodologies (section 7.10.2), and Feynman-Kac particle models (section 9.5 and section 9.6).

Other textbooks and lecture notes which can be useful for supplemental reading in probability and applications during this lecture are:

- P. Billingsley. *Convergence of Probability Measures*, Wiley (1999).
- L. Devroye. *Non-Uniform Random Variate Generation*, Springer (1986).
- W. Feller. *An Introduction to Probability Theory*, Wiley (1971).
- G.R. Grimmett, D.R. Stirzaker, *One Thousand Exercises in Probability*, OUP (2001).
- A. Shiryaev. *Probability*, Graduate Texts in Mathematics, Springer (2013).
- P. Del Moral, B. Remillard, S. Rubenthaler. *Introduction aux Probabilités*. Ellipses Edition [in French] (2006).

We also refer to some online resources such as the review of probability theory by T. Tao, the Wikipedia article on Brownian Motion, and a Java applet simulating Brownian motion.

Of course it normally takes more than an year to cover the full scope of this course with fully rigorous mathematical details. This course is designed for advanced undergraduate to master-level audiences with different educational backgrounds ranging from engineering sciences, economics, business, mathematical finance, physics, statistics to applied mathematics.

Giving a course for such a diverse audience with different backgrounds poses a significant challenge to the lecturer. To make the material accessible to all groups, the course begins with the series of illustrations in the first three chapters. This fairly gentle introduction is designed to cover a variety of subjects in a single semester course of around 40 hours, at a somehow more superficial mathematical level from the pure and applied mathematician perspective. These introductory chapters contain precise pointers to specialized chapters with more advanced mathematical material. Students with deep backgrounds in applied mathematics are invited to explore deeper into the book, while students with limited mathematical backgrounds will concentrate on the more elementary material discussed in chapter 1, chapter 2, and chapter 3.

Discrete time stochastic processes

The second type of lecture course is a pedagogical introduction to discrete time processes and Markov chain theory, a rather vast subject at the core of applied probability and many other scientific disciplines that use stochastic models. This more theoretical type course is geared towards *discrete time stochastic processes and their stability analysis.* It would essentially cover chapter 7 and chapter 8.

As in the previous lecture course, this would start with some selected motivating illustrations provided in chapter 1 and chapter 2, as well as from the material provided in section 3.1, chapter 5 and chapter 6 dedicated to Markov chain Monte Carlo methodologies.

The second part of this course will center around some selected material presented in the third part of the book dedicated to discrete time stochastic processes. Chapter 7 offers a description of the main classes of Markov chain models including nonlinear Markov chain models and their mean field particle interpretations.

After this exploration of discrete generation random processes, at least two possible different directions can be taken.

The first one is based on chapter 8 with an emphasis on the topological aspects and the stability properties of Markov chains. This type of course presents a panorama of analytical tools to study the convergence of Markov chain models when the time parameter tends to ∞. These mathematical techniques include *matrix theory, spectral and functional analysis, as well as contraction inequalities, geometric drift conditions, coupling techniques and martingale limit theorems.*

The second one, based on chapter 9, will emphasize computational aspects ranging from traditional Markov chain Monte Carlo methods and perfect sampling techniques to more advanced Feynman-Kac particle methodologies. This part or the course will cover *Metropolis-Hastings algorithms, Gibbs-Glauber dynamics, as well as the Propp and Wilson coupling from the past sampler.* The lecture would also offer *a brief introduction to more advanced particle methods and related particle Markov chain Monte Carlo methodologies.*

The course could be complemented with some selected application domains described in chapter 25, such as random walks on lattices and graphs, or urn type random walks. Another strategy would be to illustrate the theory of Markov chains with chapter 26 dedicated to iterated random functions and their applications in biology, card shuffling or fractal imaging, or with the discrete time birth and death branching processes presented in chapter 28.

Several research projects could be based on previous topics as well as on the gambling betting systems and discrete time stochastic control problems discussed in chapter 29, or on the applications in mathematical finance discussed in chapter 30.

This course is designed for master-level audiences in engineering, physics, statistics and applied probability.

Other textbooks and lecture notes which can be useful for supplemental reading in Markov chains and their applications during this lecture are:

- D. Aldous, J. Fill. *Reversible Markov Chains and Random Walks on Graphs* (1999).
- S. Meyn, R. Tweedie. *Markov Chains and Stochastic Stability*, Springer (1993).
- J. Norris. *Markov Chains*. CUP (1998).
- S.I. Resnick. *Adventures in Stochastic Processes*. Springer Birkhauser (1992).

We also refer to some online resources such as *the Wikipedia article on Markov chains*, and to the *Markov Chains chapter* in American Mathematical Society's introductory probability book, the article by D. Saupe *Algorithms for random fractals*, and the seminal and very clear article by P. Diaconis titled *The Markov Chain Monte Carlo Revolution.*

Continuous time stochastic processes

This semester-long course is dedicated to continuous time stochastic processes and stochastic differential calculus. The lecture notes would essentially cover the series of chapters 10 to 18. The detailed description of this course follows essentially the presentation of the fourth part of the book provided on page xxiv. *This course is designed for master-level audiences in engineering, mathematics and physics with some background in probability theory and elementary Markov chain theory.*

To motivate the lectures, this course would start with some selected topics presented in the first three chapters, such as the Poisson's typos discussed in section 1.7 (Poisson processes), the pinging hacker story presented in section 2.5 (piecewise deterministic stochastic processes), the lost equation discussed in section 3.3 (Brownian motion), the formal stochastic calculus derivations discussed in section 3.4 (Doeblin-Itō differential calculus), and the theory of speculation presented in section 3.5 (backward equations and financial mathematics).

The second part of the lecture would cover elementary Poisson processes, Markov chain continuous time embedding techniques, pure jump processes and piecewise deterministic models. It would also cover diffusion processes and more general jump-diffusion processes, with detailed and systematic descriptions of the infinitesimal generators, the Doeblin-Itō differential formula and the natural derivation of the corresponding Fokker-Planck equations.

The third part could be dedicated to jump diffusion processes with killing and their use to solve Cauchy problems with terminal conditions, as well as Dirichlet-Poisson problems. The course can also provide an introduction to nonlinear stochastic processes and their mean field particle interpretations, with illustrations in the context of risk analysis, fluid mechanics and particle physics.

The last part of the course would cover path-space probability measures including Girsanov's type change of probability measures and exponential type martingales, with illustrations in the analysis of nonlinear filtering processes, in financial mathematics and in rare event analysis, including in quantum Monte Carlo methodologies.

Three different applications could be discussed, depending on the background and the scientific interests of the audience (or of the lecturer):

- *Computational and statistical physics (chapter 27):*

  The material on interacting particle systems presented in section 27.3 only requires some basic knowledge on generators of pure jump processes. This topic can also be used to illustrate pure jump processes before entering into the descriptions of diffusion type processes.

  The material on molecular dynamics simulation discussed in section 27.1 only requires some knowledge on pure diffusion processes and their invariant measures. We recommend studying chapter 17 before entering into these topics.

  Section 27.2 dedicated to the Schrödinger equation is based on more advanced stochastic models such as the Feynman-Kac formulae, jump diffusion processes and exponential changes of probability measures. We recommend studying chapter 15 and chapter 18 (and more particularly section 18.1.3) before entering into this more advanced application area. Section 28.4.3 also provides a detailed discussion on branching particle interpretations of Feynman-Kac formulae.

- *Dynamic population models (chapter 28):*

  All the material discussed in this chapter can be used to illustrate jump and diffusion stochastic processes. Section 28.4.3 requires some basic knowledge on Feynman-Kac semigroups. It is recommended to study chapter 15 before entering into this topic.

- *Stochastic optimal control (section 29.3.3 and section 29.4.3)*:

  The material discussed in this chapter can also be used to illustrate jump and diffusion stochastic processes. It is recommended to start with discrete time stochastic control theory (section 29.3.1, section 29.3.2 and section 29.4.2) before entering into more sophisticated continuous time problems.

- *Mathematical finance (chapter 30):*

  The sections 30.2.4 to 30.2.7 are essentially dedicated to the Black-Scholes stochastic model. This topic can be covered with only some basic knowledge on diffusion processes and more particularly on geometric Brownian motion.

Depending on the selected application domains to illustrate the course, we also recommend for supplemental reading the following references:

- K.B. Athreya, P.E. Ney. *Branching Processes*. Springer (1972).
- D. P. Bertsekas. *Dynamic Programming and Optimal Control*, Athena Scientific (2012).
- M. Caffarel, R. Assaraf. A pedagogical introduction to quantum Monte Carlo. In *Mathematical Models and Methods for Ab Initio Quantum Chemistry.* Lecture Notes in Chemistry, eds. M. Defranceschi and C. Le Bris, Springer (2000).
- P. Del Moral. *Mean Field Simulation for Monte Carlo Integration.* Chapman & Hall/CRC Press (2013).
- P. Glasserman. *Monte Carlo methods in financial engineering*. Springer (2004).
- T. Harris. *The Theory of Branching Processes*, R-381, Report US Air Force (1964) and Dover Publication (2002).

- H. J. Kushner. *Introduction to Stochastic Control*, New York: Holt, Reinhart, and Winston (1971).
- J.M. Steele. *Stochastic Calculus and Financial Applications.* Springer (2001).
- A. Lasota, M.C. Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics.* Springer (1994).

Stochastic processes on manifolds

This advanced theoretical style course covers the fifth part of the book on stochastic diffusions on `(differentiable) manifolds`. It essentially covers the chapters 19 to 23. *This course is intended for master-level students as well as post-doctoral audiences with some background in differential calculus and stochastic processes.* The purpose of these series of lectures is to introduce the students to the basic concepts of stochastic processes on manifolds. The course is illustrated with a series of concrete applications.

In the first part of the course we review some basics tools of `differential geometry`, including `orthogonal projection` techniques, and the related first and second order `covariant derivatives` (19). The lectures offer a pedagogical introduction to the main mathematical models used in `differential geometry` such as the notions of `divergence`, `mean curvature vector`, `Lie brackets between vector fields`, `Laplacian operators`, and the `Ricci curvature` on a manifold. One of the main objectives is to help the students understand the basic concepts of deterministic and stochastic differential calculus that are relevant to the theory of stochastic processes on manifolds. The other objective is to show how these mathematical techniques apply to designing and to analyzing random explorations on constraint manifolds.

One could end this part of the course with the detailed proof of the `Bochner-Lichnerowicz formula`. This formula is pivotal in the stability analysis of diffusion processes on manifolds. It connects the second order properties of the generator of these processes in terms of the Hessian operator and the Ricci curvature of the state space.

The second part of the lecture course could be dedicated to stochastic calculus on embedded manifolds (chapter 20) and the notion of charts (a.k.a. atlases) and paramtrization spaces (chapter 21):

In a first series of lectures we introduce the Brownian motion on a manifold defined in terms of the level sets of a smooth and regular constraint function. Then, we analyze the infinitesimal generator of these models and we present the corresponding Doeblin-Itō differential calculus. We illustrate these models with the detailed description of the Brownian motion on the two-dimensional sphere and on the cylinder. These stochastic processes do not really differ from the ones discussed in chapter 14. The only difference is that they evolve in constraint manifolds embedded in the ambient space.

In another series of lectures we describe how the geometry of the constraint manifold is lifted to the parameter space (a.k.a. chart space). The corresponding geometry on the parameter space is equipped with a Riemannian scalar product. This main objective is to explore the ways to link the differential operators on the manifold in the ambient space to the Riemannian parameter space manifold. We provide some tools to compute the Ricci curvature in local coordinates and we describe the expression of Bochner-Lichnerowicz formula in Riemannian manifolds.

Several detailed examples are provided in chapter 22 which is dedicated to stochastic calculus in chart spaces. We define Brownian motion and general diffusions on Riemannian manifolds. We illustrate these stochastic models with a detailed description of the Brownian motion on the circle, the two-dimensional sphere and on the torus. Chapter 24 also offers a series of illustrations, starting with a detailed discussion of prototype manifolds such as the circle, the sphere and the torus. In each situation, we describe the Riemannian metric, the corresponding Laplacian, as well as the `geodesics`, the `Christoffel symbols` and the Ricci curvature. In the second part of the chapter we present applications in information theory. We start with a discussion on distribution and Bayesian statistical manifolds and we analyze the corresponding Fisher information metric and the Cramer-Rao lower bounds. We also discuss the Riemannian metric associated with Bolzmann-Gibbs models and multivariate normal distributions.

The course could end with the presentation of material in chapter 23 dedicated to some important analytical aspects, including the construction of geodesics and the integration on a manifold. We illustrate the impact of these mathematical objects with the design of gradient type diffusions with a Boltzmann-Gibbs invariant measure on a constraint manifold. The time discretization of these models is also presented using Metropolis-adjusted Langevin algorithms (a.k.a. MALA in engineering sciences and information theory).

This chapter also presents some analytical tools for analyzing the stability of stochastic processes on a manifold in terms of the second order properties of their generators. The gradient estimates of the Markov semigroups are mainly based on the Bochner-Lichnerowicz formula presented in chapter 19 and chapter 21. The series of exercises at the end of this chapter also provide illustrations of the stability analysis of Langevin models including projected and Riemannian Langevin processes.

Other specialized mathematical textbooks and lecture notes which can be useful for supplemental reading during this course in stochastic processes and analysis on manifolds are presented on page xxxii.

We also refer to some online resources in differential geometry such as `the Wikipedia article on Riemannian manifolds`, the very useful `list of formulae in Riemannian geometry`, and the book by M. Spivak titled *A Comprehensive Introduction to Differential Geometry* (1999). Other online resources in stochastic geometry are:

The pioneering articles by K. Itō Stochastic Differential Equations in a Differentiable Manifold. [*Nagoya Math. J.* Volume 1, pp.35–47 (1950)], and The Brownian Motion and Tensor Fields on Riemannian Manifolds [*Proc. Int. Congr. Math.*, Stockholm (1962)], and the very nice article by T. Lelièvre, G. Ciccotti, E. Vanden-Eijnden. . Projection of diffusions on submanifolds: Application to mean force computation. [*Communi. Pure Appl. Math.*, 61(3), 371-408, (2008)].

Stochastic analysis and computational techniques

There is enough material in the book to support more specialized courses on stochastic analysis and computational methodologies. A synthetic description of some of these course projects is provided below.

Stability of stochastic processes

As its name indicates, this course is concerned with the long time behavior of stochastic processes. *This course is designed for master-level audiences in mathematics and physics with some background in Markov chains theory and/or in continuous time jump diffusion processes, depending on whether the lectures are dedicated to discrete generation and/or continuous time models.*

Here again the lectures could start with some selected topics presented in the first three chapters, such as the stabilization of population subject discussed in section 2.1 (coupling techniques), the pinging hackers subject presented in section 2.5 (long time behavior of piecewise deterministic processes), and the discussion on Markov chain Monte Carlo methods (a.k.a. MCMC algorithms) provided in section 3.1.

To illustrate the lectures on discrete generation stochastic processes, the course would also benefit from the material provided in chapter 5 and chapter 6 dedicated to Markov chain Monte Carlo methodologies. Another series of illustrations based on shuffling cards problems and fractals can be found in chapter 26.

We also recommend presenting the formal stochastic calculus derivations discussed in section 3.4 (Doeblin-Itō differential calculus) if the course is dedicated to continuous time processes.

The second part of the course could be centered around the analysis toolboxes presented in chapter 8 (discrete time models) and chapter 17 (continuous time models). The objective is to introduce the students to a class of traditional approaches to analyze the long time behavior of Markov chains, including *coupling, spectral analysis, Lyapunov techniques, and contraction inequalities.* The coupling techniques can be illustrated using the discrete generation birth and death process discussed in section 28.1. The stability of continuous time birth and death models (with linear rates) can also be discussed using the explicit calculations developed in section 28.4.1.

The lectures could be illustrated by the Hamiltonian and Lagrangian molecular dynamics simulation techniques discussed in section 27.1.

The last part of the course could focus on the stability of stochastic processes on manifolds if the audience has some background on manifold-valued stochastic processes. It would essentially cover the sections 23.4 to 23.7. Otherwise, this final part of the course offers a pedagogical introduction on these models based on the material provided in chapter 19, including the detailed proof of the `Bochner-Lichnerowicz formula` and followed by gradient estimates discussed in section 23.7.

Other lecture notes and articles which can be useful for supplemental reading on Markov chains and their applications during this course are:

- S. Meyn, R. Tweedie. *Markov Chains and Stochastic Stability*, Springer (1993).
- P. Diaconis. *Something we've learned (about the Metropolis-Hasting algorithm)*, Bernoulli (2013).
- R. Douc, E. Moulines, D. Stoffer. *Nonlinear Time Series Analysis: Theory, Methods and Applications with R Examples*, Chapman & Hall/CRC Press (2014).
- M. Hairer. *Convergence of Markov Processes*. Lecture notes, Warwick University (2010).
- A. Katok, B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. CUP (1997).
- L. Saloff-Costes. *Lectures on Finite Markov Chains.* Springer (1997).
- P. Del Moral, M. Ledoux, L. Miclo. Contraction properties of Markov kernels. *Probability Theory and Related Fields,* vol. 126, pp. 395–420 (2003).
- F. Wang. *Functional Inequalities, Markov Semigroups and Spectral Theory.* Elsevier (2006).

Stochastic processes and partial differential equations

A one-quarter or one-semester course could be dedicated to probabilistic interpretations of integro-differential equations. *This course is designed for master-level audiences in applied* *probability and mathematical physics, with some background in continuous time stochastic processes and elementary integro-partial differential equation theory.*

The course could start with a review of continuous time processes, the description of the notion of generators, the Doeblin-Itō differential formula and the natural derivation of the corresponding Fokker-Planck equations. This part of the lecture could follow the formal stochastic calculus derivations discussed in section 3.4 (Doeblin-Itō differential calculus) and and the material presented in chapter 15 with an emphasis on the probabilistic interpretations of Cauchy problems with terminal conditions, as well as Dirichlet-Poisson problems (section 15.6).

The second part of the course could cover chapter 16 dedicated to nonlinear jump diffusions and the mean field particle interpretation of a class of nonlinear integro-partial differential equations (section 16.2), including Burger's equation, nonlinear Langevin diffusions, McKean-Vlasov models and Feynman-Kac semigroups.

The third part of the course could be dedicated to some selected application domains. The lectures could cover the Duncan-Zakai and the Kushner-Stratonovitch stochastic partial differential equations arising in nonlinear filtering theory (section 18.5), the Feynman-Kac description of the Schrödinger equation discussed in section 27.2, and the branching particle interpretations of the Kolmogorov-Petrovskii-Piskunov equations presented in section 28.4.3.4.

The last part of the lecture could cover the Hamilton-Jacobi-Bellman equations arising in stochastic control theory and presented in section 29.3.3.

Other textbooks and lecture notes (in alphabetical order) which can be useful for supplemental reading during this course are:

- M. Bossy, N. Champagnat. *Markov processes and parabolic partial differential equations.* Book section in R. Cont. *Encyclopedia of Quantitative Finance*, pp.1142-1159, John Wiley & Sons (2010).

- P. Del Moral. *Mean field simulation for Monte Carlo integration.* Chapman and Hall/CRC Press (2013).

- E.B. Dynkin. *Diffusions, Superdiffusions and PDEs.* AMS (2002).

- M.I. Freidlin. *Markov Processes and Differential Equations.* Springer (1996).

- H.M. Soner. Stochastic representations for nonlinear parabolic PDEs. In *Handbook of differential equations. Evolutionary Equations, volume 3.* Edited by C.M. Dafermos and E. Feireisl. Elsevier (2007).

- N. Touzi. *Optimal Stochastic Control, Stochastic Target Problems and Backward Stochastic Differential Equations.* Springer (2010).

Advanced Monte Carlo methodologies

A more applied course geared toward numerical probability and computational aspects would cover chapter 8 and the Monte Carlo techniques and the more advanced particle methodologies developed in chapter 9. Depending on the mathematical background of the audience, the course could also offer a review on the stability of Markov processes, following the description provided on page xxxviii.

Some application-oriented courses and research projects

There is also enough material in the book to support more applied courses on one or two selected application domains of stochastic processes. These lectures could cover random walk type models and urn processes (chapter 25), iterated random functions including shuffling cards, fractal models, and ancestral processes (chapter 26), computational physics and interacting particle systems (chapter 27), dynamic population models and branching processes (chapter 28), ranking and gambling betting martingale systems (chapter 29), and mathematical finance (chapter 30).

The detailed description of these course projects follows essentially the presentation of the sixth part of the book provided on page xxvi. The application domains discussed above can also be used to stimulate the development of research projects. The background requirements to enter into these topics are also discussed on page xxxvi.

We would like to thank John Kimmel for his editorial assistance, as well as for his immense support and encouragement during these last three years.

Some basic notation

We end this introduction with some probabilistic notation of current use in these lectures.

> *We could, of course, use any notation we want; do not laugh at notations; invent them, they are powerful. In fact, mathematics is, to a large extent, invention of better notations.* `Richard P. Feynman` (1918-1988).

We will use the symbol $a := b$ to define a mathematical object a in terms of b, or vice versa. We often use the letters m, n, p, q, k, l to denote integers and r, s, t, u to denote real numbers. We also use the capital letters $U, V, W X, Y, Z$ to denote random variables, and the letters u, v, w, x, y, z denote their possible outcomes.

Unless otherwise stated, S stands for some general state space model. These general state spaces and all the functions on S are assumed to be measurable; that is, they are equipped with some sigma field $\mathcal{S}$ so that the Lebesgue integral is well defined with respect to (w.r.t.) these functions (for instance $S = \mathbb{R}^d$ equipped with the sigma field generated by the open sets, as well as $\mathbb{N}^d$, $\mathbb{Z}^d$ or any other countable state space equipped with the discrete sigma field).

We also often use the letters f, g, h or F, G, H to denote functions on some state space, and μ, ν, η or $\mu(dx), \nu(dx), \eta(dx)$ for measures on some state space.

We let $\mathcal{M}(S)$ be the set of signed measures on some state space S, $\mathcal{P}(S) \subset \mathcal{M}(S)$ the subset of probability measures on the same state space S, and $\mathcal{B}(S)$ the set of bounded functions $f \ : \ x \in S \mapsto f(x) \in \mathbb{R}$.

We use the notation $dx (= dx_1 \times \ldots \times dx_k)$ to denote the Lebesgue measure on some Euclidian space $\mathbb{R}^k$, of some $k \geq 1$. For finite or countable state spaces, measures are identified to functions and we write $\mu(x), \nu(x), \eta(x)$ instead of $\mu(dx), \nu(dx), \eta(dx)$. The oscillations of a given bounded function f on some state space S are defined by $\mathrm{osc}(f) = \sup_{x,y \in S} |f(x) - f(y)|$.

We also use the proportionality sign $f \propto g$ between functions to state that $f = \lambda g$ for some $\lambda \in \mathbb{R}$.

Given a measure η on a state space S and a function f from S into $\mathbb{R}$ we set

$$\eta(f) = \int \eta(dx) f(x).$$

For multidimensional functions $f \ : \ x \in S \mapsto f(x) = (f_1(x), \ldots, f_r(x)) \in \mathbb{R}^r$, for some $r \geq 1$, we also set

$$\eta(f) = (\eta(f_1), \ldots, \eta(f_r)) \quad \text{and} \quad \eta(f^T) = (\eta(f_1), \ldots, \eta(f_r))^T$$

where a^T stands for the transpose of a vector $a \in \mathbb{R}^r$. For indicator functions $f = 1_A$, sometimes we slightly abuse notation and we set $\eta(A)$ instead of $\eta(1_A)$:

$$\eta(1_A) = \int \eta(dx) 1_A(x) = \int_A \eta(dx) = \eta(A).$$

We also consider the partial order relation between functions f_1, f_2 and measures μ_1, μ_2 given by

$$f_1 \leq f_2 \Longleftrightarrow \forall x \in S \quad f_1(x) \leq f_2(x)$$

and

$$\mu_1 \leq \mu_2 \Longleftrightarrow \forall A \in \mathcal{S} \quad \mu_1(A) \leq \mu_2(A).$$

The Dirac measure δ_a at some point $a \in S$ is defined by

$$\delta_a(f) = \int f(x) \delta_a(dx) = f(a).$$

When η is the distribution of some random variable X taking values in S, we have

$$\eta(dx) = \mathbb{P}(X \in dx) \quad \text{and} \quad \eta(f) = \mathbb{E}(f(X)).$$

For instance, the measure on $\mathbb{R}$ given by

$$\eta(dx) = \frac{1}{2} \left(\frac{1}{\sqrt{2\pi}} \ e^{-x^2/2} \ dx \right) + \frac{1}{2} \left(\frac{1}{2} \delta_0(dx) + \frac{1}{2} \delta_1(dx) \right)$$

represents the distribution of the random variable

$$X := \epsilon \ Y + (1 - \epsilon) Z$$

where (ϵ, Y, Z) are independent random variables with distribution

$$\begin{aligned} \mathbb{P}(\epsilon = 1) &= 1 - \mathbb{P}(\epsilon = 0) = 1/2 \\ \mathbb{P}(Z = 1) &= 1 - \mathbb{P}(Z = 0) = 1/2 \quad \text{and} \quad \mathbb{P}(Y \in dy) = \frac{1}{\sqrt{2\pi}} \ e^{-y^2/2} \ dy. \end{aligned}$$

For finite spaces of the form $S = \{e_1, \ldots, e_d\} \subset \mathbb{R}^d$, measures are defined by the weighted Dirac measures

$$\eta = \sum_{1 \leq i \leq d} w_i \ \delta_{e_i} \quad \text{with} \quad w_i = \eta(\{e_i\}) := \eta(e_i)$$

so that

$$\eta(f) = \int \eta(dx) f(x) = \sum_{1 \leq i \leq d} \eta(e_i) f(e_i).$$

Thus, if we identify measures and functions by the line and column vectors

$$\eta = [\eta(e_1), \ldots, \eta(e_d)] \quad \text{and} \quad f = \begin{pmatrix} f(e_1) \\ \vdots \\ f(e_d) \end{pmatrix} \tag{0.1}$$

we have

$$\eta f = [\eta(e_1), \ldots, \eta(e_d)] \begin{pmatrix} f(e_1) \\ \vdots \\ f(e_d) \end{pmatrix} = \sum_{1 \leq i \leq d} \eta(e_i) f(e_i) = \eta(f).$$

The Dirac measure δ_{e_i} is simply given by the line vector

$$\delta_{e_i} = \left[0, \ldots, 0, \underbrace{1}_{i-th}, 0, \ldots, 0 \right].$$

In this notation, probability measures on S can be interpreted as points $(\eta(e_i))_{1\leq i\leq d}$ in the $(d-1)$-dimensional simplex $\Delta_{d-1} \subset [0,1]^d$ defined by

$$\Delta_{d-1} = \left\{ (p_1, \ldots, p_d) \in [0,1]^d \; : \; \sum_{1\leq i\leq d} p_i = 1 \right\}. \tag{0.2}$$

We consider a couple of random variables (X_1, X_2) on a state space $(S_1 \times S_2)$, with marginal distributions

$$\eta_1(dx_1) = \mathbb{P}(X_1 \in dx_1) \quad \text{and} \quad \eta_2(dx_2) = \mathbb{P}(X_2 \in dx_2)$$

and conditional distribution

$$M(x_1, dx_2) = \mathbb{P}(X_2 \in dx_2 \mid X_1 = x_1).$$

For finite spaces of the form $S_1 = \{a_1, \ldots, a_{d_1}\}$ and $S_2 := \{b_1, \ldots, b_{d_2}\} \subset E = \mathbb{R}^d$, the above conditional distribution can be represented by a matrix

$$\begin{pmatrix} M(a_1, b_1) & M(a_1, b_2) & \ldots & M(a_1, b_{d_2}) \\ M(a_2, b_1) & M(a_2, b_2) & \ldots & M(a_2, b_{d_2}) \\ \vdots & \vdots & \vdots & \vdots \\ M(a_{d_1}, b_1) & M(a_2, b_2) & \ldots & M(a_{d_2}, b_{d_2}) \end{pmatrix}.$$

By construction, we have

$$\underbrace{\mathbb{P}(X_2 \in dx_2)}_{=\eta_2(dx_2)} = \int_{S_1} \underbrace{\mathbb{P}(X_1 \in dx_1)}_{\eta_1(dx_1)} \times \underbrace{\mathbb{P}(X_2 \in dx_2 \mid X_1 = x_1)}_{M(x_1, dx_2)}.$$

In other words, we have

$$\eta_2(dx_2) = \int_{S_1} \eta_1(dx_1) \; M(x_1, dx_2) := (\eta_1 M)(dx_2)$$

or in a more synthetic form $\eta_2 = \eta_1 M$.

Notice that for the finite state space model discussed above we have the matrix formulation

$$\begin{aligned} \eta_2 &= [\eta_2(b_1), \ldots, \eta_2(b_{d_2})] \\ &= [\eta_1(a_1), \ldots, \eta_1(a_{d_1})] \begin{pmatrix} M(a_1, b_1) & M(a_1, b_2) & \ldots & M(a_1, b_{d_2}) \\ M(a_2, b_1) & M(a_2, b_2) & \ldots & M(a_2, b_{d_2}) \\ \vdots & \vdots & \vdots & \vdots \\ M(a_{d_1}, b_1) & M(a_2, b_2) & \ldots & M(a_{d_2}, b_{d_2}) \end{pmatrix} \\ &= \eta_1 M. \end{aligned}$$

In this context, a matrix M with positive entries whose rows sum to 1 is also called a stochastic matrix.

Given a function f on S_2, we consider the function $M(f)$ on S_1 defined by

$$M(f)(x_1) = \int_{S_2} M(x_1, dx_2)\ f(x_2) = \mathbb{E}\left(f(X_2) \mid X_1 = x_1\right).$$

For functions $f\ :\ x \in S_2 \mapsto f(x) = (f_1(x), \ldots, f_r(x)) \in \mathbb{R}^r$ we also set

$$M(f)(x_1) = \int_{S_2} M(x_1, dx_2)\ f(x_2) = \mathbb{E}\left(f(X_2) \mid X_1 = x_1\right) = (M(f_r)(x_1), \ldots, M(f_r)(x_1))$$

or $M(f^T)(x_1) = (M(f_r)(x_1), \ldots, M(f_r)(x_1))^T$ for multidimensional functions defined in terms of column vectors.

Here again, for the finite state space model discussed above, these definitions resume to matrix operations

$$\begin{aligned} M(f) &= \begin{pmatrix} M(f)(a_1) \\ \vdots \\ M(f)(a_{d_1}) \end{pmatrix} \\ &= \begin{pmatrix} M(a_1, b_1) & M(a_1, b_2) & \ldots & M(a_1, b_{d_2}) \\ M(a_2, b_1) & M(a_2, b_2) & \ldots & M(a_2, b_{d_2}) \\ \vdots & \vdots & \vdots & \vdots \\ M(a_{d_1}, b_1) & M(a_2, b_2) & \ldots & M(a_{d_2}, b_{d_2}) \end{pmatrix} \begin{pmatrix} f(b_1) \\ \vdots \\ f(b_{d_2}) \end{pmatrix}. \end{aligned}$$

By construction,

$$\eta_1(M(f)) = (\eta_1 M)(f) = \eta_2(f) \Longleftrightarrow \mathbb{E}\left(\mathbb{E}(f(X_2)|X_1)\right) = \mathbb{E}(f(X_2)).$$

Given some matrices M, M_1 and M_2, we denote by $M_1 M_2$ the composition of the matrices M_1 and M_2, and by $M^n = M^{n-1}M = MM^{n-1}$ the n iterates of M. For $n = 0$, we use the convention $M^0 = Id$, the identity matrix on S.

We use the same integral operations for any bounded integral operator. For instance, if $Q(x_1, dx_2) := G_1(x_1) M(x_1, dx_2) G_2(x_2)$ for some bounded functions G_1 and G_2 on S_1 and S_2 we set

$$(\eta Q)(dx_2) = \int_{S_1} \eta_1(dx_1)\ Q(x_1, dx_2) \quad \text{and} \quad Q(f)(x_1) = \int_{S_2} Q(x_1, dx_2)\ f(x_2)$$

for any measure η on S_1 and any function f on S_2.

When (S, d) is equipped with a metric d, we denote by $\mathrm{Lip}(S)$ the set of Lipschitz functions f such that

$$\mathrm{lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)} < \infty$$

and $\mathrm{BLip}(S) = \mathrm{Lip}(S) \cap \mathcal{B}(S)$ the subset of bounded Lipschitz functions equipped with the norm

$$\mathrm{blip}(f) := \|f\| + \mathrm{lip}(f).$$

In Leibniz notation, the partial derivative of a smooth function f on a product space $S = \mathbb{R}^r$ w.r.t. the i-th coordinate is denoted by

$$y \mapsto \frac{\partial f}{\partial x_i}(y) = \lim_{\epsilon \downarrow 0} \epsilon^{-1}\ [f(y_1, \ldots, y_{i-1}, y_i + \epsilon, y_{i+1}, \ldots, y_r) - f(y_1, \ldots, y_{i-1}, y_i, y_{i+1}, \ldots, y_r)]$$

with $1 \leq i \leq r$. We also denote by $\frac{\partial^2 f}{\partial x_i \partial x_j}$ the second order derivatives defined as above by replacing f by the function $\frac{\partial f}{\partial x_j}$, with $1 \leq i,j \leq r$. High order derivatives $\frac{\partial^n f}{\partial x_{i_1} \ldots \partial x_{i_n}}$ are defined in the same way. We also often use the following synthetic notation

$$\partial_{x_i} f := \frac{\partial f}{\partial x_i} \qquad \partial_{x_i,x_j} f := \frac{\partial^2 f}{\partial x_i \partial x_j} \quad \text{and} \quad \partial_{x_{i_1},\ldots,x_{i_n}} f := \frac{\partial^n f}{\partial x_{i_1} \ldots \partial x_{i_n}}$$

as well as $\partial^2_{x_i} = \partial_{x_i,x_i}$, and $\partial^3_{x_i} = \partial_{x_i,x_i,x_i}$, and so on for any $1 \leq i \leq r$. When there is no confusion w.r.t. the coordinate system, we also use the shorthand

$$\partial_i = \partial_{x_i} \qquad \partial_{i,j} := \partial_{x_i,x_j} \qquad \partial^2_i := \partial_{x_i,x_i} \quad \text{and} \quad \partial_{i_1,\ldots,i_n} = \partial_{x_{i_1},\ldots,x_{i_n}}.$$

When $r = 1$ we often use the Lagrange notation

$$f' = \frac{\partial f}{\partial x} \qquad f'' = \frac{\partial^2 f}{\partial^2 x} \quad \text{and} \quad f^{(n)} = \frac{\partial^n f}{\partial^n x}$$

for any $n \geq 0$. For $n = 0$, we use the convention $f^{(0)} = f$. All of the above partial derivatives can be interpreted as operators on functional spaces mapping a smooth function f into another function with less regularity, for instance, if f is three times differentiable f' is only twice differentiable, and so on.

We also use integro-differential operators L defined by

$$L(f)(x) = \sum_{1 \leq i \leq r} a^i(x)\, \partial_i f(x) + \frac{1}{2} \sum_{1 \leq i,j \leq r} b^{i,j}(x)\, \partial_{i,j} f(x) + \lambda(x) \int (f(y) - f(x))\, M(x,dy)$$

for some functions $a(x) = (a^i(x))_{1 \leq i \leq r}$, $b(x) = (b^{i,j}(x))_{1 \leq i,j \leq r}$, $\lambda(x)$ and some Markov transition M.

The operator L maps functions which are at least twice differentiable into functions with less regularity. For instance, L maps the set of twice continuously differentiable functions into the set of continuous functions, as soon as the functions a, b, λ are continuous. Notice that L also maps the set of infinitely differentiable functions into itself, as soon as the functions a, b, λ are infinitely differentiable functions. When $a = 0 = b$, the operator L resumes to an integral operator and it maps the set of bounded integrable functions into itself, as soon as the functions a, b, λ are bounded. To avoid repetition, these operators are assumed to be defined on sufficiently smooth functions. We often use the terms "sufficiently smooth" or "sufficiently regular" functions to avoid entering into unnecessary discussions on the domain of definition of these operators.

In theoretical and computational quantum physics, the inner product and more generally dual operators on vector spaces are often represented using a bra-ket formalism introduced at the end of the 1930s by P. Dirac [108] to avoid too sophisticated matrix operations (not so developed and of current use in the beginning of the 20th century).

For finite d-dimensional Euclidian vector spaces $\mathbb{R}^d$, the bras $\prec \alpha \mid$ and the kets $\mid \beta \succ$ are simply given for any $\alpha = (\alpha_i)_{1 \leq i \leq d} \in \mathbb{R}^d$ and $\beta = (\beta_i)_{1 \leq i \leq d} \in \mathbb{R}^d$ row and column

$$\prec \alpha \mid := [\alpha_1, \ldots, \alpha_d] \quad \text{and} \quad \mid \beta \succ := \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_d \end{bmatrix} \quad \Rightarrow \prec \alpha \mid\mid \beta \succ := \prec \alpha \mid \beta \succ := \sum_{1 \leq i \leq d} \alpha_i \beta_i.$$

In much the same way, the product of bra $\prec \alpha \mid$ with a linear (matrix) operator Q corresponds to the product of the row vector by the matrix

$$\prec \alpha \mid Q := [\alpha_1, \ldots, \alpha_d] \left[\begin{array}{cccc} Q(1,1) & Q(1,2) & \ldots & Q(1,d) \\ Q(2,1) & Q(2,2) & \ldots & Q(2,d) \\ \vdots & \vdots & \ldots & \vdots \\ Q(d,1) & Q(d,2) & \ldots & Q(d,d) \end{array} \right].$$

Likewise, the product of a linear (matrix) operator Q with a ket $\mid \beta \succ$ corresponds to the product of the matrix by the column vector

$$Q \mid \beta \succ := \left[\begin{array}{cccc} Q(1,1) & Q(1,2) & \ldots & Q(1,d) \\ Q(2,1) & Q(2,2) & \ldots & Q(2,d) \\ \vdots & \vdots & \ldots & \vdots \\ Q(d,1) & Q(d,2) & \ldots & Q(d,d) \end{array} \right] \left[\begin{array}{c} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{array} \right].$$

Combining these operations, we find that

$$\prec \alpha \mid Q \mid \beta \succ = [\alpha_1, \ldots, \alpha_d] \left[\begin{array}{cccc} Q(1,1) & Q(1,2) & \ldots & Q(1,d) \\ Q(2,1) & Q(2,2) & \ldots & Q(2,d) \\ \vdots & \vdots & \ldots & \vdots \\ Q(d,1) & Q(d,2) & \ldots & Q(d,d) \end{array} \right] \left[\begin{array}{c} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{array} \right].$$

Using the vector representation (0.1) of functions f and measures η on finite state spaces $S = \{e_1, \ldots, e_d\}$, the duality formula between functions and measures takes the following form

$$\prec \eta \mid f \succ = \eta f = \sum_{1 \leq i \leq d} \eta(e_i) f(e_i) := \eta(f).$$

Likewise, for any Markov transition M from $E_1 = \{a_1, \ldots, a_{d_1}\}$ into $E_2 := \{b_1, \ldots, b_{d_2}\}$, and function f on E_2 and any measure η_1 on E_1, we have

$$\prec \eta_1 \mid M \mid f \succ = \prec \eta_1 \mid Mf \succ = \eta_1(Mf) = (\eta_1 M) f = \prec \eta_1 M \mid f \succ.$$

The bra-ket formalism is extended to differential operations by setting

$$\prec g \mid L \mid f \succ = \int g(x) \; L(f)(x) \; dx := \langle g, L(f) \rangle.$$

Given a positive and bounded function G on some state space S, we denote by Ψ_G the Boltzmann-Gibbs mapping from $\mathcal{P}(S)$ into itself, defined for any $\mu \in \mathcal{P}(S)$ by

$$\Psi_G(\mu)(dx) = \frac{1}{\mu(G)} \; G(x) \; \mu(dx). \tag{0.3}$$

In other words, for any bounded function f on S we have

$$\Psi_G(\mu)(f) = \int \Psi_G(\mu)(dx) f(x) = \frac{\mu(Gf)}{\mu(G)}.$$

We let $\mu = \mathrm{Law}(X)$ be the distribution of a random variable X and $G = 1_A$ the indicator function of some subset $A \subset S$. In this situation, we have

$$\Psi_G(\mu)(f) = \frac{\mu(Gf)}{\mu(G)} = \frac{\mathbb{E}(f(X) 1_A(X))}{\mathbb{E}(1_A(X))} = \mathbb{E}\left(f(X) \mid X \in A\right).$$

In other words, we have

$$\Psi_{1_A}(\mu) = \text{Law}(X \mid X \in A).$$

Let (X, Y) be a couple of r.v. with probability density $p(x, y)$ on $\mathbb{R}^{d+d'}$. With a slight abuse of notation, we recall that the conditional density $p(x|y)$ of X given Y is given by the Bayes' formula

$$p(x|y) = \frac{1}{p(y)}\, p(y|x)\, p(x) \quad p(y) = \int p(y|x)\, p(x)\, dx.$$

In other words,

$$\mu(dx) := p(x)dx \quad \text{and} \quad G_y(x) := p(y|x) \Rightarrow \Psi_{G_y}(\mu)(dx) = p(x|y)\, dx.$$

Given a random matrix $A = (A_{i,j})_{i,j}$, we denote by $\mathbb{E}(A) = (\mathbb{E}(A_{i,j}))_{i,j}$ the matrix of the mean values of its entries. We slightly abuse notation and we denote by 0 the null real number and the null matrix. Given some $\mathbb{R}^d$-valued random variables X, Y we denote by $\text{Cov}(X, Y)$ the covariance matrix

$$\text{Cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))'\right).$$

Sometimes, we will also use the conditional covariance w.r.t. some auxiliary random variable Z given by

$$\text{Cov}((X, Y) \mid Z) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))' \mid Z\right).$$

For one-dimensional random variables, the variance of a random variable is given by $\text{Var}(X) = \text{Cov}(X, X)$, and

$$\text{Var}(X \mid Z) = \text{Cov}((X, X) \mid Z) = \frac{1}{2}\, \mathbb{E}((X - X')^2 \mid Z), \tag{0.4}$$

where X, X' are two independent copies of X given Z. We also use the notation $\sigma(X)$ to denote the σ-algebra generated by some possibly multi-dimensional r.v. X on some state space S. In this case, for any real valued random variable Y, we recall that

$$\mathbb{E}(Y \mid \sigma(X)) = \mathbb{E}(Y \mid X)$$

which has to be distinguished from the conditional expectations w.r.t. some event, say $X \in A$, for some subset $A \in \sigma(X)$

$$\mathbb{E}(Y \mid X \in A) = \mathbb{E}(Y\, 1_A(X))/\mathbb{E}(1_A(X)).$$

We also write $X \perp Y$ when a random variable X is independent of Y

$$X \text{ independent of } Y \Longrightarrow_{def.} X \perp Y.$$

The maximum and minimum operations are denoted respectively by

$$a \vee b := \max\{a, b\} \quad a \wedge b := \min\{a, b\} \quad \text{as well as} \quad a_+ := a \vee 0 \quad \text{and} \quad a_- := -(a \wedge 0)$$

so that $a = a_+ - a_-$ and $|a| = a_+ + a_-$. We also denote by $\lfloor a \rfloor$ and $\{a\} = a - \lfloor a \rfloor$ the integer part, and respectively the fractional part, of some real number a.

We also use the Bachmann-Landau notation

$$f(\epsilon) = g(\epsilon) + \text{O}(\epsilon) \Longleftrightarrow \limsup_{\epsilon \to 0} \frac{1}{\epsilon}\, |f(\epsilon) - g(\epsilon)| < \infty$$

and

$$f(\epsilon) = g(\epsilon) + \mathrm{o}(\epsilon) \Longleftrightarrow \limsup_{\epsilon \to 0} \frac{1}{\epsilon} |f(\epsilon) - g(\epsilon)| = 0.$$

When there is no confusion, sometimes we write $\mathrm{o}(1)$ for a function that tends to 0 when the parameter $\epsilon \to 0$. We also denote by $\mathrm{O}_P(\epsilon)$ some possibly random function such that

$$\mathbb{E}\left(|\mathrm{O}_P(\epsilon)|\right) = \mathrm{O}(\epsilon).$$

We also use the traditional conventions

$$\prod_{\emptyset} = 1 \qquad \sum_{\emptyset} = 0 \qquad \inf_{\emptyset} = \infty \quad \text{and} \quad \sup_{\emptyset} = -\infty.$$

Jacques Neveu, Poster Random Models and Statistics, MAS - SMAI Conferences, Prize of the Applied and Industrial Applied Mathematics Society (SMAI), France. *Graphical conception: Timothée Del Moral.*