Bayesian Filtering and Smoothing

Second Edition

Now in its second edition, this accessible text presents a unified Bayesian treatment of the state-of-the-art filtering, smoothing, and parameter estimation algorithms for non-linear state space models. The book focuses on discrete-time state space models and carefully introduces fundamental aspects related to optimal filtering and smoothing. In particular, it covers a range of efficient non-linear Gaussian filtering and smoothing algorithms, as well as Monte Carlo-based algorithms.

This updated edition features new chapters on constructing state space models of practical systems, the discretization of continuous-time state space models, Gaussian filtering by enabling approximations, posterior linearization filtering, and the corresponding smoothers. Coverage of key topics is expanded, including extended Kalman filtering and smoothing, and parameter estimation.

The book's practical, algorithmic approach assumes only modest mathematical prerequisites, suitable for graduate and advanced undergraduate students. Many examples are included, with the MATLAB and Python code available online, enabling readers to implement the algorithms in their own projects.

SIMO SÄRKKÄ is Associate Professor in the Department of Electrical Engineering and Automation at Aalto University, Finland. His research interests center on state estimation and stochastic modeling, and he has authored two books (2013 & 2019) on these topics. He is Fellow of ELLIS, Senior Member of IEEE, a recipient of multiple paper awards, and has been chair of MLSP and FUSION conferences.

LENNART SVENSSON is Professor in the Department of Electrical Engineering at Chalmers University of Technology, Gothenberg. His research focuses on non-linear filtering, deep learning, and tracking in particular. He has organized a massive open online course on multiple object tracking and received paper awards at the International Conference on Information Fusion in 2009, 2010, 2017, 2019, and 2021.

INSTITUTE OF MATHEMATICAL STATISTICS TEXTBOOKS

IMS Textbooks give introductory accounts of topics of current concern suitable for advanced courses at master's level, for doctoral students and for individual study. They are typically shorter than a fully developed textbook, often arising from material created for a topical course. Lengths of 100–290 pages are envisaged. The books typically contain exercises.

In collaboration with the International Society for Bayesian Analysis (ISBA), selected volumes in the IMS Textbooks series carry the "with ISBA" designation at the recommendation of the ISBA editorial representative.

Other Books in the Series (*with ISBA)

1. *Probability on Graphs*, by Geoffrey Grimmett
2. *Stochastic Networks*, by Frank Kelly and Elena Yudovina
3. *Bayesian Filtering and Smoothing*, by Simo Särkkä
4. *The Surprising Mathematics of Longest Increasing Subsequences*, by Dan Romik
5. *Noise Sensitivity of Boolean Functions and Percolation*, by Christophe Garban and Jeffrey E. Steif
6. *Core Statistics*, by Simon N. Wood
7. *Lectures on the Poisson Process*, by Günter Last and Mathew Penrose
8. *Probability on Graphs (Second Edition)*, by Geoffrey Grimmett
9. *Introduction to Malliavin Calculus*, by David Nualart and Eulália Nualart
10. *Applied Stochastic Differential Equations*, by Simo Särkkä and Arno Solin
11. **Computational Bayesian Statistics*, by M. Antónia Amaral Turkman, Carlos Daniel Paulino, and Peter Müller
12. *Statistical Modelling by Exponential Families*, by Rolf Sundberg
13. *Two-Dimensional Random Walk: From Path Counting to Random Interlacements*, by Serguei Popov
14. *Scheduling and Control of Queueing Networks*, by Gideon Weiss
15. *Principles of Statistical Analysis: Learning from Randomized Experiments*, by Ery Arias-Castro
16. *Exponential Families in Theory and Practice*, by Bradley Efron

Bayesian Filtering and Smoothing

SECOND EDITION

SIMO SÄRKKÄ
Aalto University, Finland

LENNART SVENSSON
Chalmers University of Technology, Gothenberg

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi - 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence

www.cambridge.org
Information on this title: www.cambridge.org/9781108926645

DOI: 10.1017/9781108917407

First published 2013
Second edition 2023

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Särkkä, Simo, author. | Svensson, Lennart, 1976- author.
Title: Bayesian filtering and smoothing / Simo Särkkä and Lennart Svensson.
Description: Second edition. | New York : Cambridge University Press, 2023. |
Series: Institute of Mathematical Statistics textbooks | Revised
edition of: Bayesian filtering and smoothing / Simo Sarkka. 2013. |
Includes bibliographical references and index.
Identifiers: LCCN 2022058412 | ISBN 9781108926645 (paperback)
Subjects: LCSH: Bayesian statistical decision theory. | Filters
(Mathematics) | Smoothing (Statistics)
Classification: LCC QA279.5 .S27 2023 | DDC 519.5/42–dc23/eng20230323
LC record available at https://lccn.loc.gov/2022058412

ISBN 978-1-108-92664-5 Paperback

Contents

Preface

The aim of this book is to give a concise introduction to non-linear Kalman filtering and smoothing, particle filtering and smoothing, and to the related parameter estimation methods. Although the book is intended to be an introduction, the mathematical ideas behind all the methods are carefully explained, and a mathematically inclined reader can get quite a deep understanding of the methods by reading the book. The book is purposely kept relatively short for quick reading.

The book is mainly intended for advanced undergraduate and graduate students in applied mathematics, computer science, and electrical engineering. However, the book is also suitable for researchers and practitioners (engineers) who need a concise introduction to the topic on a level that enables them to implement or use the methods. Readers are assumed to have a background in linear algebra, vector calculus, and Bayesian inference, and MATLAB or Python programming skills.

As implied by the title, the mathematical treatment of the models and algorithms in this book is Bayesian, which means that all the results are treated as being approximations to certain probability distributions or their parameters. Probability distributions are used both to represent uncertainties in the models and to model the physical randomness. The theories of non-linear filtering, smoothing, and parameter estimation are formulated in terms of Bayesian inference, and both the classical and recent algorithms are derived using the same Bayesian notation and formalism. This Bayesian approach to the topic is far from new and was pioneered by Stratonovich in the 1950s and 1960s – even before Kalman's seminal article in 1960. Thus the theory of non-linear filtering has been Bayesian from the beginning (see Jazwinski, 1970).

The main additions to the second edition of the book are the chapters on how to construct state space models of practical systems along with coverage of the iterated extended Kalman filters and smoothers, generalized statistical linear regression based filters and smoothers, and posterior linearization filters and smoothers. These additions have also resulted in a slight reordering of the material related to the coverage of Gaussian filters and smoothers. Methods for Bayesian estimation in discrete state systems, including, for example, the Viterbi algorithm, are now also covered.

Chapter 1 is a general introduction to the idea and applications of Bayesian filtering and smoothing. The purpose of Chapter 2 is to briefly review the basic concepts of Bayesian inference as well as the basic numerical methods used in Bayesian computations. Chapter 3 starts with a step-by-step introduction to recursive Bayesian estimation by demonstrating how to recursively solve a linear regression problem. The transition to Bayesian filtering and smoothing theory is explained by extending and generalizing the problem. The first Kalman filter of the book is also encountered in this chapter.

Chapters 4 and 5 are concerned with practical modeling with state space models. In particular, Chapter 4 is concerned with transforming continuous-time models of tracking models into discrete-time state space models that are compatible with the discrete-time estimation methods considered in this book, as well as examples of dynamic models. Chapter 5 proceeds to augment the models with linear, non-linear, Gaussian, and non-Gaussian measurement models and explains how certain classes of machine learning and signal processing models can be recast as state space models.

The Bayesian filtering theory starts in Chapter 6 where we derive the general Bayesian filtering equations and, as their special case, the celebrated Kalman filter, along with discrete state Bayesian filters. Taylor series-based non-linear extensions of the Kalman filter, the extended Kalman filter (EKF), and iterated extended Kalman filter (IEKF) are presented in Chapter 7. After that, Chapter 8 starts by introducing the moment matching-based general Gaussian filter algorithm, and the Gauss–Hermite Kalman filter (GHKF), cubature Kalman filter (CKF), unscented Kalman filter (UKF), and higher order cubature/unscented Kalman filters are then derived as special cases of it.

Chapter 9 introduces a different perspective and reformulates all the Gaussian filters in terms of enabling linearizations. The presentation starts with statistical linearization and the statistically linearized filter (SLF), and proceeds to statistical linear regression (SLR) and the related filters, which turn out to recover and extend all the Gaussian filters covered in the previous chapters. By further extending the concept of enabling linearizations, Chapter 10 introduces the posterior linearization filter (PLF), which generalizes the concept of iterated Gaussian filtering. Sequential Monte Carlo

(SMC)-based particle filters (PF) are explained in Chapter 11 by starting from the basic SIR filter and ending with Rao–Blackwellized particle filters (RBPF).

Chapter 12 starts with a derivation of the general (fixed-interval) Bayesian smoothing equations and then continues to a derivation of the Rauch–Tung–Striebel (RTS) smoother as their special case. In that chapter, we also present methods for smoothing in discrete state systems, including the Viterbi algorithm. The extended RTS smoother (ERTSS) and the iterated extended RTS smoother (IERTSS) are presented in Chapter 13. The general Gaussian smoothing framework is presented in Chapter 14, and the Gauss–Hermite RTS smoother (GHRTSS), cubature RTS smoother (CRTSS), unscented RTS smoother (URTSS), and higher order cubature/unscented RTS smoothers are derived as its special cases. The chapter then proceeds to the iterated posterior linearization smoother (IPLS), which generalizes the concept of iterated Gaussian smoothing.

In Chapter 15 we start by showing how the basic SIR particle filter can be used to approximate the smoothing solutions with a minor modification. We then introduce the numerically superior backward-simulation particle smoother and the reweighting (or marginal) particle smoother. Finally, we discuss the implementation of Rao–Blackwellized particle smoothers.

Chapter 16 is an introduction to parameter estimation in state space models concentrating on optimization and expectation-maximization (EM)-based computation of maximum likelihood (ML) and maximum a posteriori (MAP) estimates, as well as on Markov chain Monte Carlo (MCMC) methods. We start by presenting the general methods and then show how Kalman filters and RTS smoothers, non-linear Gaussian filters and RTS smoothers, and finally particle filters and smoothers, can be used to compute or approximate the quantities needed in the implementation of parameter estimation methods. This leads to, for example, classical EM algorithms for state space models, as well as to particle EM and particle MCMC methods. We also discuss how Rao–Blackwellization can sometimes be used to help parameter estimation.

Chapter 17 is an epilogue where we give general advice on selecting different methods for different purposes. We also discuss and give references to various technical points and related topics that are important but did not fit into this book.

Each of the chapters ends with a range of exercises that give the reader hands-on experience in implementing the methods and selecting the appropriate method for a given purpose. The MATLAB and Python source code needed in the exercises as well as much other material can be found on the book's web page.

We are grateful to many people who carefully checked the book and gave many valuable suggestions for improving the text. It is not possible to include all of them, but we would like to specifically mention Arno Solin, Robert Piché, Juha Sarmavuori, Thomas Schön, Pete Bunch, Isambi S. Mbalawata, Adrien Corenflos, Fatemeh Yaghoobi, Hany Abdulsamad, Jakob Lindqvist, and Lars Hammarstrand. We are also grateful to Jouko Lampinen, Aki Vehtari, Jouni Hartikainen, Ville Väänänen, Heikki Haario, Simon Godsill, Ángel García-Fernández, Filip Tronarp, Toni Karvonen, and various others for research co-operation that led to improvement of our understanding of the topic as well as to the development of some of the methods that now are explained in this book. We would also like to thank the editors of Cambridge University Press for their original suggestion for the publication of the book. We are also grateful to our families for their support and patience during the writing of this book.

Simo and Lennart

Symbols and Abbreviations

General Notation

$a, b, c, x, t, \alpha, \beta$	Scalars
$\mathbf{a}, \mathbf{f}, \mathbf{s}, \mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\beta}$	Vectors
$\mathbf{A}, \mathbf{F}, \mathbf{S}, \mathbf{X}, \mathbf{Y}$	Matrices
$\mathbb{A}, \mathbb{F}, \mathbb{S}, \mathbb{X}, \mathbb{Y}$	Spaces or sets

Notational Conventions

$\mathbf{A}^\mathsf{T}$	Transpose of matrix
$\mathbf{A}^{-1}$	Inverse of matrix
$\mathbf{A}^{-\mathsf{T}}$	Inverse of transpose of matrix
$[\mathbf{A}]_i$	ith column of matrix $\mathbf{A}$
$[\mathbf{A}]_{ij}$	Element at ith row and jth column of matrix $\mathbf{A}$
$\lvert a \rvert$	Absolute value of scalar a
$\lvert \mathbf{A} \rvert$	Determinant of matrix $\mathbf{A}$
$\mathrm{d}\mathbf{x}/\mathrm{d}t$	Time derivative of $\mathbf{x}(t)$
$\frac{\partial g_i(\mathbf{x})}{\partial x_j}$	Partial derivative of g_i with respect to x_j
$(a_1, \ldots, a_n)$	Column vector with elements $a_1, \ldots, a_n$
$(a_1 \; \cdots \; a_n)$	Row vector with elements $a_1, \ldots, a_n$
$(a_1 \; \cdots \; a_n)^\mathsf{T}$	Column vector with elements $a_1, \ldots, a_n$
$\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}$	Gradient (column vector) of scalar function g
$\frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}$	Jacobian matrix of vector-valued function $\mathbf{x} \mapsto \mathbf{g}(\mathbf{x})$
$\mathrm{Cov}[\mathbf{x}]$	Covariance $\mathrm{Cov}[\mathbf{x}] = \mathrm{E}[(\mathbf{x} - \mathrm{E}[\mathbf{x}])\,(\mathbf{x} - \mathrm{E}[\mathbf{x}])^\mathsf{T}]$ of the random variable $\mathbf{x}$
$\mathrm{Cov}[\mathbf{x}, \mathbf{y}]$	Cross-covariance $\mathrm{Cov}[\mathbf{x}, \mathbf{y}] = \mathrm{E}[(\mathbf{x} - \mathrm{E}[\mathbf{x}])\,(\mathbf{y} - \mathrm{E}[\mathbf{y}])^\mathsf{T}]$ of the random variables $\mathbf{x}$ and $\mathbf{y}$
$\mathrm{diag}(a_1, \ldots, a_n)$	Diagonal matrix with diagonal values $a_1, \ldots, a_n$

$\sqrt{\mathbf{P}}$	Matrix such that $\mathbf{P} = \sqrt{\mathbf{P}}\,\sqrt{\mathbf{P}}^\top$
$\mathbf{P}^{-1/2}$	Alternative notation for $[\sqrt{\mathbf{P}}]^{-1}$
$\mathrm{E}[\mathbf{x}]$	Expectation of $\mathbf{x}$
$\mathrm{E}[\mathbf{x} \mid \mathbf{y}]$	Conditional expectation of $\mathbf{x}$ given $\mathbf{y}$
$\int f(\mathbf{x})\,\mathrm{d}\mathbf{x}$	Integral of $f(\mathbf{x})$ over the space $\mathbb{R}^n$
$\int_a^b g(t)\,\mathrm{d}t$	Integral of $g(t)$ over the interval $t \in [a, b]$
$p(\mathbf{x})$	Probability density of continuous random variable $\mathbf{x}$ or probability of discrete random variable $\mathbf{x}$
$p(\mathbf{x} \mid \mathbf{y})$	Conditional probability density or conditional probability of $\mathbf{x}$ given $\mathbf{y}$
$p(\mathbf{x}) \propto q(\mathbf{x})$	$p(\mathbf{x})$ is proportional to $q(\mathbf{x})$, that is, there exists a constant c such that $p(\mathbf{x}) = c\, q(\mathbf{x})$ for all values of $\mathbf{x}$
$\mathrm{tr}\,\mathbf{A}$	Trace of matrix $\mathbf{A}$
$\mathrm{Var}[x]$	Variance $\mathrm{Var}[x] = \mathrm{E}[(x - \mathrm{E}[x])^2]$ of the scalar random variable x
$x \gg y$	x is much greater than y
$x_{i,k}$	ith component of vector $\mathbf{x}_k$
$\mathbf{x} \sim p(\mathbf{x})$	Random variable $\mathbf{x}$ has the probability density or probability distribution $p(\mathbf{x})$
$\mathbf{x} \triangleq \mathbf{y}$	$\mathbf{x}$ is defined to be equal to $\mathbf{y}$
$\mathbf{x} \approx \mathbf{y}$	$\mathbf{x}$ is approximately equal to $\mathbf{y}$
$\mathbf{x} \simeq \mathbf{y}$	$\mathbf{x}$ is assumed to be approximately equal to $\mathbf{y}$
$\mathbf{x}_{0:k}$	Set or sequence containing the vectors $\{\mathbf{x}_0, \dots, \mathbf{x}_k\}$
$\dot{\mathbf{x}}$	Time derivative of $\mathbf{x}(t)$

Symbols

α	Parameter of the unscented transform or a pendulum angle
α_i	Acceptance probability in an MCMC method
$\bar{\alpha}_*$	Target acceptance rate in an adaptive MCMC method
β	Parameter of the unscented transform, or a parameter of a dynamic model
γ	Step size in line search
$\Gamma(\cdot)$	Gamma function
$\delta(\cdot)$	Dirac delta function, or steering angle
$\delta\mathbf{x}$	Difference of $\mathbf{x}$ from the mean $\delta\mathbf{x} = \mathbf{x} - \mathbf{m}$
Δt	Sampling period
Δt_k	Length of time interval $\Delta t_k = t_{k+1} - t_k$

ε_k	Measurement error at time step k
$\boldsymbol{\varepsilon}_k$	Vector of measurement errors at time step k
θ	A parameter or heading angle
$\boldsymbol{\theta}$	Vector of parameters
$\boldsymbol{\theta}_k$	Vector of parameters at time step k
$\boldsymbol{\theta}^{(n)}$	Vector of parameters at iteration n of the EM-algorithm
$\boldsymbol{\theta}^{(i)}$	Vector of parameters at iteration i of an MCMC-algorithm
$\boldsymbol{\theta}^*$	Candidate point in an MCMC-algorithm
$\hat{\boldsymbol{\theta}}^{\mathrm{MAP}}$	Maximum a posteriori (MAP) estimate of parameter $\boldsymbol{\theta}$
κ	Parameter of the unscented transform or auxiliary variable
λ	Parameter of the unscented transform or the Poisson distribution, or regularization parameter
λ_0	Parameter of the Poisson distribution
$\boldsymbol{\Lambda}$	Noise covariance in (generalized) statistical linear regression
$\boldsymbol{\Lambda}^{(i)}$	Noise covariance on the ith iteration of posterior linearization
$\boldsymbol{\Lambda}_k$	Noise covariance in a Gaussian enabling approximation of a dynamic model
$\boldsymbol{\mu}$	Mean of Student's t-distribution
$\boldsymbol{\mu}^{+,(i-1)}$	Predicted mean from iteration $i-1$
$\boldsymbol{\mu}^{(i)}$	Predicted measurement model mean at iteration i
$\boldsymbol{\mu}_k^-$	Predicted dynamic model mean in a statistical linear regression filter
$\boldsymbol{\mu}_k^+$	Predicted measurement model mean in a statistical linear regression filter
$\boldsymbol{\mu}_k$	Predicted mean of measurement $\mathbf{y}_k$ at time step k
$\boldsymbol{\mu}_k^-(\mathbf{x}_{k-1})$	Conditional mean moment of a dynamic model
$\boldsymbol{\mu}_k(\mathbf{x}_k)$	Conditional mean moment of a measurement model
$\boldsymbol{\mu}_k^{-(i)}$	Predicted dynamic model mean for sigma point i at time step k
$\boldsymbol{\mu}_k^{+,(i-1)}$	Predicted mean from iteration $i-1$
$\boldsymbol{\mu}_k^{(i)}$	Predicted measurement model mean for sigma point i at time step k or predicted measurement model mean at ith iteration
$\boldsymbol{\mu}_{\mathrm{G}}$	Mean in generalized statistical linear regression approximation
$\boldsymbol{\mu}_{\mathrm{L}}$	Mean in the linear (Taylor series-based) approximation
$\boldsymbol{\mu}_{\mathrm{M}}$	Mean in the Gaussian moment matching approximation
$\boldsymbol{\mu}_{\mathrm{Q}}$	Mean in the quadratic approximation
$\boldsymbol{\mu}_{\mathrm{R}}$	Mean in the statistical linear regression approximation
$\boldsymbol{\mu}_{\mathrm{R}}(\mathbf{x})$	Conditional mean moment in statistical linear regression

$\boldsymbol{\mu}_{\mathrm{S}}$	Mean in the statistical linearization approximation
$\boldsymbol{\mu}_{\mathrm{U}}$	Mean in the unscented approximation
ν	Degrees of freedom in Student's t-distribution
ξ	Unit Gaussian random variable
$\xi^{(i)}$	ith scalar unit sigma point
$\boldsymbol{\xi}$	Vector of unit Gaussian random variables
$\boldsymbol{\xi}^{(i)}$	ith unit sigma point vector
$\boldsymbol{\xi}^{(i_1,\ldots,i_n)}$	Unit sigma point in the multivariate Gauss–Hermite cubature
π	Constant $\pi = 3.14159265358979323846\ldots$
$\pi(\cdot)$	Importance distribution or linearization distribution
$\pi_k^{\mathbf{f}}(\mathbf{x}_k)$	Linearization distribution of a dynamic model
$\pi_k^{\mathbf{h}}(\mathbf{x}_k)$	Linearization distribution of a measurement model
$\boldsymbol{\Pi}$	Transition matrix of a hidden Markov model
$\Pi_{i,j}$	Element (i, j) of transition matrix $\boldsymbol{\Pi}$
ρ	Probability
σ^2	Variance
σ_i^2	Variance of noise component i
$\boldsymbol{\Sigma}$	Auxiliary matrix needed in the EM-algorithm
$\boldsymbol{\Sigma}_i$	Proposal distribution covariance in the Metropolis algorithm
τ	Time
φ	Direction angle
$\varphi_k(\boldsymbol{\theta})$	Energy function at time step k
$\Phi(\cdot)$	A function returning the lower triangular part of its argument or cumulative density of the standard Gaussian distribution
$\boldsymbol{\Phi}$	An auxiliary matrix needed in the EM-algorithm
ω	Angular velocity
$\boldsymbol{\Omega}_k$	Noise covariance in a Gaussian enabling approximation of a measurement model
$\boldsymbol{\Omega}_k$	Noise covariance at the ith iteration of a posterior linearization filter
$\mathbf{a}$	Action in decision theory or a part of a mean vector
$\mathbf{a}_{\mathrm{o}}$	Optimal action
$\mathbf{a}_k$	Offset in an affine dynamic model or constant term in a statistical linear regression approximation
$\mathbf{a}(\cdot)$	Non-linear drift function in a stochastic differential equation or acceleration
A	Resampling index
A_i	Resampling index

$\mathbf{A}$	Dynamic model matrix in a linear time-invariant model, the lower triangular Cholesky factor of a covariance matrix, the upper left block of a covariance matrix, a coefficient matrix in statistical linearization, or an arbitrary matrix
$\mathbf{A}^{(i)}$	Coefficient matrix at the ith iteration of posterior linearization
$\mathbf{A}_k$	Dynamic model matrix (i.e., transition matrix) of the jump from step k to step $k+1$, or approximate transition matrix in a statistical linear regression approximation
$\mathbf{A}_\mathbf{x}$	Jacobian matrix of $\mathbf{a}(\mathbf{x})$
b_k	Binary value in the Gilbert–Elliot channel model
$\mathbf{b}$	The lower part of a mean vector, the offset term in statistical linearization, or an arbitrary vector
$\mathbf{b}^{(i)}$	Offset term at the ith iteration of posterior linearization
$\mathbf{b}_k^{(i)}$	Offset term at ith iteration
$\mathbf{b}_k$	Dynamic bias vector or offset in an affine measurement model or constant term in a statistical linear regression approximation
$\mathrm{Be}(\cdot)$	Bernoulli distribution
$\mathbf{B}$	Lower right block of a covariance matrix, an auxiliary matrix needed in the EM-algorithm, or an arbitrary matrix
c	Scalar constant
c_k	Clutter (i.e., outlier) indicator
C	Arbitrary scalar constant
$C(\cdot)$	Cost or loss function
$\mathbf{C}$	The upper right block of a covariance matrix, an auxiliary matrix needed in the EM-algorithm, or an arbitrary matrix
$\mathbf{C}_k$	Cross-covariance matrix in a non-linear Kalman filter
$\mathbf{C}_\mathrm{G}$	Cross-covariance in the generalized statistical linear regression approximation
$\mathbf{C}_\mathrm{L}$	Cross-covariance in the linear (Taylor series-based) approximation
$\mathbf{C}_\mathrm{M}$	Cross-covariance in the Gaussian moment matching approximation
$\mathbf{C}_\mathrm{Q}$	Cross-covariance in the quadratic approximation
$\mathbf{C}_\mathrm{R}$	Cross-covariance in the statistical linear regression approximation
$\mathbf{C}_\mathrm{S}$	Cross-covariance in the statistical linearization approximation

$\mathbf{C}_{\mathrm{U}}$	Cross-covariance in the unscented approximation
d	Positive integer, usually dimensionality of the parameters
d_i	Order of a monomial
$\mathrm{d}t$	Differential of time variable t
$\mathrm{d}\mathbf{x}$	Differential of vector $\mathbf{x}$
$\mathbf{D}$	Derivative of the Cholesky factor, an auxiliary matrix needed in the EM-algorithm, or an arbitrary matrix
$\mathbf{D}_k$	Cross-covariance matrix in a non-linear RTS smoother or an auxiliary matrix used in derivations
$\mathbf{e}_i$	Unit vector in the direction of the coordinate axis i
$\tilde{\mathbf{e}}$	Noise term in statistical linear regression
$\tilde{\mathbf{e}}_k$	Noise term in an enabling Gaussian approximation of a dynamic model at time step k
$\mathbf{f}(\cdot)$	Dynamic transition function in a state space model
$F[\cdot]$	An auxiliary functional needed in the derivation of the EM-algorithm
$\mathbf{F}_{\mathbf{x}}(\cdot)$	Jacobian matrix of the function $\mathbf{x} \mapsto \mathbf{f}(\mathbf{x})$
$\mathbf{F}$	Feedback matrix of a continuous-time linear state space model
$\mathbf{F}_{\mathbf{xx}}^{(i)}(\cdot)$	Hessian matrix of $\mathbf{x} \mapsto f_i(\mathbf{x})$
g	Gravitation acceleration
$g(\cdot)$	An arbitrary function
$g_i(\cdot)$	An arbitrary function
$\mathbf{g}(\cdot)$	An arbitrary vector-valued function
$\mathbf{g}(t)$	Vector of forces
$\mathbf{g}^{-1}(\cdot)$	Inverse function of $\mathbf{g}(\cdot)$
$\tilde{\mathbf{g}}(\cdot)$	Augmented function with elements $(\mathbf{x}, \mathbf{g}(\cdot))$
$\mathbf{G}_k$	Gain matrix in an RTS smoother
$\mathbf{G}_{\mathbf{x}}(\cdot)$	Jacobian matrix of the function $\mathbf{x} \mapsto \mathbf{g}(\mathbf{x})$
$\mathbf{G}_{\mathbf{xx}}^{(i)}(\cdot)$	Hessian matrix of $\mathbf{x} \mapsto g_i(\mathbf{x})$
$\mathbf{h}(\cdot)$	Measurement model function in a state space model
$H_p(\cdot)$	pth order Hermite polynomial
$\mathbf{H}$	Measurement model matrix in a linear Gaussian model, or a Hessian matrix
$\mathbf{H}_k$	Measurement model matrix at time step k in a linear Gaussian or affine model, or approximate measurement model matrix in a statistical linear regression approximation
$\mathbf{H}_k^{(i)}$	Measurement model matrix at the ith iteration of posterior linearization filter

$\mathbf{H}_{\mathbf{x}}(\cdot)$	Jacobian matrix of the function $\mathbf{x} \mapsto \mathbf{h}(\mathbf{x})$
$\mathbf{H}^{(i)}_{\mathbf{xx}}(\cdot)$	Hessian matrix of $\mathbf{x} \mapsto h_i(\mathbf{x})$
i	Integer-valued index variable
$I_{\max}$	Number of iterations in iterated methods
$\mathbf{I}$	Identity matrix
$I_i(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)})$	An integral term needed in the EM-algorithm
j	Integer-valued index variable
$\mathbf{J}(\cdot)$	Jacobian matrix
k	Time step number
$\mathbf{K}^{(i)}$	Gain at iteration i of iterated posterior linearization
$\mathbf{K}_k$	Gain matrix of a Kalman/Gaussian filter
$\mathbf{K}^{(i)}_k$	Gain matrix at the ith iteration of an iterated filter at time step k
L	Positive constant
$L(\cdot)$	Negative logarithm of distribution
$L_{\mathrm{GN}}(\cdot)$	Gauss–Newton objective function
$\mathbf{L}$	Noise coefficient (i.e., dispersion) matrix of a continuous-time linear state space model
$\mathcal{L}(\cdot)$	Likelihood function
m	Dimensionality of a measurement, mean of the univariate Gaussian distribution, a mass, number of sigma points, or loop counter
$\mathbf{m}$	Mean of a Gaussian distribution
$\tilde{\mathbf{m}}$	Mean of an augmented random variable
$\mathbf{m}^{(i)}$	Mean at iteration i of iterated posterior linearization
$\mathbf{m}_k$	Mean of a Kalman/Gaussian filter at the time step k
$\mathbf{m}^{(i)}_k$	Mean at the ith iteration of a posterior linearization filter, mean of the Kalman filter in the particle i of RBPF at time step k
$\mathbf{m}^{(i)}_{0:T}$	History of means of the Kalman filter in the particle i of RBPF
$\tilde{\mathbf{m}}_k$	Augmented mean at time step k, an auxiliary variable used in derivations, or linearization point
$\mathbf{m}^{-}_k$	Predicted mean of a Kalman/Gaussian filter at time step k, just before the measurement $\mathbf{y}_k$
$\mathbf{m}^{-(i)}_k$	Predicted mean of the Kalman filter in the particle i of RBPF at time step k
$\tilde{\mathbf{m}}^{-}_k$	Augmented predicted mean at the time step k

$\mathbf{m}_k^{\mathrm{s}}$	Mean computed by a Gaussian (RTS) smoother for the time step k
$\mathbf{m}_{0:T}^{\mathrm{s}}$	Trajectory of the smoother means from a Gaussian (RTS) smoother
$\mathbf{m}_{0:T}^{\mathrm{s},(i)}$	Trajectory of means at smoother iteration i or history of means of the RTS smoother in the particle i of RBPS
$\mathbf{m}_\pi$	Expected value of $\mathbf{x} \sim \pi(\mathbf{x})$
$\mathbf{m}_k^{\mathbf{f}}$	Mean of $\pi_k^{\mathbf{f}}$
$\mathbf{m}_k^{\mathbf{h}}$	Mean of $\pi_k^{\mathbf{h}}$
M	Constant in rejection sampling
n	Positive integer, usually the dimensionality of the state
n'	Augmented state dimensionality in a non-linear transform
n''	Augmented state dimensionality in a non-linear transform
N	Positive integer, usually the number of Monte Carlo samples
$\mathrm{N}(\cdot)$	Gaussian distribution (i.e., normal distribution)
$O_{i,j}$	Element (i, j) of emission matrix $\mathbf{O}$
$\mathbf{O}$	Emission matrix of a hidden Markov model
p	Order of a Hermite polynomial
p_0	State-switching probability in the Gilbert–Elliot channel model
p_1	State-switching probability in the Gilbert–Elliot channel model
p_2	State-switching probability in the Gilbert–Elliot channel model
$p_{j,k}^-$	Predictive distribution for the discrete state $x_k = j$
$p_{j,k}$	Filtering distribution for the discrete state $x_k = j$
$p_{j,k}^{\mathrm{s}}$	Smoothing distribution for the discrete state $x_k = j$
P	Variance of the univariate Gaussian distribution
$\mathrm{Po}(\cdot)$	Poisson distribution
$\mathbf{P}$	Covariance of the Gaussian distribution
$\tilde{\mathbf{P}}$	Covariance of an augmented random variable
$\mathbf{P}^{\mathrm{xy},(i-1)}$	Predicted cross-covariance from iteration $i-1$ in iterated posterior linearization
$\mathbf{P}^{\mathrm{y},(i-1)}$	Predicted covariance from iteration $i-1$ in iterated posterior linearization
$\mathbf{P}^{(i)}$	Covariance at iteration i of iterated posterior linearization
$\mathbf{P}_k$	Covariance of a Kalman/Gaussian filter at time step k

$\mathbf{P}_k^{(i)}$	Covariance at iteration i of iterated posterior linearization filter, covariance of the Kalman filter in the particle i of RBPF at time step k
$\mathbf{P}_{0:T}^{(i)}$	History of covariances of the Kalman filter in the particle i of RBPF
$\tilde{\mathbf{P}}_k$	Augmented covariance at time step k or an auxiliary variable used in derivations
$\mathbf{P}_k^-$	Predicted covariance of a Kalman/Gaussian filter at the time step k just before the measurement $\mathbf{y}_k$
$\tilde{\mathbf{P}}_k^-$	Augmented predicted covariance at time step k
$\mathbf{P}_k^{-(i)}$	Predicted covariance of the Kalman filter in the particle i of RBPF at time step k
$\mathbf{P}_k^{\mathrm{s}}$	Covariance computed by a Gaussian (RTS) smoother for the time step k
$\mathbf{P}_{0:T}^{\mathrm{s}}$	Trajectory of smoother means from a Gaussian (RTS) smoother
$\mathbf{P}_{0:T}^{\mathrm{s},(i)}$	Trajectory of smoother covariances from iteration i of an iterated smoother or history of covariances of the RTS smoother in the particle i of RBPS
$\mathbf{P}_k^{\mathrm{x}}$	Predicted dynamic model covariance in a statistical linear regression filter
$\mathbf{P}_k^{\mathrm{xx}}$	Predicted dynamic model cross-covariance in a statistical linear regression filter
$\mathbf{P}_k^{\mathrm{xy}}$	Predicted measurement model cross-covariance in a statistical linear regression filter
$\mathbf{P}_k^{\mathrm{y}}$	Predicted measurement model covariance in a statistical linear regression filter
$\mathbf{P}_k^{\mathrm{x}}(\mathbf{x}_{k-1})$	Conditional covariance moment for a dynamic model
$\mathbf{P}_k^{\mathrm{y}}(\mathbf{x}_k)$	Conditional covariance moment for a measurement model
$\mathbf{P}_k^{x,(i)}$	Predicted dynamic model covariance in SPCMKF for sigma point i on time step k
$\mathbf{P}_k^{y,(i)}$	Predicted measurement model covariance in SPCMKF for sigma point i on time step k
$\mathbf{P}_k^{\mathrm{xy},(i-1)}$	Predicted cross-covariance from iteration $i-1$ in the iterated posterior linearization filter
$\mathbf{P}_k^{\mathrm{y},(i-1)}$	Predicted covariance from iteration $i-1$ in the iterated posterior linearization filter
$\mathbf{P}_\pi$	Covariance of $\mathbf{x} \sim \pi(\mathbf{x})$
$\mathbf{P}_k^{\mathbf{f}}$	Covariance of $\pi_k^{\mathbf{f}}$

$\mathbf{P}_k^{\mathbf{h}}$	Covariance of $\pi_k^{\mathbf{h}}$
q_0	Smaller probability of error in the Gilbert–Elliot channel model
q_1	Larger probability of error in the Gilbert–Elliot channel model
q^{c}	Spectral density of a white noise process
q_i^{c}	Spectral density of component i of a white noise process
$q(\cdot)$	Proposal distribution in the MCMC algorithm or an arbitrary distribution in the derivation of the EM-algorithm
$q^{(n)}$	Distribution approximation on the nth step of the EM-algorithm
$\mathbf{q}$	Gaussian random vector
$\mathbf{q}_k$	Gaussian process noise
$\tilde{\mathbf{q}}_k$	Euler–Maruyama approximation-based Gaussian process noise
Q	Variance of scalar process noise
$\mathcal{Q}_k^{(\cdot)}$	Sigma point of the process noise $\mathbf{q}_k$
$\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)})$	An auxiliary function needed in the EM-algorithm
$\mathbf{Q}$	Covariance of the process noise in a time-invariant model
$\mathbf{Q}_k$	Covariance of the process noise at the jump from step k to $k+1$
$\tilde{\mathbf{Q}}_k$	Euler–Maruyama approximation-based covariance of the process noise
$\mathbf{Q}^{\mathrm{c}}$	Spectral density matrix of (vector-valued) white noise
r	Distance to the center of rotation
r_k	Scalar Gaussian measurement noise
$\mathbf{r}_k$	Vector of Gaussian measurement noises
$\mathbf{r}_j(\cdot)$	Residual term in the Gauss–Newton objective function
R	Variance of scalar measurement noise
$\mathcal{R}_k^{(\cdot)}$	Sigma point of the measurement noise $\mathbf{r}_k$
$\mathbf{R}$	Covariance matrix of the measurement in a time-invariant model or the covariance-related parameter in Student's t-distribution
$\mathbf{R}_k$	Covariance matrix of the measurement at the time step k
$\mathbb{R}$	Space of real numbers
$\mathbb{R}^n$	n-dimensional space of real numbers
$\mathbb{R}^{n \times m}$	Space of real $n \times m$ matrices
s	Speed, generic integration variable, or temporary variable
s_k	Regime signal in the Gilbert–Elliot channel model

$s_{i,x}$	x-coordinate of radar i
$s_{i,y}$	y-coordinate of radar i
S	Number of backward-simulation draws
$\mathrm{St}(\cdot)$	Student's t-distribution
$\mathbf{S}^{(i)}$	Covariance at iteration i of iterated posterior linearization
$\mathbf{S}_k$	Innovation covariance of a Kalman/Gaussian filter at time step k
$\mathbf{S}_k^{(i)}$	Innovation covariance at the ith iteration of an iterated filter at time step k
$\mathbf{S}_{\mathrm{G}}$	Covariance in the generalized statistical linear regression
$\mathbf{S}_{\mathrm{L}}$	Covariance in the linear (Taylor series-based) approximation
$\mathbf{S}_{\mathrm{M}}$	Covariance in the Gaussian moment matching approximation
$\mathbf{S}_{\mathrm{Q}}$	Covariance in the quadratic approximation
$\mathbf{S}_{\mathrm{R}}$	Covariance in the statistical linear regression approximation
$\mathbf{S}_{\mathrm{R}}(\mathbf{x})$	Conditional covariance moment in statistical linear regression
$\mathbf{S}_{\mathrm{S}}$	Covariance in the statistical linearization approximation
$\mathbf{S}_{\mathrm{U}}$	Covariance in the unscented approximation
t	Time variable $t \in [0, \infty)$
t'	Another time variable $t' \in [0, \infty)$
$t^{(i)}$	Cumulative sum in resampling
t_k	Time of the step k (usually time of the measurement y_k)
T	Index of the last time step or the final time of a time interval
$\mathcal{T}_k$	Sufficient statistics
u	Scalar (random) variable
$\mathbf{u}_k$	Latent (non-linear) variable in a Rao–Blackwellized particle filter or smoother, or deterministic input to a dynamic model
$\mathbf{u}_k^{(i)}$	Latent variable value in particle i
$\mathbf{u}_{0:k}^{(i)}$	History of latent variable values in particle i
$U(\cdot)$	Utility function
$\mathrm{U}(\cdot)$	Uniform distribution
$\upsilon^{(i)}$	Random variable
υ_k	Bernoulli sequence in the Gilbert–Elliot channel model
$\upsilon_k^{(i)}$	Unnormalized weight in a SIR particle filter-based likelihood evaluation
$\mathbf{v}_k$	Innovation vector of a Kalman/Gaussian filter at time step k
$\tilde{\mathbf{v}}_k$	Noise term in an enabling Gaussian approximation of a measurement model on time step k
$\mathbf{v}_k^{(i)}$	Innovation vector at ith iteration of iterated extended Kalman filter at time step k

V	Volume of space
$V_k(x_k)$	Value function at time step k of the Viterbi algorithm
$\mathbb{V}$	Region in space (e.g., $\mathbb{V} = [-1, 1]$)
$w^{(i)}$	Normalized weight of the particle i in importance sampling
$\tilde{w}^{(i)}$	Weight of the particle i in importance sampling
$w^{*(i)}$	Unnormalized weight of the particle i in importance sampling
$w_k^{(i)}$	Normalized weight of the particle i at time step k of a particle filter
$w_{k\|n}^{(i)}$	Normalized weight of a particle smoother
w_i	Weight i in a regression model
$\mathbf{w}_k$	Vector of weights at time step k in a regression model
$\mathbf{w}(t)$	Gaussian white noise process
W	Weight in the cubature or unscented approximation
W_i	ith weight in a sigma point approximation
$W_i^{(\mathrm{m})}$	Mean weight of the unscented transform
$W_i^{(\mathrm{c})}$	Covariance weight of the unscented transform
$W_{i_1,\ldots,i_n}$	Weight in multivariate Gauss–Hermite cubature
x	Scalar random variable or state, sometimes regressor variable, or a generic scalar variable
$\mathbf{x}$	Random variable or state
$\hat{\mathbf{x}}$	Estimate of $\mathbf{x}$ or nominal $\mathbf{x}$
$\mathbf{x}^{(i)}$	ith Monte Carlo draw from the distribution of $\mathbf{x}$
$\mathbf{x}_k$	State at time step k
$\mathbf{x}_k^*$	Optimal state at time step k
$\mathbf{x}_k^{(i)}$	ith iterate of state estimate for time step k in iterated extended Kalman filter or smoother, or ith Monte Carlo sample of state in MCKF
$\hat{\mathbf{x}}_k^{(i)}$	Predicted ith Monte Carlo sample of the state in MCKF (before prediction)
$\mathbf{x}_k^{-(i)}$	Predicted ith Monte Carlo sample of the state in MCKF (after prediction)
$\mathbf{x}(t)$	State at (continuous) time t
$\tilde{\mathbf{x}}_k$	Augmented state at time step k
$\mathbf{x}_{0:k}$	Set containing the state vectors $\{\mathbf{x}_0, \ldots, \mathbf{x}_k\}$
$\mathbf{x}_{0:k}^{(i)}$	The history of the states in the particle i
$\mathbf{x}_{0:T}^*$	Optimal state trajectory