Statistics and Econometrics for Finance

Yong Zeng
Shu Wu *Eds.*

State-Space Models

Applications in Economics and Finance

Springer

Statistics and Econometrics for Finance

For further volumes:
http://www.springer.com/series/10377

Yong Zeng • Shu Wu
Editors

State-Space Models

Applications in Economics and Finance

Springer

Editors

Yong Zeng
Department of Mathematics and Statistics
University of Missouri at Kansas City
Kansas City, Missouri, USA

Shu Wu
Department of Economics
The University of Kansas
Lawrence, Kansas, USA

ISBN 978-1-4614-7788-4 ISBN 978-1-4614-7789-1 (eBook)
DOI 10.1007/978-1-4614-7789-1
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013943592

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Since the seminal papers of Kalman (1960, 1961) and the early development in the field of engineering, state-space models have become an increasingly important tool for research in finance and economics in recent years. This book is a collection of contributed papers that reflect this new phase of theoretical developments of state-space models and their applications in economics and finance. We hope that the breadth of research shared in this volume will serve as an inspiration and a valuable reference for future users of state-space models in these two areas.

A generic state-space model consists of two equations. One describes how observable economic variables relate to potentially unobservable *state variables*, and the other describes how state variables evolve over time. Both state and observable variables can be either discrete- or continuous-time stochastic processes.

This book is divided into four parts. In Parts I and II, we mainly consider discrete-time state-space models. Let Y_t be an $n \times 1$ observable variable and X_t be a $k \times 1$ state variable. Then, a state-space model can be written as follows:

$$Y_t = f(X_t, \theta, \varepsilon_t) \quad (1)$$

$$X_{t+1} = g(X_t, \theta, \eta_{t+1}) \quad (2)$$

where θ is an $m \times 1$ vector of model parameters, and ε_t and η_t are independent and identically distributed random shocks (or noises). $f(\cdot)$ and $g(\cdot)$ are $n \times 1$ and $k \times 1$ (non)linear functions, respectively. Because of its flexibility, many discrete-time models in economics and finance can be represented in this state-space form. These models include autoregressive moving average (ARMA) models, regression models with time-varying coefficients, dynamic factor models (DFM), models with stochastic volatility (SV), regime-switching models, and hidden Markov models (HMM). In these models, the parameter θ is typically unknown. Therefore, one central question in many applications is to estimate θ and to conduct statistical inferences of the model (e.g., hypothesis testing). Researchers are also highly interested in obtaining unbiased and efficient estimates of the underlying state variables X_t.

In Parts III and IV, we include models where the state variable, $X(t)$, can be a continuous-time finite-state Markov chain as that in hidden Markov or regime-switching models of Chaps. 8–11. $X(t)$ can also be a general continuous-time Markov process described by a stochastic differential equation as those in Chaps. 12–15. The Markov processes include a geometric Brownian motion (GBM) and a jump-diffusion process with regime switching among others. The observation process, $Y(t)$, can be a continuous-path process (Chaps. 10 and 11), an equally spaced time series (Chap. 14), or an irregularly spaced point process (Chaps. 8, 13, and 15). These models have found many applications in finance such as pricing of credit risk, optimal trading rules and hedging, optimal annuity purchasing or dividend policies, optimal learning in financial markets, and modeling (ultra) high-frequency data.

Below, we provide a more detailed description of each part.

Part I includes three chapters on *Particle Filtering and Parameter Learning in Nonlinear State-Space Models*. The introduction of particle filters has had a major impact on the development of nonlinear and non-Gaussian state-space models. This technique has expanded the range of practical applicability of state-space models to cases with high-dimension state space.

In Chap. 1, Tze Leung Lai and Vibhav Bukkapatanam provide a review of the estimation of the latent state variables using particle filters with known or unknown model parameters. They also present a new adaptive particle filter that uses a computationally efficient Markov Chain Monte Carlo estimate of the posterior distribution of the state-space model parameters in conjunction with sequential state estimation. They describe several applications in finance and economics, including frailty models of portfolio default probabilities, SV models with contemporaneous price and volatility jumps, and hidden Markov models for high-frequency transaction data.

In Chap. 2, Maria Paula Rios and Hedibert Freitas Lopes explore kernel smoothing and conditional sufficient statistics extensions of the auxiliary particle filters (Pitt and Shepard 1999) and bootstrap filters (Gordon, Salmond and Smith 1993). Using simulated data from SV models with Markov switching, they show that the Liu-West particle filter degenerates and has the largest Monte Carlo error, while their auxiliary particle filter extended with sufficient statistics (APF + SS) has a much better performance. Their APF + SS filter takes advantage of recursive sufficient statistics that are sequentially tracked and whose behavior resembles that of a latent state with conditionally deterministic updates. They also assess the performance of the APF + SS filter in sequential estimation in examples with real data.

In Chap. 3, Alexandre J. Chorin, Matthias Morzfeld, and Xuemin Tu review the implicit particle filter. The key idea is to concentrate the particles on the high-probability regions of the target probability density function (pdf) so that the number of particles required for a good approximation of the pdf remains manageable even if the state space has high dimensions. They explain how this idea is implemented, discuss special cases of practical importance, and show the relations of the implicit particle filter with other data assimilation methods. They further illustrate the method with examples such as SV, stochastic Lorenz attractor, stochastic

Kuramoto–Sivashinsky equation, and data assimilations.

Part II includes four chapters on the application of *Linear State-Space Models in Macroeconomics and Finance*.

In Chap. 4, Yulei Luo, Jun Nie, and Eric Young explicitly solve a linear-quadratic macroeconomic model under model uncertainty (due to concerns of model misspecification) and state uncertainty (due to limited information constraint). They show that the model can be mapped to a state-space representation that can be used to quantify the key parameters of model uncertainty. They demonstrate through examples how this framework can be used to study a range of interesting questions in macroeconomics and international finance such as explaining current account dynamics and resolving the international consumption puzzle.

In Chap. 5, Pym Manopimoke estimates a state-space model of the inflation dynamics in Hong Kong. The model allows her to decompose Hong Kong inflation into a stochastic trend and a stationary cycle component that can be driven by both domestic and foreign economic variables such as output gaps. This empirical model is consistent with economic theories of inflation and output, offering new insight into the determination of trend and cyclical inflation in Hong Kong. This is an example of the power of state-space models in empirical macroeconomic research.

In recent years, vector autoregression models (VARs) have become a primary tool for investigating dynamic relationship between multiple economic variables. One challenge, however, is that such relationships are often evolving over time as a result of shifts in government policies or structural changes in the economy. In Chap. 6, Taeyoung Doh and Michael Connolly show that the state-space representation is a useful tool to estimate VARs with time-varying coefficients and/or SV. They show that these models can better capture the changing relationships between important macroeconomic variables.

Chapter 7 is an application to finance. Jun Ma nd Mark Wohar use a state-space model to address one important issue regarding sources of stock market volatility. They argue that the existing empirical studies have focused on point estimation and lack robust statistical inference. The authors show that the small signal-to-noise ratio has made the market data contain too little useful information for researchers to reach robust conclusions about the relative importance of different sources of stock market volatility.

Part III includes five chapters on *Hidden Markov Models* (*HMM*), *Regime Switching*, and *Mathematical Finance*.

Chapters 8 and 9 are on hidden Markov models and their applications to finance. In Chap. 8, Robert Elliott and Tak Kuen Siu discuss an intensity-based model of portfolio credit risk using a collection of hidden Markov-modulated single jump processes. The model is a dynamic version of a frailty model casted in state-space form, able to describe dependent default risks among firms that are exposed to a common hidden dynamic frailty factor. The authors develop filtering equations and filter-based estimates of the model in recursive forms. They also obtain the joint default probability of reference entities in a credit portfolio as well as the variance dynamics for both observations and hidden states. In Chap. 9, Xiaojing Xi and Rogemar Mamon develop a weak hidden Markov model (WHMM) for the term structure of interest rates where the means and volatilities of bond yields are governed by a second-order Markov chain in discrete time. The authors use the multivariate filtering technique in conjunction with the EM algorithm to estimate the model parameters. They assess the goodness of fit of the model based on out-of-sample forecasts and apply AIC to determine the optimal number of regimes in their model. They apply the model to a data set of daily Treasury yields in the USA. The empirical results show that their WHMM outperforms the standard HMM in terms of out-of-sample forecasts.

Chapters 10 and 11 are applications of regime-switching models to insurance risk and optimal trading rule. Models with regime switching usually don't have analytical solutions to the associated stochastic control problems. In Chap. 10, Zhuo Jin and George Yin discuss numerical methods for solving stochastic optimization problems involving regime-switching models. They propose a numerical solution to the system of HJB equations based on Markov chain approximation. They show how these regime-switching models can be applied to analyze optimal annuity purchasing and optimal dividend payment strategy problems. In Chap. 11, Eunju Sohn and Qing Zhang study an optimal trading rule problem where the underlying asset price is governed by a mean-reverting process with regime switching. The investor's objective is to buy and sell the asset so as to maximize the overall return. The authors consider the case in which the jump rates of the Markov chain can go to infinite. They study the asymptotic properties of the limit value functions and establish a limiting problem which is easier to solve. They show that the solution to the limiting problem can be used to construct a trading rule that is nearly optimal.

In Chap. 12, Mingming Wang and Allanus Tsoi discuss Constant Proportion Portfolio Insurance (CPPI) problem with jump diffusion. They also consider the associated problem of hedging using both the PDE/PIDE and martingale approaches. In particular, they consider the mean-variance hedging problem when the contingent claim is a function of the CPPI portfolio value.

Part IV includes three chapters on *Nonlinear State-Space Models for High-Frequency Financial Data*.

In Chap. 13, combining classical Kyle and Glosten–Milgrom models, Yoonjung Lee proposes a new state-space modeling framework under asymmetric information. The model is able to describe the interactions among some important variables in financial markets such as the price impact of a trade, the duration between trades, and the degree of information asymmetry. In the model, a private signal is partially revealed through trades, while new public information arrives continuously at the market. In order to set a competitive price that rationally incorporates these two sources of information, the market maker utilizes Bayesian learning. The author derives the corresponding nonlinear filtering equation using anticipative Girsanov transformation. She further proves the existence and uniqueness of the ask and bid prices using an SPDE approach. The pricing rule depends on the actual sequence of order arrivals, not just the total number of buy/sell orders. The price impact of a trade tends to decrease when the duration between trades gets longer. The speed at which the information gets incorporated into the price depends on the quality of the private signal and the trading rate of informed traders.

Chapter 14 is concerned with volatility estimation and prediction. A popular approach is to use the high-frequency data to estimate volatilities and then fit a low-frequency AR volatility model for forecasting. While the empirical performance of this approach is good, there is a lack of theoretical foundation. In this chapter, Yazhen Wang and Xin Zhang show that, for rather general underlying price and volatility processes, the realized volatility estimators approximately follow a heterogeneous autoregressive model, hence providing theoretical justifications of the popular approach. An important feature of the model is that the two- or multi-scaled realized volatility estimators employed are based on a state-space model, where the prices from high-frequency transactions may include market microstructure noise.

Chapter 15 is concerned with estimating models for ultra-high frequency data. The class of models has a random-arrival-time state-space form that explicitly accommodates market microstructure noises in asset price. Although the model is able to capture stylized facts of tick data, the nonlinear state-space model structure makes parameter estimation a challenge. Cai Zhu and James Huang apply particle Markov Chain Monte Carlo (PMCMC) method to estimate a couple models when the underlying intrinsic value processes follow a GBM or a jump-diffusion process. They show that the PMCMC method is able to yield reasonable estimates of the model parameters and further discuss numeric methods that are able to enhance the efficiency of the algorithm.

We would like to express our gratitude to all the contributors of the book chapters for their efforts in making their research accessible to a wide range readers. We hope the book can lead to more interdisciplinary research among economists, mathematicians, and statisticians. We also would like to thank Yaozhong Hu of University of Kansas, Neng Wang of Columbia University, and Zhenxiao Wu of National University of Singapore for their generous help during the preparation of this book.

We want to thank Brian Foster, a former Springer editor, for his enthusiasm and help in the early stage of this book project. We also want to thank Hannah Bracken, Marc Strauss, Nicolas Philipson, and William Curtis of Springer New York. Their continuing support and commitment throughout the project are highly appreciated. We owe a special thanks to Hannah Bracken for her superb editorial assistance.

Yong Zeng's research is supported by National Science Foundation under the grants of TG-DMS100019 and DMS-1228244 and a grant from the University of Missouri Research Board in 2011. He gratefully acknowledges those grant supports. Finally, we would like to thank our families for their continuous support, which makes the completion of this project possible.

Kansas City, MO Yong Zeng
Lawrence, KS Shu Wu

Contents

List of Contributors

Vibhav Bukkapatanam
Department of Management Science and Engineering, Stanford University, Stanford, CA, USA, e-mail: vibhav@stanford.edu

Alexandre J. Chorin
Department of Mathematics, University of California Lawrence, Berkeley National Laboratory, Berkeley, CA, USA,
e-mail: chorin@math.berkeley.edu

Michael Connolly
Federal Reserve Bank of Kansas City, Kansas City, MO, USA,
e-mail: Michael.Connolly@kc.frb.org

Taeyoung Doh
Federal Reserve Bank of Kansas City, Kansas City, MO, USA,
e-mail: Taeyoung.Doh@kc.frb.org

Robert J. Elliott
Haskayne School of Business, University of Calgary, Calgary, AB, Canada
School of Mathematical Sciences, University of Adelaide, Adelaide, SA, Australia
Centre for Applied Financial Studies, University of South Australia, Adelaide, SA, Australia, e-mail: relliott@ucalgary.ca

Jian Hui Huang
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, e-mail: majhuang@inet.polyu.edu.hk

Zhou Jin
Center for Actuarial Studies, Department of Economics, The University of Melbourne, VIC, Australia, e-mail: zjin@unimelb.edu.au

Tze Leung Lai
Department of Statistics, Stanford University, Stanford, CA, USA,
e-mail: lait@stanford.edu

Yoonjung Lee
Boston, MA, USA, e-mail: yoonjung.lee.lin@gmail.com

Hedibert Freitas Lopes
Booth School of Business, University of Chicago, Chicago, IL, USA,
e-mail: hlopes@chicagobooth.edu

Yulei Luo
Department of Economics, The University of Hong Kong, Pokfulam, Hong Kong, China, e-mail: yluo@econ.hku.hk

Jun Ma
Department of Economics, Finance and Legal Studies, University of Alabama, Tuscaloosa, AL, USA, e-mail: jma@cba.ua.edu

Rogemar Mamon
Department of Statistics and Actuarial Sciences, University of Western Ontario, London, ON, Canada, e-mail: rmamon@stats.uwo.ca

Pym Manopimoke
Department of Economics, University of Kansas, Lawrence, KS, USA,
e-mail: pymm@ku.edu

Matthias Morzfeld
Lawrence Berkeley National Laboratory, Berkeley, CA, USA,
e-mail: mmo@math.lbl.gov

Jun Nie
Federal Reserve Bank of Kansas City, Kansas City, MO, USA,
e-mail: jun.nie@kc.frb.org

Maria Paula Rios
Booth School of Business, University of Chicago, Chicago, IL, USA,
e-mail: maria@chicagobooth.edu

Tak Kuen Siu
Cass Business School, City University London, London, UK,
e-mail: Ken.Siu.1@city.ac.uk

Eunju Sohn
Department of Science and Mathematics, Columbia College of Chicago, Chicago, USA, e-mail: esohn@colum.edu

Allanus Tsoi
Department of Mathematics, University of Missouri, Columbia, MO, USA,
e-mail: tsoia@missouri.edu

Xuemin Tu
Department of Mathematics, University of Kansas, Lawrence, KS, USA,
e-mail: xtu@math.ku.edu

Mingming Wang
School of Insurance, The University of International Business and Economics, Beijing, China, e-mail: wiryming@gmail.com

Yazhen Wang
Department of Statistics, University of Wisconsin-Madison, Madison, WI, USA, e-mail: yzwang@stat.wisc.edu

Xiaojing Xi
Department of Applied Mathematics, University of Western Ontario, London, ON, Canada, e-mail: xxi2@uwo.ca

Mark E. Wohar
Department of Economics, University of Nebraska at Omaha, Omaha, NE, USA, e-mail: mwohar@unomaha.edu

George Yin
Department of Mathematics, Wayne State University, Detroit, MI, USA, e-mail: gyin@math.wayne.edu

Eric R. Young
Department of Economics, University of Virginia, Charlottesville, VA, USA, e-mail: ey2d@virginia.edu

Qing Zhang
Department of Mathematics, University of Georgia, Athens, GA, USA, e-mail: qingz@math.uga.edu

Xin Zhang
Department of Statistics, University of Wisconsin-Madison, Madison, WI, USA, e-mail: zhangxin@stat.wisc.edu

Cai Zhu
Department of Finance, Business School, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, e-mail: czhuaa@ust.hk

Part I
Particle Filtering and Parameter Learning in Nonlinear State-Space Models

Chapter 1
Adaptive Filtering, Nonlinear State-Space Models, and Applications in Finance and Econometrics

Tze Leung Lai and Vibhav Bukkapatanam

1.1 Introduction

The Kalman filter, which is applicable to linear Gaussian models, and its modifications such as extended Kalman filters, Gaussian sum filters, and unscented Kalman filters for nonlinear state-space models are widely used in engineering. Without relying on local linearization techniques or functional approximations, particle filters are able to handle a large class of nonlinear non-Gaussian state-space models and have become increasingly popular in engineering applications in the past decade. Filtering in state-space models involves sequential computation of the posterior distribution of the latent state x_t given observations $y_1, \dots, y_t$. Smoothing involves the estimation of the hidden state x_t given observations $y_1, \dots, y_n$, with $1 \le t \le n$. More details are given in Sect. 1.2.

State-space models typically involve unknown parameters that have to be estimated from the data by either maximum likelihood or Bayesian methods. Replacing these unknown parameters in a particle filter by their sequential estimates leads to an adaptive particle filter; see Liu and West (2001) in [31], Storvik (2002) in [38], Carvalho et al. (2010) in [9], Polson et al. (2008) in [35], and Andrieu et al. (2010) in [2]. Section 1.3 reviews existing methods for sequential parameter estimation in state-space models. Section 1.4 describes a new adaptive filter that combines a novel Markov Chain Monte Carlo (MCMC) scheme for sequential parameter estimation with an efficient particle filter to estimate the state x_t. An important advantage of the new approach is that it yields a consistent estimate of the Monte Carlo standard error.

T.L. Lai (✉)
Department of Statistics, Stanford University, Stanford, CA, USA
e-mail: lait@stanford.edu

V. Bukkapatanam
Department of Management Science and Engineering, Stanford University, Stanford, CA, USA
e-mail: vibhav@stanford.edu

Y. Zeng and S. Wu (eds.), *State-Space Models: Applications in Economics and Finance*, Statistics and Econometrics for Finance 1, DOI 10.1007/978-1-4614-7789-1_1,

Particle filters have powerful and far-reaching applications in state-space models in finance and econometrics, some of which are described in Sect. 1.5. Section 1.6 gives some concluding remarks.

1.2 Particle Filters in Nonlinear State-Space Models

A general state-space model, also called a hidden Markov model (HMM), is defined by the evolution and observation density functions

$$\begin{aligned} x_t|x_{t-1} &\sim f(x_t|x_{t-1},\theta) \\ y_t|x_t &\sim g(y_t|x_t,\theta) \end{aligned} \tag{1.1}$$

with respect to measures μ and ν, respectively, where θ is a vector of parameters of the model. In the Bayesian formulation, the initial state x_0 has prior density $f(x_0|\theta)$, and θ has the prior density $\pi(\theta)$ with respect to some measure on the parameter space. In the HMM, x_t is the latent state and y_t is the observed data at time t. The filtering problem is to sequentially estimate the posterior distribution $p(x_t,\theta|\mathscr{Y}_t)$, where $\mathscr{Y}_t = (y_1, y_2, \ldots, y_t)$ is the set of observations up to time t. Particle filters approximate this posterior density by using a set of particles $(x_t,\theta)^{(i)}$ with weights $\widetilde{w}_t^{(i)}$ $(i = 1, 2, \ldots, N)$ summing to 1, so that

$$p^N(x_t,\theta|\mathscr{Y}_t)(\cdot) = \sum_{i=1}^{N} \widetilde{w}_t^{(i)} \delta_{(x_t,\theta)^{(i)}}(\cdot) \tag{1.2}$$

in which δ_z is the Dirac delta function with point mass at z. We now describe different approaches for sampling x_s $(s \le t)$ sequentially to form the particle filter in (1.2) when θ is fixed in advance. The weights $\widetilde{w}_t^{(i)}$ in (1.2) can be converted to $w_t^{(i)} = 1/N$ by a resampling step that samples with replacement N particles from $\{(x_t,\theta)^{(i)} : 1 \le i \le N\}$ with respective weights $\widetilde{w}_t^{(i)}$, as in the original proposal by Gordon et al. (1993) in [23] reviewed below. This is called *bootstrap resampling*. An alternative resampling scheme is introduced in Sect. 1.2.3. Resampling can help to mitigate the degeneracy of particles, which will be discussed in Sects. 1.3 and 1.4. Since resampling is optional and also can be used occasionally, we do not include it in the description of the basic algorithm in Sect. 1.2.1. Moreover, we assume the parameter vector to be known and therefore omit it in the rest of this section.

1.2.1 Bootstrap Filter

Originally proposed by Gordon et al. (1993) in [23], the bootstrap filter chooses samples from the prior distribution of the states. The Bayesian update equation for the posterior in a general filter can be written as

$$p(x_t|\mathscr{Y}_t) = \frac{g(y_t|x_t)p(x_t|\mathscr{Y}_{t-1})}{\int g(y_t|x_t')p(x_t'|\mathscr{Y}_{t-1})d\mu(x_t')} \tag{1.3}$$

where

$$p(x_t|\mathscr{Y}_{t-1}) = \int f(x_t|x_{t-1})p(x_{t-1}|\mathscr{Y}_{t-1})d\mu(x_{t-1}) \tag{1.4}$$

The bootstrap filter samples from the filtering distribution $p(x_t|\mathscr{Y}_{t-1})$ by first propagating the particles which approximate $p(x_{t-1}|\mathscr{Y}_{t-1})$ using the evolution density $f(x_t|x_{t-1})$, and reweighing the particles thus obtained by the likelihood ratio weights, as summarized in Algorithm 1.

Algorithm 1 : Bootstrap Filter

Initialize particles $\{x_0^{(i)}\}_{i=1}^N$ and the corresponding weights $\{\widetilde{w}_0^{(i)} = 1/N\}_{i=1}^N$
for t=1, 2, ..., T **do**
  for i=1, 2, ..., N **do**
    a) Propagate particle $x_{t-1}^{(i)}$ to $x_t^{(i)}$ using the evolution density $f(x_t|x_{t-1}^{(i)})$
    b) Update particle weights according to $\widetilde{w}_t^{(i)} \propto w_{t-1}^{(i)}\, g(y_t|x_t^{(i)})$
  end for
end for

1.2.2 Auxiliary Particle Filter

The auxiliary particle filter proposed by Pitt and Shephard (1999) in [33] generates samples from the filtering distribution with density function

$$p(x_t, x_{t-1}|\mathscr{Y}_t) \propto p(x_t|y_t, x_{t-1})p(y_t|x_{t-1})p(x_{t-1}|\mathscr{Y}_{t-1}) \tag{1.5}$$

Particles approximating $p(x_{t-1}|\mathscr{Y}_{t-1})$ are first resampled using weights proportional to the predictive density $p(y_t|x_{t-1})$, and the resampled particles are propagated forward using $p(x_t|y_t, x_{t-1})$. This is convenient only if $p(x_t|y_t, x_{t-1})$ is not difficult to sample from and $p(y_t|x_{t-1})$ is easily available, which is often not the case. Accordingly Pitt and Shephard in [33] suggest to replace $p(y_t|x_{t-1})$ by $p(y_t|\lambda(x_{t-1}))$, in which $\lambda(x_{t-1})$ is the mean, median, or mode of the distribution of x_t given x_{t-1}, and to propagate the resampled particles by sampling from the proposal density $f(x_t|x_{t-1})$, instead of directly from $p(x_t|x_{t-1}, y_t)$. The method is summarized in Algorithm 2.

Algorithm 2 : Auxiliary Particle Filter

Initialize particles $\{x_0^{(i)}\}_{i=1}^N$ and the corresponding weights $\{w_0^{(i)} = 1/N\}_{i=1}^N$
for t=1, 2, ..., T **do**
  for i=1, 2, ..., N **do**
    a) Resample particle $\tilde{x}_{t-1}^{(i)}$ from $\{x_{t-1}^{(1)}, \ldots, x_{t-1}^{(N)}\}$ using the weight $\tilde{w}_t^{(i)} \propto p(y_t|\lambda(x_{t-1}^{(i)}))$
    b) Propagate particle $\tilde{x}_{t-1}^{(i)}$ to $\tilde{x}_t^{(i)}$ using $f(\tilde{x}_t|\tilde{x}_{t-1}^{(i)})$
    c) Resample $x_t^{(i)}$ from $\{\tilde{x}_t^{(1)}, \ldots, \tilde{x}_t^{(N)}\}$ using the weight $w_t^{(i)} \propto g(y_t|\tilde{x}_t^{(i)})/g(y_t|\lambda(x_{t-1}^{(i)}))$
  end for
end for

1.2.3 *Residual Bernoulli Resampling*

Bootstrap resampling has been described in the paragraph preceding Sect. 1.2.1. In fact, the name *bootstrap filter* in Sect. 1.2.1 came from bootstrap resampling that Gordon et al. in [23] used to convert weighted particles to particles with equal weights. *Residual Bernoulli resampling* has been proposed as an alternative to bootstrap resampling and has been shown to often lead to smaller variance for the associated particle filter than the bootstrap resampling scheme. The method is summarized in Algorithm 3.

Algorithm 3 : Residual Resampling Scheme

Input: A set of particles $\{(\tilde{w}_t^{(i)}, \tilde{\mathscr{X}}_t^{(i)}), i = 1, 2, \ldots, M\}$
Output: A new set of particles $\{(\frac{1}{M}, \mathscr{X}_t^{(i)}), i = 1, 2, \ldots, M\}$
Set $R = \sum_{i=1}^M \lfloor M\tilde{w}_t^{(i)} \rfloor$
for i=1, 2, ..., M **do**
  – Set $\hat{w}_t^{(i)} = \frac{M\tilde{w}_t^{(i)} - \lfloor M\tilde{w}_t^{(i)} \rfloor}{M-R}$
end for
for j=1, 2, ..., M **do**
  – Sample $\hat{N}_j \sim \mathtt{mult}(M-R, \hat{w}_t^{(1)}, \hat{w}_t^{(2)}, \ldots, \hat{w}_t^{(M)})$
  – Set $N_j = \lfloor M\tilde{w}_n^{(j)} \rfloor + \hat{N}_j$
  – Set $\mathscr{X}_t^{(j)} = \tilde{\mathscr{X}}_t^{(N_j)}$
end for

1.3 Particle Filters with Sequential Parameter Estimation

An important problem which has been studied extensively in recent filtering literature is that of joint parameter estimation and filtering for general state-space models. Traditional methods which incorporate the parameters as part of the latent state vector suffer from severe degeneracy problems due to the absence of state evolution dynamics for the subvector of latent states representing parameters. Methods to address this issue have been considered in [2, 9, 31, 35, 38], and [34]. They are summarized below.

1.3.1 Liu and West's Filter

Liu and West (2001) in [31] suggest to use a kernel smoothing approximation to the posterior density $p(\theta|\mathscr{Y}_{t-1})$ of the unknown parameter θ via a mixture of multivariate normals and to combine it with an auxiliary particle filter described in Algorithm 2. Let $\{x_{t-1}^{(i)}, \theta_{t-1}^{(i)}\}_{i=1}^{N}$ be a set of particles with weights $w_{t-1}^{(i)}$ $(i = 1,\ldots,N)$, which approximate $p(x_{t-1}, \theta|\mathscr{Y}_{t-1})$. They approximate the posterior density for θ by

$$p(\theta|\mathscr{Y}_{t-1}) = \sum_{j=1}^{N} w_{t-1}^{(j)} N(\theta; m^{(j)}, \Sigma_{t-1})$$

where

$$m^{(j)} = a\theta_{t-1}^{(j)} + (1-a)\bar{\theta}$$

$$\bar{\theta} = \sum_{j=1}^{N} \frac{\theta_{t-1}^{(j)}}{N}$$

$$\Sigma_{t-1} = (1-a^2) \sum_{j=1}^{N} \frac{(\theta_{t-1}^{(j)} - \bar{\theta})(\theta_{t-1}^{(j)} - \bar{\theta})'}{N}$$

The constant a measures the extent of shrinkage of the individual $\theta_{t-1}^{(j)}$ to the overall mean $\bar{\theta}$. It is a tuning parameter whose choice is discussed in [31]. The mixture approximation generates new samples from the current posterior and attempts to avoid particle degeneracy. The method is summarized in Algorithm 4.

Algorithm 4 : Liu and West's Filter

Output: The filtering particles $\{(x_t^{(i)}, \theta_t^{(i)})\}_{i=1}^{N}$ and the parameter posterior $p(\theta|\mathscr{Y}_t)$, $t = 1,\ldots,T$
for t=1,2,...,T **do**
  for i=1,2,...,N **do**
    a) Resample $(\tilde{x}_{t-1}, \tilde{\theta}_{t-1})^{(i)}$ from $\{(x_{t-1}^{(j)}, \theta_{t-1}^{(j)})\}_{j=1}^{N}$ with weights $w_t^{(i)} \propto p(y_t|\lambda(x_{t-1}^{(i)}), m^{(i)})$
    b) Propagate $\tilde{\theta}_{t-1}^{(i)}$ to $\hat{\theta}_t^{(i)}$ using $N(\cdot; m^{(i)}, \Sigma_{t-1})$
    c) Propagate $\tilde{x}_{t-1}^{(i)}$ to $\hat{x}_t^{(i)}$ using $p(x_t|\tilde{x}_{t-1}^{(i)}, \tilde{\theta}_t^{(i)})$
    d) Resample $(x_t, \theta_t)^{(i)}$ from $\{(\hat{x}_t, \hat{\theta}_t)^{(j)}\}_{j=1}^{N}$ with weights $w_t^{(i)} \propto p(y_t|\hat{x}_t^{(i)}, \hat{\theta}_t^{(i)}) / p(y_t|\lambda(\tilde{x}_{t-1}^{(i)}), m^{(i)})$
  end for
end for

1.3.2 Storvik's Filter

Storvik (2002) in [38] considered sequential parameter estimation for a class of state-space models in which the posterior parameter density $p(\theta|\mathscr{X}_t, \mathscr{Y}_t)$ can be written as $p(\theta|s_t)$, where s_t is a low-dimensional set of sufficient statistics that can be recursively updated by $s_t = \mathscr{S}(s_{t-1}, x_t, y_t)$. The method is summarized in Algorithm 5.

Algorithm 5 : Storvik's Filter

Output: The filtering particles $\{(x_t^{(i)}, \theta_t^{(i)})\}_{i=1}^N$ and the parameter posterior $p(\theta|s_t)$, $t = 1, \ldots, T$
for t=1,2,...,T **do**
  for i=1,2,...,N **do**
    a) Propagate $x_{t-1}^{(i)}$ to $\tilde{x}_t^{(i)}$ using $q(x_t|\tilde{x}_{t-1}^{(j)}, \theta, \mathscr{Y}_t)$
    b) Resample $(x_t, s_{t-1})^{(i)}$ from $\{(\tilde{x}_t, s_{t-1})^{(j)}\}_{j=1}^N$ with weights $w_t^{(i)} \propto \frac{p(y_t|\tilde{x}_{t-1}^{(i)}, \theta) p(\tilde{x}_t^{(i)}|x_{t-1}^{(i)}, \theta)}{q(\tilde{x}_t^{(i)}|x_{t-1}^{(i)}, \theta, \mathscr{Y}_t)}$
    c) Compute sufficient statistics $s_t^{(i)} = \mathscr{S}(s_{t-1}^{(i)}, x_t^{(i)}, y_t)$
    d) Sample $\theta_t^{(t)}$ from $p(\theta|s_t^{(i)})$
  end for
end for

1.3.3 Particle Learning

Carvalho et al. (2010) in [9] propose the *particle learning* method which utilizes a resample-propagate scheme similar to auxiliary particle filters and show that the proposed method outperforms the filter of Liu and West of [31] in some comparative studies. Assuming the availability of conditional sufficient statistics s_t to represent the posterior of the parameter vector θ, and conditional sufficient statistics $s_{t,x}$ recursive state and parameter updates. The particles are now represented at each time by $z_t^{(i)}(x_t, s_t, s_{t,x}, \theta)^{(i)}$ and the Bayesian updating equation can be written as

$$p(z_t|\mathscr{Y}_t) = \int p(s_t|x_t, s_{t-1}, y_t) p(x_t|z_{t-1}, y_t) p(z_{t-1}|\mathscr{Y}_t) dx_t dz_{t-1} \tag{1.6}$$

where

$$p(z_{t-1}|\mathscr{Y}_t) \propto p(\mathscr{Y}_t|z_{t-1}) p(z_{t-1}|\mathscr{Y}_{t-1}) \tag{1.7}$$

Denoting the updating formulas for s_t and $s_{t,x}$ by $s_t = \mathscr{S}(s_{t-1}, x_t, y_t)$ and $s_{t,x} = \mathscr{K}(s_{t-1,x}, \theta, y_t)$, Algorithm 6 summarizes the resampling and propagation steps in their filter.

Algorithm 6 : Particle Learning

Output: The filtering density approximated by particles $z_t^{(i)}$ and the parameter posterior $p(\theta|s_t)$
for t=1,2,. . . ,T **do**
1) Resample: $\hat{z}_{t-1}^{(i)}$ from the particles $\{z_{t-1}^{(j)}\}_{j=1}^{N}$ using weights $w_t^{(i)} \propto p(y_t|z_{t-1}^{(i)})$
2) Propagate: $\hat{x}_{t-1}^{(i)}$ to $x_t^{(i)}$ using the distribution $p(x_t|\hat{z}_{t-1}^{(i)}, y_t)$
3) Propagate: Parameter sufficient statistics $s_t^{(i)} = \mathscr{S}(\hat{s}_{t-1}^{(i)}, x_t^{(i)}, y_t)$
4) Propagate: Sample $\theta^{(i)}$ from $p(\theta|s_t^{(i)})$
5) Propagate: State sufficient statistics $s_{t,x}^{(i)} = \mathscr{K}(\hat{s}_{t-1,x}^{(i)}, \theta^{(i)}, y_t)$
end for

1.3.4 Particle MCMC

Hybrid methods that combine particle filters with MCMC schemes have been considered in the literature. Important recent developments in this direction are [2, 35], and [34]. Polson et al. (2008) in [35] use a rolling window MCMC algorithm that approximates the target posterior distribution by a mixture of lag-k smoothing distributions. They recast the filtering problem as a sequence of smaller smoothing problems which can be solved using standard MCMC approaches as in [7] and [8]. They exploit, whenever possible, a sufficient statistic structure as in [19] and [38] to perform parameter updates and develop an algorithm with linear computational cost. Andrieu et al. (2010) in [2] introduce the particle MCMC (PMCMC) methods to perform inference on the unknown parameter vector θ. Pitt et al. (2012) in [34] provide further analytic results on PMCMC and show that using auxiliary particle filters in PMCMC schemes may help reduce computation time.

1.4 A New Approach to Adaptive Particle Filtering

In this section, we describe a new adaptive filtering technique, recently introduced in [11, 12], for joint parameter and latent state filtering in particle filters, which provides substantial improvement over previous approaches. The authors in [12] propose an efficient MCMC method to estimate the posterior distribution of the parameters, which can be used in conjunction with traditional particle filter methods that assume the parameters to be known. Bukkapatanam et al. (2012) in [5] provide further development of the methodology and its applications to economics and finance, which will be summarized in Sect. 1.5.

Chan and Lai (2012a) in [11] begin by considering the case where the parameter vector θ is known so that it can be omitted from the notation for particle filters, as in Sects. 1.2.1–1.2.3. They consider more generally the estimation of $\psi_T = \mathbb{E}[\psi(\mathscr{X}_T)|\mathscr{Y}_T]$ instead of $\mathbb{E}[\psi(x_T)|\mathscr{Y}_T]$, where $\mathscr{X}_T = (x_1, \ldots, x_T)$. When bootstrap resampling is performed at every stage, they show that the bootstrap filter estimate $\widehat{\psi}_T$ of $\psi_T = \mathbb{E}[\psi(\mathscr{X}_T)|\mathscr{Y}_T]$ has a martingale representation