

UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

Regime Switching Polynomial Diffusion Models and Topological Hidden Markov Model

Permalink

<https://escholarship.org/uc/item/6sc2k4z5>

ISBN

9798186380757

Author

Xu, Minxing

Publication Date

2026-06-04

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

Regime Switching Polynomial Diffusion Models and Topological Hidden Markov Model

A thesis submitted in partial satisfaction
of the requirements for the degree

Master of Arts
in
Statistics

by

Minxing Xu

Committee in charge:

Professor Gareth W. Peters, Chair
Professor Alexander Shkolnik
Professor Andrew Carter

June 2026

The thesis of Minxing Xu is approved.

Professor Alexander Shkolnik

Professor Andrew Carter

Professor Gareth W. Peters, Committee Chair

May 2026

Regime Switching Polynomial Diffusion Models and Topological Hidden Markov Model

Copyright © 2026

by

Minxing Xu

To My Family

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Gareth W. Peters, for his supervision, continuous support, guidance, and encouragement throughout the past several years. His high academic standards, broad vision in research, and constant motivation have profoundly influenced my development as a researcher. I am especially grateful for the time, patience, and effort he devoted to discussions, feedback, and collaborative work throughout this thesis and related research projects.

I would also like to sincerely thank my committee members, Professor Alexander Shkolnik and Professor Andrew Carter, for their valuable comments, support, and willingness to serve on my thesis committee.

Special thanks are extended to Professor Ian Duncan for his support, guidance, and encouragement throughout my academic journey, as well as for providing valuable opportunities related to research, industry exposure, and academic conferences.

I am also grateful to my collaborators, Professor Pavel V. Shevchenko, Zimo Zhu, and Peilun He, for many insightful discussions, collaborations, and shared research ideas throughout the development of this work.

Finally, and most importantly, I would like to express my greatest gratitude to my parents and family for their unconditional love, support, encouragement, and sacrifices throughout my education and life. Their belief in me has always been my strongest source of motivation.

Curriculum Vitæ

Minxing Xu

Education

- 2026 M.A. in Statistics (Expected), University of California, Santa Barbara.
- 2024 B.S. in Statistics and Data Science, University of California, Santa Barbara.

Publications

Peters, Gareth W; **Xu, Minxing**; Zhu, Zimo; Shevchenko, Pavel V. Regime-Switching Polynomial Diffusions via Topological Hidden Markov Model Inference using Onsager–Machlup Functionals for Asset Pricing. *SSRN Working Paper, 2026. Available at: <https://ssrn.com/abstract=6130426>*

Preprints

Peters, Gareth W; **Xu, Minxing**; Zhu, Zimo; He, Peilun. *Asynchronous Regime-Switching Multivariate CIR Spot-Rate Models via Onsager–Machlup Topological HMM Inference*. Preprint, 2026.

Abstract

Regime Switching Polynomial Diffusion Models and Topological Hidden Markov Model

by

Minxing Xu

This thesis studies regime-switching polynomial diffusion models and their statistical inference, and proposes a novel approach that incorporates forward-looking information via the Onsager–Machlup (OM) functional.

Classical likelihood-based inference relies on pointwise observations and does not incorporate forward-looking information. In this thesis, we propose a pathwise inference approach based on the OM functional, which provides a continuous-time analogue of likelihood for diffusion trajectories and enables most-probable trajectory estimation. As a result, the proposed framework allows for more timely detection of regime changes and yields more accurate reconstruction of the underlying diffusion process.

Building on this idea, we extend the Topological Hidden Markov Model (THMM) framework to regime-switching polynomial diffusion models. This formulation allows inference to remain computationally simple while preserving the use of standard Hidden Markov Model algorithms, including forward–backward filtering, Baum–Welch estimation, and Viterbi decoding, with suitable modifications.

The proposed framework is evaluated through simulation studies and empirical analyses of financial time series, including volatility process and spot-rate modeling applications, with emphasis on both mathematical rigor and economic interpretability.

Notation

- For all diffusion models, both univariate and multivariate state processes are denoted using non-bold notation. Bold symbols are used only for vector-valued quantities arising in the asynchronous regime-switching setting.
- Continuous-time quantities are indexed by t , while discrete-time sampled quantities are indexed by k .
- Unless otherwise specified, all polynomial diffusion models in this thesis are formulated under the physical probability measure $\mathbb{P}$. Exceptions include the affine interest-rate model section, which is formulated under the risk-neutral measure $\mathbb{Q}$ for arbitrage-free pricing, and the Heston option-pricing simulations, which are also performed under $\mathbb{Q}$.

Notation	Description
Ω	Sample space.
$\mathcal{F}$	Sigma-algebra on Ω .
$\{\mathcal{F}_t\}_{t \geq 0}$	Filtration satisfying the usual conditions.
$\mathcal{F}_t^X$	Filtration generated by the observed process.
$\mathbb{P}$	Physical probability measure.
$\mathbb{Q}$	Risk-neutral probability measure.
W_t	Standard Brownian motion.
X_t	State diffusion process.
$X_t^{(j)}$	j th latent factor component.
E	State space of the diffusion process.
$b(x)$	Drift coefficient.
$\sigma(x)$	Diffusion coefficient.
$\Sigma(x)$	Diffusion matrix in multivariate models.
$a(x) = \sigma(x)\sigma(x)^\top$	Local covariance matrix.
$\mathcal{A}$	Infinitesimal generator.
$\text{Pol}_m(E)$	Polynomial space of degree at most m .
S_t	Latent regime process.
$\mathcal{S} = \{1, \dots, L\}$	Finite regime state space.
$Q = (q_{ij})$	Generator matrix of the continuous-time Markov chain.
$\Pi = (\pi_{ij})$	Discrete-time transition matrix sampled at step Δ .
π_{ij}	One-step regime transition probability.
ν_i	Initial regime probability.
Δ	Observation step size.
X_k	Discretely sampled diffusion state X_{t_k} .
S_k	Discretely sampled regime state S_{t_k} .
$\mathcal{S}_{\text{joint}}$	Joint regime state space in the asynchronous model.
$p_i(x_k, x_{k+1})$	Regime-specific transition density.
$R_k(i)$	Euler residual under regime i .
$\Gamma(x)$	Local covariance matrix in multivariate diffusion.
$\rho = (\rho_{ij})$	Correlation matrix of Brownian motions.
$P(t, T)$	Zero-coupon bond price at time t with maturity T .
r_t	Instantaneous short rate.
$R_t^* = \int_0^t r_s ds$	Integrated short rate.

Table 0.1: Summary of notation used throughout the thesis.

Notation	Description
$C^2(E)$	Space of twice continuously differentiable functions on E .
$C_b^2(E)$	Space of bounded twice continuously differentiable functions with bounded derivatives.
$C_c^2(E)$	Space of twice continuously differentiable functions with compact support.
$\text{Pol}_m(E)$	Space of polynomials on E of degree at most m .
α_0, α_j	Affine short-rate coefficients.
$y_t(\tau)$	Zero-coupon yield with maturity τ .
$x_t^{(j)}$	Yield-spread factor.
$\tilde{x}_t^{(j)}$	Shifted positive yield-spread factor.
κ, θ, σ	CIR mean-reversion, long-run mean, and volatility parameters.
$\kappa_i, \theta_i, \sigma_i$	Regime-specific CIR parameters.
$\kappa_i^{(\eta)}, \theta_i^{(\eta)}, \sigma_i^{(\eta)}$	CIR parameters for latent factor η under regime i .
$P(\Delta) = e^{Q\Delta}$	Continuous-time Markov transition matrix.
$C_i(x_k, x_{k+1})$	Local Onsager–Machlup cost under regime i .
$w_i(x_k, x_{k+1})$	OM-based transition weight.
$\varphi(t)$	Candidate diffusion trajectory.
$\dot{\varphi}(t)$	Time derivative of the candidate trajectory.
$\mathcal{L}(\varphi, \dot{\varphi})$	Onsager–Machlup Lagrangian.
$I(u, v)$	Onsager–Machlup functional ratio.
$I[\varphi]$	Onsager–Machlup functional associated with trajectory φ .
$\delta_k(i)$	Viterbi dynamic programming score.
$\psi_k(i)$	Viterbi backpointer variable.
$\alpha_k(i)$	Forward probability.
$\beta_k(i)$	Backward probability.
$\gamma_k(i)$	Smoothed posterior regime probability.
$\xi_k(i, j)$	Posterior transition probability between regimes.
Θ	Collection of model parameters.
$\nabla^n f(x)$	n -th order derivative tensor of the function f at x .
$\partial_{x_i} f(x)$	Partial derivative of f with respect to x_i .

Table 0.2: Summary of notation used throughout the thesis.

Figures

Figure Label	Caption / Description
Figure 7.1	One-month expected-move band for the simulated asset price.
Figure 7.2	Standard HMM: Viterbi decoding with true parameters in the challenging simulation setting.
Figure 7.3	Challenging case under the Standard HMM: Baum–Welch parameter estimation.
Figure 7.4	Challenging case under the Standard HMM: Baum–Welch confidence intervals.
Figure 7.5	Challenging case under the OM-functional THMM: Viterbi decoding with true parameters.
Figure 7.6	Challenging case under the OM-functional THMM: Baum–Welch parameter estimation.
Figure 7.7	Challenging case under the OM-functional THMM: Baum–Welch confidence intervals.
Figure 7.8	Three-dimensional surface of daily US Treasury yields across maturities.
Figure 7.9	CIR-HMM regime decoding results for Treasury yield spreads. Blue and red curves denote filtered probabilities for the two latent regimes, while black markers indicate the Viterbi-decoded state sequence.
Figure 7.10	Five-day-ahead prediction intervals under the CIR-HMM model for Treasury yield spreads. The shaded grey region denotes the predictive confidence interval, while coloured segments indicate the decoded latent regime.

Tables

Table 0.3: Summary of tables in the simulation study section.

Table Label	Caption / Description
Table 7.1	Simulation parameter configuration for the two-regime Heston model. The parameter values are intentionally chosen to be relatively similar across regimes in order to create a difficult regime-classification setting and evaluate the inference under weak regime separation.

Contents

Curriculum Vitae	vi
Abstract	vii
1 Introduction	1
1.1 Introduction	1
2 Polynomial Diffusion Models	5
2.1 Probabilistic Setup	5
2.2 Polynomial Diffusions	6
2.3 Existence and Uniqueness	11
2.4 Properties of Polynomial Diffusions	14
2.5 Discretization	16
2.6 Examples of Polynomial Diffusions	19
3 Regime-Switching Polynomial Diffusion Models	23
3.1 Univariate Regime-Switching Model	23
3.2 Asynchronous Multivariate Regime-Switching Model	28
4 Financial Applications on Polynomial Diffusion model	34
4.1 Arbitrage-Free Pricing Framework	34
4.2 Affine Stochastic Volatility Models	35
4.3 Multivariate Spot-Rate Model	38
5 Hidden Markov Model and Inference	45
5.1 Problem Statements for HMM	45
5.2 Classical Hidden Markov Model	47
5.3 Parameter Estimation via Baum–Welch	47
5.4 Regime Decoding via Viterbi	49
5.5 Two-Stage Marginal Estimation and Covariance Reconstruction	50
5.6 Limitations of Classical HMM and Motivation for THMM	53

6	Topological Hidden Markov Model	54
6.1	Onsager–Machlup Functional	54
6.2	Topological Hidden Markov Models	61
6.3	THMM Inference	63
7	Simulation and Experiments	66
7.1	Simulation Studies THMM	66
7.2	Regime Estimation on Spot-Rate Modelling	77
8	Conclusion and Future Studies	82
A	Existence and Uniqueness of the CIR Process	85
A.1	Auxiliary Lemmas	85
A.2	Existence and Uniqueness	86

Chapter 1

Introduction

1.1 Introduction

Continuous-time stochastic models play a central role in modern financial mathematics, quantitative finance, and stochastic inference. Many important financial quantities, including equity prices, stochastic volatility, interest rates, credit spreads, and yield-curve factors, are naturally modeled through stochastic differential equations driven by diffusion processes Black and Scholes (1973); Merton (1973); Vasicek (1977); Cox et al. (1985). Diffusion-based models provide a mathematically tractable framework for describing randomness evolving continuously through time while preserving important probabilistic and analytical structures.

One of the major developments in continuous-time finance was the introduction of affine diffusion models and affine term-structure models Duffie et al. (2003); Filipović (2009); Dai and Singleton (2000); Filipović and Larsson (2016). Affine models are widely used because they admit tractable transform representations and often lead to closed-form pricing formulas for financial derivatives and fixed-income securities. Classical examples include the Vasicek model Vasicek (1977), the Cox–Ingersoll–Ross (CIR) model Cox et al.

(1985), and the Heston stochastic volatility model Heston (1993). These models capture important empirical features such as mean reversion, stochastic volatility, and positive interest-rate dynamics.

More recently, polynomial diffusion models have emerged as a broad generalization of affine processes Cuchiero et al. (2012); Filipović and Larsson (2016); Larsson (2017) . Polynomial diffusions preserve finite-dimensional polynomial spaces under the action of their infinitesimal generators, leading to tractable moment representations and efficient computational methods. This framework includes many classical financial models as special cases while allowing substantially greater modeling flexibility.

Despite their analytical tractability, classical diffusion models often fail to capture structural changes and regime-dependent behavior observed in empirical financial time series. Financial markets exhibit persistent periods of high and low volatility, abrupt transitions between economic environments, and changing dependence structures across time Hamilton (1989); Guidolin (2011); Ang and Bekaert (2002); Buffington and Elliott (2002); Kalimipalli and Susmel (2004) . Such phenomena motivate the introduction of regime-switching stochastic models, in which the underlying diffusion dynamics depend on an unobserved latent state process evolving according to a finite-state Markov chain Yin and Zhu (2010) .

Regime-switching diffusion models combine continuous stochastic evolution with discrete structural transitions. In these models, different regimes correspond to distinct economic or financial environments, such as recession and expansion periods, low- and high-volatility markets, or alternative monetary-policy regimes. Regime-switching extensions of stochastic volatility models and interest-rate models have proven particularly useful for capturing volatility clustering, yield-curve shifts, and changing market risk premia Guidolin (2011); Gray (1996) .

The introduction of latent regime processes naturally connects regime-switching diffu-

sions with Hidden Markov Models (HMMs) Baum and Petrie (1966); Baum et al. (1970); Elliott et al. (1995) . In the discrete-time setting, HMMs provide a probabilistic framework for modeling observations generated conditionally on an unobserved Markov chain. Classical HMM inference relies on efficient recursive algorithms such as the forward–backward algorithm, the Viterbi algorithm Viterbi (1967) , and the Baum–Welch expectation maximization procedure Baum et al. (1970) . These methods have been widely applied in signal processing, econometrics, finance, and machine learning.

When diffusion models are discretely observed, the regime-switching diffusion framework inherits a hidden Markov structure through the latent regime process. However, inference becomes significantly more challenging because transition densities of nonlinear diffusions are often unavailable in closed form or computationally expensive to evaluate. In high-dimensional systems, exact likelihood-based inference can become particularly difficult.

To address these challenges, pathwise approaches based on the Onsager–Machlup (OM) functional have attracted increasing attention Onsager and Machlup (1953); Machlup and Onsager (1953); Zeitouni and Dembo (1989) . The Onsager–Machlup functional may be interpreted as a continuous-time analogue of a negative log-likelihood for diffusion trajectories. It provides a probabilistic description of the most likely paths followed by stochastic systems and plays an important role in stochastic analysis, path-space inference, and filtering. The OM functional should be distinguished from Freidlin–Wentzell large-deviation theory Freidlin and Wentzell (1984): the latter controls exponential probabilities for small-noise diffusions on a large-deviation scale, whereas the OM functional characterises the relative likelihood of small tubes around smooth paths and is the appropriate tool for most-probable-path analysis.

Recent developments by Kashlak et al. (2023) introduced the Topological Hidden Markov Model (THMM), which extends classical hidden Markov methodology to infinite-

dimensional and topological settings by replacing classical emission densities with Onsager–Machlup-type pathwise weights. This construction allows hidden Markov inference to be performed directly on diffusion trajectories without requiring explicit transition densities. The resulting framework combines latent Markov structure with continuous-time pathwise information and is particularly attractive for nonlinear diffusion systems.

Motivated by these developments, this thesis studies regime-switching polynomial diffusion models and their associated hidden Markov representations. Particular emphasis is placed on stochastic volatility models, multivariate spot-rate dynamics, and asynchronous regime-switching structures arising in fixed-income markets. We develop a framework in which regime-dependent polynomial diffusions are coupled with hidden Markov inference and Onsager–Machlup pathwise representations.

The remainder of this thesis is organized as follows. Chapter 2 introduces polynomial diffusion theory and affine stochastic models. Chapter 3 develops regime-switching polynomial diffusion models and their hidden Markov representations. Chapter 4 studies stochastic volatility and spot-rate applications, including multivariate CIR yield-curve dynamics. Chapter 5 develops the classical Hidden Markov Model framework, including forward–backward recursions, Baum–Welch estimation, and Viterbi decoding for regime-switching diffusion models. Chapter 6 introduces Onsager–Machlup functionals and the Topological Hidden Markov Model framework together with the corresponding pathwise inference methodology. Finally, Chapter 7 discusses applications, numerical experiments, and future research directions.

Chapter 2

Polynomial Diffusion Models

2.1 Probabilistic Setup

We consider a filtered probability space

$$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}),$$

where Ω is the sample space, $\mathcal{F}$ is a σ -algebra, $\{\mathcal{F}_t\}_{t \geq 0}$ is a filtration satisfying the usual conditions, and $\mathbb{P}$ is a probability measure.

Let $W_t = (W_t^{(1)}, \dots, W_t^{(d)})^\top$ be a d -dimensional standard Brownian motion defined on this filtered probability space and adapted to $\{\mathcal{F}_t\}_{t \geq 0}$.

Let $d \in \mathbb{N}$ denote the dimension of the state space, and let $E \subseteq \mathbb{R}^d$ be a closed subset. In many applications, E is a basic closed semialgebraic set, which denote the state space of the process. In general, polynomial diffusions may be defined on a broad class of state spaces.

We assume that the solution process $\{X_t\}_{t \geq 0}$ takes values in E almost surely.

We consider an E -valued stochastic process $\{X_t\}_{t \geq 0}$, adapted to $\{\mathcal{F}_t\}$, satisfying the

stochastic differential equation

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t, \quad (2.1)$$

where $b : E \rightarrow \mathbb{R}^d$ and $\sigma : E \rightarrow \mathbb{R}^{d \times d}$ are measurable functions.

We assume that the solution remains in the state space E , that is, $X_t \in E$ almost surely for all $t \geq 0$. In particular, the state space E characterizes the support of the process (X_t) .

We define the diffusion matrix $a(x) := \sigma(x)\sigma(x)^\top \in \mathbb{S}_+^d$, where $\mathbb{S}_+^d$ denotes the set of $d \times d$ symmetric positive semidefinite matrices.

The dynamics (2.1) define a general class of diffusion processes. In the following, we restrict attention to a specific subclass, namely polynomial diffusions. Also, in this thesis, we primarily focus on polynomial diffusions defined on subsets of $\mathbb{R}_+^d$ motivated by applications in interest-rate and stochastic volatility modeling.

2.2 Polynomial Diffusions

Polynomial diffusions form an important subclass of diffusion processes whose infinitesimal generators preserve finite-dimensional polynomial spaces. This structural property provides substantial analytical tractability, particularly for moment calculations, transition representations, filtering procedures, and regime-switching extensions.

The theory of polynomial diffusions extends beyond standard Euclidean spaces and may be formulated on a broad class of admissible state spaces. Examples include Euclidean spaces, positive orthants, compact intervals, simplices, spheres, smooth manifolds, and spaces of positive semidefinite matrices. In many financial applications, the relevant state space is a subset of $\mathbb{R}_+^d$, reflecting positivity constraints associated with interest

rates, volatility processes, and covariance dynamics.

2.2.1 Infinitesimal Generator

Infinitesimal generator. Let $C^2(E)$ denote the space of twice continuously differentiable functions on E . For $f \in C^2(E)$, the infinitesimal generator $\mathcal{A}$ associated with the diffusion process $\{X_t\}_{t \geq 0}$ is defined by

$$\mathcal{A}f(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}[f(X_t) \mid X_0 = x] - f(x)}{t}, \quad (2.2)$$

whenever the limit exists.

Under suitable regularity conditions, the generator admits the representation

$$\begin{aligned} \mathcal{A}f(x) &= b(x)^\top \nabla f(x) + \frac{1}{2} \text{Tr} \left(a(x) \nabla^2 f(x) \right) \\ &= \sum_{i=1}^d b_i(x) \partial_{x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \partial_{x_i} \partial_{x_j} f(x). \end{aligned} \quad (2.3)$$

The infinitesimal generator plays a central role in the analysis of diffusion processes, as many probabilistic properties of the process may be characterized through the action of $\mathcal{A}$ on suitable function spaces.

2.2.2 Polynomial Spaces and Polynomial Generators

Definition 2.2.1 (Polynomial Spaces). *For each integer $m \geq 0$, let $\text{Pol}_m(\mathbb{R}^d)$ denote the vector space of real-valued polynomials on $\mathbb{R}^d$ of degree at most m . The restriction of these polynomials to the state space E defines $\text{Pol}_m(E)$. We further define*

$$\text{Pol}(E) := \bigcup_{m=0}^{\infty} \text{Pol}_m(E).$$

The spaces $\text{Pol}_m(E)$ are finite-dimensional vector spaces, which forms one of the key structural features underlying the tractability of polynomial diffusion models.

Definition 2.2.2 (Polynomial Generator Larsson (2017)). *The infinitesimal generator $\mathcal{A}$ is said to be polynomial on E if*

$$\mathcal{A}(\text{Pol}_m(E)) \subseteq \text{Pol}_m(E), \quad \forall m \geq 0. \quad (2.4)$$

The polynomial closure property implies that the action of the generator preserves finite-dimensional polynomial spaces. Consequently, conditional moments of the diffusion process can often be represented through finite-dimensional linear systems.

2.2.3 Polynomial Diffusions

Definition 2.2.3. *An E -valued diffusion process $\{X_t\}_{t \geq 0}$ is called a polynomial diffusion on E if its infinitesimal generator $\mathcal{A}$ satisfies $\mathcal{A}(\text{Pol}_m(E)) \subseteq \text{Pol}_m(E), \forall m \geq 0$.*

Polynomial coefficient characterization. A sufficient condition for the polynomial property is that the drift coefficients satisfy $b_i \in \text{Pol}_1(E)$, while the covariance coefficients satisfy $a_{ij} \in \text{Pol}_2(E)$, for all $i, j \in \{1, \dots, d\}$. Under these conditions, the generator maps polynomials to polynomials without increasing degree.

Remark 2.2.4. *The diffusion coefficient $\sigma(x)$ itself need not be polynomial. Instead, the polynomial property depends on the covariance structure $a(x) = \sigma(x)\sigma(x)^\top$.*

This distinction is particularly important for square-root diffusions such as the CIR process, where $\sigma(x) = \sigma\sqrt{x}$ is not polynomial, while the covariance coefficient $a(x) = \sigma^2 x$ remains polynomial.

Example 2.2.5 (Typical state spaces for polynomial diffusions). *Polynomial diffusions are frequently defined on constrained state spaces reflecting positivity, boundedness, or geometric structure. Typical examples include*

- Euclidean space $E = \mathbb{R}^d$,
- nonnegative orthant $E = \mathbb{R}_+^d$,
- unit cube $E = [0, 1]^d$,
- unit simplex $E = \Delta^{d-1} = \{x \in [0, 1]^d : x_1 + \cdots + x_d = 1\}$,
- unit sphere $E = \mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$,
- unit ball $E = B^d = \{x \in \mathbb{R}^d : |x| \leq 1\}$.

These state spaces arise naturally in applications. For example, nonnegative orthants appear in interest-rate and stochastic volatility models such as the CIR process, simplex-valued diffusions arise in probability and population models, while positive semidefinite matrix-valued diffusions appear in Wishart stochastic covariance models.

Definition 2.2.6 (Affine Polynomial Diffusion). *An E -valued diffusion process $\{X_t\}_{t \geq 0}$ is called an affine diffusion if its drift vector $b : E \rightarrow \mathbb{R}^d$ and covariance matrix $a : E \rightarrow \mathbb{S}_+^d$ are affine functions of the state variable, namely*

$$b(x) = b_0 + \sum_{i=1}^d x_i b_i, \quad a(x) = a_0 + \sum_{i=1}^d x_i a_i, \quad (2.5)$$

where $b_0, \dots, b_d \in \mathbb{R}^d$, $a_0, \dots, a_d \in \mathbb{S}_+^d$. Equivalently, the infinitesimal generator satisfies $\mathcal{APol}_1(E) \subseteq \text{Pol}_1(E)$.

Remark 2.2.7. *Every affine diffusion is a polynomial diffusion, since affine coefficients imply that the infinitesimal generator maps polynomials of degree m into polynomials of*

degree at most m . However, the converse is generally false. Polynomial diffusions allow covariance coefficients of degree up to two, whereas affine diffusions restrict both drift and covariance coefficients to be affine functions of the state variable. Their importance in finance comes from the exponential-affine transform property, which yields tractable closed-form or semi-closed-form expressions for bond prices, characteristic functions, and derivative valuation formulas.

2.2.4 Martingale Problem Characterization

Polynomial diffusions may be characterized through the martingale problem associated with their infinitesimal generators. This formulation is useful because it describes the process through its generator rather than through a particular stochastic differential equation.

Definition 2.2.8 (Martingale Problem). *Let $E \subseteq \mathbb{R}^d$ be a closed state space, let $\mathcal{A}$ be a linear operator with domain $\mathcal{D}(\mathcal{A})$, and let ϖ be a probability measure on E . The martingale problem $\text{MP}(E, \mathcal{A}, \varpi)$ consists of finding an E -valued process $\{X_t\}_{t \geq 0}$ such that $X_0 \sim \varpi$, and for every $f \in \mathcal{D}(\mathcal{A})$, the process*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s) ds \quad (2.6)$$

is a local martingale.

The martingale problem formulation provides a probabilistic characterization of diffusion processes through their infinitesimal generators. In the context of polynomial diffusions, this is especially useful because the generator acts invariantly on polynomial spaces, yielding tractable moment dynamics and finite-dimensional representations.

Remark 2.2.9. *Existence of a solution to $\text{MP}(E, \mathcal{A}, \varpi)$ provides weak existence of the*

associated process. If the solution law is unique for every initial distribution ϖ , then the martingale problem is said to be well posed, and this corresponds to uniqueness in law.

2.3 Existence and Uniqueness

The polynomial property describes the algebraic action of the infinitesimal generator on polynomial spaces. However, it does not by itself guarantee that a stochastic process with the prescribed generator exists, nor that such a process is unique. Existence and uniqueness therefore require additional analytical conditions.

Although classical SDE theory provides strong existence and pathwise uniqueness under Lipschitz-type conditions, such assumptions are often too restrictive for polynomial diffusions arising in financial applications. For example, square-root diffusions such as the CIR process fail to satisfy global Lipschitz conditions near the boundary. Moreover, the preservation of constrained state spaces is not automatic in the standard SDE framework.

2.3.1 Weak Existence

Definition 2.3.1 (Positive Maximum PrincipleLarsson (2017)). *The operator $\mathcal{A}$ is said to satisfy the positive maximum principle on E , denoted by $\text{PMP}(E)$, if for every $f \in C_b^2(\mathbb{R}^d)$ and every $x \in E$ such that*

$$f(x) = \max_{y \in E} f(y) \geq 0,$$

one has $\mathcal{A}f(x) \leq 0$.

The positive maximum principle is a necessary compatibility condition between the generator and the state space. If an E -valued solution to the martingale problem exists for every initial point, then the generator must satisfy $\text{PMP}(E)$. Conversely, together with a

non-explosion approximation condition, the positive maximum principle is sufficient for weak existence. We state the result in the pure diffusion setting used in this thesis.

Theorem 2.3.2 (Weak Existence via the Martingale Problem Larsson (2017)). *Suppose that the operator $\mathcal{A}$ satisfies the following conditions:*

1. $\mathcal{A}$ satisfies the positive maximum principle on E ;
2. for every $f \in C_c^2(\mathbb{R}^d)$, the function $\mathcal{A}f$ is continuous on E and vanishes at infinity;
3. There exists a sequence $(f_n)_{n \geq 1} \subset C_c^2(\mathbb{R}^d)$ such that $\sup_{y \in E} (|f_n(y)| + |\mathcal{A}f_n(y)|) < \infty$ and $f_n(x) \rightarrow 1$, $\mathcal{A}f_n(x) \rightarrow 0$, $x \in E$.

Then, for every probability measure ϖ on E , there exists an E -valued process $\{X_t\}_{t \geq 0}$ with $X_0 \sim \varpi$ that solves the martingale problem $\text{MP}(E, \mathcal{A}, \varpi)$.

The theorem establishes weak existence because the martingale problem does not require a fixed probability space or a prescribed Brownian motion. Instead, it guarantees the existence of a probability law under which the coordinate process has generator $\mathcal{A}$ and remains in the state space E .

Remark 2.3.3. *Condition (3) appearing in Theorem 2.3.2 is mild and is satisfied in many polynomial diffusion settings. In particular, if the state space E is compact, the condition holds automatically by choosing functions f_n that equal one on a neighborhood of E .*

For pure diffusion models, a sufficient growth condition is

$$|b(x)| + \|a(x)\|^{1/2} \leq C(1 + |x|), \quad x \in E, \quad (2.7)$$

for some constant $C > 0$.

Many classical polynomial diffusions, including Ornstein–Uhlenbeck and Cox–Ingersoll–Ross models, satisfy such growth conditions together with the positive maximum principle. Consequently, the martingale problem framework provides a convenient existence theory for the models considered in this thesis.

2.3.2 Uniqueness

Suppose that weak existence holds for $\text{MP}(E, \mathcal{A}, \varpi)$; see Theorem 2.3.2. Our goal is to establish uniqueness in law for the martingale problem. More precisely, we aim to show that if X and $\tilde{X}$ are two solutions of $\text{MP}(E, \mathcal{A}, \varpi)$, then $\mathcal{L}(X) = \mathcal{L}(\tilde{X})$. Equivalently, the martingale problem $\text{MP}(E, \mathcal{A}, \varpi)$ is unique in law.

Theorem 2.3.4 (Uniqueness in law for polynomial diffusions Larsson (2017)). *Let $E \subseteq \mathbb{R}^d$ be a closed state space, and let $\mathcal{A}$ be the generator of a polynomial diffusion on E . Fix an initial law ϖ on E . Suppose that the following conditions hold:*

1. *Weak existence holds for $\text{MP}(E, \mathcal{A}, \varpi)$; see Theorem 2.3.2.*
2. *The generator preserves polynomial spaces, that is, $\mathcal{APol}_m(E) \subseteq \text{Pol}_m(E) \forall m \in \mathbb{N}_0$.*
3. *The joint moment problem for the finite-dimensional distributions is determinate.*

Then $\text{MP}(E, \mathcal{A}, \varpi)$ is unique in law.

2.3.3 Application to Polynomial Diffusions

For the models in Example 2.2.5 on constrained domains, existence and uniqueness in law follow once the generator satisfies the PMP together with the regularity and boundary compatibility conditions of Theorem 2.3.2 and Theorem 2.3.4. The diffusion coefficient must vanish sufficiently rapidly at the boundary, while the drift remains inward-pointing to keep the process inside the admissible state space.

As a representative example, consider the CIR process Cox et al. (1985)

$$dX_t = \kappa(\theta - X_t) dt + \sigma\sqrt{X_t} dW_t, \quad X_t \in \mathbb{R}_+. \quad (2.8)$$

The polynomial property follows immediately from $b(x) = \kappa(\theta - x) \in \text{Pol}_1(\mathbb{R}_+)$ and $a(x) = \sigma^2 x \in \text{Pol}_1 \subset \text{Pol}_2(\mathbb{R}_+)$. The PMP holds on $\mathbb{R}_+$ because $a(0) = 0$ and $b(0) = \kappa\theta \geq 0$, so the drift points inward at the boundary. The linear growth condition of Remark 2.3.3 is also satisfied, so the assumptions of Theorem 2.3.2 and Theorem 2.3.4 hold regardless of whether the Feller condition is satisfied. Hence $\text{MP}(\mathbb{R}_+, \mathcal{A}, \varpi)$ admits a unique solution in law. A complete proof, including verification of all conditions and a discussion of boundary behavior under the Feller condition, is given in Appendix A.

2.4 Properties of Polynomial Diffusions

Polynomial diffusions possess several structural properties that make them analytically and computationally tractable. The key feature underlying these properties is the polynomial closure of the infinitesimal generator. This allows moment dynamics, conditional expectations, and transition representations to be characterized through finite-dimensional linear algebra.

2.4.1 Polynomial Closure

By Definition 2.2.3, the infinitesimal generator preserves polynomial spaces, the action of the generator on each space $\text{Pol}_m(E)$ remains finite-dimensional. Consequently, if $f \in \text{Pol}_m(E)$, then $\mathcal{A}f \in \text{Pol}_m(E)$. This invariance property is the main source of tractability for polynomial diffusion models. Unlike general diffusion processes, where the generator may produce highly nonlinear functions, polynomial diffusions preserve finite-dimensional

polynomial structures.

2.4.2 Moment Tractability

An important consequence of the polynomial closure property is the tractability of conditional moments.

If $f \in \text{Pol}_m(E)$, then the conditional expectation $\mathbb{E}[f(X_t) \mid X_0 = x]$ remains an element of $\text{Pol}_m(E)$ for all $t \geq 0$. Consequently, conditional moments of polynomial functions remain polynomial functions of the current state.

In particular, for monomials of the form $f(x) = x^\alpha$ with $\alpha \in \mathbb{N}_0^d$, the conditional moments $\mathbb{E}[X_t^\alpha \mid X_0 = x]$ can be computed within a finite-dimensional polynomial space.

This finite-dimensional structure is one of the principal sources of analytical tractability for polynomial diffusion models. In particular, moments of asset prices, interest rates, and stochastic volatility processes can often be represented through closed-form expressions or finite-dimensional linear systems.

2.4.3 Finite-Dimensional Polynomial Generator

The polynomial closure property implies that the infinitesimal generator admits a finite-dimensional matrix representation on each polynomial space.

Let

$$H_m(x) = (h_1(x), \dots, h_N(x))^\top$$

be a basis vector for $\text{Pol}_m(E)$, where $N = \dim(\text{Pol}_m(E)) < \infty$.

Since $\mathcal{A}h_j \in \text{Pol}_m(E)$ for each basis function h_j , there exists a matrix $Q_m \in \mathbb{R}^{N \times N}$ such that

$$\mathcal{A}H_m(x) = Q_m H_m(x). \tag{2.9}$$

The matrix Q_m is called the generator matrix associated with $\mathcal{A}$ on $\text{Pol}_m(E)$.

This representation converts the infinite-dimensional action of the generator into a finite-dimensional linear algebra problem. As a result, many quantities of interest can be computed through matrix exponentiation.

The generator matrix representation yields explicit formulas for conditional moments. Using standard results from the Kolmogorov backward equation, by the representation (2.9), one obtains

$$\mathbb{E}[H_m(X_t) \mid X_0 = x] = e^{tQ_m} H_m(x). \quad (2.10)$$

Hence, conditional moments of polynomial functions can be computed via finite-dimensional matrix exponentials.

This property is one of the major advantages of polynomial diffusion models. It allows moments and expectations to be evaluated without directly solving partial differential equations or performing large-scale Monte Carlo simulation.

2.5 Discretization

In practice, the continuous-time diffusion process is observed only at discrete time points. We therefore introduce a discrete observation scheme together with the associated transition structure and Euler approximation.

Discrete observation grid. Let $0 = t_0 < t_1 < \dots < t_N = T$ be an observation grid over the time interval $[0, T]$. For simplicity, we consider a regular discretization $t_k = k\Delta, k = 0, \dots, N$, where

$$\Delta = \frac{T}{N} > 0$$

denotes the observation step size. We define the sampled processes $X_k := X_{t_k}, S_k := S_{t_k}$.

Euler approximation. Consider the diffusion process

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t. \quad (2.11)$$

The Euler–Maruyama approximation over the interval $[t_k, t_{k+1}]$ is given by

$$X_{k+1} = X_k + b(X_k)\Delta + \sigma(X_k)\sqrt{\Delta} Z_k, \quad (2.12)$$

where $Z_k \sim N(0, I_d)$ are independent standard Gaussian random vectors. Equivalently,

$$X_{k+1} - X_k \approx b(X_k)\Delta + \sigma(X_k)\Delta W_k,$$

where $\Delta W_k = W_{t_{k+1}} - W_{t_k} \sim N(0, \Delta I_d)$.

The Euler scheme is obtained by freezing the drift and diffusion coefficients locally over the interval $[t_k, t_{k+1}]$.

Strong convergence Under suitable regularity assumptions including existence of a unique strong solution together with linear growth conditions, the Euler–Maruyama approximation converges to the true diffusion process.

For discrete grid, we denote the sampled exact process on the discrete observation grid. We denote by X_k^Δ the Euler approximation constructed with time step Δ .

Strong convergence concerns pathwise approximation accuracy between the numerical approximation and the sampled exact process. The Euler scheme is said to converge strongly with order $\gamma_s > 0$ if

$$\left(\mathbb{E} \left[\max_{0 \leq k \leq N} \|X_k^\Delta - X_k\|^2 \right] \right)^{1/2} = O(\Delta^{\gamma_s}), \quad \Delta \rightarrow 0. \quad (2.13)$$

For the Euler–Maruyama scheme, $\gamma_s = \frac{1}{2}$.

Strong convergence is particularly important for trajectory-dependent quantities and pathwise inference procedures. In the present work, this is relevant for Onsager–Machlup functionals and Topological Hidden Markov Models, where local path behavior and trajectory reconstruction play a central role.

Weak convergence Weak convergence concerns approximation of distributions and expectations. The Euler scheme is said to converge weakly with order $\gamma_w > 0$ if, for sufficiently smooth test functions f ,

$$|\mathbb{E}[f(X_N^\Delta)] - \mathbb{E}[f(X_N)]| = O(\Delta^{\gamma_w}), \quad \Delta \rightarrow 0. \quad (2.14)$$

For the Euler–Maruyama scheme, $\gamma_w = 1$.

Weak convergence is primarily relevant for likelihood approximation, moment estimation, and transition-density approximation, since these depend mainly on the distributional properties of the discretized process rather than exact sample paths.

Transition density approximation. Under the Euler approximation, the conditional distribution of X_{k+1} given $X_k = x_k$ is approximately Gaussian:

$$X_{k+1} \mid X_k = x_k \sim N(x_k + b(x_k)\Delta, a(x_k)\Delta). \quad (2.15)$$

Accordingly, the one-step transition density is approximated by

$$p^\Delta(x_k, x_{k+1}) = \frac{1}{(2\pi\Delta)^{d/2} \det(a(x_k))^{1/2}} \quad (2.16)$$

$$\times \exp\left(-\frac{1}{2\Delta} R_k(i)^\top a(x_k)^{-1} R_k(i)\right), \quad (2.17)$$

where $R_k(i) = x_{k+1} - x_k - b(x_k)\Delta$ denotes the local Euler residual.

When Δ is fixed, we also write for simplicity

$$p(x_k, x_{k+1}) := p^\Delta(x_k, x_{k+1}).$$

Remark 2.5.1. *The Euler–Maruyama discretization is a first-order numerical scheme that is widely used due to its simplicity and the fact that the resulting conditional transition distribution is Gaussian under the discretized dynamics. However, higher-order discretization schemes, such as the Milstein scheme Kloeden and Platen (1992), may provide improved strong convergence rates and reduced discretization bias.*

Remark 2.5.2. *As $\Delta \rightarrow 0$, the Euler approximation converges to the true diffusion process in both the weak and strong senses under suitable regularity conditions. Consequently, smaller mesh sizes Δ generally reduce discretization error, although at the expense of increased computational cost; see Kloeden and Platen (1992).*

2.6 Examples of Polynomial Diffusions

In this section, we present several classical examples of polynomial diffusions. These examples illustrate the general structure introduced in the previous sections and demonstrate that many widely used stochastic models in finance fall within the polynomial diffusion framework.

2.6.1 Ornstein–Uhlenbeck Process

The Ornstein–Uhlenbeck (OU) process is defined on the state space $E = \mathbb{R}$ by the SDE

$$dX_t = \kappa(\theta - X_t) dt + \sigma dW_t, \tag{2.18}$$

where $\kappa > 0$ is the mean-reversion rate, $\theta \in \mathbb{R}$ is the long-term mean, and $\sigma > 0$ is the volatility parameter. The drift and diffusion coefficients are given by $b(x) = \kappa(\theta - x)$, $a(x) = \sigma^2$. Hence, $b \in \text{Pol}_1(E)$ and $a \in \text{Pol}_0(E)$, implying that the OU process is a polynomial diffusion.

The infinitesimal generator is

$$\mathcal{A}f(x) = \kappa(\theta - x)f'(x) + \frac{1}{2}\sigma^2 f''(x),$$

which maps polynomials to polynomials of at most the same degree.

The OU process is Gaussian, mean-reverting, and admits a closed-form solution. It serves as a fundamental building block for more complex polynomial diffusion models.

2.6.2 Geometric Brownian Motion

The geometric Brownian motion (GBM) is defined on the state space $E = \mathbb{R}_+$ by the SDE

$$dX_t = bX_t dt + \sigma X_t dW_t, \tag{2.19}$$

where $b \in \mathbb{R}$ and $\sigma > 0$. The coefficients are $b(x) = bx$, $a(x) = \sigma^2 x^2$, so that $b \in \text{Pol}_1(E)$ and $a \in \text{Pol}_2(E)$, and hence GBM is a polynomial diffusion.

The generator is

$$\mathcal{A}f(x) = bx f'(x) + \frac{1}{2}\sigma^2 x^2 f''(x).$$

GBM is widely used in financial modeling, particularly in the Black–Scholes framework for asset prices. Unlike the OU process, it is not mean-reverting and exhibits multiplicative noise.

2.6.3 Cox–Ingersoll–Ross Process

The Cox–Ingersoll–Ross (CIR) process is defined on the state space $E = \mathbb{R}_+$ by the SDE

$$dX_t = \kappa(\theta - X_t) dt + \sigma\sqrt{X_t} dW_t, \quad (2.20)$$

with parameters $\kappa > 0$, $\theta > 0$, and $\sigma > 0$. The drift and diffusion coefficients satisfy $b(x) = \kappa(\theta - x)$, $a(x) = \sigma^2 x$. Thus, $b \in \text{Pol}_1(E)$ and $a \in \text{Pol}_1(E)$, implying that the CIR process belongs to the class of polynomial diffusions, despite the non-polynomial form of $\sigma(x)$.

The generator is given by

$$\mathcal{A}f(x) = \kappa(\theta - x)f'(x) + \frac{1}{2}\sigma^2 x f''(x).$$

The CIR process is mean-reverting and strictly non-negative under the Feller condition

$$2\kappa\theta \geq \sigma^2.$$

It is widely used as a short-rate model in interest rate theory.

2.6.4 Heston Stochastic Volatility Model

The Heston stochastic volatility model is defined on the state space $E = \mathbb{R}_+ \times \mathbb{R}_+$ by the system of SDE

$$\begin{aligned} dP_t &= \mu P_t dt + \sqrt{V_t} P_t dW_t^{(1)}, \\ dV_t &= \kappa(\theta - V_t) dt + \sigma\sqrt{V_t} dW_t^{(2)}, \end{aligned} \quad (2.21)$$

where $\mu \in \mathbb{R}$ is the asset drift, $\kappa > 0$ is the mean-reversion rate, $\theta > 0$ is the long-run variance level, and $\sigma > 0$ is the volatility-of-volatility parameter. The Brownian motions

$W_t^{(1)}$ and $W_t^{(2)}$ may be correlated. Let $X_t = (P_t, V_t)^\top$. The corresponding drift vector is given by

$$b(x) = \begin{pmatrix} \mu p \\ \kappa(\theta - v) \end{pmatrix}, \quad x = (p, v)^\top,$$

while the covariance matrix takes the form

$$a(x) = \begin{pmatrix} vp^2 & \rho\sigma vp \\ \rho\sigma vp & \sigma^2 v \end{pmatrix},$$

where $\rho \in [-1, 1]$ denotes the instantaneous correlation coefficient. The drift coefficients are polynomial functions of degree at most one, while the covariance coefficients are polynomial functions of degree at most two. Consequently, the Heston model is a two-dimensional polynomial diffusion. The Heston model is widely used in mathematical finance because it captures stochastic volatility, volatility clustering, and leverage effects through the correlation structure between the asset price and its variance process.

Chapter 3

Regime-Switching Polynomial Diffusion Models

Many financial and economic time series exhibit structural changes associated with different economic environments, such as periods of high and low volatility or shifts in mean-reversion dynamics. A natural approach to capturing such behavior is to allow the parameters of a stochastic model to depend on an unobserved regime process.

Following the work of Hamilton (1989) and subsequent developments such as Yin and Zhu (2010), regime-switching models introduce a latent state variable that evolves according to a finite-state Markov chain. In the present chapter, we combine this idea with the polynomial diffusion framework developed in the previous chapter.

3.1 Univariate Regime-Switching Model

3.1.1 Univariate Regime-Switching Polynomial Diffusion

Definition 3.1.1 (Univariate Regime-Switching Polynomial Diffusion). *Let*

$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, and let

$E \subseteq \mathbb{R}$ be a closed state space. Let W_t be a standard one-dimensional Brownian motion adapted to $\{\mathcal{F}_t\}_{t \geq 0}$. Let $\{S_t\}_{t \geq 0}$ be a latent $\{\mathcal{F}_t\}_{t \geq 0}$ -adapted process taking values in the finite state space $\mathcal{S} = \{1, \dots, L\}$, independent of $\{W_t\}_{t \geq 0}$. The process $\{S_t\}$ is latent in the sense that the observer has access only to the filtration $\{\mathcal{F}_t^X\}_{t \geq 0}$ generated by $\{X_t\}_{t \geq 0}$.

Let $\{X_t\}_{t \geq 0}$ be an E -valued process adapted to $\{\mathcal{F}_t\}_{t \geq 0}$ and satisfying

$$dX_t = b_{S_t}(X_t) dt + \sigma_{S_t}(X_t) dW_t, \quad (3.1)$$

where for each regime $i \in \mathcal{S}$, $b_i : E \rightarrow \mathbb{R}$ and $\sigma_i : E \rightarrow \mathbb{R}$ are the regime-specific drift and diffusion coefficients. Equivalently,

$$b_{S_t}(x) = \sum_{i=1}^L b_i(x) \mathbf{1}_{\{S_t=i\}}, \quad \sigma_{S_t}(x) = \sum_{i=1}^L \sigma_i(x) \mathbf{1}_{\{S_t=i\}}.$$

Polynomial Diffusion Structure. For each regime $i \in \mathcal{S}$, define the local variance coefficient $a_i(x) = \sigma_i(x)^2$. We assume the polynomial diffusion condition

$$b_i \in \text{Pol}_1(E), \quad a_i \in \text{Pol}_2(E), \quad i = 1, \dots, L.$$

Hence, conditional on a fixed regime $S_t = i$, the process $\{X_t\}_{t \geq 0}$ evolves as a polynomial diffusion with regime-specific generator $\mathcal{A}_i$, satisfying $\mathcal{A}_i(\text{Pol}_m(E)) \subseteq \text{Pol}_m(E)$ for all $m \geq 0$.

3.1.2 Markov Chain Dynamics for Regime Process

We assume that the regime process $\{S_t\}_{t \geq 0}$ is a time-homogeneous continuous-time Markov chain with infinitesimal generator matrix $Q = (q_{ij})_{i,j=1}^L$.

For $i \neq j$, $q_{ij} \geq 0$ denotes the transition intensity from regime i to regime j , and

$$q_{ii} = - \sum_{j \neq i} q_{ij},$$

so that each row of Q sums to zero.

For any time increment $\Delta > 0$, the continuous-time transition probability matrix is $P(\Delta) = \exp(Q\Delta)$. Its entries are $p_{ij}(\Delta) = \mathbb{P}(S_{t+\Delta} = j \mid S_t = i)$.

Since $\{S_t\}$ is time-homogeneous, the sampled process $\{S_k\}_{k=0}^N$ is a discrete-time Markov chain with one-step transition matrix

$$\Pi^\Delta = (\pi_{ij}^\Delta)_{i,j=1}^L, \quad \pi_{ij}^\Delta = \mathbb{P}(S_{k+1} = j \mid S_k = i) = [e^{Q\Delta}]_{ij}. \quad (3.2)$$

When Δ is fixed, we simplify notation by writing $\Pi := \Pi^\Delta$, $\pi_{ij} := \pi_{ij}^\Delta$.

Transition density approximation. Under Euler approximation, we denote the transition density for each regime with changing regime by

$$p_i^\Delta(x_k, x_{k+1}) = \frac{1}{(2\pi\Delta)^{1/2} a_i(x_k)^{1/2}} \quad (3.3)$$

$$\times \exp\left(-\frac{1}{2\Delta} R_k(i)^\top a_i(x_k)^{-1} R_k(i)\right), \quad (3.4)$$

where $R_k(i) = x_{k+1} - x_k - b_i(x_k)\Delta$ denotes the Euler residual associated with regime i .

When Δ is fixed, we also write $p_i(x_k, x_{k+1}) := p_i^\Delta(x_k, x_{k+1})$.

3.1.3 Hidden Markov Structure

Conditional Markov structure. The sampled process satisfies the conditional Markov property

$$\mathbb{P}(X_{k+1} \in A \mid X_{0:k}, S_{0:k}) = \mathbb{P}(X_{k+1} \in A \mid X_k, S_k), \quad (3.5)$$

for all measurable sets $A \subseteq E$.

More precisely, the model satisfies the conditional independence relations

$$\{X_{k+1} \perp (X_{0:k-1}, S_{0:k-1})\} \mid \{(X_k, S_k)\}, \quad (3.6)$$

$$\{S_{k+1} \perp X_{0:k}\} \mid \{S_k\}. \quad (3.7)$$

The first relation states that, conditional on the current state and regime (X_k, S_k) , the next state X_{k+1} is independent of the entire past history. The second relation states that the regime process evolves according only to its current regime state.

Hidden Markov Model Representation Hence, the pair

$$\{(X_k, S_k)\}_{k=0}^N$$

admits the standard hidden Markov model structure, where:

- the latent Markov chain $\{S_k\}$ represents the hidden regime process;
- the observed process $\{X_k\}$ corresponds to the discretely sampled diffusion trajectory;
- and the regime-specific transition densities p_i play the role of emission densities.

3.1.4 Joint Law Representation

Let $x_{0:N} = (x_0, \dots, x_N)$, $s_{0:N} = (s_0, \dots, s_N)$ denote the observed state path and latent regime path on the discrete grid.

The complete-data likelihood factorizes as

$$p(x_{0:N}, s_{0:N}) = \varpi_{s_0} \prod_{k=0}^{N-1} \pi_{s_k s_{k+1}} p_{s_k}(x_k, x_{k+1}), \quad (3.8)$$

where $\varpi_i = \mathbb{P}(S_0 = i)$ denotes the initial regime distribution.

The observed-data likelihood is obtained by marginalizing over all possible latent regime paths:

$$\begin{aligned} p(x_{0:N}) &= \sum_{s_{0:N} \in \mathcal{S}^{N+1}} p(x_{0:N} \mid s_{0:N}) p(s_{0:N}) \\ &= \sum_{s_{0:N} \in \mathcal{S}^{N+1}} \varpi_{s_0} \prod_{k=0}^{N-1} \pi_{s_k s_{k+1}} p_{s_k}(x_k, x_{k+1}). \end{aligned} \quad (3.9)$$

This representation has the same factorization structure as a hidden Markov model, with latent regime process $\{S_k\}$ and regime-dependent emission densities p_i .

3.1.5 Existence and Uniqueness under Regime Switching

The existence and uniqueness properties of the regime-switching polynomial diffusion follow from the corresponding properties of the underlying polynomial diffusions established in Theorem 2.3.2 and Theorem 2.3.4.

Since the latent regime process $\{S_t\}_{t \geq 0}$ is piecewise constant, the coefficients of the diffusion remain fixed between successive regime-switching times. Conditional on a fixed regime state, the dynamics therefore reduce locally to a standard polynomial diffusion with regime-specific coefficients (b_i, σ_i) .

By assumption, each regime-specific diffusion satisfies the conditions required for weak existence and uniqueness in law on the state space E . Hence, the global regime-switching process may be constructed by concatenating the local fixed-regime diffusion solutions across successive switching intervals.

Consequently, provided that each regime-specific polynomial diffusion is well posed on E , the regime-switching diffusion (3.1) is also well defined and admits a unique solution in law.

3.2 Asynchronous Multivariate Regime-Switching Model

In many financial applications, different components of a multivariate state process are driven by distinct economic forces and may transition between regimes at different times. For example, in fixed-income markets, short-end rates react rapidly to monetary-policy shocks, while long-end rates adjust more gradually through inflation expectations and term premia. A synchronous specification, which requires all components to share a common regime process, is therefore overly restrictive. The asynchronous extension developed in this section allows each component to switch regimes independently, yielding a more flexible and empirically realistic framework.

Let

$$\mathbf{X}_t = (X_t^{(1)}, \dots, X_t^{(d)})^\top \in E \subseteq \mathbb{R}^d$$

be a d -dimensional state process, and let $\mathbf{W}_t = (W_t^{(1)}, \dots, W_t^{(d)})^\top$ be a d -dimensional standard Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$.

3.2.1 Model Definition

In the asynchronous setting, each component $\eta = 1, \dots, d$ of the state process is governed by its own latent regime process. For each component η , let

$$S_t^{(\eta)} \in \mathcal{S}_\eta, \quad \mathcal{S}_\eta = \{1, \dots, L_\eta\},$$

be an independent continuous-time Markov chain with generator matrix $Q^{(\eta)} = (q_{ij}^{(\eta)})$.

The joint regime vector is

$$\mathbf{S}_t = (S_t^{(1)}, \dots, S_t^{(d)}) \in \mathcal{S}_{\text{joint}} := \prod_{\eta=1}^d \mathcal{S}_\eta.$$

Definition 3.2.1 (Asynchronous Multivariate Regime-Switching Polynomial Diffusion).

For a generic joint regime configuration $\mathbf{i} = (i_1, \dots, i_d) \in \mathcal{S}_{\text{joint}}$, the asynchronous multivariate regime-switching diffusion is defined by

$$d\mathbf{X}_t = \mathbf{b}_{\mathbf{S}_t}(\mathbf{X}_t) dt + \boldsymbol{\Sigma}_{\mathbf{S}_t}(\mathbf{X}_t) d\mathbf{W}_t, \quad (3.10)$$

where, conditional on $\mathbf{S}_t = \mathbf{i}$, $\mathbf{b}_{\mathbf{i}} : E \rightarrow \mathbb{R}^d$ is the regime-specific drift vector and $\boldsymbol{\Sigma}_{\mathbf{i}} : E \rightarrow \mathbb{R}^{d \times d}$ is the regime-specific diffusion matrix. The associated local covariance matrix is

$$\mathbf{a}_{\mathbf{i}}(\mathbf{x}) = \boldsymbol{\Sigma}_{\mathbf{i}}(\mathbf{x}) \boldsymbol{\Sigma}_{\mathbf{i}}(\mathbf{x})^\top.$$

Polynomial Diffusion Structure. For each fixed joint regime $\mathbf{i} \in \mathcal{S}_{\text{joint}}$, we assume the polynomial diffusion condition

$$b_{\mathbf{i},\eta} \in \text{Pol}_1(E), \quad a_{\mathbf{i},\eta\bar{\eta}} \in \text{Pol}_2(E), \quad \eta, \bar{\eta} = 1, \dots, d.$$

Hence, conditional on a fixed joint regime $\mathbf{S}_t = \mathbf{i}$, the process $\{\mathbf{X}_t\}_{t \geq 0}$ evolves as a multivariate polynomial diffusion with regime-specific generator $\mathcal{A}_{\mathbf{i}}$, satisfying $\mathcal{A}_{\mathbf{i}}(\text{Pol}_m(E)) \subseteq \text{Pol}_m(E)$ for all $m \geq 0$.

3.2.2 Product-State Markov Chain Structure

Since the component-specific regime processes $\{S_t^{(\eta)}\}_{\eta=1}^d$ are assumed mutually independent, the sampled joint regime process $\{\mathbf{S}_k\}_{k=0}^N$ is a discrete-time Markov chain on $\mathcal{S}_{\text{joint}}$ with product-form transition probabilities

$$\mathbb{P}(\mathbf{S}_{k+1} = \mathbf{j} \mid \mathbf{S}_k = \mathbf{i}) = \prod_{\eta=1}^d \pi_{i_{\eta}j_{\eta}}^{(\eta)}, \quad (3.11)$$

where

$$\pi_{i_{\eta}j_{\eta}}^{(\eta)} := \mathbb{P}\left(S_{k+1}^{(\eta)} = j_{\eta} \mid S_k^{(\eta)} = i_{\eta}\right) = [e^{Q^{(\eta)}\Delta}]_{i_{\eta}j_{\eta}}$$

is the one-step transition probability of the η th component chain.

The product structure in (3.11) is the key computational advantage of the asynchronous model. Although the joint state space has cardinality $|\mathcal{S}_{\text{joint}}| = \prod_{\eta=1}^d L_{\eta}$, the factorization avoids estimating an unrestricted $|\mathcal{S}_{\text{joint}}|^2$ -dimensional transition matrix, reducing instead to d separate smaller estimation problems.

3.2.3 Discretization and Transition Density

On the discrete grid $t_k = k\Delta$, conditional on the joint regime $\mathbf{S}_k = \mathbf{i}$, define the Euler residual

$$\mathbf{R}_k(\mathbf{i}) := \mathbf{x}_{k+1} - \mathbf{x}_k - \mathbf{b}_{\mathbf{i}}(\mathbf{x}_k) \Delta. \quad (3.12)$$

Under the Euler–Maruyama scheme, the conditional distribution of $\mathbf{X}_{k+1}$ is exactly

Gaussian:

$$\mathbf{X}_{k+1} \mid (\mathbf{X}_k = \mathbf{x}_k, \mathbf{S}_k = \mathbf{i}) \sim N(\mathbf{x}_k + \mathbf{b}_i(\mathbf{x}_k)\Delta, \mathbf{a}_i(\mathbf{x}_k)\Delta). \quad (3.13)$$

The corresponding regime-specific one-step transition density is

$$p_i^\Delta(\mathbf{x}_k, \mathbf{x}_{k+1}) = \frac{1}{(2\pi\Delta)^{d/2} \det(\mathbf{a}_i(\mathbf{x}_k))^{1/2}} \exp\left(-\frac{1}{2\Delta} \mathbf{R}_k(\mathbf{i})^\top \mathbf{a}_i(\mathbf{x}_k)^{-1} \mathbf{R}_k(\mathbf{i})\right), \quad (3.14)$$

where $\mathbf{R}_k(\mathbf{i})$ is as defined in (3.12). When Δ is fixed, we write $p_i(\mathbf{x}_k, \mathbf{x}_{k+1}) := p_i^\Delta(\mathbf{x}_k, \mathbf{x}_{k+1})$.

3.2.4 Conditional Markov Structure

The sampled asynchronous model satisfies the conditional Markov property

$$\mathbb{P}(\mathbf{X}_{k+1} \in A \mid \mathbf{X}_{0:k}, \mathbf{S}_{0:k}) = \mathbb{P}(\mathbf{X}_{k+1} \in A \mid \mathbf{X}_k, \mathbf{S}_k), \quad (3.15)$$

for every measurable set $A \subseteq E$, together with the regime independence condition

$$\mathbb{P}(\mathbf{S}_{k+1} = \mathbf{j} \mid \mathbf{S}_{0:k}, \mathbf{X}_{0:k}) = \mathbb{P}(\mathbf{S}_{k+1} = \mathbf{j} \mid \mathbf{S}_k).$$

Under componentwise independence, the right-hand side factorizes as in (3.11). Hence the pair $\{(\mathbf{X}_k, \mathbf{S}_k)\}_{k=0}^N$ admits a hidden Markov model structure, with $\{\mathbf{S}_k\}$ as the latent chain and p_i^Δ as the regime-specific emission density.

3.2.5 Joint Law Representation

Let $\mathbf{x}_{0:N} = (\mathbf{x}_0, \dots, \mathbf{x}_N)$ and $\mathbf{s}_{0:N} = (\mathbf{s}_0, \dots, \mathbf{s}_N)$ denote the observed multivariate path and latent joint regime path on the discrete grid. Let $\varpi_i := \mathbb{P}(\mathbf{S}_0 = \mathbf{i})$ denote the initial distribution over $\mathcal{S}_{\text{joint}}$.

The complete-data joint law factorizes as

$$p(\mathbf{x}_{0:N}, \mathbf{s}_{0:N}) = \varpi_{\mathbf{s}_0} \prod_{k=0}^{N-1} \pi_{\mathbf{s}_k \mathbf{s}_{k+1}} p_{\mathbf{s}_k}(\mathbf{x}_k, \mathbf{x}_{k+1}), \quad (3.16)$$

where, under the independent component-chain assumption,

$$\pi_{\mathbf{ij}} = \mathbb{P}(\mathbf{S}_{k+1} = \mathbf{j} \mid \mathbf{S}_k = \mathbf{i}) = \prod_{\eta=1}^d \pi_{i_\eta j_\eta}^{(\eta)}.$$

The marginal law of $\mathbf{X}_{0:N}$ is obtained by summing over all latent regime paths:

$$p(\mathbf{x}_{0:N}) = \sum_{\mathbf{s}_{0:N} \in \mathcal{S}_{\text{joint}}^{N+1}} \varpi_{\mathbf{s}_0} \prod_{k=0}^{N-1} \pi_{\mathbf{s}_k \mathbf{s}_{k+1}} p_{\mathbf{s}_k}(\mathbf{x}_k, \mathbf{x}_{k+1}). \quad (3.17)$$

Remark 3.2.2 (Likelihood interpretation). *When the model parameters*

$$\Theta = (\{\varpi_{\mathbf{i}}\}, \{Q^{(\eta)}\}, \{\mathbf{b}_{\mathbf{i}}, \Sigma_{\mathbf{i}}\})$$

are treated as unknown, the right-hand side of (3.16) is interpreted as the complete-data likelihood

$$L(\Theta \mid \mathbf{x}_{0:N}, \mathbf{s}_{0:N})$$

The joint law and the likelihood function are distinguished throughout.

This representation has the same hidden Markov factorization structure as the univariate case in Chapter 2, but the latent state now belongs to the product space $\mathcal{S}_{\text{joint}}$.

3.2.6 Existence and Uniqueness

Existence and uniqueness of the asynchronous multivariate regime-switching diffusion follow by applying Theorem 2.3.2 and Theorem 2.3.4 to each fixed joint-regime diffusion

separately.

Since $\{\mathbf{S}_t\}_{t \geq 0}$ is piecewise constant, on each interval between successive switching times the dynamics reduce to a standard multivariate polynomial diffusion with fixed coefficients $\mathbf{b}_i$ and Σ_i . By assumption, each such fixed-regime diffusion satisfies the conditions of Theorem 2.3.2 and Theorem 2.3.4 on E . The global solution is then obtained by concatenating these local solutions across switching times.

Consequently, provided that each fixed joint-regime diffusion is well posed on E , the asynchronous multivariate regime-switching diffusion (3.10) is also well defined and admits a unique solution in law.

Chapter 4

Financial Applications on Polynomial Diffusion model

4.1 Arbitrage-Free Pricing Framework

Modern continuous-time financial modelling is typically formulated under the principle of absence of arbitrage. A financial market is called arbitrage-free if it is impossible to generate riskless profits without initial investment. The mathematical foundation for arbitrage-free pricing is provided by the Fundamental Theorem of Asset Pricing.

Theorem 4.1.1 (Fundamental Theorem of Asset Pricing). *Let $X = (X^{(1)}, \dots, X^{(d)})$ denote the discounted price process of Filipović and Larsson (2016)d liquidly traded assets. Then, up to suitable technical conditions,*

$$X \text{ is arbitrage-free} \iff \exists \mathbb{Q} \sim \mathbb{P} \text{ such that } X \text{ is a local } \mathbb{Q}\text{-martingale.}$$

The probability measure $\mathbb{Q}$ is called a risk-neutral measure or an equivalent local martingale measure.

Under the risk-neutral measure $\mathbb{Q}$, discounted asset prices evolve as local martingales. Consequently, derivative prices may be represented as conditional expectations of discounted future payoffs. Let P_T be the payoff of a contingent claim at maturity T . Then its arbitrage-free price process satisfies

$$P_t = \mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_s ds \right) P_T \middle| \mathcal{F}_t \right]. \quad (4.1)$$

The main challenge in continuous-time finance is therefore the construction of stochastic processes for which the conditional expectations appearing in (4.1) can be computed efficiently. Affine and polynomial diffusions provide an important class of models for this purpose because their transform and moment structures admit tractable finite-dimensional representations.

4.2 Affine Stochastic Volatility Models

Unless otherwise specified, all models in this section are defined under the physical measure $\mathbb{P}$.

4.2.1 Heston Model

Stochastic volatility models extend the classical Black–Scholes framework by allowing the volatility process itself to evolve randomly over time. Such models are capable of capturing important empirical features of financial markets, including volatility clustering, heavy tails, leverage effects, and time-varying implied volatility surfaces.

A classical example is the Heston stochastic volatility model Heston (1993). Let $X_t = (Y_t, V_t)^\top$ with $V_t > 0$ representing the instantaneous variance process. The Heston

dynamics are

$$dP_t = \mu P_t dt + \sqrt{V_t} P_t dW_t^{(1)}, \quad (4.2)$$

$$dV_t = \kappa(\theta - V_t) dt + \xi \sqrt{V_t} dW_t^{(2)}. \quad (4.3)$$

where $(W_t^{(1)}, W_t^{(2)})$ is a correlated two-dimensional Brownian motion satisfying $d\langle W^{(1)}, W^{(2)} \rangle_t = \rho dt$.

The variance process follows a CIR-type square-root diffusion with mean-reversion speed κ , long-run variance level θ , and volatility-of-volatility parameter ξ .

The affine property follows from the fact that the drift vector is affine in the state variable and the covariance matrix depends affinely on V_t . Consequently, characteristic functions and option prices may be computed efficiently using Fourier-transform techniques.

More generally, affine stochastic volatility models form an important bridge between continuous-time asset-pricing theory and tractable derivative valuation methods. Their analytical tractability, combined with flexible volatility dynamics, has made them one of the most widely used model classes in quantitative finance.

4.2.2 Regime-Switching Stochastic Volatility Models

Regime-switching stochastic volatility models combine continuous diffusion dynamics with discrete structural changes governed by a latent regime process. Such models are particularly useful for capturing volatility clustering, market crises, abrupt structural breaks, and changing dependence patterns observed in financial time series.

Let $S_t \in \mathcal{S} = \{1, \dots, L\}$ denote a finite-state continuous-time Markov chain representing different market environments. Conditional on the current regime, the asset price and volatility dynamics evolve according to regime-dependent stochastic differential

equations Peters et al. (2026).

The regime-switching extension allows subsets of the model parameters to depend on the latent regime process S_t . Several important specifications arise naturally.

Case 1: Regime Switching in Asset Prices.

In the first specification, only the drift of the asset-price process depends on the latent regime:

$$dX_t = \mu_{S_t} X_t dt + \sqrt{V_t} X_t dW_t^{(1)}. \quad (4.4)$$

Under this formulation, the volatility dynamics remain unchanged across regimes, while the expected return process varies according to market conditions. Such a model may capture bull and bear market states or macroeconomic growth regimes.

Case 2: Regime Switching in Volatility.

A second specification allows the variance dynamics to vary across regimes:

$$dV_t = \kappa_{S_t} (\theta_{S_t} - V_t) dt + \xi_{S_t} \sqrt{V_t} dW_t^{(2)}. \quad (4.5)$$

In this setting, the latent regime affects the long-run variance level, mean-reversion speed, and volatility-of-volatility. The model is capable of describing transitions between calm and stressed market conditions, where volatility persistence and fluctuation intensity may change substantially.

Case 3: Joint Regime Switching in Asset Prices and Volatility.

A more general formulation allows both the asset-price and variance dynamics to depend on the latent regime:

$$dP_t = \mu_{S_t} P_t dt + \sqrt{V_t} P_t dW_t^{(1)}, \quad (4.6)$$

$$dV_t = \kappa_{S_t} (\theta_{S_t} - V_t) dt + \xi_{S_t} \sqrt{V_t} dW_t^{(2)}. \quad (4.7)$$

This specification allows simultaneous structural changes in both returns and volatility dynamics. Different regimes may correspond to distinct market environments characterized by different expected returns, volatility persistence, and volatility shock magnitudes.

From the polynomial-diffusion perspective, each fixed-regime model is locally affine in the state variables. Conditional on a fixed regime, the joint process (P_t, V_t) evolves according to a standard stochastic volatility diffusion. Globally, the regime-switching model may therefore be viewed as a piecewise affine stochastic volatility system driven by a latent Markov chain.

4.3 Multivariate Spot-Rate Model

We now introduce the multivariate CIR spot-rate model used in this work. The model is formulated under a risk-neutral probability measure $\mathbb{Q}$, so that arbitrage-free prices of fixed-income claims can be represented as conditional expectations of discounted future payoffs.

Let $\mathbf{X}_t = (X_t^{(1)}, \dots, X_t^{(d)})^\top \in \mathbb{R}_+^d$ denote a vector of positive rate factors, where the η -th component $X_t^{(\eta)}$ corresponds to a yield-curve segment factor associated with maturity region η , for $\eta = 1, \dots, d$.

4.3.1 Affine Short-Rate Mapping

The instantaneous short rate is specified as an affine function of the positive rate factors:

$$r_t = \alpha_0 + \boldsymbol{\alpha}^\top \mathbf{X}_t, \quad \alpha_0 \geq 0, \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)^\top \in \mathbb{R}_+^d. \quad (4.8)$$

When each factor $X_t^{(\eta)}$ is a yield spread across a pair of maturities, the coefficients $(\alpha_0, \boldsymbol{\alpha})$ admit a closed-form derivation from the CIR term-structure equations. Under the regime-switching specification, these coefficients become regime-dependent,

$$r_t = \alpha_{0,S_t} + \boldsymbol{\alpha}_{S_t}^\top \mathbf{X}_t, \quad (4.9)$$

where $S_t \in \{1, \dots, L\}$ is the latent regime at time t and the pair $(\alpha_{0,S_t}, \boldsymbol{\alpha}_{S_t})$ is determined by the individual CIR process via Cochrane–Orcutt regression Cochrane and Orcutt (1949).

The integrated short rate is

$$R_t^* = \int_0^t r_s ds. \quad (4.10)$$

In classical affine term-structure models, bond prices admit the exponential-affine representation

$$P(t, T) = \mathbb{E}^\mathbb{Q} \left[e^{-(R_T^* - R_t^*)} \mid \mathcal{F}_t \right]. \quad (4.11)$$

However, because the state process contains asynchronous regimes and copula-based dependence, closed-form exponential-affine bond prices are generally unavailable. The bond price is therefore defined by the risk-neutral valuation formula (4.1).

4.3.2 CIR Short-Rate Model

Univariate CIR Model. The Cox–Ingersoll–Ross short-rate model Cox et al. (1985) is defined on $E = \mathbb{R}_+$ by

$$dX_t = \kappa(\theta - X_t) dt + \sigma\sqrt{X_t} dW_t, \quad (4.12)$$

where $\kappa > 0$ is the mean-reversion speed, $\theta > 0$ is the long-run mean, and $\sigma > 0$ is the volatility coefficient. The infinitesimal generator of the CIR process is

$$\mathcal{A}f(x) = \kappa(\theta - x)f'(x) + \frac{1}{2}\sigma^2 x f''(x). \quad (4.13)$$

Feller Condition. A standard sufficient condition for strict positivity of the CIR process is the Feller condition Cox et al. (1985)

$$2\kappa\theta \geq \sigma^2. \quad (4.14)$$

When (4.14) holds, the boundary at zero is inaccessible and $X_t > 0$ almost surely for all $t \geq 0$. If the condition is violated, the process may reach zero but remains non-negative under the CIR dynamics.

Multivariate CIR Model. In vector form, the multivariate CIR factor dynamics are

$$d\mathbf{X}_t = \boldsymbol{\mu}(\mathbf{X}_t) dt + \boldsymbol{\Sigma}(\mathbf{X}_t) d\mathbf{W}_t, \quad (4.15)$$

where $\mathbf{W}_t = (W_t^{(1)}, \dots, W_t^{(d)})^\top$ is a d -dimensional standard Brownian motion, and the drift vector is

$$\boldsymbol{\mu}(\mathbf{x}) = (\kappa^{(1)}(\theta^{(1)} - x^{(1)}), \dots, \kappa^{(d)}(\theta^{(d)} - x^{(d)}))^\top, \quad \mathbf{x} = (x^{(1)}, \dots, x^{(d)})^\top. \quad (4.16)$$

Define the diagonal volatility matrix

$$\mathbf{D}(\mathbf{x}) = \text{diag}\left(\sigma^{(1)}\sqrt{x^{(1)}}, \dots, \sigma^{(d)}\sqrt{x^{(d)}}\right). \quad (4.17)$$

Let $\boldsymbol{\rho} = (\rho_{\eta\tilde{\eta}})_{1 \leq \eta, \tilde{\eta} \leq d}$ be a positive semidefinite correlation matrix describing instantaneous dependence across factors η and $\tilde{\eta}$. The diffusion matrix is

$$\boldsymbol{\Sigma}(\mathbf{x}) = \mathbf{D}(\mathbf{x}) \boldsymbol{\rho}^{1/2}, \quad (4.18)$$

so that the local covariance matrix is

$$\mathbf{a}(\mathbf{x}) = \boldsymbol{\Sigma}(\mathbf{x}) \boldsymbol{\Sigma}(\mathbf{x})^\top = \mathbf{D}(\mathbf{x}) \boldsymbol{\rho} \mathbf{D}(\mathbf{x}). \quad (4.19)$$

Equivalently, the $(\eta, \tilde{\eta})$ covariance entry satisfies

$$a_{\eta\tilde{\eta}}(\mathbf{x}) = \rho_{\eta\tilde{\eta}} \sigma^{(\eta)} \sigma^{(\tilde{\eta})} \sqrt{x^{(\eta)} x^{(\tilde{\eta})}}, \quad \eta, \tilde{\eta} = 1, \dots, d. \quad (4.20)$$

The marginal dynamics of the η -th factor are

$$dX_t^{(\eta)} = \kappa^{(\eta)}(\theta^{(\eta)} - X_t^{(\eta)}) dt + \sigma^{(\eta)} \sqrt{X_t^{(\eta)}} dW_t^{(\eta)}, \quad \eta = 1, \dots, d, \quad (4.21)$$

which is a one-dimensional CIR diffusion with parameters $(\kappa^{(\eta)}, \theta^{(\eta)}, \sigma^{(\eta)})$.

4.3.3 Yield-Spread Factor Construction

Let $y_t(\tau)$ denote the continuously compounded zero-coupon yield observed at time t with maturity τ . In empirical fixed-income applications, latent spot-rate factors are often constructed from yield spreads across different maturities in order to capture the slope and curvature dynamics of the yield curve.

We construct three spread factors interpreted as the short-end, belly, and long-end yield-curve factors, respectively Litterman and Scheinkman (1991); Diebold and Li (2006):

- $x_t^{(1)}$ captures the short-end slope and monetary-policy dynamics;
- $x_t^{(2)}$ captures medium-term curvature and business-cycle expectations;
- $x_t^{(3)}$ captures long-horizon inflation expectations and term premia.

Since CIR-type diffusions require strictly positive state variables, observed spreads are shifted prior to estimation. For each $\eta = 1, \dots, d$, define

$$\tilde{x}_t^{(\eta)} = x_t^{(\eta)} + c_\eta, \quad c_\eta > 0, \quad (4.22)$$

where c_η is chosen such that $\tilde{x}_t^{(\eta)} > 0$ throughout the observation period. The latent factor process entering the model is then

$$\mathbf{X}_t = (X_t^{(1)}, \dots, X_t^{(d)})^\top := (\tilde{x}_t^{(1)}, \dots, \tilde{x}_t^{(d)})^\top \in \mathbb{R}_+^d. \quad (4.23)$$

4.3.4 Asynchronous Regime-Switching Spot-Rate Model

We now extend the multivariate CIR specification to an asynchronous regime-switching framework. The key idea is that each yield-curve factor $X_t^{(\eta)}$ is governed by its own la-

tent regime process, allowing short-end, belly, and long-end dynamics to switch regimes independently in response to different economic conditions.

Component-Specific Regime Processes. For each factor $\eta = 1, \dots, d$, let

$$S_t^{(\eta)} \in \mathcal{S}_\eta, \quad \mathcal{S}_\eta = \{1, \dots, L^{(\eta)}\}, \quad (4.24)$$

be an independent continuous-time Markov chain with generator matrix $Q^{(\eta)} = (q_{ij}^{(\eta)})_{i,j=1}^{L_\eta}$. The regime index $i, j \in \mathcal{S}_\eta$ denotes a specific regime state of factor η , while the superscript (η) identifies the factor. The CIR parameters of factor η under regime i are denoted $(\kappa_i^{(\eta)}, \theta_i^{(\eta)}, \sigma_i^{(\eta)})$.

The joint regime vector is

$$\mathbf{S}_t = (S_t^{(1)}, \dots, S_t^{(d)}) \in \mathcal{S}_{\text{joint}} := \prod_{\eta=1}^d \mathcal{S}_\eta. \quad (4.25)$$

Regime-Dependent CIR Dynamics. Conditional on the regime state $S_t^{(\eta)} = i$, the η -th latent factor follows the regime-dependent CIR diffusion

$$dX_t^{(\eta)} = \kappa_i^{(\eta)} (\theta_i^{(\eta)} - X_t^{(\eta)}) dt + \sigma_i^{(\eta)} \sqrt{X_t^{(\eta)}} dW_t^{(\eta)}, \quad \eta = 1, \dots, d, \quad i \in \mathcal{S}_\eta. \quad (4.26)$$

The Brownian motions $\{W_t^{(\eta)}\}_{\eta=1}^d$ may be correlated across maturity segments:

$$d\langle W^{(\eta)}, W^{(\tilde{\eta})} \rangle_t = \rho_{\eta\tilde{\eta}} dt, \quad \eta, \tilde{\eta} = 1, \dots, d, \quad (4.27)$$

where $\boldsymbol{\rho} = (\rho_{\eta\tilde{\eta}})$ is a positive semidefinite correlation matrix shared across all regime configurations.

Affine Short-Rate Aggregation. For each fixed joint regime configuration $\mathbf{i} = (i^{(1)}, \dots, i^{(d)}) \in \mathcal{S}_{\text{joint}}$, the drift of the η -th component is affine in $X_t^{(\eta)}$ and the covariance entry $a_{\eta\bar{\eta}}$ is a polynomial of degree at most two in the state variables; see (4.20). Consequently, the asynchronous regime-switching multivariate CIR model remains within the class of polynomial diffusions conditional on the latent regime process, as established in Chapter 2.

The resulting model provides a flexible framework for capturing heterogeneous yield-curve movements, monetary-policy transitions, and regime-dependent fixed-income dynamics, estimation of this model will be explained in section 5.5.

Chapter 5

Hidden Markov Model and Inference

The regime-switching diffusion model developed in Chapter 3 admits a hidden Markov structure, where the latent regimes $\{S_k\}$ evolve according to a Markov chain and the observed process $\{X_k\}$ follows regime-dependent transition dynamics. The joint likelihood decomposes into local contributions: transition probabilities π_{ij} capture regime dynamics, while transition densities $p_i(X_k, X_{k+1})$ describe how the state evolves under each regime.

In this chapter, we outline the main statistical inference problems addressed in this thesis, present the classical HMM solutions via Baum–Welch Baum and Petrie (1966) and Viterbi Viterbi (1967) algorithms, and develop a two-stage estimation procedure for the asynchronous multivariate setting. We close by motivating the need to extend beyond classical HMM inference, which provides the foundation for the Topological Hidden Markov Model framework introduced in Chapter 6.

5.1 Problem Statements for HMM

The main problems considered in this chapter 5 and chapter 6 are summarized below:

Problem 5.1.1 (Static Parameter Estimation). *Given discrete observations $\{X_k\}_{k=0}^N$ and model parameters $\Theta = \{\{\varpi_i\}, \{\pi_{ij}\}, \{b_i, \sigma_i\}\}$, the marginal likelihood is*

$$p(X_{0:N} \mid \Theta) = \sum_{S_{0:N} \in \mathcal{S}^{N+1}} p(X_{0:N}, S_{0:N} \mid \Theta), \quad (5.1)$$

where $S_{0:N}$ denotes the latent regime sequence. The goal is to find $\hat{\Theta}$ that maximises this marginal likelihood. The Baum–Welch algorithm iteratively solves this problem, yielding estimates of both the transition matrix Π and the regime-specific diffusion parameters.

Problem 5.1.2 (Optimal Regime Sequence Decoding). *Conditioned on parameters Θ , the most probable sequence of regimes is*

$$\hat{S}_{0:N} = \arg \max_{S_{0:N}} p(S_{0:N} \mid X_{0:N}, \Theta). \quad (5.2)$$

The Viterbi algorithm computes this efficiently via dynamic programming followed by backtracking.

Problem 5.1.3 (Most Probable Trajectory via the OM Functional). *Given the decoded regime sequence $\hat{S}_{0:N}$, the most probable continuous path of X_t is obtained by minimising the Onsager–Machlup (OM) functional*

$$I[\phi] = \frac{1}{2} \int_0^T \frac{(\dot{\phi}(t) - b_{\hat{S}_t}(\phi(t)))^2}{a_{\hat{S}_t}(\phi(t))} dt + \frac{1}{2} \int_0^T b'_{\hat{S}_t}(\phi(t)) dt, \quad (5.3)$$

where the second term is the Onsager–Machlup divergence correction, accounting for the distortion of probability mass induced by the drift field, consistent with equation (6.7). The minimiser $\varphi^*(t)$ represents the most likely trajectory conditional on the regime sequence, linking discrete regime inference with continuous path optimisation. This problem is treated in Chapter 6.

5.2 Classical Hidden Markov Model

We briefly recall the standard HMM setup before developing the regime-switching diffusion specialisation.

Let $\{S_k\}_{k=0}^N$ be an unobserved regime process taking values in $\mathcal{S} = \{1, \dots, L\}$ and evolving as a time-homogeneous discrete-time Markov chain with initial distribution $\varpi_i = \mathbb{P}(S_0 = i)$ and one-step transition probabilities $\pi_{ij} = \mathbb{P}(S_{k+1} = j \mid S_k = i)$.

Let $\{X_k\}_{k=0}^N$ denote the observed process. Conditional on $S_k = i$, the emission density of X_{k+1} given X_k is $p_i(X_k, X_{k+1})$. In the regime-switching diffusion setting, this is the regime-specific Euler transition density from Section 3.1.

The complete-data likelihood factorises as

$$p(X_{0:N}, S_{0:N}) = \varpi_{S_0} \prod_{k=0}^{N-1} \pi_{S_k S_{k+1}} p_{S_k}(X_k, X_{k+1}), \quad (5.4)$$

and the observed-data likelihood is obtained by marginalising over all latent paths:

$$p(X_{0:N}) = \sum_{S_{0:N} \in \mathcal{S}^{N+1}} \varpi_{S_0} \prod_{k=0}^{N-1} \pi_{S_k S_{k+1}} p_{S_k}(X_k, X_{k+1}). \quad (5.5)$$

5.3 Parameter Estimation via Baum–Welch

The Baum–Welch algorithm is an Expectation–Maximisation (EM) procedure that solves Problem 5.1.1.

5.3.1 Forward and Backward Variables

Define the *forward variable*

$$\alpha_k(j) = p(X_{0:k}, S_k = j), \quad (5.6)$$

which represents the joint probability of observing the trajectory up to time k and being in regime j . It satisfies the recursion

$$\alpha_0(j) = \varpi_j, \quad \alpha_k(j) = \left(\sum_{i=1}^L \alpha_{k-1}(i) \pi_{ij} \right) p_j(X_{k-1}, X_k). \quad (5.7)$$

Define the *backward variable*

$$\beta_k(j) = p(X_{k+1:N} \mid S_k = j), \quad (5.8)$$

representing the likelihood of future observations given the current regime. It satisfies the recursion

$$\beta_N(j) = 1, \quad \beta_k(j) = \sum_{\ell=1}^L \pi_{j\ell} p_\ell(X_k, X_{k+1}) \beta_{k+1}(\ell). \quad (5.9)$$

5.3.2 Posterior Probabilities (E-step)

Using the forward and backward variables, we define:

- The *smoothed posterior* (soft regime assignment at time k):

$$\gamma_k(j) = p(S_k = j \mid X_{0:N}) = \frac{\alpha_k(j) \beta_k(j)}{\sum_{i'=1}^L \alpha_k(i') \beta_k(i')}. \quad (5.10)$$

- The *pairwise posterior* (expected number of transitions from i to j):

$$\xi_k(i, j) = p(S_k = i, S_{k+1} = j \mid X_{0:N}) = \frac{\alpha_k(i) \pi_{ij} p_j(X_k, X_{k+1}) \beta_{k+1}(j)}{\sum_{i'=1}^L \sum_{j'=1}^L \alpha_k(i') \pi_{i'j'} p_{j'}(X_k, X_{k+1}) \beta_{k+1}(j')}. \quad (5.11)$$

5.3.3 Parameter Updates (M-step)

Parameters are updated by maximising the expected complete-data log-likelihood:

- *Initial distribution:*

$$\varpi_j^{(r+1)} = \gamma_0(j). \quad (5.12)$$

- *Transition probabilities:*

$$\pi_{ij}^{(r+1)} = \frac{\sum_{k=0}^{N-1} \xi_k(i, j)}{\sum_{k=0}^{N-1} \gamma_k(i)}. \quad (5.13)$$

- *Diffusion parameters:*

$$\Theta_j^{(r+1)} \in \arg \max_{\Theta_j} \sum_{k=0}^{N-1} \gamma_{k+1}(j) \log p_{\Theta_j}(X_k, X_{k+1}). \quad (5.14)$$

The EM algorithm alternates between estimating latent regimes (E-step) and updating model parameters (M-step) until convergence.

5.4 Regime Decoding via Viterbi

The Viterbi algorithm solves Problem 5.1.2 by computing the most probable latent regime sequence via dynamic programming.

5.4.1 Dynamic Programming Formulation

Define the *Viterbi variable*

$$\delta_k(j) = \max_{S_{0:k-1}} p(X_{0:k}, S_k = j), \quad (5.15)$$

representing the probability of the most likely partial path ending in regime j at time k , together with the *backpointer*

$$\psi_k(j) = \arg \max_{i \in \mathcal{S}} [\delta_{k-1}(i) \pi_{ij}]. \quad (5.16)$$

5.4.2 Recursion

Initialisation:

$$\delta_0(j) = \varpi_j, \quad j \in \mathcal{S}. \quad (5.17)$$

Forward recursion for $k = 1, \dots, N$:

$$\delta_k(j) = \max_{i \in \mathcal{S}} [\delta_{k-1}(i) \pi_{ij}] p_j(X_{k-1}, X_k). \quad (5.18)$$

Termination:

$$\hat{S}_N = \arg \max_{j \in \mathcal{S}} \delta_N(j). \quad (5.19)$$

Backtracking for $k = N - 1, \dots, 0$:

$$\hat{S}_k = \psi_{k+1}(\hat{S}_{k+1}). \quad (5.20)$$

The output $\hat{S}_{0:N}$ is the optimal regime path solving Problem 5.1.2.

5.5 Two-Stage Marginal Estimation and Covariance Reconstruction

In the multivariate asynchronous regime-switching setting of Section 3.2, direct joint inference is computationally intractable: if each component $\eta = 1, \dots, d$ admits $L^{(\eta)}$ regimes, the joint regime space has size $\prod_{\eta=1}^d L^{(\eta)}$, rendering standard HMM inference infeasible.

We therefore adopt a two-stage estimation strategy based on marginal inference and subsequent reconstruction of the covariance structure.

5.5.1 Model Structure and Justification

Let $X_t = (X_t^{(1)}, \dots, X_t^{(d)})^\top \in \mathbb{R}_+^d$ satisfy the asynchronous regime-switching diffusion

$$dX_t = \mu_{S_t}(X_t) dt + \Sigma_{S_t}(X_t) dW_t, \quad (5.21)$$

where $S_t = (S_t^{(1)}, \dots, S_t^{(d)})$ denotes the asynchronous joint regime process.

We assume:

- The component-wise regime processes $\{S_t^{(\eta)}\}$ are mutually independent;
- Cross-sectional dependence arises only through the Brownian motion, via a shared correlation matrix $\rho = (\rho_{\eta\bar{\eta}})$;
- There is no lagged cross-dependence in the drift or diffusion coefficients.

Under these assumptions, the dependence structure is purely contemporaneous and does not affect the temporal evolution of individual regime processes. Consequently, regime inference can be performed independently for each marginal component.

5.5.2 Stage 1: Marginal Estimation

For each component $\eta = 1, \dots, d$, we treat $X_t^{(\eta)}$ as a univariate regime-switching CIR diffusion. Conditional on $S_t^{(\eta)} = i$, the dynamics are

$$dX_t^{(\eta)} = \kappa_i^{(\eta)} (\theta_i^{(\eta)} - X_t^{(\eta)}) dt + \sigma_i^{(\eta)} \sqrt{X_t^{(\eta)}} dW_t^{(\eta)}, \quad i \in \mathcal{S}_\eta. \quad (5.22)$$

Applying the Baum–Welch algorithm of Section 5.3 and the Viterbi decoder of Section 5.4 to each marginal separately yields:

- regime-specific parameter estimates $\hat{\kappa}_i^{(\eta)}, \hat{\theta}_i^{(\eta)}, \hat{\sigma}_i^{(\eta)}$;

- decoded regime sequences $\widehat{S}_k^{(\eta)}$ for each component η .

5.5.3 Stage 2: Standardised Residuals and Covariance Reconstruction

Given the Stage 1 estimates, we construct discrete-time residuals on the grid $t_k = k\Delta$:

$$R_k = X_{k+1} - X_k - \mu_{\widehat{S}_k}(X_k) \Delta, \quad (5.23)$$

where $\widehat{S}_k = (\widehat{S}_k^{(1)}, \dots, \widehat{S}_k^{(d)})$ is the component-wise decoded regime vector.

Under the diffusion model, $R_k \approx \Sigma_{\widehat{S}_k}(X_k) \Delta W_k$ with $\Delta W_k \sim \mathcal{N}(0, \Delta \rho)$. We isolate the dependence structure by standardising:

$$\widehat{\varepsilon}_k = \Sigma_{\widehat{S}_k}(X_k)^{-1} R_k \approx \Delta W_k, \quad (5.24)$$

so that $\text{Cov}(\widehat{\varepsilon}_k) \approx \Delta \rho$.

The correlation matrix is estimated from the standardised residuals:

$$\widehat{\rho} = \frac{1}{N} \sum_{k=0}^{N-1} \frac{\widehat{\varepsilon}_k \widehat{\varepsilon}_k^\top}{\Delta}, \quad (5.25)$$

followed by normalisation to enforce unit diagonal. Finally, the local covariance matrix is reconstructed as

$$\widehat{a}_{S_k}(X_k) = \Sigma_{\widehat{S}_k}(X_k) \widehat{\rho} \Sigma_{\widehat{S}_k}(X_k)^\top, \quad (5.26)$$

separating marginal volatility effects (captured by Σ) from cross-sectional dependence (captured by ρ).

5.6 Limitations of Classical HMM and Motivation for THMM

The classical Baum–Welch and Viterbi procedures described above rely on the availability of explicit regime-dependent transition densities $p_i(X_k, X_{k+1})$. Under the Euler–Maruyama approximation, these are Gaussian and admit closed-form expressions; see Section 3.1.

However, this approach has two fundamental limitations. Classical HMM inference evaluates the likelihood of X_{k+1} given X_k using only the two endpoints of each observed increment. It does not exploit the local path structure of the trajectory, including the velocity $\dot{X}_t$ and higher-order behaviour of the sample path. As a result, regime transitions that manifest through path dynamics rather than through level shifts may not be well detected.

These limitations motivate the Topological Hidden Markov Model (THMM) framework, in which the pointwise transition density $p_i(X_k, X_{k+1})$ is replaced by an energy-based surrogate derived from the Onsager–Machlup (OM) functional. This surrogate scores the local compatibility of the observed path increment with each candidate regime, incorporating continuous-time path information while preserving the computational structure of classical HMM inference. The THMM framework is developed in Chapter 6.

Chapter 6

Topological Hidden Markov Model

This chapter develops the Topological Hidden Markov Model (THMM) framework ? and its application to regime-switching polynomial diffusion models. We first introduce the Onsager–Machlup (OM) functional as a pathwise analogue of the log-likelihood for diffusion trajectories. We then construct the THMM by replacing the classical emission density with an OM-based transition weight, and finally develop the THMM inference procedures that address Problems 5.1.1 and 5.1.2 from Chapter 5, together with the most-probable-path problem (Problem 5.1.3).

6.1 Onsager–Machlup Functional

In infinite-dimensional spaces, there is generally no analogue of Lebesgue measure and therefore no classical probability density function. To address this issue, the THMM framework replaces the standard likelihood density with the Onsager–Machlup (OM) functional. The OM functional measures the relative asymptotic probability of small neighbourhoods around two paths under a Gaussian measure ?.

6.1.1 Abstract Definition

Let γ be a Gaussian measure on a metric space $\mathcal{X}$. The Onsager–Machlup functional is defined through the ratio

$$I(u, v) = \lim_{\varepsilon \downarrow 0} \frac{\gamma(K(u, \varepsilon))}{\gamma(K(v, \varepsilon))}, \quad u, v \in \mathcal{X}, \quad (6.1)$$

where $K(u, \varepsilon)$ denotes the closed ball of radius ε centred at u . Intuitively, this functional compares the relative probability that a Gaussian process remains close to one trajectory versus another as the neighbourhood size shrinks to zero from above.

Definition 6.1.1 (Cameron–Martin Space). *The Cameron–Martin space $H(\gamma)$ is the reproducing kernel Hilbert space associated with the Gaussian measure γ . More precisely, $H(\gamma)$ consists of all elements $h \in \mathcal{X}$ such that the shifted measure $\gamma(\cdot - h)$ is absolutely continuous with respect to γ , equipped with the norm $\|\cdot\|_{H(\gamma)}$ induced by the Cameron–Martin inner product. For a Gaussian measure on a function space, $H(\gamma)$ typically consists of paths with one additional degree of differentiability relative to the support of γ .*

Within the Cameron–Martin space $H(\gamma)$, the OM functional admits the representation

$$\lim_{\varepsilon \downarrow 0} \frac{\gamma(V_\varepsilon + h)}{\gamma(V_\varepsilon + k)} = \exp\left(\frac{1}{2}\|\pi_q k\|_{H(\gamma)}^2 - \frac{1}{2}\|\pi_q h\|_{H(\gamma)}^2\right), \quad (6.2)$$

where π_q denotes the orthogonal projection associated with the seminorm structure of the underlying space. Using this formulation, the THMM replaces the standard emission density with an OM-based emission function of the form

$$b_i(O_t) = \exp\left(-\frac{1}{2}\|\pi_q(O_t - h_i)\|_{H(\gamma)}^2\right), \quad (6.3)$$

where $h_i \in H(\gamma)$ represents the latent state mean associated with regime i . The OM functional therefore acts as an infinite-dimensional analogue of a Gaussian log-likelihood, allowing Hidden Markov Models to be extended from finite-dimensional Euclidean observations to functional data and stochastic process trajectories.

6.1.2 OM Functional for Diffusion Processes

For simplicity, we introduce the OM functional in the univariate setting. Consider a one-dimensional Itô diffusion process

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t, \quad a(x) = \sigma(x)^2, \quad (6.4)$$

where $b : \mathbb{R} \rightarrow \mathbb{R}$ is the drift, $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ is the diffusion coefficient, and W_t is a standard Brownian motion.

For an absolutely continuous path $\varphi : [0, T] \rightarrow \mathbb{R}$, the probability that the diffusion remains within a small tube around φ satisfies

$$\mathbb{P}(X_t \text{ stays within a small neighbourhood of } \varphi(t)) \approx \exp\left(-\int_0^T \mathcal{L}(\varphi(t), \dot{\varphi}(t)) dt\right) \quad (6.5)$$

as the tube radius tends to zero Onsager and Machlup (1953). The Lagrangian for the univariate case is

$$\mathcal{L}(x, v) = \frac{(v - b(x))^2}{2a(x)} + \frac{1}{2}b'(x), \quad (6.6)$$

where the first term measures the squared deviation of the path velocity from the drift scaled by the local variance, and the second term accounts for the distortion of probability mass induced by the drift field.

The associated Onsager–Machlup functional is

$$I[\varphi] = \int_0^T \mathcal{L}(\varphi(t), \dot{\varphi}(t)) dt = \frac{1}{2} \int_0^T \frac{(\dot{\varphi}(t) - b(\varphi(t)))^2}{a(\varphi(t))} dt + \frac{1}{2} \int_0^T b'(\varphi(t)) dt, \quad (6.7)$$

which serves as a pathwise analogue of the negative log-likelihood. φ is defined as 6.8. Paths that closely follow the deterministic drift $b(\cdot)$ incur a low cost, while paths that deviate significantly from the drift are heavily penalised. Minimising $I[\varphi]$ yields the most probable path connecting two states, in direct analogy with the principle of least action in classical mechanics.

Remark 6.1.2. *The OM functional extends naturally to the multivariate setting by replacing the scalar variance $a(x)$ with the local covariance matrix $a(x) = \sigma(x)\sigma(x)^\top$ and the scalar squared deviation with the quadratic form $(v - b(x))^\top a(x)^{-1}(v - b(x))$; see the second paper for details. In this thesis, we work primarily in the univariate setting for clarity of exposition.*

Remark 6.1.3 (Regularity of OM Functional for CIR Diffusions). *The Lagrangian (6.6) and the OM functional (6.7) involve the term $a(\varphi(t))^{-1}$, which is well defined only when the local variance $a(\varphi(t))$ is strictly positive along the path. For the CIR process, however, $a(x) = \sigma^2 x$ vanishes at $x = 0$, so $a(\varphi(t))^{-1}$ becomes singular as $\varphi(t) \rightarrow 0$. Consequently, $I[\varphi]$ in (6.7) is meaningful only for strictly positive paths satisfying $\varphi(t) > 0$ for all $t \in [0, T]$.*

The Feller condition $2\kappa\theta \geq \sigma^2$ provides the required guarantee: when it holds, a CIR process started from $X_0 > 0$ never reaches zero almost surely, so every realised sample path satisfies $X_t > 0$ for all $t \geq 0$, and the OM functional is well defined on those paths.

A more rigorous derivation applies the Lamperti transform Madsen and Lindström (2010) $Y_t = 2\sqrt{X_t}/\sigma$, which maps the CIR diffusion to a unit-diffusion process $dY_t = b_Y(Y_t) dt + dW_t$. The OM functional can then be derived in Y -space, where the standard

non-degenerate theory applies, and subsequently transformed back to the original coordinate. In all CIR applications in this thesis the Feller condition is assumed to hold, which justifies the direct X -space formulation used throughout.

6.1.3 Discretisation and Local OM Cost

To obtain a tractable approximation of (6.7), we replace the time derivative by the forward difference

$$\dot{\varphi}(t_k) \approx \frac{\varphi(t_{k+1}) - \varphi(t_k)}{\Delta}, \quad (6.8)$$

and approximate the integrals by Riemann sums over the discrete grid $t_k = k\Delta$.

For a realised transition from X_k to X_{k+1} , define the *Euler residual*

$$R_k = X_{k+1} - X_k - b(X_k)\Delta. \quad (6.9)$$

The corresponding *local Onsager–Machlup cost function* $C : E \times E \rightarrow \mathbb{R}$ is defined by

$$C(X_k, X_{k+1}) := \frac{R_k^2}{2\Delta a(X_k)} + \frac{\Delta}{2} b'(X_k). \quad (6.10)$$

The continuous OM functional is thus approximated by the additive path cost

$$I[\varphi] \approx \sum_{k=0}^{N-1} C(X_k, X_{k+1}). \quad (6.11)$$

If the increment agrees well with the drift prediction $b(X_k)\Delta$, the residual R_k is small and the cost is low. If the increment is unlikely under the diffusion dynamics, the cost is high. Hence C acts as a local negative log-likelihood surrogate for one observed transition.

Remark 6.1.4 (Numerical Stability of the Discrete OM Cost for CIR Processes). *For*

CIR-type processes where $a(X_k) = \sigma^2 X_k$, the local OM cost (6.10) contains the term

$$\frac{R_k^2}{2\Delta a(X_k)} = \frac{R_k^2}{2\Delta \sigma^2 X_k},$$

which diverges as $X_k \rightarrow 0$. Moreover, the plain Euler–Maruyama scheme can in principle generate non-positive values of X_k , rendering $a(X_k)^{-1}$ undefined. Three safeguards are adopted in this work.

1. **Feller condition.** All regime-specific CIR parameters satisfy $2\kappa_i\theta_i \geq \sigma_i^2$, ensuring the true process remains strictly positive almost surely.
2. **Positive shift.** Observed yield spreads are shifted prior to estimation via $\tilde{x}_t^{(\eta)} = x_t^{(\eta)} + c_\eta$, $c_\eta > 0$, so that $X_k > 0$ holds for all observed grid points.
3. **Regularisation floor.** In numerical implementation a small lower bound $\varepsilon > 0$ is applied, replacing $a(X_k)$ by $\max(a(X_k), \varepsilon)$ when necessary.

From a theoretical standpoint, the mathematically consistent procedure is to discretise the OM functional in the Lamperti-transformed Madsen and Lindström (2010) coordinate $Y_k = 2\sqrt{X_k}/\sigma$ and map back to the original scale, which avoids the singularity at zero entirely. The direct X -space discretisation used here is justified by the positivity guarantees provided by points (i) and (ii) above.

Theorem 6.1.5 (OM Action for the Regime-Switching Heston Model). *Let $\varphi_P : [0, T] \rightarrow (0, \infty)$ and $\varphi_V : [0, T] \rightarrow (0, \infty)$ be absolutely continuous candidate paths with almost everywhere derivatives $\dot{\varphi}_P, \dot{\varphi}_V$. Define the regime- i OM action (negative log-emission up to an additive constant independent of (φ_P, φ_V)) by*

$$\mathcal{L}_i[\varphi_P, \varphi_V] = \frac{1}{2} \int_0^T \left[\frac{(\dot{\varphi}_P - \mu_i \varphi_P)^2}{V_i \varphi_P^2} + \frac{(\dot{\varphi}_V - \kappa_i(\theta_i - \varphi_V))^2}{\xi_i^2 \varphi_V} \right] dt, \quad (6.12)$$

where V_i denotes the regime- i variance level. Then $\mathcal{L}_i[\varphi_P, \varphi_V]$ can be computed numerically from discretely observed paths $\{(\varphi_P(t_k), \varphi_V(t_k))\}_{k=0}^N$ on the grid $0 = t_0 < \dots < t_N = T$ via the Riemann-sum approximation

$$\mathcal{L}_i[\varphi_P, \varphi_V] \approx \frac{1}{2} \sum_{k=0}^{N-1} \Delta_k \left[\frac{(\Delta\varphi_{P,k}/\Delta_k - \mu_i \varphi_{P,k})^2}{V_i \varphi_{P,k}^2} + \frac{(\Delta\varphi_{V,k}/\Delta_k - \kappa_i(\theta_i - \varphi_{V,k}))^2}{\xi_i^2 \varphi_{V,k}} \right], \quad (6.13)$$

where $\Delta_k = t_{k+1} - t_k$, $\Delta\varphi_{P,k} = \varphi_P(t_{k+1}) - \varphi_P(t_k)$, $\Delta\varphi_{V,k} = \varphi_V(t_{k+1}) - \varphi_V(t_k)$, and $\varphi_{P,k}, \varphi_{V,k}$ are left-endpoint values on each interval.

Proof: The expression (6.12) is the quadratic energy (action) associated with the two-dimensional regime- i Heston diffusion, written in the form $dX_t = b_i(X_t) dt + \Sigma_i(X_t) dW_t$ with $X_t = (P_t, V_t)^\top$, drift

$$b_i(P, V) = (\mu_i P, \kappa_i(\theta_i - V))^\top, \quad (6.14)$$

and diagonal diffusion matrix

$$\Sigma_i(P, V) = \begin{pmatrix} P\sqrt{V_i} & 0 \\ 0 & \xi_i\sqrt{V} \end{pmatrix}. \quad (6.15)$$

For an absolutely continuous candidate path $\varphi(t) = (\varphi_P(t), \varphi_V(t))^\top$, the OM Lagrangian integrand is proportional (up to additive constants not depending on φ) to the squared norm of the control term $\dot{\varphi}(t) - b_i(\varphi(t))$ measured in the inverse diffusion metric:

$$\frac{1}{2} (\dot{\varphi}(t) - b_i(\varphi(t)))^\top (\Sigma_i(\varphi(t)) \Sigma_i(\varphi(t))^\top)^{-1} (\dot{\varphi}(t) - b_i(\varphi(t))). \quad (6.16)$$

Since $\Sigma_i \Sigma_i^\top = \text{diag}(V_i \varphi_P^2, \xi_i^2 \varphi_V)$, its inverse is $\text{diag}((V_i \varphi_P^2)^{-1}, (\xi_i^2 \varphi_V)^{-1})$, which yields exactly the integrand in (6.12). Finally, approximating time derivatives by forward dif-

ferences $\dot{\varphi}_P(t_k) \approx \Delta\varphi_{P,k}/\Delta_k$, $\dot{\varphi}_V(t_k) \approx \Delta\varphi_{V,k}/\Delta_k$, and the integral by a Riemann sum gives (6.13).

6.2 Topological Hidden Markov Models

We now use the OM cost developed in Section 6.1 to construct a Topological Hidden Markov Model for regime-switching diffusion processes.

6.2.1 OM-Based Transition Weights

Under the Onsager–Machlup interpretation, the likelihood of a path is proportional to the exponential of the negative action functional. Using (6.11), this motivates the *local transition weight*

$$w(X_k, X_{k+1}) \propto \exp(-C(X_k, X_{k+1})). \quad (6.17)$$

The quantity $w(X_k, X_{k+1})$ is not necessarily a normalised transition density. Instead, it is an energy-based surrogate likelihood that scores how compatible the observed transition is with the diffusion dynamics.

In practice, one may also use the *Gaussian Euler approximation*

$$C^{\text{Gauss}}(X_k, X_{k+1}) = \frac{R_k^\top a(X_k)^{-1} R_k}{2\Delta} + \frac{1}{2} \log \det a(X_k), \quad (6.18)$$

which corresponds to the negative log-density of the Euler Gaussian transition up to constants independent of the model parameters.

6.2.2 THMM for Regime-Switching Diffusions

We now introduce the latent regime process $S_k \in \mathcal{S} = \{1, \dots, L\}$. Each regime corresponds to a different diffusion specification: conditional on regime i , the state process evolves according to

$$dX_t = b_i(X_t) dt + \sigma_i(X_t) dW_t, \quad (6.19)$$

with local variance $a_i(x) = \sigma_i(x)^2$.

For regime i , define the *regime-specific Euler residual*

$$R_k^{(i)} = X_{k+1} - X_k - b_i(X_k)\Delta, \quad (6.20)$$

the *regime-specific local OM cost* $C_i : E \times E \rightarrow \mathbb{R}$,

$$C_i(X_k, X_{k+1}) := \frac{(R_k^{(i)})^2}{2\Delta a_i(X_k)} + \frac{\Delta}{2} b_i'(X_k), \quad (6.21)$$

and the *regime-specific OM transition weight*

$$w_i(X_k, X_{k+1}) \propto \exp(-C_i(X_k, X_{k+1})). \quad (6.22)$$

This cost allows each candidate regime to assign a different pathwise score to the same observed transition. Regime inference is then performed by comparing which regime produces the lowest accumulated OM cost along the observed trajectory.

Joint law representation. The OM-based construction yields a THMM in which the classical emission density is replaced by a pathwise transition weight. The THMM joint

likelihood surrogate is defined up to proportionality as

$$L(X_{0:N}, S_{0:N}) \propto \varpi_{S_0} \prod_{k=0}^{N-1} \pi_{S_k S_{k+1}} w_{S_k}(X_k, X_{k+1}) \quad (6.23)$$

$$\propto \varpi_{S_0} \prod_{k=0}^{N-1} \pi_{S_k S_{k+1}} \exp(-C_{S_k}(X_k, X_{k+1})). \quad (6.24)$$

This representation preserves the additive structure required for forward–backward recursion, EM estimation, and Viterbi decoding, while avoiding explicit evaluation of transition densities that may be unavailable or computationally expensive.

6.3 THMM Inference

The THMM framework addresses Problems 5.1.1, 5.1.2, and 5.1.3 from Chapter 5 by replacing the classical emission density $p_i(X_k, X_{k+1})$ with the OM transition weight $w_i(X_k, X_{k+1})$ throughout.

6.3.1 Modification of the Classical HMM Algorithms

The substitution

$$p_i(X_k, X_{k+1}) \longrightarrow w_i(X_k, X_{k+1}) \quad (6.25)$$

is the *only* modification to the classical algorithms. The Markov transition probabilities π_{ij} , the initial regime distribution ϖ_i , and the dynamic programming structure of the forward–backward and Viterbi algorithms all remain unchanged.

Forward–backward recursion. The forward recursion (5.7) becomes

$$\alpha_k(j) = \left(\sum_{i=1}^L \alpha_{k-1}(i) \pi_{ij} \right) w_j(X_{k-1}, X_k). \quad (6.26)$$

The backward recursion and all posterior quantities $\gamma_k(j)$, $\xi_k(i, j)$ are obtained analogously by replacing $p_j(X_k, X_{k+1})$ with $w_j(X_k, X_{k+1})$.

Viterbi decoding. The Viterbi recursion from Section 5.4 becomes

$$\delta_k(j) = \max_{i \in \mathcal{S}} [\delta_{k-1}(i) \pi_{ij}] w_j(X_{k-1}, X_k). \quad (6.27)$$

M-step update. In the EM algorithm, the E-step is unchanged in form. The M-step replaces the weighted log-density objective in (5.14),

$$\sum_{k=0}^{N-1} \gamma_k(j) \log p_j(X_k, X_{k+1}), \quad (6.28)$$

by the weighted OM objective

$$- \sum_{k=0}^{N-1} \gamma_k(j) C_j(X_k, X_{k+1}). \quad (6.29)$$

Hence, the THMM preserves the computational structure of the classical HMM but replaces density-based inference by cost-based pathwise inference.

6.3.2 Most Probable Path: Solving Problem 5.1.3

Given the decoded regime sequence $\widehat{S}_{0:N}$ from the THMM Viterbi algorithm, Problem 5.1.3 asks for the most probable continuous path conditional on the regimes. This is solved by minimising the OM functional along each fixed-regime segment:

$$\varphi^* = \arg \min_{\varphi} \sum_{k=0}^{N-1} C_{\widehat{S}_k}(\varphi(t_k), \varphi(t_{k+1})), \quad (6.30)$$

subject to $\varphi(t_0) = X_0$. The minimiser links the discrete regime inference with continuous-time path optimisation, providing a trajectory reconstruction that is consistent with both the observed data and the inferred regime dynamics.

6.3.3 Interpretation

The classical HMM relies on explicit transition densities, whereas the THMM uses the OM functional to measure the local compatibility of a path segment with a given regime. In this sense, the OM functional provides a geometric and pathwise alternative to density-based emissions, making regime inference feasible even when exact transition densities are intractable.

The key insight is that the OM cost $C_i(X_k, X_{k+1})$ summarises the entire local path information into a scalar penalty that replaces the role of the log-transition-density $\log p_i(X_k, X_{k+1})$. Regimes for which the observed increment closely follows the drift prediction $b_i(X_k)\Delta$ receive low cost and high weight; regimes inconsistent with the local dynamics are penalised. This pathwise scoring provides a more principled and geometrically motivated basis for regime inference in nonlinear diffusion systems than the Euler Gaussian approximation alone.

Chapter 7

Simulation and Experiments

This section provides controlled experimental case studies to compare classical HMM inference (Baum–Welch estimation and Viterbi decoding) with the proposed topological HMM (THMM) modifications incorporating Onsager–Machlup (OM) based emissions. We focus on a two-regime switching Heston model in which the latent market state alternates between a calm regime and a turbulent regime for the volatility. Using both a baseline and a challenging configuration, we evaluate regime recovery accuracy, parameter estimation stability, and the implications for volatility diagnostics and option-implied quantities.

7.1 Simulation Studies THMM

This section presents controlled simulation studies designed to evaluate whether the proposed THMM framework can recover latent volatility regimes from asset-price data generated under a regime-switching Heston model. We compare two inferential approaches:

- *Standard HMM*: classical Baum–Welch and Viterbi inference based on pointwise

Gaussian emission likelihoods;

- *THMM*: the proposed framework replacing emission densities with Onsager–Machlup pathwise costs.

Performance is assessed along three dimensions: regime classification accuracy via Viterbi decoding, parameter recovery via Baum–Welch estimation, and uncertainty quantification via confidence intervals.

7.1.1 Data Simulation

We first generate a latent regime sequence $\{S_k\}_{k=0}^N$, where $S_k \in \{1, 2\}$ represents the underlying market volatility state (e.g., calm or turbulent regime). The regime process evolves according to a two-state Markov chain with transition matrix $\Pi = (\pi_{ij})$.

Conditional on the realized regime path $\{S_k\}$, we then simulate the asset-price process $\{P_k\}$ and latent variance process $\{V_k\}$ under regime-dependent Heston dynamics. For each regime i , the Heston model is characterized by the parameter set $(\mu_i, \kappa_i, \theta_i, \xi_i, \rho)$.

Using these regime-dependent parameters, we simulate the discrete-time financial process $\{P_k, V_k\}$ on a daily grid via an Euler-type discretization scheme. The resulting asset-price process $\{P_k\}$ serves as the observable financial time series, while the variance process $\{V_k\}$ remains latent and represents the hidden volatility state of the market.

To construct option-implied volatility diagnostics, we additionally generate a grid of European option prices across multiple strikes and delta levels using the Heston pricing model.

Latent Regime Process We consider the two-regime Heston model in (4.7) with regime process $S_k \in \{1, 2\}$ evolving as a discrete-time Markov chain with transition

matrix

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix}, \quad \pi_{i1} + \pi_{i2} = 1.$$

In both configurations, we set $\pi_{11} = \pi_{22} = 0.99$ and $\pi_{12} = \pi_{21} = 0.01$, so that regimes will not switch frequently. We simulate $N = 252$ daily observations, where each trading day is associated with a single latent regime state.

Heston Parameters The regime-dependent Heston parameters differ across experimental configurations. In the baseline setting, regimes are well separated and easy to differentiate, see paper Peters et al. (2026), whereas in the challenging setting the parameter values are more similar, making regime identification substantially more difficult. The parameter values for challenging case are given in Table 7.1.

Table 7.1: Challenging simulation parameters (Calm vs. Turbulent).

Regime	μ_i	κ_i	θ_i	ξ_i	ρ_i	p_0	v_0
Calm ($i = 1$)	0.50	5	0.10	0.05	-0.90	100	0.10
Turbulent ($i = 2$)	0.50	1	0.20	0.05	-0.90	100	0.10

Simulation Procedure We simulate $N = 252$ daily observations over a one-year horizon ($\Delta = 1/252$), with $p_0 = 100$ and $v_0 = 0.10$. Conditional on regime $S_k = i$, the state (P_{k+1}, V_{k+1}) is updated via the log-Euler scheme

$$P_{k+1} = P_k \exp\left((\mu_i - \frac{1}{2}V_k)\Delta + \sqrt{V_k}\Delta Z_k^{(1)}\right), \quad (7.1)$$

$$V_{k+1} = \max\left\{v_{\min}, V_k + \kappa_i(\theta_i - V_k)\Delta + \xi_i\sqrt{\max(V_k, 0)}\Delta Z_k^{(2)}\right\}, \quad (7.2)$$

where $(Z_k^{(1)}, Z_k^{(2)})^\top \sim N(\mathbf{0}, \mathbf{R})$ with $\mathbf{R}_{12} = \rho$, and $v_{\min} > 0$ is a small positivity floor.

Option Surface Construction under the Risk-Neutral Measure. To construct option-implied volatility diagnostics and expected-move bands, we additionally generate a synthetic option surface under the risk-neutral probability measure $\mathbb{Q}$. While the latent volatility dynamics and regime inference procedures are formulated under the physical measure $\mathbb{P}$, option prices must be computed under $\mathbb{Q}$ to ensure arbitrage-free valuation.

Under the risk-neutral measure, the asset-price dynamics satisfy

$$dP_t = (r - q)P_t dt + \sqrt{V_t} P_t dW_t^{(1),\mathbb{Q}}, \quad (7.3)$$

where r denotes the risk-free interest rate and q denotes the dividend yield. In this example, we take $r = 0.02$ and $q = 0$. The variance process evolves according to the Heston stochastic volatility dynamics

$$dV_t = \kappa_i^{\mathbb{Q}}(\theta_i^{\mathbb{Q}} - V_t) dt + \xi_i \sqrt{V_t} dW_t^{(2),\mathbb{Q}}, \quad (7.4)$$

with instantaneous correlation $d\langle W^{(1),\mathbb{Q}}, W^{(2),\mathbb{Q}} \rangle_t = \rho_i dt$.

For each simulated trading day t_k , we construct a grid of European option prices across multiple strike levels

$$K_{k,h} = P_k \times \mathcal{M}_h,$$

where $\{\mathcal{M}_h\}$ denotes a predefined relative-moneyness grid. In this simulation study, we use $\mathcal{M}_h \in [0.85, 1.15]$, with grid spacing 0.01. Equivalently, the strike grid is given by $\mathcal{M}_h = 0.85, 0.86, \dots, 1.15$. Option maturities are fixed at a one-month horizon $T = 30/252$ for simplicity.

Remark 7.1.1 (Physical and Risk-Neutral Measures in Option Construction). *The option-implied quantities constructed in this section are generated under the risk-neutral probability measure $\mathbb{Q}$, since arbitrage-free option pricing requires discounted asset prices*

to evolve as local martingales under an equivalent martingale measure. 4.1.1

In contrast, the observed historical asset-price process and realized volatility series arise under the physical probability measure $\mathbb{P}$, which governs the real-world dynamics of the market. Consequently, the implied volatility surface extracted from option prices under $\mathbb{Q}$ need not coincide with realized or historical volatility observed under $\mathbb{P}$. This discrepancy is commonly referred to as the realized variance disagreement Chong and Todorov (2026), and reflects compensation demanded by investors for bearing volatility risk.

Accordingly, the regime inference and statistical calibration procedures developed in this thesis are performed primarily under the physical measure $\mathbb{P}$, using historical asset-price and volatility observations. The risk-neutral measure $\mathbb{Q}$ is introduced separately for the construction of synthetic option surfaces, implied-volatility diagnostics, and expected-move volatility bands.

7.1.2 Volatility Proxy Construction

Volatility Proxy Three volatility series enter the inference procedure:

- **True variance** V_k : known in simulation, used only for diagnostic comparison.
- **Realized variance** RV_k : intraday sum-of-squares estimator; noisy proxy for V_k .
- **Smoothed realized variance** $\widetilde{RV}_k$: LOWESS-smoothed version of RV_k , reducing high-frequency noise while preserving low-frequency regime structure Alizadeh et al. (2002).

Both RV_k and $\widetilde{RV}_k$ are constructed volatility proxies derived from the simulated price process. Inference is performed on $\widetilde{RV}_k$ as the primary volatility input. Therefore the

intraday simulation yields intraday log-returns $\{r_{k,m}\}_{m=1}^{n_{\text{intraday}}}$ and daily realized variance $\text{RV}_k = \sum_m r_{k,m}^2$.

Expected Move Volatility Band

To provide an option-implied reference for the volatility scale, we compute a one-month expected move band from the Heston implied volatility surface. For each day t , two strikes (K_t^+, K_t^-) are determined by

$$\Delta_{\text{Heston}}^{\text{call}}(K_t^+) = 0.1, \quad \Delta_{\text{Heston}}^{\text{put}}(K_t^-) = -0.1,$$

defining an 80% option-implied price interval. The corresponding implied volatility levels $\sigma_t^\pm$ and variance endpoints $V_{t+T}^\pm = (\sigma_t^\pm)^2$ are then connected to the current instantaneous variance V_t via the linear path

$$\phi_t^\pm(\tau) = V_t + \frac{V_{t+T}^\pm - V_t}{T} \tau, \quad \tau \in [0, T], \quad (7.5)$$

with $T = 30/252$ (one month). These upper and lower variance trajectories serve as option-market benchmarks against which the inferred regime dynamics are compared; see Figure 7.1.

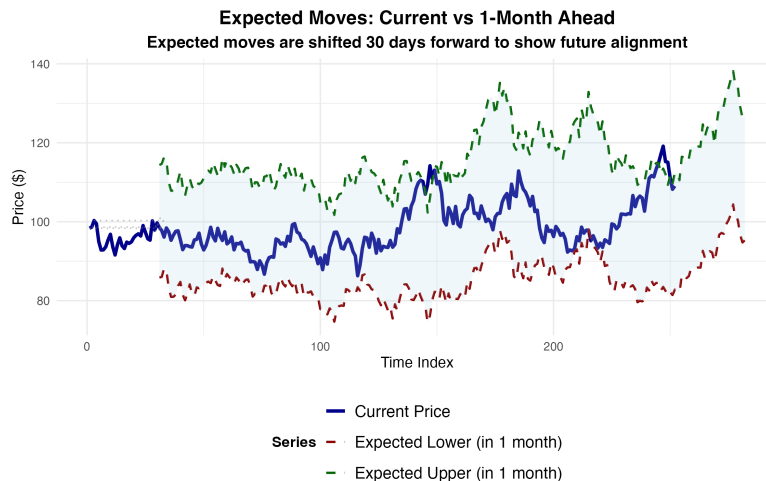


Figure 7.1: One-month expected-move volatility band for simulated asset price.

This section provides controlled experimental case studies to compare classical HMM inference (Baum–Welch estimation and Viterbi decoding) with the proposed topological HMM (THMM) modifications incorporating Onsager–Machlup (OM) based emissions. We focus on a two-regime switching Heston model in which the latent market state alternates between a calm regime and a turbulent regime for the volatility. Using both a baseline and a challenging configuration, we evaluate regime recovery accuracy, parameter estimation stability, and the implications for volatility diagnostics and option-implied quantities.

7.1.3 Results and Discussion

We now report inference performance, we compare two inferential approaches: (a) classical Hidden Markov Model (HMM) inference based on pointwise emission likelihoods, referred to as the Standard approach, and (b) the proposed Topological HMM (THMM) incorporating Onsager–Machlup functional–based path likelihoods.

Performance is assessed along three complementary dimensions: (i) regime classifi-

cation accuracy and timing via Viterbi decoding; (ii) segmentation stability, including run-lengths and transition coherence; and (iii) parameter recovery and uncertainty quantification via Baum–Welch estimation and associated confidence intervals. The confidence intervals begin as 5-day-ahead predictions immediately after each detected regime change. Subsequently, a new 5-day-ahead confidence interval is generated every 10 time steps. When regime switching occurs too rapidly, likely due to poor parameter estimation, the confidence intervals fail to form properly. Together, these diagnostics allow us to evaluate not only pointwise accuracy, but also the structural coherence of inferred regime dynamics.

Viterbi decoding with true parameters.

The plot presents regimes with similar parameter values between regimes, making inference substantially more difficult. Figure 7.2 shows that even with true parameters, the Standard HMM exhibits frequent regime switching and reduced segmentation stability. Regime classification accuracy deteriorates markedly, with short run-lengths and delayed detection of true regime changes.

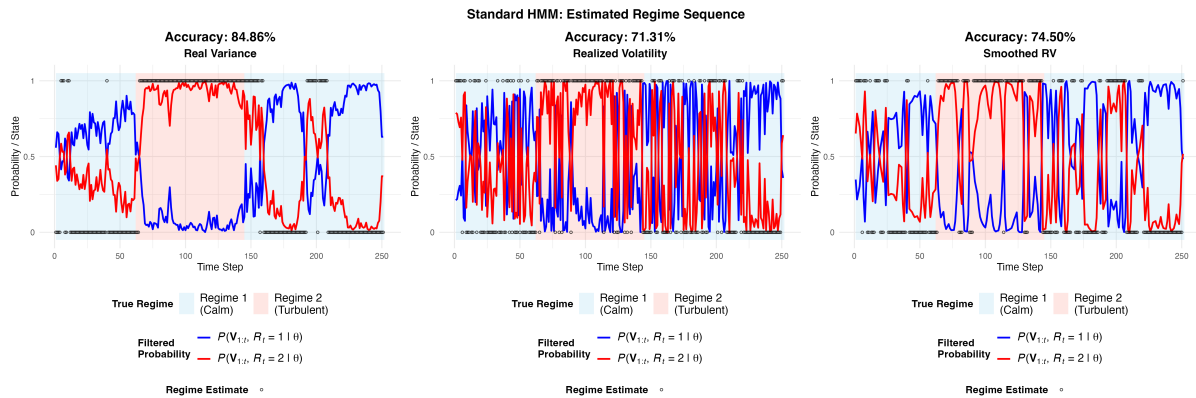


Figure 7.2: Standard HMM: Viterbi decoding with true parameters.

Baum–Welch parameter estimation.

Figure 7.3 reports Baum–Welch estimates under the Standard model. Parameter recovery is unstable, with substantial mixing between regimes and diminished interpretability of regime-specific dynamics. This highlights the limitations of pointwise emission likelihoods when regimes are weakly separated and observations are noisy.

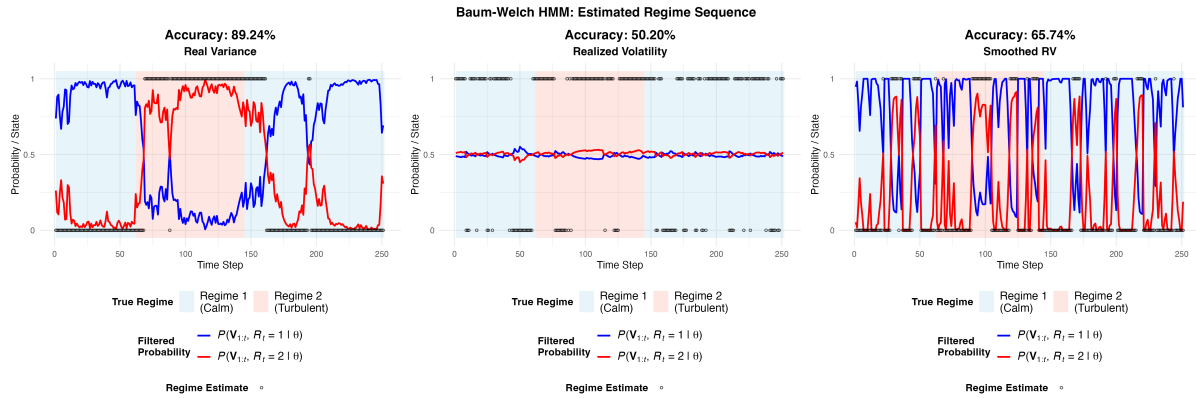


Figure 7.3: Challenging (Standard): Baum–Welch parameter estimation.

Baum–Welch confidence intervals.

As shown in Figure 7.4, confidence intervals widen substantially in the challenging case, often overlapping across regimes. This reflects both parameter uncertainty and frequent regime misclassification, reducing the practical usefulness of the inferred model.

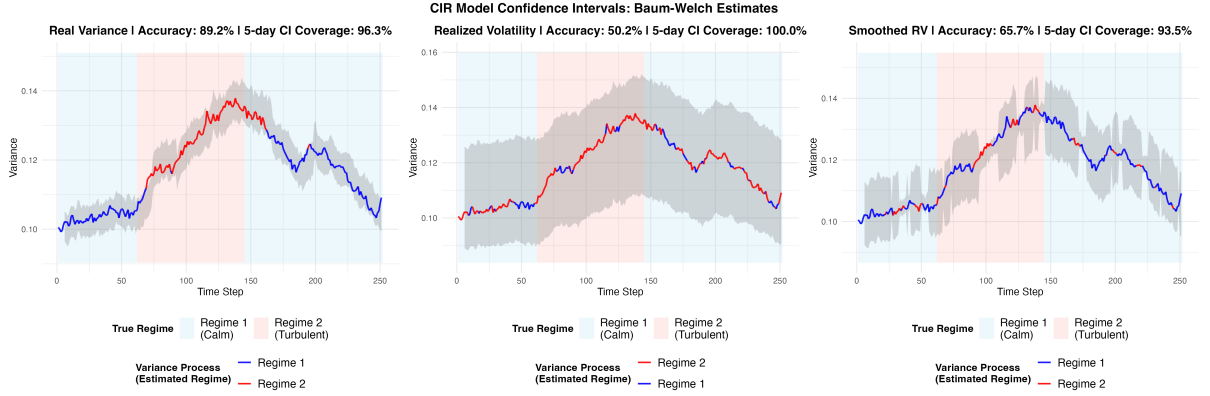


Figure 7.4: Challenging (Standard): Baum-Welch confidence intervals.

7.1.4 Challenging Case Experiment: OM-Functional THMM

Viterbi decoding with true parameters.

Figure 7.5 demonstrates that the OM-functional THMM substantially improves regime decoding in the challenging configuration. Although regimes remain difficult to distinguish, the inferred regime path exhibits longer run-lengths and improved alignment with the true latent states, indicating that pathwise likelihood information provides critical regularization.

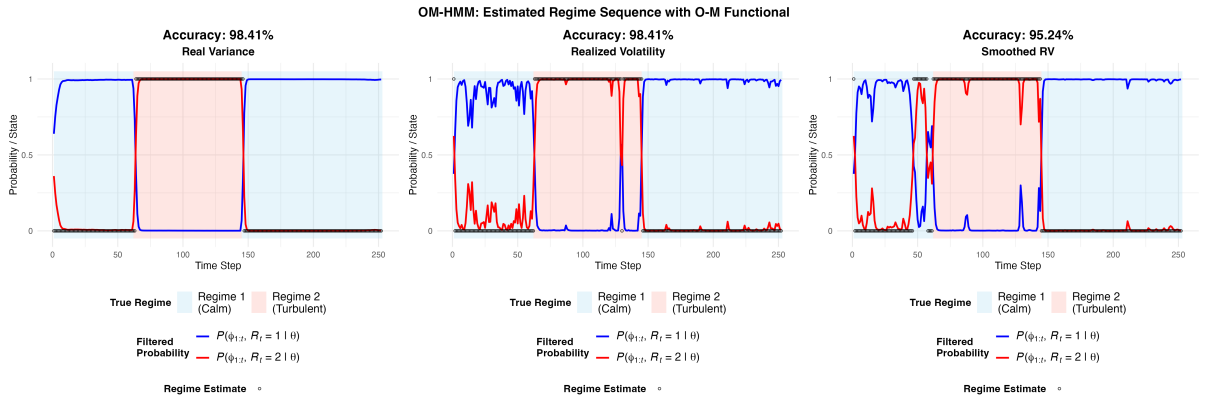


Figure 7.5: Challenging (OM): Viterbi decoding with true parameters.

Baum–Welch parameter estimation.

Figure 7.6 shows that OM-adjusted Baum–Welch estimation yields markedly more stable parameter trajectories than the Standard approach. While perfect regime separation is not attainable in this setting, the THMM mitigates excessive regime switching and improves interpretability of the inferred dynamics.

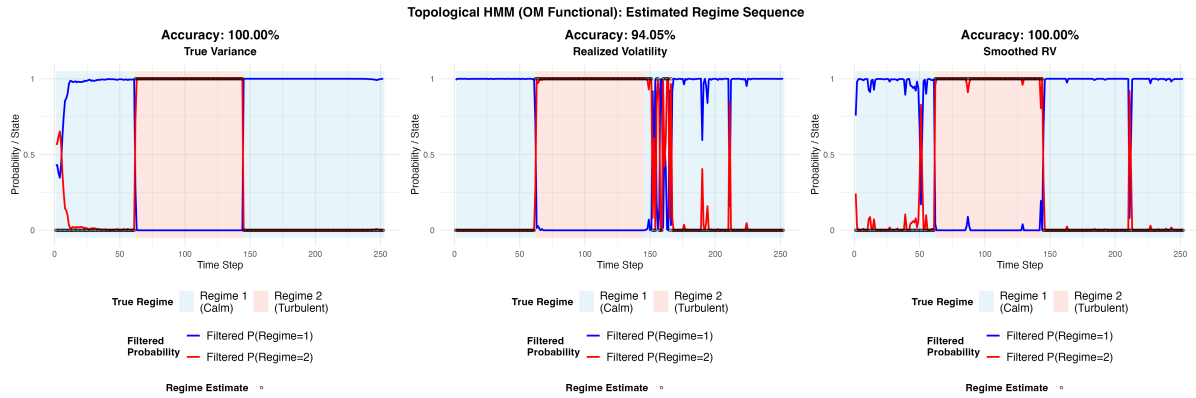


Figure 7.6: Challenging (OM): Baum–Welch parameter estimation.

Baum–Welch confidence intervals.

Finally, Figure 7.7 indicates that confidence intervals under the OM-functional framework remain narrower and more structured than their Standard counterparts. This confirms that incorporating OM-based path likelihoods enhances both regime detection and uncertainty quantification, particularly in empirically realistic scenarios where regimes are weakly separated.

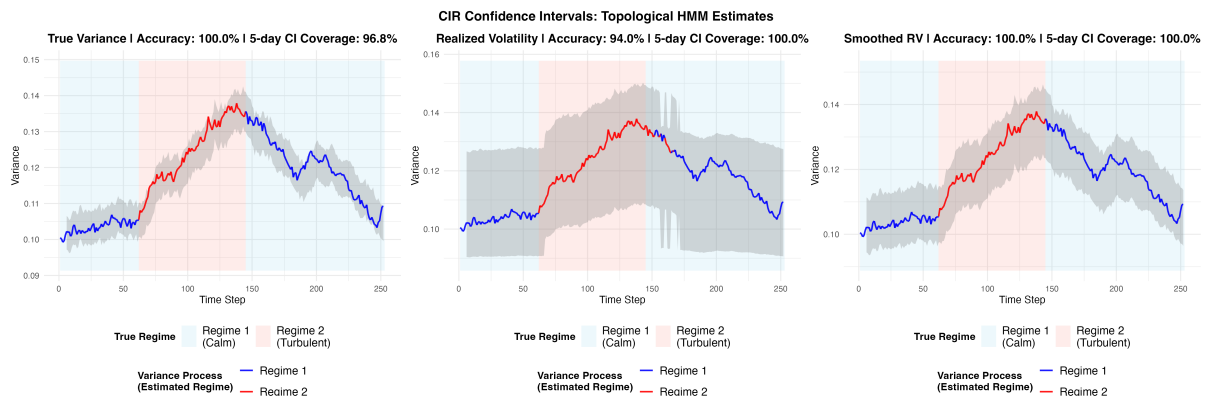


Figure 7.7: Challenging (OM): Baum–Welch confidence intervals.

7.2 Regime Estimation on Spot-Rate Modelling

In this section, we study regime estimation for US Treasury spot-rate dynamics under the asynchronous regime-switching CIR framework developed in Section 4.3. The estimation procedure follows the marginal Baum–Welch and Viterbi decoding framework introduced in Section 5.5.

Our objective is not to impose a single common latent regime on the entire yield curve, but rather to allow different segments of the curve to evolve under distinct regime configurations. Economically, short-end, intermediate, and long-end rates react to different macroeconomic and monetary-policy forces, and hence their regime transitions may occur at different times.

7.2.1 Empirical Setting for Treasury Spot-Rate Modelling

We consider Treasury yields across standard maturities, measured in years, $\{\frac{1}{12}, \frac{1}{4}, \frac{1}{2}, 1, 2, 3, 5, 7, 10, 20, 30\}$.

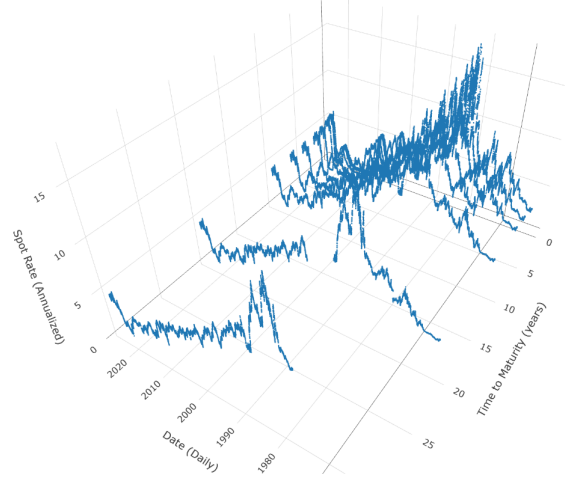


Figure 7.8: 3-D surface of daily US Treasury yields across maturities.

From the full dataset, we retain the most recent using daily US Treasury yield Board of Governors of the Federal Reserve System (US) (2024) data spanning the period from November 18, 2022 to November 22, 2024. $N_{\text{emp}} = 2 \times 252$ daily observations, corresponding approximately to two years of trading data.

To reduce the strong common level effect in yields and focus on relative term-structure dynamics, we construct spread processes between adjacent maturity segments:

$$x_t^{(1)} = y_t(2Y) - y_t(1Y), \quad (7.6)$$

$$x_t^{(2)} = y_t(5Y) - y_t(2Y), \quad (7.7)$$

$$x_t^{(3)} = y_t(10Y) - y_t(5Y), \quad (7.8)$$

representing short-end, intermediate, and long-end slope components of the Treasury curve.

Each spread process is modelled independently using a two-regime CIR-HMM specification, allowing regime persistence and volatility structure to differ across maturity segments. Throughout, we adopt the following labelling convention:

- Regime 1 (blue): calm or persistent regime, characterised by low volatility and stable mean-reversion;
- Regime 2 (red): turbulent or stressed regime, characterised by elevated volatility and faster mean-reversion.

7.2.2 Regime Decoding Results

Figure 7.9 presents the filtered regime probabilities together with the Viterbi-decoded regime sequence for each spread process.

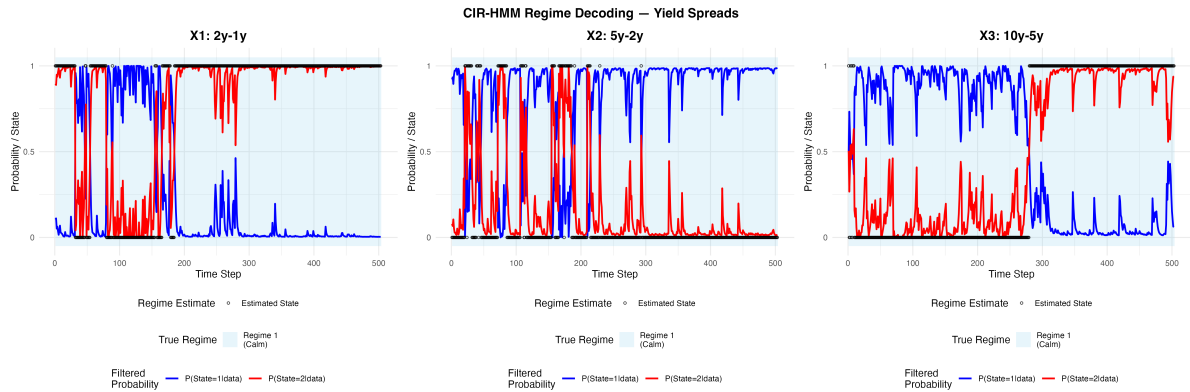


Figure 7.9: CIR-HMM regime decoding results for Treasury yield spreads. The blue and red curves denote filtered probabilities for the two latent regimes, while the black markers indicate the Viterbi-decoded state sequence.

Several features are apparent. First, the short-end spread X_1 exhibits relatively frequent switching behaviour, indicating higher sensitivity to short-term monetary-policy expectations and liquidity conditions. In contrast, the long-end spread X_3 displays substantially more persistent regime behaviour, reflecting slower-moving macroeconomic and term-premium dynamics.

Second, the decoded regimes are not synchronized across the three maturity segments.

Regime transitions in the short-end spread often occur earlier than those in the intermediate or long-end spreads, supporting the asynchronous regime-switching hypothesis developed in Section 4.3.

Finally, the filtered probabilities are frequently close to either zero or one, indicating strong regime separation and stable inference under the CIR-HMM specification.

7.2.3 Prediction and Confidence Interval Performance

To evaluate predictive performance, we construct 5-day-ahead prediction intervals using the estimated CIR-HMM dynamics. Figure 7.10 displays the realised spread paths together with the associated predictive confidence intervals.

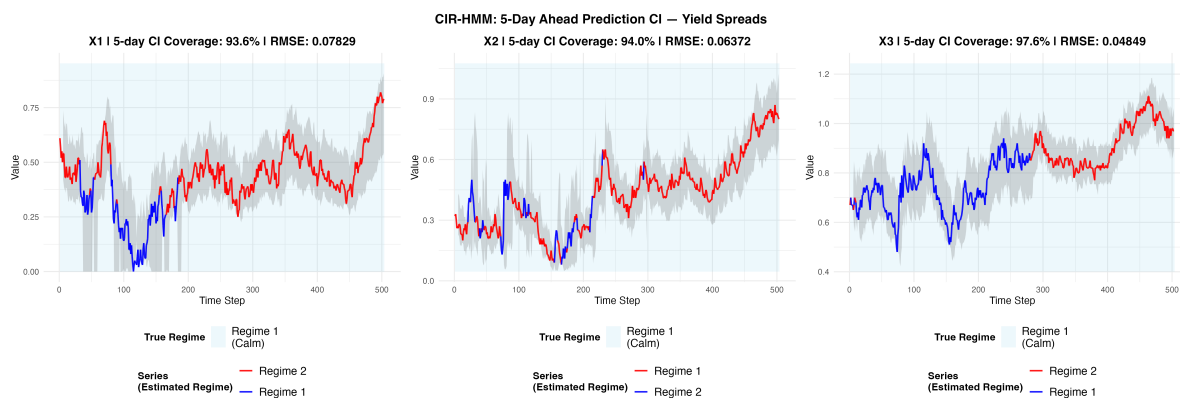


Figure 7.10: Five-day-ahead prediction intervals under the CIR-HMM model for Treasury yield spreads. The shaded grey region denotes the predictive confidence interval, while coloured segments indicate the decoded latent regime.

The empirical coverage rates are approximately 93.6%, 94.0%, and 97.6% for X_1 , X_2 , and X_3 , respectively, demonstrating that the model captures the conditional variability of the spread processes reasonably well. Moreover, prediction uncertainty expands during periods of regime transition, particularly in the short-end and intermediate segments. This reflects the increased volatility and parameter uncertainty associated with stressed

or rapidly changing market conditions. The long-end spread process exhibits the lowest RMSE among the three components, suggesting that longer-maturity slope dynamics are more persistent and easier to forecast under the regime-switching CIR specification.

An economically important feature of the empirical results is the apparent propagation of turbulent regimes across the term structure. Turbulent regimes first emerge in short-maturity spreads, reflecting the rapid response of short-end rates to monetary-policy expectations, inflation shocks, and liquidity conditions. Regime transitions then propagate toward intermediate maturities as markets adjust their expectations regarding future business-cycle dynamics and policy paths. Finally, long-maturity spreads react more gradually, since they are primarily driven by slower-moving components such as long-run inflation expectations and term premia. Overall, the results support an asynchronous regime-switching interpretation of the term structure, where macroeconomic shocks diffuse progressively from the short end toward longer maturities rather than affecting the entire yield curve simultaneously.

Chapter 8

Conclusion and Future Studies

This thesis studied regime-switching polynomial diffusion models and their associated hidden Markov representations, with particular emphasis on pathwise inference through the Onsager–Machlup (OM) functional. The work combined polynomial diffusion theory, regime-switching stochastic systems, and Topological Hidden Markov Models (THMMs) within a unified continuous-time inference framework motivated by applications in financial mathematics and stochastic modeling.

We first introduced the general framework of polynomial diffusions and reviewed their main structural properties, including polynomial closure, moment tractability, and finite-dimensional generator representations. These properties provide an analytically tractable class of diffusion processes that naturally includes many classical financial models such as the Ornstein–Uhlenbeck process, the CIR process, and the Heston stochastic volatility model.

Building on this foundation, we developed both univariate and asynchronous multivariate regime-switching polynomial diffusion models. In the asynchronous setting, different components of the multivariate system were allowed to evolve under distinct latent regime processes, providing a flexible framework for modeling heterogeneous structural

changes across financial factors. The hidden Markov structure associated with the latent regime dynamics was derived together with the corresponding likelihood factorization and Euler transition approximations.

The thesis then introduced Onsager–Machlup functionals as continuous-time pathwise analogues of negative log-likelihoods for diffusion trajectories. This formulation allows inference to incorporate local path behavior and trajectory information beyond classical pointwise likelihood methods. By replacing transition-density-based emissions with OM-based pathwise costs, we developed a Topological Hidden Markov Model framework suitable for regime-switching diffusion systems. An important feature of the proposed construction is that it preserves computationally tractable inference procedures, allowing standard hidden Markov algorithms such as forward–backward filtering, Baum–Welch estimation, and Viterbi decoding to remain applicable within the pathwise setting.

The proposed methodology was investigated through simulation studies and financial applications involving stochastic volatility and multivariate spot-rate models. Numerical results demonstrated that the OM-based framework improves regime separation, enhances sensitivity to structural changes, and provides more stable identification of latent regimes. In particular, the multivariate CIR spot-rate applications illustrated how pathwise hidden Markov inference may be combined with affine term-structure modeling to recover latent rate dynamics and regime-dependent yield-curve behavior.

Future work will focus on extending the proposed asynchronous regime-switching CIR spot-rate framework toward derivative-consistent calibration and higher-dimensional term-structure modeling. In particular, the framework may be combined with swaption and cap/floor implied volatility information to incorporate forward-looking market expectations into the OM-based THMM inference procedure. Additional directions include the study of multivariate dependence through Wishart-type diffusions, stochastic correlation structures, and scalable filtering methods for high-dimensional regime-switching

systems.

Overall, the results of this thesis demonstrate that the combination of polynomial diffusion theory, regime-switching dynamics, and Onsager–Machlup pathwise inference provides a flexible and tractable framework for continuous-time stochastic modeling. The proposed methodology contributes both theoretically and computationally to the study of hidden Markov inference in nonlinear diffusion systems and offers several promising directions for future research in stochastic analysis, financial mathematics, and statistical learning.

Appendix A

Existence and Uniqueness of the CIR Process

This appendix establishes existence and uniqueness in law for the martingale problem $\text{MP}(\mathbb{R}_+, \mathcal{A}, \varpi)$ associated with the CIR generator on $E = \mathbb{R}_+$. Throughout, we fix parameters $\kappa, \sigma \geq 0$ and $\theta \geq 0$ with $\kappa\theta \geq 0$. The infinitesimal generator of the CIR process

$$dX_t = \kappa(\theta - X_t) dt + \sigma\sqrt{X_t} dW_t, \quad X_0 \sim \varpi \text{ on } \mathbb{R}_+, \quad (\text{A.1})$$

is given by

$$\mathcal{A}f(x) = \kappa(\theta - x) f'(x) + \frac{\sigma^2}{2} x f''(x), \quad x \in \mathbb{R}_+. \quad (\text{A.2})$$

We work in the pure-diffusion setting ($\nu = 0$).

A.1 Auxiliary Lemmas

We first record two structural properties of $\mathcal{A}$ that are used to verify the hypotheses of the existence and uniqueness theorems.

Lemma A.1.1 (Polynomial structure). *The operator $\mathcal{A}$ is polynomial on $\mathbb{R}_+$, that is, $\mathcal{A}\text{Pol}_n(\mathbb{R}_+) \subseteq \text{Pol}_n(\mathbb{R}_+)$ for every $n \in \mathbb{N}$.*

Proof: According to the characterisation of Proposition 3.6 in Larsson (2017), it suffices in the pure-diffusion case to verify that the drift and diffusion coefficients satisfy

$$b(x) = \kappa(\theta - x) \in \text{Pol}_1(\mathbb{R}_+) \quad \text{and} \quad a(x) = \sigma^2 x \in \text{Pol}_2(\mathbb{R}_+).$$

Both conditions are immediate. To confirm directly, let $p(x) = \sum_{k=0}^n c_k x^k \in \text{Pol}_n(\mathbb{R}_+)$. A direct computation gives

$$\mathcal{A}p(x) = \kappa\theta \sum_{k=1}^n k c_k x^{k-1} - \kappa \sum_{k=1}^n k c_k x^k + \frac{\sigma^2}{2} \sum_{k=2}^n k(k-1) c_k x^{k-1},$$

and every term has degree at most n , so $\mathcal{A}p \in \text{Pol}_n(\mathbb{R}_+)$.

Lemma A.1.2 (Positive maximum principle). *The operator $\mathcal{A}$ satisfies $\text{PMP}(\mathbb{R}_+)$ in the sense of Definition 2.3.1.*

Proof: Let $f \in C_b^2(\mathbb{R})$ and suppose $x_* \in \mathbb{R}_+$ satisfies $f(x_*) = \max_{y \in \mathbb{R}_+} f(y) \geq 0$. We consider two cases according to whether x_* lies in the interior or at the boundary of $\mathbb{R}_+$.

If $x_* > 0$, then x_* is an interior maximum, so the standard first- and second-order optimality conditions yield $f'(x_*) = 0$ and $f''(x_*) \leq 0$. Consequently,

$$\mathcal{A}f(x_*) = \kappa(\theta - x_*) \underbrace{f'(x_*)}_{=0} + \frac{\sigma^2}{2} x_* \underbrace{f''(x_*)}_{\leq 0} \leq 0.$$

If $x_* = 0$, then $f(y) \leq f(0)$ for all $y > 0$, so

$$f'(0) = \lim_{y \downarrow 0} \frac{f(y) - f(0)}{y} \leq 0.$$

Since $a(0) = \sigma^2 \cdot 0 = 0$, we obtain

$$\mathcal{A}f(0) = \kappa\theta f'(0) \leq 0,$$

where the inequality uses $\kappa\theta \geq 0$ and $f'(0) \leq 0$.

In both cases $\mathcal{A}f(x_*) \leq 0$, so $\text{PMP}(\mathbb{R}_+)$ holds.

Remark A.1.3. *The two conditions $a(0) = 0$ and $b(0) = \kappa\theta \geq 0$ are the essential boundary requirements. The first ensures the diffusion vanishes at $\{0\}$, and the second ensures the drift points inward, so that both geometric mechanisms cooperate to keep the process in $\mathbb{R}_+$.*

A.2 Existence and Uniqueness

Theorem A.2.1 (Existence and uniqueness of the CIR process). *For every probability measure ϖ on $\mathbb{R}_+$, the martingale problem $\text{MP}(\mathbb{R}_+, \mathcal{A}, \varpi)$ admits a unique solution in law.*

Proof: Existence. We verify the three hypotheses of Theorem 2.3.2.

Hypothesis (1) requires $\text{PMP}(\mathbb{R}_+)$, which holds by Lemma A.1.2. Hypothesis (2), that $\mathcal{A}f$ is continuous on $\mathbb{R}_+$ and vanishes at infinity for every $f \in C_c^2(\mathbb{R})$, follows immediately from the continuity of the coefficients b and a and the compact support of f . For hypothesis (3), since $\nu = 0$ it suffices by the linear growth condition stated in

Remark 2.3.3 to exhibit a constant $C > 0$ such that $|b(x)| + \|a(x)\|^{1/2} \leq C(1 + |x|)$ for all $x \in \mathbb{R}_+$. Indeed,

$$|b(x)| = |\kappa(\theta - x)| \leq \kappa|\theta| + \kappa x \leq (\kappa|\theta| + \kappa)(1 + x),$$

and, using $\sqrt{x} \leq 1 + x$ for $x \geq 0$,

$$\|a(x)\|^{1/2} = \sigma\sqrt{x} \leq \sigma(1 + x).$$

Setting $C = \kappa|\theta| + \kappa + \sigma$ yields the required bound. All three hypotheses of Theorem 2.3.2 are therefore satisfied, and existence follows.

Uniqueness. We verify the three hypotheses of Theorem 2.3.4.

Hypothesis (1) is weak existence, which was just established. Hypothesis (2) requires that $\mathcal{A}$ preserves polynomial spaces, which holds by Lemma A.1.1. For hypothesis (3), we must show that the joint moment problem for the finite-dimensional distributions is determinate. The moment formula (Theorem 3.12 of Larsson (2017)), since $\mathcal{A}$ is polynomial, and iterating the formula shows that every joint moment $\mathbb{E}[X_{t_1}^{\alpha_1} \cdots X_{t_m}^{\alpha_m}]$ is uniquely determined by $\mathcal{A}$ and ϖ . The marginal of any solution at time t is a scaled non-central chi-squared distribution, whose moment generating function is finite in a neighbourhood of the origin. By Carleman's condition, the moments of a measure on $\mathbb{R}_+$ uniquely determine it whenever $\sum_{n=1}^{\infty} (m_n)^{-1/(2n)} = \infty$ with $m_n = \mathbb{E}[X_t^n]$. The linear growth bound above and Lemma 3.20 of Larsson (2017) together imply $m_n \leq (C')^n n!$ for some $C' > 0$, from which Carleman's condition follows. Hence all three hypotheses of Theorem 2.3.4 hold, and uniqueness in law follows.

Remark A.2.2 (Feller condition). *The classical Feller condition $2\kappa\theta \geq \sigma^2$ determines the accessibility of the boundary $\{0\}$: when it holds, $\{0\}$ is of entrance-not-exit type and $X_t > 0$ almost surely for all $t \geq 0$ whenever $X_0 > 0$; when it fails, the process may reach zero but remains non-negative. Theorem A.2.1 holds regardless of whether the Feller condition is satisfied, since neither Lemma A.1.2 nor the linear growth bound depends on it.*

Bibliography

- Alizadeh, S., Brandt, M. W., and Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3):1047–1091.
- Ang, A. and Bekaert, G. (2002). Regime switches in interest rates. *Journal of Business and Economic Statistics*, 20(2):163–182.
- Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *Annals of Mathematical Statistics*, 37(6):1554–1563.
- Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *Annals of Mathematical Statistics*, 41(1):164–171.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.
- Board of Governors of the Federal Reserve System (US) (2024). Market yield on u.s. treasury securities. <https://fred.stlouisfed.org/>. Retrieved from FRED, Federal Reserve Bank of St. Louis.
- Buffington, J. and Elliott, R. J. (2002). American options with regime switching. *International Journal of Theoretical and Applied Finance*, 5(5):497–514.
- Chong, C. H. and Todorov, V. (2026). Realized variance disagreement. Available at SSRN 6249218.
- Cochrane, D. and Orcutt, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, 44(245):32–61.
- Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.
- Cuchiero, C., Keller-Ressel, M., and Teichmann, J. (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4):711–740.

- Dai, Q. and Singleton, K. J. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55(5):1943–1978.
- Diebold, F. X. and Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364.
- Duffie, D., Filipović, D., and Schachermayer, W. (2003). Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053.
- Elliott, R. J., Aggoun, L., and Moore, J. B. (1995). *Hidden Markov Models: Estimation and Control*. Springer, New York.
- Filipović, D. (2009). *Term-Structure Models: A Graduate Course*. Springer.
- Filipović, D. and Larsson, M. (2016). Polynomial diffusions and applications in finance. *Finance and Stochastics*, 20(4):931–972.
- Freidlin, M. I. and Wentzell, A. D. (1984). *Random Perturbations of Dynamical Systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften*. Springer, New York.
- Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1):27–62.
- Guidolin, M. (2011). Markov switching models in empirical finance. Working Paper 415, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University. Accessed May 22, 2026.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.
- Kalimipalli, M. and Susmel, R. (2004). Regime-switching stochastic volatility and short-term interest rates. *Journal of Empirical Finance*, 11(3):309–329.
- Kashlak, A. B., Loliencar, P., and Heo, G. (2023). Topological hidden markov models. *Journal of Machine Learning Research*, 24(340):1–49.
- Kloeden, P. E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics*. Springer.
- Larsson, M. (2017). Notes on polynomial and affine jump-diffusions. Lecture notes, ETH Zurich.
- Litterman, R. and Scheinkman, J. (1991). Common factors affecting bond returns. *The Journal of Fixed Income*, 1(1):54–61.

- Machlup, S. and Onsager, L. (1953). Fluctuations and irreversible process. ii. systems with kinetic energy. *Physical Review*, 91(6):1512–1515.
- Madsen, H. and Lindström, E. (2010). From state-dependent diffusion to constant diffusion in stochastic differential equations by the lamperti transform. Technical report, Technical University of Denmark.
- Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183.
- Onsager, L. and Machlup, S. (1953). Fluctuations and irreversible processes. *Physical Review*, 91(6):1505–1512.
- Peters, G., Xu, M., Zhu, Z., and Shevchenko, P. V. (2026). Regime-switching polynomial diffusions via topological hidden markov model inference using onsager-machlup functionals for asset pricing. *SSRN Electronic Journal*. Available at SSRN: <https://ssrn.com/abstract=6130426>.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.
- Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2):260–269.
- Yin, G. and Zhu, C. (2010). *Hybrid Switching Diffusions: Properties and Applications*. Springer.
- Zeitouni, O. and Dembo, A. (1989). On the onsager–machlup functional of diffusion processes around non- c^1 curves. *Annals of Probability*, 17(3):1037–1054.