Stochastic Differential Equations and Strict Local Martingales

Lisha Qiu

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2018

ABSTRACT

Stochastic Differential Equations and Strict Local Martingales

Lisha Qiu

In this thesis, we address two problems arising from the application of stochastic differential equations (SDEs). The first one pertains to the detection of asset bubbles, where the price process solves an SDE. We combine the strict local martingale model together with a statistical tool to instantaneously check the existence and severity of asset bubbles through the asset's historical price process. Our approach assumes that the price process of interest is a CEV process. We relate the exponent parameter in the CEV process to an asset bubble by studying the future expectation and the running maximum of the CEV process. The detection of asset bubbles then boils down to the estimation of the exponent. With a dynamic linear regression model, inference on the exponent can be carried out using historical price data. Estimation of the volatility and calibration of the parameters in the dynamic linear regression model are also studied. When using SDEs in practice, for example, in the detection of asset bubbles, one often would like to simulate its paths using the Euler scheme to study the behavior of the solution. The second part of this thesis focuses on the convergence property of the Euler scheme under the assumption that the coefficients of the SDE are locally Lipschitz and that the solution has no finite explosion. We prove that if a numerical scheme converges uniformly on any compact time set (UCP) in probability with a certain rate under the globally Lipschitz condition, then when the globally Lipschitz condition is replaced with a locally Lipschitz one plus a no finite explosion condition, UCP convergence with the same rate holds. One contribution of this thesis is the proof of $\sqrt{n}$-weak convergence of the asymptotic normalized error process. The limit error process is also provided. We further study the boundedness for the second moment of the weak limit process and its running maximum under both the globally Lipschitz and the locally Lipschitz conditions. The convergence of the Euler scheme in the sense of approximating expectations of functionals is also studied under the locally Lipschitz condition.

Table of Contents

List of Figures

List of Tables

Acknowledgments

First I want to express my deepest gratitude to Professor Philip Protter, who has been a great advisor and mentor. I benefitted greatly from his wisdom in academic research, received tremendous help during our meetings. He has given me a lot of invaluable suggestions in both research and life. I could not imagine having a better advisor than Professor Protter, who allowed me to me find my own research interests and helped me to learn how to think independently. At the same time, he was always there to support me whenever I need his help. I also feel extremely grateful for his confidence and patience with me and the encouragement he has given me through out all my years as a graduate student. As a international student, the kindness he displayed to me helped me feel more at home in the department. Taking his class was also a very unforgettable and enjoyable experience, fulfilled with nice and clear explanations of math and good stories from the history of mathematics.

I would also like thank my committee members, Professor Bodhisattva Sen, Professor Marcel Nutz, Professor Jaime San Martin, and Professor Paul Glasserman for their kind help and insightful comments. I wish to thank all the professors and staff in the statistics department, especially Dood Kalicharan and Anthony Cruz for their consistent help.

I wish to give thanks to Professor Jaime San Martin. I worked closely with him in Columbia for one week, and he offered generous and important help to me on my second project, now the third chapter in the thesis. I also would like to thank Professor Denis Talay, discussions with him and his comments were very valuable. Thanks to professor Michael Sobel, who was always ready to help me during my years as a PhD student.

I would like to thank my dearest friends from Columbia and USTC, Jian Lu, Phyllis Wan, Jing Zhang, Yuanjun Gao, Chengliang Tang, Xinle Liu, Aditi Dandapani, Richard Neuberg, Wenkai Pan, Johannes Heiny, Haolei Wen, Yuting Ma, Rohit Patra, Yuyan Wang, Ziyuan Bai, Roseline Bilina-Falafala, Minghan Yan, Yi Zhang, Wenhao Liu and Qun Cao. You all are so warm and kind to me, making New York like a second home to me. I also want to thank Li Liu, Ruixin Zhao, Wensong Bai, Kevin Ge, Chenguang Hu and Jenny Chen. I am honored to be one of the first members for the NY Guqin Association with you and I enjoyed all the Yaji with you, which enriched my PhD life.

Lastly, and most importantly, I want to thank my parents Zhida Qiu and Jianhong Du, my grandmothers Xiuying Yuan and Huiying Song. Without their unconditional love and support, this work would never have been done. I would like to thank again to my grandma Xiuying, your love is the most precious treasure in my life.

To my parents

Chapter 1

Introduction

In this thesis, we study the application of SDEs in asset bubble detection and numerical schemes for solving SDEs taking the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x_0 \in \mathbb{R}, \tag{1.0.1}$$

where W is a standard Brownian motion. We assume $\mu(x)$ and $\sigma(x)$ satisfy the Engelbert-Schmidt conditions [19], which ensures a unique weak solution to the SDE above. In practice, for various purposes, one often requires that $\mu(x)$ and $\sigma(x)$ are of linear growth at most or globally Lipschitz. For example, in modeling asset price process with SDEs, assuming at most linear growth of coefficients leads to X being a martingale under a risk neutral measure. Under the globally Lipschitz assumption, numerical methods such as the Euler scheme for solving SDEs, have been proved to have different types of convergence properties. In this thesis, we go beyond the linear growth world. Chapters 2 and 3 are based on two working papers. Chapter 2 studies the detection of asset bubbles using SDEs and Chapter 3 studies the limit distribution of the asymptotic error process from the Euler scheme.

In Chapter 2, we focus on using SDEs to detect asset bubbles. From Mijatovic and Urusov [59], that when the coefficient $\sigma(x)$ goes to infinity with at least order $n^{1+\gamma}, \forall\gamma > 0$ as x goes to infinity, X is a strict local martingale under a risk neutral measure. If X is the price process of an asset, this indicates the existence of an asset bubble, (see Jarrow, Kchia and Protter[41], Jarrow, Protter and Shimbo [43],[44],[45], and Protter [66] for details). In Section 1 of Chapter 2, we review existing methods for detecting asset bubbles. In Section 2, we give definitions for the fundamental price process and the bubble process. Then we introduce the concept of a strict local martingale and describe how it can be used to model asset bubbles in a finite time horizon. Section 3 studies the constant elasticity variance (CEV) model, our main model for the stock price process. We show that the exponent parameter of the CEV model is linked to financial bubbles in two respects. First, when the exponent exceeds a fixed threshold, it indicates the existence of a bubble. Secondly, a larger exponent parameter leads to a more extreme asset bubble in the sense of smaller future expectation of the asset price process and a higher probability for the running maximum to hit a large value in a certain time range. In Section 4, we propose the usage of the dynamic linear regression method with a prior distribution on the noise terms to estimate the exponent parameter in the CEV model. We also address the problem of volatility estimation and focus on the Florens-Zmirou estimators. The details of using the Monte Carlo method to estimate the parameters is also studied in Section 4. Section 5 illustrates the asset bubble detection technique by applying it to several stocks from the alleged internet dot-com bubble.

In short, we have two innovations. One is that instead of assuming volatility $\sigma(x)$ as a functional of the price stays the same for the observation time period, we allow it to change over time to gain more model flexibility. The second is that we link the exponent parameter in the CEV model to the existence and severity of asset bubbles and propose a statistical method to estimate the exponent parameter instantaneously using historical price data up to the current time.

The motivation for the work in Chapter 3 started from the work in Chapter 2. It is known that a price process has an asset price bubble if and only if it is a nonnegative strict local martingale under a risk neutral measure. To study price processes with bubbles, we would like to simulate the paths of SDEs (1.0.1) with $\mu = 0$, and $\sigma(x)$ being superlinear as their solutions can be nonnegative strict local martingales. Since a nonnegative strict local martingale is a supermartingale, its future expectation is decreasing with time. However, when we use the Euler scheme to solve SDE (1.0.1) with $\mu = 0$, the numerical solution is a true martingale. The future expectation of a martingale does not decrease with time indicating that the Euler scheme diverges in L^1. In existing work on numerical schemes, the globally Lipschitz or at most linear growth condition is often assumed to ensure different types of convergence, including L^p convergence. The example of the nonnegative strict local martingale is one piece of evidence that in practice, this is often too stringent a requirement to meet. Many Brownian motion driven SDEs used in applications have coefficients which are only Lipschitz on a compact sets, for example, superlinear continuous functions. But the solutions to such SDEs can be arbitrarily large. This leads us to ask while L^1 convergence does not hold for the Euler scheme with only the locally Lipschitz condition, can other types of convergence hold? e.g., convergence in probability or convergence in distribution? And if so, at what rate does the Euler scheme converge? This is answered in Chapter 3. In Section 3 of Chapter 3, we prove that if a numerical scheme converges uniformly in probability on any compact time interval with a certain rate under the globally Lipschitz condition, the same result holds when the globally Lipschitz condition is replaced with a locally Lipschitz one and a no finite explosion condition. Convergence in probability for the Euler and the Milstein schemes are studied as examples. Beginning with Section 4, we focus on the Euler scheme. We prove the sequence of the normalized error process from the Euler scheme with normalizing coefficient as $\sqrt{n}$ is relatively compact. Furthermore, by proving uniqueness of the limit process, we have the asymptotic normalized error process converges in distribution. The limit error process is also provided as a solution to an SDE. Section 5 turns to the study of the the second moment of the limit error process and its running maximum. In Section 6 of Chapter 3, we give an upper bound for the rate of convergence in the sense of approximating expectations of functionals of the Euler scheme under the locally Lipschitz assumption.

Chapter 4 summarizes the results in this thesis and presents several interesting ideas for future work.

Chapter 2

Detecting Asset Bubbles with the Strict Local Martingale Model

Abstract

Using local martingales to model asset price processes, the detection of asset price bubbles is equivalent to detecting whether or not the price process is a strict local martingale under a risk neutral measure. In this paper, we model asset price processes with the CEV (constant elasticity of variance) model with time varying parameters. Some mathematical properties of the CEV processes are studied and linked to the severity or size of asset bubbles. The dynamic linear regression method is described for instantaneously detecting asset bubbles by estimating the exponent parameter in the CEV processes from historical asset price data. Applications in detecting asset bubbles in the dot-com bubble era are presented.

Keywords: stochastic differential equation, strict local martingale, asset bubble, constant elasticity of variance model, dynamic linear regression

2.1 Introduction

Financial bubbles have a long history. The first documented bubble is Tulipmania, which occurred in Amsterdam in the 17th century. More recently, from 2000 to 2002, the U.S. market experienced the dot-com bubble and in 2008 the housing bubble. It is of intrinsic interest to investigate the causes of financial bubbles, and there is a wealth of economic literature on the subject ([20],[21],[24],[25], [31],[34],[53],[58],[71],[72],[74],[79]). However studying the causes of financial bubbles is not the purpose of this paper. Instead, we aim to propose a statistical model to analyze historical prices and to determine instantaneously whether or not a bubble is occurring and how extreme the bubble is, regardless of its origins.

The detection of financial bubbles has received a great deal of attention. Instead of presenting a thorough survey, we discuss five existing methods for bubble detection. The first is proposed in the papers of Jarrow and Madan [42], Gilles [27], and Gilles and Lerov [28]. To explain the method, we require the technical concept of a price operator. Let $\psi = (\Delta, E^{\nu})$ denote the payoff of an asset, where Δ represents the asset's cumulative dividend process and E^{ν} is a nonnegative random variable representing the asset's terminal payoff at fixed time ν. The market price operator $F_t : \psi \Rightarrow R^+$ is a function mapping from payoff ψ to a nonnegative price. The existence of asset bubbles is linked to the concept of countable additivity for price operators, meaning the price system is the sum or integral of the values of the individual components. The main theorem used to detect asset bubbles is that for a fixed time t, a price operator F_t is countably additive, if and only if bubbles do not exist.

The second approach is that of Caballero et al [12], who, as described by Phillips et al [62], proposed a simple general equilibrium model without monetary factors, but with goods that may be partially securitized.

The third approach builds upon the second approach using the recursive implementation of a right-side unit root test. Let x_t be log stock price or log dividend:

$$x_t = \mu_x + \delta x_{t-1} + \Sigma_{j=1}^{J} \psi_j \Delta x_{t-j} + \epsilon_{x,t}, \quad \epsilon_{x,t} \sim N(0, \sigma_x^2) \tag{2.1.1}$$

The detection of financial bubbles reduces to the problem of testing whether or not $\delta > 1$, and the model (2.1.1) is estimated repeatedly, using subsets of the sample data incremented by one observation at each time.

The fourth approach is due to Sornette and co-authors ([8], [73], [69], [35] and [46]). The key feature of their financial model is that log-periodic oscillations appear in the price of the asset before the critical date t_c, which is the bubble crash date. Let p be the price process before the critical date. The price process evolves as

$$p(t) \approxeq p_c - \frac{\kappa}{\beta}[B_0(t-t_0)^{\beta} + B_1(t_c-t)^{\beta} cos[\omega log(t_c-t) - \phi]].$$

The crash as a point process has intensity

$$h(t) \approxeq B_0(t_c-t)^{-\alpha} + B_1(t_c-t)^{-\alpha} cos[\omega log(t_c-t) - \psi'].$$

Then a bubble exists when the crash hazard rate accelerates with time.

Recent work on asset price bubbles is based on arbitrage-free martingale pricing technology for Brownian driven asset price processes. Using this framework, Jarrow, Kchia and Protter[41], Jarrow, Protter and Shimbo [43],[44],[45], and Protter [66] propose a new approach. The key theorem they use is that for a nonnegative price process, the existence of a bubble in a finite time horizon equals that the price process being a strict local martingale under a risk neutral measure. Given the price process of a risky asset that follows a stochastic differential equation under a risk neutral measure taking the form

$$dX_t = \sigma(X_t)dB_t, \quad B_t \text{ is a standard Brownian motion,}$$

the condition on $\sigma(x)$ such that X_t is a strict local martingale is given in Mijatovic and Urusov [59]. The main method they used in [41] to estimate $\sigma(x)$ is by reproducing kernel spaces, which extends the estimator for $\sigma(X)$ with the values where the price process X is observed.

This paper extends the fifth approach. In Section 2, we give definitions for the fundamental price process and the bubble process. Then we introduce the concept of a strict local martingale and how it can be used to model the asset bubbles in a finite time horizon. Section 3 studies the constant elasticity variance (CEV) model, our main model for the stock price process. We show that the exponent parameter of the CEV model is linked to financial bubbles in two respects. First, when the exponent exceeds a fixed threshold, it indicates the existence of a bubble. Secondly, a larger exponent parameter leads to a more extreme asset bubble in the sense of smaller future expectation of the asset price process and a higher probability for the running maximum to hit a large value in a certain time range. In Section 4, we propose the usage of the dynamic linear regression method with a prior distribution on the noise terms to estimate the exponent parameter in the CEV model. We also address the problem of volatility estimation and focus on the Florens-Zmirou estimators. The details of using the Monte Carlo method to estimate the parameters is also studied in Section 4. Section 5 illustrates the asset bubble detection technique by applying it to several stocks from the alleged internet dot-com bubble.

In short, we have two innovations. One is that instead of assuming volatility $\sigma(x)$ as a functional of the price stays the same for the observation time period, we allow it to change over time to gain more model flexibility. The second is that we link the exponent parameter in the CEV model to the existence and severity of asset bubbles and propose a statistical method to estimate the exponent parameter instantaneously using historical price data up to the current time.

2.2 A Mathematical Definition of Bubbles

Following Protter [66], we begin with a complete probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$ satisfying the usual hypotheses, (see Protter [?] for details of usual hypotheses). Let $r = (r_t)_{t\geqslant 0}$ be at least progressively measurable, and denote the instantaneous defaultfree spot interest rate. Define A_t as

$$A_t = \exp(\int_0^t r_u d_u),$$

which is the time t value of a money market account. We work on the time interval $[0, T^\star]$, where $T^\star$ is assumed to be a finite fixed time in this paper. The assumption that $T^\star$ is finite is reasonable, since we can always make $T^\star$ large enough to include the lifetime of a risky asset that we are interested in. Let τ be the lifetime of the risky asset, where τ is a stopping time and $\tau < T^\star$. Let D_t be the dividend process, and S_t the nonnegative price process of the risky asset. Both D_t and S_t are assumed to be semimartingales. Since S_t has càdlàg paths [1], it represents the price process ex-cash flow. By ex-cash flow, we mean that the price at time t is after all dividends have been paid, including the time t dividend. Let $\Delta \in \mathcal{F}_t$ be the time τ terminal payoff or liquidation value of the asset, and $\Delta \geqslant 0$. Finally, let W be the wealth process associated with the market price of the risky asset plus accumulated cash flows. Then,

$$W_t = I_{t<\tau} S_t + A_t \int_0^{t\wedge\tau} \frac{1}{A_u} dD_u + \frac{A_t}{A_\tau} \Delta I_{\tau\leqslant t}, \tag{2.2.1}$$

with all the cash flows invested in the money market account. In the standard setting, one often assumes that the market is arbitrage-free in the sense of *No Free Lunch with Vanishing Risk (NFLVR)*, (see Delbaen and Schachermayer [17] for details). *NFLVR* guarantees the existence of a local martingale measure Q, with the same null sets as P (we write $Q \sim P$), such that under Q, the wealth process W is a local martingale. Q is often called a risk neutral measure. We use this risk neutral measure to calculate the market's fundamental value for the risky asset; this should be the best guess for the future discounted cash flows, given one's knowledge at the present time.

[1] Càdlàg paths refer to paths that are right continuous with left limits almost surely.

Definition The market's ***fundamental value process*** for the risky asset is defined as

$$S_t^\star = \mathbb{E}_Q\left(\int_t^\tau \frac{1}{A_u} dD_u + \frac{\Delta}{A_\tau} I_{\tau \leqslant T^\star} | \mathcal{F}_t\right) A_t.$$

With this definition, we are now able to define the bubble process.

Definition Let $\tilde{S}_t$ be the difference between the market price process and the fundamental price process:

$$\tilde{S}_t = S_t - S_t^\star.$$

Then the process $\tilde{S}_t$ is called the ***bubble process***.

Protter [66] studied the characterization of the bubble process $\tilde{S}_t$ under the assumption that the stock pays no dividends and the interest rate is 0. The bubble process $\tilde{S}_t$ is a nonnegative process and a price bubble exists when $\tilde{S}_t$ is not identically 0. Protter [66] identified three possible ways in which a bubble can occur:

Theorem 2.2.1. *(Protter) The existences of a bubble in an asset's price occurs only under 3 possibilities.*

1. If $\mathbb{P}(\tau = \infty) > 0$, then β_t is a local martingale (which could be a uniformly integrable martingale).

2. If τ is unbounded, but with $\mathbb{P}(\tau < \infty) = 1$, then β_t is a local martingale, but not a uniformly integrable martingale.

3. If τ is a bounded stopping time, β_t is a strict $Q-$local martingale.

For a proof, see Protter [66]. Of the three situations above, the third is most interesting, since we are working on the compact time interval. As seen in Theorem 2.2.1, a bubble exists on $[0, T]$, if and only if the price process S_t is a strict local martingale, and a bubble does not exist if and only if S is a martingale under a risk neutral measure. Let us make the reasonable assumption that S is the unique strong solution of a SDE

$$\begin{aligned} dS_t &= \sigma(S_t)dW_t + \mu(S_t, \nu_t)dt, \\ d\nu_t &= s(\nu_t)dB_t + g(\nu_t)dt, \end{aligned} \tag{2.2.2}$$

where W_t and B_t are two independent standard Brownian motions. This gives a model for S_t in the context of an incomplete market. Under mild conditions on σ and μ, (see Protter [66]) , there exists a risk neutral measure Q under which (2.2.2) reduces to

$$dS_t = \sigma(S_t)dW_t, \tag{2.2.3}$$

and the SDE (2.2.3) does not depend on which risk neutral measure is chosen, though there might be infinitely many of them. The condition under which S_t is a strict local martingale is studied in Mijatovic and Urusov [59].

Theorem 2.2.2. *(Mijatovic and Urusov) Let W be a Brownian motion and suppose S follows the SDE below under a measure Q*

$$dS_t = \sigma(S_t)dW_t, \quad S_0 > 0, \tag{2.2.4}$$

where $\sigma : (0, +\infty) \to \mathbb{R}$ is a Borel function satisfying the Engelbert-Schmidt conditions

$$\begin{aligned} &\sigma(x) \neq 0, \quad \forall x \in J, \quad J = (0, +\infty), \\ &\frac{1}{\sigma^2} \in L^1_{loc}(J) \quad \textit{is integrable on compact subsets of } J. \end{aligned}$$

*Assume that S is stopped at its hitting time of 0. Then S is a strict local martingale if and only if for some $a > 0$, the following condition **H** holds*

$$\textbf{\textit{H:}} \quad \int_a^\infty \frac{x}{\sigma^2(x)} dx < \infty. \tag{2.2.5}$$

Theorem 2.2.2 indicates that for a price process S following (2.2.4) under a risk neutral measure, detecting the existence of a price bubble is equivalent to checking the condition

$\boldsymbol{H}$, which can be realized when $\sigma(x)$ is known or properly estimated.

2.2.1 An Example of a Strict Local Martingale

It is not so straightforward to see that a nonnegative diffusion without a drift term can be a strict local martingale whose future expectation strictly decreases with time. We give one famous example, the inverse Bessel process. Let W be a standard three dimensional Brownian motion starting from the point (1,0,0). Define a process X by

$$X_t = \frac{1}{\|W_t\|}, \quad t \geqslant 0.$$

Then X is a nonnegative process with finite values almost surely, since, with probability 1, W never hits the origin. It is known that an alternate representation for the inverse Bessel process is as a solution to a SDE of the form

$$X_t = X_t^2 dB_t, \quad X_0 = 1,$$

where B_t is a standard Brownian motion. This is a special case of the CEV process, the model assumed in later parts of this chapter. Applying Theorem 2.2.2, X_t is a strict local martingale. The mean of X_t can be calculated explicitly as $\mathbb{E}(X_t) = 2\phi(\frac{1}{\sqrt{t}}) - 1$, where ϕ is the cumulative distribution function of the standard normal distribution. Thus, $\mathbb{E}(X_t)$ is decreasing with time and $\lim_{t \to +\infty} \mathbb{E}(X_t) = 0$.

2.3 The Constant Elasticity Variance (CEV) Model and Its Relation to Asset Bubbles

2.3.1 The CEV Model

It is always assumed herein that the asset price process follows the SDE (2.2.2). In order to apply Theorem 2.2.2, we impose a parametric form restriction on $\sigma(x)$ such that we can estimate it. If we assume $\sigma(x)$ is a power function, then CEV model developed by John Cox [13] is obtained. The CEV model is the main model used here for detecting asset bubbles. Under the model, the following deterministic relationship between price and volatility

$$\sigma(S_t, t) = \delta S_t^{\frac{\theta}{2}}, \quad \theta > 0,$$

is assumed to hold. Here the exponent parameter is in the fraction form for notational convenience in later sections. Then the price process follows the SDE below

$$dS_t = \mu S_t dt + \delta S_t^{\frac{\theta}{2}} dW_t, \quad \mu = r - q, S_{t_0} = s_0, \tag{2.3.1}$$

where W_t is a standard brownian motion, r is the interest rate and q is the dividend rate. It is known that there exists a risk neutral measure Q, under which the drift term in (2.3.1) is removed, and S follows

$$dS_t = \delta S_t^{\frac{\theta}{2}} dW_t. \tag{2.3.2}$$

By applying condition $\boldsymbol{H}$ in Theorem 2.2.2, we have that for $\theta > 2$, the price process S_t is a strict local martingale and an asset bubble exists. When $\theta \leq 2$, S is a martingale and there is no asset bubble. Here we assume the price process follows the CEV model. The model is often used by practitioners in the financial industry, especially for modeling equities and commodities. Besides the wide acceptance and simplicity of the CEV model, the exponent parameter θ is related to the severity or scale of asset bubbles, which is discussed subsequently.

Theorem 2.2.2 indicates that the behavior of $\sigma(x)$ when $x \to \infty$ determines whether S_t is a martingale or not, thus the existence of an asset bubble. In the CEV model, $\sigma(x)$ is restricted to be a power function, but it is able to capture the tail behavior of a wider range of diffusion functions, namely those which are continuous and locally bounded. This can be shown by applying the Stone-Weierstrass approximation theorem.

Theorem 2.3.1 (Stone-Weierstrass)**.** *Let X be a compact Hausdorff space $\mathcal{C}(X, R)$, which* *is the space of all bounded continuous real functions defined on X, and let A be a closed subalgebra which separates points and contains a non-zero constant function of $\mathcal{C}(X, R)$. Then A equals $\mathcal{C}(X, R)$.*

Take X as any closed interval $[a, b]$, and let A be the closure of all functions of the form $p(x) = a_0 + a_1 x^{b_1} \dots a_n x^{b_n}$, where $a_i \in \mathbb{R}\backslash\{0\}$ and $b_i \in \mathbb{R}^+$. The Stone-Weierstrass theorem states that every continuous function defined on a closed interval $[a, b]$ can be uniformly approximated as closely as desired by functions in the form of $p(x)$. And for $p(x) = a_0 + a_1 x^{b_1} \dots a_n x^{b_n}$, with b_n as the largest power coefficient, its tail behavior is the same as $\tilde{p}(x) = a_n x^{b_n}$.Therefore in terms of analyzing the tail behavior of $p(x)$ for detecting asset bubbles, the CEV model with a simple power function form for $\sigma(x)$ is a good choice. If the coefficients of the smaller exponents are exceptionally large, the bubble affect could be masked. But if all the coefficients are reasonable, the CEV model assumption should work.

2.3.2 Future Expectations under the CEV model

Assuming that a price process S follows the CEV model, the mathematical properties of S have been extensively studied in the situation that S is a martingale under a risk neutral measure. However to the best of our knowledge, there are not many papers considering the possible existence of asset bubbles while applying the CEV model. Theorem 2.2.2 implies that under the CEV model, S being a strict local martingale if and only if $\theta > 2$. In practice, it is often assumed that $\theta \leq 2$ to ensure that S is a martingale, thus excluding the possibility of asset bubbles.

In this subsection, we are going to study the mathematical properties of S with the assumption that $\theta > 2$ i.e., under the assumption that an asset bubble exists. Two features of the price process S that are of interest are the future expectation $\mathbb{E}(S_t)$ and the running maximum $S_T^* = \sup\limits_{0 \leq s \leq T} S_s$. They are not only related to the return and risk of the price process, but also closely linked to the behavior of asset bubbles, as a price process with a bubble has a decreasing future expectation and a high possibility of a soaring period, hitting a relatively high value. It is well known that if S is a nonnegative strict local martingale, one has $\mathbb{E}(S_s|\mathcal{F}_t) < S_t, \forall s > t > 0$. The decreasing expectation of a nonnegative strict local martingale causes the bubble process $\tilde{S}$ not to be identically 0, and also implies that the expected future return of an asset with a bubble is negative under a risk neutral measure. We will prove that if S follows the CEV model with $\theta > 2$, larger values of θ result in more extreme asset bubbles. By more extreme bubbles, we mean a lower value of $\mathbb{E}(S_t)$ and higher probability of hitting a high price level in a short time range.

To study $\mathbb{E}(S_t)$, we relate S_T to the distribution of a noncentral chi-squared distribution as in Emmanuel and MacBeth [16]. The expectation of a CEV process was studied in S.Mark [57]. However the form provided by Mark was incorrect for the case $\theta > 2$. Proposition 2.3.1 below corrects this.

Proposition 2.3.1. *Assume that a nonnegative process S_t follows the SDE (2.3.1). Define a, b, c, λ by*

$$\begin{aligned}
a &= \frac{1}{2}\delta^2(\theta-2)^2, \\
b &= \mu(2-\theta), \\
c &= \frac{1}{2}\delta^2(\theta^2-3\theta+2), \\
\lambda &= \frac{2{s_0}^{2-\theta}be^{b\tau}}{a(e^{b\tau}-1)}, \\
df &= \frac{2c}{a}
\end{aligned}$$

where $\tau = T - t_0 > 0$. Let $F[x; k, d]$ be the cumulative distribution function of a noncentral chi-squared distribution with degree of freedom k (k is a positive real number, and can be non integer) and the noncentrality parameter d. Then, when $\theta > 2$,

$$\mathbb{E}(S_T|S_{t_0} = s_0) = s_0 e^{\mu\tau} F\Big[\lambda; df-2, 0\Big],$$

Proof. We first show that $Z = \lambda e^{-b\tau}(\frac{S_T}{s_0})^{2-\theta}$ follows a noncentral chi-squared distribution

$\chi^2_{df,\lambda}$. Let $Y_t = S_t^{2-\theta}$. Applying Itô's formula to Y_t gives

$$dY_t = [\mu(2-\theta)Y_t + \frac{1}{2}\delta^2(\theta^2 - 3\theta + 2)]dt + \delta(2-\theta)\sqrt{Y_t}dW_t,\ Y_{t_0} = s_0^{2-\theta}.$$

Let $p(y_t, t)$ be the probability density function of process Y at time $t > t_0$ with value y_t. Then by the Fokker-Planck equation, $p(y_t, t)$ satisfies the partial deferential equation

$$\frac{\partial}{\partial t}p(y_t,t) = -\frac{\partial}{\partial y}\big[\mu_Y(y_t,t)p(y_t,t)\big] + \frac{\partial^2}{\partial y^2}\big[D_Y(y_t,t)p(y_t,t)\big], \tag{2.3.3}$$

where the drift coefficient μ_Y and the diffusion coefficient D_Y are

$$\mu_Y(y_t,t) = \mu(2-\theta)y_t + \frac{1}{2}\delta^2(\theta^2 - 3\theta + 2),$$
$$D_Y(y_t,t) = \frac{1}{2}\delta^2(\theta-2)^2 y_t.$$

Using Feller's result (see [22]), given the initial condition that $Y_{t_0} = s_0^{2-\theta}$, the Laplace transform of the probability density $p(y_T, T)$ is

$$\omega_Y(\tau, s) = \Big[\frac{b}{sa(e^{b\tau}-1)+b}\Big]^{c/a} \exp\Big\{\frac{-s_0^{2-\theta}sbe^{b\tau}}{sa(e^{b\tau}-1)+b}\Big\}, \quad \tau = T - t_o.$$

To get the density function $p(y_t, t)$, instead of applying the inverse Laplace transformation to $\omega_Y(\tau, s)$, we relate $\omega_Y(\tau, s)$ to the noncentral chi-squared distribution. Since $Z = \lambda e^{-b\tau}(\frac{S_T}{s_0})^{2-\theta} = \frac{2bS_T^{2-\theta}}{a(e^{b\tau}-1)}$, the Laplace transform of the probability density of Z is

$$\omega_Z(\tau, s) = \omega_Y(\tau, \frac{2bs}{a(e^{b\tau}-1)}) = (1+2s)^{-\frac{c}{a}} \exp\Big\{\frac{2s_0^{2-\theta}be^{b\tau}}{a(e^{b\tau}-1)(1+2s)}\Big\}.$$

It can be shown that $\omega_Z(\tau, s)$ equals the Laplace transform of a noncentral chi-squared distribution, which is

$$w_X(s) = (1+2s)^{-\frac{df}{2}} \exp(\frac{\lambda}{1+2s}), \quad X \sim \chi^2_{df,\lambda}.$$

Since the Laplace transform uniquely determines a probability distribution, one has $Z \sim$

$\chi^2_{df,\lambda}$. Let $p[z;df,\lambda]$ be the density function for the distribution $\chi^2_{(df,\lambda)}$,

$$f_{\chi^2_{(df,\lambda)}}(z) = p(z;df,\lambda) = \frac{1}{2}e^{\frac{z+\lambda}{2}}(\frac{z}{\lambda})^{\frac{df}{4}-\frac{1}{2}}\mathbb{I}_{\frac{df}{2}-1}(\sqrt{\lambda z}), \tag{2.3.4}$$

where $\mathbb{I}_\nu(\cdot)$ is the modified Bessel function of the first kind of order ν. Then by $S_T = s_0 e^{\mu\tau}(\frac{Z}{\lambda})^{\frac{1}{2-\theta}}$, and (2.3.4)

$$\begin{aligned}\mathbb{E}^\theta(S_T|\mathcal{F}_{t_0}) &= \int_0^{+\infty} s_0 e^{\mu\tau}(\frac{z}{\lambda})^{\frac{1}{2-\theta}} p(z;df,\lambda)dz \\ &= s_0 e^{\mu\tau}\int_0^{+\infty} p(\lambda;df,z)dz\end{aligned}$$

Using the property of density of noncentral chi-squared distributions (see S.Mark [57]),

$$\int_y^{+\infty} p(z;\nu,k)dk = F(z;\nu-2,y), \tag{2.3.5}$$

we have

$$\mathbb{E}^\theta(S_T|\mathcal{F}_{t_0}) = s_0 e^{\mu\tau}F(\lambda;df-2,0).$$

□

With the explicit form of $\mathbb{E}^\theta(S_T)$ given in Proposition 2.3.1 under the CEV model, we are able to compare future expectations with different values of θ. The goal is to show that θ is related to the size of asset bubbles by the study of future expectations and the running maximum of S_t.

Proposition 2.3.2. *Let S_1 and S_2 follow SDEs*

$$\begin{aligned}&dS_1(t) = \delta S_1^{\frac{\theta_1}{2}}(t)dW_1(t),\\ &dS_2(t) = \delta S_2^{\frac{\theta_2}{2}}(t)dW_2(t), \quad \textit{for } t\in[t_0,T]\\ &S_1(t_0) = S_2(t_0) = s_0 > 0,\end{aligned}$$

where W_1 and W_2 are two standard Brownian motions. Let $\sigma_{ir} = \delta s_0^{\frac{\theta_i-2}{2}}$, $i=1,2$, and

$L(\theta, \sigma_r, \tau) = \max\{(\theta-2)^{-\frac{1}{2}}\sigma_r^{-2}, (\theta-2)^{-2}\sigma_r^{-2}\}$. *Then, if* $\theta_1 > \theta_2 > 2$, $s_0 \geq 1$ *and* $\tau = T - t_0 < \min\{L(\theta_1, \sigma_{2r}, s_0), L(\theta_2, \sigma_{1r}, s_0)\}$,

$$\mathbb{E}[\mathbf{S_1(T)}] < \mathbb{E}[\mathbf{S_2(T)}] < \mathbf{s_0}.$$

Proof. Let S_t follow the SDE (2.3.2) with $\theta > 2$ and initial value $s_0 > 0$. By Proposition 2.3.1, and using the same notation there,

$$\mathbb{E}^{\theta}(S_T) = F\Big[\lambda; df - 2, 0\Big] s_0.$$

Let $d = \frac{1}{\theta-2}$ and $C = \frac{2}{\delta^2\tau}$. Then

$$\begin{aligned} U &= 1 - F\Big[\lambda; df - 2, 0\Big] = 1 - F\Big[2Cs_0^{-\frac{1}{d}}d^2; 2d, 0\Big] \\ &= \int_{2Cs_0^{-\frac{1}{d}}d^2}^{+\infty} \frac{1}{\Gamma(d)}\frac{1}{2^d} z^{d-1}e^{-z/2}dz = \int_{Cs_0^{-\frac{1}{d}}d^2}^{+\infty} \frac{1}{\Gamma(d)} y^{d-1}e^{-y}dy. \end{aligned}$$

To study the monotonicity of U as a function of d, we examine $\frac{\partial U}{\partial d}$,

$$\begin{aligned} \Gamma(d)\frac{\partial U}{\partial d} &= -\Big(Cs_0^{-\frac{1}{d}}d^2\Big)^{d-1} e^{-Cs_0^{-\frac{1}{d}}d^2}\Big(Cs_0^{-\frac{1}{d}}\log(s_0) + 2dCs_0^{-\frac{2}{d}}\Big) \\ &\quad + \int_{Cs_0^{-\frac{1}{d}}d^2}^{+\infty} y^{d-1}\log(y)e^{-y}dy - \frac{\Gamma'(d)}{\Gamma(d)}\int_{Cs_0^{-\frac{1}{d}}d^2}^{+\infty} y^{d-1}e^{-y}dy. \end{aligned}$$

Let $H = Cs_0^{-\frac{1}{d}}d^2$. Then,

$$\Gamma(d)\frac{\partial U}{\partial d} = -H^d e^{-H}\Big(\frac{\log(s_0)}{d^2} + \frac{2}{d}\Big) + \int_H^{+\infty} y^{d-1}\log(y)e^{-y}dy - \frac{\Gamma'(d)}{\Gamma(d)}\int_H^{+\infty} y^{d-1}e^{-y}dy.$$

Notice that

$$\psi(d) = \frac{\Gamma'(d)}{\Gamma(d)}, \quad \text{is the polygamma function of order 0.}$$

Rewriting $H^d e^{-H}$ as

$$H^d e^{-H} = \int_H^{+\infty} y^d e^{-y} - d y^{d-1} e^{-y} dy,$$

Then

$$\frac{\partial U}{\partial d} = \frac{1}{\Gamma(d)} \int_H^{+\infty} y^{d-1} e^{-y} \Big[\log(y) - \psi(d) + \Big(\frac{\log s_0}{d^2} + \frac{2}{d} \Big)(d-y) \Big] dy. \tag{2.3.6}$$

Equation (2.3.6) indicates that when s_0 and H are large enough, $\frac{\partial U}{\partial d} \leq 0$. However it is hard to solve $\frac{\partial U(d)}{\partial d} \leq 0$ explicitly. As a compromise, we try to find a sufficient condition to guarantee $\frac{\partial U}{\partial d} < 0$. With the assumption that $s_0 \geq 1$ and by approximating $\psi(d)$ from its expansion (see [2])

$$\psi(d) = -\gamma - \sum_{k=0}^{\infty} \Big(\frac{1}{d+k} - \frac{1}{k+1} \Big), \quad \gamma = 0.5772\ldots,$$

we have when $y \geq H \geq \max\{2d^{\frac{3}{2}}, 2\}$,

$$\log(y) - \psi(d) + \Big(\frac{\log s_0}{d^2} + \frac{2}{d} \Big)(d-y) < 0,$$

which leads to $\frac{\partial U}{\partial d} < 0$. With the definition of H and C, one gets $H \geq \max\{2d^{\frac{3}{2}}, 2\}$ is equivalent to $\tau \leq \min\Big\{ \frac{s_0^{-\frac{1}{d}} d^{\frac{1}{2}}}{\delta^2}, \frac{s_0^{-\frac{1}{d}} d^2}{\delta^2} \Big\}$. Thus when $s_0 \geq 1$, $\tau \leq \min\Big\{ \frac{s_0^{-\frac{1}{d}} d^{\frac{1}{2}}}{\delta^2}, \frac{s_0^{-\frac{1}{d}} d^2}{\delta^2} \Big\}$, $\mathbb{E}^\theta(S_T) = s_0(1-U)$ is a increasing function of d, therefore also a decreasing function of θ conditioning on $\theta > 2$. Since $\sigma_r = \delta s_0^{\frac{\theta-2}{2}}$, the proof concludes. □

The reason that a nonnegative strict local martingale has an expectation that decreases with time is that there is some probability mass escaping to infinity. Theorem 2.2.2 suggests that a type 3 bubble only occurs when $\sigma(S) \to \infty$ too fast when $S \to \infty$. Thus we expect that a larger value of θ leads to a more extreme bubble. One attribute of a more extreme bubble is relatively smaller future expectation. However Proposition 2.3.2 suggests the situation is more complicated, that we actually require some conditions on the starting value and time. The intuition behind the conditions is that when S is very small, a larger θ actually shrinks the volatility $\sigma(S)$ more. An extreme case would be $S < 1$, when $\theta \to +\infty$, $\sigma(S) \to 0$, $\mathbb{E}^{\theta}(S_T)$ barely decreases with time. The condition on t is for the same reason. When t is large enough, S_t is likely to be a small value, as $\mathbb{E}^{\theta}(S_t)$ is decreasing towards 0 when $\theta > 2$, and a large value of θ shrinks volatility when S is smaller than 1.

In Proposition 2.3.2, the conditions on s_0 and τ are usually satisfied in practice. For example, when $2 < \theta < 3$, $s_0 > 1, \sigma_r < 0.5$, which is usually the case in the stock markets, large values of θ lead to smaller future expectations within 4 years. In later sections, we shall use real trading data to estimate θ. And the that $\theta > 3$ rarely occurs, and the lifetime of a bubble generally stays within 4 years. To better illustrate how $\mathbb{E}^{\theta}(S_T|\mathcal{F}_{t_0})/S_{t_0}$ changes with different θ, we present an example. Let $\delta = 0.5$, $S_{t_0} = 10$, θ range from 2 to 4.8, and τ range from a quarter of a year to 2 years. For different values of θ and $\tau = T - t_0$, we calculate the ratio $\mathbb{E}^{\theta}(S_T|\mathcal{F}_{t_0})/S_{t_0}$. Each row in Table 2.3.1 illustrates that $\mathbb{E}^{\theta}(S_T|\mathcal{F}_{t_0})/S_{t_0}$ is decreasing with

Table 2.3.1: Future expectation of CEV processes

	τ =1/4	τ =1/2	τ =3/4	τ =1	τ =5/4	τ =3/2	τ =7/4	τ =2
$\theta = 2.0$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\theta = 2.4$	1.000	1.000	0.999	0.993	0.973	0.939	0.893	0.840
$\theta = 2.8$	0.621	0.353	0.237	0.175	0.137	0.111	0.093	0.080
$\theta = 3.2$	0.135	0.078	0.056	0.044	0.037	0.031	0.028	0.025
$\theta = 3.6$	0.054	0.035	0.027	0.023	0.020	0.018	0.016	0.015
$\theta = 4.0$	0.032	0.023	0.018	0.016	0.014	0.013	0.012	0.011
$\theta = 4.4$	0.023	0.017	0.015	0.013	0.012	0.011	0.010	0.010
$\theta = 4.8$	0.019	0.014	0.013	0.011	0.010	0.010	0.009	0.009

$\mathbb{E}^{\theta}(S_T|\mathcal{F}_{t_0})/S_{t_0}$ with $S_{t_0} = 10, \delta = 0.5$.

time τ, except that when $\theta = 2$. Each column illustrates that as θ increases from 2 to 4.8, $\mathbb{E}^{\theta}(S_T|\mathcal{F}_{t_0})/S_{t_0}$ is decreasing. Figure 2.3.1 also illustrates how $\mathbb{E}^{\theta}(S_T|\mathcal{F}_{t_0})/S_{t_0}$ changes with τ and θ. In the left plot of Figure 2.3.1, each curve represents $\mathbb{E}^{\theta}(S_T|\mathcal{F}_{t_0})/S_{t_0}$ as a function of τ, for a fixed θ. At first curves corresponding to larger values of θ decrease more rapidly than curves with smaller values of θ. But when time τ is large enough, curves begin to cross, which means the result that CEV processes with a larger value for θ produce smaller future expectations. However, this occurs only for very large τ when θ is well bounded from above. In our example, for the curves with $\theta \leq 4$, (which is usually the case in stock markets), there is no crossing when $\tau < 10$ years.

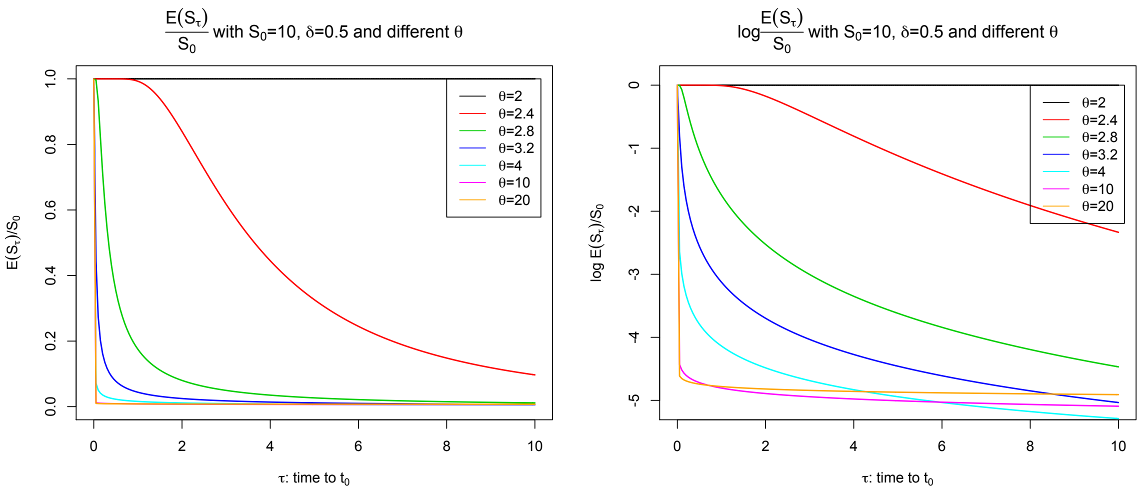

Figure 2.3.1: Future expectation of CEV processes

Linking Proposition 2.3.2 to the fundamental price process and bubble process, we have

Theorem 2.3.2. *Let $S_1(t)$ and $S_2(t)$ be the nonnegative price processes of two stocks. Assume that S_1 and S_2 pay no dividends and the spot interest rate is 0. Let Q_1 and Q_2 be risk neutral measures under which $S_1(t)$, $S_2(t)$ are local martingales that follows the SDEs below*

$$dS_1(t) = \delta S_1^{\frac{\theta_1}{2}}(t) dW_1(t), \;\; S_1(t_0) = s_0,$$
$$dS_2(t) = \delta S_2^{\frac{\theta_2}{2}}(t) dW_2(t), \;\; S_2(t_0) = s_0, \quad \text{for } t \in [t_0, \; T].$$

Assume that asset bubbles do not exist for either $S_1(t)$ or $S_2(t)$ after time T. Let $S_1^\star$, $S_2^\star$ be the fundamental value processes of the assets calculated under Q_1 and Q_2, and let $\tilde{S}_1(t) = S_1(t) - S_1^*(t)$, $\tilde{S}_2(t) = S_1(t) - S_1^*(t)$, be the bubble processes. Let $\sigma_{ir} = \delta s_0^{\frac{\theta-2}{2}}$, $i = 1, 2$, L is defined as in Proposition 2.3.2. Then, if $\theta_1 > \theta_2 > 2$, $s_0 \geq 1$, $\tau = T - t_0 < min\{L(\theta_1, \sigma_{2r}, s_0), L(\theta_2, \sigma_{1r}, s_0)\}$:

(i) for the two fundamental value processes at time t_0,

$$S_1^\star(t_0) < S_2^\star(t_0);$$

(ii) for the two bubble processes at time t_0,

$$\tilde{S}_1(t_0) > \tilde{S}_2(t_0).$$

Proof. (i) Assume S is a price process which has a bubble when $t \in [t_0,\ T]$, and no bubble after T until liquidation time. Since there is no bubble after T, the stock price S_T should match the fundamental price $S^\star_T$. Under the assumption of no dividends and an interest rate of 0, under the price process's risk neutral measure Q,

$$S^\star_{t_0} = \mathbb{E}_Q(S^\star_T|\mathcal{F}_{t_0}) = \mathbb{E}_Q(S_T|\mathcal{F}_{t_0}).$$

With the assumptions on $s_0,\ \tau,\ \theta_1$ and θ_2, we can apply Proposition 2.3.2 to obtain

$$S^\star_1(t_0) = E_{Q_1}(S_1(T)|F_{t_0}) < E_{Q_2}(S_2(T)|F_{t_0}) = S^\star_2(t_0).$$

(ii) From (i),

$$\begin{aligned}\tilde{S}_1(t_0) &= S_1(t_0) - S^\star_1(t_0),\\ \tilde{S}_2(t_0) &= S_2(t_0) - S^\star_2(t_0),\\ \tilde{S}_1(t_0) &> \tilde{S}_2(t_0).\end{aligned}$$

□

2.3.3 Running Maximum under the CEV Model

The running maximum of a price process is interesting to look at, since when a bubble occurs, the price process is likely to have a temporary boom and hit relatively high levels in a short time. The running maximum of a price process is also linked to the problem of pricing look back options. Under the CEV model, the Laplace transform is given in V.Linetsky [52], D.Davydov and V.Linetsky [15]. However we are not able to obtain the explicit form of the distribution for $S^*_T = \max\{S_t : t_0 \leq t \leq T\}$ when $\theta > 2$, since the inverse Laplace transformation is not easy to calculate. However if T is an exponentially distributed random variable, one obtains an explicit form for the distribution of S_T^*.

Proposition 2.3.3. *Let S be a process satisfying the SDE (2.3.1). Let τ be a random time independent of S and has exponential distribution $P(\tau > t) = e^{-\lambda t}, \lambda > 0$ and $T = t_0 + \tau$. Let $S_T^* = \max\{S_t : t_0 \le t \le T\}$ be the running maximum at time T. Then for any $M > s_0$, define ν, a and b by*

$$\nu = \frac{1}{\theta - 2},\ a = \frac{2\nu}{\delta} s_0^{-\frac{1}{2\nu}},\ b = \frac{2\nu}{\delta} M^{-\frac{1}{2\nu}}.$$

Then

$$P(S_T^* > M) = \frac{a^{-\nu} K_\nu(a\sqrt{2\lambda})}{b^{-\nu} K_\nu(b\sqrt{2\lambda})}, \tag{2.3.7}$$

where K_ν is the modified Bessel function of the second kind.

Proof. Let $R_t = \frac{2\nu}{\delta} S_t^{-\frac{1}{2\nu}}$. Then by Itô's formula, R_t follows the SDE

$$dR_t = (\frac{1}{2} + \nu)\frac{1}{R_t} dt - dW_t, R_{t_0} = a.$$

Thus R_t is a Bessel process with order $\nu = \frac{1}{\theta - 2}$. For $a,\ b > 0$, $\tau_{a,b}^{(\nu)}$ is the first hitting time to b for R_t. The Laplace transform of the hitting time of a Bessel process has been well studied. We link the probability distribution of the running maximum of CEV process to its corresponding Bessel process hitting time:

$$P_{S_0}(S_T^* > M) = P(\tau_{a,b} < t_0 + \tau) = \int_0^{+\infty} P(\tau_{a,b}^{(\nu)} < t_0 + t) e^{-\lambda t} dt = \mathbb{E}[e^{-\lambda(\tau_{a,b}^{(\nu)} - t_0)}].$$

From the result of A.N.Borodin and P.Salminen [?], we have

$$\mathbb{E}[e^{-\lambda(\tau_{a,b}^{(\nu)} - t_0)}] = \frac{a^{-\nu} K_\nu(a\sqrt{2\lambda})}{b^{-\nu} K_\nu(b\sqrt{2\lambda})}.$$

∎

With Proposition 2.3.3, we are able to have some comparison results on the running maximum of CEV processes with different values of θ.

Theorem 2.3.3. *Let* S_1 *and* S_2 *be defined in the same way as in Proposition* 2.3.2. *Assume* $\theta_1 > \theta_2 > 2$, $s_0 \geq 1$. *Then*

$$\lim_{t \to t_0+} \frac{P(S_1^*(t) > M)}{P(S_2^*(t) > M)} = +\infty, \ \forall M > s_0. \tag{2.3.8}$$

Further if $\tau = T - t_0 < \min\{L(\theta_1, \sigma_{2r}, s_0), L(\theta_2, \sigma_{1r}, s_0)\}$, *where* L *and* $\sigma_{ir}, \ i = 1, 2$ *are the same as in Proposition* 2.3.2, *then*

$$\lim_{M \to \infty} \frac{P(S_1^*(T) > M)}{P(S_2^*(T) > M)} = \frac{s_0 - \mathbb{E}[S_1(T)]}{s_0 - \mathbb{E}[S_2(T)]} > 1. \tag{2.3.9}$$

Proof. Let τ be a random time independent of S_1 and S_2 with exponential distribution $P(\tau > t) = e^{-\lambda t}, \ \lambda > 0$, and let $T = t_0 + \tau$,

$$\lim_{t \to t_0+} \frac{P(S_1^*(t) > M)}{P(S_2^*(t) > M)} = \lim_{\lambda \to +\infty} \frac{P(S_1^*(T) > M)}{P(S_2^*(T) > M)}.$$

For the Bessel function of second kind K_ν, it is known that

$$\lim_{z \to +\infty} \frac{K_\nu(z)}{\sqrt{\frac{\pi}{2z}} e^{-z}} = 1.$$

Let $a_{\nu,M} = \frac{2\nu}{\delta} s_0^{-\frac{1}{2\nu}}$, $b_{\nu,M} = \frac{2\nu}{\delta} M^{-\frac{1}{2\nu}}$ and $\nu_1 = \frac{1}{\theta_1 - 2}$, $\nu_2 = \frac{1}{\theta_2 - 2}$. From Proposition 2.3.3,

$$\lim_{t \to t_0+} \frac{P(S_1^*(t) > M)}{P(S_2^*(t) > M)} = \lim_{\lambda \to +\infty} \sqrt{\frac{a_{\nu_2} b_{\nu_1}}{b_{\nu_2} a_{\nu_1}}} e^{\sqrt{2\lambda}(b_{\nu_1} - a_{\nu_1} - b_{\nu_2} - a_{\nu_2})}. \tag{2.3.10}$$

Since

$$\frac{\partial^2 (b_{\nu,M} - a_{\nu,M})}{\partial \nu \partial M} = -\frac{1}{2\nu^2} \log M M^{-\frac{1}{2\nu} - 1} < 0, \ M > s_0 \geq 1$$

one gets

$$\frac{\partial(b_{\nu,M} - a_{\nu,M})}{\partial\nu} < \frac{\partial(b_{\nu,M} - a_{\nu,M})}{\partial\nu}\bigg|(M = s_0) = 0.$$

Thus as $\nu_1 < \nu_2$, we get $(b_{\nu_1,M} - a_{\nu_1,M}) - (b_{\nu_2,M} - a_{\nu_2,M}) > 0$. Together with (2.3.10), we conclude that (2.3.8) holds. To prove (2.3.11), we first state one known result in Madan and Yor [55]. For a non-negative local martingale X starting from $t = 0$ with value x_0, there is the identity

$$x_0 - \mathbb{E}(X_t) = \lim_{K\to+\infty} KP(X_t^* > K).$$

Thus

$$\lim_{M\to\infty} \quad \frac{P(S_1^*(T) > M)}{P(S_2^*(T) > M)} = \frac{s_0 - \mathbb{E}[S_1(T)]}{s_0 - \mathbb{E}[S_2(T)]}. \tag{2.3.11}$$

With the assumptions on τ and s_0, Proposition 2.3.2 applies and $\frac{s_0-\mathbb{E}[S_1(T)]}{s_0-\mathbb{E}[S_2(T)]} < 1$. □

We use an example to illustrate how $P(S_T^* > M)$ changes with θ, with $\tau = T - t_0$ exponentially distributed. Set the starting value as 10 and $M = 40$, we vary λ and θ. Table 2.3.2 and Figure 2.3.2 illustrate how $P(S_T^* > M)$ changes with different λ and θ. In Figure 2.3.2, each line represents $P(S_T^* > M)$ as a function of $\frac{1}{\lambda} = \mathbb{E}(\tau)$, for a fixed θ. We can see that CEV processes with larger values of θ where $\theta > 2$, have larger probabilities to hit a large value from below in a short time as proved in Theorem 2.3.3 and illustrated in Figure 2.3.2. As indicated by Figure 2.3.2, we also see that when time range is sufficiently long, the CEV processes with larger values of θ, conditioning on $\theta > 2$, has smaller probabilities of hitting a large value, which can be proved using Hamana and Matsumoto's result in [30], interested readers can refer to Appendix A.1.

Table 2.3.2: Distribution of the running maximum for CEV processes

	$\lambda = 0.001$	$\lambda = 0.01$	$\lambda = 0.1$	$\lambda = 1$	$\lambda = 2$	$\lambda = 5$	$\lambda = 10$
$\theta = 2.1$	0.248	0.230	0.137	0.015	0.004	0.000	0.000
$\theta = 2.3$	0.248	0.236	0.164	0.034	0.012	0.002	0.000
$\theta = 2.5$	0.249	0.239	0.185	0.059	0.028	0.006	0.001
$\theta = 3.0$	0.249	0.244	0.216	0.128	0.091	0.045	0.020
$\theta = 4.0$	0.248	0.245	0.234	0.202	0.185	0.156	0.128
$\theta = 10$	0.243	0.240	0.237	0.233	0.231	0.229	0.227
$\theta = 20$	0.239	0.237	0.235	0.233	0.233	0.232	0.231
$\theta = 40$	0.236	0.235	0.234	0.233	0.232	0.232	0.232

$P(S_T^* > M)$ with $P(T - \tau > t) = e^{-\lambda t}$, $s_0 = 10,\ M = 40,\ \delta = 0.5$

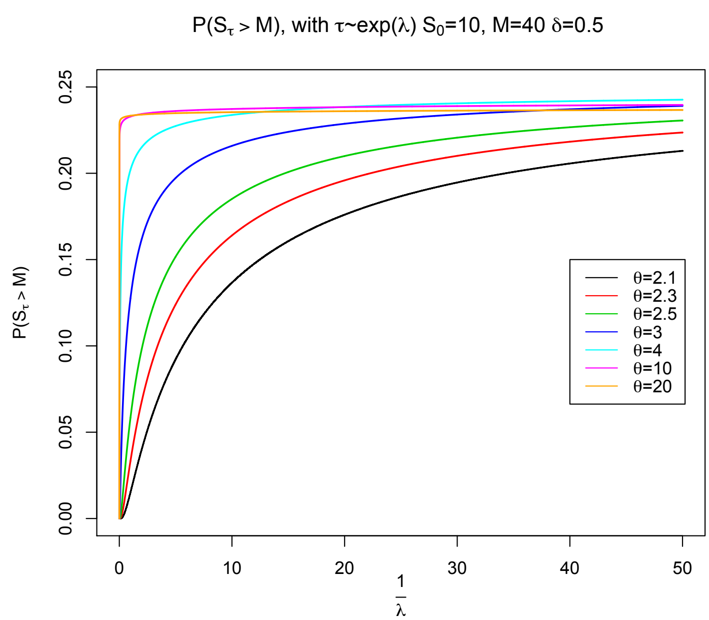

Figure 2.3.2: Distribution of the running maximum of CEV processes

2.4 Detecting Bubbles Under the CEV Model

2.4.1 Statistical Formulation

If the asset price process follows the general form (2.2.2), Theorem 2.2.2 indicates that the $\sigma(x)$ as a function of the current price is critical for the existence of asset bubbles. Under the assumption that the functional form σ does not change in the time range of interest, Jarrow, Kchia, and Protter [43] used reproducing kernel space methods to estimate $\sigma(x)$ and check condition $\boldsymbol{H}$ in Theorem 2.2.2 to detect asset bubbles. In this paper we are interested in detecting the existence and severity of asset bubbles. We allow the functional form $\sigma(X_t)$ to change over the time range $[0, T]$, but stay the same within a discretization of time with size ΔT. Letting $T_0 = 0,\ T_n = T,\ \Delta T = T_i - T_{i-1},\ i = 1 \dots n$, the price process follows the SDE

$$dX_t = \sum_{i=1}^{n} \sigma_i(X_t) I_{\{t \in (T_{i-1}, T_i]\}} dW_t. \tag{2.4.1}$$

The form (2.4.1) is also used in detecting the lifetime of bubble in Protter, Obayashi and Wang [67]. Assuming that each $\sigma_i(S)$ satisfies the Engelbert-Schmidt conditions and applying Theorem 2.2.2, the asset bubble detection problem is transformed into finding those $i's$ with $\sigma_i(x)$ satisfying condition $\boldsymbol{H}$ in Theorem 2.2.2. The advantage of using a parametric form for the volatility function $\sigma(x)$ is that once the parameters have been estimated, we know the tails. We choose a family of volatility functions to include the cases where S is either a martingale or a strict local martingale under its risk neutral measure. Here we use power functions $\sigma_i(S) = a_i S^{b_i}$, which results in the CEV process, where a_i and b_i are unknown parameters. Our goal is to detect the potential asset bubble time region, and the severity of the bubble; these could be achieved by making inference on the parameters a_i and b_i.By applying condition $\boldsymbol{H}$ in Theorem 2.2.2, we find that when $b_i > 1$, S is a strict local martingale, which implies the existence of an asset bubble, and when $b_i \leq 1$, there is no asset bubble. Let $a_i > 0, b_i > 0$, it is known that if $S_0 > 0$, S_t is positive almost surely. Since $\sigma_i(S) = a_i S^{b_i}$, if we take the log on both sides, we have $log(\sigma_i(S)) = log(a_i) + b_i log(S)$.

We therefore construct a statistical model with time varying parameters $log(a_i)$ and b_i

$$\log(\hat{\sigma}_i(S_i)) = \log(a_i) + b_i \log(S_i) + \nu_i,$$

where ν_i is an error with $\mathbb{E}(\nu_i | \log(S_i)) = 0$. The terms $log(S_i)$ and $log(\hat{\sigma}_i(S_i))$ can both be obtained by using the historical stock price.

2.4.2 Estimating Volatility

In order to apply Theorem 2.2.2 to detect asset bubbles, we need to estimate the volatility $\sigma(x)$ at fixed prices. Using the price and volatility pairs $(x, \hat{\sigma}^2(x))$, we can estimate the parameters in $\sigma(x)$. To estimate the volatility at fixed prices, we use the Florens-Zmirou's estimator:

$$\hat{\sigma}_n^2(x) = \frac{\sum_{i=0}^{n} I_{\{|S_{t_i} - x| < h_n\}} n(S_{t_{i+1}} - S_{t_i})^2}{T \sum_{i=0}^{n} I_{\{|S_{t_i} - x| < h_n\}}}, \tag{2.4.2}$$

where S_{t_i} are n observations on $[0, T]$, with $t_{i+1} - t_i = \frac{T}{n}$, and h_n is a sequence of positive real numbers converging to 0. The following theorem can be found in Florens-Zmirou [18].

Theorem 2.4.1. *Assume σ is bounded, strictly positive, and has three continuous and bounded derivatives. If $(h_n)_{n \geq 1}$ satisfies that $nh_n \to \infty$ and $nh_n^3 \to 0$, then $\sqrt{N_x^n}(\frac{S_n(x)}{\sigma^2(x)} - 1)$ converges in distribution to $\sqrt{2}Z$, where Z is a standard normal random variable and $N_x^n = \sum_{i=0}^{n} I_{\{|S_{t_i} - x| < h_n\}}$.*

Remark: For the Florens-Zmirou's estimator, in the theorem σ needs to be bounded. In practice, conditioning on the price history, the maximum of the price is always finite, σ for that path is finite as well.

For each trading day, using the intraday trading prices we estimate $\sigma(\bar{x})$, where $\bar{x}$ is the daily average of the trading prices. Assuming that the intraday fluctuation is relatively small, we can take

$$\hat{\sigma}^2(\bar{x}) = \frac{\sum_{i=0}^{n}(S_{t_{i+1}} - S_{t_i})^2}{\Delta t N}, \tag{2.4.3}$$

where N is the number of price observations in a day. By Theorem 2.4.1, the distribution of $\hat{\sigma}^2(\bar{x})$ is approximately $N(\sigma^2(\bar{x}), 2\sigma^4(\bar{x}))$. Note that the the variance varies with $\bar{x}$, which is undesirable when estimating σ. We apply the delta method to resolve this issue. We take the log transform to $\hat{\sigma}^2(\bar{x})$, then

$$\log \hat{\sigma}^2(\bar{x}) \sim N\Big(\log \sigma^2(\bar{x}), \frac{\sigma^4(\bar{x})}{N} \Big[\frac{\partial \log \sigma^2(x)}{\partial \sigma^2(x)} \Big| \bar{x} \Big] \Big),$$

which leads to

$$\log \hat{\sigma}^2(\bar{x}) \sim N\Big(\log \sigma^2(\bar{x}), \frac{2}{N} \Big). \tag{2.4.4}$$

There are several reasons for choosing one day as our time range. Taking a time range finer than a day will necessitate the consideration of the higher volatility associated with the market opening and closing times. Further, the use of a finer range will result in fewer data points if the sampling frequency does not change. At the same time, the sampling frequency can not go too high in order to avoid microstructure noise. Taking a time range greater than one day will include the overnight effect, which is not desirable.

For day k, we have N historical prices $X_{k,1}, X_{k,2} \ldots X_{k,N}$ with regular sampling $T_{k,i} = T_{k,i-1} + \Delta T$, where ΔT is the sampling time gap. Let $\bar{x}_k$ be the average price of the observations $X_{k,1}, X_{k,2} \ldots X_{k,N}$ on k_{th} day, and $\hat{\sigma}^2(\bar{x}_k)$ is obtained using (2.4.3). With all the pairs $(\bar{x_k}, \hat{\sigma_k}^2(\bar{x}_k))$, we can estimate $\sigma_k(x)$.

2.4.3 The Dynamic Linear Regression Model

As discussed previously, an asset bubble exists for those $i's$ with $b_i > 1$ in the regression

$$log(\hat{\sigma}_t(S_t)) = log(a_t) + b_t log(S_t) + \nu_t,\ t = 1, 2 \dots$$

where ν_t is the noise term, $\log(S_t)$ and $\log(\hat{\sigma}_t(S_t))$ are known from observations of the stock price process. For notational convenience, let $\alpha_t = log(a_t)$, $\beta_t = b_t$, $\theta_t = \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix}$, $Y_t = log(\hat{\sigma}_t(S_t))$ and $F_t = \begin{bmatrix} 1 & log(S_t) \end{bmatrix}$. We further assume $\alpha_{t+1}|\alpha_t \sim N(\alpha_t, w_1)$ and $\beta_{t+1}|\beta_t \sim N(\beta_t, w_2)$. A dynamic linear model is specified by a normal prior $\theta_0 \sim N_p(m_0, C_0)$ together with a pair of equations for each time $t \geqslant 1$,

$$Y_t = F_t\theta_t + \nu_t, \nu_t \sim N(0, V_t),\ \text{observation equation;}$$

$$\theta_t = G_t\theta_{t-1} + \omega_t, \omega_t \sim N_p(0, W_t),\ \text{state equation,}$$

where $G_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $F_t = \begin{bmatrix} 1 & log(S_t) \end{bmatrix}$, and $(\nu_t)_{t\geqslant 1}$, $(\omega_t)_{t\geqslant 1}$ are two independent sequences of independent Gaussian random vectors with mean zero and variance matrices $V_t = \begin{bmatrix} v \end{bmatrix}$ and $W_t = \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix}$. Consider the dynamic linear model specified above, and let $\theta_{t-1}|y_{1:t-1} \sim N(m_{t-1}, C_{t-1})$. Then the following statements from Kalman Filtering hold.

(i)The one-step-ahead predictive distribution of θ_t given $y_{1:t-1}$ is Gaussian, with mean and covariance matrix

$$a_t = \mathbb{E}(\theta_t|y_{1:t-1}) = G_t m_{t-1},\ R_t = Var(\theta_t|y_{1:t-1}) = G_t C_{t-1} G_t' + W_t.$$

(ii)The one-step-ahead predictive distribution of Y_t given $y_{1:t-1}$ is Gaussian, with mean and covariance matrix

$$f_t = \mathbb{E}(Y_t|y_{1:t-1}) = F_t a_t,\ R_t = Var(Y_t|y_{1:t-1}) = F_t R_t F_t' + V_t.$$

(iii)The distribution of θ_t given $y_{1:t}$ is Gaussian, with mean and covariance matrix

$$m_t = \mathbb{E}(\theta_t|y_{1:t}) = a_t + R_t F_t' Q_t^{-1} e_t,\ C_t = Var(\theta_t|y_{1:t}) = R_t - R_t F_t' Q_t^{-1} R_t,$$

where $e_t = Y_t - f_t$ is the forecast error. We are most interested in the last step of deriving the distribution of θ_t given $y_{1:t}$, because it gives the distribution of b_t given the price process until the t_{th} trading day. However the distribution $\theta_t|y_{1:t}$ requires calculating the 3 steps above recursively.

2.4.3.1 Estimating the Predictive Covariance Matrix W_t

For the Kalman Filter decribded above, we have some knowledge of the matrix V_t from the property of Florens-Zmirou's estimator for volatility. However we do not know what $W_t = \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix}$ is. We can not use the maximum likelihood as the maximizer does not exist, even though we already assume it is a constant matrix. One solution is to put a conjugate prior, then use the posterior mean of W. We put independent inverse gamma distribution priors on w_1, w_2,

$$w_1 \sim Inv\_gamma(a_1, b_1),\ \text{with density } \pi(w_1) = \frac{b_1^{a_1}}{\Gamma(a_1)} w_1^{-a_1-1} \exp(\frac{-b_1}{w_1});$$
$$w_2 \sim Inv\_gamma(a_2, b_2),\ \text{with density } \pi(w_2) = \frac{b_2^{a_2}}{\Gamma(a_2)} w_2^{-a_2-1} \exp(\frac{-b_2}{w_2}).$$

With the above prior distributions for w_1 and w_2, we can easily compute the conditional distribution $\pi(w_1, w_2|D_T, \theta_{0...T})$ as

$$\pi(w_1, w_2|D_T, \theta_{0...T}) \sim \pi(w_1)\pi(w_2)\pi(\theta_0)\Pi_{t=1}^{T}\pi(Y_t|\theta_t)\Pi_{t=1}^{T}\pi(\theta_t|\theta_{t-1})$$

$$\propto w_1^{-(\frac{T}{2}+a_1+1)} \exp\left\{-\frac{2b_1 + \sum_{t=1}^{T}(\alpha_t - \alpha_{t-1})^2}{2w_1}\right\} w_2^{-(\frac{T}{2}+a_2+1)} \exp\left\{-\frac{2b_2 + \sum_{t=1}^{T}(\beta_t - \beta_{t-1})^2}{2w_2}\right\},$$

where $D_T = (Y_1, \ldots, T_T)$. Let $\tilde{\theta} = \theta_{0...T}$,

$$w_1|D_T, \tilde{\theta} \sim Inv\_gamma(a_1 + T/2, b_1 + \sum_{t=1}^{T}(\alpha_t - \alpha_{t-1})^2/2), \tag{2.4.5}$$

$$w_2|D_T, \tilde{\theta} \sim Inv\_gamma(a_2 + T/2, b_2 + \sum_{t=1}^{T}(\beta_t - \beta_{t-1})^2/2). \tag{2.4.6}$$

To estimate w_1 and w_2, we use the sample mean from their posterior distribution. We sample from the joint distribution $\pi(w_1, w_2, \tilde{\theta}|D_T)$ using the Gibbs sampler instead of directly from $\pi(w_1, w_2|D_T)$, as the later is not easy to calculate explicitly. Gibbs sampling from $\pi(w_1, w_2, \theta_{0...T}|D_T)$ requires iteratively simulating from the full conditional distributions $\pi(\theta_{0:T}|D_T, w_1, w_2)$, $\pi(w_1, w_2|D_T, \theta_{0...T})$ and $\pi(w_1, w_2|D_T, \theta_{0...T})$. We describe the steps of the Gibbs sampling procedure used here:

1. Initializer: set $w_1 = w_1^{(0)}, w_2 = w_2^{(0)}$,

2. For $i = 1 \ldots N$:
   (i)Draw $\theta_{0:T}^{i}$ from $\pi(\theta_{0:T}|D_T, w_1 = w_1^{(i-1)}, w_2 = w_2^{(i-1)})$,
   (ii)Draw w_1 from $\pi(w_1|D_T, \theta_{0...T} = \theta_{0...T}^{(i)})$,
   (ii)Draw w_2 from $\pi(w_2|D_T, \theta_{0...T} = \theta_{0...T}^{(i)}))$.

For step 2(ii) and 2(iii), we already have the explicit form from (2.4.5). For step 2(i), we introduce the forward filtering backward sampling method (FFBS).

1. Run the Kalman Filter as described in the previous section;

2. Draw θ_T from the distribution $N(m_T, C_T)$;

3. For $t = T-1, T-2\ldots 0$, draw θ_t from $\pi(\theta_t|D_t, \theta_{t+1}, \theta_{t+2}\ldots\theta_T) = N(h_t, H_t)$, where

$$h_t = m_t + C_t G'_{t+1} R^{-1}_{t+1}(\theta_{t+1} - a_{t+1}),$$
$$H_t = C_t - C_t G'_{t+1} R^{-1}_{t+1} G_{t+1} C_t.$$

After sampling from $\pi(w_1, w_2, \theta_{0\ldots T}|D_T)$, we use the sample means as the estimators for w_1 and w_2. This approach also solves the problem of filtering, smoothing and forecasting at the same time. However, unlike the Kalman Filter, it is not designed for recursive inference. If new observations of y_t become available, one has to run a new Markov Chain all over again for sampling, which is computationally costly. One solution is to use part of the data to estimate w_1 and w_2 in the beginning, and as commonly used in practice, treat the estimates $\hat{w_1}$, $\hat{w_2}$ as if they were true values of the parameters in applying filtering and smoothing recursions. For a simulation result on the procedure described in this section in estimating the time varying parameters in the CEV model, see Appendix A.1.

2.5 Real Data Examples

2.5.1 Dot-com Bubbles

The dot-come bubble happened roughly during 1997 to 2000 in the US stock markets. It was characterized by a rapid rise in the stock markets, fueled by investments in Internet-based companies. The effects of the bubble bursting were that several companies went bankrupt and many other struggling companies became acquired or merged with other companies, with negative consequences such as employee layoffs and delays in the development of potential technologies. We pick several representative stocks to show how our model works for detecting asset bubbles instantaneously and we examine the severity of the bubbles by estimating the exponent parameter in the CEV model.

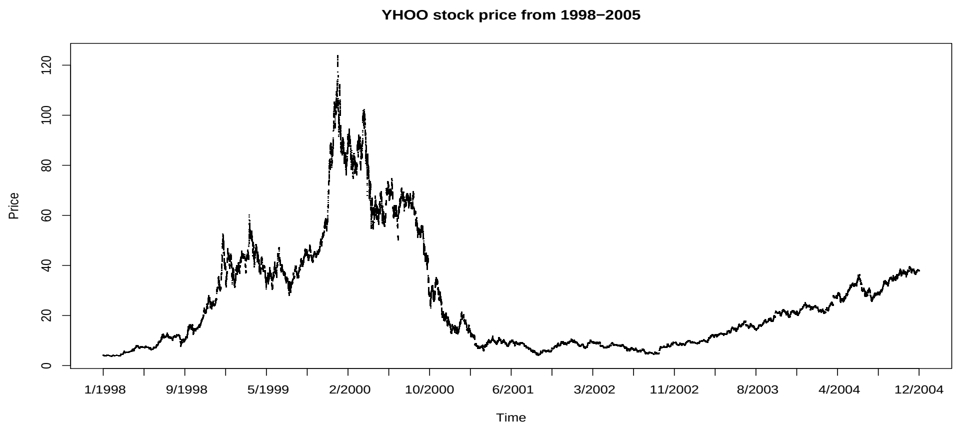

Figure 2.5.1: Historical YHOO stock price.

2.5.1.1 Stock YHOO

Yahoo!'s stock price skyrocketed during the dot-com bubble and closed at an all-time high in 2000; however, after the dot-com bubble burst, it reached an all-time low in 2001. The price history of the stock symbol YHOO is shown in Figure 2.5.1. Applying the dynamic linear regression described in the previous section on historical trading data of symbol YHOO, we have our instantaneous estimated $\hat{\beta}_t$ in the sense that $\hat{\beta}_t$ is obtained using the price process records up to the time t, together with its 95% confidence interval plotted in Figure 2.5.2 . As illustrated in Figure 2.5.2, $\hat{\beta}_t > 1$ for the time range roughly from Janurary 1998 to June 2001, we can conclude that the stock was in bubble in that time range. In Figure 2.5.3, the price is colored red, if $\hat{\beta}_t > 1$ for the corresponding day. The time region colored red well captured the soaring and crashing period of the stock. For better illustration, we also take the log of the price process in Figure 2.5.4.

2.5.1.2 Stock INSP

Founded in the mid-1990s amid the beginnings of the dot-com boom, InfoSpace went public in 1998 and the company's stock quickly soared. At its peak in early 2000, InfoSpace stock was worth more than 1,000 dollars per share. It dropped drastically after the dot-com bubble burst and fluctuated around its initial offering price of roughly15 dollars. The history of

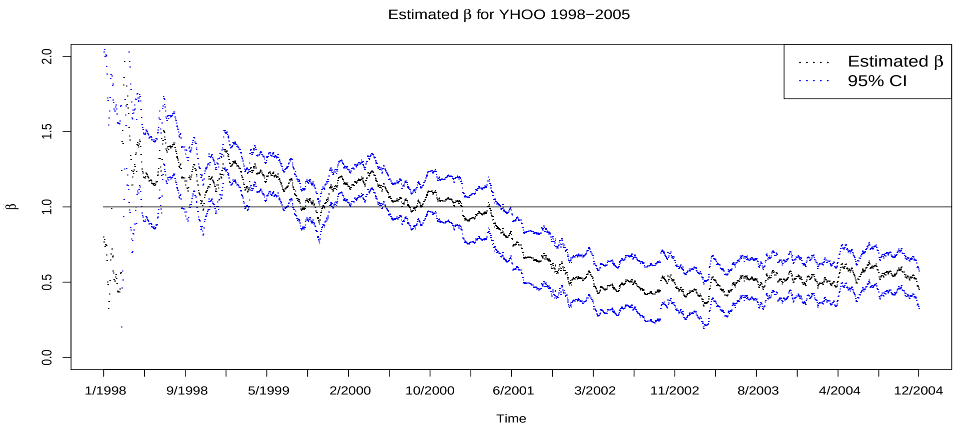

Figure 2.5.2: Instantaneous estimation $\hat{\beta}$ for YHOO.

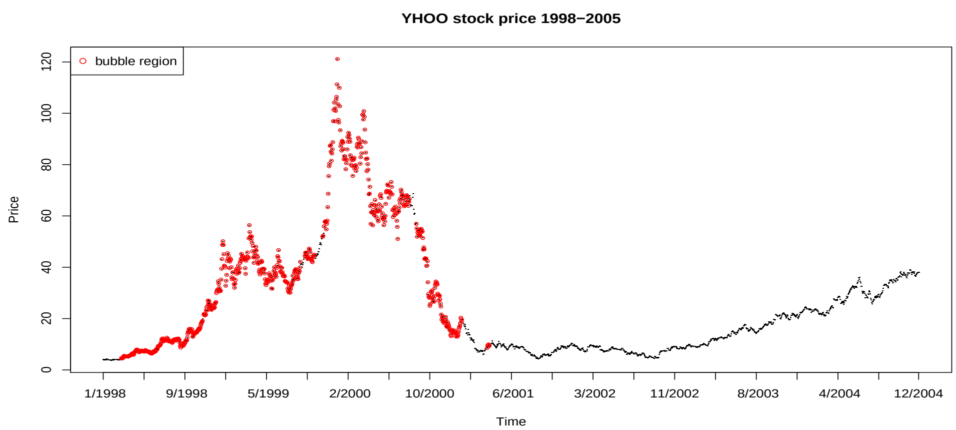

Figure 2.5.3: Historical price for YHOO.

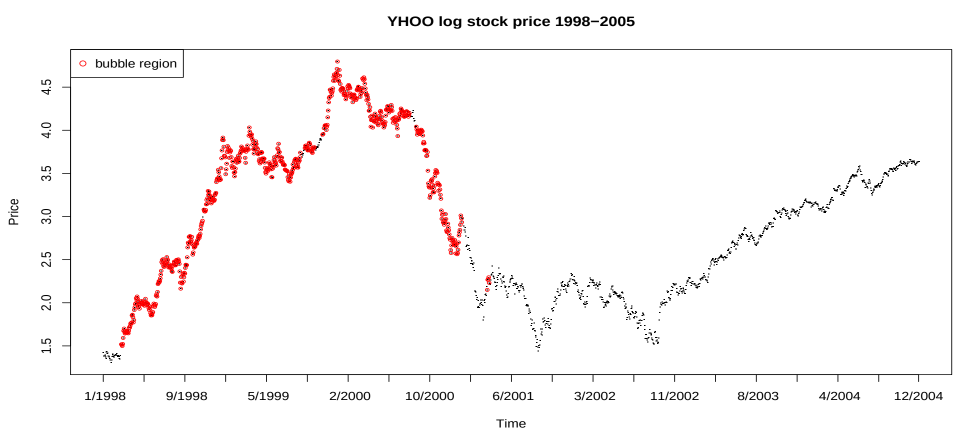

Figure 2.5.4: Historical log price for YHOO.

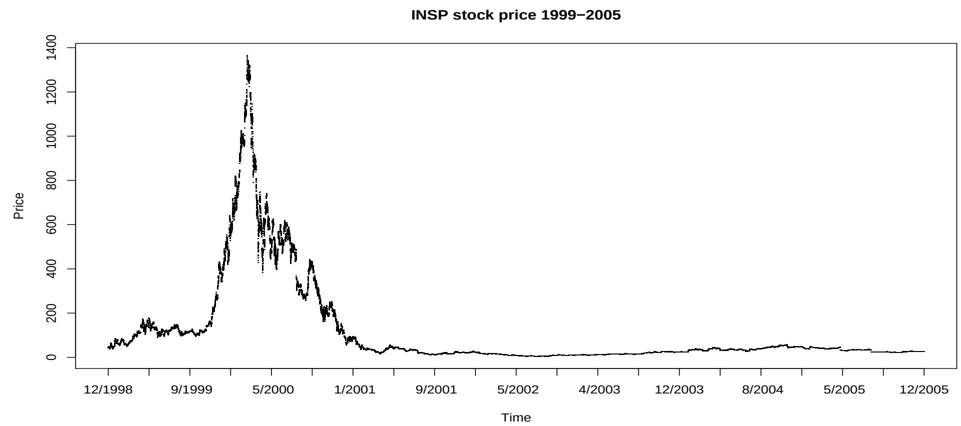

Figure 2.5.5: Historical price for INSP.

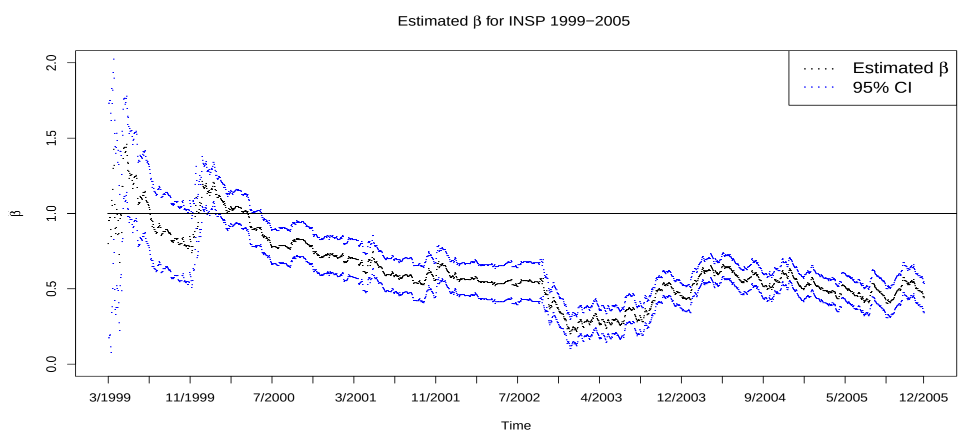

Figure 2.5.6: Instantaneous estimation $\hat{\beta}$ for INSP.

INSP stock price is shown in Figure 2.5.5. Applying the dynamic linear regression described in the previous section to the historical trading data for INSP, we obtain the instantaneous estimated $\hat{\beta}$, together with its 95% confidence interval as plotted in Figure 2.5.6 . In Figure 2.5.7 we color the price in red for stock INSP, if on that day the estimation $\hat{\beta} > 1$. For better illustration, we also take the log scale to the price process in Figure 2.5.8. As illustrated both in Figure 2.5.6 and Figure 2.5.7, the asset bubble in INSP which occurred shortly after it was issued was successfully detected. The extreme soaring period starting from the end of 1999 to the early 2000 is also well captured as a time range with bubbles.

With the dynamic regression model, we are also able to obtain the estimated distribution

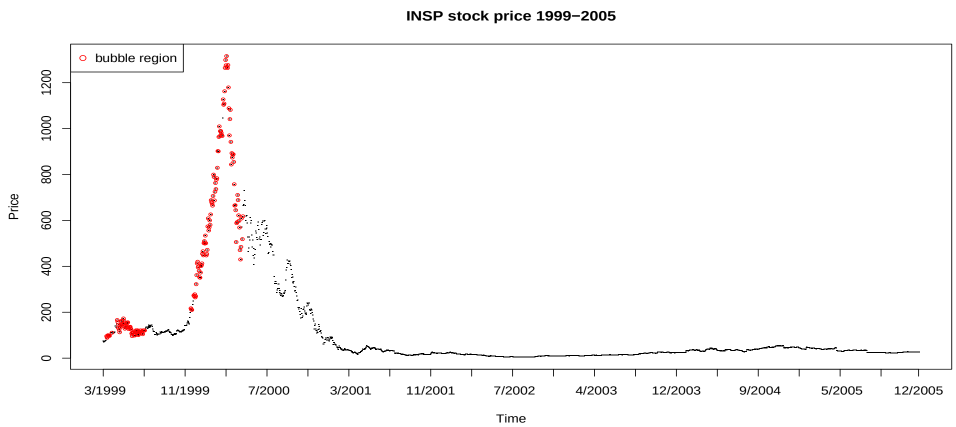

Figure 2.5.7: Historical price for INSP.

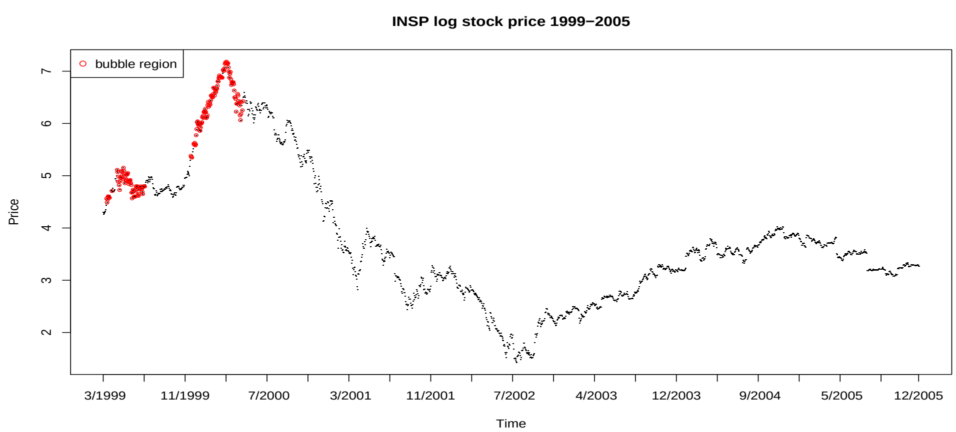

Figure 2.5.8: Historical log price for INSP.

for expected future return of a stock after time Δt, conditioning on the size of the current bubble will last for Δt. For details see Appendix A.1.

2.5.2 Details of the Data Analysis

The dataset is provided by Wharton Research Data Services. The original dataset is the recorded trade price. We only use the data during the market open time, from 9:30am to 16:00pm. To avoid microstructure noise in estimating volatility, we sample at the frequency of every 10min, which give us 40 price records per day. The Gaussian prior we used for the dynamic linear regression is $\theta_0 \sim N_2(m_0, C_0)$. We do not have much information on what θ_0 should be before analyzing the historical data, thus we chose a relatively flat prior, $m_0 = \begin{bmatrix} -5 \\ 0.8 \end{bmatrix}$ and $C_0 = \begin{bmatrix} 9 & 0 \\ 0 & 0.5 \end{bmatrix}$. For the variance of the noise term in the observation equation, we set $V = 2/(40-1)$, where 40 is the number of price observations each trading day. To estimate the covariance matrix $W = \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix}$ in the state equation, we first use the beginning 6 month price and estimated volatility data, roughly 120 pairs of data points. The inverse gamma distribution priors on W are $w_1 \sim Inv\_gamma(20, 0.001)$ and $w_2 \sim Inv\_gamma(20, 0.001)$. To run the Gibbs sampling procedure, the initial values for w_1 and w_2 are both 0.01. We notice that after roughly 50 iterations of sampling, the Markov chain for Gibbs sampling of w_1 and w_2 is already stable. Thus we run 2000 iterations and drop the first 50 pairs, taking the average of the remaining w_1, w_2 sample pairs as the estimates.

Chapter 3

The Asymptotic Error Distribution for the Euler Scheme with Locally Lipschitz Coefficients

Abstract

In traditional work on numerical schemes for solving stochastic differential equations (SDEs), it is usually assumed that the coefficients are globally Lipschitz. This assumption has been used to establish a powerful analysis of the numerical approximations of the solutions of stochastic differential equations. In practice, however, the globally Lipschitz assumption on the coefficients is on occasion too stringent a requirement to meet. Some Brownian motion driven SDEs used in applications have coefficients that are Lipschitz only on compact sets. Reflecting the importance of the locally Lipschitz case, it has been well studied in recent years, yet some simple to state, fundamental results remain unproved. We attempt to fill these gaps in this paper, establishing both a rate of convergence, but also we find the asymptotic normalized error process of the error process arising from a sequence of approximations. The result is analogous to the original result of this type, established in [51] back in 1991. This result was improved in 1998 in [38], and recently(2009) it was

partially extended in [60]. As we indicate, the results in our paper provide the basis of a statistical analysis of the error; in this spirit we give conditions for a finite variance.

Keywords: stochastic differential equation, locally Lipschitz, convergence in probability, Euler scheme, normalized error process, weak convergence

3.1 Introduction

We investigage the numerical solution of a one-dimensional stochastic differential equation (SDE) of the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,\ 0 \le t \le T,\ X_0 = x_0 \in \mathbb{R}. \tag{3.1.1}$$

Here $X_t \in \mathbb{R}$ for each t, $\mu, \sigma : \mathbb{R} \to \mathbb{R}$ are coefficient functions, and W is a one dimensional standard Brownian motion. We assume the initial value $x_0 \in \mathbb{R}$ is non-random. For background information about SDEs, we refer to Chapter 5 of Protter [65], Chapter 9 of Revuz and Yor [70] and Chapter 5 of Karatzas and Shreve [47].

In applications, one would often like to solve (3.1.1) numerically, as an explicit solution is usually not obtainable. This is often done in low dimensions using PDE methods that require heavy computational complexity. Hence, in practice, it is advisable to solve (3.1.1) with the simple Euler scheme. (See the survey paper of Talay [76] for a discussion of this issue). Our primary objective is to study uniform convergence in probability and weak convergence of the normalized error process for the Euler scheme under locally Lipschitz and no finite explosion assumptions on (3.1.1). Note that the result on uniform convergence in probability in this paper is not restricted to the Euler scheme, but applicable to all numerical schemes satisfying some mild assumptions.

The use of the Euler scheme to solve Brownian motion driven SDEs is already well studied. A number of treatments impose conditions on μ and σ in (3.1.1), and in particular a globally Lipschitz condition and/or a linear growth condition is imposed. We list some of the works here. For the rate of convergence of the expectation of functionals, see Talay and Tubaro [77]; for the rate of convergence of the distribution function, see Bally and Talay [5]; for the rate of convergence of the density, see Bally and Talay [6]; for error analysis, see Bally and Talay [4]; for an Euler scheme when one has irregular coefficients and Hölder continuous coefficients see Yan [81], and in this regard see also Bass-Pardoux [7]; for complete reviews, see Talay [78] and Kloeden-Platen [48]. In two interesting recent papers M. Bossy et al [11], [9] have studied a modified (symmetrized) Euler scheme to handle solutions of the Cox-Ingersoll-Ross type (CIR), but for equations where the diffusive coefficient is of the form $|x|^\alpha$ for $\frac{1}{2} \leq \alpha < 1$, which are of course locally Lipschitz.

There is also some work on numerical schemes for solving SDEs not tied to Brownian motion, but rather driven by semimartingales with jumps. The case of SDEs driven by Brownian motion and Lebesgue measure can be found in Kurtz and Protter [51] where the convergence in distribution of the normalized Euler scheme is first studied. L^p estimates of the Euler scheme error were given by Kohatsu-Higa and Protter [49]. Protter and Talay [64] also studied the Euler scheme for SDEs driven by Lévy processes. Jacod and Protter [38] obtained a (to date) definitive result about the asymptotic error distributions for the Euler scheme solving SDEs driven by a vector of semimartingales. More recent work has focused on numerical schemes to solve SDEs under relaxed conditions on the coefficients, to wit the locally Lipschitz condition replaces the customary Lipschitz condition. Under the locally Lipschitz hypothesis, the Euler scheme may diverge in the strong sense of convergence, such as L^p. The L^p convergence, or more correctly the lack of it, is studied in Hutzenthaler, Jentzen and Kloeden [36]. To obtain convergence results for the Euler scheme under the locally Lipschitz condition, additional assumptions are assumed in existing work. Examples of attempts are assuming the existence of a Lyapunov function, or a one sided Lipschitz condition and finite moments of the true solution and the numerical solution (see [29],[32],[56]).

Convergence in probability for Euler-type schemes in general still holds, see Hutzenthaler and Jentzen [37] and the citations therein. Under the condition that μ, σ are continuously differentiable ($\mathcal{C}^1$) and grow at most linearly, Kurtz and Protter [51] obtained the limit distribution for the asymptotic normalized error process for the Euler scheme. Neuenkirch and Zähle [60] generalized the result of Kurtz and Protter by assuming the solution never leaves an open set in finite time and that the coefficients are $\mathcal{C}^1$.

In this paper, we study the limit distribution for the asymptotic normalized error process with only a locally Lipschitz assumption plus no finite time explosions, and σ in (3.1.1) being bounded away from 0. By relaxing the $\mathcal{C}^1$ and linear growth hypotheses to the assumption that the coefficient need only be locally Lipschitz, we are able to deal with coefficients that may have super linear growth, and their derivatives may have poor smoothness properties, or may not even exist.

Some locally Lipschitz coefficients lead to well defined stochastic differential equations, but only because the solution remains always positive. This is the case for example with the CIR type processes. The Euler scheme approximations, however, need not be defined, since for example we might be taking the square root of a negative quantity at some steps. For these situations, we can use a nice trick due to Bossy at al [9, 11] where the Euler scheme is replaced by what is known as a symmetrized Euler scheme. This keeps the approximations positive, too. Our results apply for these schemes as well, since they are "local", which is our rubric for the types of schemes we utilize in this chapter.

This chapter is organized as follows. Section 2 briefly reviews existing work. In Section 3, we prove that if a numerical scheme converges uniformly in probability on any compact time interval with a certain rate under the globally Lipschitz condition, then the same result holds when the globally Lipschitz condition is replaced with a locally Lipschitz condition and a no finite time explosion condition. The Euler and Milstein schemes are studied as examples. From Section 4 on, we focus on the Euler scheme. We prove that the sequence of the error process for the Euler scheme normalized by $\sqrt{n}$ is relatively compact. Furthermore, by proving uniqueness of the limit process, we show the normalized error process converges in law. The limit error process is also provided as a solution to an SDE. This is not surprising, given the results of [51]. Section 5 turns to a study on the the second moment of the weak limit process and its running maximum. In the last section, we give an upper bound for the rate of weak convergence for the approximating expectation of functionals for the Euler scheme.

3.2 A Brief Review of Existing Work

3.2.1 The Globally Lipschitz Case

We start with some notation. For a discretization of the time interval $[0,T]$ with discretization size $\frac{T}{n}$, let $n(t) = [\frac{nt}{T}]\frac{T}{n}$, the nearest left time grid point for t. For all $g : [0,T] \to \mathbb{R}$, define

$$\Delta g_t^{(n)} = g(t) - g(n(t)). \tag{3.2.1}$$

The continuous Euler scheme for solving SDE (3.1.1) is defined by

$$X_t^{E,n} = X_{n(t)}^{E,n} + \sigma(X_{n(t)}^{E,n})\Delta W_t^{(n)} + \mu(X_{n(t)}^{E,n})\Delta t^{(n)},\ X_0^{E,n} = X_0, \tag{3.2.2}$$

and the continuous Milstein scheme is defined by

$$\begin{aligned} X_t^{M,n} &= X_{n(t)}^{M,n} + \sigma(X_{n(t)}^{n})\Delta W_t^{(n)} + \mu(X_{n(t)}^{n})\Delta t^{(n)} + \frac{1}{2}\sigma(X_{n(t)}^{M,n})\sigma'(X_{n(t)}^{M,n})[(\Delta W_t^{(n)})^2 - \Delta t^{(n)}] \\ &\quad + \frac{1}{2}\mu(X_{n(t)}^{M,n})\mu'(X_{n(t)}^{M,n})(\Delta t^{(n)})^2, \quad X_0^{M,n} = X_0. \end{aligned}$$

Without further specification, in this chapter, X^n represents the numerical solution from the continuous Euler scheme with step size $\frac{T}{n}$ on $[0,T]$.

The convergence properties of the Euler scheme for solving SDEs with globally Lipschitz coefficients have been widely studied. The following Theorem is from the review paper of Talay [78].

Theorem 3.2.1. *Consider the SDE (*3.1.1*) and suppose the coefficients μ, σ are globally Lipschitz. Let X_n be the numerical solution from the Euler scheme, then* $\forall\ 0 < T < \infty$

(i) X^n *converges to* X *pathwise almost surely and*

$$\forall \alpha < \frac{1}{2},\ n^{\alpha} \sup_{0\le t\le T} |X_t - X_t^n| \to 0,\ a.s.;$$

(ii) There exists a positive constant C*, increasing with* T*, such that,*

$$\sup_{0\le t\le T} \mathbb{E}(|X_t - X_t^n|^2) \le \frac{C}{n};$$

(iii) For $f \in \mathcal{C}^{\infty}$*, and at most polynomial growth, there exists a positive constant* D*, increasing with* T*, such that,*

$$|\mathbb{E}f(X_T) - \mathbb{E}f(X_T^n)| \le \frac{D}{n}.$$

For the asymptotic distribution for the normalized error process, Kurtz and Protter [51] proved that

Theorem 3.2.2 (Kurtz and Protter)**.** *Consider the SDE (*3.1.1*) and suppose the coefficients* μ, σ *are* C^1 *and bounded. Let* X_n *be the numerical solution from the Euler scheme, and* $U^n = \sqrt{n}(X^n - X)$*. Then* U^n *converges in law to a limiting process* U*, which is the unique solution to the following linear equation:*

$$U_t = \int_0^t \mu'(X_s)U_s ds + \int_0^t \sigma'(X_s)U_s dW_s + \frac{\sqrt{2}}{2}\int_0^t \sigma(X_s)\sigma'(X_s)dB_s,\ U_0 = 0.$$

where B_t *is a standard Brownian motion independent of* W_t *in a extended space.*

3.2.2 The Locally Lipschitz Case

Hutzenthaler, Jentzen and Kloeden [36] have studied the strong and weak divergence of the Euler Scheme for solving SDEs with coefficients that are not globally Lipschitz.

Theorem 3.2.3 (Hutzenthaler, Jentzen and Kloeden)**.** *Consider the SDE (*3.1.1*) and sup-*

pose $\mathbb{P}(\sigma(X) \neq 0) > 0$ and let $C \geq 1,\ \beta > \alpha > 1$ be constants such that

$$\max(\,|\mu(x)|\,,|\sigma(x)|) \geq \frac{|x|^{\beta}}{C} \text{ and } \min(\,|\mu(x)|\,,|\sigma(x)|) \leq C|x|^{\alpha}$$

for all $|x| \geq C$. Then, if the exact solution X satisfies $\mathbb{E}[|X_T|^p] < \infty$ for one $p \in [1, \infty)$,

$$\lim_{n\to\infty} \mathbb{E}[|X_T - X_T^n|^p] = \infty \text{ and } \lim_{n\to\infty} \mathbb{E}[|X_T^n|^p] = \infty,$$

where X^n is the numerical solution for solving SDE (3.1.1) from the Euler scheme.

To ensure different types of convergence of the Euler scheme with locally Lipschitz condition, additional assumptions are required. For convergence in probability we refer to Marion, Mao and Renshaw [56]. The authors deal with the multidimensional case.

$$dX_t = \hat{\mu}(X_t)dt + \hat{\sigma}(X_t)dB_t,\ X_0 = x_0, \tag{3.2.3}$$

where $x = (x_1, \dots x_d), \hat{\mu} = (\mu_1(x), \dots \mu_d(x)), \hat{\sigma}(x) = (\sigma_{ij}(x)_{d\times m}),\ B$ is a m-dimensional Brownian motion defined on a given complete probability space $(\Omega, \mathcal{F}, P)$ with a filtration satisfying the usual conditions. The solution takes value in $G \subseteq R^d$.

Theorem 3.2.4 (Marion, Mao and Renshaw)**.** *Consider SDE (3.2.3). Let G be an open set of R^d, $x_0 \in G$ and the solution $X_t \in G$. Let X^n be the numerical solution from the Euler scheme. Suppose the following conditions are satisfied*

(i) $\hat{\mu}(X_t)$ and $\hat{\sigma}(X_t)$ are locally Lipschitz;

(ii) there exits a C^2-function $V : G \to R_+$ such that $\{x \in G : V(x) \leq r\}$ is compact for any $r > 0$;

(iii) $LV(x) \leq K(1 + V(x))$ where

$$LV(x) = V_x(x)\hat{\mu}(x) + \frac{1}{2} trace\Big[\hat{\sigma}^T(x)V_{xx}\hat{\sigma}(x)\Big];$$

(iv) there exists a positive constant $K_3(\mathcal{D})$ such that for all $x, y \in D$

$$\left|V(x) - V(y)\right| \vee \left|V_x(x) - V_x(y)\right| \vee \left|V_{xx}(x) - V_{yy}(y)\right| \leq K_3(\mathcal{D})|x - y| \,.$$

Assume the Euler scheme is well defined in this case. Then for any $\epsilon,\ \delta > 0$, there exists $n' > 0$ such that

$$\mathbb{P}_n(\delta) = \mathbb{P}(\sup_{0 \leq t \leq T} |X^n - X|^2 \geq \delta) \leq \epsilon, \tag{3.2.4}$$

provided that $n > n'$ and the initial value $x_0 \in G$.

In [56], the authors also provided an upper bound for $P_n(\delta)$. It is later mentioned in Hutzenthaler and Jentzen [37], to ensure convergence in probability, one only needs the locally Lipschitz condition on $\hat{\mu}$ and $\hat{\sigma}$ and the solution $X_t \in G$ exists on $[0, T]$. Convergence almost surely for the Euler scheme is studied in Gyöngy [29] who takes up the multi-dimensional case with time dependent coefficient $\hat{\mu}, \hat{\sigma}$. We present the result for the case that $\hat{\mu}$ and $\hat{\sigma}$ not depending on time.

Theorem 3.2.5 (Gyöngy)**.** *Consider SDE (*3.2.3*). Let D be an open set of $\mathbb{R}^d$, $x_0 \in D$ and the solution $X_t \in D$. Let X^n be the numerical solution solving (*3.2.3*) from the Euler scheme. Suppose the following conditions are satisfied:*

(i) $\hat{\mu}$, $\hat{\sigma}$ are locally Lipschitz and there exists an increasing sequence of bounded domains $\{D_k\}_{k=1}^{\infty}$ such that $\cup_{k=1}^{\infty} D_k = D$, and for every $k, t \in [0, k]$

$$\sup_{x \in D_k} \left|\hat{\mu}(x)\right| \leq M_k; \ \sup_{x \in D_k} \left|\hat{\sigma}(x)\right|^2 \leq M_k,$$

where M_k is a constant;

(ii) there exists a nonnegative function $V \in C^2$ such that

$$LV(x) \leq V(x), \forall t \in [0, T], x \in D,$$

$$V_k(T) := \inf_{x \in \partial D_k, t \leq T} V(x) \to \infty,$$

as $k \to \infty$ for every finite T, where $M = M(T)$ is a constant, ∂D_k denotes the boundary of D_k, and L is the differential operator

$$LV(x) = V_x(x)\hat{\mu}(x) + \frac{1}{2}trace\Big[\hat{\sigma}^T(x)V_{xx}\hat{\sigma}(x)\Big].$$

Then for every $\gamma < \frac{1}{2}$ and $T > 0$, there is a finite random variable η such that

$$\sup_{0 \le t \le T} |X^n - X|^2 \le n^{-\gamma} \ a.s..$$

Strong convergence is studied in Higham, Mao and Stuart [32] who consider the one dimensional case:

Theorem 3.2.6 (Higham, Mao and Stuart)**.** *Consider the SDE (*3.1.1*). Let X^n be the numerical solution solving (*3.1.1*) from the Euler scheme. If μ and σ are locally Lipschitz and for some $p > 2$ there is constant A such that*

$$\mathbb{E}\Big[\sup_{0 \le t \le T} |X_t|^p\Big] \vee \mathbb{E}\Big[\sup_{0 \le t \le T} |X_t^n|^p\Big] \le A,$$

then

$$\lim_{n \to \infty} \mathbb{E}\Big[\sup_{0 \le t \le T} |X_t^n - X_t|^2\Big] = 0.$$

3.3 Convergence in Probability

Under the globally Lipschitz condition, most of the proposed numerical schemes including the Euler and Milstein schemes have been proved to converge uniformly in probability at a finite time point. Fortunately, the same result can be extended to the locally Lipschitz case if one also adds a no finite time explosion condition. To prove this, we need a localization technique. Let us start with some notation.

Notation 3.3.1. *Given a process Z, we denote by $\mathbb{T}^m(Z) = \inf\{t \ge 0 : |Z_t| > m\}$. Also,* *we denote by $Z^{\mathbb{T}}$ the stopped process.*

In what follows, we denote by $X = X(x_0, \mu, \sigma, W)$ the unique solution of the SDE (3.1.1), where the coefficients μ, σ are assumed regular enough to have a unique strong solution (for example locally Lipschitz). For every $m \geq 1$ consider $\mu^{(m)}$ a continuous modification of μ such that $\mu(x) = \mu^{(m)}(x)$ for $|x| \leq m$, $\mu^{(m)}(x) = \mu(m+1)$ for $x \geq m+1$ and $\mu^{(m)}(x) = \mu(-m-1)$ for $x \leq -m-1$. In case the numerical procedure assumes that μ is $\mathcal{C}^k$ (or Lipschitz) we interpolate $\mu^{(m)}$ on $(-m-1, -m) \cup (m, m+1)$ in such a way that $\mu^{(m)}$ is also $\mathcal{C}^k$ (respectively Lipschitz). Similarly, we denote by $\sigma^{(m)}$ a modification of σ. Given a numerical procedure ϕ, we denote by $(X^{\phi,n})_n = (X^{\phi,n}(x_0, \mu, \sigma, W))_n$ the associated sequence of approximations. We remove the dependence on ϕ in $X^{\phi,n}$ when there is no possible confusion. Note that we use the same Brownian motion for every n. This numerical procedure is assumed **local** in the following sense. Assume that $\mu = \tilde{\mu}, \sigma = \tilde{\sigma}$ on the interval $[-m, m]$, where $|x_0| < m$. Then for all n and for $\mathbb{T} = \mathbb{T}^m(X^{\phi,n}(x_0, \mu, \sigma, W))$ it holds

$$(X^{\phi,n}(x_0, \mu, \sigma, W))^{\mathbb{T}} = (X^{\phi,n}(x_0, \tilde{\mu}, \tilde{\sigma}, W))^{\mathbb{T}}$$

almost surely. In particular $\mathbb{T}^m(X^{\phi,n}(x_0, \mu, \sigma, W)) = \mathbb{T}^m(X^{\phi,n}(x_0, \tilde{\mu}, \tilde{\sigma}, W))$ a.s.. This hypothesis is satisfied, for example, by the Euler and Milstein schemes. On the other hand, if (μ, σ) and $(\tilde{\mu}, \tilde{\sigma})$ are regular, the associated solutions satisfy

$$(X(x_0, \mu, \sigma, W))^{\mathbb{T}} = (X(x_0, \tilde{\mu}, \tilde{\sigma}, W))^{\mathbb{T}}$$

almost surely for $\mathbb{T} = \mathbb{T}^m(X(x_0, \mu, \sigma, W))$. Again, we have $\mathbb{T}^m(X(x_0, \mu, \sigma, W)) = \mathbb{T}^m(X(x_0, \tilde{\mu}, \tilde{\sigma}, W))$ a.s.. Now, we present Theorem 3.3.1.

Theorem 3.3.1. *Assume that a numerical scheme ϕ is well defined and local. Let X_t^n be the numerical solution using ϕ for the SDE (3.1.1) on $[0, T]$. If X_t^n converges in probability uniformly on $[0, T]$, with order $\alpha > 0$, that is $\forall\ C > 0$*

$$\mathbb{P}(n^\alpha \sup_{0 \leq t \leq T} |X_t^n - X_t| > C) \to 0, \quad \textit{as } n \to +\infty, \tag{3.3.1}$$

given μ, σ are globally Lipschitz, then (3.3.1) *also holds when the globally Lipschitz condition is replaced with a locally Lipschitz condition and a no finite time explosion condition.*

Proof. In what follows, to avoid overly burdensome notation, we denote by

$$X = X(x_0, \mu, \sigma, W),\ X^n = X^n(x_0, \mu, \sigma, W),$$

$$Y^{(m)} = X(x_0, \mu^{(m)}, \sigma^{(m)}, W),\ Y^{n,(m)} = X^n(x_0, \mu^{(m)}, \sigma^{(m)}, W)\,.$$

Now, define

$$\mathscr{X}_n = \left\{ \omega : n^\alpha \sup_{0 \le t \le T} |X_t^n - X_t| \le C \right\},\ \mathscr{Y}_{n,(m)} = \left\{ \omega : n^\alpha \sup_{0 \le t \le T} |Y_t^{n,(m)} - Y_t^{(m)}| \le C \right\}.$$

We also consider $\mathbb{T} = \mathbb{T}^m(X), \mathcal{S} = \mathbb{T}^{m-1}(X)$. It is clear that $X^{\mathcal{S}} = (Y^{(m)})^{\mathcal{S}}$ (actually they are equal up to time $\mathbb{T}$). Since the numerical procedure is local, also $\mathcal{U} = \mathbb{T}^m(X^n) = \mathbb{T}^m(Y^{n,(m)})$ and

$$(X^n)^{\mathcal{U}} = (Y^{n,(m)})^{\mathcal{U}}\,.$$

Consider m large enough such that $|x_0| < m - 1$ and n large enough such that $C/n^\alpha < 1$. For these values of n, m, we show that a.s.

$$\mathscr{Y}_{n,(m)} \bigcap \{\mathcal{S} > T\} \subset \mathscr{X}_n\,.$$

Indeed, on the set $\{\mathcal{S} > T\}$ the two processes $X, Y^{(m)}$ agree on $[0, T]$ a.s.. In particular, we have that $\sup_{0 \le t \le T} |Y_t^{(m)}| = \sup_{0 \le t \le T} |X_t| \le m - 1$ a.s. On the other hand on the set $\mathscr{Y}_{n,(m)} \bigcap \{\mathcal{S} > T\}$ we have a.s.

$$\sup_{0 \le t \le T} |Y_t^{n,(m)}| \le m - 1 + \frac{C}{n^\alpha} < m\,.$$

That is $\mathbb{T}^m(Y^{n,(m)}) > T$ a.s.. Since ϕ is local, we deduce that on $[0, T]$ the processes X^n

and $Y^{n,(m)}$ agree a.s. and therefore on $\mathscr{Y}_{n,(m)} \bigcap \{\mathcal{S} > T\}$

$$\sup_{0 \le t \le T} |X^n - X_t| = \sup_{0 \le t \le T} |Y^{n,(m)} - Y_t^{(m)}| \le \frac{C}{n^\alpha}$$

holds also a.s., proving the desired inclusion. This shows that the inequality

$$\mathbb{P}\left(n^\alpha \sup_{0 \le t \le T} |X_t^n - X_t| > C\right) \le \mathbb{P}\left(n^\alpha \sup_{0 \le t \le T} |Y_t^{n,(m)} - Y_t^{(m)}| > C\right) + \mathbb{P}(\mathbb{T}^{m-1}(X) \le T)\,.$$

Now, given $\epsilon > 0$ choose m large such that $\mathbb{P}(\mathbb{T}^{m-1}(X) \le T) \le \epsilon/2$. For that m, according to the hypothesis of the Theorem, there exists $n_0 = n_0(m, \epsilon)$, such that for all $n \ge n_0$

$$\mathbb{P}\left(n^\alpha \sup_{0 \le t \le T} |Y_t^{n,(m)} - Y_t^{(m)}| > C\right) \le \frac{\epsilon}{2}\,,$$

giving the result. □

Taking $\alpha = 0$ in Theorem 3.3.1, we immediately see that if a numerical scheme converges in probability uniformly on compact time intervals for solving SDEs with the globally Lipschitz coefficients, then the same result also holds under the locally Lipschitz condition and no finite time explosion condition.

We illustrate the application of Theorem 3.3.1 using the two most widely used numerical schemes, the Euler scheme and the Milstein scheme. Under the globally Lipschitz condition, it is well known that the continuous Euler and Milstein schemes converge in probability uniformly on compact time intervals with any order between $[0, \frac{1}{2})$ and $[0, 1)$ respectively. Interested readers can refer to [75] and [81] for details. Then applying Theorem 3.3.1 leads to the following corollary.

Corollary 3.3.1. *Consider SDE (*3.1.1*), assume μ, σ are locally Lipschitz and the solution has no finite time explosion and the continuous Euler scheme $X^{E,n}$ and the continuous Milstein scheme $X^{M,n}$ are well defined for solving (*3.1.1*). Then $X^{E,n}$ and $X^{M,n}$, converge* *in probability uniformly to* X *on* $[0,T]$*. Moreover* $\forall\ \gamma \in (0, \frac{1}{2}]$*, we have*

$$\mathbb{P}(n^{\frac{1}{2}-\gamma} \sup_{0 \le t \le T} \Big| X^{E,n} - X_t \Big| > C) \to 0, \quad as\ n \to +\infty,$$
$$\mathbb{P}(n^{1-\gamma} \sup_{0 \le t \le T} \Big| X^{M,n} - X_t \Big| > C) \to 0, \quad as\ n \to +\infty.$$

3.4 Asymptotic Error Distribution for the Euler Scheme

In this section, we are going to prove that the asymptotic normalized error process from the Euler scheme converges in distribution with rate $\sqrt{n}$, under the locally Lipschitz and no finite time explosion assumption. In the previous section, the localization technique used in the proof for Theorem 3.3.1 transfers the locally Lipschitz case into the globally Lipschitz case. The localization technique will be used in this section as well, and we present Proposition 3.4.1 to make future proofs concise when applying this technique.

Proposition 3.4.1. *Consider the SDE (*3.1.1*), assume that* μ *and* σ *are locally Lipschitz and the solution* X *has no finite time explosion. For* $m > |x_0|$*, define* $Y^{(m)}$ *as in the proof of Theorem* 3.3.1*. Let* X^n*,* $Y^{n,(m)}$ *be the numerical solutions from a numerical scheme* ϕ*. Assume* ϕ *is well defined, local and converges uniformly in probability on compact time interval. Then* $\forall\ 0 < T < \infty$*,*

$$\lim_{m \to \infty} \sup_n \mathbb{P}\Big(X \neq Y^{(m)}\ or\ X^n \neq Y^{n,(m)} \Big) = 0,\ \ on\ [0,T].$$

Proof. Since the Euler scheme is local,

$$\mathbb{T}^m(X) = \mathbb{T}^m(X^{(m)}),\ \mathbb{T}^m(X^n) = \mathbb{T}^m(X^{n,(m)}). \tag{3.4.1}$$

Thus, we have on $[0,T]$

$$\mathbb{P}\Big(X \neq Y^{(m)} \text{ or } X^n \neq Y^{n,(m)} \Big) = \mathbb{P}\Big(\mathbb{T}^m(X) < T \text{ or } \mathbb{T}^m(X^n) < T \Big). \tag{3.4.2}$$

Since X has no finite time explosion, $\forall\epsilon > 0$, there exists $m_1 = m_1(\epsilon)$ large enough so that

$$\mathbb{P}(\mathbb{T}^{m_1}(X) \leq T) < \frac{\epsilon}{3}.$$

By the uniform convergence in probability of X^n on $[0, T]$, there exists $n' = n'(\epsilon)$ such that $\forall n > n'$, we have

$$\mathbb{P}(\sup_{0\leq s\leq T} |X^n_s - X_s| \geq 1) < \frac{\epsilon}{3}$$

and

$$\mathbb{P}(\mathbb{T}^{m_1+1}(X^n) \leq T) \leq \mathbb{P}(\mathbb{T}^{m_1}(X) \leq T) + \mathbb{P}(\sup_{0\leq s\leq T} |X^n_s - X_s| \geq 1) < \frac{2}{3}\epsilon.$$

Hence when $n > n'$,

$$\mathbb{P}(\mathbb{T}^{m_1+1}(X^n) \leq T \text{ or } \mathbb{T}^{m_1+1}(X) \leq T) \leq \mathbb{P}(\mathbb{T}^{m_1}(X) \leq T) + \mathbb{P}(\mathbb{T}^{m_1+1}(X^n) \leq T) < \epsilon.$$

Now for $n \leq n'$, take $Q^\star_t = \max\left\{X^*_t, X^{1,*}_t, X^{2,*}_t \ldots X^{n',*}_t\right\}$, where X^* indicates the running maximum of the absolute process. As the true solution X and the Euler scheme numerical solutions have no explosion on $[0, T]$, and n' is finite, the process $Q^\star_t$ has no explosion either. We can always find $m = m(n') > m_1 + 1$ large enough so that for the hitting time of $Q^\star_t$, $\mathbb{P}(\mathbb{T}^m(Q^\star_t) \leq T) \leq \epsilon$. Together with (3.4.2), the proof concludes. □

Remark 3.4.1. *Proposition* 3.4.1 *also holds for the multidimensional case with the same techniques used in the proof.*

Remark 3.4.2. *Kurtz and Protter [51] obtained the weak limit for the sequence of normalized error process for the Euler scheme under the condition that* μ, σ *are* $\mathcal{C}^1$ *and of at most linear growth. Proposition* 3.4.1 *implies that the condition can be replaced with* μ, σ *are* $\mathcal{C}^1$ *and at most linear growth condition on any compact set plus the solution has no finite time explosion. This generalization can be found in Neuenkirch and Zähle [60].*

Remark 3.4.3. *In Yan [?], the $\mathcal{C}^2$ and at most growth hypotheses used for obtaining for the weak limit of the sequence of normalized error process for the Mistein scheme, can be replaced with $\mathcal{C}^2$ and at most linear growth conditions on any compact set, plus that the solution has no finite time explosion.*

Now we turn to prove a weak convergence result for the normalized error of the Euler scheme with the locally Lipschitz assumption. Define Z^n as follows

$$Z_t^{n11} = \int_0^t \sqrt{n}\Delta s^{(n)} ds,\ Z_t^{n12} = \int_0^t \sqrt{n}\Delta s^{(n)} dW_s,$$
$$Z_t^{n21} = \int_0^t \sqrt{n}\Delta W_s^{(n)} ds,\ Z_t^{n22} = \int_0^t \sqrt{n}\Delta W_s^{(n)} dW_s.$$

Our goal is to prove convergence in distribution for the asymptotic error process from the Euler scheme at the rate $\sqrt{n}$, which requires Z^n to converge in distribution.

Proposition 3.4.2. *The sequence Z^n is tight and converges in distribution to Z under the uniform topology on compact time set, where Z is independent of W and $Z^{1,1} = Z^{1,2} = Z^{2,1} = 0$, $\sqrt{2}Z^{2,2}$ is a standard Brownian motion.*

Proposition 3.4.2 is implied by Theorem 5.1 in Jacod and Protter [38].

Proposition 3.4.3. *Consider SDE (3.1.1), and assume that $\mu(x), \sigma(x)$ are both Lipschitz and bounded. Let X^n be the numerical solution to (3.1.1) on $[0,T]$ from the continuous Euler scheme with step size $\frac{T}{n}$. Then the sequence of normalized error processes $U_n = \sqrt{n}(X^n - X)$ is relatively compact.*

Proof. It has been proved that Z^n are good sequences (see [50] for the definition of a good sequence). From Proposition 3.4.2, $Z^n \Rightarrow Z$, where $\Rightarrow$ denotes convergence in distribution under the uniform topology on a compact time set. The limit process Z is independent of W and $Z^{1,1} = Z^{1,2} = Z^{2,1} = 0$, $Z^{2,2}$ is mean zero Brownian motion with $\mathbb{E}[(Z_t^{2,2})^2] = \frac{t}{2}$. By Corollary 3.3.1, we also have $(X^n, Z^n) \Rightarrow (X, Z)$. By the definition of a continuous Euler scheme, X^n can also be represented as

$$X_t^n = \int_0^t \mu(X_{n(s)}^n)ds + \int_0^t \sigma(X_{n(s)}^n)dW_s.$$

Then

$$\begin{aligned} U_t^n &= \sqrt{n}(X_t^n - X_t) \\ &= \int_0^t \sqrt{n}\{\mu(X_{n(s)}^n) - \mu(X_s)\}ds + \int_0^t \sqrt{n}\{\sigma(X_{n(s)}^n) - \sigma(X_s)\}dW_s. \end{aligned}$$

For $x \neq y$, define functions $g, h : \mathbb{R}^2 \to \mathbb{R}$ as

$$g(x,y) = \frac{\mu(x) - \mu(y)}{x - y}, \quad h(x,y) = \frac{\sigma(x) - \sigma(y)}{x - y}.$$

Since μ, σ are Lipschitz, $g(x,y)$ and $h(x,y)$ are bounded. Now we separate the error process into two terms $U^n = U^{1,n} + U^{2,n}$, where

$$\begin{aligned} U_t^{1,n} &= \int_0^t \sqrt{n}\{\mu(X_{n(s)}^n) - \mu(X_s)\}ds \\ &= \int_0^t \sqrt{n}\{\mu(X_s^n) - \mu(X_s)\}ds - \int_0^t \sqrt{n}\{\mu(X_s^n) - \mu(X_{n(s)}^n)\}ds \\ &= \int_0^t g(X_s^n, X_s)U_t^n ds - \int_0^t \frac{\mu(X_s^n) - \mu(X_{n(s)}^n)}{X_s - X_{n(s)}^n}(X_s^n - X_{n(s)}^n)\sqrt{n}ds. \end{aligned}$$

Note that $X_s^n - X_{n(s)}^n = \mu(X_{n(s)}^n)\Delta s^{(n)} + \sigma(X_{n(s)}^n)\Delta W_s^{(n)}$. Then

$$U_t^{1,n} = \int_0^t g(X_{n(s)}^n, X_s)U_t^n - g(X_s^n, X_{n(s)}^n)\left\{\mu(X_{n(s)}^n)\sqrt{n}\Delta s^{(n)} + \sigma(X_{n(s)}^n)\sqrt{n}\Delta W_s^{(n)}\right\}ds.$$

Similarly,

$$U_t^{2,n} = \int_0^t h(X_{n(s)}^n, X_s)U_t^n - h(X_s^n, X_{n(s)}^n)\left\{\mu(X_{n(s)}^n)\sqrt{n}\Delta s^{(n)} + \sigma(X_{n(s)}^n)\sqrt{n}\Delta W_s^{(n)}\right\}dW_s.$$

For notational convenience, define $\tilde{f}^n$ as

$$\tilde{f}^n = \left[g(X_s^n, X_s), g(X_s^n, X_{n(s)}^n), h(X_s^n, X_s), h(X_s^n, X_{n(s)}^n)\right].$$

If μ, σ are also assumed to be continuously differentiable, as in Kurtz and Protter [51], then $\tilde{f}^n$ converges weakly uniformly to $[\mu'(X), \mu'(X), \sigma'(X), \sigma'(X)]$ on $[0, T]$. By results on weak convergence of stochastic integrals in Kurtz and Protter [50], U^n converges weakly uniformly on $[0, T]$ as well.

However, here $\sigma,\ \mu$ are only assumed to be Lipschitz and bounded, hence their derivatives might not be continuous or not even exist. This would cause $\tilde{f}^n$ to fail to converge weakly. Fortunately, by the boundedness of $\tilde{f}^n$, applying weak convergence techniques in [50] would give relative compactness of U^n under the uniform topology, which is shown in the following steps.

By Prokhorov's Theorem which states that tightness is equivalent to relative compactness in our case, $\tilde{f}^n$ is also relatively compact. Then for every subsequence of $\tilde{f}^n$, there exists a further subsubsequence n_k such that $\tilde{f}^{n_k}$ converges weakly uniformly on $[0, T]$. It is also known that $(X_{n(.)}^n, X^n, \sqrt{n}Z^n) \Rightarrow (X, X, Z)$, and the sequence is a good sequence (see [51] for details). Then, we can assume on $[0, T]$,

$$\begin{aligned}&\left[f^{n_k,1}, f^{n_k,2}, f^{n_k,3}, f^{n_k,4}, X_{n_k(.)}^{n_k}, X^{n_k}, Z^{n_k,1}, Z^{n_k,2}, Z^{n_k,3}, Z^{n_k,4}\right] \\ &\Rightarrow [G, \tilde{G}, H, \tilde{H}, X, X, 0, 0, 0, \frac{\sqrt{2}}{2}B].\end{aligned}$$

Since Z^n is a good sequence and μ, σ are bounded, then, by proof of Theorem 3.5 in Kurtz and Protter [51], $U^{n_k} \Rightarrow R$ on $[0, T]$, where

$$R_t = \int_0^t G_t R_t ds + \int_0^t H_t R_t dW_s + \frac{\sqrt{2}}{2} \int_0^t \sigma(X_t) \tilde{H}_t dB_s. \tag{3.4.3}$$

Thus every subsequence of $U^n = \sqrt{n}(X^n - X)$ has a subsubsequence that converges weakly uniformly on $[0, T]$, implying that U^n is relatively compact. □

Remark 3.4.4. *Our next theorem is similar to results in [51, 60] but with two important differences: We do not assume the coefficients are $\mathcal{C}^1$, but only that they are locally Lipschitz; We do not assume a linear growth condition, but rather assume only locally Lipschitz combined with no finite explosions in finite time. As a simple example, this allows for the consideration of coefficients of the form $\sigma(x) = x^\gamma$, with $\gamma > 1$. In Economics, such coefficients are known as CEV (= Constant Elasticity of Variance). Usually γ is assumed to be less than or equal to one, but here we lay the groundwork to consider $\gamma > 1$ on a practical level.*

Theorem 3.4.1. *Consider the SDE (3.1.1), assume that μ, σ are locally Lipschitz and that the solution X has no finite time explosion. Further assume that $\sigma(x)$ is non-negative and bounded from below by some $d \in \mathbb{R}^+$ on any compact set. Let $\mu'(x), \sigma'(x)$ equal the derivatives of μ, σ at x when the derivatives exist; and that they equal 0, when the derivatives at a point x do not exist.*

Let X^n be numerical solution from the continuous Euler scheme, and $U_n = \sqrt{n}(X^n - X)$ be the normalized error process. Then for all $0 < T < \infty$, U_n converges weakly uniformly on $[0, T]$ to U, where U satisfies

$$U_t = \int_0^t \mu'(X_s)U_s ds + \int_0^t \sigma'(X_s)U_s dW_s + \frac{\sqrt{2}}{2}\int_0^t \sigma(X_s)\sigma'(X_s)dB_s,\ U_0 = 0, \tag{3.4.4}$$

where B is a standard Brownian motion and is independent of W.

Proof. Use Proposition 3.4.1 and apply the localization technique, we can assume μ, σ are bounded and globally Lipschitz and there exists $d > 0$, such that for all $x \in \mathbb{R}$ we have $|\sigma(x)| > d$ without loss of generality. In what follows we denote by K a constant that bounds $|\mu|, |\sigma|$ and the Lipschitz constants of μ, σ.

Define $g(x, y), h(x, y)$ as in Proposition 3.4.3. By Proposition 3.4.2

$$(Z^{n11}, Z^{n12}, Z^{n21}, Z^{n22}) \Rightarrow Z = (0, 0, 0, \frac{\sqrt{2}}{2}B),\ \text{on } [0, T]. \tag{3.4.5}$$

B is a standard Brownian motion and is independent of W. Proposition 3.4.3 shows U_t^n is relatively compact. Thus for any subsequence n', there exists a subsubsequence n'_k of n' and a process R in $\mathcal{C}[0,T]$, such that $U_t^{n'_k} \Rightarrow R$. SDE (3.4.4) has unique weak solution as it satisfies the Engelbert-Schmidt conditions, (see [19] for details). To prove $U^n \Rightarrow U$, it is sufficient to prove that R is a weak solution to SDE (3.4.4).

Because $(U^{n'_k}, X, W, Z^n) \Rightarrow (R, X, W, Z)$, by the almost sure representation theorem (Theorem 1.10.4 on page 59 of van der Vaart and Wellner [80]), there exists a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ and a sequence of processes $\tilde{Y}_k$ and Y, with $\mathcal{L}(Y^k) = \mathcal{L}(U^{n'_k}, X, W, Z^n)$ for all $k \geq 1$, such that $\mathcal{L}(Y) = \mathcal{L}(R, X, W, Z)$, and $Y^k \stackrel{a.s.}{\to} Y$ uniformly on $[0,T]$. If we could prove that the first element of Y is a weak solution to SDE (3.4.4), it follows immediately that R is also a weak solution to (3.4.4).

Thus, without loss of generality, we assume $(U^n, X, W, Z^n) \stackrel{a.s.}{\to} (R, X, W, Z)$ as $n \to \infty$ and we try to prove R is a weak solution to (3.4.4). In particular, for fixed $T > 0$ we remove $\mathscr{A}_T$ a set of probability $\mathbb{P}(\mathscr{A}_T) = 0$, such that for all $\omega \in \mathscr{A}_T^c$ and uniform in $[0,T]$, we have the convergence

$$(U^n, X, W, Z^n) \to (R, X, W, Z).$$

We first present one known result for the continuous Euler scheme under the condition that μ, σ are globally Lipschitz, stated here as (3.4.6). The proof of (3.4.6) can be found in Kloeden [48], proof of Theorem 10.2.2.

$$\sup_n \mathbb{E} \sup_{0<s\leq T} |U_s^n|^2 < \infty. \tag{3.4.6}$$

Since $U^n \stackrel{a.s.}{\to} R$ on $[0,T]$, by Fatou's lemma, we also have

$$\mathbb{E} \sup_{0<s\leq T} |R_s|^2 < \infty. \tag{3.4.7}$$

From the definition of U^n, we have $U^n = \sqrt{n}(X^n - X) = U^{1,n} + U^{2,n}$, where $U^{1,n}, U^{2,n}$ are the same as in proof of Lemma 3.4.3. Since the Lipschitz condition implies differentiability almost everywhere, we can find subset A of $\mathbb{R}$ with Lebesgue measure 0 such that both μ

and σ are differentiable on $\mathbb{R} \cap A^c$. Define $I_1 = I_{\{s:X_s \in A^c\}}$ and $I_2 = I_{\{s:X_s \in A\}}$. We analyze the following terms, $i = 1, 2$,

$$
\begin{aligned}
G_t^{ni1} &= \int_0^t I_i \ \{g(X_{n(s)}^n, X_s)U_t^n - \mu'(X_t)R_t\}ds, \\
G_t^{ni2} &= \int_0^t I_i \ g(X_s^n, X_{n(s)}^n)\mu(X_{n(s)}^n)\sqrt{n}\Delta s^{(n)}ds, \\
G_t^{ni3} &= \int_0^t I_i \ g(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n)\sqrt{n}\Delta W_s^{(n)}ds, \\
F_t^{ni1} &= \int_0^t I_i \ \{h(X_{n(s)}^n, X_s)U_t^n - \sigma'(X_t)R_t\}dW_s, \\
F_t^{ni2} &= \int_0^t I_i \ h(X_s^n, X_{n(s)}^n)\mu(X_{n(s)}^n)\sqrt{n}\Delta s^{(n)}dW_s, \\
F_t^{ni3} &= \int_0^t I_i \ h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n)\sqrt{n}\Delta W_s^{(n)}dW_s - \frac{\sqrt{2}}{2}\int_0^t I_i\sigma(X_t)\sigma'(X_t)dB_s.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\sum_{i=1}^{2}\sum_{j=1}^{3}(G^{nij} + F^{nij}) \\
&= U^n - \Big\{\int_0^t \mu'(X_t)R_t ds + \int_0^t \sigma'(X_s)R_s dW_s + \frac{\sqrt{2}}{2}\int_0^t \sigma(X_s)\sigma'(X_s)dB_s\Big\}.
\end{aligned}
\tag{3.4.8}
$$

Our goal is to show that each term of G^{nij}, F^{nij} converges to a 0 process on $[0, T]$ in distribution.

Consider term G^{n11}. Since $(X^n, X_{n(.)}^n, U^n) \overset{a.s.}{\to} (X, X, R)$ as $n \to \infty$, and μ differentiable on A^c, for each $\omega \in \mathscr{A}_T^c$,

$$
I_{\{X_t \in A^c\}} g(X_{n(t)}^n, X_t)U_t^n \to I_{\{X_t \in A^c\}}\mu'(X_t)R_t, \quad \text{pointwise in } t.
$$

Since U^n is a continuous process and the convergence is uniform, it's limit R will be continuous as well, and moreover

$$
R_T^*(\omega) = \sup_{0 \le s \le T} |R_s|(\omega) < \infty, \ \sup_n \sup_{0 < s \le T} |U_s^n|(\omega) < \infty.
$$

By the globally Lipschitz condition on μ,

$$\Big|g(X^n_{n(t)}, X_t)U^n_t - \mu'(X_t)R_t\Big| \leq K(\sup_{0<s\leq T}|U^n_s| + \sup_{0<s\leq T}|R_s|).$$

Applying the dominated convergence theorem, G^{n11} converges to 0 uniformly almost surely on $[0,T]$.

Consider term G^{n21}. By (3.4.6) and (3.4.7) there exists $C_1 > 0$

$$\begin{aligned}\mathbb{E}(G^{n21}) &\leq \mathbb{E}\Big[\int_0^T I_{\{s:X_s\in A\}} K\Big\{|U^n_s| + |R_s|\Big\}ds\Big] \\ &\leq K\Big(\mathbb{E}\Big[\big(\sup_{0<s\leq T}|U^n_s| + \sup_{0<s\leq T}|R_s|\big)^2\Big]\Big)^{\frac{1}{2}}\Big(\mathbb{E}\Big[\int_0^T I_{\{s:X_s\in A\}}ds\Big]\Big)^{\frac{1}{2}} \\ &\leq C_1\mathbb{E}\Big[\int_0^T I_{\{s:X_s\in A\}}ds\Big].\end{aligned} \tag{3.4.9}$$

By Corollary 3.8 in Chap 7 of Revuz and Yor [70], let $\mathbb{T}^a, \mathbb{T}^b$ be hitting time of X_s and $a < b$, then

$$\mathbb{E}\Big[\int_0^{\mathbb{T}^a\wedge\mathbb{T}^b} I_{\{s:X_s\in A\}}ds\Big] = \int_0^{\mathbb{T}^a\wedge\mathbb{T}^b} G_I(x_0,y)I_{\{y\in A\}}m(dy),$$

where

$$\begin{aligned} s(x) &= \int_c^x \exp\Big(-\int_c^y 2\mu(z)\sigma^{-2}(z)dz\Big)dy, \quad \forall c\in\mathbb{R}; \\ G_I &= \begin{cases} \frac{(s(x)-s(a))(s(b)-s(y))}{s(b)-s(a)}, & a\leq x\leq y\leq b, \\ \frac{(s(y)-s(a))(s(b)-s(x))}{s(b)-s(a)}, & a\leq y\leq x\leq b, \\ 0, & otherwise; \end{cases} \\ m(dx) &= \frac{2}{s'(x)\sigma^2(x)}dx. \end{aligned}$$

Recall that s is the scale function, G_I is the Green function and $m(dy)$ is the speed measure.

By the boundedness of μ, σ, we have $G_I(x_0, y)$ and $\frac{2}{s'(x)\sigma^2(x)}$ are bounded. Since A has

Lebesgue measure 0

$$\mathbb{E}\Big[\int_0^{T\wedge\mathbb{T}^a\wedge\mathbb{T}^b} I_{\{s:X_s\in A\}}ds\Big] \le \mathbb{E}\Big[\int_0^{\mathbb{T}^a\wedge\mathbb{T}^b} I_{\{s:X_s\in A\}}ds\Big] = 0.$$

Let $a \to -\infty, b \to \infty$, and apply Fatou's lemma,

$$\mathbb{E}\Big[\int_0^T I_2 ds\Big] = \mathbb{E}\Big[\int_0^T I_{\{s:X_s\in A\}}ds\Big] = 0 \tag{3.4.10}$$

Together with (3.4.9), we have $G^{n21} = 0$.

Consider F^{n11}, from the Burkholder-Davis-Gundy inequality, there exists $C_3 > 0$ s.t.

$$\begin{aligned}\mathbb{E}\Big[\sup_{0<s\le T}\Big|F_t^{n11}\Big|\Big] &\le C\,\mathbb{E}\Big[\Big(\int_0^T I_1\Big(h(X^n_{n(s)}, X_s)U_s^n - \sigma'(X_s)R_s\Big)^2 ds\Big)^{\frac{1}{2}}\Big]\\ &\le C_3\mathbb{E}\Big[\Big(\int_0^T I_1\Big\{h(X^n_{n(s)}, X_s)(U_t^n - R_t)\Big\}^2 ds\Big)^{\frac{1}{2}}\Big]\\ &+ C_3\mathbb{E}\Big[\Big(\int_0^T I_1\Big\{R_t(h(X^n_{n(s)}, X_s) - \sigma'(X_t))\Big\}^2 ds\Big)^{\frac{1}{2}}\Big].\end{aligned} \tag{3.4.11}$$

Consider the first term on the right side of (3.4.11). Since $|h| \le K$,

$$\mathbb{E}\Big[\Big(\int_0^T I_1\Big\{h(X^n_{n(s)}, X_s)(U_s^n - R_s)\Big\}^2 ds\Big)^{\frac{1}{2}}\Big] \le K\,T\,\mathbb{E}\Big(\sup_{0<s\le T}|U_s^n - R_s|\Big).$$

On the one hand $\sup_{0<s\le T}|U_s^n - R_s| \overset{a.s.}{\to} 0$. On the other by (3.4.6) and (3.4.7), we get

$$\sup_n\Big(\mathbb{E}\Big[\sup_{0<s\le T}|U_s^n - R_s|^2\Big]\Big) < \infty,$$

which gives a uniform integrability condition to ensure

$$\lim_{n\to\infty}\mathbb{E}\Big(\sup_{0<s\le T}|U_s^n - R_s|\Big) = 0.$$

Thus the first term on the right side of (3.4.11) converges to 0. For the second term, an application of the Hölder's inequality gives

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T I_1\Big\{R_s(h(X^n_{n(s)},X_s)-\sigma'(X_s))\Big\}^2 ds\Big)^{\frac{1}{2}}\Big] \\
&\le \mathbb{E}\Big[\sup_{0<s\le T}|R_s|\Big(\int_0^T I_1\Big\{h(X^n_{n(s)},X_s)-\sigma'(X_s)\Big\}^2 ds\Big)^{\frac{1}{2}}\Big] \\
&\le \Big(\mathbb{E}\Big[\sup_{0<s\le T}|R_s|^2\Big]\Big)^{\frac{1}{2}}\Big(\mathbb{E}\Big[\int_0^T I_1\Big\{h(X^n_{n(s)},X_s)-\mu'(X_s)\Big\}^2 ds\Big]\Big)^{\frac{1}{2}}.
\end{aligned}$$

For each $\omega \in \mathscr{A}_T^c$, we have $I_{\{t:X_t\in A^c\}}h(X^n_{n(t)},X_t) \overset{a.s.}{\to} I_{\{t:X_t\in A^c\}}\sigma'(X_t)$ pointwise in t, and $\Big|h(X^n_{n(s)},X_s)\Big|, |\mu'(X_s)|$ are uniformly bounded by K.

From the dominated convergence theorem,

$$\lim_{n\to\infty}\mathbb{E}\Big(\int_0^T I_{\{s:X_s\in\mathbb{R}\cap A^c\}}\Big\{h(X^n_{n(s)},X_s)-\sigma'(X_s)\Big\}^2 ds\Big)=0.$$

With (3.4.7), we have the second term of right side of (3.4.11) also converges to 0. Thus

$$\lim_{n\to\infty}\mathbb{E}\sup_{0<s\le t}\Big|F_t^{n11}\Big|=0.$$

For the term F^{n21}, we would like to prove

$$\lim_{n\to\infty}\mathbb{E}\sup_{0<s\le T}\Big|F_s^{n21}\Big|=0. \tag{3.4.12}$$

Similarly to the analysis of F^{n11}, to prove (3.4.12) we are only left to prove

$$\lim_{n\to\infty}\mathbb{E}\Big(\int_0^T I_2\Big\{h(X^n_{n(s)},X_s)-\sigma'(X_s)\Big\}^2 ds\Big)=0,$$

which is implied by (3.4.10) and boundedness of $\Big|h(X^n_{n(s)},X_s)\Big|, |\sigma'(X_s)|$.

Consider the terms $G^{n12}, G^{n13}, G^{n22}, G^{n23}, F^{n12}, F^{n22}$ all of which converge to the constant process 0 almost surely uniformly on $[0,T]$ because g,h,μ,σ are bounded and $(Z^{n11},Z^{n12},Z^{n21}) \overset{a.s.}{\to} (0,0,0)$ uniformly on $[0,T]$.

For dealing with the last two terms F^{n13} and F^{n23}, we first define $\tilde{F}^{n13}$ as

$$\tilde{F}_t^{n13} = \int_0^t I_1 \left\{h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n) - \sigma'(X_s)\sigma(X_s)\right\}\sqrt{n}\Delta W_s^{(n)} dW_s.$$

From the Burkholder-Davis-Gundy inequality, there exists $C_3 > 0$ such that

$$\mathbb{E}[\sup_{0<s\leq T}\left|\tilde{F}_s^{n13}\right|] \leq C_3\ \mathbb{E}\Big[\Big(\int_0^T I_1\big(h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n) - \sigma'(X_s)\sigma(X_s)\big)^2(\sqrt{n}\Delta W_s^{(n)})^2 ds\Big)^{\frac{1}{2}}\Big].$$

Applying Cauchy-Schwarz inequality to the right side, there exists $C_4',\ C_4 > 0$ s.t.

$$\begin{aligned}
&\mathbb{E}[\sup_{0<s\leq T}\left|\tilde{F}_s^{n13}\right|] \\
&\leq C_4'\mathbb{E}\Big[\Big(\int_0^T I_1\big(h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n) - \sigma'(X_s)\sigma(X_s)\big)^4 ds\Big)^{\frac{1}{4}}\Big(\int_0^T I_1\big(\sqrt{n}\Delta W_s^{(n)}\big)^4 ds\Big)^{\frac{1}{4}}\Big] \\
&\leq C_4'\Big[\mathbb{E}\Big(\int_0^T I_1\big(\sqrt{n}\Delta W_s^{(n)}\big)^4 ds\Big)^{\frac{1}{2}}\Big]^{\frac{1}{2}}\Big[\mathbb{E}\Big(\int_0^T I_1\big(h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n) - \sigma'(X_s)\sigma(X_s)\big)^4 ds\Big)^{\frac{1}{2}}\Big]^{\frac{1}{2}} \\
&\leq C_4\Big[\mathbb{E}\Big(\int_0^T I_1\big(h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n) - \sigma'(X_s)\sigma(X_s)\big)^4 ds\Big)^{\frac{1}{2}}\Big]^{\frac{1}{2}}.
\end{aligned}$$

Since h, σ, σ' are bounded, by the dominated convergence theorem,

$$\lim_{n\to\infty}\mathbb{E}\Big[\sup_{0<s\leq t}\left|\tilde{F}_t^{n13}\right|\Big] = \lim_{n\to\infty}\mathbb{E}\Big(\int_0^T I_1\big(h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n) - \sigma'(X_s)\sigma(X_s)\big)^4 ds\Big)^{\frac{1}{2}} = 0.$$

Thus $\tilde{F}^{n13} \xrightarrow{L^1} 0$ uniformly on $[0, T]$. We define $\bar{F}^{n13}$ as

$$\bar{F}_t^{n13} = \int_0^T I_1\ \sigma'(X_s)\sigma(X_s)dZ^{n22} - \int_0^T I_1\ \sigma'(X_s)\sigma(X_s)dB_s. \tag{3.4.13}$$

Since $Z^{n22} \xrightarrow{a.s.} B_s$ uniformly on $[0, T]$ and Z^{n22} is a good sequence, the result on convergence in probability of stochastic integrals in Protter and Kurtz [50] leads to $\bar{F}^{n13} \xrightarrow{p} 0$ uniformly on $[0, T]$. As $F^{n13} = \tilde{F}^{n13} + \bar{F}^{n13}$, $F^{n13} \xrightarrow{p} 0$.

For the last term F^{n23}, applying the Burkholder-Davis-Gundy inequality first, then using the same technique as in bounding $\tilde{F}^{n13}$, together with (3.4.10), give

$$\lim_{n\to\infty}\mathbb{E}\Big[\sup_{0<s\leq T}\Big|\int_0^T I_2\ h(X_s^n, X_{n(s)}^n)\sigma(X_{n(s)}^n)\sqrt{n}\Delta W_s^{(n)}dW_s\Big|\Big] = 0,$$
$$\lim_{n\to\infty}\mathbb{E}\Big[\sup_{0<s\leq T}\Big|\int_0^T I_2\sigma(X_s)\sigma'(X_s)dB_s\Big|\Big] = 0.$$

Thus $F^{n23} \stackrel{L^1}{\to} 0$ uniformly on $[0,T]$. Each of the G and F terms converges to 0 uniformly on $[0,T]$ either almost surely or in L^1 or in probability. Then, by (3.4.8), $U^n \stackrel{p}{\to} \tilde{R}$ uniformly on $[0,T]$, where

$$\tilde{R}_t = \int_0^t \mu'(X_s)R_s ds + \int_0^t \sigma'(X_s)R_s dW_s + \frac{\sqrt{2}}{2}\int_0^t \sigma(X_s)\sigma'(X_s)dB_s.$$

Since also $U^n \stackrel{a.s}{\to} R$ on $[0,T]$, the two limits must equal each other, and R follows

$$R_t = \int_0^t \mu'(X_t)R_t ds + \int_0^t \sigma'(X_s)R_s dW_s + \frac{\sqrt{2}}{2}\int_0^t \sigma(X_s)\sigma'(X_s)dB_s. \tag{3.4.14}$$

This conclues the proof. □

Remark 3.4.5. *Both in Kurtz and Protter [51] and Neuenkirch and Zähle [60], μ and σ are assumed to be $\mathcal{C}^1$. Since Lipschitz continuity does not imply differentiability, the key part in proof of Theorem* 3.4.1 *is to show that the time the weak limit error process spends on the set where μ and σ are not differentiable has Lebesgue measure 0.*

3.5 Study of The Normalized Limit Error Process

With the weak limit of normalized error process for the Euler scheme being derived, we are interested to further analyze its properties. Though Kurtz and Protter [51] derived the form of the normalized error process of the Euler scheme under the condition that the coefficients are $\mathcal{C}^1$ and bounded, its properties have barely been studied in previous work. In this section, we focus on the mean, variance and martingality of the limit error process under the globally Lipschitz condition. The locally Lipschitz case is more complicated and is studied through examples as well.

3.5.1 The Globally Lipschitz Case

Theorem 3.5.1. *When μ and σ are globally Lipschitz, for the normalized error process $U_n = \sqrt{n}(X^n - X)$ from the continuous Euler scheme, there exists $0 < C_t < \infty$, where C_t increasing with t, such that*

$$\mathbb{E}[U_t^2] \leq \mathbb{E}[U_t^{*2}] \leq C_t,$$

where $U_t^* = \sup_{0 \leq s \leq t} |U_s|$. Furthermore when $\mu' = 0$, U is a square integrable martingale.

Proof. Since $U^n \Rightarrow U$ uniformly on $[0,T]$, we have $\forall\, t \in [0,T]$, $U_t^{n*} \Rightarrow U_t^*$. When μ and σ are both globally Lipschitz, from Kloeden [48] proof of Theorem 10.2.2, there exists a C_t, increasing with t, such that

$$\sup_n \mathbb{E}[(U_t^{n*})^2] < C_t. \tag{3.5.1}$$

Without loss of generality we can assume there exists a subsequence $(U^{n_k})^2 \overset{a.s.}{\to} U^2$ uniformly on $[0,T]$. Since $(U^{n_k})^2 \geq 0$, from Fatou's lemma

$$\mathbb{E}[U_t^2] \leq \mathbb{E}[U_t^{*2}] \leq \liminf_{k \to \infty} \mathbb{E}[(U_t^{n_k *})^2] \leq C_t.$$

When $\mu' = 0$, there is no drift term in (3.4.4). Thus U is a local martingale. We also have a bound for the expectation of the quadratic variation of U_t. Since μ, σ are globally Lipschitz, it is known that $\mathbb{E}(X_t^2) < \infty, \forall\, t \in [0,T]$. Let K be the Lipschitz coefficients for μ and σ, then

$$\begin{aligned}\mathbb{E}(\langle U, U\rangle_t) &= \mathbb{E}\Big[\int_0^t \{\sigma'^2(X)U_s^2 + \sigma'^2(X)\sigma^2(X_s)\}ds\Big] \\ &\leq \int_0^t K^2 \mathbb{E}(U_s^2)ds + \int_0^t K^4 \mathbb{E}(X_s^2)ds < \infty.\end{aligned}$$

U_s is a local martingale with finite expected quadratic variation. From Corollary 3 in page 73 in Protter [65], we conclude it is a martingale when $\mu' = 0$. □

3.5.2 The Locally Lipschitz Case and Examples

3.5.2.1 The Inverse Bessel Process

When μ and σ are only locally Lipschitz, the finiteness of the second moment of the corresponding U_t may not hold. Theorem 3.5.1 can not be extended to the locally Lipschitz plus no finite time time case. One example is the inverse Bessel process, which is a solution to the SDE

$$dX_t = X_t^2 dW_t, \quad X_0 > 0.$$

The coefficient $\sigma(x) = x^2$ is locally Lipschitz and X has no finite explosion. From Theorem 3.4.1, the error process $U_t^n = \sqrt{n}(X_t^n - X_t)$ converges in distribution uniformly to U_t on $[0, T]$. U_t is solution to

$$dU_t = 2X_t U_t dW_t + \sqrt{2} X_t^3 dB_t,$$

where B is a Brownian motion independent of W.

$$\begin{aligned}\mathbb{E}(U_t^2) &= \mathbb{E}\Big(2\int_0^t X_s U_s dW_s + \sqrt{2}\int_0^t X_s^3 dB_s\Big)^2 \\ &= 4\mathbb{E}\Big(\int_0^t X_s U_s dW_s\Big)^2 + 2\mathbb{E}\Big(\int_0^t X_s^6 ds\Big)\end{aligned}$$

Since the inverse Bessel process can also be represented as the inverse of the norm of a three dimensional Brownian motion starting from $(1, 0, 0)$, its explicit distribution can be obtained (for example see [23]). A calculation shows if $X_0 > 0$, then $\forall t > 0$, $\mathbb{E}X_t^6 = \infty$. This gives $\mathbb{E}(U_t^2) = \infty$ and $\mathbb{E}(U_t^{*2}) = \infty$. This indicates that under the locally Lipschitz condition, the asymptotic distribution for the normalized error process might have larger tail probability than in the globally Lipschitz case.

3.5.2.2 The CIR process

There are also examples with μ and σ only locally Lipschitz, U_t still has finite second moment. We look at the Cox-Ingersoll-Ross model (or CIR model) which is often used to describe the evolution of interest rates. The CIR process follows the SDE

$$dX_t = (a - bX_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 > 0, a > 0. \tag{3.5.2}$$

The coefficient function $\sigma\sqrt{X_t}$ is only locally Lipschitz. The true solution to (3.5.2) remains always positive, but the numerical solution from the Euler scheme may go negative. Thus the Euler scheme is not well defined for solving (3.5.2). We use the same trick due to Bossy at al [9], replacing the Euler scheme by a symmetrized Euler scheme. Let U^n be the sequence of normalized error from the symmetrized Euler scheme solving (3.5.2). By Theorem 2.2 in Berkaoui, Bossy and Diop [9], there exists a C_t, increasing with t, such that

$$\sup_n \mathbb{E}[(U_t^{n*})^2] < C_t, \tag{3.5.3}$$

if the following condition holds

$$\frac{\sigma^2}{8}\Big(\frac{2a}{\sigma^2} - 1\Big)^2 > \mathcal{K}(8), \text{ with } \mathcal{K}(p) = \max\{b(4p-1), (2\sigma(2p-1))^2\}.$$

Since the symmetrized Euler scheme is local and the true solution never hits 0 or ∞ in finite time, it can be shown that $U^n \Rightarrow U$ as $n \to \infty$ on any finite time interval. The weak limit U has the same form as in Theorem 3.4.1, it solves the SDE below.

$$dU_t = -bU_t dt + \frac{\sigma U_t}{2\sqrt{X_t}}dW_t + \frac{\sqrt{2}}{2}\sigma^2 dB_t.$$

With (3.5.3), applying Fatou's Lemma, we have

$$\mathbb{E}[U_t^2] \leq \mathbb{E}[U_t^{*2}] \leq C_t.$$

The inverse Bessel and CIR examples show that the finiteness of the second moment of the normalized error process for the Euler scheme (or modified Euler scheme in order for the scheme to be well defined) under the locally Lipschitz situation is more complicated than the globally Lipschitz situation.

3.6 Approximation of Expectations of Functionals

In applications, the convergence of expectations of functionals (also called weak convergence in existing literature) of the Euler scheme is important. To avoid confusion, in this section weak convergence means the convergence of expectations of functionals unless further specified. We are interested in the rate of convergence for $\mathbb{E}[g(X_T^n)] - \mathbb{E}[g(X_T)]$ to 0, as n goes to infinity. When μ and σ are only assumed to be locally Lipchitz, inferred from Hutzenthaler, Jentzen and Kloeden [36], even for g with linear growth, weak convergence in the sense of expectations of functionals may not hold. As a compromise, in this section we assume g is Lipschitz and bounded, and give upper bound for the weak convergence rate with the no finite explosion condition and some other mild conditions on the SDE (3.1.1). Before we deal with the locally Lipschitz case, we need the following Proposition.

In Proposition 3.6.1, inequality (3.6.1) can be inferred from Kloeden [48] page 343 proof of Theorem 10.2.2 in chapter 10, or Theorem 4.4 in H. Desmond and X.Mao [32].

Proposition 3.6.1. *Consider SDE (*3.1.1*), if μ and σ are globally Lipschitz with Lipschitz coefficient as K, then for all $T > 0$ there exists $c > 0$ not depending on K increasing with T, such that for all $n \in \mathbb{N}^+$,*

$$\mathbb{E}[\sup_{0 \leq s \leq T} (X_s^n - X_s)^2] \leq \frac{\exp(cK)}{n}, \tag{3.6.1}$$

and for all $\gamma \in [0, \frac{1}{2})$

$$\mathbb{P}(\sup_{0 \le s \le T} |X^n_s - X_s| > n^{-\gamma}) \le \exp(cK) n^{-1+2\gamma}. \tag{3.6.2}$$

Proof. From Theorem 4.4 in H. Desmond and X.Mao [32], with the globally Lipschitz condition, for any $\delta > 0$ there exists universal constant C and A independent of n such that

$$\mathbb{E}[\sup_{0 \le s \le T} (X^n_s - X_s)^2] \le \frac{C}{n}(K^2 + 1)\exp\{4K(T+4)\} + \frac{A}{n}.$$

Thus (3.6.1) holds. Applying the Chebyshev's inequality gives (3.6.2). □

Theorem 3.6.1. *Consider the SDE (*3.1.1*). If* μ, σ *are locally Lipschitz and we assume the Lipschitz constant has at most polynomial growth with exponent* $a \in \mathbb{R}^+$*, that is, for all* $x, y \in \mathbb{R}$

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \le (\max\{|x|, |y|\} + K)^a \, |x - y|,$$

where K *is a constant. We assume there exists* $\kappa, \nu > 0$*, such that for all* $x > 0$

$$\mathbb{P}(X^*_T > x) \le \kappa x^{-\nu}.$$

Then, there exists a finite constant $C = C(\kappa, \nu, K, a, |x_0|)$ *such that for any Lipschitz and bounded function* g *and for all* $n > 1$

$$\left|\mathbb{E}g(X^n_T) - \mathbb{E}g(X_T)\right| < C(\|g\|_\infty + G + 1)\Big(\log n\Big)^{-\frac{\nu}{a}},$$

where G *is the Lipschitz constant of* g*.*

Proof. For a fixed $m > |x_0|$, define $\mu^{(m)}, \sigma^{(m)}$ and $Y^{(m)}$ as in Proposition 3.3.1. Let $\mathbb{T}^m(Z) = \inf\{t \ge 0 : |Z_t| > m\}$. Since the Euler scheme is local,

$$\mathbb{T}^m(X(x_0, \mu, \sigma, W)) = \mathbb{T}^m(X(x_0, \mu^{(m)}, \sigma^{(m)}, W)),$$
$$\mathbb{T}^m(X^n(x_0, \mu, \sigma, W)) = \mathbb{T}^m(X^n(x_0, \mu^{(m)}, \sigma^{(m)}, W)).$$

Let $\theta^m = \mathbb{T}^{m+2}(X(x_0,\mu,\sigma,B)) \wedge \mathbb{T}^{m+2}(X^n(x_0,\mu,\sigma,B))$. Then

$$\left|\mathbb{E}g(X_T^n) - \mathbb{E}g(X_T)\right| \leq \left|\mathbb{E}g(X_{T\wedge\theta^m}^n) - \mathbb{E}g(X_{T\wedge\theta^m})\right| + 2\|g\|_\infty \mathbb{P}(\theta^m < T)$$

Let G be a Lipschitz constant for g. Then by (3.6.1) in Proposition 3.6.1, and since the Lipschitz constant of μ,σ on $[-(m+2), m+2]$ is bounded by $(m+K+2)^a$

$$\left|\mathbb{E}g(X_{T\wedge\theta^m}^n) - \mathbb{E}g(X_{T\wedge\theta^m})\right| \leq G(\mathbb{E}|X_{T\wedge\theta^m}^n - X_{T\wedge\theta^m}|^2)^{\frac{1}{2}} \leq G\exp\left(\frac{c}{2}(m+K+2)^a\right)n^{-\frac{1}{2}},$$

where c is the constant given in Proposition 3.6.1. By the distribution assumption on X^* and (3.6.2) in Proposition 3.6.1 with $\gamma = 0$,

$$\begin{aligned}
&\mathbb{P}(\theta^m < T) \leq \mathbb{P}(\mathbb{T}^{m+2}(X(x_0,\mu,\sigma,W)) < T) + \mathbb{P}(\mathbb{T}^{m+2}(X^n(x_0,\mu,\sigma,W)) < T)\\
&\leq 2P(\mathbb{T}^{m+1}(X) \leq T) + \mathbb{P}(\mathbb{T}^{m+1}(X) > T,\, \mathbb{T}^{m+2}(X^n) < T)\\
&= 2P(\mathbb{T}^{m+1}(X) \leq T) + \mathbb{P}(\mathbb{T}^{m+1}(X^{(m+2)}) > T,\, \mathbb{T}^{m+2}(X^{n,(m+2)}) < T)\\
&\leq 2P(\mathbb{T}^{m+1}(X) < T) + \exp(c(m+K+2)^a)n^{-1}\\
&\leq 2\kappa(m+1)^{-\nu} + \exp(c(m+K+2)^a)n^{-1}.
\end{aligned}$$

For $m \geq 0$, we have $(m+1)^{-\nu} \leq (K+2)^\nu(m+K+2)^{-\nu}$. Thus, we get

$$\begin{aligned}
\left|\mathbb{E}g(X_T^n) - \mathbb{E}g(X_T)\right| \leq\ & 2\|g\|_\infty\left[2\kappa\,(K+2)^\nu(m+K+2)^{-\nu} + \exp(c(m+K+2)^a)n^{-1}\right]\\
&+G\exp\left(\tfrac{c}{2}(m+K+2)^a\right)n^{-\frac{1}{2}}.
\end{aligned}$$

Take $n = (m+K+2)^{2\nu}\exp(c(m+K+2)^a)$ to get

$$\left|\mathbb{E}g(X_T^n) - \mathbb{E}g(X_T)\right| \leq (2\|g\|_\infty[(K+2)^\nu 2\kappa + 1] + G)(m+K+2)^{-\nu}.$$

Notice that

$$\log n \leq c(m+K+2)^a + 2\nu\log(m+K+2) \leq (c+2\nu/a)(m+K+2)^a,$$

which implies that

$$|\mathbb{E}g(X_T^n) - \mathbb{E}g(X_T)| \leq (2\|g\|_\infty[(K+2)^\nu 2\kappa + 1] + G)\left(\frac{\log n}{c + 2\nu/a}\right)^{-\frac{\nu}{a}}.$$

The result follows from this estimation. Finally, notice we have assumed $m \geq |x_0|$, which imposes that $n \geq n_0$ has to be large enough, for example

$$\log n_0 \geq c(|x_0| + K + 2)^a + 2\nu \log(|x_0| + K + 2).$$

□

As an example we consider the constant elasticity of variance process which follows the following SDE

$$dS_t = bS_t^\beta dW_t, \ S_0 > 0$$

When $\beta > 1$, the solution to the above SDE is strict local martingale and is used for detecting asset bubbles. By a result of A. N. Borodin and P. Salminen [10], chapter 4.6, $\forall x > S_0, \ T > 0$

$$P(S_T^* > x) < \frac{S_0}{x}.$$

Thus, there exists a constant $C > 0$, such that for all $g : \mathbb{R} \to \mathbb{R}$, bounded and Lipschitz and for all $n > 1$

$$|\mathbb{E}g(S_T^n) - \mathbb{E}g(S_T)| < C(\|g\|_\infty + G + 1)\big(\log n\big)^{-\frac{1}{(\beta-1)}}$$

Remark 3.6.1. *The above example of the CEV process for $\beta > 1$ illustrates the weakness of the result of Theorem* 3.6.1. *The rate of convergence is so slow as to be essentially useless in practice. It is our hope that future research will illustrate methods that will permit a more practically useful analysis of the rate of convergence. This seems far away at this point.*

Chapter 4

Conclusion

In this thesis, we address two problems related to stochastic differential equations of the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x_0 \in \mathbb{R},$$

where W is a one dimensional standard Brownian motion. Chapter 2 treats an important problem in financial markets, how to detect the existence of an asset bubble and judge its severity instantaneously with observations from a price process. To address these problems, we propose an approach combining the use of SDEs and dynamic linear regression. The CEV model with time varying parameters is used to model asset price processes with potential bubbles. When using the CEV model, we show that the exponent parameter is linked to the existence and severity of an asset bubble. An asset bubble exists if and only if the exponent parameter is larger than a fixed threshold. Conditioning on the existence of an asset bubble, we prove that a larger exponent indicates the price process has smaller future expectation and a greater probability for the running maximum to hit a large value in a certain time range. We use the dynamic linear regression model to estimate the exponent parameter in the CEV model, which allows the existence and severity of asset bubbles to be checked instantaneously with historical intraday price observations. The approach is illustrated with examples from the Dot-com bubble era. Under the CEV model, $\sigma(x)$ is restricted to be a power function. In future work, it would be useful to obtain a comparison result on the future expectation of X under a risk neutral measure with a more a general form of $\sigma(x)$.

Chapter 3 focuses on the convergence property of the Euler scheme solving SDEs under general assumptions, namely, locally Lipschitz coefficients and no finite explosion. Existing work often assumes the globally Lipschitz condition, which is too stringent in practice. We have shown that if a numerical scheme converges in probability uniformly on any compact time set (UCP), with a certain rate under the global Lipschitz condition, then the UCP convergence with same rate holds when the globally Lipschitz condition is replaced with a locally Lipschitz condition plus a no finite explosion condition. For the Euler scheme, we prove the $\sqrt{n}$ rate of convergence in distribution for the asymptotic normalized error process. The limit error process is derived as a solution to an SDE. In addition, we further study the boundedness of the second moment of the limit error process and its running maximum. We also study the weak convergence for the Euler scheme in the sense of expectation of bounded and Lipschitz functions. We show that with some mild conditions, the weak convergence is at least at the rate of $(\log n)^{\alpha}$, where $\alpha > 0$ depends on the SDE. In chapter 3, we consider the case of one dimensional Brownian motion driven SDEs. Generalizing the results to the multi-dimensional case and to semimartingale driven SDEs with jumps are left for future work. With the locally Lipschitz assumption, the Euler scheme may diverge in L^p, even if the p^{th} moment of the solution exists. How to modify the Euler scheme so that it will converge in L^p is also left for future work.

Bibliography

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions.* U.S. Government Printing Office: p378, 1964.

[2] G. Arfken, H. Weber, *Mathematical methods for physicists*, Academic Press, 2001.

[3] J. Arnulf and P. E. Kloeden, The numerical approximation of stochastic partial differential equations. *Milan Journal of Mathematics* 77.1: 205-244, 2009.

[4] V. Bally and D. Talay, The Euler scheme for stochastic differential equations: error analysis with Malliavin calculus. *Mathematics and computers in simulation* 38.1-3: 35-41, 1995.

[5] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations. *Probability theory and related fields* 104.1: 43-60, 1996.

[6] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations: II. Convergence rate of the density. *Monte Carlo Methods and Applications* 2.2: 93-128, 1996.

[7] R. F. Bass and E. Pardoux, Uniqueness for diffusions with piecewise constant coefficients. *Probability Theory and Related Fields* 76.4: 557-572, 1987.

[8] K. Bastiaensen, P. Cauwels, Z. Q. Jiang, D. Sornette, R. Woodard, and W. X. Zhou, Bubble diagnosis and prediction of the 2005-2007 and 2008-2009 Chinese stock market bubbles. *Journal of Economic Behavior and Organization*, 74, 149-162, 2010.

[9] A. Berkaoui, M. Bossy and A. Diop, Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics*, 12, 1-11, 2008.

[10] A.N. Borodin and P. Salminen, *Handbook of Brownian Motion-Facts and Formulae, Second Edition*, Birkhäuser, 2015.

[11] M. Bossy, A. Diop. An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^a$, a in $[1/2, 1)$ (Doctoral dissertation, INRIA).

[12] R. Caballero, E. Fahri and P. O. Gourinchas, Financial crash, commodity prices and global imbalances. *Brookings Papers on Economic Activity*, Fall, 155, 2008.

[13] J. Cox, Notes on option pricing I: Constant elasticity of diffusions. Unpublished draft, Stanford University, 1975.

[14] A. M. G. Cox and D. G. Hobson, Local martingales, bubbles and option prices. *Finance and Stochastics*, 9 (4), 477-492, 2005.

[15] D. Davydov and V. Linetsky, The valuation and hedging of path-dependent options under the CEV process. *Management Sci.* 47:949-965, 2001.

[16] D. C. Emanuel, and J. D. MacBeth, Further Results of the Constant Elasticity of Variance Call Option Pricing Model. *Journal of Financial and Quantitative Analysis*, 4 : 533-553, 1982.

[17] F. Delbaen, W. Schachermayer, A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, 300(1), 463-520, 1994.

[18] D. Florens-Zmirou, On estimating the diffusion coefficient from discrete observations. *Journal of applied probability*, 790-804,1993 Dec;30(4).

[19] H. Engelbert and W. Schimidt, On one-dimensional stochastic differential equations with generalized drift. *Stochastic Differential Systems Filtering and Control*: 143-155, 1985.

[20] G. Evans, A test for speculative bubbles in the sterling-dollar exchange rate:1981-84, *American Economic Review*, 76 (4), 621-636, 1986.

[21] J. D. Farmer, and D. Foley, The economy needs agent-based modelling. *Nature*, 460, 680-681, 2009.

[22] W. Feller, Two Singular Diffusion Problems. *Annals of Mathematics*, 54, 173-182, 1951.

[23] H. Föllmer and P. Protter, Local martingales and filtration shrinkage. *ESAIM: Probability and Statistics* 15: S25-S38, 2011.

[24] K. Froot and M. Obstfeld, Intrinsic Bubbles: the case of stock prices, *American Economic Review*, 81 (5), 1189 - 1214, 1991.

[25] J. K. Galbraith, *A short history of financial euphoria*, Penguin Books,New York, 1993.

[26] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D.B. Rubin, *Baysian data analysis (3rd ed.)*, Springer, 2014.

[27] C. Gilles, Charges as equilibrium prices and asset bubbles. *Journal of Mathematical Economics*, 18, 155 - 167, 1988.

[28] C. Gilles and S. F. LeRoy, Bubbles and charges, *International Economic Review*, 33 (2), 323 - 339, 1992.

[29] I, Gyöngy. A note on Euler's approximations. *Potential Analysis*, May 1;8(3):205-16, 1998 .

[30] Y. Hamana, and H. Matsumoto, The probability distributions of the first hitting times of Bessel processes, Trans. Amer. Math. Soc. 365, no. 10, 5237-5257, 2013.

[31] S. Heston, M. Loewenstein, and G. A. Willard, Options and bubbles, *Review of Financial Studies*. 20(2):359-90, 2007.

[32] D. J. Higham, D. J. Mao and A.M.Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM Journal on Numerical Analysis* 40.3: 1041-1063, 2002.

[33] U. Homm, J. Breitung, Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics*, 10(1):198-231, 2012.

[34] H. Hong, J. Scheinkman, and W. Xiong, Advisors and Asset Prices: A Model of the Origins of Bubbles, *Journal of Financial Economics* 89, 268-287, 2008.

[35] A. Hüsler, D. Sornette, C. H. Hommes. Super-exponential bubbles in lab experiments: evidence for anchoring over-optimistic expectations on price. *Journal of Economic Behavior and Organization.* 92:304-16, 2013.

[36] M. Hutzenthaler, A. Jentzen and P. E. Kloeden, Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. Vol. 467. No. 2130.* The Royal Society, 2011.

[37] M. Hutzenthaler and A. Jentzen, Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. Vol. 236. No. 1112. American Mathematical Society, 2015.

[38] J. Jacod and P. Protter, Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability*: 267-307, 1998.

[39] J. Jacod, Non-parametric Kernel Estimation of the Coefficient of a Diffusion. *Scandinavian Journal of Statistics*, 27, 83-96, 2000.

[40] J. Jacod and P. Protter, *Discretization of processes*. Vol. 67. Springer Science and Business Media, 2011.

[41] J. Jarrow, Y. Kchia, P. Protter, How to detect an asset bubble. *SIAM Journal on Financial Mathematics.* 2, 839-865, 2011.

[42] R. Jarrow and D. Madan, Arbitrage, martingales, and private monetary value, *Journal of Risk*, 3 (1), 2000.

[43] R. Jarrow, P. Protter, An introduction to financial asset pricing. *Handbooks in Operations Research and Management Science*, 15, .13-69, 2007.

[44] R. Jarrow, P. Protter and K. Shimbo, Asset Price Bubbles in Complete Markets. *Advances in Mathematical Finance*, Springer-Verlag, M.C. Fu et al, editors, 97-122, 2007.

[45] R. Jarrow, P. Protter and K. Shimbo, Asset Price Bubbles in Incomplete Markets, *Mathematical Finance*, 20, 145-185, 2010.

[46] A. Johansen, D. Sornette D. Modeling the stock market prior to large crashes. *The European Physical Journal B-Condensed Matter and Complex Systems*. 1;9(1):167-74, 1999.

[47] I. Karatzas, S. Shreve, *Brownian motion and stochastic calculus, 2nd edn.* Springer, NewYork, 2012.

[48] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations springer-verlag*, New York,1999.

[49] A. Kohatsu-Higa, and P. Protter, The Euler scheme for SDE's driven by semimartingales. *Pitman research notes in mathematics series*: 141-141, 1994.

[50] T. G. Kurtz and P. Protter, Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability*: 1035-1070, 1991.

[51] T. G. Kurtz and P. Protter, Wong-Zakai corrections, random evolutions and numerical schemes for SDE's. *Stochastic Analysis*: 331-346, 1991.

[52] V. Linetsky, Lookback options and diffusion times: A spectral expansion approach. *Finance and Stochastics* 8:373-398, 2004.

[53] M. Loewenstein and G. A. Willard, Rational equilibrium asset-pricing bubbles in continuous trading models. *Journal of Economic Theory*, 91 (1), 17- 58, 2000.

[54] M. Loewenstein and G. A. Willard, Local martingales, arbitrage and viability: free snacks and cheap thrills. *Journal of Economic Theory*, 16, 135 - 161, 2000.

[55] D. B. Madan and M. Yor, Ito's integrated formula for strict local martingales, *Lecture notes in mathematics-springer verlag*, 1;1874:157, 2006 Jan.

[56] G. Marion X. Mao and E. Renshaw, Convergence of the Euler scheme for a class of stochastic differential equation. *International Mathematical Journal* 1.1: 9-22, 2002.

[57] S. Mark, Computing the constant elasticity of variance option pricing formula. *The Journal of Finance* 44.1: 211-21 9, 1989.

[58] R. Meese, Testing for bubbles in exchange markets: A case of sparkling rates. *Journal of Political Economy*, 94 (2), 345-373, 1986.

[59] A. Mijatovic, M. Urusov, On the martingale property of certain local martingales. *Probability Theory and Related Fields* 152,1-30, 2012.

[60] A. Neuenkirch and H. Zähle, Asymptotic error distribution of the Euler method for SDEs with non-Lipschitz coefficients. *Monte Carlo Methods and Applications*, 15(4), 333-351, 2009.

[61] G. Petris, S. Petrone, and P. Campagnoli, *Dynamic Linear Models with R*. Springer: New York, 2009.

[62] P. C. B. Phillips and J. Yu, Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics*, 2, 455-491, 2011.

[63] C. Pigorsch, U. Pigorsch, and I. Popov, Volatility Estimation Based on High-Frequency Data. *Handbook of Computational Finance*, Springer. 2010.

[64] P. Protter and D. Talay, The Euler scheme for Lévy driven stochastic differential equations. *The Annals of Probability* 25.1: 393-423, 1997.

[65] P. Protter, *Stochastic integration and differential equations, version 2.1, 2nd edn.* Springer, Heidelberg, 2005.

[66] P. Protter, A mathematical theory of financial bubbles. In *Paris-Princeton Lectures on Mathematical Finance 2013* (pp. 1-108). Springer International Publishing, 2013.

[67] O. Yoshiki, P. Protter, and S. Yang. The lifetime of a financial bubble. *Mathematics and Financial Economics* 11.1, 45-62, 2017.

[68] L. R. Rabiner, B. H. Juang . An introduction to hidden Markov models. *IEEE ASSP Magazine*: 4-15, 1986.

[69] R. Rebib, D. Sornette, R. Woodard and W. Yan, Detection of crashes and rebounds in major equity markets, *International Journal of PortfolioAnalysis Management*, 1(1), 59-79, 2012.

[70] D. Revuz and M. Yor, *Continuous martingales and Brownian motion.* Vol. 293. Springer Science and Business Media, 2013.

[71] J. Scheinkman and W. Xiong, Overconfidence and speculative bubbles, *Journal of Political Economy*, 111 (6), 1183-1219, 2003.

[72] J. Scheinkman and W. Xiong, Heterogeneous beliefs, speculation and trading in financial markets, *Paris-Princeton Lecture Notes on Mathematical Finance 2003*, Lecture Notes in Mathematics, 1847, 217-250, 2004.

[73] D. Sornette, R. Woodard, W. Yan, and W. X. Zhou, Clarifications to questions and criticisms on the Johansen-Ledoit-Sornette financial bubble model. *Physica A: Statistical Mechanics and its Applications*, 392(19), 4417-4428, 2013.

[74] M. Swayne, How to detect a gold bubble, in eHow Money, November 28, 2010, at the URL http://www.ehow.com/how 7415180 detect-goldbubble.htm

[75] D. Talay, Résolution trajectorielle et analyse numérique des équations différentielles stochastiques. Stochastics: *An International Journal of Probability and Stochastic Processes* 9.4 : 275-306, 1983.

[76] D. Talay, Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE, and applications. Filtering and control of random processes. Springer Berlin Heidelberg, 294-313, 1984.

[77] D. Talay and L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic analysis and applications* 8.4 : 483-509, 1990.

[78] D. Talay, Simulation and numerical analysis of stochastic differential systems: a review. (Doctoral dissertation, INRIA), 1990.

[79] J. Tirole, Asset bubbles and overlapping generations, *Econometrica*, 53(5), 1071 - 1100, 1985.

[80] A. W. van der Vart and J. A. Wellner, *Weak convergence and empirical processes: with applications to statistics*. Springer Science and Business Media, 2013.

[81] L. Yan, The Euler scheme with irregular coefficients. *The Annals of Probability* 30.3: 1172-1194, 2002.

Appendices

Appendix A

First Appendix Section

A.1 Option Pricing and Auxiliary Results for the CEV Model (a correction of S.Mark [57])

The following theorem is a correction of the call option pricing formula in S.Mark [57].

Theorem A.1.1. *Assume that a nonnegative process S_t follows the SDE (2.3.1). We use the same notations as in Proposition 2.3.1. Let $y = \lambda e^{-b\tau}(\frac{E}{s_0})^{2-\theta}$. The European call price with exercise price E is*

$$C = s_0 e^{-a\tau}\Big\{F(\lambda; df-2, 0) - F(\lambda; df-2, 2y)\Big\} - Ee^{-r\tau}F(2x; df, 2y).$$

Proof. The European call price with exercise price E is

$$\begin{aligned}
C &= e^{-r\tau}\int_E^{+\infty}(S_T - E)f(S_T, T; S_{t_0}, t_0)dS_T \\
&= e^{-r\tau}\int_E^{+\infty} S_T f(S_T, T; S_{t_0}, t_0)dS_T - Ee^{-r\tau}\int_E^{+\infty} f(S_T, T; S_{t_0}, t_0)dS_T \\
&= A - B
\end{aligned}$$

We deal with the two parts separately

$$\begin{aligned} A &= e^{-r\tau} \int_E^{+\infty} S_T f(S_T, T; S_{t_0}, t_0) dS_T \\ &= e^{-r\tau} \int_y^{+\infty} s_0 e^{\mu\tau} (\frac{z}{\lambda})^{\frac{1}{2-\theta}} p(z; df, \lambda) dz \\ &= s_0 e^{-a\tau} \int_y^{+\infty} p(\lambda; df, z) dz. \end{aligned}$$

By equation (2.3.5)

$$A = s_0 e^{-a\tau} \Big\{ F(\lambda; df - 2, 0) - F(2x; df - 2, 2y) \Big\}.$$

For the term B,

$$\begin{aligned} B &= E \int_E^{+\infty} f(S_T, T; S_{t_0}, t_0) dS_T \\ &= E e^{-r\tau} \int_y^{+\infty} p(z; df, \lambda) dz \\ &= E e^{-r\tau} F(\lambda; df, 2y). \end{aligned}$$

Thus

$$\begin{aligned} C &= A - B \\ &= s_0 e^{-a\tau} \Big\{ F(\lambda; df - 2, 0) - F(\lambda; df - 2, 2y) \Big\} - E e^{-r\tau} F(\lambda; df, 2y). \end{aligned}$$

□

From Proposition 2.3.1, we obtain the following two corollaries.

Corollary A.1.1. *Assume S follows the SDE*

$$dS = \delta S^{\frac{\theta}{2}} dW, \quad with \quad S(t_0) = s_0 > 0. \tag{A.1.1}$$

If $\theta > 2$, S_T converges to 0 in L^1 as T goes to infinity.

Proof. From Proposition 2.3.1,

$$\mathbb{E}^\theta(S_T) = F[\frac{4s_0^{2-\theta}}{\delta^2(2-\theta)^2\tau}; \frac{2}{\theta-2}, 0]s_0,\ \tau = T - t_0.$$

Then

$$\lim_{\tau\to+\infty} F[\frac{4s_0^{2-\theta}}{\delta^2(2-\theta)^2\tau}; \frac{2}{\theta-2}, 0] = F[0; \frac{2}{\theta-2}, 0] = 0.$$

This concludes the proof. □

Corrollary(A.1.1) indicates that if an asset bubble lasts long enough, the price process will converge to 0 in L^1 as time goes to infinity.

Corollary A.1.2. *Assume S follows the SDE* (A.1.1)*. Then*

$$\lim_{\theta\to+\mathbf{2}} \mathbb{E}^\theta_{\mathbf{s_0}}(\mathbf{S_T}) = \mathbf{s_0}, \tag{A.1.2}$$

and

$$\textit{when } s_0 < 1, \quad \lim_{\theta\to+\infty} \mathbb{E}^\theta_{\mathbf{s_0}}(\mathbf{S_T}) = \mathbf{s_0},$$
$$\textit{when } s_0 \geq 1, \quad \lim_{\theta\to+\infty} \mathbb{E}^\theta_{\mathbf{s_0}}(\mathbf{S_T}) = \mathbf{1}.$$

Proof. Using Theorem 2.3.1,

$$\mathbb{E}^\theta(S_T) = F[\frac{4S_t^{2-\theta}}{\delta^2(2-\theta)^2\tau}; \frac{2}{\theta-2}, 0]s_0.$$

Note that

$$\lim_{\theta\to+2} 1 - F[\frac{4s_0^{2-\theta}}{\delta^2(2-\theta)^2\tau}; \frac{2}{\theta-2}, 0] \leq \lim_{\theta\to+2} \frac{\frac{2}{\theta-2}}{\frac{4s_0^{2-\theta}}{\delta^2(2-\theta)^2\tau}} = 0.$$

Thus (A.1.2) holds.

When $s_0 < 1$,

$$\lim_{\theta \to +\infty} \frac{2}{\theta - 2} = 0, \text{ and } \lim_{\theta \to +\infty} \frac{4 s_0^{2-\theta}}{\delta^2 (2-\theta)^2 \tau} = +\infty.$$

Thus,

$$\lim_{\theta \to +\infty} F[\frac{4 s_0^{2-\theta}}{\delta^2 (2-\theta)^2 \tau}; \frac{2}{\theta - 2}, 0] = 1$$

We use the same notation as in Proposition **??** for the case $s_0 > 1$. Since

$$F[\frac{4 s_0^{2-\theta}}{\delta^2 (2-\theta)^2 \tau}; \frac{2}{\theta - 2}, 0] = \int_0^{C s_0^{-\frac{2}{d}} d^2} \frac{1}{\Gamma(d)} y^{d-1} e^{-y} dz,$$

$$\begin{aligned}
\lim_{\theta \to +\infty} \mathbb{E}_{s_0}^{\theta}(S_T) &= \lim_{d \to +0} s_0 \int_0^{C s_0^{-\frac{1}{d}} d^2} \frac{1}{\Gamma(d)} y^{d-1} e^{-y} dz \\
&= \lim_{d \to +0} s_0 \int_0^{C s_0^{-\frac{1}{d}} d^2} d y^{d-1} dz \\
&= \lim_{d \to +0} s_0 (C s_0^{-\frac{1}{d}} d^2)^d = 1.
\end{aligned}$$

This concludes the proof. □

We also study the tail behavior of the CEV processes. When $0 < \theta < 2$,

$$P(S_T > A) = 1 - F[\frac{2CA^{2-\theta}}{(\theta-2)^2}; \frac{2}{2-\theta}, \frac{C S_t^{2-\theta}}{(\theta-2)^2}]. = O(\exp(-\frac{2CA^{2-\theta}}{(\theta-2)^2}))$$

Thus, the distribution of S_T is light tailed and has finite moments of all order. When $\theta > 2$, it can be shown that

$$P(S_T > A) = 1 - F[\frac{2CA^{2-\theta}}{(\theta-2)^2}; 2 + \frac{2}{2-\theta}, \frac{2C S_t^{2-\theta}}{(\theta-2)^2}] = O(A^{1-\theta}).$$

Thus, the distribution of S_T has a heavy tail with power law $1 - \theta$.

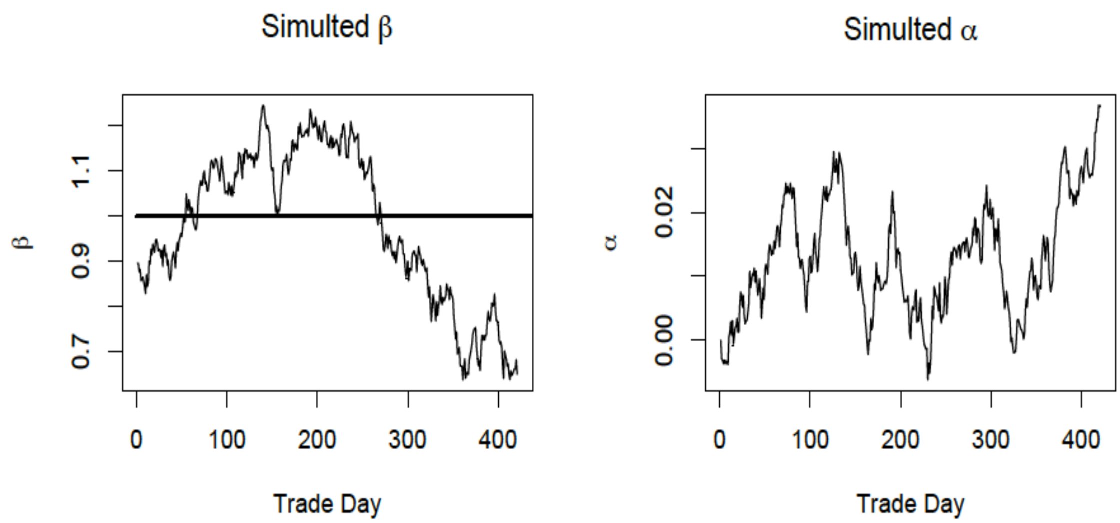

Figure A.2.1: Simulated α and β.

A.2 Simulation Result for the Dynamic Regression and the CEV Model

In this section, we simulate the paths of the price process under the assumption that the price process follows the CEV model with parameters α_n and β_n that vary each day, but are constant within a day. Under the assumptions of the Kalman Filter, the series α_n and β_n are both AR(1). Let α_n, β_n be the values of the coefficients for the n_{th} trading day. Suppose there is no trading time gap between two consecutive trading days. For the n_{th} trading day, the trading time interval is $[T_{n-1}, T_n)$. We simulate β_n and α_n in the following way,

$$\alpha_{n+1} = \alpha_n + \epsilon_1, \quad \epsilon_1 \sim N(0, 0.002^2), \tag{A.2.1}$$

$$\beta_{n+1} = \beta_n + \epsilon_2, \quad \epsilon_2 \sim N(0, 0.02^2). \tag{A.2.2}$$

In Figure (A.2.1), it shows that the simulated β_n stay above 1 in the time range around the 50_{th} to the 250_{th} day. The price process under a risk measure follows the CEV model with time varying parameters

$$dS_t = e^{\alpha_n} S_t^{\beta_n} dW_t,\ t \in [T_{n-1}, T_n) \tag{A.2.3}$$

We use the Euler Method to simulate the paths of the price process S

$$S_{t_{k+1}+\Delta t} = S_{t_k} + e^{\alpha_{n_k}} S_t^{\beta_{n_k}} (W_{t_{k+1}} - W_{t_k}),\ t_k \in [T_{n_k-1}, T_{n_k}) \tag{A.2.4}$$

Based on the simulated α_n and β_n, we simulate 500 sample paths from the CEV model (A.2.3) using the Euler scheme (A.2.4). We set the starting value as 40. When doing the simulation, we retain those sample paths whose maximum value is at least 60 in order to resemble a real bubble. For each simulated price process, we use the dynamic linear model to achieve real time instantaneous estimation for the parameters of the underlying CEV process. Figure (A.2.2) illustrates β_n for $n = 1 \ldots 400$, the average of the instantaneous estimates for all 500 simulated price processes for 400 days, as well as the upper and lower 5% quantiles of $\hat{\beta_n}$. The average of $\hat{\beta}$ is below 1 at the beginning of the true bubble time range, because only a small part of the price data used to estimate β for those days is from the bubble region. Since the statistical method is getting information from the whole past history up to the current time, the recent time bubble is masked by the initial time range which is not in bubble. As illustrated in Figure A.2.2, the average of the instantaneous estimates of β_n for all 500 simulated price processes is quite close to the underlying true value of β_n. We plot the first 6 simulated price processes and apply the dynamic linear regression model to the processes. The true bubble regions for each price process are between the two black vertical lines. In Figure A.2.3 the red dots indicate the regions with instantaneous estimate $\hat{\beta} > 1$ from the dynamic linear regression model. We regard the days as marked with red dots as the days detected as days with bubbles. For day k, $\hat{\beta}_{k,j}$ is the instantaneous estimate for β using the $j_t h$ simulated price process. Bubble detecting rate for day k is calculated as $\frac{1}{500}\Sigma I_{\hat{\beta}_{k,j}>1}$. When the underlying β is larger than 1, the statistical procedure can only detect it with a certain probability. Figure A.2.4 indicates that the probability at time t is higher when the time range with bubbles before t is longer and the value of the true β is

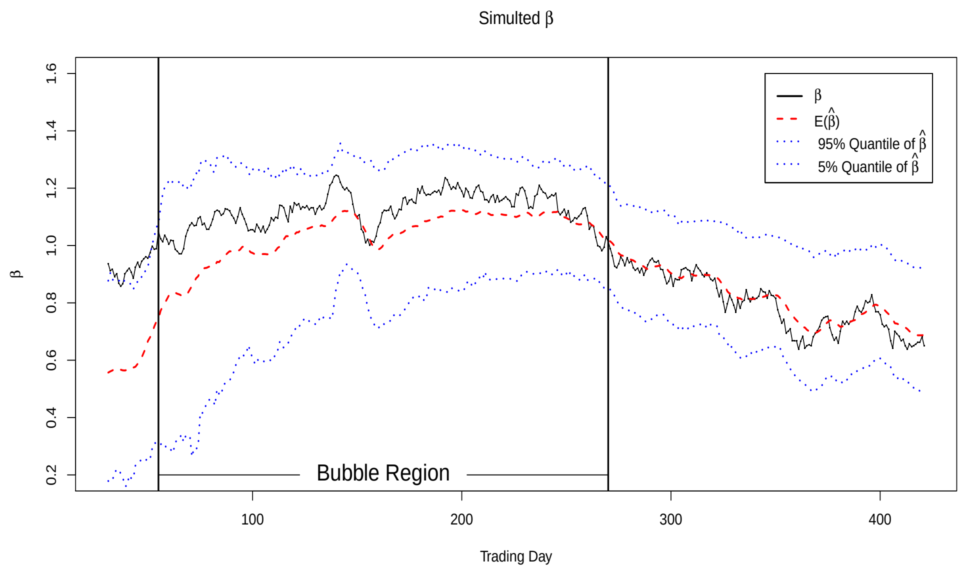

Figure A.2.2: Average of $\hat{\beta}_n$ for simulated price processes with upper, lower 5% quantiles.

further away above 1.

A.3 Further Study of the Stock YHOO and Stock INSP

Using the formula (2.3.1), for each t we sample α_t and β_t from the posterior distribution with observations up to time t, according to the time varying regression model . With each sample of α_t and β_t from the posterior distribution, we could calculate $\mathbb{E}_{\alpha_t,\beta_t}(\frac{S_{t+\Delta t}}{S_t}|\mathcal{F}_t)$, assuming $\alpha = \alpha_t$ and $\beta = \beta_t$. The heat maps in Figure A.3.1 and Figure A.3.2 illustrate the posterior distribution of $\mathbb{E}_{\alpha_t,\beta_t}(\frac{S_{t+\Delta t}}{S_t}|\mathcal{F}_t)$ when Δt is a quarter of a year and 1 year respectively, for stock YHOO. Figure A.3.3 and Figure A.3.4 illustrate the posterior distribution of $\mathbb{E}_{\alpha_t,\beta_t}(\frac{S_{t+\Delta t}}{S_t}|\mathcal{F}_t)$ when Δt is a quarter of a year and 1 year respectively, for stock INSP. From the heat maps, a darker color indicates a smaller future expectation relative to the current starting value. We can see the evolvement of severity of asset bubbles for YHOO and INSO from 1998 to 2001, which is the dot-com bubble era.

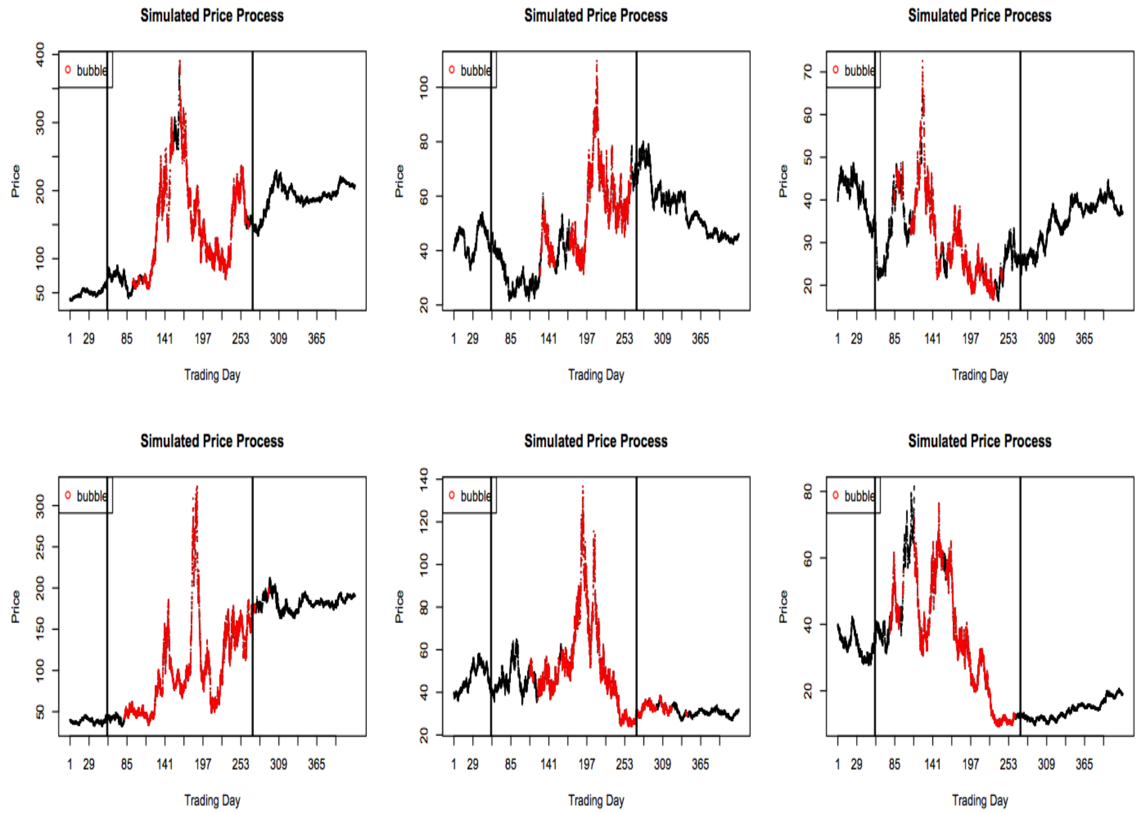

Figure A.2.3: Simulated sample paths

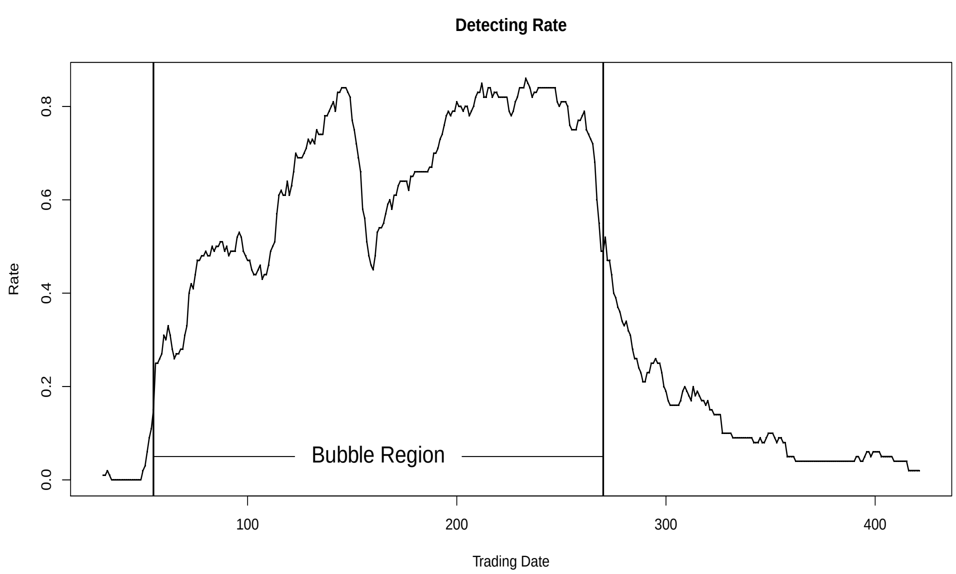

Figure A.2.4: Bubble detection rate from simulation

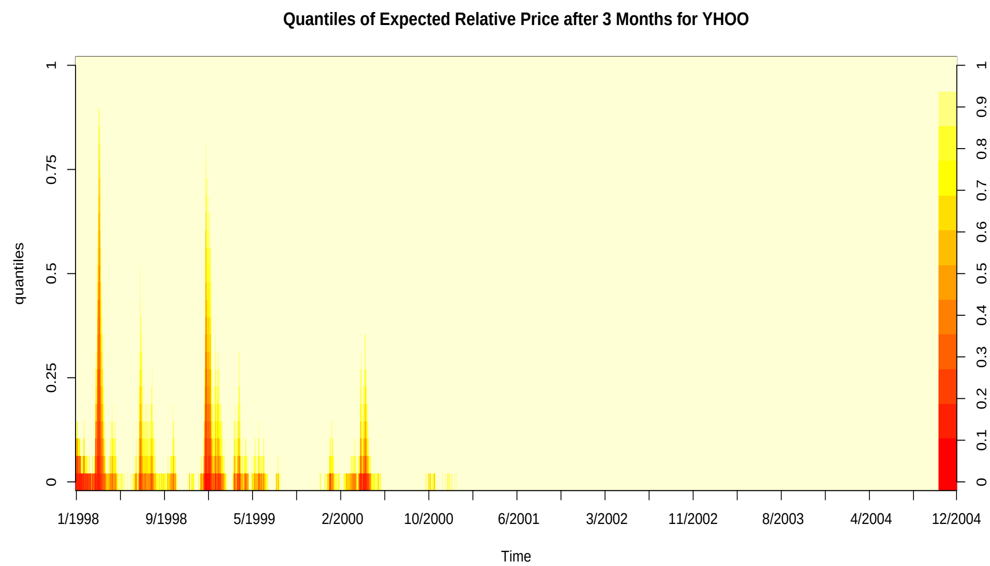

Figure A.3.1: Distribution of the expected relative price after 3 months for YHOO

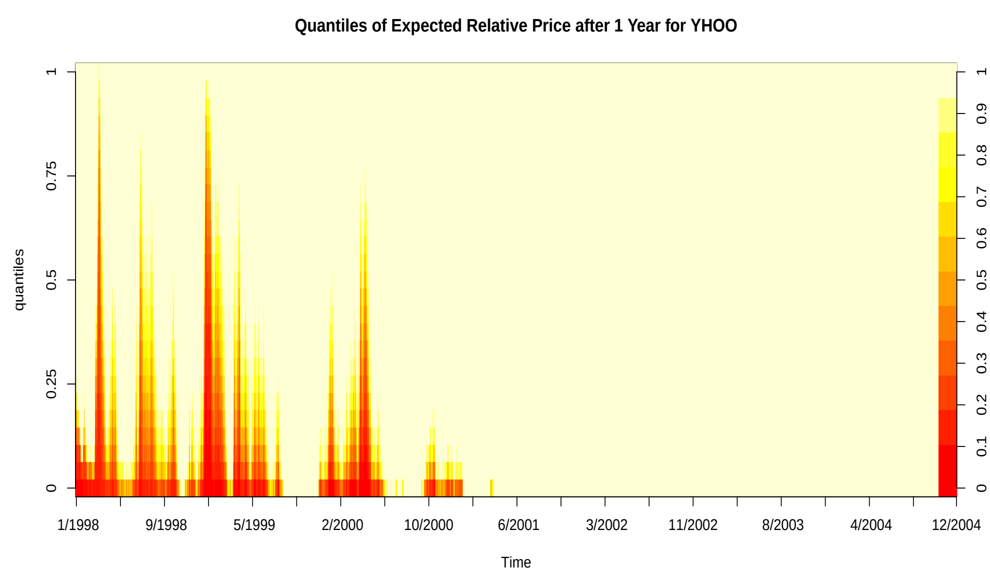

Figure A.3.2: Distribution of the expected relative price after 1 year for YHOO

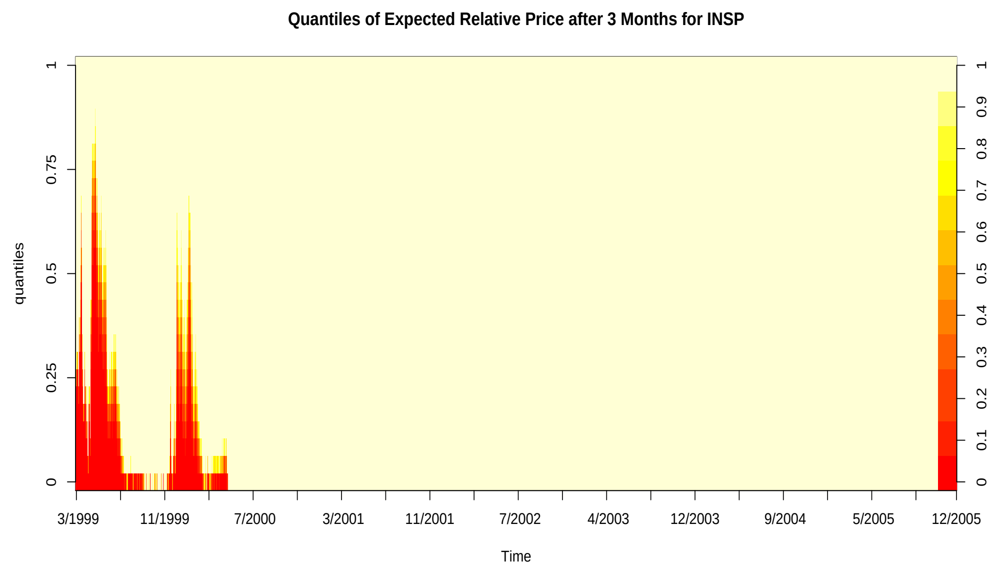

Figure A.3.3: Distribution of the expected relative price after 3 months for INSP

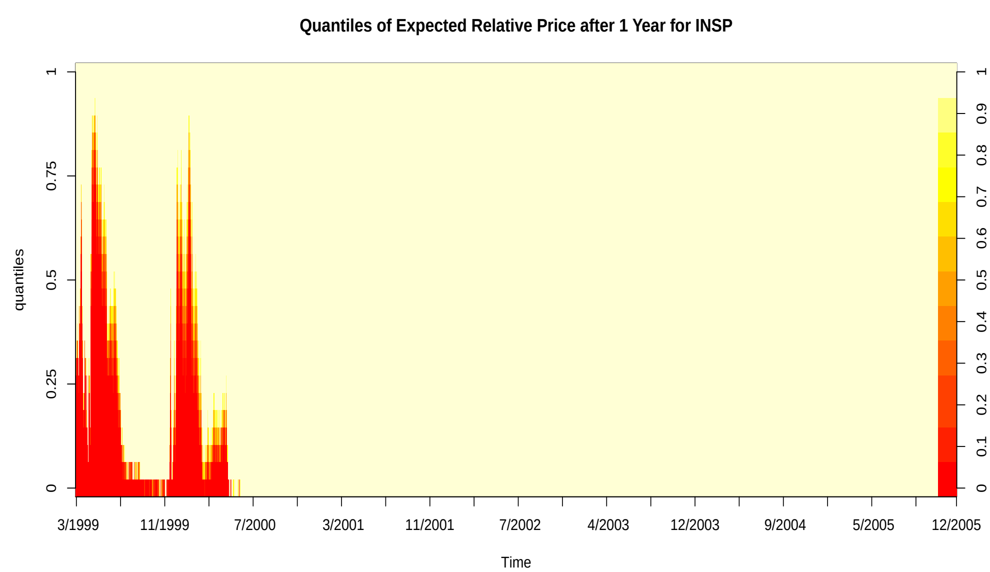

Figure A.3.4: Distribution of the expected relative price after 1 year for INSP

A.4 The Running Maximum of CEV Processes over A Long Time Range

This section uses the same notation as in Proposition 2.3.3. Y.Hamana and H.Matsumoto [30] provided asymptotic approximation for the cumulative probability function of the hitting time of a Bessel process. We can easily use the result to get asymptotic approximation for $P(S_T^* > M)$, when $\theta > 2$, and $t \to \infty$:

$$
\begin{aligned}
P_\theta(S_T^* > M) &= 1 - P(\tau_{a,b}^{(\nu)} > t) \\
&= (\frac{b}{a})^{2\nu} - (\frac{b^3}{2a})^{\nu}\left\{(\frac{a}{b})^{\nu} - (\frac{b}{a})^{\nu}\right\}\frac{1}{\Gamma(1+\nu)t^{\nu}} + o(\frac{1}{t^{\nu}}) \\
&= \frac{s_0}{M} - (\frac{s_0}{2M})^{1/2}M^{-1}(\frac{2\nu}{\delta\theta})^{2\nu}\left\{(\frac{M}{s_0})^{1/2} - (\frac{s_0}{M})^{1/2}\right\}\frac{1}{\Gamma(1+\nu)t^{\nu}} + o(\frac{1}{t^{\nu}}) \\
&= \tilde{P}_\theta(S_T^* > M) + o(\frac{1}{t^{\nu}}) \\
\tilde{P}_\theta(S_T^* > M) &= \frac{s_0}{M} - (\frac{s_0}{2M})^{1/2}M^{-1}(\frac{2\nu}{\delta})^{2\nu}\left\{(\frac{M}{s_0})^{1/2} - (\frac{s_0}{M})^{1/2}\right\}\frac{1}{\Gamma(1+\nu)t^{\nu}}
\end{aligned}
$$

When $\theta_1 > \theta_2 > 2$ and $t \geq \frac{4e^2}{\delta^2(\theta_1-1)(\theta_1-2)}$ for $\forall M > s_0$,

$$
\tilde{P_{\theta_1}}(S_T^* > M) < \tilde{P_{\theta_2}}(S_T^* > M) \tag{A.4.1}
$$

Appendix B

Second Appendix Section

B.1 Proof of Weak Convergence for Z^{n22} in Proposition 3.4.2

Let Y be a continuous one dimensional local martingale. Let C be the quadratic variation process. Let Z^n be

$$D_t^n = \int_0^t \Delta Y^{(n)} dY, \tag{B.1.1}$$

where $\Delta Y^{(n)}$ is defined as in (3.2.1). Jacod and Protter [38] proved the following theorem:

Theorem B.1.1 (Jacod and Protter)**.** *If the quadratic variation process for Y satisfies*

$$C_t = \int_0^t c_s ds, \; \int_0^1 |c_s|^2 \, ds \leq \infty,$$

then the sequence $\sqrt{n} D^n$ converges in distribution to a process D given by

$$D_t = \frac{1}{\sqrt{2}} \int_0^t c_s db,$$

where B is a standard Brownian motion defined on an extension of the space on which Y is defined and independent of Y.

Consider Z_t^{n22} in Proposition 3.4.2. Let $Y = W$, where W is a standard one dimensional

Brownian motion, then $c_s = 1$ in this case. By Theorem B.1.1, $\sqrt{2}Z_t^{n22} \Rightarrow B$ on $[0, T]$, where B is a standard Brownian motion independent of W.

B.2 Simulation From the Euler Scheme

In order to illustrate the difference between the limit normalized error processes for the globally Lipschitz and the locally Lipschitz cases, we use the geometric Brownian motion and inverse Bessel process as a simulation example. Let $S^{(1)}$ and $S^{(2)}$ follow the SDEs below:

$$\begin{aligned} dX_t &= X_t dW_t,\ S_0 = 1, \\ d\tilde{X}_t &= \tilde{X}_t^2 dW_t,\ \tilde{X}_0 = 1. \end{aligned} \tag{B.2.1}$$

Then X is a geometric Brownian motion and $\tilde{X}$ is an inverse Bessel process. Take $T = 1$ and set the number of discretizations to 1000 for the Euler scheme for solving (B.2.1). The normalized error processes are obtained using the Euler scheme again on the SDE (3.4.4) in Theorem 3.4.1, substituting X with X^n and taking $\sigma(x) = x$ and x^2 respectively:

$$\bar{U}_i^n = \sigma'(X_{i-1})\bar{U}_{i-1}^n \Delta W_i + \frac{\sqrt{2}}{2}\sigma'(X_{i-1})\sigma(X_{i-1})\Delta B_i,\ i = 1 \dots n.$$

Figure B.2.1 illustrates the 50 simulated paths for the limit normalized error processes. As shown in Figure B.2.1, the simulated normalized error process from the inverse Bessel process has a higher probability of high spikes than the one for the geometric Brownian motion. As proved in Theorem 3.5.1, the limit normalized error process for the globally Lipschitz case has finite first moment for its running maximum. The same result does not hold for the case of the inverse Bessel process as the SDE it follows has coefficients that are only locally Lipschitz. A better way to compare the scale of the simulated normalized error processes is to divide them by the absolute simulated value of the underlying processes that the Euler scheme approximates. Let $\tilde{U}_t^n = \frac{\bar{U}_t^n}{|X_t^n|}$; Figure B.2.2 illustrates the 50 simulated paths for $\tilde{U}^n$. We also use the kernel estimation method to estimate the maximum normalized error up to time $T = 1$ from the Euler schemes for the two processes. From Figure B.2.3, the

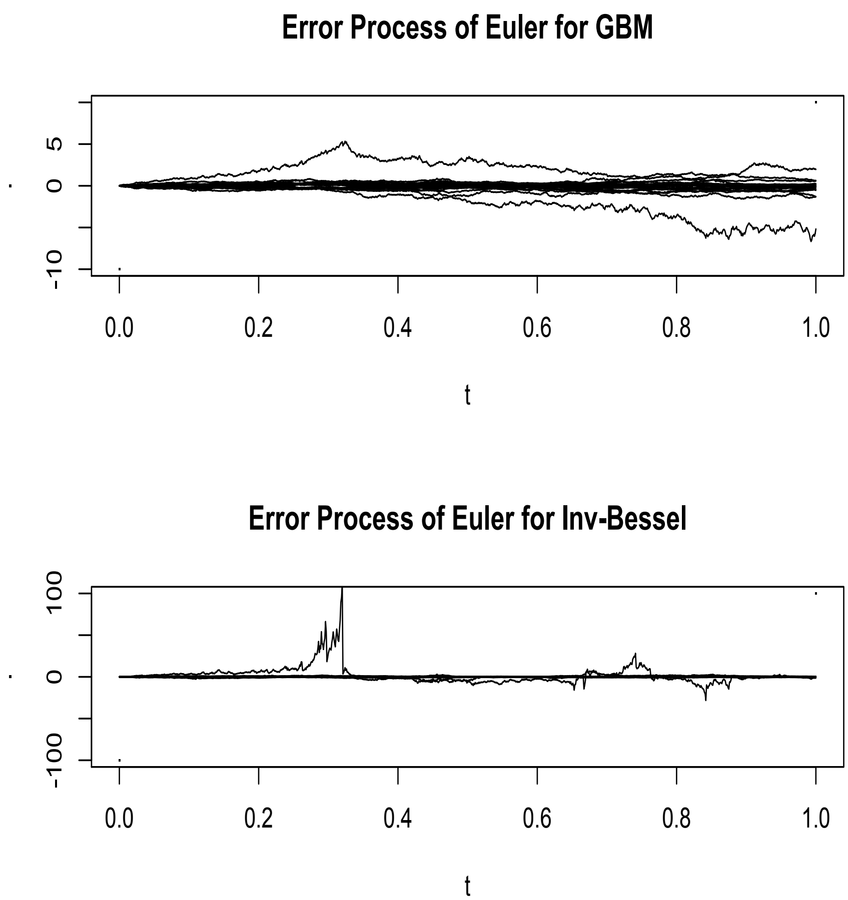

Figure B.2.1: Simulated normalized error process for the Euler scheme

estimated density for the running maximum of the normalized error process is more heavy tailed for the situation of geometric Brownian motion than for the inverse Bessel process.

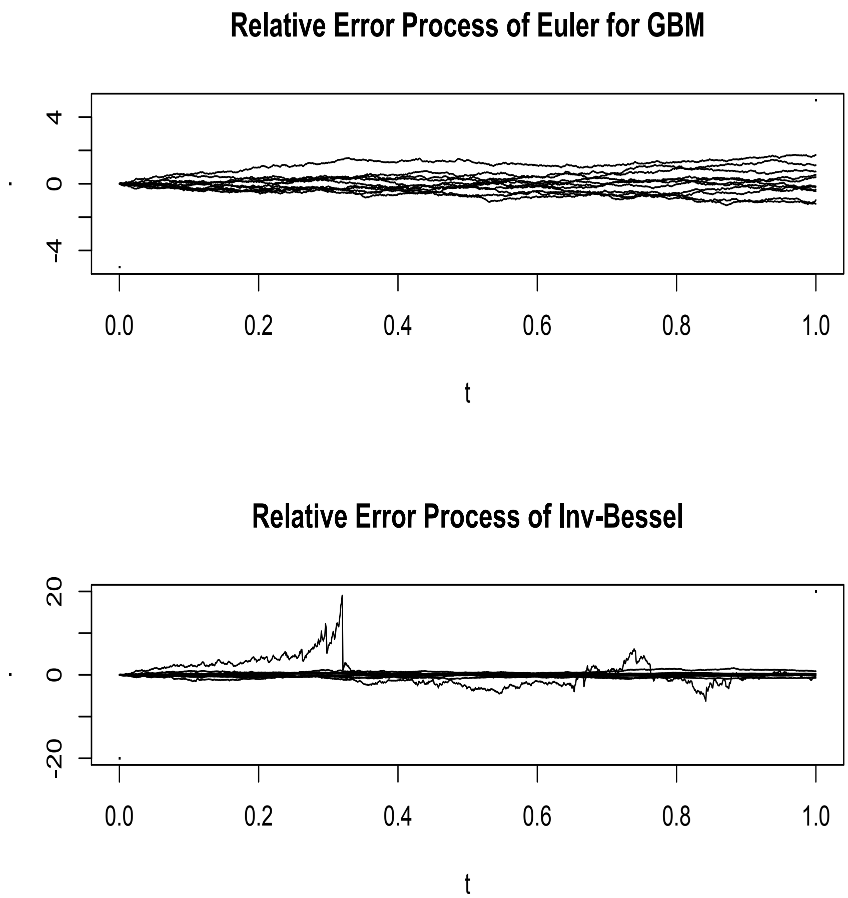

Figure B.2.2: Simulated relative normalized error process for the Euler scheme

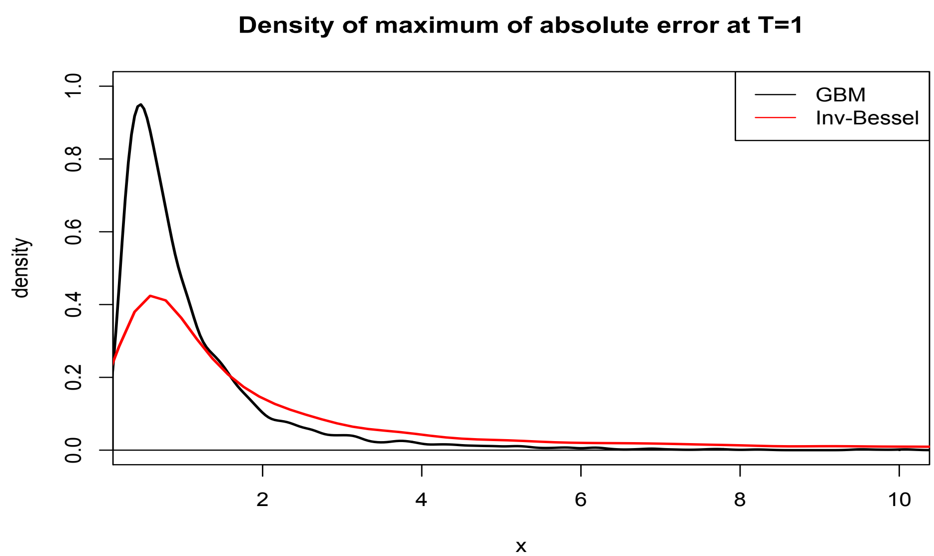

Figure B.2.3: Estimation of the density for the running maximum of the normalized error process