PATH-DEPENDENT VOLATILITY MODELS

ANTOINE JACQUIER AND CHLOÉ LACOMBE

Abstract. We provide a thorough analysis of the path-dependent volatility model introduced by Guyon [13], proving existence and uniqueness of a strong solution, characterising its behaviour at boundary points, and deriving large deviations estimates. We further develop a numerical algorithm in order to jointly calibrate S&P 500 and VIX market data.

1. Introduction

Stochastic volatility models have been used extensively over the past three decades in order to reproduce particular features of market data, on Equities, FX and Fixed Income markets, both under the historical measure and for pricing purposes. Most of the literature and the models used in practice are based on a Markovian assumption for the underlying process, essentially for mathematical convenience, as PDE techniques and Monte Carlo schemes are more readily available then. However, recent models have departed from this Markovian confinement and have shown to provide extremely accurate fit to market data. One approach considers instantaneous volatility driven by fractional Brownian motion, giving rise to the rough volatility generation and its numerous descendents. A less strodden, yet very intuitive, path, originally introduced by Engle [7] and Bollerslev [2] in the early 1980s suggested to consider models where volatility depends on the past history of the stock price process. Their approach, though, was under the historical measure, and Duan [5] investigated these discrete-time models in the context of option pricing. With this in mind, Hobson and Rogers [15] extended this approach to continuous time, suggesting that instantaneous volatility depends on exponentially weighted moments of the stock price. Contrary to stochastic volatility models (rough or not), the market here is complete. Hobson and Rogers [15] showed that such models generate implied volatility smiles and skews consistent with market data. Further results investigated some theoretical properties of these models, in particular [20] proving existence and uniqueness of strong solutions. This path has recently been given new highlights by Guyon [12], who concentrated on the following setup for the stock price process S:

$$\frac{\mathrm{d}S_t}{S_t} = \sigma(t, S_t, Y_t) l(t, S_t)\mathrm{d}W_t, \quad S_0 := s_0 > 0,$$

where W is a standard Brownian motion, Y an adapted process and $l(\cdot)$ the leverage function ensuring (similar to local volatility models) that European options are fully recovered. Inspired by Hobson and Rogers [15], Guyon [13] suggested to choose Y as an exponentially weighted moving average of the stock price. Not only does this model calibrate perfectly to the observed smile, but the diffusion map $\sigma(\cdot)$ can be chosen in such a way that joint calibration with VIX data becomes feasible, a notoriously hard task so far. We take up on this challenge set by Guyon, and analyse theoretical and empirical properties of this class of models. In particular, we provide a full characterisation of the behaviour of the process at its boundaries and derive precise small-time large deviations asymptotics. Only scarce related results for systems with memory are available in the literature, for example by Azencott, Geiger and Ott [1] for systems with finite discrete-time delay or Ma, Ren, Touzi and Zhang [18] for non-Markovian stochastic differential equations with random coefficients.

Date: January 15, 2020.

2010 *Mathematics Subject Classification.* 41A60, 60F10, 60G15.

Key words and phrases. Path-dependent volatility, large deviations, implied volatility asymptotics.

The authors would like to thank Alexander Kalinin for his insightful remarks about the behaviour of solutions at the boundary and Julien Guyon for introducing us to this exciting problem.

The structure of the paper is as follows. In Section 2, we set the notations and present the model. Section 3 gathers the main theoretical results, proving existence and uniqueness of a strong solution (Section 3.1), deriving the stationary distribution (Section 3.2) and presenting pathwise large deviations estimates in Section 3.3, from which implied volatility asymptotics follow readily. Finally, in Section 4 we construct a numerical scheme to estimate the coefficients of the system for a joint calibration to S&P 500 and VIX market data.

2. Set up and notations

The goal of this project is to develop and analyse a model describing the relationship between the VIX index and the VVIX, a volatility of volatility index. Figure 1 below shows a scatter plot of one versus the other over a five-year period. The approximate linear relationship highlighted by the least-square regression fit was first noted by Guyon [13], and we follow his recommendations here. The underlying

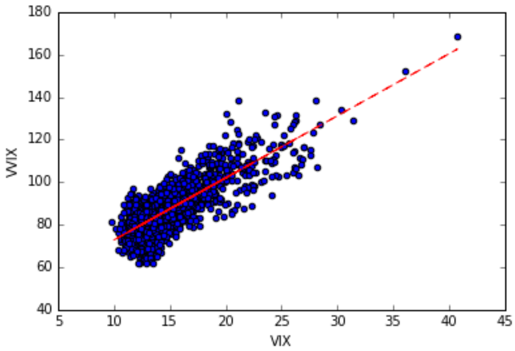


Figure 1. Historical VVIX vs historical VIX (13/4/12 - 8/5/17). *Source: CBOE data.*

process S, describing the evolution of the S&P index follows the general dynamics

$$\frac{\mathrm{d}S_t}{S_t} = \sigma(Y_t)\mathrm{d}W_t, \quad S_0 := s_0 > 0$$

for some given Brownian motion W generating a filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$, where $\sigma : \mathbb{R}_+^* \to \mathbb{R}$ is non anticipative. Following Guyon [13] and Hobson and Rogers [15], we assume that the process Y is adapted to $\mathcal{F}$ and is a function of the past history of the stock S, making the latter non-Markovian, in the sense

$$Y_t := \frac{S_t}{\overline{S}_t^h}, \quad \text{for } t \in \mathcal{T}, \qquad \text{where } \overline{S}_t^h := \frac{1}{h}\int_{-\infty}^t \exp\left\{-\frac{t-u}{h}\right\} S_u \mathrm{d}u \tag{1}$$

is the exponential weighted moving average (EWMA) of the stock price process. Here, $\mathcal{T} = [0, T]$ for some fixed time horizon T. The constant $h > 0$, denoting the length of the time window, is left unspecified for now. Using Itô's formula, we can summarise the dynamics for the couple (S, Y) as

$$\begin{aligned} \mathrm{d}S_t &= S_t\sigma(Y_t)\mathrm{d}W_t, & S_0 = s_0 > 0, \\ \mathrm{d}Y_t &= b(Y_t)\mathrm{d}t + \widetilde{\sigma}(Y_t)\mathrm{d}W_t, & Y_0 = y_0 > 0, \end{aligned} \tag{2}$$

with $b(y) := \frac{1}{h}y(1-y)$ and $\widetilde{\sigma}(y) := y\sigma(y)$ for $y > 0$. Guyon [13] showed that, for the linear relationship between the VIX and the VVIX to hold, one needs to consider a diffusion coefficient of the form

$$\sigma(x) := -\frac{\alpha}{\beta} + \gamma x^{-\beta}, \tag{3}$$

with $\alpha, \beta, \gamma > 0$. In that case, $\widetilde{\sigma}$ is null at $y_\sigma := \left(\frac{\beta\gamma}{\alpha}\right)^{1/\beta}$, and

$$\widetilde{\sigma}(0) = \begin{cases} \text{not defined}, & \text{if } \beta > 1, \\ 0, & \text{if } \beta < 1, \\ \gamma, & \text{if } \beta = 1. \end{cases}$$

Before diving into the asymptotic behaviour of the process, we first concentrate on the existence and uniqueness of a strong solution for (2), depending on the values of α, β and γ, and classify the singular points 0, y_σ and infinity.

3. Theoretical analysis of the path-dependent volatility model

3.1. **Boundary conditions.** We start the analysis of the model by a detailed study of the behaviour of the system (2) at the boundary points 0, y_σ and ∞. Recall [3, Definition 2.3] that a point $d \in \mathbb{R}$ is called singular if $\frac{1+|b|}{\tilde{\sigma}^2}$ is not locally integrable in a neighbourhood of d. A point that is not singular is called regular. A singular point is further called isolated if it admits a deleted neighbourhood consisting of regular points. Introduce the hitting times $\tau_x := \inf\{t \geq 0 : Y_t = x\}$ and $\tau_{x,y} := \min(\tau_x, \tau_y)$. The precise classification of boundary points is rather delicate and we only give below an informal help, and refer the interested reader to [3, Sections 2.3 and 2.4] for full and precise details:

Definition 3.1. Let $a > 0$.

- A point x is said to have right-type 0, and we write $x \in \mathcal{B}^a_{0+}$, if for any $y_0 \in [x, x+a]$, there exists a unique solution defined up to $\tau_{x,a}$. This solution reaches x with strictly positive probability, $\mathbb{E}(\tau_{x,x+a})$ is finite, and $\mathbb{P}\left(Y_{\tau_{x,x+a}} = 0\right) > 0$.
- A point x is said to have right-type 1, and we write $x \in \mathcal{B}^a_{1+}$, if for any $y_0 \in [x, x+a]$, there exists a unique solution defined up to $\tau_{x,a}$. This solution reaches x with strictly positive probability. Any solution may only leave x in the left direction.
- A point x is said to have right-type 3 if for any $y_0 \in (x, x+a]$, there exists a unique solution defined up to τ_{x+a}. This solution never reaches x and $\mathbb{E}(\tau_{x+a})$ is finite. We write $x \in \mathcal{B}^a_{3+}$.
- Infinity is called a recurrent boundary point if a solution cannot explode there. Moreover, if there are no singular points between y_0 and a point $z < y_0$, then the solution reaches z almost surely.

The left-types $\mathcal{B}^a_{1-}$ and $\mathcal{B}^a_{3-}$ are defined similarly.

The following proposition, proved in Appendix A.1, characterises the behaviour of the solution to (2) at the origin and at infinity.

Proposition 3.2. *Infinity is a recurrent boundary, and the origin is regular if and only if $\beta > \frac{1}{2}$, and*

- *for $\beta > \frac{1}{2}$, zero is an exit, non-entrance boundary point and belongs to $\mathcal{B}^{y_\sigma}_{0+}$;*
- *for $0 \leq \beta \leq \frac{1}{2}$, zero is an exit, non-entrance boundary point and belongs to $\mathcal{B}^{y_\sigma}_{1+}$.*

The following theorem, proved in Section A.2, is more involved and fully characterises the solution to the SDE (2).

Theorem 3.3. *y_σ is an isolated singular point, and for $a > 0$, the following holds:*

	Left boundary of $y_\sigma > y_0$	*Right boundary of $y_\sigma < y_0$*
$y_\sigma < 1$	$\mathcal{B}^a_{1-}$	$\mathcal{B}^a_{3+}$
$y_\sigma \geq 1$	$\mathcal{B}^a_{3-}$	$\mathcal{B}^a_{1+}$

3.2. **Stationary distribution and pricing PDE.** We now prove that the process Y introduced in (2) admits a stationary distribution and we derive the pricing PDE associated to (2).

3.2.1. *Stationary distribution.* We recall [9, Section 3.2] that a process $(Y_t)_{t>0}$ is ergodic if it admits a unique, stationary distribution Π, and for any measurable bounded function g, the almost sure limit

$$\lim_{t\uparrow\infty} \frac{1}{t}\int_0^t g(Y_s)\mathrm{d}s = \int g(y)\Pi(\mathrm{d}y)$$

holds. If it exists, an ergodic solution must satisfy $\mathcal{L}^*\Pi = 0$, where $\mathcal{L}^*$ is the adjoint of the infinitesimal generator $\mathcal{L}$, defined [9, Section 1.5.3] via the equality

$$\int \Psi(\xi)\mathcal{L}\phi(\xi)\mathrm{d}\xi = \int \phi(\xi)\mathcal{L}^*\Psi(\xi)\mathrm{d}\xi, \tag{4}$$

for any rapidly decaying smooth test functions ϕ and Ψ.

Proposition 3.4. *The infinitesimal generator $\mathcal{L}_Y$ of Y in* (2) *and its dual $\mathcal{L}_Y^*$ read*

$$(\mathcal{L}_Y f)(y) = \frac{1}{h}y(1-y)\frac{\partial}{\partial y}+\frac{1}{2}y^2\sigma^2(y)\frac{\partial^2}{\partial y^2} \qquad \text{and} \qquad (\mathcal{L}_Y^* f)(y) = -\frac{\partial}{\partial y}\left(\frac{1}{h}y(1-y)f(y)\right)+\frac{1}{2}\frac{\partial^2}{\partial y^2}\left(y^2\sigma^2(y)f(y)\right).$$

Proof. Using (4) and integration by part, we obtain the expression for $\mathcal{L}^*$. Given $y \in \mathbb{R}$ and $f, g : \mathbb{R} \to \mathbb{R}$ twice continuously differentiable functions with bounded derivatives, we have

$$\begin{aligned}
\langle f, \mathcal{L}_Y^* g\rangle &= \int_{\mathbb{R}} f(y)\left[-\frac{\partial}{\partial y}\left(\frac{1}{h}y(1-y)g(y)\right)+\frac{1}{2}\frac{\partial^2}{\partial y^2}\left(y^2\sigma^2(y)g(y)\right)\right]\mathrm{d}y,\\
&= -\int_{\mathbb{R}} \frac{\partial f}{\partial y}(y)\left\{-\frac{1}{h}y(1-y)g(y)+\frac{1}{2}\frac{\partial}{\partial y}\left(y^2\sigma^2(y)g(y)\right)\right\}\mathrm{d}y,\\
&= \frac{1}{h}\int_{\mathbb{R}} \frac{\partial f}{\partial y}(y)y(1-y)g(y)\mathrm{d}y - \frac{1}{2}\int_{\mathbb{R}} \frac{\partial f}{\partial y}(y)\frac{\partial}{\partial y}\left(y^2\sigma^2(y)g(y)\right)\mathrm{d}y,\\
&= \frac{1}{h}\int_{\mathbb{R}} \frac{\partial f}{\partial y}(y)y(1-y)g(y)\mathrm{d}y + \frac{1}{2}\int_{\mathbb{R}} \frac{\partial^2 f}{\partial y^2}(y)y^2\sigma^2(y)g(y)\mathrm{d}y,\\
&= \int_{\mathbb{R}} g(y)\left[\frac{1}{h}y(1-y)\frac{\partial f}{\partial y}(y)+\frac{1}{2}y^2\sigma^2(y)\frac{\partial^2 f}{\partial y^2}(y)\right]\mathrm{d}y,\\
&= \langle \mathcal{L}_Y f, g\rangle.
\end{aligned}$$

The rapidly decaying smooth test functions f and g ensure that the boundary terms in the integration by parts are equal to zero. □

In our setting, the process Y in (2) admits at least one stationary distribution, which can be proved easily following the arguments in [21].

Proposition 3.5. *The SDE* (2) *admits a stationary distribution Φ. However, since the map σ is not bounded away from zero, Φ might not be the unique solution to the Poisson equation $\mathcal{L}^*\Phi = 0$.*

Remark 3.6. For $f : \mathbb{R} \to \mathbb{R}$, finding the explicit solution of the Poisson equation is tedious. Indeed,

$$(\mathcal{L}_Y^* f)(y) = -\frac{\partial}{\partial y}\left(\frac{1}{h}y(1-y)f(y)\right)+\frac{1}{2}\frac{\partial^2}{\partial y^2}\left(y^2\sigma^2(y)f(y)\right) = 0,$$

is equivalent to

$$\begin{aligned}
&\frac{1}{2}y^2 f''(y)\left[\left(\frac{\alpha}{\beta}\right)^2 - \frac{2\alpha\gamma}{\beta}y^{-\beta} + \gamma^2 y^{-2\beta}\right]\\
&+ yf'(y)\left[2\left\{\left(\frac{\alpha}{\beta}\right)^2 - \frac{\alpha\gamma}{\beta}(2-\beta)y^{-\beta} + (1-\beta)\gamma^2 y^{-2\beta}\right\} - \frac{1}{h}(1-y)\right]\\
&+ f(y)\left[\left(\frac{\alpha}{\beta}\right)^2 + \frac{\alpha\gamma}{\beta}(1-\beta)(2\alpha+\beta)y^{-\beta} + \gamma^2(\beta-1)\left(\alpha^2(\beta-1)+\beta\right)y^{-2\beta} - \frac{1}{h}(1-2y)\right] = 0,
\end{aligned}$$

with the constraint $\int_{\mathbb{R}} f(y)\mathrm{d}y = 1$. This is a highly non-linear problem, which does not admit any explicit solution in general.

3.2.2. *Pricing PDE.* Consider an option with payoff $H(S_T, Y_T)$ at expiry T, and denote its price $P(t, S_t, Y_t)$ at time $t \le T$. Since the market is complete, we can construct a self-financing and riskless replicating portfolio consisting of the derivative itself and $-(SP_S + YP_Y)/S$ units of the underlying asset S to find the corresponding pricing PDE.

Proposition 3.7. *Under the risk-neutral measure $\mathbb{Q}$, the pricing PDE associated to* (2) *is*

$$\left(\partial_t + rs\partial_S + \left(\frac{1-y}{h}+r\right)y\partial_y + \frac{s^2}{2}\sigma^2(y)\partial_{ss} + sx\sigma^2(y)\partial_{sy} + \frac{y^2}{2}\sigma^2(y)\partial_{yy} - r\right)P(t,s,y) = 0, \tag{5}$$

for all $(t,s,y) \in [0,T) \times (0,\infty)^2$, with terminal condition $P(T,s,y) = H(s,y)$.

Equation (5) can be rewritten as $(\mathcal{L}_Y + \mathcal{L}_1 + \mathcal{L}_2)\,P = 0$, with $\mathcal{L}_Y$ defined in Proposition 3.4 and

$$\begin{cases} \mathcal{L}_1 := sy\sigma^2(y)\dfrac{\partial^2}{\partial s\partial y} + ry\dfrac{\partial}{\partial y}, \\ \mathcal{L}_2 := \dfrac{\partial}{\partial t} + \dfrac{s^2}{2}\sigma^2(y)\dfrac{\partial^2}{\partial s^2} + rs\dfrac{\partial}{\partial s} - r. \end{cases}$$

The operator $\mathcal{L}_2$ is the Black-Scholes infinitesimal generator with volatility $\sigma(y)$. Unfortunately, this pricing PDE does not admit an obvious explicit solution. Approximate solutions can be found by expanding the solution using perturbation methods, as developed in [9], but we leave this to future endeavours.

3.3. **Small-time asymptotics.** We now investigate the small-time behaviour of the solution to (2) using large deviations techniques. Consider the log-stock price process $X := \log S$. For $\varepsilon > 0$ and $t \in \mathcal{T}$, introduce the small-time rescaling $(X_t^\varepsilon, Y_t^\varepsilon) := (X_{\varepsilon t}, Y_{\varepsilon t})$. The model becomes

$$\begin{cases} \mathrm{d}X_t^\varepsilon = -\dfrac{\varepsilon}{2}\sigma^2(Y_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\sigma(Y_t^\varepsilon)\mathrm{d}W_t, & X_0^\varepsilon := x_0 = \log(S_0), \\ \mathrm{d}Y_t^\varepsilon = \varepsilon b(Y_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\widetilde{\sigma}(Y_t^\varepsilon)\mathrm{d}W_t, & Y_0^\varepsilon = y_0 > 0. \end{cases} \tag{6}$$

Let $\overline{\mathcal{H}}$ denote the space of absolutely continuous functions starting at the origin, with square integrable derivatives, such that

$$\overline{\mathcal{H}} := \left\{ f : \mathcal{T} \to \mathbb{R} \text{ such that } f = \int g(s)\mathrm{d}s \text{ on } \mathcal{T} \text{ for some } g \in L^2(\mathcal{T}), \text{ and } \inf_{t\in\mathcal{T}} f_t \geq \frac{1}{\alpha}\log\left(1 - y_0^\beta \frac{\alpha}{\beta\gamma}\right)\right\}.$$

Remark 3.8. When $y_0 \geq y_\sigma$, the condition

$$\inf_{t\in\mathcal{T}} f_t \geq \frac{1}{\alpha}\log\left(1 - y_0^\beta\frac{\alpha}{\beta\gamma}\right), \tag{7}$$

is automatically satisfied and the space $\overline{\mathcal{H}}$ boils down to the usual Cameron-Martin space. When $y_0 < y_\sigma$, Condition (7) is needed to ensure that the solution of the controlled ODE introduced below is positive.

The main result here is the following theorem, proved in Appendix A.5, which states a pathwise large deviations principle for the log-stock price process. With $\mathrm{x}_0 := (x_0, y_0)$, introduce the map $\mathrm{I}^{X,Y}$ on $\mathcal{C}(\mathcal{T}, \mathbb{R}\times\mathbb{R}_+^*)$ by

$$\mathrm{I}^{X,Y}(\mathrm{g}) := \inf\left\{\Lambda(f), f \in \overline{\mathcal{H}}, \mathcal{S}^{\mathrm{x}_0}(f) = \mathrm{g}\right\},$$

where Λ is the usual rate function of the standard Brownian motion:

$$\Lambda(f) := \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \left\|\dot{f}_t\right\|^2 \mathrm{d}t, & \text{if } f \in \mathcal{H}, \\ +\infty, & \text{otherwise,} \end{cases}$$

and $\mathcal{S}^{\mathrm{x}_0}(f)$ on $\mathcal{T}$ is the solution to the controlled ODE $\dot{\mathrm{g}}_t = (\sigma(\mathrm{g}_t), \widetilde{\sigma}(\mathrm{g}_t))^\top \cdot (\dot{f}_t, \dot{f}_t)$, starting from $\mathrm{g}_0 = \mathrm{x}_0$.

Theorem 3.9. *The rescaled* log*-stock price process X^ε satisfies a pathwise large deviations principle on $\mathcal{C}(\mathcal{T}, \mathbb{R})$ as ε tends to zero with speed ε and rate function*

$$\mathrm{I}^X(g) := \inf\left\{\mathrm{I}^{X,Y}(\mathrm{h}), \mathrm{h} := (g, l), l \in \mathcal{C}(\mathcal{T}, \mathbb{R}_+^*), l_0 = y_0\right\}, \tag{8}$$

The proof of the theorem relies on first obtaining a large deviations principle for the rescaled process Y^ε, which we state below (and defer its proof to Appendix A.4). Similarly to above, denote $\mathcal{S}^y(f)(t)$ the solution to the controlled ODE $\dot{g}_t = \widetilde{\sigma}(g_t)\dot{f}_t$, with $g_0 = y$.

Proposition 3.10. *The rescaled process Y^ε satisfies a pathwise large deviations principle on $\mathcal{C}(\mathcal{T}, \mathbb{R}_+^*)$ as ε tends to zero with speed ε and rate function*

$$\mathrm{I}^Y(g) := \inf\left\{\Lambda(f), f \in \overline{\mathcal{H}}, \mathcal{S}^{y_0}(f) = g\right\}. \tag{9}$$

Large deviations have been used extensively in Mathematical Finance in order to derive asymptotic behaviour of the implied volatility [10]. The latter, $\Sigma_t(k)$, is the unique non-negative solution to $\mathcal{C}_{BS}(t, \mathrm{e}^k, \Sigma_t(k)) = \mathcal{C}_{obs}(t, \mathrm{e}^k)$, with $\mathcal{C}_{obs}(t, \mathrm{e}^k)$ the observed Call option prices with maturity t and strike e^k, and $\mathcal{C}_{BS}$ is the Call option price in the Black-Scholes model.

Corollary 3.11. *The small-time asymptotic behaviour of the implied volatility is given by*

$$\lim_{t\downarrow 0}\Sigma_t(k) = \begin{cases} \dfrac{k^2}{2}\left(\inf\limits_{y\geq k}\mathrm{I}^X(g)|_{g(1)=y}\right)^{-1}, & \textit{if } k>0,\\ \dfrac{k^2}{2}\left(\inf\limits_{y\leq k}\mathrm{I}^X(g)|_{g(1)=y}\right)^{-1}, & \textit{if } k<0.\end{cases}$$

Proof. We only consider $k>0$, the other case being symmetric. It follows from [11, Corollary 7.1] that taking $\varepsilon = t$, from Theorem 3.9, X satisfies a large deviations principle with speed t and rate function I^X:

$$\lim_{t\downarrow 0} t\log\mathbb{P}(X\geq k) = -\inf_{y\geq k}\mathrm{I}^X(g)|_{g(1)=y}.$$

In the Black Scholes model, we have the following small-time implied volatility behaviour:

$$\lim_{t\downarrow 0} t\Sigma_t^2(k)\log\mathbb{P}(X\geq k) = -\frac{k^2}{2},$$

and the result follows from [11, Corollary 7.1]. □

4. Numerical estimation of the coefficients $(\alpha,\beta\,\gamma)$

We now devise an algorithm, using VIX and VVIX data, to find the optimal parameters (α,β,γ) appearing in the definition of the diffusion function σ in (3). We consider the continuously monitored formula for the VIX:

$$\mathrm{VIX}_t^2 := \mathbb{E}\left[\frac{1}{\Delta}\int_t^{t+\Delta}\mathrm{d}\langle X_s, X_s\rangle\mathrm{d}s \mid \mathcal{F}_t\right] = \frac{1}{\Delta}\mathbb{E}\left[\int_t^{t+\Delta}\sigma^2(Y_u)\mathrm{d}u\mid\mathcal{F}_t\right],$$

where Δ is equal to 30 days. In order to simulate the VIX on $[0,T]$, we must first approximate the integral and the conditional expectation, and we denote $\widehat{\mathrm{VIX}}_t$ its approximation: along a discretisation grid $\mathcal{U}_t := (u_i)_{i=0,\dots,N}$ with $u_i := t + i\frac{\Delta}{N}$, assume that $\sigma(Y_u)$ is constant on each interval, and write

$$\mathrm{VIX}_t^2 = \frac{1}{\Delta}\sum_{i=0}^{N-1}\int_{u_i}^{u_{i+1}}\mathbb{E}_t\left[\sigma^2(Y_u)\right]\mathrm{d}u \simeq \frac{1}{\Delta}\sum_{i=0}^{N-1}\mathbb{E}_t\left[\sigma^2(Y_{u_i})\right](u_{i+1}-u_i) = \frac{1}{N}\sum_{i=0}^{N-1}\mathbb{E}_t\left[\sigma^2(Y_{u_i})\right]$$

The second step is to approximate the conditional expectation. Since the process Y is Markovian, then $\mathbb{E}\left[\sigma^2(Y_{u_i})\mid\mathcal{F}_t\right] = \mathbb{E}\left[\sigma^2(Y_{u_i})\mid Y_t\right]$, for any $u_i\in\mathcal{U}_t$. We thus approximate the conditional expectation using the empirical mean, by simulating M paths $(Y^{(1)},\cdots,Y^{(M)})$ of Y, so that

$$\frac{1}{N}\sum_{i=0}^{N-1}\mathbb{E}_t\left[\sigma^2(Y_{u_i})\right] \simeq \frac{1}{NM}\sum_{j=1}^{M}\sum_{i=0}^{N-1}\sigma^2(Y_{u_i}^{(j)}) =: \widehat{\mathrm{VIX}}_t^2. \tag{10}$$

Note that, in order to simulate $\widehat{\mathrm{VIX}}_t^2$, one first needs to obtain the starting point Y_t, from which all the points $(Y_{u_i})_{i=1,\dots,N}$ can then be determined. Since $Y_t := \frac{S_t}{\overline{S}_t^h}$ from (1), with

$$\overline{S}_t^h := \frac{1}{h}\int_{-\infty}^t \mathrm{e}^{-\frac{t-v}{h}}S_v\mathrm{d}v \simeq \frac{1}{h}\sum_{i=0}^{I-1}\int_{v_i}^{v_{i+1}}\mathrm{e}^{-\frac{t-v}{h}}S_v\mathrm{d}v \simeq \frac{\mathrm{e}^{-t/h}}{h}\Delta_t\sum_{i=0}^{I-1}\mathrm{e}^{v_i/h}S_{v_i},$$

with $\mathcal{V}_t := (v_i)_{i=0,\dots,I}$ a time-discretisation of $(-\infty,t]$, such that v_0 is the first time for which we can observe the S&P 500, and $v_I = t$; we further set the grid size to be $v_{i+1}-v_i =: \Delta_t$ for all $i=0,\dots,I$. In our simulations, we will consider v_0 to be 25/04/2008. Empirically, in order for this approximation to be as precise as possible, one needs Δ_t to be as small as possible. However here, dealing with daily data for the S&P500 imposes $\Delta_t = 1$ day. We leave an extended analysis of this with high-frequency data to future work. Hence

$$Y_t \simeq \frac{S_t}{\overline{S}_t^h} \simeq \frac{S_t}{\frac{\mathrm{e}^{-t/h}}{h}\Delta_t\sum_{i=0}^{I-1}\mathrm{e}^{v_i/h}S_{v_i}} \simeq \frac{h\mathrm{e}^{t/h}S_t}{\Delta_t\sum_{i=0}^{I-1}\mathrm{e}^{v_i/h}S_{v_i}}. \tag{11}$$

Note that for any $t\in[0,T]$, the two time grids $\mathcal{V}_t$ and $\mathcal{U}_t$ do not overlap, the former taking into account time before t, the latter time after t. The algorithm to approximate $\left(\mathrm{VIX}_t^2\right)_{t\in[0,T]}$ is as follows:

Algorithm:

(i) Compute $(Y_t)_{t\in[0,T]}$ using (11), with S obtained from S&P 500 data;

(ii) For $t \in [0,T]$, using Y_t from (i), simulate the auxiliary process $(\widehat{Y}_s)_{s>t}$ on the grid $\mathcal{U}_t$ as

$$\widehat{Y}_{u_{i+1}} := \psi(\widehat{Y}_{u_i}, W_{u_i})\mathbb{1}_{\{\psi(\widehat{Y}_{u_i}, W_{u_i})>D\}},$$

where

$$\psi(x,w) := x\left(1 + \frac{\Delta_t}{h}(1-x) + \sigma(x)\Delta w\right),$$

with $\Delta_t := \frac{1}{252}$ and $\Delta W_{u_i} := W_{u_i+1} - W_{u_i} \simeq \frac{1}{\sqrt{252}}\mathcal{N}(0,1)$, and obtain $\frac{1}{N}\sum_{i=0}^{N-1}\sigma^2(\widehat{Y}_{u_i})$;

(iii) Repeat step (ii) M times to obtain $\widehat{\mathrm{VIX}}_t^2$ from (10);

(iv) Repeat steps (ii)-(iii) on the grid to obtain $\widehat{\mathrm{VIX}}^2$ on $\mathcal{V}_T \cap [0,T]$.

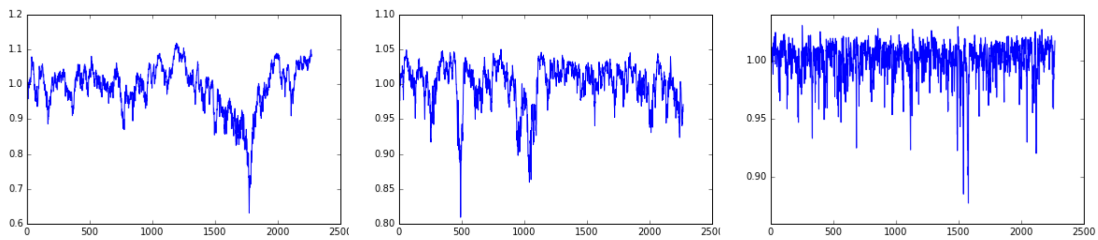
Remark 4.1. In the Euler scheme in step (ii) of the algorithm, we introduced the threshold D to prevent the simulations from 'exploding' when Y approaches zero. We could alternatively apply the threshold on $\sigma(Y_t)$, but we leave the precise analysis of the (weak or strong) convergence of our discretisation for future work, hints thereabout can be picked from [17].

4.1. **Simulations.** We use daily data to simulate VIX^2, between January 2nd, 2013 and May 8th, 2017 (corresponding to $T = 1200$ days). The simulations are run for $h \in \{180, 30, 5\}$ days with different parameters (α, β, γ), first taken arbitrarily, but optimised over later. Table 1 summarises them:

	$h = 180$	$h = 30$	$h = 5$
α	0.9	2.1	5
β	1.5	1.2	6
γ	0.8	1.9	1.7
N	500	500	500
M	500	500	500
D	0	0.01	1

TABLE 1. Parameters used in the simulations

Note first from (2) that the process is mean reverting to 1. The smaller the h, the larger the mean reversion $1/h$, which is clearly visible in Figure 2.

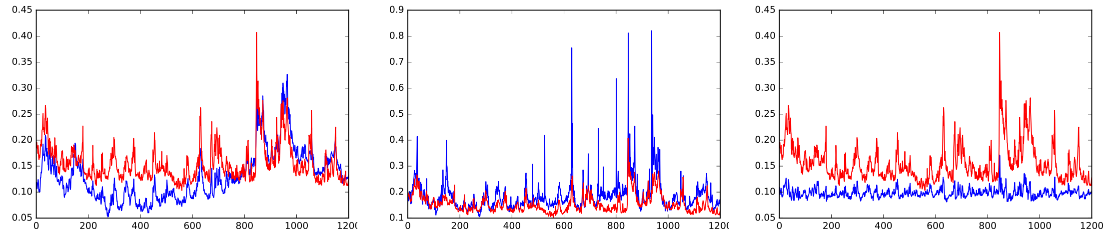


(A) $h = 180$ days (B) $h = 30$ days (C) $h = 5$ days

FIGURE 2. Process Y simulated using the Euler scheme between 2/1/2013 and 8/5/2017.

Using the algorithm above, the simulations for the VIX^2, respectively for $h = 180$, 30 and 5 days, are depicted in Figure 3.

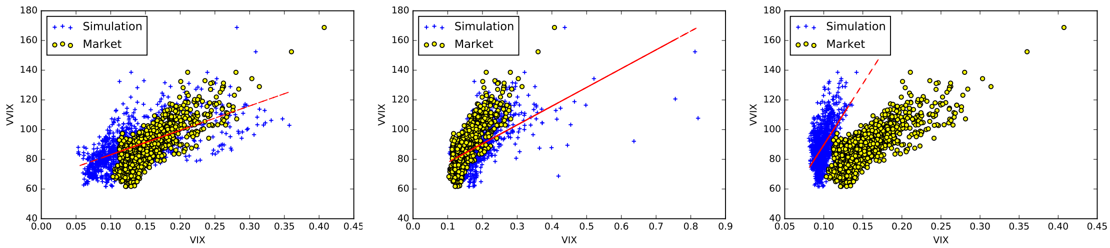
In all cases, the simulated VIX^2 seems to have jumps similar to those visible in the historical VIX^2. However, when $h = 5$ days, the variations of the simulated VIX^2 are very small compared to the historical one. In order to obtain similar magnitudes, for $h = 5$ days, we had to consider a threshold $D = 1$, meaning



(A) $h = 180$ days (B) $h = 30$ days (C) $h = 5$ days

FIGURE 3. Simulated (blue) and historical (red) VIX, between 2/1/2013 and 8/5/2017.

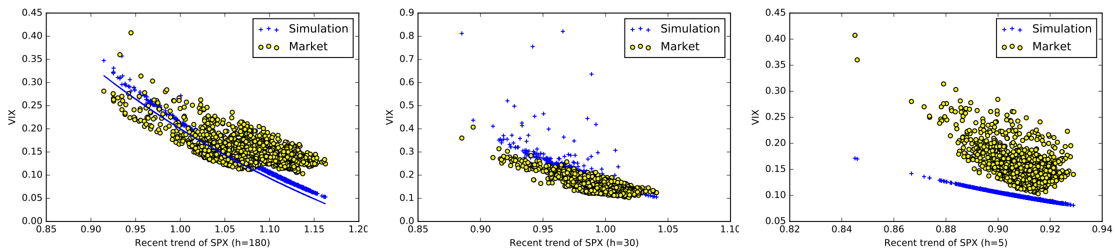
that many simulated paths were discarded, thus creating a clear bias. We also chose $(\alpha, \beta, \gamma) = (5, 6, 1.5)$, giving unrealistic volatility values, for example $\sigma(1) = 0.66$ and confirming that the representation (3) works best for $h > \Delta = 30$. In order to analyse the results, we plot in Figure 4 the simulated VIX and the historical VIX against the historical VVIX as well as against the recent trend of S&P 500, that is, the process Y simulated using its definition $Y_t := \frac{S_t}{S_t^h}$ for $t \in [0, T]$ and $h \in \{5, 30, 180\}$.



(A) $h = 180$ days (B) $h = 30$ days (C) $h = 5$ days

FIGURE 4. Simulated (blue crosses) and historical (yellow circles) VIX against historical VVIX, between 2/1/13 and 8/5/17. The red line is the linear regression fit.

In the first two figures (for $h = 180$ and 30 days), it seems to fit the regression of the simulated data as well. However, for $h = 5$ days, the simulated values clearly follow a different trend. This might be explained by the presence of a large threshold D in the simulations, and by the fact that the approximation for σ using the VIX data is only valid for $h > \Delta$. Finally, Figure 5 shows the simulated VIX and historical VIX against the recent trend of S&P 500. The latter are computed using the definition of Y changing the window h to adapt it to each case. We observe that the model does not fit market data for $h = 5$, which is consistent with the assumption that the representation of the function σ is only valid for $h > \Delta$.



(A) $h = 180$ days (B) $h = 30$ days (C) $h = 5$ days

FIGURE 5. Simulated (blue crosses) and historical (yellow circles) VIX against the recent trend of S&P500, between 2/1/13 and 8/5/17. *Source: CBOE VIX Index.*

4.2. **Optimisation of coefficients.** The calibration of the model is performed by minimising the least squared error (LSE)

$$\mathrm{LSE}(\alpha,\beta,\gamma) := \sum_{t\in\mathcal{V}_T\cap[0,T]} \left|\mathrm{VIX}_t - \widehat{\mathrm{VIX}}_t(\alpha,\beta,\gamma)\right|^2$$

over the coefficients (α,β,γ). Table 2 summarises the results of the optimisation, indicating the optimised parameters as well as the initial guesses and the number of runs. The optimal coefficients are close to the initial parameters $(\alpha_0,\beta_0,\gamma_0)$. This is due to starting with good guesses, in order to reduce the actual computation time, quite involved in the numerics (which can be seen from the small number of simulations $(N,M)=(100,200)$ considered).

	$h=180$	$h=30$	$h=5$
α_0	0.9	1.9	3.1
β_0	1.5	1.2	8.1
γ_0	0.8	1.8	0.5
α^*	0.88	1.96	3.15
β^*	1.52	1.21	8.01
γ^*	0.83	1.76	0.48
T	500	200	100
N	100	200	100
M	200	100	200
D	0	0.01	0

Table 2. Initial and calibrated parameters

Appendix A. Additional Proofs

A.1. **Proof of Proposition 3.2.** We start with the behaviour at the origin. Regarding the first statement, for $\varepsilon>0$,

$$\int_{-\varepsilon}^{\varepsilon}\frac{1+|b(x)|}{\widetilde{\sigma}^2(x)}\mathrm{d}x \le 2\left[1+\max_{x\in[-\varepsilon,\varepsilon]}|b(x)|\right]\int_0^{\varepsilon}\frac{\mathrm{d}x}{\widetilde{\sigma}^2(x)}, \le 2\frac{\left[1+\max_{x\in[-\varepsilon,\varepsilon]}|b(x)|\right]}{\gamma^2+\delta}\int_0^{\varepsilon}\frac{\mathrm{d}x}{x^{2(1-\beta)}},$$

as there exists $\delta\in(-\gamma^2,\infty)$ and $\overline{x}>0$ such that for $x\le\overline{x}$, $\gamma^2-2\frac{\alpha\gamma}{\beta}x^\beta+\frac{\alpha^2}{\beta^2}x^{2\beta}\ge\gamma^2+\delta>0$. This integral is finite if and only if $\beta>\frac{1}{2}$. Now, for $x>0$, we have

$$\lim_{x\downarrow 0^+}\frac{b(x)}{\frac{1}{h}x} = \lim_{x\downarrow 0^+}(1-x) = 1 \qquad\text{and}\qquad \lim_{x\downarrow 0^+}\frac{\widetilde{\sigma}(x)}{\gamma x^{1-\beta}} = \lim_{x\downarrow 0^+}\left(1-\frac{\alpha}{\beta\gamma}x^\beta\right) = 1.$$

We can then apply [3, Theorem 5.3] for $\overline{\gamma}:=2\beta-1\neq-1$ as $\beta>0$, and conclude.

Regarding the behaviour at infinity, for $x>0$, we have, as $\beta>0$

$$\lim_{x\uparrow\infty}\frac{b(x)}{-\frac{1}{h}x^2} = 1 \qquad\text{and}\qquad \lim_{x\uparrow\infty}\frac{\widetilde{\sigma}(x)}{-\frac{\alpha}{\beta}x} = \lim_{x\uparrow\infty}\left\{1-\frac{\beta\gamma}{\alpha}x^{-\beta}\right\} = 1.$$

We can then apply [3, Theorem 5.7] for $\overline{\gamma}=0>-1$ and $\overline{\lambda}:=-\frac{\beta^2}{h\alpha^2}<0$ and ∞ has type A.

A.2. **Proof of Theorem 3.3(i).** We first show that y_σ is an isolated point. For $\varepsilon>0$,

$$\begin{aligned}\int_{y_\sigma}^{y_\sigma+\varepsilon}\frac{1+|b(x)|}{\widetilde{\sigma}^2(x)}\mathrm{d}x &\ge \int_{y_\sigma}^{y_\sigma+\varepsilon}\frac{\mathrm{d}x}{\widetilde{\sigma}^2(x)} = \int_{y_\sigma}^{y_\sigma+\varepsilon}\frac{\mathrm{d}x}{x^2\left(-\frac{\alpha}{\beta}+\gamma x^{-\beta}\right)^2}\\ &\ge K_\varepsilon\int_{y_\sigma}^{y_\sigma+\varepsilon}\frac{\mathrm{d}x}{\left(-\frac{\alpha}{\beta}x^\beta+\gamma\right)^2} = \frac{K_\varepsilon}{\gamma^2}\int_0^{\varepsilon}\frac{\mathrm{d}x}{\left(1-\frac{\alpha}{\beta\gamma}x_\sigma^\beta\left(1+\frac{x}{y_\sigma}\right)^\beta\right)^2}\end{aligned} \tag{12}$$

with $K_\varepsilon := (y_\sigma+\varepsilon)^{-2(1-\beta)}$ for $\beta\in(0,1)$, $K_\varepsilon := y_\sigma^{2(\beta-1)}$ for $\beta>1$ and $K_\varepsilon := 1$ for $\beta=1$. Introduce the following quantities:

$$\left\{\begin{array}{ll} A_\beta & := \dfrac{(\beta-1)(\beta-2)}{2}, \\ B_\beta & := A_\beta\left[A_\beta - \dfrac{\beta-3}{3}\right], \\ C_\beta & := A_\beta\left[-\dfrac{(\beta-3)(\beta-4)}{12} + \dfrac{(\beta-1)(\beta-2)(\beta-3)}{3} - A_\beta^2\right]. \end{array}\right. \tag{13}$$

A straightforward Taylor expansion around the origin then yields

$$\left[1-\frac{\alpha}{\beta\gamma}x_\sigma^\beta\left(1+\frac{x}{y_\sigma}\right)^\beta\right]^{-2} = \frac{y_\sigma^2}{\beta^2x^2}\left\{\left[1-\frac{2A_\beta x}{y_\sigma}+(2B_\beta+A_\beta^2)\frac{x^2}{y_\sigma^2}\right]+\left[2(C_\beta-A_\beta B_\beta)\frac{x^3}{y_\sigma^3}+o(x^3)\right]\right\},$$

which is not integrable around zero. We thus conclude about the right behaviour of y_σ by noting that the last term in (12) diverges. Using similar arguments,

$$\int_{y_\sigma-\varepsilon}^{y_\sigma}\frac{1+|b(x)|}{\widetilde{\sigma}^2(x)}\mathrm{d}x \geq \int_{y_\sigma-\varepsilon}^{y_\sigma}\frac{\mathrm{d}x}{\widetilde{\sigma}^2(x)} = \int_{y_\sigma-\varepsilon}^{y_\sigma}\frac{\mathrm{d}x}{x^2\left(-\frac{\alpha}{\beta}+\gamma x^{-\beta}\right)^2} \geq \overline{K}_\varepsilon\int_{y_\sigma-\varepsilon}^{y_\sigma}\frac{\mathrm{d}x}{\left(-\frac{\alpha}{\beta}x^\beta+\gamma\right)^2}=\infty,$$

with $\overline{K}_\varepsilon := y_\sigma^{-2(1-\beta)}$ for $\beta\in(0,1)$, $\overline{K}_\varepsilon := (y_\sigma-\varepsilon)^{2(\beta-1)}$ for $\beta>1$ and $\overline{K}_\varepsilon := 1$ for $\beta=1$.

Introduce, for $a>0$ and $x\in(0,a]$, the following functions:

$$\rho(x) := \exp\left(\int_x^a\frac{2\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y\right),\quad s(x) := \left\{\begin{array}{ll}\displaystyle\int_0^x\rho(y)\mathrm{d}y, & \text{if } y_\sigma\geq 1,\\ \displaystyle-\int_x^a\rho(y)\mathrm{d}y, & \text{if } y_\sigma<1,\end{array}\right.\quad\text{and}\quad \varphi(x) := \frac{1+|\overline{b}(x)|}{\overline{\sigma}^2(x)}. \tag{14}$$

We rely here on the analysis in [3, Section 2.3]. The key ingredient is the following lemma:

Lemma A.1. *The following hold:*

$$\begin{array}{lllll} \displaystyle\lim_{x\downarrow 0}\rho(x)=0, & \displaystyle\int_0^a\rho(y)\mathrm{d}y<\infty, & \displaystyle\int_0^a\frac{\varphi(x)}{\rho(x)}\mathrm{d}x=\infty, & \displaystyle\int_0^a\varphi(x)s(x)\mathrm{d}x<\infty, & \textit{if } y_\sigma\geq 1,\\ \displaystyle\lim_{x\downarrow 0}\rho(x)=\infty, & \displaystyle\int_0^a\rho(y)\mathrm{d}y=\infty, & & \displaystyle\int_0^a\frac{\varphi(x)}{\rho(x)}|s(x)|\mathrm{d}x<\infty, & \textit{if } y_\sigma<1.\end{array}$$

Introduce now the process $Z := (Y-y_\sigma)$, satisfying the SDE

$$\mathrm{d}Z_t = \overline{b}(Z_t)\mathrm{d}t+\overline{\sigma}(Z_t)\mathrm{d}W_t,\qquad Z_0 := y_0-y_\sigma>0.$$

with $\overline{b}(x) := b(x+y_\sigma)$ and $\overline{\sigma}(x) := \widetilde{\sigma}(x+y_\sigma)$ for $x>0$. Armed with Lemma A.1, the right behaviour of Z at the origin, which corresponds to the right behaviour for the original process Y at y_σ. The case $y_\sigma\geq 1$ follows from [3, Theorem 2.13], whereas the case $y_\sigma<1$ follows from [3, Theorem 2.16].

Proof of Lemma A.1. We start with the limiting behaviour of the function ρ and its integral. A straightforward Taylor expansion around the origin yields

$$\left[1-\frac{\alpha}{\beta\gamma}x_\sigma^\beta\left[1+\frac{y}{y_\sigma}\right]^\beta\right]^{-2} = \frac{y_\sigma^2}{\beta^2y^2}\left[\left[1-2A_\beta\frac{y}{y_\sigma}+(2B_\beta+A_\beta^2)\frac{y^2}{y_\sigma^2}\right]+\left[2(C_\beta-A_\beta B_\beta)\frac{y^3}{y_\sigma^3}+o(x^3)\right]\right], \tag{15}$$

with $A_\beta, B_\beta, C_\beta$ introduced in (13). Introduce $K := \frac{2x_\sigma^{2\beta+1}(1-y_\sigma)}{h\beta^2\gamma^2}\leq 0$ as $y_\sigma\geq 1$ and $K_\beta^a := -\frac{1}{a}-\frac{2A_\beta}{y_\sigma}\log(a)+\frac{2B_\beta+A_\beta^2}{y_\sigma^2}a$.

When $y_\sigma\geq 1$, using (15), we obtain the asymptotic behaviour, as x approaches zero,

$$\frac{2\overline{b}(x)}{\overline{\sigma}^2(x)} = \frac{K}{x^2}\left(1-2A_\beta\frac{x}{y_\sigma}+(2B_\beta+A_\beta^2)\left(\frac{x}{y_\sigma}\right)^2+2(C_\beta-A_\beta B_\beta)\left(\frac{x}{y_\sigma}\right)^3+o(x^3)\right). \tag{16}$$

As the expansion is uniform on $[x,a]$, one obtains

$$\begin{aligned}\rho(x) &= \exp\left(2\int_x^a \frac{\bar{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y\right) = \exp\left(\frac{K}{x}\left(1+\frac{2A_\beta}{y_\sigma}x\log x\right)\right)\exp\left(KK_\beta^a - \frac{2B_\beta + A_\beta^2}{y_\sigma^2}Kx + o(x)\right)\\ &= \exp\left(\frac{K}{x} + KK_\beta^a\right)x^{\frac{2A_\beta K}{y_\sigma}}(1+o(x)).\end{aligned} \tag{17}$$

Since $K \leq 0$, $\rho(x)$ tends to zero as x tends to zero from above, and $\int_0^a \rho(x)\mathrm{d}x$ is finite.

In the case $y_\sigma < 1$, the expansion (17) is still valid, albeit with $K \geq 0$. Therefore ρ explodes at the origin and $\int_0^a \rho(y)\mathrm{d}y$ is infinite.

Now, it is straightforward to see that

$$\int_0^a \frac{1+|\bar{b}(x)|}{\rho(x)\overline{\sigma}^2(x)}\mathrm{d}x \geq \int_0^a \frac{\mathrm{d}x}{\rho(x)\widetilde{\sigma}^2(x+y_\sigma)} \geq \int_0^a \frac{\mathrm{d}x}{\widetilde{\sigma}^2(x+y_\sigma)},$$

which is clearly infinite., because ρ is bounded above by 1 on $(0,a]$ and from the asymptotic behaviour of the integrand around the origin, analysed in (12).

We now prove the last statement of the lemma, and start with the case $y_\sigma \geq 1$. With this, $\int_0^a \rho(y)\mathrm{d}y$ is finite and $s(x) = \int_0^x \rho(y)\mathrm{d}y$ for $x \in (0,a]$. Using (17), we can write the asymptotic behaviour of $s(\cdot)$ around the origin by integrating the asymptotic behaviour of $\rho(\cdot)$ around zero. Classical asymptotic expansions for integrals [19, Chapter 3.3] (note that the leading contribution arises at the right boundary of the integration domain) yields, after the change of variable $y \mapsto zx$,

$$s(x) = \int_0^x \rho(y)\mathrm{d}y = \mathrm{e}^{KK_\beta^a}x^{\frac{2A_\beta K}{y_\sigma}+1}\int_0^1 \exp\left(\frac{K}{zx}\right)z^{\frac{2A_\beta K}{y_\sigma}}\mathrm{d}z = \mathrm{e}^{KK_\beta^a}x^{\frac{2A_\beta K}{y_\sigma}+1}\exp\left\{\frac{K}{x}\right\}\left(-\frac{1}{K}x + o(x)\right),$$

as x tends to zero. Combining this with (15) and (17), we obtain

$$\frac{s(x)}{\rho(x)\overline{\sigma}^2(x)} = \frac{x_\sigma^{2\beta}}{\beta^2\gamma^2}x\left(-\frac{1}{K}x + o(x)\right)\frac{1}{x^2}\left(1+o(1)\right) = -\frac{x_\sigma^{2\beta}}{K\beta^2\gamma^2}\left(1+o(1)\right),$$

which is integrable on $(0,a]$ and concludes the proof, using the fact that

$$\int_0^a \frac{1+|\bar{b}(x)|}{\rho(x)\overline{\sigma}^2(x)}s(x)\mathrm{d}x \leq \left(1 + \max_{x\in(0,a]}|\bar{b}(x)|\right)\int_0^a \frac{s(x)}{\rho(x)\overline{\sigma}^2(x)}\mathrm{d}x < \infty.$$

Consider now the case where $y_\sigma < 1$. $\int_0^a \frac{1+|\bar{b}(x)|}{\rho(x)\overline{\sigma}^2(x)}|s(x)|\mathrm{d}x \leq \left[1+\max_{(0,a]}|\bar{b}|\right]\int_0^a \frac{|s(x)|}{\rho(x)\overline{\sigma}^2(x)}\mathrm{d}x$. Since both $|s(x)|$ and $\rho(x)\overline{\sigma}^2(x)$ diverge to infinity at the origin, we need to study the behaviour of the integrand around zero to be able to conclude. We already proved that $\int_0^a \rho(x)\mathrm{d}x$ is infinite for $y_\sigma < 1$. Hence, for $x \in (0,a]$, we obtain $s(x) := -\int_x^a \rho(y)\mathrm{d}y$. For $\delta > 0$ such that $x < \delta < a$ and $x > 0$, $\int_x^a \rho(y)\mathrm{d}y = \int_x^\delta \rho(y)\mathrm{d}y + \int_\delta^a \rho(y)\mathrm{d}y$. The second integral exists as the integral of a continuous function over a closed interval in $\mathbb{R}_+$. Regarding the first one, classical asymptotic expansions for integrals and (17), yield, after the change of variable $y \mapsto zx$,

$$\int_x^\delta \rho(y)\mathrm{d}y = \mathrm{e}^{KK_\beta^a}x^{1+\frac{2A_\beta K}{y_\sigma}}\int_1^{\delta/x}\exp\left(\frac{K}{xz}\right)z^{\frac{2A_\beta K}{y_\sigma}}\mathrm{d}z = \mathrm{e}^{KK_\beta^a}x^{1+\frac{2A_\beta K}{y_\sigma}}\mathrm{e}^{\frac{K}{x}}\left(\frac{x}{K}+o(x)\right),$$

and the asymptotic behaviour of the integrand around the origin becomes

$$\frac{|s(x)|}{\rho(x)\overline{\sigma}^2(x)} = \frac{x_\sigma^{2\beta}}{\beta^2\gamma^2}\left(\frac{x^2}{K}+o(x^2)\right)\frac{1}{x^2}(1+o(1)) = \frac{x_\sigma^{2\beta}}{\beta^2\gamma^2 K}(1+o(1)),$$

which is integrable at the origin, and the claim is proved. □

A.3. **Proof of Theorem 3.3(ii).** The strategy to prove the proposition is similar, albeit with different computations, to the previous case. Introduce, for $a > 0$ and $x \in (0,a]$,

$$\rho(x) := \exp\left(\int_x^a \frac{2\bar{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y\right) \qquad \text{and} \qquad s(x) := \begin{cases}\displaystyle\int_0^x \rho(y)\mathrm{d}y, & \text{if } y_\sigma < 1,\\ \displaystyle -\int_x^a \rho(y)\mathrm{d}y, & \text{if } y_\sigma \geq 1.\end{cases} \qquad \text{and} \qquad \varphi(x) := \frac{1+|\bar{b}(x)|}{\overline{\sigma}^2(x)}.$$

In order to determine the left behaviour of y_σ, we first state the following lemmas. They will enable us to decide which theorem to use from [3, Section 2.3].

Lemma A.2.

$$\begin{array}{llll} \lim_{x\downarrow 0^+} \rho(x) = \infty, & \int_0^a \rho(y)\mathrm{d}y = \infty, & & \int_0^a \frac{\varphi(x)}{\rho(x)}|s(x)|\mathrm{d}x < \infty, \quad \text{if } y_\sigma \geq 1, \\ \lim_{x\downarrow 0^+} \rho(x) = 0, & \int_0^a \rho(y)\mathrm{d}y < \infty, & \int_0^a \frac{\varphi(x)}{\rho(x)}\mathrm{d}x = \infty, & \int_0^a \varphi(x)s(x)\mathrm{d}x < \infty, \quad \text{if } y_\sigma < 1. \end{array}$$

Similarly to before, introduce the process $Z := y_\sigma - Y$ satisfying the SDE

$$\mathrm{d}Z_t = \overline{b}(Z_t)\mathrm{d}t + \overline{\sigma}(Z_t)\mathrm{d}W_t, \qquad Z_0 := y_\sigma - y_0 > 0,$$

as well as the maps $\overline{b}(x) := -b(y_\sigma - x)$ and $\overline{\sigma}(x) := -\widetilde{\sigma}(y_\sigma - x)$ for $x > 0$. With Lemma A.2, we obtain the left behaviour of Z at the origin, corresponding to the left behaviour of y_σ for Y. The case $y_\sigma < 1$ follows from [3, Theorem 2.13], and the case $y_\sigma \geq 1$ from [3, Theorem 2.16], and Theorem 3.3(ii) follows.

Proof of Lemma A.2. A straightforward Taylor expansion around the origin yields

$$\left[1 - \frac{\alpha x_\sigma^\beta}{\beta\gamma}\left[1 - \frac{y}{y_\sigma}\right]^\beta\right]^{-2} = \frac{y_\sigma^2}{\beta^2 y^2}\left\{\left[1 + \frac{2A_\beta y}{y_\sigma} + (2B_\beta + A_\beta^2)\frac{y^2}{y_\sigma^2}\right] + \left[2(A_\beta B_\beta - C_\beta)\frac{y^3}{y_\sigma^3} + o(x^3)\right]\right\}, \tag{18}$$

with $A_\beta, B_\beta, C_\beta$ defined in (13). We start with the behaviour of the function ρ and its integrated version.

Consider first the case $y_\sigma \geq 1$. We split the range of possibilities into three possible intervals for a:

(i) If $a < y_\sigma - 1$, then $y_\sigma - x \geq y_\sigma - a > 1$ and b is negative on $[y_\sigma - a, y_\sigma - x]$. Then, for $x \in (0, a]$,

$$\begin{aligned} \int_x^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y = \int_x^a \frac{-b(y_\sigma - y)}{\widetilde{\sigma}^2(y_\sigma - y)}\mathrm{d}y &= -\int_x^a \frac{b(y_\sigma - y)\mathrm{d}y}{(y_\sigma - y)^{2(1-\beta)}\gamma^2\left(1 - \frac{\alpha}{\beta\gamma}x_\sigma^\beta(1 - \frac{y}{y_\sigma})^\beta\right)^2} \\ &= \frac{1}{h}\int_x^a \frac{(y_\sigma - y)^{2\beta-1}(y_\sigma - y - 1)}{\gamma^2\left(1 - \frac{\alpha}{\beta\gamma}x_\sigma^\beta(1 - \frac{y}{y_\sigma})^\beta\right)^2}\mathrm{d}y \\ &\geq \frac{(y_\sigma - a)^{2\beta-1}(y_\sigma - a - 1)}{\gamma^2 h}\int_x^a \frac{\mathrm{d}y}{\left(1 - \frac{\alpha}{\beta\gamma}x_\sigma^\beta(1 - \frac{y}{y_\sigma})^\beta\right)^2}, \end{aligned}$$

as $\min_{y\in[x,a]}\left[(y_\sigma - y)^{2\beta-1}(y_\sigma - y - 1)\right] = (y_\sigma - a)^{2\beta-1}(y_\sigma - a - 1) > 0$. Indeed the map $y \mapsto y^{2\beta-1}(y-1)$ is increasing on $[y_\sigma - a, y_\sigma - x]$ because $y_\sigma - a > 1$. Noting that (18) is uniform on $[x, a]$, we obtain, as x approaches zero

$$\exp\left\{\frac{2(y_\sigma - a)^{2\beta-1}(y_\sigma - a - 1)}{\gamma^2 h}\int_x^a \frac{\mathrm{d}y}{\left(1 - \frac{\alpha}{\beta\gamma}x_\sigma^\beta(1 - \frac{y}{y_\sigma})^\beta\right)^2}\right\} = \exp\left\{\frac{\overline{K}}{x} + \overline{K}\overline{K}_\beta^a\right\}x^{-\frac{2A_\beta\overline{K}}{y_\sigma}}(1 + o(x)),$$

with $\overline{K} := \frac{(y_\sigma - a)^{2\beta-1}(y_\sigma - a - 1)y_\sigma^2}{h\beta^2\gamma^2} > 0$ and $\overline{K}_\beta^a := -\frac{1}{a} + \frac{2A_\beta}{y_\sigma}\log(a) + \frac{2B_\beta + A_\beta^2}{y_\sigma^2}a$, and therefore $\lim_{x\downarrow 0}\rho(x) = \infty$ and $\int_0^a \rho(x)\mathrm{d}x = \infty$.

(ii) If $y_\sigma - 1 \leq a < y_\sigma$, then for $x \in (0, a]$,

$$\int_x^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y = \int_x^{y_\sigma - 1} \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y + \int_{y_\sigma - 1}^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y.$$

Similarly to (i), one can prove that $\lim_{x\downarrow 0}\int_x^{y_\sigma - 1}\frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y = \infty$. Then, on $(y_\sigma - 1, a]$, $\overline{\sigma}$ does not go to zero and is continuous, thus bounded; similarly, $\overline{b}$ is negative and continuous, hence bounded on $(y_\sigma - 1, a]$. Therefore $\int_{y_\sigma - 1}^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y < \infty$, for $x \in (0, a]$, and

$$\lim_{x\downarrow 0}\rho(x) = \exp\left(\int_{y_\sigma - 1}^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y\right)\exp\left(\int_x^{y_\sigma - 1} \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y\right) = \infty.$$

(iii) If $y_\sigma \le a$, then for $x \in (0,a]$,

$$\int_x^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y = \int_x^{y_\sigma - 1} \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y + \int_{y_\sigma-1}^{y_\sigma} \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y + \int_{y_\sigma}^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y.$$

Similarly to (ii), it is easy to show that $\int_{y_\sigma-1}^{y_\sigma} \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y = -\int_0^1 \frac{b(y)}{\widetilde{\sigma}^2(y)}\mathrm{d}y \le 0$. Then, as $\overline{b}$ is positive on $(y_\sigma, a]$, $\int_{y_\sigma}^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y \ge 0$, and $\lim_{x\downarrow 0}\rho(x) = \infty$.

Consider now the case $y_\sigma < 1$. Using (18), we have the asymptotic behaviour, as x approaches zero

$$\frac{2\overline{b}(x)}{\overline{\sigma}^2(x)} = \frac{\widetilde{K}}{x^2}\left(1 + 2A_\beta \frac{x}{y_\sigma} + (2B_\beta + A_\beta)\left(\frac{x}{y_\sigma}\right)^2 + 2(a_\beta B_\beta - C_\beta)\left(\frac{x}{y_\sigma}\right)^3 + o(x^3)\right), \tag{19}$$

with $\widetilde{K} := \frac{2x_\sigma^{2\beta+1}(y_\sigma - 1)}{h\beta^2\gamma^2} < 0$ as $y_\sigma < 1$. Since the expansion is again uniform on $[x,a]$, we obtain

$$\rho(x) = \exp\left\{2\int_x^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y\right\} = \exp\left\{\frac{\widetilde{K}}{x}\left(1 - \frac{2A_\beta}{y_\sigma}x\log x\right)\right\}\exp\left\{\widetilde{K}\overline{K}_\beta^a - \frac{2B_\beta + A_\beta^2}{y_\sigma^2}\widetilde{K}x + o(x)\right\},$$

$$= \exp\left\{\frac{\widetilde{K}}{x} + \widetilde{K}\overline{K}_\beta^a\right\} x^{-\frac{2A_\beta\widetilde{K}}{y_\sigma}}(1 + o(x)). \tag{20}$$

Since $\widetilde{K} < 0$, we easily deduce that $\lim_{x\downarrow 0}\rho(x) = 0$, and $\int_0^a \rho(x)\mathrm{d}x$ is finite.

The middle statement in the lemma is straightforward. When $x \in (0,a]$, $\int_x^a \frac{\overline{b}(y)}{\overline{\sigma}^2(y)}\mathrm{d}y = -\int_{y_\sigma-a}^{y_\sigma - x} \frac{b(y)}{\widetilde{\sigma}^2(y)}\mathrm{d}y$. Since $0 < y_\sigma - a < y_\sigma - x < 1$, b is positive on $[y_\sigma - a, y_\sigma - x]$ and the above integral is therefore negative. Hence, ρ is bounded by 1 on $(0,a]$, and

$$\int_0^a \frac{1 + |\overline{b}(x)|}{\overline{\sigma}^2(x)}\mathrm{d}x \ge \int_0^a \frac{\mathrm{d}x}{\widetilde{\sigma}^2(y_\sigma - x)} = \infty,$$

using (18), which concludes the proof.

The final integrals in the lemma are rather delicate to analyse. We start with the case $y_\sigma < 1$. As $\int_0^a \rho(x)\mathrm{d}x < \infty$, $s(x) = \int_0^x \rho(y)\mathrm{d}y$. Using (20), one can obtain the asymptotic behaviour of $s(\cdot)$ around zero by integrating the asymptotic behaviour of $\rho(\cdot)$ around 0. Classical asymptotic expansions for integrals (note that the leading contribution arises at the right boundary of the integration domain) yields, after the change of variable $y \mapsto xz$,

$$\begin{aligned} s(x) &= \int_0^x \rho(y)\mathrm{d}y = \mathrm{e}^{\widetilde{K}\overline{K}_\beta^a} x^{1-\frac{2A_\beta\widetilde{K}}{y_\sigma}} \int_0^1 \exp\left(\frac{\widetilde{K}}{xz}\right) z^{\frac{-2A_\beta\widetilde{K}}{y_\sigma}}\mathrm{d}z, \\ &= \mathrm{e}^{\widetilde{K}\overline{K}_\beta^a} x^{1-\frac{2A_\beta\widetilde{K}}{y_\sigma}} \exp\left(\frac{\widetilde{K}}{x}\right)\left(-\frac{x}{\widetilde{K}} + o(x)\right), \quad \text{as } x \text{ tends to zero.} \end{aligned} \tag{21}$$

Therefore, combining (18), (20) and (21), we obtain

$$\frac{s(x)}{\rho(x)\overline{\sigma}^2(x)} = x\left(-\frac{1}{\widetilde{K}}x + o(x)\right)\frac{x_\sigma^{2\beta}}{\beta^2\gamma^2}\frac{1}{x^2}(1 + o(1)) = -\frac{x_\sigma^{2\beta}}{\beta^2\gamma^2\widetilde{K}}(1 + o(1)),$$

which is integrable on $(0,a]$ and concludes the proof.

We now move on to the case where $y_\sigma \ge 1$. In that case,

$$\int_0^a \frac{1 + |\overline{b}(x)|}{\rho(x)\overline{\sigma}^2(x)}|s(x)|\mathrm{d}x \le \left[1 + \max_{(0,a]}|\overline{b}|\right]\int_0^a \frac{|s(x)|}{\rho(x)\overline{\sigma}^2(x)}\mathrm{d}x.$$

As $\lim_{x\downarrow 0}|s(x)| = \infty$ and $\lim_{x\downarrow 0}\rho(x)\overline{\sigma}^2(x) = \infty$, one needs to study the behaviour of the integrand around zero to conclude. As $\int_0^a \rho(x)\mathrm{d}x = \infty$, for $x \in (0,a]$, $s(x) := -\int_x^a \rho(y)\mathrm{d}y$. For $\delta > 0$ such that $x < \delta < a$ and $x > 0$, $\int_x^a \rho(y)\mathrm{d}y = \int_x^\delta \rho(y)\mathrm{d}y + \int_\delta^a \rho(y)\mathrm{d}y$. Note that the second integral is convergent as the integral of a continuous function over a closed interval of $\mathbb{R}$.

Classical asymptotic expansions for integrals [19, Chapter 3.3] and (20), yield, after mapping $y \mapsto zx$,

$$\int_x^\delta \rho(y)\mathrm{d}y = \mathrm{e}^{\widetilde{K}\widetilde{K}_\beta^a} x^{1-\frac{2A_\beta \widetilde{K}}{y_\sigma}} \int_1^{\delta/x} exp\left(\frac{\widetilde{K}}{xz}\right) z^{-\frac{2A_\beta \widetilde{K}}{y_\sigma}}\mathrm{d}z = \mathrm{e}^{\widetilde{K}\widetilde{K}_\beta^a} x^{1-\frac{2A_\beta \widetilde{K}}{y_\sigma}} \mathrm{e}^{\frac{\widetilde{K}}{x}} \left(\frac{x}{\widetilde{K}} + o(x)\right),$$

and the asymptotic behaviour of the integrand around the origin is given by

$$\frac{|s(x)|}{\rho(x)\overline{\sigma}^2(x)} = \frac{x_\sigma^{2\beta}}{\beta^2\gamma^2}\left(\frac{x^2}{\widetilde{K}} + o(x^2)\right)\frac{1+o(1)}{x^2} = \frac{x_\sigma^{2\beta}}{\beta^2\gamma^2\widetilde{K}}(1+o(1)),$$

which is integrable at the origin, and concludes the proof. □

A.4. **Proof of Proposition 3.10.** As the process Y^ε is in $\mathbb{R}_+^*$ instead of $\mathbb{R}$, we adapt the proof of [22, Theorem 2.9] to show that it satisfies a large deviations principle with speed ε and rate function I^Y. Note first that for $y_\sigma > 0$, and in both cases $y_0 \geq y_\sigma$ and $y_0 < y_\sigma$, the function $\widetilde{\sigma}$ is locally Lipschitz continuous on $\mathbb{R}_+^*$. Furthermore, for $f \in \overline{\mathcal{H}}$, the Picard-Lindelöf Theorem implies that the controlled ODE $\dot{g}_t = \widetilde{\sigma}(g_t)\dot{f}_t$, with $g_0 = x$ admits the solution

$$\mathcal{S}^x(f)(t) = \left(\frac{\beta\gamma}{\alpha}\right)^{\frac{1}{\beta}}\left[\mathrm{e}^{-\alpha\int_0^t \dot{f}_u \mathrm{d}u}\left(x^\beta\frac{\alpha}{\beta\gamma} - 1\right) + 1\right]^{1/\beta}, \qquad \text{for } t \in \mathcal{T}, x > 0.$$

This formulation requires the term $\left[\mathrm{e}^{-\alpha\int_0^t \dot{f}_u \mathrm{d}u}\left(x^\beta\frac{\alpha}{\beta\gamma} - 1\right) + 1\right]$ to be positive for all $x > 0$:

- If $y_0 \geq y_\sigma$, then $y_0^\beta\frac{\alpha}{\beta\gamma} - 1 \geq 0$ and $\mathcal{S}^{y_0}(f)$ is positive on $\mathcal{T}$;
- If $y_0 < y_\sigma$, then $y_0^\beta\frac{\alpha}{\beta\gamma} - 1 < 0$ and $\mathcal{S}^{y_0}(f)$ is positive on $\mathcal{T}$ if and only if Condition (7) holds.

The crucial step in [22, Theorem 2.9] is [22, Theorem 2.7], which states that if $\sqrt{\varepsilon}W$ is close to $f \in \overline{\mathcal{H}}$, then Y^ε should be close to $\mathcal{S}^{y_0}(f)$, the solution of the controlled ODE. The case of bounded and globally Lipschitz coefficients follows directly from [22, Theorem 2.7]. In order to deal with locally Lipschitz coefficients here, we need to localise. For $0 < r_\sigma < r_b$, the functions

$$\overline{b}(x) := \begin{cases} b(x), & \|x\| \leq r_b, \\ b\left(\frac{rx}{\|x\|}\right), & \|x\| > r_b, \end{cases} \qquad \text{and} \qquad \overline{\sigma}(x) := \begin{cases} \widetilde{\sigma}(x), & \|x\| > r_\sigma, \\ \widetilde{\sigma}\left(\frac{rx}{\|x\|}\right), & \|x\| \leq r_\sigma. \end{cases}$$

are bounded and globally Lipschitz continuous on $\mathbb{R}_+^*$, and clearly $\varepsilon\overline{b}(\cdot)$ converges uniformly to zero on $\mathbb{R}_+^*$ as ε goes to zero. Moreover, given $\eta \in (r_b, r_\sigma)$, there exists $0 < r \leq \min\{\eta - r_\sigma, \eta - r_b\}$ such that the δ-tube around $\mathcal{S}^{y_0}(f)$ is contained in $B_r(\eta)$. In order for this radius r to exist, one simply needs to make sure that the solution $\mathcal{S}^{y_0}(f)$ of the controlled ODE never reaches zero (explosion is impossible as infinity is recurrent), which is obvious when $y_0 \geq y_\sigma$, and guaranteed by Condition (7) when $y_0 < y_\sigma$.

Denote $\overline{Y}^\varepsilon$ the solution to $\mathrm{d}\overline{Y}_t^\varepsilon = \varepsilon\overline{b}(\overline{Y}_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\overline{\sigma}(\overline{Y}_t^\varepsilon)\mathrm{d}W_t$ with $\overline{Y}_0^\varepsilon = y_0 > 0$. Then the two sequences $(\overline{Y}^\varepsilon)_{\varepsilon>0}$ and $(Y^\varepsilon)_{\varepsilon>0}$ are identical on $B_r(\eta)$. Thus, for each $\delta, \lambda > 0$, there exist $\xi, \zeta > 0$ such that, for all $f \in \overline{\mathcal{H}}$ with $\Lambda(f) \leq \lambda$ and $x \in B_\xi(y_0)$,

$$\mathbb{P}\left[\|Y^\varepsilon - \mathcal{S}^{y_0}(f)\|_\infty > \delta, \|\sqrt{\varepsilon}W - f\|_\infty \leq \zeta\right] = \mathbb{P}\left[\|\overline{Y}^\varepsilon - \mathcal{S}^{y_0}(f)\|_\infty > \delta, \|\sqrt{\varepsilon}W - f\|_\infty \leq \zeta\right].$$

The constants $\xi, \zeta > 0$ are such that [22, Theorem 2.7] is satisfied. Hence, for each $R, \delta, \lambda > 0$, there exist $\zeta, \xi, \varepsilon_0 > 0$ such that, for all $f \in \overline{\mathcal{H}}$ with $\Lambda(f) \leq \lambda$, $x \in B_\xi(y_0)$, $\varepsilon \leq \varepsilon_0$,

$$\mathbb{P}\left[\|Y^\varepsilon - \mathcal{S}^{y_0}(f)\|_\infty > \delta, \|\sqrt{\varepsilon}W - f\|_\infty \leq \zeta\right] \leq \exp\left\{-\frac{R}{\varepsilon}\right\}$$

holds, and the proof follows from [22, Theorem 2.9].

A.5. **Proof of Theorem 3.9.** To obtain a large deviations principle for X^ε, a large deviations principle for the rescaled process $\mathrm{X}^\varepsilon := (X^\varepsilon, Y^\varepsilon)$ needs to be proved.

$$\mathrm{dX}_t^\varepsilon = \varepsilon\mathrm{b}(\mathrm{X}_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\mathrm{a}(\mathrm{X}_t^\varepsilon)\mathrm{d}W_t,$$

with initial condition $\mathrm{X}_0^\varepsilon := \mathrm{x}_0 = \begin{pmatrix} \log s_0 \\ y_0 \end{pmatrix}$ and the maps $\mathrm{b}, \mathrm{a} : \mathbb{R}_+^* \to \mathbb{R}^2$ defined as

$$\mathrm{b}(\mathrm{X}_t^\varepsilon) = \begin{pmatrix} -\frac{1}{2}\sigma^2(Y_t^\varepsilon) \\ b(Y_t^\varepsilon) \end{pmatrix} \qquad \text{and} \qquad \mathrm{a}(\mathrm{X}_t^\varepsilon) = \begin{pmatrix} \sigma(Y_t^\varepsilon) \\ \widetilde{\sigma}(Y_t^\varepsilon) \end{pmatrix}.$$

These two maps are both locally Lipschitz continuous on $\mathbb{R} \times \mathbb{R}_+^*$. Solving the controlled ODE for Y^ε is sufficient to solve the controlled ODE for the process X^ε. Using the proof of Proposition 3.10, for $\mathrm{f} := (f, f)$ with $f \in \overline{\mathcal{H}}$, the controlled ODE $\dot{\mathrm{g}}_t = \mathrm{a}(\mathrm{g}_t)\dot{\mathrm{f}}_t$, with $\mathrm{g}_0 = \mathrm{x}_0$ has a solution $\mathrm{g} = \mathcal{S}^{\mathrm{x}_0}(f)$ on $\mathcal{T}$. For $y_0 > y_\sigma$, the solution $\mathcal{S}^{y_0}$ is strictly positive and $\mathcal{S}^{\mathrm{x}_0}(f)$ exists on $\mathcal{T}$ for all $f \in \overline{\mathcal{H}}$ and $\mathrm{x}_0 \in \mathbb{R} \times \mathbb{R}_+^*$. In this case, $\overline{\mathcal{H}}$ boils down to the Cameron-Martin space. For $y_0 < y_\sigma$, Condition (7) ensures that $\mathcal{S}^{y_0}$ is positive. Applying [22, Theorem 2.9], the sequence X^ε then satisfies a large deviations principle on $\mathcal{C}(\mathcal{T}, \mathbb{R} \times \mathbb{R}_+^*)$ as ε tends to zero, with speed ε and rate function

$$\mathrm{I}^{Y,X}(\mathrm{g}) := \inf\left\{\Lambda(f), f \in \overline{\mathcal{H}}, \mathcal{S}^{\mathrm{x}_0}(f) = \mathrm{g}\right\}.$$

To obtain a large deviations principle for the log-stock price X^ε, we apply the Contraction Principle [4, Theorem 4.2.1] since the projection on the first component is continuous, and the theorem follows.

References

[1] R. Azencott, B. Geiger and W. Ott. Large deviations for Gaussian diffusions with delay. *Journal of Statistical Physics*, 170(2): 254-285, 2018.

[2] T. Bollerslev. Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3): 302-327, 1986.

[3] A.S. Cherny and H-J Engelbert. Singular stochastic differential equations. Springer, 2005.

[4] A. Dembo and O. Zeitouni. Large Deviations Techniques and Applications. Springer Verlag, 1998.

[5] J.C. Duan. The GARCH option pricing model. *Mathematical Finance*, 5(1):13-32, 1995.

[6] B. Dupire. Pricing with a smile. *Risk*, 7:18-20, 1994.

[7] R. Engle. Autoregressive conditional heteroskedasticity with estimates of variance of United Kingdom. *Econometrica*, 50: 987-1008, 1982.

[8] J.P. Fouque, G. Papanicolaou and R. Sircar. Derivatives in financial markets with stochastic volatility. *Cambridge University Press*, 2000.

[9] J.P. Fouque, G. Papanicolaou, R. Sircar and K. Sølna. Multiscale stochastic volatility for equity, interest rate, and credit derivatives. *Cambridge University Press*, 2011.

[10] P. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier and J. Teichmann. Large deviations and asymptotic methods in Finance. Springer Proceedings in Mathematics and Statistics, 110, 2015.

[11] K. Gao and R. Lee. Asymptotics of implied volatility to arbitrary order. *Finance and Stochastics*, 18(2): 349-392, 2014.

[12] J. Guyon. Path-dependent volatility. *Risk*, 2014.

[13] J. Guyon. Path-dependent volatility: practical examples. *Global Derivatives Conference*, 2017.

[14] J. Guyon and P. Henry-Labordère. Being particular about calibration. *Risk*, 2012.

[15] D.G. Hobson and L.C.G. Rogers. Complete models with stochastic volatility. *Mathematical Finance*, 8(1): 27-48, 1998.

[16] S. Karlin and H.M. Taylor. A second course in stochastic processes. Academic Press, 1981.

[17] R. Lord, R. Koekkoek and D. Van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2): 177-194, 2010.

[18] J. Ma, Z. Ren, N. Touzi and J. Zhang. Large deviations for non-Markovian diffusions and path-dependent Eikonal equation. *Annales de l'IHP: Probabilités et Statistiques*, 1196-1216, 2016.

[19] P.D. Miller. Applied asymptotic analysis. American Mathematical Society, 75, 2006.

[20] S. Mohammed and T. Zhang. Anticipating stochastic differential systems with memory. *Stochastic Processes and their Applications*, 119: 2773-2802, 2009.

[21] E. Pardoux and A.Y. Veretennikov. On Poisson equation and diffusion approximation, II. *The Annals of Probability*, 31(3): 1166-1192, 2003.

[22] D. Peithmann. Large deviations and exit time asymptotics for diffusions and stochastic resonance. *Ph.D. Thesis* Humboldt University, Berlin, 2007.

[23] L.C.G. Rogers and D. Williams. Diffusions, Markov Processes and Martingales. *Cambridge Mathematical Library*, 2014.

[24] Y.F. Saporito. First-order asymptotics of path-dependent derivatives in multiscale stochastic volatility environment. *International Journal of Theoretical and Applied Finance*, 21(3): 1-22, 2018.

Department of Mathematics, Imperial College London, and Alan Turing Institute
E-mail address: a.jacquier@imperial.ac.uk

Department of Mathematics, Imperial College London
E-mail address: chloe.lacombe14@imperial.ac.uk