Asymptotic expansions of evolution equations with fast volatility

Sam D. Howison* Christoph Reisinger* Ronnie Sircar†
Zhenru Wang*

October 2023; revised October 2024

Abstract

We study Kolmogorov forward equations (KFE) and Zakai equations for diffusion processes with a fast mean-reverting stochastic volatility component. In the case of the KFE, a parabolic PDE in divergence form, we perform a matched asymptotic expansion up to first order in the small mean-reversion time. The solutions are expressed in terms of suitable PDEs with coefficients averaged over the ergodic distribution, in the spirit of extensive earlier work on the backward equation (see *J.-P. Fouque et al, CUP, 2011*). We then construct a sequence of approximations to the Zakai equation, a parabolic stochastic PDE (SPDE), and verify numerically for the first two terms weak convergence order half and order one, respectively, in the mean-reversion parameter. To this end, we give a novel numerical scheme for the original two-dimensional SPDE, which is robust in the small parameter regime, and compare derived functionals of marginals against those approximated by the solution of a sequence of homogenised one-dimensional SPDEs.

Contents

*Mathematical Institute, University of Oxford, United Kingdom (howison@maths.ox.ac.uk, christoph.reisinger@maths.ox.ac.uk, zrwang1128@gmail.com)

†ORFE Department, Princeton University, Princeton, NJ, USA (sircar@princeton.edu)

1 Introduction

In this work, we consider processes of the form

$$\begin{aligned} \mathrm{d}X_t &= \mu(X_t, Y_t)\,\mathrm{d}t + \sigma(Y_t)\Big(\rho_x\,\mathrm{d}W_t^x + \sqrt{1-\rho_x^2}\,\mathrm{d}W_t^{x,1}\Big), \quad X_0 = x_0, \\ \mathrm{d}Y_t &= \frac{\kappa}{\epsilon}(m - Y_t)\,\mathrm{d}t + \frac{g(Y_t)}{\sqrt{\epsilon}}\Big(\rho_y\,\mathrm{d}W_t^y + \sqrt{1-\rho_y^2}\,\mathrm{d}W_t^{y,1}\Big), \quad Y_0 = y_0, \end{aligned} \tag{1}$$

where $(W^x, W^y, W^{x,1}, W^{y,1})$ is a four-dimensional standard Brownian motion, W^x and W^y have correlation $\rho \in (-1,1)$, while $W^{x,1}, W^{y,1}$ are independent of each other and of (W^x, W^y); $x_0, y_0, m \in \mathbb{R}$, $\rho_x, \rho_y \in (-1,1)$, $\epsilon, \kappa > 0$ are all constant; $\mu : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ and $\sigma, g : \mathbb{R} \to \mathbb{R}_+$ given functions. For ease of notation, we introduce $\rho_{xy} = \rho_x\rho_y\rho$.

We will study the marginal distribution of (X_t, Y_t) at t, and the distribution of (X_t, Y_t) conditional on the natural filtration $\mathcal{F}_t^{x,y}$ of $W = (W^x, W^y)$ at time t, which is the reason for writing the Brownian driver in the decomposed way above. Specifically, we are interested in the setting of small ϵ, a characteristic, dimensionless reversion time of the Y-process to its mean m, and will derive equations for asymptotic expansions of the probability density function (PDF) and the conditional PDF. The former leads us to derive matched asymptotic expansions of the corresponding Kolmogorov *forward* equation (KFE, or Fokker–Planck equation), a two-dimensional parabolic PDE in divergence form, while the latter leads to expansions of a Zakai-type equation, a parabolic stochastic PDE (SPDE).

Models of the form (1) are used abundantly in financial engineering, where X describes the log price of a financial asset and $\sigma(Y_t)$ is its instantaneous (stochastic) volatility at time t. The presence of multiple time scales in market data has been documented extensively in the literature; see, e.g., [11, 10], and especially the monograph [12] and the references therein. For higher order expansions with a refined boundary layer analysis close to expiry we refer to [19] and [8], also [5] for a convergence analysis. The expansion at the level of the underlying

stochastic processes is analysed in [13] (see also the earlier discussion in the conclusions of [19]). Among more recent works, [2] performs joint asymptotic expansions of optimal investment models with fast volatility and small transaction costs, and [7] demonstrates a multiscale analysis of portfolio optimisation strategies under fast and slow volatilities.

The above Kolmogorov *backward* equations (KBE) are in non-divergence form and typically have regular (i.e., continuous) terminal data, while the Kolmogorov *forward* equations (KFE) studied here are in divergence form with Dirac delta initial data. Specifically, we will consider the model where μ is constant, and $g = \nu\sqrt{2}$ for constant ν, that is where Y an Ornstein–Uhlenbeck process, with unique ergodic distribution $\mathcal{N}(m, \nu^2)$.[1] In this case, the KFE is

$$\begin{aligned} \partial_t p^\epsilon &= \frac{1}{\epsilon}\Big(\nu^2\partial_{yy}p^\epsilon - \kappa\partial_y((m-y)p^\epsilon)\Big) + \Big(\frac{1}{2}\sigma^2\big(y\big)\partial_{xx}p^\epsilon - \mu\,\partial_x p^\epsilon\Big) \\ &\quad + \frac{1}{\sqrt{\epsilon}}\rho_{xy}\nu\sqrt{2}\partial_y(g(y)\partial_x p^\epsilon), \\ p^\epsilon(0,x,y) &= \delta(x-x_0)\otimes\delta(y-y_0). \end{aligned} \tag{2}$$

Apart from being of interest in its own right, the analysis of the forward PDE serves as preparation for that of the Zakai SPDE

$$\begin{aligned} \mathrm{d}u^\epsilon &= \frac{1}{\epsilon}\Big(\nu^2\partial_{yy}u^\epsilon - \kappa\partial_y((m-y)u^\epsilon)\Big)\,\mathrm{d}t + \Big(\frac{1}{2}\sigma^2\big(y\big)\partial_{xx}u^\epsilon - \mu\,\partial_x u^\epsilon\Big)\,\mathrm{d}t \\ &\quad + \frac{1}{\sqrt{\epsilon}}\rho_{xy}\nu\sqrt{2}\partial_y(g(y)\partial_x u^\epsilon)\,\mathrm{d}t + \rho_x\sigma(y)\partial_x u^\epsilon\,\mathrm{d}W_t^x + \rho_y\frac{\sqrt{2}\nu}{\sqrt{\epsilon}}\partial_y u^\epsilon\,\mathrm{d}W_t^y, \\ u^\epsilon(0,x,y) &= \delta(x-x_0)\otimes\delta(y-y_0). \end{aligned} \tag{3}$$

There are at least two motivations for studying (3). First, by general filtering theory (see, e.g., [1, Section 3.5]), the solution u^ϵ is the density (if it exists) of the conditional law of (X_t, Y_t) given observation of (W^x, W^y) up to time t. Second, it is the limit empirical measure of a large number N of independent realisations of (1), with independent (idiosyncratic) noise terms $W_t^{x,i}, W_t^{y,i}$, for $i = 1, \ldots, N$, replacing $W_t^{x,1}, W_t^{y,1}$, but all with the same common noise W_t^x, W_t^y (see, e.g., [20]).

This limiting equation has been used to describe the behaviour of large pools of defaultable financial entities, where the process X is replaced by one absorbed at 0 (a 'default boundary'), and the absorption of mass is interpreted as a loss to the financial system. The case of constant σ is analysed in [4] and applications to credit derivative markets are given. For a (nonlinear) SPDE model for a large pool limit of a default intensity-based credit model see e.g. [14]. More recently, an extension of the basic model in [4] to stochastic volatilities is given in [16, 17]. See also [21, 22] for different applications involving filtering of hidden Markov models with fast mean-reverting states.

We will focus particularly on the regime of small ϵ, as motivated by the empirical evidence cited above. In practical applications, one is predominantly interested in the behaviour of X, and in Y only in as much as it affects the dynamics of X. For instance, in credit risk, it is the firm log value process X which directly affects loss distributions. It is therefore desirable to derive simplified homogenised equations which allow for more efficient analytical

[1] The factor $\sqrt{2}$ is chosen to ensure the ergodic distribution has a more standard normal form, consistent with the literature (see [9, 12]).

or numerical solutions by reducing the dimensionality of the PDE or SPDE. The numerical approximation of the original two-dimensional KFE or Zakai SPDE, especially, is more costly computationally than that of the one-dimensional analogue with deterministic (e.g., constant) volatility. This is exacerbated by the presence of multiple scales, which may require a fine time mesh and fine spatial mesh in the second dimension for the stable resolution of the fast component for small parameter ϵ.

By expansion in ϵ, we will approximate p^ϵ and u^ϵ by sequences of KFEs and Zakai SPDEs, respectively, which have the advantageous feature that the fast varying volatility is replaced by its ergodic average in the differential operator, and, in the case of the SPDE, the driving noise term. Moreover, if only functionals of X (and not Y) are required, these can be computed by the solution of one-dimensional (S)PDEs, leading to an effective dimension reduction and complexity advantage.

In the context of (1), [18] consider the case of $\mu(X_t, Y_t) = r - \sigma^2(Y_t)/2$, for constant r.[2] Under certain recurrence properties of the diffusion Y, and for $\rho_{xy} = 0$, it is shown that as $\epsilon \to 0$, the stopped version of X converges in distribution to a process X^* which satisfies

$$\mathrm{d}X_t^* = (r - \langle\sigma^2\rangle/2)\,\mathrm{d}t + \sqrt{\langle\sigma^2\rangle}\Big(\rho\,\mathrm{d}W_t^x + \sqrt{1-\rho^2}\,\mathrm{d}W_t^{x,1}\Big), \quad X_0 = x_0, \tag{4}$$

where $\langle\sigma^2\rangle$ is the expectation of $\sigma^2(\cdot)$ under the invariant distribution of Y.

Moreover, weak limits of u^ϵ are shown to satisfy the SPDE

$$\mathrm{d}u^* = \Big(\frac{1}{2}\langle\sigma^2\rangle\partial_{xx}u^* - (r - \langle\sigma^2\rangle/2)\,\partial_x u^*\Big)\,\mathrm{d}t + \rho_x\langle\sigma\rangle\partial_x u^*\,\mathrm{d}W_t^x, \tag{5}$$

where $\langle\sigma\rangle$ is the expectation of $\sigma(\cdot)$ under the invariant distribution of Y.

We note that [18] allow for more general g in (1) than for the Ornstein–Uhlenbeck (O–U) process considered here. On the other hand, we want to avoid the assumption $\rho_{xy} = 0$, as the correlation between volatility and stock is an important parameter influencing the dynamic behaviour of stock price models. In particular, it is used to match the implied volatility skew in derivatives markets, while a realistic dependence of increments of the two processes is key for successful hedging. To deal with this dependence, we make more specific assumptions on the volatility process, and restrict ourselves to O–U processes, which will allow a decomposition of Y such that the SPDE is driven purely by the slow component, while fast-mean-reverting term appears explicitly in the coefficients. This helps with the construction of correction terms for a higher order expansion of the SPDE. However the function σ is general up to technical restrictions.

The main contributions and outline of the present paper are the following:

- a matched asymptotic expansion solution of the KFE (2) for small ϵ, identifying the boundary layer for small t and deriving the expansion up to order 1 for general t (Section 2);
- the heuristic derivation of a one-dimensional SPDE for a first-order correction to (5) in ϵ (Section 3);

[2] In fact, [18] allow that r and other (constant) model parameters are sampled randomly.

- numerical verification of the SPDE expansion orders (in Section 5) by novel numerical schemes for (3) and related SPDEs, with a proof of unconditional stability independent of ϵ (Section 4).

2 Perturbation analysis of the KFE

2.1 Set-up and preliminaries

We derive an expansion for the transition density function of a stochastic volatility process (X, Y) satisfying (1) with $g = \nu\sqrt{2}$ for constant ν, as the dimensionless parameter $\epsilon \to 0^+$. We assume for simplicity that σ is bounded away from zero, such that the process X takes values on all of $\mathbb{R}$ (as does the O–U process Y).

The transition density $p^\epsilon(t_0, x_0, y_0; t, x, y)$ of (X, Y) satisfies the forward Kolmogorov equation (in the variables t, x, y)

$$\partial_t p^\epsilon - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* + \mathcal{L}_2^*\right) p^\epsilon = 0, \tag{6}$$

with the initial condition

$$p^\epsilon(t_0, x_0, y_0; t_0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0). \tag{7}$$

Here the operators $\mathcal{L}_i$ and their adjoints $\mathcal{L}_i^*$ are defined by

$$\mathcal{L}_0 \cdot = \nu^2 \partial_{yy} \cdot + \kappa(m - y)\partial_y \cdot, \qquad \mathcal{L}_0^* \cdot = \nu^2 \partial_{yy} \cdot - \kappa \partial_y \left((m - y) \cdot\right), \tag{8}$$

$$\mathcal{L}_1 \cdot = \rho_{xy}\nu\sqrt{2}\sigma(y)\partial_{xy} \cdot, \qquad \mathcal{L}_1^* \cdot = \rho_{xy}\nu\sqrt{2}\partial_x\partial_y \left(\sigma(y) \cdot\right), \tag{9}$$

$$\mathcal{L}_2 \cdot = \frac{1}{2}\sigma^2(y)\partial_{xx} \cdot + \mu(x, y)\partial_x \cdot, \qquad \mathcal{L}_2^* \cdot = \frac{1}{2}\sigma^2(y)\partial_{xx} \cdot - \partial_x \left(\mu(x, y) \cdot\right). \tag{10}$$

We set $t_0 = 0$ for simplicity. We denote by Φ^Y the probability density function (PDF) of the ergodic distribution of Y,

$$\Phi^Y(y) = \frac{1}{\sqrt{2\pi\nu^2/\kappa}} \mathrm{e}^{-\kappa(y-m)^2/2\nu^2}. \tag{11}$$

We shall frequently average functions of y with respect to the measure $\Phi^Y(y)\,\mathrm{d}y$ and we use the notation $\langle \cdot \rangle = \langle \cdot, \Phi^Y \rangle$ for this average, where we denote by $\langle \cdot, \cdot \rangle$ the usual inner product on $L^2(\mathbb{R})$.

In what follows we shall repeatedly seek solutions of equations of the form

$$-\mathcal{L}_0^* u(y) = h(y), \qquad -\infty < y < \infty, \tag{12}$$

for some given h, that are integrable on $\mathbb{R}$. We note immediately that

$$-\mathcal{L}_0^* \Phi^Y = 0, \qquad -\mathcal{L}_0^* \left(\Phi^Y(y) \int^y \frac{\mathrm{d}s}{\Phi^Y(s)}\right) = 0,$$

so that these functions span the null space of $\mathcal{L}_0^*$. However, integration by parts shows that, as $y \to \infty$,

$$\begin{aligned}\int^y \frac{\mathrm{d}s}{\Phi^Y(s)} &= \text{constant} \times \int^y \mathrm{e}^{\kappa(s-m)^2/2\nu^2}\,\mathrm{d}s \\ &\sim \text{constant} \times \frac{\mathrm{e}^{\kappa(y-m)^2/2\nu^2}}{y-m}\left(1+O(1/y^2)\right),\end{aligned}$$

and hence the function $\Phi^Y(y)\int^y \mathrm{d}s/\Phi^Y(s)$ is not integrable. We shall use this fact to eliminate this solution at several points below.

The null space of $-\mathcal{L}_0$ is spanned by 1 and $\mathrm{e}^{\kappa(y-m)^2/2\nu^2}$, the latter of these being irrelevant because of its growth at infinity (it is not even integrable against $\Phi^Y(y)$). It follows from the Fredholm Alternative that integrable solutions of (12) only exist when the right-hand side satisfies the solvability condition of being orthogonal to (relevant, ie bounded) solutions of the homogeneous adjoint equation. That is, from

$$\langle 1, h\rangle = \langle 1, -\mathcal{L}_0^*\, u\rangle = \langle -\mathcal{L}_0\, 1, u\rangle = \langle 0, u\rangle = 0,$$

the necessary condition for existence of a solution of (12) is $\langle 1, h\rangle = \int_{-\infty}^{\infty} h(y)\,\mathrm{d}y = 0$. When this is satisfied, the solution is given by

$$u(y) = -\Phi^Y(y)\int^y \frac{H(s)}{\Phi^Y(s)}\,\mathrm{d}s + c\Phi^Y(y),$$

where the constant c is arbitrary (the second solution of the homogeneous equation is ruled out as noted above), and where $H(\cdot) = \int^{\cdot} h(s)\,\mathrm{d}s$ is an antiderivative of h.

2.2 Outer region: $t \gg O(\epsilon)$

Turning to the evolution of the transition density function, over timescales much longer than the mean reversion time ϵ, the volatility is effectively sampled from its ergodic distribution, as is already seen from (4), and will determine the first term of the asymptotic expansion.

We expand

$$p^\epsilon(0, x_0, y_0; t, x, y) \sim p_0(t, x, y) + \sqrt{\epsilon}p_1(t, x, y) + \epsilon p_2(t, x, y) + \epsilon^{3/2}p_3(t, x, y) + \cdots,$$

where here and henceforth we suppress the dependence on x_0 and y_0 unless it is needed.

Before proceeding, we note that, as p^ϵ is a probability density, $\iint_{\mathbb{R}^2} p^\epsilon\,\mathrm{d}x\mathrm{d}y = 1$ for all t, and that similarly $\iint_{\mathbb{R}^2} p_0\,\mathrm{d}x\mathrm{d}y = 1$ for all t (because of the initial condition), whereas $\iint_{\mathbb{R}^2} p_i\,\mathrm{d}x\mathrm{d}y = 0$ for $i > 0$. Moreover, the marginal densities

$$p^\epsilon_{X_t}(t, x) = \int_{-\infty}^{\infty} p^\epsilon(t, x, y)\,\mathrm{d}y, \qquad p^\epsilon_{Y_t}(t, y) = \int_{-\infty}^{\infty} p^\epsilon(t, x, y)\,\mathrm{d}x$$

each have their own expansions

$$p^\epsilon_{X_t}(t, x) \sim p_{X_t,0} + \sqrt{\epsilon}p_{X_t,1} + \cdots, \qquad p^\epsilon_{Y_t}(t, y) \sim p_{Y_t,0} + \sqrt{\epsilon}p_{Y_t,1} + \cdots,$$

and the first term in each expansion integrates in x (resp. in y) to 1 while the remainder integrate to zero, because each such integral is the double integral of a term in the original expansion of p^ϵ. Note, however, that it is possible for any truncated version of any of the expansions to fail to be a probability density by virtue of being negative somewhere; this is in practice invariably the case in the far tails of the distributions where a separate (large-deviations/ray-theory) expansion would be needed to accurately capture the behaviour.

Now substituting into (6) and equating coefficients of powers of ϵ leads immediately to:

$$\text{At } O(1/\epsilon): \qquad -\mathcal{L}_0^* p_0 = 0; \tag{13}$$

$$\text{At } O(1/\sqrt{\epsilon}): \qquad -\mathcal{L}_0^* p_1 = \mathcal{L}_1^* p_0; \tag{14}$$

$$\text{At } O(1): \qquad -\mathcal{L}_0^* p_2 = -\partial_t p_0 + \mathcal{L}_2^* p_0 + \mathcal{L}_1^* p_1; \tag{15}$$

$$\text{At } O(\sqrt{\epsilon}): \qquad -\mathcal{L}_0^* p_3 = -\partial_t p_1 + \mathcal{L}_2^* p_1 + \mathcal{L}_1^* p_2; \tag{16}$$

the pattern in the last two of these repeats at still higher orders.

Lowest order $O(1/\epsilon)$ We have the leading order solution

$$p_0(t,x,y) = f_0(t,x)\Phi^Y(y) + g_0(t,x)\Phi^Y(y)\int_{-\infty}^{y} \frac{\mathrm{d}s}{\Phi^Y(s)},$$

where f_0 and g_0 are unknown at this stage; however, because p^ϵ is a probability density, and must be integrable in both x and y, we have $g_0(t,x) = 0$, because the function that it multiplies is not integrable. We shall see later that $\int_{-\infty}^{\infty} f_0(t,x)\,\mathrm{d}x = 1$ and all other p_i then integrate to zero over $\mathbb{R}^2$.

We note immediately that we cannot satisfy the initial condition (7); a separate boundary-layer analysis, given in Subsection 2.3, is needed to resolve this.

At $O(1/\sqrt{\epsilon})$ From (14), we have

$$-\mathcal{L}_0^* p_1 = \mathcal{L}_1^* p_0 = \rho_{xy}\nu\sqrt{2}\left(\partial_x f_0(t,x)\right)\partial_y\left(\sigma(y)\Phi^Y(y)\right). \tag{17}$$

As the y dependence on the right-hand side integrates to zero, this equation does have an integrable solution, and it is

$$p_1(t,x,y) = -\rho_{xy}\nu\sqrt{2}\left(\partial_x f_0(t,x)\right)\Sigma(y)\Phi^Y(y) + f_1(t,x)\Phi^Y(y), \tag{18}$$

where $\Sigma(y) = \int_{-\infty}^{y}\sigma(s)\,\mathrm{d}s$ (if σ is not integrable at $-\infty$, we simply integrate from (say) zero and amend $f_1(t,x)$ accordingly). Here $f_1(t,s)$ is again unknown; the other solution of the homogeneous equation has been eliminated as above.

At $O(1)$ Now we have

$$\begin{aligned}
-\mathcal{L}_0^* p_2 &= -\partial_t p_0 + \mathcal{L}_2^* p_0 + \mathcal{L}_1^* p_1 \\
&= -\Phi^Y(y)\left(\partial_t f_0(t,x) - \frac{1}{2}\sigma^2(y)\partial_{xx} f_0(t,x) + \partial_x\left(\mu(x,y) f_0(t,x)\right)\right) \\
&\quad - 2\rho_{xy}^2\nu^2\left(\partial_{xx} f_0(t,x)\right)\partial_y\left(\sigma(y)\Sigma(y)\Phi^Y(y)\right) + \rho_{xy}\nu\sqrt{2}\left(\partial_x f_1(t,x)\right)\partial_y\left(\sigma(y)\Phi^Y(y)\right).
\end{aligned}$$

The terms in the last line satisfy the solvability condition (as, indeed, does any function that is the result of applying $\mathcal{L}_1^*$ to a function of y which vanishes at $\pm\infty$) and so we need that the terms in the middle line integrate to zero. This leads directly to

$$\partial_t f_0 = \frac{1}{2}\langle\sigma^2(\cdot)\rangle\partial_{xx} f_0 - \partial_x\left(\langle\mu(x,\cdot)\rangle f_0\right) = \langle\mathcal{L}_2^*\rangle f_0(t,x), \tag{19}$$

where we are introducing the notation $\langle\mathcal{L}_i^*\rangle$ for operators with coefficients averaged over the ergodic distribution. Using this to eliminate $\partial_t f_0(t,x)$, we see that

$$-\mathcal{L}_0^* p_2 = \left(\mathcal{L}_2^* - \langle\mathcal{L}_2^*\rangle\right) p_0 + \mathcal{L}_1^* p_1,$$

the solution of which consists of a particular solution plus a solution $f_2(t,x)\phi^Y(y)$ of the inhomogeneous equation, $f_2(t,x)$ being as yet unknown.

As expected, $p_0(t,x,y)$ is the product of the ergodic density of Y and the density of X with the stochastic parameters replaced with their means, so X and Y behave independently at this order.

We need an initial condition for $f_0(t,x)$. Given the apparent independence of X and Y at this order, we suspect that $f_0(t_0,x) = \delta(x-x_0)$, and this is confirmed by the boundary-layer analysis of the next subsection.

At $O(\sqrt{\epsilon})$ Here we have

$$-\mathcal{L}_0^* p_3 = -\partial_t p_1 + \mathcal{L}_2^* p_1 + \mathcal{L}_1^* p_2.$$

As noted above, the final term on the right-hand side automatically satisfies the solvability condition and so we fix $f_1(t,x)$ by substituting for p_1 from (18), integrating over y, and then solving

$$\begin{aligned}\partial_t f_1 - \langle\mathcal{L}_2^*\rangle f_1 &= \int_{-\infty}^{\infty}\left[\mathcal{L}_2^*\left(-\rho_{xy}\nu\sqrt{2}\left(\partial_x f_0(t,x)\right)\Sigma(y)\Phi^Y(y)\right)\right]\,\mathrm{d}y \\ &= -\rho_{xy}\nu\sqrt{2}\left(\frac{1}{2}\langle\sigma^2(\cdot)\Sigma(\cdot)\rangle\partial_{xx}\left(\partial_x f_0(t,x)\right) - \partial_x\left(\langle\mu(x,\cdot)\Sigma(\cdot)\rangle\partial_x f_0(t,x)\right)\right).\end{aligned} \tag{20}$$

As $\Phi^Y(y)$ comes out as a factor, we have an ergodic average as before. The initial condition for this problem is found via the boundary-layer analysis of the next subsection.

2.3 Boundary layer near $t = 0$

The analysis above fails when $t = O(\epsilon)$, because there is an initial layer in which Y_t transits to its ergodic distribution. The large-time limit of the boundary layer solution provides the initial conditions for the functions $f_i(t,x)$ above, via asymptotic matching.

To capture this behaviour, we rescale time via

$$t = \epsilon t'.$$

Then in the boundary layer the transition density, now denoted $p'(0,x_0,y_0;t',x,y)$, satisfies

$$\frac{1}{\epsilon}\partial_{t'}p' - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* + \mathcal{L}_2^*\right)p' = 0, \tag{21}$$

with the initial condition

$$p'(0, x_0, y_0; 0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0). \tag{22}$$

We proceed exactly as above, expanding

$$p'(0, x_0, y_0; t', x, y) \sim p'_0(t', x, y) + \sqrt{\epsilon} p'_1(t', x, y) + \epsilon p'_2(t', x, y) + \epsilon^{3/2} p'_3(t', x, y) + \cdots,$$

and again suppressing the dependence on x_0 and y_0 unless it is needed. Substituting into (21) and equating coefficients of powers of ϵ leads immediately to:

$$\text{At } O(1/\epsilon): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p'_0 = 0; \tag{23}$$

$$\text{At } O(1/\sqrt{\epsilon}): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p'_1 = \mathcal{L}_1^*\, p'_0; \tag{24}$$

$$\text{At } O(1): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p'_2 = \mathcal{L}_1^*\, p'_1 + \mathcal{L}_2^*\, p'_0; \tag{25}$$

$$\text{At } O(\sqrt{\epsilon}): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p'_3 = \mathcal{L}_1^*\, p'_2 + \mathcal{L}_2^*\, p'_1; \tag{26}$$

the pattern in the last two of these repeats at higher orders. The initial conditions for the functions p'_i are

$$p'_0(0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0), \qquad p'_i(0, x, y) = 0, \quad i = 1, 2, 3 \ldots.$$

At $O(1/\epsilon)$ We have the degenerate (because lacking in x-derivatives) parabolic equation

$$(\partial_{t'} - \mathcal{L}_0^*)\, p'_0 = 0, \qquad p'_0(0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0).$$

Now bearing in mind that for a Brownian motion W we have that $(1/\sqrt{\epsilon})W_t$ becomes a Brownian motion $W_{t'}$ under the time-change $t = \epsilon t'$, we have in law

$$\mathrm{d}Y_{t'} = \kappa\,(m - Y_{t'})\,\mathrm{d}t' + \nu\sqrt{2}\,\mathrm{d}W_{t'},$$

the marginal density of this O–U process at time t' is Normal with mean and variance

$$m(t'; y_0) = m + (y_0 - m)\mathrm{e}^{-\kappa t'}, \qquad \mathrm{var}(t') = \frac{\nu^2}{\kappa}\left(1 - \mathrm{e}^{-2\kappa t'}\right),$$

respectively. It follows that

$$p'_0(t', x, y) = \delta(x - x_0)\phi^Y(t', y),$$

where

$$\phi^Y(t', y) = \frac{1}{\sqrt{2\pi \mathrm{var}(t')}} \mathrm{e}^{-(x - m(t'; y_0))^2 / 2\mathrm{var}(t')}. \tag{27}$$

When necessary, this is interpreted in the sense of distributions. Note immediately that

$$\lim_{t' \to \infty} \phi^Y(t', y) = \Phi^Y(y)$$

as defined above. Hence the limit of $p_1(t', x, y)$ as $t' \to \infty$ is $\delta(x - x_0)\Phi(y)$ and (by asymptotic matching) this is the initial condition for $f_0(t, x)$,

$$f_0(t, x) = \delta(x - x_0). \tag{28}$$

The interpretation of this result is that, at leading order, X_t stays at its initial value x_0 while Y_t forgets its initial value and transits to its ergodic distribution. In fact, there is a small amount of diffusion of X_t, which is resolved by introducing a further (spatial) inner layer of size $O(\sqrt{\epsilon})$ around x_0. With $x = x_0 + \sqrt{\epsilon}\xi$, $p'(t', x, y) = (1/\sqrt{\epsilon})P'(t', \xi, y)$, at leading order we have

$$\partial_{t'}P' - \left(\frac{1}{2}\sigma^2(y)\partial_{\xi\xi} + \rho_{xy}\nu\sqrt{2}\sigma(y)\partial_{\xi y} + \nu^2\partial_{yy}\right)P' + \kappa\partial_y\left((m-y)P'\right) = 0, \tag{29}$$

$$P'(0, \xi, y) = \delta(\xi)\delta(y - y_0) \tag{30}$$

(note the appearance of the correlation term, brought in by a combination of its original coefficient of $1/\sqrt{\epsilon}$ and a further $1/\sqrt{\epsilon}$ from the change of variable to ξ). The solution of this equation (not, as far as we know, available in closed form) represents the slow (on the t' timescale) spreading out of the initial point mass of the marginal density of X, while Y transits to its ergodic distribution. We do not pursue this further.

At $O(1/\sqrt{\epsilon})$ The equation (24) for p_1' now becomes

$$\left(\partial_{t'} - \mathcal{L}_0^*\right)p_1' = \rho_{xy}\nu\sqrt{2}\delta'(x - x_0)\partial_y\left(\sigma(y)\phi^Y(t', y)\right). \tag{31}$$

As $t' \to \infty$, $\phi^Y(t', y) \to \Phi^Y(y)$ and hence the solution of this equation has the limiting time-independent form

$$-\rho_{xy}\nu\sqrt{2}\delta'(x - x_0)\Sigma(y)\Phi^Y(y) + c_1(x)\Phi^Y(y) \tag{32}$$

for some function $c_1(x)$ which provides the initial condition for $f_1(t, x)$, while the first term in (32) matches automatically with the corresponding part of the solution $p_1(t, x, y)$ as $t \downarrow 0$. Fortunately, we can find $c_1(x)$ without having to solve for p_1' (which we could do, using the Green's function which is, in effect, $\phi^Y(t', y)$). Integrating (31) over y, and noting that both $\mathcal{L}_0^*$ and the right-hand side integrate to zero, we find

$$\frac{\mathrm{d}}{\mathrm{d}t'}\int_{-\infty}^{\infty} p_1'(t', x, y)\,\mathrm{d}y = 0$$

(this is essentially the orthogonality we used in the outer region) and hence the integral is equal to its initial value, namely zero. It follows that (32) must also integrate to zero. Bearing in mind that $\Phi^Y(y)$ is a probability density and so integrates to 1, we have

$$\begin{aligned} c_1(x) &= -\rho_{xy}\nu\sqrt{2}\delta'(x - x_0)\int_{-\infty}^{\infty}\Sigma(y)\Phi^Y(y)\,\mathrm{d}y \\ &= -\rho_{xy}\nu\sqrt{2}\delta'(x - x_0)\langle\Sigma\rangle \\ &= f_1(0, x). \end{aligned} \tag{33}$$

This is the initial condition for (20).

In summary, we have that $p_0(t, x, y) = f_0(t, x)\Phi^Y(y)$, where Φ^Y is given by (11) and f_0 solves (19) with initial datum (28); for μ which is constant in x, f_0 is simply a normal density. At $O(\sqrt{\epsilon})$, p_1 is given by (18) with f_1 satisfying (20) with initial condition (33). Note that $p_1(0, x, y) \neq 0$, confirming the need for the inner region.

2.4 A global approximation and correction equation

As we know the density of the O–U process Y to be $\Phi^Y(t/\epsilon, y)$, from (27), we can define an approximation globally in time as

$$p_{0,\epsilon}(t,x,y) = \Phi^Y(t/\epsilon, y) f_0(t,x), \tag{34}$$

which has the correct initial datum $p_0^\epsilon(0,x,y) = \delta(x-x_0) \otimes \delta(y-y_0)$, the exact marginal density for Y_t, and is correct to leading order in ϵ in both the inner and outer layer. By insertion, we see directly that $p_{0,\epsilon}$ from (34) satisfies

$$\partial_t p_{0,\epsilon} - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \langle \mathcal{L}_2^* \rangle \right) p_{0,\epsilon} = 0. \tag{35}$$

Taking the difference with (6), we have

$$\partial_t (p^\epsilon - p_{0,\epsilon}) - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* + \mathcal{L}_2^* \right)(p^\epsilon - p_{0,\epsilon}) = \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* p_{0,\epsilon} + (\langle \mathcal{L}_2^* \rangle - \mathcal{L}_2^*) p_{0,\epsilon}. \tag{36}$$

Replacing the operator on the left-hand side in (36) by that in (35), we define a correction to p^ϵ by $p_{0,\epsilon} + p_{1,\epsilon}$, where

$$\partial_t p_{1,\epsilon} - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \langle \mathcal{L}_2^* \rangle \right) p_{1,\epsilon} = \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* p_{0,\epsilon} + (\langle \mathcal{L}_2^* \rangle - \mathcal{L}_2^*) p_{0,\epsilon}, \qquad p_{1,\epsilon}(0,x,y) = 0. \tag{37}$$

This approach will be useful for the SPDE in the next section.

3 Perturbation analysis of the Zakai SPDE

3.1 Set-up and preliminaries

The matched asymptotic expansion analysis for the KFE (2) shows the different ansatz needed for small times (the 'inner layer' in subsection 2.3), where the process Y transits from its initial value to the stationary distribution, and for all times after this initial transit (the 'outer layer' in 2.2). These two expressions can be reconciled by using the analytical exact form for the marginal law of Y and an expansion for the effect of Y on X only (see subsection 2.4).

For the SPDE (3), the presence of a fast driving process correlated to a slow driving process creates extra difficulties for a direct asymptotic expansion. In a formal expansion of the SPDE similar to that in Section 2 for the KFE, the appearance of $O(1/\sqrt{\epsilon})$ terms multiplying both W_t^y and the $\partial_x \partial_y$ terms makes it difficult to transfer the multiple scales expansion from the KFE to the Zakai SPDE. We note that [18] has to restrict the derivation of the limiting SPDE for $\epsilon \to 0$ (determining the zero order term) to the case $\rho = 0$.

To avoid this last issue and take advantage of a global approximation, we split Y into a component U which has the correct instantaneous correlation with X and a component $Y^\dagger$ which is independent of X. We then study the joint dynamics of X and $Y^\dagger$ conditional on

(W^x, W^y), keeping track of the dynamics of $(U, Y^\dagger)$ and its law exactly, while we approximate the generator of X by an expansion.

Specifically, we introduce a process U as the (strong) solution to

$$\mathrm{d}U_t = -\frac{\kappa}{\epsilon} U_t \,\mathrm{d}t + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} \rho_y \,\mathrm{d}W_t^y, \qquad U_0 = 0. \tag{38}$$

Then $Y^\dagger := Y - U - m$ satisfies

$$\mathrm{d}Y_t^\dagger = -\frac{\kappa}{\epsilon} Y_t^\dagger \,\mathrm{d}t + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} \sqrt{1-\rho_y^2} \,\mathrm{d}W_t^{y,1}, \qquad Y_0^\dagger = y_0 - m,$$

where $W^{y,1}$ is independent of W^x and $W^{x,1}$, and W^x, W^y have correlation ρ. We set in the following $m = 0$ without loss of generality, as it simply results in a constant shift of the OU process; this can be accounted for in X by a horizontal translation of the function σ.

We will therefore study the two-dimensional Zakai SPDE for $(X, Y^\dagger)$, describing the marginal probability distribution of $(X_t, Y_t^\dagger)$ conditional on the natural filtration $\mathcal{F}_t^{x,y}$ of $W = (W^x, W^y)$,

$$\begin{aligned} \mathrm{d}v &= \frac{1}{\epsilon}\Big(\nu^2(1-\rho_y^2)\partial_{yy} v + \kappa \partial_y (yv)\Big)\,\mathrm{d}t \\ &\qquad + \Big(\frac{1}{2}\sigma^2\big(y + U_t\big)\partial_{xx} v - \mu\,\partial_x v\Big)\,\mathrm{d}t - \rho_x \sigma\big(y+U_t\big)\partial_x v \,\mathrm{d}W_t^x \\ &= \Big(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \widetilde{\mathcal{L}}_2^*\Big) v \,\mathrm{d}t - \rho_x \sigma(y+U_t)\partial_x v \,\mathrm{d}W_t^x, \\ v(0,x,y) &= \delta(x-x_0) \otimes \delta(y-y_0), \end{aligned} \tag{39}$$

where, in analogy to earlier,

$$\widetilde{\mathcal{L}}_0^* \cdot \;= \nu^2(1-\rho_y^2)\partial_{yy}\;\cdot\; - \kappa \partial_y (y\,\cdot), \qquad \widetilde{\mathcal{L}}_2^* \cdot \;= \frac{1}{2}\sigma^2\big(y+U_t\big)\partial_{xx}\;\cdot\; - \mu\,\partial_x \cdot,$$

and where $\nu > 0,\ \kappa > 0,\ \rho_x \in (-1,1),\ 0 < \epsilon \ll 1$ are fixed constants, $\sigma : \mathbb{R} \to \mathbb{R}_+$ is a real-valued function of which we will specify later any conditions needed. For simplicity, we consider $\mu \in \mathbb{R}$ a given constant.

We aim to find an expansion of the solution v to the SPDE (39) as $\epsilon \to 0$.

3.2 Zero order term

Following [18], and in line with the findings of Section 2, we introduce the following SPDE akin (5), by averaging the coefficients over the stationary distribution,

$$\begin{aligned} \mathrm{d}v_0^x &= \Big(\frac{1}{2}\langle\sigma^2\rangle \partial_{xx} v_0^x - \mu\,\partial_x v_0^x\Big)\mathrm{d}t - \rho_x \langle\sigma\rangle \partial_x v_0^x \,\mathrm{d}W_t^x \\ &= \langle \mathcal{L}_2^* \rangle v_0^x \,\mathrm{d}t - \rho_x \langle\sigma\rangle \partial_x v_0^x \,\mathrm{d}W_t^X, \\ v_0^x(0,x) &= \delta(x - x_0), \end{aligned} \tag{40}$$

where $\langle\sigma\rangle$ and $\langle\sigma^2\rangle$ are again the averages over the ergodic distribution and the operator $\langle\widetilde{\mathcal{L}}_2^*\rangle$ is defined as

$$\langle\widetilde{\mathcal{L}}_2^*\rangle\,\cdot\;:=\frac{1}{2}\langle\sigma^2\rangle\partial_{xx}\,\cdot\,-\mu\,\partial_x\,\cdot\,. \tag{41}$$

The SPDE (40) has the analytic solution

$$\begin{aligned} v_0^x(t,x) &= \Psi(t,x) && (42)\\ &:= f_0(t,x-\rho_x\langle\sigma\rangle W_t^x) && (43)\\ &= \frac{1}{\sqrt{2\pi\big(\langle\sigma^2\rangle-\rho_x^2\langle\sigma\rangle^2\big)t}}\exp\bigg(-\frac{(x-x_0-\mu t-\rho_x\langle\sigma\rangle W_t^x)^2}{2\big(\langle\sigma^2\rangle-\rho_x^2\langle\sigma\rangle^2\big)t}\bigg), \end{aligned}$$

where f_0 has been introduced earlier as solution to (19).

Now we include the initial transient of the processes $Y^\dagger$ and U to their stationary distribution. The marginal density of the O–U process $Y^\dagger$ at time t is known to be $\Phi^\dagger(t/\epsilon,y)$ with

$$\Phi^\dagger(t',y)=\frac{1}{\sqrt{2\pi\sigma_\dagger^2(t')}}\exp\bigg(-\frac{\big(y-\mu_\dagger(t')\big)^2}{2\sigma_\dagger^2(t')}\bigg), \tag{44}$$

where $\mu_\dagger(t')$ and $\sigma_\dagger^2(t')$ have the form

$$\mu_\dagger(t')=y_0\mathrm{e}^{-\kappa t'/\epsilon},\quad \sigma_\dagger^2(t')=\frac{(1-\rho_y^2)\nu^2}{\kappa}\left(1-\mathrm{e}^{-2\kappa t'/\epsilon}\right), \tag{45}$$

and it satisfies the PDE

$$\begin{aligned} \partial_t\Phi=\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^*\Phi&=\frac{1}{\epsilon}\left(\nu^2(1-\rho_y^2)\partial_{yy}\Phi+\kappa\partial_y\left(y\Phi\right)\right),\\ \Phi(0,y)&=\delta(y-y_0). \end{aligned}$$

Next, we follow the principle in Subsection 2.4 to define an approximation $v_{0,\epsilon}$ for which we track U and keep the marginal density of $Y^\dagger$ exact, but approximate the density of X given $(Y^\dagger,U)$, and hence the joint density. Therefore, we consider the equation

$$\begin{aligned} \mathrm{d}v_{0,\epsilon}&=\frac{1}{\epsilon}\left(\nu^2(1-\rho_y^2)\partial_{yy}v_{0,\epsilon}+\kappa\partial_y\left(yv_{0,\epsilon}\right)\right)\mathrm{d}t\\ &\qquad+\left(\frac{1}{2}\langle\sigma^2\rangle\partial_{xx}v_{0,\epsilon}-\partial_x(\mu v_{0,\epsilon})\right)\mathrm{d}t-\rho_x\langle\sigma\rangle\partial_x v_{0,\epsilon}\,\mathrm{d}W_t^x \qquad (46)\\ &=\left(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^*+\langle\widetilde{\mathcal{L}}_2^*\rangle\right)v_{0,\epsilon}\,\mathrm{d}t-\rho_x\langle\sigma\rangle\partial_x v_{0,\epsilon}\,\mathrm{d}W_t^x. \end{aligned}$$

Moreover, equation (46) has the closed-form solution

$$v_{0,\epsilon}(t,x,y)=\Psi(t,x)\Phi^\dagger(t/\epsilon,y), \tag{47}$$

where $\Psi(t,x)$ is defined in (42), and $\Phi^\dagger(t,y)$ is defined in (44). Note that $v_{0,\epsilon}$ satisfies the correct initial condition

$$v_{0,\epsilon}(0,x,y)=\delta(x-x_0)\otimes\delta(y-y_0).$$

We will later show numerically that $v_{0,\epsilon}-v\to 0$ in a weak sense as $\epsilon\to 0$.

3.3 Correction terms

The goal of this section is to construct successively better approximations. We first derive an inhomogeneous SPDE for $v - v_{0,\epsilon}$. Taking the difference between (39) and (46), we obtain

$$\begin{aligned} \mathrm{d}(v - v_{0,\epsilon}) &= \Big(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \widetilde{\mathcal{L}}_2^*\Big)(v - v_{0,\epsilon})\,\mathrm{d}t - \rho_x\sigma(y + U_t)\partial_x(v - v_{0,\epsilon})\,\mathrm{d}W_t^x \\ &\quad - \frac{1}{2}\Big(\langle\sigma^2\rangle - \sigma^2(y + U_t)\Big)\partial_{xx}v_{0,\epsilon}\,\mathrm{d}t + \rho_x\big(\langle\sigma\rangle - \sigma(y + U_t)\big)\partial_x v_{0,\epsilon}\,\mathrm{d}W_t^x, \qquad (48) \\ (v - v_{0,\epsilon})(0, x, y) &= 0. \end{aligned}$$

Intuitively, the effect of the terms in the second line in (48) is expected to be small for small ϵ, as the fast process U averages the terms involving $\sigma(\cdot + U_t)$ and $\sigma^2(\cdot + U_t)$ over the stationary distribution of U, while the presence of the dominant term $\widetilde{\mathcal{L}}_0^*$ on the left-hand side effects an additional averaging over the stationary distribution of $Y^\dagger$; the combined effect is an averaging over the stationary distribution of Y over timescales of order 1, so that the right-hand side, and hence the solution $v - v_{0,\epsilon}$, will be small.

Similar to before, we define $v_{1,\epsilon}$ as the leading order approximation to $v - v_{0,\epsilon}$,

$$\begin{aligned} \mathrm{d}v_{1,\epsilon} &= \Big(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \langle\widetilde{\mathcal{L}}_2^*\rangle\Big)v_{1,\epsilon}\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_{1,\epsilon}\,\mathrm{d}W_t^x \\ &\quad - \frac{1}{2}\Big(\langle\sigma^2\rangle - \sigma^2(y + U_t)\Big)\partial_{xx}v_{0,\epsilon}\,\mathrm{d}t + \rho_x\big(\langle\sigma\rangle - \sigma(y + U_t)\big)\partial_x v_{0,\epsilon}\,\mathrm{d}W_t^x, \qquad (49) \\ v_{1,\epsilon}(0, x, y) &= 0. \end{aligned}$$

In this definition, coming from (48), which describes the exact error, we have approximated $\mathcal{L}_2^*$ by $\langle\widetilde{\mathcal{L}}_2^*\rangle$ and σ by $\langle\sigma\rangle$ in the first line.

Given $v_{0,\epsilon}$ in (46), and $v_{1,\epsilon}$ in (49), we can recursively find higher order corrections as follows:

$$\begin{aligned} \mathrm{d}v_{n+1,\epsilon} &= \Big(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \langle\widetilde{\mathcal{L}}_2^*\rangle\Big)v_{n+1,\epsilon}\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_{n+1,\epsilon}\,\mathrm{d}W_t^x \\ &\quad - \frac{1}{2}\Big(\langle\sigma^2\rangle - \sigma^2(y + U_t)\Big)\partial_{xx}v_{n,\epsilon}\,\mathrm{d}t + \rho_x\big(\langle\sigma\rangle - \sigma(y + U_t)\big)\partial_x v_{n,\epsilon}\,\mathrm{d}W_t^x, \qquad (50) \\ v_{n+1,\epsilon}(0, x, y) &= 0. \end{aligned}$$

3.4 Approximation to the marginal density of X

One would hope that the expansion in subsection 3.3 allows a computationally more efficient approximation to the solution than solving the original two-dimensional SPDE directly numerically.

This is not clear when considering (49) directly, as the solution still depends on x and y. If we were to drop $\widetilde{\mathcal{L}}_0^*$ in (49), justified after the initial transient, the SPDE is parametrised by y, so essentially two-dimensional if the solution for all x and y is needed.

However, in practical applications (see, e.g., Section 5), one is often only interested in the marginal law of X, characterised by the marginal density

$$v^x(t, x) = \int_{-\infty}^{\infty} v(t, x, y)\,\mathrm{d}y.$$

Integrating (46) over y gives the zero order approximation

$$v_0^x = \int_{-\infty}^{\infty} v_{0,\epsilon}(t,x,y)\,\mathrm{d}y = \Psi(t,x),$$

where $\Psi(t,x)$ is given by (42).

As for the first order term $v_{1,\epsilon}$, we similarly define v_1^x as

$$v_1^x = \int_{-\infty}^{\infty} v_{1,\epsilon}(t,x,y)\,\mathrm{d}y.$$

If we assume $\lim_{y\to\pm\infty} v_{1,\epsilon}(t,x,y) = 0$, $\lim_{y\to\pm\infty} \partial_y v_{1,\epsilon}(t,x,y) = 0$, it follows that

$$\int_{-\infty}^{\infty} \widetilde{\mathcal{L}}_0^* v_{1,\epsilon}\,\mathrm{d}y = 0.$$

Moreover, from (47) and (44),

$$\begin{aligned}
\int_{-\infty}^{\infty} \sigma^2(y+U_t)\partial_{xx} v_{0,\epsilon}(t,x,y)\,\mathrm{d}y &= \int_{-\infty}^{\infty} \sigma^2(y+U_t)\partial_{xx}\Psi(t,x)\Phi^\dagger(t,y)\,\mathrm{d}y \\
&= \partial_{xx}\Psi(t,x)\int_{-\infty}^{\infty} \sigma^2(y+U_t)\Phi_\infty^\dagger(y)\,\mathrm{d}y + o(\epsilon^r) \\
&= \langle \sigma^2(\cdot+U_t)\rangle_\dagger \partial_{xx}\Psi(t,x) + o(\epsilon^r), \qquad \forall r>0, \\
\int_{-\infty}^{\infty} \sigma(y+U_t)\partial_x v_{0,\epsilon}(t,x,y)\,\mathrm{d}y &= \int_{-\infty}^{\infty} \sigma(y+U_t)\partial_x\Psi(t,x)\Phi_\infty^\dagger(y)\,\mathrm{d}y + o(\epsilon^r) \\
&= \langle \sigma(\cdot+U_t)\rangle_\dagger \partial_x\Psi(t,x) + o(\epsilon^r), \qquad \forall r>0,
\end{aligned} \tag{51}$$

for fixed $t>0$, and where $\Phi_\infty^\dagger = \lim_{t\to\infty}\Phi^\dagger(t/\epsilon,\cdot) = \lim_{\epsilon\to 0}\Phi^\dagger(t/\epsilon,\cdot)$, the stationary density of $Y^\dagger$, $\langle\cdot\rangle_\dagger$ the average over that distribution, and noting from (44), (45) that convergence of $\Phi^\dagger(t/\epsilon,\cdot)$ to $\Phi_\infty^\dagger$ is exponential in ϵ.

Integrating (49) over y yields an SPDE for a first order approximation $v_{1,\epsilon}^x(t,x)$ to $v^x(t,x)$,

$$\begin{aligned}
\mathrm{d}v_{1,\epsilon}^x &= \langle\widetilde{\mathcal{L}}_2^*\rangle v_{1,\epsilon}^x\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_{1,\epsilon}^x\,\mathrm{d}W_t^x \\
&\quad - \frac{1}{2}\left(\langle\sigma^2\rangle - \langle\sigma^2(\cdot+U_t)\rangle_\dagger\right)\partial_{xx}\Psi(t,x)\,\mathrm{d}t \\
&\quad + \rho_x\left(\langle\sigma\rangle - \langle\sigma(\cdot+U_t)\rangle_\dagger\right)\partial_x\Psi(t,x)\,\mathrm{d}W_t^x, \\
v_{1,\epsilon}^x(0,x) &= 0.
\end{aligned} \tag{52}$$

4 Numerical schemes for the Zakai SPDEs

In this section, we present numerical schemes for the SPDEs introduced in the previous section. We will use these in Section 5 to test the accuracy of the expansion solutions.

The most challenging equation among those considered is the original SPDE (39), repeated here for convenience,

$$\begin{aligned} \mathrm{d}v = \frac{1}{\epsilon}\Big(\nu^2(1-\rho_y^2)\partial_{yy}v + \kappa\partial_y(yv)\Big)\,\mathrm{d}t \\ + \Big(\frac{1}{2}\sigma^2\big(y+U_t\big)\partial_{xx}v - \mu\,\partial_x v\Big)\,\mathrm{d}t - \rho_x\sigma\big(y+U_t\big)\partial_x v\,\mathrm{d}W_t^x, \qquad (53) \\ v(0,x,y) = \delta(x-x_0)\otimes\delta(y-y_0), \end{aligned}$$

$$\mathrm{d}U_t = -\frac{\kappa}{\epsilon}U_t\,\mathrm{d}t + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}}\rho_y\,\mathrm{d}W_t^y, \qquad U_0 = 0. \qquad (54)$$

A principal difficulty in solving this two-dimensional SPDE arises from the fact that we seek numerical solutions which are stable and accurate uniformly across all ϵ.

We will also be solving the SPDE (48) to determine the error of the zero-order approximation, and the SPDE (49) for the first order correction, by straightforward modifications of the scheme for (39).

The zero-order marginal approximation in x, v_0^x, is given in analytic form by Ψ in (42), while the correction term $v_{1,\epsilon}^x$ is given by (52) and can be found by a one-dimensional SPDE scheme.

We apply a Milstein ADI scheme to the SPDE (53), and an Euler scheme to the SDE (54), taking care to maintain uniform stability and accuracy for small ϵ. We achieve this by a semi-implicit approximation of the Zakai SPDE and an approximation of the SDEs on a time mesh that scales with ϵ, as detailed below.

4.1 Simulation of the OU–process

We simulate the Ornstein–Uhlenbeck process (54) with timestep $k\epsilon$. It was found empirically in [6] that the simulated process using the Euler–Maruyama scheme then has a strong error independent of ϵ. This does not change the total computational effort significantly since the cost of simulating U is typically much smaller than solving the SPDE for v.

The discrete-time approximation of (U_t) is thus generated by

$$\begin{aligned} &\widehat{U}_n = \widehat{U}_{n-1} - \kappa\,k\,\widehat{U}_{n-1} + \nu\sqrt{2}\rho_y\Big(W_{t_n}^y - W_{t_{n-1}}^y\Big), \\ &\widehat{U}_0 = 0, \end{aligned} \qquad (55)$$

where $n = 1, 2, \ldots$, $t_n - t_{n-1} = k\epsilon$.

In practice, we first generate the bivariate standard normal random variables $(Z_{n,x}, Z_{n,y})$ with correlation ρ, where $n = 1, 2, \ldots, NN_\epsilon$, $N_\epsilon = 1/\epsilon$, and where we assume for simplicity $1/\epsilon \in \mathbb{Z}$. We generate

$$\widehat{U}_n = \widehat{U}_{n-1} - \frac{\kappa}{\epsilon}\,k\,\widehat{U}_{m-1} + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}}\rho_y\sqrt{k}\Big(Z_{n,y} - Z_{n-1,y}\Big),$$

where $\widehat{U}_0 = 0$ and $n = 1, 2, \ldots, NN_\epsilon$. Then we take

$$U^n = \widehat{U}_{nN_\epsilon}, \qquad n = 1, 2, \cdots, N, \qquad (56)$$

as the approximation to the process U_t at time $t = nk$, i.e., at the n-th time step on the coarser time mesh with width k, on which we will approximate the SPDE.

The Brownian increment of W^x over a "large" timestep k is thus

$$W^x_{nk} - W^x_{(n-1)k} = \sqrt{k}\widetilde{Z}_{n,x} \coloneqq \sum_{i=1}^{N_\epsilon} \sqrt{k\epsilon} Z_{N_\epsilon(n-1)+i,x},$$

which has the correct correlation ρ with W^y. Hence, we get

$$\widetilde{Z}_{n,x} = \sqrt{\epsilon} \sum_{i=1}^{N_\epsilon} Z_{N_\epsilon(n-1)+i,x}.$$

To simplify the notation, we write $Z_{n,x}$ instead of $\widetilde{Z}_{n,x}$ in the following.

4.2 Approximation of the one-dimensional SPDE

The scheme we use for the marginal SPDE for $v^x_{1,\epsilon}(t,x)$ from (52) is an adaptation of the schemes proposed in [23, 24]. We consider a mesh $\mathcal{X} = (x_i)_{-I \le i < I}$ for some integer I, $h_x > 0$ given, and define an approximation $V^n_{i,1,x}$ to $v^x_{1,\epsilon}(t_n, x_i)$ by

$$\begin{aligned}
&\left(I - \frac{\langle\sigma^2\rangle}{2}\frac{k}{h_x^2}D_{xx} + \mu\frac{k}{2h_x}D_x\right)V^{n+1}_{1,x} \\
&= \left(I - \rho_x\langle\sigma\rangle\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x + \rho_x^2\langle\sigma\rangle^2\frac{k(Z^2_{n,x}-1)}{2h_x^2}D_{xx}\right)V^n_{1,x} \\
&\quad - \frac{1}{2}\Big(\langle\sigma^2\rangle - \langle\sigma^2(\cdot + U^n)\rangle_\dagger\Big)k\Psi_{xx}(nk,\mathcal{X}) + \rho_x\Big(\langle\sigma\rangle - \langle\sigma(\cdot+U^n)\rangle_\dagger\Big)\Psi_x(nk,\mathcal{X})\sqrt{k}Z_{n,x} \\
&\qquad + \frac{1}{2}\rho_x^2\Big(\langle\sigma\rangle - \langle\sigma(\cdot+U_t)\rangle_\dagger\Big)^2\Psi_{xx}(nk,\mathcal{X})k(Z^2_{n,x}-1),
\end{aligned} \tag{57}$$

where the operators D_x and D_{xx} are defined as the standard finite difference matrices with

$$(D_xV)_i = V_{i+1} - V_{i-1}, \qquad (D_{xx}V)_i = V_{i+1} - 2V_i + V_{i-1},$$

U^n is found from (56), and where $\Psi_x(nk,\mathcal{X})$ and $\Psi_{xx}(nk,\mathcal{X})$ are the value of the functions applied on the mesh X, i.e., $\Psi_x(nk,\mathcal{X})$ is the vector of values $\Psi_x(nk,x_i)$, with Ψ from (42).

The scheme is semi-implicit to ensure stability in L_2 irrespective of the step sizes. The terms in the first line of (57) hence come from an implicit finite difference discretisation of the operator $\langle\mathcal{L}_2^*\rangle$; the second line contains an Euler–Maruyama term for the Brownian integral, and the Milstein correction for strong first order in k; the last two lines use the exact expressions of Ψ and its derivatives in the inhomogeneous terms and a Milstein approximation to the Brownian integral.

4.3 Approximation of the two-dimensional SPDEs

Original SPDE

We approximate the SPDE (39) with an alternating direction implicit (ADI) scheme of the operators $\widetilde{\mathcal{L}}_0^*$ and $\widetilde{\mathcal{L}}_2^*$, and a Milstein approximation of the Brownian integral.

We use a spatial mesh with uniform spacing h_x, $h_y > 0$, and, for $T > 0$ fixed, N time steps of size $k = T/N$. Let $V^n_{i,j}$ be the approximation to $v(nk, ih_x, jh_y)$, $n = 1, \dots, N$, $i, j \in \mathbb{Z}$.

Adapting the schemes from [25] to our setting,

$$\begin{aligned}&\left(I - \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy} - \frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\right)\left(I - \frac{\sigma^2(\mathcal{Y}+U^{n+1})}{2}\frac{k}{h_x^2}D_{xx} + \mu\frac{k}{2h_x}D_x\right)V^{n+1}\\ &= \left(I - \rho_x\sigma(\mathcal{Y}+U^n)\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x + \rho_x^2\sigma^2(\mathcal{Y}+U^n)\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\right)V^n,\end{aligned} \tag{58}$$

where U^n is from (56), and $D_.$, $D_{..}$ standard finite difference matrices defined by

$$\begin{aligned}(D_xV)_{i,j} &= V_{i+1,j} - V_{i-1,j}, & (D_yV)_{i,j} &= V_{i,j+1} - V_{i,j-1},\\ (D_{xx}V)_{i,j} &= V_{i+1,j} - 2V_{i,j} + V_{i-1,j}, & (D_{yy}V)_{i,j} &= V_{i,j+1} - 2V_{i,j} + V_{i,j-1},\\ (D_{xy}V)_{i,j} &= V_{i+1,j+1} - V_{i-1,j+1} - V_{i+1,j-1} + V_{i-1,j-1}.\end{aligned}$$

Moreover, $\mathcal{Y}$ is the diagonal matrix such that each element of the diagonal corresponds to a mesh point $(x_i, y_j) = (ih_x, jh_y)$, ordered the same way as V, so that for instance, by slight abuse of notation, $\sigma(\mathcal{Y} + U^n)$ is a diagonal matrix where the entry corresponding to point (ih_x, jh_y) is $(\sigma(\mathcal{Y} + U^n))_{i,j} = \sigma(y_j + U^n)$. V^0 is an approximation of the initial condition to the SPDE (39).

The implicit treatment of $\widetilde{\mathcal{L}}_2^*$ and particularly $\widetilde{\mathcal{L}}_0^*$ is important for stability for all mesh parameters and especially for all ϵ, as demonstrated by Proposition 4.1 below. The ADI factorisation allows an efficient solution of the implicit scheme by a sequence of tridiagonal systems.

Proposition 4.1. *Provided that*

$$|\rho_x| \le \frac{1}{\sqrt[4]{2}}\frac{\inf_{y\in\mathbb{R}}\sigma(y)}{\sup_{y\in\mathbb{R}}\sigma(y)}, \tag{59}$$

the scheme (58) *is stable in the* ℓ_2*-norm,* $|V|_2^2 := \sum_{i,j} V_{i,j}^2$. *Specifically, for all* $\epsilon > 0$, $h_x, h_y > 0$ *and* k, N *with* $kN = T$, $k\epsilon y_{\max}^2 \le \nu^2(1-\rho_y^2)$, *we have*

$$\mathbb{E}|V^n|_2^2 \le \exp\left(\frac{T\epsilon\, y_{\max}^2}{\nu^2(1-\rho_y^2)}\right)|V^0|_2^2. \tag{60}$$

Proof. We consider the discrete-continuous Fourier pair

$$V^n_{l,j} = \int_{-\pi}^{\pi} \widetilde{V}^n_j(\omega)\mathrm{e}^{\mathrm{i}\omega l}\,\mathrm{d}\omega, \qquad \widetilde{V}^n_j(\omega) = \frac{1}{2\pi}\sum_{l=-\infty}^{\infty} V^n_{l,j}\mathrm{e}^{-\mathrm{i}\omega l}.$$

By insertion and standard algebraic manipulations,

$$\left((I - kL_y)\widetilde{V}^{n+1}(\omega)\right)_j = \frac{\widetilde{L}^{\mathrm{ex},n}_{x,j}}{\widetilde{L}^{\mathrm{im},n}_{x,j}}\widetilde{V}^n_j(\omega), \tag{61}$$

where

$$\left(L_y\widetilde{V}^n\right)_j = \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{1}{h_y^2}\left(\widetilde{V}^n_{j+1} - 2\widetilde{V}^n_j + \widetilde{V}^n_{j-1}\right) - \frac{\kappa}{\epsilon}\frac{1}{2h_y}\left(y_{j+1}\widetilde{V}^n_{j+1} - y_{j-1}\widetilde{V}^n_{j-1}\right)$$

$$\widetilde{L}^{\mathrm{ex},n}_{x,j} = 1 - \mathrm{i}\,\rho_x\sigma(y_j + U^n)\frac{\sqrt{k}}{h_x}\sin(\omega)Z_n - 2\rho_x^2\frac{k}{h_x^2}\sin^2(\omega/2)(Z_n^2 - 1),$$

$$\widetilde{L}^{\mathrm{im},n}_{x,j} = 1 + 2\sigma^2(y_j + U^{n+1})\frac{k}{h_x^2}\sin^2(\omega/2) + \mathrm{i}\frac{k}{h_x}\sin(\omega).$$

Multiplying the left-hand side of (61) by $\widetilde{V}_j^{n+1,*}$, with $*$ denoting the complex conjugate, summing over j and carrying out summation by parts,

$$-\sum_j\left(L_y\widetilde{V}^{n+1}\right)_j\widetilde{V}_j^{n+1,*} = \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{|\widetilde{V}^{n+1}_{j+1} - \widetilde{V}^{n+1}_j|^2}{h_y^2} + \frac{\kappa}{\epsilon}y_j\widetilde{V}_j^{n+1}\frac{(\widetilde{V}^{n+1}_{j+1} - \widetilde{V}^{n+1}_{j-1})^*}{2h_y}.$$

Writing in the last term $\widetilde{V}^{n+1}_{j+1} - \widetilde{V}^{n+1}_{j-1} = (\widetilde{V}^{n+1}_{j+1} - \widetilde{V}^{n+1}_j) + (\widetilde{V}^{n+1}_j - \widetilde{V}^{n+1}_{j-1})$, shifting the index in the summation, apply Young's inequality as

$$\frac{1}{2}(y_j\widetilde{V}_j^{n+1} + y_{j+1}\widetilde{V}^{n+1}_{j+1})\frac{(\widetilde{V}^{n+1}_{j+1} - \widetilde{V}^{n+1}_j)^*}{h_y} \geq$$
$$-\frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{|\widetilde{V}^{n+1}_{j+1} - \widetilde{V}^{n+1}_j|^2}{h_y^2} - \frac{\epsilon\,|y_j\widetilde{V}_j^{n+1} + y_{j+1}\widetilde{V}^{n+1}_{j+1}|^2}{8\nu^2(1-\rho_y^2)}.$$

Upon a further application of Young's inequality and insertion,

$$\begin{aligned}\sum_j\left((I - kL_y)\widetilde{V}^{n+1}\right)_j\widetilde{V}_j^{n+1,*} &\geq \sum_j|\widetilde{V}_j^{n+1}|^2 - k\sum_j\frac{\epsilon\,|y_j\widetilde{V}_j^{n+1}|^2}{2\nu^2(1-\rho_y^2)}\\ &\geq \left(1 - k\frac{\epsilon\,y_{\max}^2}{2\nu^2(1-\rho_y^2)}\right)\sum_j|\widetilde{V}_j^{n+1}|^2. \qquad (62)\end{aligned}$$

To estimate the right-hand side of (61),

$$\begin{aligned}\mathbb{E}[|\widetilde{L}^{\mathrm{ex},n}_{x,j}|^2|\mathcal{F}_{t_n}] &= 1 + \rho_x^2\sigma^2(y_j + U^n)\frac{k}{h_x^2}\sin^2(\omega) + 8\rho_x^4\frac{k^2}{h_x^4}\sin^4(\omega/2)\\ &\leq \left(1 + 2\sqrt{2}\sup_{y\in\mathbb{R}}\sigma^2(y)\frac{k}{h_x^2}\sin^2(\omega/2)\right)^2,\\ |\widetilde{L}^{\mathrm{im},n}_{x,j}|^2 &\geq \left(1 + 2\inf_{y\in\mathbb{R}}\sigma^2(y)\frac{k}{h_x^2}\sin^2(\omega/2)\right)^2.\end{aligned}$$

Hence, assuming (59),

$$\mathbb{E}\left[\frac{|\widetilde{L}^{\mathrm{ex},n}_{x,j}|^2}{|\widetilde{L}^{\mathrm{im},n}_{x,j}|^2}\middle|\mathcal{F}_{t_n}\right] \leq 1.$$

On the right-hand side of (61), we have

$$\sum_j \frac{\widetilde{L}^{\mathrm{ex},n}_{x,j}}{\widetilde{L}^{\mathrm{im},n}_{x,j}} \widetilde{V}^n_j \widetilde{V}^{n+1,*}_j \leq \frac{1}{2}\sum_j \left|\frac{\widetilde{L}^{\mathrm{ex},n}_{x,j}}{\widetilde{L}^{\mathrm{im},n}_{x,j}}\right|^2 |\widetilde{V}^n_j|^2 + \frac{1}{2}\sum_j |\widetilde{V}^{n+1}_j|^2. \tag{63}$$

From (61), (62) and (63),

$$\left(1 - k\frac{\epsilon\, y^2_{\max}}{2\nu^2(1-\rho_y^2)}\right)\sum_j \mathbb{E}[|\widetilde{V}^{n+1}_j|^2|\mathcal{F}_{t_n}] \leq \frac{1}{2}\sum_j |\widetilde{V}^n_j|^2 + \frac{1}{2}\sum_j \mathbb{E}[|\widetilde{V}^{n+1}_j|^2|\mathcal{F}_{t_n}].$$

Rearranging, using Parseval's identity, and $1 - \delta/2 \geq \exp(-\delta)$ for any $0 \leq \delta \leq 1$, we obtain (60) by induction over n. □

The proof of Proposition 4.1 is more complicated than similar results in the literature, in that

- we seek stability with a constant independent of ϵ,
- the coefficients depend on the fast process Y, and
- the coefficients are variable in y.

We address this by a combination of a Fourier transform in x to take advantage of the constant coefficients as in a standard von Neumann stability analysis, and an energy-type argument for the y direction as is common in the finite element and finite difference literature.

Remark 4.1. *Note that* (59) *could be replaced by the simpler condition* $\sqrt{2}\rho_x^2 \leq 1$ *if the term* $\sigma^2(\mathcal{Y} + U^{n+1})$ *on the left-hand side of* (58) *was replaced by* $\sigma^2(\mathcal{Y} + U^n)$*. The need to use the crude bounds on* σ *comes from the fact that we cannot control the difference between* U^{n+1} *and* U^n*, especially for small* ϵ*.*

Zero-order approximation (in ϵ)

For the zero order term $v_{0,\epsilon}$, we denote the corresponding numerical solution as V_0, with the scheme

$$\begin{aligned}&\left(I - \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy} - \frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\right)\left(I - \frac{\langle\sigma^2\rangle}{2}\frac{k}{h_x^2}D_{xx} + \mu\frac{k}{2h_x}D_x\right)V_0^{n+1}\\ &= \left(I - \rho_x\langle\sigma\rangle\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x + \rho_x^2\langle\sigma\rangle^2\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\right)V_0^n,\end{aligned} \tag{64}$$

with notation as earlier. Note that the closed-form solution to V_0 is (47), and $\langle\sigma\rangle$, $\langle\sigma^2\rangle$ will be computed analytically for specific choices of $\sigma(\cdot)$ in the next section.

To determine the error (in ϵ) of the zero-order approximation, we can directly solve the SPDE for $v - v_{0,\epsilon}$ from (48). Denoting the solution as $\widetilde{V}_0$, we have the scheme

$$\begin{aligned}
&\Big(I - \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy} - \frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\Big)\Big(I - \frac{\sigma^2(\mathcal{Y}+U^{n+1})}{2}\frac{k}{h_x^2}D_{xx} + \mu\frac{k}{2h_x}D_x\Big)\widetilde{V}_0^{n+1} \\
&= \Big(I - \rho_x\sigma(\mathcal{Y}+U^n)\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x + \rho_x^2\sigma^2(Y+U^n)\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\Big)\widetilde{V}_0^{n+1} \\
&\quad - \frac{1}{2}\Big(\langle\sigma^2\rangle - \sigma^2(\mathcal{Y}+U_t)\Big)k\big(\partial_{xx}V_0^n\big) + \rho_x\Big(\langle\sigma\rangle - \sigma(\mathcal{Y}+U_t)\Big)\big(\partial_x V_0^n\big)\sqrt{k}Z_{n,x} \\
&\quad + \frac{1}{2}\rho_x^2\Big(\langle\sigma\rangle - \sigma(\mathcal{Y}+U_t)\Big)^2\big(\partial_{xx}V_0^n\big)k(Z_{n,x}^2-1),
\end{aligned} \tag{65}$$

where we use the analytic solution (47) to compute $\partial_x V_0^n = (\partial_x v_{0,\epsilon}(t_n, x_i, y_j))_{i,j}$, and similarly for $\partial_{xx}V_0$, and the initial condition $\widetilde{V}_0^0 = 0$.

First-order correction (in ϵ)

Similarly, the scheme for the approximation of the first-order term (49), denoted by V_1, is

$$\begin{aligned}
&\Big(I - \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy} - \frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\Big)\Big(I - \frac{\langle\sigma^2\rangle}{2}\frac{k}{h_x^2}D_{xx} + \mu\frac{k}{2h_x}D_x\Big)V_1^{n+1} \\
&= \Big(I - \rho_x\langle\sigma\rangle\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x + \rho_x^2\langle\sigma\rangle^2\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\Big)V_1^n \\
&\quad - \frac{1}{2}\Big(\langle\sigma^2\rangle - \sigma^2(\mathcal{Y}+U_t)\Big)k\big(\partial_{xx}V_0^n\big) + \rho_x\Big(\langle\sigma\rangle - \sigma(\mathcal{Y}+U_t)\Big)\big(\partial_x V_0^n\big)\sqrt{k}Z_{n,x} \\
&\quad + \frac{1}{2}\rho_x^2\Big(\langle\sigma\rangle - \sigma(\mathcal{Y}+U_t)\Big)^2\big(\partial_{xx}V_0^n\big)k(Z_{n,x}^2-1),
\end{aligned} \tag{66}$$

with zero initial condition, $V_1^0 = 0$. We also use the analytic solution for $\partial_x V_0$ and $\partial_{xx}V_0$ in the scheme (66).

In the computations below, we make some further specifications.

5 Numerical results

In this section, we illustrate the convergence of the expansion for the model (39) by way of numerical tests. For illustration, we use the following Gaussian distribution as initial condition:

$$v(0,x,y) = \Psi\big(t_0, x\,; W_{t_0}^x = 0\big)\Phi^\dagger(t_0, y), \tag{67}$$

where $t_0 > 0$ is a fixed constant, and Ψ, Φ are defined in (42) and (44), respectively. We choose this smooth initial condition so that when we numerically approximate $v_{1,\epsilon}$ in (49), the approximation is stable when taking the partial derivatives of $v_{0,\epsilon}$.[3]

[3]If we initialised with Dirac data, further stabilisation may be needed, and the analysis in ℓ_2 would not carry over; see the Fourier analysis in [15, 25] for schemes for simpler SPDEs with Dirac initial data.

Moreover, we can still obtain a closed-form solution to $v_{0,\epsilon}$ in (46),

$$v_{0,\epsilon}(T,x,y) = \Psi\big(t_0+T,x\,;W^x_{t_0}=0\big)\Phi^\dagger(t_0+T,y). \tag{68}$$

For analytical convenience, we further specify $\sigma(x)=\exp(\alpha x)$, where $\alpha>0$, such that $\exp(Y)$ follows an exponential OU process. This is also a popular stochastic volatility model in investment banks. Then we have by direct integration

$$\langle\sigma\rangle = \int_{-\infty}^{\infty}\frac{\kappa}{\nu\sqrt{2\pi}}\exp\Big(\alpha x-\frac{\kappa x^2}{2\nu^2}\Big)\,\mathrm{d}x = \exp\Big(\frac{\alpha^2\nu^2}{2\kappa}\Big), \qquad \langle\sigma^2\rangle = \exp\Big(\frac{4\alpha^2\nu^2}{2\kappa}\Big), \tag{69}$$

and

$$\langle\sigma(\cdot+U_t)\rangle_\dagger = \exp\left(\frac{\alpha^2\nu^2(1-\rho_y^2)}{2\kappa}+\alpha U_t\right),$$
$$\langle\sigma^2(\cdot+U_t)\rangle_\dagger = \exp\left(\frac{4\alpha^2\nu^2(1-\rho_y^2)}{2\kappa}+2\alpha U_t\right),$$

where $\langle\cdot\rangle$ denotes as earlier the average over the ergodic distribution of Y, and $\langle\cdot\rangle_\dagger$ over that of $Y^\dagger$ only.

As a base case, we choose the parameters $T=1,\ x_0=y_0=2,\ \mu=0.05,\ \rho_x=0.3,\ \rho_y=0.2,\ \rho=0.5,\ \kappa=0.2,\ \nu=0.5,\ \alpha=0.1$, and t_0 in the initial condition (67) as 0.2. We then vary ϵ and estimate the contribution of $v_{0,\epsilon}$ and $v_{1,\epsilon}$ to expected functionals of the solution. Later on, we will also test the effect of different parameters, in particular negative ρ, and different ratios of κ and ν.

For the computations, we truncate the domain to $[-10,10]\times[-10,10]$, chosen large enough that the effect of truncation with zero Dirichlet boundary conditions on the solution was found negligible for the tested parameter values.

To study the convergence $\epsilon\to 0$, we consider the linear functional

$$P_T(v) = \int_0^\infty\int_{-\infty}^\infty v(T,x,y)\,\mathrm{d}y\,\mathrm{d}x, \qquad T<\infty. \tag{70}$$

There are two motivations for this. First, convergence in distribution of $\mathbb{P}[X_T\in\mathcal{I}|W^x,W^y]$ for intervals $\mathcal{I}$ is theoretically supported by [18] (albeit for the process with absorption at $x=0$). Second, P_T models the aggregate loss in credit portfolio models as e.g. in [4, 3, 16, 18], and is therefore of practical interest.

To approximate $P_T(v)$ in (70), we use the trapezoidal rule for the numerical integration.

5.1 Weak convergence of $P_T(v_{0,\epsilon})$

We first analyse the numerical convergence of $\mathbb{E}[P_T(v)-P_T(v_{0,\epsilon})]$, where P_T is the functional from (70). Since P_T is linear, we have $P_T(v)-P_T(v_{0,\epsilon})=P_T(v-v_{0,\epsilon})$. We use the scheme (65) to directly approximate $v-v_{0,\epsilon}$, and estimate $\mathbb{E}[P_T(v-v_{0,\epsilon})]$ by standard Monte Carlo sampling as detailed below. While simulating $v-v_{0,\epsilon}$ from (48), compared to simulating v in (39) and $v_{0,\epsilon}$ in (46) separately, lead to the same $\mathbb{E}[P_T(v-v_{0,\epsilon})]$ for $h_x,h_y,k\to 0$, the former choice has computational savings and is hence faster by a constant factor.[4]

[4]Using the same Brownian paths for v and $v_{0,\epsilon}$ leads to a variance reduction and less paths are needed.

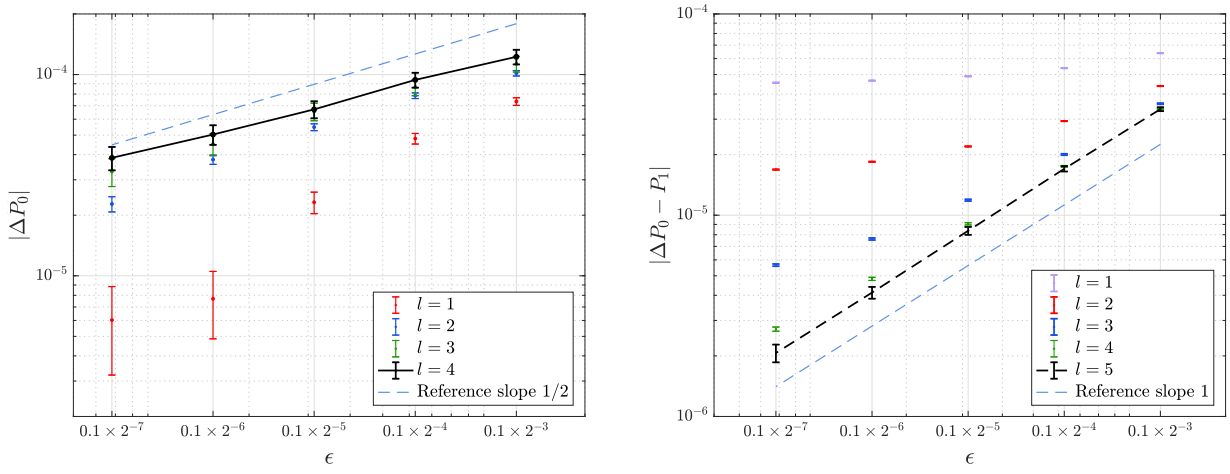
Given M Brownian paths, the Monte Carlo estimator for $\mathbb{E}[P_T(v - v_{0,\epsilon})]$ is

$$\Delta\widehat{P}_0 = \frac{1}{M}\sum_{i=1}^{M} P_T(\widetilde{V}_0^{(i)}),$$

where $\widetilde{V}_0^{(i)}$ is the numerical solution to $v - v_{0,\epsilon}$ for the i-th path of (W^x, W^y) in (65).

For each ϵ, the numerical error between $\Delta\widehat{P}_0$ and $\mathbb{E}[P_T(v - v_{0,\epsilon})]$ consists of discretisation error in h and k (bias), and Monte Carlo noise (variance). If $h,\ k \to 0$, and $M \to \infty$, $\Delta\widehat{P}_0$ is expected to converge to $\mathbb{E}[P_T(v - v_{0,\epsilon})]$, and we treat $\Delta\widehat{P}_0$ as the weak error in ϵ for small $h,\ k$ and large M.

Figure 1(a) shows the convergence to zero of $\Delta\widehat{P}_0$, with $\epsilon = 0.1\times 2^{-3},\ 0.1\times 2^{-4}, \ldots, 0.1\times 2^{-7}$, and error bars with 3 standard deviations. For each ϵ, we let $h_x = h_y = 2^{-l}$, and $k = 0.5\cdot 4^{-l}$, where $l = 1,\ 2,\ 3,\ 4$. We use $M = 10^6$ for $l = 1,\ 2$, and $M = 10^5$ for $l = 3,\ 4$, as the computational cost for finer meshes is large. For $l = 4$ with $M = 10^5$, the run time of the Matlab code is up to around 72 hours with 36 cores in parallel (speed 2300 RPMs, RAM 768GB, Linux system). We take the results from $l = 4$ as numerical approximation to $\mathbb{E}[P_T(v - v_{0,\epsilon})]$, shown as the black solid line in the loglog plot in Figure 1(a). One can identify from the figure that the slope is slightly less than 1/2 (see the dashed line), and linear regression gives a slope of 0.4237. A plausible reason is that for $l = 4$ the error in h and k is not small enough to be neglected. We deduce empirically that the weak error $\mathbb{E}[P_T(v - v_{0,\epsilon})]$ is $O(\sqrt{\epsilon})$.



(a) $\Delta\widehat{P}_0$. (b) $\Delta\widehat{P}_0 - \widehat{P}_1$.

Figure 1: Weak convergence of $\Delta\widehat{P}_0$ and $\Delta\widehat{P}_0 - \widehat{P}_1$.

5.2 Weak convergence of first order approximation

To verify $P_T(v_{1,\epsilon})$ is indeed the leading order approximation to $P_T(v) - P_T(v_{0,\epsilon})$, we further exhibit $\Delta\widehat{P}_0 - \widehat{P}_1$, which is the Monte Carlo estimator for $P_T(v - v_{0,\epsilon} - v_{1,\epsilon})$. Note that we have used the same Brownian paths for $\Delta\widehat{P}_0$ and $\widehat{P}_1$ to reduce the variance.

Figure 1(b) shows the loglog plot of the convergence of $\mathbb{E}[\Delta\widehat{P}_0 - \widehat{P}_1]$ to zero, with respect to ϵ, with the error bars being three standard deviations away. Similar to previous tests, here we stop at $l = 5$ with 1000 Monte Carlo samples, and linear regression yields a fitted slope of 1.0092.

We can thus deduce empirically that $P_T(v_{1,\epsilon})$ is the leading order approximation to $P_T(v - v_{0,\epsilon})$, with $P_T(v - v_{0,\epsilon} - v_{1,\epsilon}) = O(\epsilon)$.

5.3 Convergence of $\mathbb{E}[P_T(v_{1,\epsilon})]$

We now analyse $\mathbb{E}[P_T(v_{1,\epsilon})]$, where $v_{1,\epsilon}$ satisfies the SPDE (49), and compare it to $P_T^x(v_1^x)$, where v_1^x is the solution to the marginal SPDE (52), and $P_T^x(\cdot)$ is defined by

$$P_T^x(v^x) = \int_0^\infty v^x(T,x)\,\mathrm{d}x, \qquad T < \infty.$$

We expect these two values should be approximately the same.

From the derivation of the SPDE for v_1^x in Section 3.4, when we integrate over the y-dimension in (51), we replace $\Phi^Y(T,y)$ in the analytic solution (47) by the invariant distribution $\Phi_\infty^Y(y)$, which yields a more concise analytical form. This does not change the convergence order, as for $\epsilon \to 0$, $\Phi^Y(T,y)$ converges to $\Phi_\infty^Y(y)$ exponentially fast. Hence, to make sure $P_T(v_{1,\epsilon})$ gives the same results as $P_T(v_1^x)$, we also use $v_0(T,x,y) = \Psi(T,x)\Phi_\infty^Y(y)$ as the analytic solution for the zero order term $v_{0,\epsilon}$ in the schemes (65) and (66).

Given M Brownian paths, we define Monte Carlo estimators for $\mathbb{E}[P_T(v_{1,\epsilon})]$ and $\mathbb{E}[P_T(v_1^x)]$ by

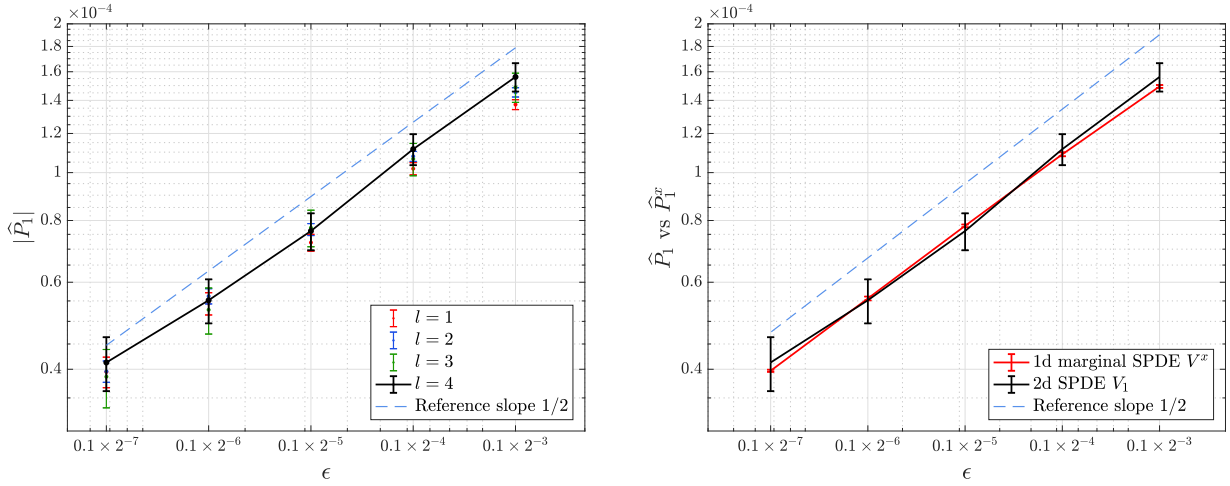
$$\widehat{P}_1 = \frac{1}{M}\sum_{i=1}^M P_T(V_1^{(i)}), \qquad \widehat{P}_1^x = \frac{1}{M}\sum_{i=1}^M P_T\big((V_1^x)^{(i)}\big), \tag{71}$$

where V_1 obeys the scheme (66), and V_1^x the scheme (57).

Figure 2(a) shows $\widehat{P}_1$, with $\epsilon = 0.1\times 2^{-3},\ 0.1\times 2^{-4}, \ldots, 0.1\times 2^{-7}$, and the error bars with 3 standard deviations. Similar to Figure 1(a), for each ϵ, we let $h_x = h_y = 2^{-l}$, and $k = 0.5\cdot 4^{-l}$, where $l = 1,\ 2,\ 3,\ 4$. We use $M = 10^6$ for $l = 1,\ 2$, and $M = 10^5$ for $l = 3,\ 4$. We take the results from $l = 4$ as numerical approximation to $\mathbb{E}[P_T(v_{1,\epsilon})]$, shown as the black solid line in the loglog plot in Figure 2(a), comparing it to a dashed line with slope 1/2.

We further include the approximation $\widehat{P}_1^x$ from the marginal SPDE (52), and compare $\widehat{P}_1^x$ with $\widehat{P}_1$ in Figure 2(b). For $\widehat{P}_1^x$, we use a multilevel Monte Carlo method, with prescribed 1% relative error (the ratio between root mean-square error and true value). We do not give any details on the multilevel construction here (see, e.g., [24]), and only note that, unlike approximating $\mathbb{E}[P_T(v)]$ by a standard Monte Carlo method, and discretising with mesh size $h_l = h_0\times 2^{-l}$ and timestep $k_l = k_0\times 4^{-l}$, approximating by MLMC requires a good coupling between fine path and coarse path. Therefore, when we apply MLMC to estimate $\mathbb{E}[P_T(v_1^x)]$, the timestep is set as $k_l = k_0\times\epsilon\times 4^{-l}$, proportional to ϵ.

Linear regression yields that the fitted slope for $\widehat{P}_1$ is 0.4855, and for $\widehat{P}_1^x$ it is 0.4792. This is consistent with the finding from the previous section that inclusion of the first order term approximately cancels the error of the zero order approximation, which is of order $\sqrt{\epsilon}$. Moreover, since v_1^x is the solution to a one-dimensional SPDE, the computational cost



(a) $\widehat{P}_1$ defined in (71).

(b) Comparing $\widehat{P}_1$ with $\widehat{P}_1^x$ from (71).

Figure 2: Weak convergence of $\widehat{P}_1$ for $h_x = h_y = 2^{-l}$, $k = 0.5{\cdot}4^{-l}$, $\epsilon = 0.1{\times}2^{-3}, \cdots, 0.1{\times}2^{-7}$.

is much lower for $\mathbb{E}[P_T(v_1^x)]$ with the same accuracy than for $\mathbb{E}[P_T(v_{1,\epsilon})]$, which shows the benefit of our asymptotic expansion.

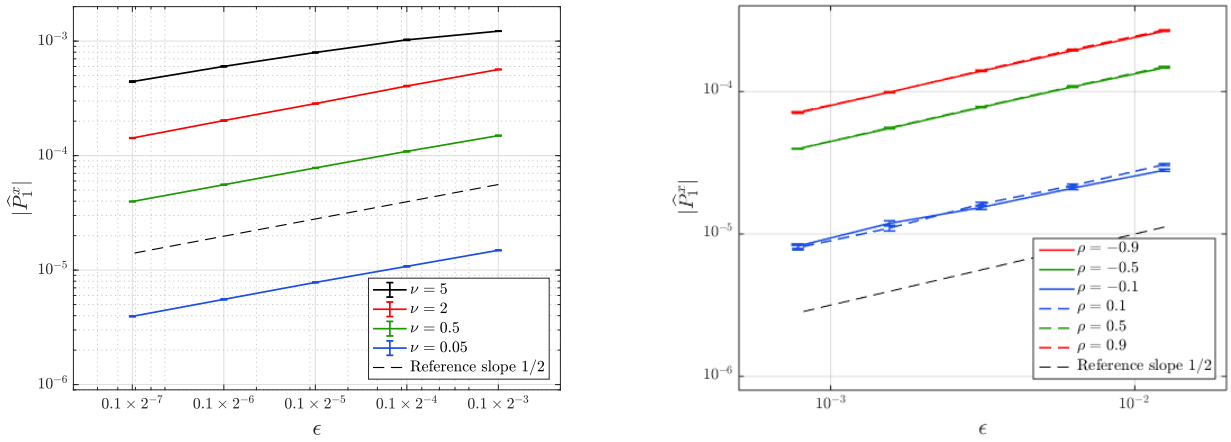
5.4 Parameter studies

Finally, we test the effect of the ratio between ν and κ, and of the correlation ρ. Figure 3(a) shows the effect of different ν/κ, by varying $\nu \in \{0.05, 0.5, 2, 5\}$, while keeping other parameters fixed. We choose the numerical parameters to ensure that the relative error is below 1%. We can see from Figure 3(a) that $|\widehat{P}_1^x|$ increases as ν/κ increases. From the raw data we found approximately that for fixed ϵ, $|\widehat{P}_1^x(\nu/\kappa;\epsilon)| = O(\nu/\kappa)$, which is consistent with the previous tests where ϵ varies, through the scaling relationship $(\nu/\sqrt{\epsilon})/(\kappa/\epsilon) = \sqrt{\epsilon}\,\nu/\kappa$. A point to note is that the values for $\nu = 5$ are negative, whereas the others are positive.

Figure 3(b) shows the effect of changing the parameter ρ, where we make sure to keep the relative error less than 5%. We can see from Figure 3(b) that $|\widehat{P}_1^x|$ increases as $|\rho|$ increases; inspection of the raw data shows that $\widehat{P}_1^x$ is positive when ρ is positive, and vice versa. This effect is similar to the asymptotic expansion of the backward PDE for the stochastic volatility model [9], where the leading order correction term is proportional to the correlation between the two Brownian motions involved.

References

[1] Bain, A., and Crisan, D. *Fundamentals of Stochastic Filtering*, vol. 3. Springer, 2009.



(a) $\widehat{P}_1^x$ with different ν.

(b) $\widehat{P}_1^x$ with different ρ.

Figure 3: Comparing $\widehat{P}_1^x$ with regard to ν and ρ .

[2] Bichuch, M., and Sircar, R. Optimal investment with transaction costs and stochastic volatility part I: Infinite horizon. *SIAM Journal on Control and Optimization 55*, 6 (2017), 3799–3832.

[3] Bujok, K., and Reisinger, C. Numerical valuation of basket credit derivatives in structural jump-diffusion models. *Journal of Computational Finance 15*, 4 (2012), 115–158.

[4] Bush, N., Hambly, B. M., Haworth, H., Jin, L., and Reisinger, C. Stochastic evolution equations in portfolio credit modelling. *SIAM Journal on Financial Mathematics 2*, 1 (2011), 627–664.

[5] Conlon, J. G., and Sullivan, M. G. Convergence to Black–Scholes for ergodic volatility models. *European Journal of Applied Mathematics 16*, 3 (2005), 385–409.

[6] Cozma, A. S., and Reisinger, C. Simulation of conditional expectations under fast mean-reverting stochastic volatility models. In *International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing* (2020), Springer, pp. 223–240.

[7] Fouque, J.-P., and Hu, R. Multiscale asymptotic analysis for portfolio optimization under stochastic environment. *Multiscale Modeling & Simulation 18*, 3 (2020), 1318–1342.

[8] Fouque, J.-P., Lorig, M., and Sircar, R. Second order multiscale stochastic volatility asymptotics: stochastic terminal layer analysis and calibration. *Finance and Stochastics 20*, 3 (2016), 543–588.

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

[10] Fouque, J.-P., Papanicolaou, G., Sircar, R., and Solna, K. Multiscale stochastic volatility asymptotics. *Multiscale Modeling & Simulation 2*, 1 (2003), 22–42.

[11] Fouque, J.-P., Papanicolaou, G., Sircar, R., and Solna, K. Singular perturbations in option pricing. *SIAM Journal on Applied Mathematics 63*, 5 (2003), 1648–1665.

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

[13] Fukasawa, M. Asymptotic analysis for stochastic volatility: martingale expansion. *Finance and Stochastics 15* (2011), 635–654.

[14] Giesecke, K., Spiliopoulos, K., Sowers, R., and Sirignano, J. Large portfolio asymptotics for loss from default. *Mathematical Finance 25*, 1 (2015), 77–114.

[15] Giles, M. B., and Reisinger, C. Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance. *SIAM Journal on Financial Mathematics 3*, 1 (2012), 572–592.

[16] Hambly, B., and Kolliopoulos, N. Stochastic evolution equations for large portfolios of stochastic volatility models. *SIAM Journal on Financial Mathematics 8*, 1 (2017), 962–1014.

[17] Hambly, B., and Kolliopoulos, N. Stochastic PDEs for large portfolios with general mean-reverting volatility processes. *arXiv preprint arXiv:1906.05898* (2019).

[18] Hambly, B., and Kolliopoulos, N. Fast mean-reversion asymptotics for large portfolios of stochastic volatility models. *Finance and Stochastics 24*, 3 (2020), 757–794.

[19] Howison, S. Matched asymptotic expansions in financial engineering. *Journal of Engineering Mathematics 53* (2005), 385–406.

[20] Kurtz, T. G., and Xiong, J. Particle representations for a class of nonlinear SPDEs. *Stochastic Processes and their Applications 83*, 1 (1999), 103–126.

[21] Papanicolaou, A. Nonlinear filters for hidden Markov models of regime change with fast mean-reverting states. *Multiscale Modeling & Simulation 10*, 3 (2012), 906–935.

[22] Papanicolaou, A., and Spiliopoulos, K. Filtering the maximum likelihood for multiscale problems. *Multiscale Modeling & Simulation 12*, 3 (2014), 1193–1229.

[23] Reisinger, C. Mean-square stability and error analysis of implicit time-stepping schemes for linear parabolic SPDEs with multiplicative Wiener noise in the first derivative. *International Journal of Computer Mathematics 89*, 18 (2012), 2562–2575.

[24] Reisinger, C., and Wang, Z. Analysis of multi-index Monte Carlo estimators for a Zakai SPDE. *Journal of Computational Mathematics 36*, 2 (2018), 202–236.

[25] Reisinger, C., and Wang, Z. Stability and error analysis of an implicit Milstein finite difference scheme for a two-dimensional Zakai SPDE. *BIT Numerical Mathematics 59*, 4 (2019), 987–1029.