MULTISCALE MODEL. SIMUL.
Vol. 23, No. 1, pp. 486–513
© 2025 Society for Industrial and Applied Mathematics

ASYMPTOTIC EXPANSIONS OF EVOLUTION EQUATIONS WITH FAST VOLATILITY*

SAM D. HOWISON†, CHRISTOPH REISINGER†,
RONNIE SIRCAR‡, AND ZHENRU WANG†

Abstract. We study Kolmogorov forward equations (KFEs) 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., *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*, Cambridge, University Press, 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 homogenized one-dimensional SPDEs.

Key words. matched asymptotic expansions, stochastic volatility, fast mean-reversion, forward Kolmogorov equation, Zakai equation

MSC codes. 60H15, 35R60, 35C20, 65C30, 65M99

DOI. 10.1137/23M1610720

1. Introduction.

In this work, we consider processes of the form

$$\begin{aligned} \mathrm{d}X_t^\epsilon &= \mu(X_t^\epsilon, Y_t^\epsilon)\,\mathrm{d}t + \sigma(Y_t^\epsilon)\Big(\rho_x\,\mathrm{d}W_t^x + \sqrt{1-\rho_x^2}\,\mathrm{d}W_t^{x,1}\Big), \quad X_0 = x_0, \\ \mathrm{d}Y_t^\epsilon &= \frac{\kappa}{\epsilon}(m - Y_t^\epsilon)\,\mathrm{d}t + \frac{g(Y_t^\epsilon)}{\sqrt{\epsilon}}\Big(\rho_y\,\mathrm{d}W_t^y + \sqrt{1-\rho_y^2}\,\mathrm{d}W_t^{y,1}\Big), \quad Y_0 = y_0, \end{aligned} \tag{1.1}$$

where (W^x, W^y) is a two-dimensional Brownian motion with correlation $\rho \in (-1,1)$, while $W^{x,1}, W^{y,1}$ are Brownian motions that 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 $\sigma, g : \mathbb{R} \to \mathbb{R}_+$ are given functions. For ease of notation, we introduce $\rho_{xy} = \rho_x\rho_y\rho$.

We will study the distribution of $(X_t^\epsilon, Y_t^\epsilon)$, and the distribution of $(X_t^\epsilon, Y_t^\epsilon)$ 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 scale of the Y-process to its mean m, and we 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).

*Received by the editors October 23, 2023; accepted for publication (in revised form) October 14, 2024; published electronically February 12, 2025.
https://doi.org/10.1137/23M1610720
†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 08544 USA (sircar@princeton.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Models of the form (1.1) are used abundantly in financial engineering, where X^ϵ describes the log price of a financial asset and $\sigma(Y_t^\epsilon)$ 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., [14, 13], and especially the monograph [15] and the references therein. These works derive asymptotic expansions for Kolmogorov *backward* equations (KBEs) arising as pricing equations for financial derivatives on such underlying assets with fast mean-reverting stochastic volatility. For higher order expansions with a refined boundary layer (in time) analysis close to expiry, we refer the reader to [22] and [11], and also to [7] for a convergence analysis.

The KBEs are in nondivergence form and typically have regular (i.e., continuous) terminal data. The KFEs studied in this paper are in divergence form. Asymptotic expansions for such KFEs have been derived in [23] for smooth initial data, while we will consider Dirac delta initial data, corresponding to the important case where the initial value of the stochastic process is known.

The expansion at the level of the underlying stochastic processes is analyzed in [16] (see also the earlier discussion in the conclusions of [22]). Among more recent works, [3] performs joint asymptotic expansions of optimal investment models with fast volatility and small transaction costs, and [9, 10] provide a multiscale analysis of portfolio optimization strategies under fast and slow volatilities.

For concreteness, we will consider the model where μ is constant, and $g = \nu\sqrt{2}$ for constant ν, that is, where Y is an Ornstein–Uhlenbeck (OU)_process, with unique ergodic distribution $\mathcal{N}(m, \nu^2)$.[1] In this case, the KFE is

$$\begin{aligned} \partial_t p^\epsilon &= \frac{1}{\epsilon}\left(\nu^2\partial_{yy}p^\epsilon - \kappa\partial_y((m-y)p^\epsilon)\right) + \left(\frac{1}{2}\sigma^2(y)\partial_{xx}p^\epsilon - \mu\,\partial_x p^\epsilon\right) \\ &\quad + \frac{1}{\sqrt{\epsilon}}\rho_{xy}\nu\sqrt{2}\partial_y(\sigma(y)\partial_x p^\epsilon), \\ p^\epsilon(0,x,y) &= \delta(x-x_0)\otimes\delta(y-y_0). \end{aligned} \tag{1.2}$$

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

$$\begin{aligned} \mathrm{d}u^\epsilon &= \frac{1}{\epsilon}\left(\nu^2\partial_{yy}u^\epsilon - \kappa\partial_y((m-y)u^\epsilon)\right)\mathrm{d}t + \left(\frac{1}{2}\sigma^2(y)\partial_{xx}u^\epsilon - \mu\,\partial_x u^\epsilon\right)\mathrm{d}t \\ &\quad + \frac{1}{\sqrt{\epsilon}}\rho_{xy}\nu\sqrt{2}\partial_y(\sigma(y)\partial_x u^\epsilon)\,\mathrm{d}t - \rho_x\sigma(y)\partial_x u^\epsilon\,\mathrm{d}W_t^x - \rho_y\frac{\sqrt{2}\nu}{\sqrt{\epsilon}}\partial_y u^\epsilon\,\mathrm{d}W_t^y, \\ u^\epsilon(0,x,y) &= \delta(x-x_0)\otimes\delta(y-y_0). \end{aligned} \tag{1.3}$$

There are at least two motivations for studying (1.3). First, by general filtering theory (see, e.g., [2, section 3.5]), the solution u^ϵ is the density (if it exists) of the conditional law of $(X_t^\epsilon, Y_t^\epsilon)$ 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 realizations of (1.1), with independent (idiosyncratic) noise terms $W_t^{x,i}, W_t^{y,i}$, for $i = 1,\dots,N$, 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., [24]).

This limiting equation has been used to describe the behavior 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 analyzed in [6] and applications to credit

[1] The factor $\sqrt{2}$ is chosen to ensure the ergodic distribution has a more standard normal form, consistent with the literature (see [12, 15]).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

derivative markets are given. For a (nonlinear) SPDE model for a large pool limit of a default intensity-based credit model see, e.g., [17]. More recently, an extension of the basic model in [6] to stochastic volatilities is given in [19, 20]. See also [26, 27] 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 behavior of X^ϵ, and in Y^ϵ 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 homogenized equations which allow for more efficient analytical 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^ϵ (and not Y^ϵ) 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 (1.1), [21] considers the case of $\mu(X_t^\epsilon, Y_t^\epsilon) = r - \sigma^2(Y_t^\epsilon)/2$ for constant r.[2] Under certain recurrence properties of the diffusion Y^ϵ, and for $\rho_{xy} = 0$, it is shown that as $\epsilon \to 0$, the stopped version of X^ϵ converges in distribution to a process X^* which satisfies

$$\mathrm{d}X_t^* = (r - \langle\sigma^2\rangle/2)\,\mathrm{d}t + \sqrt{\langle\sigma^2\rangle}\,\mathrm{d}B_t, \quad X_0^* = x_0, \tag{1.4}$$

where $\langle\sigma^2\rangle$ is the expectation of $\sigma^2(\cdot)$ under the invariant distribution of Y^ϵ. Note that in (1.4) we were able to replace $\rho W_t^x + \sqrt{1-\rho^2}W_t^{x,1}$ from (1.1) by a standard Brownian motion B_t because we study convergence in distribution, whereas this is not possible when we consider conditional laws u^ϵ. It is shown in [21] that weak limits of u^ϵ satisfy the SPDE

$$\mathrm{d}u^* = \left(\frac{1}{2}\langle\sigma^2\rangle\partial_{xx}u^* - (r - \langle\sigma^2\rangle/2)\,\partial_x u^*\right)\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x u^*\,\mathrm{d}W_t^x, \tag{1.5}$$

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

We note that [21] allows for more general g in (1.1) than for the OU 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 behavior 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 OU processes, which will allow a decomposition of Y^ϵ such that the SPDE is driven purely by the slow component, while the fast-mean-reverting term appears explicitly

[2] In fact, [21] allows that r and other (constant) model parameters are sampled randomly.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

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 approximation the KFE (1.2) for small ϵ, which differs from [23] by considering Dirac initial data (and hence a more intricate behavior in an initial layer) and the equation posed on the whole plane (section 2);
- the heuristic derivation of a one-dimensional SPDE for a first-order correction to (1.5) in ϵ (section 3);
- numerical verification of the SPDE expansion orders (in section 5) by novel numerical schemes for (1.3) and related SPDEs, with a proof of unconditional stability independent of ϵ (section 4).

Sections 2 and 3 provide formal asymptotic analysis of the two problems studied here, the novelty being the matched asymptotics at the initial layer. A rigorous convergence result along the lines of [11] is beyond the scope of the current paper. We expect it is certainly feasible, but involving detailed technical estimates which we relegate to future work.

2. Perturbation analysis of the KFE.

2.1. Summary of structure and leading-order approximations. In this section, for later reference and comparison, we summarize the asymptotic structure as $\epsilon \to 0^+$ of the transition density function of a stochastic volatility process (X^ϵ, Y^ϵ) satisfying (1.1) with $g = \nu\sqrt{2}$ for constant ν. Such problems have been studied for some time and in particular our results are closely related to those in [23]. These authors consider "slow" and "fast" processes which live in connected, compact Riemannian manifolds of dimension r and d, respectively, and the driving Brownian motion has dimension p; the drift and volatility coefficient matrices are arbitrary C^∞ functions of all the state variables. Importantly, their SDEs have a two-scale structure which includes ours as a special case (their small parameter ε is the same as our $\sqrt{\epsilon}$). It is shown in [23] that the transition density function, which satisfies the forward Kolmogorov parabolic PDE (FKE for short), has an asymptotic behavior as $\varepsilon \to 0^+$ in which, roughly speaking, the density of the "fast" (volatility) variable evolves on a more rapid timescale than that of the "slow" variable, leading to a structure comprising (1) a small initial layer in which the marginal density of the fast variable evolves to its invariant distribution, followed by (2) an outer region in which the slow process evolves with all volatility dependence replaced with suitable averages with respect to the invariant measure of the fast variable. They also provide a "composite" expansion, which is valid in both regions.

Although our SDEs are a special case of those in [23], our analysis differs in several respects. We solve on $\mathbb{R} \times \mathbb{R}$, which is not compact and leads to some variations in the application of the Fredholm alternative. We consider a delta-function as the initial data for the FKE, rather than the C^∞ data of [23], leading to distributional solutions of the FKE in the initial layer away from the initial value of the SDEs; this in turn necessitates an "inner-inner" region near that point. Because we have some constant coefficients, we are able to obtain the invariant distribution of the fast variable explicitly. Finally, we highlight the correlation term in the FKE explicitly (rather than implicitly as in [23]); although this does not alter the analysis as compared with that in [23], the corresponding term in our later analysis of the SPDE does require special treatment, as described in section 3.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

The transition density $p^\epsilon(t_0=0,x_0,y_0;t,x,y)$ of (X^ϵ,Y^ϵ) satisfies the FKE

$$\partial_t p^\epsilon - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* + \mathcal{L}_2^*\right) p^\epsilon = 0, \tag{2.1}$$

with the initial condition

$$p^\epsilon(0,x_0,y_0;0,x,y) = \delta(x-x_0)\otimes\delta(y-y_0). \tag{2.2}$$

Here the operators $\mathcal{L}_i$ and their adjoints $\mathcal{L}_i^*$ are defined by[3]

$$\mathcal{L}_0 = \nu^2\partial_{yy}(\,\cdot\,) + \kappa(m-y)\partial_y(\,\cdot\,), \qquad \mathcal{L}_0^* = \nu^2\partial_{yy}(\,\cdot\,) - \kappa\partial_y\left((m-y)\cdot\,\right), \tag{2.3}$$

$$\mathcal{L}_1 = \rho_{xy}\nu\sqrt{2}\sigma(y)\partial_{xy}(\,\cdot\,), \qquad \mathcal{L}_1^* = \rho_{xy}\nu\sqrt{2}\partial_x\partial_y\left(\sigma(y)\cdot\,\right), \tag{2.4}$$

$$\mathcal{L}_2 = \frac{1}{2}\sigma^2(y)\partial_{xx}(\,\cdot\,) + \mu(x,y)\partial_x(\,\cdot\,), \qquad \mathcal{L}_2^* = \frac{1}{2}\sigma^2(y)\partial_{xx}(\,\cdot\,) - \partial_x\left(\mu(x,y)\cdot\,\right). \tag{2.5}$$

We denote by Φ^Y the probability density function (PDF) of the ergodic distribution of Y^ϵ

$$\Phi^Y(y) = \frac{1}{\sqrt{2\pi\nu^2/\kappa}}\mathrm{e}^{-\kappa(y-m)^2/2\nu^2} \tag{2.6}$$

and the average of a functions of y with respect to the measure $\Phi^Y(y)\,\mathrm{d}y$ by $\langle\,\cdot\,\rangle = \langle\,\cdot\,,\Phi^Y\rangle$, using the usual inner product on $L^2(\mathbb{R})$.

As set out in detail in Appendix A, the asymptotic structure has four regions, each with its associated expansion in powers of $\epsilon^{\frac{1}{2}}$. The regions, and the first one or two terms of the expansions of the corresponding solution to the FKE, are as follows:

1. An "inner-inner" region, in which $x-x_0=O(\sqrt{\epsilon})$ and $t=O(\epsilon)$ while y is not rescaled; to leading-order the transition density takes the form $(1/\sqrt{\epsilon})p_0''(t':=t/\epsilon, x':=(x-x_0)/\sqrt{\epsilon}, y)$ satisfying (A.30), (A.31), namely

$$\partial_{t'}p_0'' - \left(\frac{1}{2}\sigma^2(y)\partial_{x'x'} + \rho_{xy}\nu\sqrt{2}\sigma(y)\partial_{x'y} + \nu^2\partial_{yy}\right)p_0'' + \kappa\partial_y\left((m-y)p_0''\right) = 0, \tag{2.7}$$

$$p_0''(0,x',y) = \delta(x')\delta(y-y_0). \tag{2.8}$$

   This is a fully parabolic problem and the solution is smooth. Note that [23] only considers smooth initial data and therefore does not have this region.

2. An initial layer in which $t=O(\epsilon)$, $|x-x_0|\gg O(\sqrt{\epsilon})$ but is otherwise not rescaled, and y is also not rescaled; at leading order, the solution takes the form $p'(t'=t/\epsilon,x,y)\sim p_0'(t',x,y)+\sqrt{\epsilon}p_1'(t',x,y)+O(\epsilon)$. The leading-order term is

$$p_0'(t',x,y) = \delta(x-x_0)\phi^Y(t',y),$$

   where, with $m(t')$ and $\mathrm{var}(t')$ defined as in (A.24),

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

   is the evolving marginal density of Y^ϵ with X^ϵ fixed at x_0. In this initial layer, the marginal density of Y^ϵ rapidly transits to its invariant distribution

[3]Note that [23] uses $\mathcal{L}_1$ where we use $\mathcal{L}_0$, and so on.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

while the marginal density of X^ϵ remains concentrated at x_0 (resolved by the inner-inner layer). The next term satisfies (A.27), namely

$$(\partial_{t'} - \mathcal{L}_0^*)\, p_1' = \rho_{xy}\nu\sqrt{2}\,(\partial_x \delta(x-x_0))\,\partial_y\left(\sigma(y)\phi^Y(t',y)\right), \qquad p_1'(0,x,y) = 0. \tag{2.9}$$

3. The outer region in which no variable is rescaled, and we expand the solution to the FKE as $p^\epsilon(t,x,y) \sim p_0(t,x,y) + \sqrt{\epsilon}p_1(t,x,y) + O(\epsilon)$. Then $p_0(t,x,y) = f_0(t,x)\Phi^Y(y)$, where from (A.15)

$$\begin{aligned} \partial_t f_0 &= \frac{1}{2}\langle\sigma^2(\cdot)\rangle\partial_{xx} f_0 - \partial_x\left(\langle\mu(x,\cdot)\rangle f_0\right) = \langle\mathcal{L}_2^*\rangle f_0(t,x), \\ f_0(0,x) &= \delta(x-x_0). \end{aligned} \tag{2.10}$$

In this equation, all coefficients are averaged with respect to the invariant measure $\Phi^Y(y)\,\mathrm{d}y$. Moreover, as in (A.14), (A.16), and (A.29),

$$p_1(t,x,y) = -(\rho_{xy}\sqrt{2}/\nu)\,(\partial_x f_0(t,x))\,\Sigma(y)\Phi^Y(y) + f_1(t,x)\Phi^Y(y), \tag{2.11}$$

where $\mathrm{d}\Sigma(y)/\mathrm{d}y = \sigma(y)$ and $f_1(t,x)$ satisfies

$$\begin{aligned} \partial_t f_1 - \langle\mathcal{L}_2^*\rangle\, f_1 = -(\rho_{xy}\sqrt{2}/\nu)\Big(&\frac{1}{2}\langle\sigma^2(\cdot)\Sigma(\cdot)\rangle\partial_{xx}\left(\partial_x f_0(t,x)\right) \\ &- \partial_x\left(\langle\mu(x,\cdot)\Sigma(\cdot)\rangle\partial_x f_0(t,x)\right)\Big), \end{aligned} \tag{2.12}$$

with initial condition

$$f_1(0,x) = -(\rho_{xy}\sqrt{2}/\nu)\,(\partial_x\delta(x-x_0))\,\langle\Sigma\rangle. \tag{2.13}$$

4. A far field, in which a large-deviations approach is needed; we do not treat this here.

More details, including higher-order terms, are given in Appendix A.

2.2. A global approximation and correction equation. As we know the density of the OU process Y^ϵ to be $\Phi^Y(t/\epsilon, y)$, from (A.25), we can define an approximation globally in time as

$$p_{0,\epsilon}(t,x,y) = \Phi^Y(t/\epsilon,y) f_0(t,x), \tag{2.14}$$

which has the correct initial datum $p_{0,\epsilon}(0,x,y) = \delta(x-x_0)\otimes\delta(y-y_0)$ and the exact marginal density for Y_t^ϵ and is correct to leading order in ϵ in both the inner and outer layers. However, we note that it is not valid in the (small) inner-inner layer, where the correct leading-order description, $p_0''(t',x',y)$, satisfies (A.30), (A.31); the first of these is not satisfied (to the same order in ϵ) by (2.14) (although both give approximations of $\delta(x-x_0)\otimes\delta(y-y_0)$ as $\epsilon\to 0^+$, they are different approximations). Our global approximation is the same as equation (3.1) of [23], allowing for the way they study the boundary layer correction (which is what we now do); but the caveat for the smoothing of the initial data is not needed in their case.

By insertion, we see directly that $p_{0,\epsilon}$ from (2.14) satisfies

$$\partial_t p_{0,\epsilon} - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \langle\mathcal{L}_2^*\rangle\right) p_{0,\epsilon} = 0. \tag{2.15}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Taking the difference with (A.1), we have

(2.16)
$$\partial_t(p^\epsilon - p_{0,\epsilon}) - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* + \mathcal{L}_2^*\right)(p^\epsilon - p_{0,\epsilon}) = \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* p_{0,\epsilon} - (\langle\mathcal{L}_2^*\rangle - \mathcal{L}_2^*)p_{0,\epsilon}.$$

Replacing the operator on the left-hand side in (2.16) by that in (2.15), we obtain an inhomogeneous SPDE with averaged coefficients on the left-hand side,

(2.17)
$$\partial_t p_{1,\epsilon} - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \langle\mathcal{L}_2^*\rangle\right)p_{1,\epsilon} = \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* p_{0,\epsilon} - (\langle\mathcal{L}_2^*\rangle - \mathcal{L}_2^*)p_{0,\epsilon}, \qquad p_{1,\epsilon}(0,x,y) = 0.$$

Considering $p_{1,\epsilon}$ as an approximation to $p^\epsilon - p_{0,\epsilon}$, we could define an approximation to p^ϵ by $p_{0,\epsilon} + p_{1,\epsilon}$.

We do not pursue this further here, but 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 (1.2) shows the different ansatz needed for small times (the "inner layer" in subsection A.3), where the process Y^ϵ transits from its initial value to the stationary distribution, and for all times after this initial transit (the "outer layer" in subsection A.2). These two expressions can be reconciled by using the analytical exact form for the marginal law of Y^ϵ and an expansion for the effect of Y^ϵ on X^ϵ only (see subsection 2.2).

For the SPDE (1.3), the presence of a fast driving process at a timescale $O(1/\epsilon)$ correlated to a driving process with an $O(1)$ timescale 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 [21] 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^ϵ into a component U^ϵ which has the correct instantaneous correlation with X^ϵ and a component $Y^{\epsilon,\dagger}$ which is independent of X^ϵ. We then study the joint dynamics of X^ϵ and $Y^{\epsilon,\dagger}$ conditional on (W^x, W^y), keeping track of the dynamics of $(U^\epsilon, Y^{\epsilon,\dagger})$ and its law exactly, while we approximate the generator of X^ϵ by an expansion.

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

(3.1)
$$\mathrm{d}U_t^\epsilon = -\frac{\kappa}{\epsilon}U_t^\epsilon\,\mathrm{d}t + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}}\rho_y\,\mathrm{d}W_t^y, \qquad U_0 = 0.$$

Then $Y^{\epsilon,\dagger} := Y^\epsilon - U^\epsilon - m$ satisfies

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

where $W^{y,1}$ is independent of W^x and $W^{x,1}$, and W^x, W^y have correlation ρ. We recall here and for future reference the elementary fact that if $Z \sim N(0,\rho_y^2)$ and $Z^\dagger \sim N(0,1-\rho_y^2)$ are independent, then $Z + Z^\dagger \sim N(0,1)$. 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^ϵ by a horizontal translation of the function σ.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

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

$$\begin{aligned} \mathrm{d}v^\epsilon &= \frac{1}{\epsilon}\Big(\nu^2(1-\rho_y^2)\partial_{yy}v^\epsilon + \kappa\partial_y(yv^\epsilon)\Big)\,\mathrm{d}t \\ &\quad + \left(\frac{1}{2}\sigma^2\left(y+U_t^\epsilon\right)\partial_{xx}v^\epsilon - \mu\,\partial_x v^\epsilon\right)\mathrm{d}t - \rho_x\sigma\big(y+U_t^\epsilon\big)\partial_x v^\epsilon\,\mathrm{d}W_t^x \\ &= \left(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \widetilde{\mathcal{L}}_2^*\right)v^\epsilon\,\mathrm{d}t - \rho_x\sigma(y+U_t^\epsilon)\partial_x v^\epsilon\,\mathrm{d}W_t^x, \\ v^\epsilon(0,x,y) &= \delta(x-x_0)\otimes\delta(y-y_0), \end{aligned} \tag{3.2}$$

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, and $\sigma : \mathbb{R} \to \mathbb{R}_+$ is a real-valued function for 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^ϵ to the SPDE (3.2) as $\epsilon \to 0$.

3.2. Zero order term. Following [21], and in line with the findings of section 2, we introduce the following SPDE akin to (1.5), by averaging the coefficients over the stationary distribution:

$$\begin{aligned} \mathrm{d}v_0^x &= \left(\frac{1}{2}\langle\sigma^2\rangle\partial_{xx}v_0^x - \mu\,\partial_x v_0^x\right)\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_0^x\,\mathrm{d}W_t^x \\ &= \langle\widetilde{\mathcal{L}}_2^*\rangle v_0^x\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_0^x\,\mathrm{d}W_t^x, \\ v_0^x(0,x) &= \delta(x-x_0), \end{aligned} \tag{3.3}$$

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

$$\langle\widetilde{\mathcal{L}}_2^*\rangle\cdot := \frac{1}{2}\langle\sigma^2\rangle\partial_{xx}\,\cdot\, - \mu\,\partial_x\,\cdot\,. \tag{3.4}$$

The SPDE (3.3) has the analytic solution $v_0^x(t,x) = \Psi(t,x)$, where

$$\Psi(t,x) := \frac{1}{\sqrt{2\pi\big(\langle\sigma^2\rangle - \rho_x^2\langle\sigma\rangle^2\big)t}}\exp\left(-\frac{(x-x_0-\mu t-\rho_x\langle\sigma\rangle W_t^x)^2}{2\big(\langle\sigma^2\rangle-\rho_x^2\langle\sigma\rangle^2\big)t}\right), \tag{3.5}$$

which can be checked directly by an application of Itô's lemma.

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

$$\Phi^\dagger(t',y) = \frac{1}{\sqrt{2\pi\sigma_\dagger^2(t')}}\exp\left(-\frac{\big(y-\mu_\dagger(t')\big)^2}{2\sigma_\dagger^2(t')}\right), \tag{3.6}$$

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

$$\mu_\dagger(t') = y_0\mathrm{e}^{-\kappa t'}, \quad \sigma_\dagger^2(t') = \frac{(1-\rho_y^2)\nu^2}{\kappa}\left(1-\mathrm{e}^{-2\kappa t'}\right) \tag{3.7}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

(which are $\mu_\dagger(t') = m(t')$ with mean 0 and $\sigma_\dagger^2(t') = (1-\rho_y^2)\,\mathrm{var}(t')$ from equation A.24), and it satisfies the PDE

$$\begin{aligned}\partial_{t'}\Phi^\dagger &= \widetilde{\mathcal{L}}_0^*\Phi = \left(\nu^2(1-\rho_y^2)\partial_{yy}\Phi^\dagger + \kappa\partial_y\left(y\Phi^\dagger\right)\right),\\ \Phi^\dagger(0,y) &= \delta(y-y_0).\end{aligned}$$

Next, we follow the principle in subsection 2.2 to define an approximation v_0^ϵ for which we track U^ϵ and keep the marginal density of $Y^{\epsilon,\dagger}$ exact, but approximate the density of X given $(Y^{\epsilon,\dagger}, U^\epsilon)$, and hence the joint density. Therefore, we consider the equation

$$\begin{aligned}\mathrm{d}v_0^\epsilon &= \frac{1}{\epsilon}\left(\nu^2(1-\rho_y^2)\partial_{yy}v_0^\epsilon + \kappa\partial_y\left(yv_0^\epsilon\right)\right)\mathrm{d}t\\ &\quad + \left(\frac{1}{2}\langle\sigma^2\rangle\partial_{xx}v_0^\epsilon - \partial_x(\mu v_0^\epsilon)\right)\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_0^\epsilon\,\mathrm{d}W_t^x\\ &= \left(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \langle\widetilde{\mathcal{L}}_2^*\rangle\right)v_0^\epsilon\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_0^\epsilon\,\mathrm{d}W_t^x.\end{aligned} \tag{3.8}$$

Moreover, (3.8) has the closed-form solution

$$v_0^\epsilon(t,x,y) = \Psi(t,x)\Phi^\dagger(t/\epsilon,y), \tag{3.9}$$

where $\Psi(t,x)$ is defined as in (3.5), and $\Phi^\dagger(t,y)$ is defined as in (3.6). Note that v_0^ϵ 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 $\mathbb{E}\big[P(v_0^\epsilon) - P(v^\epsilon)\big] \to 0$ as $\epsilon\to 0$ for P some linear functionals.

We note the analogy of (3.9) with (2.14), where Ψ satisfies the SPDE (3.3) and f_0 the PDE (2.10). Due to the constant coefficients of $\langle\widetilde{\mathcal{L}}_2^*\rangle$, the solution to (3.3) can be expressed as a solution to (2.10), but with the x argument shifted by the Brownian driver and the variance reduced appropriately to account for the second order Itô term. The separable structure is retained at zeroth order.

3.3. Correction terms. The goal of this section is to construct successively better approximations. We extend the procedure from subsection 2.2 from the KFE to the Zakai SPDE.

We first derive an inhomogeneous SPDE for $v^\epsilon - v_0^\epsilon$, by taking the difference between (3.2) and (3.8),

$$\begin{aligned}\mathrm{d}(v^\epsilon - v_0^\epsilon) &= \left(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \widetilde{\mathcal{L}}_2^*\right)(v^\epsilon - v_0^\epsilon)\,\mathrm{d}t - \rho_x\sigma(y+U_t^\epsilon)\partial_x(v^\epsilon - v_0^\epsilon)\,\mathrm{d}W_t^x\\ &\quad - \frac{1}{2}\left(\langle\sigma^2\rangle - \sigma^2(y+U_t^\epsilon)\right)\partial_{xx}v_0^\epsilon\,\mathrm{d}t + \rho_x\left(\langle\sigma\rangle - \sigma(y+U_t^\epsilon)\right)\partial_x v_0^\epsilon\,\mathrm{d}W_t^x,\\ (v^\epsilon - v_0^\epsilon)(0,x,y) &= 0.\end{aligned} \tag{3.10}$$

Intuitively, the effect of the terms in the second line in (3.10) is expected to be small for small ϵ, as the fast process U averages the terms involving $\sigma(\cdot+U_t^\epsilon)$ and $\sigma^2(\cdot+U_t^\epsilon)$ over the stationary distribution of U^ϵ, 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

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$Y^{\epsilon,\dagger}$; the combined effect is an averaging over the stationary distribution of Y^ϵ over timescales of order 1, so that the right-hand side, and hence the solution $v^\epsilon - v_0^\epsilon$, will be small.

Similar to subsection 2.2, we consider (3.10), which describes the exact error, and approximate $\mathcal{L}_2^*$ by $\langle\widetilde{\mathcal{L}}_2^*\rangle$ and σ by $\langle\sigma\rangle$ in the first line,

(3.11)
$$\begin{aligned} \mathrm{d}v_1^\epsilon = &\left(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \langle\widetilde{\mathcal{L}}_2^*\rangle\right) v_1^\epsilon\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_1^\epsilon\,\mathrm{d}W_t^x \\ &- \frac{1}{2}\left(\langle\sigma^2\rangle - \sigma^2(y+U_t^\epsilon)\right)\partial_{xx}v_0^\epsilon\,\mathrm{d}t + \rho_x\big(\langle\sigma\rangle - \sigma(y+U_t^\epsilon)\big)\partial_x v_0^\epsilon\,\mathrm{d}W_t^x, \\ v_1^\epsilon(0,x,y) = &\,0. \end{aligned}$$

We conjecture that v_1^ϵ is the leading order approximation to $v^\epsilon - v_0^\epsilon$, which we verify numerically later, and will therefore consider $v_0^\epsilon + v_1^\epsilon$ as an approximation to v^ϵ.

Given v_0^ϵ in (3.8), and v_1^ϵ in (3.11), we can recursively find higher order corrections as follows:

(3.12)
$$\begin{aligned} \mathrm{d}v_{n+1}^\epsilon = &\left(\frac{1}{\epsilon}\widetilde{\mathcal{L}}_0^* + \langle\widetilde{\mathcal{L}}_2^*\rangle\right) v_{n+1}^\epsilon\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_{n+1}^\epsilon\,\mathrm{d}W_t^x \\ &- \frac{1}{2}\Big(\langle\sigma^2\rangle - \sigma^2(y+U_t^\epsilon)\Big)\partial_{xx}v_n^\epsilon\,\mathrm{d}t + \rho_x\big(\langle\sigma\rangle - \sigma(y+U_t^\epsilon)\big)\partial_x v_n^\epsilon\,\mathrm{d}W_t^x, \\ v_{n+1}^\epsilon(0,x,y) = &\,0. \end{aligned}$$

3.4. Approximation to the marginal density of X. In practical applications (see, e.g., section 5), one is often only interested in the marginal law of X, characterized by the marginal density

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

We will show here how the expansion in subsection 3.3 allows a computationally more efficient approximation to v^x than solving the original two-dimensional SPDE directly numerically and integrating out y.

Considering (3.11) directly, the solution v_0^ϵ depends on x and y. If we were to drop $\widetilde{\mathcal{L}}_0^*$ in (3.11), justified after the initial transient, the SPDE is parametrized by y, so is essentially two-dimensional if the solution for all x and y is needed.

However, integrating (3.8) 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 (3.5). It is worth emphasizing that v_0^ϵ, given by (3.8), depends on ϵ through $\Phi^\dagger$ only, which is normalized, and therefore the y-integral of v_0^ϵ is independent of ϵ.

We now want to use (3.11) to define a correction to v_0^x by integrating over 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

$$\int_{-\infty}^{\infty} \widetilde{\mathcal{L}}_0^* v_1^\epsilon\,\mathrm{d}y = 0.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Moreover, from (3.9),

$$\int_{-\infty}^{\infty} \sigma^2(y+U_t^\epsilon)\partial_{xx}v_0^\epsilon(t,x,y)\,\mathrm{d}y = \partial_{xx}\Psi(t,x)\int_{-\infty}^{\infty}\sigma^2(y+U_t^\epsilon)\Phi^\dagger(t,y)\,\mathrm{d}y.$$

We get from Theorem 6.1 in [25] that

$$\left|\int_{-\infty}^{\infty}\sigma^2(y+U_t^\epsilon)\Phi^\dagger(t,y)\,\mathrm{d}y - \langle\sigma^2(\cdot+U_t^\epsilon)\rangle_\dagger\right| \le B(x_0,y_0,U_t^\epsilon)\beta^{t/\epsilon} \tag{3.13}$$

for some $0<\beta<1$. We assume in the following that we can find $B<\infty$ above independent of U_t^ϵ (which is possible by retracing the proof in [25] using suitable assumptions on σ).

Then for $t \ge c\epsilon^r$ for fixed $c>0$ and some $0<r<1$, the right-hand side of (3.13) is of order ϵ^p for any $p>0$, so that we get

$$\int_{-\infty}^{\infty}\sigma^2(y+U_t)\partial_{xx}v_0^\epsilon(t,x,y)\,\mathrm{d}y = \langle\sigma^2(\cdot+U_t^\epsilon)\rangle_\dagger\partial_{xx}\Psi(t,x) + o(\epsilon^p) \qquad \forall p>0.$$

Similarly,

$$\int_{-\infty}^{\infty}\sigma(y+U_t^\epsilon)\partial_x v_0^\epsilon(t,x,y)\,\mathrm{d}y = \langle\sigma(\cdot+U_t)\rangle_\dagger\partial_x\Psi(t,x) + o(\epsilon^p) \qquad \forall p>0.$$

Integrating (3.11) over y, replacing the y-integral of v_1^ϵ by $v_1^{x,\epsilon}$ in the first line and making the approximations above in the second line, yields the SPDE

$$\begin{aligned}\mathrm{d}v_1^{x,\epsilon} &= \langle\widetilde{\mathcal{L}}_2^*\rangle v_1^{x,\epsilon}\,\mathrm{d}t - \rho_x\langle\sigma\rangle\partial_x v_1^{x,\epsilon}\,\mathrm{d}W_t^x \\ &\quad - \frac{1}{2}\left(\langle\sigma^2\rangle - \langle\sigma^2(\cdot+U_t^\epsilon)\rangle_\dagger\right)\partial_{xx}\Psi(t,x)\,\mathrm{d}t \\ &\quad + \rho_x\left(\langle\sigma\rangle - \langle\sigma(\cdot+U_t^\epsilon)\rangle_\dagger\right)\partial_x\Psi(t,x)\,\mathrm{d}W_t^x, \\ v_1^{x,\epsilon}(0,x) &= 0.\end{aligned} \tag{3.14}$$

We will refer to $v_0^x(t,x)+v_1^{x,\epsilon}(t,x)$ as a first order approximation to $v^{x,\epsilon}(t,x)$.[4]

A rigorous analysis of the convergence and its orders is beyond the scope of this paper, and we restrict ourselves to a computational analysis in section 4.

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 (3.2), repeated here for convenience:

$$\begin{aligned}\mathrm{d}v^\epsilon &= \frac{1}{\epsilon}\Big(\nu^2(1-\rho_y^2)\partial_{yy}v^\epsilon + \kappa\partial_y(yv^\epsilon)\Big)\,\mathrm{d}t \\ &\quad + \left(\frac{1}{2}\sigma^2\big(y+U_t^\epsilon\big)\partial_{xx}v^\epsilon - \mu\,\partial_x v^\epsilon\right)\mathrm{d}t - \rho_x\sigma\big(y+U_t^\epsilon\big)\partial_x v^\epsilon\,\mathrm{d}W_t^x,\end{aligned} \tag{4.1}$$

$$v^\epsilon(0,x,y) = \delta(x-x_0)\otimes\delta(y-y_0), \tag{4.2}$$

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

[4]For $t \le c\epsilon^r$, the difference between averaging over $\Phi^\dagger$ and its limit is not small; however, this only applies over a time interval of length $O(\epsilon^r)$, so that we conjecture without proof that the effect on the solution is of order $O(\epsilon^r)$ for any $0<r<1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

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. There is, naturally, a vast body of literature on numerical methods designed specifically for multiscale problems. For a survey of a general approach of blending "micro" and "macro" models, the heterogeneous multiscale method, we refer the reader to [1]. Our focus here is not on the design of such a method but rather to approximate the multiscale equation by our expansion into terms that can be determined by solving single scale problems.

We will also be solving the SPDE (3.10) to determine the error of the zero-order approximation, and the SPDE (3.11) for the first order correction, by straightforward modifications of the scheme for (3.2).

The zero-order marginal approximation in x, v_0^x, is given in analytic form by Ψ in (3.5), while the correction term $v_1^{x,\epsilon}$ is given by (3.14) and can be found by a one-dimensional SPDE scheme.

We apply a Milstein alternating direction implicit (ADI) scheme to the SPDE (4.1), and an Euler scheme to the SDE (4.2), 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 briefly illustrate the difficulties in simulating the fast OU process (4.2) and review pertinent literature. We then propose a simple solution for our purposes.

A direct Euler–Maruyama approximation of (U_t^ϵ) with step size k is generated by

$$\begin{aligned} \widehat{U}_n &= \widehat{U}_{n-1} - \frac{\kappa}{\epsilon}\, k\, \widehat{U}_{n-1} + \frac{\nu\sqrt{2}}{\epsilon}\rho_y\Big(W^y_{t_n} - W^y_{t_{n-1}}\Big), \\ \widehat{U}_0 &= 0, \end{aligned} \tag{4.3}$$

where $n = 1, 2, \ldots$, $t_n - t_{n-1} = k > 0$ constant. Mimicking the derivation of the well-known solution of (4.2), we can write

$$\widehat{U}_n = \Big(1 - \frac{\kappa}{\epsilon}\, k\Big)^n \widehat{U}_0 + \frac{\nu\sqrt{2}}{\epsilon}\rho_y \sum_{j=0}^{n-1} \Big(1 - \frac{\kappa}{\epsilon}\, k\Big)^{n-j} \Big(W^y_{t_j} - W^y_{t_{j-1}}\Big).$$

It is clear that for fixed k and small enough ϵ, in particular where $\epsilon \ll \kappa k$, the scheme does not produce a suitable approximation to U^ϵ even in a qualitative sense. This can be partially mitigated by an implicit version of the scheme, where the "integrating factors" would be replaced by $(1 + \kappa k/\epsilon)^{-n}$, which correctly vanish as $\epsilon \to 0$ as in the continuous equation. Schemes where the limits $k \to 0$ and $\epsilon \to 0$ are exchangeable have been termed "asymptotic preserving"; see [4] for a discussion. In this special case, the solution to (4.2) in terms of an integral over the driving Brownian motion enables an exact weak simulation scheme as in [4], which could then be used in the simulation of the slow component.

In the present context of the Zakai equation, e.g., in a filtering application or a large pool model with common noise, one may be interested in the solution conditional on the path of (W^x, W^y), which requires a strong approximation.

Instead, we simulate the OU process (4.2) by the Euler–Maruyama scheme with timestep $k\epsilon$. It was found empirically in [8] that the simulated process then has a strong error independent of ϵ. This does not change the total computational effort significantly since the cost of simulating U^ϵ is typically much smaller than solving the SPDE for v^ϵ.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

In practice, we first generate independent standard normal random variables $Z_{m,x}$ for $m = 1, 2, \ldots, M_\epsilon = N/\epsilon$, where we assume for simplicity $1/\epsilon \in \mathbb{Z}$. We generate

$$\widehat{U}_m = \widehat{U}_{m-1} - \kappa\, k\, \widehat{U}_{m-1} + \nu\sqrt{2}\rho_y \sqrt{k} Z_{m-1,y},$$

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

$$U^n = \widehat{U}_{n\epsilon^{-1}}, \qquad n = 1, 2, \ldots, N, \tag{4.4}$$

as the approximation to the process U^ϵ at time $t = nk$, i.e., at the nth 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

$$W^x_{nk} - W^x_{(n-1)k} = \sqrt{k} Z_{n,x},$$

where $Z_{n,x}$ is chosen to have the correct correlation ρ with $\sum_{i=1}^{\epsilon^{-1}} Z_{\epsilon^{-1}(n-1)+i,y}$.

4.2. Approximation of the one-dimensional SPDE. The scheme we use for the marginal SPDE for $v_1^{\epsilon^x(t,x)}$ from (3.14) is an adaptation of the schemes proposed in [28, 29]. 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 $v_1^{\epsilon^x(t_n,x_i)}$ by

$$\begin{aligned}
&\left(I - \frac{\langle\sigma^2\rangle}{2}\frac{k}{h_x^2}D_{xx} + \mu\frac{k}{2h_x}D_x\right) V^{n+1}_{1,x} \\
&\quad = \left(I - \rho_x\langle\sigma\rangle\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x + \rho_x^2\langle\sigma\rangle^2\frac{k(Z^2_{n,x}-1)}{2h_x^2}D_{xx}\right) V^n_{1,x} \\
&\qquad - \frac{1}{2}\Big(\langle\sigma^2\rangle - \langle\sigma^2(\cdot + U^n)\rangle_\dagger\Big)k\Psi_{xx}(nk,\mathcal{X}) + \rho_x\Big(\langle\sigma\rangle - \langle\sigma(\cdot+U^n)\rangle_\dagger\Big)\Psi_x(nk,\mathcal{X})\sqrt{k}Z_{n,x} \\
&\qquad + \frac{1}{2}\rho_x^2\Big(\langle\sigma\rangle - \langle\sigma(\cdot+U^n)\rangle_\dagger\Big)^2\Psi_{xx}(nk,\mathcal{X})k(Z^2_{n,x}-1),
\end{aligned} \tag{4.5}$$

where Ψ is given by (3.5), $\Psi_x(nk,\mathcal{X})$ and $\Psi_{xx}(nk,\mathcal{X})$ are the values of the derivatives applied on the mesh X, i.e., $\Psi_x(nk,\mathcal{X})$ is the vector of values $\Psi_x(nk,x_i)$, the operators D_x and D_{xx} are defined as the standard finite difference matrices with

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

and U^n is found from (4.4).

The scheme is semi-implicit to ensure stability in L_2 irrespective of the step sizes. The terms in the first line of (4.5) hence come from an implicit finite difference discretization 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 (3.2) 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.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

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 $V_{i,j}^n$ be the approximation to $v(nk, ih_x, jh_y)$, $n = 1, \dots, N$, $i, j \in \mathbb{Z}$.

Adapting the schemes from [30] to our setting, we have

(4.6)
$$\left(I - \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy} - \frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\right)\left(I - \frac{\sigma^2(\mathcal{Y}+U^{n+1})}{2}\frac{k}{h_x^2}D_{xx} + \mu\frac{k}{2h_x}D_x\right)V^{n+1}$$
$$= \left(I - \rho_x\sigma(\mathcal{Y}+U^n)\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x + \rho_x^2\sigma^2(\mathcal{Y}+U^n)\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\right)V^n,$$

where U^n is from (4.4), and $D_\cdot$, $D_{\cdot\cdot}$ are standard finite difference matrices defined by

$$(D_xV)_{i,j} = V_{i+1,j} - V_{i-1,j}, \qquad (D_yV)_{i,j} = V_{i,j+1} - V_{i,j-1},$$
$$(D_{xx}V)_{i,j} = V_{i+1,j} - 2V_{i,j} + V_{i-1,j}, \qquad (D_{yy}V)_{i,j} = V_{i,j+1} - 2V_{i,j} + V_{i,j-1}.$$

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 (3.2).

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 factorization allows an efficient solution of the implicit scheme by a sequence of tridiagonal systems.

PROPOSITION 4.1. *Provided that*

$$|\rho_x| \le \frac{1}{\sqrt[4]{2}}\frac{\inf_{y\in\mathbb{R}}\sigma(y)}{\sup_{y\in\mathbb{R}}\sigma(y)}, \tag{4.7}$$

the scheme (4.6) *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*

$$\mathbb{E}|V^n|_2^2 \le \exp\left(\frac{T\epsilon y_{\max}^2}{\nu^2(1-\rho_y^2)}\right)|V^0|_2^2. \tag{4.8}$$

Proof. We consider the discrete-continuous Fourier pair

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

By insertion and standard algebraic manipulations,

$$\left((I - kL_y)\widetilde{V}^{n+1}(\omega)\right)_j = \frac{\widetilde{L}_{x,j}^{\mathrm{ex},n}}{\widetilde{L}_{x,j}^{\mathrm{im},n}}\widetilde{V}_j^n(\omega), \tag{4.9}$$

where

$$\left(L_y\widetilde{V}^n\right)_j = \frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{1}{h_y^2}\left(\widetilde{V}_{j+1}^n - 2\widetilde{V}_j^n + \widetilde{V}_{j-1}^n\right) - \frac{\kappa}{\epsilon}\frac{1}{2h_y}\left(y_{j+1}\widetilde{V}_{j+1}^n - y_{j-1}\widetilde{V}_{j-1}^n\right),$$
$$\widetilde{L}_{x,j}^{\mathrm{ex},n} = 1 - \mathrm{i}\,\rho_x\sigma(y_j+U^n)\frac{\sqrt{k}}{h_x}\sin(\omega)Z_n - 2\rho_x^2\frac{k}{h_x^2}\sin^2(\omega/2)(Z_n^2-1),$$
$$\widetilde{L}_{x,j}^{\mathrm{im},n} = 1 + 2\sigma^2(y_j+U^{n+1})\frac{k}{h_x^2}\sin^2(\omega/2) + \mathrm{i}\frac{k}{h_x}\sin(\omega).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Multiplying the left-hand side of (4.9) by $\widetilde{V}_j^{n+1,*}$, with $*$ denoting the complex conjugate, summing over j, and carrying out summation by parts, we obtain

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

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

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

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

$$\begin{aligned}
\sum_j \left((I - kL_y)\widetilde{V}^{n+1}\right)_j \widetilde{V}_j^{n+1,*} &\geq \sum_j |\widetilde{V}_j^{n+1}|^2 - k \sum_j \frac{\epsilon |y_j \widetilde{V}_j^{n+1}|^2}{2\nu^2(1-\rho_y^2)} \\
&\geq \left(1 - k\frac{\epsilon\, y_{\max}^2}{2\nu^2(1-\rho_y^2)}\right) \sum_j |\widetilde{V}_j^{n+1}|^2. \qquad (4.10)
\end{aligned}$$

To estimate the right-hand side of (4.9), we have

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

Hence, assuming (4.7), we get

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

On the right-hand side of (4.9), we have

$$\sum_j \frac{\widetilde{L}_{x,j}^{\mathrm{ex},n}}{\widetilde{L}_{x,j}^{\mathrm{im},n}} \widetilde{V}_j^n \widetilde{V}_j^{n+1,*} \leq \frac{1}{2} \sum_j \left|\frac{\widetilde{L}_{x,j}^{\mathrm{ex},n}}{\widetilde{L}_{x,j}^{\mathrm{im},n}}\right|^2 |\widetilde{V}_j^n|^2 + \frac{1}{2} \sum_j |\widetilde{V}_j^{n+1}|^2. \qquad (4.11)$$

From (4.9), (4.10), and (4.11),

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

Rearranging, using Parseval's identity, and $1 - \delta/2 \geq \exp(-\delta)$ for any $0 \leq \delta \leq 1$, we obtain (4.8) by induction over n. ∎

The proof of Proposition 4.1 is more complicated than earlier results in [29, 30], in that

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

- we seek stability with a constant independent of ϵ,
- the coefficients depend on the fast process Y^ϵ, 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 4.2. The condition $k\epsilon y_{\max}^2 \le \nu^2(1-\rho_y^2)$ is not practically restrictive as long as ρ_y is bounded away from 1 and -1, because k, the time step size, will be chosen sufficiently small and we are moreover interested in the regime of small ϵ, while reasonable truncation bounds $y_{\max}$ on the variance process, as well as the parameter ν^2, are of order $O(1)$.

We also note that (4.7) could be replaced by the simpler condition $\sqrt{2}\rho_x^2 \le 1$ if the term $\sigma^2(\mathcal{Y}+U^{n+1})$ on the left-hand side of (4.6) 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^ϵ, which satisfies (3.8), we denote the corresponding numerical solution as V_0, with the scheme

$$\begin{aligned}(4.12)\quad &\left(I-\frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy}-\frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\right)\left(I-\frac{\langle\sigma^2\rangle}{2}\frac{k}{h_x^2}D_{xx}+\mu\frac{k}{2h_x}D_x\right)V_0^{n+1}\\ &\quad=\left(I-\rho_x\langle\sigma\rangle\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x+\rho_x^2\langle\sigma\rangle^2\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\right)V_0^n,\end{aligned}$$

with notation as earlier. Note that the closed-form solution to V_0 is (3.9), 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 $v^\epsilon - v_0^\epsilon$ from (3.10). Denoting the solution as $\widetilde{V}_0$, we have the scheme

$$\begin{aligned}(4.13)\quad &\left(I-\frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy}-\frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\right)\left(I-\frac{\sigma^2(\mathcal{Y}+U^{n+1})}{2}\frac{k}{h_x^2}D_{xx}+\mu\frac{k}{2h_x}D_x\right)\widetilde{V}_0^{n+1}\\ &=\left(I-\rho_x\sigma(\mathcal{Y}+U^n)\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x+\rho_x^2\sigma^2(Y+U^n)\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\right)\widetilde{V}_0^{n+1}\\ &\quad-\frac{1}{2}\Big(\langle\sigma^2\rangle-\sigma^2(\mathcal{Y}+U^n)\Big)k\big(\partial_{xx}V_0^n\big)+\rho_x\Big(\langle\sigma\rangle-\sigma(\mathcal{Y}+U^n)\Big)\big(\partial_xV_0^n\big)\sqrt{k}Z_{n,x}\\ &\quad+\frac{1}{2}\rho_x^2\Big(\langle\sigma\rangle-\sigma(\mathcal{Y}+U^n)\Big)^2\big(\partial_{xx}V_0^n\big)k(Z_{n,x}^2-1),\end{aligned}$$

where we use the analytic solution (3.9) to compute $\partial_xV_0^n=(\partial_xv_0^\epsilon(t_n,x_i,y_j))_{i,j}$, 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 v_1^ϵ from (3.11), denoted by V_1, is

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

(4.14)
$$\begin{aligned}
&\left(I-\frac{\nu^2(1-\rho_y^2)}{\epsilon}\frac{k}{h_y^2}D_{yy}-\frac{\kappa}{\epsilon}\frac{k}{2h_y}D_y\mathcal{Y}\right)\left(I-\frac{\langle\sigma^2\rangle}{2}\frac{k}{h_x^2}D_{xx}+\mu\frac{k}{2h_x}D_x\right)V_1^{n+1}\\
&\quad=\left(I-\rho_x\langle\sigma\rangle\frac{\sqrt{k}Z_{n,x}}{2h_x}D_x+\rho_x^2\langle\sigma\rangle^2\frac{k(Z_{n,x}^2-1)}{2h_x^2}D_{xx}\right)V_1^n\\
&\qquad-\frac{1}{2}\Big(\langle\sigma^2\rangle-\sigma^2(\mathcal{Y}+U^n)\Big)k\big(\partial_{xx}V_0^n\big)+\rho_x\Big(\langle\sigma\rangle-\sigma(\mathcal{Y}+U^n)\Big)\big(\partial_xV_0^n\big)\sqrt{k}Z_{n,x}\\
&\qquad+\frac{1}{2}\rho_x^2\Big(\langle\sigma\rangle-\sigma(\mathcal{Y}+U^n)\Big)^2\big(\partial_{xx}V_0^n\big)k(Z_{n,x}^2-1),
\end{aligned}$$

with zero initial condition, $V_1^0=0$. We also use the analytic solution for ∂_xV_0 and $\partial_{xx}V_0$ in the scheme (4.14).

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 (3.2) by way of numerical tests. For illustration, we use the following Gaussian distribution as initial condition:

$$v^\epsilon(0,x,y)=\Psi\big(t_0,x\ ;W_{t_0}^x=0\big)\Phi^\dagger(t_0,y), \tag{5.1}$$

where $t_0>0$ is a fixed constant, and Ψ, Φ are defined as in (3.5) and (3.6), respectively. We choose this smooth initial condition so that when we numerically approximate v_1^ϵ in (3.11), the approximation is stable when taking the partial derivatives of v_0^ϵ.[5]

Moreover, we can still obtain a closed-form solution to v_0^ϵ in (3.8):

$$v_0^\epsilon(T,x,y)=\Psi\big(t_0+T,x\ ;W_{t_0}^x=0\big)\Phi^\dagger(t_0+T,y). \tag{5.2}$$

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

(5.3)
$$\langle\sigma\rangle=\int_{-\infty}^{\infty}\frac{\kappa}{\nu\sqrt{2\pi}}\exp\left(\alpha x-\frac{\kappa x^2}{2\nu^2}\right)\mathrm{d}x=\exp\left(\frac{\alpha^2\nu^2}{2\kappa}\right),\qquad\langle\sigma^2\rangle=\exp\left(\frac{4\alpha^2\nu^2}{2\kappa}\right),$$

and

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

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 (5.1) as 0.2. We then vary ϵ and estimate the contribution of v_0^ϵ and v_1^ϵ to expected

[5]If we initialized with Dirac data, further stabilization may be needed, and the analysis in ℓ_2 would not carry over; see the Fourier analysis in [18, 30] for schemes for simpler SPDEs with Dirac initial data.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

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 $[-10,10]\times[-10,10]$ and set zero Dirichlet boundary conditions. The domain was chosen large enough experimentally so that the effect on the results from further enlarging the domain was found negligible for the tested parameter values.

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

$$P_T(v^\epsilon) = \int_0^\infty \int_{-\infty}^\infty v^\epsilon(T,x,y)\,\mathrm{d}y\,\mathrm{d}x, \qquad T<\infty. \tag{5.4}$$

There are two motivations for this. First, convergence in distribution of $\mathbb{P}[X_T^\epsilon \in \mathcal{I}|W^x,W^y]$ for intervals $\mathcal{I}$ is theoretically supported by [21] (albeit for the process with absorption at $x=0$). Second, P_T models the aggregate loss in credit portfolio models as, e.g., in [6, 5, 19, 21], and is therefore of practical interest.

To approximate $P_T(v^\epsilon)$ in (5.4), we use the trapezoidal rule for the numerical integration.

5.1. Weak convergence of $P_T(v_0^\epsilon)$. We first analyze the numerical convergence of $\mathbb{E}[P_T(v^\epsilon)-P_T(v_0^\epsilon)]$, where P_T is the functional from (5.4). Since P_T is linear, we have $P_T(v^\epsilon)-P_T(v_0^\epsilon)=P_T(v^\epsilon-v_0^\epsilon)$. We use the scheme (4.13) to directly approximate $v^\epsilon-v_0^\epsilon$, and estimate $\mathbb{E}[P_T(v^\epsilon-v_0^\epsilon)]$ by standard Monte Carlo sampling as detailed below. While simulating $v-v_0^\epsilon$ from (3.10), compared to simulating v^ϵ in (3.2) and v_0^ϵ in (3.8) separately, leads to the same $\mathbb{E}[P_T(v^\epsilon-v_0^\epsilon)]$ for $h_x,h_y,k\to 0$, the former choice has computational savings and is hence faster by a constant factor.[6]

Given M Brownian paths, the Monte Carlo estimator for $\mathbb{E}[P_T(v^\epsilon-v_0^\epsilon)]$ 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 $v^\epsilon-v_0^\epsilon$ for the ith path of (W^x,W^y) in (4.13).

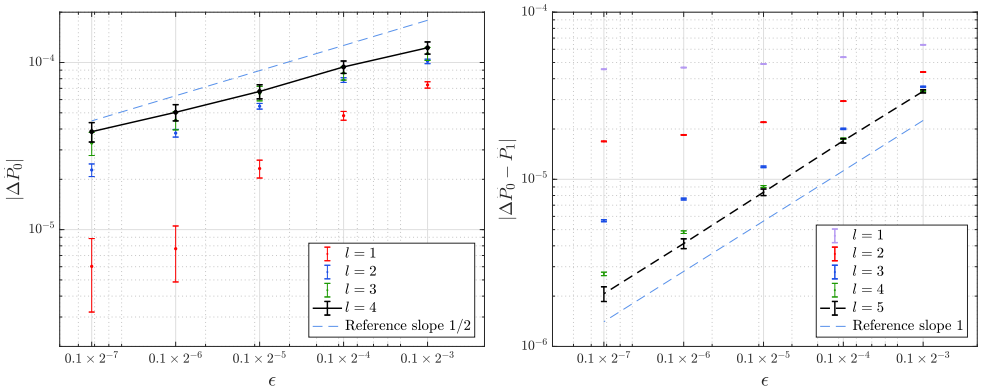
For each ϵ, the numerical error between $\Delta\widehat{P}_0$ and $\mathbb{E}[P_T(v^\epsilon-v_0^\epsilon)]$ consists of discretization error in h and k (bias), and Monte Carlo noise (variance). If $h_x,h_y,k\to 0$, and $M\to\infty$, $\Delta\widehat{P}_0$ is expected to converge to $\mathbb{E}[P_T(v^\epsilon-v_0^\epsilon)]$, and we treat $\Delta\widehat{P}_0$ as the weak error in ϵ for small h_x,h_y,k and large M.

Figure 1(a) shows the convergence to zero of $\Delta\widehat{P}_0$, with $\epsilon = 0.1\times 2^{-3},\ 0.1\times 2^{-4},\dots,0.1\times 2^{-7}$, 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 $\mathbb{E}[P_T(v^\epsilon-v_0^\epsilon)]$, 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). Indeed, linear regression gives a slope of 0.4237, with a 95% confidence interval of $[0.3948, 0.4526]$, which suggests that the empirically observed order is lower than 0.5 in a statistically significant sense. A plausible reason is that in this case for $l=4$ the finite difference error in $h_x=h_y$ and k is not small enough to be neglected, which is supported by the observation that the

[6] Using the same Brownian paths for v^ϵ and v_0^ϵ leads to a variance reduction, and fewer paths are needed.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy



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

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

results for slightly coarser meshes (represented by different colors in the figure) differ noticeably. We conjecture that the asymptotic weak error $\mathbb{E}[P_T(v - v_0^\epsilon)]$ is $O(\sqrt{\epsilon})$, which is also in line with the observations in section 5.3.

5.2. Weak convergence of first order approximation. To verify $P_T(v_1^\epsilon)$ is indeed the leading order approximation to $P_T(v^\epsilon) - P_T(v_0^\epsilon)$, we further exhibit $\Delta \widehat{P}_0 - \widehat{P}_1$, which is the Monte Carlo estimator for $P_T(v^\epsilon - v_0^\epsilon - v_1^\epsilon)$. 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 $P_T(v^\epsilon - v_0^\epsilon)$, with $P_T(v^\epsilon - v_0^\epsilon - v_1^\epsilon) = O(\epsilon)$.

5.3. Convergence of $\mathbb{E}[P_T(v_1^\epsilon)]$. We now analyze $\mathbb{E}[P_T(v_1^\epsilon)]$, where v_1^ϵ satisfies the SPDE (3.11), and compare it to $P_T^x(v_1^{x,\epsilon})$, where $v_1^{x,\epsilon}$ is the solution to the marginal SPDE (3.14), 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 that these two values should be approximately the same.

From the derivation of the SPDE for $v_1^{x,\epsilon}$ in section 3.4, when we integrate over the y-dimension, we replace $\Phi^Y(T,y)$ in the analytic solution (3.9) 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 $P_T(v_1^{x,\epsilon})$, 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^ϵ in the schemes (4.13) and (4.14).

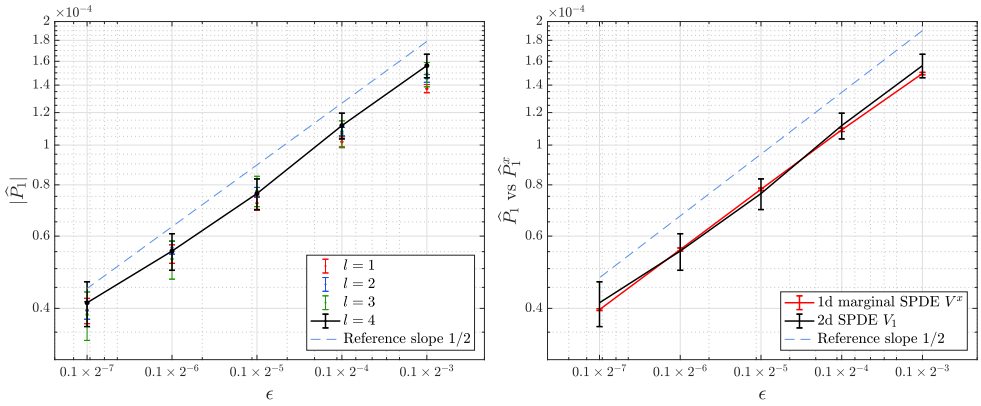
Given M Brownian paths, we define Monte Carlo estimators for $\mathbb{E}[P_T(v_1^\epsilon)]$ and $\mathbb{E}[P_T(v_1^{x,\epsilon})]$ by

$$\widehat{P}_1 = \frac{1}{M}\sum_{i=1}^M P_T(V_1^{(i)}), \qquad \widehat{P}_1^x = \frac{1}{M}\sum_{i=1}^M P_T\big((V_1^x)^{(i)}\big), \tag{5.5}$$

where V_1 obeys the scheme (4.14), and V_1^x the scheme (4.5).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy



(a) $\widehat{P}_1$ defined in (5.5). (b) Comparing $\widehat{P}_1$ with $\widehat{P}_1^x$ from (5.5).

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

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 (3.14), 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., [29]), and only note that, unlike approximating $\mathbb{E}[P_T(v^\epsilon)]$ by a standard Monte Carlo method, and discretizing with mesh size $(h_x = h_y =)\ h_l = h_0 \times 2^{-l}$ and timestep $(k =)\ k_l = k_0 \times 4^{-l}$, approximating by MLMC requires a good coupling between fine path and coarse path. Therefore, when we apply MLMC to estimate $\mathbb{E}[P_T(v_1^{x,\epsilon})]$, the timestep is set as $k_l = k_0 \times \epsilon \times 4^{-l}$, proportional to ϵ.

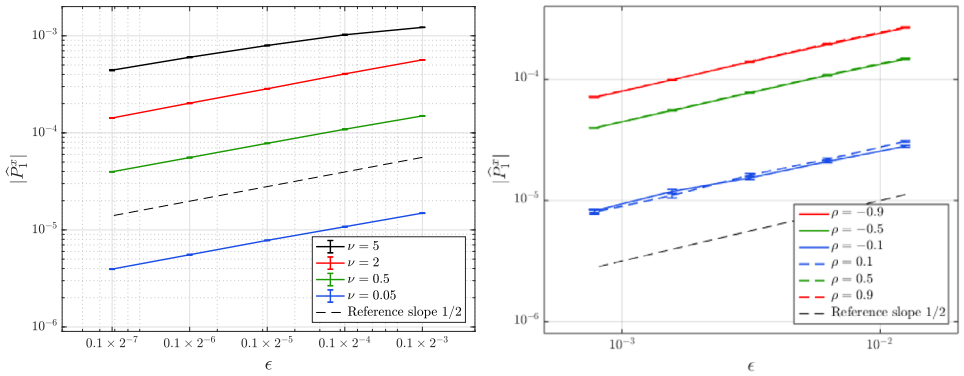
Linear regression yields that the fitted slope for $\widehat{P}_1$ is 0.4855, and for $\widehat{P}_1^x$ it is 0.4792. This is consistent with the finding from the previous section that inclusion of the first order term approximately cancels the error of the zero order approximation, which is of order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

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



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

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

and vice versa. This effect is similar to the asymptotic expansion of the backward PDE for the stochastic volatility model [12], where the leading order correction term is proportional to the correlation between the two Brownian motions involved.

Appendix A. In this appendix we give details of the results summarized in section 2, also pointing out where our analysis differs from that in [23]. We derive an expansion for the transition density function of a stochastic volatility process (X, Y) satisfying (1.1) with $g = \nu\sqrt{2}$ for constant ν, as $\epsilon \to 0^+$. 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 OU process Y); to avoid technical complications we make the undemanding assumption that $\sigma(y)\mathrm{e}^{-ay^2} \to 0$ as $y \to \pm\infty$ for all $a > 0$. This ensures that all averages of $\sigma(y)$ and $\Sigma(y)$ that we encounter below exist.

A.1. Preliminaries. The transition density $p^\epsilon(t_0, x_0, y_0; t, x, y)$ of (X, Y) satisfies the FKE

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

with the initial condition

$$p^\epsilon(t_0, x_0, y_0; t_0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0). \tag{A.2}$$

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

$$\mathcal{L}_0 = \nu^2\partial_{yy}(\,\cdot\,) + \kappa(m - y)\partial_y(\,\cdot\,), \qquad \mathcal{L}_0^* = \nu^2\partial_{yy}(\,\cdot\,) - \kappa\partial_y\left((m - y)\cdot\,\right), \tag{A.3}$$

$$\mathcal{L}_1 = \rho_{xy}\nu\sqrt{2}\sigma(y)\partial_{xy}(\,\cdot\,), \qquad \mathcal{L}_1^* = \rho_{xy}\nu\sqrt{2}\partial_x\partial_y\left(\sigma(y)\cdot\,\right), \tag{A.4}$$

$$\mathcal{L}_2 = \frac{1}{2}\sigma^2(y)\partial_{xx}(\,\cdot\,) + \mu(x, y)\partial_x(\,\cdot\,), \qquad \mathcal{L}_2^* = \frac{1}{2}\sigma^2(y)\partial_{xx}(\,\cdot\,) - \partial_x\left(\mu(x, y)\cdot\,\right). \tag{A.5}$$

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

$$\Phi^Y(y) = \frac{1}{\sqrt{2\pi\nu^2/\kappa}}\mathrm{e}^{-\kappa(y-m)^2/2\nu^2}. \tag{A.6}$$

Averages of functions of y with respect to the measure $\Phi^Y(y)\,dy$ are denoted by $\langle\,\cdot\,\rangle = \langle\,\cdot\,, \Phi^Y\rangle$ for this average, where $\langle\,\cdot\,,\,\cdot\,\rangle$ is the usual inner product on $L^2(\mathbb{R})$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

We shall see many equations of the form

$$-\mathcal{L}_0^* u(y) = h(y), \qquad -\infty < y < \infty, \tag{A.7}$$

for some given h, needing solutions that are integrable on $\mathbb{R}$; this entails some considerations not needed by [23], who work in a compact setting. In fact, the only integrable solution of $-\mathcal{L}_0^* u(y) = 0$ are constant multiples of $\Phi^Y(y)$; the linearly independent solution is not integrable.[7] We 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 is 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 (A.7) only exist when the right-hand side satisfies the solvability condition of being orthogonal to (relevant, i.e., 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 (A.7) is $\langle 1, h\rangle = \int_{-\infty}^{\infty} h(y)\,\mathrm{d}y = 0$. When this is satisfied, the solution is given by

$$u(y) = -\Phi^Y(y) \int^y \frac{H(s)}{\nu^2 \Phi^Y(s)}\,\mathrm{d}s + c\Phi^Y(y), \tag{A.8}$$

where the constant c is arbitrary, and where $H(\cdot) = \int^{\cdot} h(s)\,\mathrm{d}s$.

We describe the spatial decomposition from large to small (in contrast with the summary in section 2).

A.2. Outer region: $t \gg O(\epsilon)$. We begin with timescales much longer than the mean-reversion time ϵ; the volatility is effectively sampled from its ergodic distribution, as is already seen from (1.4). This determines 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. 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

[7] This solution is

$$\Phi^Y(y) \int^y \frac{\mathrm{d}s}{\Phi^Y(s)},$$

and integration by parts shows that, as $y \to \infty$,

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

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

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

each have their own expansions

$$p_X^\epsilon(t,x) \sim p_{X,0} + \sqrt{\epsilon} p_{X,1} + \cdots, \qquad p_Y^\epsilon(t,y) \sim p_{Y,0} + \sqrt{\epsilon} p_{Y,1} + \cdots,$$

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 positive everywhere; 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 behavior.

Now substituting into (A.1) and equating coefficients of powers of ϵ leads immediately to

(A.9) At $O(1/\epsilon)$: $-\mathcal{L}_0^* p_0 = 0;$

(A.10) At $O(1/\sqrt{\epsilon})$: $-\mathcal{L}_0^* p_1 = \mathcal{L}_1^* p_0;$

(A.11) At $O(1)$: $-\mathcal{L}_0^* p_2 = -\partial_t p_0 + \mathcal{L}_2^* p_0 + \mathcal{L}_1^* p_1;$

(A.12) At $O(\sqrt{\epsilon})$: $-\mathcal{L}_0^* p_3 = -\partial_t p_1 + \mathcal{L}_2^* p_1 + \mathcal{L}_1^* p_2;$

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 $g_0(t,x) = 0$. We 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 (A.2); a separate boundary-layer analysis, given in subsection A.3, is needed to resolve this.

At $O(1/\sqrt{\epsilon})$. From (A.10), we have

$$-\mathcal{L}_0^* p_1 = \mathcal{L}_1^* p_0 = \rho_{xy}\nu\sqrt{2}\,(\partial_x f_0(t,x))\,\partial_y\left(\sigma(y)\Phi^Y(y)\right). \tag{A.13}$$

As the y dependence on the right-hand side integrates to zero, this equation does have an integrable solution, given by (A.8) as

$$p_1(t,x,y) = -(\rho_{xy}\sqrt{2}/\nu)\,(\partial_x f_0(t,x))\,\Sigma(y)\Phi^Y(y) + f_1(t,x)\Phi^Y(y), \tag{A.14}$$

where $\Sigma(y)$ is any antiderivative of $\sigma(y)$. Here $f_1(t,x)$ is again unknown.

At $O(1)$. Now we have

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

The terms in the last line, being derivatives of functions that vanish at $\pm\infty$, satisfy the solvability condition, and so the terms in the middle line integrate to zero, giving

$$\partial_t f_0 = \frac{1}{2}\langle\sigma^2(\cdot)\rangle\partial_{xx} f_0 - \partial_x\left(\langle\mu(x,\cdot)\rangle f_0\right) = \langle\mathcal{L}_2^*\rangle f_0(t,x), \tag{A.15}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

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

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

satisfied by 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 that the density function is, to leading order, a product, X and Y are independent at this order, suggesting that $f_0(t_0,x) = \delta(x - x_0)$, a fact which is confirmed by the boundary-layer analysis of the next subsection.

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

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

Once again, solvability requires that the right-hand side is orthogonal to 1 over $\mathbb{R}$; this lets us fix f_1. This is automatically true for the term $L_1^* p_2$. Substituting for p_1 from (A.14) into the terms $-\partial_t p_1 + \mathcal{L}_2^* p_1$, integrating over y (the integral is finite by our assumption on $\sigma(y)$), and setting the result equal to zero gives the desired equation for $f_1(t,x)$:

(A.16)

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

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.

A.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 behavior, 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

$$\frac{1}{\epsilon}\partial_{t'} p' - \left(\frac{1}{\epsilon}\mathcal{L}_0^* + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1^* + \mathcal{L}_2^*\right) p' = 0, \tag{A.17}$$

with the initial condition

$$p'(0,x_0,y_0;0,x,y) = \delta(x - x_0) \otimes \delta(y - y_0). \tag{A.18}$$

We proceed as above, expanding

$$p'(0,x_0,y_0;t',x,y) \sim p_0'(t',x,y) + \sqrt{\epsilon} p_1'(t',x,y) + \epsilon p_2'(t',x,y) + \epsilon^{3/2} p_3'(t',x,y) + \cdots.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Substituting into (A.17) and equating coefficients of powers of ϵ leads immediately to

$$\text{(A.19)} \qquad \text{At } O(1/\epsilon): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p_0' = 0;$$

$$\text{(A.20)} \qquad \text{At } O(1/\sqrt{\epsilon}): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p_1' = \mathcal{L}_1^*\, p_0';$$

$$\text{(A.21)} \qquad \text{At } O(1): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p_2' = \mathcal{L}_1^*\, p_1' + \mathcal{L}_2^*\, p_0';$$

$$\text{(A.22)} \qquad \text{At } O(\sqrt{\epsilon}): \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p_3' = \mathcal{L}_1^*\, p_2' + \mathcal{L}_2^*\, p_1';$$

the pattern in the last two of these repeats at higher orders. The initial conditions are

$$p_0'(0,x,y) = \delta(x-x_0) \otimes \delta(y-y_0), \qquad p_i'(0,x,y) = 0, \quad i = 1,2,3\ldots.$$

These comprise the (trivial) expansion in ϵ of the initial data for the full problem, combined with a matching condition with the solution in the inner-inner region (described in the next section).

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

$$\text{(A.23)} \qquad (\partial_{t'} - \mathcal{L}_0^*)\, p_0' = 0, \qquad p_0'(0,x,y) = \delta(x-x_0) \otimes \delta(y-y_0).$$

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 density of this OU process at time t' is normal with mean and variance

$$\text{(A.24)} \qquad m(t';y_0) = m + (y_0 - m)\mathrm{e}^{-\kappa t'}, \qquad \mathrm{var}(t') = \frac{\nu^2}{\kappa}\left(1 - \mathrm{e}^{-2\kappa t'}\right),$$

respectively. Its density is

$$\text{(A.25)} \qquad \phi^Y(t',y) = \frac{1}{\sqrt{2\pi \mathrm{var}(t')}}\mathrm{e}^{-(x-m(t';y_0))^2/2\mathrm{var}(t')},$$

which satisfies $(\partial_{t'} - \mathcal{L}_0^*)\,\phi^Y(t',y) = 0, \quad \phi^Y(0,y) = \delta(y-y_0)$. It follows that the unique solution to (A.23) is

$$p_0'(t',x,y) = \delta(x-x_0)\phi^Y(t',y).$$

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

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

as defined above. Hence the limit of $p_0(t',x,y)$ as $t' \to \infty$ is $\delta(x-x_0)\Phi(y)$ and (by asymptotic matching) this is the initial condition for $f_0(t,x)$:

$$\text{(A.26)} \qquad f_0(t,x) = \delta(x-x_0).$$

Thus, 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.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

At $O(1/\sqrt{\epsilon})$. The equation (A.20) for p_1' now becomes

$$(\partial_{t'} - \mathcal{L}_0^*)\, p_1' = \rho_{xy}\nu\sqrt{2}\,(\partial_x \delta(x - x_0))\,\partial_y\left(\sigma(y)\phi^Y(t', y)\right). \tag{A.27}$$

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

$$-\rho_{xy}\nu\sqrt{2}\,(\partial_x \delta(x - x_0))\,\Sigma(y)\Phi^Y(y) + c_1(x)\Phi^Y(y) \tag{A.28}$$

for an as yet unknown function $c_1(x)$, which, as we see below, provides the initial condition for $f_1(t, x)$, while the first term in (A.28) 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 (A.27) 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 (A.28) must also integrate to zero. Bearing in mind that $\Phi^Y(y)$ is a probability density and so integrates to 1, we have

$$\begin{aligned} c_1(x) &= -(\rho_{xy}\sqrt{2}\nu)\,(\partial_x \delta(x - x_0)) \int_{-\infty}^{\infty} \Sigma(y)\Phi^Y(y)\,\mathrm{d}y \\ &= -(\rho_{xy}\sqrt{2}/\nu)\,(\partial_x \delta(x - x_0))\,\langle\Sigma\rangle \\ &= f_1(0, x). \end{aligned} \tag{A.29}$$

This is the initial condition for (A.16).

A.4. Inner-inner layer. The delta function in (A.25) signals that a further rescaling is needed (not treated in [23], as their initial data is C^∞ in x and y). 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 $x = x_0 + \sqrt{\epsilon}x'$, $p'(t', x, y) = (1/\sqrt{\epsilon})p''(t', x', y)$, at leading order in an expansion $p'' \sim p_0'' + O(\sqrt{\epsilon})$, we have

$$\partial_{t'} p_0'' - \left(\frac{1}{2}\sigma^2(y)\partial_{x'x'} + \rho_{xy}\nu\sqrt{2}\sigma(y)\partial_{x'y} + \nu^2\partial_{yy}\right) p_0'' + \kappa\partial_y\left((m - y)p_0''\right) = 0, \tag{A.30}$$

$$p_0''(0, x', y) = \delta(x')\delta(y - y_0) \tag{A.31}$$

(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 x'). 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.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

REFERENCES

[1] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden, *The heterogeneous multiscale method*, Acta Numer., 21 (2012), pp. 1–87, https://doi.org/10.1017/S0962492912000025.

[2] A. Bain and D. Crisan, *Fundamentals of Stochastic Filtering*, Stoch. Model. Appl. Probab. 60, Springer, 2009.

[3] M. Bichuch and R. Sircar, *Optimal investment with transaction costs and stochastic volatility part I: Infinite horizon*, SIAM J. Control Optim., 55 (2017), pp. 3799–3832, https://doi.org/10.1137/140983859.

[4] C.-E. Bréhier and S. Rakotonirina-Ricquebourg, *On asymptotic preserving schemes for a class of stochastic differential equations in averaging and diffusion approximation regimes*, Multiscale Model. Simul., 20 (2022), pp. 118–163, https://doi.org/10.1137/20M1379836.

[5] K. Bujok and C. Reisinger, *Numerical valuation of basket credit derivatives in structural jump-diffusion models*, J. Comput. Finance, 15 (2012), pp. 115–158, https://doi.org/10.21314/JCF.2012.249.

[6] N. Bush, B. M. Hambly, H. Haworth, L. Jin, and C. Reisinger, *Stochastic evolution equations in portfolio credit modelling*, SIAM J. Financial Math., 2 (2011), pp. 627–664, https://doi.org/10.1137/100796777.

[7] J. G. Conlon and M. G. Sullivan, *Convergence to Black–Scholes for ergodic volatility models*, European J. Appl. Math., 16 (2005), pp. 385–409.

[8] A. S. Cozma and C. Reisinger, *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, Springer, 2020, pp. 223–240.

[9] J.-P. Fouque and R. Hu, *Multiscale asymptotic analysis for portfolio optimization under stochastic environment*, Multiscale Model. Simul., 18 (2020), pp. 1318–1342, https://doi.org/10.1137/19M1245967.

[10] J.-P. Fouque, R. Hu, and R. Sircar, *Sub- and supersolution approach to accuracy analysis of portfolio optimization asymptotics in multiscale stochastic factor markets*, SIAM J. Financial Math., 13 (2022), pp. 109–128, https://doi.org/10.1137/21M1428625.

[11] J.-P. Fouque, M. Lorig, and R. Sircar, *Second order multiscale stochastic volatility asymptotics: Stochastic terminal layer analysis and calibration*, Finance Stoch., 20 (2016), pp. 543–588, https://doi.org/10.1007/s00780-016-0298-y.

[12] J.-P. Fouque, G. Papanicolaou, and K. R. Sircar, *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press, 2000.

[13] J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna, *Multiscale stochastic volatility asymptotics*, Multiscale Model. Simul., 2 (2003), pp. 22–42, https://doi.org/10.1137/030600291.

[14] J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna, *Singular perturbations in option pricing*, SIAM J. Appl. Math., 63 (2003), pp. 1648–1665, https://doi.org/10.1137/S0036139902401550.

[15] J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna, *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*, Cambridge University Press, 2011.

[16] M. Fukasawa, *Asymptotic analysis for stochastic volatility: Martingale expansion*, Finance Stoch., 15 (2011), pp. 635–654, https://doi.org/10.1007/s00780-010-0136-6.

[17] K. Giesecke, K. Spiliopoulos, R. Sowers, and J. Sirignano, *Large portfolio asymptotics for loss from default*, Math. Finance, 25 (2015), pp. 77–114, https://doi.org/10.1111/mafi.12011.

[18] M. B. Giles and C. Reisinger, *Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance*, SIAM J. Financial Math., 3 (2012), pp. 572–592, https://doi.org/10.1137/110841916.

[19] B. Hambly and N. Kolliopoulos, *Stochastic evolution equations for large portfolios of stochastic volatility models*, SIAM J. Financial Math., 8 (2017), pp. 962–1014, https://doi.org/10.1137/17M111715X.

[20] B. Hambly and N. Kolliopoulos, *Stochastic PDEs for large portfolios with general mean-reverting volatility processes*, Probab. Uncertain. Quant. Risk., 9 (2024), pp. 263–300, https://doi.org/10.3934/puqr.2024013.

[21] B. Hambly and N. Kolliopoulos, *Fast mean-reversion asymptotics for large portfolios of stochastic volatility models*, Finance Stoch., 24 (2020), pp. 757–794, https://doi.org/10.1007/s00780-020-00422-7.

[22] S. Howison, *Matched asymptotic expansions in financial engineering*, J. Eng. Math., 53 (2005), pp. 385–406, https://doi.org/10.1007/s10665-005-7716-z.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[23] R. Khasminskii and G. Yin, *Limit behavior of two-time-scale diffusions revisited*, J. Differential Equations, 212 (2005), pp. 85–113, https://doi.org/10.1016/j.jde.2004.08.013.
[24] T. G. Kurtz and J. Xiong, *Particle representations for a class of nonlinear SPDEs*, Stochastic Process. Appl., 83 (1999), pp. 103–126, https://doi.org/10.1016/S0304-4149(99)00024-1.
[25] S. P. Meyn and R. L. Tweedie, *Stability of Markovian processes* III*: Foster–Lyapunov criteria for continuous-time processes*, Adv. Appl. Probab., 25 (1993), pp. 518–548, https://doi.org/10.2307/1427522.
[26] A. Papanicolaou, *Nonlinear filters for hidden Markov models of regime change with fast mean-reverting states*, Multiscale Model. Simul., 10 (2012), pp. 906–935, https://doi.org/10.1137/110819937.
[27] A. Papanicolaou and K. Spiliopoulos, *Filtering the maximum likelihood for multiscale problems*, Multiscale Model. Simul., 12 (2014), pp. 1193–1229, https://doi.org/10.1137/140952648.
[28] C. Reisinger, *Mean-square stability and error analysis of implicit time-stepping schemes for linear parabolic SPDEs with multiplicative Wiener noise in the first derivative*, Int. J. Comput. Math., 89 (2012), pp. 2562–2575, https://doi.org/10.1080/00207160.2012.718762.
[29] C. Reisinger and Z. Wang, *Analysis of multi-index Monte Carlo estimators for a Zakai SPDE*, J. Comput. Math., 36 (2018), pp. 202–236, https://doi.org/10.4208/jcm.1612-m2016-0681.
[30] C. Reisinger and Z. Wang, *Stability and error analysis of an implicit Milstein finite difference scheme for a two-dimensional Zakai SPDE*, BIT, 59 (2019), pp. 987–1029, https://doi.org/10.1007/s10543-019-00761-8.

Downloaded 02/24/25 to 82.17.189.10 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.