



# Stability and error analysis of an implicit Milstein finite difference scheme for a two-dimensional Zakai SPDE

Christoph Reisinger<sup>1</sup> · Zhenru Wang<sup>1</sup>

Received: 10 December 2018 / Accepted: 5 June 2019  
© The Author(s) 2019

## Abstract

This article proposes an implicit finite difference scheme for a two-dimensional parabolic stochastic partial differential equation of Zakai type. The scheme is based on a Milstein approximation to the stochastic integral and an alternating direction implicit discretisation of the elliptic term. Mean-square stability and  $L_2$ -convergence of first order in time and second order in space are proven by Fourier analysis, in the presence of Dirac initial data. Numerical tests confirm these findings empirically.

**Keywords** Stochastic partial differential equations · Milstein scheme · Stochastic finite differences · Splitting schemes · Mean-square stability ·  $L_2$ -convergence

**Mathematics Subject Classification** 65T50 · 60H15 · 65N06 · 65N12

## 1 Introduction

The analysis and numerical computation of Zakai equations and other types of stochastic partial differential equation (SPDE) have been extensively studied in recent years. A general form of Zakai equation (see [1,17]) is given by

$$\begin{aligned} dv(t, x) = & \left( \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(x)v(t, x)] - \sum_{i=1}^d \frac{\partial}{\partial x_i} [b_i(x)v(t, x)] \right) dt \\ & - \sum_{l=1}^m \sum_{i=1}^d \frac{\partial}{\partial x_i} (\gamma_{i,l}(x)v(t, x)) dM_t^l, \end{aligned} \quad (1.1)$$

---

Communicated by David Cohen.

---

✉ Zhenru Wang  
zhenru.wang@maths.ox.ac.uk

Christoph Reisinger  
christoph.reisinger@maths.ox.ac.uk

<sup>1</sup> Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road, Oxford OX2 6GG, UK

where  $M_t = (M_t^1, \dots, M_t^m)'$  is an  $m$ -dimensional standard Brownian motion,  $a$  is a  $d \times d$  matrix-valued function,  $b$  is a  $\mathbb{R}^d$ -valued function, and  $\gamma = (\gamma_{i,j}(x))$  is a  $d \times m$  matrix-valued function. This Zakai equation arises from a nonlinear filtering problem: given an  $m$ -dimensional observation process  $M$  and a  $d$ -dimensional signal process  $Z$ , the goal is to estimate the conditional distribution of  $Z$  given  $M$ . If  $Z$  satisfies

$$Z_t = Z_0 + \int_0^t \beta(Z_s) ds + \int_0^t \sigma(Z_s) dB_s + \int_0^t \gamma(Z_s) dM_s, \quad (1.2)$$

where  $B$  is a  $d$ -dimensional standard Brownian motion independent of  $M$ ,  $\sigma$  is a  $d \times d$ -matrix valued function, and  $\beta$  is a  $\mathbb{R}^d$ -valued function, then under appropriate conditions, the conditional distribution function of  $Z$  given  $M$  has a density  $v(t, \cdot) \in L_2(\mathbb{R}^d) = \{f : \int_{\mathbb{R}^d} |f(x)|^2 dx < \infty\}$  almost surely (see Corollary 3.1 in [29]), and, from Theorem 3.1 in [29],  $v$  satisfies (1.1) in a weak sense with

$$a = \sigma \sigma^\top + \gamma \gamma^\top, \quad b = \beta. \quad (1.3)$$

Moreover, the solution  $v$  to (1.1) can be interpreted as the density—if it exists—of the limit empirical measure  $v_t = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \delta_{Z_t^i}$  for

$$Z_t^i = Z_0 + \int_0^t \beta(Z_s^i) ds + \int_0^t \sigma(Z_s^i) dB_s^i + \int_0^t \gamma(Z_s^i) dM_s, \quad (1.4)$$

where  $B^i$  and  $N$  are independent Brownian motions, independent of  $M$ , and the rest as above.

There are two major approaches to the numerical approximation of the Zakai equation. One is by simulating the particle system (1.4) with Monte Carlo method, for instance as in [10–12, 17]. The other approach is to directly solve the Zakai SPDE by spatial approximation methods and time stepping schemes, coupled again with Monte Carlo sampling, which is the subject of this paper.

Within this second class of methods, several schemes were developed in earlier works for different types of SPDEs, including finite differences [14, 18–20], finite elements [27, 37], and stochastic Taylor schemes [24, 25], but these were restricted to types of SPDEs not including Zakai equations of the type (1.1).

More recently, methods have been developed and analysed for parabolic SPDEs of the generic form

$$dv = \mathcal{L}v dt + G(v) dM_t, \quad (1.5)$$

where  $\mathcal{L}$  is a second order elliptic differential operator, and  $G$  is a functional mapping  $v$  onto a linear operator from martingales  $M$  into a suitable function space.

Under suitable regularity, for equations of type (1.5), mean-square convergence of order 1/2 is shown for an Euler semi-discretisation in [30] for square-integrable (not necessarily continuous), infinite-dimensional martingale drivers. In contrast, [3] allows only for continuous martingales but prove convergence of higher order in space and up to 1 in time, in  $L^p$  and almost surely, for a Milstein scheme and spatial Galerkin

approximation of sufficiently high order; this is extended to advection-diffusion equations with possibly discontinuous martingales in [2].

Giles and Reisinger [16] use an explicit Milstein finite difference approximation to the solution of the following one-dimensional SPDE, a special case of (1.1) for  $d = 1$  and constant coefficients,

$$dv = -\mu \frac{\partial v}{\partial x} dt + \frac{1}{2} \frac{\partial^2 v}{\partial x^2} dt - \sqrt{\rho} \frac{\partial v}{\partial x} dM_t, \quad (t, x) \in (0, T) \times \mathbb{R}, \quad (1.6)$$

where  $T > 0$ ,  $M$  is a standard Brownian motion, and  $\mu$  and  $0 \leq \rho < 1$  are real-valued parameters. This is extended in [35] to an approximation of (1.6) with an implicit method on the basis of the  $\sigma$ - $\theta$  time-stepping scheme, where the finite variation parts of the double stochastic integral are taken implicit. This is further applied in [36] to Multi-index Monte Carlo estimation of expectations of a functional of the solution.

A finite difference scheme for a filtered jump-diffusion process resulting in a stochastic integro-differential equation is studied in [13], where convergence of order 1 in space and 1/2 in time, in  $L_2$  and  $L_\infty$  in space, is proven for an Euler time stepping scheme.

The theoretical results in this paper are an extension from those in [16] to the multi-dimensional case. They are more specific than those in [13] in that we analyse only the case of constant coefficient local SPDEs. In contrast to [2,4], we consider only finite-dimensional Brownian motions, as is relevant in our applications. But we specifically include the case of Dirac initial data and extend the results to a practically attractive, semi-implicit alternating direction implicit factorisation in the context of the Milstein scheme.

We want to allow for Dirac initial data because they correspond to the natural situation where all particles in (1.4) start from the same initial position, or a filtering problem with known current state  $Z_0$  in (1.2).

Specifically, we study first the two-dimensional stochastic partial differential equation

$$\begin{aligned} dv = & -\mu_x \frac{\partial v}{\partial x} dt - \mu_y \frac{\partial v}{\partial y} dt + \frac{1}{2} \left( \frac{\partial^2 v}{\partial x^2} + 2\sqrt{\rho_x \rho_y \rho_{xy}} \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 v}{\partial y^2} \right) dt \\ & - \sqrt{\rho_x} \frac{\partial v}{\partial x} dM_t^x - \sqrt{\rho_y} \frac{\partial v}{\partial y} dM_t^y, \end{aligned} \quad (1.7)$$

for  $x, y \in \mathbb{R}$ ,  $0 < t \leq T$ , where  $\mu_x, \mu_y$  and  $0 \leq \rho_x, \rho_y < 1$ ,  $-1 \leq \rho_{xy} \leq 1$  are real-valued parameters, subject to the Dirac initial data

$$v(0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0), \quad (1.8)$$

with  $x_0$  and  $y_0$  given. It is derived from the special case where the signal processes  $Z = (X, Y)'$  satisfies (1.2) with

$$\beta = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \quad \sigma = \begin{bmatrix} \sqrt{1 - \rho_x} & 0 \\ 0 & \sqrt{1 - \rho_y} \end{bmatrix}, \quad \gamma = \begin{bmatrix} \sqrt{\rho_x} & 0 \\ \sqrt{\rho_y \rho_{xy}} & \sqrt{\rho_y (1 - \rho_{xy}^2)} \end{bmatrix}.$$

A classical result states that, for a class of SPDEs including (1.7), with initial condition in  $L_2$ , there exists a unique solution  $v \in L_2(\Omega \times (0, T), \mathcal{F}, L_2(\mathbb{R}^2))$  (see [28]). This does not include Dirac initial data (1.8), but in fact, the solution to (1.7) and (1.8) can be found analytically, similar to the heat equation, by solving the SDE (3.2) below (compared to an ODE in the case of the heat equation) in the Fourier space and transforming back, which yields a smooth (in  $x$  and  $y$ ) function

$$v(T, x, y) = \frac{\exp\left(-\frac{(x-x_0-\mu_x T-\sqrt{\rho_x}M_T^x)^2}{2(1-\rho_x)T} - \frac{(y-y_0-\mu_y T-\sqrt{\rho_y}M_T^y)^2}{2(1-\rho_y)T}\right)}{2\pi\sqrt{(1-\rho_x)(1-\rho_y)T}}. \quad (1.9)$$

The availability of a closed-form solution in this case helps us check the validity of our numerical scheme and its convergence rate, whereas the scheme itself is more widely applicable.

For the SPDE (1.7), we consider both explicit and implicit Milstein schemes. We study the mean-square stability, and the strong convergence of the second moment. This can give us an error bound for the expected error. The advantage over the simpler Euler scheme is that the strong convergence order is improved from  $1/2$  to  $1$ . As expected, we find that the explicit scheme is stable in the mean-square sense only under a strong CFL-type condition on the timestep,  $k \leq Ch^2$  for timestep  $k$ , mesh size  $h$ , and a constant  $C$ , while the implicit scheme is mean-square stable under the very mild and somewhat unusual CFL condition  $k \leq C|\log h|^{-1}$  (provided also some constraints on  $\rho_x, \rho_y, \rho_{xy}$ ).

We therefore focus on the implicit scheme, for which we prove first order convergence in the timestep and second order in the spatial mesh size. The analysis is made more difficult by the Dirac initial datum compared to, say,  $L_2$  initial data. We adapt the approach used in [7] for the heat equation by studying the convergence for different wave number regions in Fourier space and then assemble the contributions to the error by the inverse transform.

Furthermore, we use an Alternating Direction Implicit (ADI) scheme to approximately factorise the discretisation matrix for the implicit elliptic part in (1.7). This concept is well established for PDEs (see, e.g., [9,23,34]). It is well known that in the multi-dimensional case standard implicit schemes result in sparse banded linear systems, which cannot be solved by direct elimination in a computational cost which scales linearly with the number of unknowns like in the one-dimensional, tridiagonal case. An alternative to advanced iterative linear solvers such as multigrid methods is to reduce the large sparse linear system approximately to a sequence of tridiagonal linear systems, which are computationally easier to handle, by ADI factorisation. To our knowledge, the present work is the first application of ADI to SPDEs. We show that the ADI approximation is also mean-square stable under the same conditions as the original implicit scheme and has the same convergence order.

We note that published analysis of ADI schemes for parabolic PDEs in the presence of mixed spatial derivative terms is currently restricted to constant coefficients (through the use of von Neumann stability analysis; see e.g. [39]). Notwithstanding this, the empirical evidence overwhelmingly suggests that the conclusions drawn there extend to most cases of variable coefficients.

We give a natural extension of the proposed scheme to the SPDE (1.1). In that case, additional iterated stochastic integrals (the *Lévy area*) appear in the Milstein approximation. The efficient, accurate simulation has been studied in the context of SDEs in [15, 26] and invariably leads to relatively complicated schemes. As the computational effort in the context of the SPDE (1.1) is dominated by the matrix computations from the finite difference scheme, it is sufficient to perform a simple approximation of the stochastic integrals  $\int_t^{t+k} (W_s - W_t) dB_s$ , for correlated Brownian motions  $W$  and  $B$ , by simple Euler integration with step  $k^2$ , without adversely affecting the convergence.

As a specific application, we approximate the equation

$$\begin{aligned} du = & \left[ \kappa_1 u - \left( r_1 - \frac{1}{2}y - \xi_1 \rho_{3,1} \rho_{2,1} \right) \frac{\partial u}{\partial x} - \left( \kappa_1 (\theta_1 - y) - \xi_1^2 \right) \frac{\partial u}{\partial y} \right. \\ & + \frac{1}{2} y \frac{\partial^2 u}{\partial x^2} + \xi_1 \rho_{3,1} \rho_{2,1} y \frac{\partial^2 u}{\partial x \partial y} + \frac{\xi_1^2}{2} y \frac{\partial^2 u}{\partial y^2} \Big] dt \\ & - \rho_{1,1} \sqrt{y} \frac{\partial u}{\partial x} dW_t - \xi_1 \rho_{2,1} \frac{\partial}{\partial y} (\sqrt{y} u) dB_t, \end{aligned} \quad (1.10)$$

taken from [21], with the scheme presented in this paper. Although our analysis (based on Fourier transforms) does not directly apply in this case, the scheme preserves first order convergence in time and second order convergence in space in our numerical tests.

The sharp estimates we derive give a precise description of the error given Dirac initial data. This is achieved by a Fourier analysis originating from Carter and Giles in [7], where they estimated the error arising from explicit and implicit approximations of the constant-coefficient 1-d convection–diffusion equation with Dirac initial data.

Moreover, the stability and error analysis in the constant coefficient case provide an accurate description of the local behaviour at  $(t, x) = (0, x_0)$  in the variable coefficient case.<sup>1</sup> This connection is established by “freezing” the coefficients at the point  $(0, x_0)$  in a more general SPDE of the form (1.7), such that we obtain

$$\begin{aligned} dv_0(t, x) = & \left( \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x_0) \frac{\partial^2}{\partial x_i \partial x_j} v_0(t, x) - \sum_{i=1}^d b_i(x_0) \frac{\partial}{\partial x_i} v_0(t, x) \right) dt \\ & - \sum_{i=1}^d \left( \sum_{l=1}^m \gamma_{i,l}(x_0) \right) \frac{\partial}{\partial x_i} v_0(t, x) dM_t^l. \end{aligned} \quad (1.11)$$

It can be seen by elementary calculus that (1.11) in the case  $d = 2$  can be brought in the form (1.7) by a linear coordinate transformation, noting that  $a - \gamma \gamma^\top$  is symmetric positive definite by (1.3). This reduction of a broader class of problems to a test equation with a known solution is in the spirit of the well-established paradigm of stability analysis for ODEs described in [22].

We give a numerical justification for this reduction in Sect. 6, by presenting numerical tests for the SPDEs (1.10) and (1.11) to compare the error at  $(x_0, y_0)$  between

<sup>1</sup> Note that for simplicity the coefficients in (1.7) have no  $t$ -dependence.

the original SPDE and SPDE with frozen coefficients at  $(x_0, y_0)$ . We find that the behaviour is extremely close, especially for small time intervals, which confirms the broader usefulness of the method.

Summarising, the novel contributions of this paper are as follows. We

- Give a rigorous stability and error analysis for a Milstein finite difference scheme for the SPDE (1.7) in terms of  $L_2$  in probability, pointwise as well as  $L_2$  in space, deriving sharp leading order error terms;
- Derive pointwise errors for Dirac initial data, which reveal a mild instability for large implicit timesteps and small spatial mesh sizes in this case, not seen in previous studies for  $L_2$  data;
- Extend the analysis to an alternating direction implicit (ADI) factorisation, which, to our knowledge, is the first application of an ADI scheme to stochastic PDEs;
- Propose a modification for the more general equation (1.1) through sub-simulation of the Lévy area, which is empirically shown to be of first order.

The rest of this article is structured as follows. We define the approximation schemes in Sect. 2. Then we analyse the mean-square stability and  $L_2$ -convergence in Sects. 3 and 4 in the constant coefficient case of (1.7). Section 5 shows numerical experiments confirming the above findings. Section 6 extends the scheme to variable coefficients as in (1.1) and presents tests for the example (1.10). Section 7 offers conclusions and directions for further research.

## 2 Approximation and main results

### 2.1 Semi-implicit Milstein finite difference scheme

First, we introduce the numerical scheme to the SPDE (1.7), repeated here for convenience,

$$\begin{aligned} dv = & -\mu_x \frac{\partial v}{\partial x} dt - \mu_y \frac{\partial v}{\partial y} dt + \frac{1}{2} \left( \frac{\partial^2 v}{\partial x^2} + 2\sqrt{\rho_x \rho_y} \rho_{xy} \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 v}{\partial y^2} \right) dt \\ & - \sqrt{\rho_x} \frac{\partial v}{\partial x} dM_t^x - \sqrt{\rho_y} \frac{\partial v}{\partial y} dM_t^y, \end{aligned}$$

with Dirac initial  $v(0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0)$ . We use a spatial grid 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  $v(nk, ih_x, jh_y)$ ,  $n = 1, \dots, N$ ,  $i, j \in \mathbb{Z}$ , where  $i_0 := \lfloor x_0/h_x \rfloor$ ,  $j_0 := \lfloor y_0/h_y \rfloor$ , the closest integers to  $x_0/h_x$  and  $y_0/h_y$ . We approximate  $v(0, x, y)$  by

$$V_{i,j}^0 = h_x^{-1} h_y^{-1} \delta_{(i_0, j_0)} = \begin{cases} h_x^{-1} h_y^{-1}, & i = i_0, j = j_0, \\ 0, & \text{otherwise.} \end{cases} \quad (2.1)$$

To improve the accuracy of the approximation of  $v$  in the present case of Dirac initial data, we subsequently choose  $h_x$  and  $h_y$  such that  $x_0/h_x$  and  $y_0/h_y$  are integers and therefore  $x_0$  and  $y_0$  are on the grid.

Extending the implicit Euler scheme in [35] for the 1-d case, it is natural to take the drift term

$$-\mu_x \frac{\partial v}{\partial x} - \mu_y \frac{\partial v}{\partial y} + \frac{1}{2} \left( \frac{\partial^2 v}{\partial x^2} + 2\sqrt{\rho_x \rho_y} \rho_{xy} \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 v}{\partial y^2} \right)$$

implicit, and the terms driven by  $M^x$  and  $M^y$  explicit. We will prove later that in this way we obtain better stability (compare Proposition 3.2 to Theorem 3.1). For computational simplicity, in the following, we take the mixed derivative term therein explicit. This is also in preparation for the ADI splitting schemes we will study later.

Using such a semi-implicit Euler scheme, the SPDE (1.7) can be approximated by

$$\begin{aligned} V^{n+1} = & V^n - \frac{\mu_x k}{2h_x} D_x V^{n+1} - \frac{\mu_y k}{2h_y} D_y V^{n+1} + \frac{k}{2h_x^2} D_{xx} V^{n+1} + \frac{k}{2h_y^2} D_{yy} V^{n+1} \\ & + \sqrt{\rho_x \rho_y} \rho_{xy} \frac{k}{4h_x h_y} D_{xy} V^n - \frac{\sqrt{\rho_x k} Z_{n,x}}{2h_x} D_x V^n - \frac{\sqrt{\rho_y k} \tilde{Z}_{n,y}}{2h_y} D_y V^n, \end{aligned}$$

where

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

and  $\tilde{Z}_n^y = \rho_{xy} Z_n^x + \sqrt{1 - \rho_{xy}^2} Z_n^y$ , with  $Z_n^x, Z_n^y \sim N(0, 1)$  being independent normal random variables. To achieve a higher order of convergence, we introduce the Milstein scheme. Integrating (1.7) over the time interval  $[nk, (n+1)k]$ ,

$$\begin{aligned} v(nk+k, x, y) = & v(nk, x, y) + \int_{nk}^{nk+k} \left( -\mu_x \frac{\partial v}{\partial x} - \mu_y \frac{\partial v}{\partial y} \right. \\ & + \frac{1}{2} \frac{\partial^2 v}{\partial x^2} + \sqrt{\rho_x \rho_y} \rho_{xy} \frac{\partial^2 v}{\partial x \partial y} + \frac{1}{2} \frac{\partial^2 v}{\partial y^2} \Big) ds \\ & - \int_{nk}^{nk+k} \sqrt{\rho_x} \frac{\partial v}{\partial x} dM_s^x - \int_{nk}^{nk+k} \sqrt{\rho_y} \frac{\partial v}{\partial y} dM_s^y. \end{aligned}$$

In the Euler scheme, we approximate all integrands by their value at time  $nk$  or  $(n+1)k$ , which is a zero-order expansion in time. By contrast, in the Milstein scheme, we use a first-order expansion for the stochastic integrals, such that we approximate  $v(s, x, y) \approx v(nk, x, y)$  for  $nk < s < (n+1)k$  in the first integral and

$$\begin{aligned} v(s, x, y) \approx & v(nk, x, y) - \sqrt{\rho_x} \frac{\partial v}{\partial x}(nk, x, y)(M_s^x - M_{nk}^x) \\ & - \sqrt{\rho_y} \frac{\partial v}{\partial y}(nk, x, y)(M_s^y - M_{nk}^y) \end{aligned}$$

in the second and third. We denote  $nk$  as  $t$ , and it follows

$$\begin{aligned}
 & - \int_t^{t+k} \sqrt{\rho_x} \frac{\partial v}{\partial x}(s, x, y) \, dM_s^x - \int_t^{t+k} \sqrt{\rho_y} \frac{\partial v}{\partial y}(s, x, y) \, dM_s^y \\
 & \approx -\sqrt{\rho_x} \frac{\partial v}{\partial x}(t, x, y) \Delta M_n^x - \sqrt{\rho_y} \frac{\partial v}{\partial y}(t, x, y) \Delta M_n^y \\
 & \quad + \rho_x \frac{\partial^2 v}{\partial x^2}(t, x, y) \int_t^{t+k} (M_s^x - M_t^x) \, dM_s^x \\
 & \quad + \rho_y \frac{\partial^2 v}{\partial y^2}(t, x, y) \int_t^{t+k} (M_s^y - M_t^y) \, dM_s^y + \sqrt{\rho_x \rho_y} \frac{\partial^2 v}{\partial x \partial y}(t, x, y) \\
 & \quad \left( \int_t^{t+k} (M_s^x - M_t^x) \, dM_s^y + \int_t^{t+k} (M_s^y - M_t^y) \, dM_s^x \right),
 \end{aligned}$$

where

$$\Delta M_n^x = M_{t+k}^x - M_t^x = \sqrt{k} Z_n^x, \quad \Delta M_n^y = M_{t+k}^y - M_t^y = \sqrt{k} \tilde{Z}_n^y.$$

From standard Itô calculus, we have

$$\begin{aligned}
 \int_t^{t+k} (M_s^x - M_t^x) \, dM_s^x &= \frac{1}{2} \left( (\Delta M_n^x)^2 - k \right), \\
 \int_t^{t+k} (M_s^y - M_t^y) \, dM_s^y &= \frac{1}{2} \left( (\Delta M_n^y)^2 - k \right), \\
 \int_t^{t+k} (M_s^x - M_t^x) \, dM_s^y + \int_t^{t+k} (M_s^y - M_t^y) \, dM_s^x &= \Delta M_n^x \Delta M_n^y - \rho_{xy} k.
 \end{aligned}$$

We see that the mixed-derivative terms cancel, and we derive the implicit Milstein scheme as follows,

$$\begin{aligned}
 & \left( I + \frac{\mu_x k}{2h_x} D_x + \frac{\mu_y k}{2h_y} D_y - \frac{k}{2h_x^2} D_{xx} - \frac{k}{2h_y^2} D_{yy} \right) V^{n+1} \\
 &= \left( I - \frac{\sqrt{\rho_x k} Z_{n,x}}{2h_x} D_x - \frac{\sqrt{\rho_y k} \tilde{Z}_{n,y}}{2h_y} D_y + \frac{\rho_x k (Z_{n,x}^2 - 1)}{8h_x^2} D_x^2 \right. \\
 & \quad \left. + \frac{\rho_y k (\tilde{Z}_{n,y}^2 - 1)}{8h_y^2} D_y^2 + \frac{\sqrt{\rho_x \rho_y} k Z_{n,x} \tilde{Z}_{n,y}}{4h_x h_y} D_{xy} \right) V^n.
 \end{aligned} \tag{2.2}$$

To facilitate its implementation, we combine the scheme with an Alternating Direction Implicit (ADI) factorisation, which has been introduced in [34] for parabolic PDEs to approximately factorise the system matrix by matrices which correspond to derivatives in individual directions and which can thus more easily be inverted, while the consistency order is maintained. Applying this principle to the implicit terms on the left-hand side of (2.2),



$$\begin{aligned}
 & \left( I + \frac{\mu_x k}{2h_x} D_x + \frac{\mu_y k}{2h_y} D_y - \frac{k}{2h_x^2} D_{xx} - \frac{k}{2h_y^2} D_{yy} \right) V \\
 &= \left( I + \frac{\mu_x k}{2h_x} D_x - \frac{k}{2h_x^2} D_{xx} \right) \left( I + \frac{\mu_y k}{2h_y} D_y - \frac{k}{2h_y^2} D_{yy} \right) V \\
 &\quad - k^2 \left( \frac{\mu_x}{2h_x} D_x - \frac{1}{2h_x^2} D_{xx} \right) \left( \frac{\mu_y}{2h_y} D_y - \frac{1}{2h_y^2} D_{yy} \right) V.
 \end{aligned}$$

Since the second term on right-hand side proportional to  $k^2$  is of higher order than the overall scheme (which is of first order), it is justifiable in terms of the consistency order to neglect it. This yields the ADI scheme as follows,

$$\begin{aligned}
 & \left( I + \frac{\mu_x k}{2h_x} D_x - \frac{k}{2h_x^2} D_{xx} \right) \left( I + \frac{\mu_y k}{2h_y} D_y - \frac{k}{2h_y^2} D_{yy} \right) V^{n+1} \\
 &= \left( I - \frac{\sqrt{\rho_x k} Z_{n,x}}{2h_x} D_x - \frac{\sqrt{\rho_y k} \tilde{Z}_{n,y}}{2h_y} D_y + \frac{\rho_x k (Z_{n,x}^2 - 1)}{8h_x^2} D_x^2 \right. \\
 &\quad \left. + \frac{\rho_y k (\tilde{Z}_{n,y}^2 - 1)}{8h_y^2} D_y^2 + \frac{\sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y}}{4h_x h_y} D_{xy} \right) V^n. \quad (2.3)
 \end{aligned}$$

Note that there is no substantial benefit in second order accurate splitting schemes (such as Craig–Sneyd [9] or Hundsdorfer–Verwer [23]) as the overall order is limited to 1 by the Milstein approximation to the stochastic integral.

We approximate the second derivative on the right-hand side with  $D_x^2$  and  $D_y^2$ , but the results for  $D_{xx}$  and  $D_{yy}$  would be similar.

We can also use the explicit Milstein finite difference scheme to approximate the SPDE (1.7)

$$\begin{aligned}
 V^{n+1} = & \left( I - \frac{\mu_x k + \sqrt{\rho_x k} Z_{n,x}}{2h_x} D_x - \frac{\mu_y k + \sqrt{\rho_y k} \tilde{Z}_{n,y}}{2h_y} D_y + \frac{k}{2h_x^2} D_{xx} + \frac{k}{2h_y^2} D_{yy} \right. \\
 & \left. + \frac{\rho_x k (Z_{n,x}^2 - 1)}{8h_x^2} D_x^2 + \frac{\rho_y k (\tilde{Z}_{n,y}^2 - 1)}{8h_y^2} D_y^2 + \frac{\sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y}}{4h_x h_y} D_{xy} \right) V^n, \quad (2.4)
 \end{aligned}$$

but, as we will see, this scheme is stable only under a restrictive condition on the timestep.

## 2.2 Main convergence results

The following theorems describe the mean-square stability and convergence of the implicit finite difference scheme (2.2) and the ADI scheme (2.3). We make the following assumption:

**Assumption 2.1** Let  $0 \leq \rho_x, \rho_y < 1$ ,  $-1 \leq \rho_{xy} \leq 1$  such that

$$2\rho_x^2(1 + 2|\rho_{xy}|) < 1, \quad (2.5a)$$

$$2\rho_y^2(1 + 2|\rho_{xy}|) < 1, \quad (2.5b)$$

$$2\rho_x\rho_y(3\rho_{xy}^2 + 2|\rho_{xy}| + 1) < 1. \quad (2.5c)$$

In Sect. 3 we show that Assumption 2.1 is a sufficient condition for stability of the schemes (2.2) and (2.3).<sup>2</sup> If  $\rho_{xy} = 0$ , these conditions reduce to  $2\rho_x^2 \leq 1$  and  $2\rho_y^2 \leq 1$ , which is analogous to the condition for mean-square stability in the 1-dimensional case in [35]. In the worst case,  $|\rho_{xy}| = 1$ , sufficient conditions are  $\rho_x, \rho_y \leq 1/\sqrt{6}$ , and  $\rho_x\rho_y \leq 1/12$ .

First, we recall the setting of our numerical schemes. For  $T > 0$  fixed, we discretise  $[0, T]$  by  $N$  steps, and let  $k = T/N$  be the timestep. We discretise space  $\mathbb{R}^2$  with mesh sizes  $h_x$  and  $h_y$ .

The following theorem shows the convergence of the implicit Milstein difference scheme (2.2). The constant  $\theta \in (0, 1)$  therein is determined by the parameters  $\rho_x, \rho_y$  and  $\rho_{xy}$ . The proof of Lemma 4.5 gives an explicit value; though not sharp, this is sufficient to highlight the divergence for  $h_x, h_y \rightarrow 0$  when  $k$  is fixed.

**Theorem 2.1** *Let  $T > 0$ ,  $k = T/N$ ,  $h_x > 0$  and  $h_y > 0$  be mesh sizes. Then, under Assumption 2.1, there exists  $\theta \in (0, 1)$ , independent of  $h_x, h_y$  and  $k$ , such that the implicit Milstein finite difference scheme (2.2) has the error expansion*

$$\begin{aligned} V_{i,j}^N - v(T, x_i, y_j) &= k E_1(T, x_i, y_j) + h_x^2 E_2(T, x_i, y_j) \\ &\quad + h_y^2 E_3(T, x_i, y_j) + \theta^N h_x^{-2} E_4(T, x_i, y_j) \\ &\quad + \theta^N h_y^{-2} E_5(T, x_i, y_j) \\ &\quad + o(k, h_x^2, h_y^2, \theta^N h_x^{-2}, \theta^N h_y^{-2}) R(T, x_i, y_j), \end{aligned} \quad (2.6)$$

where  $x_i = ih_x$ ,  $y_j = jh_y$ ,  $E_1, \dots, E_5$ , and  $R$  are random variables with bounded first and second moments, all independent of  $h_x, h_y$  and  $k$ .

**Proof** See Sect. 4. □

**Remark 2.1** In the setting of Theorem 2.1 with  $\hat{h} = \min\{h_x, h_y\}$ , if

$$\theta^{\frac{T}{k}} \leq 2^{-C_0} \cdot \hat{h}^{4+\beta},$$

for some  $\beta, C_0 > 0$  independent of  $h_x, h_y$  and  $k$ , or, equivalently,

$$k \leq \frac{T \log_2(\theta^{-1})}{C_0 + (4 + \beta) \log_2(\hat{h}^{-1})}, \quad (2.7)$$

then the implicit Milstein scheme (2.2) has the error expansion

<sup>2</sup> We do not believe it to be necessary, due to estimates made in the derivation and as evidenced by numerical tests.

$$\begin{aligned} V_{i,j}^N - v(T, x_i, y_j) &= k E_1(T, x_i, y_j) + h_x^2 E_2(T, x_i, y_j) \\ &\quad + h_y^2 E_3(T, x_i, y_j) + o(k, h_x^2, h_y^2) R(T, x_i, y_j). \end{aligned}$$

**Corollary 2.1** *Under the conditions of Theorem 2.1 and Remark 2.1, the error of the implicit Milstein scheme (2.2) at time  $T$  satisfies, for all  $i, j \in \mathbb{Z}$ ,*

$$\sqrt{\mathbb{E}[|V_{i,j}^N - v(T, ih_x, jh_y)|^2]} = O(h_x^2) + O(h_y^2) + O(k). \quad (2.8)$$

For the ADI discretisation scheme (2.3), a similar convergence result holds.

**Theorem 2.2** *Under the conditions of Remark 2.1, the error of the ADI scheme (2.3) has the same order as for the implicit Milstein scheme,*

$$\begin{aligned} V_{i,j}^N - v(T, x_i, y_j) &= k E_1(T, x_i, y_j) + h_x^2 E_2(T, x_i, y_j) \\ &\quad + h_y^2 E_3(T, x_i, y_j) + o(k, h_x^2, h_y^2) R(T, x_i, y_j), \end{aligned}$$

where  $E_1, E_2, E_3$  and  $R$  are random variables with bounded first and second moments.

**Proof** See Sect. 4.  $\square$

Theorems 2.1 and 2.2 state the convergence pointwise in space and  $L_2$  in probability. The discrete  $L_2$  error in space and (continuous) in probability is defined as

$$\begin{aligned} &\sqrt{\mathbb{E}\left[\left|\sqrt{\sum_{i,j} [V_{i,j}^N - v(T, ih_x, jh_y)]^2} h_x h_y\right|^2\right]} \\ &= \sqrt{\sum_{i,j} \mathbb{E}[|V_{i,j}^N - v(T, ih_x, jh_y)|^2] h_x h_y}, \end{aligned} \quad (2.9)$$

where the term within the outer  $|\cdot|$  is the discrete spatial  $L_2$  norm. Applying Parseval's theorem to (2.9), we get:

**Corollary 2.2** *Under the conditions of Theorem 2.1, the discrete  $L_2$  error in space and  $L_2$  in probability of the implicit Milstein scheme (2.2) at time  $T$ , defined in (2.9), satisfies,*

$$\begin{aligned} &\sqrt{\sum_{i,j} \mathbb{E}[|V_{i,j}^N - v(T, ih_x, jh_y)|^2] h_x h_y} \\ &= O(h_x^2) + O(h_y^2) + O(k) + O(\theta^N h_x^{-1/2} h_y^{-1/2}). \end{aligned} \quad (2.10)$$

*If the initial condition lies in  $L_2$ , then*

$$\sqrt{\sum_{i,j} \mathbb{E}[|V_{i,j}^N - v(T, ih_x, jh_y)|^2] h_x h_y} = O(h_x^2) + O(h_y^2) + O(k).$$

**Proof** See Sect. 4.  $\square$

### 3 Fourier analysis of mean-square stability

Recall the SPDE (1.7),

$$\begin{aligned} dv = & \left[ -\mu_x \frac{\partial v}{\partial x} - \mu_y \frac{\partial v}{\partial y} + \frac{1}{2} \left( \frac{\partial^2 v}{\partial x^2} + 2\sqrt{\rho_x \rho_y} \rho_{xy} \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 v}{\partial y^2} \right) \right] dt \\ & - \sqrt{\rho_x} \frac{\partial v}{\partial x} dM_t^x - \sqrt{\rho_y} \frac{\partial v}{\partial y} dM_t^y. \end{aligned} \quad (3.1)$$

Define the Fourier transform pair

$$\begin{aligned} \tilde{v}(t, \xi, \eta) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v(t, x, y) e^{-i\xi x - i\eta y} dx dy, \\ v(t, x, y) &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{v}(t, \xi, \eta) e^{i\xi x + i\eta y} d\xi d\eta. \end{aligned}$$

The Fourier transform of (3.1) yields

$$d\tilde{v} = - \left( (i\mu_x \xi + i\mu_y \eta + \frac{1}{2} \xi^2 + \sqrt{\rho_x \rho_y} \rho_{xy} \xi \eta + \frac{1}{2} \eta^2) dt + i\sqrt{\rho_x} \xi dM_t^x + i\sqrt{\rho_y} \eta dM_t^y \right) \tilde{v}, \quad (3.2)$$

subject to the initial data  $\tilde{v}(0) = e^{-i\xi x_0 - i\eta y_0}$ . For the remainder of the analysis, we take  $\mu_x = \mu_y = 0$ . This does not alter the results (see Remark 2.3 in [35] for the 1d case).

The solution to (3.2) is

$$\tilde{v}(t) = X(t) e^{-i\xi x_0 - i\eta y_0}, \quad (3.3)$$

where

$$X(t) = \exp \left( -\frac{1}{2} (1 - \rho_x) \xi^2 t - \frac{1}{2} (1 - \rho_y) \eta^2 t - i\xi \sqrt{\rho_x} M_t^x - i\eta \sqrt{\rho_y} M_t^y \right). \quad (3.4)$$

From this we see that  $\mathbb{E}[|X(t)|^2]$  goes to zero exponentially as  $t \rightarrow \infty$ , as well as  $\mathbb{E}[\|\tilde{v}(t, \cdot, \cdot)\|_{L_2}^2]$  and, by isometry,  $\mathbb{E}[\|v(t, \cdot, \cdot)\|_{L_2}^2]$ . The latter holds not only for Dirac initial data, but also for initial data in  $L_2$ . We therefore say that  $v = 0$  is a mean-square stable equilibrium solution.

For the numerical solution, we can use a discrete-continuous Fourier decomposition (note that we approximate  $v(nk, ih_x, jh_y)$  by  $V_{i,j}^n$ )

$$V_{i,j}^0 = \frac{1}{4\pi^2 h_x h_y} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \tilde{V}^0(u, v) e^{i((i-i_0)u + (j-j_0)v)} du dv,$$

where  $i_0 = x_0/h_x$ ,  $j_0 = y_0/h_y$ , and

$$\tilde{V}^0(u, v) = h_x h_y \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} V_{i,j}^0 e^{i(-(i-i_0)u - (j-j_0)v)}.$$

From (2.1),  $V_{i,j}^0 = h_x^{-1}h_y^{-1}\delta_{(i_0, j_0)}$ , we have  $\tilde{V}^0(u, v) = 1$  for all  $(u, v) \in \mathbb{R}^2$ . Similarly for  $n$ -th time-step,

$$\begin{aligned} V_{i,j}^n &= \frac{1}{4\pi^2 h_x h_y} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \tilde{V}^n(u, v) e^{i((i-i_0)u + (j-j_0)v)} du dv \\ &= \frac{1}{4\pi^2} \int_{-\frac{\pi}{h_y}}^{\frac{\pi}{h_y}} \int_{-\frac{\pi}{h_x}}^{\frac{\pi}{h_x}} \tilde{V}^n(\xi, \eta) e^{i((i-i_0)\xi h_x + (j-j_0)\eta h_y)} d\xi d\eta. \end{aligned} \quad (3.5)$$

In the last step, we integrate by substitution,  $\xi = u/h_x$ ,  $\eta = v/h_y$ .

By analogy with the theoretical solution  $\tilde{v}(t) = X(t)\tilde{v}(0)$ , we make the ansatz

$$\tilde{V}^n(\xi, \eta) = X_n(\xi, \eta)\tilde{V}^0(\xi, \eta), \quad (3.6)$$

but as  $\tilde{V}^0(\xi, \eta) = 1$  we simply have  $\tilde{V}^n(\xi, \eta) = X_n(\xi, \eta)$ . We can regard  $X_n(\xi, \eta)$  as the numerical approximation to  $X(nk)$  in (3.4).

We say that the equilibrium solution zero of the scheme is asymptotically mean-square stable, provided for any  $(\xi, \eta) \in [-\pi/h_x, \pi/h_x] \times [-\pi/h_y, \pi/h_y]$ ,

$$\lim_{n \rightarrow \infty} \mathbb{E} [|X_n(\xi, \eta)|^2] = 0. \quad (3.7)$$

This concept has been defined in the context of systems of SDEs in Definition 2.2, 3., in [5], and we apply it here to a fixed wave number in the Fourier domain. A generalisation to SPDEs is analysed in [31] (see Definition 2.1 therein). We will show convergence in  $L_2$  (for fixed  $T$ ) directly under the same conditions; see [35] for mean-square stability and convergence of a 1-d parabolic SPDE. If (3.7) holds without any restriction between  $h_x$ ,  $h_y$  and  $k$ , we call it unconditionally stable. This leads to three conditions summarised in Assumption 2.1, as shown by the following.

**Theorem 3.1** *The implicit Milstein finite difference scheme (2.2) is unconditionally stable in the mean-square sense of (3.7) provided Assumption 2.1 holds.*

**Proof** By inserting (3.5) and (3.6) in (2.2), we have

$$\begin{aligned} X_{n+1}(\xi, \eta) &= \frac{1}{1 - (a_x + a_y)k} \left( 1 - ic_x \sqrt{\rho_x k} Z_{n,x} - ic_y \sqrt{\rho_y k} \tilde{Z}_{n,y} \right. \\ &\quad + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) \\ &\quad \left. + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y} \right) X_n(\xi, \eta), \end{aligned} \quad (3.8)$$

where

$$a_x = -\frac{2 \sin^2 \frac{\xi h_x}{2}}{h_x^2}, \quad b_x = -\frac{\sin^2 \xi h_x}{2h_x^2}, \quad c_x = \frac{\sin \xi h_x}{h_x}, \quad d = -\frac{\sin \xi h_x \sin \eta h_y}{h_x h_y}, \quad (3.9a)$$

$$a_y = -\frac{2 \sin^2 \frac{\eta h_y}{2}}{h_y^2}, \quad b_y = -\frac{\sin^2 \eta h_y}{2 h_y^2}, \quad c_y = \frac{\sin \eta h_y}{h_y}. \quad (3.9b)$$

Given the time-homogeneity of (3.8) and noting that the bracketed term in (3.8) and  $X_n$  are independent,  $\mathbb{E}|X_{n+1}|^2/\mathbb{E}|X_n|^2$  is independent of  $n$ . Hence, to ensure mean-square stability, it is necessary and sufficient that for any  $(\xi, \eta)$  we have

$$\mathbb{E} \left| \frac{1 - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y}}{1 - k(a_x + a_y)} \right|^2 < 1.$$

This is equivalent to

$$\begin{aligned} & \mathbb{E} \left[ \left( 1 + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y} \right)^2 \right. \\ & \quad \left. + \left( c_x \sqrt{\rho_x k} Z_{n,x} + c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} \right)^2 \right] \\ & < (1 - k(a_x + a_y))^2. \end{aligned} \quad (3.10)$$

Note that  $\tilde{Z}_n^y = \rho_{xy} Z_n^x + \sqrt{1 - \rho_{xy}^2} Z_n^y$ , with  $Z_n^x, Z_n^y \sim N(0, 1)$  being independent normal random variables, hence

$$\mathbb{E}[Z_{n,x}^2 \tilde{Z}_{n,y}^2] = 1 + 2\rho_{xy}^2, \quad \mathbb{E}[Z_{n,x} \tilde{Z}_{n,y}] = \rho_{xy}, \quad \mathbb{E}[Z_{n,x}^3 \tilde{Z}_{n,y}] = \mathbb{E}[Z_{n,x} \tilde{Z}_{n,y}^3] = 3\rho_{xy},$$

and

$$\begin{aligned} & \mathbb{E} \left[ \left( 1 + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y} \right)^2 \right] \\ & = 1 + b_x^2 \rho_x^2 k^2 \mathbb{E}[(Z_{n,x}^2 - 1)^2] + b_y^2 \rho_y^2 k^2 \mathbb{E}[(\tilde{Z}_{n,y}^2 - 1)^2] + d^2 \rho_x \rho_y k^2 \mathbb{E}[Z_{n,x}^2 \tilde{Z}_{n,y}^2] \\ & \quad + 2b_x \rho_x k \mathbb{E}[Z_{n,x}^2 - 1] + 2b_y \rho_y k \mathbb{E}[\tilde{Z}_{n,y}^2 - 1] + 2d \sqrt{\rho_x \rho_y k} \mathbb{E}[Z_{n,x} \tilde{Z}_{n,y}] \\ & \quad + 2b_x b_y \rho_x \rho_y k^2 \mathbb{E}[(Z_{n,x}^2 - 1)(\tilde{Z}_{n,y}^2 - 1)] + 2b_x d \rho_x \sqrt{\rho_x \rho_y k^2} \mathbb{E}[(Z_{n,x}^2 - 1) Z_{n,x} \tilde{Z}_{n,y}] \\ & \quad + 2b_y d \rho_y \sqrt{\rho_x \rho_y k^2} \mathbb{E}[(\tilde{Z}_{n,y}^2 - 1) Z_{n,x} \tilde{Z}_{n,y}] \\ & = 1 + 2b_x^2 \rho_x^2 k^2 + 2b_y^2 \rho_y^2 k^2 + d^2 \rho_x \rho_y (1 + 2\rho_{xy}^2) k^2 + 2d \sqrt{\rho_x \rho_y} \rho_{xy} k + 4b_x b_y \rho_x \rho_y \rho_{xy}^2 k^2 \\ & \quad + 4b_x d \rho_x \sqrt{\rho_x \rho_y} \rho_{xy} k^2 + 4b_y d \rho_y \sqrt{\rho_x \rho_y} \rho_{xy} k^2, \end{aligned}$$

$$\text{and } \mathbb{E} \left[ (c_x \sqrt{\rho_x k} Z_{n,x} + c_y \sqrt{\rho_y k} \tilde{Z}_{n,y})^2 \right] = c_x^2 \rho_x k + c_y^2 \rho_y k + 2c_x c_y \sqrt{\rho_x \rho_y} \rho_{xy} k.$$

Note that  $c_x c_y + d = 0$ ,  $4b_x b_y = d^2$ , and

$$\begin{aligned} b_x &= \cos^2 \frac{\xi h_x}{2} a_x, \quad b_y = \cos^2 \frac{\eta h_y}{2} a_y, \quad c_x^2 = -2 \cos^2 \frac{\xi h_x}{2} a_x, \\ c_y^2 &= -2 \cos^2 \frac{\eta h_y}{2} a_y, \quad d^2 = 4 \cos^2 \frac{\xi h_x}{2} \cos^2 \frac{\eta h_y}{2} a_x a_y. \end{aligned}$$

Therefore

$$\begin{aligned}
 & \mathbb{E} \left[ \left( 1 + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y} k Z_{n,x} \tilde{Z}_{n,y} \right)^2 \right. \\
 & \quad \left. + \left( c_x \sqrt{\rho_x} k Z_{n,x} + c_y \sqrt{\rho_y} k \tilde{Z}_{n,y} \right)^2 \right] \\
 &= 1 + 2 \cos^4 \frac{\xi h_x}{2} a_x^2 \rho_x^2 k^2 + 2 \cos^4 \frac{\eta h_y}{2} a_y^2 \rho_y^2 k^2 \\
 & \quad + 4 \cos^2 \frac{\xi h_x}{2} \cos^2 \frac{\eta h_y}{2} \rho_x \rho_y (1 + 3 \rho_{xy}^2) a_x a_y k^2 \\
 & \quad + 4 d \sqrt{\rho_x \rho_y} \rho_{xy} k^2 \left( \cos^2 \frac{\xi h_x}{2} a_x \rho_x + \cos^2 \frac{\eta h_y}{2} a_y \rho_y \right) \\
 & \quad - 2 \cos^2 \frac{\xi h_x}{2} a_x \rho_x k - 2 \cos^2 \frac{\eta h_y}{2} a_y \rho_y k \\
 &\leq 1 + 2 \cos^4 \frac{\xi h_x}{2} a_x^2 \rho_x^2 k^2 + 2 \cos^4 \frac{\eta h_y}{2} a_y^2 \rho_y^2 k^2 \\
 & \quad + 4 \cos^2 \frac{\xi h_x}{2} \cos^2 \frac{\eta h_y}{2} \rho_x \rho_y (1 + 3 \rho_{xy}^2) a_x a_y k^2 \\
 & \quad + 4 |\rho_{xy}| k^2 \left( \cos^2 \frac{\xi h_x}{2} a_x \rho_x + \cos^2 \frac{\eta h_y}{2} a_y \rho_y \right)^2 \\
 & \quad - 2 \cos^2 \frac{\xi h_x}{2} a_x \rho_x k - 2 \cos^2 \frac{\eta h_y}{2} a_y \rho_y k \\
 &\leq 1 + 2 \rho_x^2 (1 + 2 |\rho_{xy}|) a_x^2 k^2 + 2 \rho_y^2 (1 + 2 |\rho_{xy}|) a_y^2 k^2 \\
 & \quad + 4 \rho_x \rho_y (3 \rho_{xy}^2 + 2 |\rho_{xy}| + 1) a_x a_y k^2 \\
 & \quad - 2 \cos^2 \frac{\xi h_x}{2} a_x \rho_x k - 2 \cos^2 \frac{\eta h_y}{2} a_y \rho_y k,
 \end{aligned}$$

and  $(1 - k(a_x + a_y))^2 = 1 + k^2 a_x^2 - 2ka_x + k^2 a_y^2 - 2ka_y + 2k^2 a_x a_y$ .

One sufficient condition for (3.10) to hold is

$$\begin{aligned}
 & 2 \rho_x^2 (1 + 2 |\rho_{xy}|) a_x^2 k^2 - 2 \cos^2 \frac{\xi h_x}{2} a_x \rho_x k < k^2 a_x^2 - 2ka_x, \\
 & 2 \rho_y^2 (1 + 2 |\rho_{xy}|) a_y^2 k^2 - 2 \cos^2 \frac{\eta h_y}{2} a_y \rho_y k < k^2 a_y^2 - 2ka_y, \\
 & 4 \rho_x \rho_y (3 \rho_{xy}^2 + 2 |\rho_{xy}| + 1) a_x a_y k^2 \leq 2k^2 a_x a_y.
 \end{aligned}$$

Replacing  $a_x, a_y$  with their expressions in (3.9), the above is equivalent to

$$\begin{aligned}
 & \frac{k}{h_x^2} \left( 2 \rho_x^2 (1 + 2 |\rho_{xy}|) - 1 \right) \sin^2 \frac{\xi h_x}{2} + \rho_x \cos^2 \frac{\xi h_x}{2} < 1, \\
 & \frac{k}{h_y^2} \left( 2 \rho_y^2 (1 + 2 |\rho_{xy}|) - 1 \right) \sin^2 \frac{\eta h_y}{2} + \rho_y \cos^2 \frac{\eta h_y}{2} < 1, \\
 & 2 \rho_x \rho_y (3 \rho_{xy}^2 + 2 |\rho_{xy}| + 1) \leq 1.
 \end{aligned}$$

A sufficient condition for this is that  $\rho_x, \rho_y, \rho_{xy}$  satisfy (2.5).  $\square$

Now we prove the stability of the ADI scheme.

**Proposition 3.1** *Under Assumption 2.1, mean-square stability (3.7) also holds for the ADI scheme (2.3).*

**Proof** By insertion in (2.3), we have

$$\begin{aligned} X_{n+1}(\xi, \eta) = & \frac{1}{(1 - a_x k)(1 - a_y k)} \left( 1 - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} \right. \\ & + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) \\ & \left. + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y} \right) X_n(\xi, \eta). \end{aligned} \quad (3.11)$$

As in the proof of Theorem 3.1, we need  $\mathbb{E}|X_{n+1}|^2 / \mathbb{E}|X_n|^2 < 1$  to achieve stability. Since  $|1 - (a_x + a_y)k| \leq |(1 - a_x k)(1 - a_y k)|$  for  $a_x, a_y \leq 0$ , the stability also holds for the ADI scheme.  $\square$

**Proposition 3.2** *The explicit Milstein (finite difference) scheme (2.4) is stable in the mean-square sense provided*

$$\frac{k}{h_x^2} \leq (2 + 2\rho_x^2 + 2\rho_x \rho_y + (3\rho_x + \rho_y + 4\rho_x^2 + 4\rho_x \rho_y)|\rho_{xy}| + 6\rho_x \rho_y \rho_{xy}^2)^{-1}, \quad (3.12a)$$

$$\frac{k}{h_y^2} \leq (2 + 2\rho_y^2 + 2\rho_x \rho_y + (\rho_x + 3\rho_y + 4\rho_y^2 + 4\rho_x \rho_y)|\rho_{xy}| + 6\rho_x \rho_y \rho_{xy}^2)^{-1}. \quad (3.12b)$$

**Proof** By insertion in (2.4), we have

$$\begin{aligned} X_{n+1}(\xi, \eta) = & \left( 1 + (a_x + a_y)k - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) \right. \\ & \left. + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y} \right) X_n(\xi, \eta). \end{aligned} \quad (3.13)$$

To ensure mean-square stability in this case, similar to before, we need

$$\begin{aligned} \mathbb{E} \left[ \left| 1 + (a_x + a_y)k - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) \right. \right. \\ \left. \left. + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y} \right|^2 \right] < 1. \end{aligned}$$



For simplicity, we denote  $u = |\sin \frac{\xi h_x}{2}|$ ,  $v = |\sin \frac{\eta h_y}{2}|$ , then we have

$$\begin{aligned}
 & \mathbb{E} \left[ \left| 1 + (a_x + a_y)k - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) \right. \right. \\
 & \quad \left. \left. + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y} k Z_{n,x} \tilde{Z}_{n,y} \right|^2 \right] \\
 &= 1 + (a_x^2 k^2 + 2a_x k + 2b_x^2 \rho_x^2 k^2 + c_x^2 \rho_x k) + (a_y^2 k^2 + 2a_y k + 2b_y^2 \rho_y^2 k^2 + c_y^2 \rho_y k) \\
 & \quad + (2a_x a_y k^2 + \rho_x \rho_y (1 + 3\rho_{xy}^2) d^2 k^2) \\
 & \quad + 2d(a_x + 2b_x \rho_x + a_y + 2b_y \rho_y) \sqrt{\rho_x \rho_y} \rho_{xy} k^2 \\
 &\leq 1 - 4 \frac{k}{h_x^2} u^2 \left( 1 - \rho_x (1 - u^2) - \frac{k}{h_x^2} u^2 (1 + 2\rho_x^2) \right) \\
 & \quad - 4 \frac{k}{h_y^2} v^2 \left( 1 - \rho_y (1 - v^2) - \frac{k}{h_y^2} v^2 (1 + 2\rho_y^2) \right) \\
 & \quad + 8 \frac{k^2}{h_x^2 h_y^2} u^2 v^2 \left( 1 + 2\rho_x \rho_y (1 + 3\rho_{xy}^2) \right) + 8 |\rho_{xy}| \left( \frac{k}{h_x^2} u^2 \rho_x + \frac{k}{h_y^2} v^2 \rho_y \right) \\
 & \quad \times \left( \frac{k}{h_x^2} u^2 (1 + 2\rho_x) + \frac{k}{h_y^2} v^2 (1 + 2\rho_y) \right) \\
 &\leq 1 - 4 \frac{k}{h_x^2} u^2 \left[ 1 - \rho_x (1 - u^2) - \frac{k}{h_x^2} u^2 \left( 2 + 2\rho_x^2 + 2\rho_x \rho_y \right. \right. \\
 & \quad \left. \left. + (3\rho_x + \rho_y + 4\rho_x^2 + 4\rho_x \rho_y) |\rho_{xy}| + 6\rho_x \rho_y \rho_{xy}^2 \right) \right] \\
 & \quad - 4 \frac{k}{h_y^2} v^2 \left[ 1 - \rho_y (1 - v^2) - \frac{k}{h_y^2} v^2 \left( 2 + 2\rho_y^2 + 2\rho_x \rho_y + (\rho_x + 3\rho_y + 4\rho_y^2 \right. \right. \right. \\
 & \quad \left. \left. + 4\rho_x \rho_y) |\rho_{xy}| + 6\rho_x \rho_y \rho_{xy}^2 \right) \right] \\
 &< 1, \quad \text{for all } 0 \leq u, v < 1.
 \end{aligned}$$

This leads to the two sufficient conditions in (3.12).  $\square$

It follows that if  $\rho_{xy} = 0$ , the stability conditions are

$$\begin{aligned}
 \frac{k}{h_x^2} &\leq (2 + 2\rho_x^2 + 2\rho_x \rho_y)^{-1}, \\
 \frac{k}{h_y^2} &\leq (2 + 2\rho_y^2 + 2\rho_x \rho_y)^{-1}.
 \end{aligned}$$

So it is sufficient that  $k/h_x^2 \leq 1/6$ , and  $k/h_y^2 \leq 1/6$ . If  $|\rho_{xy}| = 1$ , the worst case in (3.12), the stability conditions are

$$\begin{aligned}\frac{k}{h_x^2} &\leq (2 + 3\rho_x + \rho_y + 6\rho_x^2 + 12\rho_x\rho_y)^{-1}, \\ \frac{k}{h_y^2} &\leq (2 + \rho_x + 3\rho_y + 6\rho_y^2 + 12\rho_x\rho_y)^{-1}.\end{aligned}$$

So it is sufficient to ensure  $k/h_x^2 \leq 1/24$ , and  $k/h_y^2 \leq 1/24$ .

## 4 Fourier analysis of $L_2$ -convergence

Extending the analysis in [7] for the standard 1-d (deterministic) heat equation to our 2-d SPDE, we compare the numerical solution to the exact solution in Fourier space first by splitting the Fourier domain into two wave number regions. Assume  $p$  is a constant satisfying  $0 < p < \frac{1}{4}$ . Then we define the low wave number region by

$$\Omega_{\text{low}} = \{(\xi, \eta) : |\xi| \leq \min\{h_x^{-2p}, k^{-p}\} \text{ and } |\eta| \leq \min\{h_y^{-2p}, k^{-p}\}\},$$

and the high wave number region by

$$\begin{aligned}\Omega_{\text{high}} &= \{(\xi, \eta) : |\xi| > \min\{h_x^{-2p}, k^{-p}\} \text{ or } |\eta| \\ &> \min\{h_y^{-2p}, k^{-p}\}\} \cap [-\pi h_x^{-1}, \pi h_x^{-1}] \times [-\pi h_y^{-1}, \pi h_y^{-1}].\end{aligned}$$

Note that both  $X_n$  and  $X(nk)$  are functions of  $\xi$  and  $\eta$ . The idea of the convergence proof is that  $X_n$  is a good approximation to  $X(nk)$  in the low wave region, and they both damp exponentially in the high wave region.

**Lemma 4.1** For  $(\xi, \eta) \in \Omega_{\text{low}}$ , we have

$$\begin{aligned}X_N - X(T) &= X(T) \cdot \left( h_x^2 f_1(\xi) + h_y^2 f_2(\eta) + k f_3(\xi, \eta) \right. \\ &\quad \left. + o(k, h_x^2, h_y^2) \cdot \varphi(T, h_x \xi, h_y \eta) \right),\end{aligned}$$

where  $f_1(\xi)$ ,  $f_2(\eta)$ ,  $f_3(\xi, \eta)$ ,  $\varphi(T, h_x \xi, h_y \eta)$  are random variables such that after multiplication by  $X(T)$ , the integral over  $\Omega_{\text{low}}$  has bounded first and second moments independent of  $N$ .

**Proof** See Sect. 4.1. □

**Lemma 4.2** Under Assumption 2.1, there exists  $C > 0$  independent of  $h_x$ ,  $h_y$ , and  $k$ , such that

$$\sqrt{\mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) d\xi d\eta \right|^2 \right]} \leq C h_x^{-2\theta^N} + C h_y^{-2\theta^N},$$

where  $0 < \theta < 1$  is independent of  $h_x$ ,  $h_y$ , and  $k$ .

**Proof** See Sect. 4.2. □

The following proof of Theorem 2.1 now deduces mean square convergence of the implicit finite difference scheme (2.2).

**Proof of Theorem 2.1** By Lemma 4.1 and Lemma 4.2, the inverse Fourier transform gives

$$\begin{aligned} V_{i,j}^N - v(T, ih_x, jh_y) &= \frac{1}{4\pi^2} \iint_{\Omega_{\text{low}} \cup \Omega_{\text{high}}} (X_N - X(T)) e^{i((i-i_0)\xi h_x + (j-j_0)\eta h_y)} d\xi d\eta + o(k) \\ &= k E_1(T, x_i, y_j) + h_x^2 E_2(T, x_i, y_j) + h_y^2 E_3(T, x_i, y_j) + \theta^N h_x^{-2} E_4(T, x_i, y_j) \\ &\quad + \theta^N h_y^{-2} E_5(T, x_i, y_j) + o(k, h_x^2, h_y^2, \theta^N h_x^{-2}, \theta^N h_y^{-2}) R(T, x_i, y_j), \end{aligned}$$

where  $x_i = ih_x$ ,  $y_j = jh_y$ ,  $E_1, \dots, E_5$ , and  $R$  are random variables with bounded first and second moments,  $N = T/k$ , and  $0 < \theta < 1$  is independent of  $h_x, h_y$  and  $k$ . □

Next we give a proof of Corollary 2.2, the  $L_2$  convergence in space and probability of the implicit finite difference scheme (2.2).

**Proof of Corollary 2.2** We apply Parseval's theorem to  $V_{i,j}^N - v(T, ih_x, jh_y)$  and its Fourier transform. It follows

$$\begin{aligned} \sum_{i,j} \left| V_{i,j}^N - v(T, ih_x, jh_y) \right|^2 h_x h_y &= \iint |\tilde{v}(0, \xi, \eta)|^2 |X(T, \xi, \eta) - X_N(\xi, \eta)|^2 d\xi d\eta + O(h_x^4) + O(h_y^4). \end{aligned}$$

In Lemma 4.1, we have proved  $|X(T, \xi, \eta) - X_N(\xi, \eta)| = X(T)(1 + O(h_x^2) + O(h_y^2) + O(k))$  for  $(\xi, \eta)$  in the low wave region. In Lemma 4.2, we have proved  $|X(T, \xi, \eta) - X_N(\xi, \eta)|^2 \leq C\theta^N$ , for some  $C > 0$  and  $\theta \in (0, 1)$ ,  $(\xi, \eta)$  in the high wave region.

As for Dirac initial datum,  $|\tilde{v}(0, \xi, \eta)| = 1$ , we have

$$\begin{aligned} \iint |X(T, \xi, \eta) - X_N(\xi, \eta)|^2 d\xi d\eta &= \iint_{\Omega_{\text{low}}} |X(T, \xi, \eta) - X_N(\xi, \eta)|^2 d\xi d\eta + \iint_{\Omega_{\text{high}}} |X(T, \xi, \eta) - X_N(\xi, \eta)|^2 d\xi d\eta \\ &= O(h_x^2) + O(h_y^2) + O(k) + O(h_x^{-1} h_y^{-1} \theta^N). \end{aligned}$$

For initial data in  $L_2$ ,  $\iint |\tilde{v}(0, \xi, \eta)|^2 d\xi d\eta < \infty$ , and therefore

$$\iint_{\Omega_{\text{high}}} |\tilde{v}(0, \xi, \eta)|^2 |X(T, \xi, \eta) - X_N(\xi, \eta)|^2 d\xi d\eta = o(k^r) \quad \text{for any } r > 0,$$

and consequently

$$\iint |X(T, \xi, \eta) - X_N(\xi, \eta)|^2 d\xi d\eta = O(h_x^2) + O(h_y^2) + O(k).$$

□

#### 4.1 Low wave number region (proof of Lemma 4.1)

For the low wave region, we consider the case where both  $\xi$ ,  $\eta$  are small. It follows from (3.4) that the exact solution of  $X(t_{n+1})$  given  $X(t_n)$  is

$$X(t_{n+1}) = X(t_n) \exp \left( -\frac{1}{2}(1-\rho_x)\xi^2 k - \frac{1}{2}(1-\rho_y)\eta^2 k - i\xi\sqrt{\rho_x k}Z_{n,x} - i\eta\sqrt{\rho_y k}\tilde{Z}_{n,y} \right),$$

where  $M_{t_{n+1}}^x - M_{t_n}^x \equiv \sqrt{k}Z_{n,x}$ ,  $M_{t_{n+1}}^y - M_{t_n}^y \equiv \sqrt{k}\tilde{Z}_{n,y}$  are the Brownian increments.

Now we consider  $X_n$ , the numerical approximation of  $X(nk)$ . Let

$$X_{n+1} = C_n X_n,$$

where

$$C_n = \exp \left( -\frac{1}{2}(1-\rho_x)\xi^2 k - \frac{1}{2}(1-\rho_y)\eta^2 k - i\xi\sqrt{\rho_x k}Z_{n,x} - i\eta\sqrt{\rho_y k}\tilde{Z}_{n,y} + e_n \right), \quad (4.1)$$

and  $e_n$  is the logarithmic error between the numerical solution and the exact solution introduced during  $[nk, (n+1)k]$ . Aggregating over  $N$  time steps, at  $t_N = kN = T$ ,

$$X_N = X(T) \exp \left( \sum_{n=0}^{N-1} e_n \right), \quad (4.2)$$

where

$$\begin{aligned} X(T) = \exp \left( -\frac{1}{2}(1-\rho_x)\xi^2 T - \frac{1}{2}(1-\rho_y)\eta^2 T \right. \\ \left. - i\xi\sqrt{\rho_x k} \sum_{n=0}^{N-1} Z_{n,x} - i\eta\sqrt{\rho_y k} \sum_{n=0}^{N-1} \tilde{Z}_{n,y} \right) \end{aligned}$$

is the exact solution at time  $T$ .

From (4.1), we have

$$e_n = \log C_n + \frac{1}{2}(1-\rho_x)\xi^2 k + \frac{1}{2}(1-\rho_y)\eta^2 k + i\xi\sqrt{\rho_x k}Z_{n,x} + i\eta\sqrt{\rho_y k}\tilde{Z}_{n,y},$$

hence

$$\begin{aligned} \sum_{n=0}^{N-1} e_n &= \sum_{n=0}^{N-1} \log C_n + \frac{1}{2}(1 - \rho_x)\xi^2 T + \frac{1}{2}(1 - \rho_y)\eta^2 T + i\xi\sqrt{\rho_x k} \sum_{n=0}^{N-1} Z_{n,x} \\ &\quad + i\eta\sqrt{\rho_y k} \sum_{n=0}^{N-1} \tilde{Z}_{n,y}. \end{aligned}$$

From (3.8),  $C_n$  has the form

$$C_n = \frac{1 - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y}}{1 - (a_x + a_y)k},$$

where

$$\begin{aligned} a_x &= -\frac{\xi^2}{2} \cdot \frac{\sin^2 \frac{\xi h_x}{2}}{\left(\frac{\xi h_x}{2}\right)^2} = -\frac{\xi^2}{2} + \frac{\xi^4}{24} h_x^2 + O(h_x^4 \xi^6), \\ a_y &= -\frac{\eta^2}{2} \cdot \frac{\sin^2 \frac{\eta h_y}{2}}{\left(\frac{\eta h_y}{2}\right)^2} = -\frac{\eta^2}{2} + \frac{\eta^4}{24} h_y^2 + O(h_y^4 \eta^6), \\ b_x &= -\frac{\xi^2}{2} \cdot \frac{\sin^2 \xi h_x}{\xi^2 h_x^2} = -\frac{\xi^2}{2} + \frac{\xi^4}{6} h_x^2 + O(h_x^4 \xi^6), \\ b_y &= -\frac{\eta^2}{2} \cdot \frac{\sin^2 \eta h_y}{\eta^2 h_y^2} = -\frac{\eta^2}{2} + \frac{\eta^4}{6} h_y^2 + O(h_y^4 \eta^6), \\ c_x &= \xi \cdot \frac{\sin \xi h_x}{\xi h_x} = \xi - \frac{\xi^3}{6} h_x^2 + O(h_x^4 \xi^5), \\ c_y &= \eta \cdot \frac{\sin \eta h_y}{\eta h_y} = \eta - \frac{\eta^3}{6} h_y^2 + O(h_y^4 \eta^5), \\ d &= -\xi \eta \cdot \frac{\sin \xi h_x \sin \eta h_y}{\xi h_x \eta h_y}. \end{aligned}$$

Note that  $c_x c_y + d = 0$ ,  $b_x + \frac{1}{2} c_x^2 = 0$ ,  $b_y + \frac{1}{2} c_y^2 = 0$ , then one can derive by Taylor expansion (by lengthy, but elementary calculations),

$$\begin{aligned} \log C_n &= -i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} \\ &\quad + (a_x + a_y - b_x \rho_x - b_y \rho_y)k + \left(b_x + \frac{1}{2} c_x^2\right) \rho_x k Z_{n,x}^2 \\ &\quad + \left(b_y + \frac{1}{2} c_y^2\right) \rho_y k \tilde{Z}_{n,y}^2 + (c_x c_y + d) \sqrt{\rho_x \rho_y k} Z_{n,x} \tilde{Z}_{n,y} \\ &\quad + O((|\xi| + |\eta|)^3 k \sqrt{k}) \cdot i \phi_1(Z_{n,x}, \tilde{Z}_{n,y}) \\ &\quad + O((|\xi| + |\eta|)^4 k^2) \cdot \phi_2(Z_{n,x}, \tilde{Z}_{n,y}) + o((|\xi| + |\eta|)^4 k^2) \end{aligned}$$

$$\begin{aligned}
&= -ic_x\sqrt{\rho_x k}Z_{n,x} - ic_y\sqrt{\rho_y k}\tilde{Z}_{n,y} + (a_x + a_y - b_x\rho_x - b_y\rho_y)k \\
&\quad + O((|\xi| + |\eta|)^3 k\sqrt{k}) \cdot i\phi_1(Z_{n,x}, \tilde{Z}_{n,y}) + O((|\xi| + |\eta|)^4 k^2) \\
&\quad \cdot \phi_2(Z_{n,x}, \tilde{Z}_{n,y}) + o((|\xi| + |\eta|)^4 k^2),
\end{aligned}$$

where  $\phi_1(\cdot, \cdot)$  is an odd and  $\phi_2(\cdot, \cdot)$  an even degree polynomial. Therefore

$$\begin{aligned}
\sum_{n=0}^{N-1} e_n &= \sum_{n=0}^{N-1} \log C_n + \frac{1}{2}(1 - \rho_x)\xi^2 T + \frac{1}{2}(1 - \rho_y)\eta^2 T \\
&\quad + i\xi\sqrt{\rho_x k} \sum_{n=0}^{N-1} Z_{n,x} + i\eta\sqrt{\rho_y k} \sum_{n=0}^{N-1} \tilde{Z}_{n,y} \\
&= i(\xi - c_x)\sqrt{\rho_x k} \sum_{n=0}^{N-1} Z_{n,x} + i(\eta - c_y)\sqrt{\rho_y k} \sum_{n=0}^{N-1} \tilde{Z}_{n,y} \\
&\quad + \left(a_x + a_y - b_x\rho_x - b_y\rho_y + \frac{1 - \rho_x}{2}\xi^2 + \frac{1 - \rho_y}{2}\eta^2\right)T \\
&\quad + O((|\xi| + |\eta|)^3 k\sqrt{k}) \cdot i \sum_{n=0}^{N-1} \phi_1(Z_{n,x}, \tilde{Z}_{n,y}) \\
&\quad + O((|\xi| + |\eta|)^4 k^2) \cdot \sum_{n=0}^{N-1} \phi_2(Z_{n,x}, \tilde{Z}_{n,y}) + o((|\xi| + |\eta|)^4 k),
\end{aligned}$$

so we have

$$\begin{aligned}
\exp\left(\sum_{n=0}^{N-1} e_n\right) &= \exp\left(\left(a_x + a_y - b_x\rho_x - b_y\rho_y + \frac{1 - \rho_x}{2}\xi^2 + \frac{1 - \rho_y}{2}\eta^2\right)T\right) \\
&\quad \cdot \exp\left(i(\xi - c_x)\sqrt{\rho_x k} \sum_{n=0}^{N-1} Z_{n,x} + i(\eta - c_y)\sqrt{\rho_y k} \sum_{n=0}^{N-1} \tilde{Z}_{n,y}\right) \\
&\quad + O((|\xi| + |\eta|)^3 k\sqrt{k}) \cdot i \sum_{n=0}^{N-1} \phi_1(Z_{n,x}, \tilde{Z}_{n,y}) \\
&\quad + O((|\xi| + |\eta|)^4 k^2) \cdot \sum_{n=0}^{N-1} \phi_2(Z_{n,x}, \tilde{Z}_{n,y}) + o((|\xi| + |\eta|)^4 k).
\end{aligned}$$

Here

$$\begin{aligned}
&\exp\left(\left(a_x + a_y - b_x\rho_x - b_y\rho_y + \frac{1 - \rho_x}{2}\xi^2 + \frac{1 - \rho_y}{2}\eta^2\right)T\right) \\
&= 1 + \frac{\xi^4}{24}h_x^2(1 - 4\rho_x)T + \frac{\eta^4}{24}h_y^2(1 - 4\rho_y)T + O(\xi^6 h_x^4) + O(\eta^6 h_y^4),
\end{aligned}$$

and

$$\begin{aligned}
 & \exp \left( i(\xi - c_x) \sqrt{\rho_x k} \sum_{n=0}^{N-1} Z_{n,x} + i(\eta - c_y) \sqrt{\rho_y k} \sum_{n=0}^{N-1} \tilde{Z}_{n,y} \right. \\
 & \quad + O((|\xi| + |\eta|)^3 k \sqrt{k}) \cdot i \sum_{n=0}^{N-1} \phi_1(Z_{n,x}, \tilde{Z}_{n,y}) \\
 & \quad \left. + O((|\xi| + |\eta|)^4 k^2) \cdot \sum_{n=0}^{N-1} \phi_2(Z_{n,x}, \tilde{Z}_{n,y}) \right) \\
 & = 1 + i(\xi - c_x) \sqrt{\rho_x k} \sum_{n=0}^{N-1} Z_{n,x} + i(\eta - c_y) \sqrt{\rho_y k} \sum_{n=0}^{N-1} \tilde{Z}_{n,y} \\
 & \quad - \frac{1}{2} (\xi - c_x)^2 \rho_x k \left( \sum_{n=0}^{N-1} Z_{n,x} \right)^2 \\
 & \quad - \frac{1}{2} (\eta - c_y)^2 \rho_y k \left( \sum_{n=0}^{N-1} \tilde{Z}_{n,y} \right)^2 + O((|\xi| + |\eta|)^3 k \sqrt{k}) \cdot i \sum_{n=0}^{N-1} \hat{\phi}_1(Z_{n,x}, \tilde{Z}_{n,y}) \\
 & \quad + O((|\xi| + |\eta|)^4 k^2) \cdot \sum_{n=0}^{N-1} \hat{\phi}_2(Z_{n,x}, \tilde{Z}_{n,y}) + o((|\xi| + |\eta|)^4 k),
 \end{aligned}$$

where  $\hat{\phi}_1(\cdot, \cdot)$  is a polynomial function with odd degree, and  $\hat{\phi}_2(\cdot, \cdot)$  are with even degree, and

$$\begin{aligned}
 \mathbb{E} \left[ X(T) \sum_n \hat{\phi}_1(Z_{n,x}, \tilde{Z}_{n,y}) \right] &= O(k^{-\frac{1}{2}}) \exp \left( -\frac{1}{2} (\xi^2 + \eta^2 + 2\xi\eta\sqrt{\rho_x\rho_y\rho_{xy}}) T \right), \\
 \mathbb{E} \left[ X(T) \sum_n \hat{\phi}_2(Z_{n,x}, \tilde{Z}_{n,y}) \right] &= O(k^{-1}) \exp \left( -\frac{1}{2} (\xi^2 + \eta^2 + 2\xi\eta\sqrt{\rho_x\rho_y\rho_{xy}}) T \right), \\
 \mathbb{E} \left| X(T) \sum_n \hat{\phi}_1(Z_{n,x}, \tilde{Z}_{n,y}) \right|^2 &= O(k^{-1}) \exp \left( -(\xi^2 + \eta^2 + 2\xi\eta\sqrt{\rho_x\rho_y\rho_{xy}}) T \right), \\
 \mathbb{E} \left| X(T) \sum_n \hat{\phi}_2(Z_{n,x}, \tilde{Z}_{n,y}) \right|^2 &= O(k^{-2}) \exp \left( -(\xi^2 + \eta^2 + 2\xi\eta\sqrt{\rho_x\rho_y\rho_{xy}}) T \right).
 \end{aligned}$$

Hence we have in the low wave number region,

$$\begin{aligned}
 X_N - X(T) &= X(T) \cdot \left( \exp \left( \sum_{n=0}^{N-1} e_n \right) - 1 \right) \\
 &= X(T) \cdot \left\{ \frac{i}{6} \sqrt{\rho_x} \xi^3 h_x^2 M_T^x + \frac{1}{24} (1 - 4\rho_x) \xi^4 h_x^2 T \right.
 \end{aligned}$$

$$\begin{aligned}
& + \frac{i}{6} \sqrt{\rho_y} \eta^3 h_y^2 \tilde{M}_T^y + \frac{1}{24} (1 - 4\rho_y) \eta^4 h_y^2 T + O((|\xi| + |\eta|)^3 k \sqrt{k}) \\
& \cdot i \sum_{n=0}^{N-1} \hat{\phi}_1(Z_{n,x}, \tilde{Z}_{n,y}) + O((|\xi| + |\eta|)^4 k^2) \cdot \sum_{n=0}^{N-1} \hat{\phi}_2(Z_{n,x}, \tilde{Z}_{n,y}) \\
& + o(k, h_x^2, h_y^2) \cdot \varphi(T, h_x \xi, h_y \eta) \Big\},
\end{aligned}$$

where  $\varphi(T, h_x \xi, h_y \eta)$  is a random variable with bounded moments.

**Remark 4.1** We can derive the exact leading order term by taking the inverse Fourier transform. For instance, the leading order error in  $h_x$  is

$$\left( -\frac{1}{6} \sqrt{\rho_x} M_T^x \frac{\partial^3}{\partial x^3} v(T, x, y) + \frac{1}{24} (1 - 4\rho_x) T \frac{\partial^4}{\partial x^4} v(T, x, y) \right) \cdot h_x^2,$$

and similar for  $h_y$  (replacing ‘x’ by ‘y’); the leading order error in  $k$  can be found by the same technique but is significantly lengthier and hence omitted.

## 4.2 High wave number region (proof of Lemma 4.2)

Now we consider the case when either  $\xi$  or  $\eta$  is large.

First we calculate the upper bound of  $\mathbb{E}[|X_N(\xi, \eta)|^2]$ . To simplify the proof, we take  $h_x = h_y = h$ , and the case where  $h_x \neq h_y$  is similar. Write  $\lambda = \frac{k}{h^2}$ .

**Lemma 4.3** For  $(\xi, \eta) \notin \Omega_{low}$ ,

$$\mathbb{E}[|X_N(\xi, \eta)|^2] \leq |X_0|^2 \left( 1 - 4\beta \frac{\lambda (\sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}) + \lambda^2 (\sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2})^2}{(1 + 2\lambda (\sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}))^2} \right)^N,$$

where

$$\begin{aligned}
\beta = \min \{ & 1 - \rho_x, 1 - \rho_y, 1 - 2\rho_x^2(1 + 2|\rho_{xy}|), 1 - 2\rho_y^2(1 + 2|\rho_{xy}|), \\
& 1 - 2\rho_x \rho_y (1 + 2|\rho_{xy}| + 3\rho_{xy}^2) \} \in (0, 1).
\end{aligned}$$

**Proof** By (3.8), we have  $X_N = X_0 \prod_{n=0}^{N-1} C_n$ , where

$$C_n = \frac{1 - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y} k Z_{n,x} \tilde{Z}_{n,y}}{1 - (a_x + a_y)k},$$

$$\begin{aligned}
a_x &= -\frac{2 \sin^2 \frac{\xi h_x}{2}}{h_x^2}, & b_x &= -\frac{\sin^2 \xi h_x}{2h_x^2}, & c_x &= \frac{\sin \xi h_x}{h_x}, & d &= -\frac{\sin \xi h_x \sin \eta h_y}{h_x h_y}, \\
a_y &= -\frac{2 \sin^2 \frac{\eta h_y}{2}}{h_y^2}, & b_y &= -\frac{\sin^2 \eta h_y}{2h_y^2}, & c_y &= \frac{\sin \eta h_y}{h_y}.
\end{aligned}$$



Then

$$\begin{aligned} \mathbb{E}[|C_n|^2] \leq & 1 - \frac{4 \frac{k}{h^2} \sin^2 \frac{\xi h}{2} \left( 1 - \rho_x \cos^2 \frac{\xi h}{2} + \frac{k}{h^2} \sin^2 \frac{\xi h}{2} (1 - 2\rho_x^2(1 + 2|\rho_{xy}|)) \right)}{\left[ 1 + 2 \frac{k}{h^2} \left( \sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2} \right) \right]^2} \\ & - \frac{4 \frac{k}{h^2} \sin^2 \frac{\eta h}{2} \left( 1 - \rho_y \cos^2 \frac{\eta h}{2} + \frac{k}{h^2} \sin^2 \frac{\eta h}{2} (1 - 2\rho_y^2(1 + 2|\rho_{xy}|)) \right)}{\left[ 1 + 2 \frac{k}{h^2} \left( \sin^2 \frac{\eta h}{2} + \sin^2 \frac{\xi h}{2} \right) \right]^2} \\ & - \frac{8 \frac{k^2}{h^4} \sin^2 \frac{\xi h}{2} \sin^2 \frac{\eta h}{2} \left( 1 - 2\rho_x \rho_y (1 + 2|\rho_{xy}| + 3\rho_{xy}^2) \right)}{\left[ 1 + 2 \frac{k}{h^2} \left( \sin^2 \frac{\eta h}{2} + \sin^2 \frac{\xi h}{2} \right) \right]^2}. \end{aligned}$$

Denote  $\lambda = \frac{k}{h^2}$ ,  $a = \sin^2 \frac{\xi h}{2}$ ,  $b = \sin^2 \frac{\eta h}{2}$ . It follows that

$$\begin{aligned} \mathbb{E}[|C_n|^2] \leq & 1 - \frac{4\lambda a \left( 1 - \rho_x \cos^2 \frac{\xi h}{2} + \lambda a (1 - 2\rho_x^2(1 + 2|\rho_{xy}|)) \right)}{(1 + 2\lambda(a + b))^2} \\ & - \frac{4\lambda b \left( 1 - \rho_y \cos^2 \frac{\eta h}{2} + \lambda b (1 - 2\rho_y^2(1 + 2|\rho_{xy}|)) \right)}{(1 + 2\lambda(a + b))^2} \\ & - \frac{8\lambda^2 ab \left( 1 - 2\rho_x \rho_y (1 + 2|\rho_{xy}| + 3\rho_{xy}^2) \right)}{(1 + 2\lambda(a + b))^2}. \end{aligned}$$

By Assumption 2.1,

$$\begin{aligned} 0 \leq \rho_x < 1, \quad 0 \leq \rho_y < 1, \quad 0 < 1 - 2\rho_x^2(1 + 2|\rho_{xy}|) \leq 1, \\ 0 < 1 - 2\rho_y^2(1 + 2|\rho_{xy}|) \leq 1, \quad 0 < 1 - 2\rho_x \rho_y (1 + 2|\rho_{xy}| + 3\rho_{xy}^2) \leq 1, \end{aligned}$$

we write

$$\begin{aligned} \beta &= \min \left\{ 1 - \rho_x, 1 - \rho_y, 1 - 2\rho_x^2(1 + 2|\rho_{xy}|), \right. \\ &\quad \left. 1 - 2\rho_y^2(1 + 2|\rho_{xy}|), 1 - 2\rho_x \rho_y (1 + 2|\rho_{xy}| + 3\rho_{xy}^2) \right\} \in (0, 1), \\ d &= a + b = \sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}. \end{aligned}$$

Consequently,

$$\begin{aligned} \mathbb{E}[|C_n|^2] &\leq 1 - \frac{4\lambda a(\beta + \lambda a\beta) + 4\lambda b(\beta + \lambda b\beta) + 8\lambda^2 ab\beta}{(1 + 2\lambda(a + b))^2} \\ &= 1 - \frac{4\beta(\lambda d + \lambda^2 d^2)}{(1 + 2\lambda d)^2}. \end{aligned}$$

Therefore we have

$$\begin{aligned}\mathbb{E}[|X_N|^2] &= |X_0|^2 \prod_{n=0}^{N-1} \mathbb{E}[|C_n|^2] \\ &\leq |X_0|^2 \left(1 - 4\beta \frac{\lambda \left(\sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}\right) + \lambda^2 \left(\sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}\right)^2}{\left(1 + 2\lambda \left(\sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}\right)\right)^2}\right)^N.\end{aligned}\quad (4.3)$$

□

Then we consider two scenarios:  $0 < \lambda < 1$  and  $\lambda \geq 1$ . For  $0 < \lambda < 1$ ,

$$\Omega_{\text{high}} = \left\{ (\xi, \eta) : |\xi| > h^{-2p}, \text{ or } |\eta| > h^{-2p} \right\} \cap [-\pi h^{-1}, \pi h^{-1}] \times [-\pi h^{-1}, \pi h^{-1}].$$

**Lemma 4.4** For  $0 < \lambda < 1$  (i.e.,  $k < h^2$ ),

$$\mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) \, d\xi \, d\eta \right|^2 \right] = o(h^r), \quad \forall r > 0.$$

**Proof** Note that

$$X(T) = \exp \left( -\frac{1}{2}(1 - \rho_x)\xi^2 T - \frac{1}{2}(1 - \rho_y)\eta^2 T - i\xi\sqrt{\rho_x}M_T^x - i\eta\sqrt{\rho_y}M_T^y \right).$$

Then

$$\begin{aligned}\mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) \, d\xi \, d\eta \right|^2 \right] \\ &< 4\pi^2 h^{-2} \mathbb{E} \left[ \iint_{\Omega_{\text{high}}} \left| X(T, \xi, \eta) - X_N(\xi, \eta) \right|^2 \, d\xi \, d\eta \right] \\ &\leq 8\pi^2 h^{-2} \iint_{\Omega_{\text{high}}} \mathbb{E} \left[ |X_N(\xi, \eta)|^2 + |X(T, \xi, \eta)|^2 \right] \, d\xi \, d\eta \\ &= 8\pi^2 h^{-2} \iint_{\Omega_{\text{high}}} \mathbb{E} \left[ |X_N(\xi, \eta)|^2 \right] \, d\xi \, d\eta + f_0(k),\end{aligned}$$

where

$$\begin{aligned}f_0(k) &= 8\pi^2 h^{-2} \iint_{\Omega_{\text{high}}} e^{-(1-\rho_x)\xi^2 T - (1-\rho_y)\eta^2 T} \, d\xi \, d\eta \\ &= 8\pi^2 h^{-2} \int_0^{\pi/h} \int_{h^{-2p}}^{\pi/h} e^{-(1-\rho_x)\xi^2 T - (1-\rho_y)\eta^2 T} \, d\xi \, d\eta\end{aligned}$$

$$\begin{aligned}
 & + 8\pi^2 h^{-2} \int_{h^{-2p}}^{\pi/h} \int_0^{\pi/h} e^{-(1-\rho_x)\xi^2 T - (1-\rho_y)\eta^2 T} d\xi d\eta \\
 & - 8\pi^2 h^{-2} \int_{h^{-2p}}^{\pi/h} \int_{h^{-2p}}^{\pi/h} e^{-(1-\rho_x)\xi^2 T - (1-\rho_y)\eta^2 T} d\xi d\eta \\
 & \leq C \cdot h^{-2+2p} (e^{-(1-\rho_x)Th^{-2p}} + e^{-(1-\rho_y)Th^{-2p}}) = o(h^r), \quad \forall r > 0.
 \end{aligned}$$

Denote  $d = \sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}$ . From (4.3) and  $\lambda = k/h^2$ ,

$$\begin{aligned}
 \mathbb{E}[|X_N|^2] & \leq |X_0|^2 \left( 1 - \frac{4\beta(d + \lambda d^2)}{(1 + 2\lambda d)^2} \cdot \frac{T}{Nh^2} \right)^N \\
 & < |X_0|^2 \exp \left( - \frac{4\beta(d + \lambda d^2)T}{(1 + 2\lambda d)^2} \cdot h^{-2} \right).
 \end{aligned}$$

In this case, as at least one of  $\xi$  and  $\eta$  belongs to  $(h^{-2p}, \pi/h)$ , we have

$$d = a + b = \sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2} \geq \sin^2 \frac{h^{1-2p}}{2} = \frac{h^{2-4p}}{4} - \frac{h^{4-8p}}{48} + O(h^{5-10p}).$$

Therefore,

$$\begin{aligned}
 \mathbb{E}[|X_N|^2] & < |X_0|^2 \exp \left( - \frac{4\beta(d + \lambda d^2)T}{(1 + 2\lambda d)^2} \cdot h^{-2} \right) \\
 & < |X_0|^2 \exp(-4\beta d T h^{-2}) < |X_0|^2 \exp(-\beta T h^{-4p}) = o(h^r).
 \end{aligned}$$

As a result,

$$\begin{aligned}
 & 8\pi^2 h^{-2} \iint_{\Omega_{\text{high}}} \mathbb{E}[|X_N(\xi, \eta)|^2] d\xi d\eta \\
 & < 16\pi^4 |X_0|^2 h^{-4} \exp(-\beta T h^{-4p}) = o(h^r), \quad \forall r > 0.
 \end{aligned}$$

Therefore, for all  $r > 0$ ,

$$\begin{aligned}
 & \mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) d\xi d\eta \right|^2 \right] \\
 & < 8\pi^2 h^{-2} \iint_{\Omega_{\text{high}}} \mathbb{E}[|X_N(\xi, \eta)|^2] d\xi d\eta + f_0(k) = o(h^r).
 \end{aligned}$$

□

For  $\lambda \geq 1$ , we further separate the domain  $\Omega_{\text{high}}$  into a middle wave region

$$\Omega_{\text{high}}^1 = \left\{ (|\xi|, |\eta|) \in [k^{-p}, k^{-\frac{1}{2}}] \times [0, k^{-\frac{1}{2}}] \cup [0, k^{-\frac{1}{2}}] \times [k^{-p}, k^{-\frac{1}{2}}] \right\}, \quad (4.4)$$

and a high wave region

$$\Omega_{\text{high}}^2 = \left\{ (|\xi|, |\eta|) \in \left[ k^{-\frac{1}{2}}, \pi/h \right] \times [0, \pi/h] \cup [0, \pi/h] \times \left[ k^{-\frac{1}{2}}, \pi/h \right] \right\}, \quad (4.5)$$

such that  $\Omega_{\text{high}} = \Omega_{\text{high}}^1 \cup \Omega_{\text{high}}^2$ .

**Lemma 4.5** *For  $\lambda \geq 1$  (i.e.,  $k \geq h^2$ ), there exists  $\theta \in (0, 1)$  independent of  $h$  and  $k$  such that*

$$\mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) \, d\xi d\eta \right|^2 \right] \leq Ch^{-4} \theta^N.$$

**Proof**

$$\begin{aligned} & \mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) \, d\xi d\eta \right|^2 \right] \\ & \leq 2 \mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}^1} X(T, \xi, \eta) - X_N(\xi, \eta) \, d\xi d\eta \right|^2 \right] \\ & \quad + 2 \mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}^2} X(T, \xi, \eta) - X_N(\xi, \eta) \, d\xi d\eta \right|^2 \right] \\ & < 8k^{-1} \mathbb{E} \left[ \iint_{\Omega_{\text{high}}^1} \left| X(T, \xi, \eta) - X_N(\xi, \eta) \right|^2 \, d\xi d\eta \right] \\ & \quad + 8\pi^2 h^{-2} \mathbb{E} \left[ \iint_{\Omega_{\text{high}}^2} \left| X(T, \xi, \eta) - X_N(\xi, \eta) \right|^2 \, d\xi d\eta \right] \\ & \leq 16k^{-1} \iint_{\Omega_{\text{high}}^1} \mathbb{E} \left[ |X_N(\xi, \eta)|^2 + |X(T, \xi, \eta)|^2 \right] \, d\xi d\eta \\ & \quad + 16\pi^2 h^{-2} \iint_{\Omega_{\text{high}}^2} \mathbb{E} \left[ |X_N(\xi, \eta)|^2 + |X(T, \xi, \eta)|^2 \right] \, d\xi d\eta \\ & = 16k^{-1} \iint_{\Omega_{\text{high}}^1} \mathbb{E} \left[ |X_N(\xi, \eta)|^2 \right] \, d\xi d\eta \\ & \quad + 16\pi^2 h^{-2} \iint_{\Omega_{\text{high}}^2} \mathbb{E} \left[ |X_N(\xi, \eta)|^2 \right] \, d\xi d\eta + f_1(k), \end{aligned}$$

where

$$\begin{aligned} f_1(k) &= 16k^{-1} \iint_{\Omega_{\text{high}}^1} e^{-(1-\rho_x)\xi^2 T - (1-\rho_y)\eta^2 T} \, d\xi d\eta \\ & \quad + 16\pi^2 h^{-2} \iint_{\Omega_{\text{high}}^2} e^{-(1-\rho_x)\xi^2 T - (1-\rho_y)\eta^2 T} \, d\xi d\eta \end{aligned}$$

$$\leq C \cdot k^{p-1} \exp(-\beta T k^{-p}) + C \cdot k h^{-2} \exp(-2\beta T k^{-1}).$$

Denote  $d = \sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}$ , from (4.3),

$$\mathbb{E}[|X_N|^2] \leq |X_0|^2 \left( 1 - \frac{4\beta(\lambda d + \lambda^2 d^2)}{(1 + 2\lambda d)^2} \right)^N.$$

For  $\lambda \geq 1$ , and  $(\xi, \eta) \in \Omega_{\text{high}}^1$ ,

$$\begin{aligned} \mathbb{E}[|X_N|^2] &\leq |X_0|^2 \left( 1 - \frac{4\beta(d + \lambda d^2)}{(1 + 2\lambda d)^2} \cdot \frac{T}{N h^2} \right)^N \\ &< |X_0|^2 \exp \left( - \frac{4\beta(d + \lambda d^2)T}{(1 + 2\lambda d)^2} \cdot h^{-2} \right). \end{aligned}$$

In this case, as at least one of  $\xi$  and  $\eta$  belongs to  $(k^{-p}, k^{-1/2})$ , we have

$$\begin{aligned} d = a + b &= \sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2} \\ &\geq \sin^2 \frac{k^{-p} h}{2} = \frac{k^{-2p} h^2}{4} - \frac{k^{-4p} h^4}{48} + O(k^{-5p} h^5). \\ \mathbb{E}[|X_N|^2] &< |X_0|^2 \exp \left( - \frac{4\beta(d + \lambda d^2)T}{(1 + 2\lambda d)^2} \cdot h^{-2} \right) \\ &< |X_0|^2 \exp(-4\beta d T h^{-2}) < |X_0|^2 \exp(-\beta T k^{-2p}). \end{aligned}$$

Since  $|\Omega_{\text{high}}^1| < 4k^{-1}$ ,

$$16k^{-1} \iint_{\Omega_{\text{high}}^1} \mathbb{E}[|X_N(\xi, \eta)|^2] d\xi d\eta < 64|X_0|^2 k^{-2} \exp(-\beta T k^{-2p}).$$

For  $\lambda \geq 1$ , and  $(\xi, \eta) \in \Omega_{\text{high}}^2$ ,  $d = \sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2} \in [\sin^2 \frac{1}{2\sqrt{\lambda}}, 2]$ ,

$$\begin{aligned} \max_d \left( 1 - \frac{4\beta(\lambda d + \lambda^2 d^2)}{(1 + 2\lambda d)^2} \right) &= 1 - \beta \min_d \left( 1 - \frac{1}{(1 + 2\lambda d)^2} \right) \\ &= 1 - \beta \left( 1 - \max_d \frac{1}{(1 + 2\lambda d)^2} \right). \end{aligned}$$

As

$$\max_d \frac{1}{(1 + 2\lambda d)^2} = \frac{1}{(1 + 2\lambda d)^2} \Big|_{d=\sin^2 \frac{1}{2\sqrt{\lambda}}}$$

$$= \frac{1}{(1 + 2\lambda \sin^2 \frac{1}{2\sqrt{\lambda}})^2} \leq \frac{1}{(1 + 2 \sin^2 \frac{1}{2})^2} < 0.5,$$

we have

$$\begin{aligned} 1 - \frac{4\beta(\lambda d + \lambda^2 d^2)}{(1 + 2\lambda d)^2} &\leq 1 - \beta \left( 1 - \frac{1}{(1 + 2\lambda \sin^2 \frac{1}{2\sqrt{\lambda}})^2} \right) \\ &= 1 - \beta + \frac{\beta}{(1 + 2\lambda \sin^2 \frac{1}{2\sqrt{\lambda}})^2} < 1 - \frac{1}{2}\beta. \end{aligned}$$

Denote  $\theta_0 := 1 - \frac{1}{2}\beta \in (0, 1)$ , then

$$\mathbb{E}[|X_N|^2] \leq |X_0|^2 \left( 1 - \frac{4\beta(\lambda d + \lambda^2 d^2)}{(1 + 2\lambda d)^2} \right)^N \leq |X_0|^2 \cdot \theta_0^N.$$

So

$$16\pi^2 h^{-2} \iint_{\Omega_{\text{high}}^2} \mathbb{E}[|X_N(\xi, \eta)|^2] d\xi d\eta < 64\pi^4 |X_0|^2 h^{-4} \theta_0^N.$$

Hence

$$\begin{aligned} \mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) d\xi d\eta \right|^2 \right] \\ \leq C k^{p-1} \exp(-\beta T k^{-p}) + C k h^{-2} \exp(-2\beta T k^{-1}) \\ + 64|X_0|^2 k^{-2} \exp(-\beta T k^{-2p}) + 64\pi^4 |X_0|^2 h^{-4} \theta_0^N. \end{aligned} \quad (4.6)$$

As the first three terms in (4.6) are terms of higher order than  $h^4 \theta^N$ , we have, for  $\lambda \geq 1$ ,

$$\mathbb{E} \left[ \left| \iint_{\Omega_{\text{high}}} X(T, \xi, \eta) - X_N(\xi, \eta) d\xi d\eta \right|^2 \right] \leq C h^{-4} \theta_0^N.$$

Letting  $\theta = \sqrt{\theta_0}$  the result follows.  $\square$

### 4.3 Convergence of the ADI scheme (proof of Theorem 2.2)

Theorem 2.2 states that the error of the ADI method (2.3) has the same order as the implicit Milstein scheme (2.2). We now give a proof as follows.

**Proof of Theorem 2.2** Let

$$X_{n+1} = C_n X_n,$$

where

$$C_n \equiv \exp \left( -\frac{1}{2}(1 - \rho_x)\xi^2 k - \frac{1}{2}(1 - \rho_y)\eta^2 k - i\xi\sqrt{\rho_x k}Z_{n,x} - i\eta\sqrt{\rho_y k}\tilde{Z}_{n,y} + e_n \right),$$

and  $e_n$  is the logarithmic error between the numerical solution and the exact solution introduced during  $[nk, (n + 1)k]$ . From (3.11),  $C_n$  has the form

$$C_n = \frac{1 - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y} k Z_{n,x} \tilde{Z}_{n,y}}{(1 - a_x k)(1 - a_y k)}.$$

In the low wave region, the numerical solutions are close to the exact solutions. We get from Taylor expansion that

$$\begin{aligned} X_N - X(T) = X(T) \cdot \left\{ \frac{i}{6} \sqrt{\rho_x} \xi^3 h_x^2 M_T^x + \frac{1}{24} (1 - 4\rho_x) \xi^4 h_x^2 T \right. \\ \left. + \frac{i}{6} \sqrt{\rho_y} \eta^3 h_y^2 \tilde{M}_T^y + \frac{1}{24} (1 - 4\rho_y) \eta^4 h_y^2 T \right. \\ \left. + i k \sqrt{k} \sum_{n=0}^{N-1} \hat{\phi}_1(Z_{n,x}, \tilde{Z}_{n,y}) + k^2 \sum_{n=0}^{N-1} \hat{\phi}_2(Z_{n,x}, \tilde{Z}_{n,y}) + o(k, h_x^2, h_y^2) \right\}. \end{aligned}$$

In the high wave region, we have

$$X_N = X_0 \prod_{n=0}^{N-1} \frac{1 - i c_x \sqrt{\rho_x k} Z_{n,x} - i c_y \sqrt{\rho_y k} \tilde{Z}_{n,y} + b_x \rho_x k (Z_{n,x}^2 - 1) + b_y \rho_y k (\tilde{Z}_{n,y}^2 - 1) + d \sqrt{\rho_x \rho_y} k Z_{n,x} \tilde{Z}_{n,y}}{(1 - a_x k)(1 - a_y k)}.$$

Then

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}[X_N] &= X_0 \exp \left( -\frac{1}{2}(\xi^2 u + \eta^2 v + \frac{1}{2}\xi^2 \eta^2 uvk \right. \\ &\quad \left. + 2\xi\eta \frac{\sin \xi h_x \sin \eta h_y}{\xi h_x \eta h_y} \sqrt{\rho_x \rho_y \rho_{xy}}) T \right), \\ \lim_{N \rightarrow \infty} \mathbb{E}[|X_N|^2] &\leq |X_0|^2 \exp \left( -\frac{1}{2}\xi^2 u T (1 - \rho_x + \frac{1}{4}\xi^2 uk (1 - 2\rho_x (1 + \rho_{xy}))) \right. \\ &\quad \left. - \frac{1}{2}\eta^2 v T (1 - \rho_y + \frac{1}{4}\eta^2 vk (1 - 2\rho_y (1 + \rho_{xy}))) \right. \\ &\quad \left. - \frac{1}{4}\xi^2 \eta^2 uvk T (1 - \rho_{xy} (1 + \rho_{xy} + 3\rho_{xy}^2)) \right), \end{aligned}$$

where  $u = \sin^2 \frac{h_x \xi}{2} / (\frac{h_x \xi}{2})^2$ ,  $v = \sin^2 \frac{h_y \eta}{2} / (\frac{h_y \eta}{2})^2$ . By the same reasoning as for the implicit scheme (2.2), the integration over the high wave region is of higher order than  $h_x^2$  and  $h_y^2$  given condition (2.7). Then the inverse Fourier transform gives the result.  $\square$

## 5 Numerical tests

In this section, we illustrate the stability and convergence results from the previous section by way of empirical tests.

Unless stated otherwise, we choose parameters  $T = 1$ ,  $x_0 = y_0 = 2$ ,  $\mu_x = \mu_y = 0.0809$ ,  $\rho_x = \rho_y = 0.2$ ,  $\rho_{xy} = 0.45$ . For the computations, we truncate the domain to  $[-8, 12] \times [-8, 12]$ , chosen large enough such that the effect of zero Dirichlet boundary conditions on the solution is negligible.

Figure 1a shows the numerical solution for one Brownian path, with  $h_x = h_y = 2^{-4}$ ,  $k = 2^{-10}$ . Figure 1b plots the pointwise error between the Milstein-ADI approximation (2.3) and the analytic solution (1.9).

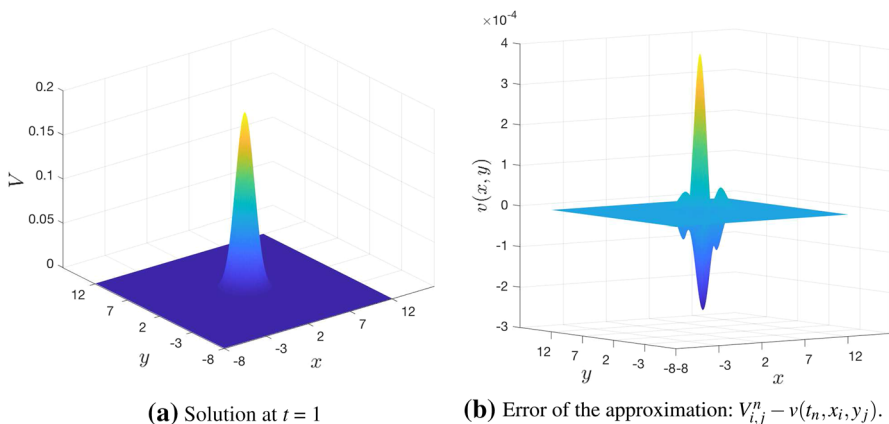
Figure 2 verifies the  $L_2$ -convergence order in  $h$  and  $k$  from (2.10). We approximate the error by

$$\begin{aligned} & \left( \sum_{i,j} h_x h_y E_L [|V_{i,j}^N - v(T, x_i, y_j)|^2] \right)^{1/2} \\ &= \left( \sum_{i,j} \frac{h_x h_y}{L} \sum_{l=1}^L |V_{i,j}^N(M^{(l)}) - v(T, x_i, y_j; M^{(l)})|^2 \right)^{1/2}, \end{aligned} \quad (5.1)$$

where  $M^{(l)}$  are independent Brownian motions and  $E_L$  is the empirical mean with  $L = 100$  samples.

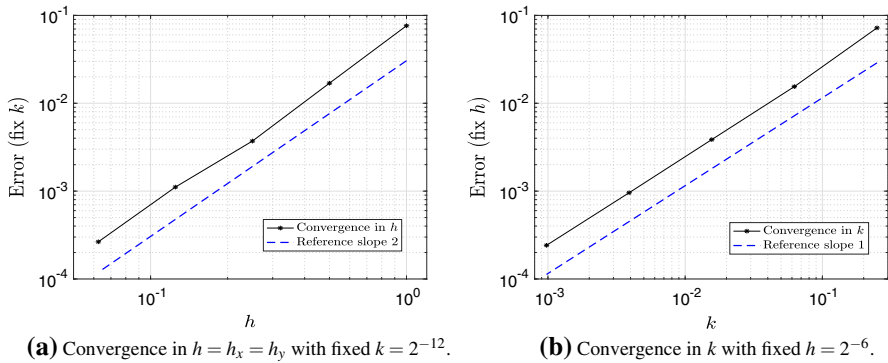
Here, Fig. 2a shows the convergence in  $h = h_x = h_y$  with  $k = 2^{-12}$  small enough, which demonstrates second order convergence in  $h$ . Figure 2b shows the convergence in  $k$  with  $h = h_x = h_y = 2^{-6}$  small enough to ensure sufficient accuracy of the spatial approximation. One can clearly observe first order convergence in  $k$ .

In Fig. 3, we illustrate the dependence of the approximation error in the  $L_2$ -norm on the correlation parameters. The error increases as a function of  $\rho_x$  and  $\rho_y$ . The

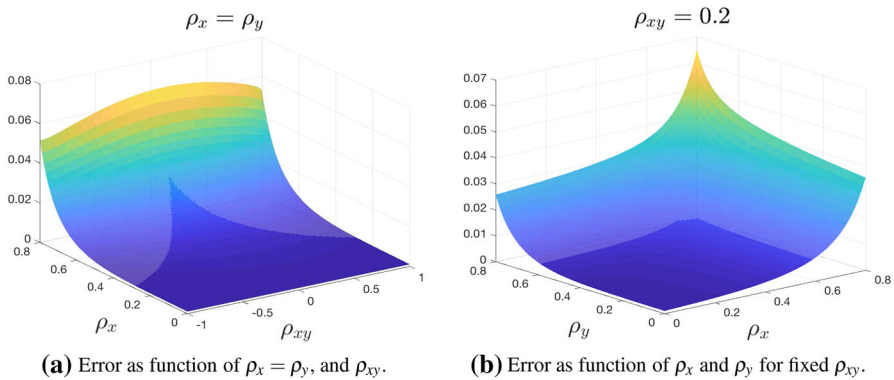


**Fig. 1** Numerical approximation to the SPDE (1.7) with the Milstein-ADI approximation (2.3) for one Brownian path, and the pointwise error to the analytic solution (1.9)





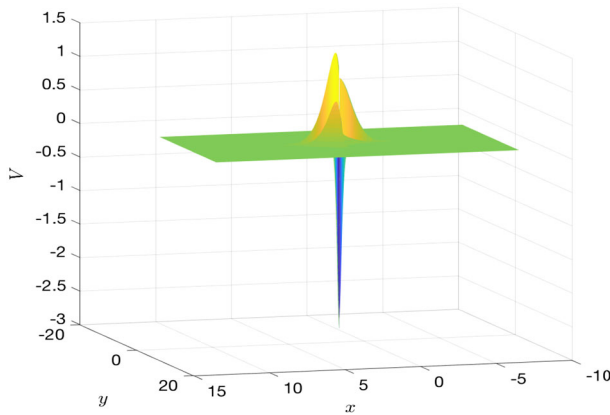
**Fig. 2** Verifies the  $L_2$ -convergence order in  $h$  and  $k$  from (2.10), with coarsest level  $h = 1$ ,  $k = 1/4$  and finest level  $h = 2^{-5}$ ,  $k = 2^{-10}$ , using 100 Monte Carlo samples



**Fig. 3**  $L_2$  error in space as function of correlation parameters for fixed  $h_x = h_y = 2^{-3}$  and  $k = 2^{-9}$ , for a fixed path. The dark blue areas correspond to the stability region from Assumption 2.1

error for  $\rho_x = \rho_y \leq 0.3$  (see Fig. 3a) varies between roughly  $10^{-3}$  and  $3 \cdot 10^{-3}$ , the error being smallest for  $\rho_x = \rho_y = 0$  (the PDE case), and largest for large  $\rho_x = \rho_y$  and  $\rho_{xy}$  between 0.1 and 0.4. For larger  $\rho_x$  and  $\rho_y$ , the error increases sharply. The stability region from Assumption 2.1 is marked in dark blue, which shows that stable results are obtained even outside the region where mean-square stability is proven. We found problems only for  $\rho_x = \rho_y \geq 0.8$ . This discrepancy is partly due to the fact that Assumption 2.1 is sufficient, but not necessary, as some of the estimates are not sharp. Figure 3b shows a similar behaviour when varying  $\rho_x$  and  $\rho_y$  independently for fixed  $\rho_{xy}$ .

Figure 4 shows the singular behaviour of the solution for large  $k$  and small  $h$ , as predicted by Theorem 2.1. Figure 5 investigates the behaviour of the error in this regime further, with  $h_x = h_y = h$  in Fig. 5a, and  $h_y = 2^{-1}$  fixed,  $h_x = h$  in Fig. 5b, letting  $h$  go to zero. We calculate the  $L_2$  error in space, and compare different scenarios under the same Brownian path. The top black line shows the error with  $k = 2^{-2}$  fixed, and  $\rho_x = \rho_y = 0.6$ ,  $\rho_{xy} = 0.1$ . We can see that as  $h$  goes to zero, the error diverges with rate  $h^{-1/2}$  (choosing  $h_y = 2^{-1}$  fixed enables more refinements in  $h_x$  to show



**Fig. 4** Unstable solution for the Milstein-ADI scheme (2.3) for one Brownian path,  $k = 2^{-2}$  and  $h_x = h_y = 2^{-9}$ ,  $\rho_x = \rho_y = 0.6$ ,  $\rho_{xy} = 0.1$

the asymptotic behaviour better). The blue line second from top plots the error with  $k = 2^{-4}$  fixed instead. Note that in this case the error will eventually diverge for  $h$  going to zero, but this is not visible yet for this level of  $h$ . The next red line plots the error for  $\rho_x = \rho_y = 0$ , with  $k = 2^{-2}$  fixed. Then the SPDE (1.7) becomes a PDE and divergence does not appear. Finally, the bottom blue dotted line plots the error for the SPDE (1.7) with initial condition

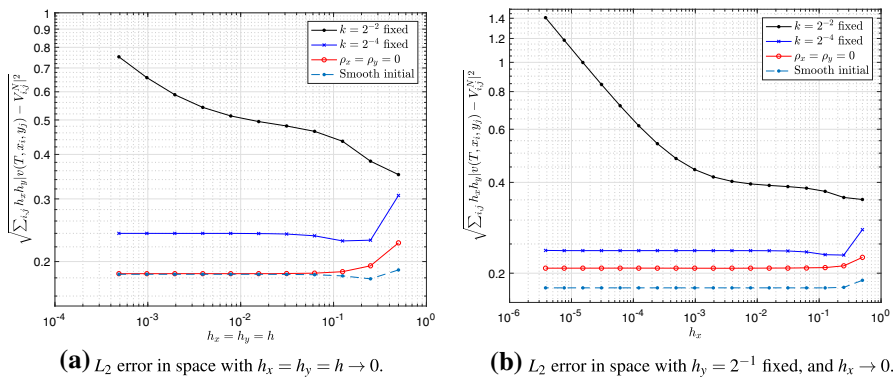
$$v(0, x, y) = \frac{1}{2\pi\sqrt{(1-\rho_x)(1-\rho_y)}} \exp\left(-\frac{(x-x_0-\mu_x)^2}{2(1-\rho_x)} - \frac{(y-y_0-\mu_y)^2}{2(1-\rho_y)}\right). \quad (5.2)$$

For this smooth initial condition, the solution does not diverge for large  $k$  and small  $h$ . Hence, this verifies that the divergence is a result of the interplay of singular data and stochastic terms only, as shown in Corollary 2.2. We emphasise that the instability is so mild that it is only visible in artificial numerical tests, while for reasonably small  $k$ , in particular for  $k \sim h^2$  as would be chosen in practice, no instabilities occur.

## 6 An extended scheme and tests for a more general SPDE

To approximate the general Zakai SPDE (1.1),

$$\begin{aligned} dv(t, x) = & \left( \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(x)v(t, x)] - \sum_{i=1}^d \frac{\partial}{\partial x_i} [b_i(x)v(t, x)] \right) dt \\ & - \sum_{l=1}^m \sum_{i=1}^d \frac{\partial}{\partial x_i} (\gamma_{i,l}(x)v(t, x)) dM_t^k, \end{aligned}$$



**Fig. 5** Comparing the  $L_2$  error in space with fixed  $k$  and letting  $h \rightarrow 0$  under various scenarios:  $k = 2^{-2}$ ;  $k = 2^{-4}$ ;  $k = 2^{-2}$  and  $\rho_x = \rho_y = 0$ ;  $k = 2^{-2}$  and smooth initial condition (5.2), all with a same Brownian path

by the Milstein scheme, we approximate in the last term

$$v(s) \approx v(t) - \sum_{l=1}^m \sum_{i=1}^d \frac{\partial}{\partial x_i} \left( \gamma_{i,l}(x) v(t, x) \right) (M_l(s) - M_l(t)),$$

then

$$\begin{aligned} & - \sum_{l=1}^m \sum_{i=1}^d \int_t^{t+k} \frac{\partial}{\partial x_i} \left( \gamma_{i,l}(x) v(s, x) \right) dM_l(s) \\ & = - \sum_{l=1}^m \sum_{i=1}^d \int_t^{t+k} \frac{\partial}{\partial x_i} \left( \gamma_{i,l}(x) \left[ v(t) - \sum_{p=1}^m \sum_{j=1}^d \frac{\partial}{\partial x_j} \left( \gamma_{j,p}(x) v(t, x) \right) (M_p(s) - M_p(t)) \right] \right) dM_l(s) \\ & = - \sum_{l=1}^m \sum_{i=1}^d \frac{\partial}{\partial x_i} \left( \gamma_{i,l}(x) v(t, x) \right) \Delta M_l^n + \sum_{l=1}^m \sum_{p=1}^m \sum_{i=1}^d \frac{\partial}{\partial x_i} \\ & \quad \times \left( \gamma_{il}(x) \sum_{j=1}^d \frac{\partial}{\partial x_j} \left( \gamma_{jp}(x) v(t, x) \right) \right) \int_t^{t+k} M_p(s) - M_p(t) dM_l(s). \end{aligned}$$

The corresponding ADI implicit Milstein scheme is

$$\prod_{i=1}^d \left( I + \frac{k}{2h_i} D_i b_i(X) - \frac{1}{2} \frac{k}{h_i^2} D_{ii} a_{ii}(X) \right) V^{n+1}$$

$$\begin{aligned}
&= \left\{ I + \frac{1}{2} \sum_{i \neq j} \frac{k}{4h_i h_j} D_i D_j a_{ij}(X) - \sum_{l=1}^m \Delta M_l^n \sum_{i=1}^d \frac{1}{2h_i} D_i \gamma_{i,l}(X) \right. \\
&\quad + \sum_{l=1}^m \sum_{p=1}^m \left( \int_{nk}^{(n+1)k} (M_p(s) - M_p(t)) dM_l(s) \right) \\
&\quad \left. \times \sum_{i=1}^d \left( \frac{1}{2h_i} D_i \gamma_{i,l}(X) \right) \sum_{j=1}^d \frac{1}{2h_j} D_j \gamma_{j,p}(X) \right\} V^n.
\end{aligned}$$

Here,  $\{D_i\}_{1 \leq i \leq d}$  are first order difference operators, and  $\{D_{ij}\}_{1 \leq i, j \leq d}$  are second order difference operators,  $X$  is the vector of mesh points ordered the same way as  $V$ , and, by slight abuse of notation, we denote by  $a(X)$ ,  $b(X)$ ,  $\gamma(X)$  the diagonal matrices such that each element of the diagonal corresponds to the function evaluated at the corresponding mesh point.

Notice the presence of an iterated Itô integral  $\int_t^{t+k} (M_p(s) - M_s(t)) dM_l(s)$ , called Lévy area, is common in multi-dimensional Milstein schemes. It has been proved in [8,33] that there is no way to achieve a better order of strong convergence than for the Euler scheme by using solely the discrete increments of the driving Brownian motions. An efficient algorithm for the approximate simulation of the Lévy area has been proposed in [38], building on earlier work in [15,26] and based on an approximation of the distribution of the tail-sum in a truncated infinite series representation derived from the characteristic functions of these integrals, achieving the complexity  $\varepsilon^{-3/2}$  of sampling a single path to obtain a strong error  $\varepsilon$ . Malham and Wiese proposed another algorithm for the simulation of the Lévy area conditioned on the endpoints, in [32], representing the Lévy area by an infinite weighted sum of independent logistic random variables, further reducing the complexity to  $\varepsilon^{-1} |\log(\varepsilon^{-1})|$ .

However, the algorithms mentioned above are fairly complex. What we do instead is to further divide the interval  $[t, t+k]$  into  $O(k^{-1})$  steps and perform a simple Euler approximation to the Lévy area. Note that we always use the timestep  $k$  for the SPDE computation. Overall, this still leads to first order convergence in time, and second order in space. To balance the leading order error, the optimal choice is  $O(k) = O(h^2) = O(\varepsilon)$ . Therefore, the estimate of the Lévy area in each time-step increases the computation time by  $O(k^{-1}) = O(\varepsilon^{-1})$  for one step, whereas the matrix calculation for each time step is also  $O(\varepsilon^{-1})$ , and hence the order of the total complexity does not change.

Moreover, the path simulation including the Lévy areas can be performed separately beforehand using vectorisation, leading to further speed-up.

Now we apply this method to an SPDE from [21],

$$\begin{aligned}
du = & \left[ \kappa_1 u - \left( r_1 - \frac{1}{2} y - \xi_1 \rho_{3,1} \rho_{1,1} \rho_{2,1} \right) \frac{\partial u}{\partial x} - \left( \kappa_1 (\theta_1 - y) - \xi_1^2 \right) \frac{\partial u}{\partial y} + \frac{1}{2} y \frac{\partial^2 u}{\partial x^2} \right. \\
& \left. + \xi_1 \rho_{3,1} \rho_{1,1} \rho_{2,1} y \frac{\partial^2 u}{\partial x \partial y} + \frac{\xi_1^2}{2} y \frac{\partial^2 u}{\partial y^2} \right] dt - \rho_{1,1} \sqrt{y} \frac{\partial u}{\partial x} dW_t - \xi_1 \rho_{2,1} \frac{\partial}{\partial y} (\sqrt{y} u) dB_t,
\end{aligned} \tag{6.1}$$

with Dirac initial  $u(0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0)$ . This SPDE models the limit empirical measure of a large portfolio of defaultable assets in which the asset value processes are modelled by Heston-type stochastic volatility models with common and idiosyncratic factors in both the asset values and the variances, and default is triggered by hitting a lower boundary.

Similar to before, we implement the SPDE (6.1) with a Milstein ADI scheme:

$$\begin{aligned} A_x A_y U^{n+1} = & \left( (1 + \kappa_1 k) I - \frac{k}{4h_x} \xi_1 \rho_{1,1} \rho_{2,1} \rho_3 D_x - \frac{\sqrt{k} Z_{n,x}}{2h_x} \rho_{1,1} \sqrt{Y} D_x \right. \\ & + \frac{k(Z_{n,x}^2 - 1)}{8h_x^2} \rho_{1,1}^2 Y D_x^2 + \frac{k}{4h_x} \xi_1 \rho_{1,1} \rho_{2,1} Z_{n,x} \tilde{Z}_{n,y} \left( D_x + \frac{1}{h_y} Y D_{xy} \right) \\ & + \frac{1}{4h_x} \xi_1 \rho_{2,1} \rho_{1,1} \left( \int_t^{t+k} (W_s - W_t) dB_s \right) D_x \Big) U^n \\ & - \frac{\sqrt{k} \tilde{Z}_{n,y}}{2h_y} \xi_1 \rho_{2,1} D_y (\sqrt{Y} U^n) \\ & + \frac{k(\tilde{Z}_{n,y}^2 - 1)}{8h_y^2} \xi_1^2 \rho_{2,1}^2 D_y \left( \sqrt{Y} (D_y (\sqrt{Y} U^n)) \right), \end{aligned} \quad (6.2)$$

where the notation for  $Y$  follows the same principle as above for  $X$ ,  $\tilde{Z}_n^y = \rho_{xy} Z_n^x + \sqrt{1 - \rho_{xy}^2} Z_n^y$ , with  $Z_n^x, Z_n^y \sim N(0, 1)$  independent normal random variables, and

$$\begin{aligned} A_x &= I + \frac{k}{2h_x} \left( r_1 - \xi_1 \rho_3 \rho_{1,1} \rho_{2,1} \right) D_x - \frac{k}{4h_x} Y D_x - \frac{k}{2h_x^2} Y D_{xx}, \\ A_y &= I + \frac{k}{2h_y} (\kappa_1 \theta_1 - \xi_1^2) D_y - \frac{k}{2h_y} \kappa_1 Y D_y - \frac{k}{2h_y^2} \xi_1^2 Y D_{yy}. \end{aligned}$$

Furthermore, we also implement the SPDE (6.1) with coefficients frozen at  $(x_0, y_0)$ ,

$$\begin{aligned} du = & \left[ \kappa_1 u - \left( r_1 - \frac{1}{2} y_0 - \xi_1 \rho_3 \rho_{1,1} \rho_{2,1} \right) \frac{\partial u}{\partial x} - \left( \kappa_1 (\theta_1 - y_0) - \xi_1^2 \right) \frac{\partial u}{\partial y} + \frac{1}{2} y_0 \frac{\partial^2 u}{\partial x^2} \right. \\ & + \xi_1 \rho_3 \rho_{1,1} \rho_{2,1} y_0 \frac{\partial^2 u}{\partial x \partial y} + \frac{\xi_1^2}{2} y_0 \frac{\partial^2 u}{\partial y^2} \Big] dt \\ & - \rho_{1,1} \sqrt{y_0} \frac{\partial u}{\partial x} dW_t - \xi_1 \rho_{2,1} \frac{\partial}{\partial y} (\sqrt{y_0} u) dB_t, \end{aligned} \quad (6.3)$$

with Dirac initial datum  $u(0, x, y) = \delta(x - x_0) \otimes \delta(y - y_0)$ . We can apply a Milstein ADI scheme similar to (2.2),

$$\begin{aligned}
& \left( I + \frac{k}{2h_x} \left( r_1 - \xi_1 \rho_3 \rho_{1,1} \rho_{2,1} - \frac{1}{2} y_0 \right) D_x - \frac{k}{2h_x^2} y_0 D_{xx} \right) \\
& \left( I + \frac{k}{2h_y} \left( \kappa_1 (\theta_1 - y_0) - \xi_1^2 \right) D_y - \frac{k}{2h_y^2} \xi_1^2 y_0 D_{yy} \right) U^{n+1} \\
& = \left( (1 + \kappa_1 k) I - \frac{\sqrt{k} Z_{n,x}}{2h_x} \rho_{1,1} \sqrt{y_0} D_x + \frac{k(Z_{n,x}^2 - 1)}{8h_x^2} \rho_{1,1}^2 y_0 D_x^2 \right. \\
& \quad - \frac{\sqrt{k} \tilde{Z}_{n,y}}{2h_y} \xi_1 \rho_{2,1} \sqrt{y_0} D_y + \frac{k(\tilde{Z}_{n,y}^2 - 1)}{8h_y^2} \xi_1^2 \rho_{2,1}^2 y_0 D_y^2 \\
& \quad \left. + \frac{k Z_{n,x} \tilde{Z}_{n,y}}{4h_x h_y} \xi_1 \rho_{1,1} \rho_{2,1} D_{xy} \right) U^n.
\end{aligned} \tag{6.4}$$

We will test the behaviour of (6.3) at  $(x_0, y_0)$  and compare with the SPDE (1.7).

We choose parameters  $T = 1$ ,  $x_0 = 2$ ,  $y_0 = 1.4$ ,  $r = 0.05$ ,  $\xi = 0.5$ ,  $\theta = 0.4$ ,  $\kappa = 2$ ,  $\rho_{1,1} = 0.3$ ,  $\rho_{2,1} = 0.2$ ,  $\rho_3 = 0.5$ . We truncate the domain to  $[-3, 7] \times [0, 1.5]$  for the scheme (6.2), and  $[-3, 7] \times [-1, 1.5]$  for the scheme (6.4), sufficiently large in this setting.

Figure 6a shows the density for a single Brownian path, with  $k = 2^{-8}$ ,  $h_x = 5/16$ , and  $h_y = 1/80$ .

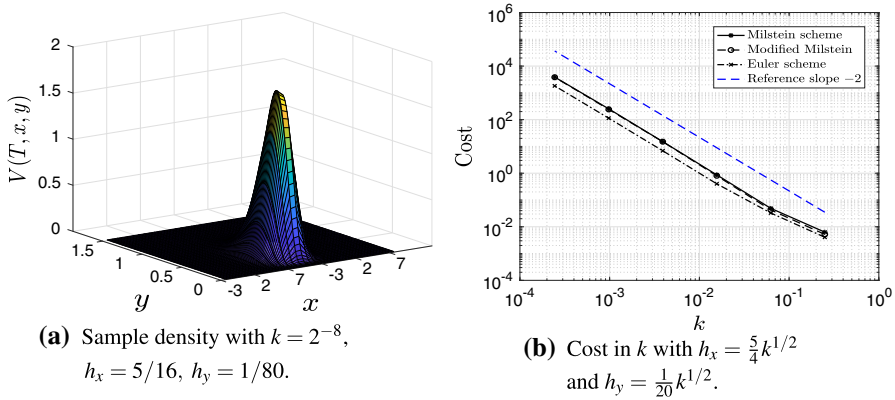
Figure 6b compares the computational cost under different time-stepping schemes: the Milstein scheme, a “modified” Milstein scheme, and the Euler scheme. Here, for the Milstein scheme we approximate the Lévy area by sub-timestepping as explained above, while in the “modified” Milstein scheme we drop  $\int_t^{t+k} (W_s - W_t) dB_s$  but keep the one-dimensional iterated integrals as they are known analytically. We expect that the latter will lead to a worse convergence in time (for non-zero  $\xi_1$ ,  $\rho_{2,1}$ ,  $\rho_{1,1}$ ), which is verified in Fig. 7b.

In Fig. 6b, from a coarsest mesh with  $h_x = 0.625$ ,  $h_y = 0.025$ , and  $k = 0.25$ , we keep decreasing the time-step  $k$  by a factor of 4, and the spatial mesh width by a factor of 2. This shows that the cost, measured by time elapsed in simulating one path, increases by a factor of 16, demonstrating that these three schemes result in the same order of complexity.

Figure 7 verifies the  $L_2$  convergence order in  $h$  and  $k$ . In absence of an exact solution, we compute a proxy to the error in  $h_x$ ,  $h_y$  by

$$\begin{aligned}
& \left( \sum_{i,j} h_x h_y E_L \left[ \left| V_{2i,2j}^N \left( k, \frac{h_x}{2}, \frac{h_y}{2} \right) - V_{i,j}^N(k, h_x, h_y) \right|^2 \right] \right)^{1/2} \\
& = \left( \sum_{i,j} \frac{h_x h_y}{L} \sum_{l=1}^L \left| V_{2i,2j}^N \left( k, \frac{h_x}{2}, \frac{h_y}{2}; M^{(l)} \right) - V_{i,j}^N(k, h_x, h_y; M^{(l)}) \right|^2 \right)^{1/2},
\end{aligned}$$

where  $V_{i,j}^N(k, h_x, h_y)$  is the numerical approximation to  $v(T, ih_x, jh_y)$  with mesh widths  $h_x$  and  $h_y$ ,  $V_{2i,2j}^N(k, h_x/2, h_y/2)$  uses a fine mesh with  $h_x/2$  and  $h_y/2$ , and  $E_L$  is the empirical mean with  $L$  samples of the Brownian path  $M$ . Both share the same



**Fig. 6** Single path realisation of the density to the SPDE (6.1) and its numerical cost (CPU time in sec)

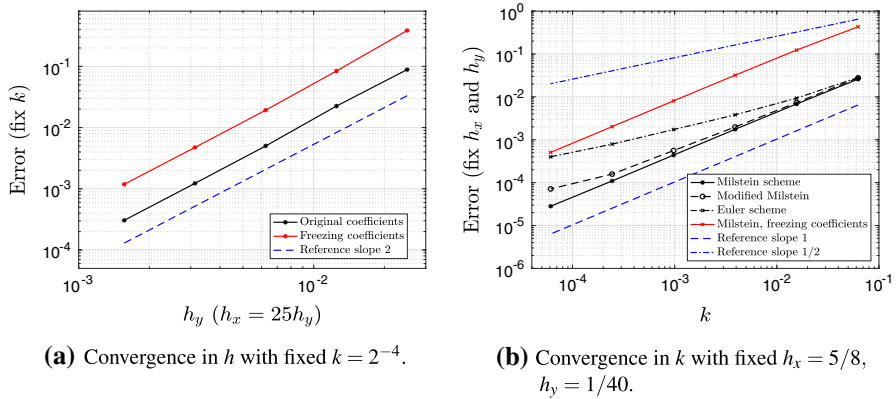
Brownian path and same time step  $k$ , thus the univariate error in  $k$  cancels and we should see the correct convergence order in  $h_x$  and  $h_y$ . We could alternatively solve on a very fine mesh and treat the result as the exact solution to test the convergence as in (5.1). It is clear from the construction that the observed convergence order will be the same. However, since the computation of a highly accurate solution is extremely expensive, and such a numerical reference solution will introduce a bias, we do not perform this test here. Here, Fig. 7a shows the convergence in  $h = h_x = h_y$  with  $k = 2^{-4}$  fixed and  $L = 100$ , which demonstrates second order convergence in  $h$ .

Similarly, we study the error in  $k$  by using the difference between two solutions with same mesh size, same Brownian path, but different time-steps,

$$\begin{aligned} & \left( \sum_{i,j} h_x h_y E_L [ |V_{i,j}^N(k, h_x, h_y) - V_{i,j}^{2N}(k/2, h_x, h_y)|^2 ] \right)^{1/2} \\ &= \left( \sum_{i,j} \frac{h_x h_y}{L} \sum_{l=1}^L \left| V_{i,j}^N(k, h_x, h_y; M^{(l)}) - V_{i,j}^{2N}(k/2, h_x, h_y; M^{(l)}) \right|^2 \right)^{1/2}. \end{aligned}$$

Figure 7b shows the convergence in  $k$  with fixed  $h_x = 5/8$ ,  $h_y = 1/40$ , under the Milstein scheme, “modified” Milstein scheme, and Euler scheme. The timestep  $k$  decreases by a factor of 4 from one level to the next. We also plot two blue dashed lines with slope 1/2 and 1 as reference. One can clearly observe first order convergence in  $k$  for the Milstein scheme, and half order convergence in  $k$  for the Euler scheme. As for the “modified” Milstein scheme, although it appears to converge with first order on coarse levels (due to dominance of the terms which converge with first order for this level of accuracy), the asymptotic order is seen to be lower.

In Fig. 7, we also compare the convergence of the scheme (6.4) for the frozen coefficient SPDE (6.3) with the scheme (6.2) for the SPDE (6.1). We can see that the scheme (6.4) has second order convergence in  $h$ , and first order convergence in  $k$ , as expected.



**Fig. 7** Test the convergence order in  $(h_x, h_y)$  and  $k$  for the schemes (6.2) and (6.4), in  $L_2$  in space and probability, using 100 paths

We now illustrate more carefully the behaviour at  $(x_0, y_0)$  over time. Figure 8a shows one path realisation of the solution  $U(t, x_0, y_0)$  with the scheme (6.2) as a function of  $t$ , and  $\hat{U}(t, x_0, y_0)$  with the scheme (6.4). The mesh sizes and time step are  $h_x = 5 \times 2^{-8}$ ,  $h_y = 0.2 \times 2^{-8}$ ,  $k = 4^{-8}$ . Note that the plots below start from  $t = 0.015$  in Fig. 8a and  $t = 0.03$  in Fig. 8b rather than zero, because of the Dirac initial datum, however, even under this singular initial condition,  $U(t, x_0, y_0)$  and  $\hat{U}(t, x_0, y_0)$  are very close over an initial time period. Figure 8b shows the error of the schemes (6.2) and (6.4) for one realisation of the path, where we use  $h_x = 5/16$ ,  $h_y = 1/80$ ,  $k = 1/256$  for “coarse” solutions  $U_c$  and  $\hat{U}_c$ , and we treat the fine approximations  $U$  and  $\hat{U}$  as the exact solution. The errors for the SPDEs (6.1) and (6.3) track each other closely.

In Fig. 9, we use another approach to show the error by taking the difference between the solutions with mesh sizes and time steps  $(h_x, h_y, k)$  and  $(h_x/2, h_y/2, k/4)$ . We show the error both for one path, as well as  $L_2$  in probability  $(E_L[|U_f(t, x_0, y_0) - U_c(t, x_0, y_0)|^2])^{1/2}$ , where  $E_L$  is the empirical mean with  $L = 100$  Monte Carlo samples.

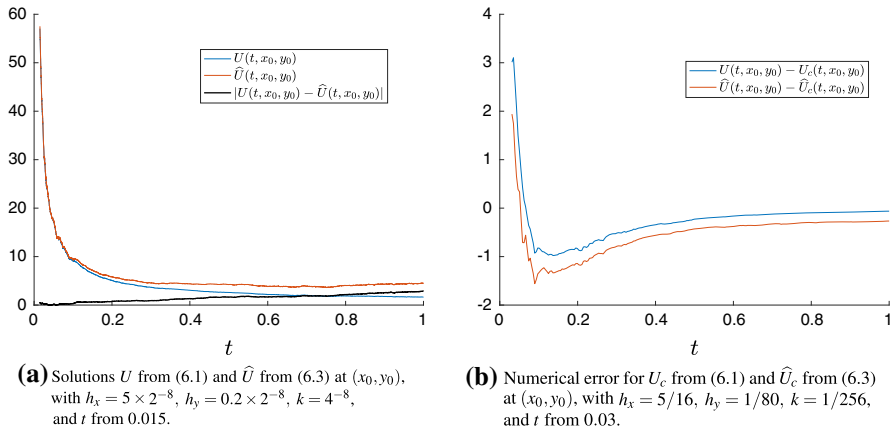
## 7 Conclusions

We studied a two-dimensional parabolic SPDE arising from a filtering problem. We proved mean-square stability and convergence for a semi-implicit Milstein discretisation scheme in terms of  $L_2$  in probability, and pointwise as well as  $L_2$  in space (as defined previously). To reduce the complexity, we also implemented an ADI version of the scheme, and provided corresponding convergence results.

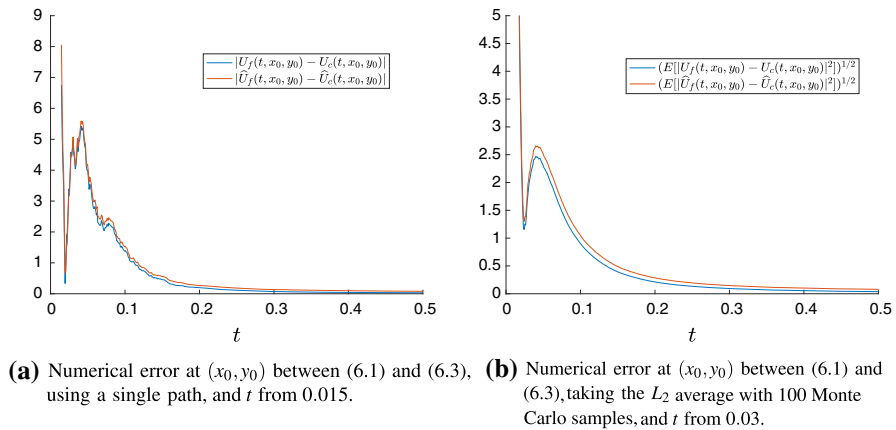
Further research is needed to analyse almost sure convergence, which is of interest for filtering applications and does not follow directly from our analysis.

Another open question is a complete analysis of the numerical approximation of initial-boundary value problems (as opposed to problems posed on  $\mathbb{R}^d$ ) for the considered SPDE, when the regularity at the boundary is lost. For example, for the 1-d SPDE with constant coefficients on the half-line,





**Fig. 8** Single path realisation of the solution at  $(x_0, y_0)$  over time, of SPDE (6.1) with variable coefficients and (6.3) with frozen coefficients



**Fig. 9** Numerical error at  $(x_0, y_0)$  by taking difference between  $(h_x, h_y, k)$  and  $(h_x/2, h_y/2, k/4)$  to the schemes (6.2) and (6.4), where  $h_x = 5/32$ ,  $h_y = 1/160$ ,  $k = 1/1024$

$$dv = -\mu \frac{\partial v}{\partial x} dt + \frac{1}{2} \frac{\partial^2 v}{\partial x^2} dt - \sqrt{\rho} \frac{\partial v}{\partial x} dM_t, \quad (t, x) \in (0, T) \times \mathbb{R}^+,$$

$$v(t, 0) = 0,$$

with initial condition  $v(0, \cdot) \in H^1$ , the second derivative can be unbounded, i.e.,  $v(t, \cdot) \notin H^2$ . This and more general forms have been studied in [6, 28]. In such cases, the assumptions on Galerkin approximations in papers previously mentioned such as in [2, 4] are not established in the literature, hence a new approach for the numerical analysis is to be developed.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution,

and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Bain, A., Crisan, D.: *Fundamentals of Stochastic Filtering*, vol. 60 of stochastic modelling and applied probability. Springer, New York, NY (2009)
2. Barth, A., Lang, A.: Milstein approximation for advection–diffusion equations driven by multiplicative noncontinuous martingale noises. *Appl. Math. Optim.* **66**(3), 387–413 (2012)
3. Barth, A., Lang, A.:  $L^p$  and almost sure convergence of a Milstein scheme for stochastic partial differential equations. *Stoch. Process. Appl.* **123**(5), 1563–1587 (2013)
4. Barth, A., Lang, A., Schwab, C.: Multilevel Monte Carlo method for parabolic stochastic partial differential equations. *BIT Numer. Math.* **53**(1), 3–27 (2013)
5. Buckwar, E., Kelly, C.: Towards a systematic linear stability analysis of numerical methods for systems of stochastic differential equations. *SIAM J. Numer. Anal.* **48**(1), 298–321 (2010)
6. Bush, N., Hambly, B.M., Haworth, H., Jin, L., Reisinger, C.: Stochastic evolution equations in portfolio credit modelling. *SIAM J. Financ. Math.* **2**(1), 627–664 (2011)
7. Carter, R., Giles, M.B.: Sharp error estimates for discretizations of the 1D convection-diffusion equation with Dirac initial data. *IMA J. Numer. Anal.* **27**(2), 406–425 (2007)
8. Clark, J.M.C., Cameron, R.J.: The maximum rate of convergence of discrete approximations for stochastic differential equations. In: Grigelionis, B. (ed.) *Stochastic Differential Systems Filtering and Control*, pp. 162–171. Springer, Berlin (1980)
9. Craig, I.J.D., Sneyd, A.D.: An alternating-direction implicit scheme for parabolic equations with mixed derivatives. *Comput. Math. Appl.* **16**(4), 341–350 (1988)
10. Crisan, D.: Exact rates of convergence for a branching particle approximation to the solution of the Zakai equation. *Ann. Probab.* **31**(2), 693–718 (2003)
11. Crisan, D., Lyons, T.: A particle approximation of the solution of the Kushner–Stratonovitch equation. *Probab. Theory Relat. Fields* **115**(4), 549–578 (1999)
12. Crisan, D., Xiong, J.: Approximate McKean–Vlasov representations for a class of SPDEs. *Stoch. Int. J. Probab. Stoch. Process.* **82**(1), 53–68 (2010)
13. Daretiotis, K., Leahy, J.M.: Finite difference schemes for linear stochastic integro–differential equations. *Stoch. Process. Appl.* **126**(10), 3202–3234 (2016)
14. Davie, A.M., Gaines, J.: Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations. *Math. Comput.* **70**(233), 121–134 (2001)
15. Gaines, J.G., Lyons, T.J.: Random generation of stochastic area integrals. *SIAM J. Appl. Math.* **54**(4), 1132–1146 (1994)
16. Giles, M.B., Reisinger, C.: Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance. *SIAM J. Financ. Math.* **3**(1), 572–592 (2012)
17. Gobet, E., Pages, G., Pham, H., Printems, J.: Discretization and simulation of the Zakai equation. *SIAM J. Numer. Anal.* **44**(6), 2505–2538 (2006)
18. Gyöngy, I.: Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space–time white noise I. *Potential Anal.* **1**(9), 1–25 (1998)
19. Gyöngy, I.: Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space–time white noise II. *Potential Anal.* **11**(1), 1–37 (1999)
20. Gyöngy, I., Nualart, D.: Implicit scheme for stochastic parabolic partial differential equations driven by space–time white noise. *Potential Anal.* **7**(4), 725–757 (1997)
21. Hambly, B., Kollitopoulos, N.: Stochastic evolution equations for large portfolios of stochastic volatility models. *SIAM J. Financ. Math.* **8**(1), 962–1014 (2017)
22. Higham, D.J., Trefethen, L.N.: Stiffness of ODEs. *BIT Numer. Math.* **33**(2), 285–303 (1993)
23. Hundsdorfer, W., Verwer, J.G.: *Numerical Solution of Time-Dependent Advection–Diffusion–Reaction Equations*, vol. 33 of Springer series in computational mathematics. Springer, Berlin, Heidelberg (2003)
24. Jentzen, A., Kloeden, P.: The numerical approximation of stochastic partial differential equations. *Milan J. Math.* **77**(1), 205–244 (2009)
25. Jentzen, A., Kloeden, P.: Taylor expansions of solutions of stochastic partial differential equations with additive noise. *Ann. Probab.* **38**(2), 532–569 (2010)

26. Kloeden, P.E., Platen, E., Wright, I.W.: The approximation of multiple stochastic integrals. *Stoch. Anal. Appl.* **10**(4), 431–441 (1992)
27. Kruse, R.: Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise. *IMA J. Numer. Anal.* **34**(1), 217–251 (2014)
28. Krylov, N.V., Rozovskii, B.L.: Stochastic evolution equations. *J. Sov. Math.* **16**(4), 1233–1277 (1981)
29. Kurtz, T.G., Xiong, J.: Particle representations for a class of nonlinear SPDEs. *Stoch. Process. Appl.* **83**(1), 103–126 (1999)
30. Lang, A.: Mean square convergence of a semidiscrete scheme for SPDEs of Zakai type driven by square integrable martingales. In: International Conference on Computational Science, no. 1 in *Procedia Computer Science*, pp. 1615–1623 (2010)
31. Lang, A., Petersson, A., Thalhammer, A.: Mean-square stability analysis of approximations of stochastic differential equations in infinite dimensions. *BIT Numer. Math.* **57**(4), 963–990 (2017)
32. Malham, S.J.A., Wiese, A.: Efficient almost-exact Lévy area sampling. *Stat. Probab. Lett.* **88**, 50–55 (2014)
33. Müller-Gronbach, T.: Strong approximation of systems of stochastic differential equations. Ph.D. thesis, Darmstadt University of Technology (2002)
34. Peaceman, D.W., Rachford Jr., H.H.: The numerical solution of parabolic and elliptic differential equations. *J. Soc. Ind. Appl. Math.* **3**(1), 28–41 (1955)
35. Reisinger, C.: Mean-square stability and error analysis of implicit time-stepping schemes for linear parabolic SPDEs with multiplicative Wiener noise in the first derivative. *Int. J. Comput. Math.* **89**(18), 2562–2575 (2012)
36. Reisinger, C., Wang, Z.: Analysis of multi-index Monte Carlo estimators for a Zakai SPDE. *J. Comput. Math.* **36**(2), 202–236 (2018)
37. Walsh, J.B.: Finite element methods for parabolic stochastic PDE's. *Potential Anal.* **23**(1), 1–43 (2005)
38. Wiktorsson, M.: Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions. *Ann. Appl. Probab.* **11**(2), 470–487 (2001)
39. Wyna, M., Hout, K.J.: Convergence of the Modified Craig-Sneyd scheme for two-dimensional convection-diffusion equations with mixed derivative term. *J. Comput. Appl. Math.* **296**, 170–180 (2016)