A Two-Step Numerical Scheme in Time for Surface Quasi Geostrophic Equations Under Location Uncertainty

Camilla Fiorini, Pierre-Marie Boulvard, Long Li, Etienne Mémin

A Two-Step Numerical Scheme in Time for Surface Quasi Geostrophic Equations Under Location Uncertainty

Camilla Fiorini, Pierre-Marie Boulvard, Long Li, and Etienne Mémin

Abstract In this work we consider the surface quasi-geostrophic (SQG) system under location uncertainty (LU) and propose a Milstein-type scheme for these equations, which is then used in a multi-step method. The SQG system considered here consists of one stochastic partial differential equation, which models the stochastic transport of the buoyancy, and a linear operator linking the velocity and the buoyancy. In the LU setting, the Euler-Maruyama scheme converges with weak order 1 and strong order 0.5. Our aim is to develop higher order schemes in time, based on a Milstein-type scheme in a multi-step framework. First we compared different kinds of Milstein schemes. The scheme with the best performance is then included in the two-step scheme. Finally, we show how our two-step scheme decreases the error in comparison to other multi-step schemes.

1 Introduction

The main aim of the modelling under location uncertainty (LU) consists in simulating on coarse meshes an enriched system mimicking a high resolution deterministic chaotic dynamics. Such LU models allow one to recover phenomena such as backscattering, dissipation and reorganisation on very coarse meshes. Furthermore, it provides a natural framework for uncertainty quantification analysis [14]. The LU framework, first introduced in [11], is based on the decomposition of the Lagrangian velocity into two components: a large-scale smooth component and

C. Fiorini (✉)
Laboratoire M2N, Conservatoire National des Arts et Métiers, Paris, France
e-mail: camilla.fiorini@lecnam.net

P.-M. Boulvard
Inria Paris, Equipe ANGE, Paris, France

Inria Rennes - Bretagne Atlantique, Équipe FLUMINANCE, Rennes, France

L. Li · E. Mémin
Inria Rennes - Bretagne Atlantique, Équipe FLUMINANCE, Rennes, France

B. Chapron et al. (eds.), *Stochastic Transport in Upper Ocean Dynamics*,
Mathematics of Planet Earth 10, https://doi.org/10.1007/978-3-031-18988-3_5

a small-scale fast oscillating one. This decomposition leads to a stochastic transport operator, and one can, in turn, develop the stochastic version of classical fluid-dynamics systems derived from the Navier–Stokes equations. SQG in particular consists of one stochastic partial differential equation (SPDE), which models the stochastic transport of the buoyancy, and a linear operator relating the velocity and the buoyancy:

$$\begin{cases} \mathrm{d}b_t = \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla b_t)\mathrm{d}t - \boldsymbol{v}^*\cdot\nabla b_t\mathrm{d}t - \nabla b_t\cdot\boldsymbol{\sigma}\mathrm{d}\mathbf{B}_t, \\ b_t = N(-\Delta)^{1/2}\psi, \\ \mathbf{u} = \nabla^{\perp}\psi, \end{cases} \tag{1}$$

where b_t is the buoyancy at time t, $\mathbf{u}$ the large-scale smooth velocity, N a constant depending on the vertical oscillation frequency of the buoyancy and a Coriolis parameter, $\mathbf{B}$ a Wiener process, ψ the stream function and $\boldsymbol{v}^* = \boldsymbol{u} - \frac{1}{2}\nabla\cdot\boldsymbol{a} + \boldsymbol{\sigma}\nabla\cdot\boldsymbol{\sigma}$ is a corrected velocity associated with the effect of the noise inhomogeneity on the advected variables. The spatial correlations of the noise are given through an integral kernel operator $\boldsymbol{\sigma}$ (here assumed deterministic and symmetric for sake of simplicity), and the variance matrix, $\boldsymbol{a}$, given by the matrix kernel of the operator $\boldsymbol{\sigma}\boldsymbol{\sigma}$ provides a local measure of the noise strength. For more details on the derivation of this system, see [10, 13]. In the rest of this work we will mainly focus on the first equation, and the last two will be condensed in $\mathbf{u} = \mathcal{H}(b)$. Concerning the modelling of the noise, we use the equivalent convenient spectral definition:

$$\boldsymbol{\sigma}\mathrm{d}\mathbf{B}_t = \sum_m \boldsymbol{\varphi}^m \mathrm{d}\beta_t^m,$$

where $\beta^m = \beta^m(t)$ are independent one-dimensional standard Brownian motions and $\boldsymbol{\varphi}^m = [\varphi_x^m, \varphi_y^m]^T(\mathbf{x})$ are basis functions. The number of terms involved in the sum is in theory infinite, but in numerical application a truncation is considered. In the definition of the numerical schemes we will thus assume that it is a finite sum. For the computation of the basis functions, two strategies are possible: an offline strategy, where they are defined from the eigenfunctions of an empirical covariance tensor built from high-resolution data as described in [10, 13]; of strategies, where the functions are updated during the simulation and in this case they are a function of the buoyancy b. With this representation, the variance tensor reads:

$$\boldsymbol{a} = \sum_m \boldsymbol{\varphi}^m(\boldsymbol{\varphi}^m)^T.$$

2 Numerical Schemes

In this section we derive a two-step numerical scheme in time for the SQG system under LU (SQG-LU). We compare this scheme to other multi-step schemes for the SPDE, in particular the ones developed in [5] and [4], and show how our scheme improves the precision. Concerning discretisation in space, standard spectral

methods are used: the linear terms are treated in the Fourier space, whilst the nonlinear terms are discretised in the physical space.

The derivation of the time scheme consists of two steps: first, we derive a class of Milstein schemes for SQG-LU and we empirically verify their convergence, then a two-step scheme is proposed.

2.1 Derivation of a Milstein Scheme

To design the Milstein schemes, we consider the integral form of the SPDE in (1), namely

$$b_t = b_{t_0} + \int_{t_0}^{t} \left(\frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla b_s) - \boldsymbol{v}^* \cdot \nabla b_s \right) \mathrm{d}s - \int_{t_0}^{t} \sum_m \nabla b_s \cdot \boldsymbol{\varphi}^m \mathrm{d}\beta_s^m, \tag{2}$$

and we can define the following functions:

$$f(b_t, t) = \frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla b_t) - \boldsymbol{v}^* \cdot \nabla b_t \qquad \text{and} \qquad g^m(b_t, t) = -\nabla b_t \cdot \boldsymbol{\varphi}^m. \tag{3}$$

We can now use the functional extension of the Itô formula [3] for both f and g to write their differential forms:

$$\begin{aligned} f(b_t, t) = f(b_{t_0}, t_0) + \int_{t_0}^{t} \frac{\partial f}{\partial s}(b_s, s) \mathrm{d}s + \int_{t_0}^{t} \frac{\partial f}{\partial b}(b_s, s) \mathrm{d}b_s \\ + \frac{1}{2} \int_{t_0}^{t} \frac{\partial^2 f}{\partial b^2}(b_s, s) \mathrm{d}\langle b, b \rangle_s \end{aligned} \tag{4}$$

$$\begin{aligned} g^m(b_t, t) = g^m(b_{t_0}, t_0) + \int_{t_0}^{t} \frac{\partial g^m}{\partial s}(b_s, s) \mathrm{d}s + \int_{t_0}^{t} \frac{\partial g^m}{\partial b}(b_s, s) \mathrm{d}b_s \\ + \frac{1}{2} \int_{t_0}^{t} \frac{\partial^2 g^m}{\partial b^2}(b_s, s) \mathrm{d}\langle b, b \rangle_s \end{aligned} \tag{5}$$

We remark that, since the basis $\boldsymbol{\varphi}^m$ is constant in time then so is $\boldsymbol{a}$ and the functions f and g^m do not depend explicitly on time, therefore $\partial f / \partial t = \partial g^m / \partial t = 0$.

Concerning the first derivatives with respect to b, it has to be interpreted as a Fréchet derivative. The Fréchet derivative of an operator F is the bounded linear operator $DF(\overline{x})$ which satisfies the following relation:

$$\lim_{\|h\| \to 0} \frac{\|F(\overline{x} + h) - F(\overline{x}) - DF(\overline{x})h\|}{\|h\|} = 0, \tag{6}$$

which implies that for a linear operator $DF(\overline{x})h = F(h)$. We start for g and use the fact that ∇ is a linear operator:

$$\frac{\partial g}{\partial b}(\overline{b})b = -\nabla b \cdot \boldsymbol{\varphi}^m - \nabla b \cdot \frac{\partial \boldsymbol{\varphi}^m}{\partial b}. \tag{7}$$

If the basis is computed offline, $\boldsymbol{\varphi}^m$ does not depend on b and therefore the second term in (7) is zero. If the basis is computed online and $\boldsymbol{\varphi}^m$ does depend on b, we can rewrite the second term of the sum by components and, using the chain rule, one has:

$$\nabla b \cdot \frac{\partial \boldsymbol{\varphi}^m}{\partial b} = \frac{\partial b}{\partial x}\frac{\partial \varphi_x^m}{\partial b} + \frac{\partial b}{\partial y}\frac{\partial \varphi_y^m}{\partial b} = \nabla \cdot \boldsymbol{\varphi}^m. \tag{8}$$

For the second term of f, i.e. $\boldsymbol{v}^* \cdot \nabla b$, the same considerations are valid. To compute the derivative of the first term of f, we remark that it is a composition and product of three operators, two of which are linear. We can define:

$$F_1(\boldsymbol{h}) = \frac{1}{2}\nabla \cdot \boldsymbol{h}, \quad F_2(b) = \boldsymbol{a}(b), \quad F_3(b) = \nabla b. \tag{9}$$

Using the chain rule and the linearity of F_1 and F_3 one has:

$$\begin{aligned} D\Big(F_1\big(F_2(b)F_3(b)\big)\Big)b &= DF_1\big(F_2(b)F_3(b)\big)\big(DF_2(b)F_3(b) + F_2(b)DF_3(b)\big)b \\ &= F_1\big(F_3(b)DF_2(b)b + F_2(b)F_3(b)\big) \\ &= \frac{1}{2}\nabla \cdot \left(\frac{\partial \boldsymbol{a}}{\partial b}\nabla b + \boldsymbol{a}\nabla b\right). \end{aligned} \tag{10}$$

Finally, with the same considerations used above, we remark that we can write $(\partial \boldsymbol{a}/\partial b)\nabla b = \nabla \cdot \boldsymbol{a}$. Therefore:

$$\frac{\partial f}{\partial b}(\overline{b})b = f(b) + \frac{1}{2}\nabla \cdot \nabla \cdot \boldsymbol{a} - \nabla \cdot \boldsymbol{v}^*, \quad \frac{\partial g^m}{\partial b}(\overline{b})b = g^m(b) - \nabla \cdot \boldsymbol{\varphi}^m. \tag{11}$$

As for the Itô covariation bracket, one has:

$$\langle b, b\rangle_t = \Big\langle \int_{t_0}^{\cdot} \sum_m g^m(b_s, s)\mathrm{d}\beta_s^m, \int_{t_0}^{\cdot} \sum_k g^k(b_\tau, \tau)\mathrm{d}\beta_\tau^k \Big\rangle_t = \int_{t_0}^{t} \left(\sum_m g^m(b_s, s)\right)^2 \mathrm{d}s$$

We now suppose to be in either one of the following cases:

- the basis functions $\boldsymbol{\varphi}^m$ (and therefore $\boldsymbol{a}$) do not depend on b and $\nabla \cdot \boldsymbol{v}^* = 0$,
- the basis functions $\boldsymbol{\varphi}^m$ depend on b but are such that $\nabla \cdot \boldsymbol{v}^* = \nabla \cdot \nabla \cdot \boldsymbol{a} = \nabla \cdot \boldsymbol{\sigma} = \nabla \cdot \boldsymbol{\varphi}^m = 0$.

It can be noticed that the first case corresponds to a noise defined from external high-resolution data (and thus that does not depend on the solution) while the second case boils down to impose an incompressibility condition constraint on the large scale component, $\nabla \cdot \boldsymbol{u} = 0$, that is indeed often considered in practice with particular scaling of the noise [1, 2]. With these assumptions, we have then:

$$\frac{\partial f}{\partial b} = \frac{\partial^2 f}{\partial b^2} = f, \quad \frac{\partial g^m}{\partial b} = \frac{\partial^2 g^m}{\partial b^2} = g^m. \tag{12}$$

We can now replace all these expressions into (4) and (5), and then (4) and (5) into (2). Keeping only the terms of order one or lower, we obtain:

$$b_t = b_{t_0} + f(b_{t_0})\Delta t + \sum_m g^m(b_{t_0})\Delta\beta^m + \int_{t_0}^t \int_{t_0}^s \sum_{m,k} g^m(g^k(b_\tau))\mathrm{d}\beta_\tau^k \mathrm{d}\beta_s^m, \tag{13}$$

where $\Delta t = t - t_0$ and $\Delta\beta^m = \beta_t^m - \beta_{t_0}^m$. We define the following quantities:

$$G^{m,k} := g^m(g^k(b_{t_0})), \quad I^{m,k} := \int_{t_0}^t \int_{t_0}^s \mathrm{d}\beta_\tau^k \mathrm{d}\beta_s^m,$$

then the double iterated Itô integral in (13) can be approximated as follows:

$$\sum_{m,k} G^{m,k} I^{m,k} = \sum_{m,k} G^{m,k} \frac{I^{m,k} + I^{k,m}}{2} + G^{m,k} \frac{I^{m,k} - I^{k,m}}{2}.$$

The first symmetric term can be computed analytically from Itô integration by part formulae, $I^{m,k} + I^{k,m} = \Delta\beta^m \Delta\beta^k - \delta_{m,k}\Delta t$, however the second antisymmetric term $(I^{m,k} - I^{k,m})/2 =: A_{t_0,t}^{m,k}$ cannot and it is known as the Lévy area.

2.1.1 Lévy Area Simulation

In this subsection, we briefly introduce the methods we used to simulate the Lévy area. More details can be found in [6, 8], where these methods were proposed. The first method to simulate the Lévy area will be referred to as the weak approximation in the rest of this work: in this method, we simulate a random variable that has the same moments as the Lévy area. The second method, which will be referred to as the conditional method, is a recursive method: the time interval (t_0, t) is recursively split into two subintervals of the same length, and the two following relations are used:

$$A_{t_0,t}^{m,k} = A_{t_0,u}^{m,k} + A_{u,t}^{m,k} + \frac{1}{2}\Big((\beta_u^m - \beta_{t_0}^m)(\beta_t^k - \beta_u^k) - (\beta_u^k - \beta_{t_0}^k)(\beta_t^m - \beta_u^m)\Big) \tag{14}$$

$$\mathbb{E}[A_{t_0,t}|\mathbf{B}_t - \mathbf{B}_{t_0}] = 0.$$

For more details on these two methods, see [7]. Finally, we consider a third approach, where we neglect the Lévy area. We remark that this approach is exact if $G^{m,k} = G^{k,m}$, which is not the case here.

2.2 Multi-Step Schemes

We next propose a two-step scheme in which the Milstein method is used as the prediction step and the Euler method is adopted as the correction step, it reads:

$$\begin{cases} b_t^* = b_{t_0} + f(b_{t_0}, \boldsymbol{u}_{t_0})\Delta t + \sum_m g^m(b_{t_0})\Delta\beta^m + \sum_{m,k} G^{m,k}\left(S_{t_0,t}^{m,k} + \tilde{A}_{t_0,t}^{m,k}\right) \\ \boldsymbol{u}_t^* = \boldsymbol{\mathcal{H}}(b_t^*) \\ b_t = \frac{1}{2}b_{t_0} + \frac{1}{2}\left(b_t^* + f(b_t^*, \boldsymbol{u}_t^*)\Delta t + \sum_m g^m(b_t^*)\Delta\beta^m\right) \end{cases} \tag{15}$$

where $S_{t_0,t}^{m,k} := (\Delta\beta^m\Delta\beta^k - \delta_{m,k}\Delta t)/2$ and $\tilde{A}_{t_0,t}^{m,k}$ is one of the approximations of the Lévy area described in the previous subsection. This scheme will be referred to as SRK2-EM (EM stands for Euler-Milstein not for Euler-Maruyama) in the rest of the paper.

In the next section, we first analyse the results of the Milstein schemes with the different Lévy area approximations in order to select the best one. Then, we compare our multi-step scheme to two other multi-step schemes developed in [5] and [4]. We briefly recall them here. The first one, based on a third order Runge-Kutta scheme, (SSPRK3) [5], is:

$$\begin{cases} b^{(1)} = b_{t_0} + f_s(b_{t_0}, \boldsymbol{u}_{t_0})\Delta t + \sum_m g^m(b_{t_0})\Delta\beta^m \\ \boldsymbol{u}^{(1)} = \boldsymbol{\mathcal{H}}(b^{(1)}) \\ b^{(2)} = \frac{3}{4}b_{t_0} + \frac{1}{4}\left(b^{(1)} + f_s(b^{(1)}, \boldsymbol{u}^{(1)})\Delta t + \sum_m g^m(b^{(1)})\Delta\beta^m\right) \\ \boldsymbol{u}^{(2)} = \boldsymbol{\mathcal{H}}(b^{(2)}) \\ b_t = \frac{1}{3}b_{t_0} + \frac{2}{3}\left(b^{(2)} + f_s(b^{(2)}, \boldsymbol{u}^{(2)})\Delta t + \sum_m g^m(b^{(2)})\Delta\beta^m\right) \end{cases} \tag{16}$$

where $f_s = f - \boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}b)/2$ denotes the modified drift under Stratonovich integral. The second one, relies on Euler-Heun method [4] equally for Stratonovich integral, reads:

$$\begin{cases} b^{(1)} = b_{t_0} + f_s(b_{t_0}, \boldsymbol{u}_{t_0})\Delta t + \sum_m g^m(b_{t_0})\Delta\beta^m \\ \boldsymbol{u}^{(1)} = \boldsymbol{\mathcal{H}}(b^{(1)}) \\ b_t = \frac{1}{2}b_{t_0} + \frac{1}{2}\left(b^{(1)} + f_s(b^{(1)}, \boldsymbol{u}^{(1)})\Delta t + \sum_m g^m(b^{(1)})\Delta\beta^m\right) \end{cases} \tag{17}$$

3 Numerical Results

In this section we show some numerical results. First, the effect of the different approximations of the Lévy area is studied on the Milstein scheme. Then, the multi-step scheme is assessed and compared to the ones already proposed in the literature. We focus on two variations of one specific test case plotted in Fig. 1: the initial condition (left) consists of two warm elliptical anticyclones on the bottom of the domain and two cold elliptical cyclones on the top. After one day under moderate noise (centre), the four structures have rotated of approximately 45^o. After one day under strong noise (right) the nonlinearity of the dynamic is more noticeable. One can find all the configuration details used for these simulations in Chapter 6 of [10] for the moderate noise configuration. For the strong noise, all the basis functions $\boldsymbol{\varphi}^m$ are multiplied by a factor 10.

We will use the following abbreviations for the different numerical schemes

- Euler: Euler-Maruyama scheme.
- Milstein-0: Milstein scheme without the Lévy area.
- Milstein-weak: Milstein scheme with the weak approximation of the Lévy area.
- Milstein-cond-n: Milstein scheme with the conditional approximation of the Lévy area. Here n stands for the number of times the interval is recursively split (cf. (14)).
- SRK2-EM: scheme (15) with $\tilde{A}^{m,k}_{t_0,t} = 0$.
- SSPRK3: scheme (16).
- Heun: scheme (17).

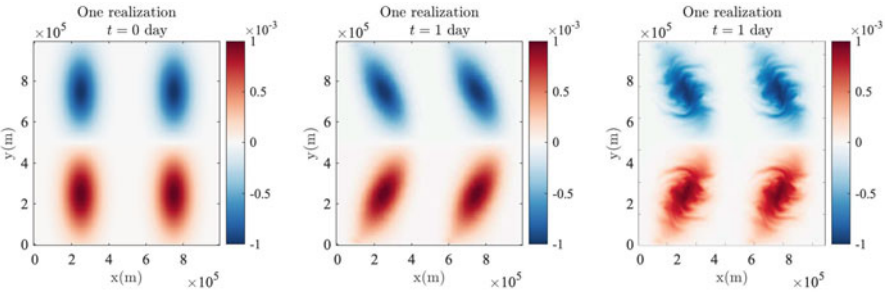

Fig. 1 Euler-Maruyama simulation of system (1) on a 128 × 128 spatial grid

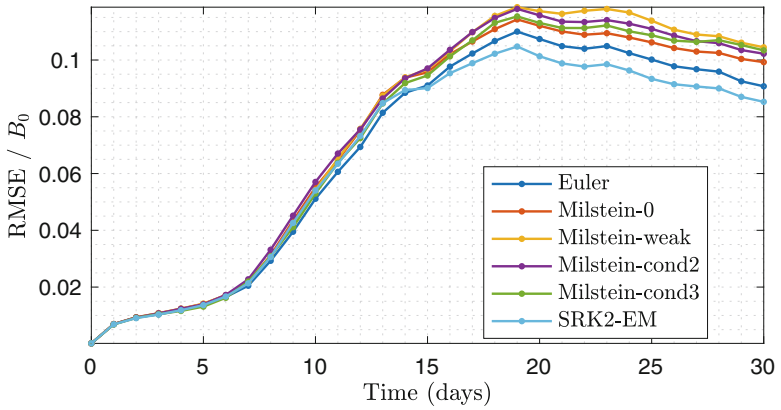

Fig. 2 RMSE (normalised by the amplitude of buoyancy $B_0 = 10^{-3}$ m/s^2) of different schemes during 30 days of simulation under moderate noise

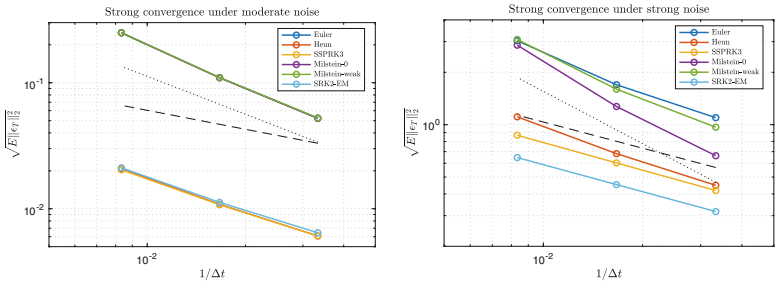

Fig. 3 Convergence of different schemes under weak and strong noise. Order 1 in dotted black, order 0.5 in dashed black

In Figs. 2 and 3 one can see the difference among the Euler-Maruyama scheme and all the Milstein schemes proposed. In Fig. 2 we plot for each scheme for a period of 30 day the root mean squared error (RMSE), defined as:

$$\text{RMSE} = \frac{1}{|\Omega|}\mathbb{E}\Big[\big\|b_h - b\big\|^2_{L^2(\Omega)}\Big]^{1/2}, \tag{18}$$

where Ω denotes the spatial domain, b_h is the numerical solution of stochastic system (1), and b stands for the reference solution downsampled from a high-resolution deterministic simulation (recall that the aim of the stochastic setting is to reproduce on coarse grid high-resolution deterministic simulations). The downsampling procedure consists of a first low-pass filtering performed in the Fourier domain and a subsequent subsampling operation. The expectations are estimated from 30 of realization. These results are obtained with a Δt twice as small for the Euler scheme with respect to the other schemes. One can observe that Milstein-0 performs slightly better than the other Milstein schemes.

In Fig. 3, we show the rate of strong convergence γ of all the schemes discussed, under weak and strong noise. Since the exact solution is unknown, we use the following method [15] to estimate γ, for a sufficiently small Δt:

$$\gamma \simeq \log_2\left(\frac{e_1}{e_2}\right), \text{ with } e_i := \mathbb{E}\left[\left\| b_h\left(T, \frac{\Delta t}{2^{i-1}}\right) - b_h\left(T, \frac{\Delta t}{2^i}\right)\right\|^2_{L^2(\Omega)}\right]^{1/2},$$

where $b_h(T, \Delta t)$ is the numerical solution at the final time T obtained with a time step Δt. It is important to underline that in order for this method to work, the Brownian trajectories must be fixed. We applied this method for time steps 30, 60, 120, 240, hence obtaining two estimates for γ. Is is important to remark that the value of the time steps is given in seconds and the time-scale of the studied phenomenon is of the order of one day. For reference, the CFL condition for this problem at the initial time would give a time step around 300 s. The smallest time step we considered to obtain this estimate is ten times smaller than this. As one can see from Fig. 3, under weak noise all the one-step schemes provide almost identical results and all the multi-step schemes are very similar. It is hard to distinguish among the different numerical schemes proposed. In particular, for the considered span of time steps, the error of the Euler scheme under moderate noise displays a linear trend and the prevailing convergence order in this case is one. The reason of that is explained in Appendix.

Under strong noise, it is easier to see the differences among the schemes. Milstein-weak is a slight improvement on the Euler-Maruyama, but its rate of convergence is far from 1. Milstein-0 has the highest rate of convergence among all the schemes.

In conclusion, Milstein-0 seem to perform better than the other Milstein schemes. Furthermore, it is less computationally demanding. For these reasons, we built our two-step scheme based on Milstein-0.

In Fig. 3 we also compare the multi-step schemes mentioned above: they all have a similar behaviour, with a rate of convergence $0.5 \leq \gamma \leq 1$, but a much smaller error when compared to the one-step schemes. In particular, the two-step scheme proposed in this work (SRK2-EM in the figures) yields the smallest error of all for this test case. The SRK2-EM schemes also yields the smallest RMSE (cf. Fig. 2).

4 Conclusion and Perspectives

The Milstein schemes analysed in this work improve the numerical results, in particular when used in a multi-step framework. The Lévy area does not seem to play a key role in these test cases, which allows us to drastically reduce the computational costs. It must be pointed out that under weak noise, all the schemes tested provide very similar results. Some ongoing and future work include the understanding of the (non) importance of the Lévy area and whether this is related to the test case, the equations, or other factors.

Appendix: Convergence of Euler-Maruyama Scheme Under Moderate Noise

To study the behaviour of our system under moderate noise, we use the formalism of [12]; in particular, we write our system in the following generic form:

$$dX_t = a(x,t)dt + \epsilon b(x,t)dW_t + \epsilon^2 c(x,t)dt, \quad t \in [0,T] \tag{19}$$

with a, b, c, being jointly L^2-measurable in (x,t), Lipschitz, bounded linear-growth functions in x.

Let $Y_\cdot^\delta$ be an Euler-Maruyama integration scheme for $X_\cdot$ with integration step δ. Then we may prove in a similar fashion to theorem 4.5.4 in [9] that:

1. $\mathbb{E}[X_t]^2 \leq C, \quad \forall t \in [0,T]$
2. $\mathbb{E}\left[|X_{t+\delta} - Y_{t+\delta}^\delta| \middle| X_{t+\delta} = x\right] \leq K(x)(\delta + \sqrt{\epsilon}\sqrt{\delta} + \circ(\delta))$.

Using this and the Lipschitziannity of the coefficients in (19), we may prove a result, to some extent similar to theorem 2.1 in [12], namely that

$$\mathbb{E}\left[\sup_{t_0 \leq t \leq T} |X_t - Y_t^\delta| \middle| X_{t_0} = x\right] \leq K'(x)(\delta + \sqrt{\epsilon}\sqrt{\delta} + \circ(\delta)). \tag{20}$$

In light of this estimate, we may interpret the convergence rate displayed in Fig. 3 as a case where δ is not small enough when compared to ϵ so that $\sqrt{\epsilon}\sqrt{\delta}$ does not necessarily prevail over δ which is evidenced by the linear rate of convergence.

References

1. W. Bauer, P. Chandramouli, B. Chapron, L. Li, and E. Mémin. Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach. *Journal of Physical Oceanography*, 50(4):983–1003, 2020.
2. R. Brecht, L. Li, W. Bauer, and E. Mémin. Rotating shallow water flow under location uncertainty with a structure-preserving discretization. *J. of Advances in Modeling of Earth Systems*, 13(12), 2021.
3. R. Cont and D. Fournie. A functional extension of the Itō formula. *Comptes Rendus Mathematique*, 348(1–2):57–61, 2010.
4. C. Cotter, D. Crisan, D. Holm, W. Pan, and I. Shevchenko. Modelling uncertainty using stochastic transport noise in a 2-layer quasi-geostrophic model. *Foundations of Data Science*, 2(2):173, 2020.
5. C. Cotter, D. Crisan, D. D. Holm, W. Pan, and I. Shevchenko. Numerically modeling stochastic lie transport in fluid dynamics. *Multiscale Modeling & Simulation*, 17(1):192–232, 2019.
6. G. Flint and T. Lyons. Pathwise approximation of sdes by coupling piecewise abelian rough paths. *arXiv preprint arXiv:1505.01298*, 2015.
7. J. Foster. Lévy area simulation. https://github.com/james-m-foster/levy-area-simulation/blob/master/levy_area_simulation.ipynb, 2019.

8. T. L. J. Foster and H. Oberhauser. An optimal polynomial approximation of Brownian motion. *SIAM Journal on Numerical Analysis*, 58(3):1393–1421, 2020.
9. P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer-Verlag Inc, Berlin; New York, 1995.
10. L. Li. *Stochastic modeling and numerical simulation of ocean dynamics*. PhD thesis, Université Rennes 1, 2021.
11. E. Mémin. Fluid flow dynamics under location uncertainty. *Geophysical & Astrophysical Fluid Dynamics*, 108(2):119–146, 2014.
12. G. N. Milstein and M. V. Tretyakov. Numerical methods in the weak sense for stochastic differential equations with small noise. *SIAM J. Numer. Anal.*, 34(6):2142–2167.
13. V. Resseguier, L. Li, G. Jouan, P. Derian, E. Mémin, and B. Chapron. New trends in ensemble forecast strategy: uncertainty quantification for coarse-grid computational fluid dynamics. *Archives of Computational Methods in Engineering*, 28(1):1886–1784, 2020.
14. V. Resseguier, A. M. Picard, E. Mémin, and B. Chapron. Quantifying truncation-related uncertainties in unsteady fluid dynamics reduced order models. *SIAM/ASA Journal on Uncertainty Quantification*, 9(3):1152–1183, 2021.
15. K. Schmitz A. and W. T. Shaw. Measure order of convergence without an exact solution, Euler vs Milstein scheme. *International Journal of Pure and Applied Mathematics*, 24(3):365–381, 2005.