PRECONDITIONING FOR THE HIGH-ORDER SAMPLING OF THE INVARIANT DISTRIBUTION OF PARABOLIC SEMILINEAR SPDES

CHARLES-EDOUARD BRÉHIER, ADRIEN BUSNOT LAURENT, ARNAUD DEBUSSCHE, AND GILLES VILMART

ABSTRACT. For a class of ergodic parabolic semilinear stochastic partial differential equations (SPDEs) with gradient structure, we introduce a preconditioning technique and design high-order integrators for the approximation of the invariant distribution. The preconditioning yields improved temporal regularity of the dynamics while preserving the invariant distribution and allows the application of postprocessed integrators. For the semilinear heat equation driven by space-time white noise in dimension 1, we obtain new temporal integrators with orders 1 and 2 for sampling the invariant distribution with a minor overcost compared to the standard semilinear implicit Euler method of order 1/2. Numerical experiments confirm the theoretical findings and illustrate the efficiency of the approach.

Keywords: stochastic partial differential equations, invariant distribution, ergodicity, Gibbs distribution, space-time white noise, postprocessed integrator.

AMS subject classification (2020): 60H35, 60H15, 37M25.

1. INTRODUCTION

The goal of the present paper is to estimate numerically integrals in infinite dimension of the form

$$\int_H \varphi d\mu_\star = \lim_{t \to \infty} \mathbb{E}\big[\varphi(X(t))\big],$$

where the probability distribution $\mu_\star$ is defined on a separable Hilbert space H, $\varphi : H \to \mathbb{R}$ is a function with sufficient regularity, and $X = \big(X(t)\big)_{t \geqslant 0}$ is an ergodic H-valued Markov process with unique invariant distribution $\mu_\star$. Precisely, we consider stochastic evolution equations of the form

$$(1) \qquad dX(t) = AX(t)dt + F(X(t))dt + dW(t),$$

where $\big(W(t)\big)_{t \geqslant 0}$ is a cylindrical Wiener process, and the nonlinear term $F : H \to H$ is a Lipschitz continuous function on H. We assume a gradient structure: one has

$$F(x) = -DV(x), \quad \forall\, x \in H,$$

for a given potential $V : H \to \mathbb{R}$. The linear operator A is assumed to be self-adjoint, negative, and to be such that $Q = (-A)^{-1}$ is a trace-class operator. The class of problems (1) encompasses in particular parabolic semilinear stochastic partial differential equations (SPDEs),

$$(2) \qquad \frac{\partial x(t,z)}{\partial t} = \frac{\partial^2 x(t,z)}{\partial z^2} + f\big(x(t,z)\big) + \dot{W}(t,z), \quad t > 0,\ z \in (0,1),$$

driven by space-time white noise, with homogeneous Dirichlet boundary conditions, and for a smooth and Lipschitz nonlinearity $f : \mathbb{R} \to \mathbb{R}$. In this case, one has $H = L^2(0,1)$, and for all $x \in H$, $V(x) = -\int_0^1 U(x(z))dz$, where $U' = f$. In the absence of nonlinearity $F = 0$, the probability distribution $\mu_\star$ coincides with the Gaussian distribution $\nu = \mathcal{N}(0, \frac{1}{2}Q)$ with mean zero and covariance operator $\frac{1}{2}Q$.

Under technical assumptions, the invariant distribution $\mu_\star$ is absolutely continuous with respect to the Gaussian distribution ν with a Gibbs density,

$$d\mu_\star(x) = Z^{-1} \exp\big(-2V(x)\big)d\nu(x), \tag{3}$$

where $Z = \int_H e^{-2V(x)}d\nu(x)$ is the normalization constant. The aim of this paper is to design numerical integrators for sampling the invariant distribution $\mu_\star$ with high order of convergence with respect to the discretisation timestep in spite of the low regularity of the solution of the dynamics.

Consider an integrator given by a discrete-time Markov process

$$Y_{n+1}^{\Delta t} = \psi^{\Delta t}(Y_n^{\Delta t}), \tag{4}$$

where $\psi^{\Delta t} : H \to H$ is a random mapping depending on the time-step size Δt. Assuming ergodicity of the Markov process $Y_n^{\Delta t}$ with the invariant distribution $\mu^{\Delta t}$, where we recall that $\mu^{\Delta t} \neq \mu_\star$ in general, we define the error for the invariant distribution

$$\epsilon(\varphi, \Delta t) = \int_H \varphi d\mu^{\Delta t} - \int_H \varphi d\mu_\star,$$

for sufficiently smooth test functions $\varphi : H \to \mathbb{R}$. Note that a spatial discretization is also required in practice, however we mostly focus on the temporal discretization error in this article. The discretization is said to be of order q in time for the approximation of the invariant distribution, if for any sufficiently smooth function $\varphi : H \to \mathbb{R}$, there exists $C(\varphi) \in (0, \infty)$ such that for any time-step size Δt, one has

$$|\epsilon(\varphi, \Delta t)| \leqslant C(\varphi)\Delta t^q.$$

Recall that the discretization is said to be of weak order p if for any time $T \in (0, \infty)$ and any sufficiently smooth function $\varphi : H \to \mathbb{R}$, there exists $C(T, \varphi) \in (0, \infty)$ such that for any time-step size $\Delta t = T/N$ where $N \in \mathbb{N}$, one has

$$\big|\mathbb{E}[\varphi(Y_N^{\Delta t})] - \mathbb{E}[\varphi(X(N\Delta t))]\big| \leqslant C(T, \varphi)\Delta t^p.$$

Under mild assumptions, the order q for the invariant distribution is larger than the weak order p, i.e. $q \geqslant p$. Indeed, in the weak error estimates above, one is able to show that $C(T, \varphi)$ may be chosen independent of T.

The main contribution of this article is to propose and analyse integrators of order $q = 1$ and $q = 2$ for the approximation of the invariant distribution. This is a non-trivial result as for parabolic semilinear SPDEs, the weak order for the linear implicit Euler scheme is $p = \frac{1}{2}^-$ (which is consistent with the $\frac{1}{4}^-$ Hölder regularity of the trajectories). We refer for instance to [5, 9, 12, 13, 15, 29, 34, 35] for works on weak convergence of numerical schemes applied to semilinear parabolic SPDEs. Concerning the numerical approximation of the invariant distributions of such SPDEs, preliminary works are [8, 16], where uniform in time weak error estimates have been obtained, for temporal and full discretization respectively. We also refer to more recent works [10, 28, 31, 49, 60, 67] for similar results, in particular for equations with non-globally Lipschitz continuous nonlinearities. The article [24] provides central limit theorems instead of weak error estimates. The recent work [11] deals with the same type of Gibbs invariant distributions as in this manuscript, with the introduction and analysis of a modified Euler scheme which preserves regularity of trajectories.

Finally, the question of the numerical approximation of invariant distributions has been studied for other types of SPDEs: Schrödinger equations [27, 45, 46], Maxwell equations [25, 26], damped wave equations [53, 23], Burgers equations [7], Navier–Stokes equations [40, 69], Burgers–Huxley equations [66, 68], Cahn–Hilliard equations [36].

To the best of our knowledge, no higher-order method with a rigorous analysis is known in the literature for semi-linear SPDEs of the form (1). We mention the first attempt [17] on the construction of higher-order methods for the approximation of the invariant distribution, where

new schemes were designed. Numerical evidence suggested that one may achieve order $q \geqslant 1$, and in some cases, order $q = \frac{3}{2}^{-}$ has been proved in the simplified case of a linear function F.

The construction of the higher-order methods presented in this article is based on a preconditioning technique. Applying a suitable preconditioning operator $\mathcal{P}$ allows us to modify the dynamics (1) without altering the invariant distribution (3). In general, one obtains the preconditioned dynamics,

$$dY^{\mathcal{P}}(t) = \mathcal{P}\Big(AY^{\mathcal{P}}(t) + F(Y^{\mathcal{P}}(t))\Big)dt + \mathcal{P}^{\frac{1}{2}}dW(t).$$

In this paper, we shall focus on the choice $\mathcal{P} = (-A)^{-1} = Q$, which appears already in [30, 63, 42] in the context of Markov Chain Monte-Carlo (MCMC) methods or for the sampling of conditioned diffusions using SPDE techniques [41, 43, 44]. There has recently been a growing interest for preconditioning techniques for finite dimensional stochastic systems, in the context of high-dimensional Langevin dynamics in molecular dynamics [6, 47] and in particular for overdamped Langevin dynamics, see [63, 64] and recently [21, 32, 57, 58, 59], where preconditioning permits to enhance the convergence to equilibrium in the transient phase. Under mild assumptions, the stochastic evolution equation

$$dY(t) = -Y(t)dt + QF(Y(t))dt + Q^{\frac{1}{2}}dW(t)$$

is also ergodic, with the same invariant distribution $\mu_{\star}$ as the original dynamics. Note that the preservation of the invariant dynamics when preconditioning the dynamics crucially relies on the gradient structure condition $F = -DV$. As will be clearly explained below, the preconditioning technique permits to increase the temporal regularity of the trajectories, while the spatial regularity is not modified. Indeed, Y is solution to an infinite dimensional stochastic evolution equation driven by a trace-class noise, thus trajectories of Y are $\frac{1}{2}^{-}$ Hölder continuous. This permits to benefit from standard techniques that have been extensively studied in recent years for finite dimensional SDEs, in particular for the design of high order schemes for the invariant distribution (non-Markovian schemes [54], postprocessing [1, 65], modified equations [2, 20, 52]), and for the design of variance reduction techniques (Multilevel Monte-Carlo methods (MLMC) [38, 39], perturbations [3, 37, 56]).

While stochastic Runge-Kutta methods are a popular approach for the design of high-order weak methods for SDEs [22], the use of postprocessing techniques [17, 65] yields cheap efficient methods with high-order specifically for the invariant distribution, i.e. they can be used to increase the order q without modifying the weak order p. Given an integrator (4), a postprocessor is a random mapping $\overline{\psi}^{\Delta t} : H \to H$, which defines a perturbation $\overline{Y}_n^{\Delta t}$ of the integrator output $Y_n^{\Delta t}$, such that $(Y_n^{\Delta t}, \overline{Y}_n^{\Delta t})$ is a Markov chain and

$$\overline{Y}_n^{\Delta t} = \overline{\psi}^{\Delta t}(Y_n^{\Delta t}).$$

With a good choice of postprocessor, for finite dimensional ergodic SDEs it is shown in [65] that the order for the invariant distribution of $\overline{Y}_n^{\Delta t}$ can be higher than the one of $Y_n^{\Delta t}$. Moreover, as the postprocessor correction is only applied at the last step of the method (hence the name postprocessing), the accuracy is improved with negligible additional computational cost.

In this work, we propose in particular the following new postprocessed method where $\overline{Y}_n$ yields second order for the invariant distribution $\mu_{\star}$ of the original dynamics (1):

$$Y_{n+1} = Y_n + \Delta t G(Y_n + \frac{1}{2}\Delta W_n^Q) + \Delta W_n^Q,$$

$$\overline{Y}_n = Y_n + \frac{1}{2}\Delta W_n^Q,$$

where $G(y) = -y + QF(y)$. The above method is an extension of the second order Leimkuhler–Matthews method [54, 55] for preconditioned stochastic partial differential equations. That method has been first introduced in the finite dimensional setting in the context of overdamped Langevin

dynamics. It has originally been proposed as a non-Markovian method and has then been refor-mulated as a postprocessed method in [65]. It was also recently extended for nonlinear diffusion coefficient in [18, 21]. Compared with the processed integrator introduced in [17], where no pre-conditioning is employed, one is able to provide rigorous justification for achieving the higher order $q = 2$ for the approximation of the invariant distribution.

Let us present the main results of this article. Section 3 presents the application of the precon-ditioning technique in the context of stochastic evolution equations. Proposition 3.1 shows that the preconditioning does not modify the invariant distribution in the considered setting (see also Lemma 3.3 for a more general version). In the error analysis, one employs solutions to auxiliary Kolmogorov and Poisson equations, Proposition 3.5 provides some regularity properties of solutions to Kolmogorov equations. In Section 4, we consider the standard explicit Euler scheme

$$Y_{n+1} = Y_n + \Delta t G(Y_n) + \Delta W_n^Q$$

applied to the preconditioned stochastic evolution equation, but the postprocessing technique has not been applied yet. We show for that method that $q = 1$, see Theorem 4.5. In fact, weak error estimates with order $p = 1$, which are uniform with respect to time T, are established, and one obtains $q = p = 1$ letting T go to ∞. Note that obtaining methods of order $q = 1$ is already an improvement compared with methods applied to the original dynamics, which only have order $\frac{1}{2}-$. Next, in Section 5, we show that the method proposed above, which combines the preconditioning and the postprocessing techniques, achieves order $q = 2$, see Theorem 5.4. We first exhibit order conditions for a more general class of methods, and then show that they are satisfied for the considered scheme. Finally, numerical experiments are presented in Section 6, in order to illustrate the main results and compare different methods.

2. Setting

We present in this section the standard tools for the numerical analysis of SPDEs and refer to the monographs [33, 48, 50, 61] for further details. Let H be a separable, infinite dimensional, Hilbert space. The inner product and the norm are denoted by $\langle \cdot, \cdot \rangle$ and $|\cdot|$ respectively. The space of bounded linear operators from H to H is denoted by $\mathcal{L}(H)$, and the associated norm is denoted by $\|\cdot\|_{\mathcal{L}(H)}$. The following standard convention is used in this article. Let $\varphi : H \to \mathbb{R}$ be a function of class $\mathcal{C}^2$. For all $x \in H$, the first-order derivative $D\varphi(x)$ and the second-order derivative $D^2\varphi(x)$ are interpreted as elements of H and of $\mathcal{L}(H)$ respectively, owing to the Riesz theorem.

Definition 2.1. *Let* $(e_k)_{k\in\mathbb{N}}$ *be a complete orthonormal system of* H *and* $(q_k)_{k\in\mathbb{N}}$ *be a sequence of positive real numbers, such that* $\sum_{k=1}^{\infty} q_k < \infty$. *Set*

$$Qx = \sum_{k\in\mathbb{N}} q_k \langle x, e_k \rangle e_k , \quad x \in H$$

$$-Ax = \sum_{k\in\mathbb{N}} \lambda_k \langle x, e_k \rangle e_k , \quad x \in D(A),$$

with $\lambda_k = q_k^{-1}$ *for all* $k \in \mathbb{N}$, *and* $D(A) = \left\{ x \in H; \ \sum_{k=1}^{\infty} \lambda_k^2 \langle x, e_k \rangle^2 < \infty \right\}$. *Finally, let* $\nu = \mathcal{N}(0, \frac{1}{2}Q)$ *be the centered Gaussian distribution on* H, *with covariance operator* $\frac{1}{2}Q$.

Note that Q is a self-adjoint, nonnegative and trace-class $(\text{Tr}(Q) = \sum_{k=1}^{\infty} \langle Qe_k, e_k \rangle = \sum_{k=1}^{\infty} q_k < \infty)$, thus ν is a well-defined Gaussian distribution on H. Recall that ν is the distribution of the H-valued random variable $\sum_{k\in\mathbb{N}} \frac{\sqrt{q_k}}{\sqrt{2}} \xi_k e_k$, where $(\xi_k)_{k\in\mathbb{N}}$ are independent standard real-valued Gaussian random variables, $\xi_k \sim \mathcal{N}(0, 1)$. Without loss of generality, assume that the sequence $(q_k)_{k\in\mathbb{N}}$ is non-increasing. The sequence $(\lambda_k)_{k\in\mathbb{N}}$ is then non-decreasing, and $\lambda_k \underset{k\to\infty}{\to} \infty$. Observe that $\|Q\|_{\mathcal{L}(H)} = q_1 = \lambda_1^{-1}$. For example, let $H \in L^2(0, 1)$ and for all $k \in \mathbb{N}$, set $e_k(\cdot) = \sqrt{2}\sin(k\pi\cdot)$ and

$\lambda_k = \pi^2 k^2$. Then, the operator A given in Definition 2.1 is the Laplace operator with homogeneous Dirichlet boundary conditions, as described in (2).

For $\alpha \in (0,1)$, we define the unbounded linear operator $(-A)^\alpha$ as follows:

$$(-A)^\alpha x = \sum_{k \in \mathbb{N}} \langle x, e_k \rangle \lambda_k^\alpha e_k,$$

for all $x \in D((-A)^\alpha) = \left\{ x \in H; \sum_{k \in \mathbb{N}} \lambda_k^{2\alpha} \langle x, e_k \rangle^2 < \infty \right\}$. These operators are helpful in quantifying the spatial regularity of the stochastic processes defined below. Let

$$(5) \qquad \alpha_{\sup} = \sup \left\{ \alpha; \operatorname{Tr}\left((-A)^{2\alpha} Q\right) = \sum_{k=1}^\infty q_k^{1-2\alpha} < \infty \right\}.$$

We assume that $\alpha_{\sup} > 0$. In particular, $\alpha_{\sup} = 1/4$ for the Laplace operator with homogeneous Dirichlet boundary conditions. Note that $\alpha_{\sup} \leqslant 1/2$.

Let $(\beta_k)_{k \in \mathbb{N}}$ be a sequence of independent standard real-valued Wiener processes, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfies the usual conditions. For all $t \geqslant 0$, set

$$W(t) = \sum_{k \in \mathbb{N}} \beta_k(t) e_k,$$

$$W^Q(t) = \sum_{k \in \mathbb{N}} \sqrt{q_k} \beta_k(t) e_k.$$

On the one hand, $\left(W(t)\right)_{t \geqslant 0}$ is a cylindrical Wiener process and does not take values in H. On the other hand, $\left(W^Q(t)\right)_{t \geqslant 0}$ is a Q-Wiener process and takes values in H. It is straightforward to check that ν is the unique invariant distribution of the H-valued Ornstein-Uhlenbeck processes X^{OU} and Y^{OU} given by

$$X^{\mathrm{OU}}(t) = e^{tA} X^{\mathrm{OU}}(0) + \int_0^t e^{(t-s)A} dW(s) , \quad Y^{\mathrm{OU}}(t) = e^{-t} Y^{\mathrm{OU}}(0) + \int_0^t e^{s-t} dW^Q(s),$$

where $e^{tA} x = \sum_{k \in \mathbb{N}} e^{-t\lambda_k} \langle x, e_k \rangle e_k$ for all $t \geqslant 0$. Note that the stochastic convolution is well-defined with values in H:

$$\mathbb{E}| \int_0^t e^{(t-s)A} dW(s)|^2 = \sum_{k=1}^\infty \mathbb{E}| \int_0^t e^{-\lambda_k(t-s)} d\beta_k(s)|^2 = \sum_{k=1}^\infty \frac{1 - e^{-2\lambda_k t}}{2\lambda_k} < \infty.$$

The Ornstein-Uhlenbeck processes X^{OU} and Y^{OU} are solutions of the stochastic evolution equations, driven by additive noise,

$$dX^{\mathrm{OU}}(t) = AX^{\mathrm{OU}}(t)dt + dW(t) , \quad dY^{\mathrm{OU}}(t) = -Y^{\mathrm{OU}}(t)dt + dW^Q(t),$$

respectively.

The main question studied in this article is the estimation of $\int \varphi d\mu_\star$, using numerical integrators for associated ergodic dynamics, where the probability distribution $\mu_\star$ is given below.

Definition 2.2. *Let $V : H \to \mathbb{R}$ be a function of class $\mathcal{C}^2$ with bounded second order derivatives. We define*

$$d\mu_\star(x) = Z^{-1} \exp\left(-2V(x)\right) d\nu(x),$$

where $\nu = \mathcal{N}(0, \frac{1}{2}Q)$, and $Z = \int_H e^{-2V(x)} d\nu(x)$.

Observe that V is globally Lipschitz continuous, thus its growth is at most linear. By the Fernique theorem, Z is finite and thus $\mu_\star$ is well-defined. More generally, it has finite moments of all orders, precisely, for all $\kappa \in \mathbb{N}$,

$$\int |x|^\kappa d\mu_\star(x) < \infty.$$

We then define the nonlinear operators $F, G : H \to H$ as follows:
$$F(x) = -DV(x), \quad G(x) = -x + QF(x).$$

To prove ergodicity and analyze the speed of convergence to equilibrium for the continuous and discrete-time processes studied below, the following assumption gives a convenient sufficient condition. We mention that this assumption could be weakened for the study of a nonglobally Lipschitz nonlinearity F as studied in [14], but this is out of the scope of the present paper.

Assumption 1. *Assume that*
$$\text{Lip}(F) = \sup_{x_1 \neq x_2 \in H} \frac{|F(x_2) - F(x_1)|}{|x_2 - x_1|} < q_1^{-1} = \lambda_1 = \min_{k \in \mathbb{N}} \lambda_k.$$

For example, let $f : \mathbb{R} \to \mathbb{R}$ be a bounded, Lipschitz continuous mapping and $H = L^2(0, 1)$. The Nemytskii operator $F : H \to H$ is defined for all $x \in H$ by
$$(6) \qquad\qquad\qquad F(x)(\cdot) = f(x(\cdot)).$$
Then $F(x) = -DV(x)$ satisfies Assumption 1, where
$$V(x) = \int_0^1 U(x(z))dz, \quad U' = -f.$$

To conclude this section, let us introduce a finite dimensional discretization in space, which shall prove useful for the analysis given that a number of objects are not well-posed directly in the infinite-dimensional Hilbert space H. Let spectral Galerkin projection operators $\left(\pi^K\right)_{K \in \mathbb{N}}$ be defined as follows:
$$(7) \qquad\qquad\qquad \pi^K x = \sum_{k=1}^{K} \langle x, e_k \rangle e_k.$$

Let $H^K = \text{span}\{e_1, \ldots, e_K\}$ denote the range of π^K, with $\dim H^K = K$. We define the probability distribution $\mu_\star^K$ on H^K by
$$(8) \qquad\qquad\qquad d\mu_\star^K(x) = \frac{1}{Z^K} \exp\left(-2V(x)\right) d\nu^K(x),$$

where $\nu^K = \mathcal{N}(0, \frac{1}{2}Q^K)$, with $Q^K = \pi^K Q \pi^K = \sum_{k=1}^{K} \frac{\langle \cdot, e_k \rangle}{\lambda_k} e_k$, and $Z^K = \int_{H^K} e^{-2V(x)} d\nu^K(x)$. Note that ν^K is the push-forward measure of ν by π_K. Observe that, for any smooth test function $\varphi : H \to \mathbb{R}$, one has
$$\int_{H^K} \varphi \circ \pi^K d\mu_\star^K = \frac{\int_{H^K} \varphi \circ \pi^K e^{-V} d\nu^K}{\int_{H^K} e^{-V} d\nu^K} = \frac{\int_H \varphi \circ \pi^K e^{-V \circ \pi^K} d\nu}{\int_H e^{-V \circ \pi^K} d\nu} \underset{K \to \infty}{\to} \int_H \varphi d\mu_\star.$$

In addition, let $V^K(x) = V(x)$ for all $x \in H^K$. Then, one has the identities $-DV^K(x) = -\pi^K DV(\pi^K x) = \pi^K F(\pi^K x) =: F^K(x)$ for all $x \in H^K$. Note also that
$$\sup_{K \in \mathbb{N}} \int_{H^K} |x^K| d\mu_\star^K(x_K) < \infty, \quad \int_H |x| d\mu_\star(x) < \infty.$$

3. Preconditioning of ergodic SPDEs

In this section, we introduce the preconditioned dynamics that will permit to derive high-order methods in spite of the low-regularity of the original dynamics. A crucial ingredient in the analysis is to consider dynamics projected on finite dimensional spaces H^K with finite dimension K arbitrarily large and to derive estimates with uniform constants w.r.t. K. In particular, in the remaining of the paper, C denotes a generic constant independent of K.

Let $\mathcal{P}$ be a bounded, self-adjoint, linear operator that commutes with A. This means that $\mathcal{P}x = \sum_{k \in \mathbb{N}} p_k \langle x, e_k \rangle e_k$, and we assume that $p_k > 0$ for all $k \in \mathbb{N}$, and $\sup_{k \in \mathbb{N}} p_k < \infty$. Let $\left(Y^{\mathcal{P}}(t)\right)_{t \geqslant 0}$ be defined as the solution of

$$(9) \qquad dY^{\mathcal{P}}(t) = \mathcal{P}\Big(AY^{\mathcal{P}}(t) + F(Y^{\mathcal{P}}(t))\Big)dt + \mathcal{P}^{\frac{1}{2}}dW(t).$$

For $K \in \mathbb{N}$, we denote $\left(Y^{\mathcal{P},K}\right)_{t \geqslant 0}$ the solution of spectral Galerkin method of (9),

$$(10) \qquad dY^{\mathcal{P},K}(t) = \mathcal{P}\Big(AY^{\mathcal{P},K}(t) + F^K(Y^{\mathcal{P},K}(t))\Big)dt + dW^{\mathcal{P},K}(t), \quad W^{\mathcal{P},K} = \pi^K W^{\mathcal{P}},$$

where $F^K = \pi^K \circ F \circ \pi^K$ and $Y^{\mathcal{P},K}(0) = \pi^K Y^{\mathcal{P}}(0)$. Observe that the projection operator π^K defined in (7) commutes with A and $\mathcal{P}$. Note that choosing $\mathcal{P} = I$, problem (9) then reduces to the original problem (1) without any preconditioning. Its spectral Galerkin projection is

$$(11) \qquad dX^K(t) = AX^K(t)dt + F^K(X^K(t))dt + dW^K(t), \quad X^K(0) = \pi^K X(0).$$

3.1. **Preconditioning preserves the invariant distribution.** A natural choice of preconditioning is $\mathcal{P} = Q = (-A)^{-1}$ which yields the modified problems

$$(12) \qquad dY(t) = \big(-Y(t) + QF(Y(t))\big)dt + dW^Q(t),$$

$$(13) \qquad dY^K(t) = \big(-Y^K(t) + QF^K(Y^K(t))\big)dt + dW^{Q,K}(t), \quad Y^K(0) = \pi^K Y(0).$$

Compared with (1), we observe that the stochastic evolution equation (12) is not a parabolic stochastic PDE anymore as the unbounded diffusion operator A has been removed via preconditioning. However, contrary to the cylindrical noise in (1), the stochastic evolution equation (12) is driven by a trace-class noise, since $\mathrm{Tr}(Q) < \infty$. Since $G = -I + QF$ and $G^K = -I + QF^K$ are Lipschitz continuous under Assumption 1, it is straightforward to check that, for any initial condition $y_0 \in H$, both equations (12)-(13) admit a unique solution $\left(Y(t)\right)_{t \geqslant 0}$, defined at all positive times, with $Y(0) = y_0$.

Proposition 3.1. *Assume Assumption 1 and let Y, Y^K be the solutions of the preconditioned stochastic evolution equations (12)-(13). Then, $\mu_\star$ (respectively $\mu_\star^K$) is the unique invariant distribution for (1) and (12) (respectively for (11) and (13)). Moreover, exponential mixing holds: there exists $C > 0$ such that for all $T \geqslant 0$ and all globally Lipchitz $\varphi : H \to \mathbb{R}$, one has*

$$\left| \mathbb{E}[\varphi(Y(T))] - \int \varphi \, d\mu_\star \right| \leqslant C\mathrm{Lip}(\varphi)e^{-(1 - \frac{\mathrm{Lip}(F)}{\lambda_1})T}(1 + \mathbb{E}|Y(0)|),$$

$$\sup_{K \in \mathbb{N}} \left| \mathbb{E}[\varphi(Y^K(T))] - \int \varphi \, d\mu_\star^K \right| \leqslant C\mathrm{Lip}(\varphi)e^{-(1 - \frac{\mathrm{Lip}(F)}{\lambda_1})T}(1 + \mathbb{E}|Y(0)|).$$

Remark 3.2. *The convergence speed estimate of Proposition 3.1 is to be compared with the estimate for the original problem (1), where, under Assumption 1, the solution X satisfies*

$$\left| \mathbb{E}[\varphi(X(T)) - \int \varphi \, d\mu_\star \right| \leqslant C\mathrm{Lip}(\varphi)e^{-(\lambda_1 - \mathrm{Lip}(F))T}(1 + \mathbb{E}|X(0)|).$$

Under Assumption 1, one has the following estimates: there exists $C > 0$ such that

$$(14) \qquad \sup_{t \geqslant 0} \mathbb{E}|Y(t)| \leqslant C(1 + \mathbb{E}|Y(0)|),$$

$$(15) \qquad \sup_{K \in \mathbb{N}, t \geqslant 0} \mathbb{E}|Y^K(t)| \leqslant C(1 + \mathbb{E}|Y(0)|).$$

Proposition 3.1 is then a straightforward corollary of Lemma 3.3 below.

Lemma 3.3. *Let Assumption 1 be satisfied, $\mu_\star$ and $\mu_\star^K$ be given by Definition 2.2 and equation* (8), *and assume that $\lambda_1 \sup_{k \in \mathbb{N}} p_k \leqslant \inf_{k \in \mathbb{N}} (\lambda_k p_k)$. Then, there exist $C, \gamma > 0$ such that, for every globally Lipschitz test function $\varphi : H \to \mathbb{R}$ and for all $t \geqslant 0$,*

$$\left| \mathbb{E}[\varphi(Y^\mathcal{P}(t))] - \int \varphi d\mu_\star \right| \leqslant C\mathrm{Lip}(\varphi) e^{-\gamma t} \big(1 + \mathbb{E}|Y^\mathcal{P}(0)| \big),$$

$$\sup_{K \in \mathbb{N}} \left| \mathbb{E}[\varphi(Y^{\mathcal{P},K}(t))] - \int \varphi d\mu_\star^K \right| \leqslant C\mathrm{Lip}(\varphi) e^{-\gamma t} \big(1 + \mathbb{E}|Y^\mathcal{P}(0)| \big),$$

and $\mu_\star$ (resp. $\mu_\star^K$) is the unique invariant distribution of (9) *(resp.* (10)*).*

Note that the condition in Lemma 3.3 is satisfied for $\mathcal{P} = I$ and $\mathcal{P} = (-A)^{-1}$. In the latter case, $\lambda_1 \sup_{k \in \mathbb{N}} p_k = 1 = \inf_{k \in \mathbb{N}} (\lambda_k p_k)$.

Proof of Lemma 3.3. Let $K \in \mathbb{N}$ and consider the spectral Galerkin method $\big(Y^{\mathcal{P},K}\big)_{t \geqslant 0}$ defined in (10). First, note that $\mu_\star^K$ is the unique invariant distribution of (10) on H^K. Indeed, let $\tilde{Y}^{\mathcal{P},K}(t) = \mathcal{P}^{-\frac{1}{2}} Y^{\mathcal{P},K}(t)$ (note that $\mathcal{P}$ is invertible on H^K), and observe that this auxiliary process solves the SDE (on H^K)

$$d\tilde{Y}^{\mathcal{P},K}(t) = \mathcal{P}A\tilde{Y}^{\mathcal{P},K}(t)dt + \mathcal{P}^{\frac{1}{2}} F^K(\mathcal{P}^{\frac{1}{2}}\tilde{Y}^{\mathcal{P},K}(t))dt + dW^K(t),$$

where $W^K(t) = \pi^K W(t)$ is a standard Wiener process in H^K. Note that the drift coefficient satisfies, for all $y \in H^K$, the identity

$$\mathcal{P}Ay + \mathcal{P}^{\frac{1}{2}} F^K(\mathcal{P}^{\frac{1}{2}}y) = -\frac{1}{2} D\big(|(-A)^{\frac{1}{2}} \mathcal{P}^{\frac{1}{2}}(\cdot)|^2\big)(y) - D\tilde{V}^{\mathcal{P},K}(y)$$

with $\tilde{V}^{\mathcal{P},K}(y) = V(\pi^K \mathcal{P}^{\frac{1}{2}} y) = V^K(\mathcal{P}^{\frac{1}{2}}y)$. It is a standard result that the unique invariant distribution of $\tilde{Y}^{\mathcal{P},K}$ has density (with respect to Lebesgue measure on H^K) which is proportional to

$$\exp\Big(-|(-A)^{\frac{1}{2}} \mathcal{P}^{\frac{1}{2}} y|^2 - 2\tilde{V}^{\mathcal{P},K}(y)\Big).$$

Since $Y^{\mathcal{P},K}(t) = \mathcal{P}^{\frac{1}{2}} \tilde{Y}^{\mathcal{P},K}(t)$, this identity and the expression of $\tilde{V}^{\mathcal{P},K}$ then imply that the probability distribution with density (with respect to Lebesgue measure on H^K) which is proportional to

$$\exp\Big(-|(-A)^{\frac{1}{2}} y|^2 - 2V^K(y)\Big)$$

is the unique invariant distribution of (9). It is straightforward to check that this probability distribution is $\mu_\star^K$.

Second, let $Y_1^{\mathcal{P},K}$ and $Y_2^{\mathcal{P},K}$ be two solutions of (10), driven by the same noise, and starting from two different initial conditions y_1^K and y_2^K. Let $\delta Y^K(t) = Y_2^{\mathcal{P},K}(t) - Y_1^{\mathcal{P},K}(t)$, and $\delta F^K(t) = F^K(Y_2^{\mathcal{P},K}(t)) - F^K(Y_1^{\mathcal{P},K}(t))$. Then $dR(t) = \mathcal{P}A\delta Y^K(t)dt + \mathcal{P}\delta F^K(t)$, and

$$\begin{aligned}
\frac{1}{2} \frac{d|\delta Y^K(t)|^2}{dt} &= \langle \mathcal{P}A\delta Y^K(t), \delta Y^K(t) \rangle + \langle \mathcal{P}\delta F^K(t), \delta Y^K(t) \rangle \\
&\leqslant -\inf_{1 \leqslant k \leqslant K} \lambda_k p_k |\delta Y^K(t)|^2 + \sup_{1 \leqslant k \leqslant K} p_k |\delta F^K(t)||\delta Y^K(t)| \\
&\leqslant \Big(-\inf_{k \in \mathbb{N}} \lambda_k p_k + \sup_{k \in \mathbb{N}} p_k \mathrm{Lip}(F)\Big)|\delta Y^K(t)|^2.
\end{aligned}$$

Applying the Gronwall lemma yields

$$|\delta Y^K(t)| \leqslant e^{-\gamma(\mathcal{P},F)t} |\delta Y^K(0)|,$$

8

with $\gamma(\mathcal{P}, F) = \inf_{k \in \mathbb{N}} \lambda_k p_k - \sup_{k \in \mathbb{N}} p_k \mathrm{Lip}(F) > 0$, for all $t \geqslant 0$, and all initial conditions y_1^K and y_2^K. Since φ is globally Lipschitz continuous, for all $y^K \in H^K$, one has

$$
\left| \mathbb{E}_{y^K}[\varphi(Y^{\mathcal{P},K}(t))] - \int \varphi d\mu_\star^K \right| = \left| \mathbb{E}_{y^K}[\varphi(Y^{\mathcal{P}}(t))] - \mathbb{E}_{\mu^K}[\varphi(Y^{\mathcal{P}}(t))] \right|
$$

$$
\leqslant \mathrm{Lip}(\varphi) \iint_{H^K \times H^K} \mathbb{E}|\delta Y^K(t)| d\mu_\star^K(y_2^K) d\delta_{y^K}(y_K^1)
$$

$$
\leqslant \mathrm{Lip}(\varphi) e^{-\gamma(\mathcal{P},F)t} \int_{H^K} |\cdot - y^K| d\mu_\star^K.
$$

Finally, taking to the limit $K \to \infty$ concludes the proof. $\qquad\square$

In the sequel the preconditioning operator $\mathcal{P}$ is chosen as $\mathcal{P} = (-A)^{-1}$ (unless specified otherwise), and thus the stochastic evolution equation (1) is considered. This choice is motivated by a practical reason: in many situations, (a discrete version of) the operator A is known, and it is not required to compute numerically its inverse $\mathcal{P} = (-A)^{-1}$ and its square root $\mathcal{P}^{\frac{1}{2}}$. Instead, in the numerical integrators presented below, computing the LU and Cholesky decompositions is sufficient. The choice $\mathcal{P} = (-A)^{-1}$ is also motivated by the following theoretical property, stated below: the trajectories of the solution $\big(Y(t)\big)_{t \geqslant 0}$ are almost surely α-Hölder continuous for all $\alpha \in (0, \frac{1}{2})$. This property follows from the application of the Kolmogorov-Centzov regularity criterion, and the estimate

$$
\left(\mathbb{E}|Y(t) - Y(s)|^q \right)^{\frac{1}{q}} \leqslant C|t - s|^{\frac{1}{2}}, \ \forall s, t \in [0, T].
$$

That estimate is established using the trace-class property of the noise ($\mathrm{Tr}(Q) < \infty$). The computations are straightforward and are thus left for the readers.

Remark 3.4. *On the one hand, this regularity is much better than for trajectories of the original SPDE (1), which are only $\frac{\alpha}{2}$-Hölder continuous for such parameters $\alpha \in (0, \frac{1}{2})$. On the other hand, this temporal regularity is the same as for solutions of finite dimensional SDEs. As a consequence, it is natural to expect that the order of convergence with respect to the time-step size is increased, and that one can design higher order methods. Note that the spatial regularity of trajectories is unchanged by the preconditioning procedure. This is consistent with the fact that $\big(X(t)\big)_{t \geqslant 0}$ and $\big(Y(t)\big)_{t \geqslant 0}$ have the same invariant distribution $\mu_\star$. Note also that applying the spectral Galerkin discretization to the preconditioned equation (12) yields the same process as when applying the preconditioning technique after the spectral Galerkin discretization of the original equation (12). As a consequence, applying the preconditioning technique does not modify the error due to spatial discretization.*

3.2. Kolmogorov and Poisson equations for the preconditioned dynamics. Let $M > 0$, $\beta \in [0, 2\alpha_{\sup})$ (defined in (5)), $H^\beta = D((-A)^{\beta/2})$, and $|x|_\beta = |(-A)^{\beta/2} x|$, we say that $F \colon H \to H$ is in $\mathfrak{X}_P^{M,\beta}$ if for all K, $F^K \colon H^{K,\beta} \to H^K$ is M times differentiable and if there exists $C > 0$, $\kappa > 0$ independent of K such that for all $1 < m \leqslant M$,

$$
|D^m F^K(x)(h_1, \ldots, h_m)| \leqslant C(1 + |x|_\beta^\kappa)|h_1|_\beta \ldots |h_m|_\beta, \quad x, h_1, \ldots, h_m \in H^\beta.
$$

Similarly, we say that $\varphi \colon H \to \mathbb{R}$ is in $\mathcal{C}_P^{M,\beta}$ if for all K, $\varphi^K \colon H^{K,\beta} \to \mathbb{R}$ is M times differentiable and if there exists $C > 0$, $\kappa > 0$ independent of K such that for all $1 \leqslant m \leqslant M$,

$$
|D^m \varphi^K(x)(h_1, \ldots, h_m)| \leqslant C(1 + |x|_\beta^\kappa)|h_1|_\beta \ldots |h_m|_\beta, \quad x, h_1, \ldots, h_m \in H^\beta.
$$

For instance for the one-dimensional Laplacian, observe that the Nemytskii operator (6) for $f(x) = -x + \cos(x)$ satisfies $F \in \mathfrak{X}_P^{M,\beta}$ for M arbitrarily large and some $\beta \in (0, 1/4)$.

9

Define the auxiliary functions v and Ψ as follows, where for the sake of brevity, we omit the index K of the spectral Galerkin projection in the notation: for all $t \geqslant 0$ and $y \in H^K$,

$$(16) \qquad v(t,y) = \mathbb{E}_y[\varphi(Y^K(t))],$$

and

$$(17) \qquad \Psi(y) = \int_0^\infty \left(v(t,y) - \int \varphi d\mu_\star^K\right) dt.$$

Note that the function Ψ is well-defined owing to Proposition 3.1.

Introduce the infinitesimal generator $\mathcal{L}$ associated with the preconditioned dynamics (12): if $\varphi : H \to \mathbb{R}$ is twice differentiable,

$$\mathcal{L}\varphi(y) = D\varphi(y).G^K(y) + \sum_{k=1}^K q_k D^2\varphi(y).(e_k, e_k).$$

On the one hand, v is the solution of the Kolmogorov equation

$$(18) \qquad \frac{\partial}{\partial t}v(t,y) = \mathcal{L}v(t,y) = Dv(t,y).G^K(y) + \frac{1}{2}\sum_{k=1}^K q_k D^2 v.(e_k, e_k),$$

with initial condition $v(0, \cdot) = \varphi$, where $Dv(t,y)$ and $D^2 v(t,y)$ denote spatial derivatives of v. On the other hand, Ψ is the solution of the Poisson equation with $x \in H^K$,

$$(19) \qquad -\mathcal{L}\Psi(x) = \varphi(x) - \int \varphi d\mu_\star^K.$$

The first and second order spatial derivatives of v satisfy the estimates given by Proposition 3.5 below.

Proposition 3.5. *Assume Assumption 1 with $\gamma = 1 - \frac{\mathrm{Lip}(F)}{\lambda_1} > 0$. Consider $F \in \mathfrak{X}_P^{M,\beta}$, $v(t,y)$ defined in (16), and $\varphi \in \mathcal{C}_P^{M,\beta}$. Then, there exists $C > 0$, such that for all $t \geqslant 0$ and all $y \in H$, $h_1, \ldots, h_m \in H^\beta$, one has*

$$(20) \qquad |D^m v(t,y)(h_1, \ldots, h_m)| \leqslant C(1 + |y|_\beta^\kappa)e^{-\gamma t}|h_1|_\beta \ldots |h_m|_\beta.$$

Note that the estimates (20) justify that Ψ is of class $\mathcal{C}^2$ and Ψ is solution of the Poisson equation (19).

Proof. For the sake of clarity, we present the proof of (20) for $M = 1, 2, 3$. The estimates are proven analogously for higher M.

Case $M = 1$: One has

$$Dv(t,y).h = \mathbb{E}_y[D\varphi(Y^K(t)).\eta^h(t)]$$

with

$$\frac{d\eta^h(t)}{dt} = -\eta^h(t) + QDF^K(Y^K(t)).\eta^h(t), \quad \eta^h(0) = h.$$

By the assumption on φ:

$$|Dv(t,y).h| \leqslant C\mathbb{E}[|(-A)^\beta \eta^h(t)|].$$

Decompose

$$\tilde{\eta}^h(t) = \eta^h(t) - e^{-t}h,$$

with

$$\frac{1}{2}\frac{d\tilde{\eta}^h(t)}{dt} = -\tilde{\eta}^h(t) + QDF^K(Y^K(t)).\tilde{\eta}^h(t) + QDF^K(Y^K(t)).e^{-t}h.$$

Using an energy estimate in the norm $|(-A)^{\frac{1}{2}}\cdot| = |Q^{-\frac{1}{2}}\cdot|$, one obtains

$$\frac{1}{2}\frac{d|(-A)^{\frac{1}{2}}\eta^h(t)|^2}{dt} = -|(-A)^{\frac{1}{2}}\tilde{\eta}^h(t)|^2 + \langle DF^K(Y^K(t)).\tilde{\eta}^h(t), \tilde{\eta}^h(t)\rangle$$
$$+ \langle DF^K(Y^K(t)).e^{-t}h, \tilde{\eta}^h(t)\rangle$$
$$\leqslant -|(-A)^{\frac{1}{2}}\tilde{\eta}^h(t)|^2 + \mathrm{Lip}(F)|\tilde{\eta}^h(t)|^2 + \epsilon|\tilde{\eta}^h(t)|^2 + \frac{1}{4\epsilon}e^{-2t}|DF^K(Y^K(t)).h|^2$$
$$\leqslant -(\gamma - \epsilon)|(-A)^{\frac{1}{2}}\tilde{\eta}^h(t)|^2 + C_\epsilon e^{-2t}|h|^2,$$

where $\epsilon > 0$ can be chosen arbitrarily small.

Using the Gronwall inequality (observe that $\tilde{\eta}^h(0) = 0$) one obtains

$$|(-A)^{\frac{1}{2}}\tilde{\eta}^h(t)|^2 \leqslant C_\epsilon \int_0^t e^{-2(\gamma-\epsilon)(t-s)}e^{-2s}ds|h|^2 \leqslant C_\epsilon e^{-2(\gamma-\epsilon)t}|h|^2.$$

Using the upper bound $\beta \leqslant \frac{1}{2}$, and recalling that $\eta^h(t) = e^{-t}h + \tilde{\eta}^h(t)$, one obtains

$$|(-A)^\beta\eta^h(t)| \leqslant e^{-t}|(-A)^\beta h| + C|(-A)^{\frac{1}{2}}\tilde{\eta}^h(t)|$$
$$\leqslant e^{-t}|(-A)^\beta h| + Ce^{-(\gamma-\epsilon)t}|h|$$
$$\leqslant Ce^{-(\gamma-\epsilon)t}|(-A)^\beta h|,$$

which is the same as above (up to γ being replaced by $\gamma - \epsilon$). We then obtain

$$|Dv(t,y).h| \leqslant Ce^{-\gamma t}|(-A)^\beta h|.$$

Case $M = 2$: One has

$$D^2v(t,y).(h_1,h_2) = \mathbb{E}_y[D\varphi(Y^K(t)).\zeta^{h_1,h_2}(t)] + \mathbb{E}_y[D^2\varphi(Y^K(t)).(\eta^{h_1}(t), \eta^{h_2}(t))],$$

with $\zeta^{h_1,h_2}(0) = 0$ and

$$\frac{d\zeta^{h_1,h_2}(t)}{dt} = -\zeta^{h_1,h_2}(t) + QDF^K(Y^K(t)).\zeta^{h_1,h_2}(t) + QD^2F^K(Y^K(t)).(\eta^{h_1}(t), \eta^{h_2}(t)).$$

Using the assumption on φ and the upper bound obtained in the previous case $M = 1$:

$$\left|\mathbb{E}[D^2\varphi(Y^K(t)).(\eta^{h_1}(t), \eta^{h_2}(t))]\right| \leqslant Ce^{-2\gamma t}|(-A)^\beta h_1||(-A)^\beta h_2|.$$

To prove upper bounds for $|(-A)^\beta\zeta^{h_1,h_2}(t)|$, the second strategy described in the case $M = 1$ is employed:

$$\frac{1}{2}\frac{d|(-A)^{\frac{1}{2}}\zeta^{h_1,h_2}(t)|^2}{dt} = -|(-A)^{\frac{1}{2}}\zeta^{h_1,h_2}(t)|^2 + \langle DF^K(Y^K(t)).\zeta^{h_1,h_2}(t), \zeta^{h_1,h_2}(t)\rangle$$
$$+ \langle D^2F^K(Y^K(t)).(\eta^{h_1}(t), \eta^{h_2}(t)), \zeta^{h_1,h_2}(t)\rangle$$
$$\leqslant -\gamma|(-A)^{\frac{1}{2}}\zeta^{h_1,h_2}(t)|^2 + |D^2F^K(Y^K(t)).(\eta^{h_1}(t), \eta^{h_2}(t))||\zeta^{h_1,h_2}(t)|$$
$$\leqslant -(\gamma - \epsilon)|(-A)^{\frac{1}{2}}\zeta^{h_1,h_2}(t)|^2 + C_\epsilon|(-A)^\beta\eta^{h_1}(t)||(-A)^\beta\eta^{h_2}(t)|$$
$$\leqslant -(\gamma - \epsilon)|(-A)^{\frac{1}{2}}\zeta^{h_1,h_2}(t)|^2 + C_\epsilon e^{-2\gamma t}|(-A)^\beta h_1|^2|(-A)^\beta h_2|^2.$$

Applying the Gronwall inequality gives for all $t \geqslant 0$

$$|(-A)^{\frac{1}{2}}\zeta^{h_1,h_2}(t)|^2 \leqslant C_\epsilon e^{-2(\gamma-\epsilon)t}|(-A)^\beta h_1|^2|(-A)^\beta h_2|^2.$$

In addition, one has $\beta \leqslant \frac{1}{2}$, therefore one obtains

$$|(-A)^\beta\zeta^{h_1,h_2}(t)| \leqslant C_\epsilon e^{-(\gamma-\epsilon)t}|(-A)^\beta h_1||(-A)^\beta h_2|.$$

Finally one obtains
$$\left|D^2 v(t,y).(h_1,h_2)\right| \leqslant Ce^{-\gamma t}|(-A)^\beta h_1||(-A)^\beta h_2|.$$

Case $M = 3$: One has
$$
\begin{aligned}
D^3 v(t,y).(h_1,h_2,h_3) = {} & \mathbb{E}_y[D\varphi(Y^K(t)).\xi^{h_1,h_2,h_3}(t)] \\
& + \mathbb{E}_y[D^2\varphi(Y^K(t)).(\eta^{h_1}(t),\zeta^{h_2,h_3}(t))] \\
& + \mathbb{E}_y[D^2\varphi(Y^K(t)).(\eta^{h_2}(t),\zeta^{h_3,h_1}(t))] \\
& + \mathbb{E}_y[D^2\varphi(Y^K(t)).(\eta^{h_3}(t),\zeta^{h_1,h_2}(t))] \\
& + \mathbb{E}_y[D^3\varphi(Y^K(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))]
\end{aligned}
$$

with $\xi^{h_1,h_2,h_3}(0) = 0$ and
$$
\begin{aligned}
\frac{d\xi^{h_1,h_2,h_3}(t)}{dt} = {} & -\xi^{h_1,h_2,h_3}(t) + QDF^K(Y^K(t)).\xi^{h_1,h_2,h_3}(t) \\
& + QD^2 F^K(Y^K(t)).(\eta^{h_1}(t),\zeta^{h_2,h_3}(t)) \\
& + QD^2 F^K(Y^K(t)).(\eta^{h_2}(t),\zeta^{h_3,h_1}(t)) \\
& + QD^2 F^K(Y^K(t)).(\eta^{h_3}(t),\zeta^{h_1,h_2}(t)) \\
& + QD^3 F^K(Y^K(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t)).
\end{aligned}
$$

It is only necessary to deal with the term $\mathbb{E}_y[D\varphi(Y^K(t)).\xi^{h_1,h_2,h_3}(t)]$, since all the other terms in the expression of $D^3 v(t,y).(h_1,h_2,h_3)$ are upper bounded by $Ce^{-\gamma t}|(-A)^\beta h_1||(-A)^\beta h_2||(-A)^\beta h_3|$.

Using the same strategy as above, one obtains
$$\frac{1}{2}\frac{d|(-A)^{\frac{1}{2}}\xi^{h_1,h_2,h_3}(t)|^2}{dt} \leqslant -|(-A)^{\frac{1}{2}}\xi^{h_1,h_2,h_3}(t)|^2 + Ce^{-3\gamma t}|(-A)^\beta h_1||(-A)^\beta h_2||(-A)^\beta h_3|,$$

and applying the Gronwall inequality gives for all $t \geqslant 0$
$$|(-A)^{\frac{1}{2}}\xi^{h_1,h_2,h_3}(t)|^2 \leqslant C_\epsilon e^{-2(\gamma-\epsilon)t}|(-A)^\beta h_1|^2|(-A)^\beta h_2|^2|(-A)^\beta h_3|^2.$$

In addition, one has $\beta \leqslant \frac{1}{2}$, therefore one obtains
$$|(-A)^\beta \xi^{h_1,h_2,h_3}(t)| \leqslant C_\epsilon e^{-(\gamma-\epsilon)t}|(-A)^\beta h_1||(-A)^\beta h_2||(-A)^\beta h_3|.$$

Finally, we obtain the estimate
$$\left|D^3 v(t,y).(h_1,h_2,h_3)\right| \leqslant Ce^{-\gamma t}|(-A)^\beta h_1||(-A)^\beta h_2||(-A)^\beta h_3|,$$

which concludes the proof. $\qquad\square$

Corollary 3.6. *Under the assumptions of Proposition 3.5 with $\beta = 0$, for all $t_2 \geqslant t_1 \geqslant 0$ and all $y, h \in H$, one has*
$$\tag{21} |Dv(t_2,y).h - Dv(t_1,y).h| \leqslant C(1+|y|)\sqrt{t_2 - t_1}e^{-\gamma' t_1}|h|.$$

Proof. It remains to prove (21). Let $y, h \in H$ and $0 \leqslant t_1 \leqslant t_2$, then one has
$$
\begin{aligned}
Dv(t_2,y).h - Dv(t_1,y).h = {} & \mathbb{E}_y\big[\big(D\varphi(Y^K(t_2)) - D\varphi(Y^K(t_1))\big).\eta^h(t_2)\big] \\
& + \mathbb{E}_y\big[D\varphi(Y^K(t_1)).\big(\eta^h(t_2) - \eta^h(t_1)\big)\big].
\end{aligned}
$$

12

On the one hand, since by assumption φ is twice differentiable, one has

$$\left|\mathbb{E}_y\big[\big(D\varphi(Y^K(t_2)) - D\varphi(Y^K(t_1))\big).\eta^h(t_2)\big]\right| \leqslant e^{-\gamma t_2}|h|\mathbb{E}_y|Y^K(t_2) - Y^K(t_1)|$$
$$\leqslant \sqrt{t_2 - t_1}(1 + \sup_{t\geqslant 0}\mathbb{E}_y[|Y^K(s)|])e^{-\gamma t_1}|h|.$$

The last inequality follows from the identity $Y^K(t_2) - Y^K(t_1) = \int_{t_1}^{t_2} G^K(Y^K(s))ds + W^Q(t_2) - W^Q(t_1)$, from the equality $\mathbb{E}|W^Q(t_2) - W^Q(t_1)|^2 = \mathrm{Tr}(Q)(t_2 - t_1)$, and the moment bounds (14) combined with the global Lipschitz continuity of F.

On the other hand, using the identity $\eta^h(t) = e^{-t}h + \int_0^t e^{s-t}DF^K(Y^K(s)).\eta^h(s)ds$, the global Lipschitz continuity of F, and the Gronwall estimate $|\eta^h(t)| \leqslant e^{-\gamma t}|h|$, one obtains, for all $0 \leqslant t_1 \leqslant t_2$,

$$\mathbb{E}_y\big[D\varphi(Y^K(t_1)).\big(\eta^h(t_2) - \eta^h(t_1)\big)\big] \leqslant C|\eta^h(t_2) - \eta^h(t_1)|$$
$$\leqslant \big(e^{-t_1} - e^{-t_2}\big)|h|$$
$$+ \int_{t_1}^{t_2} e^{s-t_2}\big|QDF^K(Y^K(s)).\eta^h(s)\big|ds$$
$$+ \int_0^{t_1} \big(e^{s-t_1} - e^{s-t_2}\big)\big|QDF^K(Y^K(s)).\eta^h(s)\big|ds$$
$$\leqslant e^{-t_1}(t_2 - t_1)|h| + C(t_2 - t_1)\sup_{s\geqslant t_1}|\eta^h(s)|$$
$$+ C(t_2 - t_1)\int_0^{t_1} e^{-(t_1-s)}e^{-\gamma s}|h|ds$$
$$\leqslant Ce^{-\gamma' t_1}(t_2 - t_1)|h|.$$

Gathering the estimates then concludes the proof of (21). $\qquad\square$

4. High order one methods for sampling the invariant distribution

In order to discretize the stochastic evolution equation (12), the simplest numerical method is the explicit Euler scheme:

$$(22) \qquad\qquad Y_{n+1} = Y_n + \Delta t G(Y_n) + \Delta W_n^Q,$$

where $\Delta W_n^Q = W^Q(t_{n+1}) - W^Q(t_n)$, $t_n = n\Delta t$. More generally, we also consider the θ-method, with $\theta \in [0,1]$:

$$(23) \qquad\qquad Y_{n+1} = \frac{1 - (1-\theta)\Delta t}{1 + \theta\Delta t}Y_n + \frac{\Delta t}{1 + \theta\Delta t}QF(Y_n) + \frac{1}{1 + \theta\Delta t}\Delta W_n^Q.$$

This includes the semilinear implicit Euler method for $\theta = 1$, the Crank-Nicholson scheme for $\theta = \frac{1}{2}$, and the explicit Euler method (22) for $\theta = 0$.

4.1. Moment bounds and ergodicity.
The ergodicity of the numerical scheme $\big(Y_n\big)_{n\geqslant 0}$ defined by (22), for sufficiently small time step size Δt, is ensured by Lemma 4.1 below.

Lemma 4.1. *Let Assumption 1 be satisfied. Let $\big(Y_n\big)_{n\geqslant 0}$ be given by (22). Then, there exists $C > 0$, such that*

$$(24) \qquad\qquad \sup_{\Delta t\in(0,1]}\sup_{n\geqslant 0}\mathbb{E}|Y_n|^2 \leqslant C(1 + \mathbb{E}|Y_0|^2).$$

13

Moreover, if $\Delta t \in (0,1]$, then $\big(Y_n\big)_{n \geqslant 0}$ is an ergodic discrete-time Markov process, with unique invariant distribution denoted by $\mu^{\Delta t}$. One has the estimate

$$\sup_{\Delta t \in (0,1]} \int |y|^2 d\mu^{\Delta t}(y) \leqslant C.$$

Finally, ifs $\gamma(\Delta t) = 1 - \Delta t(1 - q_1 \mathrm{Lip}(F)) \in (0,1)$, then for any globally Lipschitz test function $\varphi : H \to \mathbb{R}$ and all $n \in \mathbb{N}$, one has

$$\Big| \mathbb{E}[\varphi(Y_n)] - \int \varphi d\mu^{\Delta t} \Big| \leqslant \mathrm{Lip}(\varphi)\gamma(\Delta t)^n(1 + \mathbb{E}[|Y_0|]).$$

Proof of Lemma 4.1. To prove the moment bounds (24), introduce the auxiliary processes given by

$$Z_{n+1} = (1 - \Delta t)Z_n + \Delta W_n^Q \ , \quad Z_0 = 0 \ , \quad R_n = Y_n - Z_n.$$

First, one has

$$\mathbb{E}|Z_{n+1}|^2 = |1 - \Delta t|^2 \mathbb{E}|Z_n|^2 + \Delta t \mathrm{Tr}(Q).$$

Thus

$$\sup_{n \geqslant 0} \mathbb{E}|Z_n|^2 = \Delta t \mathrm{Tr}(Q) \sum_{m=0}^{\infty} |1 - \Delta t|^{2m} \leqslant \mathrm{Tr}(Q).$$

Second, one has

$$R_{n+1} = (1 - \Delta t)R_n + \Delta t Q F(R_n + Z_n),$$

and the global Lipschitz continuity of F yields

$$|R_{n+1}| \leqslant (|1 - \Delta t| + \Delta t q_1 \mathrm{Lip}(F))|R_n| + C\Delta t(1 + |Z_n|).$$

Using $\Delta t \in (0,1]$, one has

$$\gamma(\Delta t) = 1 - \Delta t(1 - q_1 \mathrm{Lip}(F)) < 1.$$

One thus obtains

$$\big(\mathbb{E}|R_n|^2\big)^{\frac{1}{2}} \leqslant \gamma(\Delta t)^n \big(\mathbb{E}|R_0|^2\big)^{\frac{1}{2}} + C\Delta t \sum_{m=0}^{n-1} \gamma(\Delta t)^m \sup_{m \geqslant 0} \big(\mathbb{E}|Z_m|^2\big)^{\frac{1}{2}}$$
$$\leqslant \big(\mathbb{E}|Y_0|^2\big)^{\frac{1}{2}} + C\Delta t \frac{1}{1 - \gamma(\Delta t)}$$
$$\leqslant \big(\mathbb{E}|Y_0|^2\big)^{\frac{1}{2}} + C\frac{1}{1 - q_1 \mathrm{Lip}(F)}.$$

Gathering the estimates concludes the proof of the moment bounds (24).

It remains to establish the ergodicity of the discrete-time Markov process $\big(Y_n\big)_{n \geqslant 0}$. Let us first check the uniqueness of invariant distributions. Let $\big(Y_n^1\big)_{n \in \mathbb{N}_0}$ and $\big(Y_n^2\big)_{n \in \mathbb{N}_0}$ be two solutions of (22), driven by the same noise, and starting from two different initial conditions y^1 and y^2. Let $\delta Y_n = Y_n^2 - Y_n^1$. Then, for all $n \geqslant 0$,

$$|\delta Y_{n+1}| \leqslant |1 - \Delta t||\delta Y_n| + \Delta t \big| Q F(Y_n^2) - Q F(Y_n^1)\big|$$
$$\leqslant \gamma(\Delta t)|\delta Y_n|.$$

Thus for all $n \geqslant 0$,

$$|\delta Y_n| \leqslant \gamma(\Delta t)^n |y^2 - y^1|,$$

with $\gamma(\Delta t) \in (0,1)$. This yields uniqueness of the invariant distribution.

The existence of an invariant distribution is obtained using the Krylov-Bogoliubov criterion. Indeed, let $\alpha \in (0, \alpha_{\sup})$. Then $\mathrm{Tr}\big((-A)^{2\alpha}Q\big) < \infty$, moreover the set $\{x \in H;\ |(-A)^\alpha x| \leqslant M\}$ is compact in H for all $M > 0$. Assume that $Y_0 = 0$. Then, for all $n \in \mathbb{N}$,

$$Y_n = \Delta t \sum_{m=0}^{n-1} (1 - \Delta t)^{n-1-m} QF(Y_m) + \sum_{m=0}^{n-1} (1 - \Delta t)^{n-1-m} \Delta W_m^Q.$$

Since $(-A)^\alpha Q$ is a bounded linear operator ($\alpha \leqslant 1$), using the Lipschitz continuity of F, the moment bounds (24) and the equality $\mathbb{E}|(-A)^\alpha \Delta W_m^Q|^2 = \Delta t \mathrm{Tr}\big((-A)^{2\alpha}Q\big)$, one obtains

$$\big(\mathbb{E}|(-A)^\alpha Y_n|^2\big)^{\frac{1}{2}} \leqslant \Delta t \sum_{m=0}^{n-1} (1 - \Delta t)^{n-1-m} C(1 + \big(\mathbb{E}|Y_m|^2\big)^{\frac{1}{2}})$$

$$+ \Big(\mathrm{Tr}\big((-A)^{2\alpha}Q\big)\Delta t \sum_{m=0}^{n-1} (1 - \Delta t)^{2(n-1-m)}\Big)^{\frac{1}{2}}$$

$$\leqslant C + \mathrm{Tr}\big((-A)^{2\alpha}Q\big)^{\frac{1}{2}}.$$

As a consequence, we deduce $\sup_{n \in \mathbb{N}} \mathbb{E}|(-A)^\alpha Y_n|^2 < \infty$ if $\alpha \in (0, \alpha_{\sup})$. Applying the Krylov-Bogoliubov criterion then concludes the proof of the existence of an invariant distribution $\mu^{\Delta t}$, and Lemma 4.1 is proved. $\square$

Remark 4.2. *Lemma 4.1 remains valid for other integrators, such as the θ-method (23) or the method (30), assuming the timestep Δt small enough. Moreover, if $\theta \geqslant \frac{1}{2}$, then there is no such timestep restriction to obtain moment bounds and ergodicity.*

4.2. **The Gaussian case.** In this section, we detail the Gaussian case for the θ-method (23). We have the following result.

Proposition 4.3. *Assume that $F = 0$ and that $\Delta t < \frac{2}{1-2\theta}$ if $\theta < 1/2$. Then $\big(Y(t)\big)_{t \geqslant 0}$ and $\big(Y_n\big)_{n \geqslant 0}$ are Gaussian processes with invariant distributions ν, $\nu^{\Delta t, \theta}$ satisfying*

$$\big| \int \varphi d\nu^{\Delta t, \theta} - \int \varphi d\nu \big| \leqslant \sup \|D^2\varphi(y)\|_{\mathcal{L}(H)} \frac{1 - \sigma_\theta(\Delta t)^2}{2} \mathrm{Tr}(Q),$$

where

$$\sigma_\theta(\Delta t)^2 = \frac{2}{2 + (2\theta - 1)\Delta t}.$$

In particular, for $\theta = \frac{1}{2}$, the bias for the invariant distribution vanishes.

Proof. It is straightforward to check that the numerical invariant distribution $\mu^{\Delta t, \theta} = \nu^{\Delta t, \theta}$ is a centered Gaussian distribution, with covariance $\frac{\sigma_\theta(\Delta t)^2}{2}Q$. If $\theta = \frac{1}{2}$, then $\sigma_\theta(\Delta t)^2 = 1$, thus ν is the invariant distribution of the discrete-time process $\big(Y_n\big)_{n \geqslant 0}$.

Assume now that $\theta \in [0, 1] \backslash \{\frac{1}{2}\}$, then one has

$$\sigma_\theta(\Delta t)^2 = 1 + \mathrm{O}(\Delta t).$$

If $\varphi : H \to \mathbb{R}$ is of class $\mathcal{C}^2$, with bounded second order derivative, then

$$(25) \qquad \big| \int \varphi d\nu^{\Delta t, \theta} - \int \varphi d\nu \big| \leqslant \frac{\mathrm{Tr}(Q)}{2} \sup_{y \in H} \|D^2\varphi(y)\|_{\mathcal{L}(H)} \big|\sigma(\Delta t)^2 - 1\big| = \mathrm{O}(\Delta t).$$

Indeed, assume without loss of generality that $\sigma_\theta(\Delta t)^2 < 1$. Let $Y^{\Delta t, \theta} \sim \mathcal{N}(0, \frac{\sigma_\theta(\Delta t)^2}{2}Q) = \nu^{\Delta t, \theta}$ and $R^{\Delta t, \theta} \sim \mathcal{N}(0, \frac{1 - \sigma_\theta(\Delta t)^2}{2}Q)$ be two independent Gaussian random variables. Then the sum

$Y^{\Delta t,\theta} + R^{\Delta t,\theta}$ has distribution $\mathcal{N}(0, \frac{1}{2}Q) = \nu$, which yields

$$
\begin{aligned}
\big| \int \varphi d\nu^{\Delta t,\theta} - \int \varphi d\nu \big| &= \big| \mathbb{E}[\varphi(Y^{\Delta t,\theta})] - \mathbb{E}[\varphi(Y^{\Delta t,\theta} + R^{\Delta t,\theta})] \big| \\
&= \big| \mathbb{E}[\varphi(Y^{\Delta t,\theta})] - \mathbb{E}[\varphi(Y^{\Delta t,\theta} + R^{\Delta t,\theta})] - \mathbb{E}[\langle D\varphi(Y^{\Delta t,\theta}), R^{\Delta t,\theta} \rangle] \big| \\
&\leqslant \sup \ \|D^2\varphi(y)\|_{\mathcal{L}(H)} \mathbb{E}[|R^{\Delta t,\theta}|^2] \\
&= \sup \ \|D^2\varphi(y)\|_{\mathcal{L}(H)} \frac{1 - \sigma_\theta(\Delta t)^2}{2} \mathrm{Tr}(Q).
\end{aligned}
$$

The case $\sigma_\theta(\Delta t)^2 > 1$ is treated by a similar argument. $\qquad\square$

Proposition 4.4. *The Gaussian distributions ν and $\nu^{\Delta t,\theta}$ are singular, as soon as $\sigma_\theta(\Delta t)^2 \neq 1$. As a consequence, when $\Delta t \to 0$, the convergence does not hold in total variation distance, more precisely*

$$
(26) \qquad d_{\mathrm{TV}}(\nu^{\Delta t,\theta}, \nu) = \frac{1}{2} \sup_{\varphi \in \mathcal{B}_b(H,\mathbb{R}), \|\varphi\|_\infty \leqslant 1} \big| \int \varphi d\nu^{\Delta t,\theta} - \int \varphi d\nu \big| = 1,
$$

where $\mathcal{B}_b(H,\mathbb{R})$ denotes the space of bounded measurable functions from H to $\mathbb{R}$.

Proof. Assume without loss of generality that $\sigma(\Delta t)^2 > 1$. Let $Y^{\Delta t,\theta} \sim \nu^{\Delta t,\theta}$ and $Y \sim \nu$, then by the Law of Large Numbers, almost surely

$$
\frac{1}{\sqrt{K}} \sum_{k=1}^{K} \Big(\frac{|\langle Y^{\Delta t,\theta}, e_k \rangle|^2}{q_k} - 1 \Big) \underset{K \to \infty}{\sim} \sqrt{K} \frac{\sigma_\theta(\Delta t)^2 - 1}{2} \underset{K \to \infty}{\to} \infty,
$$

whereas, by the Central Limit Theorem, one has the convergence in distribution

$$
\frac{1}{\sqrt{K}} \sum_{k=1}^{K} \Big(\frac{|\langle Y, e_k \rangle|^2}{q_k} - 1 \Big) \underset{K \to \infty}{\Longrightarrow} \mathcal{N}(0, \frac{1}{2}).
$$

Consider the following family of bounded continuous functions:

$$
\varphi_{\epsilon,K} = 2 \min \Big(1, \exp\big(-\frac{\epsilon}{\sqrt{K}} \sum_{k=1}^{K} \big(\frac{|\langle \cdot, e_k \rangle|^2}{q_k} - 1 \big) \big) \Big) - 1,
$$

with arbitrary $\epsilon > 0$ and $K \in \mathbb{N}$. Then letting first $K \to \infty$, then $\epsilon \to 0$, one has

$$
\sup_{\varphi \in \mathcal{C}_b^0(H,\mathbb{R}), \|\varphi\|_\infty \leqslant 1} \big| \int \varphi d\nu^{\Delta t,\theta} - \int \varphi d\nu \big| \geqslant \lim_{\epsilon \to 0} \lim_{K \to \infty} \big| \int \varphi_{\epsilon,K} d\nu^{\Delta t,\theta} - \int \varphi_{\epsilon,K} d\nu \big| = 2,
$$

where $\mathcal{C}_b^0(H,\mathbb{R})$ denotes the space of bounded continuous functions from H to $\mathbb{R}$. $\qquad\square$

4.3. **Error analysis of the bias for the invariant distribution.** The objective of this section is to show that the method (22) has order one of convergence for sampling the invariant distribution. We rely on the following additional assumption, that is naturally satisfied if F is of $(\beta, 2)$-regularity.

Assumption 2. *The mapping $G : H \to H$ is twice Fréchet differentiable, with bounded first and second order derivatives.*

Assumption 2 is a strong assumption that simplifies the convergence analysis, but does not include for instance the case of the Nemytskii operator (6) for $f(x) = -x + \cos(x)$. In Section 5, we shall consider instead alternate assumptions for higher order estimates for the invariant distribution.

Theorem 4.5. *Let Assumptions 1 and 2 be satisfied, and consider the integrator given by (22). Assume that $\varphi : H \to \mathbb{R}$ is twice Fréchet differentiable, such that*

$$\sup_{y \in H} |D\varphi(y)| + \sup_{y \in H} \|D^2\varphi(y)\|_{\mathcal{L}(H)} \leqslant 1.$$

Then, there exists $C > 0$ such that for all $\Delta t \in (0, 1]$, one has

$$(27) \qquad \Big| \int \varphi \, d\mu^{\Delta t} - \int \varphi \, d\mu_\star \Big| \leqslant C\Delta t.$$

Theorem 4.5 is a generalization to the non-Gaussian case of the error estimate (25). The need to consider sufficiently regular test functions φ has been justified by the non-convergence result (26). The main tool of the proof is the auxiliary function v defined by (16), which solves the Kolmogorov equation (18). The regularity estimates stated in Proposition 3.5 play a crucial role.

Proof of Theorem 4.5. Let $\gamma' \in (0, \gamma)$ be fixed. To prove that (27) holds, it suffices to establish the following weak error estimate: there exists $C > 0$ such that for all $\Delta t \in (0, \Delta t_0]$ and all $N \in \mathbb{N}$, one has

$$\big| \mathbb{E}[\varphi(Y_N)] - \mathbb{E}[\varphi(Y(t_N))] \big| \leqslant C\Delta t,$$

with $t_N = N\Delta t$. Passing to the limit $N \to \infty$ then yields (27).

The strategy employed to study the weak error of numerical schemes, both in finite and infinite dimension is as follows. First the weak error is rewritten in terms of the solution v of the Kolmogorov equation. Second a telescoping sum argument yields the following error expansion:

$$E[\varphi(Y_N)] - \mathbb{E}[\varphi(Y(t_N))] = \mathbb{E}[v(0, Y_N)] - \mathbb{E}[v(t_N, Y_0)]$$

$$= \sum_{n=0}^{N-1} \mathbb{E}\big[v(t_N - t_{n+1}, Y_{n+1}) - v(t_N - t_n, Y_n) \big].$$

In order to apply the Itô formula to estimate each term of the sum above, it is convenient to introduce the following continuous-time auxiliary process: for all $n \in \mathbb{N}_0$ and all $t \in [t_n, t_{n+1}]$, set

$$\tilde{Y}(t) = Y_n + (t - t_n)G(Y_n) + \big(W^Q(t) - W^Q(t_n) \big).$$

Observe that $\tilde{Y}(t_n) = Y_n$ for all $n \in \mathbb{N}_0$, and that for $t \in [t_n, t_{n+1}]$, one has

$$d\tilde{Y}(t) = G(Y_n)dt + dW^Q(t).$$

Owing to the moment bounds (24) and using the Lipschitz continuity of F, the following moment bounds for the auxiliary process $\tilde{Y}$ is obtained: if $\Delta t \leqslant \Delta t_0$, one has

$$(28) \qquad \sup_{t \geqslant 0} \mathbb{E}|\tilde{Y}(t)| \leqslant C(1 + \sup_{n \geqslant 0} \mathbb{E}|Y_n|) \leqslant C(1 + \mathbb{E}|Y_0|).$$

For each $n \in \mathbb{N}_0$, introduce the auxiliary infinitesimal generator $\tilde{\mathcal{L}}_n^{\Delta t}$, such that

$$\tilde{\mathcal{L}}_n^{\Delta t}\varphi(y) = \langle G(Y_n), D\varphi(y) \rangle + \frac{1}{2}\text{Tr}\big(QD^2\varphi(y) \big).$$

Applying the Itô formula and the fact that v solves the Kolmogorov equation (18), for all $n \in \{0, \dots, N-1\}$, one has

$$\mathbb{E}\big[v(t_N - t_{n+1}, Y_{n+1}) - v(t_N - t_n, Y_n) \big] = \int_{t_n}^{t_{n+1}} \mathbb{E}\big[\big(\tilde{\mathcal{L}}_n^{\Delta t} - \partial_t \big)v(t_N - t, \tilde{Y}(t)) \big]dt$$

$$= \int_{t_n}^{t_{n+1}} \mathbb{E}\big[\big(\tilde{\mathcal{L}}_n^{\Delta t} - \mathcal{L} \big)v(t_N - t, \tilde{Y}(t)) \big]dt$$

$$= \int_{t_n}^{t_{n+1}} \mathbb{E}\big[\langle \frac{G(Y_n) - G(\tilde{Y}(t))}{1 + \theta\Delta t}, Dv(t_N - t, \tilde{Y}(t)) \rangle \big]dt.$$

17

We introduce the following decomposition of the error:

$$\epsilon_n^1 = \int_{t_n}^{t_{n+1}} \mathbb{E}\big[\langle G(Y_n) - G(\tilde{Y}(t)), Dv(t_N - t_n, Y_n)\rangle\big]dt,$$

$$\epsilon_n^2 = \int_{t_n}^{t_{n+1}} \mathbb{E}\big[\langle G(Y_n) - G(\tilde{Y}(t)), Dv(t_N - t_n, \tilde{Y}(t)) - Dv(t_N - t_n, Y_n)\rangle\big]dt,$$

$$\epsilon_n^3 = \int_{t_n}^{t_{n+1}} \mathbb{E}\big[\langle G(Y_n) - G(\tilde{Y}(t)), Dv(t_N - t, \tilde{Y}(t)) - Dv(t_N - t_n, \tilde{Y}(t))\rangle\big]dt.$$

To deal with the first error term ϵ_n^1, introduce the auxiliary function $\Psi_n = \langle G(\cdot), Dv(t_N - t_n, Y_n)\rangle$. Using a conditioning argument and applying the Itô formula, one obtains

$$\mathbb{E}\big[\langle G(Y_n) - G(\tilde{Y}(t)), Dv(t_N - t_n, Y_n)\rangle \mid Y_n\big] = \mathbb{E}\big[\Psi_n(Y_n) - \Psi_n(\tilde{Y}_n(t)) \mid Y_n\big]$$

$$= \int_{t_n}^{t} \mathbb{E}\big[\tilde{\mathcal{L}}_n^{\Delta t}\Psi_n(\tilde{Y}(s)) \mid Y_n\big]ds.$$

Owing to Assumption 2 and to Proposition 3.5, it is then straightforward to check that the first and second order derivatives of Ψ_n satisfy

$$|D\Psi_n(y).h| \leqslant Ce^{-\gamma(t_N - t_n)}|h| \quad , \quad |D^2\Psi_n(y).(h,k)| \leqslant Ce^{-\gamma(t_N - t_n)}|h||k|.$$

As a consequence, using the moment bounds (28),

$$|\epsilon_n^1| \leqslant C(1 + \mathbb{E}|Y_0|)\Delta t^2 e^{-\gamma(t_N - t_n)}.$$

To deal with the second error term ϵ_n^2, it suffices to use Proposition 3.5 and the Lipschitz continuity of F, to obtain

$$|\epsilon_n^2| \leqslant Ce^{-\gamma'(t_N - t_n)} \int_{t_n}^{t} \mathbb{E}|\tilde{Y}(t) - Y_n|^2 dt \leqslant C(1 + \mathbb{E}|Y_0|^2)\Delta t^2 e^{-\gamma'(t_N - t_n)}.$$

To get the last inequality above, observe that for all $t \in [t_n, t_{n+1}]$,

(29) $$\mathbb{E}|\tilde{Y}(t) - Y_n|^2 = (t - t_n)^2\mathbb{E}|G(Y_n)|^2 + (t - t_n)\text{Tr}(Q) \leqslant C(1 + |y|^2)\Delta t,$$

using the moment bounds (24).

To deal with the third error term ϵ_n^3, it suffices to use the estimates (21) (see Proposition 3.5) and (29): indeed, one obtains

$$|\epsilon_n^3| \leqslant C \int_{t_n}^{t_{n+1}} e^{-\gamma'(t_N - t)}(t - t_n)(1 + \mathbb{E}|Y_0|^2)ds,$$

using the moment bounds (24).

Gathering the estimates for ϵ_n^1, ϵ_n^2 and ϵ_n^3 then yields

$$\left|\mathbb{E}[\varphi(Y_N)] - \mathbb{E}[\varphi(Y(t_N))]\right| \leqslant \sum_{n=0}^{N-1} \left|\mathbb{E}\big[v(t_N - t_{n+1}, Y_{n+1}) - v(t_N - t_n, Y_n)\big]\right|$$

$$\leqslant \sum_{n=0}^{N-1} \big(|\epsilon_n^1| + |\epsilon_n^2| + |\epsilon_n^3|\big)$$

$$\leqslant C(1 + \mathbb{E}|Y_0|^2)\Delta t \int_0^{t_N} e^{-\gamma'(t_N - t)}dt$$

$$\leqslant C(1 + \mathbb{E}|Y_0|^2)\Delta t \int_0^{\infty} e^{-\gamma't}dt.$$

This concludes the proof of Theorem 4.5. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5. High order convergence analysis

The aim of this section is to create, under stronger regularity assumptions, high order methods for sampling $\mu_\star$, given by (3). We consider one step integrators with postprocessors,

$$Y_{n+1}^{\Delta t} = \psi^{\Delta t}(Y_n^{\Delta t}), \quad \overline{Y}_n^{\Delta t} = \overline{\psi}^{\Delta t}(Y_n^{\Delta t}).$$

We consider a vector field $F \in \mathfrak{X}_P^{4,\beta}$ and test functions $\varphi \in \mathcal{C}_P^{6,\beta}$. Recall that the dynamics $\big(Y(t)\big)_{t \geqslant 0}$ to be discretized is given by the preconditioned SPDE (12). This process is ergodic and mixing, and its unique invariant distribution is $\mu_\star$. Thus, we assume a similar property for the integrator. We consider the following assumptions.

Assumption 3. *For some $\Delta t_0 > 0$, if $\Delta t \in (0, \Delta t_0]$, the Markov process $\big(Y_n^{\Delta t}\big)_{n \geqslant 0}$ is ergodic and mixing. Its unique invariant distribution, denoted by $\mu^{\Delta t}$, satisfies*

$$\mathbb{E}[\varphi(Y_n)] \underset{n \to \infty}{\to} \int \varphi d\mu^{\Delta t},$$

for all bounded and continuous test functions $\varphi : H \to \mathbb{R}$, and any arbitrary initial condition $y \in H$.

Assumption 4. *The integrator has bounded moments of all order: for any $\kappa \in \mathbb{N}$, there exists $C > 0$ such that*

$$\sup_{\Delta t \in (0, \Delta t_0]} \sup_{n \geqslant 0} \mathbb{E}[|Y_n^{\Delta t}|_\beta^\kappa] \leqslant C(1 + |Y_0|_\beta^\kappa).$$

Note that Assumption 4 provides moment bounds for the numerical invariant distribution $\mu^{\Delta t}$:

$$\sup_{\Delta t \in (0, \Delta t_0]} \int |y|_\beta^\kappa d\mu^{\Delta t}(y) \leqslant C.$$

In Section 5.1, we perform an analysis for a general scheme relying on Assumptions 3 and 4, while in Section 5.2, we focus in particular on the following method (30) of second order for the invariant distribution. It corresponds to the Leimkuhler-Matthews method [54, 55] applied to the preprocessed dynamics (12) and reformulated as a postprocessed integrator [65],

$$
\begin{aligned}
(30) \qquad Y_{n+1} &= Y_n + \Delta t G(Y_n + \frac{1}{2}\Delta W_n^Q) + \Delta W_n^Q, \\
\overline{Y}_n &= Y_n + \frac{1}{2}\Delta W_n^Q,
\end{aligned}
$$

and on its spectral discretisation

$$
\begin{aligned}
Y_{n+1}^K &= Y_n^K + \Delta t G^K(Y_n^K + \frac{1}{2}\Delta W_n^{Q,K}) + \Delta W_n^{Q,K}, \\
\overline{Y}_n^K &= Y_n^K + \frac{1}{2}\Delta W_n^{Q,K}.
\end{aligned}
$$

Note that, under Assumption 1, the scheme (30) naturally satisfies Assumptions 3 and 4, which is shown by using the same techniques as in the proof of the moment bounds (24) from Lemma 4.1 (see also [62, Lemma 2.2]).

5.1. Convergence analysis for high order methods.
The objective of this section is to exhibit order conditions to have methods of order 2.

First, Lemma 5.1 is an alternative approach compared to Section 4 to check that a method is of order 1.

Lemma 5.1. *Under Assumptions 3 and 4, assume that, for any $\varphi \in \mathcal{C}_P^{4,\beta}$, one has the Taylor-type expansion*

$$\mathbb{E}\big[\varphi\big(\psi^{\Delta t}(y)\big)\big] - \varphi(y) = \Delta t \mathcal{L}\varphi(y) + \Delta t^2 R_1(\varphi, y, \Delta t),$$

for all Δt assumed small enough, with $\displaystyle\sup_{\Delta t \in (0, \Delta t_0]} |R_1(\varphi, y, \Delta t)| \leqslant C(1 + |y|_\beta^\kappa)$ for some $p \in \mathbb{N}$ and $C > 0$. Then, for $\Delta t \to 0$ and $\varphi \in \mathcal{C}_P^{4,\beta}$,

$$(31) \qquad \int \varphi d\mu^{\Delta t} - \int \varphi d\mu_\star = \mathrm{O}(\Delta t).$$

Proof of Lemma 5.1. Since $\mu^{\Delta t}$ is the invariant distribution of the Markov chain $\big(Y_n^{\Delta t}\big)_{n \geqslant 0}$, one has the following equality: for all $\Delta t \in (0, \Delta t_0]$,

$$(32) \qquad \int \Big(\mathbb{E}\big[\varphi\big(\psi^{\Delta t}(y)\big)\big] - \varphi(y)\Big) d\mu^{\Delta t}(y) = 0.$$

Then, one obtains

$$0 = \int \mathcal{L}\varphi(y) d\mu^{\Delta t}(y) + \Delta t \int R_1(\varphi, y, \Delta t) d\mu^{\Delta t}(y).$$

Choosing $\varphi = \Psi$ then yields

$$\int \varphi d\mu^{\Delta t} - \int \varphi d\mu_\star = \Delta t \int R_1(\varphi, y, \Delta t) d\mu^{\Delta t}(y) = \mathrm{O}(\Delta t),$$

owing to the moment bounds for the numerical invariant distribution $\mu^{\Delta t}$. $\qquad\square$

Using the same approach then yields the following result.

Proposition 5.2. *Under Assumptions 3 and 4, assume that, for any $\varphi \in \mathcal{C}_P^{6,\beta}$, one has the expansion*

$$\mathbb{E}\big[\varphi\big(\psi^{\Delta t}(y)\big)\big] - \varphi(y) = \Delta t \mathcal{L}\varphi(y) + \Delta t^2 \mathcal{A}_1 \varphi(y) + \Delta t^3 R_2(\varphi, y, \Delta t),$$

for all Δt assumed small enough, where $\mathcal{A}_1 \varphi \in \mathcal{C}_P^{2,\beta}$ and the following estimate is satisfied for some $p \in \mathbb{N}$ and $C > 0$,

$$\sup_{\Delta t \in (0, \Delta t_0]} \big(|\mathcal{A}_1\varphi(y)| + |R_2(\varphi, y, \Delta t)|\big) \leqslant C(1 + |y|_\beta^\kappa),$$

where $p \in \mathbb{N}$ and $C > 0$. Then, if $\varphi \in \mathcal{C}_P^{6,\beta}$, one has for $\Delta t \to 0$,

$$(33) \qquad \int \varphi d\mu^{\Delta t} - \int \varphi d\mu_\star = \Delta t \int \mathcal{A}_1 \Psi(y) d\mu_\star(y) + \mathrm{O}(\Delta t^2),$$

where Ψ is given by the Poisson equation (19). Moreover, assume further that for any $\varphi \in \mathcal{C}_P^{6,\beta}$, one has the error expansion

$$\mathbb{E}\big[\varphi\big(\overline{\psi}^{\Delta t}(y)\big)\big] - \varphi(y) = \Delta t \overline{\mathcal{A}}_1 \varphi(y) + \Delta t^2 \overline{R}_2(\varphi, y, \Delta t),$$

with $\overline{\mathcal{A}}_1 \varphi \in \mathcal{C}_P^{4,\beta}$, $\overline{\mathcal{A}}_1 1 = 0$, and the estimate for some $p \in \mathbb{N}$ and $C > 0$,

$$\sup_{\Delta t \in (0, \Delta t_0]} \big(|\overline{\mathcal{A}}_1\varphi(y)| + |\overline{R}_2(\varphi, y, \Delta t)|\big) \leqslant C(1 + |y|_\beta^\kappa).$$

Then, if $\varphi \in \mathcal{C}_P^{6,\beta}$, one has

$$(34) \qquad \int \varphi d\overline{\mu}^{\Delta t} - \int \varphi d\mu_\star = \Delta t \int \big(\mathcal{A}_1 + [\mathcal{L}, \overline{\mathcal{A}}_1]\big) \Psi(y) d\mu_\star(y) + \mathrm{O}(\Delta t^2),$$

where $[\mathcal{L}, \overline{\mathcal{A}}_1] = \mathcal{L}\overline{\mathcal{A}}_1 - \overline{\mathcal{A}}_1 \mathcal{L}$ denotes the commutator of the operators $\mathcal{L}$ and $\overline{\mathcal{A}}_1$.

Proof of Proposition 5.2. The proof of (33) follows the same steps as the proof of (31): using the identity (32) with $\varphi = \Psi$, one obtains

$$\int \varphi d\mu^{\Delta t} - \int \varphi d\mu_\star = \Delta t \int \mathcal{A}_1 \Psi d\mu^{\Delta t} + \mathrm{O}(\Delta t^2)$$

$$= \Delta t \int \mathcal{A}_1 \Psi d\mu_\star + \mathrm{O}(\Delta t^2)$$

using (31) for the function $\mathcal{A}_1 \Psi \in \mathcal{C}_P^{4,\beta}$ in the last step. To prove (34), observe that

$$\int \varphi d\overline{\mu}^{\Delta t} - \int \varphi d\mu^{\Delta t} = \int \Big(\mathbb{E}[\varphi(\overline{\psi}^{\Delta t}(y))] - \varphi(y) \Big) d\mu^{\Delta t}(y)$$

$$= \Delta t \int \overline{\mathcal{A}}_1 \varphi d\mu^{\Delta t} + \mathrm{O}(\Delta t^2)$$

$$= \Delta t \int \overline{\mathcal{A}}_1 \varphi d\mu_\star + \mathrm{O}(\Delta t^2),$$

using (31) for the function $\overline{\mathcal{A}}_1 \varphi \in \mathcal{C}_P^{4,\beta}$ in the last step. Combined with (33), this yields

$$\int \varphi d\overline{\mu}^{\Delta t} - \int \varphi d\mu_\star = \Delta t \int \big(\mathcal{A}_1 \Psi - \overline{\mathcal{A}}_1 \varphi \big) d\mu_\star + \mathrm{O}(\Delta t^2)$$

$$= \Delta t \int \big(\mathcal{A}_1 - \overline{\mathcal{A}}_1 \mathcal{L} \big) \Psi d\mu_\star + \mathrm{O}(\Delta t^2)$$

$$= \Delta t \int \big(\mathcal{A}_1 + [\mathcal{L}, \overline{\mathcal{A}}_1] \big) \Psi(y) d\mu_\star(y) + \mathrm{O}(\Delta t^2),$$

using $\varphi = -\mathcal{L}\Psi + \int \varphi d\mu_\star$, the assumption that $\overline{\mathcal{A}}_1 1 = 0$, and, in the last step, the property $\int \mathcal{L}\varphi d\mu_\star = 0$. $\square$

As an immediate consequence of Proposition 5.2, one may state algebraic conditions for order 2 as follows.

Corollary 5.3. *Under the Assumptions of Proposition 5.2 and if $\int \mathcal{A}_1 \varphi d\mu_\star = 0$ for all $\varphi \in \mathcal{C}_P^{6,\beta}$, then, for all $\varphi \in \mathcal{C}_P^{6,\beta}$, one has*

$$\int \varphi d\mu^{\Delta t} - \int \varphi d\mu_\star = \mathrm{O}(\Delta t^2).$$

Alternatively, assume that $\int \big(\mathcal{A}_1 + [\mathcal{L}, \overline{\mathcal{A}}_1] \big) \varphi d\mu_\star = 0$ for all $\varphi \in \mathcal{C}_P^{6,\beta}$. Then, for all $\varphi \in \mathcal{C}_P^{6,\beta}$, one has

$$\int \varphi d\overline{\mu}^{\Delta t} - \int \varphi d\mu_\star = \mathrm{O}(\Delta t^2).$$

5.2. Construction of second order methods. The main result of this section is the following convergence theorem of order two for the new method (30).

Theorem 5.4. *Let the integrator with postprocessing be given by* (30). *Under Assumption 1 and assuming $F \in \mathfrak{X}_P^{4,\beta}$, for all $\varphi \in \mathcal{C}_P^{6,\beta}$ and all Δt small enough, the following second order estimate for sampling the invariant measure of* (1) *holds,*

$$\int \varphi d\overline{\mu}^{\Delta t} - \int \varphi d\mu_\star = \mathrm{O}(\Delta t^2).$$

Moreover, in the Gaussian case $F = 0$, the method is exact: $\overline{\mu}^{\Delta t} = \nu$.

The proof of Theorem 5.4 relies on Corollary 5.3 and the lemmas below on the weak Taylor expansion of method and its postprocessor (30), in the spirit of the analysis of [4, Sect. 8] and [65, Sect. 7] in the finite dimensional context. A key ingredient is to perform the calculations for the finite-dimensional dynamics in H^K, and then take the limit $K \to \infty$.

Remark 5.5. *The analysis could be extended for the creation of methods of any high order for the invariant distribution in the spirit of [4] in the finite dimensional context. As the calculations become tedious, we mention that the algebraic formalism of exotic Butcher series [19, 20, 51, 52] could be extended to the present context and used for simplifying the calculations of the Taylor expansions.*

Lemma 5.6. *Under the assumptions of Theorem 5.4 and for all $\varphi \in \mathcal{C}_P^{6,\beta}$, the local weak expansion of the integrator (30) is*

$$\mathbb{E}(\varphi(Y_{n+1}^K)|Y_n^K = x) = \varphi(x) + \Delta t \mathcal{L}\varphi(x) + \Delta t^2 \mathcal{A}_1 \varphi(x) + \Delta t^3 R_2(\varphi, x, \Delta t)$$

with

$$\mathcal{A}_1 \varphi = \frac{1}{2} D^2 \varphi.(G^K, G^K) + \frac{1}{2} \sum_{i=1} q_i D^3 \varphi.(e_i, e_i, G^K) + \frac{1}{8} \sum_{i,j=1} q_i q_j D^4 \varphi.(e_i, e_i, e_j, e_j)$$

$$+ \frac{1}{8} D\varphi.\sum_{i=1} q_i D^2 G^K.(e_i, e_i) + \frac{1}{2} \sum_{i=1} q_i D^2 \varphi.(DG^K.e_i, e_i),$$

where $\left(e_k\right)_{k\in\mathbb{N}}$ is the complete orthonormal system of H considered in Definition 2.1 and the remainder satisfies the bound

$$\mathbb{E}|R_2(\varphi, x, \Delta t)| \leqslant C(1 + |x|_\beta^\kappa).$$

Proof. Following the lines of [65, Sect. 7] (where the q_i coefficients were not considered), we perform a Taylor expansion of $\varphi(Y_{n+1})$ around Y_n up to order 6 and a Taylor expansion of $G(Y_{n+1})$ up to order 4, which yields the desired expansion with the remainder:

$$R_2(\varphi, x, \Delta t) = \mathbb{E} \int_0^1 \frac{(1-\theta)^3}{3!} D\varphi(x).D^4 G^K(x + \frac{\theta}{2}\Delta W_N^Q).(\frac{\frac{1}{2}\Delta W_N^Q}{\sqrt{\Delta t}})^4 d\theta$$

$$+ \mathbb{E} \int_0^1 \frac{(1-\theta)^5}{5!} D^6 \varphi(x + \theta \Delta Y_n^K).(\frac{\Delta Y_n^K}{\sqrt{\Delta t}})^6 d\theta,$$

where

$$\Delta Y_n^K = Y_{n+1}^K - Y_n^K = \Delta t G^K(Y_n^K + \frac{1}{2}\Delta W_n^Q) + \Delta W_n^Q.$$

We first observe using (5) and $G = -x + QF(x)$:

$$\mathbb{E}|\Delta W_n^Q|_\beta^\kappa = \mathbb{E}|(-A)^{\beta/2}Q^{1/2}\Delta W_n|^\kappa \leqslant C\Delta t^{\kappa/2}, \quad |G(x)|_\beta \leqslant C(1 + |x|_\beta^\kappa).$$

Thus, the increment satisfies

$$\mathbb{E}|\frac{\Delta Y_n^K}{\sqrt{\Delta t}}| \leqslant C(1 + |x|_\beta^\kappa).$$

The bound on the remainder is a consequence of the regularity assumptions $\varphi \in \mathcal{C}_P^{6,\beta}$ and $F \in \mathfrak{X}_P^{4,\beta}$. $\square$

Lemma 5.7. *Under the assumptions of Theorem 5.4 and for all $\varphi \in \mathcal{C}_P^{6,\beta}$, the local weak expansion of the postprocessor integrator (30) is*

$$\mathbb{E}(\varphi(\overline{Y}_n^K)|Y_n^K = x) = \varphi(x) + \Delta t \overline{\mathcal{A}}_1 \varphi(x) + \Delta t^2 \overline{R}_1(\varphi, x, \Delta t),$$

where

$$\overline{\mathcal{A}}_1 \varphi = \frac{1}{8} \sum_{i=1} q_i D^2 \varphi.(e_i, e_i),$$

and the remainder satisfies the bound

$$\mathbb{E}[|\overline{R}_1(\varphi, x, \Delta t)|] \leqslant C(1 + |x|_\beta^\kappa).$$

Proof. The remainder is given by

$$\overline{R}_1(\varphi, x, \Delta t) = \mathbb{E}\int_0^1 \frac{(1-\theta)^3}{3!} D^4\varphi(x + \theta\Delta W_n^Q).(\frac{\Delta W_n^Q}{\sqrt{\Delta t}})^4 d\theta.$$

and the proof concludes analogously to that of Lemma 5.6. $\qquad\square$

Lemma 5.8. *The operators $\mathcal{A}_1$ and $\overline{\mathcal{A}}_1$ from Lemma 5.6 and 5.7 satisfy the second order condition of Proposition 5.2:*

$$\langle\mathcal{A}_1\varphi + [\mathcal{L}, \overline{\mathcal{A}}_1]\varphi\rangle = 0.$$

Proof. Considering the notations $\langle\varphi\rangle_{\mu_\star} = \int_H \varphi(x)d\mu_\star(x)$, $G_j^K = \langle G^K, e_j\rangle$, $\partial_j\varphi = \varphi'e_j$, and using repeatedly the integration by parts formula,

$$\langle\partial_j\varphi_1\,\varphi_2\rangle_{\mu_\star} = \left\langle -\varphi_1\partial_j\varphi_2 - 2\varphi_1\varphi_2 q_j^{-1}G_j^K\right\rangle_{\mu_\star}$$

a calculation yields

$$\left\langle\sum_{i,j=1}q_iq_jD^4\varphi.(e_i, e_i, e_j, e_j)\right\rangle_{\mu_\star} = \left\langle -2\sum_{i=1}q_iD^3\varphi.(e_i, e_i, G^K)\right\rangle_{\mu_\star},$$

$$\left\langle\sum_{i=1}q_iD^3\varphi.(e_i, e_i, G^K)\right\rangle_{\mu_\star} = \left\langle -\sum_{i=1}q_iD^2\varphi.(DG^K.e_i, e_i) - 2D^2\varphi.(G^K, G^K)\right\rangle_{\mu_\star}.$$

We deduce

$$\langle\mathcal{A}_1\varphi\rangle = \left\langle\frac{1}{8}\sum_{i=1}q_iD\varphi.(D^2G^K.(e_i, e_i)) + \frac{1}{4}\sum_{i=1}q_iD^2\varphi.(DG^K.e_i, e_i)\right\rangle$$

Using the calculation for the commutator operator,

$$[\mathcal{L}, \overline{\mathcal{A}}_1]\varphi = -\frac{1}{8}\sum_{i=1}q_iD\varphi.(D^2G^K.(e_i, e_i)) - \frac{1}{4}\sum_{i=1}q_iD^2\varphi.(DG^K.e_i, e_i),$$

this concludes the proof. $\qquad\square$

Proof of Theorem 5.4. It remains to check the statement that in the Gaussian case $F = 0$, one finds $\overline{\mu}^{\Delta t} = \nu$ for all Δt. First, if $(Y_n)_{n\geqslant 0}$ is given by (30), then

$$Y_{n+1} = (1 - \Delta t)Y_n + (1 - \frac{\Delta t}{2})\Delta W_n^Q.$$

Thus, if $\Delta t < 1$, the numerical invariant distributions are Gaussian: $\mu^{\Delta t} = \mathcal{N}(0, \frac{\sigma(\Delta t)^2}{2}Q)$, with

$$\sigma(\Delta t)^2 = \frac{2\Delta t(1 - \frac{\Delta t}{2})^2}{1 - (1 - \Delta t)^2} = 1 - \frac{\Delta t}{2},$$

and $\overline{\mu}^{\Delta t} = \mathcal{N}(0, \frac{\sigma(\Delta t)^2 + \frac{\Delta t}{2}}{2}Q) = \mathcal{N}(0, \frac{1}{2}Q) = \nu$. $\qquad\square$

Remark 5.9. *We emphasize that alternatively to method (30), one can simply consider weak order 2 integrators applied to the preconditioned dynamics (12), for instance, the Runge-Kutta 2 (RK2) method with two evaluations of G per timestep,*

(35)
$$\hat{Y}_n = Y_n + \Delta tG(Y_n) + \Delta W_n^Q,$$
$$Y_{n+1} = Y_n + \frac{\Delta t}{2}\big(G(Y_n) + G(\hat{Y}_n)\big) + \Delta W_n^Q.$$
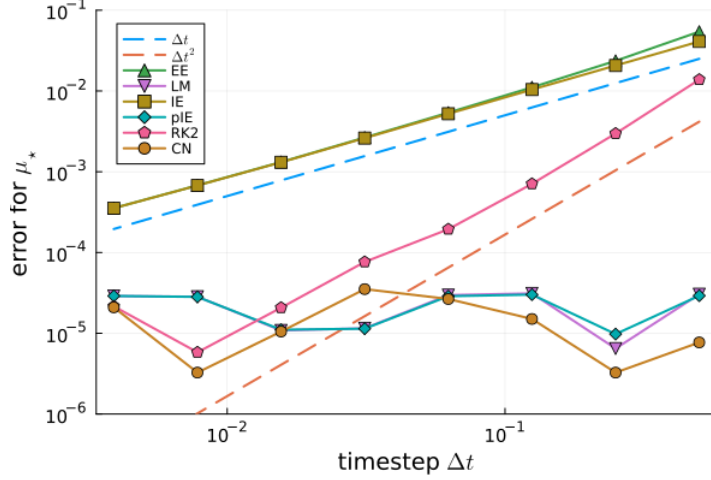
FIGURE 1. Error for sampling the invariant distribution of the preconditioned SPDE (12) in the spatial domain $[0, 1]$ with $A = \Delta$ and $F(x) = 0$ using the explicit and implicit Euler methods EE and IE ((23) with $\theta = 0$ and $\theta = 1$), their postprocessed counterparts the Leimkuhler-Matthews method LM (30) and pIE (36), the RK2 method (35) and the Crank-Nicholson scheme CN (23) with $\theta = \frac{1}{2}$, with the test function $\varphi(x) = \exp(-\|x\|_{L^2}^2)$ and $M = 10^8$ trajectories.

Note, however, that, in contrast to method (30), method (35) is not exact for the invariant distribution in the Gaussian case $F = 0$. Alternatively, one can also consider the postprocessed integrator from [17] based on the semilinear implicit Euler method

(36)
$$Y_{n+1} = \frac{1}{1 + \Delta t} Y_n + \frac{\Delta t}{1 + \Delta t} Q F \left(Y_n + \frac{1}{2(1 + \Delta t)} \Delta W_n^Q \right) + \frac{1}{1 + \Delta t} \Delta W_n^Q,$$
$$\overline{Y}_n = Y_n + \frac{1}{2(1 + \frac{\Delta t}{2})^{\frac{1}{2}}} \Delta W_n^Q.$$

An advantage of method (36), with only one evaluation of F per timestep, is that it has no timestep restriction for stability and it is exact for the invariant distribution in the Gaussian case $F = 0$.

6. NUMERICAL EXPERIMENTS

Let us now illustrate numerically the theoretical results. We give error curves for the new methods of the article for the preconditioned SDPE (12) to confirm the theoretical orders of convergence. Then, we observe the effect of different preconditioning on the order of convergence and introduce an alternative preconditioner based on Chebyshev polynomials.

We first consider the postprocessed explicit and implicit Euler methods (30)-(36) for solving the preconditioned SPDE (12) in the spatial domain $[0, 1]$ with $A = \Delta$. We apply a standard finite difference approximation of the Laplacian with the stepsize $\Delta x = 0.02$, the final time $T = 10$ which is sufficient to reach equilibrium, and different timesteps Δt to observe the rates of convergence. We consider $M = 10^8$ trajectories (using 450 CPUs) and plot an estimate of the error for the invariant distribution by using the standard mean estimator:

$$\overline{J}_{\Delta t} = \frac{1}{M} \sum_{m=1}^{M} \varphi(Y_N^{(m)}) \approx \mathbb{E}[\varphi(Y_N)],$$
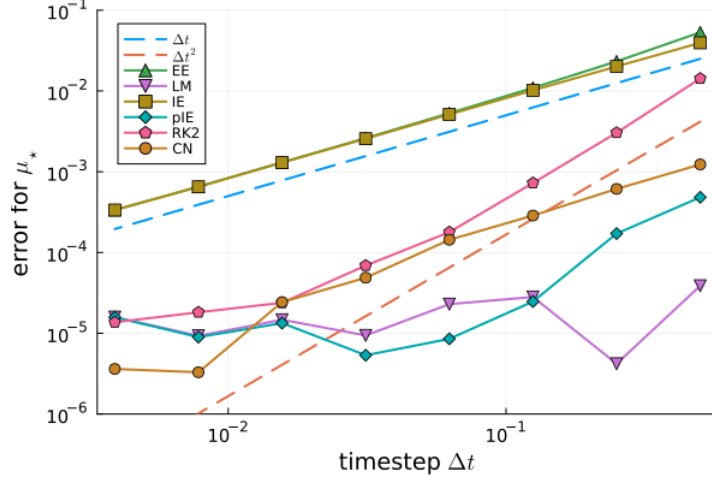
24

FIGURE 2. Error for sampling the invariant distribution of the preconditioned SPDE (12) in the spatial domain $[0,1]$ with $A = \Delta$ and $F(x) = -x + \cos(x)$ using the explicit and implicit Euler methods EE and IE ((23) with $\theta = 0$ and $\theta = 1$), their postprocessed counterparts the Leimkuhler-Matthews method LM (30) and pIE (36), the RK2 method (35) and the Crank-Nicholson scheme CN (23) with $\theta = \frac{1}{2}$, with the test function $\varphi(x) = \exp(-\|x\|_{L^2}^2)$ and $M = 10^8$ trajectories.

where $Y_N^{(m)}$ is the m-th realisation of the integrator at time T and we consider the test function $\varphi(x) = \exp(-\|x\|_{L^2}^2)$. We compare this approximation with a reference value for $\int_H \varphi d\mu_\star$ given by the approximation $\overline{J}_{2^{-9}}$ with the Leimkuhler-Matthews method (30). We emphasize that the approach could be combined with variance reduction techniques, such as Multi-Level Monte-Carlo methods, though this is out of the scope of the present paper.

We observe in Figure 1 the Gaussian case where $F = 0$. The integrators EE, IE, and RK2 have the expected order of convergence for the invariant distribution. Moreover, the other integrators have no bias for the invariant distribution and we observe only the Monte-Carlo error. We repeat the same experiment, this time with the Lipschitz nonlinearity $F(x) = -x + \cos(x)$. We observe the predicted order for the invariant distribution of the methods in Figure 2, which illustrates Theorem 5.4. It is remarkable that the Leimkuhler-Matthews scheme with preprocessing (30) yields the best accuracy among the tested methods.

Let us now observe the behaviour of a given scheme for approximating the preconditioned dynamics (9) with $\mathcal{P} = (-A)^{-\alpha}$ for different $\alpha \in [0,1]$ and the nonlinearity $F(x) = -x + \cos(x)$. We apply the following time-discretisation with the same finite difference approximation as previously:

$$(37) \qquad Y_{n+1} = (I - \Delta t \mathcal{P} A)^{-1}(Y_n + \Delta t \mathcal{P} F(Y_n) + \Delta W_n^{\mathcal{P}}).$$

The reference values are given by the integrator (37) with $\Delta_t = 2^{-8}$ and $\alpha = 1$. We observe on Figure 3 that the preconditioner improves the order of convergence from $1/2$ to 1 as α goes from 0 to 1, to emphasise the advantage of the preconditioner enhancing the convergence rate, here illustrated with the method (37).
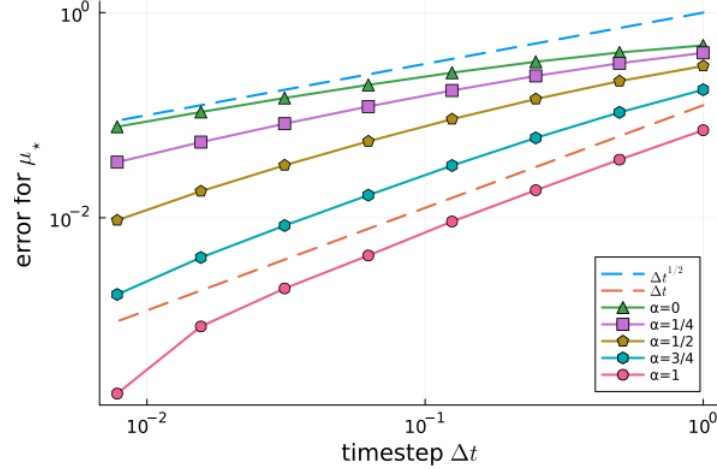
FIGURE 3. Error of the integrator (37) for sampling the invariant distribution of the preconditioned SPDE (9) in the spatial domain $[0,1]$ with $A = \Delta$, $F(x) = -x + \cos(x)$, $\mathcal{P} = (-A)^{-\alpha}$ for different α, the test function $\varphi(x) = \exp(-\|x\|_{L^2}^2)$ and $M = 10^7$ trajectories.

REFERENCES

[1] A. Abdulle, I. Almuslimani, and G. Vilmart. Optimal explicit stabilized integrator of weak order 1 for stiff and ergodic stochastic differential equations. *SIAM/ASA J. Uncertain. Quantif.*, 6(2):937–964, 2018.

[2] A. Abdulle, D. Cohen, G. Vilmart, and K. C. Zygalakis. High weak order methods for stochastic differential equations based on modified equations. *SIAM J. Sci. Comput.*, 34(3):A1800–A1823, 2012.

[3] A. Abdulle, G. A. Pavliotis, and G. Vilmart. Accelerated convergence to equilibrium and reduced asymptotic variance for Langevin dynamics using Stratonovich perturbations. *C. R. Math. Acad. Sci. Paris*, 357(4):349–354, 2019.

[4] A. Abdulle, G. Vilmart, and K. C. Zygalakis. High order numerical approximation of the invariant measure of ergodic SDEs. *SIAM J. Numer. Anal.*, 52(4):1600–1622, 2014.

[5] A. Andersson, R. Kruse, and S. Larsson. Duality in refined Sobolev-Malliavin spaces and weak approximation of SPDE. *Stoch. Partial Differ. Equ. Anal. Comput.*, 4(1):113–149, 2016.

[6] R. B. Best and G. Hummer. Coordinate-dependent diffusion in protein folding. *Proc. Natl. Acad. Sci. U.S.A.*, 107(3):1088–1093, 2010.

[7] S. Boyaval, S. Martel, and J. Reygner. Finite-volume approximation of the invariant measure of a viscous stochastic scalar conservation law. *IMA J. Numer. Anal.*, 42(3):2710–2770, 2022.

[8] C.-E. Bréhier. Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise. *Potential Anal.*, 40(1):1–40, 2014.

[9] C.-E. Bréhier. Influence of the regularity of the test functions for weak convergence in numerical discretization of SPDEs. *J. Complexity*, 56:101424, 15, 2020.

[10] C.-E. Bréhier. Approximation of the invariant distribution for a class of ergodic SPDEs using an explicit tamed exponential Euler scheme. *ESAIM Math. Model. Numer. Anal.*, 56(1):151–175, 2022.

[11] C.-E. Bréhier. Analysis of a modified regularity-preserving Euler scheme for parabolic semilinear SPDEs: total variation error bounds for the numerical approximation of the invariant distribution. *Found. Comput. Math.*, 25(2):511–586, 2025.

[12] C.-E. Bréhier and A. Debussche. Kolmogorov equations and weak order analysis for SPDEs with nonlinear diffusion coefficient. *J. Math. Pures Appl. (9)*, 119:193–254, 2018.

[13] C.-E. Bréhier and L. Goudenège. Weak convergence rates of splitting schemes for the stochastic Allen-Cahn equation. *BIT*, 60(3):543–582, 2020.

[14] C.-E. Bréhier and L. Goudenège. High order sampling of the invariant distribution of SPDEs with non-globally Lipschitz coefficients. *In preparation*, 2025.

[15] C.-E. Bréhier, M. Hairer, and A. M. Stuart. Weak error estimates for trajectories of SPDEs under spectral Galerkin discretization. *J. Comput. Math.*, 36(2):159–182, 2018.

[16] C.-E. Bréhier and M. Kopec. Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme. *IMA J. Numer. Anal.*, 37(3):1375–1410, 2017.

[17] C.-E. Bréhier and G. Vilmart. High Order Integrator for Sampling the Invariant Distribution of a Class of Parabolic Stochastic PDEs with Additive Space-Time Noise. *SIAM J. Sci. Comput.*, 38(4):A2283–A2306, 2016.

[18] E. Bronasco. *Algebraic structures and numerical methods for invariant measure sampling of Langevin dynamics.* PhD thesis, 2025. PhD thesis, Univ. Geneva.

[19] E. Bronasco. Exotic B-series and S-series: algebraic structures and order conditions for invariant measure sampling. *Found. Comput. Math.*, 25(1):271–301, 2025.

[20] E. Bronasco and A. Laurent. Hopf algebra structures for the backward error analysis of ergodic stochastic differential equations. *Submitted*, 2024. arXiv:2407.07451, (47 pages).

[21] E. Bronasco, B. Leimkuhler, D. Phillips, and G. Vilmart. Efficient Langevin sampling with position-dependent diffusion. *Submitted*, 2025. arXiv:2501.02943, (31 pages).

[22] K. Burrage and P. M. Burrage. Order conditions of stochastic Runge-Kutta methods by B-series. *SIAM J. Numer. Anal.*, 38(5):1626–1646, 2000.

[23] M. Cai, C. Chen, J. Hong, and T. Zhou. Long-time dynamics of stochastic wave equation with dissipative damping and its full discretization: Exponential ergodicity and strong law of large numbers. *arXiv preprint arXiv:2402.01137*, 2024.

[24] C. Chen, T. Dang, J. Hong, and T. Zhou. CLT for approximating ergodic limit of SPDEs via a full discretization. *Stochastic Process. Appl.*, 157:1–41, 2023.

[25] C. Chen, J. Hong, L. Ji, and G. Liang. Invariant measures of stochastic Maxwell equations and ergodic numerical approximations. *J. Differential Equations*, 416:1899–1959, 2025.

[26] C. Chen, J. Hong, and G. Liang. Long-time weak convergence analysis of a semi-discrete scheme for stochastic Maxwell equations. *arXiv preprint arXiv:2403.09293*, 2024.

[27] C. Chen, J. Hong, and X. Wang. Approximation of invariant measure for damped stochastic nonlinear Schrödinger equation via an ergodic numerical scheme. *Potential Anal.*, 46(2):323–367, 2017.

[28] Z. Chen, S. Gan, and X. Wang. A full-discrete exponential Euler approximation of the invariant measure for parabolic stochastic partial differential equations. *Appl. Numer. Math.*, 157:135–158, 2020.

[29] D. Conus, A. Jentzen, and R. Kurniawan. Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients. *Ann. Appl. Probab.*, 29(2):653–716, 2019.

[30] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. MCMC methods for functions: modifying old algorithms to make them faster. *Statist. Sci.*, 28(3):424–446, 2013.

[31] J. Cui, J. Hong, and L. Sun. Weak convergence and invariant measure of a full discretization for parabolic SPDEs with non-globally Lipschitz coefficients. *Stochastic Process. Appl.*, 134:55–93, 2021.

[32] T. Cui, X. Tong, and O. Zahm. Optimal Riemannian metric for Poincaré inequalities and how to ideally precondition Langevin dymanics. *Preprint*, 2024. arXiv:2404.02554 (28 pages).

[33] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

[34] A. Debussche. Weak approximation of stochastic partial differential equations: the nonlinear case. *Math. Comp.*, 80(273):89–117, 2011.

[35] A. Debussche and J. Printems. Weak order for the discretization of the stochastic heat equation. *Math. Comp.*, 78(266):845–863, 2009.

[36] N. Deng, Y. Wang, and W. Cao. Ergodicity and invariant measure approximation of the stochastic Cahn-Hilliard equation via an explicit fully discrete scheme. *arXiv preprint arXiv:2512.01621*, 2025.

[37] A. B. Duncan, T. Lelièvre, and G. A. Pavliotis. Variance reduction using nonreversible Langevin samplers. *J. Stat. Phys.*, 163(3):457–491, 2016.

[38] M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.

[39] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numer.*, 24:259–328, 2015.

[40] N. E. Glatt-Holtz and C. F. Mondaini. Long-term accuracy of numerical approximations of SPDEs with the stochastic Navier-Stokes equations as a paradigm. *IMA J. Numer. Anal.*, 45(3):1648–1742, 2025.

[41] M. Hairer, A. Stuart, and J. Voß. Sampling conditioned diffusions. In *Trends in stochastic analysis*, volume 353 of *London Math. Soc. Lecture Note Ser.*, pages 159–185. Cambridge Univ. Press, Cambridge, 2009.

[42] M. Hairer, A. M. Stuart, and S. J. Vollmer. Spectral gaps for a Metropolis-Hastings algorithm in infinite dimensions. *Ann. Appl. Probab.*, 24(6):2455–2490, 2014.

[43] M. Hairer, A. M. Stuart, and J. Voss. Analysis of SPDEs arising in path sampling. II. The nonlinear case. *Ann. Appl. Probab.*, 17(5-6):1657–1706, 2007.

[44] M. Hairer, A. M. Stuart, J. Voss, and P. Wiberg. Analysis of SPDEs arising in path sampling. I. The Gaussian case. *Commun. Math. Sci.*, 3(4):587–603, 2005.

[45] J. Hong and X. Wang. *Invariant measures for stochastic nonlinear Schrödinger equations*, volume 2251 of *Lecture Notes in Mathematics*. Springer, Singapore, 2019. Numerical approximations and symplectic structures.

[46] J. Hong, X. Wang, and L. Zhang. Numerical analysis on ergodic limit of approximations for stochastic NLS equation via multi-symplectic scheme. *SIAM J. Numer. Anal.*, 55(1):305–327, 2017.

[47] G. Hummer. Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J. of Phys.*, 7:34 – 34, 2005.

[48] A. Jentzen and P. E. Kloeden. *Taylor approximations for stochastic partial differential equations*, volume 83 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[49] Y. Jiang and X. Wang. Uniform-in-time weak error estimates of explicit full-discretization schemes for SPDEs with non-globally lipschitz coefficients. *arXiv preprint arXiv:2504.21364*, 2025.

[50] R. Kruse. *Strong and weak approximation of semilinear stochastic evolution equations*, volume 2093 of *Lecture Notes in Mathematics*. Springer, Cham, 2014.

[51] A. Laurent. *Algebraic Tools and Multiscale Methods for the Numerical Integration of Stochastic Evolutionary Problems*. PhD thesis, University of Geneva, 2021.

[52] A. Laurent and G. Vilmart. Exotic aromatic B-series for the study of long time integrators for a class of ergodic SDEs. *Math. Comp.*, 89(321):169–202, 2020.

[53] Z. Lei, C.-E. Bréhier, and S. Gan. Numerical approximation of the invariant distribution for a class of stochastic damped wave equations. *J. Comput. Math.*, 43(4):976–1015, 2025.

[54] B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express. AMRX*, 2013(1):34–56, 2013.

[55] B. Leimkuhler, C. Matthews, and M. V. Tretyakov. On the long-time integration of stochastic gradient systems. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 470(2170):20140120, 16, 2014.

[56] T. Lelièvre, F. Nier, and G. A. Pavliotis. Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2):237–274, 2013.

[57] T. Lelièvre, G. A. Pavliotis, G. Robin, R. Santet, and G. Stoltz. Optimizing the diffusion coefficient of overdamped Langevin dynamics. *to appear in Math. Comp.*, 2026. arXiv:2404.12087 (76 pages).

[58] T. Lelièvre, R. R. Santet, and G. Stoltz. Improving sampling by modifying the effective diffusion. *Preprint*, 2024. arXiv:2410.00525 (33 pages).

[59] T. Lelièvre, R. Santet, and G. Stoltz. Unbiasing Hamiltonian Monte Carlo algorithms for a general Hamiltonian function. *Found. Comput. Math. (published electronically, to appear)*, 2024. arXiv:2303.15918, 63 pages.

[60] Z. Liu. Numerical ergodicity of stochastic Allen-Cahn equation driven by multiplicative white noise. *Commun. Math. Res.*, 41(1):30–44, 2025.

[61] G. J. Lord, C. E. Powell, and T. Shardlow. *An introduction to computational stochastic PDEs*. Cambridge Texts in Applied Mathematics. Cambridge University Press, New York, 2014.

[62] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer-Verlag, Berlin, 2004.

[63] L. Rey-Bellet and K. Spiliopoulos. Improving the convergence of reversible samplers. *J. Stat. Phys.*, 164(3):472–494, 2016.

[64] G. O. Roberts and O. Stramer. Langevin diffusions and Metropolis-Hastings algorithms. volume 4, pages 337–357. Methodol. Comput. Appl. Probab., 2002. International Workshop in Applied Probability (Caracas, 2002).

[65] G. Vilmart. Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.*, 37(1):A201–A220, 2015.

[66] X. Wang, S. Wu, J. Huang, and J. Zheng. Numerical ergodicity of a full-discrete exponential Euler scheme for stochastic Burgers-Huxley equation. *Math. Methods Appl. Sci.*, 48(4):5202–5234, 2025.

[67] Y. Wang and W. Cao. Approximation of the invariant measure for stochastic Allen-Cahn equation via an explicit fully discrete scheme. *arXiv preprint arXiv:2408.00953*, 2024.

[68] Y. Wang, S. Zhang, and W. Cao. Approximation of the invariant measure for stochastic Burgers-Huxley equation via an explicit fully discrete scheme. *Discrete Contin. Dyn. Syst. Ser. S*, 18(11):3532–3559, 2025.

[69] S. Wu, X. Wang, and J. Huang. Numerical ergodicity of two dimensional stochastic Navier-Stokes equations with Gaussian noise. *Comput. Math. Appl.*, 141:15–32, 2023.

Universite de Pau et des Pays de l'Adour, E2S UPPA, CNRS, LMAP, Pau, France
Email address: Charles-Edouard.Brehier@univ-pau.fr

Univ Rennes, INRIA (MINGuS), IRMAR - UMR 6625, ENS Rennes, France
Email address: Adrien.Busnot-Laurent@inria.fr

Univ Rennes, CNRS, IRMAR - UMR 6625, ENS Rennes, France
Email address: Arnaud.Debussche@ens-rennes.fr

Section of Mathematics, University of Geneva, Switzerland
Email address: Gilles.Vilmart@unige.ch