HIGH ORDER INTEGRATOR FOR SAMPLING THE INVARIANT DISTRIBUTION OF A CLASS OF PARABOLIC STOCHASTIC PDES WITH ADDITIVE SPACE-TIME NOISE[*]

CHARLES-EDOUARD BRÉHIER[†] AND GILLES VILMART[‡]

Abstract. We introduce a time-integrator to sample with high order of accuracy the invariant distribution for a class of semilinear stochastic PDEs (SPDEs) driven by an additive space-time noise. Combined with a postprocessor, the new method is a modification with negligible overhead of the standard linearized implicit Euler–Maruyama method. We first provide an analysis of the integrator when applied for SDEs (finite dimension), where we prove that the method has order 2 for the approximation of the invariant distribution, instead of 1. We then perform a stability analysis of the integrator in the semilinear SPDE context, and we prove in a linear case that a higher order of convergence is achieved. Numerical experiments, including the semilinear heat equation driven by space-time white noise, confirm the theoretical findings and illustrate the efficiency of the approach.

Key words. stochastic partial differential equations, postprocessor, invariant measure, ergodicity, space-time white noise

AMS subject classifications. 60H15, 60H35, 37M25

DOI. 10.1137/15M1021088

1. Introduction. We introduce an efficient integrator for the sampling of the invariant probability distribution of a class of semilinear parabolic stochastic PDEs (SPDEs) with additive noise written as an abstract stochastic evolution equation (in the sense of [11])

$$(1.1) \qquad du(t) = (Au(t) + F(u(t))) \, dt + dW^Q(t), \quad u(0) = u_0.$$

Its solution $u(t)$ takes values in a separable infinite dimensional Hilbert space $\mathcal{H}$ with initial condition u_0 (assumed deterministic for simplicity). We assume that $-A$ is a positive unbounded self-adjoint linear operator with an associated sequence of positive eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \ldots$ and an associated complete orthonormal family of eigenvectors $e_1, e_2, \ldots$. The coefficient $F : \mathcal{H} \to \mathcal{H}$ is a Lipschitz continuous nonlinearity, and it is assumed to derive from a continuously differentiable potential function $V : \mathcal{H} \to \mathbb{R}$, i.e., $F = -DV$. Finally, we assume that $\left(W^Q(t)\right)_{t \in \mathbb{R}^+}$ is a Q-Wiener process on $\mathcal{H}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ fulfilling the usual conditions; the covariance operator $Q : \mathcal{H} \to \mathcal{H}$ is a bounded, nonnegative self-adjoint linear operator, such that $Qe_i = q_i e_i$ for some bounded sequence of real numbers $(q_i)_{i \in \mathbb{N}^*}$, where we use the notation $\mathbb{N}^* = \{1, 2, 3, \ldots\}$. We assume the following trace condition:

$$(1.2) \qquad \overline{s} = \sup \left\{ s \in (0,1) \; : \; \mathrm{Trace}\left((-A)^{-1+s}Q\right) < +\infty \right\} > 0.$$

[†]Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France (brehier@math.univ-lyon1.fr).

[‡]Université de Genève, Section de mathématiques, 2-4 rue du Lièvre, CP 64, 1211 Genève 4, Switzerland (Gilles.Vilmart@unige.ch).

Then, there exists a unique mild solution of (1.1) on $\mathbb{R}^+$ (see [11, Chapter 7]), i.e., an $\mathcal{H}$-valued continuous stochastic process $\big(u(t)\big)_{t \in \mathbb{R}^+}$ which satisfies

$$(1.3) \qquad u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))ds + \int_0^t e^{(t-s)A}dW^Q(s).$$

This abstract setting includes the stochastic semilinear heat equation where $A = \Delta = \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2}$ in $\mathcal{H} = L^2(\mathcal{D})$ for some open smooth bounded domain $\mathcal{D} \subset \mathbb{R}^d$,

$$(1.4) \qquad \frac{\partial u}{\partial t}(x,t) = \Delta u(x,t) + f(u(x,t)) + \frac{\partial W^Q}{\partial t}(x,t),$$

with homogeneous Dirichlet boundary conditions on $\partial \mathcal{D}$. In (1.4), the coefficient $f : \mathbb{R} \to \mathbb{R}$ is a smooth, Lipschitz function; the associated *Nemytskii* coefficient $F : u \in \mathcal{H} \mapsto f \circ u \in \mathcal{H}$ satisfies $F(u) = -DV(u)$ for all $u \in \mathcal{H}$, with $V(u) = -\int_0^1 \langle F(\theta u), u \rangle_{\mathcal{H}} d\theta$.

If $d = 1$, one can consider space-time white noise in (1.4), (i.e., with the identity covariance operator $Q = I$; this choice yields $\bar{s} = 1/2$ in (1.2)). When $d > 1$, a nontrivial covariance operator $Q \neq I$ is required, yielding noise which is white in time and colored in space. Notice also that (1.2) is automatically satisfied with $\bar{s} = 1$ for a trace-class noise with $\mathrm{Trace}(Q) < \infty$.

To study the long-time behavior of the process u, we make the additional assumption that F in (1.1) is Lipschitz continuous with constant $\mathrm{L} > 0$ such that $\mathrm{L} < \lambda_1 = \min_{p \in \mathbb{N}^*} \lambda_p$. Then (see, e.g., section 8.6 in [10]) (1.1) admits a unique invariant distribution μ_∞. This means that for all (smooth and Lipschitz) test functions $\phi : \mathcal{H} \to \mathbb{R}$, and for all initial conditions u_0,

$$\lim_{T \to \infty} \text{a. s.} \frac{1}{T} \int_0^T \phi(u(t))dt = \int_{\mathcal{H}} \phi(y)d\mu_\infty(y),$$

where the notation $\lim$ a. s. means that the limit holds with probability 1. Moreover (see, e.g., section 6.3 in [10]) $u(t)$ converges in law to μ_∞ exponentially fast in the following sense: for all $t > 0$, and any test function ϕ,

$$(1.5) \qquad \left| \mathbb{E}\phi(u(t)) - \int_{\mathcal{H}} \phi(y)d\mu_\infty(y) \right| \leq C(\phi, u_0)e^{-(\lambda_1 - L)t},$$

where $C(\phi, u_0)$ is independent of t. It is a standard approach to take advantage of the estimate (1.5) to compute ergodic integrals of the form $\int_{\mathcal{H}} \phi(y)d\mu_\infty(y)$. To do so, in practice one needs to rely on a discretization of the evolution equation. We thus now present two implicit-explicit time-discretization schemes.

Linearized implicit Euler method. We first consider the simplest numerical scheme, which is referred to as the linearized implicit Euler method in this article. Given a constant timestep size $h > 0$, it is defined by $v_{n+1} = v_n + hAv_{n+1} + hF(v_n) + \sqrt{h}\xi_n^Q$, equivalently,

$$(1.6) \qquad v_{n+1} = J_1 \left(v_n + hF(v_n) + \sqrt{h}\xi_n^Q \right),$$

where $v_0 = u(0) = u_0$, $J_1 = (I - hA)^{-1}$, and $\xi_n^Q = h^{-1/2}\big(W^Q((n+1)h) - W^Q(nh)\big)$. For a fixed final time $T > 0$, it is known that the scheme (1.6) applied to (1.1) has weak order of accuracy q for all $q < \bar{s}$, i.e., it satisfies for all h small enough,

$$(1.7) \qquad |\mathbb{E}(\phi(v_n)) - \mathbb{E}(\phi(u(t_n)))| \leq C(u_0, \phi, T)h^q$$

for all $t_n = nh \leq T$, where $C(u_0, \phi, T)$ is independent of n, h.

The weak order of accuracy (1.7) has been analyzed in [14] for (1.6) applied to the heat equation in the case $F = 0$ and in [12] in the semilinear case; see also [39] for the case of colored noise in dimension $d > 1$ and [23] where different techniques are used. Concerning the approximation of the invariant distribution of the heat equation, it is proved in [4] that for $d = 1$ and $Q = I$, one has for all time t_n, similarly to (1.5), the exponential convergence property

$$(1.8) \qquad \left| \mathbb{E}(\phi(v_n)) - \int_{\mathcal{H}} \phi(y) d\mu_\infty(y) \right| \leq K(u_0, \phi)e^{-\lambda t_n} + C(\phi)h^r$$

for all $r < \overline{s}$ with $\overline{s} = 1/2$, where $C(\phi), K(u_0, \phi)$ are independent of n and h, and λ is independent of ϕ, h, n. The proof is based on the analysis of the weak approximation using the tools of [12] (expansion of the error using the backward Kolmogorov equation, Malliavin calculus), with a proof that constants $C(u_0, \phi, T)$ in (1.7) can be chosen independent of the final time T, in the spirit of [35].

New method. The contribution of this paper is the introduction of a modification, denoted u_n, of the standard Euler scheme (1.6), together with a postprocessor, denoted $\overline{u}_n$, which permits us to achieve the estimate (1.8) with the higher order $r < \overline{s} + 1$ instead of $r < \overline{s}$, when in (1.8) v_n is replaced with $\overline{u}_n$. Precisely, the postprocessed method is given by two sequences $(u_n)_{n \geq 0}$ and $(\overline{u}_n)_{n \geq 0}$ in $\mathcal{H}$ defined by

$$(1.9) \quad u_{n+1} = J_1\left(u_n + hF\left(u_n + \frac{1}{2}\sqrt{h}J_2\xi_n^Q\right) + \frac{\sqrt{2}-1}{2}\sqrt{h}J_2\xi_n^Q\right) + \frac{3-\sqrt{2}}{2}\sqrt{h}J_2\xi_n^Q,$$

$$(1.10) \quad \overline{u}_n = u_n + \frac{1}{2}\sqrt{h}J_3\xi_n^Q,$$

where we introduce additional operators J_2, J_3 as follows:

$$J_2 = \left(I - \frac{3-\sqrt{2}}{2}hA\right)^{-1}, \quad J_3QJ_3^T = \left(I - \frac{h}{2}A\right)^{-1}Q.$$

Note that J_3 is not determined uniquely by the equation above: one may define $J_3 = (I - \frac{h}{2}A)^{-1/2}$ or use a Cholesky decomposition. We emphasize the importance of the postprocessing operation (1.10): indeed, the order of convergence to the invariant distribution μ_∞ is not improved if in (1.8) v_n is replaced with u_n, instead of $\overline{u}_n$. Notice that the new scheme (1.9)–(1.10) has a negligible overcost compared to (1.6). First, the postprocessor (1.10) need not be computed at each timestep, but it can be computed only once at the end of each numerical trajectory. Second, computing $J_2\xi_n^Q$ at each timestep is the only additional calculation. Third, we prove that the constant $\lambda > 0$ in (1.8) can be chosen with the same size for both methods: thus in terms of cost (number of timesteps required to achieve a given accuracy on the left-hand side of (1.8)), our new integrator performs better than the standard scheme (1.6).

The fact that the long-time weak order of accuracy r in (1.8) can be made strictly larger than the short-time accuracy q in (1.7) is not surprising: this is known for SDEs (i.e., in finite dimension in the terminology of this paper); see [26, 24, 1, 38] in the context of Brownian dynamics and [3, 2] in the context of Langevin dynamics, where integrators with low weak order, typically $q = 1$ in (1.7), are shown to achieve a high order $r > q$ over long times (1.8) for sampling the invariant measure. Inspired by these recent advances, the popular technique of processing for deterministic differential equations [8] was recently extended to the stochastic context in [38] and serves as a

crucial ingredient to derive the new method proposed in this paper. In our context, this idea is to enhance the accuracy of the modified numerical method u_n in (1.9) by applying a suitable change of variables $u_n \mapsto \overline{u}_n$ defined in (1.10).

Alternatively, note that high order integrators in the strong sense (approximation of the trajectories instead of the distribution) for parabolic problems of the form (1.1) are proposed in [22, 20, 21]; however, these schemes belong to the class of exponential integrators, while the method proposed in this paper avoids the computation of matrix exponentials. We also mention another natural integrator for (1.1), which is the (stochastic) trapezoidal method (also known as the Crank–Nicolson method),

$$(1.11) \qquad u_{n+1} = u_n + \frac{h}{2} A(u_n + u_{n+1}) + h F(u_n) + \sqrt{h} \xi_n^Q.$$

However, in contrast to (1.6), the scheme (1.11) is not L-stable, a desirable property for severely stiff problems, already in the deterministic literature [15]. This makes exponential convergence estimates of the form (1.8) not true in general for the scheme (1.11).

Finally, since the space $\mathcal{H}$ is infinite dimensional, a space discretization scheme is required in practice, e.g., finite differences or finite elements—see section 4. In this article, we only focus on the time-discretization issue. We mention [27], where a postprocessing technique is applied to improve the spatial discretization in the strong sense.

Outline and main results. This paper is organized as follows. In section 2, we explain the derivation of a generalized nonlinear version of the integrator (1.9)–(1.10) in the context of finite dimensional SDEs, where many standard analysis tools are available compared to the SPDE context. We prove that the new integrator has order 2 for the approximation of the invariant distribution of nonlinear ergodic SDEs, instead of order 1 for the standard Euler scheme. Section 3 is devoted to the analysis of the integrator (1.9)–(1.10) for the SPDEs (1.1): we detail the abstract Hilbert space setting (section 3.1), we show the stability and ergodicity properties of the integrator (section 3.2), we prove in a simplified linear case the improved order of accuracy $\overline{s}+1$ of the new method for SPDEs, in contast to the order $\overline{s}$ for the standard Euler scheme (section 3.3), and we show that the proposed scheme exhibits the correct spatial regularity of the invariant distribution in terms of Sobolev-like spaces (section 3.4). Finally, section 4 is dedicated to numerical experiments which confirm the theoretical findings and illustrate the efficiency of the new method.

2. New high order integrator: derivation and analysis in finite dimension. In the context of SDEs in finite dimension $N \in \mathbb{N}^*$, we consider the more general case of a nonlinear system of the form

$$(2.1) \qquad dX(t) = \big(f_1(X(t)) + f_2(X(t))\big) dt + \sigma dW^Q(t), \qquad X(0) = X_0,$$

with solution $X(t)$ in $\mathbb{R}^N$. The nonlinearities $f_1, f_2 : \mathbb{R}^N \to \mathbb{R}^N$ are smooth and Lipschitz functions such that $f_1(x) + f_2(x) = f_0(x) = -\nabla V_0(x)$ for some potential function $V_0 : \mathbb{R}^N \to \mathbb{R}$, where f_1 is a term to be treated implicitly and f_2 is a term to be treated explicitly. The initial condition X_0 is assumed deterministic for simplicity. We also define the Q-Wiener process $W^Q(t) = Q^{1/2}W(t)$, where Q now denotes an $N \times N$ symmetric positive definite matrix and $W(t)$ is a standard N-dimensional Wiener process, and $\sigma > 0$ is a fixed constant.

We introduce the following new implicit-explicit scheme for sampling with high order 2 the invariant measure of (2.1),

$$
X_{n+1} = X_n + h f_1 \left(X_{n+1} + \frac{-2 + \sqrt{5}}{2} J_{n,2} \sigma \sqrt{h} \xi_n^Q \right) + h f_2 \left(X_n + \frac{1}{2} J_{n,2} \sigma \sqrt{h} \xi_n^Q \right)
$$
$$
+ \left(\frac{1 - \sqrt{2} + \sqrt{5}}{2} J_{n,1}^{-1} + \frac{1 + \sqrt{2} - \sqrt{5}}{2} \right) J_{n,2} \sigma \sqrt{h} \xi_n^Q,
$$

(2.2) $\overline{X}_n = X_n + \frac{1}{2} J_{n,3} \sigma \sqrt{h} \xi_n^Q,$

where $J_{n,1}, J_{n,2}, J_{n,3}$ are given by

$$
J_{n,1} = (I - h f_1'(X_n))^{-1}, \quad J_{n,2} = \left(I - \frac{3 - \sqrt{2}}{2} h f_1'(X_n) \right)^{-1}, \quad J_{n,3} Q J_{n,3}^T = \left(I - \frac{h}{2} f_1'(X_n) \right)^{-1}.
$$

We emphasize that the matrix inverses $J_{n,1}, J_{n,2}, J_{n,3}$ are used only in the notation to define the scheme but should not be computed in practice. Indeed, in practical implementations, an LU decomposition should be used in place of computing matrix inverses. Moreover, in the semilinear case (2.3), this decomposition needs to be done only once and may be used for all further iterations.

Remark 2.1. After a spatial discretization with finite differences or finite elements of the SPDE (1.1), in general one arrives at a system of stiff SDEs in $\mathbb{R}^N$ with large dimension N of the form (2.1),

(2.3) $$dX(t) = AX(t)dt + f(X(t))dt + \sigma dW^Q(t).$$

where assume that $f(x) = -\nabla V(x)$ for some $V : \mathbb{R}^N \to \mathbb{R}$. It corresponds to the special case $f_1(x) = Ax$, where $-A$ now denotes an $N \times N$ symmetric positive definite matrix. In this case, the above scheme (2.2) simplifies to the integrator (1.9). Moreover, in the special case $f_1 = 0$, the method (2.2) reduces to

(2.4) $$X_{n+1} = X_n + h f \left(X_n + \frac{1}{2} \sigma \sqrt{h} \xi_n \right) + \sigma \sqrt{h} \xi_n, \qquad \overline{X}_n = X_n + \frac{1}{2} \sigma \sqrt{h} \xi_n,$$

a method which was first proposed in [24] and analyzed in [25], although it was constructed in another manner using a non-Markovian formulation.

For simplicity, we assume for the rest of this section that $Q = I$ is the identity matrix, without losing generality, applying the appropriate change of variable in (fixed) finite dimension. Section 2.1 is devoted to a stability analysis in the case of an Ornstein–Uhlenbeck process: we prove L-stability and exactness results, which play an important role in the efficiency of the integrator for semilinear SPDEs. Section 2.2 contains details on the construction of the method in order to satisfy the above properties and to achieve order 2 of accuracy for sampling the invariant measure.

2.1. Stability analysis. For the study of the stability of stochastic integrators applied to stiff SDEs with additive noise of the form (2.1), a widely used test problem is the following scalar SDE problem (Ornstein–Uhlenbeck process):

(2.5) $$dX = -\lambda X dt + \sigma dW(t),$$

where $\lambda, \sigma > 0$ are fixed constants. Notice that this test equation provides only a useful insight but no rigorous general conclusion on the numerical long-time behavior for nonlinear problems with additive noise; see [6] and references therein. A rigorous analysis of the proposed scheme in our semilinear SPDE context is presented in section 3.2. Notice also that other test equations are used in the literature in the case of multiplicative noise; see [33, 19, 34, 7, 37, 32, 5]. Consider a one-step method of the form

$$(2.6) \qquad X_{n+1} = \mathcal{A}(z)X_n + \mathcal{B}(z)\sqrt{h}\sigma\xi_n, \qquad z = -\lambda h,$$

where h is the stepsize, $\mathcal{A}(z), \mathcal{B}(z)$ are analytic functions, and $\xi_n \sim \mathcal{N}(0,1)$ are independent Gaussian random variables. The SDE (2.5) is ergodic with the unique invariant measure a Gaussian with mean zero and variance $\frac{\sigma^2}{2\lambda}$, i.e., with density $\rho_\infty(x) = \sqrt{\frac{\lambda}{\pi\sigma^2}}\exp(-\frac{\lambda}{\sigma^2}x^2)$. Indeed, for any initial condition $X_0 = x$, the exact solution is a Gaussian random variable, with $\lim_{t\to\infty}\mathbb{E}(X(t)) = 0$ and $\lim_{t\to\infty}\mathbb{E}(|X(t)|^2) = \frac{\sigma^2}{2\lambda}$. The second-order moment $\mathbb{E}(|X_n|^2)$ remains bounded as $n \to \infty$ if $|\mathcal{A}(z)| < 1$, which corresponds to the mean-square stability condition. The condition $|\mathcal{A}(z)| \leq 1$ for all z with negative real part is called A-stability in the deterministic literature [15]. This is a desirable property of numerical integrators for stiff problems, because it permits us to avoid a severe timestep size restriction. For $|\mathcal{A}(z)| < 1$ (stability condition), we obtain

$$(2.7) \qquad \lim_{n\to\infty}\mathbb{E}(|X_n|^2) = \frac{\sigma^2}{2\lambda}\mathcal{R}(z), \quad \text{where } \mathcal{R}(z) = \frac{-2z\mathcal{B}(z)^2}{1 - \mathcal{A}(z)^2}.$$

We see that the method is exact (for the approximation of the invariant distribution) if and only if $\mathcal{R}(z) = 1$ for all z; this is the case for instance, for the trapezoidal method (1.11), which is such that $\mathcal{A}(z) = \frac{1+z/2}{1-z/2}$, $\mathcal{B}(z) = \frac{1}{1-z/2}$. However, in addition to A-stability, a desirable property of Runge–Kutta methods for very stiff problems is L-stability [15], namely, $\mathcal{A}(\infty) = 0$; this is not satisfied by the trapezoidal method, for which $\mathcal{A}(\infty) = -1$. The following proposition states that for Runge–Kutta type methods, where $\mathcal{A}(z), \mathcal{B}(z)$ are rational functions, L-stability (i.e., $\mathcal{A}(\infty) = 0$) is incompatible with the exactness for the invariant distribution, i.e., $\mathcal{R}(z) \equiv 1$.

PROPOSITION 2.1. *Consider a method of the form* (2.6) *where* $\mathcal{A}(z), \mathcal{B}(z)$ *are rational functions. If the method samples exactly the invariant distribution of* (2.5) *(i.e.,* $\mathcal{R}(z) \equiv 1$*), then* $|\mathcal{A}(\infty)| = 1$.

Proof. If $\mathcal{A}(z)$ is a rational function with $|\mathcal{A}(\infty)| \neq 1$, there exist a constant $C \in \mathbb{R}$ and an odd integer $k \in \mathbb{Z}$ such that $\mathcal{R}(z) \sim Cz^k$ for $z \to \infty$, and thus $\mathcal{R}(z) \not\equiv 1$. □

However, as shown below, the above barrier for L-stable Runge–Kutta methods can be circumvented by applying an appropriate postprocessor to (2.6) of the form

$$(2.8) \qquad \overline{X}_n = \mathcal{C}(z)X_n + \mathcal{D}(z)\sqrt{h}\sigma\xi_n, \qquad z = -\lambda h.$$

We will see that this feature serves as a crucial ingredient in section 3 to achieve high order in the SPDE case.

PROPOSITION 2.2. *The method* (2.2) *applied to the SDE test problem* (2.5) *with* $f_1(x) = -\lambda x$ *and* $f_2(x) = 0$ *is such that the scheme* $X_n \mapsto X_{n+1}$ *is L-stable and* $\overline{X}_n$ *is exact for sampling the invariant measure, i.e., for all test functions ϕ and all timesteps h,* $\lim \text{a. s.}_{M\to\infty} \frac{1}{M+1}\sum_{n=0}^{M} \phi(\overline{X}_n) = \lim_{n\to+\infty}\mathbb{E}(\phi(\overline{X}_n)) = \int_\mathbb{R} \phi(y)\rho_\infty(y)dy.$

Proof. The method is L-stable because it has the same stability function $\mathcal{A}(z) = \frac{1}{1-z}$ as the linearized implicit Euler method (1.6). A calculation using the above notation yields

$$(2.9) \qquad \lim_{n \to \infty} \mathbb{E}(\overline{X}_n^2) = \frac{\sigma^2}{2\lambda} \overline{\mathcal{R}}(z), \quad \text{where } \overline{\mathcal{R}}(z) = \mathcal{C}(z)^2 \mathcal{R}(z) - 2z\mathcal{D}(z)^2,$$

with $z = -\lambda h$ and $\mathcal{R}(z)$ given by (2.7). Noting that $\mathcal{A}(z) = \mathcal{B}(z) = (1-z)^{-1}$, $\mathcal{C}(z) = 1$, $\mathcal{D}(z) = \frac{1}{2}(1-z/2)^{-1/2}$ yields $\overline{\mathcal{R}}(z) \equiv 1$, which proves that the method is exact. $\square$

Remark 2.2. Relaxing the assumption that the matrices A and Q commute, the scheme (2.2) can be adapted to remain exact for sampling the invariant measure when applied to the linear problem $dX = AX dt + \sigma Q^{1/2} dW$ without nonlinearity ($f_1(x) = Ax, f_2(x) = 0$). In fact, the definition of matrix J_3 should be modified and obtained solving a system of Lyapunov equations.

2.2. Construction of the integrator (2.2). We explain in this section how we construct method (2.2) with high order of accuracy for the invariant measure by using the idea of postprocessing from [38]. Consider a system of SDEs in $\mathbb{R}^N$ of the form

$$(2.10) \qquad dX(t) = f_0(X(t))dt + \sigma dW(t), \qquad X(0) = X_0,$$

where $\sigma > 0$ and $\big(W(t)\big)_{t \in \mathbb{R}^+}$ is a standard d-dimensional Wiener process.

Assumption 2.1. We assume the following:
1. f_0 is of class C^∞, with bounded derivatives of any order, and there exists a potential function $V_0 : \mathbb{R}^N \to \mathbb{R}$ such that $f_0 = -\nabla V_0$;
2. there exist $C, \beta > 0$ such that for all $x \in \mathbb{R}^N$, $x^T f_0(x) \leq -\beta x^T x + C$.

Then, system (2.10) is ergodic with a unique invariant measure μ_∞ (see, e.g., [18]) given by

$$\mu_\infty(dy) = \rho(y)dy \quad \text{with } \rho(y) = \frac{1}{Z} \exp\left(-\frac{2}{\sigma^2} V_0(y)\right).$$

Moreover the solution $X(t)$ of (2.10) satisfies an exponential ergodicity property analogous to (1.5).

The following theorem is the main result in [38] where postprocessed integrators for SDEs are introduced. It permits us to improve the accuracy of a method of weak order q to order $r = q + 1$ for the invariant measure. Notice that this theorem remains valid for general classes of SDEs with additive or multiplicative noise in multiple dimensions. The error estimates in [38] rely on classical results from Talay and Tubaro [36] and Milstein [28] based on the backward Kolmogorov equation (see also [29, Chapters 2.2, 2.3]). In this section, we denote by $\mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$ the set of C^∞ functions whose derivatives up to any order have a polynomial growth. We denote by $\mathcal{L}$ the infinitesimal generator of (2.1) where we set $f_0 = f_1 + f_2$: for any test function $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$,

$$(2.11) \qquad \mathcal{L}\phi = f_0 \cdot \nabla\phi + \frac{\sigma^2}{2}\Delta\phi,$$

where $\nabla\phi$ denotes the gradient of ϕ and $\Delta\phi$ the Laplacian of ϕ.

THEOREM 2.3 (see [38, Theorem 4.1]).   *Under Assumption* 2.1, *let* X_n *be an ergodic numerical solution of* (2.10) *with bounded moments of any order* $M \in \mathbb{N}^*$, *i.e.,*

$$\mathbb{E}(\|X_n\|^M) \leq C_M \tag{2.12}$$

for all $n \geq 0$, *where* C_M *is independent of* h, n. *Assume further that the scheme has local weak order* $q \geq 1$, *i.e., it satisfies for all initial condition* $X_0 = x$ *and all* h *sufficiently small,*

$$|\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(h)))| \leq C(x, \phi)h^{q+1} \tag{2.13}$$

for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$, *where* $x \mapsto C(x, \phi)$ *has a polynomial growth with respect to* x.
    Assume also that the numerical solution admits a weak Taylor series expansion of the form

$$\mathbb{E}(\phi(X_1)) = \phi(x) + h\mathcal{A}_0\phi(x) + h^2\mathcal{A}_1\phi(x) + \cdots \tag{2.14}$$

for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$, *where* $\mathcal{A}_i : \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R}) \to \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$, $i = 0, 1, 2, \ldots$, *are linear differential operators with smooth coefficients. Let* G_n *denote independent and identically distributed (i.i.d.) random maps in* $\mathbb{R}^N$, *independent of* $\{X_j\}_{j \leq n}$, *with* $\overline{X}_n = G_n(X_n)$ *having bounded moments of any order, and satisfying a weak Taylor expansion of the form*

$$\mathbb{E}(\phi(G_n(x))) = \phi(x) + h^q\overline{\mathcal{A}}_q\phi(x) + \mathcal{O}(h^{q+1}) \tag{2.15}$$

for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$, *where the constant in* $\mathcal{O}$ *has a polynomial growth with respect to* x. *Assuming further*

$$(\mathcal{A}_q + [\mathcal{L}, \overline{\mathcal{A}}_q])^*\rho = 0,$$

where we use the commutator notation $[A, B] = AB - BA$. *Then the postprocessor* $\overline{X}_n = G_n(X_n)$ *yields an approximation of order* $r = q + 1$ *for the invariant measure,*

$$\left| \lim_{M \to \infty} \text{a. s. } \frac{1}{M+1} \sum_{n=0}^M \phi(\overline{X}_n) - \int_{\mathbb{R}^N} \phi(y)d\mu_\infty(y) \right| \leq C(\phi)h^{q+1}, \tag{2.16}$$

$$\left| \mathbb{E}(\phi(\overline{X}_n)) - \int_{\mathbb{R}^N} \phi(y)d\mu_\infty(y) \right| \leq K(\phi, x)e^{-\lambda t_n} + C(\phi)h^{q+1} \tag{2.17}$$

for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$ *with* $t_n = nh$, *where* $\lambda, K(\phi, x), C(\phi)$ *are independent of* n *and* h *assumed small enough.*

Remark 2.3. We emphasize that the boundedness of moment condition (2.12) can be easily proved for Runge–Kutta type methods, such as the proposed method (2.2), following the methodology of [28] (see also [29, Chapter 2.2]) using the global Lipschitz continuity of the SDE fields. We also refer to [38], where this assumption is discussed in the context of postprocessed integrators for SDEs. Notice also that the Lipschitz condition on the SDE fields and the ergodicity assumption on the numerical method could be relaxed using the concept of rejecting exploding trajectories, as introduced in [31]; see also [30] in the context of ergodic SDEs.

We are now in position to state the main result of this section. We show that the new method (2.2) satisfies the assumptions of Theorem 2.3 with $q = 1$ and thus has order $r = 2$ of accuracy for the invariant measure of ergodic SDEs.

THEOREM 2.4. *Under Assumption* 2.1, *consider the method* (2.2) *with postprocessor* $\overline{X}_n$. *Assume that* X_n *is ergodic when applied to the system* (2.1). *Then* $\overline{X}_n$ *has order* 2 *of accuracy for the invariant measure, precisely,*

$$(2.18) \quad \left| \lim_{M \to \infty} \text{a. s.} \frac{1}{M+1} \sum_{k=0}^{M} \phi(\overline{X}_k) - \int_{\mathbb{R}^N} \phi(y) d\mu_\infty(y) \right| \leq C(\phi) h^2,$$

$$(2.19) \quad \left| \mathbb{E}(\phi(\overline{X}_n)) - \int_{\mathbb{R}^N} \phi(y) d\mu_\infty(y) \right| \leq K(\phi, x) e^{-\lambda t_n} + C(\phi) h^2$$

for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$ *with* $t_n = nh$, *where* $\lambda, K(\phi, x), C(\phi)$ *are independent of* n *and* h *assumed small enough.*

The proof of Theorem 2.4 relies on the following lemma, where conditions of order 2 of accuracy are derived for a perturbation for the linearized Euler method. The proof of this lemma is postponed to the appendix.

LEMMA 2.5. *Consider the following modification of the linearized Euler scheme for* (2.1):

$$Y_{n+1} = Y_n + h f_1 \left(Y_{n+1} + a_1 \sigma \sqrt{h} \xi_n \right) + h f_2 (Y_n + a_2 \sigma \sqrt{h} \xi_n) + (I + a_3 h f_1'(Y_n)) \sigma \sqrt{h} \xi_n,$$

$$(2.20)$$
$$\overline{Y}_n = Y_n + b_1 h f_1(\overline{Y}_n) + b_2 h f_2(Y_n) + c \sigma \sqrt{h} \xi_n.$$

where $a_1, a_2, a_3, b_1, b_2, c$ *are fixed real coefficients. If the conditions*

$$a_1^2 + 2a_1 + b_1 - c^2 = 0, \qquad a_1 + a_3 + \frac{1}{4} + b_1 - c^2 = 0, \qquad a_2^2 + b_2 - c^2 = 0,$$

$$(2.21) \quad -\frac{1}{4} + a_2 + b_2 - c^2 = 0, \qquad (b_2 - b_1)[f_2, f_1] = 0,$$

hold, then, assuming the ergodicity of Y_n *in* (2.20), *the postprocessed scheme* (2.20) *satisfies the assumptions of Theorem* 2.3 *with* $q = 1$, *and* $\overline{Y}_n$ *has order* 2 *of accuracy for the invariant measure, i.e., it satisfies* (2.18), (2.19) *(with* $\overline{X}_n$ *replaced by* $\overline{Y}_n$).

Note that the Lie bracket $[f_2, f_1] = f_2' f_1 - f_1' f_2$ involved in the second-order conditions (2.21) vanishes only when the flows associated to the fields f_1, f_2 commute, which is not true in general. We thus impose $b_1 = b_2$. Still, the system (2.21) has infinitely many solutions. Setting $b_1 = b_2 = 0$ for simplicity of the postprocessor, two solutions remain. Choosing the solution which minimizes the absolute value of a_1 and a_3, we obtain the following choice of coefficients for the order 2 scheme (2.20):

$$(2.22) \qquad a_1 = -a_3 = \frac{-2 + \sqrt{5}}{2}, \qquad a_2 = c = \frac{1}{2}, \qquad b_1 = b_2 = 0.$$

Proof of Theorem 2.4. It is sufficient to prove that the method (2.2) satisfies the same expansions (2.14) and (2.15) with $q = 1$ as method (2.20), (2.22) (with the same differential operators $\mathcal{A}_0 = \mathcal{L}$, $\mathcal{A}_1$, and $\overline{\mathcal{A}}_1$). Indeed, applying Theorem 2.3 with $q = 1$ to the scheme (2.2), we deduce that method (2.2) also has second order of accuracy for the invariant measure, which concludes the proof of Theorem 2.4.

To recover the scheme (2.2) from (2.20), (2.22), in the first line of (2.20) one has to replace ξ_n with $J_{n,2}\xi_n = (I + \mathcal{O}(h))\xi_n$ in the arguments of f_1, f_2 and also to substitute $I + a_3 h f_1'(x)$ with

$$\left(\frac{1 - \sqrt{2} + \sqrt{5}}{2} J_{n,1}^{-1} + \frac{1 + \sqrt{2} - \sqrt{5}}{2}\right) J_{n,2} = I + a_3 h f_1'(x) + \mathcal{O}(h^2);$$

in the second line of (2.20), one has to replace ξ_n with $J_{n,3}\xi_n = (I + \mathcal{O}(h))\xi_n$. We obtain that the difference between one step of (2.2) and one step of (2.20), (2.22) with initial condition $X_0 = Y_0 = x$ has the form $X_1 - Y_1 = R(x)\xi h^{5/2} + \mathcal{O}(h^3)$. Using $\mathbb{E}(\xi) = 0$, we deduce $\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(Y_1)) = \mathcal{O}(h^3)$, while $\mathbb{E}(\phi(\overline{X}_0)) - \mathbb{E}(\phi(\overline{Y}_0)) = \mathcal{O}(h^2)$. □

It can be seen from the proof of Theorem 2.4 that the operator $J_{n,2}$ in front of ξ_n and the operator $J_{n,3}$ in the definition of the method (2.2) have no influence on its order 2 of accuracy for the invariant measure in finite dimension. In infinite dimension, however, these operators play an important role for the well-posedness, the stability, and the accuracy of the scheme in the SPDE case presented in section 3.

3. Analysis in the SPDE case.

3.1. Abstract setting and assumptions. The state space in the SPDE case is an infinite dimensional separable Hilbert space $\mathcal{H}$, for which we denote by $\langle \cdot, \cdot \rangle$ the scalar product and by $|\cdot|$ the associated norm. Consider the linear operator A involved in the parabolic SPDE (1.1). Recall that we assume that $-A$ is an unbounded self-adjoint linear operator with eigenvalues $0 < \lambda_1 \leq \cdots \leq \lambda_p \leq \lambda_{p+1} \leq \cdots$, such that $\lambda_p \to +\infty$, when $p \to +\infty$, and associated normalized eigenvectors e_p (such that $Ae_p = -\lambda_p e_p$), which form a complete orthonormal system in $\mathcal{H}$.

For any $s \in \mathbb{R}^+$, we classically define the unbounded linear operator $(-A)^{s/2}$ from $\mathcal{H}$ to $\mathcal{H}$ and its domain $\mathcal{H}^s \subset \mathcal{H}$ as follows:

$$(-A)^{s/2}u = \sum_{p=1}^{+\infty} \langle u, e_p \rangle \lambda_p^{s/2} e_p \quad \text{for all } u \in \mathcal{H}^s = \left\{ u \in \mathcal{H} : |u|_s^2 = \sum_{p=1}^{+\infty} |\langle u, e_p \rangle|^2 \lambda_p^s < +\infty \right\}.$$

We also define the bounded linear operator $(-A)^{-s/2}$ and the semigroup $\left(e^{tA}\right)_{t \in \mathbb{R}^+}$, both as linear operators from $\mathcal{H}$ to $\mathcal{H}$, by

$$(-A)^{-s/2}u = \sum_{p=1}^{+\infty} \langle u, e_p \rangle \lambda_p^{-s/2} e_p, \qquad e^{tA}u = \sum_{p=1}^{+\infty} \exp(-t\lambda_p)\langle u, e_p \rangle e_p.$$

It is straightforward that for any $t \in (0, +\infty)$ we have $e^{tA} \in \mathcal{L}(\mathcal{H}, \mathcal{H})$—the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{H}$, endowed with the norm denoted by $\| \cdot \|$—with $\|e^{tA}\| \leq \exp(-\lambda_1 t)$. Moreover, the following regularization property holds true: $\|(-A)^{s/2}e^{tA}\| \leq \frac{C_s}{t^{s/2}}$, where $C_s^2 = \sup_{r \in \mathbb{R}^+} \exp(-2r)r^s \in (0, +\infty)$.

Covariance operator. We assume that Q is a bounded, nonnegative, self-adjoint linear operator from $\mathcal{H}$ to $\mathcal{H}$, which satisfies $Qe_p = q_p e_p$ for any $p \in \mathbb{N}^*$, where the eigenvalues $(q_p)_{p \in \mathbb{N}^*}$ form a bounded sequence of nonnegative real numbers. We assume that condition (1.2) is satisfied. The Q-Wiener process in (1.1) is then defined as follows: for any $t \geq 0$,

$$(3.1) \qquad W^Q(t) = \sum_{p=1}^{+\infty} \sqrt{q_p}\beta_p(t)e_p,$$

where $(\beta_p)_{p\in\mathbb{N}^*}$ is a sequence of independent standard scalar Wiener processes on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Notice that the operators A and Q commute: $AQu = QAu$ for any $u \in \mathcal{H}^2$. This property is often assumed in the literature and simplifies the analysis of the order of convergence made in section 3.3. Nevertheless several arguments (especially Proposition 3.1 and the results of section 3.2) do not require this property; in particular, the scheme remains well-defined in the noncommuting case.

Nonlinearity. The nonlinear coefficient F is assumed to be a Lipschitz continuous function from $\mathcal{H}$ to $\mathcal{H}$, with Lipschitz constant L satisfying the dissipation condition

$$(3.2) \qquad L = \sup_{u^1 \neq u^2 \in \mathcal{H}} \frac{|F(u^1) - F(u^2)|}{|u^1 - u^2|} < \min_{p \in \mathbb{N}^*} \lambda_p = \lambda_1.$$

This condition ensures ergodicity of the continuous-time process (Proposition 3.1) and of its time-discretized approximations (Proposition 3.3). A typical example of such a Lipschitz function on $\mathcal{H} = L^2(\mathcal{D})$—where $\mathcal{D}$ is an open smooth bounded domain in $\mathbb{R}^d$—is the Nemytskii operator $F : u \mapsto f \circ u$, where $f : \mathbb{R} \to \mathbb{R}$ is a globally Lipschitz function. Note that $F = -DV$ is the derivative of the potential function $V : L^2(\mathcal{D}) \to \mathbb{R}$ where $V(u) = -\int_0^1 \int_\mathcal{D} u(x) f(\theta u(x)) dx d\theta$.

Under the above hypotheses, and assuming the trace condition (1.2), the process $\left(u(t)\right)_{t\in\mathbb{R}^+}$ takes values in $\mathcal{H}^s$ for any $s < \overline{s}$, and we recall without proof the following result of exponential convergence to a unique invariant distribution; see, e.g., [10] for general results and [13, section 3.1.1].

PROPOSITION 3.1. *Assume* (1.2) *and the above hypotheses. Then, the process* $(u(t))_{t\in\mathbb{R}^+}$ *solution of* (1.1) *admits a unique invariant probability distribution* μ_∞ *on* $\mathcal{H}$. *Moreover for all* $s < \overline{s}$,

$$\int_\mathcal{H} |u|_s^2 \mu_\infty(du) < +\infty,$$

and for all $\phi : \mathcal{H} \to \mathbb{R}$ *Lipschitz continuous, and all* $t > 0$,

$$\left| \mathbb{E}\left[\phi(u(t))\right] - \int_\mathcal{H} \phi(v)\mu_\infty(dv) \right| \leq C(\phi, u_0) e^{-(\lambda_1 - L)t},$$

where $C(\phi, u_0)$ *is independent of* t.

Condition (3.2) is crucial for the proof of the uniqueness of the numerical invariant distributions established in the next section: we compare the solutions starting from different initial conditions and driven with the same noise process and show an exponential contraction similar to the result of Proposition 3.1. Notice that weaker conditions than (3.2) are known in the literature (see, e.g., [13] and references therein) to ensure the ergodicity and the exponential convergence of (1.1)—for instance, when F is bounded and Lipschitz continuous with no size restriction on L.

3.2. Stability and ergodicity of the integrator for SPDEs. In this section, we prove the existence and uniqueness of invariant distributions for the time-discretized processes defined by the numerical method, for any timestep size $h > 0$, in a general setting. Notice that the results of this section do not require the gradient assumption $F = -DV$. The results are analogous to classical results for the θ-method in the context of stiff SDEs [6, Theorem 3.1] and for the linearized implicit Euler method (1.9) in the context of SPDEs as studied, e.g., in [4, Remark 4.8].

It is a key observation here to exploit that the sequence $(u_n, \bar{u}_{n-1})_{n \in \mathbb{N}}$ defining the new scheme (1.9), (1.10) is a Markov chain on the product space $\mathcal{H} \times \mathcal{H}$. The initial condition $(u_0, \bar{u}_{-1})$ is given by $u_0 = u(0) = u_0$ and an arbitrary $\bar{u}_{-1} \in \mathcal{H}$ which plays no role in the dynamics, since $(u_{n+1}, \bar{u}_n)$ depends only on u_n and ξ_n^Q, not on $\bar{u}_{n-1}$.

In the following proposition, we state uniform bounds—with respect to $n \in \mathbb{N}$ and $h \in (0, 1)$—on first-order moments for the norm $|\cdot|_s$ for $s < \bar{s}$.

PROPOSITION 3.2. *Assume the hypotheses of section* 3.1 *and consider the scheme* (1.9), (1.10). *For all* $s \in [0, \bar{s})$, *assuming* $u_0 \in \mathcal{H}^s$, *there exists a constant* $C_s \in (0, +\infty)$ *such that for all* $h \in (0, 1)$,

$$\sup_{n \in \mathbb{N}} \mathbb{E}|u_n|_s \leq C_s(1 + |u_0|_s), \qquad \sup_{n \in \mathbb{N}} \mathbb{E}|\bar{u}_n|_s \leq C_s(1 + |u_0|_s).$$

Proof. Thanks to (1.2), $\mathrm{Trace}(J_i Q J_i) < +\infty$ for $i \in \{1, 2, 3\}$, and thus u_n and $\bar{u}_n$ are well-defined in $\mathcal{H}$ for all $n \in \mathbb{N}$. The contributions of the drift part and of the stochastic perturbation are treated separately: we introduce the auxiliary process $(\ell_n)_{n \in \mathbb{N}}$, as the solution of the equation

$$\ell_{n+1} = J_1 \ell_n + \sigma \sqrt{h} \left(\frac{\sqrt{2}-1}{2} J_1 + \frac{3-\sqrt{2}}{2} I \right) J_2 \xi_n^Q$$

with $\ell_0 = 0$. Set $d_n = u_n - \ell_n$ for all $n \in \mathbb{N}$; then $d_0 = u_0$ and

$$d_{n+1} = J_1 d_n + h J_1 F\left(d_n + \ell_n + \frac{1}{2} J_2 \sigma \sqrt{h} \xi_n^Q \right).$$

The quantity ℓ_n satisfies for all $n \in \mathbb{N}^*$ the identity

$$\ell_n = \sigma \sqrt{h} \left(\frac{\sqrt{2}-1}{2} J_1 + \frac{3-\sqrt{2}}{2} I \right) J_2 J_1^{-1} \sum_{k=0}^{n-1} (J_1)^{n-k} \xi_k^Q.$$

Observe that $\left(\frac{\sqrt{2}-1}{2} J_1 + \frac{3-\sqrt{2}}{2} I \right) J_2 J_1^{-1}$ is a bounded linear operator from $\mathcal{H}^s$ to $\mathcal{H}^s$ with norm less than $\frac{2}{3-\sqrt{2}}$ for $s \in [0, \bar{s})$. Since the Gaussian random variables $(\xi_k^Q)_{k \in \mathbb{N}}$ are independent, for all $0 \leq s \leq \bar{s}$,

$$\mathbb{E}|\ell_n|_s^2 \leq \frac{2\sigma^2 h}{3-\sqrt{2}} \sum_{k=0}^{n-1} \mathbb{E}\left|(J_1)^{n-k} \xi_k^Q\right|_s^2 \leq \frac{2\sigma^2 h}{3-\sqrt{2}} \sum_{k=1}^{+\infty} \sum_{p=1}^{+\infty} \frac{q_p \lambda_p^s}{(1+\lambda_p h)^{2k}}$$

$$\leq \frac{2\sigma^2}{3-\sqrt{2}} \sum_{p=1}^{+\infty} q_p \lambda_p^{s-1} \frac{\lambda_p h}{(1+\lambda_p h)^2 - 1} \leq \frac{\sigma^2}{3-\sqrt{2}} \mathrm{Trace}\left((-A)^{-1+s} Q\right).$$

Now thanks to (3.2), straightforward computations show that

$$\mathbb{E}|d_n| \leq \frac{1+hL}{1+\lambda_1 h} \mathbb{E}|d_{n-1}| + \frac{h}{1+\lambda_1 h} |F(0)| + \frac{Lh}{1+\lambda_1 h} \left(\mathbb{E}|\ell_{n-1}| + \mathbb{E}\left|\frac{1}{2} J_2 \sigma \sqrt{h} \xi_{n-1}^Q\right| \right)$$

$$\leq \frac{(1+hL)^n}{(1+\lambda_1 h)^n} |u_0| + \frac{|F(0)|}{\lambda_1 - L} + \frac{L}{\lambda_1 - L} \left(\sup_{k \in \mathbb{N}} \mathbb{E}|\ell_k| + \frac{1}{3-\sqrt{2}} \sigma \left(\mathrm{Trace}\left((-A)^{-1} Q\right)\right)^{1/2} \right)$$

with $\frac{(1+hL)}{(1+\lambda_1 h)} \leq 1 - \frac{\lambda_1 - L}{1+\lambda_1 h} h \leq \exp\left(-\frac{(\lambda_1 - L)}{1+\lambda_1 h} h\right) < 1$.

As a consequence the claim follows for $(u_n)_{n\in\mathbb{N}}$ in the case $s = 0$. In particular, for some constant $C \in (0, +\infty)$ it comes that $\sup_{n\in\mathbb{N}} \mathbb{E}\big|F(u_n + \frac{1}{2}J_2\sigma\sqrt{h}\xi_n^Q)\big| < C(1 + |u_0|)$.

The case $s \in (0, \bar{s})$ is treated using the estimates of Lemma 3.2 in [4]: for $n \in \mathbb{N}$

$$
\begin{aligned}
\mathbb{E}\big|d_n\big|_s &= \mathbb{E}\left|(-A)^{s/2}(J_1)^n u_0 + h\sum_{k=0}^{n-1}(-A)^{s/2}(J_1)^{n-k}F\left(u_k + \frac{1}{2}J_2\sigma\sqrt{h}\xi_k^Q\right)\right| \\
&\leq \frac{1}{(1+\lambda_1 h)^n}|u_0|_s + hC(1+u_0|)\sum_{k=1}^{n}\left(\frac{\mathbb{1}_{kh\leq 1}}{(kh)^{s/2}} + \mathbb{1}_{kh>1}\frac{C_s}{(1+\lambda_1 h)^{k-\lfloor 1/h\rfloor}}\right) \\
&\leq C_s(1+|u_0|_s).
\end{aligned}
$$

Finally, for $n \in \mathbb{N}$,

$$
\mathbb{E}\left|\frac{1}{2}J_3\sigma\sqrt{h}\xi_n^Q\right|_s^2 = \frac{\sigma^2 h}{4}\text{Trace}\left(\left(I - \frac{h}{2}A\right)^{-1}(-A)^s Q\right) \leq \frac{\sigma^2}{2}\text{Trace}\big((-A)^{-1+s}Q\big).
$$

This concludes the proof, since $\bar{u}_n = u_n + \frac{1}{2}J_3\sigma\sqrt{h}\xi_n^Q$.   □

We now state the following existence and uniqueness result for the invariant distribution of the Markov chain $(u_n, \bar{u}_{n-1})_{n\in\mathbb{N}}$. We prove that the convergence is exponentially fast, in contrast to the trapezoidal method (1.11) which is not L-stable and for which such exponential estimate does not hold in general for stiff problems (even in finite dimension).

PROPOSITION 3.3. *Assume the hypotheses of section 3.1. For any $h \in (0,1)$, the $\mathcal{H}\times\mathcal{H}$-valued Markov chain $(u_n, \bar{u}_{n-1})_{n\in\mathbb{N}}$ admits a unique invariant distribution M_∞^h in $\mathcal{H}\times\mathcal{H}$, with marginals in $\mathcal{H}$ denoted by μ_∞^h and $\bar{\mu}_\infty^h$, respectively.*

Moreover, the convergence of the distributions to equilibrium is exponentially fast: for all Lipschitz test function $\varphi : \mathcal{H} \to \mathbb{R}$, and all $t_n = nh$,

$$
\left|\mathbb{E}\varphi(u_n) - \int_\mathcal{H}\varphi d\mu_\infty^h\right| + \left|\mathbb{E}\varphi(\bar{u}_n) - \int_\mathcal{H}\varphi d\bar{\mu}_\infty^h\right| \leq C(\varphi, |u_0|)\exp\left(-\frac{(\lambda_1 - L)}{1+\lambda_1 h}t_n\right).
$$

Proof. Existence. The semigroup $(P^n)_{n\in\mathbb{N}}$ on $\mathcal{H}\times\mathcal{H}$ generated by the Markov chain $(u_n, \bar{u}_{n-1})_{n\in\mathbb{N}}$ satisfies the Feller property: for any $n \in \mathbb{N}$, for any bounded continuous test function $\phi : \mathcal{H}\times\mathcal{H} \to \mathbb{R}$, the map $(u_0, \bar{u}_{-1}) \mapsto P^n\phi(u_0, \bar{u}_{-1}) = \mathbb{E}\phi(u_n, \bar{u}_{n-1})$ is continuous. The claim then follows from the standard Krylov–Bogoliubov criterion (see section 3.1 in [10]): given an arbitrary initial condition $(u_0, \bar{u}_{-1}) \in \mathcal{H}\times\mathcal{H}$, if M_n denotes the law of $(u_n, \bar{u}_{n-1})$, then

- $\left(\frac{1}{n+1}\sum_{k=0}^{n}\mathrm{M}_k\right)_{n\in\mathbb{N}}$ is a *tight* sequence of probability distributions on $\mathcal{H}\times\mathcal{H}$— as a consequence of Proposition 3.2 combined with the Markov inequality, and of the observation that for any $s \in (0, \bar{s})$ and any $R > 0$, the set $\{|u|_s \leq R, |\bar{u}|_s \leq R\}$ is a compact subset of $\mathcal{H}\times\mathcal{H}$;
- every subsequence limit point M is an invariant distribution for the semigroup.

Uniqueness. Consider two initial conditions $u_0^1, u_0^2 \in \mathcal{H}$, as well as $\bar{u}_{-1}^1, \bar{u}_{-1}^2 \in \mathcal{H}$ and the associated processes $(u_n^i)_{n\in\mathbb{N}}$ and $(\bar{u}_n^i)_{n\in\mathbb{N}}$, for $i = 1, 2$, defined by (1.9), (1.10), and driven by a unique noise process $(\xi_n^Q)_{n\in\mathbb{N}}$.

Then by Lipschitz continuity of F, and using the cancellations of several noise terms, computations similar to those of the proof of Proposition 3.2 yield for any

$n \in \mathbb{N}$ the almost sure contraction property

$$\left|\overline{u}_n^1 - \overline{u}_n^2\right| = \left|u_n^1 - u_n^2\right| \le \frac{1+Lh}{1+\lambda_1 h}\left|u_{n-1}^1 - u_{n-1}^2\right| \le \exp\left(-\frac{(\lambda_1 - L)}{1+\lambda_1 h}t_n\right)\left|u_0^1 - u_0^2\right|.$$

Finally, taking $(u_0^2, \overline{u}_{-1}^2)$ random, independent of the noise process $\left(\xi_n^Q\right)_{n\in\mathbb{N}}$ and distributed according to an ergodic invariant distribution M_∞^h gives the exponential convergence and the uniqueness properties. $\qquad\square$

3.3. Analysis of the order of convergence: A simplified linear case. It is shown in [4] (for $d = 1$ and $\overline{s} = 1/2$, associated with space-time white noise $Q = I$), that the standard linearized implicit Euler scheme (1.9) has order $r = 1/2 - \varepsilon$ for all $\varepsilon \in (0, 1/2)$ for the approximation of the invariant distribution μ_∞ of (1.1). In this section, we show that the postprocessed scheme has the improved order of convergence $\overline{s} + 1 - \varepsilon$. Since the techniques from section 2 do not extend straightforwardly to the SPDE case, we only focus on a simplified case, where the nonlinear coefficient F is replaced with a bounded linear operator. Numerical experiments of section 4 show that the higher order is preserved for various examples of nonlinearities F.

In addition to the hypotheses of section 3.1, assume that the coefficient F is given by a linear mapping: for any $u \in \mathcal{H}$, $F(u) = Bu$, where $B \in \mathcal{L}(\mathcal{H})$ satisfies $Be_p = -b_p e_p$ for all $p \in \mathbb{N}^*$, with real eigenvalues $b_p \in (-\lambda_1, \lambda_1)$ (due to condition (3.2)):

$$(3.3) \qquad du(t) = Au(t)dt + Bu(t)dt + \sigma dW^Q(t), \quad u(0) = u_0.$$

In this situation, the components $\langle u(t), e_p \rangle$ in the basis $\{e_p\}_{p\in\mathbb{N}^*}$ of the solution $\left(u(t)\right)_{t\in\mathbb{R}^+}$ of the SPDE (1.1) are independent Ornstein–Uhlenbeck processes. Similarly, the components in the basis $\{e_p\}_{p\in\mathbb{N}^*}$ of the discrete-time processes $\left(v_n\right)_{n\in\mathbb{N}}$ (resp., $\left(u_n\right)_{n\in\mathbb{N}}$, resp., $\left(\overline{u}_n\right)_{n\in\mathbb{N}}$) are also independent processes.

As a consequence, explicit expressions for the invariant distributions μ_∞, ν_∞^h, μ_∞^h, and $\overline{\mu}_\infty^h$ are available: they are centered Gaussian probability measures on $\mathcal{H}$,

$$\mu_\infty = \mathcal{N}\left(0, Q_\infty\right), \quad Q_\infty = \frac{\sigma^2}{2}Q(-A-B)^{-1},$$

$$\nu_\infty^h = \mathcal{N}\left(0, Q_{\infty,\nu}^h\right), \quad Q_{\infty,\nu}^h = \frac{\sigma^2}{2}Q(-A-B)^{-1}\left(I - h\frac{A-B}{2}\right)^{-1},$$

$$\mu_\infty^h = \mathcal{N}\left(0, Q_\infty^h\right), \quad \overline{\mu}_\infty^h = \mathcal{N}\left(0, \overline{Q}_\infty^h\right).$$

The expressions for Q_∞^h and $\overline{Q}_\infty^h$ being more complicated are displayed in (3.7) below.

The following lemma is a key elementary tool in order to exhibit the order of convergence as $h \to 0$ of the approximating measures toward μ_∞.

LEMMA 3.4. *Let $\pi_j = \mathcal{N}(0, Q_j)$, $j \in \{1, 2\}$, be two Gaussian probability distributions on the Hilbert space $\mathcal{H}$. Assume that for all $p \in \mathbb{N}^*$, $j \in \{1, 2\}$ $Q_j e_p = q_{j,p} e_p$. Let $\varphi \in \mathcal{C}^2(\mathcal{H}, \mathbb{R})$ satisfy that $\sup_{v\in\mathcal{H}} \|D^2\varphi(v)\| < +\infty$. Then*

$$(3.4) \qquad \left|\int_{\mathcal{H}} \varphi d\pi_2 - \int_{\mathcal{H}} \varphi d\pi_1\right| \le \frac{\sup_{v\in\mathcal{H}}\|D^2\varphi(v)\|}{2}\sum_{p=1}^{+\infty}|q_{2,p} - q_{1,p}|.$$

Remark 3.1. Assume that $q_{2,p} \ge q_{1,p}$. Then the above Lemma 3.4 yields optimal orders of convergence. Indeed choosing the test function $\varphi_{\mathrm{opt}}(u) = \exp(-|u|^2)$,

Lemma 9.5 in [23] yields that

$$\int_{\mathcal{H}} \varphi_{\mathrm{opt}} d\pi_1 - \int_{\mathcal{H}} \varphi_{\mathrm{opt}} d\pi_2 \geq \frac{\mathrm{Trace}(Q_2 - Q_1)}{\exp(6\mathrm{Trace}(Q_2))}.$$

This means that the quantity $\mathrm{Trace}(Q_2 - Q_1)$ also provides a lower bound for the error between the invariant distributions π_1 and π_2.

Proof of Lemma 3.4. Let $(\gamma_p)_{p\in\mathbb{N}^*}$ and $(\delta_p)_{p\in\mathbb{N}^*}$ be two independent sequences of i.i.d. standard real valued Gaussian random variables, centered and with variance 1. Set $X_j = \sum_{p\in\mathbb{N}^*} \sqrt{q_{j,p}}\gamma_p e_p$, and $R_j = \sum_{p\in\mathbb{N}^*} \sqrt{\max((-1)^j(q_{1,p} - q_{2,p}), 0)}\delta_p e_p$, for $j \in \{1, 2\}$. Observe that $X_j \sim \pi_j$ and that $X_1 + R_1$ and $X_2 + R_2$ have the same Gaussian distribution. This yields

$$\left| \int_{\mathcal{H}} \varphi d\pi_2 - \int_{\mathcal{H}} \varphi d\pi_1 \right| = \left| \mathbb{E}[\varphi(X_2)] - \mathbb{E}[\varphi(X_1)] \right|$$

$$= \left| \mathbb{E}[\varphi(X_2)] - \mathbb{E}[\varphi(X_2 + R_2)] + \mathbb{E}[\varphi(X_1 + R_1)] - \mathbb{E}[\varphi(X_1)] \right|$$

$$\leq \left| \mathbb{E}[\varphi(X_2 + R_2)] - \mathbb{E}[\varphi(X_2)] - \mathbb{E}[D\varphi(X_2).R_2] \right|$$

$$+ \mathbb{E}\left| [\varphi(X_1 + R_1)] - \mathbb{E}[\varphi(X_1)] - \mathbb{E}[D\varphi(X_1).R_1] \right|.$$

Indeed, $\mathbb{E}[D\varphi(X_j).R_j] = 0$, because X_j and R_j are independent and $\mathbb{E}[R_j] = 0$. Using a second-order Taylor expansion, we deduce

$$\left| \int_{\mathcal{H}} \varphi d\pi_2 - \int_{\mathcal{H}} \varphi d\pi_1 \right| \leq \frac{\sup_{v\in\mathcal{H}} \|D^2\varphi(v)\|}{2}(\mathbb{E}|R_1|^2 + \mathbb{E}|R_2|^2),$$

where we note $\mathbb{E}|R_1|^2 + \mathbb{E}|R_2|^2 = \sum_{p=1}^{+\infty} |q_{2,p} - q_{1,p}|$. $\qquad\square$

We now explain how Lemma 3.4 permits us to find the order of convergence of $\overline{\mu}_\infty^h$ to μ_∞. We define, for $p \in \mathbb{N}^*$, the component processes $u_\cdot(p)$ and $\overline{u}_\cdot(p)$ by projecting on the eigenvector e_p: for any $n \in \mathbb{N}$, $u_n(p) = \langle u_n, e_p \rangle$, $\overline{u}_{n-1}(p) = \langle \overline{u}_{n-1}, e_p \rangle$. Then (1.9), (1.10) applied to (3.3) is rewritten as a system of independent equations, decoupled with respect to $p \in \mathbb{N}^*$,

$$(3.5) \qquad \begin{aligned} u_{n+1}(p) &= \mathcal{A}(-\lambda_p h, -b_p h)u_n(p) + \sigma\sqrt{h}\sqrt{q_p}\mathcal{B}(-\lambda_p h, -b_p h)\xi_{n,p}, \\ \overline{u}_n(p) &= \mathcal{C}(-\lambda_p h)u_n(p) + \sigma\sqrt{h}\sqrt{q_p}\mathcal{D}(-\lambda_p h)\xi_{n,p}, \end{aligned}$$

where $\sqrt{q_p}\xi_{n,p} = \langle \xi_n^Q, e_p \rangle$: thus $(\xi_{n,p})_{n\in\mathbb{N}, p\in\mathbb{N}^*}$ are independent standard Gaussian random variables, and the rational functions $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ satisfy for any $z \in (-\infty, 0)$ and $\beta \in (-1, \min(1, |z|))$,

(3.6)

$$\mathcal{A}(z, \beta) = \frac{1+\beta}{1-z} \quad \mathcal{B}(z, \beta) = \frac{1 + \frac{\beta}{2} - \frac{3-\sqrt{2}}{2}z}{(1-z)(1 - \frac{3-\sqrt{2}}{2}z)}, \quad \mathcal{C}(z) = 1, \quad \mathcal{D}(z) = \frac{1}{2(1 - z/2)^{1/2}}.$$

Since for $\beta \in (-1, \min(1, |z|))$ the stability condition $|\mathcal{A}(z, \beta)| < 1$ is satisfied, straightforward computations yield

$$(3.7) \qquad \begin{aligned} Q_\infty^h e_p &= \lim_{n\to+\infty} \mathbb{E}|u_n(p)|^2 e_p = \frac{\sigma^2 q_p}{2(\lambda_p + b_p)}\mathcal{R}(-\lambda_p h, -b_p h)e_p, \\ \overline{Q}_\infty^h e_p &= \lim_{n\to+\infty} \mathbb{E}|\overline{u}_n(p)|^2 e_p = \frac{\sigma^2 q_p}{2(\lambda_p + b_p)}\overline{\mathcal{R}}(-\lambda_p h, -b_p h)e_p, \end{aligned}$$

where $\mathcal{R}(z,\beta) = \frac{-2(z+\beta)\mathcal{B}(z,\beta)^2}{1-\mathcal{A}(z,\beta)^2}$ and

$$\overline{\mathcal{R}}(z,\beta) = \mathcal{C}(z)^2\mathcal{R}(z,\beta) - 2(z+\beta)\mathcal{D}(z)^2 = 1 + \beta z\frac{P_1(z)\beta + P_2(z)}{(2+\beta-z)P_3(z)}$$

with polynomial functions $P_1(z) = 10 - 4\sqrt{2} - (11 - 6\sqrt{2})z$, $P_2(z) = 20 - 8\sqrt{2} - (44 - 24\sqrt{2})z + (11 - 6\sqrt{2})z^2$, and $P_3(z) = (2-z)(2-(3-\sqrt{2})z)^2$.

The following estimate $\overline{\mathcal{R}}(z,\beta) = 1 + \mathcal{O}(z\beta)$ as $z,\beta \to 0$ is crucial to obtain an improved order for the convergence of $\overline{\mu}_\infty^h$ to μ_∞ when $h \to 0$. It is not surprising because the scheme samples exactly the invariant measure of (3.3) in both cases $A = 0$ or $B = 0$ (as already shown in Proposition 2.2), equivalently $\overline{\mathcal{R}}(z,0) = \overline{\mathcal{R}}(0,\beta) = 1$ for all z,β.

LEMMA 3.5. *For all $z \leq 0$ and all $\beta \in (-1, \min(1, |z|))$, we have*

$$\left|1 - \overline{\mathcal{R}}(z,\beta)\right| \leq |z\beta|\frac{(15 - 6\sqrt{2})}{4(1-z)^2}.$$

Proof. Observe that for all $z \leq 0$, we have $P_2(z) \geq P_1(z) \geq 0$ and $(2+\beta-z) > 0$, $P_3(z) > 0$. Since $(2+\beta-z)^{-1} \leq (1-z)^{-1}$, we obtain

$$\left|1 - \overline{\mathcal{R}}(z,\beta)\right| \leq |z\beta|\frac{P_1(z) + P_2(z)}{(1-z)P_3(z)}.$$

The estimate then follows by observing that $(1-z)\frac{P_1(z)+P_2(z)}{P_3(z)}$ is an increasing function of $z \leq 0$, with maximum at $z = 0$, given by $\frac{P_1(0)+P_2(0)}{P_3(0)} = \frac{(15-6\sqrt{2})}{4}$. $\qquad\square$

We are in position to state our main convergence result, which yields the order of convergence $r = s + 1$ for any $s < \overline{s}$ for the invariant distribution with postprocessing $\overline{\mu}_\infty^h$.

THEOREM 3.6. *Consider the method (1.9)–(1.10) applied to (3.3). Let $\varphi \in \mathcal{C}^2$ $(\mathcal{H}, \mathbb{R})$ such that $\sup_{v\in\mathcal{H}} \|D^2\varphi(v)\| < +\infty$. For any $s < \overline{s}$ there exists $C_s \in (0, +\infty)$ such that for any $h \in (0, 1/L)$ we have*

$$(3.8) \qquad \left|\int_{\mathcal{H}} \varphi d\mu_\infty - \int_{\mathcal{H}} \varphi d\overline{\mu}_\infty^h\right| \leq C_s \sup_{v\in\mathcal{H}} \|D^2\varphi(v)\|\sigma^2 h^{1+s}.$$

Moreover, if $B = 0$ (i.e., $b_p = 0$ for all $p \in \mathbb{N}^*$), then the method is exact: $\overline{\mu}_\infty^h = \overline{\mu}_\infty$.

Proof. Thanks to Lemma 3.4, it is sufficient to control

$$\sum_{p=1}^{+\infty} \left|\langle(Q_\infty - \overline{Q}_\infty^h)e_p, e_p\rangle\right| = \sum_{p=1}^{+\infty} \frac{\sigma^2 q_p}{2(\lambda_p + b_p)}\left|\overline{\mathcal{R}}(-\lambda_p h, -b_p h) - 1\right|$$

$$\leq C\lambda_1\sigma^2 \sum_{p=1}^{+\infty} \frac{q_p|b_p|}{\lambda_p}\frac{\lambda_p h^2}{(1+\lambda_p h)^2}$$

$$\leq C\lambda_1^2\sigma^2 \sum_{p=1}^{+\infty} q_p\lambda_p^{-1+s}\frac{(\lambda_p h)^{1-s}h^{1+s}}{(1+\lambda_p h)^2}$$

$$\leq C_s\lambda_1^2\sigma^2\text{Trace}\big((-A)^{-1+s}Q\big)h^{1+s},$$

which gives the order of convergence $s + 1$ for all $s < \overline{s}$. Moreover, it is clear that $\overline{\mathcal{R}}(z,0) = 1$ for any $z \leq 0$, so that if $b_p = 0$ for all $p \in \mathbb{N}^*$, then $\overline{Q}_\infty^h = Q_\infty$. $\qquad\square$

Remark 3.2. We see in Lemma 3.5 that the error is zero if $z = 0$ or $\beta = 0$. This is related to Proposition 2.2, which shows that the error of the postprocessed method is zero in the linear case when $\beta = 0$. This feature permits us to gain one power of h and thus one order of accuracy in the proof above. In contrast, notice that the standard linearized implicit Euler scheme has the lower-order s for all $s < \bar{s}$. Indeed, under the hypotheses of Theorem 3.6, then for all h small enough, $Q_\infty - Q_{\infty,\nu}^h$ is nonnegative and using Lemma 3.4 we only need to control

$$
\begin{aligned}
\text{Trace}\big(Q_\infty - Q_{\infty,\nu}^h\big) &= \frac{\sigma^2}{2} \sum_{p=1}^{+\infty} \frac{q_p}{\lambda_p + b_p} \frac{(\lambda_p + b_p)h}{2 + \lambda_p h + b_p h} \\
&= \frac{\sigma^2 h^s}{2} \sum_{p=1}^{+\infty} q_p \lambda_p^{-1+s} \frac{\lambda_p}{\lambda_p + b_p} \frac{\frac{\lambda_p + b_p}{\lambda_p^s} h^{1-s}}{2 + \lambda_p h + b_p h} \\
&\leq \frac{\sigma^2 h^s}{2} C_s \text{Trace}\big((-A)^{-1+s} Q\big).
\end{aligned}
$$

Thus for all $s < \bar{s}$ that there exists $C_s \in (0, +\infty)$ such that for all h small enough

$$
\tag{3.9}
\left| \int_{\mathcal{H}} \varphi d\mu_\infty - \int_{\mathcal{H}} \varphi d\nu_\infty^h \right| \leq C_s \sup_{v \in \mathcal{H}} \| D^2 \varphi(v) \| \sigma^2 h^s.
$$

It is also possible to prove that the above Theorem 3.6 (resp., (3.9)) gives optimal order of convergence, namely, that (3.8) does not hold true for all $h > 0$ if $s > \bar{s} + 1$ (resp., (3.9) does not hold true for any $h > 0$ if $s > \bar{s}$). This fact is also supported by the numerical simulations of Figure 3 in section 4.

3.4. Spatial regularity analysis. We show in this section that the postprocessed method yields a solution which has the same regularity in space as the exact solution, in contrast to the standard linearized implicit Euler method, which yields a solution that is too smooth. The action of the postprocessing thus not only increases the order of the convergence but also provides a qualitatively better approximation with the correct regularity. For all Borel probability measure μ on $\mathcal{H}$, we define its regularity, denoted $\text{reg}(\mu) \in \mathbb{R} \cup \{-\infty, +\infty\}$, by the supremum of s such that the norm $|\cdot|_s$ of $\mathcal{H}^s$ is square-integrable with respect to μ:

$$
\tag{3.10}
\text{reg}(\mu) = \sup \left\{ s \in \mathbb{R}, \int_{\mathcal{H}} |u|_s^2 \mu(du) < \infty \right\}.
$$

The interpretation in terms of random variables is the following. For a random variable v with values in $\mathcal{H}$, denoting $\mathbb{P}_v$ its probability law, we have the identity $\mathbb{E}(|v|_s^2) = \int_{\mathcal{H}} |u|_s^2 \mathbb{P}_v(du)$ which is a finite quantity if $s < \text{reg}(\mathbb{P}_v)$ and $+\infty$ if $s > \text{reg}(\mathbb{P}_v)$. Notice that instead of quantifying the regularity in terms of the Sobolev space $\mathcal{H}^s$, one could state similar results in terms of Hölder regularity. We refer to [9] for such a study of the θ-method applied to the stochastic heat equation with finite differences.

We focus for simplicity on the case $F = 0$, $\sigma = 1$, and with the initial condition $u_0 = 0$, but we emphasize that the extension to the general semilinear situation is straightforward, the regularity being determined only by the stochastic terms in our setting. First, notice that for the exact solution $u(t)$, the regularity parameter is

$\operatorname{reg}(\mathbb{P}_{u(t)}) = \operatorname{reg}(\mu_\infty) = \overline{s}$ for all $t > 0$ (see also Proposition 3.1). Indeed, using (1.3), and the Itô formula yields

$$\mathbb{E}|u(t)|_s^2 = \sum_{p=1}^{+\infty} \frac{q_p}{2\lambda_p^{1-s}} \left(1 - \exp(-2\lambda_p t)\right) \begin{cases} < +\infty & \text{if } s < \overline{s}, \\ = +\infty & \text{if } s > \overline{s}. \end{cases}$$

The following proposition shows that at the discrete-time level, the standard linearized implicit Euler method v_n in (1.6) and the method without processing u_n in (1.9) have the regularity parameter $\overline{s}+1$, whereas the postprocessor $\overline{u}_n$ in (1.10) has the correct regularity parameter $\overline{s}$.

PROPOSITION 3.7. *Consider* (1.1) *with* $F = 0$, $\sigma = 1$, $u_0 = 0$ *and assume* (1.2). *Then, for all* $h > 0$ *and all* $n \in \mathbb{N}^*$, $\operatorname{reg}(\mathbb{P}_{v_n}) = \operatorname{reg}(\mathbb{P}_{u_n}) = \operatorname{reg}(\nu_\infty^h) = \operatorname{reg}(\mu_\infty^h) = \overline{s}+1$, *whereas* $\operatorname{reg}(\mathbb{P}_{\overline{u}_n}) = \operatorname{reg}(\overline{\mu}_\infty^h) = \overline{s}$.

Proof. Inspecting the proof of Proposition 3.2, we have $u_n = \ell_n$ and for all $s < \overline{s} + 1, h > 0, n \in \mathbb{N}^*$

$$\mathbb{E}|\ell_n|_{s+1}^2 \leq \frac{C}{h} \sum_{p=1}^{+\infty} q_p \lambda_p^s \frac{(\lambda_p h)^2}{(1 + \lambda_p h)^2 - 1} \leq \frac{C}{h} \operatorname{Trace}\left((-A)^{s-1} Q\right);$$

this yields $\operatorname{reg}(\mathbb{P}_{u_n}) = \operatorname{reg}(\mu_\infty^h) \geq \overline{s} + 1$. The reverse inequality is obtained with a similar lower bound. The proof for the standard linearized implicit Euler scheme v_n is similar. Now, adding the postprocessing, $\overline{u}_n = u_n + \frac{1}{2}J_3\sigma\sqrt{h}\xi_n^Q$, we obtain using the definition of J_3,

$$\mathbb{E}|J_3\sigma\sqrt{h}\xi_n^Q|_s^2 = h\operatorname{Trace}\left(\left(I - \frac{h}{2}A\right)^{-1}(-A)^s Q\right)$$

$$= \sum_{p=1}^{+\infty} \frac{q_p}{\lambda_p^{1-s}} \frac{\lambda_p h}{1 + \frac{1}{2}\lambda_p h} \begin{cases} < +\infty & \text{if } s < \overline{s}, \\ = +\infty & \text{if } s > \overline{s}. \end{cases}$$

The term $\frac{1}{2}J_3\sigma\sqrt{h}\xi_n^Q$ thus has exactly the same regularity $\overline{s}$ as the exact solution. This concludes the proof of $\operatorname{reg}(\mathbb{P}_{\overline{u}_n}) = \operatorname{reg}(\overline{\mu}_\infty^h) = \overline{s}$. $\square$

4. Numerical experiments. In this section, we compare numerically the performances of the new postprocessed method (2.2) with the standard linearized implicit Euler method (1.6), and the trapezoidal method (1.11), in both finite and infinite dimensions.

We consider first in Figure 1 the scalar nonlinear SDE (2.3) with dimension $N = 1$, parameters $A = -1, \sigma = 1$, and the initial condition $X(0) = 0$. Taking $f_1(x) = Ax$ and $f_2(x) = f(x)$ in (2.2), we consider the nonlinearities $f(x) = -x - \sin(x)$ and $f(x) = -2x - x^3$, respectively, and we compute the averages over 10^{10} independent trajectories with final time $T = 1$ and compare for many timestep sizes the accuracy for $\mathbb{E}(\exp(-X(T)^2)) = \int_{-\infty}^{+\infty} \exp(-x^2)\rho(x)dx$. The final time $T = 1$ is chosen large enough so that the equilibrium is reached and the exponentially decaying term $e^{-\lambda T}$ in (2.19) is negligible. In the left picture of Figure 1 where the nonlinearity $f(x) = -x - \sin(x)$ is Lipschitz, we observe as shown in Theorem 2.4 the expected order 2 of convergence for the new method, while the standard methods exhibit order 1 of convergence (see the reference lines with slopes $1, 2$). Although our analysis in section 2.2 applies only to globally Lipschitz vector fields, we observe that the excellent
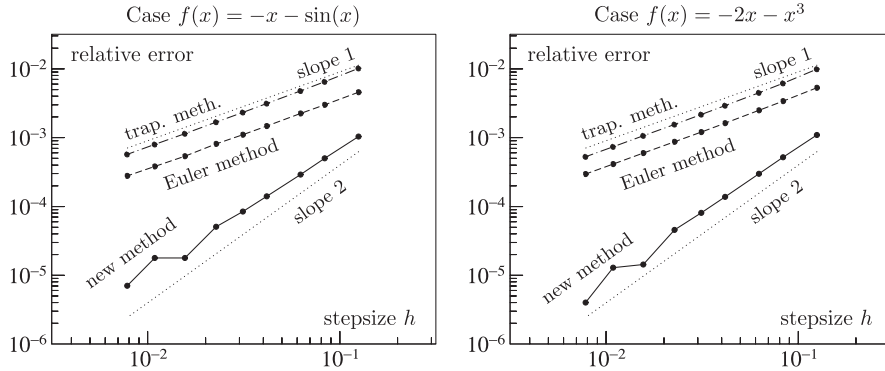
FIG. 1. *Comparison of the new method (solid lines) with the standard linearized implicit Euler method (dashed lines) and the trapezoidal method (dashed-dotted lines) for the scalar SDE (2.3) (dimension $N = 1$) with $A = -1, \sigma = 1$, nonlinearity $f(x)$. Error for $\mathbb{E}(\exp(-X(T)^2))$ at final time $T = 1$ versus the stepsize h, where $1/h = 8, 12, 16, 24, 32, 44, 64, 92, 128$. Averages over 10^{10} samples.*

performances of the new method persist also in the example with the non-Lipschitz nonlinearity $f(x) = -2x - x^3$ and the globally bounded test function $\phi(x) = \exp(-x^2)$ (right picture of Figure 1).

We next consider a standard finite difference approximation $U_j(t) \simeq u(j\Delta x, t)$ of the one-dimensional heat equation (1.4) with zero Dirichlet boundary conditions on a uniform grid with size $\Delta x = 1/(N + 1)$. This yields the following system of SDEs in dimension N:

$$dX(t) = \frac{1}{\Delta x^2} \begin{pmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 \end{pmatrix} \begin{pmatrix} X^1(t) \\ X^2(t) \\ \vdots \\ X^N(t) \end{pmatrix} dt + \begin{pmatrix} f(X^1(t)) \\ f(X^2(t)) \\ \vdots \\ f(X^N(t)) \end{pmatrix} dt + \frac{1}{\sqrt{\Delta x}} \begin{pmatrix} dW^1(t) \\ dW^2(t) \\ \vdots \\ dW^N(t) \end{pmatrix},$$

where $W^1, \ldots, W^N$ are independent one-dimensional standard Wiener processes.

Considering $N = 100$ grid points for the space discretization, we take the initial condition $u(x, 0) = \sin(2\pi x)$ and plot in Figure 2 a sample trajectory on the time interval $(0, 1)$ for the standard linearized implicit Euler method (1.6) and the new method with postprocessor $\overline{X}_n$, using the same sets of generated random numbers for both methods. We observe that the solution of the standard linearized implicit Euler method (Figure 2(a)) is qualitatively too smooth compared to the new method (Figure 2(b)) with the postprocessor applied at each timestep, which corroborates the statement of Proposition 3.7 in section 3.4. Notice that the trajectory for the new method without applying the postprocessor would look very similar to that of the standard linearized implicit Euler method. The spatial regularity observed in Figure 2(b) is the same as the one of a diffusion process driven by Brownian motion, conditioned to be zero at initial and final times. This property is not surprising; see [17, 16] for a study of the link between the law of a conditioned diffusion and invariant distributions of SPDEs.

We finally plot in Figure 3 the error at final time $T = 1$ for the quantity $\mathbb{E}(\exp(-\|u\|_{L^2(0,1)}^2))$, where we use the approximation $\|u\|_{L^2(0,1)}^2 \simeq \Delta x \sum_{j=1}^{N}(X^j)^2$. We use arbitrarily the initial condition $u(x, 0) = 0$ and we compute the averages
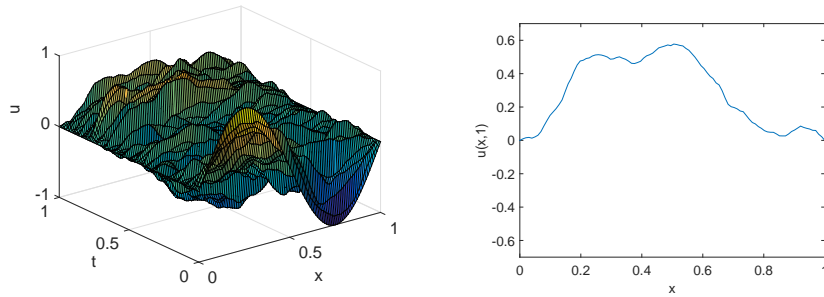
(a) Standard linearized Euler method, sample trajectory $u(x,t)$ and corresponding profile at final time $t = 1$.
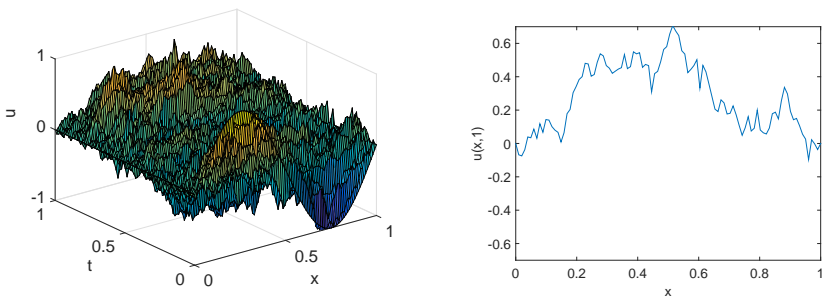
(b) New method with postprocessor, sample trajectory $u(x,t)$ and corresponding profile at final time $t = 1$.

FIG. 2. *Samples of realization of the stochastic nonlinear heat equation* (1.4) *with $f(u) = -u - \sin(u)$ using the standard linearized implicit Euler method and the new method with postprocessor. With $N = 100$ space grid points and timestep size $h = 1/100$.*

over 10^9 independent trajectories so that the Monte Carlo errors become negligible. The reference solution is computed using the new method with stepsize $h = 1/512$. We consider, respectively, the cases of various nonlinearities $f(u) = 0$, $f(u) = -u$, $f(u) = -u - \sin(u)$, $f(u) = -2u - u^3$. We observe in all cases an order of convergence $1/2$ for the standard linearized implicit Euler method, while the trapezoidal method has order 1. In contrast to the standard linearized implicit Euler method, the new method with postprocessor has a much better accuracy by a factor $15-250$ in the range of stepsizes considered. It has a zero bias for $f(u) = 0$ and order $3/2$ in the linear case $f(u) = -u$, as proved in Theorem 3.6. We observe that the order of convergence persists in the nonlinear case $f(u) = -u - \sin(u)$, and the excellent accuracy persists in the non-Lipschitz case $f(u) = -2u - u^3$, although an order reduction can be observed.

Appendix.

Proof of Lemma 2.5. We consider the variant (2.20) of the linearized implicit Euler method. A straightforward calculation yields the following weak expansion for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$:

$$(4.1) \qquad \mathbb{E}(\phi(Y_1)|Y_0 = x) = \phi(x) + h\mathcal{L}\phi(x) + h^2\mathcal{A}_1\phi(x) + \mathcal{O}(h^3),$$
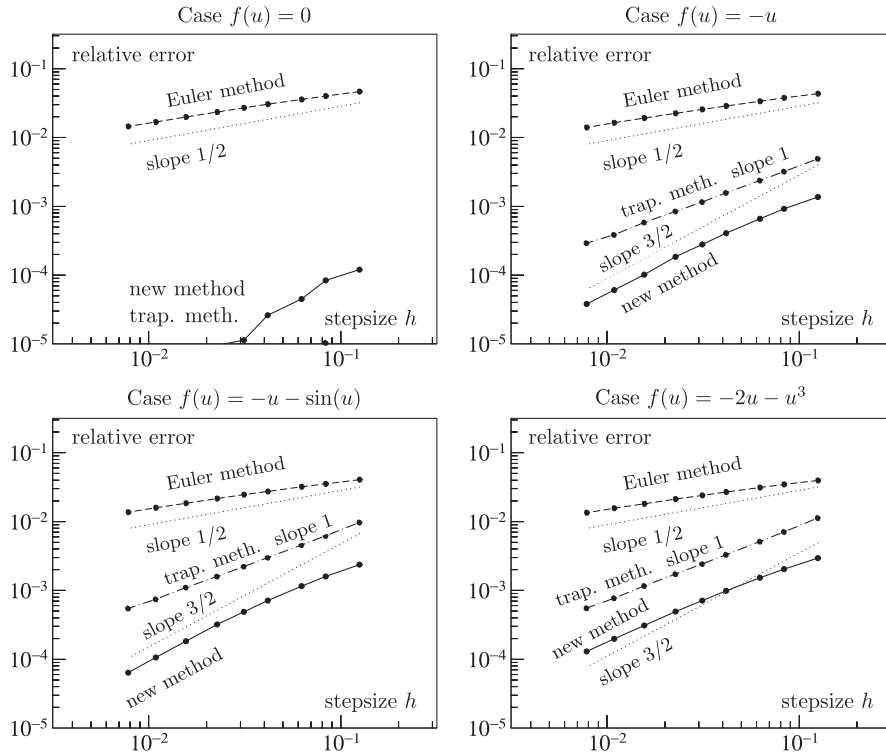
FIG. 3.   *Comparison of the new method (solid lines) with the standard linearized implicit Euler method (dashed lines) and the trapezoidal method (dashed-dotted lines) for the stochastic heat equation (1.4) with nonlinearity $f(u)$ discretized in space with $N = 100$ grid points. Error for $\mathbb{E}(\exp(-\|u(T)\|^2))$ at final time $T = 1$ versus the stepsize h, where $1/h = 8, 12, 16, 24, 32, 44, 64, 92, 128$. Averages over 10^9 samples.*

where the constant symbolized by $\mathcal{O}$ is independent of h but depends on ϕ and depends on x with a polynomial growth. Here, the fourth-order linear differential operator $\mathcal{A}_1$ is given by[1]

$$
\mathcal{A}_1 \phi = \frac{1}{2} \phi''(f_0, f_0) + \frac{\sigma^2}{2} \sum_{i=1}^{N} \phi'''(e_i, e_i, f_0) + \frac{\sigma^4}{8} \sum_{i,j=1}^{N} \phi^{(4)}(e_i, e_i, e_j, e_j) + \phi' f_1' f_0
$$

$$
+ \left((a_1 + 1)^2\right) \frac{\sigma^2}{2} \phi' \sum_{i=1}^{N} f_1''(e_i, e_i) + \left(a_1 + 1 + a_3\right) \sigma^2 \sum_{i=1}^{N} \phi''(f_1' e_i, e_i)
$$

$$
+ a_2^2 \frac{\sigma^2}{2} \phi' \sum_{i=1}^{N} f_2''(e_i, e_i) + a_2 \sigma^2 \sum_{i=1}^{N} \phi''(f_2' e_i, e_i),
$$

where $e_1, \ldots, e_N$ denotes the canonical basis of $\mathbb{R}^N$.

For any $\psi \in \mathcal{C}_P^\infty(\mathbb{R}^N, \mathbb{R})$, we denote $\langle \psi \rangle = \int_{\mathbb{R}^N} \psi(y) \mu_\infty(dy) = \int_{\mathbb{R}^N} \psi(y) \rho(y) dy$.

---

[1] We denote $\phi'(x) : \mathbb{R}^N \to \mathbb{R}$ the first derivative of ϕ at point $x \in \mathbb{R}^N$, $\phi''(x)$ the second derivative (a symmetric bilinear form on $\mathbb{R}^N \times \mathbb{R}^N$), $\phi'''(x)$ the third derivative (a symmetric trilinear form), etc.

By integration by parts, and using the identity $\nabla \rho = \frac{2}{\sigma^2}\rho f_0$, we obtain

$$\langle \phi''(f_0, f_0)\rangle = \langle -\phi'(f_0' f_0 + (\operatorname{div} f_0)f_0 + \tfrac{2}{\sigma^2}\|f_0\|^2 f_0)\rangle,$$

$$\langle \sigma^2 \textstyle\sum_i \phi'''(f_0, e_i, e_i)\rangle = \langle \phi'(\sigma^2 \textstyle\sum_i f_0''(e_i, e_i) + 4f_0'f_0 + 2(\operatorname{div} f_0)f_0 + \tfrac{4}{\sigma^2}\|f_0\|^2 f_0)\rangle,$$

$$\left\langle \sigma^2 \textstyle\sum_{ij} \phi^{(4)}(e_i, e_i, e_j, e_j)\right\rangle = \langle -\textstyle\sum_i 2\phi'''(f_0, e_i, e_i)\rangle,$$

$$(4.2) \quad \langle \sigma^2 \textstyle\sum_i \phi''(f_1' e_i, e_i)\rangle = \langle -\phi'(\sigma^2 \textstyle\sum_i f_1''(e_i, e_i) + 2f_1' f_0)\rangle;$$

see [1, 38] for examples of such calculations. Then

$$\langle \mathcal{A}_1 \phi\rangle = \Bigg\langle \Big(a_1^2 + 2a_1\Big)\frac{\sigma^2}{2}\phi' \sum_{i=1}^N f_1''(e_i, e_i) + \Big(\frac{1}{4} + a_1 + a_3\Big)\sigma^2 \sum_{i=1}^N \phi''(f_1' e_i, e_i)$$

$$+ a_2^2 \frac{\sigma^2}{2}\phi' \sum_{i=1}^N f_2''(e_i, e_i) + \Big(-\frac{1}{4} + a_2\Big)\sigma^2 \sum_{i=1}^N \phi''(f_2' e_i, e_i)\Bigg\rangle.$$

We note that the postprocessor $Y_n \mapsto \overline{Y}_n$ in (2.20) satisfies (2.15) with $q = 1$ and $\overline{\mathcal{A}}_1 \phi = b_1 \phi' f_1 + b_2 \phi' f_2 + \frac{c^2}{2}\sigma^2 \Delta \phi,$

$$[\mathcal{L}, \overline{\mathcal{A}}_1]\phi = (b_1 - b_2)\phi'(f_2' f_1 - f_1' f_2) - \frac{c^2}{2}\sigma^2 \phi' \sum_{i=1}^N f_0''(e_i, e_i) - c^2 \sigma^2 \sum_{i=1}^N \phi''(f_0' e_i, e_i).$$

We deduce the following expression for $\langle \mathcal{A}_1 \phi + [\mathcal{L}, \overline{\mathcal{A}}_1]\phi\rangle$, which is the key quantity to compute in Theorem 2.3. Summing up, we obtain

$$\langle \mathcal{A}_1 \phi + [\mathcal{L}, \overline{\mathcal{A}}_1]\phi\rangle$$

$$= \Bigg\langle \Big(a_1^2 + 2a_1 + b_1 - c^2\Big)\frac{\sigma^2}{2}\phi' \sum_{i=1}^N f_1''(e_i, e_i) + \Big(a_1 + a_3 + \frac{1}{4} + b_1 - c^2\Big)\sigma^2 \sum_{i=1}^N \phi''(f_1' e_i, e_i)$$

$$+ \Big(a_2^2 + b_2 - c^2\Big)\frac{\sigma^2}{2}\phi' \sum_{i=1}^N f_2''(e_i, e_i) + \Big(-\frac{1}{4} + a_2 + b_2 - c^2\Big)\sigma^2 \sum_{i=1}^N \phi''(f_2' e_i, e_i)$$

$$+ (b_1 - b_2)\phi'(f_2' f_1 - f_1' f_2)\Bigg\rangle.$$

The above quantity is zero by assumption for all test function ϕ. This means $(\mathcal{A}_1 + [\mathcal{L}, \overline{\mathcal{A}}_1])^* \rho = 0$. Applying Theorem 2.3 with $q = 1$ to the scheme (2.20) then yields for $\overline{Y}_n$ in (2.20) a method of weak order two for the invariant measure. This concludes the proof of Lemma 2.5. □

REFERENCES

[1] A. ABDULLE, G. VILMART, AND K. C. ZYGALAKIS, *High order numerical approximation of the invariant measure of ergodic SDEs*, SIAM J. Numer. Anal., 52 (2014), pp. 1600–1622.

[2] A. ABDULLE, G. VILMART, AND K. C. ZYGALAKIS, *Long time accuracy of Lie-Trotter splitting methods for Langevin dynamics*, SIAM J. Numer. Anal., 53 (2015), pp. 1–16.

[3] N. BOU-RABEE AND H. OWHADI, *Long-run accuracy of variational integrators in the stochastic context*, SIAM J. Numer. Anal., 48 (2010), pp. 278–297.

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

[5] E. Buckwar and C. Kelly, *Towards a systematic linear stability analysis of numerical methods for systems of stochastic differential equations*, SIAM J. Numer. Anal., 48 (2010), pp. 298–321.

[6] E. Buckwar, M. G. Riedler, and P. E. Kloeden, *The numerical stability of stochastic ordinary differential equations with additive noise*, Stoch. Dyn., 11 (2011), pp. 265–281.

[7] K. Burrage, P. Burrage, and T. Tian, *Numerical methods for strong solutions of stochastic differential equations: An overview*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 460 (2004), pp. 373–402.

[8] J. C. Butcher, *The effective order of Runge-Kutta methods*, in Proceedings of Conference on the Numerical Solution of Differential Equations, J. L. Morris, ed., Lecture Notes in Math. 109, Springer, New York, 1969, pp. 133–139.

[9] Y. Chong and J. B. Walsh, *The roughness and smoothness of numerical solutions to the stochastic heat equation*, Potential Anal., 37 (2012), pp. 303–332.

[10] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite-Dimensional Systems*, London Math. Soc. Lecture Note Ser. 229, Cambridge University Press, Cambridge, UK, 1996.

[11] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, 2nd ed., Encyclopedia Math. Appl. 152, Cambridge University Press, Cambridge, UK, 2014.

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

[13] A. Debussche, *Ergodicity results for the stochastic Navier-Stokes equations: An introduction*, in Topics in Mathematical Fluid Mechanics, Lecture Notes in Math. 2073, Springer, New York, 2013, pp. 23–108.

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

[15] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I. *Nonstiff Problems*, 2nd ed., Springer Ser. Comput. Math. 8, Springer, Berlin, 1993.

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

[17] 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 (2005), pp. 587–603.

[18] R. Hasminskii, *Stochastic Stability of Differential Equations*, Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands, 1980.

[19] D. J. Higham, *Mean-square and asymptotic stability of the stochastic theta method*, SIAM J. Numer. Anal., 38 (2000), pp. 753–769.

[20] A. Jentzen, *Higher order pathwise numerical approximations of SPDEs with additive noise*, SIAM J. Numer. Anal., 49 (2011), pp. 642–667.

[21] A. Jentzen, P. Kloeden, and G. Winkel, *Efficient simulation of nonlinear parabolic SPDEs with additive noise*, Ann. Appl. Probab., 21 (2011), pp. 908–950.

[22] A. Jentzen and P. E. Kloeden, *Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 465 (2009), pp. 649–667.

[23] A. Jentzen and R. Kurniawan, *Weak Convergence Rates for Euler-Type Approximations of Semilinear Stochastic Evolution Equations with Nonlinear Diffusion Coefficients*, arXiv: 1501.03539, 2015.

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

[25] 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 2014.

[26] T. Li, A. Abdulle, and W. E, *Effectiveness of implicit methods for stiff stochastic differential equations*, Commun. Comput. Phys., 3 (2008), pp. 295–307.

[27] G. J. Lord and T. Shardlow, *Postprocessing for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal., 45 (2007), pp. 870–889.

[28] G. Milstein, *Weak approximation of solutions of systems of stochastic differential equations*, Theory Probab. Appl., 30 (1986), pp. 750–766.

[29] G. Milstein and M. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Sci. Comput., Springer-Verlag, Berlin, 2004.

[30] G. Milstein and M. Tretyakov, *Computing ergodic limits for Langevin equations*, Phys. D, 229 (2007), pp. 81–95.

[31] G. N. Milstein and M. V. Tretyakov, *Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients*, SIAM J. Numer. Anal., 43 (2005), pp. 1139–1154.

[32] A. Rathinasamy and K. Balachandran, *Mean-square stability of second-order Runge-Kutta methods for multi-dimensional linear stochastic differential systems*, J. Comput. Appl. Math., 219 (2008), pp. 170–197.

[33] Y. Saito and T. Mitsui, *Stability analysis of numerical schemes for stochastic differential equations*, SIAM J. Numer. Anal., 33 (1996), pp. 2254–2267.

[34] Y. Saito and T. Mitsui, *Mean-square stability of numerical schemes for stochastic differential systems*, Vietnam J. Math., 30 (2002), pp. 551–560.

[35] D. Talay, *Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE and applications*, in Filtering and Control of Random Processes Lecture Notes in Control and Inform. Sci. 61, Springer, New York, 1984, pp. 294–313.

[36] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl., 8 (1990), pp. 483–509.

[37] A. Tocino, *Mean-square stability of second-order Runge-Kutta methods for stochastic differential equations*, J. Comput. Appl. Math., 175 (2005), pp. 355–367.

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

[39] X. Wang and S. Gan, *Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise*, J. Math. Anal. Appl., 398 (2013), pp. 151–169.