Cubature methods for stochastic (partial) differential equations in weighted spaces

**Philipp Dörsek · Josef Teichmann ·
Dejan Velušček**

Abstract The cubature on Wiener space method, a high-order weak approximation scheme, is established for SPDEs in the case of unbounded characteristics and unbounded test functions. We first describe a recently introduced flexible functional analytic framework, so called weighted spaces, where Feller-like properties hold. A refined analysis of vector fields on weighted spaces then yields optimal convergence rates of cubature methods for stochastic partial differential equations of Da Prato–Zabczyk type. The ubiquitous stability for the local approximation operator within the functional analytic setting is proved for SPDEs, however, in the infinite dimensional case we need a newly introduced technical assumption on weak symmetry of the cubature formula. Computational results for a cubature discretization of a spatially extended stochastic FitzHugh–Nagumo model, an SPDE model from mathematical biology, are shown, illustrating the applicability of our theory.

P. Dörsek · J. Teichmann (✉)
D-MATH, ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland
e-mail: jteichma@math.ethz.ch

P. Dörsek
e-mail: philipp.doersek@gmail.com

D. Velušček
Faculty of Mathematics and Physics, University of Ljubljana,
Jadranska 19, 1000 Ljubljana, Slovenia
e-mail: dejan.veluscek@fmf.uni-lj.si

1 Introduction

Cubature on Wiener space, a realization of the abstract KLV high order method after Kusuoka [21] and Lyons and Victoir [24], is a weak approximation scheme for stochastic differential equations. Significant advantages in comparison to other weak approximation schemes such as Taylor methods, see Kloeden and Platen [20], are that it respects the geometry of the problem, and that at least theoretically, it is possible to reach arbitrarily high rates of convergence without requiring the calculation of higher derivatives, see Lyons and Victoir [24, Theorem 2.4, Prop. 2.5]. The concrete construction of such cubature paths of high order is still quite difficult, see Gyurkó and Lyons [15] for cubature formulas (on the Lie algebra level) up to degree 11 for a single driving Brownian motion. Cubature schemes provide a time-discretization approximating the unknown expected value of a functional of the solution process of the SPDE by an expectation of an iteratively constructed function on a high-dimensional discrete product space. Often a direct evaluation of the functional on the discrete probability space is too expensive, therefore, several methods to speed-up the evaluation of cubature schemes such as recombination, see Litterer and Lyons [23] and Schmeiser et al. [31], or tree-based branching, see Crisan and Lyons [5], have emerged. Alternatively, the functionals can be evaluated using Monte Carlo or Quasi Monte Carlo algorithms on the discrete product probability space.

The combination with Quasi Monte Carlo integration algorithms makes high-order weak approximation schemes a particularly interesting alternative to standard multi-level Monte Carlo schemes. Indeed, multi-level Monte Carlo schemes lead to complexity estimates of order $\mathcal{O}(\epsilon^{-2})$, i.e., a number of operations of order ϵ^{-2} is necessary to reach accuracy ϵ, see Giles [13] and Giles and Szpruch [14]. In contrast, QMC evaluations of weak, high-order approximation schemes of order k lead to complexity estimates of order (almost) $\mathcal{O}(\epsilon^{-1-1/k})$, as long as the QMC integration yields optimal convergence (this in turn also depends on the dimension of integration space, which is moderate for high order methods). We therefore believe that it is worth analysing the functional analytic framework of cubature schemes in depth, i.e., we aim to construct a pool of Banach spaces of test functions and Banach spaces of characteristics flexible enough to embed relevant problems from practice.

In this work, we shall relax the regularity assumptions of the cubature method, similarly as was done in Dörsek [9] and Dörsek and Teichmann [11,12] for the splitting approach of Ninomiya and Victoir [26]. Consider a stochastic differential equation on $\mathbb{R}^n$ in its Stratonovich form,

$$\mathrm{d}X_t^x = \sum_{j=0}^d V_j(X_t^x) \circ \mathrm{d}B_t^j . \tag{1.1}$$

In this article $(B_t^j)_{t\geq 0}$, for $j = 1, \ldots, d$, always denotes the j-th component of a d-dimensional Brownian motion, running time is encoded for convenience by the zeroth component $B_t^0 = t$, for $t \geq 0$, and usually $V_j : \mathbb{R}^n \to \mathbb{R}^n$, for $j = 0, \ldots, d$ denote vector fields. A deterministic initial value is supposed, i.e. $X_0^x = x$, for $x \in \mathbb{R}^n$. All initial work was based on the fundamental assumption that the vector fields

$V_j : \mathbb{R}^n \to \mathbb{R}^n$ are bounded and C^∞-bounded. This is also a typical assumption in other approximation methods for stochastic differential equations, e.g., in Talay and Tubaro [32]. Some success in relaxing these assumptions, which are rarely satisfied in practical problems, was achieved for approximations of the splitting type in Alfonsi [1] and Tanaka and Kohatsu-Higa [33]. While the first one focuses on the CIR process and the second one on Lévy driving noise, it was recognised in both works that polynomially bounded test functions are the correct context for problems with Lipschitz continuous vector fields.

Another approach was taken in Dörsek and Teichmann [11]. There, splitting schemes were analyzed on general weighted spaces, allowing in particular the approximation of Da Prato–Zabczyk stochastic partial differential equations where the drift part, the infinitesimal generator of a strongly continuous semigroup on the infinite dimensional state space, is not even continuous.

All these approaches profited from the special structure of splitting schemes, as there the stability or power boundedness of the discrete approximation operator can be shown by investigating every part separately. Instead, we follow a similar idea as was applied to the stochastic Navier–Stokes equations in Dörsek [10]. We extend the results of Bayer and Teichmann [2], where strong conditions are imposed on the vector fields, to more general coefficients and test functions. This allows us to obtain methods of order higher than 2 without having to resort to extrapolation, see Blanes and Casas [3] and Oshima et al. [27]. The weighted spaces developed originally in Röckner and Sobol [29] and used for the numerical analysis of weak approximation methods in Dörsek and Teichmann [11] are a suitable tool for our needs, and we provide a refined analysis of the vector fields defined on these spaces. This allows us to do a Taylor expansion of the cubature approximations to compute the local approximation order.

We use two different approaches for proving stability. In the finite dimensional case with sufficiently smooth vector fields, the Gronwall inequality yields the claim in a straightforward manner under a reasonable assumption of compatibility between the vector fields and the weight function. In the infinite dimensional case, we apply the method of the moving frame from Teichmann [35]. This leads to time dependent vector fields that are nonsmooth in the time component. As this makes a Taylor expansion impossible, we introduce a *weak symmetry condition* on cubature paths, a technical assumption usually satisfied by cubature schemes. This allows us to obtain stability not only for Da Prato–Zabczyk equations with pseudocontractive generator, but also for stochastic differential equations on infinite dimensional state spaces where the vector fields depend roughly, i.e., continuously, but not differentiably, on time.

As a numerical example, we consider a *cubature discretization of degree $m = 5$* corresponding to weak convergence order 2 of a stochastic FitzHugh–Nagumo model, extending the stochastic model from Tang et al. [34] spatially using methods from Tuckwell [36]. We formulate the necessary functional analytic setting, and show results of the numerical simulation of two quantities of interest that correspond to purely noise induced state changes. Our computations establish the effectivity of the proposed algorithm. It is worth pointing out that the cubature discretization of stochastic FitzHugh–Nagumo is *locally deterministic*, i.e., we can perform local random walk like studies within this model setting that approximate the original model of high

order 2. This might be very useful in applications. To the best of our knowledge such high order random walk approximations of FitzHugh–Nagumo have been unknown.

There are many successful discretisation schemes for stochastic partial differential equations that can be applied to the stochastic FitzHugh–Nagumo model and similar equations. Jentzen and Kloeden [17] give an overview of strong and pathwise schemes. Weak approximation schemes are more difficult. Recently, it was proved in Debussche [8] that an implicit Euler scheme converges with weak rate almost 1/2 for equations driven by space–time white noise, doubling the corresponding strong rate of convergence; see also the references in Debussche [8] for more background on weak approximation schemes for stochastic partial differential equations with space–time white noise. In contrast, we restrict ourselves to finite-dimensional driving noise, but obtain the same optimal high order weak convergence as for finite-dimensional state spaces. Additionally we obtain by the cubature method itself deterministic construction recipes of typical trajectories of the model.

The paper is organised as follows. Sections 2 and 3 contain an exposition of the theory of $\mathcal{B}^\psi$ spaces, originally introduced in Röckner and Sobol [29], and explain how directional derivatives can be analysed in this setting, extending results from Dörsek and Teichmann [11,12]. In Sect. 4, we prove stability of cubature schemes in $\mathcal{B}^\psi$ spaces under various assumptions on the vector fields, using a technical assumption of weak symmetry of the underlying cubature paths in the infinite-dimensional case. Section 5 is devoted to the convergence proofs of cubature schemes. Finally, in Sect. 6, we present numerical results for a spatially extended stochastic FitzHugh–Nagumo model.

2 $\mathcal{B}^\psi$ spaces

We recall the following definition of spaces of functions with controlled growth, see also Röckner and Sobol [29] for their use in the construction of the solution of martingale problems in infinite dimension, and Dörsek [9,10] and Dörsek and Teichmann [11,12] for their application to the analysis of splitting schemes for stochastic partial differential equations.

Definition 2.1 Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space, and $\varphi\colon X \to (0, \infty)$ be bounded from below by some $\delta > 0$. For a Banach space $(Y, \|\cdot\|_Y)$, we set

$$\mathrm{B}^\varphi(X; Y) := \left\{ f\colon X \to Y\colon \sup_{x \in X} \varphi(x)^{-1} \|f(x)\|_Y < \infty \right\}, \tag{2.1}$$

endowed with the φ-norm

$$\|f\|_\varphi := \sup_{x \in X} \varphi(x)^{-1} \|f(x)\|_Y. \tag{2.2}$$

Let $k \geq 0$. If $\varphi = (\varphi_j)_{j=0,\ldots,k}$, $\varphi_j \colon X \to (0,\infty)$ bounded from below by some $\delta > 0$, $j = 0, \ldots, k$, we set

$$\mathrm{B}_k^\varphi(X; Y) := \left\{ f \in \mathrm{C}^k(X; Y) \colon \sup_{x \in X} \varphi_j(x)^{-1} \|D^j f(x)\|_{L_j(X;Y)} < \infty \right.$$
$$\left. \text{for } j = 0, \ldots, k \right\}. \tag{2.3}$$

$\mathrm{B}_k^\varphi(X; Y)$ is endowed with the norm

$$\|f\|_{\varphi,k} := \|f\|_{\varphi_0} + \sum_{j=1}^k |f|_{\varphi_j,j}, \tag{2.4}$$

where the seminorms $|\cdot|_{\varphi_j,j}$ are given by

$$|f|_{\varphi_j,j} := \sup_{x \in X} \varphi_j(x)^{-1} \|D^j f(x)\|_{L_j(X;Y)}. \tag{2.5}$$

Here, $L_j(X; Y)$ denotes the space of bounded multilinear forms $a \colon X^j \to Y$, and is endowed with the norm

$$\|a\|_{L_j(X;Y)} := \sup_{\|h_i\| \leq 1, i=1,\ldots,j} \|a(h_1, \ldots, h_j)\|_Y. \tag{2.6}$$

For simplicity, we set $L_0(X; Y) := Y$; we remark that $L_1(X; Y)$ is the space of bounded linear operators $X \to Y$, and in this case, the above norm is the usual operator norm. If $Y = \mathbb{R}$, we define $\mathrm{B}^\varphi(X) := \mathrm{B}^\varphi(X; \mathbb{R})$ and $\mathrm{B}_k^\varphi(X) := \mathrm{B}_k^\varphi(X; \mathbb{R})$.

Definition 2.2 Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space. A function φ is called *admissible weight function* if and only if $\varphi \colon X \to (0,\infty)$ is such that $K_R := \{x \in X \colon \varphi(x) \leq R\}$ is weak-$*$ compact for all $R > 0$ and bounded from below by some $\delta > 0$.

It is called *D-admissible weight function* if and only if it is an admissible weight function and for every $x \in X$, there exists some $R > 0$ such that $B_\varepsilon(x) \subset K_R$ for some $\varepsilon > 0$, where $B_\varepsilon(x) := \{y \in X \colon \|y - x\|_X \leq \varepsilon\}$ is the closed ε-ball around x.

It is called *C-admissible weight function* if and only if φ is bounded from below by some $\delta > 0$, weak-$*$ lower semicontinuous, and if for every $x \in X$, there exists some $\varepsilon > 0$ such that φ is bounded on $B_\varepsilon(x)$.

Remark 2.3 We do not require C-admissible weight functions to be admissible. However, φ is D-admissible if and only if it is admissible and C-admissible.

Theorem 2.4 *Let $k \in \mathbb{N}$, and assume that $\varphi = (\varphi_j)_{j=0,\ldots,k}$ is a vector of C-admissible weight functions. Then, $\mathrm{B}_k^\varphi(X; Y)$ is a Banach space.*

Proof Let $(f_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in this space. It is clear that f_n admits a pointwise limit f. Moreover, it follows that for every $x \in X$ and every closed ε-ball $B_\varepsilon(x)$, $f_n|_{B_\varepsilon(x)}$ are Cauchy sequences in $C^k(B_\varepsilon(x); Y)$. But this entails that $f|_{B_\varepsilon(x)} \in C^k(B_\varepsilon(x); Y)$. As differentiability is a local property, we see that $f \in C^k(X; Y)$. The necessary estimates for f and its derivatives are now easy to see. $\square$

Definition 2.5 Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space, its predual being W, $X = W^*$, and $(Y, \|\cdot\|_Y)$ a Banach space. The space of *bounded smooth cylindrical functions* is defined by

$$\mathcal{A}(X, Y) := \Big\{ f \colon X \to Y \colon f = g(\langle \cdot, w_1 \rangle, \ldots, \langle \cdot, w_n \rangle) \\ \text{for some } g \in C_b^\infty(\mathbb{R}^n; Y), \\ w_i \in W, i = 1, \ldots, n, n \in \mathbb{N} \Big\}. \tag{2.7}$$

Here, $\langle \cdot, \cdot \rangle$ denotes the dual pairing of X and W and $C_b^\infty(\mathbb{R}^n; Y)$ the space of infinitely often differentiable functions from Y to $\mathbb{R}$ that are bounded with all their derivatives bounded. For $Y = \mathbb{R}$, we set $\mathcal{A}(X) := \mathcal{A}(X, \mathbb{R})$.

Definition 2.6 Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space and $(Y, \|\cdot\|_Y)$ be a Banach space. Let ψ be an admissible weight function on X.

The space $\mathcal{B}^\psi(X; Y)$ is the closure of $\mathcal{A}(X, Y)$ in $B^\psi(X; Y)$. For $Y = \mathbb{R}$, we set $\mathcal{B}^\psi(X) := \mathcal{B}^\psi(X; \mathbb{R})$.

Remark 2.7 [11, Theorem 4.2] shows that our definition of $\mathcal{B}^\psi(X)$ here agrees with our earlier definition from Dörsek and Teichmann [11, Definition 2.2]. Due to Dörsek and Teichmann [11, Theorem 2.7], the functions in $\mathcal{B}^\psi(X)$ are characterized by the property that both $f|_{K_R} \in C((K_R)_{w*})$ and

$$\lim_{R \to \infty} \sup_{x \in X \setminus K_R} \psi(x)^{-1} |f(x)| = 0. \tag{2.8}$$

Definition 2.8 Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space and $(Y, \|\cdot\|_Y)$ be a Banach space. Let $\psi = (\psi_j)_{j=0,\ldots,k}$ with ψ_j D-admissible weight functions for $j = 0, \ldots, k$. The space $\mathcal{B}_k^\psi(X; Y)$ is the closure of $\mathcal{A}(X, Z)$ in $B_k^\psi(X; Y)$. For $Y = \mathbb{R}$, we set $\mathcal{B}_k^\psi(X) := \mathcal{B}_k^\psi(X; \mathbb{R})$. In particular, by Theorem 2.4, it follows that $\mathcal{B}_k^\psi(X)$ is a separable Banach space.

One essential property of $\mathcal{B}^\psi(X)$ spaces is that the dual space of this separable Banach space is a well understood space of Radon measures, such as in the case of $C_0(X)$ for locally compact spaces X. The following result follows from the theory of Röckner and Sobol [29] (see also Dörsek and Teichmann [11] and [12]).

Proposition 2.9 (Riesz representation for $\mathcal{B}^\psi(X)$) *Let $\ell \colon \mathcal{B}^\psi(X) \to \mathbb{R}$ be a continuous linear functional. Then, there exists a finite signed Radon measure μ on X such that*

$$\ell(f) = \int_X f(x) \mu(\mathrm{d}x) \quad \text{for all } f \in \mathcal{B}^\psi(X). \tag{2.9}$$

Furthermore,

$$\int_X \psi(x)|\mu|(\mathrm{d}x) = \|\ell\|_{L(\mathcal{B}^\psi(X),\mathbb{R})}, \tag{2.10}$$

where $|\mu|$ denotes the total variation measure of μ.

As every such measure defines a continuous linear functional on $\mathcal{B}^\psi(X)$, this completely characterizes the dual space of $\mathcal{B}^\psi(X)$. This allows for the introduction of the *generalized Feller property*, such that we can speak about strongly continuous semigroups on spaces of functions with growth controlled by ψ, in particular functions which are in general unbounded.

Let $(P_t)_{t\geq 0}$ be a family of bounded linear operators $P_t \colon \mathcal{B}^\psi(X) \to \mathcal{B}^\psi(X)$ with the following properties:

(F1) $P_0 = I$, the identity on $\mathcal{B}^\psi(X)$,
(F2) $P_{t+s} = P_t P_s$ for all $t, s \geq 0$,
(F3) for all $f \in \mathcal{B}^\psi(X)$ and $x \in X$, $\lim_{t\to 0+} P_t f(x) = f(x)$,
(F4) there exist a constant $C \in \mathbb{R}$ and $\varepsilon > 0$ such that for all $t \in [0, \varepsilon]$, $\|P_t\|_{L(\mathcal{B}^\psi(X))} \leq C$,
(F5) P_t is positive for all $t \geq 0$, that is, for $f \in \mathcal{B}^\psi(X)$, $f \geq 0$, we have $P_t f \geq 0$.

Alluding to Kallenberg [19, Chapter 17], such a family of operators will be called a *generalized Feller semigroup*. This is justified by the following result, which is a direct consequence of Lebesgue's dominated convergence theorem with respect to the measure existing due to Riesz representation. Its proof is given in Dörsek and Teichmann [11, Theorem 3.2] and [12, Corollary 4].

Proposition 2.10 *Let $(P_t)_{t\geq 0}$ satisfy* **F1** *to* **F4**. *Then, $(P_t)_{t\geq 0}$ is strongly continuous on $\mathcal{B}^\psi(X)$, that is,*

$$\lim_{t\to 0+} \|P_t f - f\|_\psi = 0 \quad \text{for all } f \in \mathcal{B}^\psi(X). \tag{2.11}$$

3 Vector fields and directional derivatives

When we ask for convergence rates we have to specify large enough sets of test functions within the basic $\mathcal{B}^\psi(X)$-spaces. For this purpose we need to analyze directional derivatives and their functional analytic behavior. This can be done within the setting of $\mathcal{B}_k^\psi(X; Y)$ spaces.

Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space. Given $(Z, \|\cdot\|_Z)$ the dual space of another separable Banach space that is embedded in X, we derive conditions on $V \colon Z \to X$ such that the *directional derivative $g \in \mathcal{B}_{k-1}^{\hat{\psi}}(Z)$, where*

$$g(z) := Df(z)(V(z)) \quad \text{for } z \in Z \tag{3.1}$$

and $f \in \mathcal{B}_k^\psi(X)$. Here, $\psi = (\psi_j)_{j=0,\dots,k}$ and $\hat{\psi} = (\hat{\psi}_j)_{j=0,\dots,k-1}$ are vectors of D-admissible weight functions on X and Z, respectively.

We shall assume that $V \in B_{k-1}^{\varphi}(Z; X)$ for some vector $\varphi = (\varphi_j)_{j=0,\dots,k-1}$ of C-admissible weight functions on Z. Then, V is $k-1$ times continuously Fréchet differentiable. As $f \in C^k(X)$, the Leibniz rule yields

$$D^j g(z)(h_1, \dots, h_j) = \sum_{i=0}^{j} \frac{1}{i!(j-i)!} \sum_{\sigma \in \mathcal{S}_j} g_{j,i}(z, h_{\sigma_1}, \dots, h_{\sigma_j}), \quad j = 0, \dots, k-1.$$

(3.2)

Here, $\mathcal{S}_j$ denotes the symmetric group with j elements, and

$$g_{j,i}(z, h_1, \dots, h_j) := D^{i+1} f(z)(h_1, \dots, h_i, D^{j-i} V(z)(h_{i+1}, \dots, h_j)).$$ (3.3)

In particular, if we assume that for some constant $C > 0$,

$$\hat{\psi}_j(z) \geq C^{-1} \sum_{i=0}^{j} \binom{j}{i} \psi_{i+1}(z) \varphi_{j-i}(z) \quad \text{for } j = 0, \dots, k-1,$$ (3.4)

it follows that $g \in B_{k-1}^{\hat{\psi}}(Z)$.

It is not so straightforward to prove that g can also be approximated by functions in $\mathcal{A}(X)$, which would imply $g \in \mathcal{B}_{k-1}^{\hat{\psi}}(Z)$. In Dörsek [9], a general theory for multiplication operators on $\mathcal{B}^{\psi}$ spaces is derived. Here, we take a different route, focusing on the problem at hand. The following definition is essential.

Definition 3.1 Given a Banach space $(X, \|\cdot\|_X)$ and the dual space $(Z, \|\cdot\|_Z)$ of a separable Banach space. Let $V \in B_k^{\varphi}(Z; X)$ with φ a given vector of C-admissible weight functions on Z. We say that $V \in \mathcal{C}_k^{\varphi}(Z; X)$ if and only if for every $y \in X^*$, there exists a constant $C_{V,y} > 0$ such that for each $R > 0$, there exists a sequence $(v_n)_{n \in \mathbb{N}} \subset \mathcal{A}(Z)$ (depending on $R > 0$ and $y \in X^*$) with $\sup_{n \in \mathbb{N}} \|v_n\|_{\varphi,k} \leq C_{V,y}$ such that, with $v := y \circ V$,

$$\lim_{n \to \infty} \|v - v_n\|_{C^k(B_R(0))} = 0.$$ (3.5)

Here, $B_R(0)$ is the closed ball of radius R in Z, and

$$\|g\|_{C^k(B_R(0))} := \sum_{j=0}^{k} \sup_{z \in B_R(0)} \|D^j g(z)\|_{L_j(Z)}.$$ (3.6)

Remark 3.2 It is clear that vector fields such as those from Dörsek and Teichmann [12, Sect. 3.1.2] satisfy the above assumption. More generally, if Z is a Hilbert space and is compactly embedded into a larger Hilbert space Y such that $y \circ V$ can be extended to a smooth mapping $Y \to \mathbb{R}$ lying in $C_b^k(Y; \mathbb{R})$ for all $y \in X^*$, then the above assumption is satisfied, i.e., $V \in \mathcal{C}_k^{\varphi}(Z; X)$ for every vector φ of C-admissible

weight functions on Z. Indeed, the extension of $y \circ V$ and its derivatives are continuous on Y, whence uniformly continuous on the compact set $B_R(0)$. Let us fix a sequence of increasing finite-dimensional, orthogonal projections $\pi_n \to \mathrm{id}_Y$ converging strongly to the identity: composing the extension of $y \circ V$ with π_n yields a pointwise converging, equicontinuous sequence of cylindrical function on $B_R(0)$, which is—up to a smoothing argument—the desired assertion.

Comparable arguments are also used in Dörsek and Teichmann [12, Theorem 5] and Dörsek [9, Theorem 2.39]. In particular, this implies that Nemytskii operators are included in our setup if Z is a Sobolev space of sufficiently smooth functions, see also Dörsek [9, Example 2.48].

This definition should also be compared to the form of the multiplicative noise suggested in Debussche [8, Remark 2.3]. It is similar in spirit to the definition of $\mathcal{C}_k^{\varphi}(H; H)$, as there, A is assumed to be a negative self-adjoint operator with a compact inverse. Hence, if we consider a single component of the noise, $x \mapsto \tilde{\sigma}((-A)^{-1/4}x)$, with $\tilde{\sigma} \colon H \to H$ a C^3-function with derivatives bounded up to order 3, it satisfies our assumptions given above and hence lies in $\mathcal{C}_3^{\varphi}(H)$ with $\varphi_0(x) := (1 + \|x\|_H^2)^{1/2}$ and $\varphi_j(x) := 1$, $j \geq 1$.

Theorem 3.3 *Fix $k \geq 1$. Let $\psi = (\psi_i)_{i=0,\dots,k}$ be a vector of D-admissible weight functions on X, and $\hat{\psi} = (\hat{\psi}_j)_{j=0,\dots,k-1}$ a vector of D-admissible and $\varphi = (\varphi_j)_{j=0,\dots,k-1}$ a vector of C-admissible weight functions on Z. Suppose (3.4).*

Then, the Lie derivative $\mathcal{L} \colon \mathcal{C}_{k-1}^{\varphi}(Z; X) \times \mathcal{B}_k^{\psi}(X) \to \mathcal{B}_{k-1}^{\hat{\psi}}(Z)$ defined through

$$\mathcal{L}(V, f)(z) := \mathcal{L}_V f(z) := Df(z)(V(z)) \tag{3.7}$$

is a bilinear, bounded operator.

Remark 3.4 Clearly, $V \in \mathcal{C}_k^{\varphi}(Z; X)$ is necessary if $\mathcal{L}_V f \in \mathcal{B}_k^{\hat{\psi}}(Z)$ is supposed to hold for $f \in \mathcal{B}_{k+1}^{\psi}(X)$ for a sufficiently large class of weight functions ψ. Indeed, choose $\psi_0(x) := \rho(\|x\|_X)$ with some increasing, left continuous and superlinear function ρ, and ψ_j arbitrary D-admissible weight functions on X. Then, $f := y \in \mathcal{B}_{k+1}^{\psi}(X)$ for all $y \in X^*$. Hence, $\mathcal{L}_V f(z) = y(V(z))$, and $y \circ V \in \mathcal{B}_k^{\hat{\psi}}(Z)$ implies $V \in \mathcal{C}_k^{\varphi}(Z; X)$.

Proof The claimed boundedness of $\mathcal{L}$ was remarked above, and follows straight away from (3.4).

Hence, we only need to prove that $\mathcal{L}_V f \in \mathcal{B}_{k-1}^{\hat{\psi}}(Z)$ for given $V \in \mathcal{C}_{k-1}^{\varphi}(Z; X)$ and $f = g(\langle \cdot, w_1 \rangle, \dots, \langle \cdot, w_n \rangle) \in \mathcal{A}(X)$; the result then follows from a density argument. Fix $\varepsilon > 0$. We shall construct $g_\varepsilon \in \mathcal{A}(Z)$ such that $\|\mathcal{L}_V f - g_\varepsilon\|_{\hat{\psi},k} < C\varepsilon$ with some constant $C > 0$ independent of ε.

Choose a dual set of vectors $(\zeta_i)_{i=1,\dots,n} \subset Z$ of $(w_i)_{i=1,\dots,n}$, i.e., $\langle \zeta_i, w_j \rangle = \delta_{ij}$. Let $Z_n := \mathrm{span}\{\zeta_i : i = 1, \dots, n\}$, and define $\pi \colon X \to Z_n$ by $\pi x := \sum_{i=1}^n \langle x, w_i \rangle \zeta_i$. Then, $f \circ \pi = f$, and

$$\mathcal{L}_V f(z) = \sum_{i=1}^n Df(z)(\zeta_i)\langle V(z), w_i \rangle. \tag{3.8}$$

Clearly, $w_i \in X^*$, and thus by Definition 3.1, there exists $C_V := \max_{i=1,\dots,n} C_{V,w_i} > 0$ such that for each $R > 0$, we can find $v_{R,\varepsilon}^i \in \mathcal{A}(Z)$ with $\|v_{R,\varepsilon}^i\|_{\varphi,k-1} \leq C_V$ and

$$\|w_i \circ V - v_{R,\varepsilon}^i\|_{\mathbf{C}^{k-1}(B_R(0))} < \varepsilon, \qquad (3.9)$$

where $B_R(0)$ denotes the closed ball in Z. Setting $g_\varepsilon := \sum_{i=1}^n Df(\cdot)(\zeta_i)v_{R,\varepsilon}^i \in \mathcal{A}(Z)$, it follows that with a constant $C_f > 0$ independent of $R > 0$,

$$\|\mathcal{L}_V f - g_\varepsilon\|_{\mathbf{C}^{k-1}(B_R(0))} < C_f \varepsilon. \qquad (3.10)$$

Choose $R_\varepsilon > 0$ large enough such that $\psi_j(z) > \varepsilon^{-1}$ for $\|z\|_Z > R_\varepsilon$. This is possible as the embedding $Z \to X$ is continuous. Hence, as f and all its derivatives are bounded,

$$\hat{\psi}_j(z)^{-1}\|D^j \mathcal{L}_V f(z)\|_{L_j(Z)} < C_f \varepsilon \quad \text{for } \|z\|_Z > R_\varepsilon, \quad j = 0, \dots, k-1, \quad (3.11)$$

where C_f is independent of ε. Furthermore,

$$\hat{\psi}_j(z)^{-1}\|D^j g_\varepsilon(z)\|_{L_j(Z)} \leq C_{f,V} \varepsilon \quad \text{for } \|z\|_Z > R_\varepsilon, \quad j = 0, \dots, k-1, \quad (3.12)$$

where $C_{f,V} > 0$ depends on f and V, but not on ε or R_ε. Plugging the results together proves the claim. $\qquad\square$

Let us consider two special cases.

Corollary 3.5 *Let $(H, \|\cdot\|_H)$ be a Hilbert space, $(Z, \|\cdot\|_Z)$ a continuously embedded Hilbert space. Define the D-admissible weight functions $\psi_j(x) := \cosh(\|x\|_H)$ on H and $\hat{\psi}_j(x) := \cosh(\|x\|_Z)$ on Z and the C-admissible weight functions $\varphi_j(x) := 1$ on Z, $j \geq 0$. Then, for every $k \geq 0$, the mapping*

$$\mathcal{L} : \mathcal{C}_{k-1}^\varphi(Z; X) \times \mathcal{B}_k^\psi(X) \to \mathcal{B}_{k-1}^{\hat{\psi}}(Z), \quad (V, f) \mapsto \mathcal{L}_V f, \qquad (3.13)$$

given by $\mathcal{L}_V f(x) := Df(x)V(x)$, is bounded and bilinear.

Remark 3.6 If $Z = H$, this has the simple interpretation that bounded vector fields map cosh-weighted spaces into themselves.

Proof This is straightforward from Theorem 3.3, as the $\hat{\psi}_j$ defined there is only a multiple of $\hat{\psi}_j$ in this case. $\qquad\square$

The following special case is very useful in the analysis of stochastic partial differential equations of Da Prato–Zabczyk type.

Corollary 3.7 *Let $(H, \|\cdot\|_H)$ be a Hilbert space, $(Z, \|\cdot\|_Z)$ a continuously embedded Hilbert space. Fix $n \in \mathbb{N}$. Define the D-admissible weight functions $\psi_j(x) := (1 + \|x\|_H^2)^{(n-j)/2}$ on H and $\hat{\psi}_j(x) := (1 + \|x\|_Z^2)^{(n-j)/2}$ on Z, $j = 0, \dots, n-1$, and the*

C-admissible weight functions $\varphi_0(x) := (1 + \|x\|_Z^2)^{1/2}$ and $\varphi_j(x) := 1$ on Z, $j \in \mathbb{N}$. Then, for $k \leq n - 1$, the mapping

$$\mathcal{L} \colon \mathcal{C}_{k-1}^\varphi(Z; X) \times \mathcal{B}_k^\psi(X) \to \mathcal{B}_{k-1}^{\hat\psi}(Z), \quad (V, f) \mapsto \mathcal{L}_V f, \qquad (3.14)$$

given by $\mathcal{L}_V f(x) := Df(x)V(x)$, is bounded and bilinear.

Remark 3.8 This means that linearly bounded vector fields $Z \to X$ with bounded derivatives (hence also Lipschitz continuous) map polynomially bounded functions to polynomially bounded functions, with the same weights. In particular, if the linear operator $A \colon \operatorname{dom} A \subset H \to H$ is densely defined and closed, then $V_A \in \mathcal{C}_k^\varphi(\operatorname{dom} A; H)$ for all $k \geq 0$, where φ is defined as in Corollary 3.7, $\operatorname{dom} A$ is endowed with the operator norm, and $V_A(x) := Ax$ for $x \in \operatorname{dom} A$.

Proof Calculating

$$(1 + \|x\|_Z^2)^{(n-1)/2}(1 + \|x\|_Z^2)^{1/2} + \sum_{i=0}^j \binom{j}{i}(1 + \|x\|_Z^2)^{(n-i-1)/2}$$

$$\leq C\hat\psi_j(x), \qquad (3.15)$$

the claim again follows from an application of Theorem 3.3.   □

4 Stability of cubature schemes

We shall now prove stability of cubature on Wiener space in the setting of weighted spaces. Consider from now on the following setup. Let on $[0,1]$ be given Lipschitz-continuous paths $(\omega_i^{(1)})_{i=1,\ldots,N}$, $\omega_i^{(1)}(s) = (\omega_i^{(1),j}(s))_{j=0,\ldots,d}$, $\omega_i^{(1),0}(s) = s$ starting at 0, and weights $(\lambda_i)_{i=1,\ldots,N}$ of a cubature on Wiener space of degree $m \geq 1$ for a d-dimensional Brownian motion, i.e., for all multi-indices $(j_1, \ldots, j_k)$ with $k + \#\{i \colon j_i = 0\} \leq m$ and a d-dimensional Brownian motion $(B_t^j)_{j=1,\ldots,d, t \geq 0}$,

$$\mathbb{E}\left[\int \cdots \int_{0 \leq s_1 \leq \cdots \leq s_k \leq 1} \circ dB_{s_1}^{j_1} \cdots \circ dB_{s_k}^{j_k}\right] \qquad (4.1)$$

$$= \sum_{i=1}^N \lambda_i \int \cdots \int_{0 \leq s_1 \leq \cdots \leq s_k \leq 1} d\omega_i^{(1),j_1}(s_1) \cdots \circ d\omega_i^{(1),j_k}(s_k).$$

Here, we have set $B_t^0 := t$ and $\circ dB_t^0 := dt$ for ease of notation. For a general time interval $[0, \Delta t]$, we set

$$\omega_i^{(\Delta t),0}(s) := s \quad \text{and} \quad \omega_i^{(\Delta t),j}(s) := \sqrt{\Delta t}\,\omega_i^{(1),j}(s/\Delta t), \quad j = 1, \ldots, d, \qquad (4.2)$$

so that $(\omega_i^{(\Delta t)})_{i=1,\dots,N}$ and $(\lambda_i)_{i=1,\dots,N}$ define a cubature formula on Wiener space of degree m on $[0, \Delta t]$. The approximation of the Markov semigroup $(P_t)_{t\geq 0}$, given by $P_t f(x) := \mathbb{E}[f(X_t^x)]$ for a function $f: H \to \mathbb{R}$, where $(X_t^x)_{t\geq 0}$ solves the Stratonovich stochastic differential equation

$$\mathrm{d}X_t^x = \sum_{j=0}^{d} V_j(X_t^x) \circ \mathrm{d}B_t^j, \quad X_0^x = x, \tag{4.3}$$

on some state space H, then reads

$$P_t f(x) \approx Q_{(t/n)}^n f(x), \tag{4.4}$$

where the one step approximation operator is defined by

$$Q_{(\Delta t)} f(x) := \sum_{i=1}^{N} \lambda_i f\left(X_{\Delta t}^x\left(\omega_i^{(\Delta t)}\right)\right), \tag{4.5}$$

with $X_t^x(\omega_i^{(\Delta t)})$ the solution of the problem

$$\mathrm{d}X_s^x(\omega_i^{(\Delta t)}) = \sum_{j=0}^{d} V_j\left(X_s^x\left(\omega_i^{(\Delta t)}\right)\right) \mathrm{d}\omega_i^{(\Delta t),j}(s), \quad X_0^x\left(\omega_i^{(\Delta t)}\right) = x. \tag{4.6}$$

Under certain smoothness assumptions on the vector fields V_j, $j = 0, \dots, d$, and the test function f, we expect that

$$|P_t f(x) - Q_{(t/n)}^n f(x)| \leq C n^{-(m-1)/2}, \tag{4.7}$$

where the constant $C > 0$ can depend on f, V_j, $j = 0, \dots, d$, and $x \in H$. For the case H finite-dimensional and f and V_j bounded and C^∞-bounded, $j = 0, \dots, d$, it is known that C depends on the supremum norms of f and its derivatives, but not on $x \in H$, see Lyons and Victoir [24]. For more background on the method, see Bayer and Teichmann [2], Crisan and Ghazali [4] and Lyons and Victoir [24]. An alternative approach can be found in Kusuoka [21,22]. Its implementation as a splitting method is given in Ninomiya and Victoir [26], see also Alfonsi [1], Ninomiya and Ninomiya [25], and Tanaka and Kohatsu-Higa [33].

Our strategy is as follows. First, we consider the finite dimensional case. Here, the analysis is straightforward. Afterwards, we turn to the infinite dimensional setting. Here, our aim is to prove stability for Da Prato–Zabczyk equations with pseudodissipative generator. We prove first the auxiliary result in Theorem 4.4, which might be of independent interest. The method of the moving frame then yields first Theorem 4.7, and the Szőkefalvi–Nagy theorem allows us to conclude in Corollary 4.8.

4.1 Finite dimensional state space

Given a Stratonovich SDE on $\mathbb{R}^n$,

$$dX_t^x = \sum_{j=0}^{d} V_j(X_t^x) \circ dB_t^j, \quad X_0^x = x, \tag{4.8}$$

we let the local discretisation of $P_t f(x) := \mathbb{E}[f(X_t^x)]$ be defined by

$$Q_{(\Delta t)} f(x) := \sum_{i=1}^{N} \lambda_i f\left(X_{\Delta t}^x\left(\omega_i^{(\Delta t)}\right)\right), \tag{4.9}$$

where $X_t^x(\omega_i^{(\Delta t)})$ is the solution of the problem

$$dX_s^x(\omega_i^{(\Delta t)}) = \sum_{j=0}^{d} V_j\left(X_s^x\left(\omega_i^{(\Delta t)}\right)\right) d\omega_i^{(\Delta t),j}(s), \quad X_0^x(\omega_i^{(\Delta t)}) = x. \tag{4.10}$$

Theorem 4.1 *Let ψ be an admissible weight function on $\mathbb{R}^n$, and assume that*

$$|V_i V_j \psi(x)| + |V_i \psi(x)| \leq C\psi(x) \quad for\ i = 0, \ldots, d\ and\ j = 1, \ldots, d, \tag{4.11}$$

where we require that all the necessary derivatives are well-defined.
Then, there exists a constant $\tilde{C} > 0$ independent of $0 < \Delta t \leq T$ such that

$$Q_{(\Delta t)} \psi(x) \leq \exp(\tilde{C} \Delta t)\psi(x). \tag{4.12}$$

Proof We define the intermediate operator

$$Q_{(\Delta t, s)} f(x) := \sum_{i=1}^{N} \lambda_i f\left(X_s^x\left(\omega_i^{(\Delta t)}\right)\right) \quad \text{for } s \in [0, t] \tag{4.13}$$

and note that $Q_{(\Delta t)} = Q_{(\Delta t, \Delta t)}$. The definition of the iteration step yields

$$\psi(X_s^x(\omega_i^{(\Delta t)})) = \psi(x) + \sum_{j=0}^{d} \int_0^s V_j\left(X_r^x\left(\omega_i^{(\Delta t)}\right)\right) d\omega_i^{(\Delta t),j}(r)$$

$$= \psi(x) + \int_0^s V_0 \psi\left(X_r^x\left(\omega_i^{(\Delta t)}\right)\right) dr + \sum_{j=1}^{d} V_j \psi(x)\omega_i^{(\Delta t),j}(s)$$

$$+ \sum_{j=1}^{d} \sum_{k=0}^{d} \int_0^s \int_0^r V_k V_j \psi\left(X_q^x\left(\omega_i^{(\Delta t)}\right)\right) d\omega_i^{(\Delta t),k}(q) d\omega_i^{(\Delta t),j}(r).$$

$$\tag{4.14}$$

By (4.11),

$$\int_0^s V_0 \psi \left(X_r^x \left(\omega_i^{(\Delta t)} \right) \right) dr \leq C \int_0^s \psi \left(X_r^x \left(\omega_i^{(\Delta t)} \right) \right) dr. \qquad (4.15)$$

Furthermore, as $|\omega_i^{(\Delta t),j}(s)| \leq \hat{C}(\Delta t)^{1/2}$ and $|\frac{\partial}{\partial s}\omega_i^{(\Delta t),j}(s)| \leq \hat{C}(\Delta t)^{-1/2}$ with some constant $\hat{C} > 0$ independent of $0 < \Delta t \leq T$ (notice the case $j = 0$), Fubini's theorem yields

$$\int_0^s \int_0^r V_k V_j \psi \left(X_q^x \left(\omega_i^{(\Delta t)} \right) \right) d\omega_i^{(\Delta t),k}(q) d\omega_i^{(\Delta t),j}(r)$$

$$\leq C \int_0^s |\omega_i^{(\Delta t),j}(s) - \omega_i^{(\Delta t),j}(q)| \psi \left(X_q^x \left(\omega_i^{(\Delta t)} \right) \right) \left| \frac{\partial}{\partial q} \omega_i^{(\Delta t),k}(q) \right| dq$$

$$\leq 2\hat{C}^2 C \int_0^s \psi \left(X_q^x \left(\omega_i^{(\Delta t)} \right) \right) dq. \qquad (4.16)$$

Thus, we see that for some $C' > 0$,

$$Q_{(\Delta t,s)}\psi(x) = \sum_{i=1}^N \lambda_i \psi \left(X_s^x \left(\omega_i^{(\Delta t)} \right) \right) \qquad (4.17)$$

$$\leq \psi(x) + \sum_{j=1}^d V_j \psi(x) \sum_{i=1}^N \lambda_i \omega_i^{(\Delta t),j}(s) + C' \int_0^s Q_{(\Delta t,r)}\psi(x) dr.$$

Defining $\alpha_{\Delta t,s}(x) := \sum_{j=1}^d V_j \psi(x) \sum_{i=1}^N \lambda_i \omega_i^{(\Delta t),j}(s)$, the Gronwall inequality yields

$$Q_{(\Delta t,s)}\psi(x) \leq \psi(x) + \alpha_{\Delta t,s}(x) + \int_0^s \left(\psi(x) + \alpha_{\Delta t,r}(x) \right) C' \exp(C'(s-r)) dr. \qquad (4.18)$$

Note that $\alpha_{\Delta t,\Delta t}(x) = 0$ by the equality $\sum_{i=1}^N \lambda_i \omega_i^{(\Delta t),j}(\Delta t) = 0$. Furthermore,

$$\alpha_{\Delta t,r}(x) \leq C\sqrt{\Delta t}\psi(x) \leq \frac{C}{2}(1 + \Delta t)\psi(x) \leq \frac{C}{2}\exp(\Delta t)\psi(x). \qquad (4.19)$$

This proves

$$Q_{(\Delta t)}\psi(x) = Q_{(\Delta t,\Delta t)}\psi(x) \leq \psi(x) \left(1 + \left(1 + \frac{C}{2}\exp(\Delta t) \right) (\exp(C'\Delta t) - 1) \right)$$

$$\leq \exp(\tilde{C}\Delta t)\psi(x), \qquad (4.20)$$

where $\tilde{C} := \max(C/2(1 + C') + 1, C' + 1)$, which is the required estimate. $\qquad \square$

4.2 Time-dependent stochastic ordinary differential equations on Hilbert space

Let H be a Hilbert space, and consider the nonautonomous stochastic ordinary differential equation

$$\mathrm{d}X_t^x = \sum_{j=0}^{d} V_j(t, X_t^x) \circ \mathrm{d}B_t^j, \quad X_0^x = x, \tag{4.21}$$

on H. We define cubature approximations of (4.21) by

$$\mathrm{d}X_s^{t,x}\left(\omega_i^{(\Delta t)}\right) = \sum_{j=0}^{d} V_j\left(t + s, X_s^{t,x}\left(\omega_i^{(\Delta t)}\right)\right) \mathrm{d}\omega_i^{(\Delta t),j}(s), \quad X_0^{t,x} = x, \tag{4.22}$$

and the approximation operator by

$$Q_{(\Delta t)}^t f(x) := \sum_{i=1}^{N} \lambda_i f\left(X_{\Delta t}^{t,x}\left(\omega_i^{(\Delta t)}\right)\right). \tag{4.23}$$

Definition 4.2 A cubature formula $(\omega_i^{(\Delta t)}, \lambda_i)_{i=1,\ldots,N}$ is called *symmetric* if for every $i \in \{1, \ldots, N\}$, there exists some $i' \in \{1, \ldots, N\}$ such that $\lambda_i = \lambda_{i'}$ and

$$\omega_i^{(\Delta t),j}(s) = -\omega_{i'}^{(\Delta t),j}(s) \quad \text{for all } s \in [0, \Delta t] \text{ and } j = 1, \ldots, d. \tag{4.24}$$

It is called *weakly symmetric* if for $j = 1, \ldots, d$,

$$\sum_{i=1}^{N} \lambda_i \omega_i^{(\Delta t),j}(s) = 0 \quad \text{for} \quad s \in [0, \Delta t]. \tag{4.25}$$

Remark 4.3 In words symmetry means that with every cubature path ω also its *space*-component-wise negative, i.e., $(s, -\omega^1(s), \ldots, -\omega^d(s))$, for $s \in [0, \Delta t]$ is among the cubature paths. Notice that time is *not* reversed. Clearly, all symmetric cubature formulas are also weakly symmetric. Note that many known cubature formulas are actually symmetric. Moreover, a non-symmetric cubature formula can be made symmetric by adding the space-component-wise negatives of the paths with the same weights to it and dividing all weights by two. This will at most double the number of paths. Thus, if we use a cubature formula with a small number of paths in high dimensions, we can also find a symmetric cubature formula with this property. Also notice that we only need weak symmetry of cubature formulas for our considerations, but we introduced symmetry for convenience.

Theorem 4.4 *Suppose that the cubature formula used in the definition of $Q^t_{(\Delta t)}$ is weakly symmetric. Let ψ be an admissible weight function on H and suppose*

$$\|D\psi(x)\| \leq C_1(1 + \|x\|^2)^{-1/2}\psi(x) \quad and \qquad (4.26)$$
$$\|D^2\psi(x)\| \leq C_1(1 + \|x\|^2)^{-1}\psi(x) \qquad (4.27)$$

with some constant $C_1 > 0$, Furthermore, assume that for some constant $C_2 > 0$,

$$\|V_j(t, x)\| \leq C_2(1 + \|x\|^2)^{1/2} \quad for \ j = 0, \ldots, d, x \in X and \ t \in [0, T], \quad (4.28)$$

and that $x \mapsto V_j(t, x)$ is continuously differentiable with derivative bounded uniformly in $t \in [0, T]$ for $j = 1, \ldots, d$.

Then, there exists a constant $\tilde{C} > 0$ such that for all $t \in [0, T]$ and $\Delta t \in [0, T - t]$,

$$Q^t_{(\Delta t)}\psi(x) \leq \exp(\tilde{C}t)\psi(x) \quad for \ all \ x \in H. \qquad (4.29)$$

Remark 4.5 The above result is remarkable as we do not assume that the vector fields V_j are differentiable with respect to t. This is also the reason why we cannot directly use the ideas in Theorem 4.1 to conclude.

Proof Define the intermediate approximation for $s \in [0, \Delta t]$ by

$$Q^t_{(\Delta t, s)}f(x) := \sum_{i=1}^{N} \lambda_i f\left(X^{t,x}_s\left(\omega^{(\Delta t)}_i\right)\right). \qquad (4.30)$$

As in the proof of Theorem 4.1, we note that $Q^t_{(\Delta t, \Delta t)} = Q^t_{(\Delta t)}$. For $0 \leq s \leq \Delta t$,

$$\psi\left(X^{t,x}_s\left(\omega^{(\Delta t)}_i\right)\right) =$$
$$\psi(x) + \sum_{j=0}^{d}\int_0^s D\psi\left(X^{t,x}_r\left(\omega^{(\Delta t)}_i\right)\right)V_j\left(t+r, X^{t,x}_r\left(\omega^{(\Delta t)}_i\right)\right)d\omega^{(\Delta t),j}_i(r). \quad (4.31)$$

Consider $g_j(r, x) := D\psi(x)V_j(t + r, x)$. Then,

$$g_j(\rho, X^{t,x}_r(\omega^{(\Delta t)}_i)) = g_j(\rho, x)$$
$$+ \sum_{k=0}^{d}\int_0^r D_x g_j\left(\rho, X^{t,x}_q\left(\omega^{(\Delta t)}_i\right)\right)V_k\left(t+q, X^{t,x}_q\left(\omega^{(\Delta t)}_i\right)\right)d\omega^{(\Delta t),k}_i(q). \quad (4.32)$$

From (4.26) to (4.28), we obtain that for $0 \leq s \leq \Delta t \leq T$,

$$|g_0(r, x)| = |D\psi(x)V_0(t + r, x)| \leq \|D\psi(x)\| \cdot \|V_0(t + r, x)\|$$
$$\leq C_1 C_2 \psi(x). \qquad (4.33)$$

We argue in a similar manner for $D_x g_j(r, x) V_k(t+q, x)$, $j = 1, \ldots, d$, $k = 0, \ldots, d$, to obtain that for $0 \leq q \leq r \leq \Delta t$,

$$
\begin{aligned}
|D_x g_j(r, x) V_k(t + q, x)| &= |D^2 \psi(x)(V_j(t + r, x), V_k(t + q, x)) \\
&\quad + D\psi(x) D_x V_j(t + r, x) V_k(t + q, x)| \\
&\leq C_1 (C_2^2 + C_2) \psi(x).
\end{aligned}
\tag{4.34}
$$

An application of Fubini's theorem just as in the proof of Theorem 4.1 gives

$$
\begin{aligned}
&\psi \left(X_s^{t,x} \left(\omega_i^{(\Delta t)} \right) \right) \\
&= \psi(x) + \int_0^s g_0 \left(r, X_r^{t,x} \left(\omega_i^{(\Delta t)} \right) \right) dr + \sum_{j=1}^d \int_0^s g_j(r, x) d\omega_i^{(\Delta t), j}(r) \\
&\quad + \sum_{j=1}^d \sum_{k=0}^d \int_0^s \int_0^r D_x g_j \left(r, X_q^{t,x} \left(\omega_i^{(\Delta t)} \right) \right) V_k \left(t + q, X_q^{t,x} \left(\omega_i^{(\Delta t)} \right) \right) d\omega_i^{(\Delta t), k}(q) d\omega_i^{(\Delta t), j}(r) \\
&\leq \psi(x) + C' \int_0^s \psi \left(X_r^x \left(\omega_i^{(\Delta t)} \right) \right) dr + \sum_{j=1}^d \int_0^s g_j(r, x) d\omega_i^{(\Delta t), j}(r)
\end{aligned}
\tag{4.35}
$$

with a constant $\tilde{C} > 0$ depending on C_1 and C_2, where we apply that $\Delta t \leq T$. As from the weak symmetry of the cubature paths,

$$
\begin{aligned}
\sum_{i=1}^N \lambda_i \sum_{j=1}^d \int_0^s g_j(r, x) d\omega_i^{(\Delta t), j}(r) &= \sum_{j=1}^d \int_0^s g_j(r, x) d \left(\sum_{i=1}^N \lambda_i \omega_i^{(\Delta t), j}(r) \right) \\
&= 0,
\end{aligned}
\tag{4.36}
$$

we obtain

$$
Q_{(\Delta t, s)} \psi(x) \leq \psi(x) + \tilde{C} \int_0^s Q_{(\Delta t, r)} \psi(x) dr.
\tag{4.37}
$$

An application of Gronwall's lemma yields $Q_{(\Delta t)} \psi(x) \leq \exp(\tilde{C} \Delta t) \psi(x)$, which proves the result. □

Remark 4.6 It is clear that the given assumptions on the vector fields and the weight function are not the only ones possible. Instead, we could also require the vector fields to be bounded uniformly in $t \in [0, T]$, and allow the weight function to satisfy $\|D\psi(x)\| + \|D^2\psi(x)\| \leq C\psi(x)$. While the situation above corresponds to polynomially growing weight functions and linearly bounded vector fields, this variant corresponds to exponentially growing weight functions and bounded vector fields, see also Corollaries 3.5 and 3.7.

Such an approach might be more appropriate when dealing with exponentials of stochastic processes such as Lévy processes, which are ubiquitous in applications in mathematical finance, as they ensure nonnegativity in a simple manner and allow us to work on the natural scale of the problem.

4.3 Da Prato–Zabczyk equations

Suppose now that

$$dX_t^x = AX_t^x dt + \sum_{j=0}^{d} V_j(X_t^x) \circ dB_t^j, \quad X_0^x = x, \tag{4.38}$$

is a stochastic partial differential equation of Da Prato–Zabczyk type on some Hilbert space H, see Da Prato and Zabczyk [6,7] for a comprehensive exposition of the theory of such equations. Here, solutions are understood in the mild sense,

$$X_t^x = \exp(tA)x + \sum_{j=0}^{d} \int_0^t \exp((t-s)A)V_j(X_t^x) \circ dB_s^j, \tag{4.39}$$

and we also define the cubature discretisations in the mild sense,

$$X_t^x\left(\omega_i^{(\Delta t)}\right) = \exp(tA)x + \sum_{j=0}^{d} \int_0^t \exp((t-s)A)V_j\left(X_t^x\left(\omega_i^{(\Delta t)}\right)\right) d\omega_i^{(\Delta t),j}(s). \tag{4.40}$$

Again, the approximation of the Markov semigroup $P_t f(x) := \mathbb{E}[f(X_t^x)]$ is given by

$$Q_{(\Delta t)} f(x) := \sum_{i=1}^{N} \lambda_i f\left(X_{\Delta t}^x\left(\omega_i^{(\Delta t)}\right)\right). \tag{4.41}$$

Here we do only treat Da Prato–Zabczyk equations driven by finitely many Brownian motions, since there are no generically infinite dimensional cubature approaches known for Da Prato–Zabczyk equations driven by infinitely many Brownian motions. Certainly approximation arguments would always work but are not considered in this article.

Theorem 4.7 *Suppose that A is the generator of a group $S_t = \exp(tA)$, $t \in \mathbb{R}$, with growth estimate $\|S_t\| \leq M \exp(\lambda|t|)$ for $t \in \mathbb{R}$ and for some numbers $M \geq 1$ and $\lambda \in \mathbb{R}$ (which always exist!), and that the cubature formula used in the definition of $Q_{(\Delta t)}$ is weakly symmetric. Let ψ be an admissible weight function on H. Suppose that for some $C > 0$ depending on A, ψ, and V_j, $j = 1, \ldots, d$, $\psi(S_t x) \leq \exp(Ct)\psi(x)$*

for all $x \in H$ and $t > 0$,

$$\|D\psi(x)\| \leq C(1 + \|x\|^2)^{-1/2}\psi(x), \tag{4.42a}$$

$$\|D^2\psi(x)\| \leq C(1 + \|x\|^2)^{-1}\psi(x), \quad and \tag{4.42b}$$

$$\|V_j(x)\| \leq C(1 + \|x\|^2)^{1/2} \quad for \ j = 0, \ldots, d, \tag{4.42c}$$

and that V_j is continuously differentiable with derivative bounded by C for $j = 1, \ldots, d$.

Then, for any $T > 0$, there exists a constant $\tilde{C} > 0$ depending on C and T such that for every $\Delta t \in [0, T]$, the operator $Q_{(\Delta t)}$ satisfies

$$Q_{(\Delta t)}\psi(x) \leq \exp(\tilde{C}\Delta t)\psi(x) \quad for \ all \ x \in H. \tag{4.43}$$

Proof We apply the method of the moving frame from [35]. This yields that $X_t^x = S_t Y_t^x$, where $(Y_t^x)_{t\geq 0}$ satisfies the Hilbert space stochastic ordinary differential equation

$$dY_t^x = \sum_{j=0}^{d} \tilde{V}_j(t, Y_t^x) \circ dB_t^j, \quad Y_0^x = x, \tag{4.44}$$

with $\tilde{V}_j(t, y) = S_{-t}V_j(S_t y)$. Thus, rewriting the cubature discretisations of $(X_t^x)_{t\geq 0}$ using $(Y_t^x)_{t\geq 0}$,

$$dY_s^x\left(\omega_i^{(\Delta t)}\right) = \sum_{j=0}^{d} \tilde{V}_j\left(s, Y_s^x\left(\omega_i^{(\Delta t)}\right)\right) d\omega_i^{(\Delta t),j}(s), \tag{4.45}$$

we see that, if we define

$$\tilde{Q}_{(\Delta t)}f(y) := \sum_{i=1}^{N} \lambda_i f\left(Y_{\Delta t}^y\left(\omega_i^{(\Delta t)}\right)\right) \tag{4.46}$$

for $f \colon H \to \mathbb{R}$, then $Q_{(\Delta t)}h(x) = \tilde{Q}_{(\Delta t)}g(x)$, where $g(y) := h(S_{\Delta t}y)$. In particular,

$$Q_{(\Delta t)}\psi(x) = \tilde{Q}_{(\Delta t)}(\psi \circ S_{\Delta t})(x) \leq \exp(C\Delta t)\tilde{Q}_{(\Delta t)}\psi(x), \tag{4.47}$$

where we apply the assumptions on ψ and the positivity of $\tilde{Q}_{(\Delta t)}$.

But now, we are in the situation of Theorem 4.4: The estimates for ψ are clear by assumption, and for $\tilde{V}_j(s, y)$, we note that, as $s \in [0, T]$,

$$\|\tilde{V}_j(s, y)\| = \|S_{-s}V_j(S_s y)\| \leq CM^2\exp(2\lambda T)(1 + \|y\|^2)^{1/2} \tag{4.48}$$

$$and \ \|D_y\tilde{V}_j(s, y)\| = \|S_{-s}D_yV_j(S_s y)S_s\| \leq CM^2\exp(2\lambda T) \quad for \ j = 1, \ldots, d. \tag{4.49}$$

An appeal to Theorem 4.4 yields

$$\tilde{Q}_{(\Delta t)}\psi(x) \leq \exp(\tilde{C}\Delta t)\psi(x) \tag{4.50}$$

with some constant $\tilde{C} > 0$, and the result follows.   □

The Szőkefalvi–Nagy theorem now allows us to obtain a corresponding result for pseudocontractive semigroups.

Corollary 4.8 *Suppose that A is the generator of a semigroup of pseudocontractions $S_t = \exp(tA)$, $t \geq 0$. Let $\psi(x) = \rho(\|x\|^2)$ with some increasing and left continuous function $\rho\colon [0, \infty) \to (0, \infty)$ (see also Dörsek and Teichmann [11, Example 4.1]) which satisfies $\rho(Cu) \leq C\rho(u)$ for all $u \geq 0$ and $C > 0$, and which is twice differentiable and satisfies*

$$\rho'(u) \leq C(1+u)^{-1}\rho(u) \quad and \quad \rho''(u) \leq C(1+u)^{-2}\rho(u). \tag{4.51}$$

Furthermore, assume that $\|V_j(x)\| \leq C(1 + \|x\|^2)^{1/2}$ for $j = 0, \ldots, d$, and that V_j is continuously differentiable with bounded derivative for $j = 1, \ldots, d$, and that the cubature formula used in the definition of $Q_{(\Delta t)}$ is weakly symmetric.

Then, for any $T > 0$, there exists a constant $\tilde{C} > 0$ such that for every $\Delta t \in [0, T]$, the operator $Q_{(\Delta t)}$ satisfies

$$Q_{(\Delta t)}\psi(x) \leq \exp(\tilde{C}\Delta t)\psi(x) \quad for all \ x \in H. \tag{4.52}$$

Proof Assume without loss of generality that $(S_t)_{t \geq 0}$ is a semigroup of contractions. By the Szőkefalvi–Nagy theorem [28, p. 452, Théorème IV], we see that we can find a Hilbert space $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ containing H as a closed subspace and a strongly continuous group $(\mathcal{S}_t)_{t \in \mathbb{R}}$ of unitary mappings such that $S_t = \pi \mathcal{S}_t$, where $\pi\colon \mathcal{H} \to H$ is the orthogonal projection.

Define $\psi_{\mathcal{H}}(y) := \rho(\|y\|_{\mathcal{H}}^2)$ and $V_j^{\mathcal{H}}(y) := V_j(\pi y)$, then it is easy to see that the assumptions of Theorem 4.7 are satisfied. The results of Teichmann [35] prove that the solution of

$$X_t^{\mathcal{H}, y} = \mathcal{S}_t y + \sum_{j=0}^{d} \int_0^t \mathcal{S}_{t-s} V_j^{\mathcal{H}}\left(X_s^{\mathcal{H}, y}\right) \circ \mathrm{d}B_s^j \tag{4.53}$$

satisfies $X_t^x = \pi X_t^{\mathcal{H}, x}$, and similarly for the cubature approximations. Setting

$$Q_{(\Delta t)}^{\mathcal{H}} f(y) := \sum_{i=1}^{N} \lambda_i f\left(X_{\Delta t}^{\mathcal{H}, y}\left(\omega_i^{(\Delta t)}\right)\right), \tag{4.54}$$

Theorem 4.7 yields that $Q^{\mathcal{H}}_{(\Delta t)}\psi_{\mathcal{H}}(y) \leq \exp(\tilde{C}\Delta t)\psi_{\mathcal{H}}(y)$ with some constant $\tilde{C} > 0$, and from $\psi_{\mathcal{H}}(x) = \psi(x)$ for $x \in H$ we obtain that for $x \in H$,

$$Q_{(\Delta t)}\psi(x) = \sum_{i=1}^{N} \lambda_i \rho\left(\|\pi X^{\mathcal{H},x}_{\Delta t}\left(\omega_i^{(\Delta t)}\right)\|^2\right) \leq \sum_{i=1}^{N} \lambda_i \rho\left(\|X^{\mathcal{H},x}_{\Delta t}\left(\omega_i^{(\Delta t)}\right)\|^2_{\mathcal{H}}\right)$$

$$= Q^{\mathcal{H}}_{(\Delta t)}\psi_{\mathcal{H}}(x) \leq \exp(\tilde{C}\Delta t)\psi_{\mathcal{H}}(x) = \exp(\tilde{C}\Delta t)\psi(x). \tag{4.55}$$

The result is thus proved. $\qquad\square$

5 Convergence estimates of cubature schemes

We are now ready to prove rates of convergence for cubature on Wiener space on weighted spaces. We shall only prove these results in the infinite-dimensional setting; corresponding results in finite dimensions are obtained in a similar manner.

Let H be a Hilbert space and A the infinitesimal generator of a strongly continuous semigroup of pseudocontractions on H. Fix $\ell_0 \in \mathbb{N}$. For $\ell = 0, \ldots, \ell_0$, let H_ℓ be subspaces of H endowed with Hilbert norms $\|\cdot\|_{H_\ell}$, $H_0 = H$, such that for $\ell = 0, \ldots, \ell_0 - 1$, $H_{\ell+1} \subset H_\ell$ and $A\colon H_{\ell+1} \to H_\ell$ is a bounded linear operator. On H_ℓ, we define D-admissible weight functions

$$\psi_\ell^s(x) := \left(1 + \|x\|^2_{H_\ell}\right)^{s/2}, \quad s \geq 1, \quad \ell = 0, \ldots, \ell_0, \quad \psi^s := \psi_0^s, \tag{5.1}$$

and the functions

$$\varphi_{\ell,0}(x) := \left(1 + \|x\|^2_{H_\ell}\right)^{1/2}, \quad \varphi_{\ell,j}(x) := 1, \quad j \geq 1. \tag{5.2}$$

Define the vectors of weight functions $\psi_\ell^{(n)} := (\psi_\ell^{n-j})_{j=0,\ldots,k}$, $k < n$, and $\varphi_\ell := (\varphi_{\ell,j})_{j=0,\ldots,k}$.

Assumption 5.1 The vector fields satisfy

$$V_j \in \mathcal{C}_k^{\varphi_\ell}(H_\ell, H_\ell) \quad \text{for } j = 0, \ldots, d \text{ and } \ell = 0, \ldots, \ell_0. \tag{5.3}$$

Remark 3.8 shows that $A \in \mathcal{C}_k^{\varphi_{\ell+1}}(H_{\ell+1}, H_\ell)$ for $\ell = 0, \ldots, \ell_0 - 1$. For $x \in H_\ell$, $\ell = 0, \ldots, \ell_0$, we can then consider the Da Prato–Zabczyk equation

$$dX_t^x = AX_t^x dt + \sum_{j=0}^{d} V_j(X_t^x) \circ dB_t^j, \quad X_0^x = x, \tag{5.4}$$

on H_ℓ. As the assumptions on the vector fields V_j essentially mean that they are Lipschitz continuous with bounded derivatives, all these equations have unique solutions, agreeing if we vary ℓ for sufficiently smooth initial conditions.

Assumption 5.2 The Markov semigroup $(P_t)_{t \geq 0}$, $P_t f(x) := \mathbb{E}[f(X_t^x)]$, is strongly continuous on $\mathcal{B}^{\psi_\ell^n}(H_\ell)$ for all $n \in \mathbb{N}$ and $\ell = 0, \ldots, \ell_0$. For some $k_0 \in \mathbb{N}$, P_t is a bounded map from $\mathcal{B}_k^{\psi_\ell^{(n)}}(H_\ell)$ into itself for $k = 0, \ldots, k_0$ and $n \in \mathbb{N}, n > k$, with norm bounded uniformly in $t \in [0, T]$ for every $T > 0$.

See also Dörsek and Teichmann [11, Sect. 4, Lemma 7.19] for sufficient conditions for these assumptions.

5.1 Taylor expansion of stochastic partial differential equations

Theorem 5.3 *Let $\ell = 1, \ldots, \ell_0$. Consider the strongly continuous semigroup $(P_t)_{t \geq 0}$ on the space $\mathcal{B}^{\psi_\ell^n}(H_\ell)$ with $n \geq 4$. Denote its generator by $(\mathcal{G}, \mathrm{dom}\,\mathcal{G})$.*

Then, $\mathcal{B}_2^{\psi_{\ell-1}^{(n)}}(H_{\ell-1}) \subset \mathrm{dom}\,\mathcal{G}$, and

$$\mathcal{G}f(x) = Df(x)(Ax) + \mathcal{L}_{V_0}f(x) + \frac{1}{2}\sum_{j=1}^{d}\mathcal{L}_{V_j}^2 f(x) \tag{5.5}$$

for $f \in \mathcal{B}_2^{\psi_{\ell-1}^{(n)}}(H_{\ell-1})$ and $x \in H_\ell$.

Proof By the Itô formula, see, e.g., Da Prato and Zabczyk [7, Theorem 7.2.1], it follows that for $f \in A(H_{\ell-1})$, we have $f \in \mathrm{dom}\,\mathcal{G}$ and f satisfies (5.5). Corollary 3.7 shows that the right hand side of (5.5) is a continuous linear operator $\mathcal{B}_2^{\psi_{\ell-1}^{(n)}}(H_{\ell-1}) \to \mathcal{B}^{\psi_\ell^n}(H_\ell)$. The closedness of $\mathcal{G}$ proves the claim. □

The next result follows directly from Corollary 3.7, together with the explicit representation in (5.5).

Corollary 5.4 *Let $k \geq 0$. Under the assumptions of Theorem 5.3, the infinitesimal generator $\mathcal{G}$ satisfies the mapping property*

$$\mathcal{G}: \mathcal{B}_{k+2}^{\psi_{\ell-1}^{(n)}}(H_{\ell-1}) \to \mathcal{B}_k^{\psi_\ell^{(n)}}(H_\ell), \quad \ell = 1, \ldots, \ell_0. \tag{5.6}$$

Induction now yields:

Corollary 5.5 *Let $j = \ell, \ldots, \ell_0$. Under the assumptions of Theorem 5.3, the powers of the infinitesimal generator $\mathcal{G}$ satisfy*

$$\mathcal{G}^j: \mathcal{B}_{k+2j}^{\psi_{\ell-j}^{(n)}}(H_{\ell-j}) \to \mathcal{B}_k^{\psi_\ell^{(n)}}(H_\ell). \tag{5.7}$$

They are given explicitly by taking the powers of (5.5).

This allows us to obtain a Taylor expansion of $P_t f$ for smooth enough f, which we will compare to the Taylor expansion of cubature approximations.

Corollary 5.6 *Let* $f \in \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}(H_{\ell-(k+1)}), k+1 \leq \ell \leq \ell_0, n \geq 2(k+2).$
 Then,

$$P_t f = \sum_{j=0}^{k} \frac{t^j}{j!} \mathcal{G}^j f + t^{k+1} R_{t,k} f, \tag{5.8}$$

where the linear operator $R_{t,k} \colon \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}(H_{\ell-(k+1)}) \to \mathcal{B}^{\psi_{\ell}^{n}}(H_{\ell})$ *satisfies*

$$\|R_{t,k} f\|_{\psi_{\ell}^{n}} \leq C_T \|f\|_{\psi_{\ell-(k+1)}^{(n)}, 2(k+1)} \quad \textit{for } t \in [0, T] \tag{5.9}$$

for a constant $C_T > 0$ *independent of* f.

5.2 Taylor expansion of cubature approximations

For a multiindex $\alpha = (i_1, \ldots, i_k)$, we define $\deg(\alpha) := k + \# \{ j = 1, \ldots, k : i_j = 0 \}$.
The empty multiindex is denoted by $\emptyset$, corresponds to $k = 0$, and satisfies $\deg(\emptyset) = 0$.
We set

$$\mathcal{A}_m := \{\alpha : \deg(\alpha) \leq m\} \quad \text{and} \quad \mathcal{A}_m^* := \mathcal{A}_m \setminus \{\emptyset, (0)\}. \tag{5.10}$$

Theorem 5.7 *Assume that the cubature formula is of degree* $m = 2k + 1$. *For* $f \in \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}(H_{\ell-(k+1)}), k+1 \leq \ell \leq \ell_0, n \geq 2(k+2),$

$$Q_{(\Delta t)} f = \sum_{j=0}^{k} \frac{(\Delta t)^j}{j!} \mathcal{G}^j f + (\Delta t)^{k+1} \hat{R}_{\Delta t, k} f, \tag{5.11}$$

where the linear operator $\hat{R}_{\Delta t, k} \colon \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}(H_{\ell-(k+1)}) \to \mathcal{B}^{\psi_{\ell}^{n}}(H_{\ell})$ *satisfies*

$$\|\hat{R}_{\Delta t, k} f\|_{\psi_{\ell}^{n}} \leq C_T \|f\|_{\psi_{\ell-(k+1)}^{(n)}, 2(k+1)} \quad \textit{for } \Delta t \in [0, T] \tag{5.12}$$

for a constant $C_T > 0$ *independent of* f.

Proof Under the assumptions on the vector fields, we can easily see that for every $f \in A(H_{\ell-(k+1)})$, we have the Taylor expansion

$$\begin{aligned} & f\left(X_{\Delta t}^x \left(\omega_i^{(\Delta t)}\right)\right) \\ & = \sum_{(i_1, \ldots, i_k) \in \mathcal{A}_m} V_{i_1} \ldots V_{i_k} f(x) I_{\Delta t}^{(i_1, \ldots, i_k)}\left(\omega_i^{(\Delta t)}\right) + \hat{R}_{\Delta t, k}^i f(x), \end{aligned} \tag{5.13}$$

where we define the iterated integrals by

$$
I_{\Delta t}^{(i_1,\ldots,i_k)} \left(\omega_i^{(\Delta t)}, g \right) \tag{5.14}
$$
$$
:= \int\limits_{0 < t_1 < \cdots < t_k < \Delta t} g \left(X_{t_1}^x \left(\omega_i^{(\Delta t)} \right) \right) \, d\omega_i^{(t),i_1}(t_1) \ldots d\omega_i^{(t),i_k}(t_k),
$$

$I_{\Delta t}^{(i_1,\ldots,i_k)}(\omega_i^{(\Delta t)}) := I_{\Delta t}^{(i_1,\ldots,i_k)}(\omega_i^{(\Delta t)}, 1)$, the remainder term $\hat{R}_{\Delta t,k}^i f$ satisfies

$$
\hat{R}_{\Delta t,k}^i f(x) = \sum_{\substack{(i_1,\ldots,i_k) \in \mathcal{A}_m \\ (i_0,i_1,\ldots,i_k) \notin \mathcal{A}_m}} I_{\Delta t}^{(i_0,\ldots,i_k)}(\omega_i^{(\Delta t)}, f_{(i_0,\ldots,i_k)}), \tag{5.15}
$$

and we set $\beta_0(x) := Ax + V_0(x)$, $\beta_j(x) := V_j(x)$, $j = 1,\ldots,d$, and $f_{(i_0,\ldots,i_k)} := \beta_{i_0} \ldots \beta_{i_k} f$, $(i_0,\ldots,i_k) \in \{0,\ldots,d\}^{k+1}$. Summing up, it is easy to see by the scaling of the cubature paths that we can find a remainder term $(\Delta t)^{k+1}$ as in the claim of the theorem with the correct estimates. To see that the initial terms have the form given, we use the degree $2k+1$ of the cubature and the explicit formula of $\mathcal{G}$ from Theorem 5.3. A density argument proves the result. $\qquad\square$

5.3 The rate of convergence

We can now present our main result.

Theorem 5.8 *For* $f \in \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}(H_{\ell-(k+1)})$, $k+1 \le \ell \le \ell_0$, $n > 2(k+1)$, $2(k+1) \le k_0$,

$$
\| P_T f - Q_{(T/n)}^n f \|_{\psi_\ell^n} \le C_T n^{-k} \| f \|_{\psi_{\ell-(k+1)}^{(n)}, 2(k+1)} \tag{5.16}
$$

with a constant C_T *independent of f.*

Proof The local estimate follows from a combination of Corollary 5.6 and Theorem 5.7. In the core of the argument lies the very definition of cubature formulas, where the expectations of iterated Brownian integrals is mimicked, see formula (4.1). The stability of $Q_{(T/n)}$ from Corollary 4.8 and the assumed invariance of $\mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}(H_{\ell-(k+1)})$ with respect to P_t prove the claim by means of the telescoping sum

$$
(P_T f - Q_{(T/n)}^n) f = \sum_{i=1}^{n-1} \left(Q_{(\frac{T}{n})} \right)^{(n-i)} \left(P_{(\frac{T}{n})} - Q_{\frac{t}{n}} \right) P_{\frac{iT}{n}} f \tag{5.17}
$$

for $f \in \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}(H_{\ell-(k+1)})$. $\qquad\square$

Example 5.9 Let $H_\ell = \mathbb{R}^N$, $\ell \geq 0$, be finite-dimensional, and assume $n \geq 5$; in the finite-dimensional setting, we do not need to consider subspaces of the state space. Then, $f_{m,i} \in \mathcal{B}_k^{\psi^{(n)}}(H)$ for all $k \geq 0$, where $f_{m,i}(y) = y_i^m$, $i = 1, \ldots, n$, $y = (y_1, \ldots, y_N)$, and $m = 1, \ldots, 4$. This implies that not only the expected value and the variance, but also the skewness and kurtosis are accurately computed by our scheme. Similarly, mixed moments are determined to high accuracy, and if n is even larger, this also holds true for higher degree moments. Such a property is very useful in risk management, where high precision in higher moments means an accurate evaluation of risk. Similar observations were made in [1] and [33].

Example 5.10 The Heath–Jarrow–Morton framework is included in our setup. As explained in [12], it is more natural to use cosh-weighted spaces instead of poly-nomially weighted spaces in this case. The more general definition of vector fields in Definition 3.1 allows us to enlarge the class of admissible equations considerably compared to [12].

6 Numerical example

The FitzHugh–Nagumo model is a popular model for the behaviour of neurons and allows the analysis of noise-induced phase changes. We consider the SODE model with multiplicative and additive noise from [34], spatially extended similarly as in [36]. The stochastic partial differential equation we analyse reads

$$\varepsilon \mathrm{d}u(t, x) = \frac{\partial^2}{\partial x^2} u(t, x) \mathrm{d}t + (u(t, x)(u(t, x) - 1)(a - u(t, x)) - bv(t, x)) \mathrm{d}t$$
$$+ 2\alpha u(t, x)(u(t, x) - 1) \circ \mathrm{d}W_t^{d-1}, \tag{6.1a}$$
$$\mathrm{d}v(t, x) = (z(u(t, x) - c) - v(t, x)) \mathrm{d}t$$
$$+ 2\beta u(t, x) \circ \mathrm{d}W_t^d + 2\gamma \sum_{j=1}^{d-2} \varphi_j(x) \circ \mathrm{d}W_t^j, \tag{6.1b}$$

complemented by Neumann boundary conditions, $\frac{\partial}{\partial x} u(t, 0) = \frac{\partial}{\partial x} u(t, 1) = 0$, $t > 0$, and initial values, $u(0, x) = u_0(x)$, $v(0, x) = v_0(x)$, $x \in (0, 1)$. This implies that there is no signal, i.e., no deterministic inhomogeneous Neumann boundary condition or forcing, so the only reason for phase changes, i.e., deflecting the system out of equilibrium, comes from the multiplicative noise terms $2\alpha u(t, x)(u(t, x) - 1) \circ \mathrm{d}W_t^{d-1}$ and $2\beta u(t, x) \circ \mathrm{d}W_t^d$ or the additive noise term $2\gamma \sum_{j=1}^{d-2} \varphi_j(x) \circ \mathrm{d}W_t^j$. Here, u denotes the fast, v the slow variable, $(W_t^1, \ldots, W_t^d)_{t \geq 0}$ is d-dimensional Brownian motion, $\varepsilon > 0$, $a \in (0, 1)$, $b, c, z, \alpha, \beta, \gamma > 0$, and $\varphi_j : [0, 1] \to \mathbb{R}$ are some smooth functions.

While the Nemytskii nonlinearities in the equation for u are non-Lipschitz, they can be cut off in a smooth manner to Lipschitz continuous functions outside a large enough interval including [0,1]. This approach seems to be reasonable, as we are interested in the behaviour of the system only for functions u with values in this interval where the stable equilibria of the corresponding deterministic system lie. Furthermore, in our

numerical examples, we never observed values for u outside $[-5 \cdot 10^{-3}, 1 + 5 \cdot 10^{-3}]$. It then follows that (6.1) has solutions lying in all Sobolev spaces $H^{\ell}(0, 1), \ell \in \mathbb{N}$, provided the initial values for u and v and the functions $\varphi_j, j = 1, \ldots, d - 2$, are regular enough. Hence, (6.1) falls into the setting of our numerical schemes: Using the notation from Sect. 5, we set

$$
H_{\ell} := \left\{ (u, v) : u, v \in H^{1+2\ell}(0, 1) \text{ with } \frac{\partial^{1+2i}}{\partial x^{1+2i}} u(0) \right.
$$
$$
\left. = \frac{\partial^{1+2i}}{\partial x^{1+2i}} u(1) = 0 \text{ for } i = 1, \ldots, \ell \right\}, \tag{6.2}
$$

$A(u, v) := (\frac{\partial^2}{\partial x^2} u, 0)$, together with Neumann boundary conditions, and

$$
V_0(u, v) := (u(u - 1)(a - u) - bv, z(u - c) - v), \tag{6.3}
$$
$$
V_j(u, v) := (0, 2\gamma \varphi_j), \quad j = 1, \ldots, d - 2, \tag{6.4}
$$
$$
V_{d-1}(u, v) := (2\alpha u(u - 1), 0), \quad \text{and} \tag{6.5}
$$
$$
V_d(u, v) := (0, 2\beta u). \tag{6.6}
$$

After cutting off V_0 and V_{d-1} as suggested above, all these vector fields satisfy Definition 3.1 by applying Runst and Sickel [30, p. 381, Theorem 2; p. 32, Theorem 1]. It follows that we have optimal weak rates of convergence for sufficiently smooth test functions. If rates of convergence are to be analysed without cutting off the vector fields, an approach similarly as in Dörsek [10] using a quickly growing weight function might be applicable. We stress that this problem, not even with cut off vector fields, can be analysed using the theory from Dörsek [10] or Dörsek and Teichmann [12], as Nemytskii operators are not contained in either setting. Similarly, the results of Bayer and Teichmann [2, Sect. 4] are not applicable because the vector fields V_{d-1} and V_d do not have any smoothing properties.

In order to gauge phase transitions, we want to determine the probability that the system ends up in one stable equilibrium, given that it started in the other. Fixing the parameters to be $\varepsilon = 0.01, a = c = 0.47, b = 0.1$, and $z = 1.0$, as in Tang et al. [34], the two stable equilibria are given approximately by $(u_1^{\infty}, v_1^{\infty}) = (.112702, -.357298)$ and $(u_2^{\infty}, v_2^{\infty}) = (.887298, .417298)$. We choose the initial value $u_0(x) := u_1^{\infty}, v_0(x) := v_1^{\infty}, x \in (0, 1)$. The two functions whose expected values we approximate are $f_i(u, v) := \exp(-\|u - u_i^{\infty}\|_{L^2(0,1)}^2 / \delta), i = 1, 2$, with $\delta = 0.05$, which serve as proxy for the probability that u ends up in the state $u_i^{\infty}, i = 1, 2$. These functions are smooth and bounded on $H^{\ell}(0, 1)$ for $\ell \geq 0$, and by the compact embedding $H^{\ell+1}(0, 1) \to H^{\ell}(0, 1), \ell \geq 0$, and Dörsek and Teichmann [12, Theorem 5], they are included in $\mathcal{B}_k^{\psi}(H^{\ell})$ for $\ell \geq 0$, where the weight functions are chosen as in Sect. 5. We set $\alpha = 0.008, \beta = 0.0001, \gamma = 0.003, d = 6$, and $\varphi_j(x) = \cos((j - 1)\pi x)$, i.e., we force the low modes, and want to determine

$$
\mathbb{E}[f_i(u(T, \cdot))], \quad i = 1, 2, \tag{6.7}
$$

for $T = 10$.

For the space discretisation, we employ a cosine expansion, writing

$$u(t, x) = \sum_{i=0}^{M-1} u_i(t) \cos(i\pi x) \quad \text{and} \quad v(t, x) = \sum_{i=0}^{M-1} v_i(t) \cos(i\pi x), \qquad (6.8)$$

which is efficient due to forcing of low modes. In the time variable, we first perform a Strang-type splitting. The split equations read

$$\varepsilon du^0(t, x) = \left(\frac{\partial^2}{\partial x^2} u^0(t, x) + u^0(t, x)(u^0(t, x) - 1)(a - u^0(t, x)) - bv^0(t, x) \right) dt,$$

$$dv^0(t, x) = (z(u^0(t, x) - c) - v^0(t, x))dt, \quad \text{and} \qquad (6.9)$$

$$\varepsilon du^1(t, x) = 2\alpha u^1(t, x)(u^1(t, x) - 1) \circ dW_t^{d-1},$$

$$dv^1(t, x) = 2\beta u^1(t, x) \circ dW_t^d + 2\gamma \sum_{j=1}^{d-2} \varphi_j(x) \circ dW_t^j. \qquad (6.10)$$

The equations for (u^1, v^1) are solved by a degree 5 cubature scheme on path level, cf. Lyons and Victoir [24] (see Gyurkó and Lyons [15] for an alternative degree 5 scheme on flow level), which is evaluated via Quasi Monte Carlo integration. The equations for (u^0, v^0) are split further into

$$\varepsilon du^{01}(t, x) = \frac{\partial^2}{\partial x^2} u^{01}(t, x)dt,$$

$$dv^{01}(t, x) = 0, \quad \text{and} \qquad (6.11)$$

$$\varepsilon du^{02}(t, x) = (u^{02}(t, x)(u^{02}(t, x) - 1)(a - u^{02}(t, x)) - bv^{02}(t, x))dt,$$

$$dv^{02}(t, x) = (z(u^{02}(t, x) - c) - v^{02}(t, x))dt. \qquad (6.12)$$

Due to the particular choice of the spatial discretisation, (6.11) can be solved analytically. As we want to use large timesteps, we choose a geometric integrator for the approximation of (6.12), the extended Störmer–Verlet scheme from Hairer et al. [16, Sect. 1.8]. The resulting non-linear equations are solved by the Newton algorithm. This is efficient, as (6.12) corresponds to entirely separated ODEs at the discretisation points, whence the non-linear equations to be solved are one-dimensional.

Finally, the split problems are concatenated using the symmetrically weighted sequential splitting, see Oshima et al. [27]. Hence, we expect to observe second order weak convergence. This is illustrated in Fig. 1, where we show the somehow more expressive relative errors with respect to the reference value (in contrast to Theorem 5.8, where errors of norms relative to the norm of the *initial value* are estimated)

$$\frac{|\mathbb{E}_{\text{app}}[f_i(u(T, \cdot))] - \mathbb{E}_{\text{ref}}[f_i(u(T, \cdot))]|}{\mathbb{E}_{\text{ref}}[f_i(u(T, \cdot))]}, \quad i = 1, 2. \qquad (6.13)$$
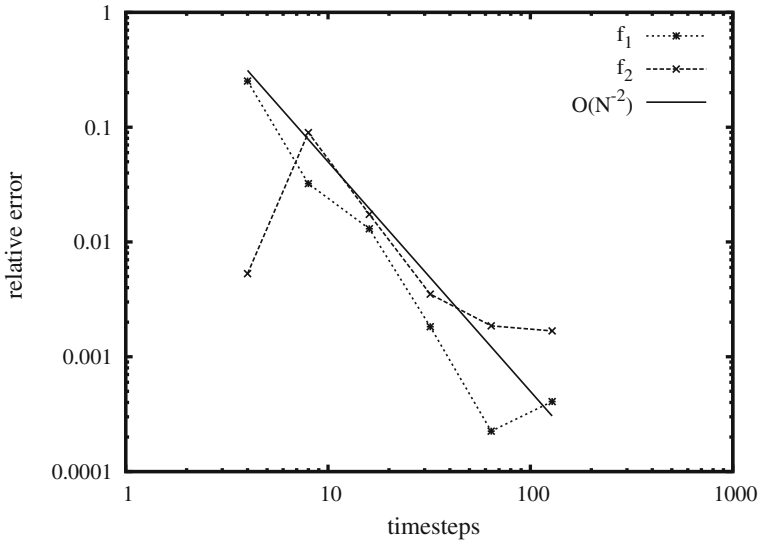
Fig. 1 Errors in weak approximation of the stochastic FitzHugh–Nagumo equation

The computations were performed using 16 cores of a ProLiant DL385 G7 with 24 AMD Opteron 6174 cores and 128 GB RAM in total. The reference values $\mathbb{E}_{\mathrm{ref}}[f_i(u(T, \cdot))]$ were obtained using $M = 128$ terms in the cosine expansions, $N = 256$ timesteps, and $K = 2^{20}$ Quasi Monte Carlo points, applying the Sobol' sequences of Joe and Kuo [18]. They were found to be $\mathbb{E}_{\mathrm{ref}}[f_1(u(T, \cdot))] \approx 0.471830$ and $\mathbb{E}_{\mathrm{ref}}[f_2(u(T, \cdot))] \approx 0.397405$. In the computation of the approximate values $\mathbb{E}_{\mathrm{app}}[f_i(u(T, \cdot))]$, we fixed $M = 32$ and $K = 2^{16}$, and used the number N of timesteps indicated in Fig. 1. The numerical results show second order convergence, perturbed probably because we compare not to an exact reference value, but to an approximate one. In particular, we obtain a relative error of less that $2 \cdot 10^{-3}$ with only 64 timesteps. This calculation takes 10 s, which establishes the effectivity of the proposed algorithm.

7 Conclusions

We considered the weak approximation of marginal distributions of stochastic (partial) differential equations. We extended the functional analytic framework of Röckner and Sobol [29], used for the numerical analysis of stochastic evolution equations in Dörsek [10] and Dörsek and Teichmann [11,12], to more general characteristics through a flexible formulation of directional derivatives in weighted spaces. This setting was then used to prove optimal rates of convergence of cubature schemes for more general equations. Results of numerical computations for an example from mathematical biology, a spatially extended stochastic FitzHugh–Nagumo model, were shown to demonstrate that our theoretical findings can be applied to practically relevant problems.

References

1. Alfonsi, A.: High order discretization schemes for the CIR process: Application to affine term structure and Heston models. Math. Comp. **79**(269), 209–237 (2010)
2. Bayer, C., Teichmann, J.: Cubature on Wiener space in infinite dimension. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **464**(2097), 2493–2516 (2008)
3. Blanes, S., Casas, F.: On the necessity of negative coefficients for operator splitting schemes of order higher than two. Appl. Numer. Math. **54**(1), 23–37 (2005)
4. Crisan, D., Ghazali, S.: On the Convergence Rates of a General Class of Weak Approximations of SDEs, Stochastic Differential Equations: Theory and Applications, vol. 2. Interdiscip. Math. Sci. World Sci. Publ., Hackensack (2007)
5. Crisan, D., Lyons, T.: Minimal entropy approximations and optimal algorithms. Monte Carlo Methods Appl. **8**(4), 343–355 (2002)
6. Da Prato, G., Zabczyk, J.: Ergodicity for Infinite-Dimensional Systems. In: London Mathematical Society Lecture Note Series, vol. 229. Cambridge University Press, Cambridge (1996)
7. Da Prato, G., Zabczyk, J.: Second Order Partial Differential Equations in Hilbert Spaces. In: London Mathematical Society Lecture Note Series, vol. 293. Cambridge University Press, Cambridge (2002)
8. Debussche, A.: Weak approximation of stochastic partial differential equations: The nonlinear case. Math. Comp. **80**(273), 89–117 (2011)
9. Dörsek, P.: Numerical Methods for Stochastic Partial Differential Equations, PhD thesis, Vienna University of Technology (2011)
10. Dörsek, P.: Semigroup splitting and cubature approximations for the stochastic Navier–Stokes equations. SIAM J. Numer. Anal. **50**(2), 729–746 (2012)
11. Dörsek, P., Teichmann, J.: A semigroup point of view on splitting schemes for stochastic (partial) differential equations, ArXiv e-prints (2010)
12. Dörsek, P., Teichmann, J.: Efficient simulation and calibration of general HJM models by splitting schemes, SIAM J. Finan. Math. **4**(1), 575–598 (2011)
13. Giles, M.B.: Multilevel Monte Carlo path simulation. Oper. Res. **56**(3), 607–617 (2008)
14. Giles, M. B., Szpruch, L.: Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation, ArXiv e-prints (2012)
15. Gyurkó, L.G., Lyons, T.J.: Efficient and practical implementations of cubature on Wiener space. In: Crisan, D. (ed.) Stochastic Analysis 2010, pp. 73–111. Springer, Berlin (2011)
16. Hairer, E., Lubich, C., Wanner, G.: Geometric numerical integration illustrated by the Störmer–Verlet method. Acta Numer. **12**, 399–450 (2003)
17. Jentzen, A., Kloeden, P.E.: The numerical approximation of stochastic partial differential equations. Milan J. Math. **77**, 205–244 (2009)
18. Joe, S., Kuo, F.Y.: Constructing Sobol' sequences with better two-dimensional projections. SIAM J. Sci. Comput. **30**(5), 2635–2654 (2008)
19. Kallenberg, O.: Foundations of Modern Probability, Probability and its Applications. Springer, New York (1997)
20. Kloeden, P.E., Platen, E.: Numerical solution of stochastic differential equations. In: Applications of Mathematics (New York), vol. 23. Springer, Berlin (1992)
21. Kusuoka, S.: Approximation of expectation of diffusion process and mathematical finance, Taniguchi Conference on Mathematics Nara '98, vol. 31, pp. 147–165. Adv. Stud. Pure Math. Math. Soc., Japan, Tokyo (2001)
22. Kusuoka, S.: Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus, Advances in mathematical economics, vol. 6, pp. 69–83. Adv. Math. Econ., Springer, Tokyo (2004)
23. Litterer, C., Lyons, T.: Cubature on Wiener Space Continued, Stochastic Processes and Applications to Mathematical Finance. World Science Publishing, Hackensack (2007)
24. Lyons, T., Victoir, N.: Cubature on Wiener space. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci **460**, 169–198 (2004)
25. Ninomiya, M., Ninomiya, S.: A new higher-order weak approximation scheme for stochastic differential equations and the Runge–Kutta method. Finance Stoch. **13**(3), 415–443 (2009)
26. Ninomiya, S., Victoir, N.: Weak approximation of stochastic differential equations and application to derivative pricing. Appl. Math. Finance **15**(1–2), 107–121 (2008)

27. Oshima, K., Teichmann, J. Velušček, D.: A new extrapolation method for weak approximation schemes with applications, Ann. Appl. Probab. (to appear) (2011)
28. Riesz, F., Sz.-Nagy, B.: Leçons d'analyse fonctionnelle, Gauthier-Villars, Paris. 3ème éd. (1955)
29. Röckner, M., Sobol, Z.: Kolmogorov equations in infinite dimensions: Well-posedness and regularity of solutions, with applications to stochastic generalized Burgers equations. Ann. Probab. **34**(2), 663–727 (2006)
30. Runst, T., Sickel, W.: Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations. In: de Gruyter Series in Nonlinear Analysis and Applications, vol. 3. Walter de Gruyter & Co., Berlin (1996)
31. Schmeiser, C., Soreff, A., Teichmann, J.: Recombination of Cubature Trees for the Weak Solution of SDEs. Unpublished note (2007)
32. Talay, D., Tubaro, L.: Expansion of the global error for numerical schemes solving stochastic differential equations. Stoch. Anal. Appl. **8**(4), 483–509 (1990)
33. Tanaka, H., Kohatsu-Higa, A.: An operator approach for Markov chain weak approximations with an application to infinite activity Lévy driven SDEs. Ann. Appl. Probab. **19**(3), 1026–1062 (2009)
34. Tang, J., Jia, Y., Yi, M., Li, J.: Multiplicative-noise-induced coherence resonance via two different mechanisms in bistable neural models. Phys. Rev. E Stat. Nonlinear Soft Matter Phys. **77**(6), 061905 (2008)
35. Teichmann, J.: Another approach to some rough and stochastic partial differential equations. Stoch. Dynam. **11**(2–3), 535–550 (2011)
36. Tuckwell, H.C.: Analytical and simulation results for the stochastic spatial FitzHugh–Nagumo model neuron. Neural Comput. **20**(12), 3003–3033 (2008)