# Strong convergence rates of probabilistic integrators for ordinary differential equations

H. C. Lie*      A. M. Stuart†      T. J. Sullivan‡

March 30, 2017

**Abstract:** Probabilistic integration of a continuous dynamical system is a way of systematically introducing model error, at scales no larger than errors inroduced by standard numerical discretisation, in order to enable thorough exploration of possible responses of the system to inputs. It is thus a potentially useful approach in a number of applications such as forward uncertainty quantification, inverse problems, and data assimilation. We extend the convergence analysis of probabilistic integrators for deterministic ordinary differential equations, as proposed by Conrad et al. (*Stat. Comput.*, 2016), to establish mean-square convergence in the uniform norm on discrete- or continuous-time solutions under relaxed regularity assumptions on the driving vector fields and their induced flows. Specifically, we show that randomised high-order integrators for globally Lipschitz flows and randomised Euler integrators for dissipative vector fields with polynomially-bounded local Lipschitz constants all have the same mean-square convergence rate as their deterministic counterparts, provided that the variance of the integration noise is not of higher order than the corresponding deterministic integrator.

**Keywords:** probabilistic numerical methods, ordinary differential equations, convergence rates, dissipative systems, Burkholder–Davis–Gundy inequalities, uncertainty quantification

**2010 Mathematics Subject Classification:** 65L20, 65C99, 37H10, 68W20

## 1 Introduction

The recent work of Conrad et al. (2016) proposed the use of probabilistic solvers for deterministic ordinary differential equations / initial value problems of the form

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) = f(u(t)), \qquad\qquad \text{for } t \geq 0, \qquad\qquad (1.1)$$

$$u(0) = u_0,$$

for a trajectory $[0, T] \ni t \mapsto u(t) \in \mathbb{R}^n$. The role of the stochasticity in such solvers is to systematically introduce and probe the model error that has been introduced by the discretisation, in order to enable thorough exploration of possible responses of the system to inputs. Probabilistic numerical solution of (1.1) is thus a potentially useful methodology in a number of applied scenarios such as forward uncertainty quantification (Smith, 2014; Sullivan, 2015), inverse problems (Kaipio and Somersalo, 2005; Stuart, 2010), and data assimilation (Law et al., 2015; Reich and Cotter, 2015). Just as with classical numerical analysis of deterministic integration schemes, it is natural to analyse the accuracy and convergence properties of probabilistic solvers for the system (1.1), in order to properly quantify the effect upon subsequent uncertainties.

---

*Institute of Mathematics, Free University of Berlin, Arnimallee 6, 14195 Berlin, Germany, hlie@math.fu-berlin.de

†Department of Computing and Mathematical Sciences, California Institute of Technology, CA 91125, United States of America, astuart@caltech.edu

‡Institute of Mathematics, Free University of Berlin, and Zuse Institute Berlin, Takustrasse 7, 14195 Berlin, Germany, sullivan@zib.de

## 1.1 Contribution and outline of the paper

One of the key results in the earlier work (Conrad et al., 2016, Theorem 2.2) was a convergence result for the error between the random values $U_k$ of a discrete-time numerical solution at discrete times $t_k := k\tau$, $\tau > 0$, and the corresponding values $u_k := u(k\tau)$ of the exact solution:

$$\max_{0 \leq k\tau \leq T} \mathbb{E}\big[\|u_k - U_k\|^2\big] \leq C\tau^{2p \wedge 2q}, \tag{1.2}$$

along with an analogous result in continuous time with the same exponent but possibly different constant. Here, loosely speaking (a more precise definition will be given at the beginning of Section 3), $\tau^q$ is the global order of accuracy of a deterministic method underlying $U_k$ and the variance of a Gaussian model $\xi_k$ for the truncation error over a time horizon $[t_k, t_{k+1}]$ of length $\tau$ scales like $\tau^{1+2p}$. This result can be interpreted as saying that the choice $p = q$ introduces the maximum amount of solution uncertainty consistent with preserving the order of accuracy of the underlying deterministic integrator.

The purpose of this paper is to extend the convergence analysis of such probabilistic numerical integrators in the following ways:

1. we extend the setting of the initial value problem (1.1) from $\mathbb{R}^n$ to a Hilbert space $\mathcal{H}$;
2. we relax the assumption that all deviations are Gaussian, and work directly with conditions on polynomial moments;
3. we bring the time supremum in (1.2) inside the expectation to yield

$$\mathbb{E}\left[\max_{0 \leq k\tau \leq T} \|u_k - U_k\|^2\right] \leq C\tau^{2p \wedge 2q}, \tag{1.3}$$

   so that the mode of convergence is now the stronger case of convergence in mean square with respect to the uniform norm on path space; and
4. the assumption that the vector field $f$ is globally Lipschitz is weakened in two ways: for integrators of arbitrary order, we consider Lipschitz flows; for Euler integrators, which have $q = 1$, we consider dissipative vector fields with polynomially-growing locally Lipschitz constant.

The outline of the paper, therefore, is as follows:

- Section 2 establishes some notation and recalls some basic auxiliary results.
- Section 3 discusses the convergence properties of randomised one-step solvers of (1.1) of arbitrary order when the true flow $\Phi^t$ is globally Lipschitz. We prove convergence with the time supremum outside the expectation in Theorem 3.4 and with the time supremum inside the expectation in Theorem 3.5.
- Section 4 considers the randomised implicit Euler method for vector fields that are locally Lipschitz with polynomially bounded local Lipschitz coefficient and that satisfy a generalised dissipativity property. The convergence result, namely mean-square-uniform convergence of order one, is split across Theorems 4.19 (discrete time) and 4.23 (continuous time).
- Section 5 adapts the arguments of Section 4 to the explicit Euler method. Well-known instabilities in the explicit Euler method, similar to those that affect the Euler–Maruyama method for stochastic differential equations, mean that the convergence result (Theorem 5.8) is limited to the case of almost surely bounded noise, which would be unnatural in an SDE context but is not unduly restrictive when the noise is to be interpreted as a model for discretisation error.

It is particularly noteworthy that the results in Sections 3–5 all show the same convergence *rate* (1.3) as the rate (1.2) shown by Conrad et al. (2016), so the bounds differ only in the constant prefactor $C$ (as well as the weaker regularity assumptions and stronger mode of convergence).

We remark that the proof techniques used in Sections 4 and 5 are very similar to those used to analyse numerical methods for SDEs (Higham et al., 2002), but since the variance of the noise is smaller than in the Brownian case, we obtain a qualitatively different behaviour in the limit as $\tau \to 0$, i.e. the deterministic process $u$.

One interpretation of the convergence results presented here is that a natural scaling, as a function of the time step $\tau > 0$, of the model for the imperfectly known truncation error is $p \geq q$, in the sense that such a scaling will not disturb the familiar convergence rate of the deterministic method. However, there is a great deal of freedom in the choice of covariance structure — and indeed non-Gaussian structure — for $\xi$. It should also be acknowledged that systematic errors should be incorporated as a non-zero mean for $\xi$. We reserve for future work the discussion of runtime calibration of the noise structure.

## 1.2 Review of probabilistic numerical methods

Continuous relationships such as ODEs and PDEs are commonplace as forward models or Bayesian likelihoods in modern statistical inverse problems (Kaipio and Somersalo, 2005; Stuart, 2010), and in particular in data assimilation algorithms with critical everyday applications such as numerical weather prediction (Law et al., 2015; Reich and Cotter, 2015). The use of a discretised solver for such forward models is usually unavoidable in practice, but introduces an additional source of uncertainty both into forward propagation of uncertainty and into subsequent inferences. While the solution to the ODE/PDE may not be random in the frequentist sense, it is nonetheless only imperfectly known though the discretised numerical solution, and probability in the subjective or Bayesian sense is one appropriate means of representing this epistemic uncertainty, particularly if the ODE/PDE solution forms part of the forward model in a Bayesian inverse problem. Failure to properly account for discretisation errors and uncertainties can lead to biased, inconsistent, and over-confident inferences (Conrad et al., 2016).

Probabilistic numerical solutions to deterministic problems such as the solution of ODEs have a long history. Modern foundations for this field were laid by the work of Diaconis (1988), O'Hagan (1992) and Skilling (1992) under the term of "Bayesian numerical analysis". More recently, such ideas have received renewed attention under the term "probabilistic numerics" (Hennig et al., 2015; Cockayne et al., 2017): the discussion of probabilistic numerical methods for ordinary differential equations given by Schober et al. (2014), Conrad et al. (2016) and Chkrebtii et al. (2016) is particularly relevant here.

Also of interest in the field of probabilistic numerics, but not directly relevant to the present work, are probabilistic numerical methods for linear algebra (Hennig, 2015), optimisation (Gonzalez et al., 2016), partial differential equations (Cockayne et al., 2016; Owhadi, 2015, 2017; Raissi and Karniadakis, 2017), and quadrature (Briol et al., 2015). In particular, Cockayne et al. (2017) sets out some axiomatic foundations for probabilistic numerical methods broadly conceived, and in particular what it means for a probabilistic numerical method to be "Bayesian".

Randomised solution of ODEs has also been studied in the context of stochastic or rough differential equations. In the case of non-autonomous ODEs driven by Carathéodory vector fields — that are locally integrable in time and continuous in space — it has been observed that randomised Euler and Runge–Kutta methods outperform their deterministic counterparts: see e.g. Stengle (1990), Jentzen and Neuenkirch (2009), Kruse and Wu (2017) and the references therein.

We note that analysing the convergence properties of numerical solutions to (1.1) in terms of the approximation error for the solution, as in (1.2) and (1.3), is very much in the spirit of classical numerical analysis. For uncertainty quantification of the discretised solution of (1.1) as a stand-alone forward problem, this viewpoint is often sufficient. However, for applications in inference, including sequential inference problems arising in data assimilation and filtering, an alternative approach is to directly examine the impact of discretisation upon the quality of later inferences, e.g. as quantified by Bayes factors (Capistrán et al., 2016). There is also the well-established literature of information-based complexity and average-case analysis, with its greater emphasis on algorithmic aspects such as computational cost and optimal accuracy for given classes of information (Novak, 1988; Ritter, 2000; Traub and Woźniakowsi, 1980; Traub et al., 1983).

# 2 Setup and notation

## 2.1 Notation

The following notation will be used throughout this article:
- $\mathcal{H}$ will denote a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$.
- $(\Omega, \mathcal{F}, \mathbb{P})$ will be a probability space sufficiently rich to serve as a common domain of definition for all the random variables and stochastic processes under consideration, and $\mathbb{E}$ denotes expectation (integration over $\Omega$) with respect to $\mathbb{P}$.
- $\stackrel{\mathrm{d}}{=}$ denotes equality in distribution of random variables.
- $C$, $C'$, etc. will denote non-negative constants whose value may change from one occurence to the next, but will always be independent of any time step $\tau > 0$ used to numerically solve the ODE of interest.
- For real numbers $a$ and $b$, $a \wedge b$ denotes their minimum and $a \vee b$ their maximum.

- Lip($\Phi$) denotes the minimal Lipschitz constant of a function $\Phi$, defined on a subset of $\mathcal{H}$ and taking values in $\mathcal{H}$, i.e. Lip($\Phi$) is the least $L \geq 0$ such that

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\| \text{ for all } x, y \in \text{domain}(\Phi).$$

- $\mathbb{N}$ denotes the natural numbers beginning with 1, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

## 2.2 Problem statement

This paper considers the numerical solution of the initial value problem (1.1), where the solution $u(t)$ at time $t$ and the initial condition $u_0$ lie in $\mathcal{H}$, and the vector field $f \colon \mathcal{H} \to \mathcal{H}$ is given. It will be assumed throughout that the vector field $f$ is sufficiently smooth that (1.1) has a unique solution up to a fixed, deterministic terminal time $T > 0$.

Let $\Phi^t \colon \mathcal{H} \to \mathcal{H}$ denote the flow map associated to the initial value problem (1.1), i.e. for each choice of initial condition $u_0 \in \mathcal{H}$, $\Phi^t(u_0) := u(t)$, the value of the solution to (1.1) at a later time $t \geq 0$. Obviously, $\Phi^0$ is the identity map, and later on it will be necessary to quantify the deviations of $\Phi^t$ from the identity for small positive $t$.

For simplicity, we consider the numerial solution of the problem (1.1) on the fixed, deterministic time interval $[0, T]$ using a constant time step $\tau > 0$ such that $K := T/\tau \in \mathbb{N}$, and define

$$t_k := k\tau \text{ for } k \in [K] := \{0, 1, \dots, K\} \tag{2.1}$$

We shall sometimes abuse notation and write $[K] = \{0, 1, \dots, K-1\}$ or $[K] = \{1, 2, \dots, K\}$. Let $u_k := u(t_k) \equiv \Phi^\tau(u_{k-1})$ denote the value of the exact solution to (1.1) at time $t_k$. We also consider a one-step numerical integration method, and hence a numerical flow map $\Psi^\tau \colon \mathcal{H} \to \mathcal{H}$, which gives an approximation $U_k$ to $u_k$ in terms of an approximation at the previous time step by $U_k := \Psi^\tau(U_{k-1})$. This setting encompasses many of the time-stepping methods in common use, such as Runge–Kutta methods of all orders. For simplicity we shall assume that $U_0 = u_0$, so that there is no initial error. We will also consider numerical approximations that provide continuous-time output $U(t)$, $t \in [0, T]$.

Classical numerical analysis seeks to quantify the accuracy of the integration method either in terms of its local order of accuracy, i.e.

$$\sup_{x \in \mathcal{H}} \|\Phi^\tau(x) - \Psi^\tau(x)\|,$$

or its global order of accuracy over the solution interval $[0, T]$, i.e.

$$\sup_{0 \leq k \leq T/\tau} \|u_k - U_k\|.$$

When the solution of (1.1) is merely a component of a larger statistical inference procedure, e.g. when it is desired to infer governing parameters or initial conditions of (1.1) from noisy partial observations of the solution at positive time, it is essential to understand the evolution of the error $u_k - U_k$ in order to arrive at accurate inferences. In this paper, we consider *random* approximations $U_k$, the quality of which will be assessed using the mean-square error, uniformly in time, i.e.

$$\mathbb{E}\left[\max_{0 \leq k\tau \leq T} \|u_k - U_k\|^2\right].$$

Indeed, we also construct probabilistic integrators that give continuous (random) output $U(t)$ for $t \in [0, T]$ and analyse the corresponding mean-square uniform error in continuous time:

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \|u(t) - U(t)\|^2\right].$$

## 2.3 Some useful inequalities

The analysis of this paper will make repeated use of several useful inequalities, which are collected here for reference. First, recall Young's inequality for real numbers $a, b \geq 0$, exponents $r, s > 0$ that are Hölder conjugate (i.e. for which $\frac{1}{r} + \frac{1}{s} = 1$), and $\delta > 0$:

$$ab \leq \frac{\delta}{r}a^r + \frac{1}{s\delta^{s/r}}b^s. \tag{2.2}$$

4

The special case $r = s = 2$ is sometimes known as the "Peter–Paul inequality". Combining that inequality with the Cauchy–Schwarz inequality in the Hilbert space $\mathcal{H}$ yields that, for all $x, y \in \mathcal{H}$ and $\delta > 0$,

$$\|x - y\|^2 \leq (1 + \delta)\|x\|^2 + (1 + \delta^{-1})\|y\|^2, \tag{2.3}$$

which will often be used either with $\delta = 1$ or $\delta = \tau$.

The following discrete-time version of Grönwall's inequality will also be useful:

**Theorem 2.1** (Discrete Grönwall inequality: Holte, 2009). *Let $(x_k)_{k \in \mathbb{N}_0}$, $(\alpha_k)_{k \in \mathbb{N}_0}$, and $(\beta_k)_{k \in \mathbb{N}_0}$ be non-negative sequences. If*

$$x_k \leq \alpha_k + \sum_{0 \leq j < k} \beta_j x_j \quad \text{for all } k \in \mathbb{N}_0, \tag{2.4}$$

*then, for all $k \in \mathbb{N}_0$,*

$$x_k \leq \alpha_k + \sum_{0 \leq j < k} \alpha_j \beta_j \prod_{j < i < k} (1 + \beta_i) \tag{2.5}$$

$$\leq \alpha_k + \sum_{0 \leq j < k} \alpha_j \beta_j \exp\left( \sum_{j < i < k} \beta_i \right). \tag{2.6}$$

*A fortiori, if $\alpha_k \leq A$ for all $k \in \mathbb{N}_0$, then*

$$x_k \leq A \prod_{0 \leq j < k} (1 + \beta_j) \leq A \exp\left( \sum_{0 \leq j < k} \beta_j \right). \tag{2.7}$$

# 3 High-order integration of Lipschitz flows

The purpose of this section is to establish a mean-square convergence result for randomised time-stepping integrators for (1.1) when the true flow is globally Lipschitz, the underlying deterministic method is a generic one with a known order of convergence, and the noise model has sufficiently small variance — these assumptions are formalised in Assumptions 3.1, 3.2, and 3.3. Under these weakened assumptions, we obtain that the randomised integrator has the same order of convergence as that obtained by Conrad et al. (2016), but in the stronger topology of uniform convergence of paths over the compact time horizon $[0, T]$, both in discrete and continuous time.

As just indicated, the first assumption is one about the exact flow map $\Phi^t$ associated to the problem (1.1). Indeed, in this section, the vector field $f$ is irrelevant and we may work purely in terms of the semigroup $(\Phi^t)_{t \geq 0}$.

**Assumption 3.1** (re: exact flow). Suppose that $f$ is smooth enough that, for $|t|$ small enough, its flow map $\Phi^t$ is globally Lipschitz with Lipschitz constant $\text{Lip}(\Phi^t) \leq 1 + C|t|$.

As is well known, Assumption 3.1 holds if the generating vector field $f$ is itself globally Lipschitz. However, Assumption 3.1 holds if, for instance, $f$ merely satisfies the one-sided Lipschitz inequality

$$\langle f(x) - f(y), x - y \rangle \leq \mu \|x - y\|^2 \tag{3.1}$$

for all $x, y \in \mathcal{H}$, for some constant $\mu \in \mathbb{R}$; in this case, a quick calculation of $\frac{d}{dt}\|u(t) - v(t)\|^2$ for trajectories $u$ and $v$ starting at initial conditions $u_0, v_0 \in \mathcal{H}$ and an application of Grönwall's inequality for differential inequalities show that $\|u(t) - v(t)\| \leq \exp(\mu|t|)\|u_0 - v_0\|$, so that $\text{Lip}(\Phi^t) \leq 1 + 2|\mu||t|$ for small enough $|t|$.

**Assumption 3.2** (re: numerical flow). Suppose that the numerical flow-map $\Psi^\tau$ has uniform local truncation error of order $q + 1$: for some constant $C \geq 0$,

$$\sup_{u \in \mathcal{H}} \|\Psi^\tau(u) - \Phi^\tau(u)\| \leq C\tau^{q+1}.$$

We now define a new randomised integrator

$$U_{k+1} := \Psi^\tau(U_k) + \xi_k(\tau),$$

where each $\xi_k(t)$, $t \in [0, \tau]$, is a continuous-time stochastic process which represents the numerical truncation error of the integrator $\Psi^\tau$. Note that this definition not only provides for forward propagation of the the numerical state $U_k$, but also continuous output via

$$U(t) := \Psi^{t-t_k}(U_k) + \xi_k(t - t_k) \quad \text{for } t \in [t_k, t_{k+1}).$$

On the other hand, if continuous output is not desired, then one can simply work with the random variables $\xi_k(\tau)$ and ignore the values of the stochastic process $(\xi_k(t))_{0 < t < \tau}$, and indeed this simplifies the analysis slightly in the sense that the following Assumption 3.3 is sufficient for discrete time but must be strengthened to (3.7) for continuous time.

**Assumption 3.3** (re: random perturbation). The $\xi_k$ are mutually independent and identically distributed mean-zero stochastic processes, and there are constants $C \geq 0$ and $p \geq 1$ such that, for all $k$ and all $t \in [0, \tau]$,

$$\mathbb{E}\big[\|\xi_k(t) \otimes \xi_k(t)\|\big] \leq Ct^{2p+1},$$

and, in particular, $\mathbb{E}\big[\|\xi_k(t)\|^2\big] \leq Ct^{2p+1}$.

As noted in the introduction, the focus of this paper is on the convergence rate of the error $e_k := u_k - U_k$ and not on, say, the covariance operator of $e_k$, though that information is also important in applications. In keeping with that focus, Assumption 3.3 and similar assumptions later in the paper are only upper bounds, and we do not actually work with the covariance operator of $\xi_k$. The precise construction of stochastic models for discretisation and truncation error is an interesting topic in its own right at the interface of numerical analysis and probability, which this paper will not address.

## 3.1 Convergence result with time supremum outside expectation

As a warm-up for the main results of this paper, which have the time supremum inside the expectation, we first prove a weaker convergence result with the time supremum outside the expectation. The following result is analogous to Theorem 2.2 of Conrad et al. (2016), but does not require the vector field $f$ to be globally Lipschitz nor $\xi$ to be Gaussian.

**Theorem 3.4** (Uniform mean-square convergence). *Suppose that Assumptions 3.1, 3.2, and 3.3 hold. Then there exist constants $C \geq 0$ such that*

$$\max_{0 \leq k \leq K} \mathbb{E}\big[\|u_k - U_k\|^2\big] \leq C\tau^{2p \wedge 2q}, \tag{3.2}$$

$$\sup_{0 \leq t \leq T} \mathbb{E}\big[\|u(t) - U(t)\|^2\big] \leq C\tau^{2p \wedge 2q}. \tag{3.3}$$

*Proof.* Let $e_k := u_k - U_k$, so that

$$e_{k+1} = \big(\Phi^\tau(u_k) - \Phi^\tau(U_k)\big) - \big(\Psi^\tau(U^K) - \Phi^\tau(U_k)\big) - \xi_k(\tau). \tag{3.4}$$

Then

$$
\begin{aligned}
\|e_{k+1}\|^2 &= \big\|\big(\Phi^\tau(u_k) - \Phi^\tau(U_k)\big) - \big(\Psi^\tau(U_k) - \Phi^\tau(U_k)\big)\big\|^2 + \|\xi_k(\tau)\|^2 \\
&\quad + 2\langle\Phi^\tau(u_k) - \Psi^\tau(U_k), \xi_k(\tau)\rangle && \text{by (3.4)} \\
&\leq (1+\tau)\|\Phi^\tau(u_k) - \Phi^\tau(U_k)\|^2 + (1 + \tau^{-1})\|\Psi^\tau(U_k) - \Phi^\tau(U_k)\|^2 \\
&\quad + \|\xi_k(\tau)\|^2 + 2\langle\Phi^\tau(u_k) - \Psi^\tau(U_k), \xi_k(\tau)\rangle && \text{by (2.3) with } \delta = \tau \\
&\leq (1+\tau)(1+C\tau)^2\|e_k\|^2 + C\tau^{1+2q} \\
&\quad + \|\xi_k(\tau)\|^2 + 2\langle\Phi^\tau(u_k) - \Psi^\tau(U_k), \xi_k(\tau)\rangle && \text{by Assumptions 3.1 and 3.2,}
\end{aligned}
$$

and so
$$\|e_{k+1}\|^2 - \|e_k\|^2 \le C\tau\|e_k\|^2 + C\tau^{1+2q} + \|\xi_k(\tau)\|^2 + 2\langle \Phi^\tau(u_k) - \Psi^\tau(U_k), \xi_k(\tau)\rangle. \tag{3.5}$$

Now apply $\mathbb{E}[\,\cdot\,]$ to both sides of (3.5). Since $\xi_k(\tau)$ is independent of $U_k$, the expected value of the inner product term in (3.5) is zero, and so

$$\begin{aligned}
\mathbb{E}\big[\|e_{k+1}\|^2 - \|e_k\|^2\big] &\le C\tau\mathbb{E}\big[\|e_k\|^2\big] + C\tau^{1+2q} + \mathbb{E}\big[\|\xi_k(\tau)\|^2\big] \\
&\le C\tau\mathbb{E}\big[\|e_k\|^2\big] + C\tau^{1+(2p\wedge 2q)} && \text{by Assumption 3.3} \\
&= C\tau\sum_{j=0}^{k-1}\mathbb{E}\big[\|e_{j+1}\|^2 - \|e_j\|^2\big] + C\tau^{1+(2p\wedge 2q)} && \text{as a telescoping sum.}
\end{aligned}$$

Hence, by the triangle inequality,

$$\big|\mathbb{E}\big[\|e_{k+1}\|^2 - \|e_k\|^2\big]\big| \le C\tau^{1+(2p\wedge 2q)} + \sum_{j=0}^{k-1}\big|\mathbb{E}\big[\|e_{j+1}\|^2 - \|e_j\|^2\big]\big|.$$

Hence, the discrete Grönwall inequality (2.7) with $A = C\tau^{1+(2p\wedge 2q)}$ and every $\beta_j = C\tau$ implies that, for all $k < K = T/\tau$,

$$\big|\mathbb{E}\big[\|e_{k+1}\|^2 - \|e_k\|^2\big]\big| \le C\tau^{1+(2p\wedge 2q)}\exp(kC\tau) \le C\tau^{1+(2p\wedge 2q)},$$

and this proves (3.2) via a telescoping sum: for $k \le K = T/\tau$,

$$\mathbb{E}\big[\|e_k\|^2\big] \le \|e_0\|^2 + \sum_{j=0}^{k-1}\big|\mathbb{E}\big[\|e_{j+1}\|^2 - \|e_j\|^2\big]\big| \le \frac{T}{\tau}C\tau^{1+(2p\wedge 2q)} = C\tau^{2p\wedge 2q}.$$

For the result in continuous time, write $e(t) := u(t) - U(t)$, so that

$$\begin{aligned}
\|e(t)\|^2 = \;&\|\Phi^{t-t_k}(u_k) - \Psi^{t-t_k}(U_k)\|^2 \\
&+ 2\langle \Phi^{t-t_k}(u_k) - \Psi^{t-t_k}(U_k), \xi_k(t-t_k)\rangle \\
&+ \|\xi_k(t-t_k)\|^2.
\end{aligned}$$

Now take the conditional expectation of both sides, conditioned upon the $\sigma$-algebra $\mathcal{F}_k$ generated by the $\xi_k$ up to time $t_k$: since $\xi_k(t-t_k)$ is independent of $U_k$, the conditional expected value of the inner product vanishes, leaving

$$\begin{aligned}
\mathbb{E}\big[\|e(t)\|^2\big|\mathcal{F}_k\big] &= \mathbb{E}\big[\|\Phi^{t-t_k}(u_k) - \Psi^{t-t_k}(U_k)\|^2\big|\mathcal{F}_k\big] + \mathbb{E}\big[\|\xi_k(t-t_k)\|^2\big|\mathcal{F}_k\big] \\
&\le \mathbb{E}\big[\|\Phi^{t-t_k}(u_k) - \Psi^{t-t_k}(U_k)\|^2\big|\mathcal{F}_k\big] + C\tau^{2p+1}.
\end{aligned}$$

Now, inequality (2.3) with $\delta = 1$ yields

$$\begin{aligned}
&\mathbb{E}\big[\|\Phi^{t-t_k}(u_k) - \Psi^{t-t_k}(U_k)\|^2\big|\mathcal{F}_k\big] \\
&\le 2\mathbb{E}\big[\|\Phi^{t-t_k}(u_k) - \Phi^{t-t_k}(U_k)\|^2\big|\mathcal{F}_k\big] + 2\mathbb{E}\big[\|\Phi^{t-t_k}(U_k) - \Psi^{t-t_k}(U_k)\|^2\big|\mathcal{F}_k\big] \\
&\le 2(1+C\tau)^2\mathbb{E}\big[\|e_k\|^2\big|\mathcal{F}_k\big] + C\tau^{2q+2}
\end{aligned}$$

by Assumptions 3.1 and 3.2. Taking unconditional expectations of both sides yields, by (3.2),

$$\mathbb{E}\big[\|e(t)\|^2\big] \le C\tau^{2p\wedge 2q},$$

as claimed. $\qquad\square$

7

## 3.2 Convergence result with time supremum inside expectation

The next result strengthens Theorem 3.4 by bringing the time supremum inside the expectation. In the discrete-time case, this can be done without additional assumptions; for the continuous-time result, we strengthen Assumption 3.3 slightly.

**Theorem 3.5** (Mean-square uniform convergence)**.** *Suppose that Assumptions 3.1, 3.2, and 3.3 hold. Then there exists $C \geq 0$ such that*

$$\mathbb{E}\left[\max_{0 \leq k \leq K} \|u_k - U_k\|^2\right] \leq C\tau^{2p \wedge 2q}. \tag{3.6}$$

*If, in addition, Assumption 3.3 is strengthened to*

$$\mathbb{E}\left[\sup_{0 \leq t \leq \tau} \|\xi_0(t)\|^2\right] \leq C\tau^{1+2p}, \tag{3.7}$$

*then*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \|u(t) - U(t)\|^2\right] \leq C\tau^{2p \wedge 2q}. \tag{3.8}$$

*Proof.* The strategy is to delay the application of expectation to (3.5) until after a time supremum has been taken, and to estimate the inner temporal maximum using the Burkholder–Davis–Gundy inequality.

We apply the operation of $\mathbb{E}\left[\max_{k \leq \ell} \cdot\right]$, where $\ell \leq K = T/\tau$, to the telescoping sum

$$\|e_k\|^2 - \|e_0\|^2 = \sum_{j=0}^{k-1}\left(\|e_{j+1}\|^2 - \|e_j\|^2\right)$$

and insert (3.5) for each $\|e_{j+1}\|^2 - \|e_j\|^2$. Then, since $e_0 := u_0 - U_0 = 0$, it follows that

$$\mathbb{E}\left[\max_{k \leq \ell} \|e_k\|^2\right] \leq \mathbb{E}\left[\max_{k \leq \ell}\sum_{j=0}^{k-1}\left(\tau C\|e_j\|^2 + C\tau^{1+2q} + \|\xi_j(\tau)\|^2\right)\right]$$

$$+ 2\mathbb{E}\left[\max_{k \leq \ell}\left\|\sum_{j=0}^{k-1}\langle\Phi^\tau(u_j) - \Psi^\tau(U_j), \xi_j(\tau)\rangle\right\|\right].$$

Hence, by the Burkholder–Davis–Gundy inequality (Ren, 2008), with $[Y]_\ell$ denoting the quadratic variation up to time $\ell$ of a process $Y_k$:

$$\mathbb{E}\left[\max_{k \leq \ell} \|e_k\|^2\right] \leq \mathbb{E}\left[\sum_{j=0}^{\ell-1}\left(\tau C\|e_j\|^2 + C\tau^{1+2q} + \|\xi_j(\tau)\|^2\right)\right]$$

$$+ C\mathbb{E}\left[[\langle\Phi^\tau(u_\bullet) - \Psi^\tau(U_\bullet), \xi_\bullet(\tau)\rangle]_{\ell-1}^{1/2}\right]$$

$$\leq CT\tau^{2p \wedge 2q} + \tau C\sum_{j-0}^{\ell-1}\mathbb{E}\left[\|e_j\|^2\right]$$

$$+ C\mathbb{E}\left[\sqrt{\max_{j \leq \ell}\|\Phi^\tau(u_j) - \Psi^\tau(U_j)\|^2 \sum_{j=0}^{\ell-1}\|\xi_j(\tau) - \xi_{j-1}(\tau)\|^2}\right]$$

$$\leq CT\tau^{2p \wedge 2q} + \tau C\sum_{j-0}^{\ell-1}\mathbb{E}\left[\|e_j\|^2\right]$$

$$+ C\mathbb{E}\left[\max_{j \leq \ell}\|\Phi^\tau(u_j) - \Psi^\tau(U_j)\|\sqrt{\sum_{j=0}^{\ell-1}\|\xi_j(\tau) - \xi_{j-1}(\tau)\|^2}\right]$$

$$\leq \tau C \sum_{j=0}^{\ell-1} \mathbb{E}\big[\|e_j\|^2\big] + CT\tau^{2p \wedge 2q} + \frac{1}{2}\mathbb{E}\Big[\max_{k\leq\ell}\|e_k\|^2\Big]$$

$$\leq \tau C \sum_{j=0}^{\ell-1} \mathbb{E}\Big[\max_{k\leq j}\|e_k\|^2\Big] + CT\tau^{2p \wedge 2q} + \frac{1}{2}\mathbb{E}\Big[\max_{k\leq\ell}\|e_k\|^2\Big].$$

Hence, by the discrete Grönwall inequality applied to the sequence $\big(\mathbb{E}\big[\max_{k\leq\ell}\|e_k\|^2\big]\big)_{\ell\geq 0}$,

$$\mathbb{E}\Big[\max_{k\leq\ell}\|e_k\|^2\Big] \leq 2CT\tau^{2p \wedge 2q}\exp(2T\tau C/\tau) = C\tau^{2p \wedge 2q},$$

and this establishes (3.6).

In continuous time, for $t \in [t_k, t_{k+1}]$, the error satisfies

$$\begin{aligned}
e(t) &= u(t) - U(t) \\
&= \Phi^{t-t_k}(u_k) - \Psi^{t-t_k}(U_k) - \xi_k(t-t_k) \\
&= \Phi^{t-t_k}(u_k) - \Phi^{t-t_k}(U_k) + \Phi^{t-t_k}(U_k) - \Psi^{t-t_k}(U_k) - \xi_k(t-t_k),
\end{aligned}$$

and so

$$\|e(t)\|^2 \leq C(1 + C(t-t_k))^2\|e_k\|^2 + C(t-t_k)^{2+2q} + C\|\xi_k(t-t_k)\|^2$$

Therefore

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\leq t\leq T}\|e(t)\|^2\Big] &= \mathbb{E}\Big[\max_{0\leq k\leq K}\sup_{t_k\leq t\leq t_{k+1}}\|e(t)\|^2\Big] \\
&\leq C\tau^{2p \wedge 2q} + C\mathbb{E}\Big[\max_{0\leq k\leq K}\sup_{t_k\leq t\leq t_{k+1}}\|\xi_k(t-t_k)\|^2\Big]
\end{aligned}$$

by the discrete-time result (3.6). Thus, it suffices to show that

$$\mathbb{E}\Big[\max_{0\leq k\leq K}\sup_{t_k\leq t\leq t_{k+1}}\|\xi_k(t-t_k)\|^2\Big] \leq C\tau^{2p}.$$

To show this, observe that

$$\begin{aligned}
\mathbb{E}\Big[\max_{0\leq k\leq K}\sup_{t_k\leq t\leq t_{k+1}}\|\xi_k(t-t_k)\|^2\Big] &\leq \mathbb{E}\Big[\sum_{k=0}^{K}\sup_{t_k\leq t\leq t_{k+1}}\|\xi_k(t-t_k)\|^2\Big] \\
&= \sum_{k=0}^{K}\mathbb{E}\Big[\sup_{t_k\leq t\leq t_{k+1}}\|\xi_k(t-t_k)\|^2\Big] \\
&= \frac{T}{\tau}\mathbb{E}\Big[\sup_{0\leq t\leq \tau}\|\xi_0(t)\|^2\Big]
\end{aligned}$$

since the $\xi_k$ are i.i.d., and the claim now follows from (3.7). $\qquad\square$

The next result shows that the strengthened Assumption 3.3, (3.7), actually holds for at least one reasonable model of truncation error, namely one that corresponds to the assumption that the true solution $u$ has one continuous derivative and a second derivative can be modelled as a Brownian process.

**Proposition 3.6.** *The strengthened Assumption 3.3, (3.7), holds for $\xi$ modelled on the time integral of Brownian motion, $\xi_0(t) := \tau^{p-1}\int_0^t B(s)\,\mathrm{d}s$, where $B$ denotes a standard $\mathcal{H}$-valued Brownian motion.*

*Proof.* Recall that, for a standard Brownian motion $B$, $\int_0^T B(t)\,\mathrm{d}t$ is Gaussian $\mathcal{N}(0, \frac{1}{3}T^3)$, and, for $c > 0$, $B(ct) \overset{\mathrm{d}}{=} c^{1/2}B(t)$. Hence, in the special case $\xi_0(t) = \tau^{p-1}\int_0^t B(s)\,\mathrm{d}s$, $\xi_0(\tau)$ is Gaussian with mean zero and variance

$$\mathbb{E}\big[\|\xi_0(\tau)\|^2\big] = \tau^{2p-2}\frac{\tau^3}{3} = \frac{\tau^{2p+1}}{3}.$$

9

Hence,

$$\xi_0(t) = \tau^{p-1} \int_0^t B(s)\,\mathrm{d}s$$

$$= \tau^{p-1} \int_0^1 B(tr)t\,\mathrm{d}r$$

$$\overset{\mathrm{d}}{=} \tau^{p-1} t^{3/2} \int_0^1 B(r)\,\mathrm{d}r,$$

and so

$$\mathbb{E}\left[\sup_{0 \leq t \leq \tau} \|\xi_0(t)\|^2\right] = \tau^{2p-2}\mathbb{E}\left[\sup_{0 \leq t \leq \tau} t^3\left(\int_0^1 B(r)\,\mathrm{d}r\right)^2\right]$$

$$= \tau^{2p+1}\mathbb{E}\left[\left(\int_0^1 B(r)\,\mathrm{d}r\right)^2\right]$$

$$= \frac{\tau^{2p+1}}{3}. \qquad \qquad \square$$

See also Takács (1993) for further discussion of the distribution and moments of time-integrated Brownian motion.

# 4 Implicit Euler integration of locally Lipschitz vector fields

This section considers the numerical integration of much less well behaved flows, but with a specific choice of integrator, namely the implicit Euler scheme. Thus, the results of this section are simultaneously more general (in terms of $f$) and more specific (in terms of the choice of $\Psi^\tau$) than those of Section 3. Again, our aim is to establish mean-square convergence in the uniform norm on paths; in view of the fact that the Euler method has order one, our aim is to show, under suitable assumptions on $f$ and $\xi_k$, that

$$\mathbb{E}\left[\sup_{0 \leq k \leq T/\tau} \|u_k - U_k\|^2\right] \leq C\tau^2,$$

and similarly in continuous time.

In this section, we shall make the following assumptions concerning the vector field $f$, namely generalised dissipativity and a polynomial growth condition on local Lipschitz constants:

**Assumption 4.1** (Generalised dissipativity)**.** Constants $\alpha \geq 0$ and $\beta \in \mathbb{R}$ are such that

$$\langle f(v), v \rangle \leq \alpha + \beta \|v\|^2 \text{ for all } v \in \mathcal{H}. \tag{4.1}$$

Because we allow for positive values of $\beta$, Assumption 4.1 is more general than the usual dissipativity property found in the literature (Humphries and Stuart, 1994, Equation (1.2)). Recent studies in numerical methods for stochastic differential equations consider constraints on the drift and diffusion of the SDEs that feature the same right-hand side above, e.g. Fang and Giles (2016) or Mao and Szpruch (2013).

Furthermore, we assume that the vector field satisfies a polynomial growth condition; this condition may be seen as a kind of local Lipschitz property (Higham et al., 2002, Assumption 4.1).

**Assumption 4.2** (Polynomial growth condition)**.** Constants $D \geq 1$ and $q \in \mathbb{N}_0$ are such that

$$\|f(a) - f(b)\|^2 \leq D(1 + \|a\|^q + \|b\|^q)\|a - b\|^2 \text{ for all } a, b \in \mathcal{H}. \tag{4.2}$$

**Remark 4.3.** Note that the case $q = 0$ corresponds to the case that the vector field is globally Lipschitz. Note also that the meaning of $q$ has changed from Section 3: $q$ no longer denotes the global order of accuracy of the underlying deterministic integrator, which is now simply 1.

As before, we denote the solution map associated to a numerical method with time step $\tau$ by $\Psi^\tau \colon \mathcal{H} \to \mathcal{H}$. Given a sequence $(\xi_k)_k$ of i.i.d. copies of some stochastic process $\xi_0$, consider a sequence $(U_k)_k$ of approximations to the sequence of states of the flow $(u(s))_{s \in [0,T]}$ evaluated at intervals of length $\tau$, where

$$U_{k+1} := \Psi^\tau(U_k) + \xi_k(\tau), \quad k \in [K]. \tag{4.3}$$

As usual, we employ the following expression for the error $e_k := u_k - U_k$ between the exact deterministic solution and the randomised approximate solution at time $t_k$:

$$e_k = \Phi^\tau(u_{k-1}) - \Phi^\tau(U_{k-1}) + \Phi^\tau(U_{k-1}) - \Psi^\tau(U_{k-1}) - \xi_{k-1}(\tau).$$

In order to obtain uniform convergence in mean square for a randomised numerical scheme of the form (4.3), we shall need to provide estimates for $\Phi^\tau(u_{k-1}) - \Phi^\tau(U_{k-1})$, and $\Phi^\tau(U_{k-1}) - \Psi^\tau(U_{k-1})$.

## 4.1 Moment bounds

The results of this section concern the implicit Euler method $\Psi^\tau \colon \mathcal{H} \to \mathcal{H}$, which is implicitly defined by

$$\Psi^\tau(v) = v + \tau f(\Psi^\tau(v)). \tag{4.4}$$

In the sequel, we shall assume that $f$ is such that we can always find a sufficiently small $\tau^* > 0$ such that (4.4) is uniquely solvable for all $0 < \tau < \tau'$; all instances of $\tau$ shall hereafter satisfy this constraint. Given a collection $(\xi_k)_k$ of random variables, and given the map $\Psi^\tau$ defined by the implicit Euler scheme (4.4), we construct a randomised implicit Euler scheme according to (4.3).

We shall use the following lemmata throughout:

**Lemma 4.4** (Generalised triangle inequality). *Given arbitrary $N \in \mathbb{N}$ and $(s_j)_{j=1}^N \subset \mathbb{R}$, it holds that*

$$\left| \sum_{j=1}^N s_j \right|^p \leq 2^{(N-1)(p-1)} \sum_{j=1}^N |s_j|^p.$$

*Proof.* We prove the desired statements by induction. For $p = 2$ and $N = 2$, the statement follows from Cauchy's inequality:

$$|a + b|^2 = |a|^2 + |b|^2 + 2ab \leq |a|^2 + |b|^2 + 2\frac{1}{2}(|a|^2 + |b|^2) = 2(|a|^2 + |b|^2).$$

Let $N = 2$. Suppose that the statement holds for $p \leq 2$. To prove the statement for $p + 1$, observe that

$$
\begin{aligned}
|a + b|^{p+1} &\leq 2^{p-1}(|a|^p + |b|^p)|a + b| \\
&\leq 2^{p-1}(|a|^{p+1} + |b|^{p+1} + |a|^p|b| + |b|^p|a|)
\end{aligned} \tag{4.5}
$$

by the inductive hypothesis. Thus, it suffices to prove that

$$|a|^p|b| \leq \frac{p}{p+1}|a|^{p+1} + \frac{1}{p+1}|b|^{p+1}, \tag{4.6}$$

since by switching $a$ and $b$ in (4.6) we may obtain

$$|a|^p|b| + |b|^p|a| \leq |a|^{p+1} + |b|^{p+1},$$

which we may substitute into (4.5) to obtain the desired conclusion. Since (4.6) follows from Young's inequality, the proof of the statement for arbitrary $p \in \mathbb{N}$ and $N = 2$ follows. Now we allow $N$ to vary. Suppose that the desired statement holds for a fixed, arbitrary $p \in \mathbb{N}$, and for some $N \leq 2$. To show it applies for $N + 1$, observe that

$$\left| \sum_{j=1}^{N+1} s_j \right|^p = \left| \sum_{j=1}^N s_j + s_{N+1} \right|^p \leq 2^{p-1} \left( \left| \sum_{j=1}^N s_j \right|^p + |s_{N+1}|^p \right).$$

By the inductive statement,

$$2^{p-1}\left(\left|\sum_{j=1}^{N}s_j\right|^p + |s_{N+1}|^p\right) \le 2^{p-1}\left(2^{(N-1)(p-1)}\sum_{j=1}^{N}|s_j|^p + |s_{N+1}|\right) \le 2^{N(p-1)}\sum_{j=1}^{N+1}|s_j|^p,$$

which completes the proof. □

Recall that $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space on which we have defined all random variables. The lemma below applies (4.11) and Lemma 4.4 in order to derive useful estimates on the moments of $\max_{i\in[K]}\|U_i\|$. We make the following assumption:

**Assumption 4.5** (Sufficiently small integration time step). There exists $\kappa > 1$ such that $\tau$ satisfies

$$0 < \tau < \min\left\{1, \frac{\kappa-1}{2\kappa|\beta|}\right\} \implies 1 < \frac{1}{1-2|\beta|\tau} \le \kappa.$$

**Lemma 4.6** (Uniform bounds on numerical solution). *Suppose that the vector field $f$ satisfies Assumptions 4.1 and 4.2, and suppose that $\tau > 0$ satisfies Assumption 4.5. For all $n \in \mathbb{N}$, the numerical solution $(U_k)_{k\in[K]}$ generated by (4.3) and (4.4) satisfies*

$$\max_{i\in[T/\tau]}\|U_i\|^{2n} \le 2^{n-1}C^n\left[1 + \tau^{-n}\left(\sum_{i=1}^{T/\tau}\|\xi_i(\tau)\|^2\right)^n\right] \tag{4.7}$$

*uniformly in $\omega \in \Omega$, where*

$$C_1 = 2e^{\kappa T(1+2|\beta|)}\max\left\{1, \|U_0\|^2 + 2\alpha\kappa T\right\}. \tag{4.8}$$

*Proof.* In what follows, we shall omit the dependence of all random variables on $\omega$, with the understanding that $\omega$ is arbitrary. Let $n \in [K]$, where $K = T/\tau \in \mathbb{N}$. From (4.3) we have, by Young's inequality (2.2),

$$\|U_{n+1}\|^2 \le (1+\tau)\|\Psi^\tau(U_n)\|^2 + (1+\tau^{-1})\|\xi_n(\tau)\|^2. \tag{4.9}$$

Taking the inner product of (4.4) with $\Psi^\tau(U_n)$, we obtain

$$\begin{aligned}
\|\Psi^\tau(U_n)\|^2 &= \langle \Psi^\tau(U_n), U_n\rangle + \tau\langle f(\Psi^\tau(U_n)), \Psi^\tau(U_n)\rangle \\
&\le \frac{1}{2}\left(\|\Psi^\tau(U_n)\|^2 + \|U_n\|^2\right) + \tau\left(\alpha + \beta\|\Psi^\tau(U_n)\|^2\right) \quad\quad \text{by (4.1)} \\
&= \|\Psi^\tau(U_n)\|^2\left(\frac{1}{2} + \beta\tau\right) + \frac{1}{2}\|U_n\|^2 + \alpha\tau.
\end{aligned}$$

Thus,

$$\|\Psi^\tau(U_n)\|^2 \le \frac{1}{1-2\beta\tau}\left(\|U_n\|^2 + 2\alpha\tau\right) \le \frac{1}{1-2|\beta|\tau}\left(\|U_n\|^2 + 2\alpha\tau\right), \tag{4.10}$$

where we used the inequality $1 - 2|\beta|\tau \le 1 + 2\beta\tau$ for the second inequality. Then (4.9) and (4.10) yield

$$\|U_n\|^2 \le \frac{1+\tau}{1-2|\beta|\tau}\left(\|U_{n-1}\|^2 + 2\alpha\tau\right) + \frac{1+\tau}{\tau}\|\xi_{n-1}(\tau)\|^2. \tag{4.11}$$

By (4.11), it follows that

$$\|U_n\|^2 - \|U_{n-1}\|^2 \le \tau\frac{1+2|\beta|}{1-2|\beta|\tau}\|U_{n-1}\|^2 + (1+\tau)\left(\frac{2\alpha\tau}{1-2|\beta|\tau} + \tau^{-1}\|\xi_{n-1}(\tau)\|^2\right).$$

Using the telescoping sum

$$\|U_n\|^2 = \|U_0\|^2 + \sum_{i=1}^{n}\left(\|U_i\|^2 - \|U_{i-1}\|^2\right)$$

12

it follows that

$$\|U_n\|^2 \leq \|U_0\|^2 + \sum_{i=1}^{n} \left( \tau \frac{1+2|\beta|}{1-2|\beta|\tau} \|U_{i-1}\|^2 + (1+\tau) \left( \frac{2\alpha\tau}{1-2|\beta|\tau} + \tau^{-1} \|\xi_{i-1}(\tau)\|^2 \right) \right).$$

Since $n \leq K \coloneqq T/\tau$, and since the right-hand side of the inequality above is nonnegative,

$$\|U_n\|^2 \leq \left[ \|U_0\|^2 + (1+\tau) \left( \frac{2\alpha T}{1-2|\beta|\tau} + \tau^{-1} \sum_{i=1}^{T/\tau} \|\xi_{i-1}(\tau)\|^2 \right) \right] + \tau \frac{(1+2|\beta|)}{1-2\|\beta\tau\|} \sum_{i=0}^{n-1} \|U_i\|^2.$$

Applying the Grönwall inequality (Theorem 2.1), yields, for all $n \in [K]$,

$$\max_{i \in [K]} \|U_i\|^2 \leq \left[ \|U_0\|^2 + (1+\tau) \left( \frac{2\alpha T}{1-2|\beta|\tau} + \tau^{-1} \sum_{i=1}^{T/\tau} \|\xi_{i-1}(\tau)\|^2 \right) \right] \exp\left( \frac{T(1+2|\beta|)}{1-2|\beta|\tau} \right)$$

$$\leq (1+\tau) \left[ \|U_0\|^2 + \frac{2\alpha T}{1-2|\beta|\tau} + \tau^{-1} \sum_{i=1}^{T/\tau} \|\xi_{i-1}(\tau)\|^2 \right] \exp\left( \frac{T(1+2|\beta|)}{1-2|\beta|\tau} \right)$$

$$\leq C_1 \left( 1 + \tau^{-1} \sum_{i=1}^{T/\tau} \|\xi_{i-1}(\tau)\|^2 \right)$$

which yields (4.7) for $n = 1$ with the desired constant $C_1$ given by (4.8). By applying Lemma 4.4, we obtain the general case. $\square$

Recall that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and that $(\xi_k)_{k \in [K]}$ are i.i.d. processes that satisfy Assumption 3.3. Given the polynomial growth condition, we shall need stronger regularity conditions on the noise.

**Assumption 4.7** $((p, R)$-regularity condition on noise). It holds that $\xi_0(t) \coloneqq \int_0^t \chi_0(s) \, ds$, where $\chi_0 \colon [0, \tau] \times \Omega \to \mathcal{H}$, and there exists $p \geq 1$, $R \in \mathbb{N}$ and $C(R) \geq 1$ such that

$$\mathbb{E}\left[ \|\xi_0(\tau)\|^r \right] \leq C(R)\tau^{r(2p+1)/2} \text{ for all } r \in \{1, \ldots, R\}.$$

**Remark 4.8.** Setting $r = 1$ in the inequality above implies that, in the limit of small $\tau$, the mean of $\xi_0(\tau)$ converges to zero. However, this does not imply that the mean of $\xi_0(\tau)$ itself must be zero.

The following lemma, like Proposition 3.6, shows that there exist processes that satisfy a strong version of Assumption 4.7.

**Lemma 4.9** (Integrated Brownian motion satisfies regularity condition). *Let $\tau > 0$ be fixed, $0 \leq t \leq \tau$, $p \geq 1$ be arbitrary, and $(B_t)_t$ be $\mathcal{H}$-valued Brownian motion. Then $\xi_0(t) \coloneqq \tau^{p-1} \int_0^t B_s \, ds$ satisfies*

$$\mathbb{E}\left[ \sup_{t \leq \tau} \|\xi_0(t)\|^r \right] \leq 4\tau^{rp+r/2} \text{ for all } r \in \mathbb{N}. \tag{4.12}$$

*Proof.* If $r = 0$, then the desired statement follows immediately. Therefore, let $p, r \geq 1$, and let $\xi_0$ be the integrated $\mathbb{P}$-Brownian motion process scaled by $\tau^{p-1}$, so that

$$\xi_0(t) \coloneqq \tau^{p-1} \int_0^t B_s \, ds.$$

Since $\xi_0(0) = 0$, we may assume without loss of generality that $t > 0$. Then

$$\|\xi_0(t)\|^r = \tau^{r(p-1)} t^r \left\| \frac{1}{t} \int_0^t B_s \, ds \right\|^r$$

$$\leq \tau^{r(p-1)} t^r \left( \frac{1}{t} \int_0^t \|B_s\|^r \, ds \right)$$

$$= \tau^{r(p-1)} t^{r-1} \int_0^t \|B_s\|^r \, ds,$$

13

where we applied Jensen's inequality to the uniform probability measure on $[0, t]$. It follows that

$$\mathbb{E}\left[\sup_{t \leq \tau} \|\xi_0(t)\|^r\right] \leq \tau^{r(p-1)}\mathbb{E}\left[\sup_{t \leq \tau} t^{r-1} \int_0^t \|B_s\|^r \, \mathrm{d}s\right]$$

$$\leq \tau^{r(p-1)}\tau^{r-1}\mathbb{E}\left[\sup_{t \leq \tau} \int_0^t \|B_s\|^r \, \mathrm{d}s\right]$$

$$\leq \tau^{rp-1} \int_0^\tau \mathbb{E}\left[\sup_{0 \leq t \leq \tau} \|B_t\|^r\right] \, \mathrm{d}s$$

$$= \tau^{rp}\mathbb{E}\left[\sup_{0 \leq t \leq \tau} \|B_t\|^r\right].$$

Above, we used the Fubini–Tonelli theorem to interchange expectation and integration with respect to $s$, and the fact that $\mathbb{E}\left[\sup_{t \leq \tau} \|B_t\|^r\right]$ is constant with respect to the variable of integration $s$.

By the Burkholder–Davis–Gundy inequality for martingales (Peškir, 1996, Equation (2.2)), the inequality

$$\mathbb{E}\left[\sup_{0 \leq t \leq \tau} \|B_t\|^r\right] \leq \frac{4-r}{2-r}\tau^{r/2}$$

is sharp for $0 < r < 2$ (Peškir, 1996). On the other hand, by Doob's martingale inequality (Revuz and Yor, 1999, Theorem II.1.7),

$$\mathbb{E}\left[\sup_{0 \leq t \leq \tau} \|B_t\|^r\right] \leq \left(\frac{r}{r-1}\right)^r \tau^{r/2}$$

holds for $r > 1$. Since $(4-r)/(2-r) = 3$ for $r = 1$ and since $f \colon (2, \infty) \to (0, \infty)$ defined by $f(r) := \left(\frac{r}{r-1}\right)^r$ is continuously differentiable and monotonically decreasing on its domain, the desired conclusion follows. $\square$

**Lemma 4.10** (Estimates on moments of sums). *Let $n \in \mathbb{N}$ be arbitrary, let $T > 0$ be fixed, let $\tau$ satisfy Assumption 4.5, and set $K := T/\tau \in \mathbb{N}$. If the regularity Assumption 4.7 holds for some $R \geq 2n$ and some $p \geq 1$, then there exists a scalar $C_2$ that does not depend on $\tau$, such that*

$$\mathbb{E}\left[\left(\sum_{k \in [K]} \|\xi_k(\tau)\|^2\right)^n\right] \leq (C_2 T)^n \tau^{2pn}. \tag{4.13}$$

The proof below indicates that an admissible value of $C_2$ is

$$C_2 := C(2n), \tag{4.14}$$

where, for any $R \in \mathbb{N}$, $C(R)$ is the constant in the bound on the $r^{\text{th}}$ absolute moment given in Assumption 4.7.

*Proof.* By the multinomial theorem and the independence of the $\xi_i$'s, we obtain

$$\mathbb{E}\left[\left(\sum_{k \in [K]} \|\xi_k(\tau)\|^2\right)^n\right] = \sum_{\sum_{k \in [K]} t_k = n} \frac{n!}{t_1! \cdots t_K!}\mathbb{E}\left[\prod_{k \in [K]} \|\xi_k(\tau)\|^{2t_k}\right]$$

$$\leq \sum_{\sum_{k \in [K]} t_k = n} \frac{n!}{t_1! \cdots t_K!} \prod_{k \in [K]} C(2t_k)\tau^{t_k(2p+1)}$$

$$\leq C_2^n \tau^{n(2p+1)} \sum_{\sum_{k \in [K]} t_k = n} \frac{n!}{t_1! \cdots t_K!} = C_2^n \tau^{n(2p+1)} \left(\frac{T}{\tau}\right)^n$$

$$= \left(C_2 T \tau^{2p}\right)^n,$$

thus proving (4.13). The inequality above follows from the constraint that $\sum_{k \in [K]} t_k = n$, and from the fact that $C(r) = 1$ for $r = 0$. $\square$

**Lemma 4.11** (Concavity of the square root). *For every $m \in \mathbb{N}$ and $x_1, \ldots, x_m \geq 0$,*

$$\sqrt{\sum_{i=1}^{m} x_i} \leq \sum_{i=1}^{m} \sqrt{x_i}. \tag{4.15}$$

*Proof.* The statement follows by induction. Since the statement holds for $m = 1$, suppose $m = 2$. Then

$$\sqrt{x_1 + x_2} \leq \sqrt{x_1} + \sqrt{x_2} \iff x_1 + x_2 \leq x_1 + x_2 + 2\sqrt{x_1 x_2},$$

where the statement on the right-hand side holds by nonnegativity of $x_1$ and $x_2$. Now suppose that the statement holds for $m \leq 2$. To prove that it holds for $m + 1$, we define $x_1' := \sum_{i=1}^{m} x_i$ and $x_2' := x_{m+1}$, apply the equivalence statement above, and use the assumption that the statement holds for $m \leq 2$. $\square$

**Proposition 4.12** (Bounds on even moments of maximum of numerical solution). *Suppose that the vector field $f$ satisfies Assumpsions 4.1 and 4.2, that the noise $\xi_0$ satisfies Assumption 4.7 for some $R \geq 2n$ and some $p \geq 1$, and that $\tau$ satisfies Assumption 4.5. Then*

$$\mathbb{E}\left[\max_{i \in [K]} \|U_i\|^{2n}\right] \leq 2^{n-1} C_1^n + (C_2 T)^n \tau^{n(2p-1)}. \tag{4.16}$$

*for some scalars $C_1, C_2 > 0$ not depending on $\tau$.*

Proposition 4.12 indicates that, if the noise $\xi_0$ has moments of orders up to $2n$, then the maximum of the absolute value of the randomised numerical solution $(U_k)_{k \in [K]}$ has moments of up to the same order.

*Proof.* Taking expectations of (4.7) and applying Lemma 4.10 yields the desired result. Furthermore, we may take $C_1$ and $C_2$ as given in (4.8) and (4.14). $\square$

We end the following section by using Proposition 4.12 to derive a sufficient condition for the maximum of the numerical solution to have a finite moment generating function on the real line.

**Theorem 4.13** (Maximum of numerical solution has finite moment generating function on $\mathbb{R}$). *Suppose that the vector field $f$ satisfies Assumptions 4.1 and 4.2, that $\tau$ satisfies Assumption 4.5, and that there is some $p \geq 1$ such that the regularity assumption (Assumption 4.7) holds for all $R \in \mathbb{N}$ and such that $(C(R))_{R \in \mathbb{N}}$ is bounded. Then the moment generating function of $\max_{i \in [K]} \|U_i\|$ is finite on $\mathbb{R}$.*

*Proof.* By applying Jensen's inequality and the concavity of the square root (Lemma 4.11), and by replacing $C_2$ in (4.16) with $\max_{r \in \mathbb{N}} C(r)$, we obtain

$$\mathbb{E}\left[\max_{i \in [K]} \|U_i\|^n\right] \leq \mathbb{E}\left[\max_{i \in [K]} \|U_i\|^{2n}\right]^{1/2} \leq (2C_1)^{n/2} + \left(T \max_{r \in \mathbb{N}} C(r) \tau^{2p-1}\right)^{n/2}.$$

By using the Taylor series of the exponential, we obtain for any $s \in \mathbb{R}$

$$\mathbb{E}\left[\exp\left(s \max_{i \in [K]} \|U_i\|\right)\right] \leq \exp\left(s\sqrt{2C_1}\right) + \exp\left(sT \max_{R \in \mathbb{N}} C(R) \tau^{p-1/2}\right),$$

as desired. $\square$

## 4.2 Convergence result in discrete time

Define the difference of displacements operator $V(\cdot, \cdot) \colon \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ by

$$V(u_0, v_0) := (\Phi^\tau(u_0) - u_0) - (\Phi^\tau(v_0) - v_0), \quad u_0, v_0 \in \mathcal{H}. \tag{4.17}$$

**Lemma 4.14** (Bounds on differences of displacements). *Given Assumption 4.2, and given arbitrary $\tau > 0$ and $u_0, v_0 \in \mathcal{H}$, it holds that there exists a strictly positive scalar $C_4(u_0, v_0)$ such that*

$$|\langle u_0 - v_0, V(u_0, v_0) \rangle| \leq 2\tau D \left(1 + \|u_0\|^q + \|v_0\|^q\right) \|u_0 - v_0\|^2 + C_4(u_0, v_0) \tau^3, \tag{4.18a}$$

$$\|V(u_0, v_0)\|^2 \leq 2\tau^2 D \left(1 + \|u_0\|^q + \|v_0\|^q\right) \|u_0 - v_0\|^2 + 2C_4(u_0, v_0) \tau^4. \tag{4.18b}$$

The proof below indicates than an admissible choice of $C_4$ is given by

$$C_4(u_0, v_0) := C_4(u_0, v_0; f, q, \tau) := 2 \left( 1 + \max_{s \in [0, \tau]} \|\Phi^s(u_0)\|^q + \max_{s \in [0, \tau]} \|\Phi^s(v_0)\|^q \right). \tag{4.19}$$

Later, we shall write $C_4(u_0) := C_4(u_0, u_0)$ for any $u_0 \in \mathcal{H}$. Note that $C_4(u_0, v_0; f, q, \tau)$ converges to $2(1 + \|u_0\|^q + \|v_0\|^q)$ in the limit of small $\tau$.

*Proof.* Recall that the solution map $\Phi^\tau$ of the initial value problem (1.1) is defined by

$$\Phi^\tau(u_0) = u_0 + \int_0^\tau f(u(s)) \, \mathrm{d}s, \quad u(0) = u_0. \tag{4.20}$$

Let $u(\cdot)$ and $v(\cdot)$ denote the solutions of (1.1) with the initial conditions $u(0) = u_0$ and $v(0) = v_0$. Then

$$V(u_0, v_0) = \int_0^\tau f(u(s)) \, \mathrm{d}s - \int_0^\tau f(v(s)) \, \mathrm{d}s.$$

Recall that, in the one-dimensional Euclidean setting, the error arising from approximating a Riemann integral over a bounded interval by the corresponding left Riemann sum may be bounded by a term that is proportional to the square of the length of the interval, where the constant of proportionality is related to the maximum of the derivative of the integrand over the interval of integration. Since the assumption of polynomial growth (Assumption 4.2) permits the estimation of finite difference ratios by a degree $q$ polynomial, this implies that we have

$$\left\| \int_0^\tau f(u(s)) \, \mathrm{d}s - \tau f(u_0) \right\| \leq D^{1/2} \left( 1 + 2 \max_{s \in [0, \tau]} \|\Phi^s(u_0)\|^q \right)^{1/2} \tau^2 =: D^{1/2} C(u_0)^{1/2} \tau^2. \tag{4.21}$$

If $f$ is itself not smooth, then we may approximate the vector field $f$ by smooth vector fields to arrive at a similar result. Therefore, by the triangle inequality,

$$\|V(u_0, v_0)\| \leq \tau \|f(u_0) - f(v_0)\| + \tau^2 [C(u_0) + C(v_0)].$$

By applying the above inequality and Cauchy's inequality, we obtain (4.18b). Inequality (4.18a) follows from the Cauchy–Schwarz inequality and the above inequality, the fact that $x^{1/2} \leq x$ for all $x \geq 1$, Young's inequality, and the constraint in Assumption 4.2 that $D \geq 1$:

$$\begin{aligned}
|\langle u_0 - v_0, V(u_0, v_0)\rangle| &\leq \tau D^{1/2} \left( 1 + \|u_0\|^q + \|v_0\|^q \right)^{1/2} \|u_0 - v_0\|^2 + [C(u_0) + C(v_0)]\tau^2 \|u_0 - v_0\| \\
&\leq \tau D \left( 1 + \|u_0\|^q + \|v_0\|^q \right) \|u_0 - v_0\|^2 + [C(u_0) + C(v_0)]\tau^2 \|u_0 - v_0\| \\
&\leq \tau D \left( 1 + \|u_0\|^q + \|v_0\|^q \right) \|u_0 - v_0\|^2 + \tau \|u_0 - v_0\|^2 + [C(u_0) + C(v_0)]^2 \tau^3 \\
&\leq \tau D \left( 2 + \|u_0\|^q + \|v_0\|^q \right) \|u_0 - v_0\|^2 + [C(u_0) + C(v_0)]^2 \tau^3.
\end{aligned}$$

Applying Cauchy's inequality and the polynomial growth assumption yields

$$\begin{aligned}
\|V(u_0, v_0)\|^2 &\leq 2\tau^2 \|f(u_0) - f(v_0)\|^2 + 2\tau^4 [C(u_0) + C(v_0)]^2 \\
&\leq 2\tau^2 D \left( 1 + \|u_0\|^q + \|v_0\|^q \right) \|u_0 - v_0\|^2 + 2\tau^4 [C(u_0) + C(v_0)]^2.
\end{aligned}$$

It remains to estimate $[C(u_0) + C(v_0)]^2$, which we do using Cauchy's inequality,

$$[C(u_0) + C(v_0)]^2 \leq 2 \left( 1 + \max_{s \in [0, \tau]} \|\Phi^s(u_0)\|^q + \max_{s \in [0, \tau]} \|\Phi^s(v_0)\|^q \right) = C_4(u_0, v_0),$$

thus yielding (4.18). $\square$

Recall that

$$e_k := u_k - U_k = \Phi^\tau(u_{k-1}) - \Phi^\tau(U_{k-1}) + \Phi^\tau(U_{k-1}) - \Psi^\tau(U_{k-1}) - \xi_{k-1}(\tau).$$

For $k \in [K]$, define the difference of displacements $V_{k-1}$ and the numerical discrepancy $W_{k-1}$ by

$$V_{k-1} := V(u_{k-1}, U_{k-1}), \qquad\qquad W_{k-1} := \Phi^\tau(U_{k-1}) - \Psi^\tau(U_{k-1}). \qquad (4.22)$$

Then (4.20), (4.3), and the preceding definitions yield that

$$e_k = e_{k-1} + V_{k-1} + W_{k-1} + \xi_{k-1}(\tau) \qquad (4.23)$$

Proposition 4.15 and Proposition 4.16 provide estimates on the expected values of $\|W_{k-1}\|^2$ and of terms that may be bounded in terms of $\|V_{k-1}\|$ respectively.

**Proposition 4.15** (Bounds on moments of numerical discrepancy). *Let $\tau > 0$ satisfy Assumption 4.5, $D \geq 1$ and $q \in \mathbb{N}_0$ be the parameters in the polynomial growth condition (Assumption 4.2), and $n \in \mathbb{N}$ be arbitrary. If the regularity assumption (Assumption 4.7) holds for some $R \geq n(2q + 2)$ and some $p \geq 1$, then for every $k - 1 \in [K]$, there exists a strictly positive scalar $C_5$ such that*

$$\mathbb{E}\big[\|W_{k-1}\|^{2n}\big] \leq C_5 \tau^{4n}. \qquad (4.24)$$

The proof below indicates that an admissible choice of $C_5$ is given by

$$C_5(k, n) := (2D)^{2n} \left[ C_4(U_k) + (1 + \|U_k\|^q + \|\Psi^\tau(U_k)\|^q)^2 \left( \|\Psi^\tau(U_k)\|^2 + \|f(0)\|^2 \right) \right]^n. \qquad (4.25)$$

where we have abused notation and written $C_4(U_{k-1}) := C_4(U_{k-1}, U_{k-1})$ for $C_4(\cdot, \cdot)$ defined in (4.19).

*Proof.* Approximating the integral in $\Phi^\tau(U_{k-1})$ by a left Riemann sum implies that

$$\|\Phi^\tau(U_{k-1}) - U_{k-1} - \tau f(U_{k-1})\|^2 \leq D \left( 1 + 2 \max_{s \in [0,\tau]} \|\Phi^s(U_{k-1})\|^q \right) \tau^4 < DC_4(U_{k-1})\tau^4, \qquad (4.26)$$

by (4.21). Therefore, by the triangle inequality and Cauchy's inequality,

$$\|W_{k-1}\|^2 \leq 2 \left( DC_4(U_{k-1})\tau^4 + \|U_{k-1} + \tau f(U_{k-1}) - U_{k-1} - \tau f(\Psi^\tau(U_{k-1}))\|^2 \right)$$
$$\leq 2\tau^2 \left( DC_4(U_{k-1})\tau^2 + \|f(U_{k-1}) - f(\Psi^\tau(U_{k-1})\|^2 \right).$$

By the polynomial growth condition (4.2),

$$\|f(U_{k-1}) - f(\Psi^\tau(U_{k-1}))\|^2 \leq D \left( 1 + \|U_{k-1}\|^q + \|\Psi^\tau(U_{k-1})\|^q \right) \|U_{k-1} - \Psi^\tau(U_{k-1})\|^2$$
$$= D \left( 1 + \|U_{k-1}\|^q + \|\Psi^\tau(U_{k-1})\|^q \right) \|\tau f(\Psi^\tau(U_{k-1}))\|^2$$
$$= \tau^2 D \left( 1 + \|U_{k-1}\|^q + \|\Psi^\tau(U_{k-1})\|^q \right) \|f(\Psi^\tau(U_{k-1}))\|^2.$$

The triangle inequality and Lemma 4.4 imply

$$\|f(\Psi^\tau(U_{k-1}))\|^2 \leq 2 \left( \|f(\Psi^\tau(U_{k-1})) - f(0)\|^2 + \|f(0)\|^2 \right)$$
$$\leq 2 \left( D \left( 1 + \|\Psi^\tau(U_{k-1})\|^q \right) \|\Psi^\tau(U_{k-1})\|^2 + \|f(0)\|^2 \right)$$
$$\leq 2D \left( 1 + \|U_{k-1}\|^q + \|\Psi^\tau(U_{k-1})\|^q \right) \left( \|\Psi^\tau(U_{k-1})\|^2 + \|f(0)\|^2 \right).$$

Thus,

$$\|W_{k-1}\|^2 \leq 4\tau^4 D^2 \left[ C_4(U_{k-1}) + (1 + \|U_{k-1}\|^q + \|\Psi^\tau(U_{k-1})\|^q)^2 \left( \|\Psi^\tau(U_{k-1})\|^2 + \|f(0)\|^2 \right) \right],$$

and, for $n \in \mathbb{N}$, Lemma 4.4 yields

$$\|W_{k-1}\|^{2n} \leq (2D\tau^2)^{2n} \left[ C_4(U_{k-1}) + (1 + \|U_{k-1}\|^q + \|\Psi^\tau(U_{k-1})\|^q)^2 \left( \|\Psi^\tau(U_{k-1})\|^2 + \|f(0)\|^2 \right) \right]^n.$$

17

Taking expectations and using (4.25) yields (4.24). It remains to show that $C_5$ as defined in (4.25) is finite. Since the regularity assumption holds for some $R \geq n(2q + 2)$, then it follows that the scalar $C_4(U_{k-1})$, which depends on $\|U_{k-1}\|^q$, has finite moments up to order $2n$, by Lemma 4.6. The remaining contributions to $C_5$ can be similarly bounded by the absolute moments of $\|U_{k-1}\|$ of up to order $n(2q+2)$, since by (4.10) and Lemma 4.4 we have

$$\|\Psi^\tau(U_{k-1})\|^{2n} \leq 2^{n-1}(1 - 2|\beta|\tau)^{-n} \left(\|U_{k-1}\|^{2n} + (2\alpha\tau)^n\right)$$

for all $n \in \mathbb{N}$. This completes the proof. $\square$

**Proposition 4.16** (Bounds on expected values of products of difference of displacements)**.** *Suppose that Assumptions 4.1, 4.2, and 4.5 hold. Let $V_{k-1}$ be defined as in (4.22), let $q \in \mathbb{N}_0$ denote the exponent in Assumption 4.2, and suppose that the regularity assumption (Assumption 4.7) holds with parameters $R = 2q + 2$ and $p \geq 1$. If*
  *(i) $q = 0$, or*
  *(ii) $q \geq 1$ and*

$$p \geq \max\left\{\frac{2}{q} + \frac{1}{2}, 1\right\} \tag{4.27}$$

*then there exists a strictly positive scalar $C' = C'(k, q)$ that remains finite in the limit of small $\tau$, such that*

$$\mathbb{E}\big[|\langle e_{k-1}, V_{k-1}\rangle|\big] \leq C'(k-1, q)\left(\tau \mathbb{E}\big[\|e_{k-1}\|^2\big] + \tau^3\right) \tag{4.28a}$$
$$\mathbb{E}\big[\|V_{k-1}\|^2\big] \leq C'(k-1, q)\left(\tau^2 \mathbb{E}\big[\|e_{k-1}\|^2\big] + \tau^4\right). \tag{4.28b}$$

*In particular, if the regularity assumption (Assumption 4.7) holds for some $R \geq 2q+2$ and some $p \geq 5/2$, then the inequalities (4.28) hold.*

**Remark 4.17.** For case (ii), it can be shown by an appropriate modification of Lemma 4.6 that the regularity assumption can be weakened to holding for $R \geq q + 3$, with the result that the condition (4.27) must be strengthened to $p \geq \max\{4/q + 1/2, 1\}$. However, since Proposition 4.15 already imposes the requirement that $R \geq 2q + 2$, the requirement in Proposition 4.16 above that $R \geq 2q$ imposes no additional constraint.

The proof below indicates that an admissible choice of the constant $C'$ for both case (i) and (ii) is given by

$$C'(k, q) := 2D\left(1 + \max_{k \in [K]} \|u_k\|^q + (2^2 C_1)^{q/2}\right) + (2^2 C_1)^{q/2} \mathbb{E}\big[\|e_k\|^4\big]^{(q+2)/(2q+2)} \tag{4.29}$$
$$+ 2^3\left(1 + \max_{s \in [0,\tau]} \|\Phi^s(u_k)\|^q + \mathbb{E}\left[\max_{s \in [0,\tau]} \|\Phi^s(U_k)\|^{2q}\right]\right)$$

where $C_1$ is defined in (4.8).

*Proof.* By the definition of $V_{k-1}$ in (4.22) and by Lemma 4.14, we have the following estimates:

$$|\langle e_{k-1}, V_{k-1}\rangle| \leq 2\tau D\left(1 + \|u_{k-1}\|^q + \|U_{k-1}\|^q\right)\|e_{k-1}\|^2 + C_4(u_{k-1}, U_{k-1})\tau^3 \tag{4.30a}$$
$$\|V_{k-1}\|^2 \leq 2\tau^2 D\left(1 + \|u_{k-1}\|^q + \|U_{k-1}\|^q\right)\|e_{k-1}\|^2 + 2C_4(u_{k-1}, U_{k-1})\tau^4. \tag{4.30b}$$

We shall prove the statement in each case separately using (4.30). Observe that the definition of $C_4$ in (4.19) and Lemma 4.4 yield

$$\mathbb{E}\big[C_4^2(u_{k-1}, U_{k-1})\big] \leq 2^2\left(1 + \max_{s \in [0,\tau]} \|\Phi^s(u_{k-1})\|^q + \mathbb{E}\left[\max_{s \in [0,\tau]} \|\Phi^s(U_{k-1})\|^{2q}\right]\right).$$

We consider cases (i) and (ii) separately. Let $k - 1 \geq 1$.
  Case (i): If $q = 0$, then the equations (4.30) immediately imply (4.28).

18

Case (ii): Let $q \in \mathbb{N}$. By the inequalities (4.30), it suffices to show that $C_4(u_{k-1}, U_{k-1})$ has finite second moment, and that there exists some scalar $C'$ not depending on $\tau$ such that $\mathbb{E}[\|U_{k-1}\|^q \|e_{k-1}\|^2]$ is bounded from above by $C'(\tau^2 + \mathbb{E}[\|e_{k-1}\|^2])$. For example, from (4.30a) we have

$$\mathbb{E}[|\langle e_{k-1}, V_{k-1}\rangle|] \leq 2\tau D\left((1 + \|u_{k-1}\|^q)\,\mathbb{E}[\|e_{k-1}\|^2] + \mathbb{E}[\|U_{k-1}\|^q \|e_{k-1}\|^2]\right) + \mathbb{E}[C_4^2(u_{k-1}, U_{k-1})]\tau^3.$$

Since the regularity assumption holds for $R \geq 2q + 2$, Proposition 4.12 implies that $\Phi^s(U_{k-1})$ has finite moment of order $2q$, and thus the finiteness of the second moment of $C_4$ follows. By (4.7) in Lemma 4.6,

$$\max_{i \in k} \|U_i\|^{2q} \leq 2^{2q-1} C_1^q \left[1 + \tau^{-q}\left(\sum_{i=0}^{k-1} \|\xi_k(\tau)\|^2\right)^q\right].$$

Therefore, using the inequality $\|U_{k-1}\|^q \leq \max_{i \leq k-1} \|U_i\|^q$, and using the concavity of the square root,

$$\mathbb{E}[\|U_{k-1}\|^q \|e_{k-1}\|^2] \leq \left(2^{2q-1} C_1^q\right)^{1/2} \mathbb{E}\left[\left(1 + \tau^{-q}\left(\sum_{i=0}^{k-2} \|\xi_i(\tau)\|^2\right)^q\right)^{1/2} \|e_{k-1}\|^2\right]$$

$$\leq \left(2^{2q-1} C_1^q\right)^{1/2}\left(\mathbb{E}[\|e_{k-1}\|^2] + \tau^{-q/2}\mathbb{E}\left[\left(\sum_{k=1}^{T/\tau} \|\xi_k(\tau)\|^2\right)^{q/2} \|e_{k-1}\|^2\right]\right).$$

Now, using Hölder's inequality and Lemma 4.10,

$$\mathbb{E}\left[\left(\sum_{k=1}^{T/\tau} \|\xi_k(\tau)\|^2\right)^{q/2} \|e_{k-1}\|^2\right]$$

$$\leq \mathbb{E}\left[\left(\sum_{k=1}^{T/\tau} \|\xi_k(\tau)\|^2\right)^{q+1}\right]^{q/(2q+2)} \mathbb{E}\left[\|e_{k-1}\|^{(4q+4)/(q+2)}\right]^{(q+2)/(2q+2)}$$

$$\leq \left((C_2 T)^{q+1}\, \tau^{2p(q+1)}\right)^{q/(2q+2)} \mathbb{E}\left[\|e_{k-1}\|^{(4q+4)/(q+2)}\right]^{(q+2)/(2q+2)}$$

$$= (C_2 T)^{q/2}\, \tau^{pq} \mathbb{E}\left[\|e_{k-1}\|^{(4q+4)/(q+2)}\right]^{(q+2)/(2q+2)}.$$

Since the regularity assumption holds for $R \geq 2q+2$ and $p \geq 1$, then by observing that $(4q+4)/(q+2) < 4$, and by applying Jensen's inequality and Lemma 4.11, we have

$$\mathbb{E}\left[\|e_{k-1}\|^{(4q+4)/(q+2)}\right]^{(q+2)/(2q+2)} \leq \mathbb{E}[\|e_k\|^4]^{1/2} \leq \left(2^{3/2}\left(\|u_k\|^4 + \mathbb{E}[\|U_k\|^4]\right)\right)^{1/2}.$$

Thus we have shown

$$\mathbb{E}\left[\left(\sum_{k=1}^{T/\tau} \|\xi_k(\tau)\|^2\right)^{q/2} \|e_{k-1}\|^2\right] \leq (C_2 T)^{q/2}\, 2^{3/2}\left(\|u_k\|^4 + \mathbb{E}[\|U_k\|^4]\right)^{1/2} \tau^{pq}. \tag{4.31}$$

The finiteness of $\mathbb{E}[\|U_k\|^4]$ follows from the fact that the regularity assumption holds up to $R \geq 2q+2 \geq 4$, and from Proposition 4.12. Thus, we have shown that

$$\mathbb{E}[\|U_{k-1}\|^q \|e_{k-1}\|^2]$$
$$\leq \left(2^{2q-1} C_1^q\right)^{1/2}\left(\mathbb{E}[\|e_{k-1}\|^2] + (C_2 T)^{q/2}\, \tau^{q(2p-1)/2}\mathbb{E}[\|e_{k-1}\|^{(4q+4)/(q+2)}]^{(q+2)/(2q+2)}\right).$$

Since (4.27) implies that $p \geq 2/q + 1/2$, which in turn is equivalent to $q(2p-1)/2 \geq 2$, we obtain (4.28a) with $C'(k-1, q)$ as given in (4.29). Applying the same argument to the estimate (4.30b) yields the estimate (4.28b). This completes the proof. $\qquad\square$

We now prove that the randomised implicit Euler method has order one. Note that for this result, we assume that the random variable $\xi_0(\tau)$ has mean zero.

**Theorem 4.18** (Uniform mean-square convergence in discrete time). *Suppose that Assumptions 4.1, 4.2, and 4.5 hold. Let $q \in \mathbb{N}_0$ denote the exponent in Assumption 4.2, and suppose that the regularity assumption (Assumption 4.7) holds with parameters $R \geq 2q + 2$ and $p \geq 1$. If*

*(i) $q = 0$, or*
*(ii) $q \geq 1$, $\mathbb{E}[\xi_0(\tau)] = 0$, and (4.27) holds,*

*then there exists a strictly positive scalar $C$ that remains bounded in the limit of small $\tau$ and does not depend on $k$, such that*

$$\max_{k \leq T/\tau} \mathbb{E}\big[\|e_k\|^2\big] \leq C\tau^2.$$

*Proof.* Given (4.23), it holds that

$$\|e_k\|^2 = \|e_{k-1}\|^2 + \|V_{k-1}\|^2 + \|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2 + 2\Big(\langle e_{k-1}, V_{k-1}\rangle + \langle e_{k-1}, W_{k-1}\rangle +$$

$$+ \langle e_{k-1}, \xi_{k-1}(\tau)\rangle + \langle V_{k-1}, W_{k-1}\rangle + \langle V_{k-1}, \xi_{k-1}(\tau)\rangle + \langle W_{k-1}, \xi_{k-1}(\tau)\rangle\Big).$$

Taking expectations, using the independence of $\xi_k(\tau)$ from $U_k$, and using the mean-zero property of $\xi_k(\tau)$ yield

$$\mathbb{E}\big[\|e_k\|^2\big] \leq \mathbb{E}\big[\|e_{k-1}\|^2 + \|V_{k-1}\|^2 + \|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2\big]$$
$$+ 2\left(\mathbb{E}\big[|\langle e_{k-1}, W_{k-1}\rangle|\big] + \mathbb{E}\big[|\langle V_{k-1}, W_{k-1}\rangle|\big] + \mathbb{E}\big[|\langle e_{k-1}, V_{k-1}\rangle|\big]\right).$$

By Young's inequality and Cauchy's inequality,

$$|\langle e_{k-1}, W_{k-1}\rangle| \leq \frac{\tau}{2}\|e_{k-1}\|^2 + \frac{1}{2\tau}\|W_{k-1}\|^2, \qquad |\langle V_{k-1}, W_{k-1}\rangle| \leq \frac{1}{2}\|V_{k-1}\|^2 + \frac{1}{2}\|W_{k-1}\|^2, \qquad (4.32)$$

and hence

$$\mathbb{E}\big[\|e_k\|^2 - \|e_{k-1}\|^2\big] \leq \mathbb{E}\left[\tau\|e_{k-1}\|^2 + 2\|V_{k-1}\|^2 + \left(2 + \frac{1}{\tau}\right)\|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2\right]$$
$$+ 2\mathbb{E}\big[|\langle e_{k-1}, V_{k-1}\rangle|\big].$$

Since $e_0 \coloneqq u_0 - U_0 = 0$, using a telescoping sum and the inequality above yields

$$\max_{i \leq n} \mathbb{E}\big[\|e_i\|^2\big]$$
$$\leq \sum_{k=1}^{n} \mathbb{E}\left[\tau\|e_{k-1}\|^2 + 2\|V_{k-1}\|^2 + \left(2 + \frac{1}{\tau}\right)\|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2 + 2|\langle e_{k-1}, V_{k-1}\rangle|\right]$$
$$\leq \sum_{k=1}^{n} \left(\mathbb{E}\big[\tau(1 + 4C')\|e_{k-1}\|^2\big] + 2C'\tau^4 + \left(2 + \frac{1}{\tau}\right)C_5\tau^4 + C(2)\tau^{2p+1} + 2C'\tau^3\right),$$

where we used Proposition 4.16, Proposition 4.15, and the regularity assumption in the second inequality. Replacing each instance of $C' = C'(k, q)$ and of $C_5 = C_5(k, 2)$ in the last inequality above with $C' = \max_{k \leq T/\tau} C'(k, q)$ and with $C_5 = \max_{k \leq T/\tau} C_5(k, 2)$ respectively, we obtain

$$\max_{i \leq T/\tau} \mathbb{E}\big[\|e_i\|^2\big] \leq T\left(4C'\tau^2 + (2\tau + 1)C_5\tau^2 + C(2)\tau^{2p}\right) + \tau(1 + 4C')\sum_{k=1}^{n} \max_{i \leq k-1} \mathbb{E}\big[\|e_i\|^2\big].$$

By Grönwall's inequality (Theorem 2.1),

$$\max_{i \leq T/\tau} \mathbb{E}\big[\|e_i\|^2\big] \leq \left[T\left(4C'\tau^2 + (2\tau + 1)C_5\tau^2 + C(2)\tau^{2p}\right)\right]\exp\left(T(1 + 4C')\right),$$

and since $p \geq 1$, the right-hand side is $O(\tau^2)$, as desired. $\qquad\square$

The next result strengthens that of Theorem 4.18 by bringing the maximum inside the expectation and not imposing the requirement that the noise $\xi_0$ have mean zero (see Remark 4.8).

**Theorem 4.19** (Mean-square uniform convergence in discrete time). *Suppose that Assumptions 4.1, 4.2, and 4.5 hold. Let $q \in \mathbb{N}_0$ denote the exponent in Assumption 4.2, and suppose that the regularity assumption (Assumption 4.7) holds with parameters $R \geq 2q + 2$ and $p \geq 1$. If*
  *(i) $q = 0$, or*
  *(ii) $q \geq 1$ and*

$$p \geq \max\left\{\frac{2}{q} + \frac{1}{2}, \frac{3}{2}\right\}, \tag{4.33}$$

*then there exists a strictly positive scalar $C$ that remains bounded in the limit of small $\tau$ and does not depend on $k$, such that*

$$\mathbb{E}\left[\max_{k \leq T/\tau} \|e_k\|^2\right] \leq C\tau^2.$$

*In particular, if the regularity assumption (Assumption 4.7) holds for some $R \geq 2q+2$ and some $p \geq 5/2$, then the conclusion holds.*

With the exception of (4.33) — which is stronger than (4.27) by the requirement that $p \geq 3/2$ instead of $p \geq 1$ — the stated hypotheses are exactly those of Proposition 4.16. This is because, in order to obtain the desired convergence result, one must ensure that the terms on an upper bound of $\mathbb{E}[\max_{k \in [K]} \|e_k\|^2]$ are of sufficiently large order in $\tau$. Of these terms, the most problematic — i.e. the ones arising due to the polynomial growth condition — are controlled by Proposition 4.16.

*Proof.* Given (4.23), it holds that

$$\|e_k\|^2 \leq \|e_{k-1}\|^2 + \|V_{k-1}\|^2 + \|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2 + 2\Big(|\langle e_{k-1}, V_{k-1}]\rangle| + |\langle e_{k-1}, W_{k-1}\rangle| +$$

$$+ |\langle e_{k-1}, \xi_{k-1}(\tau)\rangle| + |\langle V_{k-1}, W_{k-1}\rangle| + |\langle V_{k-1}, \xi_{k-1}(\tau)\rangle| + |\langle W_{k-1}, \xi_{k-1}(\tau)\rangle|\Big)$$

Using the telescoping sum $\|e_n\|^2 = \|e_0\|^2 + \sum_{k=1}^{n} \|e_k\|^2 - \|e_{k-1}\|^2$ for $n \in [K]$, and using that $u_0 = U_0$ implies $e_0 = 0$, we obtain

$$\|e_n\|^2 \leq \sum_{k=1}^{n}\Big[\|V_{k-1}\|^2 + \|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2 + 2\Big(|\langle e_{k-1}, V_{k-1}\rangle| + |\langle e_{k-1}, W_{k-1}\rangle| + |\langle e_{k-1}, \xi_{k-1}(\tau)\rangle|$$

$$+ |\langle V_{k-1}, W_{k-1}\rangle| + |\langle V_{k-1}, \xi_{k-1}(\tau)\rangle| + |\langle W_{k-1}, \xi_{k-1}(\tau)\rangle|\Big)\Big]. \tag{4.34}$$

Using Young's inequality, we have

$$|\langle e_{k-1}, W_{k-1}\rangle| \leq \frac{\tau}{2}\|e_{k-1}\|^2 + \frac{1}{2\tau}\|W_{k-1}\|^2, \qquad |\langle e_{k-1}, \xi_{k-1}(\tau)\rangle| \leq \frac{\tau}{2}\|e_{k-1}\|^2 + \frac{1}{2\tau}\|\xi_{k-1}(\tau)\|^2.$$

To $|\langle V_{k-1}, W_{k-1}\rangle|$, $|\langle V_{k-1}, \xi_{k-1}(\tau)\rangle|$, and $|\langle W_{k-1}, \xi_{k-1}(\tau)\rangle|$, we apply the Cauchy–Schwarz inequality and Young's inequality to obtain

$$\|e_n\|^2 \leq \sum_{k=1}^{n}\left[2\tau\|e_{k-1}\|^2 + 2|\langle e_{k-1}, V_{k-1}\rangle| + 3\|V_{k-1}\|^2 + \left(3 + \frac{1}{\tau}\right)\|W_{k-1}\|^2 + \left(3 + \frac{1}{\tau}\right)\|\xi_{k-1}(\tau)\|^2\right].$$

Using the nonnegativity of the summands, the fact that $3 + \tau^{-1} \leq 4\tau^{-1}$ by Assumption 4.5, and the fact that $n \leq T/\tau$, we can obtain a stronger version of the inequality above:

$$\max_{i \leq n}\|e_i\|^2 \leq \sum_{k=1}^{n}\left[2\tau \max_{i \leq k-1}\|e_i\|^2 + 2|\langle e_{k-1}, V_{k-1}\rangle| + 3\|V_{k-1}\|^2 + \frac{4}{\tau}\|W_{k-1}\|^2 + \frac{4}{\tau}\|\xi_{k-1}(\tau)\|^2\right].$$

21

We now take expectations of both sides. If either condition (i) or condition (ii) holds, then we may apply Proposition 4.16 to obtain

$$\sum_{k=1}^{n} \mathbb{E}\left[\|V_{k-1}\|^2 + |\langle e_{k-1}, V_{k-1}\rangle|\right] \leq \sum_{k=1}^{n} C'(k,q)\left(\tau \mathbb{E}\left[\max_{j\leq k-1} \|e_j\|^2\right] + \tau^3\right)$$

From Proposition 4.15, it follows that $\mathbb{E}[\|W_{k-1}\|^2] \leq C_5(k-1,2)\tau^4$, where $C_5$ is defined in (4.25). We also have $\mathbb{E}[\|\xi_{k-1}(\tau)\|^2] \leq C(2)\tau^{2p+1}$ by the regularity assumption. Combining the above estimates, and using the hypothesis that $2p \geq 3$, we obtain

$$\mathbb{E}\left[\max_{i\leq n} \|e_i\|^2\right] \leq \tilde{C}(n)\left[\sum_{k=1}^{n}\left(\tau \mathbb{E}\left[\max_{i\leq k-1} \|e_i\|^2\right] + \tau^3\right)\right] \leq \tilde{C}(n)T\tau^2 + \tilde{C}(n)\tau \sum_{k=1}^{n} \mathbb{E}\left[\max_{i\leq k-1} \|e_i\|^2\right],$$

where above,

$$\tilde{C}(n) := \max\left\{\max_{i\leq n} C'(i,q), \max_{i\leq n} C_5(i,2), C(2)\right\}.$$

Using the Grönwall inequality (Theorem 2.1), and using the fact that $n \leq T/\tau$, we obtain

$$\mathbb{E}\left[\max_{k\leq T/\tau} \|e_k\|^2\right] \leq T\tilde{C}(T/\tau)\exp(T\tilde{C}(T/\tau))\tau^2,$$

as desired. □

## 4.3 Convergence result in continuous time

In this section, we use the convergence result of Section 4.2 to obtain a similar result for the continuous-time case, using similar ideas. Since the results follow from minor modifications of the proofs of their analogues, we state the results without proof.

Given the collection of times $(t_k)_{k\in[K]}$ as defined in (2.1), we shall define for $k \in [K]$ and $t \in (t_k, t_{k+1}]$

$$u_t = \Phi^{t-t_k}(u_k) \tag{4.35}$$

$$\Phi^{t-t_k}(a) = a + \int_0^{t-t_k} f(\widetilde{u}(s))\,\mathrm{d}s, \quad \widetilde{u}(0) = a \tag{4.36}$$

$$U_t = \Psi^{t-t_k}(U_k) + \xi_k(t - t_k) \tag{4.37}$$

$$\Psi^{t-t_k}(U_k) = U_k + (t - t_k)f(\Psi^{t-t_k}(U_k)) \tag{4.38}$$

$$V_k^{t-t_k} := (\Phi^{t-t_k}(u_k) - u_k) - (\Phi^{t-t_k}(U_k) - U_k)$$
$$= \int_0^{t-t_k} f(u(t_k + s)) - f(\widetilde{u}^{(k)}(t_k + s))\,\mathrm{d}s, \quad u^{(k)}(0) = U_k \tag{4.39}$$

$$W_k^{t-t_k} := \Phi^{t-t_k}(U_k) - \Psi^{t-t_k}(U_k)$$
$$= \int_0^{t-t_k} f(\widetilde{u}^{(k)}(t_k + s))\,\mathrm{d}s - (t - t_k)f(\Psi^{t-t_k}(U_k)). \tag{4.40}$$

The error at time $t$ defined by $e_t := u_t - U_t$ satisfies

$$e_t = e_k + V_k^{t-t_k} + W_k^{t-t_k} + \xi_k(t - t_k), \tag{4.41}$$

analogously to (4.23). We have the following analogues of Propositions 4.15 and 4.16 respectively.

**Proposition 4.20** (Bounds on moments of numerical discrepancy over small time interval)**.** *Let $\tau > 0$ satisfy Assumption 4.5, $D \geq 1$ and $q \in \mathbb{N}_0$ be the parameters in the polynomial growth condition (Assumption 4.2), and $n \in \mathbb{N}$ be arbitrary. If the regularity assumption (Assumption 4.7) holds for some $R \geq n(2q+2)$ and some $p \geq 1$, then for every $k \in [K]$, there exists a strictly positive scalar $C_5$ such that*

$$\mathbb{E}\left[\sup_{t\in(t_k,t_{k+1}]} \left\|W_k^{t-t_k}\right\|^{2n}\right] \leq C_5\tau^{4n}.$$

22

**Proposition 4.21** (Bounds on expected values of products of difference of displacements over small time interval). *Suppose that Assumptions 4.1, 4.2, and 4.5 hold. Let $V_k^{t-t_k}$ be defined as in (4.39), let $q \in \mathbb{N}_0$ denote the exponent in Assumption 4.2, and suppose that the regularity assumption (Assumption 4.7) holds with parameters $R \in \mathbb{N}$ and $p \geq 1$. If*

*(i) $q = 0$ and $R \geq 2$, or*

*(ii) $q \geq 1$, $R \geq 2q + 2$, and (4.27) holds,*

*then there exists a strictly positive scalar $C'$ that remains bounded in the limit of small $\tau$ and does not depend on $k$, such that*

$$\mathbb{E}\left[\sup_{t \in (t_k, t_{k+1}]} |\langle e_k, V_k^{t-t_k} \rangle|\right] \leq C' \left(\tau \mathbb{E}\left[\|e_k\|^2\right] + \tau^3\right)$$

$$\mathbb{E}\left[\sup_{t \in (t_k, t_{k+1}]} \left\|V_k^{t-t_k}\right\|^2\right] \leq C' \left(\tau^2 \mathbb{E}\left[\|e_k\|^2\right] + \tau^4\right).$$

*In particular, if the regularity assumption (Assumption 4.7) holds for some $R \geq 2q+2$ and some $p \geq 5/2$, then the inequalities above hold.*

Before we state the continuous-time convergence result, we state the following modification of Assumption 4.7:

**Assumption 4.22** ($(p, R)$-strong regularity condition on noise). *It holds that $\xi_0(t) := \int_0^t \chi_0(s)\,\mathrm{d}s$, where $\chi_0 \colon [0, \tau] \times \Omega \to \mathcal{H}$, and there exists a $p \geq 1$ and $R \in \mathbb{N}$ such that*

$$\mathbb{E}\left[\sup_{t \in [0, \tau]} \|\xi_k(t)\|^r\right] \leq C(R)\tau^{r(2p+1)/2} \text{ for all } r \in \{1, \ldots, R\}.$$

**Theorem 4.23** (Mean-square uniform convergence in continuous time). *Suppose that Assumptions 4.1, 4.2, and 4.5 hold. Let $q \in \mathbb{N}_0$ denote the exponent in Assumption 4.2, and suppose that the strong regularity assumption (Assumption 4.22) holds with parameters $R \in \mathbb{N}$ and $p \geq 1$. If*

*(i) $q = 0$ and $R \geq 2$, or*

*(ii) $q \geq 1$, $R \geq 2q + 2$, and (4.33) holds,*

*then there exists a strictly positive scalar $C$ that remains bounded in the limit of small $\tau$ and does not depend on $k$, such that*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \|e_t\|^2\right] \leq C\tau^2.$$

*In particular, if the strong regularity assumption (Assumption 4.22) holds for some $R \geq 2q+2$ and some $p \geq 5/2$, then the conclusion holds.*

*Proof.* Except for the strong regularity assumption, the stated hypotheses are exactly those of Theorem 4.18, which we shall use to prove both cases (i) and (ii). Since $e_0 = u_0 - U_0 = 0$, it follows that

$$\sup_{0 \leq t \leq T} \|e_t\|^2 = \max_{k \in [K]} \sup_{t \in (t_k, t_{k+1}]} \|e_t\|^2.$$

By applying Lemma 4.4 to (4.41), we obtain

$$\mathbb{E}\left[\max_{k \in [K]} \sup_{t \in (t_k, t_{k+1}]} \|e_t\|^2\right] \leq 2^3 \mathbb{E}\left[\max_{k \in [K]} \left(\|e_k\|^2 + \sup_{t \in (t_k, t_{k+1}]} \left(\left\|V_k^{t-t_k}\right\|^2 + \left\|W_k^{t-t_k}\right\|^2 + \|\xi_k(t - t_k)\|^2\right)\right)\right]$$

By Theorem 4.18, the first term inside the expectation on the right-hand side of the inequality above is $O(\tau^2)$, as desired. Thus it remains to show that the other terms are all $O(\tau^2)$. We have

$$\mathbb{E}\left[\max_{k \in [K]} \sup_{t \in (t_k, t_{k+1}]} \left\|V_k^{t-t_k}\right\|^2\right] \leq \sum_{k \in [K]} \mathbb{E}\left[\sup_{t \in (t_k, t_{k+1}]} \left\|V_k^{t-t_k}\right\|^2\right] \leq C'T \left(\tau \max_{k \in [K]} \mathbb{E}\left[\|e_k\|^2\right] + \tau^3\right) \leq C'T\tau^3,$$

by applying Proposition 4.21 and Theorem 4.18, using that $\max_{k\in[K]} \mathbb{E}[\|e_k\|^2] \leq \mathbb{E}[\max_{k\in[K]} \|e_k\|^2]$, and using that $K = T/\tau$. Next, we have

$$\mathbb{E}\left[\max_{k\in[K]}\sup_{t\in(t_k,t_{k+1}]}\left\|W_k^{t-t_k}\right\|^2\right] \leq \sum_{k\in[K]}\mathbb{E}\left[\sup_{t\in(t_k,t_{k+1}]}\left\|W_k^{t-t_k}\right\|^2\right] \leq C_5 T\tau^3,$$

by applying Proposition 4.20 in the second inequality, where we have abused the notation for $C_5$ by taking the maximum over $k \in [K]$ of the prefactors $C_5$ appearing in Proposition 4.20. Finally, we have

$$\mathbb{E}\left[\max_{k\in[K]}\sup_{t\in(t_k,t_{k+1}]}\|\xi_k(t-t_k)\|^2\right] \leq \sum_{k\in[K]}\mathbb{E}\left[\sup_{t\in(t_k,t_{k+1}]}\|\xi_k(t-t_k)\|^2\right] = C(2)T\tau^{2p},$$

by using the fact that the $\xi_k$ are i.i.d. copies of $\xi_0$ and by using the strong regularity assumption (Assumption 4.22). Since the hypothesis (4.33) on $p$ implies that $2p \geq 3$, the right-hand side is $O(\tau^3)$. Putting all the estimates together yields

$$\mathbb{E}\left[\max_{k\in[K]}\sup_{t\in(t_k,t_{k+1}]}\|e_t\|^2\right] \leq 2^3\left(\mathbb{E}\left[\max_{k\in[K]}\|e_k\|^2\right] + \max\left\{C', C_5, C(2)\right\}T\tau^3\right) = O(\tau^2),$$

as desired. $\square$

# 5 Explicit Euler integration of locally Lipschitz vector fields

In this section, we adapt the arguments for implicit Euler integration for locally Lipschitz dissipative flows to explicit Euler integration, i.e. for the numerical flow map $\Psi^\tau$ defined by

$$\Psi^\tau(U_n) = U_n + \tau f(U_n). \tag{5.1}$$

We only consider the discrete-time convergence result, since the continuous-time convergence result can be obtained from this as shown in Section 4. Since the desired convergence result follows from a very similar sequence of arguments as for the section on implicit Euler integration, we shall reuse some of the results from that section and only sketch the proofs. We find that in the case of explicit Euler integration, we do not have a nice bound on the second moment of the stochastic process $(U_n)_n$ generated by (4.3) in terms of second order moments; see Proposition 5.2 below. As a result, the discrete-time convergence result requires stronger assumptions on the noise, and on the process $(U_k)_{k\in[K]}$ itself.

## 5.1 Moment bounds

The next lemma establishes an analogue of Lemma 4.6. Recall that $K = T/\tau \in \mathbb{N}$.

**Lemma 5.1** (Uniform bounds on numerical solution for explicit Euler). *Suppose that the vector field $f$ satisfies Assumptions 4.1 and 4.2, and suppose that $\tau > 0$ satisfies Assumption 4.5. For all $n \in \mathbb{N}$, the numerical solution $(U_k)_{k\in[K]}$ generated by (4.3) and (5.1) satisfies*

$$\max_{k\in[K]}\|U_k\|^{2n} \leq 4^{n-1}C_7^n\left(1 + \tau^n\max_{k\in[K]}\|U_k\|^{(q+2)n} + \tau^{-n}\left(\sum_{i=0}^{K-1}\|\xi_i(\tau)\|^2\right)^n\right) \tag{5.2}$$

*uniformly in $\omega \in \Omega$, where*

$$C_7 := \exp\left(T(5D + 2|\beta|)\right)\max\left\{3, 4DT, \|U_0\|^2 + T(2\alpha + 4D\|f(0)\|^2)\right\}. \tag{5.3}$$

*Proof.* Substituting the explicit Euler map (5.1) into the randomised numerical scheme (4.3), we obtain

$$U_{n+1} = \Psi^\tau(U_n) + \xi_n(\tau) = U_n + \tau f(U_n) + \xi_n(\tau).$$

24

Therefore,

$$\|U_{n+1}\|^2 = \|U_n\|^2 + \tau^2 \|f(U_n)\|^2 + \|\xi_n(\tau)\|^2 + 2\langle U_n, \tau f(U_n) + \xi_n(\tau)\rangle + 2\langle \tau f(U_n), \xi_n(\tau)\rangle$$
$$\leq \|U_n\|^2 + \tau^2 \|f(U_n)\|^2 + \|\xi_n(\tau)\|^2 + 2\tau\left(\alpha + |\beta| \|U_n\|^2\right) + 2\langle U_n, \xi_n(\tau)\rangle + 2\langle \tau f(U_n), \xi_n(\tau)\rangle$$
$$\leq (1 + (1 + 2|\beta|)\tau)\|U_n\|^2 + 2\tau^2 \|f(U_n)\|^2 + (2 + \tau^{-1})\|\xi_n(\tau)\|^2 + 2\tau\alpha,$$

where we used Assumption 4.1) in the first inequality and Young's inequality in the second inequality for the bounds

$$2\langle U_n, \xi_n(\tau)\rangle \leq \tau \|U_n\|^2 + \tau^{-1}\|\xi_n(\tau)\|^2, \qquad 2\langle \tau f(U_n), \xi_n(\tau)\rangle \leq \tau^2 \|f(U_n)\|^2 + \|\xi_n(\tau)\|^2.$$

By the triangle inequality and the polynomial growth assumption (Assumption 4.2), we have

$$\|f(U_n)\|^2 = \|f(U_n) - f(0) + f(0)\|^2 \leq 2D\left(1 + \|U_n\|^q\right)\|U_n\|^2 + 2\|f(0)\|^2.$$

Combining the above inequalities, and using that $\tau < 1$ and $D \geq 1$ by Assumptions 4.2 and 4.5 respectively, we have

$$\|U_{n+1}\|^2 - \|U_n\|^2 \leq \tau(1 + 2|\beta|)\|U_n\|^2 + 4\tau^2 D\left(\|U_n\|^2 + \|U_n\|^{q+2} + \|f(0)\|^2\right) + 3\tau^{-1}\|\xi_n(\tau)\|^2 + 2\tau\alpha$$
$$= \tau(5D + 2|\beta|)\|U_n\|^2 + 4\tau^2 D\|U_n\|^{q+2} + 3\tau^{-1}\|\xi_n(\tau)\|^2 + \tau(2\alpha + 4\tau D\|f(0)\|^2).$$

Using a telescoping sum and the fact that $D \geq 1$, we obtain

$$\|U_n\|^2 \leq \|U_0\|^2 + \sum_{i=0}^{n-1} \tau(5D + 2|\beta|)\|U_i\|^2 + 4\tau^2 D\|U_i\|^{q+2} + 3\tau^{-1}\|\xi_i(\tau)\|^2 + \tau(2\alpha + 4\tau D\|f(0)\|^2)$$
$$\leq \left[\|U_0\|^2 + \sum_{i=0}^{n-1} 4\tau^2 D\|U_i\|^{q+2} + 3\tau^{-1}\|\xi_i(\tau)\|^2 + \tau(2\alpha + 4\tau D\|f(0)\|^2)\right]$$
$$+ \tau(5D + 2|\beta|)\sum_{i=0}^{n-1}\|U_i\|^2.$$

Grönwall's inequality and (5.3) yield

$$\max_{k \in [K]} \|U_k\|^2 \leq \left[\|U_0\|^2 + T(2\alpha + 4D\|f(0)\|^2) + 4\tau^2 D \sum_{i=0}^{K-1}\|U_i\|^{q+2} + 3\tau^{-1}\sum_{i=0}^{K-1}\|\xi_i(\tau)\|^2\right] e^{T(5D+2|\beta|)}$$
$$\leq \left[\|U_0\|^2 + T(2\alpha + 4D\|f(0)\|^2) + 4DT\tau \max_{k \in [K]}\|U_k\|^{q+2} + 3\tau^{-1}\sum_{i=0}^{K-1}\|\xi_i(\tau)\|^2\right] e^{T(5D+2|\beta|)}$$
$$\leq C_7 \left(1 + \tau \max_{k \in [K]}\|U_k\|^{q+2} + \tau^{-1}\sum_{i=0}^{K-1}\|\xi_i(\tau)\|^2\right)$$

and thus (5.2) for the case $n = 1$, with $C_7$ defined as in (5.3). To obtain the general case, we apply the generalised triangle inequality (Lemma 4.4). $\qquad\square$

**Proposition 5.2** (Bounds on even moments of maximum of numerical solution for explicit Euler)**.** *Suppose that the vector field $f$ satisfies Assumptions 4.1 and 4.2, that the noise $\xi_0$ satisfies Assumption 4.7 for some $R \geq 2n$ and some $p \geq 1$, and that $\tau$ satisfies Assumption 4.5. Then*

$$\mathbb{E}\left[\max_{k \in [K]}\|U_k\|^{2n}\right] \leq 4^{n-1} C_7^n \left(1 + \tau^n \mathbb{E}\left[\max_{k \in [K]}\|U_k\|^{(q+2)n}\right] + \left(C_2 T \tau^{2p-1}\right)^n\right). \tag{5.4}$$

*for some scalars $C_1, C_2 > 0$ not depending on $\tau$.*

Note that we do not claim that the right-hand side of (5.4) above is finite.

*Proof.* The result follows by taking expectations of both sides of (5.2) and using the estimates of moments of sums of $\xi_i(\tau)$ from Lemma 4.10. $\qquad\square$

Modulo scalar multiplicative factors, the sole difference between Proposition 5.2 and its analogous result for implicit Euler integration is the presence of the term $\mathbb{E}[\max_{k\in[K]} \|U_k\|^{(q+2)n}]$. If $q > 0$, then this extra term implies that, even if one assumes that $\max_{k\in[K]} \|U_k\|^{q+2}$ has finite moments of all orders, it does not follow that the moment generating function of $\max_{k\in[K]} \|U_k\|$ is finite on $\mathbb{R}$.

**Theorem 5.3.** *Let $U_0$ be fixed and nonrandom, and let $q \in \mathbb{N}_0$ be the exponent of polynomial growth in Assumption 4.2. If the noise $\xi_0$ is $\mathbb{P}$-almost surely bounded, then the stochastic process $(U_k)_{k\in[K]}$ satisfies*

$$\mathbb{E}\left[\exp\left(t \max_{k\in[K]} \|U_k\|^q\right)\right] < \infty \text{ for all } t \in \mathbb{R}.$$

*Proof.* Let $b > 0$ be such that $\mathbb{P}(\|\xi_0\| \leq b) = 1$. Since the $(\xi_k)_{k\in[K]}$ are i.i.d. copies of $\xi_0$, it follows that

$$\mathbb{P}\left(\max_{k\in[K]} \|U_k\| \leq \|U_0\| + Kb\right) = 1.$$

Hence we may bound the moment generating function by $\exp(t(\|U_0\| + Kb)^q)$ for all $t \in \mathbb{R}$. $\qquad\square$

## 5.2 Convergence result: discrete time

Let $V(\cdot, \cdot)$ be the difference of displacements operator, defined as in (4.17), let $V_{k-1}$ and $W_{k-1}$ be defined as in (4.22) as

$$V_{k-1} := V(u_{k-1}, U_{k-1}), \qquad\qquad W_{k-1} := \Phi^\tau(U_{k-1}) - \Psi^\tau(U_{k-1}).$$

Recall that, by (4.23), the error $e_k := u_k - U_k$ satisfies

$$e_k = e_{k-1} + V_{k-1} + W_{k-1} + \xi_{k-1}(\tau).$$

The proof of the following proposition is much simpler than that of Proposition 4.15, because explicit Euler integration corresponds to the approximation of an integral by a left Riemann sum.

**Proposition 5.4** (Bounds on moments of numerical discrepancy for explicit Euler)**.** *Let $\tau > 0$ satisfy Assumption 4.5, $D \geq 1$ and $q \in \mathbb{N}_0$ be the parameters in the polynomial growth condition (Assumption 4.2). For every $n \in \mathbb{N}$, it holds that*

$$\mathbb{E}\left[\|W_{k-1}\|^{2n}\right] \leq D^n \mathbb{E}\left[C_4^n(U_{k-1})\right] \tau^{4n}, \tag{5.5}$$

*where $C_4(U_{k-1}) := C_4(U_{k-1}, U_{k-1})$ for $C_4(\cdot, \cdot)$ defined in (4.19).*

**Remark 5.5.** By the definition (4.19) of $C_4(\cdot, \cdot)$, a sufficient condition for $C_4(U_{k-1})$ to have finite absolute moment of order $n$ is that $U_{k-1}$ has finite absolute moment of order $n$.

*Proof of Proposition 5.4.* Recall that $W_{k-1} = \Phi^\tau(U_{k-1}) - \Psi^\tau(U_{k-1})$. By (5.1), the left-hand side of (4.26) equals $\|W_{k-1}\|^2$. Thus, we obtain (5.5) by raising both sides of (4.26) to the $n^{\text{th}}$ power and taking expectations. $\qquad\square$

Define

$$\mathcal{G} := \sigma\left(\max_{k\in[K]} \|U_k\|\right), \qquad\qquad U^* := D\left(1 + \max_{k\in[K]} \|u_k\|^q + \max_{k\in[K]} \|U_k\|^q\right). \tag{5.6}$$

Note that $\mathcal{G}$ is a proper sub-$\sigma$-algebra of the $\sigma$-algebra generated by the collection of random variables $(\xi_k(\tau))_{k\in[K]}$. Thus, $\max_{k\in[K]} \|U_k\|^q$ is measurable with respect to $\mathcal{G}$. Furthermore, note that the random variable $U^*$ is measurable with respect to $\mathcal{G}$.

**Proposition 5.6** (Bounds on conditional expectations of products of difference of displacements for explicit Euler). *Suppose that Assumptions 4.1, 4.2, and 4.5 hold. Let $V_{k-1}$ be defined as in (4.22), let $q \in \mathbb{N}_0$ denote the exponent in Assumption 4.2, and suppose that the regularity assumption (Assumption 4.7) holds for all $R \in \mathbb{N}$. If the moment generating function of $\max_{k \in [K]} \|U_k\|^q$ is finite, then*

$$\mathbb{E}\big[|\langle e_{k-1}, V_{k-1}\rangle| \big| \mathcal{G}\big] \leq 2\tau DU^* \mathbb{E}\big[\|e_{k-1}\|^2 \big| \mathcal{G}\big] + \tau^3 \mathbb{E}\big[C_4(u_{k-1}, U_{k-1}) \big| \mathcal{G}\big] \tag{5.7a}$$

$$\mathbb{E}\big[\|V_{k-1}\|^2 \big| \mathcal{G}\big] \leq 2\tau^2 DU^* \mathbb{E}\big[\|e_{k-1}\|^2 \big| \mathcal{G}\big] + 2\tau^4 \mathbb{E}\big[C_4(u_{k-1}, U_{k-1}) \big| \mathcal{G}\big], \tag{5.7b}$$

*for any $k \in [K]$, where $C_4(u_0, v_0)$ is defined in Lemma 4.14.*

*Proof.* Taking the conditional expectation of (4.30) with respect to $\mathcal{G}$ yields

$$\mathbb{E}\big[|\langle e_{k-1}, V_{k-1}\rangle| \big| \mathcal{G}\big] \leq 2\tau D\mathbb{E}\big[(1 + \|u_{k-1}\|^q + \|U_{k-1}\|^q) \|e_{k-1}\|^2 \big| \mathcal{G}\big] + \mathbb{E}\big[C_4(u_{k-1}, U_{k-1}) \big| \mathcal{G}\big]\tau^3$$

$$\mathbb{E}\big[\|V_{k-1}\|^2 \big| \mathcal{G}\big] \leq 2\tau^2 D\mathbb{E}\big[(1 + \|u_{k-1}\|^q + \|U_{k-1}\|^q) \|e_{k-1}\|^2 \big| \mathcal{G}\big] + 2\mathbb{E}\big[C_4(u_{k-1}, U_{k-1}) \big| \mathcal{G}\big]\tau^4.$$

Bounding $(1 + \|u_{k-1}\|^q + \|U_{k-1}\|^q)$ by $U^*$ and using that $U^*$ is measurable with respect to $\mathcal{G}$ yields the desired result. $\qquad\square$

**Proposition 5.7** (Bounds on expected values of products of difference of displacements for explicit Euler and sufficiently fast polynomial growth). *Suppose that Assumptions 4.1, 4.2, and 4.5 hold, with $q$ being the exponent of polynomial growth. If $q \geq 6$, the $(p, R)$-regularity assumption (Assumption 4.7) holds with parameters $R \geq 2q$ and $p \geq 1$, and the random variable $\max_{k \in [K]} \|U_k\|^{(q+2)q}$ is integrable, then the estimates (4.28) hold.*

Note that we must formulate the integrability of $\max_{k \in [K]} \|U_k\|^{(q+2)q}$ as a separate condition from that of the regularity assumption holding for $R \geq 2q$, because Proposition 5.2 does not guarantee that this random variable is integrable. Note also that, for explicit Euler integration, a suitable constant $C'$ for (4.28) is given by

$$C'(k-1, q) := 4^{q-1} C_7^q \max\left\{1, (C_2 T)^{q/2} \mathbb{E}\left[\|e_{k-1}\|^4\right]^{1/2}, \left(\mathbb{E}\left[\max_{i \in [k-1]} \|U_i\|^{(q+2)q}\right] \mathbb{E}\left[\|e_{k-1}\|^4\right]\right)^{1/2}\right\},$$

which differs from $C'$ for implicit Euler integration (as defined in (4.29)).

*Proof.* The proof proceeds in a similar fashion to the proof of case (ii) in Proposition 4.16, with the exception that we must control the expectation of the product $\|U_{k-1}\|^q \|e_{k-1}\|^2$ using Proposition 5.2 instead of Proposition 4.12. By (5.2) in Lemma 5.1,

$$\max_{k \in [K]} \|U_k\|^{2q} \leq 4^{q-1} C_7^q \left(1 + \tau^n \max_{k \in [K]} \|U_k\|^{(q+2)q} + \tau^{-q} \left(\sum_{i=0}^{K-1} \|\xi_i(\tau)\|^2\right)^q\right).$$

For simplicity, we shall omit the multiplicative factor $4^{q-1} C_7^q$ in what follows; we shall reinsert it later. Using $\|U_k\|^q \leq \max_{i \leq k} \|U_k\|^q$ and the concavity of the square root,

$$\mathbb{E}\left[\|U_{k-1}\|^q \|e_{k-1}\|^2\right] \leq \mathbb{E}\left[\left(1 + \tau^q \max_{i \in [k-1]} \|U_i\|^{(q+2)q} + \tau^{-q}\left(\sum_{i=0}^{k-1} \|\xi_i(\tau)\|^2\right)^q\right)^{1/2} \|e_{k-1}\|^2\right]$$

$$\leq \mathbb{E}\left[\|e_{k-1}\|^2\right] + \tau^{-q/2} \mathbb{E}\left[\left(\sum_{i \in [k-1]} \|\xi_i(\tau)\|^2\right)^{q/2} \|e_{k-1}\|^2\right]$$

$$+ \tau^{q/2} \mathbb{E}\left[\max_{i \in [k-1]} \|U_i\|^{(q+2)q/2} \|e_{k-1}\|^2\right].$$

The first two terms on the right-hand side of the last inequality were bounded in the proof of Proposition 4.16, and in particular (4.31); we use the assumption that the $(p, R)$ regularity condition holds for $R \geq 2q$

and $p \geq 1$ for (4.31). Since $q \geq 6$ and the random variable $\max_{k \in [K]} \|U_k\|^{(q+2)/q}$ is integrable, the terms considered thus far are finite. It remains to consider the term that does not appear in the proof of Proposition 4.12, which we do now. By Young's inequality,

$$\mathbb{E}\left[\max_{i \in [k-1]} \|U_i\|^{(q+2)q/2} \|e_{k-1}\|^2\right] \leq \left(\mathbb{E}\left[\max_{i \in [k-1]} \|U_i\|^{(q+2)q}\right] \mathbb{E}\left[\|e_{k-1}\|^4\right]\right)^{1/2}.$$

By the assumptions that $q \geq 6$ and that $\max_{k \in [K]} \|U_k\|^{(q+2)q}$ is integrable, the right-hand side is finite. We have thus shown that

$$\mathbb{E}\left[\|U_{k-1}\|^q \|e_{k-1}\|^2\right] \leq C'\left(\mathbb{E}\left[\|e_{k-1}\|^2\right] + \tau^{q(2p-1)/2} + \tau^{q/2}\right). \tag{5.9}$$

Since $q \geq 6$ and $p \geq 1$, both exponents of $\tau$ on the right-hand side are bounded from below by 3, and the desired conclusion follows from the estimates (4.30). □

We now state the discrete-time convergence result for randomised explicit Euler. We omit the case $q = 0$, since this corresponds to a globally Lipschitz vector field, and this situation is covered by the results of Section 3.

**Theorem 5.8** (Mean-square uniform convergence in discrete time). *Suppose that Assumptions 4.1, 4.2, and 4.5 hold. Let $q \in \mathbb{N}$ denote the exponent in Assumption 4.2. If $p \geq 3/2$ and either*
  *(i) $q \geq 6$, Assumption 4.7 holds for $R \geq 2q$, and $\max_{k \in [K]} \|U_k\|^{(q+2)q}$ is integrable, or*
  *(ii) the moment generating function of $\max_{k \in [K]} \|U_k\|^q$ is finite on $\mathbb{R}$,*
*then there exists a strictly positive scalar $C$ that remains bounded in the limit of small $\tau$ and does not depend on $k$, such that*

$$\mathbb{E}\left[\max_{k \in [K]} \|e_k\|^2\right] \leq C\tau^2.$$

*In particular, if Assumption 4.7 holds for all $R \in \mathbb{N}$ and $\xi_0(\tau)$ is $\mathbb{P}$-almost surely bounded, then randomised explicit Euler integration converges with order one.*

*Proof.* Note that the final statement of the theorem follows from the first, by Theorem 5.3. We proceed as in the proof of Theorem 4.19 up to the inequality

$$\max_{i \leq n} \|e_i\|^2$$
$$\leq \sum_{k=1}^{n}\left[2\tau \max_{i \leq k-1} \|e_i\|^2 + 2|\langle e_{k-1}, V_{k-1}\rangle| + 3\|V_{k-1}\|^2 + \left(3 + \frac{1}{\tau}\right)\|W_{k-1}\|^2 + \left(3 + \frac{1}{\tau}\right)\|\xi_{k-1}(\tau)\|^2\right].$$

At this point, we consider the two cases separately.

**Case (i)** The assumptions in this case permit the application of Proposition 5.7. Continuing the proof of Theorem 4.19 yields the desired result.

**Case (ii)** Taking the conditional expectation with respect to $\mathcal{G}$, using Proposition 5.6, bounding $\|e_{k-1}\|^2$ by $\max_{i \leq k-1} \|e_i\|^2$, collecting terms, and using that $\tau < 1$ by Assumption 4.5 in order to bound $\tau^2$ by $\tau$ and $(3 + \tau^{-1})$ by $4\tau^{-1}$, we obtain

$$\mathbb{E}\left[\max_{i \leq n} \|e_i\|^2\Big|\mathcal{G}\right] \leq \sum_{k=1}^{n}(2+3)2\tau DU^* \mathbb{E}\left[\max_{i \leq k-1} \|e_i\|^2\Big|\mathcal{G}\right]$$
$$+ \sum_{k=1}^{n}\left(\mathbb{E}\left[4\tau^{-1}\|W_{k-1}\|^2 + 4\tau^{-1}\|\xi_{k-1}(\tau)\|^2 + (2+3)2\tau^3 C_4(u_{k-1}, U_{k-1})\Big|\mathcal{G}\right]\right).$$

Applying Grönwall's inequality, and using the measurability of $U^*$ with respect to $\mathcal{G}$, we obtain

$$\mathbb{E}\left[\max_{i\leq n}\|e_i\|^2\Big|\mathcal{G}\right]$$

$$\leq \exp(10TDU^*)\sum_{k=1}^{K}\left(4\tau^{-1}\mathbb{E}\left[\|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2\big|\mathcal{G}\right] + 10\tau^3\mathbb{E}\left[C_4(u_{k-1},U_{k-1})\big|\mathcal{G}\right]\right)$$

$$\leq \sum_{k=1}^{K}4\tau^{-1}\mathbb{E}\left[\exp(10TDU^*)\left(\|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2\right)\big|\mathcal{G}\right]$$

$$+ 10\tau^3\mathbb{E}\left[\exp(10TDU^*)C_4(u_{k-1},U_{k-1})\big|\mathcal{G}\right].$$

Upon taking expectations and applying the towering property of conditional expectation, we obtain

$$\mathbb{E}\left[\max_{i\leq n}\|e_i\|^2\right] \leq \sum_{k=1}^{K}\mathbb{E}\left[\exp(10TDU^*)\left(\frac{4}{\tau}\left(\|W_{k-1}\|^2 + \|\xi_{k-1}(\tau)\|^2\right) + 10\tau^3 C_4(u_{k-1},U_{k-1})\right)\right].$$

it remains to show that the three resulting sums are of the correct order in $\tau$. We do this using the Cauchy–Schwarz inequality, the conclusion (5.5) of Proposition 5.4, Assumption 4.7, and Theorem 5.3:

$$4\tau^{-1}\sum_{k=1}^{K}\mathbb{E}\left[\|W_{k-1}\|^2\exp(10TDU^*)\right] \leq \frac{4}{\tau}\frac{T}{\tau}\left(C\tau^8\mathbb{E}\left[\exp\left(20TDU^*\right)\right]\right)^{1/2} = O(\tau^2)$$

$$4\tau^{-1}\sum_{k=1}^{K}\mathbb{E}\left[\|\xi_{k-1}\|^2\exp(10TDU^*)\right] \leq \frac{4}{\tau}\frac{T}{\tau}\left(C(4)\tau^{2(2p+1)}\mathbb{E}\left[\exp\left(20TDU^*\right)\right]\right)^{1/2} = O(\tau^{2p-1})$$

$$10\tau^3\sum_{k=1}^{K}\mathbb{E}\left[\exp\left(10TDU^*\right)C_4(u_{k-1},U_{k-1})\right] \leq 10T\tau^2\max_{k\in[K]}\mathbb{E}\left[C_4^2(u_{k-1},U_{k-1})\right] = O(\tau^2).$$

By the assumption that $p \geq 3/2$, it follows that all three sums are of order 2 in $\tau$, as desired. $\qquad\square$

## Acknowledgements

## References

F.-X. Briol, C. Oates, M. Girolami, and M. A. Osborne. Frank–Wolfe Bayesian quadrature: Probabilistic integration with theoretical guarantees. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 1162–1170. Curran Associates, Inc., 2015. http://papers.nips.cc/paper/5749-frank-wolfe-bayesian-quadrature-probabilistic-integration-with-theoretical-guarantee

M. A. Capistrán, J. A. Christen, and S. Donnet. Bayesian analysis of ODEs: solver optimal accuracy and Bayes factors. *SIAM/ASA J. Uncertain. Quantif.*, 4(1):829–849, 2016. doi:10.1137/140976777.

O. A. Chkrebtii, D. A. Campbell, B. Calderhead, and M. A. Girolami. Bayesian solution uncertainty quantification for differential equations. *Bayesian Anal.*, 11(4):1239–1267, 2016. doi:10.1214/16-BA1017.

J. Cockayne, C. Oates, T. J. Sullivan, and M. Girolami. Probabilistic meshless methods for partial differential equations and Bayesian inverse problems, 2016. arXiv:1605.07811.

J. Cockayne, C. Oates, T. J. Sullivan, and M. Girolami. Bayesian probabilistic numerical methods, 2017. arXiv:1702.03673.

P. R. Conrad, M. Girolami, S. Särkkä, A. M. Stuart, and K. C. Zygalakis. Statistical analysis of differential equations: introducing probability measures on numerical solutions. *Stat. Comput.*, 2016. doi:10.1007/s11222-016-9671-0.

P. Diaconis. Bayesian numerical analysis. In *Statistical Decision Theory and Related Topics, IV, Vol. 1 (West Lafayette, Ind., 1986)*, pages 163–175. Springer, New York, 1988.

W. Fang and M. B. Giles. Adaptive Euler–Maruyama method for SDEs with non-globally Lipschitz drift: Part i, finite time interval, 2016. arXiv:1609.08101.

J. Gonzalez, M. Osborne, and N. Lawrence. GLASSES: Relieving the myopia of Bayesian optimisation. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 790–799. 2016. http://jmlr.org/proceedings/papers/v51/gonzalez16b.html.

P. Hennig. Probabilistic interpretation of linear solvers. *SIAM J. Optim.*, 25(1):234–260, 2015. doi:10.1137/140955501.

P. Hennig, M. A. Osborne, and M. Girolami. Probabilistic numerics and uncertainty in computations. *P. Roy. Soc. Lond. A Mat.*, 471(2179):20150142, 2015. doi:10.1098/rspa.2015.0142.

D. J. Higham, X. Mao, and A. M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.*, 40(3):1041–1063, 2002. doi:10.1137/S0036142901389530.

J. M. Holte. Discrete Gronwall lemma and applications, 2009. http://homepages.gac.edu/ holte/publications/GronwallLemma.pdf.

A. R. Humphries and A. M. Stuart. Runge–Kutta methods for dissipative and gradient dynamical systems. *SIAM J. Numer. Anal.*, 31(5):1452–1485, 1994. doi:10.1137/0731075.

A. Jentzen and A. Neuenkirch. A random Euler scheme for Carathéodory differential equations. *J. Comput. Appl. Math.*, 224(1):346–359, 2009. doi:10.1016/j.cam.2008.05.060.

J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2005. doi:10.1007/b138659.

R. Kruse and Y. Wu. Error analysis of randomized runge–kutta methods for differential equations with time-irregular coefficients, 2017. arXiv:1701.03444.

K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation: A Mathematical Introduction*, volume 62 of *Texts in Applied Mathematics*. Springer, 2015. doi:10.1007/978-3-319-20325-6.

X. Mao and L. Szpruch. Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients. *J. Comp. Appl. Math.*, 238: 14 – 28, 2013. doi:10.1016/j.cam.2012.08.015.

E. Novak. *Deterministic and Stochastic Error Bounds in Numerical Analysis*, volume 1349 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1988. doi:10.1007/BFb0079792.

A. O'Hagan. Some Bayesian numerical analysis. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors, *Bayesian Statistics, 4: Proceedings of the Fourth Valencia International Meeting: Dedicated to the memory of Morris H. DeGroot, 1931–1989: April 15–20, 1991*, pages 345–363. Clarendon Press, 1992.

H. Owhadi. Bayesian numerical homogenization. *Multiscale Model. Simul.*, 13(3):812–828, 2015. doi:10.1137/140974596.

H. Owhadi. Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *SIAM Rev.*, 59(1):99–149, 2017. doi:10.1137/15M1013894.

G. Peškir. On the exponential Orlicz norms of stopped Brownian motion. *Studia Math.*, 117(3):253–273, 1996. http://eudml.org/doc/216255.

M. Raissi and G. E. Karniadakis. Machine learning of linear differential equations using Gaussian processes, 2017. arXiv:1701.02440.

S. Reich and C. Cotter. *Probabilistic Forecasting and Bayesian Data Assimilation*. Cambridge University Press, New York, 2015. doi:10.1017/CBO9781107706804.

Y.-F. Ren. On the Burkholder–Davis–Gundy inequalities for continuous martingales. *Statist. Probab. Lett.*, 78(17):3034–3039, 2008. ISSN 0167-7152. doi:10.1016/j.spl.2008.05.024.

D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. doi:10.1007/978-3-662-06400-9.

K. Ritter. *Average-Case Analysis of Numerical Problems*, volume 1733 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000. doi:10.1007/BFb0103934.

M. Schober, D. K. Duvenaud, and P. Hennig. Probabilistic ODE solvers with Runge–Kutta means. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Ad-

*vances in Neural Information Processing Systems 27*, pages 739–747. Curran Associates, Inc., 2014. http://papers.nips.cc/paper/5451-probabilistic-ode-solvers-with-runge-kutta-means.

J. Skilling. Bayesian solution of ordinary differential equations. In C. R. Smith, G. J. Erickson, and P. O. Neudorfer, editors, *Maximum Entropy and Bayesian Methods*, volume 50 of *Fundamental Theories of Physics*, pages 23–37. Springer, 1992. doi:10.1007/978-94-017-2219-3.

R. C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*, volume 12 of *Computational Science & Engineering*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014.

G. Stengle. Numerical methods for systems with measurable coefficients. *Appl. Math. Lett.*, 3(4):25–29, 1990. doi:10.1016/0893-9659(90)90040-I.

A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010. doi:10.1017/S0962492910000061.

T. J. Sullivan. *Introduction to Uncertainty Quantification*, volume 63 of *Texts in Applied Mathematics*. Springer, 2015. doi:10.1007/978-3-319-23395-6.

L. Takács. On the distribution of the integral of the absolute value of the Brownian motion. *Ann. Appl. Probab.*, 3(1):186–197, 1993. doi:10.1214/aoap/1177005514.

J. F. Traub and H. Woźniakowsi. *A General Theory of Optimal Algorithms*. ACM Monograph Series. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980.

J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski. *Information, Uncertainty, Complexity*. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1983.