INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 3 , Number 02, Pages 459–480

AN AXIOMATIC APPROACH TO NUMERICAL APPROXIMATIONS OF STOCHASTIC PROCESSES

HENRI SCHURZ

(Communicated by Edward J. Allen)

This paper is dedicated to emeritus professor P. Heinz Müller at TU Dresden.

Abstract. An axiomatic approach to the numerical approximation Y of some stochastic process X with values on a separable Hilbert space H is presented by means of Lyapunov-type control functions V. The processes X and Y are interpreted as flows of stochastic differential and difference equations, respectively. The main result is the proof of some extensions of well-known deterministic principle of Kantorovich-Lax-Richtmeyer to approximate solutions of initial value differential problems to the stochastic case. The concepts of invariance, smoothness of martingale parts, consistency, stability, and contractivity of stochastic processes are uniquely combined to derive efficient convergence rates on finite and infinite time-intervals. The applicability of our results is explained with drift-implicit backward Euler methods applied to ordinary stochastic differential equations (SDEs) driven by standard Wiener processes on Euclidean spaces $H = \mathbb{R}^d$ along functions such as $V(x) = \sum_{i=0}^{k} c_i x^{2i}$. A detailed discussion on an example with cubic nonlinearity from field theory in physics (stochastic Ginzburg-Landau equation) illustrates the suggested axiomatic approach.

Key Words. stochastic differential equations, numerical methods, stochastic difference equations, convergence, stability, contractivity, stochastic Kantorovich-Lax-Richtmeyer principle, Lyapunov-type functions, worst case convergence rates

1. Introduction

Many dynamic problems in Natural Sciences, Engineering, Environmental Sciences and Econometrics lead to models governed by nonlinear and dissipative stochastic ordinary and partial differential systems. These systems are explicitly solvable very rarely. Thus one has to resort to numerical approximations. In deterministic theory there are well-known principles for the approximation of their solutions in appropriate Banach spaces. Two of them are the principles of Kantorovič [17], [11] and Lax and Richtmeyer [24], [34], combining stability, consistency and convergence for well-posed problems. However, in the stochastic case, there is substantially less known about their counterparts. We are going to continue our works exhibited in [35] - [46] by establishing basic approximation principles for stochastic processes X, Y which have values in random Hilbert spaces H or Banach spaces with norms defined via subadditive pseudo-bilinear forms. As the simplest application we bear

Received by the editors July 21, 2005 and, in revised form, August 24, 2005.
2000 *Mathematics Subject Classification.* 65C20, 65C30, 65L20, 65D30, 34F05, 37H10, 60H10.
This research was supported by Southern Illinois University and Texas Tech University.

in mind the case of stochastic ordinary differential equations (SDEs) and their numerical approximations with variable step sizes. (An application to some types of stochastic partial differential equations (SPDEs) with appropriate relation between space- and time-discretization for their approximations is conceivable too, but left to future work). In this paper the time-evolution of the global discretization-error is considered without taking into account any discretization of the state space. Note that the herein suggested axiomatic approach to the analysis of numerical approximations is especially efficient within the framework of "eigenfunction approach" applied to quasilinear SPDEs.

For the description of the approximation problem we assume the following. Fix a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$ with deterministic finite time-interval $[0,T]$. Let $H = H(\omega)$ be a separable random Hilbert space with $(\mathcal{F}_t)_{0\le t\le T}$-adapted scalar product $< .,. >_H$ and real numbers as its scalars, and let μ be any nonrandom, σ-finite, positive measure on $([0,T], \mathcal{B}([0,T]))$. Here $\mathcal{B}(.)$ represents the σ-field of all Borel-sets of the inscribed set. $X = (X_t(\omega))_{0\le t\le T}$ and $Y = (Y_t(\omega))_{0\le t\le T}$ denote two $(\mathcal{F}_t)$-adapted stochastic processes on the given probability space with values in one and the same Hilbert space H. Then, obviously, the vector space

$$H_2([0,T],\mu,H) := \left\{ X = (X_t(\omega))_{0\le t\le T} : \begin{array}{l} X_t(\omega) \in H(\omega) \text{ for all times } t, \\ X_t \text{ is } (\mathcal{F}_t, \mathcal{B}(H)) - \text{measurable}, \\ X \text{ cadlag with respect to time } t, \\ \int_0^T \mathbb{E} < X_t, X_t >_H \, d\mu(t) < +\infty \end{array} \right\}$$

forms a Hilbert space with scalar product

$$< X, X >_{H_2} := \int_0^T \mathbb{E} < X_t, X_t >_H \, d\mu(t)$$

and real numbers as its scalars. The naturally induced norms are given by

$$||X||_H := \sqrt{< X, X >_H}, \quad ||X||_{H_2} := \sqrt{< X, X >_{H_2}}.$$

We are interested to tackle the approximation problem of X by Y (and also Y by X, thanks to the inherent symmetry) on this space, in particular, on the subset

$$\mathbb{D}_T = \left\{ X \in H_2([0,T],\mu,H) : \sup_{0\le t\le T} \mathbb{E} < X_t, X_t >_H < +\infty \right\}.$$

Furthermore, let $[K]_- \ge 0$ denote the negative part of K, and $[K]_+ \ge 0$ its positive part such that we have $K = [K]_+ - [K]_-$.

The paper is organized as follows. Section 2 commences with the statement of main concepts and assumptions to prove a fairly general approximation theorem for convergence rates of numerical approximations with variable step sizes. In Sections 3 and 4 we present two versions of this theorem for the most general and dissipative case. The main purpose of this paper is to publish a fairly complete proof of universal error estimates for the approximation of some Hilbert-space-valued stochastic processes while incorporating information on certain Lyapunov-function(al)s $V = V(x)$. This significantly extends the applicability of our original work [45] where we only considered the very restricted case of $V(x) = 1+||x||^2$ from practical point of view (cf. example in Section 6.2). The main theorems 3.1 and 4.1 have already been formulated in [44], but without any detailed proof-steps. Here the complete proof incorporating the role of Lyapunov-functions $V(x)$ (much more general than $V(x) = 1 + ||x||^2$) is presented by dividing it into a series of auxiliary lemmas as done in Section 5. Section 6 briefly discusses the fairly transparent case of ordinary stochastic differential equations and drift-implicit Euler methods in $\mathbb{R}^d$, including a specific example with cubic nonlinearity. Section 7 exhibits the main diagram of adequate stochastic approximation theory as a kind of summary.

2. Main concepts of numerical approximation of well-posed problems

Let $X_{s,Z}(t), Y_{s,Z}(t)$ be the one-step representations of stochastic processes X, Y evaluated at time $t \geq s$, started from $Z \in H_2([0,s],\mu,H)$ (i.e., more precisely, we have $X_{s,Z}(u) = Z_u = Y_{s,Z}(u)$ for $0 \leq u \leq s$ and $X_{s,Z}(t), Y_{s,Z}(t)$ are interpreted as the values of the stochastic processes X and Y in H at time $t \geq s$, respectively, with fixed history (memory) given by Z up to time $s \geq 0$). They are supposed to be constructable along any $(\mathcal{F}_t)$-adapted discretization of the given time-interval $[0,T]$ and could depend on a certain mesh size Δ_{max}. Assume that there are deterministic real constants $r_0, r_{sm}, r_2 \geq 0, 0 < \delta_0 \leq 1$ such that we have

(A1) **Strong $(\mathbb{D}_t)$-invariance of** X, Y, i.e. $\exists (\mathcal{F}_t)$-adapted, closed subsets $\mathbb{D}_t \subseteq H_2([0,t],\mu,H)$ such that, for all $0 \leq s < T$,

$$\mathbb{P}\{(X_{s,X(s)}(u))_{s\leq u\leq t}, (Y_{s,Y(s)}(u))_{s\leq u\leq t} \in \mathbb{D}_t : s \leq t \leq T | X(s), Y(s) \in \mathbb{D}_s\} = 1,$$

(A2) V**-Stability of** Y, i.e. $\exists$ functional $V : H_2([0,t],\mu,H) \to \mathbb{R}_+$ for all $0 \leq t \leq T$ such that $\forall Y(t) \in \mathbb{D}_t$: $\mathbb{E}\, V(Y(t)) < +\infty$, $V(Y(t))$ is $(\mathcal{F}_t)$-adapted and $\exists$ real constant K_S^Y $\forall t,h : 0 \leq h \leq \delta_0, 0 \leq t, t+h \leq T$

$$\mathbb{E}\,[V(Y_{t,Y(t)}(t+h))|\mathcal{F}_t] \;\leq\; \exp(2K_S^Y h) \cdot V(Y(t)),$$

(A3) **Mean square contractivity of** X, i.e. $\exists$ real constant K_C^X such that $\forall\, X(t), Y(t) \in \mathbb{D}_t$ (where $X(t), Y(t)$ are $(\mathcal{F}_t)$-adapted) $\forall t,h : 0 \leq h \leq \delta_0, 0 \leq t, t+h \leq T$

$$\begin{aligned}\mathbb{E}\,[||X_{t,X(t)}(t+h) \;-\; & X_{t,Y(t)}(t+h)||_H^2 | X(t), Y(t)] \\ \leq\; & \exp(2K_C^X h) \cdot ||X_{t,X(t)}(t) - X_{t,Y(t)}(t)||_H^2,\end{aligned}$$

(A4) **Mean consistency of** (X,Y) **with rate** $r_0 > 0$, i.e. $\exists$ real constant K_0^C such that $\forall\, Z(t) \in \mathbb{D}_t$ (where $(Z_u(t))_{0\leq u\leq t}$ is $(\mathcal{F}_t, \mathcal{B}(H))$-measurable) $\forall t,h : 0 \leq h \leq \delta_0, 0 \leq t, t+h \leq T$

$$||\,\mathbb{E}\,[X_{t,Z(t)}(t+h)|Z(t)] - \mathbb{E}\,[Y_{t,Z(t)}(t+h)|Z(t)]||_H \;\leq\; K_0^C \cdot \sqrt{V(Z(t))} \cdot h^{r_0},$$

(A5) **Mean square consistency of** (X,Y) **with rate** $r_2 > 0$, i.e. $\exists$ real constant K_2^C such that $\forall\, Z(t) \in \mathbb{D}_t$ $\forall t,h : 0 \leq h \leq \delta_0, 0 \leq t, t+h \leq T$

$$\left(\mathbb{E}\left[||X_{t,Z(t)}(t+h) - Y_{t,Z(t)}(t+h)||_H^2 | Z(t)\right]\right)^{1/2} \;\leq\; K_2^C \cdot \sqrt{V(Z(t))} \cdot h^{r_2}$$

(A6) **Mean square Hölder-type smoothness of diffusive (martingale) part of** X **with rate** $r_{sm} \in [0, \frac{1}{2}]$, i.e. $\exists$ real constant $K_{SM} \geq 0$ such that $\forall\, X(t), Y(t) \in \mathbb{D}_t$ (where $X(t), Y(t)$ are $(\mathcal{F}_t)$-adapted) $\forall t,h : 0 \leq h \leq \delta_0, 0 \leq t, t+h \leq T$

$$\mathbb{E}\,||M_{t,X(t)}(t+h) - M_{t,Y(t)}(t+h)||_H^2 \;\leq\; (K_{SM})^2 \cdot \mathbb{E}\,||X(t) - Y(t)||_H^2 \cdot h^{2r_{sm}}$$

$$\text{where} \quad M_{t,z}(t+h) \;=\; X_{t,z}(t+h) - \mathbb{E}\,[X_{t,z}(t+h)|\mathcal{F}_t] \;\text{ for }\; z = X(t), Y(t)\,,$$

(A7) **Interplay between consistency rates** given by $r_0 \geq r_2 + r_{sm} \geq 1.0$,

(A8) **Initial moment** V**-boundedness** $\mathbb{E}\,[V(X_0)] + \mathbb{E}\,[V(Y_0)] < +\infty$.

Stochastic approximation problems satisfying the assumptions $(A1)$ - $(A8)$ on H_2 are called **well-posed**. In the classical case of stochastic dynamics with Lipschitz-continuous vector coefficients like that of SDEs driven by Wiener processes one often takes the function $V((X_s)_{0\leq s\leq t}) = (1 + ||X_t||_H^2)^{p/2}$ or $||X_t||_H^p$ as the required functional V. Then, V plays the role of a Lyapunov function controlling the stability behavior of considered stochastic process and the smoothness condition $(A6)$ of the martingale part with $r_{sm} = 0.5$ is obviously satisfied. Of course, if $V(X) = 0$ is chosen, then X and Y must be identical and any derived convergence assertions based on above assumptions are meaningless.

Remark. Note that we need to distinguish clearly between the meaning of expressions X, X_t, $X_{t,Z(t)}(u)$ and $X(t) = (X_u(t))_{0\le u\le t}$ for the understanding of the complex set of above assumptions (analogously for Y).

3. A general approximation theorem under axioms $(A1) - (A8)$

The following fairly general approximation principle can be established. Define the **pointwise L^2-error**

$$\varepsilon_2(t) \;=\; \sqrt{\;\mathbb{E} < X_t - Y_t, X_t - Y_t >_H}$$

for the processes $X, Y \in H_2$, and the deterministic bounds

$$\Delta_{min} = \inf_{i\in\mathbb{N}} \Delta_i \;\le\; \Delta_n \;\le\; \Delta_{max} = \sup_{i\in\mathbb{N}} \Delta_i$$

on the mesh sizes Δ_n on which (at least one of) X, Y are usually based.

Theorem 3.1. *Assume that the conditions $(A1) - (A8)$ are satisfied and that $\mathbb{E}\,||X_0 - Y_0||_H^2 < K_{init}\Delta_{max}^{r_g}$.*
Then the stochastic processes $X, Y \in H_2([0,T],\mu,H)$ converge to each another with respect to the naturally induced metric

$$m(X,Y) = (< X - Y, X - Y >_{H_2})^{1/2}$$

with "worst case" convergence rate $r_g = r_2 + r_{sm} - 1.0$. More precisely, for any $\rho \neq 0$ and for any choice of deterministic step sizes Δ_i (variable or constant) with $0 < \Delta_i \le \Delta_{max} \le \delta_0$, we have the universal error estimates

$$(1)\quad \varepsilon_2(t) \;\le\; \exp((K_C^X + \rho^2)(t-s))\varepsilon_2(s) + \\ +K_I(\rho)\exp(K_S^Y t)\sqrt{\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)(t-s)) - 1}{2(K_C^X + \rho^2 - K_S^Y)}}\;\Delta_{max}^{r_g}$$

for all $0 \le s \le t \le T$, where s, t are deterministic, and

$$(2)\quad \sup_{0\le t\le T} \varepsilon_2(t) \;\le\; \exp([K_C^X + \rho^2]_+ T)\varepsilon_2(0) + \\ +K_I(\rho)\exp([K_S^Y]_+ T)\sqrt{\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)T) - 1}{2(K_C^X + \rho^2 - K_S^Y)}}\;\Delta_{max}^{r_g}$$

with appropriate constant

$$(3)\qquad K_I(\rho) \;=\; K_{max}\cdot\frac{\sqrt{(K_0^C)^2 + (K_2^C)^2[\rho^2 + (K_{SM})^2]}}{\rho}\cdot\sqrt{\mathbb{E}\,V(y_0)}$$

with $K_{max} \;=\; \exp(([K_C^X]_- + [K_S^Y]_-)\Delta_{max})$.

4. An approximation theorem for the dissipative case

The asymptotically contractive, dissipative case is covered as follows.

Theorem 4.1. *Assume that the conditions $(A1) - (A8)$ with $K_C^X < 0$ and $K_S^Y \le 0$ are satisfied on the time-interval $[0,+\infty)$ with the finite measure $\mu([0,+\infty))$, $\mathbb{E}\,||X_0 - Y_0||_H^2 < K_{init}\Delta_{max}^{r_g}$ and all constants K occurring there in $(A1)$ – $(A7)$ do not depend on the terminal times $T > 0$.*

Then the stochastic processes $X, Y \in H_2([0,+\infty),\mu,H)$ *converge to each another with respect to the naturally induced metric*

$$m(X,Y) = (< X-Y, X-Y >_{H_2})^{1/2}$$

with "worst case" convergence rate $r_g = r_2 + r_{sm} - 1.0$. *More precisely, for any choice of deterministic step sizes* Δ_i *(variable or constant) with* $0 < \Delta_i \le \Delta_{max} \le \delta_0$, *we have the universal error estimates*

$$\text{(1)} \qquad \sup_{0\le t<+\infty} \varepsilon_2(t) \;\le\; \varepsilon_2(0) + \frac{K_I(\rho)}{\sqrt{2|K_C^X + \rho^2 - K_S^Y|}}\Delta_{max}^{r_g},$$

$$\text{(2)} \qquad \lim_{t\to+\infty} \varepsilon_2(t) \;\le\; \frac{K_I(\rho)}{\sqrt{2|K_C^X + \rho^2 - K_S^Y|}}\Delta_{max}^{r_g} \quad \textit{if} \quad K_C^X + \rho^2 < 0$$

$$\text{(3)} \qquad \lim_{t\to+\infty} \varepsilon_2(t) \;=\; 0 \qquad \textit{if} \quad K_S^Y < 0, K_C^X + \rho^2 < 0$$

where

$$K_I(\rho) = \frac{\sqrt{(K_0^C)^2 + (K_2^C)^2(\rho^2 + (K_{SM})^2)}}{\rho} \cdot \sqrt{\mathbb{E}\, V(y_0)} \cdot \exp\Big((|K_C^X| + |K_S^Y|)\Delta_{max}\Big)$$

for any ρ *with* $0 < \rho^2 \le |K_C^X|$, *i.e. convergence on infinite intervals* $[0,+\infty)$ *with the "worst case" global rate* r_g *can be established on* $H_2([0,+\infty),\mu,H)$.

Using this theorem, one may establish convergence rates of numerical methods for SDEs along Lyapunov-functionals on infinite time-intervals $[0,+\infty)$. For example, the drift-implicit Euler method applied to mean square dissipative SDEs with monotone coefficients converges in the L^2-sense with the "worst case" rate $r_g = 0.5$ on infinite time-intervals $[0,+\infty)$ - a rather striking result. To verify it, one may take $H = \mathbb{D}_t = \mathbb{R}^d$, $V(x) = ||x||_d^2$ as the Euclidean norm, and shows that $r_0 \ge 1.5$, $r_2 = 1.0$ for diffusions with $r_{sm} = 0.5$. See also the discussion in Section 6.

5. Breakdown of proof steps of main theorems

It is more transparent to break down the complete proof of both theorems into the following series of auxiliary lemmas. The proof of some of them can be omitted since they are elementary, and mostly a consequence of the well-known Young's inequality (Hölder inequality) and complete inductions.

5.1. Auxiliary lemmas. Let $(\mathbb{B}, ||.||_{\mathbb{B}})$ be a Banach space with respect to the norm $||.||_{\mathbb{B}}$ and $\chi_{([l,r))}(t)$ be the characteristic function of the inscribed interval $[l,r)$.

Lemma 5.1. *Assume that* $a_i \in \mathbb{B}$ *(i=1,2,...,n). Then, for* $n \in \mathbb{N}, p \ge 1$, *we have*

$$||\sum_{i=1}^n a_i||_{\mathbb{B}}^p \;\le\; n^{p-1}\sum_{i=1}^n ||a_i||_{\mathbb{B}}^p, \qquad \sqrt[p]{||\sum_{i=1}^n a_i||_{\mathbb{B}}} \;\le\; \sum_{i=1}^n \sqrt[p]{||a_i||_{\mathbb{B}}}.$$

Proof. Suppose that $a_i \in \mathbb{B}$ are fixed. The proof directly follows from the triangle and Hölder inequality applied to the right-continuous L^p-integrable function $f : [0,n] \to \mathbb{R}_+^1$ given by

$$f(t) \;=\; \begin{cases} \displaystyle\sum_{i=1}^n ||a_i||_{\mathbb{B}} & \text{if} \quad t = n \\ \displaystyle\sum_{i=1}^n ||a_i||_{\mathbb{B}} \cdot \chi_{([i,i-1))}(t) & \text{if} \quad 0 \le t < n \end{cases}.$$

The second part is concluded from the well-known monotonicity of integrals. □

Lemma 5.2. *Let $(v(n))_{n\in\mathbb{N}}$ be a sequence of nonnegative real numbers $v(n)$ satisfying*

$$v(n+1) \leq v(n)(1+c_H(n)) + c_I(n) \quad or \quad v(n+1) \leq v(n)\exp(c_H(n)) + c_I(n)$$

with real coefficient sequences $(c_H(n))_{n\in\mathbb{N}}$ of homogeneity and $(c_I(n))_{n\in\mathbb{N}}$ of inhomogeneity. Then, for all $n \geq k, k \in \mathbb{N}$, we have

$$v(n+1) \leq v(k)\exp\Big(\sum_{l=k}^{n} c_H(l)\Big) + \sum_{l=k}^{n} c_I(l)\exp\Big(\sum_{i=l+1}^{n} c_H(i)\Big).$$

Proof. Use complete induction as done in [38]. □

Remark. We may meet the convention that $\sum_{k=n+1}^{n}(.) = 0$. The latter inequality is sometimes called the **discrete time variation-of-constants inequality**, and it is used to prove the following continous time version.

Lemma 5.3. *Let $v = v(t), -\infty < t_0 \leq t < +\infty$ be a nonnegative real–valued function which is locally absolutely Lebesgue-integrable on $[t_0, +\infty)$ (i.e. we could also use the notation $v \in L^1_{loc}([t_0,+\infty), \mathcal{B}([t_0,+\infty)), \mu)$ with Borel σ-field $\mathcal{B}([t_0,+\infty))$ and Lebesgue-measure μ). Assume that the coefficient functions $C_I = C_I(t), C_H = C_H(t) \in L^1_{loc}([t_0,+\infty), \mathcal{B}([t_0,+\infty)), \mu)$ are absolutely Lebesgue-integrable with*

$$\int_s^t C_I(u)\cdot exp\Big(\int_u^t C_H(z)\,dz\Big)\,du \;<\; +\infty$$

for all t, s with $t_0 \leq s \leq t$. Furthermore, $v(t)$ satisfies

$$\begin{aligned} v(t) &\leq v(s) + \int_s^t C_H(u)\cdot v(u)\,du + \int_s^t C_I(u)\,du \quad or \\ v(t) &\leq v(s)\cdot\exp\Big(\int_s^t C_H(u)\,du\Big) + \int_s^t C_I(u)\,du \end{aligned}$$

for all t, s with $t_0 \leq s \leq t$ and sufficiently small $|t-s|$ (say, e.g. $|t-s| \leq \delta$). Then the **continuous time variation-of-constants inequality** *holds, i.e.*

(1) $$v(t) \leq \Big(v(s) + \int_s^t C_I(u)\cdot\exp\Big(-\int_s^u C_H(z)\,dz\Big)\,du\Big)\cdot\exp\Big(\int_s^t C_H(u)\,du\Big)$$

for all t, s with $t_0 \leq s \leq t$.

Proof. First, suppose that C_I, C_H are continuous functions. Hence, they can be approximated by simple functions along any finite partition $(t_n)_{n\in\mathbb{N}}$ of a compact subinterval of $[t_0,+\infty)$. Define $c_H(n) := C_H(\tau_n)(t_{n+1}-t_n)$ and $c_I(n) := C_I(\tau_n)(t_{n+1}-t_n)$ where $\tau_n \in [t_n, t_{n+1}]$ are chosen such that the discretized function v will also satisfy the assumptions of Lemma 5.2. Then, we may use Lemma 5.2 to conclude the discretized version of the claim of Lemma 5.3. Now, one takes the limit as the mesh size of those partitions tends to 0, and we arrive at the claim of Lemma 5.3 for continuous coefficients C_H, C_I. It remains to mention the fact that the set of continuous functions is dense in L^1-spaces in order to conclude the desired assertion for L^1-integrable coefficients C_H, C_I. □

Remark. The latter two lemmas can be found in Schurz [38], and their use in [42] - [43]. As an immediate, but very helpful application of Lemma 5.3 one arrives at the following key lemma for the proof of main conclusions in the next section.

Lemma 5.4. *Let $v = v(t), -\infty < t_0 \le t < +\infty$ be a nonnegative real–valued function which is locally absolutely Lebesgue-integrable on $[t_0, +\infty)$. Assume that $v(t)$ satisfies*

$$v(t+h) \le v(t)\exp(K_H h) + K_I \exp(K_S t) h^{r_{loc}} \tag{2}$$

for all t, h with $t_0 \le t \le t+h \le T$ and $0 \le h \le \delta$ (δ sufficiently small) and a given $r_{loc} \ge 1$. Put $\hat{K}_I = K_I \exp\Big([K_S]_- \delta\Big)$. Then, for all t, s with $t_0 \le s \le t$, we have

$$\begin{aligned} & v(t) \\ \le \quad & v(s)\cdot\exp\Big(K_H(t-s)\Big) + \hat{K}_I \frac{\exp(K_S s + K_H(t-s)) - \exp(K_S t)}{K_H - K_S}\delta^{r_{loc}-1} \\ \le \quad & v(0)\cdot\exp\Big(K_H t\Big) + \hat{K}_I \frac{\exp(K_H t) - \exp(K_S t)}{K_H - K_S}\delta^{r_{loc}-1}. \end{aligned} \tag{3}$$

Proof. Suppose that $r_{loc} \ge 1$. First, note that, for $0 \le h \le \delta$, the relation

$$\exp\Big(K_S t\Big)\cdot h^{r_{loc}} \quad \le \quad \exp\Big([K_S]_- \delta\Big) \int_t^{t+h} \exp\Big(K_S u\Big)\, du \cdot \delta^{r_{loc}-1}$$

holds. Then, this result applied to the condition (2) reads as

$$\begin{aligned} v(t+h) \quad & \le \quad v(t)\exp(K_H h) + K_I \exp(K_S t) h^{r_{loc}} \\ & \le \quad v(t)\exp(K_H h) + \hat{K}_I \delta^{r_{loc}-1} \int_t^{t+h} \exp(K_S u) du \end{aligned}$$

for all t, h with $t_0 \le t \le t+h \le T$ and $0 \le h \le \delta$. The remaining proof is a straightforward application of Lemma 5.3 since its assumptions are satisfied. For the sake of completion, we evaluate the inequality (1) with identities $C_H(u) = K_H$ and $C_I(u) = \hat{K}_I \delta^{r_{loc}-1}\exp(K_S u)$. Thus, the conclusion (1) is

$$\begin{aligned} v(t) \quad & \le \quad v(s)\cdot\exp\Big(K_H(t-s)\Big) + \\ & \quad\quad + \hat{K}_I \delta^{r_{loc}-1}\cdot\exp\Big(K_H t\Big)\cdot\int_s^t \exp\Big((K_S - K_H)u\Big) du \\ & = \quad v(s)\cdot\exp\Big(K_H(t-s)\Big) + \\ & \quad\quad + \hat{K}_I \delta^{r_{loc}-1}\cdot\frac{\exp(K_S t) - \exp(K_S s + K_H(t-s))}{K_S - K_H}, \end{aligned}$$

hence the proof is completed by putting $s = 0$. □

Lemma 5.5. *For all $a, b, c \in H$ (H any given Hilbert space), we have*

$$||a-b||_H^2 = ||a-c||_H^2 + ||c-b||_H^2 + 2 < a-c, c-b >_H .$$

Proof. This can easily be verified by elementary calculation, analogously to the case $H = \mathbb{R}^d$. □

Recall that $||.||_H$ denotes the (in general random) norm defined naturally by $||a||_H^2 =< a, a >_H$ which is naturally induced by the scalar product $< ., . >_H$ of the Hilbert space H. The following lemma explains the role of the required functional V along certain moment estimates.

Lemma 5.6. *Assume that the assumptions (A1), (A2) and (A8) are satisfied. Then, for all* $0 \leq t, t+h \leq T$ *with* $(\mathcal{F}_t)$*-adapted* t, h*, we have*

$$(\mathbb{E}\,[V(Y_{0,y_0}(t+h))])^{1/2} \leq \left(\mathbb{E}\,\left[\exp(2K_S^Y h)V(Y_{0,y_0}(t))\right]\right)^{1/2} \leq \left(\mathbb{E}\,\left[\exp(2K_S^Y (t+h))V(y_0)\right]\right)^{1/2},$$

hence, for all deterministic times t *with* $0 \leq t \leq T$*, this implies that*

$$(\mathbb{E}\,[V(Y_{0,y_0}(t))])^{1/2} \leq \exp(K_S^Y t)(\mathbb{E}\,V(y_0))^{1/2}$$

and if $V(Y(t)) \geq ||Y_t||_H^2$ *for all* $t \in [0,T]$ *where* $Y(t) \in H_2([0,t],\mu,H)$ *then*

$$||Y||_{H_2([0,t],\mu,H)} \leq \left(\int_0^t \exp(2K_S^Y u)d\mu(u)\;\mathbb{E}\,V(y_0)\right)^{1/2} \leq \exp([K_S^Y]_+ T)\Big(\mu([0,T])\;\mathbb{E}\,V(y_0)\Big)^{1/2}.$$

Proof. Using elementary laws of conditional expectations leads to

$$\begin{aligned}(\mathbb{E}\,[V(Y_{0,y_0}(t+h))])^{1/2} &= (\mathbb{E}[\,\mathbb{E}\,[V(Y_{0,y_0}(t+h))|\mathcal{F}_t]])^{1/2} \\ &= \left(\mathbb{E}\left[\,\mathbb{E}\,[V(Y_{t,Y(t)}(t+h))|\mathcal{F}_t]\right]\right)^{1/2} \leq \left(\mathbb{E}\left[\,\mathbb{E}\,[\exp(2K_S^Y h)V(Y(t))|\mathcal{F}_t]\right]\right)^{1/2} \\ &= \left(\mathbb{E}\left[\exp(2K_S^Y h)V(Y_{0,y_0}(t))\right]\right)^{1/2} \leq \left(\mathbb{E}\left[\exp(2K_S^Y (t+h))V(y_0)\right]\right)^{1/2},\end{aligned}$$

hence $y(t) := ||Y||_{H_2([0,t],\mu,H)} = \int_0^T \mathbb{E}\,||Y_u||_H^2 d\mu(u)$ satisfies

$$\begin{aligned}y(t) &\leq \int_0^t \mathbb{E}\,V(Y_u)d\mu(u) \leq \int_0^t \exp(2K_S^Y u)d\mu(u)\;\mathbb{E}\,V(y_0) \\ &\leq \exp(2[K_S^Y]_+ t)\mu([0,t])\;\mathbb{E}\,V(y_0) \leq \exp(2[K_S^Y]_+ T)\mu([0,T])\;\mathbb{E}\,V(y_0)\end{aligned}$$

which trivially brings up the second statement of the above lemma. □

For $0 \leq t \leq T$, $x_0, y_0 \in H_2([0,t],\mu,H)$ where x_0, y_0 are $(\mathcal{F}_0, \mathcal{B}(H))$-measurable initial values, define the (**pointwise**) **global mean square error** $\varepsilon_2(t)$ by

$$\varepsilon_2(t) = \left(\mathbb{E}\,||X_{0,x_0}(t) - Y_{0,y_0}(t)||_H^2\right)^{1/2},$$

and the (**pointwise**) **global weak error** $\varepsilon_w(t)$

$$\varepsilon_w(t) := ||\,\mathbb{E}\,X_{0,x_0}(t) - \mathbb{E}\,Y_{0,y_0}(t)||_H.$$

Lemma 5.7. *Assume that the assumptions* $(A1), (A2), (A4)$ *and*

(A9) ***Weak contractivity of*** X*, i.e.* $\exists\, K_{WC}^X\; \forall X(t),\, Y(t) \in \mathbb{D}_t$ *(where* $X(t)$*,* $Y(t)$ *are* $(\mathcal{F}_t)$*-adapted)* $\forall t, h : 0 \leq h \leq \delta_0, 0 \leq t, t+h \leq T$ *(h deterministic)*

$$||\,\mathbb{E}\,X_{t,X(t)}(t+h) - \mathbb{E}\,X_{t,Y(t)}(t+h)||_H \leq ||\,\mathbb{E}\,X(t) - \mathbb{E}\,Y(t)||_H \exp\left(K_{WC}^X h\right)$$

are satisfied. Then, for all deterministic step sizes $0 \leq h \leq \delta_0 \leq 1$ *and for all* s, t *with* $0 \leq s \leq t \leq t+h \leq T$*, the* **global weak error** $\varepsilon_w(t)$ *satisfies*

$$\begin{aligned}\varepsilon_w(t+h) &= ||\,\mathbb{E}\,X_{0,x_0}(t+h) - \mathbb{E}\,Y_{0,y_0}(t+h)||_H \\ &\leq \exp(K_{WC}^X h)\varepsilon_w(t) + K_0^C \exp(K_S^Y t)(\mathbb{E}\,V(y_0))^{1/2} h^{r_0} \quad \textit{and} \\ \varepsilon_w(t) &\leq \exp(K_{WC}^X (t-s))\varepsilon_w(s) + \\ &\quad + \hat{K}_0 \frac{\exp(K_{WC}^X (t-s) + K_S^Y s) - \exp(K_S^Y t)}{K_{WC}^X - K_S^Y} \Delta_{max}^{r_0 - 1}\end{aligned}$$

where $\hat{K}_0 = K_0^C \exp([K_S^Y]_- \Delta_{max})(\mathbb{E}\,V(y_0))^{1/2}$*, i.e. the weak error* ε_w *is of global order* $r_0 - 1$ *(and, trivially, of local order* r_0*).*

Proof. For all $0 \leq t, t+h \leq T, 0 \leq h \leq \delta_0$, one finds

$$\begin{aligned} \varepsilon_w(t+h) &= ||\,\mathbb{E}\, X_{t,X(t)}(t+h) - \mathbb{E}\, Y_{t,Y(t)}(t+h)||_H \\ &\leq ||\,\mathbb{E}\, X_{t,X(t)}(t+h) - \mathbb{E}\, X_{t,Y(t)}(t+h)||_H + \\ &\quad +||\,\mathbb{E}\, X_{t,Y(t)}(t+h) - \mathbb{E}\, Y_{t,Y(t)}(t+h)||_H \\ &\leq ||\,\mathbb{E}\, X_{t,X(t)}(t+h) - \mathbb{E}\, X_{t,Y(t)}(t+h)||_H + \\ &\quad + \mathbb{E}\, ||\,\mathbb{E}\, [X_{t,Y(t)}(t+h)|Y(t)] - \mathbb{E}\, [Y_{t,Y(t)}(t+h)|Y(t)]||_H \\ &\leq \varepsilon_w(t)\exp(K^X_{WC}h) + K^C_0 \exp(K^Y_S t)(\mathbb{E}\, V(y_0))^{1/2}h^{r_0}. \end{aligned}$$

Now, use Lemma 5.4 to conclude the second statement. □

Remark. The result of Lemma 5.7 confirms a well-known hand-rule of deterministic-numerical analysis. By Lyapunov inequality (see [47]), we may trivially note that $\varepsilon_w(t) \leq \varepsilon_2(t)$ for all $0 \leq t \leq T$.

Lemma 5.8. *Assume that the assumptions (A1) – (A7) are satisfied. Then, for all t, h with $0 \leq h \leq \Delta \leq \delta_0$ (Δ deterministic) and $0 \leq t, t+h \leq T$, we have*

$$\mathbb{E}\, ||X_{t,Z(t)}(t+h) - Y_{t,Z(t)}(t+h)||^2_H \;\leq\; (K^C_2)^2\, \mathbb{E}\, V(Z(t))\Delta^{2r_2}$$

for any stochastic process $Z \in H_2([0,t], \mu, H)$ with $\mathbb{E}\, V(Z(t)) < +\infty$, i.e. the local mean square convergence rate $r_l \geq r_2$ can be established.

Proof. Suppose that $Z \in H_2([0,t], \mu, H)$. Then, by elementary laws of conditional expectations and using $(A5)$, for any $Z \in H_2([0,t], \mu, H)$, we have

$$\begin{aligned} \mathbb{E}\, ||X_{t,Z(t)}(t+h) &- Y_{t,Z(t)}(t+h)||^2_H \\ &= \mathbb{E}\left[\mathbb{E}\, [||X_{t,Z(t)}(t+h) - Y_{t,Z(t)}(t+h)||^2_H|Z(t)]\right] \\ &\leq \mathbb{E}\, [(K^C_2)^2 V(Z(t))h^{2r_2}] \;\leq\; (K^C_2)^2\, \mathbb{E}\, [V(Z(t))]\Delta^{2r_2}, \end{aligned}$$

as claimed for any $Z \in H_2([0,t], \mu, H)$ with $\mathbb{E}\, V(Z(t)) < +\infty$. □

Remark. Thanks to this Lemma 5.8, we know about the local convergence with worst case rate r_2 on $H_2([t, t+h], \mu, H)$. h could be chosen randomly as well. However, the requirements of a deterministic upper bound Δ on h and of deterministic rate r_2 are important ones.

A priori, but crude global mean square error estimate is found as follows.

Lemma 5.9. *Assume that the assumptions (A1) – (A8) and $r_2 \geq 1.0$ are satisfied. Then, for all $0 \leq s \leq t \leq T$ and deterministic step sizes (variable or constant) with $0 < \Delta_i \leq \min(t-s, \delta_0, 1)$, we have $\varepsilon_2(t) \leq$*

$$\exp(K^X_C(t-s))\varepsilon_2(s) + \hat{K}^C_2 \exp(K^Y_S t)\frac{\exp((K^X_C - K^Y_S)(t-s)) - 1}{K^X_C - K^Y_S}\Delta^{r_2-1}_{max},$$

where $\hat{K}^C_2 = K^C_2 \exp([K^Y_S]_- \Delta_{max})\left(\mathbb{E}\, V(y_0)\right)^{1/2}$, hence the global mean square error has at least the "worst case" convergence rate $r_2 - 1$. In particular, if $K^X_C = K^Y_S = 0$ then

$$\varepsilon_2(t) \;\leq\; \varepsilon_2(s) + K^C_2(t-s)(\mathbb{E}\, V(y_0))^{1/2}\Delta^{r_2-1}_{max}.$$

Moreover, if $K^X_C < 0, K^Y_S = 0$ then

$$\lim_{t\to+\infty} \varepsilon_2(t) \;\leq\; -\frac{K^C_2}{K^X_C}(\mathbb{E}\, V(y_0))^{1/2}\Delta^{r_2-1}_{max}\,.$$

Proof. Choose deterministic step sizes $0 < h \leq \Delta_i \leq \min(t-s,\delta_0,1)$. Using Minkowski's inequality, Lemmas 5.6 and 5.4, and elementary laws of conditional expectations, one concludes that

$$\begin{aligned}
\varepsilon_2(t + h) &\leq \left(\mathbb{E}\,||X_{t,X(t)}(t+h) - X_{t,Y(t)}(t+h)||_H^2\right)^{1/2} + \\
&\quad + \left(\mathbb{E}\,||X_{t,Y(t)}(t+h) - Y_{t,Y(t)}(t+h)||_H^2\right)^{1/2} \\
&\leq \exp(K_C^X h)\varepsilon_2(t) + \left(\mathbb{E}\,\mathbb{E}\left[||X_{t,Y(t)}(t+h) - Y_{t,Y(t)}(t+h)||_H^2 | Y(t)\right]\right)^{1/2} \\
&\leq \exp(K_C^X h)\varepsilon_2(t) + \left(\mathbb{E}\,\mathbb{E}\left[K_2^C V(Y(t)) h^{2r_2} | Y(t)\right]\right)^{1/2} \\
&\leq \exp(K_C^X h)\varepsilon_2(t) + K_2^C \exp(K_S^Y t)\,(\mathbb{E}\,V(y_0))^{1/2}\,\Delta_{max}^{r_2-1} h, \quad \text{hence}
\end{aligned}$$

$$\varepsilon_2(t) \leq$$
$$\exp(K_C^X(t-s))\varepsilon_2(s) + \hat{K}_2^C \exp(K_S^Y t) \frac{\exp((K_C^X - K_S^Y)(t-s)) - 1}{K_C^X - K_S^Y} \Delta_{max}^{r_2-1},$$

an estimate which immediately gives the claimed assertions. □

Remark. One may even get better global convergence rates than $r_2 - 1.0$ predicted by Lemma 5.9. The following lemma is needed to optimize the estimation of the global convergence rate in the stochastic case.

Lemma 5.10. *Assume that the assumptions (A1) – (A8) are satisfied. Then, for all* $t, h, \rho \in \mathbb{R} : 0 \leq h \leq \Delta \leq \delta_0$ *(*$\Delta, \rho \neq 0$ *deterministic,* $0 \leq t, t+h \leq T$*), we have* $u(t) :=$

$$\begin{aligned}
&\mathbb{E}\ < X_{t,X(t)}(t+h) - X_{t,Y(t)}(t+h), X_{t,Y(t)}(t+h) - Y_{t,Y(t)}(t+h) >_H \\
&\leq \varepsilon(t)[K_{SM}K_2^C + K_0^C]\exp(K_S^Y t)(\mathbb{E}\,V(y_0))^{1/2}\Delta^{r_2+r_{sm}} \\
&\leq \rho^2\Delta\varepsilon_2^2(t) + \frac{1}{2\rho^2}[(K_{SM}K_2^C)^2 + (K_0^C)^2]\exp(2K_S^Y t)\,\mathbb{E}\,V(y_0)\Delta^{2(r_2+r_{sm})-1}
\end{aligned}$$

for the stochastic processes $X, Y \in H_2([0,t+h],\mu,H)$ *.*

Proof. Suppose that $X, Y \in H_2([0,t+h],\mu,H)$. For $r = t+h$, define

$$\begin{aligned}
z(r) &= X_{t,X(t)}(r) - \mathbb{E}\,[X_{t,X(t)}(r)|\mathcal{F}_t] - (X_{t,Y(t)}(r) - \mathbb{E}\,[X_{t,Y(t)}(r)|\mathcal{F}_t]), \\
w(r) &= X_{t,Y(t)}(r) - Y_{t,Y(t)}(r)\,.
\end{aligned}$$

Then, by elementary calculation and properties of conditional expectations,

$$\begin{aligned}
|u(t)| &\leq |\mathbb{E}\ < z(t+h), w(t+h) >_H| + \\
&\quad + |\mathbb{E}\ < \mathbb{E}\,[X_{t,X(t)}(t+h)|\mathcal{F}_t] - \mathbb{E}\,[X_{t,Y(t)}(t+h)|\mathcal{F}_t], w(t+h) >_H| \\
&\leq (\mathbb{E}\,||z(t+h)||_H^2)^{1/2}(\mathbb{E}\,||w(t+h)||_H^2)^{1/2} + \\
&+ |\mathbb{E}(\mathbb{E}\left[< \mathbb{E}[X_{t,X(t)}(t+h)|\mathcal{F}_t] - \mathbb{E}[X_{t,Y(t)}(t+h)|\mathcal{F}_t], w(t+h) >_H |\mathcal{F}_t\right])| \\
&\leq \varepsilon_2(t)K_{SM}K_2^C(\mathbb{E}\,V(Y(t)))^{1/2}\Delta^{r_2+r_{sm}} + \\
&+ |\mathbb{E}(< \mathbb{E}[X_{t,X(t)}(t+h)|\mathcal{F}_t] - \mathbb{E}[X_{t,Y(t)}(t+h)|\mathcal{F}_t], \mathbb{E}\,[w(t+h)|\mathcal{F}_t] >_H)| \\
&\leq \varepsilon_2(t)[K_{SM}K_2^C + K_0^C]\exp(K_S^Y t)(\mathbb{E}\,V(y_0))^{1/2}\Delta^{r_2+r_{sm}} \\
&\leq \rho^2\Delta\varepsilon_2^2(t) + \frac{1}{2\rho^2}[(K_{SM}K_2^C)^2 + (K_0^C)^2]\exp(2K_S^Y t)(\mathbb{E}\,V(y_0))\Delta^{2(r_2+r_{sm})-1},
\end{aligned}$$

which gives the claimed assertions. □

5.2. The proof of Theorem 3.1. Fix the family $(\mathbb{D}_t)_{0\leq t\leq T}$ of invariant subsets $\mathbb{D}_t \subseteq H_2([0,t],\mu,H)$. Assume that the conditions $(A1)-(A8)$ are valid for $X, Y \in H_2([0,T],\mu,H)$ with corresponding local representations $X_{t,x}(t+h)$ and $Y_{t,y}(t+h)$ for any $x, y \in \mathbb{D}_0$, deterministic $h \leq \min(1, T-t, \Delta_{max}), t \in [0,T)$. Define

$$a := X_{t,X(t)}(t+h),\ b := Y_{t,Y(t)}(t+h),\ c := X_{t,Y(t)}(t+h).$$

An application of Lemma 5.5 gives

$$\varepsilon_2^2(t+h) = \mathbb{E}\,||a-b||_H^2 = \mathbb{E}\,||a-c||_H^2 + \mathbb{E}\,||c-b||_H^2 + 2\,\mathbb{E}\,<a-c, c-b>_H .$$

Therefore and thanks to Lemmas 5.6, 5.8, 5.10, we may conclude that

$$(4)\quad \begin{aligned}\varepsilon_2^2(t+h) \quad \leq \quad & \exp(2K_C^X h)\varepsilon_2^2(t) + (K_2^C)^2 \exp(2K_S^Y t)(\mathbb{E}\,V(y_0))h^{2r_2} + \\ & +2\varepsilon_2(t)[K_{SM}K_2^C + K_0^C]\exp(K_S^Y t)(\mathbb{E}\,V(y_0))^{1/2}h^{r_2+r_{sm}}.\end{aligned}$$

Now, take any real constant $\rho > 0$. Define $V_0^2 = \mathbb{E}\,V(y_0)$. Returning to (4), one arrives at

$$\begin{aligned}
&\varepsilon_2^2(t+h) \ \leq \\
&\leq \ \exp(2K_C^X h)(1+2\rho^2 h)\varepsilon_2^2(t) + \\
&+ \ \frac{(K_{SM}K_2^C)^2 + (K_0^C)^2 + (\rho K_2^C)^2}{\rho^2}\exp(2(K_S^Y t - K_C^X h))V_0^2 h^{2(r_2+r_{sm})-1} \\
&\leq \ \exp(2(K_C^X + \rho^2)h)\varepsilon_2^2(t) + \\
&+ \ \frac{(K_2^C)^2[\rho^2 + (K_{SM})^2] + (K_0^C)^2}{\rho^2}\exp(2(K_S^Y t - K_C^X h))V_0^2 h^{2(r_2+r_{sm})-1} \\
&\leq \ \exp(2(K_C^X + \rho^2)(t+h))\varepsilon_2^2(0) + \\
&\quad +K_I^2(\rho)\exp(2K_S^Y t)\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)t) - 1}{2(K_C^X + \rho^2 - K_S^Y)}\Delta_{max}^{2(r_2+r_{sm}-1)}
\end{aligned}$$

where $K_I(\rho)$ is given by (3), thanks to Lemma 5.4. Thus, by applying Lemma 5.1, we obtain

$$\begin{aligned}
&\varepsilon_2(t) \ \leq \\
&\quad \exp((K_C^X + \rho^2)t)\varepsilon_2(0) + K_I(\rho)\exp(K_S^Y t)\sqrt{\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)t) - 1}{2(K_C^X + \rho^2 - K_S^Y)}}\Delta_{max}^{r_g}
\end{aligned}$$

with "worst case" global rate $r_g \geq r_2 + r_{sm} - 1$ of mean square convergence for any $0 \leq t \leq T$ – an estimate which is particularly useful if $K_C^X + \rho^2 < 0$. Now, one may use the obtained pointwise error estimates to control the H_2-error as follows

$$\begin{aligned}
||X-Y||_{H_2([0,T],\mu,H)} \ &= \ \left(\int_0^T \mathbb{E}\,||X_t - Y_t||_H^2\, d\mu(t)\right)^{1/2} \\
(5) \qquad \leq \ & \varepsilon_2(0)\left(\int_0^T \exp(2(K_C^X + \rho^2)t)\, d\mu(t)\right)^{1/2} + \\
& +K_I(\rho)\left(\int_0^T \frac{\exp(2(K_C^X + \rho^2)t) - \exp(2K_S^Y t)}{2(K_C^X + \rho^2 - K_S^Y)}\, d\mu(t)\right)^{1/2}\Delta_{max}^{r_2+r_{sm}-1}
\end{aligned}$$

which gives the claimed rate of H_2-convergence when $\varepsilon_2(0) \leq K_{init}\Delta_{max}^{r_g}$ (Recall that μ is positive and σ-finite.). Thus, the proof is complete. □

5.3. The proof of Theorem 4.1. Suppose that $K_C^X < 0$. The proof goes similarly to that for Theorem 3.1. Take any constant parameter $\rho > 0$ satisfying $0 < \rho^2 \le |K_C^X|$. For abbreviation, define $V_0^2 = \mathbb{E}\, V(y_0)$. Returning to (4) we get

$$\begin{aligned}&\varepsilon_2^2(t+h)\\ &\quad\le\ \exp(2(K_C^X+\rho^2)h)\varepsilon_2^2(t)+\\ &\quad\quad+\frac{1}{\rho^2}[(K_2^C)^2(\rho^2+(K_{SM})^2)+(K_0^C)^2]\exp(2|K_C^X|h+2K_S^Y t)V_0^2 h^{2(r_2+r_{sm})-1}\end{aligned}$$

Applying Lemma 5.4 with (3) to the latter inequality yields

$$\begin{aligned}\varepsilon_2^2(t)\ &\le\ \exp(2(K_C^X+\rho^2)(t-s))\varepsilon_2^2(s)+\\ &+\ K_I^2(\rho)\exp(2K_S^Y t)\frac{\exp(2(K_C^X+\rho^2-K_S^Y)(t-s))-1}{2(K_C^X+\rho^2-K_S^Y)}\Delta_{max}^{2(r_2+r_{sm}-1)}.\end{aligned}$$

By taking the square root in this inequality, thanks to Lemma 5.1, we obtain the global estimate for $\varepsilon_2(t)$ as stated in (1), and, in particular

$$\begin{aligned}\varepsilon_2(t)\ &\le\ \varepsilon_2(0)\exp((K_C^X+\rho^2)t)+\\ &\quad+K_I(\rho)\sqrt{\frac{\exp(2(K_C^X+\rho^2)t)-\exp(2K_S^Y t)}{2(K_C^X+\rho^2-K_S^Y)}}\Delta_{max}^{r_2+r_{sm}-1.0} &(6)\\ &\le\ \varepsilon_2(0)\exp((K_C^X+\rho^2)t)+\frac{K_I(\rho)}{\sqrt{2|K_C^X+\rho^2-K_S^Y|}}\Delta_{max}^{r_2+r_{sm}-1.0} &(7)\\ &\le\ \varepsilon_2(0)+\frac{K_I(\rho)}{\sqrt{2|K_C^X+\rho^2-K_S^Y|}}\Delta_{max}^{r_2+r_{sm}-1.0}.\end{aligned}$$

It remains to evaluate this result. Recall that $K_C^X + \rho^2 \le 0$. Taking the supremum over all times t in the right-hand side of inequality (7) gives the estimate (1). Suppose that $K_C^X + \rho^2 < 0$. Then, taking the limit as $t \to +\infty$ in (6) confirms the estimates (2) and (3). Finally, the H_2-error is estimated similarly to the proof of Theorem 3.1 with $r_g = r_2 + r_{sm} - 1.0$. Thus, after returning to (5),

$$||X-Y||_{H_2([0,T],\mu,H)}\ \le\ \left(K_{init}+\frac{K_I(\rho)}{\sqrt{|K_C^X+\rho^2-K_S^Y|}}\right)\Big(\mu([0,+\infty])\Big)^{1/2}\Delta_{max}^{r_g}$$

which yields the claimed H_2-convergence, provided that $\varepsilon_2(0) \le K_{init}\Delta_{max}^{r_g}$, $K_S^Y \le 0$, $K_C^X+\rho^2 \le 0$ and the measure $\mu([0,+\infty]) < +\infty$. Thus, this conclusion completes the proof. □

5.4. A corollary to Theorem 4.1. A slight modification of the assumptions in Theorem 4.1 and its proof leads to an extension to the case of σ-finite measures μ as a by-product.

Corollary 5.1. *Assume that the conditions* $(A1) - (A8)$ *are satisfied on the time-interval* $[0, +\infty)$ *with σ-finite, positive measure* $\mu([0, +\infty))$,

$$\int_0^{+\infty}\Big(\exp(2(K_C^X+\rho^2)t)\ \ +\ \ \frac{\exp(2(K_C^X+\rho^2)t)-\exp(2K_S^Y t)}{2(K_C^X+\rho^2-K_S^Y)}\Big)\,d\mu(t)\ <\ +\infty$$

$$\mathbb{E}\,||X_0-Y_0||_H^2\ \le\ K_{init}\Delta_{max}^{r_g}, \tag{8}$$

for a real constant $\rho > 0$, and all constants K occurring in $(A1)$ – $(A7)$ *do not depend on the terminal times $T > 0$.*

Then the stochastic processes $X, Y \in H_2([0,+\infty),\mu,H)$ converge to each another with respect to the naturally induced metric

$$m(X,Y) \;=\; (< X-Y, X-Y >_{H_2})^{1/2} \;=\; \left(\int_0^{+\infty} \mathbb{E}\,||X_t - Y_t||_H^2 \, d\mu(t)\right)^{1/2}$$

with "worst case" convergence rate $r_g = r_2 + r_{sm} - 1.0$.

Proof. One only needs to return to the H_2-error estimate (5). This yields

$$||X \;-\; Y||_{H_2([0,+\infty),\mu,H)} \leq K_{init}\left(\int_0^{+\infty} \exp(2(K_C^X+\rho^2)t)\, d\mu(t)\right)^{1/2} \Delta_{max}^{r_g} +$$
$$+ \;\; K_I(\rho)\left(\int_0^{+\infty} \frac{\exp(2(K_C^X+\rho^2)t) - \exp(2K_S^Y t)}{2(K_C^X+\rho^2-K_S^Y)}\, d\mu(t)\right)^{1/2} \Delta_{max}^{r_g}.$$

The occurring improper integrals on the right hand side are finite thanks to the integrability property (8) of μ, and hence this fact implies the convergence with respect to naturally induced norm $||.||_{H_2([0,+\infty),\mu,H)}$ with rate $r_g = r_2 + r_{sm} - 1.0$. □

6. Example of finite-dimensional SDEs and drift-implicit Euler methods

Take $H = \mathbb{R}^d$ with the Euclidean scalar product $< x, y >_d = \sum_{i=1}^n x_i y_i$ and norm $||x||_d = \sqrt{< x, x >_d}$.

6.1. General one-sided Lipschitz case. Consider the d-dimensional system of Itô SDEs

$$dX_t \;=\; b^0(t,X_t)\,dt \;+\; \sum_{j=1}^m b^j(t,X_t)\,dW_t^j \tag{1}$$

driven by m real-valued, independent Wiener processes W^j and understood in the sense of Itô [16] for the sake of simplicity. Then it is well-known that strong solutions uniquely exist under the following one-sided conditions (e.g. see Krylov [23]). There are real constants K_{OB}, K_{OL} such that, for all $t \in [0,T]$, for all $x, y \in \mathbb{R}^d$, we have

$$b^j (j = 0,1,...,m) \text{ ... Caratheodory functions,} \tag{2}$$

$$< x, b^0(t,x) >_d + \frac{1}{2}\sum_{j=1}^m ||b^j(t,x)||_d^2 \;\leq\; K_{OB}(1+||x||_d^2), \tag{3}$$

$$<x-y, b^0(t,x) - b^0(t,y)>_d + \frac{1}{2}\sum_{j=1}^m ||b^j(t,x) - b^j(t,y)||_d^2 \;\leq\; K_{OL}||x-y||_d^2, \tag{4}$$

$$\mathbb{E}\,||X_0||_d^2 \;<\; +\infty. \tag{5}$$

The existence of unique solution can be shown by the help of Lyapunov function $V(x) = 1 + ||x||_d^2$ and more precise estimates are even found when the terms $1+||x||_d^2$ at the right side of these conditions are replaced by $V(x) = ||x||_d^2$. Moreover, the solutions X are a.s. continuous and $X \in H_2([0,T],\mu,\mathbb{R}^d)$. In contrast to the analytical theory (cf. [3], [7], [8], [9], [10], [19], [26], [30], [32], [33], [48]), fairly less is known about the convergence rates of numerical approximations for systems (1) under these more general conditions (2) – (5), cf. [5], [20], [21], [22], [28], [29], [31], [50], [43], based on stochastic Taylor-expansions [51]. There are only three major results with constant step sizes in this direction, apart from our paper. As the first, the result of Hu [13] establishes convergence rates of the drift-implicit Euler method given by

$$(6) \qquad Y_{n+1} \; = \; Y_n + b^0(t_{n+1}, Y_{n+1})\Delta_n + \sum_{j=1}^{m} b^j(t_n, Y_n)\Delta W_n^j$$

towards the exact solution under (2) - (5), with step sizes $\Delta_n = t_{n+1} - t_n$ and Wiener process increments $\Delta W_n^j = W_{t_{n+1}}^j - W_{t_n}^j$ along any discretizations

$$0 = t_0 < t_1 < t_2 < ... < t_n < t_{n+1} < ... < t_{n_T} = T < +\infty$$

with $\Delta_n \leq 1$. Second, the result of Higham, Mao and Stuart [14] proves strong mean square convergence rate 0.5 of the split step Euler Backward method on a given finite interval $[0, T]$. Third, the main result of Gyöngy [12] shows the almost sure convergence of the Euler method using Lyapunov-functions with convergence rate 1/4. In contrast to most of the publications, we are able to treat the case of variable step sizes Δ_n and infinite integration intervals $[0, +\infty)$ too (moreover, even along functions such as $V(x) = c_0 + c_2||x||_d^2$). Only few results are known from the literature when T tends to $+\infty$. Presuming equidistant discretizations, Talay [49] (for second order methods) and Mattingly, Stuart and Higham [27] (for Euler-type methods) establish convergence rates of some numerical methods with respect to invariant measures (i.e. weak convergence). In Schurz [44], [43], [45] one finds general assertions for the convergence rates related to the L^2-error for discretizations of intervals $[0, +\infty)$ with variable step size. However, there are plenty of convergence results under the obviously stronger assumptions of global Lipschitz continuity and linear-polynomially boundedness of coefficients b^j. Most of those results report on convergence with rate $r_g = 0.5$ for constant step sizes $\Delta_n = \Delta_0$ on given nonrandom intervals $[0, T]$. For an overview, see [43] and [50]. The drift-implicit Euler method has the mean square convergence rate $\gamma_2 = 0.5$ under the classical global Lipschitz-continuity conditions. That is, for the cadlag approximation process $Y = (Y_t)_{0 \leq t \leq T}$ of stochastic process $X = (X_t)_{0 \leq t \leq T}$ constructed as a step function with jumps $Y_t = Y_n$ at nondecreasing times $t = t_n \in [0, T]$ based on schemes $(Y_n)_{n \in \mathbb{N}}$ as (6), there is a constant $K_g = K_g(T)$ such that

$$(7) \qquad \begin{aligned} ||X - Y||_{H_2} \;&=\; \left(\int_0^T \mathbb{E}\, ||X_t - Y_t||_d^2 \, d\mu(t) \right)^{1/2} \\ &\leq\; (\mu([0,T]))^{1/2} \sup_{0 \leq t \leq T} \left(\mathbb{E}\, ||X_t - Y_t||_d^2 \right)^{1/2} \\ &\leq\; K_g(1 + ||X_0||_{H_2}^2 + ||Y_0||_{H_2}^2)^{1/2} (\mu([0,T]))^{1/2} \Delta^{\gamma_2} \end{aligned}$$

with the maximum step size $\Delta = \max_{n=0,1,\ldots,n_T-1} \Delta_n \leq 1$. Using our Theorem 3.1 with $V(x) = 1 + ||x||_d^2$ one can establish the same mean square convergence rate $\gamma_2 = 0.5$ for systems (1) satisfying the more general conditions (2) – (5) and discretized by "admissible" nonequidistant grids. Using Theorem 4.1, the same convergence rate $\gamma_2 = 0.5$ is maintained even for drift-implicit Euler discretizations of mean square dissipative systems of SDEs with $K_C^X < 0$, $K_S^Y \leq 0$ on infinite intervals $[0, +\infty)$ with finite μ. To confirm these results, one simply checks the behavior of (X, Y) with respect to the assumptions $(A1)$ – $(A8)$ with $V(x) = 1 + ||x||_d^2$ and $V(x) = ||x||_d^2$, respectively, in an axiomatic manner, while choosing $\rho > 0$ such that $K_C^X + \rho^2 < 0$. The latter result of L^2-convergence on infinite time-intervals can also be concluded from Corollary 5.1 under the integrability condition (8) of measure μ (e.g. with Lebesgue-measure μ, $K_C^X + \rho^2 < 0$, $K_S^Y < 0$), a fact which is not known so far.

6.2. Discussion on a specific example from physics. The master example to decribe equilibrium critical phenomena occurring in the response of a superconductor to an external magnetic field in physics (field theory) is given by the one-dimensional Itô-interpreted nonlinear test SDE

$$dX_t \;=\; (cX_t - \gamma^2 X_t^3)\,dt \;+\; \sigma X_t\,dW_t \tag{8}$$

with real constants c, σ, γ. It represents the reaction part of a stochastic Ginzburg-Landau equation without diffusion and with multiplicative noise, e.g. see [4], [6], [15], [25]. This SDE possesses the one-step integral representation

$$X_{t,x}(t+h) = x + \int_t^{t+h} (cX_{t,x}(s) - \gamma^2[X_{t,x}(s)]^3)ds + \sigma \int_t^{t+h} X_{t,x}(s)dW_s \tag{9}$$

which can be numerically integrated by the **partial-implicit Euler method**

$$Y_{n+1} = Y_n + (cY_n - \gamma^2 Y_n^2 Y_{n+1})\Delta_n + \sigma Y_n \Delta W_n \tag{10}$$

where Y_n denotes the numerical approximation of X at instant t_n along partitions $(t_n)_{n\in\mathbb{N}}$ of intervals $[0, +\infty)$. This numerical method belongs to a class of nonstandard methods due to [39] and possesses the explicit representation

$$Y_{n+1} = Y_n \left(\frac{1 + c\Delta_n + \sigma \Delta W_n}{1 + \gamma^2 Y_n^2 \Delta_n} \right) \tag{11}$$

with one-step representation

$$Y_{t,x}(t+h) = x\left(\frac{1 + ch + \sigma(W_{t+h} - W_t)}{1 + \gamma^2 x^2 h}\right) = x + x\left(\frac{(c - \gamma^2 x^2)h + \sigma(W_{t+h} - W_t)}{1 + \gamma^2 x^2 h}\right). \tag{12}$$

The set of axioms (A1) - (A8) is satisfied under $\max\{2|c|h, h\} \leq 1$. This can be seen as follows. At first, we need an auxiliary discussion on the continuity of exact solution. For this discussion, while using Dynkin formula, consider

$$\begin{aligned}
\mathbb{E}\,[|X_{t,x}(s)|^{2n}] \;&=\; |x|^{2n} + 2n\int_t^{t+h} \mathbb{E}\,[(c+\frac{\sigma^2}{2}(2n-1)(X_{t,x}(s))^{2n} - \gamma^2 (X_{t,x}(s))^{2n+2}]ds \\
&\leq\; |x|^{2n} + 2n\int_t^{t+h} (c + \frac{\sigma^2}{2}(2n-1))\,\mathbb{E}\,[(X_{t,x}(s))^{2n}]ds.
\end{aligned}$$

A standard application of Gronwall-Bellman inequality leads to

$$\mathbb{E}\,[|X_{t,x}(s)|^{2n}] \;\leq\; |x|^{2n} \exp\Big(n(2c + (2n-1)\sigma^2)(s-t)\Big) \tag{13}$$

for $s \geq t$. From these estimates, the continuity of the exact solution of (8) in mean sense becomes apparent. Derive the estimates

$$\begin{aligned}
|\,\mathbb{E}\,[X_{t,x}(s) - x]| \;&=\; \mathbb{E}\,\left[\int_t^s \Big(cX_{t,x}(u) - \gamma^2[X_{t,x}(u)]^3\Big)ds\right] \\
&\leq\; \left[|c| \max_{t\leq u\leq s} \Big(\,\mathbb{E}\,|X_{t,x}(u)|^2\Big)^{1/2} + \gamma^2 \max_{t\leq u\leq s} \Big(\,\mathbb{E}\,|X_{t,x}(u)|^6\Big)^{1/2}\right](s-t) \\
&\leq\; \Big(|c||x| \exp((c+\frac{\sigma^2}{2})_+(s-t)) + \gamma^2 |x|^3 \exp(3(c+5\frac{\sigma^2}{2})_+(s-t))\Big)(s-t), \\
&\leq\; \sqrt{V_c(x)}(s-t)
\end{aligned}$$

with appropriate functional V_c as defined below in (14), i.e. the solution of nonlinear SDE (8) is local Lipschitz-continuous in time in the mean sense. Similarly, for

$|t-s|\leq 1$, one proves local Hölder continuity with exponent 0.5 in time in mean square sense such that

$$\begin{aligned}
\mathbb{E}\,[|X_{t,x}(s)-x|^2] &= \mathbb{E}\left[\Big|\int_t^s (cX_{t,x}(u)-\gamma^2[X_{t,x}(u)]^3)du+\int_t^s \sigma X_{t,x}(u)dW_u\Big|^2\right]\\
&\leq 3c^2(s-t)^2\max_{t\leq u\leq s}\,\mathbb{E}\,[|X_{t,x}(u)|^2]+3\gamma^2(s-t)^2\max_{t\leq u\leq s}\,\mathbb{E}\,[|X_{t,x}(u)|^6]+\\
&\quad +3\sigma^2(s-t)\max_{t\leq u\leq s}\,\mathbb{E}\,[|X_{t,x}(u)|^2]\\
&\leq 3(s-t)^2(c^2+\sigma^2)|x|^2\exp\Big((2c+\sigma^2)_+(s-t)\Big)+\\
&\quad +3\gamma^2(s-t)|x|^6\exp\Big(3(2c+5\sigma^2)_+(s-t)\Big)\\
&\leq V_c(x)(s-t)
\end{aligned}$$

where $|s-t|\leq 1$ and $V_c(x)$ can be chosen as the 6th order positive polynomial

$$(14)\quad V_c(x) = \left(\sqrt{3}(|c|+|\sigma|)|x|\exp\Big((c+\frac{\sigma^2}{2})_+\Big)+\sqrt{3}|\gamma||x|^3\exp\Big(3(c+5\frac{\sigma^2}{2})_+\Big)\right)^2$$

in x. So, we may suppose that $\mathbb{E}\,[V(X_{t,x}(s))]<+\infty$ is guaranteed, which is true whenever X has finite 6th initial moments.

Axiom $(A1)$. This is obviously guaranteed since both the processes X and Y live on the entire real line without any explosions, hence H can be chosen by $H=\mathbb{R}^1$.

Axiom $(A2)$. Take $V(x)$ as defined by (15) below. Recall that $2|c|h\leq 1$ and $h\leq 1$. Square the expression (10) on both sides, take the nth power and use moment properties of Gaussian distributions $\xi_h=W_{t+h}-W_t\in\mathcal{N}(0,h)$ in order to find that

$$\begin{aligned}
&\mathbb{E}\,[(Y_{t,Y(t)}(t+h))^{2n}|\mathcal{F}_t] = \mathbb{E}\,[(Y_{t,x}(t+h))^{2n}]|_{x=Y(t)}\\
&= x^{2n}\,\mathbb{E}\left[\left(\frac{1+ch+\sigma\xi_h}{1+\gamma^2x^2h}\right)^{2n}\right]\Bigg|_{x=Y(t)}\leq x^{2n}\left(\frac{(1+ch)^2+(2n-1)\sigma^2h}{(1+\gamma^2x^2h)^2}\right)^n\Bigg|_{x=Y(t)}\\
&\leq x^{2n}\left(\frac{(1+(c+c^2h+(2n-1)\sigma^2/2)_+h)^2}{(1+\gamma^2x^2h)^2}\right)^n\Bigg|_{x=Y(t)}\\
&\leq \exp(n(2c+c^2h+(2n-1)\sigma^2)_+h)(Y(t))^{2n};
\end{aligned}$$

hence

$$\mathbb{E}\,[V(Y_{t,Y(t)}(t+h))|\mathcal{F}_t]\leq V(Y(t))\exp(4(2c+c^2h+7\sigma^2)_+h),$$

where V is any polynomial of even powers up to 8th moment. Thus, the approximation Y is V-stable with $V(x)$ defined by (15) and stability constant

$$K_S^Y\leq 2(2c+c^2\Delta_{max}+7\sigma^2)_+.$$

Axiom $(A3)$. To see mean square contractivity, we apply Itô formula (i.e. Dynkin formula) and encounter

$$\begin{aligned}
&\mathbb{E}\,[||X_{t,x}(t+h)-X_{t,y}(t+h)||^2]\\
&= ||x-y||^2+\\
&\quad +2\,\mathbb{E}\int_t^{t+h}\Big(c(X_{t,x}(s)-X_{t,y}(s))-\gamma^2([X_{t,x}(s)]^3-[X_{t,y}(s)]^3)\Big)(X_{t,x}(s)-X_{t,y}(s))ds\\
&\leq ||x-y||^2+2c\,\mathbb{E}\int_t^{t+h}||X_{t,x}(s)-X_{t,y}(s)||^2ds.
\end{aligned}$$

using monotonicity of cubic powers z^3. Note that by mean value theorem we can estimate

$$-([X_{t,x}(s)]^3 - [X_{t,y}(s)]^3)(X_{t,x}(s) - X_{t,y}(s)) = -3\eta^2(X_{t,x}(s) - X_{t,y}(s))^2 \leq 0,$$

where η is an intermediate value between $X_{t,x}(s)$ and $X_{t,y}(s)$. Applying Gronwall-Bellman Lemma leads to

$$\mathbb{E}\,[||X_{t,x}(t+h) - X_{t,y}(t+h)||^2] \quad \leq \quad ||x-y||^2 \exp(2ch).$$

It remains to substitute $x = X(t)$ and $y = Y(t)$ to compute the related conditional expectations. Thus, the solution of (8) is obviously mean square contractive with constant $K_C^X = c$.

Axiom $(A4)$. Mean consistency can be seen as follows. Subtract (12) from (9) in order to obtain

$$\begin{aligned} &\mathbb{E}\,[X_{t,x}(s) - Y_{t,x}(s)] \\ &= \mathbb{E} \int_t^{t+h} \left(cX_{t,x}(s) - x\frac{c}{1+\gamma^2x^2h} - \gamma^2[X_{t,x}(s)]^3 + \gamma^2x^3\frac{1}{1+\gamma^2x^2h} \right) ds \\ &= c\int_t^{t+h} \mathbb{E}\,[X_{t,x}(s) - x]ds + \frac{c\gamma^2x^3h^2}{1+\gamma^2x^2h} + \int_t^{t+h} \gamma^2\mathbb{E}\,[x^3 - [X_{t,x}(s)]^3]ds - \frac{\gamma^4x^5h^2}{1+\gamma^2x^2h}. \end{aligned}$$

This implies that

$$|\,\mathbb{E}\,[X_{t,x}(s) - Y_{t,x}(s)]| \quad \leq \quad \begin{cases} |c|\int_t^{t+h}|\,\mathbb{E}\,[X_{t,x}(s) - x]|ds + \frac{|\gamma c|}{2}x^2h^{3/2} + \\ +3\int_t^{t+h}\gamma^2|\,\mathbb{E}\,[\eta_{t,x}^2(s)(X_{t,x}(s) - x)]|ds + \frac{|\gamma|^3}{2}x^4h^{3/2} \end{cases}$$

where $\eta_{t,x}(s)$ is an intermediate value between $X_{t,x}(s)$ and x. Note that η^4 as convex (concave upward) function of η can be estimated by

$$\eta_{t,x}^4(s) \leq \theta(X_{t,x}(s))^4 + (1-\theta)x^4$$

where $\theta \in [0,1]$, hence

$$\eta_{t,x}^4(s) \leq \frac{(X_{t,x}(s))^4 + x^4}{2}.$$

Thus, we arrive at

$$\begin{aligned} &|\,\mathbb{E}\,[X_{t,x}(s) - Y_{t,x}(s)]| \\ &\leq |c|\int_t^{t+h}|\,\mathbb{E}\,[X_{t,x}(s) - x]|ds + \frac{|\gamma c|}{2}x^2h^{3/2} + \\ &\quad + \frac{3}{\sqrt{2}}\gamma^2\int_t^{t+h}(\,\mathbb{E}\,[X_{t,x}^4(s) + x^4])^{1/2}(\,\mathbb{E}\,[(X_{t,x}(s) - x)^2])^{1/2}ds + \frac{|\gamma|^3}{2}x^4h^{3/2} \\ &\leq |c|\sqrt{V_c(x)}\int_t^{t+h}(s-t)ds + \frac{|\gamma c|}{2}x^2h^{3/2} + \\ &\quad + \frac{3\sqrt{2}}{2}\gamma^2\sqrt{V_c(x)}x^2(1 + \exp((2c+3\sigma^2)_+h))\int_t^{t+h}(s-t)ds + \frac{|\gamma|^3}{2}x^4h^{3/2} \\ &\leq \left(\sqrt{V_c(x)}[\frac{|c|}{2} + \frac{3\sqrt{2}}{4}\gamma^2x^2(1 + \exp((2c+3\sigma^2)_+h))] + \frac{|\gamma c|}{2}x^2 + \frac{|\gamma|^3}{2}x^4 \right) h^{3/2} \\ &\leq K_0^C\sqrt{V_1(x)}h^{3/2} \end{aligned}$$

where V_c is chosen as in (14), and V_1 is defined by

$$V_1(x) = \sqrt{V_c(x)}[|c| + \frac{3\sqrt{2}}{2}\gamma^2x^2(1 + \exp((2c+3\sigma^2)_+))] + |\gamma c|x^2 + |\gamma|^3x^4$$

and $K_0^C = 1/2$. Therefore, the local mean rate can be taken at least as $r_0 = 1.5$
Axiom $(A5)$. Mean square consistency is investigated as follows. While using Cauchy-Bunyakovskii-Schwarz inequality, Young's inequality, the fact that $h \leq 1$ and Itô isometry, consider the estimate

$$\begin{aligned}
&\mathbb{E}\,[|X_{t,x}(t+h) - Y_{t,x}(t+h)|^2] \\
&= \mathbb{E}\,\Bigg|\int_t^{t+h} \left(cX_{t,x}(s) - \gamma^2[X_{t,x}(s)]^3 - \frac{xc}{1+\gamma^2x^2h} + \gamma^2\frac{x^3}{1+\gamma^2x^2h}\right) ds + \\
&\quad + \sigma\int_t^{t+h}\left(X_{t,x}(s) - \frac{x}{1+\gamma^2x^2h}\right)dW_s\Bigg|^2 \\
&\leq 6c^2\,\mathbb{E}\left|\int_t^{t+h}(X_{t,x}(s)-x)ds\right|^2 + 6\gamma^2c^2x^4\left(\frac{\gamma x h^{1/2}}{1+\gamma^2x^2h}\right)^2 h^3 + \\
&\quad +6\gamma^2\,\mathbb{E}\left|\int_t^{t+h}3\eta_{t,x}^2(s)(X_{t,x}(s)-x)ds\right|^2 + 6\gamma^6x^8\left(\frac{\gamma x h^{1/2}}{1+\gamma^2x^2h}\right)^2 h^3 + \\
&\quad +6\sigma^2\,\mathbb{E}\left|\int_t^{t+h}(X_{t,x}(s)-x)dW_s\right|^2 + 6\sigma^2\gamma^2x^4\left(\frac{\gamma x h^{1/2}}{1+\gamma^2x^2h}\right)^2 h^2 \\
&\leq 6c^2h\int_t^{t+h}\mathbb{E}\,[|X_{t,x}(s)-x|^2]ds + \frac{3}{2}\gamma^2c^2x^4h^2 + \\
&\quad +54\gamma^2h\int_t^{t+h}\mathbb{E}\,[\eta_{t,x}^4(s)(X_{t,x}(s)-x)^2]ds + \frac{3}{2}\gamma^6x^8h^2 + \\
&\quad +6\sigma^2\int_t^{t+h}\mathbb{E}\,[(X_{t,x}(s)-x)^2]ds + \frac{3}{2}\gamma^2\sigma^2x^4h^2 \\
&\leq 3c^2h^2V_c(x) + \frac{3}{2}\gamma^2c^2x^4h^2 + 4\cdot 54\gamma^2h^2|x|^6\exp(3(2c+5\sigma^2)_+h)) + \\
&\quad +\frac{3}{2}\gamma^6x^8h^2 + 3\sigma^2h^2V_c(x) + \frac{3}{2}\gamma^2\sigma^2x^4h^2 \\
&\leq K_2^C V_2(x)h^2
\end{aligned}$$

with appropriately chosen $V_2(x)$ such as

$$V_2(x) = 2(c^2+\sigma^2)V_c(x) + \gamma^2(c^2+\sigma^2)x^4 + 136\gamma^2|x|^6\exp(3(2c+5\sigma^2)_+) + \gamma^6x^8$$

involving up to $8th$ moments and $K_2^C = \sqrt{3/2}$ while presuming $h \leq 1$. Hence, the local mean square rate $r_2 = 1.0$ is verified.
Axiom $(A6)$. Mean square Hölder continuity of martingale part of solution process X of (8) with exponent $r_{sm} = 0.5$ and constant K_{SM} is established by the estimate

$$\begin{aligned}
&\mathbb{E}\,|M_{t,x}(s) - M_{t,y}(s)|^2 = \mathbb{E}\left|\sigma\int_t^{t+h}X_{t,x}(s)dW_s - \sigma\int_t^{t+h}X_{t,y}(s)dW_s\right| \\
&= \sigma^2\,\mathbb{E}\left|\int_t^{t+h}(X_{t,x}(s)-X_{t,y}(s))dW_s\right|^2 = \sigma^2\int_t^{t+h}\mathbb{E}\,|X_{t,x}(s)-X_{t,y}(s)|^2ds \\
&\leq \sigma^2|x-y|^2\int_t^{t+h}\exp(2c(s-t))ds = \sigma^2\frac{\exp(2ch)-1}{2ch}|x-y|^2h
\end{aligned}$$

while exploiting the Itô isometry relation and mean square contractivity estimate from the verification of axiom $(A3)$. Hence, we may choose

$$r_{sm} = 0.5 \quad \text{and} \quad K_{SM} \le \sigma^2 \max_{-1\le z\le 1} \frac{\exp(z)-1}{z} < +\infty.$$

Axiom $(A7)$. The interplay of local rates is satisfied since

$$r_0 = 1.5 = r_2 + r_{sm} = 1.0 + 0.5.$$

Axiom $(A8)$. Take V as any 8th order polynomial of even powers dominating V_1 and V_2, i.e.

$$V(x) \ge \max\{V_1(x), V_2(x)\} \tag{15}$$

for $x \in \mathbb{R}^1$. For example, one may choose the 8th order polynomial

$$V(x) := \begin{cases} \max\{\frac{1}{2}, 2(c^2+\sigma^2)\}V_c(x) + c^2 \\ + \max\{\frac{9}{2}\gamma^2(1+\exp((2c+3\sigma^2)_+))^2, |\gamma c|\}x^2 \\ +\gamma^2 \max\{|\gamma|, c^2+\sigma^2\}x^4 + 136\gamma^2|x|^6\exp(3(2c+5\sigma^2)_+) + \gamma^6 x^8 \end{cases} \tag{16}$$

where $V_c(x)$ is the 6th order polynomial defined by (14). It remains to require the validity of moment condition

$$\mathbb{E}\,[V(X_0)] + \mathbb{E}\,[V(Y_0)] < +\infty$$

which is obviously equivalent to the condition $\mathbb{E}\,[|X_0|^8] + \mathbb{E}\,[|Y_0|^8] < +\infty$ while taking $V(x) = \max\{V_1(x), V_2(x)\}$ or V as in (16), and that the initial values X_0 and Y_0 are independent of the naturally induced σ-algebra $\sigma(W_t : t \ge 0)$. (Note that a slight modification of above estimates leads to the relaxed requirement of 6th or even only 4th order bounded initial moments. So we do not claim that we provide the most efficient possible estimates here.)

Consequently, all axioms are fulfilled, and we may apply our basic L^2-convergence result of Theorem 3.1 to conclude the global mean square rate $r_g = 0.5$ of efficiently implementable partial-implicit Euler method (10) applied to nonlinear equations (8) along any nonrandom partitions of $[0,T]$ with maximum step size Δ_{max} such that

$$\max\{\Delta_{max}, 2|c|\Delta_{max}\} \quad \le \quad 1. \tag{17}$$

Moreover, a careful look at V_1 and V_2 while requiring

$$2c + c^2\Delta_{max} + 5\sigma^2 \quad \le \quad 0 \tag{18}$$

brings up the global convergence rate $r_g = 0.5$ along nonrandom partitions with maximum step size Δ_{max} satisfying (17) due to Theorem 4.1. This is a striking result since most of the existing papers did not investigate the dependence of leading error growth constants $K(T)$ on growing T at all (e.g. compared to those in [21], [28]). We have also shown that the partial-implicit Euler method (10) can be used to approximate the long term dynamics of nonlinear SDE (8) and their characteristics such as moments of its invariant measures with convergence rate $r_g = 0.5$ under (18). To the best of our knowledge, such a detailed and thorough analysis of properties of numerical methods for nonlinear stochastic equations as (8) is not previously known. This example can be generalized to the class of m-dissipative nonlinear SDEs (see forthcoming papers of the author).

7. A summary by an approximation diagram

The results provide an axiomatic approach to the analysis of numerical methods for stochastic processes along partitions of given time-intervals $[0,T]$ or $[0,+\infty)$ with both constant or variable step sizes. We have considered the concept of mean square convergence presuming the knowledge on Lyapunov-type functionals to control numerical stability. Thus, we also gain conclusions for the convergence of related processes in probability by checking the requirements $(A1)$ - $(A8)$ (recall that mean square convergence implies convergence in probability).

The essentials of presented approximation principles can be summarized by the following *Adequateness Diagram of Stochastic-Numerical Approximation Theory* exhibiting the interplay between the key concepts of invariance, smoothness, stability, contractivity, consistency and convergence. More precisely, under the properties of

$\mathbb{D}$-invariance of X, Y w.r.t. H_2

and

Smoothness of Diffusive (Martingale) Parts of X, Y w.r.t. H_2

one may establish

Approximative Well-posedness: **Stability** of X **Contractivity** of Y	**Consistency** of (X,Y) $\Longleftrightarrow$ $r_0 \geq r_2 + r_{sm}$ $\Updownarrow$	Approximative Well-posedness: **Stability** of Y **Contractivity** of X

Local and **Global** H_2**-Convergence** of (X,Y)
with global rate $r_g = r_2 + r_{sm} - 1.0$

This diagram describes the main crossrelations between those concepts and the basic equivalence principle in the context of stochastic approximations. As naturally expected, the concept of consistency plays the central role. Contractivity and stability property can be exchanged simultaneously if consistency holds (due to the inherent symmetry of the given class of approximation problems). Convergence is extracted from the interplay of consistency, stability and contractivity.

We have clearly seen that the Kantorovič-Lax-Richtmeyer principle "Stability and Consistency imply Convergence" holds in some modified form in numerical analysis of well-posed stochastic problems on separable Hilbert spaces too. For well-posedness of stochastic approximation problems we additionally need to presume some kind of Smoothness of Martingale Parts. Thus, our main approximation principle says that "Invariance, Smoothness of Martingale Part, Stability, Contractivity and Consistency imply Convergence" on separable Hilbert spaces.

Further illustrations with potential applications to infinite-dimensional stochastic systems are necessary. However, this rather voluminous work is left to the future. Our remaining future goal is to make the mentioned main principles come alive in conjunction with SDEs / SPDEs and their numerical analysis allowing asymptotically sharp estimates and having a large range of potential applications to several types of stochastic-numerical approximation problems in mind. In conclusion, it is recommended that previously established L^2-convergence results of numerical methods for stochastic differential equations shall be reconsidered in an axiomatic way, such as in the present investigation, especially as the leading error growth coefficients $K(T)$ do not grow in time T for all SDEs.

Acknowledgments

The author likes to express his gratitude to Linda and Edward Allen for continuously motivating support. We are also thankful to the comments of two anonymous referees. This paper is dedicated to my family and to my first academic teacher Prof. Dr. Paul Heinz Müller emerited at Dresden's University of Technology (Germany) and known for the dictionary of stochastics [52].

References

[1] E. Allen and A. Arciniega, Shooting methods for numerical solution of stochastic boundary-value problems, Stochastic Anal. Appl. 22 (2004), no. 5, 1295-1314.
[2] M.J. Anabtawi, G.S. Ladde and S. Sathananthan, Convergence and stability analysis of large-scale parabolic systems under Markovian structural perturbations I, II, Int. J. Appl. Math. 2 (2000), no. 1, 57-85, 87-111.
[3] L. Arnold, Stochastic Differential Equations, John Wiley & Sons, Inc., New York, 1974.
[4] M. Barton-Smith, Global solution for a stochastic Ginzburg-Landau equation with multiplicative noise, Stochastic Anal. Appl. 22 (2004), no. 1, 1-18.
[5] N. Bouleau and D. Lépingle, Numerical Methods for Stochastic Processes, John Wiley & Sons, Inc., New York, 1993.
[6] S. Cerrai, Stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term, Probab. Theory Related Fields 125 (2003), no. 2, 271-304.
[7] E.B. Dynkin, Markov Processes I+II, Springer-Verlag, New York, 1965.
[8] A. Friedman, Stochastic Differential Equations and Applications 1 & 2, Academic Press, New York, 1975/1976.
[9] T.C. Gard, Introduction to Stochastic Differential Equations, Marcel Dekker, Basel, 1988.
[10] I.I. Gikhman and A.V. Skorochod, Stochastische Differentialgleichungen, Akademie-Verlag, Berlin, 1971.
[11] S.K. Godunov and V.S. Ryabenkii, Difference Schemes: An Introduction to the Underlying Theory, North-Holland, Amsterdam, 1987.
[12] I. Gyöngy, A note on Euler's approximations, Potential Anal. 8 (1998), no. 3, 205-216.
[13] Y.Z. Hu, *Semi-implicit Euler-Maruyama scheme for stiff stochastic equations.* In Stochastic Analysis and Related Topics V: The Silivri Workshop (Silivri, Norway, July 18-29, 1994) (Eds. Koerezlioglu, H. et al.), Progress in Probability 38, p. 183-302, Birkhäuser, Boston, 1996.
[14] 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 (2002) 1041-1063.
[15] W. Horsthemke and R. Lefever, Noise-Induced Transitions, Springer, Berlin, 1984.
[16] K. Itô, Stochastic integral, Proc. Imp. Acad. Tokyo 20 (1944) 519-524.
[17] L.V. Kantorovič, Functional analysis and applied mathematics (in Russian), Uspehi Matem. Nauk (N.S.) 6 (1948), no. 28, 89-185.
[18] I. Karatzas and S. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, New York, 1988.
[19] R.Z. Khas'minskiĭ, Stochastic Stability of Differential Equations, Sijthoff & Noordhoff, Alphen aan den Rijn, 1980.
[20] P.E. Kloeden, E. Platen and H. Schurz, The numerical solution of nonlinear stochastic dynamical systems: a brief introduction, Int. J. Bifur. Chaos Appl. Sci. Eng. 1 (1991), no. 2, 277-286.
[21] P.E. Kloeden and E. Platen, Numerical Solution of Stochastic Differential Equations (2nd Ed.), Springer-Verlag, Berlin, 1995.
[22] P.E. Kloeden, E. Platen and H. Schurz, Numerical Solution of SDEs through Computer Experiments (3rd Corrected Printing), Springer-Verlag, Berlin, 2003.
[23] N.V. Krylov, Introduction to the Theory of Diffusion Processes, AMS, Providence, 1995.
[24] P.D. Lax and R.D. Richtmeyer, Survey of the stability of linear finite difference equations, Comm. Pure Appl. Math. 9 (1956) 267-293.
[25] S.K. Ma, Modern Theory of Critical Phenomena, Benjamin, Reading, MA, 1976.
[26] X. Mao, Stochastic Differential Equations & Applications, Horwood Publishing, Chichester, 1997.
[27] J.C. Mattingly, A.M. Stuart and D.J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, Stochastic Process. Appl. 101 (2002), no. 2, 185-232.

[28] G.N. Mil'shtein, Numerical Integration of Stochastic Differential Equations, Kluwer, Dordrecht, 1995.

[29] G.N. Mil'shtein, E. Platen and H. Schurz, Balanced implicit methods for stiff stochastic systems. SIAM J. Numer. Anal. 35 (1998), no. 3, 1010-1019.

[30] B. Øksendal, Stochastic Differential Equations, Springer-Verlag, New York, 1985.

[31] E. Pardoux and D. Talay, Discretization and simulation of stochastic differential equations, Acta Applicandae Math. 3 (1985) 23-47.

[32] P. Protter, P. Stochastic Integration and Differential Equations, Springer-Verlag, New York, 1990.

[33] D. Revuz and M. Yor, Continuous Martingales and Brownian Motion (2nd Ed.), Springer-Verlag, New York, 1994.

[34] R.D. Richtmeyer and K.W. Morton, Difference Methods for Initial-Value Problems (2nd Ed.), Interscience, New York, 1967.

[35] H. Schurz, Asymptotical mean square stability of an equilibrium point of some linear numerical solutions with multiplicative noise. Stochastic Anal. Appl. 14 (1996), no. 3, 313-354.

[36] H. Schurz, Numerical regularization for SDEs: construction of nonnegative solutions, Dynam. Systems Appl. 5 (1996), no. 1, 323-352.

[37] H. Schurz, Modelling and analysis of stochastic innovation diffusion. Z. Angew. Math. Mech. 76 (1996), Suppl. 3, I–XV, 366-369.

[38] H. Schurz, Stability, Stationarity, and Boundedness of Some Implicit Numerical Methods for Stochastic Differential Equations and Applications, Logos-Verlag, Berlin, 1997.

[39] H. Schurz, Partial- and linear-implicit numerical methods for nonlinear stochastic differential equations, Unpublished Manuscript, p. 1-29, Universidad de Los Andes, Bogotá, 1998.

[40] H. Schurz, Preservation of asymptotical laws through Euler methods for Ornstein-Uhlenbeck process, Stochastic Anal. Appl. 17 (1999), no. 3, 463-486.

[41] H. Schurz, The invariance of asymptotic laws of linear stochastic systems under discretization, Z. Angew. Math. Mech. 79 (1999), no. 6, 375-382.

[42] H. Schurz, Moment attractivity, stability and contractivity exponents of stochastic dynamical systems, Discrete Cont. Dynam. Syst. 7 (2001), no. 3, 487-515.

[43] H. Schurz, *Numerical analysis of SDEs without tears.* In Handbook of Stochastic Analysis and Applications (Eds. D. Kannan and V. Lakshmikantham), p. 237-359, Marcel Dekker, Basel, 2002.

[44] H. Schurz, *Two general theorems for the numerical approximation of stochastic processes on separable Hilbert spaces.* In Proceedings of Neural, Parallel, and Scientific Computations (Eds. M.P. Bekakos, G.S. Ladde, N.G. Medhin, M. Sambandham), Vol. 2, p. 55-58, Dynamic Publishers, Atlanta, 2002.

[45] H. Schurz, General theorems for the numerical approximation of stochastic processes on the Hilbert space $H_2([0,T],\mu,\mathbb{R}^d)$, Electron. Trans. Numer. Anal. 16 (2003) 50-69.

[46] H. Schurz, Stability of numerical methods for ordinary SDEs along Lyapunov-type and other functions with variable step sizes, Electron. Trans. Numer. Anal. 20 (2005) 27-49.

[47] A.N. Shiryaev, Probability (2nd Ed.), Springer-Verlag, Berlin, 1996.

[48] D.W. Stroock and S.R.S. Varadhan, Multidimensional Diffusion Processes, Springer-Verlag, New York, 1982.

[49] D. Talay, Second-order discretization schemes for stochastic differential systems for the computation of the invariant law, Stoch. Stoch. Reports 29 (1990) 13-36.

[50] D. Talay, *Simulation of stochastic differential systems.* In Probabilistic Methods in Applied Physics (Eds. P. Krée and W. Wedig), Springer Lecture Notes in Physics, Vol. 451, p. 54-96, Springer-Verlag, Berlin, 1995.

[51] W. Wagner and E. Platen, Approximation of Itô integral equations, *February Report*, p. 1-27, Academy of Sciences of GDR, Berlin, 1978.

[52] P.H. Müller, Dictionary of Stochastics, Probability Theory and Mathematical Statistics (in German, 5th Ed.), Akademie-Verlag, Berlin, 1991.

Department of Mathematics, Southern Illinois University, 1245 Lincoln Drive, Carbondale, IL 62901-4408, USA

E-mail: hschurz@math.siu.edu

URL: http://www.math.siu.edu/schurz/personal.html