Electronic Transactions on Numerical Analysis.
Volume 16, pp. 50-69, 2003.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

GENERAL THEOREMS FOR NUMERICAL APPROXIMATION OF STOCHASTIC PROCESSES ON THE HILBERT SPACE $H_2([0,T],\mu,\mathbb{R}^d)$*

HENRI SCHURZ†

Abstract. General theorems for the numerical approximation on the separable Hilbert space $H_2([0,T],\mu,\mathbb{R}^d)$ of cadlag, $(\mathcal{F}_t)$-adapted stochastic processes with μ-integrable second moments is presented for nonrandom intervals $[0,T]$ and positive measure μ. The use of the theorems is illustrated by the special case of systems of ordinary stochastic differential equations (SDEs) and their numerical approximation given by the drift-implicit Euler method under one-sided Lipschitz-type conditions.

Key words. stochastic-numerical approximation, stochastic Lax-Theorem, ordinary stochastic differential equations, numerical methods, drift-implicit Euler methods, balanced implicit methods.

AMS subject classifications. 65C20, 65C30, 65C50, 60H10, 37H10, 34F05.

1. Introduction and examples.

In deterministic numerical analysis of well-posed differential equations there is the well-known equivalence principle of Lax–Richtmeyer, e.g. see Godunov and Ryabenkii [5] or Strikwerda [33]. In general deterministic approximation principles combine the key concepts of consistency (a local property of the accuracy of approximations), stability (control on the error propagation in the approximation process) and that of convergence (global property of the accuracy of approximations on fixed time-intervals). One of the key statements (sometimes called the Lax principle) says that stability and consistency imply convergence of approximations to the exact solution for well-posed initial value problems in appropriate vector spaces.

In this paper we shall state and prove correspondingly generalized stochastic versions of the forementioned approximation principle on the Hilbert space $H_2([0,T],\mu,\mathbb{R}^d)$ of certain stochastic processes while avoiding a direct discussion of the notion of well-posedness of stochastic problems. There is a tight relation to the originally suggested principle for numerical analysis of elliptic equations in deterministic Hilbert spaces by Lax [17], see also Richtmeyer and Morton [20], and obviously tracing back to a general construct of Kantorovič [11, 12] according to Godunov and Ryabenkii [5] (p. 476). However, we will have to take into account the pecularity of stochastic processes. The main result can be applied to the numerical approximation of systems of SDEs with one-sided Lipschitz-continuous coefficients and leads to nontrivial results.

General notations: Let $< x,y >_d = \sum_{i=1}^d x_i y_i$ be the Euclidean scalar product for $x,y \in \mathbb{R}^d$, and $||x||_d = \sqrt{< x,x >_d}$ as the related Euclidean norm. We shall only consider $\mathbb{R}^d$-valued stochastic processes which are denoted as $X = (X_t)_{0\le t\le T}$ or $Y = (Y_t)_{0\le t\le T}$ and which are $(\mathcal{F}_t)$-adapted stochastic processes on the stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{0\le t\le T},\mathbb{P})$ which is completed with respect to all $\mathbb{P}$-null sets. $W^j = (W^j_t)_{0\le t\le T}$ are independent, real-valued standard Wiener processes on the completed stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{0\le t\le T},\mathbb{P})$. To underline the dynamic evolution depending on the initial time t and initial states $x,y\in\mathbb{R}^d$, we shall use the representations $X_{s,x}(t), Y_{s,y}(t)$ referring to the process values of X and Y starting at state x,y at time s and evaluated at time t, respectively. Let $\mathbb{E}$ denote the expectation operator with respect to the given probability measure $\mathbb{P}$. Given a strictly positive, σ-additive, deterministic measure μ on $([0,T],\mathcal{B}([0,T]))$,

*Received March 12, 2002. Accepted for publication January 31, 2003. Recommended by Richard Varga.

† Department of Mathematics, Southern Illinois University, Mailcode 4408, Carbondale, IL 62901, USA. E-mail: hschurz@math.siu.edu. This work was supported by University of Minnesota and Southern Illinois University.

ETNA
Kent State University
etna@mcs.kent.edu

where $\mathcal{B}(S)$ denotes the σ-field of Borel-measurable subsets of inscribed set S, then a numerical approximation theorem is established on closed image sets $\mathbb{D}_t \subseteq \mathbb{R}^d$ for the space $H_2 = H_2([0,T], \mu, \mathbb{R}^d)$ defined by

$$H_2([0,T],\mu,\mathbb{R}^d) := \left\{ X = (X_t)_{0\le t\le T} : \begin{array}{l} X_t \in \mathbb{R}^d \text{ is } (\mathcal{F}_t, \mathcal{B}(\mathbb{R}^d)) - \text{measurable}, \\ X \text{ cadlag process on } (\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P}), \\ \int_0^T \mathbb{E} < X_t, X_t >_d \, d\mu(t) < +\infty \end{array} \right\}$$

with real numbers as its scalars. Here $[A]_+$ represents the nonnegative part of the inscribed expression A, and $[A]_-$ the negative part such that $A = [A]_+ - [A]_-$. Furthermore, the often occurring letters K denote several real, deterministic constants.

PROPOSITION 1.1. *The space $H_2([0,T],\mu,\mathbb{R}^d)$ is a Hilbert space equipped with the scalar product $< X, Y >_{H_2} = \int_0^T \mathbb{E} < X_t, Y_t >_d \, d\mu(t)$ for a fixed, deterministic, finite, strictly positive, σ-additive measure μ on $([0,T], \mathcal{B}([0,T]))$.*

The proof is left to the reader as a relatively simple exercise, using ideas as in [32].

1.1. The example of SDEs and drift-implicit Euler method. Let $X = (X_t)_{0\le t\le T}$ satisfy the ordinary $\mathbb{R}^d$-valued stochastic differential equation (SDE)

$$dX_t = b^0(t, X_t)\, dt + \sum_{j=1}^{m} b^j(t, X_t)\, dW_t^j \tag{1.1}$$

driven by m real-valued Wiener processes W^j and understood in the sense of Itô [8] for the sake of simplicity. Then it is well-known that strong solutions uniquely exist under the following conditions. 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

$$b^j (j = 0, 1, ..., m) \text{ ... Caratheodory functions} \tag{1.2}$$

$$< x, b^0(t,x) >_d + \frac{1}{2}\sum_{j=1}^{m} ||b^j(t,x)||_d^2 \le K_{OB}(1 + ||x||_d^2) \tag{1.3}$$

$$< x - y, b^0(t,x) - b^0(t,y) >_d + \frac{1}{2}\sum_{j=1}^{m} ||b^j(t,x) - b^j(t,y)||_d^2 \le K_{OL}||x - y||_d^2 \tag{1.4}$$

$$\mathbb{E}\, ||X_0||_d^2 < +\infty. \tag{1.5}$$

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, fairly less is known about the convergence rates of numerical approximations for systems (1.1) under these conditions (1.2) – (1.5). There are two major results with constant step sizes, apart from our studies [31]. As the first, the result of Hu [7] establishes mean square convergence rates of the drift-implicit Euler method given by

$$Y_{n+1} = Y_n + b^0(t_{n+1}, Y_{n+1})\Delta_n + \sum_{j=1}^{m} b^j(t_n, Y_n)\Delta W_n^j \tag{1.6}$$

towards the exact solution under conditions (1.2) - (1.5), with the 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

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

The result of Higham, Mao and Stuart [6] additionally proves strong mean square convergence rate 0.5 of the split step Euler Backward method on a given finite interval $[0,T]$. Nothing is known from the literature when T tends to $+\infty$ to our knowledge. However, there

ETNA
Kent State University
etna@mcs.kent.edu

are plenty of convergence results under the obviously stronger assumptions of global Lipschitz continuity and linear-polynomially boundedness of the coefficients b^j. For example, see Kloeden, Platen and Schurz [14], Milstein [19], Burrage and Burrage [4], Talay [34] or Schurz [27]. 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 $Y(t_n) = Y_n$ at nondecreasing times $t_n \in [0,T]$ based on the scheme $(Y_n)_{n\in\mathbb{N}}$ as in (1.6), there is a constant $K_g = K_g(T)$ such that

$$(1.7)\quad ||X-Y||_{H_2} = \left(\int_0^T \mathbb{E}\,||X_t - Y_t||_d^2 d\mu(t)\right)^{1/2} \leq (\mu([0,T]))^{1/2} \sup_{0\leq t\leq T} (\mathbb{E}\,||X_t - Y_t||_d^2)^{1/2}$$
$$\leq K_g(1 + ||X_0||_{H_2}^2 + ||Y_0||_{H_2}^2)^{1/2}(\mu([0,T]))^{1/2}\Delta^{\gamma_2}$$

with the maximum step size $\Delta = \max_{n=0,1,\ldots,n_T-1} \Delta_n \leq 1$. Using our main theorem from below one can establish the same mean square convergence rate $\gamma_2 = 0.5$ for systems (1.1) satisfying the more general conditions (1.2) – (1.5), even for dissipative systems of SDEs on infinite intervals $[0,+\infty)$. One only needs to check the behavior of (X,Y) with respect to the assumptions stated below in an axiomatic manner.

1.2. The main assumptions of the approximation principle. In this paper we consider these main assumptions. Let $\mathbb{D}_t$ be (connected) deterministic subregions of $\mathbb{R}^d$ and K denote several real constants. In particular, K_C^X and K_S^Y may be negative or positive. There are real constants $\delta_0 \leq 1, r_0, r_{sm}, r_2 > 0$ such that we have

(A1) Strong $(\mathbb{D}_t)$-invariance of X, Y, i.e., for $Z = X$ and $Z = Y$ and $\forall s : 0 \leq s < T$

$$\mathbb{P}\left\{Z_t \in \mathbb{D}_t : s \leq t \leq T | Z_s \in \mathbb{D}_s\right\} = 1,$$

(A2) Stability of Y, i.e. $\exists K_S^Y\ \forall y \in \mathbb{D}_t\ \forall t, h : 0 \leq h \leq \delta_0, 0 \leq t, t+h \leq T$

$$\left(\mathbb{E}\left[1 + ||Y_{t,y}(t+h)||_d^2 | Y_t = y\right]\right)^{1/2} \leq \exp(K_S^Y h)\left(1 + ||y||_d^2\right)^{1/2},$$

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

$$\mathbb{E}\left[||X_{t,X(t)}(t+h) - X_{t,Y(t)}(t+h)||_d^2 | X(t), Y(t)\right] \leq \exp(2K_C^X h)||X(t) - Y(t)||_d^2,$$

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

$$||\mathbb{E}\left[X_{t,Z(t)}(t+h)|\mathcal{F}_t\right] - \mathbb{E}\left[Y_{t,Z(t)}(t+h)|\mathcal{F}_t\right]||_d \leq K_0^C(1 + ||Z(t)||_d^2)^{1/2}h^{r_0},$$

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

$$\left(\mathbb{E}\left[||X_{t,x}(t+h) - Y_{t,x}(t+h)||_d^2 | X_t = x, Y_t = x\right]\right)^{1/2} \leq K_2^C(1 + ||x||_d^2)^{1/2}h^{r_2}$$

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

$$\mathbb{E}\,||X_{t,X(t)}(t+h) - \mathbb{E}\left[X_{t,X(t)}(t+h)|\mathcal{F}_t\right] - \mathbb{E}\left[X_{t,Y(t)}(t+h)|\mathcal{F}_t\right] + X_{t,Y(t)}(t+h)||_d^2$$
$$\leq (K_{SM})^2\ \mathbb{E}\,||X(t) - Y(t)||_d^2\ h^{2r_{sm}},$$

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

ETNA
Kent State University
etna@mcs.kent.edu

1.3. Auxiliary lemmas. A series of auxiliary results is needed to prove the validity of a general approximation principle on the Hilbert space $H_2([0,T],\mu,\mathbb{R}^d)$. 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.

LEMMA 1.1. *Assume that $a_i \in \mathbb{R}^d$ (i=1,2,...,n). Then, for $n \in \mathbb{N}, p \geq 1$, we have*

$$||\sum_{i=1}^{n} a_i||_d^p \;\leq\; n^{p-1} \sum_{i=1}^{n} ||a_i||_d^p, \qquad \sqrt[p]{||\sum_{i=1}^{n} a_i||_d} \;\leq\; \sum_{i=1}^{n} \sqrt[p]{||a_i||_d}.$$

LEMMA 1.2. *Let $(v(n))_{n\in\mathbb{N}}$ be a sequence of nonnegative real numbers $v(n)$ satisfying*

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

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

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

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

LEMMA 1.3. *Let $v = v(t), -\infty < t_0 \leq t < +\infty$ be a nonnegative real–valued function which is 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 $C_I = C_I(t), C_H = C_H(t) \in L^1_{loc}([t_0,+\infty),\mathcal{B}([t_0,+\infty)),\mu)$ 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

$$v(t) \leq v(s) \;+\; \int_s^t C_I(u)\,du \;+\; \int_s^t C_H(u)\cdot v(u)\,du \quad \text{or}$$
$$v(t) \leq v(s)\cdot\exp\Big(\int_s^t C_H(u)\,du\Big) \;+\; \int_s^t C_I(u)\,du$$

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.*

$$v(t) \leq \Big(v(s) + \int_s^t C_I(u)\cdot\exp\Big(-\int_s^u C_H(z)\,dz\Big)\,du\Big)\cdot\exp\Big(\int_s^t C_H(u)\,du\Big) \tag{1.8}$$

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

Note: The latter three lemmas can be found in Schurz [21], and their use in [25] - [28]. The well-known Gronwall-Bellman lemma is included as the special case of positive, constant coefficients C_H, C_I in Lemma 1.3. As an immediate, but very helpful application of Lemma

1.3 one arrives at the following key lemma for the proof of main conclusions in the next section.

LEMMA 1.4. *Let $v = v(t), -\infty < t_0 \le t < +\infty$ be a nonnegative real–valued function which is absolutely Lebesgue-integrable on $[t_0, +\infty)$. Assume that $v(t)$ satisfies*

$$v(t+h) \le v(t)\,\exp(K_H h) \;+\; K_I \exp(K_S t) h^{r_{loc}} \tag{1.9}$$

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

$$v(t) \le v(s)\cdot\exp\Big(K_H(t-s)\Big) \;+\; \hat{K}_I \frac{\exp(K_S s + K_H(t-s)) - \exp(K_S t)}{K_H - K_S}\delta^{r_{loc}-1}$$

$$\le v(0)\cdot\exp\Big(K_H t\Big) \;+\; \hat{K}_I \frac{\exp(K_H t) - \exp(K_S t)}{K_H - K_S}\delta^{r_{loc}-1}. \tag{1.10}$$

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

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

holds. Then, the condition (1.9) reads as

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

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 the Lemma 1.3 since its assumptions are satisfied. For the sake of completion, we evaluate the inequality (1.8) 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 (1.8) is

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

LEMMA 1.5. *For all $a, b, c \in \mathbb{R}^d$, we have*

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

LEMMA 1.6. *Assume that the assumptions $(A1)$ – $(A2)$ are satisfied. Then, for all $0 \le t, t+h \le T$ with $(\mathcal{F}_t)$-adapted t, h, we have*

$$\begin{aligned} (\mathbb{E}\,[1 + ||Y_{0,y_0}(t+h)||_d^2])^{1/2} &\le \Big(\mathbb{E}\,\Big[\exp(2K_S^Y h)[1 + ||Y_{0,y_0}(t)||_d^2]\Big]\Big)^{1/2} \\ &\le \Big(\mathbb{E}\,\Big[\exp(2K_S^Y(t+h))[1 + ||y_0||_d^2]\Big]\Big)^{1/2}, \end{aligned}$$

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

$$\begin{aligned} (\mathbb{E}\,[1 + ||Y_{0,y_0}(t)||_d^2])^{1/2} &\le \exp(K_S^Y t)(1 + \mathbb{E}\,||y_0||_d^2)^{1/2} \quad \textit{and} \\ ||Y||_{H_2([0,t],\mu)} &\le \int_0^t \exp(K_S^Y u)d\mu(u)(1 + \mathbb{E}\,||y_0||_d^2)^{1/2} \\ &\le \exp([K_S^Y]_+ t)\mu([0,t])(1 + \mathbb{E}\,||y_0||_d^2)^{1/2} \\ &\le \exp([K_S^Y]_+ T)\mu([0,T])(1 + \mathbb{E}\,||y_0||_d^2)^{1/2}. \end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

Proof. Using elementary laws of conditional expectations leads to

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

which trivially brings up the second statement of the above lemma. $\diamond$

For $0 \le t \le T$, $x_0, y_0 \in H_2([0,t],\mu,\mathbb{R}^d)$ where x_0, y_0 are $(\mathcal{F}_0, \mathcal{B}(\mathbb{R}^d))$-measurable initial values, define the (pointwise) **global mean square error** $\varepsilon_2(t)$ and **global weak error** $\varepsilon_w(t)$, respectively, by

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

LEMMA 1.7. *Assume that the assumptions $(A1),(A2),(A4)$ and*

(A8) 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) $\forall t,h: 0 \le h \le \delta_0, 0 \le t, t+h \le T$ (h deterministic)

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

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

$$
\varepsilon_w(t+h) \le \exp(K_{WC}^X h)\varepsilon_w(t) + K_0^C \exp(K_S^Y t)(1+\mathbb{E}||y_0||_d^2)^{1/2} h^{r_0} \quad \textit{and}
$$

$$
\varepsilon_w(t) \le \exp(K_{WC}^X(t-s))\varepsilon_w(s) + \hat{K}_0 \frac{\exp(K_{WC}^X(t-s)+K_S^Y s) - \exp(K_S^Y t)}{K_{WC}^X - K_S^Y}\Delta_{max}^{r_0-1}
$$

where $\hat{K}_0 = K_0^C \exp([K_S^Y]_- \Delta_{max})(1+\mathbb{E}||y_0||_d^2)^{1/2}$, i.e. the weak error ε_w is of global order r_0-1 (and, trivially, of local order r_0). If $(A3)$ is valid instead of $(A8)$ then one may replace K_{WC}^X by K_C^X in the related mean square error estimates.

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

$$
\begin{aligned}
\varepsilon_w(t+h) \;&=\; ||\mathbb{E}X_{t,X(t)}(t+h) - \mathbb{E}Y_{t,Y(t)}(t+h)||_d\\
&\le ||\mathbb{E}X_{t,X(t)}(t+h) - \mathbb{E}X_{t,Y(t)}(t+h)||_d + ||\mathbb{E}X_{t,Y(t)}(t+h) - \mathbb{E}Y_{t,Y(t)}(t+h)||_d\\
&\le ||\mathbb{E}X_{t,X(t)}(t+h) - \mathbb{E}X_{t,Y(t)}(t+h)||_d +\\
&\quad +\mathbb{E}||\mathbb{E}[X_{t,Y(t)}(t+h)|Y(t)] - \mathbb{E}[Y_{t,Y(t)}(t+h)|Y(t)]||_d\\
&\le \varepsilon_w(t)\exp(K_{WC}^X h) + K_0^C \exp(K_S^Y t)(1+\mathbb{E}||y_0||_d^2)^{1/2} h^{r_0}.
\end{aligned}
$$

Now, use Lemma 1.4 to conclude the second statement. $\diamond$.

Note: By the Lyapunov inequality, we trivially note $\varepsilon_w(t) \le \varepsilon_2(t)$ for all $0 \le t \le T$.

LEMMA 1.8. *Assume that the assumptions $(A1)$ – $(A7)$ are satisfied. Then, for all t,h with $0 \le h \le \Delta \le \delta_0$ (Δ deterministic) and $0 \le t, t+h \le T$, we have*

$$
\mathbb{E}||X_{t,Z(t)}(t+h) - Y_{t,Z(t)}(t+h)||_d^2 \;\le\; (K_2^C)^2(1+\mathbb{E}||Z(t)||_d^2)\Delta^{2r_2}
$$

for any stochastic process $Z \in H_2([0,t],\mu,\mathbb{R}^d)$, i.e. the local mean square convergence rate $r_l \ge r_2$ can be established.

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

$$\begin{aligned}\mathbb{E}\,||X_{t,Z(t)}(t+h) - Y_{t,Z(t)}(t+h)||_d^2 &= \mathbb{E}\left[\mathbb{E}\,[||X_{t,Z(t)}(t+h) - Y_{t,Z(t)}(t+h)||_d^2|Z(t)]\right]\\ &\le \mathbb{E}\,[(K_2^C)^2(1+||Z(t)||_d^2)h^{2r_2}]\\ &\le \mathbb{E}\,[(K_2^C)^2(1+||Z(t)||_d^2)\Delta^{2r_2}] \;=\; (K_2^C)^2(1+\mathbb{E}\,||Z(t)||_d^2)\Delta^{2r_2}.\diamond\end{aligned}$$

Note: Therefore, we know about the local convergence with worst case rate r_2 on $H_2([t,t+h],\mu,\mathbb{R}^d)$. 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 1.9. *Assume that the assumptions $(A1)$ – $(A7)$ and $r_2 \ge 1.0$ are satisfied. Then, for all $0 \le s \le t \le T$ and deterministic step sizes with $0 < \Delta_i \le \min(t-s,\delta_0,1)$, we have*

$$\varepsilon_2(t) \le \exp(K_C^X(t-s))\varepsilon_2(s) + K_I K_2^C \exp(K_S^Y t)\frac{\exp((K_C^X - K_S^Y)(t-s)) - 1}{K_C^X - K_S^Y}\Delta_{max}^{r_2-1},$$

where $K_I = \exp([K_S^Y]_-\Delta_{max})\left(1+\mathbb{E}\,||y_0||_d^2\right)^{1/2}$, hence the mean square error has at least the "worst case" global rate r_2-1. In particular, if $K_{CS} = K_C^X = K_S^Y$ then

$$\varepsilon_2(t) \le \exp(K_{CS}(t-s))\varepsilon_2(s) + K_2^C(t-s)\exp(K_{CS}t)(1+\mathbb{E}\,||y_0||_d^2)^{1/2}\Delta_{max}^{r_2-1},$$

and if $K_C^X = K_S^Y = 0$ then

$$\varepsilon_2(t) \le \varepsilon_2(s) + K_2^C(t-s)(1+\mathbb{E}\,||y_0||_d^2)^{1/2}\Delta_{max}^{r_2-1}.$$

Moreover, if $K_C^X < 0, K_S^Y = 0$ then

$$\lim_{t\to+\infty}\varepsilon_2(t) \le -\frac{K_2^C}{K_C^X}(1+\mathbb{E}\,||y_0||_d^2)^{1/2}\Delta_{max}^{r_2-1}.$$

Proof. Choose deterministic step sizes $0 < h \le \Delta_i \le \min(t-s,\delta_0,1)$. Using Minkowski's inequality, Lemmas 1.6 and 1.4, and elementary laws of conditional expectations, one concludes that $\varepsilon_2(t+h) \le$

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

$$\varepsilon_2(t) \le \exp(K_C^X(t-s))\varepsilon_2(s) + K_I K_2^C\exp(K_S^Y t)\frac{\exp((K_C^X - K_S^Y)(t-s)) - 1}{K_C^X - K_S^Y}\Delta_{max}^{r_2-1}.$$

This implies the remaining assertions by taking the limit $K_C^X \to K_S^Y$ and $K_C^X = K_S^Y = 0$. $\diamond$

LEMMA 1.10. *Assume that the assumptions $(A1)$ – $(A7)$ are satisfied. Then, for all $t,h,\rho \in \mathbb{R} : 0 \le h \le \Delta \le \delta_0$ ($\Delta,\rho \ne 0$ deterministic, $0 \le t, t+h \le T$), we have*

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

for the stochastic processes $X, Y \in H_2([0, t+h], \mu, \mathbb{R}^d)$.

Proof. Suppose that $X, Y \in H_2([0, t+h], \mu, \mathbb{R}^d)$. For $r = t+h$, define the vectors

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

Then, by elementary calculation and properties of conditional expectations, one gets

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

2. Main results: Stochastic approximation theorems. The following fairly general approximation principle can be established. Define the pointwise error

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

for the processes $X, Y \in H_2([0, T], \mu, \mathbb{R}^d)$, 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 the approximation Y is based on.

2.1. A universal error estimate for the arbitrary case. The main result is devoted to the arbitrary case with no restrictions imposed on the sign of K_C^X, K_S^Y. In general one might refer to random step sizes Δ_i, but measurability forces us to confine to the adapted choice of step sizes. For simplicity, we confine ourselves to the case of nonrandom (i.e. deterministic, but also variable) step sizes Δ_i. Recall that the expression $(exp(\alpha t) - 1)/\alpha$ is replaced by the limit t if $\alpha = 0$.

THEOREM 2.1. *Assume that the conditions* $(A1) - (A7)$ *are satisfied.*
Then the stochastic processes $X, Y \in H_2([0, T], \mu, \mathbb{R}^d)$ *converge to each another with respect to the 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 nonrandom step sizes* Δ_i *(variable or constant) with* $0 < \Delta_i \le \Delta_{max} \le \delta_0$, *we have the universal error estimates*

$$\begin{aligned}
\varepsilon_2(t) \le\; & \exp((K_C^X + \rho^2)(t-s))\varepsilon_2(s) + \\
& +K_I(\rho) \exp(K_S^Y t) \sqrt{\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)(t-s)) - 1}{2(K_C^X + \rho^2 - K_S^Y)}}\; \Delta_{max}^{r_g}
\end{aligned} \tag{2.1}$$

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

$$\begin{aligned}
\sup_{0 \le t \le T} \varepsilon_2(t) \le\; & \exp([K_C^X + \rho^2]_+ T)\varepsilon_2(0) + \\
& +K_I(\rho) \exp([K_S^Y]_+ T) \sqrt{\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)T) - 1}{2(K_C^X + \rho^2 - K_S^Y)}}\; \Delta_{max}^{r_g}
\end{aligned} \tag{2.2}$$

ETNA
Kent State University
etna@mcs.kent.edu

with appropriate constant

$$(2.3)\qquad K_I(\rho) \; = \frac{1}{\rho}\Big((K_0^C)^2 + (K_2^C)^2[\rho^2 + (K_{SM})^2]\Big)^{1/2}(1 + \mathbb{E}\,||y_0||_d^2)^{1/2}\exp(([K_C^X]_- + [K_S^Y]_-)\Delta_{max}).$$

Proof. Fix the deterministic regions $\mathbb{D}_t \subset \mathbb{R}^d$. Assume that the conditions $(A1)$ – $(A6)$ are valid for $X, Y \in H_2([0,T],\mu,\mathbb{R}^d)$ with corresponding representations $X_{t,x}(t+h)$ and $Y_{t,y}(t+h)$ for any $x, y \in \mathbb{D}$, deterministic $h \le \min(1, T-t, \Delta)$, $t \in [0,T)$. 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 1.5 gives

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

Therefore from Lemmas 1.6, 1.8, 1.10, we may conclude that

$$(2.4)\quad \begin{aligned}\varepsilon_2^2(t+h) \le \exp(2K_C^X h)\varepsilon_2^2(t) + (K_2^C)^2\exp(2K_S^Y t)(1+\mathbb{E}\,||y_0||_d^2)h^{2r_2} + \\ +2\varepsilon_2(t)[K_{SM}K_2^C + K_0^C]\exp(K_S^Y t)(1+\mathbb{E}\,||y_0||_d^2)^{1/2}h^{r_2+r_{sm}}.\end{aligned}$$

Now, take any real constant $\rho > 0$. Returning to 2.4, one arrives at $\varepsilon_2^2(t+h) \;\le$

$$\begin{aligned}
&\le \exp(2K_C^X h)(1+2\rho^2 h)\varepsilon_2^2(t) + \\
&\quad + \frac{(K_{SM}K_2^C)^2 + (K_0^C)^2 + (\rho K_2^C)^2}{\rho^2}\exp(2(K_S^Y t + [K_C^X]_- h))(1+\mathbb{E}\,||y_0||_d^2)h^{2(r_2+r_{sm})-1} \\
&\le \exp(2(K_C^X + \rho^2)h)\varepsilon_2^2(t) + \\
&\quad + \frac{(K_2^C)^2[\rho^2 + (K_{SM})^2] + (K_0^C)^2}{\rho^2}\exp(2(K_S^Y t + [K_C^X]_-\Delta_{max}))(1+\mathbb{E}||y_0||_d^2)h^{2(r_2+r_{sm})-1} \\
&\le \exp(2(K_C^X + \rho^2)t)\varepsilon_2^2(0) + \\
&\quad + K_I^2(\rho)\exp(2K_S^Y t)\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)t) - 1}{2(K_C^X + \rho^2 - K_S^Y)}\Delta^{2(r_2+r_{sm}-1)} \text{ where}
\end{aligned}$$

where $K_I(\rho)$ is given by (2.3), thanks to Lemma 1.4. Thus, by applying Lemma 1.1, we obtain $\varepsilon_2(t) \;\le$

$$\exp((K_C^X + \rho^2)t)\varepsilon_2(0) + K_I(\rho)\exp(K_S^Y t)\sqrt{\frac{\exp(2(K_C^X + \rho^2 - K_S^Y)t) - 1}{2(K_C^X + \rho^2 - K_S^Y)}}h^{r_2+r_{sm}-1}$$

with "worst case" global rate $r_g \ge r_2 + r_{sm} - 1$ of mean square convergence for any $0 \le t \le T$ – an estimate which is particularly useful if $K_C^X + \rho^2 < 0$. □

2.2. A universal error estimate for the case $K_C^X > 0, K_S^Y \ge 0$. The expansive case is enlightened by the following slightly modified assertion taking into account the ratio $\Delta_{min}/\Delta_{max}$ of maximum Δ_{min} and minimum step sizes Δ_{max} as commonly met in variable step size algorithms. Thus, an extra proof is needed.

THEOREM 2.2. *Assume that the conditions $(A1) - (A7)$ with $K_C^X > 0$ are satisfied. Then the stochastic processes $X, Y \in H_2([0,T],\mu,\mathbb{R}^d)$ converge to each another with respect to the 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 nonrandom step sizes Δ_i (variable* *or constant) with* $0 < \Delta_{min} \le \Delta_i \le \Delta_{max} \le \delta_0$*, we have the following universal error estimates*

$$\varepsilon_2(t) \le \exp(K_A K_C^X (t-s))\varepsilon_2(s) + \tag{2.5}$$
$$+K_I(1)\exp(K_S^Y(t-s))\sqrt{\frac{\exp(2(K_A K_C^X - K_S^Y)(t-s))-1}{2(K_A K_C^X - K_S^Y)}}\;\Delta_{max}^{r_g}$$

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

$$\sup_{0\le t\le T} \varepsilon_2(t) \le \exp(K_A K_C^X T)\varepsilon_2(0) + \tag{2.6}$$
$$+K_I(1)\exp([K_S^Y]_+T)\sqrt{\frac{\exp(2(K_A K_C^X - K_S^Y)T)-1}{2(K_A K_C^X - K_S^Y)}}\;\Delta_{max}^{r_g}$$

with appropriate constants

$$K_A = \frac{\exp(2K_C^X\Delta_{max}) - 1 + 2\Delta_{max}}{2K_C^X\Delta_{max}}\frac{\Delta_{max}}{\Delta_{min}} \tag{2.7}$$
$$K_I(1) = \Big((K_0^C)^2 + (K_2^C)^2[1+(K_{SM})^2]\Big)^{1/2}(1+\mathbb{E}\,||y_0||_d^2)^{1/2}\,.$$

Proof. Let $K_C^X > 0$. Return to the estimate 2.4. Suppose that $0 < \Delta_{min} \le h \le \Delta_{max} \le 1$ for all $0 \le t, t+h \le T$. This implies that

$$\begin{aligned}
\varepsilon_2^2(t+h) &\le \exp(2K_C^X h)\varepsilon_2^2(t) + (K_2^C)^2\exp(2K_S^Y t)(1+\mathbb{E}\,||y_0||_d^2)h^{2r_2} + \\
&\quad +2h\varepsilon_2^2(t) + [(K_{SM}K_2^C)^2 + (K_0^C)^2]\exp(2K_S^Y t)(1+\mathbb{E}\,||y_0||_d^2)h^{2(r_2+r_{sm})-1} \\
&= [1 + \frac{\exp(2K_C^X h) - 1 + 2h}{2K_C^X h}2K_C^X h]\varepsilon_2^2(t) + \\
&\quad +[(K_{SM}K_2^C)^2 + (K_0^C)^2 + (K_2^C)^2]\exp(2K_S^Y t)(1+\mathbb{E}\,||y_0||_d^2)h^{2(r_2+r_{sm})-1} \\
&\le [1 + \frac{\exp(2K_C^X\Delta_{max}) - 1 + 2\Delta_{max}}{2K_C^X\Delta_{min}}2K_C^X h]\varepsilon_2^2(t) + \\
&\quad +[(K_{SM}K_2^C)^2 + (K_0^C)^2 + (K_2^C)^2]\exp(2K_S^Y t)(1+\mathbb{E}\,||y_0||_d^2)h^{2(r_2+r_{sm})-1} \\
&\le \exp\left(2K_A K_C^X h\right)\varepsilon_2^2(t) + +[K_I(1)]^2\exp(2K_S^Y t)\Delta_{max}^{2(r_2+r_{sm})-2}h
\end{aligned}$$

where the real constants $K_A, K_I(1) \ge 0$ are defined as in (2.7). Applying the Lemma 1.4 to the right hand side of latter inequality and taking the square root afterwards leads to

$$\varepsilon_2(t) \le \exp(K_A K_C^X t)\varepsilon_2(0) + K_I\exp(K_S^Y t)\sqrt{\frac{\exp(2(K_A K_C^X - K_S^Y)t)-1}{2(K_A K_C^X - K_S^Y)}}\Delta_{max}^{r_2+r_{sm}-1},$$

where the constant $K_I = K_I(1)$ is defined as above. □

2.3. A universal error estimate for the case $K_C^X = K_S^Y = 0$**.** Now, the dissipative case is discussed. Here the estimates of Theorem 2.1 turn out to be inefficient since the technical constant $\rho \neq 0$ would still occur in the related inequalities and taking the limit $\rho \to 0$ renders the inequalities to be useless. However, there is the following alternative which requires a slight modification of the proof of Theorem 2.1 too.

THEOREM 2.3. *Assume that the conditions* $(A1)-(A7)$ *with* $K_C^X = K_S^Y = 0$ *are satisfied.*
Then the stochastic processes $X, Y \in H_2([0,T], \mu, \mathbb{R}^d)$ *converge to each other with respect to the 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 nonrandom step sizes* Δ_i *(variable or constant) with* $0 < \Delta_i \leq \Delta_{max} \leq \delta_0$, *we have the universal estimate*

$$\varepsilon_2(t) \leq \varepsilon_2(s) \exp(t-s) + K_I(1) \sqrt{\frac{\exp(2(t-s)) - 1}{2}}\, \Delta_{max}^{r_g} \tag{2.8}$$

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

$$\sup_{0 \leq t \leq T} \varepsilon_2(t) \leq \varepsilon_2(0) \exp(T) + K_I(1) \sqrt{\frac{\exp(2T) - 1}{2}}\, \Delta_{max}^{r_g} \tag{2.9}$$

where $K_I(1)$ *is defined as*

$$K_I(1) = \sqrt{\Big((K_0^C)^2 + (K_2^C)^2[1 + (K_{SM})^2]\Big)(1 + \mathbb{E}\,||y_0||_d^2)}\,. \tag{2.10}$$

Furthermore, if additionally $\varepsilon_2(0) \leq K_{init}(1 + \mathbb{E}\,||y_0||_d^2)\Delta_{max}^{r_2-1}$, *the global error* ε_2 *is also controlled by the estimates*

$$\varepsilon_2(t) \leq \varepsilon_2(s) + K_2^C(T)\sqrt{|t-s|}\Delta_{max}^{r_2 + r_{sm}/2 - 1.0}, \tag{2.11}$$

$$\sup_{0 \leq t \leq T} \varepsilon_2(t) \leq \varepsilon_2(0) + K_2^C(T)\sqrt{T}\Delta_{max}^{r_2 + r_{sm}/2 - 1.0} \tag{2.12}$$

where

$$K_I^C(T) = \sqrt{[2(K_{init}^2 + K_2^C T)(K_{SM}K_2^C + K_0^C) + (K_2^C)^2](1 + \mathbb{E}\,||y_0||_d^2)}. \tag{2.13}$$

Proof. Suppose that $K_C^X = K_S^Y = 0$. Returning to (2.4), one arrives at

$$\begin{aligned}
\varepsilon_2^2(t+h) \leq \varepsilon_2^2(t) &+ (K_2^C)^2(1 + \mathbb{E}\,||y_0||_d^2)h^{2r_2} + \\
&+ 2\varepsilon_2(t)[K_{SM}K_2^C + K_0^C](1 + \mathbb{E}\,||y_0||_d^2)h^{r_2 + r_{sm}} \\
&\leq \varepsilon_2^2(t) + 2h\varepsilon_2^2(t) + K_I^2(1)h^{2(r_2 + r_{sm}) - 1.0} \\
&\leq \varepsilon_2^2(t) \cdot \exp(2h) + K_I^2(1)h^{2(r_2 + r_{sm}) - 1.0}.
\end{aligned}$$

An application of Lemma 1.4 leads to

$$\varepsilon_2^2(t) \leq \varepsilon_2^2(s) \cdot \exp(2(t-s)) + K_I^2(1)\frac{\exp(2(t-s)) - 1}{2}\Delta_{max}^{2(r_2 + r_{sm} - 1.0)}$$

which gives us the claimed estimates (2.8) and (2.9) with global rate $r_g = r_2 + r_{sm} - 1.0$ by taking square roots (Lemma 1.1). On the other hand, one may use the crude estimate

$$\varepsilon_2(t) \leq \varepsilon_2(0) + K_2^C t(1 + \mathbb{E}\,||y_0||_d^2)^{1/2}\Delta_{max}^{r_2 - 1}$$

originating from Lemma 1.9, along with the requirement

$$\varepsilon_2(0) \leq K_{init}(1 + \mathbb{E}\,||y_0||_d^2)^{1/2}\Delta_{max}^{r_2 - 1.0}.$$

Returning to (2.4), one obtains

$$\begin{aligned}\varepsilon_2^2(t+h) &\leq \varepsilon_2^2(t) + (K_2^C)^2(1+\mathbb{E}\,||y_0||_d^2)h^{2r_2} + \\ &\quad +2(K_{init}+K_2^C T)[K_{SM}K_2^C + K_0^C](1+\mathbb{E}\,||y_0||_d^2)h^{2r_2+r_{sm}-1.0} \\ &\leq \varepsilon_2^2(t) + (K_I^C(T))^2 h^{2r_2+r_{sm}-1.0}\end{aligned}$$

where

$$K_I^C(T) = \sqrt{[2(K_{init}^2 + K_2^C T)(K_{SM}K_2^C + K_0^C) + (K_2^C)^2](1+\mathbb{E}\,||y_0||_d^2)}.$$

An application of Lemma 1.4 provides the global estimate

$$\varepsilon_2^2(t) \leq \varepsilon_2^2(s) + (K_I^C(T))^2|t-s|\Delta_{max}^{2(r_2-1.0)+r_{sm}}$$

for $0 \leq s \leq t \leq T$. Taking the square root yields the second claimed group of estimates with global rate $r_g = r_2 + r_{sm}/2 - 1$, thanks to Lemma 1.1. Thus, the proof is complete. □

2.4. A universal error estimate for the case $K_C^X < 0, K_S^Y \leq 0$. It remains to consider the asymptotically contractive, dissipative case covered as follows.

THEOREM 2.4. *Assume that the conditions $(A1)-(A7)$ with $K_C^X < 0$ and $K_S^Y \leq 0$ are satisfied on the time-interval $[0,+\infty)$, all constants K occurring there in $(A1)-(A7)$ do not depend on the terminal times $T > 0$ and $\mu([0,+\infty)) < +\infty$.*
Then the stochastic processes $X, Y \in H_2([0,+\infty),\mu,\mathbb{R}^d)$ converge to each another with respect to the 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 nonrandom step sizes Δ_i (variable or constant) with $0 < \Delta_i \leq \Delta_{max} \leq \delta_0$, we have the universal error estimates

$$\sup_{0\leq t<+\infty} \varepsilon_2(t) \leq \varepsilon_2(0) + \frac{K_I(\rho)}{\sqrt{2|K_C^X + \rho^2 - K_S^Y|}}\Delta_{max}^{r_g}, \tag{2.14}$$

$$\lim_{t\to+\infty} \varepsilon_2(t) \leq \frac{K_I(\rho)}{\sqrt{2|K_C^X + \rho^2|}}\Delta_{max}^{r_g} \qquad \text{if} \quad K_S^Y = 0, \tag{2.15}$$

$$\lim_{t\to+\infty} \varepsilon_2(t) = 0 \qquad \text{if} \quad K_S^Y < 0 \quad \text{where} \tag{2.16}$$

$$K_I(\rho) = \frac{\sqrt{(K_0^C)^2 + (K_2^C)^2(\rho^2 + (K_{SM})^2)}}{\rho}\exp\Big((|K_C^X| + |K_S^Y|)\Delta_{max}\Big)(1+\mathbb{E}\,||y_0||_d^2)^{1/2}$$

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

Proof. Suppose that $K_C^X < 0$. The proof is similar to that for the Theorem 2.1. Take any constant parameter $\rho > 0$ satisfying $0 < \rho^2 < |K_C^X|$. Returning to (2.4) we get $\varepsilon_2^2(t+h) \leq$

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

Applying Lemma 1.4 with (1.10) to the latter inequality yields

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

By taking the square root in this inequality, thanks to the Lemma 1.1, we obtain the global estimate for $\varepsilon_2(t)$ (as in (2.1))

$$\varepsilon_2(t) \le \varepsilon_2(0)\exp((K_C^X+\rho^2)t) + K_I(\rho)\sqrt{\frac{\exp(2(K_C^X+\rho^2)t) - \exp(2K_S^Y t)}{2(K_C^X-\rho^2+K_S^Y)}}\Delta_{max}^{r_2+r_{sm}-1.0}$$

and hence

$$\varepsilon_2(t) \le \varepsilon_2(0)\exp((K_C^X+\rho^2)t) + \frac{K_I(\rho)}{\sqrt{2|K_C^X-\rho^2+K_S^Y|}}\Delta_{max}^{r_2+r_{sm}-1.0} \tag{2.17}$$

$$\le \varepsilon_2(0) + \frac{K_I(\rho)}{\sqrt{2|K_C^X-\rho^2+K_S^Y|}}\Delta_{max}^{r_2+r_{sm}-1.0}.$$

It remains to evaluate this result. Recall that $K_C^X+\rho^2 < 0$. Taking the supremum over all times t in the ride hand side of inequality (2.17) gives the estimate (2.14). Taking the limit as $t\to+\infty$ in (2.17) confirms the estimates (2.15) and (2.16). This completes the proof. □

3. Simple one-dimensional examples of SDEs. Some of the previous estimates turn out to be asymptotically sharp. To recognize this fact, consider the following one-dimensional examples with $\mathbb{D}=\mathbb{D}_t$ for all t, just for the sake of illustration.

3.1. Discretization of geometric Brownian motion. The geometric Brownian motion popularized by mathematical finance is governed by the Itô SDE

$$dX(t) = \alpha X(t)\,dt \;+\; \sigma\, X(t)\,dW(t)$$

with real constants α,σ. Suppose that $\mathbb{E}\,|X_0|^2 < +\infty$. It is not hard to verify that

$$2K_C^X \;=\; 2\alpha+\sigma^2$$

by the use of Itô formula applied to $V(x)=x^2$. This SDE is discretized by the **drift-implicit stochastic Theta-method**

$$Y(t_{n+1}) = Y(t_n) + [\theta Y(t_{n+1}) + (1-\theta)Y(t_n)]\alpha\Delta_n \;+\; \sigma Y(t_n)\,\Delta W_n$$

with deterministic implicitness $\theta\in\mathbb{R}^1$ and

$$\Delta_n = t_{n+1}-t_n, \quad \Delta W_n = W(t_{n+1}) - W(t_n) \in \mathcal{N}(0,\Delta_n)$$

where $\mathcal{N}(0,\Delta_n)$ denotes the standard Gaussian distribution with mean 0 and variance Δ_n. For the case of $K_C^X > 0$ one obtains very similar estimates to the standard one's for the mean square error as known in literature. However, what happens when $t\to+\infty$? From Schurz [21] it follows that $K_S^Y < 0$ whenever $K_C^X < 0$ and $\theta \ge 0.5$. Note that, for geometric Brownian motion, one can show that $K_S^Y = K_C^X = \alpha+\sigma^2/2$ satisfying

$$\forall y\in\mathbb{R}\,\forall t,h: 0\le t\le t+h\le T$$
$$\left(\mathbb{E}\left[|Y_{t,y}(t+h)|^2|Y_t=y\right]\right)^{1/2} \;\le\; \exp(K_S^Y h)|y|, \tag{$A2'$}$$

which modifies assumption $(A2)$ to $(A2')$. Then the case $K_S^Y < 0$ also implies that

$$\lim_{t\to+\infty}\varepsilon_2(t) \;=\; 0$$

ETNA
Kent State University
etna@mcs.kent.edu

which represents an asymptotically sharp estimate, thanks to a natural modification of our approximation theorem under $(A2')$ (for this fact, one may apply Lemma 1.9 as well). In contrast to that, there are equidistant step sizes for which drift-implicit Theta methods with $0 \leq \theta < 0.5$ (including the often used explicit Euler method) cannot control the magnitude of the mean square error $\varepsilon_2(t)$ as time t tends to $+\infty$ (due to exponentially growing second moments as time t advances).

3.2. Discretization of the Langevin equation. The well-known Langevin equation from statistical mechanics follows the SDE

$$dX(t) = \alpha X(t)\, dt \, + \, \sigma\, dW(t)$$

with real constants α, σ. The related stochastic process is also called as Ornstein-Uhlenbeck process. Suppose that $\mathbb{E}\,|X_0|^2 < +\infty$. It is not hard to verify that $K_C^X = \alpha$. This test SDE with additive noise can be discretized by the **drift-implicit stochastic Theta-method**

$$Y(t_{n+1}) = Y(t_n) + [\theta Y(t_{n+1}) + (1-\theta)Y(t_n)]\alpha\Delta_n \, + \, \sigma\, \Delta W_n$$

with deterministic implicitness $\theta \in \mathbb{R}$, and

$$\Delta_n = t_{n+1} - t_n, \quad \Delta W_n = W(t_{n+1}) - W(t_n) \in \mathcal{N}(0, \Delta_n).$$

In general one can easily confirm the standard estimates for the mean square error as known in the literature by our theorem. Instead of recalling them, we are particularly interested to derive estimates as the terminal time $t > 0$ tends to $+\infty$. Suppose that $K_C^X = \alpha < 0$. Then the related continuous time stochastic process $X = (X(t))_{t\geq 0}$ converges (a.s.) to the limit random variable

$$X_\infty \in \mathcal{N}(0, -\frac{\sigma^2}{2\alpha})$$

as time t tends to $+\infty$. In Schurz [23], [24] one finds that the trapezoidal method (i.e. the drift-midpoint theta-method with $\theta = 0.5$) has no mean square error at all, compared to the exact limit distribution. Other estimates for the mean square error of the stochastic Theta-methods with variable step sizes as time t tends to $+\infty$ are not known from the literature as far as the author knows. However, if $\theta \geq 0.5$ and $\alpha < 0$ then one may verify that $K_C^X = \alpha$, $K_S^Y = 0$. Also $K_{SM} = 0$, and $r_{sm} = 0.5$ since $r_0 = 2.0 \geq r_2 + r_{sm}$ with $r_2 = 1.5$. Combining the estimates of the Lemma 1.9 and Theorem 2.1 (or Theorem 2.4) with $0 < \rho^2 < |\alpha|$ (e.g. it is interesting to take $\rho^2 = |\alpha|/2$), we know that

$$\lim_{t\to+\infty} \varepsilon_2(t) \leq \min\left(\frac{K_2^C}{-\alpha}\sqrt{\Delta_{max}}, \frac{K_0^C + K_2^C}{\sqrt{2\rho^2|\alpha+\rho^2|}}\exp(|\alpha|\Delta_{max})\Delta_{max}\right)(1+\mathbb{E}\,|Y(0)|^2)^{1/2}$$

with appropriate constants $K_0^C = K_0^C(\alpha,\sigma,\theta)$ and $K_2^C = K_2^C(\alpha,\sigma,\theta)$ in the case of $\theta \geq 0.5, \alpha < 0$. Therefore, drift-implicit Theta-methods with $\theta \geq 0.5$ can maintain the global rate $\gamma_2 = 1.0$ of mean square convergence on infinite time-intervals $[0, +\infty)$, using any 'admissible discretizations' of stationary Ornstein-Uhlenbeck processes with

$$0 \, < \, \Delta_{min} \, \leq \, \Delta_n \, \leq \, \Delta_{max} \, < \, +\infty.$$

In contrast to this fact, explicit Euler methods or Theta-methods with parameter $0 < \theta < 0.5$ may already fail (due to exponentially growing second moments, which means loss of control on stability constant K_S^Y on the Hilbert space $H_2([0, +\infty), \mu, \mathbb{R}^d)$.

3.3. Discretization of a nonlinear SDE. In physical field theory one encounters Itô equations of the type

$$dX(t) = \Big(\alpha X(t) - \beta^2[X(t)]^3\Big)\,dt \,+\, \sigma\,X(t)\,dW(t)$$

with real constants α, β, σ. Suppose that $\mathbb{E}\,|X(0)|^2 < +\infty$. Then, this SDE with Lipschitz continuous diffusion (i.e. with a smoothness rate $r_{sm} \geq 0.5$) possesses a mean square contractive (in the wide sense), unique solution X with uniformly bounded second moments and mean square contractivity constant $K_C^X \leq \alpha + \sigma^2/2$. The solution X can be discretized by the **partial linear-implicit Euler method**

$$Y(t_{n+1}) = Y(t_n) \,+\, \Big(\alpha Y(t_n) - \beta^2[Y(t_n)]^2 Y(t_{n+1})\Big)\Delta_n \,+\, \sigma\,Y(t_n)\,\Delta W_n$$

or by the **partial nonlinear-implicit Euler method**

$$Y(t_{n+1}) = Y(t_n) \,+\, \Big(\alpha Y(t_n) - \beta^2[Y(t_{n+1})]^3\Big)\Delta_n \,+\, \sigma\,Y(t_n)\,\Delta W_n.$$

It is not hard to see that both methods have uniformly bounded second moments, provided that $\mathbb{E}\,|Y(0)|^2 < +\infty$. More precisely, one gets

$$\begin{aligned}\mathbb{E}\,|Y(t_{n+1})|^2 &\leq \mathbb{E}\,|Y(t_n)|^2 \cdot (1 + (2\alpha + \sigma^2 + \alpha^2\Delta_n)\Delta_n)\\ &\leq \mathbb{E}\,|Y(0)|^2 \cdot \exp\Big([2\alpha + \sigma^2 + \alpha^2 \max_{i\in\{0,1,\ldots,n\}} \Delta_i](t_{n+1} - t_0)\Big),\end{aligned}$$

hence, for the cadlag approximation $Y(t)$ based on step functions with jumps at height $Y(t_n)$ at t_n, one has

$$\sup_{t_0\leq t\leq T} \mathbb{E}\,|Y(t)|^2 \;=\; \sup_{n\in\{0,1,\ldots,n_T\}} \mathbb{E}\,|Y(t_n)|^2 \;\leq\; \mathbb{E}\,|Y(0)|^2 \cdot \exp\Big(2[K_S^Y]_+(T - t_0)\Big)$$

with stability constant

$$K_S^Y \;\leq\; \alpha + \frac{\sigma^2 + \alpha^2 \max_{i\in\{0,1,\ldots,n\}} \Delta_i}{2}$$

Furthermore, local mean and mean square consistency $(A4)$ and $(A5)$ with $\mathbb{D} = \mathbb{R}$ with rates $r_0 \geq 1.5$ and $r_2 \geq 1.0$ can be shown, provided that the initial moments $\mathbb{E}\,|X(0)|^4 + \mathbb{E}\,|Y(0)|^4 < +\infty$ and $\Delta_{min} > 0$. (For this purpose one may use the forward Euler method). Thus, the assumptions of our main theorems are satisfied. Therefore, we may conclude the global mean square convergence rate $r_g = 0.5$ which clearly exhibits a new result compared to the literature where the Lipschitz continuity of drift part is usually required. A similar result can be found for the split step Backward Euler method due to Higham, Mao and Stuart [6].

4. Numerical experiments for a nonlinear SDE. Assume that we have $\mathbb{E}\,||X_0||_2^2 < +\infty$ (more precisely, $\mathbb{E}\,[\omega^2(X_0^{(1)})^2 + (X_0^{(2)})^2] < +\infty$). We conducted numerical experiments in computing the discretizations of a *generalized Van der Pol Oscillator with multiplicative noise*, given by the Itô SDE

$$\begin{aligned}dX_t^{(1)} &= X_t^{(2)}\,dt\\ dX_t^{(2)} &= [-\omega^2 X_t^{(1)} + \gamma(1 - \mu_1(X_t^{(1)})^2 - \mu_2(X_t^{(2)})^2)X_t^{(2)}]\,dt \,+\, \sigma X_t^{(2)}\,dW_t \qquad (4.1)\end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

where all constants $\omega, \gamma, \mu_1, \mu_2, \sigma \geq 0$ are some nonnegative real numbers. This two-dimensional system satisfies the conditions (1.2) – (1.5) with respect to the norm

$$||(x,y)^T||_{2,\omega} = (\omega^2 x^2 + y^2)^{1/2}$$

with $\omega^2 > 0$, hence the unique, strong solution $X = (X^{(1)}, X^{(2)})$ exists with $X \in H_2([0,T], \mu, \mathbb{R}^d)$ for all finite $T > 0$. In fact there is a random limit cycle where all trajectories approach to (at least we suspect it whenever $\gamma\mu_2 > 0$). Assuming that $\mathbb{E}\,||X_0||_2^4 < +\infty$ and $\mu_1 = 0$, one can show that

$$K_C^X \leq \gamma + \frac{\sigma^2}{2}.$$

A computationally easily implementable, converging and stable discretization of this system is done by linear-implicit implementations of drift-implicit Euler methods. For example, take the **partial linear-implicit Euler method** ([30])

$$\text{(4.2)} \qquad Y_{n+1}^{(1)} = Y_n^{(1)} + Y_n^{(2)}\Delta_n$$
$$Y_{n+1}^{(2)} = Y_n^{(2)} + \Big(-\omega^2 Y_n^{(1)} + \gamma Y_n^{(2)} - \gamma(\mu_1 [Y_n^{(1)}]^2 + \mu_2 [Y_n^{(2)}]^2) Y_{n+1}^{(2)} \Big)\Delta_n + \sigma Y_n^{(2)} \Delta W_n.$$

Assuming that $\mathbb{E}\,||Y_0||_2^2 < +\infty$ and $\omega^2 > 0$, one can show that

$$\mathbb{E}\,||Y(t)||_2^2 \leq K_m \exp(2K_S^Y t) \mathbb{E}\,||Y_0||_2^2$$

for the cadlag approximation $Y(t)$ constructed as a step function with jumps $Y(t_n) = Y_n$ based on (4.2) with appropriate constants $K_m \leq \max(1,\omega^2)/\min(1,\omega^2)$ and

$$2K_S^Y \leq 2\gamma + \sigma^2 + \Delta_{max} \max(\gamma + \omega^2, 1 + \gamma^2 + \gamma\omega^2).$$

To see this, use a recursive estimation of the dominating Lyapunov function

$$v(n) = \mathbb{E}\left[\omega^2 (Y_n^{(1)})^2 + (Y_n^{(2)})^2 (1 + \gamma(\mu_1 (Y_{n-1}^{(1)})^2 + \mu_2 (Y_{n-1}^{(2)})^2)\Delta_n)^2\right].$$

Consequently, the assumptions of Theorem 2.1 with the slight modification of assumption $(A2)$ to

$$\exists V \in C^0(\mathbb{D}, \mathbb{R}^1_+)\, \forall y \in \mathbb{D} \,:\, \mathbb{E}\, V(y) < +\infty\, \forall t,h \,:\, 0 \leq h \leq \delta_0,\, 0 \leq t \leq t+h \leq T$$
$$\mathbb{E}\,(V(y_0))^2 < +\infty\,,$$
$$(A2'') \qquad \Big(\mathbb{E}\,[1 + ||Y_{t,y}(t+h)||_d^2 | Y_t = y]\Big)^{1/2} \leq V(y) \quad \text{and}$$
$$\Big(\mathbb{E}\,[(V(Y_{t,y}(t+h)))^2 | Y_t = y]\Big)^{1/2} \leq \exp(K_S^Y h) \cdot V(y),$$

are satisfied. Because a modification of Theorem 2.1 with the modified assumptions using Lyapunov function techniques remains valid (i.e. replace the term $1 + ||.||_d^2$ by $(V(.))^2$ whereever met in the assumptions, see a work of the author in progress, cf. [31]), we may use these methods to approximate the exact solution of (4.1) (at least with $\mu_1 = 0$) with global mean square convergence order $r_g \geq 0.5$ on any finite time interval $[0,T]$ (Note that local mean and mean square consistency $(A4)$ and $(A5)$ with $\mathbb{D} = \mathbb{R}^2$ with rates $r_0 \geq 1.5$ and $r_2 \geq 1.0$ are shown by means of the forward Euler method under the presence of multiplicative white noise, provided that $\mathbb{E}\,||X_0||_2^4 + \mathbb{E}\,||Y_0||_2^4 < +\infty$ and $\Delta_{min} > 0$).

ETNA
Kent State University
etna@mcs.kent.edu

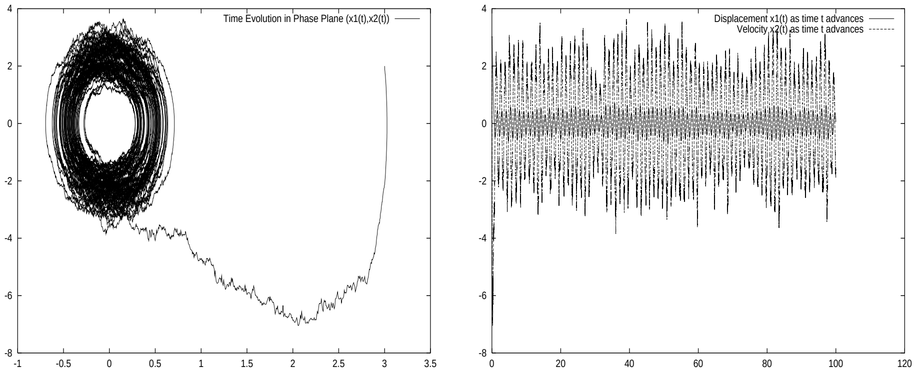

FIG. 4.1. *'Typical' trajectories approaching to the (random) limit cycle discretized by the partial-implicit Euler method applied to the generalized stochastic Van der Pol oscillator.*

In general partial linear-implicit methods such as (4.2) allow control on the numerical stability, convergence and dissipativity (existence of the approximate random limit cycle which has a support concentrated on an ellipse). In contrast to that observation, there are parameter sets where the well-known explicit Euler or Mil'shtein methods break down by showing inadequate behavior compared to that of the exact solution (a fact which can be seen best with linear systems with multiplicative white noise, cf. [23], [24]). In our experiments, we have $\mathbb{D} = \mathbb{R}^2$, $\omega = 5$, $\mu_1 = 0.05$, $\mu_2 = 0.25$, $\gamma = 1.0$, $\sigma = 0.5$ and $T = 100$ for equidistant discretization of $[0, T]$ with $\Delta = 0.001$ started at deterministic point $Y_0 = (3, 2)^T$. In the Figure 4.1 we clearly recognize that the (random) limit cycle is replicated by the partial linear-implicit numerical method (4.2) and its trajectories converge to that limit cycle, independently of its initial value. This fact is due to the inherent type of nonlinearity involving the terms with parameters $\gamma > 0, \mu_1 \geq 0, \mu_2 > 0$. In passing we note that it would be desirable to develop a general "numerical Lyapunov-technique" to treat other nonlinear oscillators than the stochastically perturbed Van der Pol oscillator.

5. Further developments and summarizing comments. Similar approximation results are true for the strong Banach spaces

$$\mathcal{H}_2^p([0,T],\mu,\mathbb{R}^d) := \left\{ X = (X_t)_{0\leq t\leq T} : \begin{array}{l} X_t \in \mathbb{R}^d \text{ is } (\mathcal{F}_t, \mathcal{B}(\mathbb{R}^d)) - \text{measurable}, \\ X \text{ cadlag process on } (\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P}), \\ \int_0^T \mathbb{E}\Big(\sup_{0\leq s\leq t} < X_s, X_s >_d\Big)^{p/2} d\mu(t) < +\infty \end{array} \right\}$$

where $p \geq 2, p \in \mathbb{R}_+$ is deterministic, which form pseudo Hilbert spaces with subadditive pseudo scalar product for $p = 2$. Of course, we have the trivial inclusion $\mathcal{H}_2^2([0,T],\mu,\mathbb{R}^d) \subseteq H_2([0,T],\mu,\mathbb{R}^d)$ for fixed positive measure μ. For this stronger setup it is crucial that a finiteness of some higher moments of involved stochastic processes is guaranteed (cf. with the case of nonlinear SDEs). Further extensions to separable Banach space-valued random processes are conceivable. There the case of separable Hilbert spaces may play a special role to extend the main results to separable Banach spaces containing those Hilbert spaces. See also author's work [31].

We also plan to incorporate "numerical Lyapunov function techniques" to extend our results to more general classes of nonlinear SDEs as already indicated by the section on our numerical experiments. Partial-implicit and split step techniques seem to be very promissing to control $(\mathbb{D}_t)$-invariance, consistency, convergence, stability and dissipativity of numerical

ETNA
Kent State University
etna@mcs.kent.edu

approximations for nonlinear SDEs. See forthcoming papers of the author. It would be desirable to discuss the maximum order bounds of those methods too, as indicated in the case of Runge-Kutta methods by Burrage and Burrage [4]. Furthermore, which error distribution as studied by Kurtz and Protter [16] do they have for nonequidistant and random partitions in the most general case? We must leave the latter two problems unanswered - as many other questions here.

This paper thoroughly follows the main principles of numerical approximation theory as discussed in the previous sections. We have clearly seen that the Kantorovič-Lax-Richtmeyer principle "Stability and Consistency imply Convergence" holds in stochastic-numerical analysis of well-posed problems too (i.e. together with some kind of Invariance, Contractivity and Smoothness of Martingale Part). The essentials of all of these 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

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

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) $\Longleftrightarrow$	Approximative Well-posedness: **Stability** of Y **Contractivity** of X

$\Updownarrow$

Global Convergence of (X, Y)

This diagram describes the main crossrelations, the fundamental equivalence principles in the context of stochastic approximations. This exhibits the point where one arrives at the heart of sophisticated numerical approximation theory for stochastic processes. Here the concept of consistency plays the central role. Contractivity and stability property can be exchanged equivalently (but simultaneously) if consistency holds (due to the inherent symmetry of the given approximation problem). Convergence is extracted from the interplay of consistency, stability and contractivity. Our remaining goal is just to make these main principles come alive in conjunction with SDEs / SPDEs and their numerical analysis in a concise course. For example, for the case of stochastic functional differential equations (SFDEs) as seen with the stochastic Pantograph equation in Baker and Buckwar [3].

The main theorems are valid for numerical approximations Y using both equidistant and variable partitions of time-intervals $[0, T]$ (so taylored for a convergence analysis of algorithms with variable step sizes). Furthermore, we may even obtain some asymptotically sharp estimates for the approximation errors and their orders in the case of linear systems integrated with constant step sizes on infinite intervals $[0, +\infty)$. The meaningfulness of the presented approximation theorems can be seen in the large range of potential application to several types of stochastic-numerical approximation problems (like to SFDEs and stochastic integro-differential equations), leading to new striking results.

Acknowledgments. The author likes to clarify that this paper is based on his (so far unpublished) technical report **1669** at IMA, University of Minnesota, Minneapolis, 1999, where

the main ideas can already be found in the context of random Banach spaces. The approximation problem for the stochastic L^1-case is sketched in our book chapter [27]. However, in both presentations one can only find fairly crude estimates for the related global approximation error. Here we are trying to have a more transparent presentation for the case of relatively simple Hilbert spaces and an easier readable text which underlines the applicability of the stochastic-numerical approximation principles in a significant manner. An axiomatic approach for Hilbert-space-valued stochastic processes is presented in a more general framework in [31]. There we continue our work by incorporating the knowledge on certain Lyapunov-type functionals.

REFERENCES

[1] L. ARNOLD, *Stochastic Differential Equations: Theory and Applications*, Krieger Publishing Company, Malabar, 1992 (reprint of the original, John Wiley and Sons, Inc. from 1974, German original, Oldenbourg Verlag from 1973).
[2] L. ARNOLD, *Random Dynamical Systems*, Springer, Berlin, 1998.
[3] C.T. BAKER and E. BUCKWAR, *Continuous θ-methods for the stochastic pantograph equation*, Electron. Trans. Numer. Anal., 11 (2000), pp. 131–151. http://etna.mcs.kent.edu/vol.11.2000/pp131-151.dir/pp131-151.pdf.
[4] K. BURRAGE AND P.M. BURRAGE, *Order conditions of stochastic Runge–Kutta methods by B-series*, SIAM J. Numer. Anal., 38 (2000), No. 5, pp. 1626–1646.
[5] S.K. GODUNOV and V.S. RYABENKII, *Difference Schemes: An Introduction to the Underlying Theory*, North-Holland, Amsterdam, 1987.
[6] 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.
[7] Y.Z. HU, *Semi-implicit Euler-Maruyama scheme for stiff stochastic equations*, in Stochastic Analysis and Related Topics V: The Silivri Workshop (held in Silivri, Norway, July 18-29, 1994), Koerezlioglu, H. et al., eds., Progr. Probab. 38, Birkhäuser, Boston, 1996, pp. 183–302.
[8] K. ITÔ, *Stochastic integral*, Proc. Imp. Acad. Tokyo, 20 (1944), pp. 519–524.
[9] K. ITÔ, *On a formula concerning stochastic differential equations*, Nagoya Math. J., 3 (1951), pp. 55–65.
[10] J. JACOD AND P. PROTTER, *Essentials of Probability*, Springer, New York, 2000.
[11] L.V. KANTOROVIČ, *Functional analysis and applied mathematics* (in Russian), Uspekhi Mat. Nauk (N.S.), 3 (1948), No. 6 (28), pp. 89–185.
[12] L.V. KANTOROVIČ, em Functional analysis and applied mathematics (in Russian), Vestnik Leningrad. Univ., 3 (1948), No. 6, pp. 3–18.
[13] I. KARATZAS AND S. SHREVE, *Brownian Motion and Stochastic Calculus*, Springer, New York, 1988.
[14] P.E. KLOEDEN, E. PLATEN AND H. SCHURZ, *Numerical Solution of SDEs Through Computer Experiments*, First edition, Springer, Berlin, 1994.
[15] A.N. KOLMOGOROV, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (in German, reprint of the 1933 edition), Springer, Berlin, 1977, also *Foundations of the Theory of Probability*, Chelsea Publishing Company, New York, 1950.
[16] T.G. KURTZ AND P. PROTTER, *Wong-Zakai corrections, random evolutions, and simulation schemes for SDE's*, in Stochastic Analysis, Proc. Conf. Honor Moshe Zakai 65th Birthday, Haifa/Isr., 1991, pp. 331–346.
[17] P.D. LAX, *A stability theorem for solutions of abstract differential equations, and its application to the study of the local behavior of solutions of elliptic equations*, Comm. Pure Appl. Math., 9 (1956), pp. 747–766.
[18] P.D. LAX and R.D. RICHTMEYER, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math., 9 (1956), pp. 267–293.
[19] G.N. MILSTEIN, *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers, Dordrecht, 1995.
[20] R.D. RICHTMEYER, R.D. and K.W. MORTON, *Difference Methods for Initial-Value Problems*, Second edition, Interscience, New York, 1967.
[21] H. SCHURZ, *Stability, Stationarity, and Boundedness of Some Implicit Numerical Methods for Stochastic Differential Equations and Applications* (original: Report No. 11, WIAS, Berlin, 1996), Logos-Verlag, Berlin, 1997.
[22] H. SCHURZ, *General principles for numerical approximation of stochastic processes on some stochastically weak Banach spaces*, Report 1669 (1999), IMA, University of Minnesota, Minneapolis, pp. 26.
[23] H. SCHURZ, *Preservation of probabilistic laws through Euler methods for Ornstein-Uhlenbeck process*, Stochastic Anal. Appl., 17 (1999), No. 3, pp. 463–486.

ETNA
Kent State University
etna@mcs.kent.edu

[24] H. SCHURZ, *The invariance of asymptotic laws of stochastic systems under discretization*, Z. Angew. Math. Mech., 79 (1999), No. 6, pp. 375–382.
[25] H. SCHURZ, *Moment attractivity, stability and contractivity exponents of stochastic dynamical systems*, Discrete Contin. Dynam. Systems, 7 (2001), No. 3, pp. 487–515.
[26] H. SCHURZ, *On moment-dissipative stochastic dynamical systems*, Dynam. Systems Appl., 10 (2001), No. 2, pp. 11–44.
[27] H. SCHURZ, *Numerical analysis of SDEs without tears*, in Handbook of Stochastic Analysis and Applications, D. Kannan and V. Lakshmikantham, eds., Marcel Dekker, Basel, 2002, pp. 237–359.
[28] H. SCHURZ, *Stability analysis of (balanced) stochastic Theta methods*, in Proceedings of Stochastic Numerics 2001, W. Petersen, ed., ETH, Zurich, 2001, pp. 1–39.
[29] H. SCHURZ, *Two general theorems for the numerical approximation of stochastic processes on separable Hilbert spaces*, in Proceedings of Neural, Parallel, and Scientific Computations, M.P. Bekakos, G.S. Ladde, N.G. Medhin and M. Sambandham, eds., Vol. 2, Dynamic Publishers, Atlanta, 2002, pp. 55–58.
[30] H. SCHURZ, *Partial and linear-implicit numerical methods for nonlinear SDEs*, Unpublished Working Paper, Universidad de Los Andes, Bogotá, 1998, pp. 29.
[31] H. SCHURZ, *An axiomatic approach to the numerical approximation of Hilbert-space valued stochastic processes*, Report M-02-007 (2002), Southern Illinois University, Carbondale, pp. 18.
[32] A.N. SHIRYAEV, *Probability*, Second edition, Springer, Berlin, 1996.
[33] J.C. STRIKWERDA, *Finite Difference Schemes and Partial Differential Equations*, Chapman & Hall, New York, 1989.
[34] D. TALAY, *Simulation of stochastic differential systems*, in Probabilistic Methods in Applied Physics, P. Krée and W. Wedig, eds. (Springer Lecture Notes in Physics 451), Springer, Berlin, 1995, pp. 54–96.