STRONG CONVERGENCE OF EULER-TYPE METHODS FOR NONLINEAR STOCHASTIC DIFFERENTIAL EQUATIONS[*]

DESMOND J. HIGHAM[†], XUERONG MAO[‡], AND ANDREW M. STUART[§]

Abstract. Traditional finite-time convergence theory for numerical methods applied to stochastic differential equations (SDEs) requires a global Lipschitz assumption on the drift and diffusion coefficients. In practice, many important SDE models satisfy only a local Lipschitz property and, since Brownian paths can make arbitrarily large excursions, the global Lipschitz-based theory is not directly relevant. In this work we prove strong convergence results under less restrictive conditions. First, we give a convergence result for Euler–Maruyama requiring only that the SDE is locally Lipschitz and that the pth moments of the exact and numerical solution are bounded for some $p > 2$. As an application of this general theory we show that an implicit variant of Euler–Maruyama converges if the diffusion coefficient is globally Lipschitz, but the drift coefficient satisfies only a one-sided Lipschitz condition; this is achieved by showing that the implicit method has bounded moments and may be viewed as an Euler–Maruyama approximation to a perturbed SDE of the same form. Second, we show that the optimal *rate* of convergence can be recovered if the drift coefficient is also assumed to behave like a polynomial.

Key words. backward Euler, Euler–Maruyama, finite-time convergence, implicit, moment bounds, nonlinearity, one-sided Lipschitz condition, split-step

AMS subject classifications. 65C30, 65C20, 65L20

PII. S0036142901389530

1. Introduction. In this paper we study the numerical solution of the stochastic differential equation (SDE)

$$(1.1) \qquad dy(t) = f(y(t))dt + g(y(t))dW(t), \quad 0 \le t \le T, \quad y(0) = y_0.$$

Here $y(t) \in \mathbb{R}^m$ for each t, and $W(t)$ is a d-dimensional Brownian motion. Thus $f : \mathbb{R}^m \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^{m \times d}$. We assume that the initial condition is chosen independently of the Wiener measure driving the equation and that all pth moments of y_0, $p > 0$, are finite. Our primary objective is to study strong convergence questions for numerical approximations in the case where f and g are not necessarily globally Lipschitz functions. Most of the existing convergence theory for numerical methods requires f and g to be globally Lipschitz; see [12, 14], for example. Recent work has studied convergence in probability [6, 8] and almost sure convergence [7], under more relaxed conditions on f and g. We focus here on strong mean square convergence, in the sense of [12, Theorem 10.6.3], which implies convergence in probability. The main result of [11] is directly related to our work; we summarize the connections at the end of this section. We also note that the work of Schurz [16] contains a number

of a priori bounds on the numerical solutions of SDEs from a nonlinear stability perspective, under structural assumptions similar to those we employ here.

In section 2, we prove that the Euler–Maruyama method converges strongly if f and g are locally Lipschitz (e.g., for $f, g \in C^1$) and the exact and numerical solution have a bounded pth moment for some $p > 2$ (Theorem 2.2). The bounded moment assumption will not, of course, hold, in general, as solutions to the SDE may explode in a finite time. In section 3 we therefore impose further assumptions on f and g to ensure that $y(t)$ has bounded moments: we assume that g is globally Lipschitz and that f satisfies a one-sided Lipschitz condition. The one-sided Lipschitz condition has proved effective in the analysis of numerical methods for deterministic problems and occurs naturally in a variety of applications [1, 3, 4, 5, 17]. For a suitably constructed "split-step" implicit variant of Euler–Maruyama we establish strong convergence (Theorem 3.3) by (a) showing that the method corresponds to Euler–Maruyama on a perturbed SDE and (b) showing that all moments of the numerical solution are bounded. We are unable to establish moment bounds for the forward Euler method and, indeed, it may not be possible to do so; the work of [7] is of importance for nonimplicit methods since, being an almost sure convergence study, it does not require moment bounds.

In section 4 we turn our attention to establishing an optimal *rate* of convergence for the split-step implicit method. We show that if f satisfies a one-sided Lipschitz condition and behaves polynomially, then Euler–Maruyama converges strongly at the optimal rate (Theorem 4.4), provided moment bounds hold. We use this to study the split-step implicit method, for which moment bounds can be found (Theorem 4.7), again by showing that the method corresponds to Euler–Maruyama on a perturbed SDE. In section 5 this result is extended to a more widely used implicit variant of Euler–Maruyama by relating the two implicit methods (Theorem 5.3). A summary is given in section 6.

It is worth mentioning at this point how our work compares with that of Hu [11]. Theorem 2.4 of [11] is a very important contribution to numerical SDE theory, being, to our knowledge, the first strong convergence result without global Lipschitz assumptions. Hu considered only the backward Euler method and derived a result with the optimal rate of convergence, and hence his work may be compared with Theorem 5.3 below. Both results assume C^1 coefficients in the SDE, a one-sided Lipschitz condition for the drift, and a global Lipschitz condition for the diffusion. Theorem 5.3 below imposes polynomial-type growth on the drift (Assumption 4.1), whereas Hu allows for a more general exponential growth. On the other hand, Theorem 5.3 and all the other results in our work deal with a very strong error measure, $\mathbb{E}\left[\sup_{0 \le t \le T} |\overline{Z}(t) - y(t)|^2\right]$, whereas [11] uses the less stringent measure $\int_0^T \mathbb{E}|\overline{Z}(t) - y(t)|^2 \, dt$. We also note that [11] uses a different continuous-time extension. Overall, Hu's result for backward Euler applies to a wider class of SDEs but controls a weaker measure of the error. The techniques of analysis are significantly different, although both hinge on establishing moment bounds for the exact and numerical solutions.

2. General result for Euler–Maruyama. Given a stepsize $\Delta t > 0$, the Euler–Maruyama (EM) method applied to (1.1) computes approximations $X_k \approx y(t_k)$, where $t_k = k\Delta t$, by setting $X_0 = y_0$ and forming

$$(2.1) \qquad X_{k+1} = X_k + \Delta t f(X_k) + g(X_k)\Delta W_k,$$

where $\Delta W_k = W(t_{k+1}) - W(t_k)$. We find it convenient to use continuous-time approximations, and hence we define $\overline{X}(t)$ by

$$\overline{X}(t) := X_k + (t - t_k)f(X_k) + g(X_k)(W(t) - W(t_k)) \quad \text{for } t \in [t_k, t_{k+1}).$$

In our analysis it will be more natural to work with the equivalent definition

$$(2.2) \qquad \overline{X}(t) := X_0 + \int_0^t f(X(s))ds + \int_0^t g(X(s))dW(s),$$

where $X(t)$ is defined by

$$(2.3) \qquad\qquad X(t) := X_k \quad \text{for } t \in [t_k, t_{k+1}).$$

Note that $\overline{X}(t_k) = X(t_k) = X_k$; that is, $\overline{X}(t)$ and $X(t)$ coincide with the discrete solution at the gridpoints. We refer to $X(t)$ and $\overline{X}(t)$ as *continuous-time extensions* of the discrete approximation $\{X_k\}$. We will study the error in $\overline{X}(t)$ in the supremum norm; this will, of course, give an immediate bound for the error in the discrete approximation.

Our first result makes the following assumption on the SDE (1.1) and the exact and numerical solutions. Here, and throughout the paper, $|\cdot|$ is used to denote both the Euclidean vector norm and the Frobenius (or trace) matrix norm.

ASSUMPTION 2.1. *For each $R > 0$ there exists a constant C_R, depending only on R, such that*

$$(2.4) \quad |f(a) - f(b)|^2 \vee |g(a) - g(b)|^2 \leq C_R|a - b|^2 \ \forall a, b \in \mathbb{R}^m \text{ with } |a| \vee |b| \leq R.$$

For some $p > 2$ there is a constant A such that

$$(2.5) \qquad\qquad \mathbb{E}\left[\sup_{0 \leq t \leq T} |\overline{X}(t)|^p\right] \vee \mathbb{E}\left[\sup_{0 \leq t \leq T} |y(t)|^p\right] \leq A.$$

Inequality (2.4) is a local Lipschitz assumption. From the mean value theorem, any f and g in C^1 will satisfy (2.4). The inequality (2.5) states that the pth moments of the exact and numerical solution are bounded for some $p > 2$. We now prove that Assumption 2.1 is sufficient to ensure strong convergence of EM.

THEOREM 2.2. *Under Assumption* 2.1, *the EM solution* (2.1) *with continuous-time extension* (2.2) *satisfies*

$$(2.6) \qquad\qquad \lim_{\Delta t \to 0} \mathbb{E}\left[\sup_{0 \leq t \leq T} |\overline{X}(t) - y(t)|^2\right] = 0.$$

Proof. First, we define

$$\tau_R := \inf\{t \geq 0 : |\overline{X}(t)| \geq R\}, \quad \rho_R := \inf\{t \geq 0 : |y(t)| \geq R\}, \quad \theta_R := \tau_R \wedge \rho_R$$

and

$$e(t) := \overline{X}(t) - y(t).$$

Recall the Young inequality: for $r^{-1} + q^{-1} = 1$

$$ab \leq \frac{\delta}{r}a^r + \frac{1}{q\delta^{q/r}}b^q \quad \forall a, b, \delta > 0.$$

We thus have for any $\delta > 0$

$$
\mathbb{E}\left[\sup_{0 \le t \le T} |e(t)|^2\right] = \mathbb{E}\left[\sup_{0 \le t \le T} |e(t)|^2 \mathbf{1}_{\{\tau_R > T, \rho_R > T\}}\right] + \mathbb{E}\left[\sup_{0 \le t \le T} |e(t)|^2 \mathbf{1}_{\{\tau_R \le T \text{ or } \rho_R \le T\}}\right]
$$

$$
\le \mathbb{E}\left[\sup_{0 \le t \le T} |e(t \wedge \theta_R)|^2 \mathbf{1}_{\{\theta_R > T\}}\right] + \frac{2\delta}{p}\mathbb{E}\left[\sup_{0 \le t \le T} |e(t)|^p\right]
$$

$$
(2.7) \qquad + \frac{1 - \frac{2}{p}}{\delta^{2/(p-2)}}\mathbb{P}\big(\tau_R \le T \text{ or } \rho_R \le T\big).
$$

Now

$$
\mathbb{P}\left(\tau_R \le T\right) = \mathbb{E}\left[\mathbf{1}_{\{\tau_R \le T\}}\frac{|\overline{X}(\tau_R)|^p}{R^p}\right] \le \frac{1}{R^p}\mathbb{E}\left[\sup_{0 \le t \le T} |\overline{X}(t)|^p\right] \le \frac{A}{R^p},
$$

using (2.5). A similar result can be derived for ρ_R so that

$$
\mathbb{P}\big(\tau_R \le T \text{ or } \rho_R \le T\big) \le \mathbb{P}\big(\tau_R \le T\big) + \mathbb{P}\big(\rho_R \le T\big) \le \frac{2A}{R^p}.
$$

Using these bounds along with

$$
\mathbb{E}\left[\sup_{0 \le t \le T} |e(t)|^p\right] \le 2^{p-1}\mathbb{E}\left[\sup_{0 \le t \le T} \big(|\overline{X}(t)|^p + |y(t)|^p\big)\right] \le 2^p A
$$

in (2.7) gives

$$
\mathbb{E}\left[\sup_{0 \le t \le T} |e(t)|^2\right] \le \mathbb{E}\left[\sup_{0 \le t \le T} |\overline{X}(t \wedge \theta_R) - y(t \wedge \theta_R)|^2\right]
$$

$$
(2.8) \qquad + \frac{2^{p+1}\delta A}{p} + \frac{(p-2)2A}{p\delta^{2/(p-2)}R^p}.
$$

Now we bound the first term on the right-hand side of (2.8) using an approach similar to a finite-time convergence proof for the globally Lipschitz case. Using

$$
y(t \wedge \theta_R) := y_0 + \int_0^{t \wedge \theta_R} f(y(s))ds + \int_0^{t \wedge \theta_R} g(y(s))dW(s),
$$

(2.2), and Cauchy–Schwarz, we have

$$
|\overline{X}(t \wedge \theta_R) - y(t \wedge \theta_R)|^2 = \left| \int_0^{t \wedge \theta_R} f(X(s)) - f(y(s))ds \right.
$$

$$
+ \left. \int_0^{t \wedge \theta_R} g(X(s)) - g(y(s))dW(s) \right|^2
$$

$$
\le 2\left[T \int_0^{t \wedge \theta_R} |f(X(s)) - f(y(s))|^2 ds \right.
$$

$$
+ \left. \left| \int_0^{t \wedge \theta_R} g(X(s)) - g(y(s))dW(s) \right|^2 \right].
$$

So, from the local Lipschitz condition (2.4) and Doob's martingale inequality [14] we have for any $\tau \leq T$

$$\mathbb{E}\left[\sup_{0 \leq t \leq \tau} |\overline{X}(t \wedge \theta_R) - y(t \wedge \theta_R)|^2\right]$$

$$\leq 2C_R(T+4)\mathbb{E}\int_0^{\tau \wedge \theta_R} |X(s) - y(s)|^2 ds$$

$$\leq 4C_R(T+4)\mathbb{E}\int_0^{\tau \wedge \theta_R} \left[|X(s) - \overline{X}(s)|^2 + |\overline{X}(s) - y(s)|^2\right] ds$$

$$\leq 4C_R(T+4)\left[\mathbb{E}\int_0^{\tau \wedge \theta_R} |X(s) - \overline{X}(s)|^2 ds + \mathbb{E}\int_0^{\tau} |\overline{X}(s \wedge \theta_R) - y(s \wedge \theta_R)|^2 ds\right]$$

$$\leq 4C_R(T+4)\left[\mathbb{E}\int_0^{\tau \wedge \theta_R} |X(s) - \overline{X}(s)|^2 ds + \int_0^{\tau} \mathbb{E}\sup_{0 \leq r \leq s} |\overline{X}(r \wedge \theta_R) - y(r \wedge \theta_R)|^2 ds\right].$$

(2.9)

To bound the first term in the parentheses on the right-hand side of (2.9), given $s \in [0, T \wedge \theta_R)$, let k_s be the integer for which $s \in [t_{k_s}, t_{k_s+1})$. Then

$$X(s) - \overline{X}(s) = X_{k_s} - \left(X_{k_s} + \int_{t_{k_s}}^s f(X(s))ds + \int_{t_{k_s}}^s g(X(s))dW(s)\right)$$

$$= -f(X_{k_s})(s - t_{k_s}) - g(X_{k_s})(W(s) - W(t_{k_s})).$$

Hence,

(2.10) $|X(s) - \overline{X}(s)|^2 \leq 2\left[|f(X_{k_s})|^2\Delta t^2 + |g(X_{k_s})|^2|W(s) - W(t_{k_s})|^2\right].$

Now, from the local Lipschitz condition (2.4), for $|y| \leq R$ we have

$$|f(y)|^2 \leq 2\left(|f(y) - f(0)|^2 + |f(0)|^2\right) \leq 2\left(C_R|y|^2 + |f(0)|^2\right),$$

and, similarly,

$$|g(y)|^2 \leq 2\left(C_R|y|^2 + |g(0)|^2\right).$$

Hence, in (2.10),

$$|X(s) - \overline{X}(s)|^2 \leq 4(C_R|X_{k_s}|^2 + |f(0)|^2 \vee |g(0)|^2)(\Delta t^2 + |W(s) - W(t_{k_s})|^2).$$

Thus, using (2.5) and the Lyapunov inequality [12]

$$\mathbb{E}\int_0^{\tau \wedge \theta_R} |X(s) - \overline{X}(s)|^2 ds$$

$$\leq \mathbb{E}\int_0^{\tau \wedge \theta_R} 4(C_R|X_{k_s}|^2 + |f(0)|^2 \vee |g(0)|^2)(\Delta t^2 + |W(s) - W(t_{k_s})|^2)ds$$

$$\leq \int_0^{\tau} 4\mathbb{E}\left[C_R|X_{k_s}|^2 + |f(0)|^2 \vee |g(0)|^2)(\Delta t^2 + |W(s) - W(t_{k_s})|^2\right] ds$$

$$\leq \int_0^T 4(C_R\mathbb{E}[|X_{k_s}|^2] + |f(0)|^2 \vee |g(0)|^2)(\Delta t^2 + m\Delta t)ds$$

$$\leq 4T(C_R A^{2/p} + |f(0)|^2 \vee |g(0)|^2)\Delta t(\Delta t + m).$$

In (2.9) we then have

$$\mathbb{E}\left[\sup_{0 \leq t \leq \tau} \left|\overline{X}(t \wedge \theta_R) - y(t \wedge \theta_R)\right|^2\right]$$

$$\leq 16C_R(T+4)T\Delta t(\Delta t + m)(C_R A^{2/p} + |f(0)|^2 \vee |g(0)|^2)$$

$$+ 4C_R(T+4)\int_0^\tau \mathbb{E}\sup_{0 \leq r \leq s}\left[\left|\overline{X}(r \wedge \theta_R) - y(r \wedge \theta_R)\right|^2\right]ds.$$

Applying the Gronwall inequality [14] we obtain

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \left|\overline{X}(t \wedge \theta_R) - y(t \wedge \theta_R)\right|^2\right] \leq C\Delta t(C_R^2 + 1)e^{4C_R(T+4)},$$

where here, and in the following, C is a universal constant independent of Δt, R, and δ. Inserting this into (2.8) gives

$$(2.11) \quad \mathbb{E}\left[\sup_{0 \leq t \leq T} |e(t)|^2\right] \leq C\Delta t(C_R^2 + 1)e^{4C_R(T+4)} + \frac{2^{p+1}\delta A}{p} + \frac{(1-\frac{2}{p})2A}{\delta^{2/(p-2)}R^p}.$$

Given any $\epsilon > 0$, we can choose δ so that $(2^{p+1}\delta A)/p < \epsilon/3$, then choose R so that

$$\frac{(1-\frac{2}{p})2A}{\delta^{2/(p-2)}R^p} < \frac{\epsilon}{3},$$

and then choose Δt sufficiently small for

$$C\Delta t(C_R^2 + 1)e^{4C_R(T+4)} < \frac{\epsilon}{3}$$

so that, in (2.11), $\mathbb{E}[\sup_{0 \leq t \leq T} |e(t)|^2] < \epsilon$, as required.   □

We remark that the proof of Theorem 2.2 is optimal in the sense that in the globally Lipschitz case ($C_R \leq C$ for all R) we may take $\delta = \Delta t$ and $R = \Delta t^{-1/(p-2)}$ in (2.11) to recover the classical finite-time convergence result

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |e(t)|^2\right] = O(\Delta t),$$

found, for example, in [12, 14].

3. Convergence with a one-sided Lipschitz condition.

3.1. The one-sided Lipschitz condition. In this section we impose further assumptions on the SDE. In section 3.2 we show that these assumptions guarantee moment bounds for $y(t)$. Theorem 2.2 requires bounds on the pth moment of the exact *and* numerical solution—a condition that is difficult to verify in practice for the method (2.1) and indeed may fail to hold. In section 3.3 we introduce an implicit version of the EM method for which moment bounds, and hence a convergence result, can be obtained.

We make the following assumptions on the SDE.

ASSUMPTION 3.1. *The functions f and g in (1.1) are C^1, and there exist constants $\mu, c > 0$ such that*

$$(3.1) \qquad\qquad \langle a - b, f(a) - f(b)\rangle \leq \mu|a-b|^2 \quad \forall a, b \in \mathbb{R}^m,$$

$$(3.2) \qquad\qquad |g(a) - g(b)|^2 \leq c|a-b|^2 \quad \forall a, b \in \mathbb{R}^m.$$

Note that we work with the case $\mu > 0$. In the deterministic setting there is a lot of attention paid to the contractive case $\mu < 0$. This case is of less interest here because, for most diffusion coefficients g, contractivity is destroyed. Hence $\mu > 0$ is a natural assumption.

It follows from Assumption 3.1 that

$$\langle f(a), a \rangle = \langle f(a) - f(0), a \rangle + \langle f(0), a \rangle \leq \mu |a|^2 + |f(0)||a| \leq \tfrac{1}{2}|f(0)|^2 + (\mu + \tfrac{1}{2})|a|^2$$

and

$$|g(a)|^2 \leq 2|g(0)|^2 + 2|g(a) - g(0)|^2 \leq 2|g(0)|^2 + 2c|a|^2.$$

This gives

$$(3.3) \qquad \langle f(a), a \rangle \vee |g(a)|^2 \leq \alpha + \beta |a|^2 \quad \forall a \in \mathbb{R}^m,$$

where

$$(3.4) \qquad \alpha := \tfrac{1}{2}|f(0)|^2 \vee 2|g(0)|^2 \quad \text{and} \quad \beta := (\mu + \tfrac{1}{2}) \vee 2c.$$

The inequality (3.3) will prove very useful in what follows. We note that from [14, Theorem 2.3.5] $f, g \in C^1$ and (3.3) ensure the existence of a unique solution to the SDE (1.1).

The inequality (3.1) in Assumption 3.1, which is known as a *one-sided Lipschitz condition*, has been exploited successfully in the deterministic numerical analysis literature [1, 3, 4, 5, 17] and in the case of SDEs has been used in [11, 15, 16]. The condition (3.3) is closely related to the *monotone* condition in [14, section 2.4]. Any f of the form $f(y) = -y^p + y$, where the integer $p \geq 3$ is odd, satisfies (3.1), and further examples can be found in [17].

3.2. Moment bounds for the SDE. We now show that under Assumption 3.1 the SDE solution has a bounded pth moment for each $p > 2$.

LEMMA 3.2. *Under Assumption* 3.1, *for each $p > 2$ there is $C = C(p, T) > 0$ such that*

$$(3.5) \qquad \mathbb{E}\left[\sup_{0 \leq t \leq T} |y(t)|^p\right] \leq C(1 + \mathbb{E}|y_0|^p).$$

Proof. Theorem 2.4.1 of [14] shows that, for $p \geq 2$, there is $C = C(p, T)$ such that

$$(3.6) \qquad \mathbb{E}|y(t)|^p \leq C[1 + \mathbb{E}|y_0|^p] \quad \forall t \in [0, T].$$

By the Itô formula

$$|y(t)|^2 = |y_0|^2 + 2\int_0^t \langle f(y(s)), y(s) \rangle ds + \int_0^t |g(y(s))|^2 ds + 2\int_0^t \langle y(s), g(y(s)) dB(s) \rangle.$$

By (3.3) we have that, for some $K = K(p)$ and $t_1 \in [0, T]$,

$$\sup_{0 \leq t \leq t_1} |y(t)|^p \leq K\left(|y_0|^p + \left\{\int_0^{t_1}[\alpha + \beta|y(s)|^2]ds\right\}^{p/2}\right.$$
$$\left. + \sup_{0 \leq t \leq t_1}\left|\int_0^t \langle y(s), g(y(s)) dB(s)\rangle\right|^{p/2}\right).$$

By property (3.6) we can take expectations to give, for a possibly different $K = K(p, T)$,

$$\mathbb{E}\left[\sup_{0 \leq t \leq t_1} |y(t)|^p\right] \leq K \left(1 + \mathbb{E}|y_0|^p + \int_0^{t_1} \mathbb{E}|y(s)|^p ds \right.$$
$$\left. + \mathbb{E}\left[\sup_{0 \leq t \leq t_1}\left|\int_0^t \langle y(s), g(y(s))dB(s)\rangle\right|^{p/2}\right]\right).$$

By the Burkholder–Davis–Gundy inequality [14], we compute that, again for redefined $K = K(p, T)$,

$$\mathbb{E}\left[\sup_{0 \leq t \leq t_1} |y(t)|^p\right] \leq K \left(1 + E|y_0|^p + \int_0^{t_1} \mathbb{E}|y(s)|^p ds \right.$$
(3.7)
$$\left. + \mathbb{E}\left[\int_0^{t_1} |y(s)|^2|g(y(s))|^2 ds\right]^{p/4}\right).$$

Next, note that, by Cauchy–Schwarz,

$$\mathbb{E}\left[\int_0^{t_1} |y(s)|^2|g(y(s))|^2 ds\right]^{p/4} \leq \mathbb{E}\left[\sup_{0 \leq s \leq t_1} |y(s)|^{p/2}\left(\int_0^{t_1} |g(y(s))|^2 ds\right)^{p/4}\right]$$
$$\leq \frac{1}{2K}\mathbb{E}\left[\sup_{0 \leq s \leq t_1} |y(s)|^p\right] + \frac{K}{2}\mathbb{E}\left[\int_0^{t_1} |g(y(s))|^2 ds\right]^{p/2}$$
$$\leq \frac{1}{2K}\mathbb{E}\left[\sup_{0 \leq s \leq t_1} |y(s)|^p\right]$$
$$+ \frac{K}{2}T^{(p-2)/2}\mathbb{E}\int_0^{t_1} (\alpha + \beta|y(s)|^2)^{p/2} ds.$$

Substituting this into (3.7) yields, again for a possibly different $K = K(p, T)$,

$$\mathbb{E}\left[\sup_{0 \leq t \leq t_1} |y(t)|^p\right] \leq K \left(1 + \mathbb{E}|y_0|^p + \int_0^{t_1} \mathbb{E}|y(s)|^p ds\right).$$

The required assertion now follows from property (3.6). □

3.3. Split-step backward Euler. We now consider the split-step backward Euler (SSBE) method, which is defined by taking $Y_0 = y_0$ and, generally,

(3.8) $$Y_k^\star = Y_k + \Delta t f(Y_k^\star),$$
(3.9) $$Y_{k+1} = Y_k^\star + g(Y_k^\star)\Delta W_k.$$

We state our convergence theorem here and then give a sequence of results that lead to a proof.

THEOREM 3.3. *Consider the SSBE* (3.8)–(3.9) *applied to the SDE* (1.1) *under Assumption* 3.1. *There exists a continuous-time extension* $\overline{Y}(t)$ *of the numerical solution (so that* $\overline{Y}(t_k) = Y_k$) *for which*

$$\lim_{\Delta t \to 0} \mathbb{E}\left[\sup_{0 \leq t \leq T} |\overline{Y}(t) - y(t)|^2\right] = 0.$$

Proof. See the end of this subsection. □

Note that (3.8) is an implicit equation that must be solved in order to obtain the intermediate approximation $Y_k^\star$. Having obtained $Y_k^\star$, adding the appropriate stochastic increment $g(Y_k^\star)\Delta W_k$ produces the next approximation Y_{k+1} in (3.9). The SSBE method reduces to the deterministic backward Euler method [5, 9] when $g \equiv 0$ and y_0 is nonrandom. The method is also studied in [15], where it is effective for inheriting ergodicity; for related reasons it is effective here in enabling the derivation of moment bounds. Another stochastic extension of the deterministic backward Euler method is considered in section 5.

Our proof of Theorem 3.3 relies on showing that SSBE has two key properties under Assumption 3.1: (a) it may be regarded as EM applied to a modified SDE of a similar form, and (b) it produces solutions with all moments bounded. The first property is established in the next lemma and corollary.

LEMMA 3.4. *Let Assumption* 3.1 *hold and suppose* $\Delta t \in (0, \Delta t_c), \Delta t_c < 1/(2\beta)$, *where* β *is defined in* (3.4). *Given* $d \in \mathbb{R}^m$ *the implicit equation*

$$(3.10) \qquad\qquad c = d + \Delta t f(c)$$

has a unique solution c. If we define the functions $F_{\Delta t}(\cdot)$, $f_{\Delta t}(\cdot)$ *and* $g_{\Delta t}(\cdot)$ *by*

$$(3.11) \qquad F_{\Delta t}(d) = c, \quad f_{\Delta t}(d) = f(F_{\Delta t}(d)), \quad \text{and} \quad g_{\Delta t}(d) = g(F_{\Delta t}(d)),$$

then $F_{\Delta t}, f_{\Delta t}, g_{\Delta t} \in C^1$, $g_{\Delta t}(\cdot) \to g(\cdot)$ *and* $f_{\Delta t}(\cdot) \to f(\cdot)$ *as* $\Delta t \to 0$ *in* C^1 *uniformly on compact sets and, for any* $a, b \in \mathbb{R}^m$,

$$(3.12) \qquad\qquad |f_{\Delta t}(a)| \le \frac{|f(a)|}{1 - \Delta t\mu},$$

$$(3.13) \qquad\qquad |F_{\Delta t}(d) - F_{\Delta t}(e)|^2 \le \frac{1}{1 - 2\Delta t\mu}|d - e|^2,$$

$$(3.14) \qquad\qquad \langle a - b, f_{\Delta t}(a) - f_{\Delta t}(b)\rangle \le \frac{\mu}{1 - 2\mu\Delta t}|a - b|^2.$$

Further, $g_{\Delta t}$ *is globally Lipschitz, and there exist* $\alpha', \beta' > 0$ *such that*

$$(3.15) \qquad\qquad \langle f_{\Delta t}(a), a\rangle \vee |g_{\Delta t}(a)|^2 \le \alpha' + \beta'|a|^2 \quad \forall a \in \mathbb{R}^m.$$

Proof. See Appendix A. □

COROLLARY 3.5. *Let Assumption* 3.1 *hold and suppose* $\Delta t \in (0, \Delta t_c), \Delta t_c < 1/(2\beta)$, *where* β *is defined in* (3.4). *Then SSBE applied to* (1.1) *is equivalent to EM applied to the modified SDE*

$$(3.16) \quad dy_{\Delta t}(t) = f_{\Delta t}(y_{\Delta t}(t))dt + g_{\Delta t}(y_{\Delta t}(t))dW(t), \qquad 0 \le t \le T, \ y_{\Delta t}(0) = y_0,$$

where $f_{\Delta t}, g_{\Delta t}$ *are defined in Lemma* 3.4.

Proof. Lemma 3.4 allows us to express the SSBE method (3.8)–(3.9) in the form

$$(3.17) \qquad\qquad Y_{k+1} = Y_k + \Delta t f_{\Delta t}(Y_k) + g_{\Delta t}(Y_k)\Delta W_k,$$

and the result is then immediate. □

Next, we show that the solution of the modified SDE (3.16) has bounded moments and converges strongly to $y(t)$.

LEMMA 3.6. *Under Assumption* 3.1, *for each* $p > 2$, *there is* $C = C(p,T) > 0$ *such that*

$$\mathbb{E}\left[\sup_{0 \le t \le T} |y_{\Delta t}(t)|^p\right] \le C(1 + \mathbb{E}|y_0|^p), \tag{3.18}$$

provided Δt *is sufficiently small. In addition,*

$$\lim_{\Delta t \to 0} \mathbb{E}\left[\sup_{0 \le t \le T} |y(t) - y_{\Delta t}(t)|^2\right] = 0. \tag{3.19}$$

Proof. It follows from Lemma 3.4 that for sufficiently small Δt the functions $f_{\Delta t}$ and $g_{\Delta t}$ satisfy (3.3) with α and β replaced by 2α and 2β. Following through the proof of Lemma 3.2, which is based entirely on (3.3), we obtain (3.18).

Now to prove (3.19) we note from Lemma 3.4 that given $R > 0$ there is a function $K_R : (0, \infty) \to (0, \infty)$ such that $K_R(\Delta t) \to 0$ as $\Delta t \to 0$ and

$$|f_{\Delta t}(u) - f(u)|^2 \vee |g_{\Delta t}(u) - g(u)|^2 \le K_R(\Delta t) \quad \forall u \in \mathbb{R}^m, |u| \le R, \tag{3.20}$$

provided Δt is sufficiently small. Also, since $f, g \in C^1$, there is a constant H_R such that

$$|f(u) - f(v)|^2 \vee |g(u) - g(v)|^2 \le H_R|u - v|^2 \quad \forall u, v \in \mathbb{R}^m, |u| \vee |v| \le R. \tag{3.21}$$

From Lemma 3.2 and (3.18) we have

$$\mathbb{E}\left[\sup_{0 \le t \le T} |y(t)|^p\right] \vee \mathbb{E}\left[\sup_{0 \le t \le T} |y_{\Delta t}(t)|^p\right] \le K := C(1 + \mathbb{E}|y_0|^p). \tag{3.22}$$

The remainder of the proof follows in a similar manner to that of Theorem 2.2. Define

$$\tau_R = \inf\{t \ge 0 : |y(t)| \ge R\}, \quad \rho_R = \inf\{t \ge 0 : |y_{\Delta t}(t)| \ge R\}, \quad \theta_R = \tau_R \wedge \rho_R.$$

For any $\delta > 0$, in the same way that (2.8) was obtained, we have

$$\mathbb{E}\left[\sup_{0 \le t \le T} |y(t) - y_{\Delta t}(t)|^2\right] \le \mathbb{E}\left[\sup_{0 \le t \le T} |y(t \wedge \theta_R) - y_{\Delta t}(t \wedge \theta_R)|^2\right]$$
$$+ \frac{2^{p+1}\delta K}{p} + \frac{(p-2)2K}{p\delta^{2/(p-2)}R^p}. \tag{3.23}$$

To bound the first term on the right-hand side of (3.23), we observe that

$$y(t \wedge \theta_R) - y_{\Delta t}(t \wedge \theta_R)$$
$$= \int_0^{t \wedge \theta_R} [f(y(s)) - f(y_{\Delta t}(s)) + f(y_{\Delta t}(s)) - f_{\Delta t}(y_{\Delta t}(s))]ds$$
$$+ \int_0^{t \wedge \theta_R} [g(y(s)) - g(y_{\Delta t}(s)) + g(y_{\Delta t}(s)) - g_{\Delta t}(y_{\Delta t}(s))]dW(s).$$

Using (3.20), (3.21), Cauchy–Schwarz, and the Doob martingale inequality, we have that, for any $\tau \le T$,

$$\mathbb{E}\left[\sup_{0 \le t \le \tau} |y(t \wedge \theta_R) - y_{\Delta t}(t \wedge \theta_R)|^2\right]$$
$$\le 4H_R(T + 4)\int_0^\tau \mathbb{E}\left[\sup_{0 \le t \le s} |y(t \wedge \theta_R) - y_{\Delta t}(t \wedge \theta_R)|^2\right] ds$$
$$+ 4T(T + 4)K_R(\Delta t).$$

So the Gronwall inequality yields

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|y(t\wedge\theta_R)-y_{\Delta t}(t\wedge\theta_R)|^2\right]\leq 4T(T+4)K_R(\Delta t)e^{4H_R(T+4)T}.$$

Inserting this into (3.23) gives

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|y(t)-y_{\Delta t}(t)|^2\right]\leq 4T(T+4)K_R(\Delta t)e^{4H_R(T+4)T}$$

(3.24)
$$+\frac{2^{p+1}\delta K}{p}+\frac{(p-2)2K}{p\delta^{2/(p-2)}R^p}.$$

The final step of the proof follows that of Theorem 2.2. $\qquad\square$

Now we show that the special structure of SSBE makes it possible for us to bound all moments of the numerical solution, under Assumption 3.1. We deal first with the discrete approximation and then with a continuous-time extension.

LEMMA 3.7. *Suppose Assumption* 3.1 *holds and let* $\Delta t\leq\Delta t_c<1/(2\beta)$, *where* β *is defined in* (3.4). *Then for each* $p\geq 2$ *there exists a* $C=C(p,T)>0$ *(independent of* Δt*) such that for the SSBE method* (3.8)–(3.9)

$$\mathbb{E}\sup_{n\Delta t\in[0,T]}|Y_n|^{2p}\leq C.$$

Proof. In the following we assume that N and M are positive integers such that $N\Delta t\leq M\Delta t\leq T$. From (3.3) and (3.8) we have

$$|Y_n^\star|^2=\langle Y_n,Y_n^\star\rangle+\Delta t\langle f(Y_n^\star),Y_n^\star\rangle$$
$$\leq\tfrac{1}{2}|Y_n|^2+\tfrac{1}{2}|Y_n^\star|^2+\Delta t(\alpha+\beta|Y_n^\star|^2).$$

Thus

(3.25)
$$|Y_n^\star|^2\leq\frac{|Y_n|^2+2\alpha\Delta t}{1-2\beta\Delta t}.$$

From (3.9) and (3.25) we have

$$|Y_{n+1}|^2\leq|Y_n|^2+\frac{2\beta\Delta t}{1-2\beta\Delta t}|Y_n|^2+\frac{2\alpha\Delta t}{1-2\beta\Delta t}$$
$$+2\langle Y_n^\star,g(Y_n^\star)\Delta W_n\rangle+|g(Y_n^\star)\Delta W_n|^2.$$

Summing, and using the notation $K=(1-2\beta\Delta t)^{-1}$, we obtain

$$|Y_N|^2\leq|Y_0|^2+2\beta\Delta tK\sum_{j=0}^{N-1}|Y_j|^2+2\alpha\Delta tNK$$
$$+2\sum_{j=0}^{N-1}\langle Y_j^\star,g(Y_j^\star)\Delta W_j\rangle+\sum_{j=0}^{N-1}|g(Y_j^\star)\Delta W_j|^2.$$

Raising both sides to the power p we have

$$\frac{1}{5^{p-1}}|Y_N|^{2p} \leq |Y_0|^{2p} + (2\beta\Delta t K)^p \left(\sum_{j=0}^{N-1}|Y_j|^2\right)^p + (2\alpha TK)^p$$

$$+ 2^p \left|\sum_{j=0}^{N-1}\langle Y_j^\star, g(Y_j^\star)\Delta W_j\rangle\right|^p + \left(\sum_{j=0}^{N-1}|g(Y_j^\star)\Delta W_j|^2\right)^p$$

$$\leq |Y_0|^{2p} + (2\beta K)^p T^{p-1}\Delta t \sum_{j=0}^{N-1}|Y_j|^{2p} + (2\alpha TK)^p$$

$$(3.26) \qquad + 2^p \left|\sum_{j=0}^{N-1}\langle Y_j^\star, g(Y_j^\star)\Delta W_j\rangle\right|^p + N^{p-1}\sum_{j=0}^{N-1}|g(Y_j^\star)\Delta W_j|^{2p}.$$

Now

$$(3.27) \qquad \mathbb{E}\left[\sup_{0\leq N\leq M}\sum_{j=0}^{N-1}|Y_j|^{2p}\right] = \sum_{j=0}^{M-1}\mathbb{E}|Y_j|^{2p}.$$

Also, letting $C = C(p,T)$ be a constant that may change line by line,

$$\mathbb{E}\left[\sup_{0\leq N\leq M}\sum_{j=0}^{N-1}|g(Y_j^\star)\Delta W_j|^{2p}\right] = \mathbb{E}\sum_{j=0}^{M-1}|g(Y_j^\star)\Delta W_j|^{2p}$$

$$\leq \sum_{j=0}^{M-1}\mathbb{E}|g(Y_j^\star)|^{2p}\mathbb{E}|\Delta W_j|^{2p}$$

$$\leq C\Delta t^p \sum_{j=0}^{M-1}\mathbb{E}[\alpha + \beta|Y_j^\star|^2]^p$$

$$\leq C\Delta t^p \sum_{j=0}^{M-1}\mathbb{E}[\alpha^p + \beta^p|Y_j^\star|^{2p}]$$

$$\leq C\Delta t^{p-1} + C\Delta t^p \sum_{j=0}^{M-1}\mathbb{E}[|Y_j|^2 + 2\alpha\Delta t]]^p$$

$$(3.28) \qquad \leq C\Delta t^{p-1} + C\Delta t^p \sum_{j=0}^{M-1}\mathbb{E}|Y_j|^{2p},$$

where we have used (3.3) and (3.25). Finally, using the Burkholder–Davis–Gundy inequality [14],

$$\mathbb{E}\left[\sup_{0\leq N\leq M}\left|\sum_{j=0}^{N-1}\langle Y_j^\star, g(Y_j^\star)\Delta W_j\rangle\right|^p\right] \leq C\mathbb{E}\left[\sum_{j=0}^{M-1}|Y_j^\star|^2|g(Y_j^\star)|^2\Delta t\right]^{p/2}$$

$$\leq C(\Delta t)^{p/2}M^{p/2-1}\mathbb{E}\sum_{j=0}^{M-1}|Y_j^\star|^p(\alpha + \beta|Y_j^\star|^2)^{p/2}$$

$$\leq C\Delta t \sum_{j=0}^{M-1}[1 + \mathbb{E}|Y_j^\star|^{2p}]$$

$$\leq C\Delta t \sum_{j=0}^{M-1} [1 + \mathbb{E}(2\alpha\Delta t + |Y_j|^2)^p]$$

(3.29)
$$\leq C + C\Delta t \sum_{j=0}^{M-1} \mathbb{E}|Y_j|^{2p}.$$

Combining (3.26)–(3.29) we obtain

$$\mathbb{E}\left[\sup_{0\leq N\leq M} |Y_N|^{2p}\right] \leq C + C\Delta t \sum_{j=0}^{M-1} \mathbb{E}|Y_j|^{2p} \leq C + C\Delta t \sum_{j=0}^{M-1} \mathbb{E}\left[\sup_{0\leq N\leq j} \mathbb{E}|Y_N|^{2p}\right].$$

Using the discrete-type Gronwall inequality (see, for example, [13]) and noting that $M\Delta t \leq T$, we obtain

$$\mathbb{E}\left[\sup_{0\leq N\leq M} |Y_N|^{2p}\right] \leq Ce^{C\Delta t M} \leq Ce^{CT},$$

and the desired result follows.   □

COROLLARY 3.8. *Suppose Assumption* 3.1 *holds and let* $\Delta t \in (0, \Delta t_c), \Delta t_c < 1/(2\beta)$, *where* β *is defined in* (3.4). *Let* $p \geq 2$. *Then there exists a continuous-time extension* $\overline{Y}(t)$ *of the SSBE solution* $\{Y_k\}$ *and a constant* $C = C(p, T) > 0$ *(independent of* Δt*) such that*

$$\mathbb{E}\sup_{0\leq t\leq T} |\overline{Y}(t)|^{2p} \leq C.$$

Proof. We know that SSBE can be regarded as EM applied to the modified SDE (3.16). Hence, we may define $\overline{Y}(t)$ using (2.2)–(2.3) with f, g replaced by $f_{\Delta t}, g_{\Delta t}$ and $X, \overline{X}, X_k$ replaced by $Y, \overline{Y}, Y_k$. By definition we have, for $t_n = n\Delta t$,

$$\overline{Y}(t_n + s) = Y_n + sf_{\Delta t}(Y_n) + g_{\Delta t}(Y_n)\Delta W_n(s), \quad s \in [0, \Delta t),$$

where

$$\Delta W_n(s) := W(t_n + s) - W(t_n).$$

However, $Y_n^\star = Y_n + \Delta t f_{\Delta t}(Y_n)$ and so, for $a = s/\Delta t$, we have

$$\overline{Y}(t_n + s) = aY_n^\star + (1-a)Y_n + g_{\Delta t}(Y_n)\Delta W_n(s), \qquad s \in [0, \Delta t).$$

Since $\Delta t \leq \Delta t_c < 1/(2\beta)$, it follows from (3.25) that

$$|\overline{Y}(t_n + s)|^2 \leq C[1 + |Y_n|^2 + |g_{\Delta t}(Y_n)\Delta W_n(s)|^2].$$

Thus

$$\sup_{0\leq t\leq T} |\overline{Y}(t)|^{2p} \leq \sup_{0\leq n\Delta t\leq T} \sup_{0\leq s\leq \Delta t} |\overline{Y}(t_n + s)|^{2p}$$

$$\leq \sup_{0\leq n\Delta t\leq T} \sup_{0\leq s\leq \Delta t} C[1 + |Y_n|^{2p} + |g_{\Delta t}(Y_n)\Delta W_n(s)|^{2p}]$$

(3.30)
$$\leq C[1 + \sup_{0\leq n\Delta t\leq T} |Y_n|^{2p} + \sup_{0\leq s\leq \Delta t} \sum_{j=0}^{N} |g_{\Delta t}(Y_j)\Delta W_j(s)|^{2p}],$$

where $0 \leq N\Delta t \leq T$. Now, using Doob's martingale inequality [14] and (3.15)

$$\mathbb{E} \sup_{0 \leq s \leq \Delta t} |g_{\Delta t}(Y_j) \Delta W_j(s)|^{2p} \leq C\mathbb{E}|g_{\Delta t}(Y_j) \Delta W_j(\Delta t)|^{2p}$$

$$\leq C\mathbb{E}|g_{\Delta t}(Y_j)|^{2p} \mathbb{E}|\Delta W_j(\Delta t)|^{2p}$$

$$\leq C\Delta t^p [1 + \mathbb{E}|Y_j|^{2p}]$$

(3.31) $$\leq C\Delta t,$$

where C is a universal constant, independent of Δt. Since $N\Delta t \leq T$, combining Lemma 3.7, (3.30), and (3.31) gives the desired result.   □

Proof of Theorem 3.3. The proof now follows from an application of the triangle inequality: the SSBE method has a solution close to the solution of an SDE with modified vector fields, and the solution of this SDE in turn is close to that of the original SDE. More precisely, we may use Corollary 3.8 to define $\overline{Y}(t)$ and bound $\mathbb{E} \sup_{0 \leq t \leq T} |\overline{Y}(t)|^p$ and Lemma 3.6 to bound $\mathbb{E} \sup_{0 \leq t \leq T} |y_{\Delta t}(t)|^p$. We also know from Lemma 3.4 that $f_{\Delta t}$ and $g_{\Delta t}$ are uniformly locally Lipschitz for small Δt. Hence, we may follow the proof of Theorem 2.2 to give

$$\lim_{\Delta t \to 0} \mathbb{E} \left[\sup_{0 \leq t \leq T} |\overline{Y}(t) - y_{\Delta t}(t)|^2 \right] = 0.$$

Combining this with (3.19) in Lemma 3.6 via the triangle inequality gives the result.   □

4. Convergence rates. In this section we show that by augmenting Assumption 3.1 with the condition that f behaves polynomially, it is possible to establish a rate of convergence. The rate is optimal, agreeing with the standard theory for the explicit EM scheme in the globally Lipschitz case. The work of [7] also yields optimal rates of almost sure convergence for the EM scheme, under conditions on the vector fields similar to ours.

ASSUMPTION 4.1. *There exist constants $D \in \mathbb{R}^+$ and $q \in \mathbb{Z}^+$ such that for all $a, b \in \mathbb{R}^m$,*

(4.1) $$|f(a) - f(b)|^2 \leq D \left(1 + |a|^q + |b|^q\right) |a - b|^2.$$

To obtain a convergence rate for EM we require the following moment bound assumption.

ASSUMPTION 4.2. *The SDE and EM solutions satisfy*

$$\mathbb{E} \sup_{0 \leq t \leq T} |y(t)|^p, \quad \mathbb{E} \sup_{0 \leq t \leq T} |X(t)|^p, \quad \mathbb{E} \sup_{0 \leq t \leq T} |\overline{X}(t)|^p < \infty \qquad \forall p \geq 1.$$

Throughout the following analysis, K and u denote generic positive real and integer constants whose values may change between occurrences. Before obtaining a convergence rate for EM, we give the following lemma.

LEMMA 4.3. *Under Assumptions* 3.1, 4.1, *and* 4.2, *for any even integer $r \geq 2$, there exists a constant $E = E(r)$ such that*

$$\sup_{0 \leq t \leq T} \mathbb{E}|X(t) - \overline{X}(t)|^r \leq E\Delta t^{r/2}.$$

Proof. Let $t \in [k\Delta t, (k+1)\Delta t)$. Then

$$
\begin{aligned}
|X(t) - \overline{X}(t)|^r &= |(t - t_k)f(X_k) + g(X_k)(W(t) - W(t_k))|^r \\
&\leq 2^r \left(\Delta t^r |f(X_k)|^r + |g(X_k)|^r |W(t) - W(t_k)|^r \right).
\end{aligned}
$$

Hence, for some $E = E(r)$,

$$
\mathbb{E}|X(t) - \overline{X}(t)|^r \leq E \left(\Delta t^r \left[1 + \mathbb{E} \sup_{0 \leq t \leq T} |X(t)|^u \right] + \left[1 + \mathbb{E} \sup_{0 \leq t \leq T} |X(t)|^u \right] (t - t_k)^{r/2} \right).
$$

Since $t - t_k \leq \Delta t$, the result follows by redefinition of E. □

THEOREM 4.4. *Under Assumptions* 3.1, 4.1, *and* 4.2 *the EM solution* (2.1) *with continuous-time extension* (2.2) *satisfies*

$$
\mathbb{E} \left[\sup_{0 \leq t \leq T} |\overline{X}(t) - y(t)|^2 \right] = O(\Delta t).
$$

Proof. Using (2.2) and

$$
(4.2) \qquad y(t) = y_0 + \int_0^t f(y(s))ds + \int_0^t g(y(s))dW(s),
$$

and letting $e(t) := y(t) - \overline{X}(t)$, the Itô formula gives

$$
\begin{aligned}
|e(t)|^2 &= \int_0^t 2\langle f(y(s)) - f(X(s)), e(s)\rangle ds + \int_0^t |g(y(s)) - g(X(s))|^2 ds \\
&\quad + M(t) \\
&\leq \int_0^t \left(2\langle f(y(s)) - f(\overline{X}(s)), e(s)\rangle + c|y(s) - X(s)|^2 \right) ds \\
&\quad + \int_0^t 2\langle f(\overline{X}(s)) - f(X(s)), e(s)\rangle ds \\
&\quad + M(t) \\
&\leq \int_0^t 2\mu |e(s)|^2 + 2c|e(s)|^2 + 2c|X(s) - \overline{X}(s)|^2 ds \\
&\quad + \int_0^t |f(\overline{X}(s)) - f(X(s))|^2 + |e(s)|^2 ds \\
&\quad + M(t) \\
&\leq (1 + 2(\mu + c)) \int_0^t |e(s)|^2 ds + \int_0^t K \left(1 + |\overline{X}(s)|^q + |X(s)|^q \right) |X(s) - \overline{X}(s)|^2 ds \\
&\quad + M(t),
\end{aligned}
$$

where

$$
M(t) = \int_0^t 2\langle e(s), (g(y(s)) - g(X(s)))dW(s)\rangle.
$$

Using Cauchy–Schwarz and Lemma 4.3 with $r = 4$

$$\mathbb{E}\left[\sup_{0\leq s\leq t}|e(s)|^2\right] \leq (1+2(\mu+c))\int_0^t \mathbb{E}|e(s)|^2 ds$$

$$+ \int_0^t K\left(\mathbb{E}\left(1+|\overline{X}(s)|^q+|X(s)|^q\right)^2 \mathbb{E}|X(s)-\overline{X}(s)|^4\right)^{1/2} ds + m(t)$$

$$\leq (1+2(\mu+c))\int_0^t \mathbb{E}|e(s)|^2 ds + K\Delta t \int_0^t \mathbb{E}\left(1+|\overline{X}(s)|^{2q}+|X(s)|^{2q}\right) ds$$

$$+ m(t)$$

$$(4.3) \qquad \leq (1+2(\mu+c))\int_0^t \mathbb{E}|e(s)|^2 ds + K\Delta t + m(t),$$

where

$$m(t) = \mathbb{E}\left[\sup_{0\leq s\leq t}|M(s)|\right].$$

From the Burkholder–Davis–Gundy inequality,

$$m(t) \leq 16\mathbb{E}\left[\int_0^t |e(s)|^2|g(y(s))-g(X(s))|^2 ds\right]^{1/2}$$

$$\leq 16\mathbb{E}\left[\sup_{0\leq s\leq t}|e(s)|^2 \int_0^t c|y(s)-X(s)|^2 ds\right]^{1/2}$$

$$\leq \tfrac{1}{2}\mathbb{E}\left[\sup_{0\leq s\leq t}|e(s)|^2\right] + 128c\mathbb{E}\int_0^t |y(s)-X(s)|^2 ds$$

$$\leq \tfrac{1}{2}\mathbb{E}\left[\sup_{0\leq s\leq t}|e(s)|^2\right] + 256c\int_0^t [\mathbb{E}|e(s)|^2 + \mathbb{E}|\overline{X}(s)-X(s)|^2] ds$$

$$\leq \tfrac{1}{2}\mathbb{E}\left[\sup_{0\leq s\leq t}|e(s)|^2\right] + 256c\int_0^t \mathbb{E}|e(s)|^2 ds + K\Delta t.$$

Hence, in (4.3),

$$\mathbb{E}\left[\sup_{0\leq s\leq t}|e(s)|^2\right] \leq 2(1+2(\mu+c)+256c)\int_0^t \mathbb{E}|e(s)|^2 ds + K\Delta t$$

$$\leq 2(1+2(\mu+c)+256c)\int_0^t \mathbb{E}\left[\sup_{0\leq r\leq s}|e(r)|^2\right] ds + K\Delta t.$$

The result follows from the Gronwall inequality. $\qquad \square$

Note that Theorem 4.4 requires moment bounds on the numerical solution (Assumption 4.2). We know that SSBE has bounded moments under Assumption 3.1, and hence we would expect to get an analogous convergence result for this method without requiring Assumption 4.2. To obtain such a result we first establish further properties of $f_{\Delta t}$ and $g_{\Delta t}$ under Assumptions 3.1 and 4.1.

LEMMA 4.5. *Under Assumptions 3.1 and 4.1, for $\Delta t \leq \Delta t_c < 1/(2\beta)$, where β is defined in (3.4), there exist constants $c', D' \in \mathbb{R}^+$ and $q' \in \mathbb{Z}^+$ such that for all $a, b \in \mathbb{R}^m$*

$$(4.4) \qquad |f_{\Delta t}(a)-f_{\Delta t}(b)|^2 \leq D'\left(1+|a|^q+|b|^q\right)|a-b|^2,$$

$$(4.5) \qquad |f(a)-f_{\Delta t}(a)|^2 \leq c'\left(1+|a|^{q'}\right)\Delta t^2,$$

$$(4.6) \qquad |g(a)-g_{\Delta t}(a)|^2 \leq c'\left(1+|a|^{q'}\right)\Delta t^2.$$

Proof. See Appendix A.   □

LEMMA 4.6. *Under Assumptions* 3.1 *and* 4.1 *the solution* $y_{\Delta t}(t)$ *of the modified SDE* (3.16) *satisfies*

$$\mathbb{E}\left[\sup_{0 \le t \le T} |y_{\Delta t}(t) - y(t)|^2\right] = O(\Delta t^2).$$

Proof. Using

$$y_{\Delta t}(t) = y_0 + \int_0^t f_{\Delta t}(y_{\Delta t}(s))ds + \int_0^t g_{\Delta t}(y_{\Delta t}(s))dW(s),$$

and (4.2), and letting $e(t) := y(t) - y_{\Delta t}(t)$, the Itô formula gives

$$
\begin{aligned}
|e(t)|^2 &= \int_0^t 2\langle f(y(s)) - f_{\Delta t}(y_{\Delta t}(s)), e(s)\rangle ds + \int_0^t |g(y(s)) - g_{\Delta t}(y_{\Delta t}(s))|^2 ds \\
&\quad + M(t) \\
&= \int_0^t 2\langle f(y(s)) - f(y_{\Delta t}(s)) + f(y_{\Delta t}(s)) - f_{\Delta t}(y_{\Delta t}(s)), e(s)\rangle ds \\
&\quad + \int_0^t |g(y(s)) - g_{\Delta t}(y_{\Delta t}(s))|^2 ds \\
&\quad + M(t) \\
&\le \int_0^t [2\mu|e(s)|^2 + |f(y_{\Delta t}(s)) - f_{\Delta t}(y_{\Delta t}(s))|^2 + |e(s)|^2]ds \\
&\quad + 2\int_0^t |g(y(s)) - g(y_{\Delta t}(s))|^2 + |g(y_{\Delta t}(s)) - g_{\Delta t}(y_{\Delta t}(s))|^2 ds \\
&\quad + M(t) \\
&\le K\int_0^t |e(s)|^2\, ds + K\Delta t^2 \int_0^t \left(1 + |y_{\Delta t}(s)|^{q'}\right)\, ds \\
&\quad + M(t),
\end{aligned}
$$

where we used Lemma 4.5 and

$$M(t) = \int_0^t 2\langle e(s), (g(y(s)) - g_{\Delta t}(y_{\Delta t}(s)))dW(s)\rangle.$$

Hence,

$$\mathbb{E}\left[\sup_{0 \le s \le t} |e(s)|^2\right] \le K\int_0^t \mathbb{E}|e(s)|^2\, ds + K\Delta t^2 \int_0^t \mathbb{E}\left[1 + |y_{\Delta t}(s)|^{q'}\right]\, ds + m(t),$$

where

$$m(t) = \mathbb{E}\left[\sup_{0 \le s \le t} |M(s)|\right].$$

Since $y_{\Delta t}(s)$ has bounded moments, we have

(4.7) $$\mathbb{E}\left[\sup_{0 \le s \le t} |e(s)|^2\right] \le K\int_0^t \mathbb{E}|e(s)|^2\, ds + K\Delta t^2 + m(t).$$

However, in the same way as in the proof of Theorem 4.4, we can show

$$m(t) \leq \tfrac{1}{2}\mathbb{E}\left[\sup_{0 \leq s \leq t}|e(s)|^2\right] + 128\mathbb{E}\int_0^t |g(y(s)) - g_{\Delta t}(y_{\Delta t}(s))|^2 ds,$$

while

$$\mathbb{E}\int_0^t |g(y(s)) - g_{\Delta t}(y_{\Delta t}(s))|^2 ds \leq 2\mathbb{E}\int_0^t |g(y(s)) - g(y_{\Delta t}(s))|^2$$
$$+ |g(y_{\Delta t}(s)) - g_{\Delta t}(y_{\Delta t}(s))|^2 ds$$
$$\leq K\int_0^t \mathbb{E}|e(s)|^2\, ds + K\Delta t^2.$$

In (4.7) we therefore have

$$\mathbb{E}\left[\sup_{0 \leq s \leq t}|e(s)|^2\right] \leq K\int_0^t \mathbb{E}|e(s)|^2\, ds + K\Delta t^2$$
$$\leq K\int_0^t \mathbb{E}\left[\sup_{0 \leq r \leq s}|e(r)|^2\right]\, ds + K\Delta t^2.$$

The result follows from the Gronwall inequality. □

We may now prove a convergence result for SSBE.

THEOREM 4.7. *Consider the SSBE method* (3.8)–(3.9) *applied to the SDE* (1.1) *under Assumptions* 3.1 *and* 4.1. *There exists a continuous-time extension* $\overline{Y}(t)$ *of the numerical solution (so that* $\overline{Y}(t_k) = Y_k$) *for which*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T}|\overline{Y}(t) - y(t)|^2\right] = O(\Delta t).$$

Proof. We know that SSBE can be regarded as EM applied to the modified SDE (3.16). Lemmas 3.6 and 3.7 and Corollary 3.8 show that $y_{\Delta t}(t)$, $Y(t)$, and $\overline{Y}(t)$ have bounded moments. Hence, copying the proof of Theorem 4.4 we may conclude that

$$\mathbb{E}\left[\sup_{0 \leq t \leq T}|\overline{Y}(t) - y_{\Delta t}(t)|^2\right] = O(\Delta t).$$

Combining this with Lemma 4.6 via the triangle inequality gives the required result. □

5. Backward Euler. The SSBE method (3.8)–(3.9) is a stochastic extension of the deterministic backward Euler method. Another, perhaps more natural, extension of backward Euler is given by $Z_0 = y_0$ and

(5.1) $$Z_{k+1} = Z_k + \Delta t f(Z_{k+1}) + g(Z_k)\Delta W_k.$$

Indeed, this implicit method has appeared frequently in the literature—it is a member of the *family of implicit Euler schemes* [12, section 12.2] or the stochastic theta method class [10] and is sometimes called the semi-implicit Euler method [2]. We will refer to the method (5.1) as simply the backward Euler (BE) method for (1.1). As mentioned at the end of section 1, our convergence result, Theorem 5.3 below, is closely related to that of [11, Theorem 2.4].

The BE method (5.1) requires an implicit equation to be solved. Under Assumption 3.1, the homotopy argument in the proof of Lemma 3.4 shows that for $2\mu\Delta t < 1$ a unique solution exists with probability one. The next lemma points out a useful connection between BE and SSBE.

LEMMA 5.1. *Let $\{Y_k\}$ and $\{Z_k\}$ denote the SSBE and BE solutions, given by* (3.8)–(3.9) *and* (5.1), *respectively. Under Assumption* 3.1, *if $Y_0 = Z_0 - \Delta t f(Z_0)$, then*

$$(5.2) \qquad Z_k = Y_k + \Delta t f_{\Delta t}(Y_k) \qquad \forall k \geq 0.$$

Proof. Let $Q_k^\star = Z_k$ and $Q_k = Z_k - \Delta t f(Z_k)$, where $\{Z_k\}$ is the BE solution (5.1). Then

$$Q_k^\star = Q_k + \Delta t f(Q_k^\star)$$

and, using (5.1),

$$Q_{k+1} = Z_{k+1} - \Delta t f(Z_{k+1}) = Q_k^\star + g(Q_k^\star)\Delta W_k.$$

Hence, $\{Q_k\}$ is precisely the SSBE solution. This gives $Y_k = Z_k - \Delta t f(Z_k)$. The relation (5.2) then follows immediately from Lemma 3.4. $\square$

Lemma 5.1 shows that the BE solution can be regarded as an $O(\Delta t)$ perturbation of the SSBE solution. We may use this relation between BE and SSBE in order to obtain a convergence result for BE via Theorem 3.3. We first deal with the perturbation to the initial data.

LEMMA 5.2. *Under Assumptions* 3.1 *and* 4.1, *if $y(t)$ and $z(t)$ are solutions of the SDE* (1.1) *with initial conditions such that*

$$\mathbb{E}|y(0)|^p, \ \mathbb{E}|z(0)|^p \leq \infty \qquad \forall p \geq 1,$$

then, for some constant M,

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

Proof. Letting $e(t) := y(t) - z(t)$ and applying the Itô formula to $|e(t)|^2$, the inequality can be obtained by following the process used in the proofs of Theorem 4.4 and Lemma 4.6. $\square$

THEOREM 5.3. *Consider the BE method* (5.1) *applied to the SDE* (1.1) *under Assumptions* 3.1 *and* 4.1. *There exists a continuous-time extension $\overline{Z}(t)$ of the numerical solution (so that $\overline{Z}(t_k) = Z_k$) for which*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |\overline{Z}(t) - y(t)|^2\right] = O(\Delta t).$$

Proof. Let $\overline{Y}(t)$ denote the continuous-time extension to SSBE defined in Theorem 3.3, with initial data $\overline{Y}(0) = y_0 - \Delta t f(y_0)$. Also, let $\overline{Z}(t) = \overline{Y}(t) + \Delta t f_{\Delta t}(\overline{Y}(t))$, so that, from Lemma 5.1, $\overline{Z}(t)$ is a continuous-time extension to the BE solution with $\overline{Z}(0) = y_0$. We let $\widehat{y}_{\Delta t}(t)$ denote the solution to (1.1), with initial data $y_{\Delta t}(0) = y_0 - \Delta t f(y_0)$.

From Lemma 5.2 we have

$$(5.3) \qquad \mathbb{E}\left[\sup_{0 \leq t \leq T} |y(t) - \widehat{y}_{\Delta t}(t)|^2\right] \leq M\Delta t^2 \mathbb{E}|f(y_0)|^2 = O(\Delta t^2).$$

Also, the SSBE convergence result in Theorem 4.7 shows that

$$(5.4) \qquad \mathbb{E}\left[\sup_{0\le t\le T}|\widehat{y}_{\Delta t}(t) - \overline{Y}(t)|^2\right] = O(\Delta t).$$

Further,

$$(5.5) \qquad \mathbb{E}\left[\sup_{0\le t\le T}|\overline{Y}(t) - \overline{Z}(t)|^2\right] \le \Delta t^2 \mathbb{E}\left[\sup_{0\le t\le T}|f_{\Delta t}(\overline{Y}(t))|^2\right] = O(\Delta t^2),$$

because, by Lemma 4.5, $f_{\Delta t}$ is polynomially bounded. Combining (5.3)–(5.5) completes the proof.   □

6. Summary. Our aim in this work was to extend strong mean square convergence theory for numerical SDE simulations beyond the realm of globally Lipschitz problems. The only previous published work in this area that we are aware of is [11]. We gave a strong convergence theorem for EM in the case where the vector fields are locally Lipschitz (e.g., C^1) and moment bounds are available. This style of analysis is useful whenever moment bounds can be established, both for the EM method and for other methods that can be shown to be "close" to EM. In general, it is not clear when such moment bounds can be expected to hold for explicit methods with $f, g \in C^1$. However, for an implicit variant of EM, we obtained bounds on all moments in the case where the diffusion coefficient is globally Lipschitz but the drift coefficient satisfies only a one-sided Lipschitz condition. Then, by interpreting the implicit method as EM applied to a modified SDE we were able to get a strong convergence result. We then considered the case where it is further assumed that f behaves like a polynomial. If all moments are bounded, then EM can be shown to converge strongly at the optimal rate, again assuming moment bounds. Moment bounds can be established for two different implicit variants of EM, allowing us to show that these implicit methods converge at the optimal rate. One of these convergence results is comparable to the main result in [11]—we use a stronger error measure but require a more restrictive assumption on the growth of the drift coefficient.

The methods of analysis could be extended to other implicit methods, such as the stochastic theta method with $\theta \in [1/2, 1]$. Such schemes, especially their split-step variants, may be of practical interest for Hamiltonian problems perturbed by damping and/or noise in the case $\theta = 1/2$.

Appendix A. Proofs of Lemmas 3.4 and 4.5.

Proof of Lemma 3.4. Existence and uniqueness for (3.10) can be proved via a contraction mapping theorem, which also establishes the C^1 smoothness of $f_{\Delta t}$ and $F_{\Delta t}$ and the convergence property of $f_{\Delta t}$; see [5, 17]. The smoothness and convergence properties of $g_{\Delta t}$ follow from $g_{\Delta t}(\cdot) = g(F_{\Delta t}(\cdot))$.

An alternative proof of uniqueness for (3.10) via a homotopy argument is given in [9, Theorem 14.2]. We repeat the homotopy construction here, as it will be used to obtain the bound (3.12).

Suppose $h = h(\tau)$ satisfies

$$(A.1) \qquad h = d + \Delta t f(h) + (\tau - 1)\Delta t f(d),$$

where τ is our homotopy parameter. For $\tau = 1$, h solves (3.10). For $\tau = 0$ we have

$$h - d = \Delta t \left(f(h) - f(d)\right)$$

and so, using Assumption 3.1,

$$|h - d|^2 = \Delta t \langle h - d, f(h) - f(d) \rangle \leq \Delta t \mu |h - d|^2.$$

Note that $\beta > \mu$ so that $2\Delta t \mu < 1$. It follows that $h = d$ is the unique solution to (A.1) when $\tau = 0$. Differentiating (A.1) with respect to τ gives

$$\dot{h} = \Delta t \frac{\partial f}{\partial y}(h)\dot{h} + \Delta t f(d).$$

So

$$(A.2) \qquad |\dot{h}|^2 - \Delta t \left\langle \dot{h}, \frac{\partial f}{\partial y}(h)\dot{h} \right\rangle = \Delta t \langle \dot{h}, f(d) \rangle.$$

Setting $a - b = \epsilon u$ in (3.1) and letting $\epsilon \to 0$, we see that

$$\left\langle u, \frac{\partial f}{\partial y}(b)u \right\rangle \leq \mu |u|^2 \quad \text{for any } u, b \in \mathbb{R}^m.$$

Hence, in (A.2),

$$|\dot{h}|^2 - \Delta t \mu |\dot{h}|^2 \leq \Delta t |\dot{h}| \, |f(d)|.$$

So

$$|\dot{h}| \leq \Delta t \frac{|f(d)|}{1 - \Delta t \mu}.$$

It follows that $h(\tau)$ exists uniquely for all $\tau > 0$ and

$$|h(1) - d| = \left| \int_0^1 \dot{h}(s) \, ds \right| \leq \Delta t \frac{|f(d)|}{1 - \Delta t \mu},$$

which establishes (3.12).

To obtain (3.13) we note that if $c^{(1)} = d^{(1)} + \Delta t f(c^{(1)})$ and $c^{(2)} = d^{(2)} + \Delta t f(c^{(2)})$, then

$$|c^{(1)} - c^{(2)}|^2 - \Delta t \langle f(c^{(1)}) - f(c^{(2)}), c^{(1)} - c^{(2)} \rangle = \langle d^{(1)} - d^{(2)}, c^{(1)} - c^{(2)} \rangle$$

and so, using Assumption 3.1,

$$(1 - \Delta t \mu)|c^{(1)} - c^{(2)}|^2 \leq \frac{1}{2}|d^{(1)} - d^{(2)}|^2 + \frac{1}{2}|c^{(1)} - c^{(2)}|^2,$$

which gives (3.13).

Next, note from the implicit definition (3.11) that $f_{\Delta t}(a)$ is equivalent to $f(a + \Delta t f_{\Delta t}(a))$. Using (3.1) we thus have

$$\langle f_{\Delta t}(a) - f_{\Delta t}(b), a + \Delta t f_{\Delta t}(a) - b - \Delta t f_{\Delta t}(b) \rangle \leq \mu |a + \Delta t f_{\Delta t}(a) - b - \Delta t f_{\Delta t}(b)|^2.$$

Hence,

$$\langle f_{\Delta t}(a) - f_{\Delta t}(b), a - b \rangle + \Delta t |f_{\Delta t}(a) - f_{\Delta t}(b)|^2 \le \mu |a - b|^2 + \; 2\mu \langle a - b, f_{\Delta t}(a) - f_{\Delta t}(b) \rangle \Delta t$$
$$+ \; \mu \Delta t^2 |f_{\Delta t}(a) - f_{\Delta t}(b)|^2,$$

and (3.14) follows.

The global Lipschitz property of $g_{\Delta t}$ follows from (3.13).

Finally, we use $f_{\Delta t}(a) = f(a + \Delta t f_{\Delta t}(a))$ and (3.3) to give

$$\langle f_{\Delta t}(a), a + \Delta t f_{\Delta t}(a) \rangle \le \alpha + \beta |a + \Delta t f_{\Delta t}(a)|^2.$$

Hence

$$(1 - 2\beta \Delta t)\langle f_{\Delta t}(a), a \rangle \le \alpha + \beta |a|^2 + [\beta \Delta t^2 - \Delta t]|f_{\Delta t}(a)|^2 \le \alpha + \beta |a|^2.$$

Since $g_{\Delta t}$ is globally Lipschitz, the inequality (3.15) follows. $\quad\square$

Proof of Lemma 4.5. Recall from Lemma 3.4 that $f_{\Delta t}(a) := f(F_{\Delta t}(a))$, where $F_{\Delta t}$ is globally Lipschitz. Hence,

$$|f_{\Delta t}(a) - f_{\Delta t}(b)|^2 = |f(F_{\Delta t}(a)) - f(F_{\Delta t}(b))|^2$$
$$\le D \left(1 + |F_{\Delta t}(a)|^q + |F_{\Delta t}(b)|^q\right) |F_{\Delta t}(a) - F_{\Delta t}(b)|^2$$
$$\le D' \left(1 + |a|^q + |b|^q\right) |a - b|^2.$$

Next, it follows from the equivalence of $f_{\Delta t}(a)$ in (3.11) and $f(a + \Delta t f_{\Delta t}(a))$ that

$$|f(a) - f_{\Delta t}(a)|^2 \le D \left(1 + |a|^q + |a + \Delta t f_{\Delta t}(a)|^q\right) \Delta t^2 |f_{\Delta t}(a)|^2.$$

From (3.12), $|f_{\Delta t}(a)| \le 2|f(a)|$, and hence we obtain (4.5). A similar argument gives (4.6). $\quad\square$

REFERENCES

[1] K. Burrage and J. C. Butcher, *Stability criteria for implicit Runge–Kutta methods*, SIAM J. Numer. Anal., 16 (1979), pp. 46–57.

[2] K. Burrage and T. Tian, *A note on the stability properties of the Euler methods for solving stochastic differential equations*, New Zealand J. Math., 29 (2000), pp. 115–127.

[3] J. C. Butcher, *A stability property of implicit Runge–Kutta methods*, BIT, 15 (1975), pp. 358–361.

[4] G. Dahlquist, *Error analysis for a class of methods for stiff non-linear initial value problems*, in Numerical Analysis, Lecture Notes in Math. 506, G. A. Watson, ed., Springer-Verlag, Berlin, 1976, pp. 60–72.

[5] K. Dekker and J. G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Equations*, North-Holland, Amsterdam, 1984.

[6] G. Fleury and P. Bernard, *Convergence of numerical schemes for stochastic differential equations*, Monte Carlo Methods Appl., 7 (2001), pp. 35–44.

[7] I. Gyöngy, *A note on Euler's approximations*, Potential Anal., 8 (1998), pp. 205–216.

[8] I. Gyöngy and N. Krylov, *Existence of strong solutions for Itô's stochastic equations via approximations*, Probab. Theory Related Fields, 105 (1996), pp. 143–158.

[9] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations* II*, Stiff and Differential-Algebraic Problems*, 2nd ed., Springer-Verlag, Berlin, 1996.

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

[11] Y. Hu, *Semi-implicit Euler-Maruyama scheme for stiff stochastic equations*, in Stochastic Analysis and Related Topics V: The Silvri Workshop, Progr. Probab. 38, H. Koerezlioglu, ed., Birkhauser, Boston, 1996, pp. 183–202.

[12] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1999.

[13] X. Mao, *Stability of Stochastic Differential Equations with Respect to Semimartingales*, Pitman Res. Notes Math. Ser. 251, Longman Scientific and Technical, Harlow, UK, 1991.

[14] X. MAO, *Stochastic Differential Equations and Applications*, Horwood, Chichester, UK, 1997.

[15] J. MATTINGLY, A. M. STUART, AND D. J. HIGHAM, *Ergodicity for SDEs and Approximations: Locally Lipschitz Vector Fields and Degenerate Noise*, Tech. Rep. 7, University of Strathclyde, Department of Mathematics, Glasgow, UK, 2001.

[16] H. SCHURZ, *Stability, Stationarity, and Boundedness of some Implicit Numerical Methods for Stochastic Differential Equations and Applications*, Logos Verlag, Berlin, 1997.

[17] A. M. STUART AND A. R. HUMPHRIES, *Dynamical Systems and Numerical Analysis*, Cambridge University Press, Cambridge, UK, 1996.