Numerical approximations to the stationary solutions of stochastic differential equations

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This article was published in the journal, SIAM Journal on Numerical Analysis [© Society for Industrial and Applied Mathematics] and the definitive version is available at: http://dx.doi.org/10.1137/100797886

Metadata Record: https://dspace.lboro.ac.uk/2134/15310

Version: Accepted for publication

Publisher: © Society for Industrial and Applied Mathematics

Please cite the published version.
This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
NUMERICAL APPROXIMATIONS TO THE STATIONARY SOLUTIONS OF STOCHASTIC DIFFERENTIAL EQUATIONS

ANDREI YEVIK * AND HUAIZHONG ZHAO †

Abstract. In this paper, we investigate the possibility of approximating the stationary solution of stochastic differential equation. We start with the random dynamical system generated by the SDE with the multiplicative noise. We prove that the pull-back flow has a stationary point. However, the stationary point is not constructible explicitly, therefore we look at the numerical approximation. We prove that the discrete time random dynamical system also has a stationary point. Finally, we prove mean-square convergence of the approximate stationary solution to the exact one as the time-step diminishes, as well as almost surely convergence when the time step is rational.

Key words. random dynamical system, stationary solution, numerical approximation

AMS subject classifications. 37H99, 60H10, 60H35

1. Introduction. The notion of stationary solution is a fundamental concept in stochastic dynamical systems. Important examples of random dynamical systems include those generated by stochastic differential equations and stochastic partial differential equations. A solid base in random dynamical systems has been laid by Kunita [15], Pesin [19] and Ruelle [20], followed by the extensive results of Arnold [1], Mohammed, Zhang and Zhao [18]. The stationary solution gives equilibrium of the stochastic system and describes the long time behavior in the pathwise sense. To find the equilibrium of a random dynamical system is therefore an important basic problem to study in order to understand the longtime behavior of the system. However, unlike in the deterministic system, the stationary solution is usually a random moving, rather than steady point in the phase space, due to the random noise pumped to the system constantly. And, due to the above fact, it is often not feasible to construct the stationary solution explicitly. In order to find the stationary solution, one can use numerical simulation to compute it. However the question to ask is whether or not you can trust what you have computed. Moreover, this is a numerical approximation of infinite time horizon, not fixed time interval that researchers have studied most. Thus, to construct the approximate solution and consider the longtime convergence of the numerical scheme is interesting, not only in numerical analysis, but also in the analysis of random dynamical systems.

For the bounded time domain researchers have obtained vast results, including the rates of convergence of algorithms, adaptive methods, complexity of approximation. Studying point-wise error, Kloeden and Platen [13] presented a scheme for constructing numerical methods of any given order of convergence at any finite time. The results on global error on finite interval include [11],[12], where Hofmann, Müller-Gronbach and Ritter presented an adaptive Euler-Maruyama method, based on $n$ evaluations of Brownian motion, which is optimal within all $n$-points methods. In [10], Higham, Mao and Stuart developed approximate methods, which did not require the usual global Lipschitz assumption on the coefficients.

The already established results for the unbounded time domain include linear...
stability analysis for the SDEs of the form

\begin{equation}
  dX_t = \lambda X_t dt + \mu X_t dW_t,
\end{equation}

where the solution tends to a nonrandom attractor (point zero) when time tends to infinity, if some assumptions on \( \lambda \) and \( \mu \) are satisfied (ref. Burrage, Burrage and Mitsui [2], Saito and Mitsui [21]).

Some results about random attractors have also been obtained. For instance, Schmalfuss [23] has shown the existence of a fixed point in an appropriate function space of random variables. This fixed point is also shown to be an attractor. Another method of looking for a steady state is first to establish the existence of a random attractor and then prove that it is a singleton set. As a result, we have a stochastic stationary solution. In this method the existence of pullback absorbing family is proved, and then the global attractor is a closed union of omega-limit sets (cf. Arnold [1], Crauel, Debussche and Flandoli [7], Schmalfuss [22]). The results that follow this method include Caraballo, Kloeden and Real [4], where random delay equations were studied and the stationary solutions of the split implicit Euler scheme were proved to converge to the exact stationary solution when the time step tends to zero. Also, in Caraballo and Kloeden [3] the stochastic evolution equations with additive noise were studied, and pathwise convergence results were proved for Galerkin approximations and implicit Euler schemes. Garrido-Atienza, Kloeden and Neuenkirch [9] studied discretization of systems with additive noise driven by fractional Brownian motion. We cannot omit also the results on nonautonomous systems by Cheban, Kloeden and Schmalfuss [6] and Kloeden and Schmalfuss [14]. The results on stationary solutions of SPDEs include Liu and Zhao [16], Mohammed, Zhang and Zhao [18], Zhang and Zhao [26], [27]. A reader should also refer to works of Mattingly [17], Caraballo, Kloeden and Schmalfuss [5].

In this paper we will prove that under certain conditions, the SDE has a random stationary solution, and we can approximate this stationary point using the explicit Euler method. More precisely, we will study the m-dimensional SDE of the form

\begin{equation}
  \frac{dX_t}{dt} = -AX_t dt + f(X_t) dt + g(X_t) dW_t,
\end{equation}

where \( X : \Omega \times \mathbb{R}^+ \to \mathbb{R}^m, f : \mathbb{R}^m \to \mathbb{R}^m, g : \mathbb{R}^m \to \mathbb{R}^{m \times k} \), \( A \) is a symmetric positive-definite \( m \times m \) matrix and \( W_t \) is a vector of \( k \) independent scalar two-sided Wiener processes on a probability space \((\Omega, \mathcal{F}, P)\). Define

\[ \mathcal{F}_s^t = \sigma \{ W_u - W_v : s \leq v, u \leq t \}, \quad \mathcal{F}_s = \mathcal{F}_s^{-\infty} = \bigvee_{s \leq t} \mathcal{F}_s^t. \]

The initial condition for equation (1.2) is \( X(t_0) = X_0 \), where \( X_0 \) is \( \mathcal{F}_0 \)-measurable random variable. Also, denote by \( \theta : \mathbb{R} \times \Omega \to \Omega \) the standard \( P \)-preserving ergodic Wiener shift on \( \Omega \):

\[ \theta_t(\omega)(s) := \omega(t+s) - \omega(t), \quad t, s \in \mathbb{R}. \]

As we see, both the drift and the diffusion terms of equation (1.2) are nonlinear, so the analysis of the infinite time horizon problem of equation (1.2) will be more challenging and useful than the analysis of the equation (1.1). The limit of the solution of equation (1.2) is not known and is random. The solution of this equation is given by the formula

\begin{equation}
  X_t = e^{-At}X_0 + e^{-At} \left( \int_0^t e^{As} g(X_s) dW_s + \int_0^t e^{As} f(X_s) ds \right).
\end{equation}
It is well-known that the SDE (1.2) under some conditions defines a random dynamical system of perfect cocycle (cf. Arnold [1], Kunita [15]). Denote

\[ X_t = X(t, \omega, X_0) . \]

In [18], Mohammed, Zhang and Zhao studied the random dynamical system based on the SDE of the form (1.2) on a Hilbert space, with additive noise. They proved that when \( f \) is bounded and Lipschitz with small coefficients, there exists a unique bistable stationary solution \( X^* (\omega) \) in the sense that

\[ X(t, \omega, X^* (\omega)) = X^* (\theta (t, \omega)) , \]

for all \( t \geq 0 \) and all \( \omega \in \Omega \). Note that in this case

\[ X(t, \theta (-t, \omega), X^* (\theta (-t, \omega))) = X^* (\omega) , \]

for all \( t \geq 0 \) and all \( \omega \in \Omega \).

In this paper we study \( X^r_t (X_0) = X(r, \theta - t \omega, X_0) \), the pullback of the original equation (1.3) with the multiplicative noise:

\[
dX^r_t = -AX^r_t dr + f (X^r_t) \, dr + g (X^r_t) \, dW^r_{t-} ,
\]

for any \( t > 0, r > 0 \), and the initial condition \( X(0, \theta - t \omega, X_0 (\theta - t \omega)) = X^r_0 = X_0 (\theta - t \omega) \), which is \( \mathcal{F}^{-t} \)-measurable random variable. The pullback concept is widely used in the literature. For more details the reader is referred to Crauel and Flandoli [8]. We will prove that under some assumptions the pullback solution converges to a fixed point in \( L^2 (\Omega) \) when the time goes to infinity. However, we are not able to construct this fixed point solution explicitly. This is why it is natural to apply the Euler numerical scheme to get the approximate stationary solution and prove that the latter one converges to the exact fixed point in \( L^2 (\Omega) \). From convergence in \( L^2 (\Omega) \) the pathwise convergence can also be established for rational time-steps. A thorough study of deterministic dynamical systems and approximating them has been presented by Stuart and Humphries [24], which helped to create this paper by comparing deterministic and random models.

We would like to point out that the existence of the stationary solution of the additive noise or multiplicative linear noise SDE has been already studied by many researchers. However, not much has been proved for the nonlinear multiplicative noise SDE. The convergence of the Euler scheme in this infinite horizon case is completely new.

2. Assumptions and some preliminary results. Denote the eigenvalues of \( A \) by \( \{ \lambda_j, j = 1, 2, \ldots, m \} \), and define \( \alpha \) to be the smallest one of them, such that

\[ 0 < \alpha = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_m . \]

Suppose the following conditions are satisfied:

**Condition 1** Global Lipschitz condition: for any \( X, Y \in \mathbb{R}^m \)

\[
|f(X) - f(Y)| \leq \beta_1 |X - Y| ,
\]

\[
|g(X) - g(Y)| \leq \beta_2 |X - Y| ,
\]

where \( \beta_1, \beta_2 > 0 \) are constants and satisfy

\[
\beta_1 + \frac{\beta_2}{2} < \alpha . \tag{2.1}
\]
From this condition it follows that for any $X,Y \in \mathbb{R}^m$, the linear growth condition also holds:

$$|f(X)| \leq \beta_1 |X| + C_1,$$

$$|g(X)| \leq \beta_2 |X| + C_2,$$

with constants $C_1, C_2 > 0$. We will also use this condition.

**Condition 2** Boundedness of initial value: there exists a constant $K^* > 0$, such that

$$E |X_0|^2 \leq K^*.$$

Note that the pullback initial point has the same bound due to the measure preserving property of $\theta_{-t}(\omega)$ for any $t > 0$:

$$E \left( |X_t|^2 \right) = E \left( |X_0|^2 \right) \leq K^*.$$

We now state some preliminary results about the matrix manipulations, that will be needed in the proof of the main results of this paper.

**Lemma 2.1.** Assume the matrix $A$ is as stated above. Then for any $t > 0$ the matrix

$$e^{-A \Delta t} - \frac{\sum_{i=0}^{p} 1}{i!} (-A \Delta t)^i$$

is positive-definite for odd $p \in \mathbb{N}$, and negative-definite for even $p \in \mathbb{N}$ and $p = 0$.

**Proof.** We start with the one-dimensional statement: for any $\alpha > 0$, any $t > 0$,

$$e^{-\alpha t} - \frac{\sum_{i=0}^{p} 1}{i!} (-\alpha t)^i$$

is positive for odd $p \in \mathbb{N}$, and negative for even $p \in \mathbb{N}$ and $p = 0$. The statement is valid for $p = 0$. Consider a function

$$f(t) := e^{-\alpha t} - \frac{\sum_{i=0}^{p+1} 1}{i!} (-\alpha t)^i,$$

and note that $f(0) = 0$. Then,

$$f'(t) = -\alpha \left( e^{-\alpha t} - \frac{\sum_{i=0}^{p} 1}{i!} (-\alpha t)^i \right).$$

If

$$e^{-\alpha t} - \frac{\sum_{i=0}^{p} 1}{i!} (-\alpha t)^i$$
is positive, then \( f'(t) \) is negative for all \( t > 0 \), and thus \( f(t) < 0 \). And vise versa. Starting from \( p = 0 \), with each consecutive integer \( p \) the sign of (2.3) will change. So, one-dimensional statement is correct.

Now, matrix exponential of \( A \) is diagonalizable:

\[
e^A = Qe^DQ^{-1},
\]

where \( Q \) is invertible and \( D \) is diagonal with eigenvalues of \( A \) as its spectrum. For any \( p \geq 0 \) we have:

\[
e^{-At} - \sum_{i=0}^{p} \frac{1}{i!} (-At)^i = Q \left( e^{-Dt} - \sum_{i=0}^{p} \frac{1}{i!} (-Dt)^i \right) Q^{-1}.
\]

Note that the matrix

\[
e^{-Dt} - \sum_{i=0}^{p} \frac{1}{i!} (-Dt)^i
\]

is diagonal with the following trace:

\[
\begin{pmatrix}
e^{-\lambda_1 t} - \sum_{i=0}^{p} \frac{1}{i!} (-\lambda_1 t)^i & 0 & \ldots & 0 \\
0 & e^{-\lambda_2 t} - \sum_{i=0}^{p} \frac{1}{i!} (-\lambda_2 t)^i & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & e^{-\lambda_m t} - \sum_{i=0}^{p} \frac{1}{i!} (-\lambda_m t)^i
\end{pmatrix}.
\]

The diagonal elements are the eigenvalues of the matrix (2.2), and according to the one-dimensional statement proved earlier, they are all positive or all negative depending on \( p \).

**Lemma 2.2.** Assume the matrix \( A \) is as stated above, and denote by \( \lambda \) the largest eigenvalue of \( A \). Then, for \( 0 < \Delta t \leq \frac{1}{\lambda} \), the matrix

\[
e^{-A\Delta t} - (I - A\Delta t)^j
\]

is positive-definite for any \( j \in \mathbb{N} \).

**Proof.** We prove this lemma by using the induction principle. For \( j = 1 \), the result follows from Lemma 2.1. Suppose now that for some \( j \) the matrix \( e^{-A\Delta t} - (I - A\Delta t)^j \) is positive-definite, and examine the matrix:

\[
e^{-A\Delta t(j+1)} - (I - A\Delta t)^{j+1} = e^{-A\Delta t} \left( e^{-A\Delta t} - (I - A\Delta t)^j \right) + \left( e^{-A\Delta t} - (I - A\Delta t) \right) (I - A\Delta t)^j.
\]

Now, all of the four terms on the right-hand side are positive-definite matrices. We also have that

\[
e^{-A\Delta t} (I - A\Delta t)^j = (I - A\Delta t)^j e^{-A\Delta t}
\]
due to the nature of the matrices \( e^{-A\Delta t} \) and \( (I - A\Delta t)^j \). Thus, we have a sum of two products of commuting positive-definite matrices, which gives us a positive definite matrix. By the induction principle, all the matrices \( e^{-A\Delta t} - (I - A\Delta t)^j \), for \( j \in \mathbb{N} \), are positive-definite. \( \square \)
Now we give some estimates of the solution of equation (1.4):

\[
X_t^r = e^{-Ar}X_0^r + e^{-Ar} \int_0^r e^{As} g(X_s^t) \, dW_{s-t} + e^{-Ar} \int_0^r e^{As} f(X_s^t) \, ds.
\]

(2.4)

Denote the \(L_p\)-norm of the solution by the following:

\[
\|X_t^r\|_p = \left( E \left( |X_t^r|^p \right) \right)^{1/p},
\]

where \(p \in \mathbb{N}\). Here for any \((d_1 \times d_2)\) matrix \(B\), \(|B|\) stands for the Euclidian (Frobenius) norm:

\[
|B| = \sqrt{\sum_{i=1}^{d_1} \sum_{j=1}^{d_2} B_{ij}^2}.
\]

**Lemma 2.3.** Assume Conditions 1 and 2 hold. Then there exists a constant \(C > 0\) such that for any \(t \geq 0, r \geq 0\), the solution of SDE (1.4) satisfies:

\[
E |X_t^r|^2 \leq C.
\]

**Proof.** First, using Ito’s formula, we have:

\[
e^{2ar} |X_t^r|^2 = |X_0^r|^2 + 2a \int_0^r e^{2as} |X_s^t|^2 \, ds - 2 \int_0^r e^{2as} (X_s^t)^T A X_s^t \, ds
\]

\[
+ 2 \int_0^r e^{2as} (X_s^t)^T f(X_s^t) \, ds + \int_0^r e^{2as} |g(X_s^t)|^2 \, ds
\]

(2.5)

\[
= 2 \int_0^r e^{2as} (X_s^t)^T (\alpha I - A) X_s^t \, ds \leq 0,
\]

(2.6)

since the matrix \(\alpha I - A\) is non-positive definite. Then from (2.5), (2.6), linear growth condition and Young’s inequality, we get:

\[
e^{2ar} E \left( |X_r^t|^2 \right) \leq ||X_0||_2^2 + 2 \int_0^r e^{2as} E \left( (X_s^t)^T f(X_s^t) \right) \, ds
\]

\[
+ \int_0^r e^{2as} E \left( |g(X_s^t)|^2 \right) \, ds
\]
\[
\begin{align*}
&\leq \|X_0\|_2^2 + 2 \int_0^r e^{2\alpha s} \left( C_1 E \left( |X_s'| \right) + \beta_1 E \left( |X_s'|^2 \right) \right) ds \\
&\quad + \int_0^r e^{2\alpha s} \left( C_2^2 + 2\beta_2 C_2 E \left( |X_s'| \right) + \beta_2^2 E \left( |X_s'|^2 \right) \right) ds \\
&\leq \|X_0\|_2^2 + 2(1 + \beta_2^2) \int_0^r e^{2\alpha s} E \left( |X_s'|^2 \right) ds \\
&\quad + (2\alpha)^{-1} C_2^2 (e^{2\alpha r} - 1) + (2\beta_1 + \beta_2^2) \int_0^r e^{2\alpha s} E \left( |X_s'|^2 \right) ds \\
&\leq K_1 + K_2 e^{2\alpha r} + K_3 \int_0^r e^{2\alpha s} E \left( |X_s'|^2 \right) ds,
\end{align*}
\]

where
\[
\begin{align*}
K_1 &= \|X_0\|_2^2 - \frac{C_2^2}{2\alpha} - \frac{(1 + \beta_2^2)C_1}{2\alpha (2\beta_1 + \beta_2^2)}, \\
K_2 &= \frac{C_2^2}{2\alpha} + \frac{(1 + \beta_2^2)\beta_2}{2\alpha (2\beta_1 + \beta_2^2)}, \\
K_3 &= 2\beta_1 + \beta_2^2 + \epsilon \left( 2\beta_1 + \beta_2^2 \right),
\end{align*}
\]

and \(\epsilon\) is such a number that
\[
\left( \beta_1 + \frac{\beta_2^2}{2} \right) (1 + \epsilon) < \alpha.
\]

This is true noting (2.1). Now applying the Gronwall inequality we have:
\[
e^{2\alpha r} E \left( |X_r'|^2 \right) \leq (K_1 + K_2) e^{K_3 r} + \frac{2\alpha K_2}{2\alpha - K_3} \left( e^{2\alpha r} - e^{K_3 r} \right).
\]

Finally,
\[
E \left( |X_r'|^2 \right) \leq \|X_0\|_2^2 + \frac{2\alpha K_2}{2\alpha - K_3}.
\]

A straightforward corollary of this lemma is that \(E (|X_r'|)\) is also bounded uniformly in \(r\). In our analysis we will also need a bound on the norm \(\|X_{t_1} - X_{t_2}\|_2\) for any fixed time \(t_1, t_2\).

**Lemma 2.4.** Assume Conditions 1 and 2 hold. Then there exist constants \(C_3 > 0, C_4 > 0\), such that for any \(t, t_1, t_2 \geq 0, t_1 \geq t_2\), the solution of SDE (1.4) satisfies:
\[
\|X_{t_1} - X_{t_2}\|_2 \leq C_3 (t_1 - t_2) + C_4 \sqrt{t_1 - t_2}.
\]

**Proof.** From (2.4), we can see that
\[
\|X_{t_1} - X_{t_2}\|_2 \leq \|X_0\|_2 \left| e^{-A_{t_1}} - e^{-A_{t_2}} \right|
\]
Secondly, applying Minkowski’s inequality, Ito’s isometry and Lemma 2.3, we get
\[ \left\| e^{-At_1} - e^{-At_2} \right\|_2 = \left\| e^{At_1} f (X^t_1) ds - e^{At_2} f (X^t_2) ds \right\|_2 + \left\| e^{At_1} f (X^t_1) ds - e^{At_2} f (X^t_2) ds \right\|_2. \]

We evaluate the terms on the right hand side separately. First note
\[ \|X_0\|_2 |e^{-At_1} - e^{-At_2}| \leq \|X_0\|_2 |A| (t_1 - t_2). \]

Secondly, applying Minkowski’s inequality, Ito’s isometry and Lemma 2.3, we get
\[ \left\| e^{At_1} f (X^t_1) ds - e^{At_2} f (X^t_2) ds \right\|_2 \leq \left\| e^{At_1} g (X^t_1) dW_s - e^{At_2} g (X^t_2) dW_s \right\|_2. \]

Again, using Minkowski’s inequality, Hölder’s inequality and Lemma 2.3, we obtain:
\[ \left\| e^{At_1} g (X^t_1) dW_s - e^{At_2} g (X^t_2) dW_s \right\|_2 \leq K_4 (t_1 - t_2) + K_5 \sqrt{t_1 - t_2}, \]

for some positive constants \( K_4, K_5. \)

Again, using Minkowski’s inequality, Hölder’s inequality and Lemma 2.3, we obtain:
\[ \left\| e^{-At_1} f (X^t_1) ds - e^{-At_2} f (X^t_2) ds \right\|_2 \leq \left\| e^{-At_1} f (X^t_1) ds \right\|_2 + \left\| e^{-At_2} f (X^t_2) ds \right\|_2 \leq \left\| e^{At_1} f (X^t_1) ds \right\|_2 + \left\| e^{At_2} f (X^t_2) ds \right\|_2 \]
\[ \leq K_6 (t_1 - t_2). \]
Combining the results, we have:
\[
\|X_{t_1}^r - X_{t_2}^r\|_2 \leq \|X_0\|_2 |A| (t_1 - t_2) + (K_4 + K_0) (t_1 - t_2) + K_5 \sqrt{t_1 - t_2}
\]
\[
\leq C_3 (t_1 - t_2) + C_4 \sqrt{t_1 - t_2},
\]
for some constants $C_3, C_4$. Note that the constants $C_3$ and $C_4$ are uniform in $t, t_1, t_2$.

\[\square\]

**Lemma 2.5.** Denote by $X_t^r$ and $Y_t^r$ two solutions of (1.4) with initial values $X_0$ and $Y_0$ respectively. Assume Condition 1 holds and Condition 2 holds for both initial values. Then, for any $t \geq 0$, any positive $\epsilon$, there exists a positive $r^*$, such that for any $r \geq r^*$:
\[
\|X_t^r - Y_t^r\|_2 \leq \epsilon.
\]

**Proof.** It is easy to see that
\[
X_t^r - Y_t^r = e^{-Ar} (X_0^r - Y_0^r) + e^{-Ar} \int_0^r e^{As} (f (X_s^r) - f (Y_s^r)) \, ds
\]
\[
+ e^{A(t-r)} \int_{t-r}^r e^{As} (g (X_{s+t}^r) - g (Y_{s+t}^r)) \, dW_s.
\]

Denote $\xi_t^r = X_t^r - Y_t^r$. Following a similar approach as in Lemma 2.3, we get:
\[
e^{2\alpha r} \|\xi_t^r\|_2^2 \leq \|\xi_0^r\|_2^2 + (2\beta_1 + \beta_2^2) \int_0^r e^{2\alpha s} \|\xi_s^r\|_2^2 \, ds,
\]
and again Gronwall's inequality gives us:
\[
\|\xi_t^r\|_2^2 \leq \|\xi_0^r\|_2^2 e^{(2\beta_1 + \beta_2^2 - 2\alpha)r}.
\]

The lemma follows immediately. \[\square\]

**Theorem 2.6.** Assume Condition 1 holds. Then there exists $X^* \in L_2 (\Omega)$, such that for any initial value $X_0 (\omega)$ satisfying Condition 2, the solution of (1.4) satisfies:
\[
\lim_{t \to \infty} \|X_t^r (X_0) - X^*\|_2 = 0.
\]

**Proof.** Condition 2 implies that $X_0 (\omega)$ belongs to a Banach space $L_2 (\Omega)$. According to Lemma 2.3, $X_t^r (X_0 (\omega))$ maps $L_2 (\Omega)$ into itself. Under Conditions 1 and 2, without assuming (2.1), the SDE (1.2) defines a random dynamical system, so that we can use the perfect cocycle property to obtain: for any $r, t, p \geq 0$ and any $X_0$
\[
X_{r+p}^t (X_0) = X (r + p, X_0, \theta_{-r-p} \omega) = X (r, X (p, X_0, \theta_{-r-p} \omega), \theta_{-t} \omega).
\]

Now, according to Lemma 2.3: $X (p, X_0, \theta_{-t-p} \omega) \in L_2 (\Omega)$. Thus, using Lemma 2.5, for any $\epsilon > 0$ there exists $r^* > 0$, such that for any $r \geq r^*, p \geq 0$:
\[
\|X_t^r (X_0) - X_{r+p}^t (X_0)\|_2 \leq \epsilon.
\]
Now take \( r = t \) and observe an arbitrary sequence of real numbers \( t_1, t_2, \ldots, t_n \), diverging to infinity. The last obtained inequality states that there exists \( n_0 > 0 \), such that for any \( i, j \geq n_0 \):
\[
\left\| X_{i}^{t_i}(X_0) - X_{j}^{t_j}(X_0) \right\| \leq \epsilon,
\]
which shows that \( X_{i}^{t_i}(X_0) \) is a Cauchy sequence, thus converges within the space \( L_2(\Omega) \). According to the convergence criterion, \( X_{i}^{t}(X_0) \) converges to some \( X^* \) in \( L_2(\Omega) \). Now, in order to show the independence of the limit from the initial point, for any \( Y_0 \) satisfying Condition 2, we observe the following:
\[
\left\| X^* - X_{i}^{t_i}(Y_0) \right\|_2 \leq \left\| X^* - X_{i}^{t_i}(X_0) \right\|_2 + \left\| X_{i}^{t_i}(X_0) - X_{j}^{t_j}(Y_0) \right\|_2.
\]
Now, sending \( t \) to infinity makes the right-hand side small enough and we obtain the convergence. 

**Remark 1.** Note that due to the invariance of the Wiener measure the following is true: for any \( t_1, t_2, t_3 \geq 0 \),
\[
\left\| X_{t_2}^{t_1} \right\|_2 = \left\| X_{t_1}^{t_3} \right\|_2.
\]
The results of Lemma 2.3, Lemma 2.4 and Lemma 2.5 could be proved assuming \( t = 0 \), and the same estimates for general \( t \geq 0 \) would follow. Theorem 2.6 of course deals with the pullback.

**3. Numerical scheme.** In the following, we divide time domain into \( N \) intervals of the length \( \Delta t \), use the equidistant discretization with mesh \( \Delta t \). We will use the simplest explicit one-step numerical method - Euler’s method. Starting from an \( \mathcal{F}^{t_0} \)-measurable random variable \( \hat{X}_0^t \) at time point zero, at each of the discretization points \((i \Delta t)\) we set the value \( \hat{X}_i^{t_i} \) with the following iteration formula:
\[
\hat{X}_{i+1}^{t_i + \Delta t} = \hat{X}_i^{t_i} - A \hat{X}_i^{t_i} \Delta t + f \left( \hat{X}_i^{t_i} \right) \Delta t + g \left( \hat{X}_i^{t_i} \right) \left( W_{i+1}^{t_i + \Delta t} - W_{i}^{t_i + \Delta t} \right),
\]
where \( i = 0, \ldots , N - 1 \).

After \( N \) iterations we can construct
\[
\hat{X}_{N \Delta t}^t = (I - A \Delta t)^N \hat{X}_0^t + \Delta t \sum_{i=0}^{N-1} (I - A \Delta t)^{N-i-1} f \left( \hat{X}_i^{t_i} \right)
\]
\[
+ \sum_{i=0}^{N-1} (I - A \Delta t)^{N-i-1} g \left( \hat{X}_i^{t_i} \right) \left( W_{i+1}^{t_i + \Delta t} - W_{i}^{t_i + \Delta t} \right).
\]
(3.1)
It is easy to verify that this scheme generates a discrete time random dynamical system. The cocycle property follows straightforward from the iteration formula. We will now prove the discrete time analogues of Lemmas 2.3, 2.5 and Theorem 2.6.

**Remark 2.** The estimates for the numerical solution as well as the convergence theorem are proved for the pullback solution, however it is possible to carry out the calculations assuming \( t = 0 \), and the estimates will be the same. The results on stationary solutions of course need the pullback.

**Lemma 3.1.** Assume Conditions 1 and 2 hold. Then there exists a constant \( \hat{C} > 0 \) such that for any \( t \geq 0 \), any integer \( M \geq 0 \) and sufficiently small \( \Delta t \), the numerical solution \( \hat{X}_{M \Delta t}^t \) defined by (3.1) satisfies:
\[
E \left| \hat{X}_{M \Delta t}^t \right|^2 \leq \hat{C}.
\]
Proof. For simplicity, denote $\tilde{X}_{i\Delta t}^t$ by $\tilde{X}_i^t$. We have:

$$(1 - \alpha \Delta t)^{-2M} |\tilde{X}_{\Delta t}^t|^2 = |\tilde{X}_0^t|^2 + \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} \left( \frac{|\tilde{X}_{i+1}^t|^2}{(1 - \alpha \Delta t)^2} - |\tilde{X}_i^t|^2 \right).$$

Notice that

$$\frac{|\tilde{X}_{i+1}^t|^2}{(1 - \alpha \Delta t)^2} = \left( \frac{(\tilde{X}_{i+1}^t)^T}{1 - \alpha \Delta t} - (\tilde{X}_i^t)^T \right) \left( \frac{\tilde{X}_{i+1}^t}{1 - \alpha \Delta t} + \tilde{X}_i^t \right)$$

$$= \left( \tilde{X}_i^T \left( \frac{I - A \Delta t}{1 - \alpha \Delta t} - I \right) + \frac{\Delta t}{1 - \alpha \Delta t} f \left( \tilde{X}_i^T \right) \right) + \frac{\Delta t}{1 - \alpha \Delta t} g \left( \tilde{X}_i^T \right) + \frac{\Delta t}{1 - \alpha \Delta t} \left( W_{(i+1)\Delta t - t} - W_{i\Delta t - t} \right) \left( W_{i\Delta t - t} - W_{(i+1)\Delta t - t} \right).$$

Note that the matrix $\left( \frac{I - A \Delta t}{1 - \alpha \Delta t} - I \right) \left( \frac{I - A \Delta t}{1 - \alpha \Delta t} + I \right)$ is non-positive definite assuming $\Delta t$ as in Lemma 2.1. Also note that for every $i$, $f \left( \tilde{X}_i^T \right)$ and $g \left( \tilde{X}_i^T \right)$ are independent from $(W_{(i+1)\Delta t - t} - W_{i\Delta t - t})$. Using this fact, the linear growth condition and Young’s inequality, we get the following estimate:

$$(1 - \alpha \Delta t)^{-2M} E \left( |\tilde{X}_{\Delta t}^t|^2 \right)$$

$$\leq E \left( |\tilde{X}_0^t|^2 \right) + \frac{2\Delta t}{(1 - \alpha \Delta t)^2} \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} E \left( \tilde{X}_i^T \left( I - A \Delta t \right) f \left( \tilde{X}_i^T \right) \right)$$

$$+ \frac{\Delta t^2}{(1 - \alpha \Delta t)^2} \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} E \left( |f \left( \tilde{X}_i^T \right)|^2 \right)$$

$$+ \frac{\Delta t}{(1 - \alpha \Delta t)^2} \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} E \left( |\tilde{X}_i^T|^2 \right)$$

$$\leq E \left( |\tilde{X}_0^t|^2 \right) + C_1^2 \Delta t^2 + C_2^2 \Delta t + C_3^2 (1 - \alpha \Delta t)^{-2M}$$

$$+ \frac{\Delta t}{(1 - \alpha \Delta t)^2} \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} E \left( \tilde{X}_i^T \left( 2\beta_1 + \beta_2 + \Delta t (\beta_1^2 + 2\beta_1 |A|) \right) \right)$$

$$+ \frac{2\Delta t}{(1 - \alpha \Delta t)^2} \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} E \left( |\tilde{X}_i^T|^2 \right) \left( C_1 + \beta_2 C_3 + \Delta t C_1 (\beta_1 + |A|) \right)$$

$$\leq K_1 + (1 - \alpha \Delta t)^{-2M} \tilde{K}_2 + \tilde{K}_3 \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} E \left( |\tilde{X}_i^T|^2 \right),$$
where
\[ \hat{K}_1 = E \left( |\hat{X}_0| \right) \]
\[ \hat{K}_2 = \frac{C_2^2 \Delta t^2 + C_2 \Delta t}{2 \alpha \Delta t - \alpha^2 \Delta t^2} + \frac{\Delta t}{2 \alpha \Delta t - \alpha^2 \Delta t^2} \left( C_1 + \beta_2 C_2 + \Delta t C_1 (\beta_1 + |A|) \right)^2 \]
\[ \hat{K}_3 = \frac{\Delta t (1 + \hat{\epsilon}) (2 \beta_1 + \beta_2^2 + \Delta t (\beta_1 + 2 |A|))}{(1 - \alpha \Delta t)^2} \]
and \( \Delta t \) and \( \hat{\epsilon} \) are chosen small enough such that
\[ (2 \beta_1 + \beta_2^2 + \Delta t (\beta_1 + 2 |A|)) (1 + \hat{\epsilon}) + \alpha^2 \Delta t < 2 \alpha. \]

Applying now discrete Gronwall’s inequality ([25]) gives us:
\[ \left( 1 - \alpha \Delta t \right)^{-2M} E \left( |\hat{X}_M| \right)^2 \leq \hat{K}_1 + \left( 1 - \alpha \Delta t \right)^{-2M} \frac{\hat{K}_2}{1 + \hat{K}_3} + \hat{K}_3 \sum_{i=1}^{M} \left( \hat{K}_1 + \hat{K}_2 + (1 - \alpha \Delta t)^{-2i} \hat{K}_2 \right) (1 + \hat{K}_3) \]

And finally,
\[ E \left( |\hat{X}_M| \right)^2 \leq \frac{\hat{K}_2}{1 + \hat{K}_3} + \hat{K}_1 \left( 1 + \hat{K}_3 \right) \left( 1 - \alpha \Delta t \right)^2 \]
\[ + \frac{\hat{K}_2 \hat{K}_3 \left( 1 - \left( 1 + \hat{K}_3 \right) \left( 1 - \alpha \Delta t \right)^2 \right)^M}{\left( 1 + \hat{K}_3 \right) \left( 1 - \left( 1 + \hat{K}_3 \right) \left( 1 - \alpha \Delta t \right)^2 \right)^M} \leq C. \]

The next lemma is the discrete analogue of Lemma 2.5. The proof is essentially the same as the proof of Lemma 3.1 and Lemma 2.5.

**Lemma 3.2.** Denote by \( \hat{X}_{M,\Delta t}^t \) and \( \hat{Y}_{M,\Delta t}^t \) the solutions of the scheme (3.1) with initial values \( \hat{X}_0^t \) and \( \hat{Y}_0^t \). Assume Condition 1 holds and assume Condition 2 holds for both initial points, \( \Delta t \) is sufficiently small as in Lemma 3.1. Then for any positive \( \epsilon \), there exists integer \( M^* \), such that for any \( M \geq M^* \):
\[ ||\hat{X}_{M,\Delta t}^t - \hat{Y}_{M,\Delta t}^t||_2 \leq \epsilon. \]

**Example 1.** We would like to provide a numerical example to illustrate the results of Lemma 3.2. Suppose that the state-space is two-dimensional and Wiener process is three-dimensional. Choose the mesh-size \( \Delta t = 0.01 \), and assume \( M = 300 \). Assume the following coefficients:
\[ A = \begin{pmatrix} 5 & 1 \\ 1 & 7 \end{pmatrix}, \]
\[ f \left( \begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x + y + \sin(xy) \\ x - y - \sin(xy) \end{pmatrix}, \]
g \left( \begin{array}{c} x \\ y \end{array} \right) = \left( \begin{array}{c} x + \cos(x-y) \\ y - x - y \\ \sin(x+y) \end{array} \right).

Note that this choice of coefficients satisfies Condition 1. We can assume \( t = 0 \), as pullback does not make difference in this lemma. We will run two simulations, in each one constructing trajectories of processes \( \hat{X} \) and \( \hat{Y} \) with different starting points. Choose the following:

\[
\hat{X}_0^1 = \left( \begin{array}{c} 2 \\ 1 \end{array} \right), \quad \hat{Y}_0^1 = \left( \begin{array}{c} 0 \\ -3 \end{array} \right), \quad \hat{X}_0^2 = \left( \begin{array}{c} -1 \\ 5 \end{array} \right), \quad \hat{Y}_0^2 = \left( \begin{array}{c} 3 \\ 4 \end{array} \right).
\]

We generate six Brownian trajectories (three independent scalar motions for each of two simulations) in the following way: \( W_0 = 0, W_{(i+1)\Delta t} = W_{i\Delta t} + \psi_{i+1}, \) where \( \psi_i \sim N(0, \sqrt{\Delta t}), \) \( i = 1, \ldots, M \). We plot the coordinates of the processes separately, so that Figure 3.1 has four graphs. As we see, the illustrations confirm the theory.

**Theorem 3.3.** Assume Condition 1 holds. Assume that \( \Delta t \) is fixed, small enough and \( t = N\Delta t \). Then there exists \( \hat{X}^* \in L_2(\Omega) \), such that for any initial value \( \hat{X}_0 \) satisfying Condition 2, the solution of the numerical scheme (3.1) satisfies:

\[
\lim_{N \to \infty} \left\| X_{N\Delta t}^t \left( \hat{X}_0^t \right) - \hat{X}^* \right\|_2 = 0.
\]

The proof of this theorem is essentially the same as the proof of Theorem 2.6. It is easy to show that the sequence \( \hat{X}_{N\Delta t}^t \) is a Cauchy sequence and therefore has a limit.

**Example 2.** We would like to provide an example to illustrate Theorem 3.3. Suppose we are working in one-dimensional space of real numbers. Choose the following coefficients:

\[ A = 5, \quad f(x) = x + \sin(x), \quad g(x) = x + \cos(x). \]
We will run two simulations, with the following meshes:

\[
\Delta t_1 = 0.01, \quad N_1^1 = 300, \quad N_2^1 = 500, \quad N_3^1 = 400,
\]
\[
\Delta t_2 = 0.05, \quad N_2^1 = 200, \quad N_2^2 = 300, \quad N_3^2 = 250.
\]

So, for instance, the first simulation will show us the values of \( \hat{X} \) (300 points), values of \( \hat{Y} \) (500 points) and values of \( \hat{Z} \) (400 points). The time-step for this scheme will be 0.01. We must consider pullback, as \( t = N \Delta t \). To get the Brownian trajectory for the negative domain we construct the positive time path, and then reflect it against point zero. We will choose the following starting points:

\[
\hat{X}_0^1 = -1, \quad \hat{Y}_0^1 = 2, \quad \hat{Z}_0^1 = 1.5, \quad \hat{X}_0^2 = 4, \quad \hat{Y}_0^2 = -3, \quad \hat{Z}_0^2 = 2.
\]

From the definition of the numerical scheme it follows that for any suitable \( N, \Delta t \) and starting point \( x \), the following is true:

\[
\hat{X} (N \Delta t, 0, -N \Delta t, \omega) (x) = \hat{X} (0, -N \Delta t, \omega) (x),
\]

where the first argument is finishing time and the second is starting time. Plotting the curves from the point \( -N \Delta t \) to point zero illustrates the idea of pullback and gives the best idea of the behavior of the processes. As we see, as time progresses, the trajectories get asymptotically close. Of course, we are mostly interested only in the final points. However, the graphs also reflect the fact that whichever (large) number of points we choose, as we move in time, the solution gets on the stationary trajectory and stays there till the fixed point. This intuitive idea come from Theorem 3.3. Of course, the attracting point is a random variable, thus different for every \( \omega \).

4. Convergence theorem. Note that the exact solution of (1.4) at time \( N \Delta t \) has the form:

\[
X_{N \Delta t} = e^{-AN \Delta t} X_0 + e^{-AN \Delta t} \int_0^{N \Delta t} e^{As} g (X_s^t) dW_s - t \\
+ e^{-AN \Delta t} \int_0^{N \Delta t} e^{As} f (X_s^t) ds.
\]

(4.1)

The approximate solution has the form (3.1), with \( \hat{X}_0^t = X_0^t \). It should not be difficult to prove that as \( \Delta t \to 0 \), for any finite horizon, the approximate solution converges
to the exact solution. In the following, we will prove that the approximate solution converges to the exact one as $\Delta t \to 0$ in infinite horizon.

**Theorem 4.1.** Assume Conditions 1,2 hold. Choose $N\Delta t = t$. If $X_t$ and $\hat{X}_t$ are the exact and the approximate solutions, given by (4.1) and (3.1) respectively, then there exists a constant $K > 0$, such that for any sufficiently small fixed $\Delta t$:

$$\limsup_{t \to \infty} \left\| X_t - \hat{X}_t \right\|_2 \leq K\sqrt{\Delta t}.$$

**Proof.** It is easy to see that for any $M \leq N$ we have:

$$X_{M\Delta t} - \hat{X}_{M\Delta t} = \left( e^{-AM\Delta t} - (I - A\Delta t)^M \right) X_0 + e^{-AM\Delta t} \int_0^{M\Delta t} e^{As} g(X_s) \, dW_s - \sum_{i=0}^{M-1} (I - A\Delta t)^{M-i-1} g(\hat{X}_{i\Delta t})(W_{(i+1)\Delta t} - W_{i\Delta t}) + e^{-AM\Delta t} \int_0^{M\Delta t} e^{As} f(X_s) \, ds - \Delta t \sum_{i=0}^{M-1} (I - A\Delta t)^{M-i-1} f(\hat{X}_{i\Delta t}).$$

(4.2)

The rest of the proof of the theorem is similar to the proof of Lemma 3.1. First note that

$$(1 - \alpha\Delta t)^{-2M} \left| X_{M\Delta t} - \hat{X}_{M\Delta t} \right|^2 = \sum_{i=0}^{M-1} (1 - \alpha\Delta t)^{-2i} \left( \left| X_{(i+1)\Delta t} - \hat{X}_{(i+1)\Delta t} \right|^2 \left(1 - \alpha\Delta t\right)^2 - \left| X_{i\Delta t} - \hat{X}_{i\Delta t} \right|^2 \right).$$

Denote

$$B_1 = \frac{1}{1 - \alpha\Delta t} \int_{i\Delta t}^{(i+1)\Delta t} \left( e^{A(s-(i+1)\Delta t)} f(X_s) - f(\hat{X}_{i\Delta t}) \right) \, ds,$$

$$B_2 = \frac{1}{1 - \alpha\Delta t} \int_{i\Delta t-t}^{(i+1)\Delta t-t} \left( e^{A(s+t-(i+1)\Delta t)} g(X_{s+t}) - g(\hat{X}_{i\Delta t}) \right) \, dW_s.$$

Then,

$$\left| X_{(i+1)\Delta t} - \hat{X}_{(i+1)\Delta t} \right|^2 \left(1 - \alpha\Delta t\right)^2 - \left| X_{i\Delta t} - \hat{X}_{i\Delta t} \right|^2 = \left( X_{i\Delta t}^T \left( e^{-A\Delta t} \frac{1}{1 - \alpha\Delta t} - I \right) \hat{X}_{i\Delta t} \right)^T \left( I - A\Delta t \frac{1}{1 - \alpha\Delta t} - I \right) + B_1^T + B_2^T \times \left( \left( e^{-A\Delta t} \frac{1}{1 - \alpha\Delta t} + I \right) X_{i\Delta t} - \left( I - A\Delta t \frac{1}{1 - \alpha\Delta t} + I \right) \hat{X}_{i\Delta t} + B_1 + B_2 \right).$$
we will estimate separately:

\[
(X_{i,\Delta t} - \hat{X}_{i,\Delta t})^T \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} - I \right) \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} + I \right) (X_{i,\Delta t} - \hat{X}_{i,\Delta t}) + B_1^T B_1 \\
+ (\hat{X}_{i,\Delta t})^T \left( \frac{e^{-A\Delta t} - I + A \Delta t}{1 - \alpha \Delta t} \right)^2 \hat{X}_{i,\Delta t} + B_2^T B_2 \\
+ 2 (X_{i,\Delta t} - \hat{X}_{i,\Delta t})^T \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} - I \right) \frac{e^{-A\Delta t} - I + A \Delta t}{1 - \alpha \Delta t} \hat{X}_{i,\Delta t} \\
+ 2 (X_{i,\Delta t})^T \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} - I \right) \frac{I - A \Delta t}{1 - \alpha \Delta t} B_1 + 2 (X_{i,\Delta t} - \hat{X}_{i,\Delta t})^T \frac{I - A \Delta t}{1 - \alpha \Delta t} B_1 \\
+ 2 \left( (X_{i,\Delta t})^T \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} - I \right) \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} + I \right) \right) B_2 + 2 B_1^T B_2.
\]

Notice that the matrix \( \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} - I \right) \left( \frac{e^{-A\Delta t}}{1 - \alpha \Delta t} + I \right) \) is non-positive definite. Now we will estimate separately:

\[
E (B_2^T B_2) = E \left( |B_2|^2 \right) \\
= \frac{1}{(1 - \alpha \Delta t)^2} E \left( \int_{i\Delta t}^{(i+1)\Delta t-t} \left| e^{A(s+t-(i+1)\Delta t)} \hat{g} (X^t_{s+t}) - g \left( \hat{X}_{i,\Delta t} \right) \right|^2 dW_s \right) \\
\leq \frac{1}{(1 - \alpha \Delta t)^2} \int_{i\Delta t}^{(i+1)\Delta t-t} \left| e^{A(s+t-(i+1)\Delta t)} \hat{g} (X^t_{s+t}) - g \left( \hat{X}_{i,\Delta t} \right) \right|^2 ds \\
\leq \frac{2(\tau + 1)}{\tau (1 - \alpha \Delta t)^2} \int_{i\Delta t}^{(i+1)\Delta t} \left| e^{A(s-(i+1)\Delta t)} - I \right|^2 E \left( |g (X^t_s)|^2 \right) ds \\
+ \frac{2(\tau + 1)}{\tau (1 - \alpha \Delta t)^2} \int_{i\Delta t}^{(i+1)\Delta t} \left| g (X^t_s) - g \left( \hat{X}_{i,\Delta t} \right) \right|^2 ds \\
+ \frac{1 + \tau}{(1 - \alpha \Delta t)^2} \int_{i\Delta t}^{(i+1)\Delta t} \left| g (X^t_s) - g \left( \hat{X}_{i,\Delta t} \right) \right|^2 ds,
\]

where \( \tau \) is a small number from applying Young’s inequality, which will be fixed later. Now, integrating and using Lemma 2.1,

\[
\int_{i\Delta t}^{(i+1)\Delta t} \left| e^{A(s-(i+1)\Delta t)} - I \right|^2 ds \\
= Tr \left( \int_{i\Delta t}^{(i+1)\Delta t} \left( e^{2A(s-(i+1)\Delta t)} + I - 2e^{A(s-(i+1)\Delta t)} \right) ds \right) \\
= Tr \left( (2A)^{-1} \left( 4 (e^{-A\Delta t} - I + A \Delta t) - (e^{-2A\Delta t} - I + 2A \Delta t) \right) \right)
\]
\[
\begin{align*}
    &\leq Tr\left( (2A)^{-1} \left( 4 \left( I - A\Delta t + \frac{1}{2} A^2 \Delta t^2 - I + A\Delta t \right) \right) \right) \\
    &\quad - Tr\left( (2A)^{-1} \left( I - 2A\Delta t + 2A^2 \Delta t^2 - \frac{8}{6} A^3 \Delta t^3 - I + 2A\Delta t \right) \right) \\
    &\leq \frac{2}{3} (\Delta t)^3 Tr\left( A^2 \right).
\end{align*}
\]

So,

\[
E\left(B_2^T B_2 \right) \leq K_7 (\Delta t)^2 + \frac{(1 + \tau) \beta_{2}^2 \Delta t}{(1 - \alpha \Delta t)^2} E\left( \left| X_{i\Delta t}^t - \hat{X}_{i\Delta t}^t \right|^2 \right).
\]

Similarly, we estimate the following term:

\[
E\left( B_1^T B_1 \right) = E\left( |B_1|^2 \right)
\]

\[
= \frac{1}{(1 - \alpha \Delta t)^2} E\left( \left( \int_{i\Delta t}^{(i+1)\Delta t} \left( e^{A(s-(i+1)\Delta t)} f \left( X_s^t \right) - f \left( \hat{X}_{i\Delta t}^t \right) \right) ds \right)^2 \right)
\]

\[
\leq \frac{1}{(1 - \alpha \Delta t)^2} \left( \int_{i\Delta t}^{(i+1)\Delta t} \left\| e^{A(s-(i+1)\Delta t)} f \left( X_s^t \right) - f \left( \hat{X}_{i\Delta t}^t \right) \right\|^2 ds \right)^2
\]

\[
\leq \frac{1}{(1 - \alpha \Delta t)^2} \left( \int_{i\Delta t}^{(i+1)\Delta t} \left| e^{A(s-(i+1)\Delta t)} - I \right| \left\| f \left( X_s^t \right) \right\| ds \right.
\]

\[
+ \int_{i\Delta t}^{(i+1)\Delta t} \left\| f \left( X_s^t \right) - f \left( \hat{X}_{i\Delta t}^t \right) \right\| ds \right.
\]

\[
+ \int_{i\Delta t}^{(i+1)\Delta t} \left\| f \left( \hat{X}_{i\Delta t}^t \right) - f \left( \hat{X}_{i\Delta t}^t \right) \right\|^2 ds \right).}
Similarly, we can estimate the expectation of the fifth term of (4.3):

\[ E \left( \left( X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \right)^T e^{-A\Delta t} \left( \frac{e^{-A\Delta t} - I + A\Delta t}{1 - \alpha\Delta t} \right) \hat{X}^t_{i\Delta t} \right) \leq K_{13} (\Delta t)^2 \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2. \]

Estimation of the expectation of the sixth term gives:

\[ E \left( (X^t_{i\Delta t})^T \left( \frac{e^{-A\Delta t} - I + A\Delta t}{1 - \alpha\Delta t} \right) B_1 \right) \]

\[ \leq \| X^t_{i\Delta t} \|_2 \left| \frac{e^{-A\Delta t} - I + A\Delta t}{1 - \alpha\Delta t} \right| B_1 \|_2 \]

\[ \leq K_9 (\Delta t)^2 - K_{10} (\Delta t)^{3/2} + \beta_1 (\Delta t) \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2 \]

\[ \leq K_{11} (\Delta t)^{5/2} + \beta_1 K_9 (\Delta t)^3 \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2. \]

Estimation of the expectation of the seventh term gives:

\[ E \left( \left( X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \right)^T \left( \frac{I - A\Delta t}{1 - \alpha\Delta t} \right) B_1 \right) \leq \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2 \left| \frac{I - A\Delta t}{1 - \alpha\Delta t} \right| B_1 \|_2 \]

\[ + \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2 |A\Delta t| \left| \frac{I - A\Delta t}{1 - \alpha\Delta t} \right| B_1 \|_2 \]

\[ \leq \frac{\beta_1 (\Delta t)^2}{(1 - \alpha\Delta t)^2} E \left( \left| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \right|^2 \right) (1 + \Delta t |A|) \]

\[ + K_{10} (\Delta t)^{3/2} \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2 (1 + \Delta t |A|). \]

Expectation of the eighth term of (4.3) is zero. Finally, expectation of the last term of (4.3) is given by:

\[ E \left( 2B_1^2 B_2 \right) \leq 2 \left( \sqrt{K_7} \Delta t + \frac{\sqrt{1 + \tau_2\beta_2 \Delta t}}{1 - \alpha\Delta t} \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2 \right) \]

\[ \times \left( \sqrt{K_8} (\Delta t)^{3/2} + \frac{\beta_1 \Delta t}{1 - \alpha\Delta t} \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2 \right) \]

\[ \leq K_{14} (\Delta t)^{5/2} + K_{15} (\Delta t)^{3/2} \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2. \]

Now we will combine all above estimates. Notice that in the above estimates, the term \( E \left( \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|^2 \right) \) has coefficients, the largest of which contains a constant multiplied by \( \Delta t \). The largest free term contains a constant multiplied by \( (\Delta t)^2 \). Choosing \( \tau \) and \( \Delta t \) sufficiently small, and again applying Young’s inequality for the term \( (\Delta t)^{3/2} \| X^t_{i\Delta t} - \hat{X}^t_{i\Delta t} \|_2 \), give the following:

\[ (1 - \alpha\Delta t)^{-2M} E \left( \| X^t_{M\Delta t} - \hat{X}^t_{M\Delta t} \|^2 \right) \]
\[
\leq \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} \left( \frac{\Delta t}{1 - \alpha \Delta t} \right)^{2} \left| X_{1}\Delta t - \hat{X}_{1}\Delta t \right|^2 + K_{16} (\Delta t)^2, \]
where the constant \(\delta\) can be chosen such that: \(2\beta_2 + \beta_2^2 + \delta < 2\alpha\). Note also that we choose \(\Delta t\) small enough such that \(2\beta_2 + \beta_2^2 + \delta + \alpha^2 \Delta t < 2\alpha\). Continuing the inequality, we get:

\[
(1 - \alpha \Delta t)^{-2M} E \left( \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} \left| X_{1}\Delta t - \hat{X}_{1}\Delta t \right|^2 \right) \leq K_{17} \Delta t.
\]

Here we apply the discrete version of Gronwall’s Lemma as in Willett and Wong [25] to get:

\[
E \left( \sum_{i=0}^{M-1} (1 - \alpha \Delta t)^{-2i} \left| X_{1}\Delta t - \hat{X}_{1}\Delta t \right|^2 \right) \leq K_{18} \Delta t.
\]

Note that the constant \(K_{18}\) does not depend on \(M, \Delta t\). Now, taking \(M = N\) and noting that \(N \Delta t = t\), we have:

\[
\lim_{t \to \infty} \sup_{t} \left| X_{t} - \hat{X}_{t} \right| \leq \sqrt{K_{18} \Delta t}. \]

**Corollary 4.2.** If \(X^*(\omega)\) and \(\hat{X}_{\Delta t}^*(\omega)\) are respectively the exact and the numerical stationary points of the solution of equation (1.4), where the numerical solution was obtained using the time-step \(\Delta t\), then

\[
\left| X^* - \hat{X}_{\Delta t} \right| \leq K_{19} \sqrt{\Delta t}.
\]

**Proof.** Since

\[
\left| X^* - \hat{X}_{\Delta t} \right| = \lim_{t \to \infty} \left| X^* - \hat{X}_{\Delta t} \right| \leq \lim_{t \to \infty} \left| X^* - X_{t} \right| + \lim_{t \to \infty} \left| X_{t} - \hat{X}_{t} \right|
\]

and

\[
\lim_{t \to \infty} \left| X^* - X_{t} \right| = 0,
\]

\[
\lim_{t \to \infty} \left| \hat{X}_{t} - \hat{X}_{\Delta t} \right| = 0,
\]

the result follows. \(\square\)

**Corollary 4.3.** If \(X^*(\omega)\) and \(\hat{X}_{\Delta t}^*(\omega)\) are respectively the exact and the numerical stationary points of the solution of equation (1.4), where the numerical solution was obtained using the time-step \(\Delta t\), then:

\[
\lim_{\Delta t \to 0} \left| X^* - \hat{X}_{\Delta t} \right| = 0.
\]
Also, for almost all $\omega \in \Omega$ and for rational $\Delta t$:
\[
\lim_{\Delta t \to 0} \left| X^* - \hat{X}^*_\Delta t \right| = 0.
\]

**Proof.** The first statement is obvious. To prove the second one, for any fixed $Z \in \mathbb{Q}$, $0 \leq Z \leq 1$, assume
\[
\Delta t = \frac{1}{(n+1)^2} + \left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right) Z.
\]
From Chebyshev inequality we have:
\[
P \left( \left| X^* - \hat{X}^*_\Delta t \right| \geq n^{-1/4} \right) \leq K_{20} \sqrt{n} \left( \frac{1}{(n+1)^2} + \left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right) Z \right),
\]
and using Borel-Cantelli lemma we get:
\[
\left| X^* - \hat{X}^*_\Delta t \right| \to 0
\]
almost surely for $n \to \infty$. Since this is true for any fixed $Z \in \mathbb{Q}$, $0 \leq Z \leq 1$, then for any sufficiently small rational $\Delta t$, $X^* - \hat{X}^*_\Delta t \to 0$ as $\Delta t \to 0$. \qed

**Acknowledgments.** The authors would like to thank the referees for useful comments.

**REFERENCES**


