STOCHASTIC EPIDEMIC MODELS DRIVEN BY STOCHASTIC ALGORITHMS WITH CONSTANT STEP

LIFENG CHEN AND JIFA JIANG*

Mathematics and Science College
Shanghai Normal University
Shanghai 200234, China

Abstract. In this paper we propose the stochastic epidemic model that relates directly to the deterministic counterpart and reveal close connections between these two models. Under the classic assumptions, the sample path of process eventually converges to the disease-free equilibrium, even though the corresponding deterministic flow converges to an endemic equilibrium. From the fact that disease can occur sporadically, we adjust the stochastic model slightly by introducing a stochastic incidence and establish precise connections between the long-run behavior of the discrete stochastic process and its deterministic flow approximation for large populations.

1. Introduction. Recently both deterministic and stochastic models are used to model biological processes involving interacting populations.

In a deterministic SIS epidemic model, individuals in the population are classified according to disease status, either susceptible or infectious. A susceptible individual, after a successful contact with an infectious individual, becomes infectious, but does not confer immunity to the disease (e.g., sexually transmitted diseases) in the population. Hence, after recovery, infected individuals return to the susceptible class.

It is well-known that growth rate in population biology equals birth rate subtracted by death rate, which implies that both birth rate and death rate are the same when population number gains its maximum. It is reasonable to assume birth rate equals death rate when population number is around its maximum.

Let $S(t)$ and $I(t)$, respectively, denote the number of susceptible and infected individuals at time $t$. We assume that there are no disease-related deaths and all newborn infants are susceptible. The classic epidemic model of SIS type with vital dynamics (births and deaths) is given by the following standard incidence form:

$$
\begin{align*}
\frac{dS}{dt} &= bN - bS - \beta \frac{S}{N} I + \gamma I, \\
\frac{dI}{dt} &= \beta \frac{S}{N} I - bI - \gamma I,
\end{align*}
$$

(1)

2010 Mathematics Subject Classification. Primary: 92Bxx; Secondary: 37Mxx.

Key words and phrases. Epidemic model, Markov chain, stochastic algorithm, invariant measure.

This work is supported by the National Natural Science Foundation of China (NSFC) under Grant No.11371252, Research and Innovation Project of Shanghai Education Committee under Grant No.14zz120, and Shanghai Gaofeng Project for University Academic Program Development.

*Corresponding author.
where $N = S(t) + I(t)$ is the total population size, parameter $b > 0$ is the natural birth rate, $\gamma > 0$ is the recovery rate, $\beta > 0$ is the contact rate. This SIS epidemic model implies that there is an inflow of newborns into the susceptible class at rate $bN$. As mentioned before, we have assumed that the birth rate is equal to the death rate. Here $(S, I) \in D_N = \{(S, I) : 0 \leq S \leq N, 0 \leq I \leq N, S + I \leq N\}.

The dynamics of model (1) is determined by the basic reproduction number $R_0 = \frac{\beta}{b+\gamma}$ (see, e.g., [9]).

**Theorem 1.1.** Let $(S(t), I(t))$ be a solution of (1).

(i) If $R_0 \leq 1$, then $P_0 = (N, 0)$ is the unique equilibrium in $D_N$ and it is globally asymptotically stable.

(ii) If $R_0 > 1$, then there is a unique positive equilibrium $P^* = (N \frac{1}{R_0}, N(1 - \frac{1}{R_0}))$ and it is globally asymptotically stable with respect to positive solutions.

Let $s(t) = \frac{S(t)}{N}$ and $i(t) = \frac{I(t)}{N}$ be the susceptible and infectious fractions, respectively. Dividing the equations (1) by the constant total population size $N$ yields

$$
\begin{align*}
\frac{ds}{dt} &= b - bs - \beta si + \gamma i, \\
\frac{di}{dt} &= \beta si - bi - \gamma i.
\end{align*}
$$

(2)

Analogously, we have

**Theorem 1.2.** Let $(s(t), i(t))$ be a solution of (2).

(i) If $R_0 \leq 1$, then there is a unique equilibrium $p_0 = (1, 0)$ in $D_1$ and it is globally asymptotically stable.

(ii) If $R_0 > 1$, then $p^* = (\frac{1}{R_0}, 1 - \frac{1}{R_0})$ is a unique positive equilibrium in $D_1 \setminus \{0\}$ and it is globally asymptotically stable for the positive initial conditions.

Here $p_0 = (1, 0)$ and $p^* = (\frac{1}{R_0}, 1 - \frac{1}{R_0})$ are called disease-free equilibrium and endemic equilibrium, respectively.

In fact, we can merely consider one independent variable $i$, because $s + i$ is asymptotic to 1 as $t \to \infty$. Hence, the model in (2) can be replaced by the following form:

$$
\frac{di}{dt} = \beta(1 - i)i - bi - \gamma i.
$$

(3)

Recent works by Li and Xu [11] have shown the precise relationship between the deterministic and stochastic SIS epidemic models. The present paper is motivated by their works. Our aim in this paper is the same as theirs but with different approach, which relies on the theory of stochastic approximations, while Li and Xu’s works rest on master equation.

The paper is organized as follows. Under the probabilistic assumptions on coefficients, the variant stochastic epidemic model is established in Section 2. The theme of this paper is presented in Section 3. We give a sufficient condition to guarantee a close connection between sample paths of the stochastic model and trajectories of the corresponding deterministic average differential equation. Section 4 gives more applications for our theoretic results.

The following notations and definitions will be used throughout. Let $M$ be a given metric space. Let $\mathcal{P}(M)$ and $C_b(M)$ denote the set of probability measures and real-valued bounded continuous functions on $M$, respectively. A probability $\mu \in \mathcal{P}(M)$ is $P$ invariant (or Markov chain $X$ invariant) if it satisfies $\mu P = \mu$, i.e., $\int_M P(x, B) \mu(dx) = \mu(B)$, $\forall B \in \mathcal{B}(M)$, where $P(x, B)$ is a Markov transition probability function of $X$ and $\mathcal{B}(M)$ is the Borel $\sigma$-algebra on $M$. 


A family \( \mathcal{L} \subset \mathcal{P}(M) \) is tight if for every \( \eta > 0 \) there exists a compact set \( K \subset M \) such that \( \mu(K) \geq 1 - \eta \) for all \( \mu \in \mathcal{L} \). When \( M \) is a Polish space (i.e., a separable complete metric space), by the Prohorov theorem, a family \( \mathcal{L} \subset \mathcal{P}(M) \) is relatively compact for the topology of weak* convergence if and only if \( \mathcal{L} \) is tight.

2. Stochastic epidemic models. We are interested in the infectious fraction \( \frac{I}{N} \), when the total population \( N \) is large enough. So it is like the biochemistry’s “large volume, small molecules” model. It is suitable to consider the discrete time stochastic model.

The total population \( N \) is fixed. Time is discrete. At any time \( k \in \mathbb{N} \), let \( S_k \) and \( I_k \) denote random variables for the number of susceptible and infected individuals at time \( k \), respectively. Since \( S_k = N - I_k \), there is only one independent random variable \( I_k \). The stochastic process \( \{ I_k : k \in \mathbb{N} \} \) has the Markov property, that is,

\[
P(I_{k+1}|I_0, \ldots, I_k) = P(I_{k+1}|I_k),
\]

for any time \( k \in \mathbb{N} \). The Markov property says that the process at time \( k+1 \) only depends on the process at current time \( k \).

We make one more assumption: at any time \( k \in \mathbb{N} \), the number of infected individuals changes by at most one during time interval \( k \to k+1 \), that is,

\[
i \to i+1, \ i \to i-1, \ i \to i.
\]

Let \( X_k = \frac{I_k}{N} \in \{0, \frac{1}{N}, \ldots, \frac{N-1}{N}, 1\} \) denote the infectious fraction at time \( k \) and it is also a random variable. The assumptions of the process \( \{ I_k : k \in \mathbb{N} \} \) imply that the process \( \{ X_k : k \in \mathbb{N} \} \) enjoys the Markov property and the random variable \( X_{k+1} - X_k \) takes values in \( \{-\frac{1}{N}, 0, \frac{1}{N}\} \). The basic object of interest is the stochastic process \( \{ X_k : k \in \mathbb{N} \} \) taking values in \( [0,1] \), particularly in its long-run behavior.

Under the probabilistic assumptions, the stochastic model can be formulated as follows. Given the time \( k \in \mathbb{N} \) and the state \( X_k = x \in [0,1] \).

The conditional probability of a new infection, \( x \to x + \frac{1}{N} \), equals the probability that an individual (called \( Q_k \)) is chosen from the susceptible class at time \( k \), which is \( 1-x \), times the probability that \( Q_k \) becomes infected at next time \( k+1 \), which consists of the probability that \( Q_k \) meets other individual, says \( A(x) \in [0,1] \) (assumed to depend on \( x \)), conditioned on such a meeting, the probability that \( Q_k \) meets the infective class, which is \( x \), times the probability \( \beta_0 \in (0,1] \) that this contact infects \( Q_k \). Hence, the conditional probability \( \mathbb{P}(X_{k+1} = x + \frac{1}{N}|X_k = x) = (1-x)x\tilde{\beta}(x) \), where \( \tilde{\beta}(x) = \beta_0 A(x) \).

The conditional probability of a recovery or death, \( x \to x - \frac{1}{N} \) is the probability that an individual (called \( Q_k \)) is chosen from the infective class at time \( k \), which is \( x \), times the probability that \( Q_k \) is susceptible or death at time \( k+1 \), i.e., it comes from the probability of recovery which is \( \tilde{\gamma}(x) \) (assumed to depend on \( x \)) or the probability of a death which is \( \tilde{\mu} \in (0,1] \). That is, \( \mathbb{P}(X_{k+1} = x - \frac{1}{N}|X_k = x) = x(\tilde{\gamma}(x) + \tilde{\mu}) \).

Finally, the conditional probability of no change (the only other possibility), \( x \to x \) is \( 1 - [(1-x)x\tilde{\beta}(x) + x(\tilde{\gamma}(x) + \tilde{\mu})] \).

For simplicity, we choose \( A(x) \) and \( \tilde{\gamma}(x) \) are both positive constants such that \( 1 - [(1-x)x\tilde{\beta}(x) + x(\tilde{\gamma}(x) + \tilde{\mu})] \) lies in the interval \([0,1]\). This implies \( \tilde{\beta} > 0 \) and \( \tilde{\gamma} + \tilde{\mu} > 0 \) are both constants. These assumptions make the stochastic epidemic model be a useful approximation to the model (3). Without causing confusion, we write \( \tilde{\beta}, \tilde{\gamma} \) and \( \tilde{\mu} \) as \( \beta, \gamma \) and \( \mu \), respectively.
In other words, for given $k \in \mathbb{N}$ and $x \in [0, 1]$, there is only three possible states $x + \frac{1}{N}$, $x - \frac{1}{N}$ and $x$, thus we have

$$P(X_{k+1} - X_k = y | X_k = x) = \begin{cases} 
\beta(1-x)x, & y = \frac{1}{N}; \\
(b+\gamma)x, & y = -\frac{1}{N}; \\
1 - [\beta(1-x)x + (b+\gamma)x], & y = 0; \\
0, & y \neq \frac{1}{N}, 0, -\frac{1}{N}.
\end{cases}$$

For each $N \in \mathbb{N}^*$, the process $X^N = \{X^N_k : k \in \mathbb{N}\}$ is a discrete-time Markov chain taking value in a finite lattice $L^N = \{0, \frac{1}{N}, \ldots, \frac{N-1}{N}, 1\}$. The properties of the process $X^N$ follow easily from Markov chain theory. The state $\{0\}$ is an absorbing state and the other states $\{\frac{1}{N}, \ldots, \frac{N-1}{N}, 1\}$ are transient. Thus it has a unique invariant measure $\mu^N = (1, 0, \ldots, 0) =: (\mu^N_0, \mu^N_1, \ldots, \mu^N_N)$. It is convenient to take $P([0, 1])$ as the common state space for $\mu^N$. In fact let $\mu^N(U) := \sum_{l \in U} \mu^N_l$, then

$$\mu^N(\cdot) = \delta_0(\cdot) \in P([0, 1]).$$

The only limit point of $\mu^N$, in the weak* topology, is the Dirac measure at point 0, i.e., $\mu^N \rightharpoonup \delta_0$, as $N \to \infty$. The Dirac measure $\delta_0$ is invariant for (3) and the support of $\delta_0$ is the disease-free equilibrium of (3).

This implies that the sample paths of $X^N$ approach the disease-free equilibrium (as $N \to \infty$), regardless of the magnitude of the basic reproduction number.

This is the most important difference between the deterministic and stochastic epidemic models.

As we know that the disease cases can occur sporadically in the host population with a small probability $\alpha$ due to environmental variability, e.g., transmission by another species, contaminated blood banks, foreign travel, etc. Therefore we make a small adjustment with a stochastic incidence into the model (see, e.g., [4, 11]): the conditional probability of a new infection also comes from the susceptible class (which is $1-x$), times the small probability $\alpha$. Because for large populations, we assume for simplicity that the conditional probability $\alpha \ll 1$ is constant.

Assume that $\alpha > 0$. Then the previous epidemic models are modified slightly, for $k \in \mathbb{N}$, $x \in L^N \subset [0, 1],$

$$P(X_{k+1} - X_k = y | X_k = x) = \begin{cases} 
\beta(1-x)x + \alpha(1-x), & y = \frac{1}{N}; \\
(b+\gamma)x, & y = -\frac{1}{N}; \\
1 - [\beta(1-x)x + (b+\gamma)x + \alpha(1-x)], & y = 0; \\
0, & y \neq \frac{1}{N}, 0, -\frac{1}{N}.
\end{cases}$$

It is easy to see from the transition diagram in Figure 1 that the Markov chain $\{X_k\}_{k \in \mathbb{N}}$ is irreducible, that is to say, there is nonzero probability of eventually going from any state to any other state. Furthermore, it follows from positive recurrence that $\{X_k\}_{k \in \mathbb{N}}$ is a finite irreducible Markov chain. Therefore from Perron-Frobenius theorem it admits a unique invariant probability measure $\mu^N = (\mu^N_0, \mu^N_1, \ldots, \mu^N_N)$, where $\mu^N_1 > 0, l = 0, \frac{1}{N}, \ldots, 1$. As total population $N$ goes to $\infty$, the variable $x$ is of course viewed as real number in $[0, 1]$. The conditional
expected value of $X_{k+1} - X_k$ is easily calculated,
\[
\mathbb{E}[X_{k+1} - X_k | X_k = x] = \frac{1}{N} [\beta(1 - x)x + \alpha(1 - x) - (b + \gamma)x] =: \frac{1}{N} F(x),
\]
where $F(x) = \beta(1 - x)x + \alpha(1 - x) - (b + \gamma)x$.

Let $F_k$ denote the $\sigma$-algebra generated by the random variables $X_0, \cdots, X_k$ and $U_{k+1} := N(X_{k+1} - \mathbb{E}[X_{k+1} | X_k])$. Manipulating the equation (4), we obtain the recursion formula
\[
X_{k+1} - X_k = \frac{1}{N} F(X_k) + X_{k+1} - \mathbb{E}[X_{k+1} | X_k] = \frac{1}{N} [F(X_k) + U_{k+1}].
\]
By the assumption that the random variable $X_{k+1} - X_k$ takes values in $\{-\frac{1}{N}, 0, \frac{1}{N}\}$,
\[
|U_k| \leq \max_{x \in [0,1]} |F(x)| + 1.
\]

It is not difficult to check that $\{U_k\}_{k \in \mathbb{N}}$ is a martingale difference sequence, i.e., $U_k$ is measurable with respect to $F_k$ and satisfies
\[
\mathbb{E}[U_{k+1} | F_k] = \mathbb{E}[U_{k+1} | X_k] = 0.
\]

Let $\epsilon = \frac{1}{N}$. Then a family $\{X_k^\epsilon\}_{k \in \mathbb{N}}$ of processes, parametrized by $\epsilon > 0$, satisfies recursion of the form
\[
X_{k+1}^\epsilon - X_k^\epsilon = \epsilon [F(X_k^\epsilon) + U_{k+1}^\epsilon],
\]
where $\{U_k^\epsilon\}_{k \in \mathbb{N}}$ is a martingale difference sequence.

Recursive processes described by (6) arise frequently in a large variety of fields such as control theory, game theory, biological modeling and elsewhere under the name stochastic approximations, which are widely studied in [2, 3, 8].

3. Rigorous relationship between the stochastic and deterministic models. We note that $\mathbb{E}[U_{k+1}^\epsilon | F_k^\epsilon] = 0$ implies $\mathbb{E}U_k^\epsilon = 0$ for $k \geq 1$. Therefore, one can think of (6) as a kind of Euler approximation scheme for numerically solving the following differential equation
\[
\frac{dx}{dt} = F(x).
\]

Here we give some conditions to make this intuition precisely.

**Hypothesis 1.** $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a Lipschitz vector field. That is, there exists a constant $L > 0$ such that
\[
\|F(x) - F(y)\| \leq L\|x - y\|, \text{ for all } x, y \in \mathbb{R}^m.
\]

The Hypothesis 1 implies that $F$ generates a flow
\[
\Phi : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m,
\]
\[
(t, x) \mapsto \Phi(t, x) = \Phi_t(x),
\]
where $t \mapsto \Phi_t(x_0)$ is the solution to (7) with initial condition $x(0) = x_0$. 

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{transition_diagram.png}
\caption{Transition diagram of Markov chain $\{X_k\}_{k \in \mathbb{N}}$, where $b(l) = \beta(1 - l^2 + \alpha(1 - l), d(l) = (b + \gamma)l$, for $l = 0, 1, \cdots, N^{-1}, 1$.}
\end{figure}
In order to analyse the behavior of \( \{X_k^t\}_{k \in \mathbb{N}} \) in terms of the behavior of the flow \( \Phi \) it is convenient to introduce the interpolated process \( \overline{X}^{\epsilon,x} : \mathbb{R}^m \to \mathbb{R}^m \) defined by

\[
\overline{X}^{\epsilon,x}(0) = x \quad \text{and} \quad \overline{X}^{\epsilon,x}(t) = X_k^t + \frac{t - k\epsilon}{\epsilon}[X_{k+1}^t - X_k^t], \quad t \in [k\epsilon, (k+1)\epsilon].
\]

**Hypothesis 2.** \( \sup_{k,\epsilon} \sqrt{\mathbb{E}\|U_k^t\|^2} =: D < \infty \).

It follows from (5) that Hypothesis 2 holds.

The following proposition is a standard averaging result for the stochastic approximation processes.

**Proposition 3.1.** Assume that Hypotheses 1 and 2 hold. Let \( K \subset \mathbb{R}^m \) be a compact set. Then for any \( T > 0 \) and \( \delta > 0 \), we have

\[
\lim_{t \to 0} \sup_{x \in K} \mathbb{P}\{ \sup_{0 \leq t \leq T} \|\overline{X}^{\epsilon,x}(t) - \Phi_t(x)\| \geq \delta \} = 0.
\]

**Proof.** Let \( t \in [0, T] \). A straightforward computation shows that \( \overline{X}^{\epsilon,x}(t) \) satisfies

\[
\overline{X}^{\epsilon,x}(t) = x + \int_0^t \left[ F(\overline{X}^{\epsilon,x}(\epsilon^{-1} s)) + U_{\lceil \frac{s}{\epsilon} \rceil + 1} \right] ds,
\]

where \( \lceil \frac{s}{\epsilon} \rceil \) denotes the integral part of \( \frac{s}{\epsilon} \). And since

\[
\Phi(t, x) = x + \int_0^t F(\Phi(\epsilon^{-1} s, x)) ds + \int_0^t [F(\Phi(s, x)) - F(\Phi(\epsilon^{-1} s, x))] ds,
\]

subtracting (9) from (8) and taking sup-norm, we obtain

\[
\sup_{0 \leq \tau \leq t} \|\overline{X}^{\epsilon,x}(\tau) - \Phi(\tau, x)\| \leq \sup_{0 \leq \tau \leq t} \int_0^\tau \|F(\overline{X}^{\epsilon,x}(\epsilon^{-1} s)) - F(\Phi(\epsilon^{-1} s, x))\| ds
\]

\[
+ \sup_{0 \leq \tau \leq t} \|\int_0^\tau U_{\lceil \frac{s}{\epsilon} \rceil + 1} ds\|
\]

\[
+ \sup_{0 \leq \tau \leq t} \int_0^\tau \|F(\Phi(s, x)) - F(\Phi(\epsilon^{-1} s, x))\| ds.
\]

Our next goal is to estimate the order of \( L^2 \) norm on the left-hand side of the above inequality. So we consider each term’s \( L^2 \) norm on the right-hand side. By Hypothesis 1, we obtain

\[
\sup_{0 \leq \tau \leq t} \int_0^\tau \|F(\overline{X}^{\epsilon,x}(\epsilon^{-1} s)) - F(\Phi(\epsilon^{-1} s, x))\| ds
\]

\[
\leq L \int_0^t \|\overline{X}^{\epsilon,x}(\epsilon^{-1} s) - \Phi(\epsilon^{-1} s, x)\| ds
\]

\[
= L \left[ \sum_{k=0}^{\lceil \frac{t}{\epsilon} \rceil - 1} \epsilon \|\overline{X}^{\epsilon,x}(k\epsilon \epsilon^{-1}) - \Phi(k\epsilon, x)\| + \int_{\lceil \frac{t}{\epsilon} \rceil \epsilon}^t \|\overline{X}^{\epsilon,x}(\epsilon^{-1} s) - \Phi(\epsilon^{-1} s, x)\| ds \right]
\]

\[
\leq L \epsilon \sum_{k=0}^{\lceil \frac{t}{\epsilon} \rceil} \sup_{0 \leq \tau \leq k\epsilon} \|\overline{X}^{\epsilon,x}(\epsilon \tau) - \Phi(\epsilon \tau, x)\|.
\]

It is easy to see that

\[
\mathbb{E}(\sup_{0 \leq \tau \leq t} \int_0^\tau \|F(\overline{X}^{\epsilon,x}(\epsilon^{-1} s)) - F(\Phi(\epsilon^{-1} s, x))\| ds)^2)^{\frac{1}{2}}
\]
exists a positive constant $2$ in the last two inequalities, $B$ follows:

$$
\text{Let } \xi \text{ as follows: where in the last inequality, we have used so-called “area principle”, which is stated for any } \xi, \text{ for any } \xi \geq i, \text{ one has }
$$

$$
|\sum_{k=i}^{\xi} f(k) - \int_i^\xi f(r)dr| \leq f(\xi).
$$

Therefore, the last inequality is a consequence of the “area principle” with $i = 0$, $\xi = t$ and $f(r) = E(\sup_{0 \leq \tau \leq t} ||X^{\tau,x}(\tau) - \Phi(\tau,x)||^2)$. We continue to estimate the second term’s $L^2$ norm on the right-hand side. A simple computation yields

$$
E(\sup_{0 \leq \tau \leq t} ||\int_0^\tau U_\tau^{\tau,x}ds||^2)^{1/2}
$$

$$
= E(\sup_{0 \leq \tau \leq t} ||\sum_{k=0}^{n-1} \epsilon U_{\tau+1} + \int_{[\tau]}^\tau U_{\tau+1}ds||^2)^{1/2}
$$

$$
\leq E(\sup_{0 \leq \tau \leq t} ||\sum_{k=0}^{n-1} \epsilon U_{\tau+1}|| + \sup_{0 \leq \tau \leq t} ||\int_{[\tau]}^\tau U_{\tau+1}ds||^2)^{1/2}
$$

$$
\leq E(\sup_{0 \leq \tau \leq t} ||\sum_{k=0}^{n-1} \epsilon U_{\tau+1}||^2)^{1/2} + E(\sup_{0 \leq \tau \leq t} ||\int_{[\tau]}^\tau U_{\tau+1}ds||^2)^{1/2}
$$

$$
\leq E(\sum_{0 \leq n \leq \tau} ||U_{n+1}||^2)^{1/2} + \{c^2 E(\sum_{n=0}^{n-1} ||U_{n+1}||^2)\}^{1/2}
$$

$$
\leq B_2 E \sum_{k=0}^{n-1} E[|U_{\tau+1}|^2]^2 + \{c^2 \sum_{n=0}^{n-1} E[|U_{\tau+1}|^2]\}^{1/2}
$$

$$
\leq \sqrt{T} B_2 D + \sqrt{c(T+1)D},
$$

where we have used Burkholder-Davis-Gundy’s inequality (see, [6]) and Hypothesis 2 in the last two inequalities, $B_2 > 0$ is a universal constant. Also, by the Lipschitz assumption of $F$ and the continuity of $\Phi(t,x)$ in a compact set $[0,T] \times K$, there exists a positive constant $2A(T)$ depending on $T$, such that $||F(\Phi(t,x))|| \leq 2A(T)$ for any $(t,x) \in [0,T] \times K$. Thus we have

$$
\sup_{0 \leq \tau \leq t} \int_0^\tau ||F(\Phi(s,x)) - F(\Phi(\epsilon_s^\tau, x))||ds
$$

$$
\leq L \int_0^t ||\Phi(s,x) - \Phi(\epsilon_s^\tau, x)||ds
$$
∀ Then Gronwall’s inequality yields that, The conclusion follows immediately from Chebyshev’s inequality.

Theorem 3.2. Let \( \text{Theorem 2.2} \), we can prove the next consequence for flow invariance theorem.

\[ \text{Proof.} \] In other words, for every \( \epsilon \) sufficiently small, we obtain

\[ \left[ \mathbb{E}(\sup_{0 \leq \tau \leq t} \| X^{\epsilon,x}(\tau) - \Phi(\tau,x) \|)^2 \right]^{1/2} \]

\[ \leq L \int_0^t \mathbb{E}(\sup_{0 \leq \tau \leq s} \| X^{\epsilon,x}(\tau) - \Phi(\tau,x) \|^2)ds + L\epsilon \int_0^t \mathbb{E}(\sup_{0 \leq \tau \leq s} \| X^{\epsilon,x}(\tau) - \Phi(\tau,x) \|^2)^{1/2}
\]

\[ + \sqrt{TB_2D} + \epsilon(T+1)D + \epsilon L(T+1)A(T), \forall t \in [0,T]. \]

In other words, for every \( \epsilon \) sufficiently small, we obtain

\[ \left[ \mathbb{E}(\sup_{0 \leq \tau \leq t} \| X^{\epsilon,x}(\tau) - \Phi(\tau,x) \|)^2 \right]^{1/2} \]

\[ \leq \frac{L}{1-\epsilon L} \int_0^t \mathbb{E}(\sup_{0 \leq \tau \leq s} \| X^{\epsilon,x}(\tau) - \Phi(\tau,x) \|^2)ds
\]

\[ + \frac{1}{1-\epsilon L} \left[ \sqrt{TB_2D} + \epsilon(T+1)D + \epsilon L(T+1)A(T) \right], \forall t \in [0,T]. \]

Then Gronwall’s inequality yields that, \( \forall x \in K \),

\[ \left[ \mathbb{E}(\sup_{0 \leq \tau \leq t} \| X^{\epsilon,x}(\tau) - \Phi(\tau,x) \|)^2 \right]^{1/2} \]

\[ \leq \frac{1}{1-\epsilon L} \left[ \sqrt{TB_2D} + \epsilon(T+1)D + \epsilon L(T+1)A(T) \right] exp\left(\frac{L}{1-\epsilon L}T\right), \forall t \in [0,T]. \]

The conclusion follows immediately from Chebyshev’s inequality.

Following the line of argument given in the proof (see, e.g., [1, Corollary 3.2] or [8, Theorem 2.2]), we can prove the next consequence for flow invariance theorem.

Theorem 3.2. Let Hypotheses 1 and 2 hold. Let \( \mu^\epsilon \) be an invariant probability measure of the Markov chain \( \{X^\epsilon_k\}_{k \in \mathbb{N}} \). Suppose that the family \( \{\mu^\epsilon\}_{\epsilon > 0} \) is tight. Then any limit point \( \mu = \lim_{\epsilon \to 0} \mu^\epsilon \) for the topology of weak* convergence is an invariant measure of \( \Phi \), i.e., \( \mu \circ \Phi_t^{-1} = \mu \) for all \( t \in \mathbb{R} \).

\[ \text{Proof.} \] Fix \( t > 0 \) and let \( n = n(t,\epsilon) := \lfloor \frac{1}{\epsilon} \rfloor \). Let \( X^{\epsilon,x} \) denote the interpolated process with initial condition \( X_0 = x \in \mathbb{R}^m \). For any \( g \in C_b(\mathbb{R}^m) \), since \( \mu^\epsilon \) is \( (P^\epsilon)^n \) invariant, we have

\[ \left| \int_{\mathbb{R}^m} g \circ \Phi_t(x) \mu^\epsilon(dx) - \int_{\mathbb{R}^m} g(x) \mu^\epsilon(dx) \right| \]

\[ = \left| \int_{\mathbb{R}^m} g \circ \Phi_t(x) \mu^\epsilon(dx) - \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} g(y)(P^\epsilon)^n(x,dy) \mu^\epsilon(dx) \right| \]

\[ = \left| \int_{\mathbb{R}^m} g \circ \Phi_t(x) \mu^\epsilon(dx) - \int_{\mathbb{R}^m} E_g(X^{\epsilon,x}(ne)) \mu^\epsilon(dx) \right| \]

\[ \leq \int_{\mathbb{R}^m} |E[g \circ \Phi_t(x) - g(X^{\epsilon,x}(ne))]| \mu^\epsilon(dx) \]
To prove the claim, we use the fact that for any 

\[ g \]

By uniform continuity of \( g \),

\[ \eta > 0. \]

Since the sequence \( \{\mu_i\} \) is tight, there exists a compact set \( K \subset \mathbb{R}^m \) such that \( \inf_{\epsilon_i} \mu_i(K) \geq 1 - \frac{\eta}{\|g\|} \) (i.e., \( \sup_{\epsilon_i} \mu_i(K^c) \leq \frac{\eta}{\|g\|} \)). Hence,

\[
\int_{\mathbb{R}^m} |g \circ \Phi_t(x) - g \circ \Phi_t(x)I_K(x)| \mu_i(dx) \leq \|g\| \frac{\eta}{\|g\|} = \eta, \quad \forall i \in \mathbb{N}.
\]

By uniform continuity of \( g \) on a compact neighborhood of \( \Phi_t(K) \) there exists \( \delta > 0 \) such that \( \forall x \in K \) and \( \|u_\Phi - \Phi_t(x)\| < \delta \) implies \( |g(u) - g(\Phi_t(x))| < \eta \). By continuity of the flow \( \Phi(\cdot) \) on a compact set \([t-1, t] \times K\) there exists \( \delta_1 := \delta_1(\delta) \in (0, \delta) \) such that \( \|\Phi_n(x) - \Phi_t(x)\| < \frac{\delta}{2} \) for any \( x \in K \) and \( \epsilon < \delta_1 \). Hence \( |g(\Phi_n(x)) - g(\Phi_t(x))| < \eta \) for any \( \epsilon < \delta_1, x \in K \). We claim that

\[
\|\Phi_t(x) - \Phi_t(x)I_K(x)\| \geq \frac{\delta}{2}.
\]

To prove the claim, we use the fact that for any \( \epsilon_i < \delta_1 \) and \( x \in K, \|\Phi_{n\epsilon_i}(x) - \Phi_{n\epsilon_i}(x)\| \geq \frac{\delta}{2} \) provided \( \|\Phi_t(x) - \Phi_{n\epsilon_i}(x)\| \geq \frac{\delta}{2} \). That is,

\[
\{\|\Phi_t(x) - \Phi_{n\epsilon_i}(x)\| \geq \frac{\delta}{2}\} \subset \{\|\Phi_{n\epsilon_i}(x) - \Phi_{n\epsilon_i}(x)\| \geq \frac{\delta}{2}\},
\]

for \( \epsilon_i < \delta_1 \) and \( x \in K \). Clearly for \( \epsilon_i < \delta_1 \), \n
\[
\|\Phi_t(x)I_K(x) - g(\Phi_{n\epsilon_i}(x))I_K(x)\| = \|\Phi_t(x)I_K(x) - g(\Phi_{n\epsilon_i}(x))I_K(x)\|_{\{\|\Phi_t(x) - \Phi_{n\epsilon_i}(x)\| < \delta\}}
\]

\[
\leq \eta + 2\|g\| \sup_{x \in K} \sup_{0 \leq s \leq t} \|\Phi_{n\epsilon_i}(s) - \Phi_{n\epsilon_i}(s)\| \geq \frac{\delta}{2}.
\]

Proposition 3.1 yields

\[
\int_{\mathbb{R}^m} |g \circ \Phi_t(x)\mu_i(dx) - \int_{\mathbb{R}^m} g(x)\mu_i(dx)| \leq \eta + \eta + 2\|g\| \sup_{x \in K} \sup_{0 \leq s \leq t} \|\Phi_{n\epsilon_i}(s) - \Phi_{n\epsilon_i}(s)\| \geq \frac{\delta}{2}.
\]

Since \( \eta > 0 \) is arbitrary and \( \mu_i \xrightarrow{w} \mu \), hence

\[
\int_{\mathbb{R}^m} g \circ \Phi_t(x)\mu(dx) = \int_{\mathbb{R}^m} g(x)\mu(dx),
\]

for any \( g \in C_b(\mathbb{R}^m) \). This shows that \( \mu \) is an invariant measure of the flow \( \Phi \).
By the Poincaré recurrence theorem (see, e.g., Mañé [12, Theorem 2.3, p. 29]), we can obtain the following consequence immediately.

**Proposition 3.3.** Assume that \( \mu \) is an invariant probability measure of the flow \( \Phi \). Let \( \text{supp}(\mu) \) denote the support of \( \mu \) and \( B(\Phi) \) be the Birkhoff’s center of \( \Phi \). Then the support of \( \mu \) is contained in the Birkhoff’s center of \( \Phi \), i.e.,

\[
\text{supp}(\mu) \subset B(\Phi),
\]

where \( B(\Phi) = \{ x \in \mathbb{R}^m : x \in \omega(x) \} \).

According to Proposition 3.3, some further investigations can be carried out by studying the properties of the Birkhoff’s center \( B(\Phi) \). Following [3, Proposition 3.2], if the flow \( \Phi \) is dissipative, then each component of \( \text{supp}(\mu) \) is internally chain transitive [5, 7].

The main result in this paper is summarized as follows.

**Theorem 3.4.** Let Hypotheses 1 and 2 hold and \( \mu^* \) be an invariant probability measure of the Markov chain \( \{ X_k \}_{k \in \mathbb{N}} \). Suppose that the family \( \{ \mu^* \}_{\epsilon > 0} \) is tight. Then any limit point \( \mu = \lim_{\epsilon \to 0} \mu^* \) for the topology of weak* convergence is an invariant measure of \( \Phi \), whose support is contained in its Birkhoff’s center.

**Remark 3.1.** Theorem 3.4 still holds if the domain \( \mathbb{R}^m \) is replaced by the unit \( m \)-cube \( [0,1]^m \) without tightness.

We now consider the ordinary differential equation that is the average equation for recursion (6)

\[
\frac{dx}{dt} = \beta(1-x)x + \alpha(1-x) - (b + \gamma)x.
\]  

\begin{align}
\frac{dx}{dt} = \beta(1-x)x + \alpha(1-x) - (b + \gamma)x.
\end{align}

\[\text{(10)}\]

**Figure 2.** Graphs of \( p = \frac{(-b-\gamma-\alpha+\beta)+\sqrt{(b+\gamma+\alpha-\beta)^2+4\alpha\beta}}{2\beta} \) for \( R_0 = 1/5 \) and \( R_0 = 2 \), where (a) \( \beta = 0.125 \), \( b = 0.125 \), \( \gamma = 0.5 \); (b) \( \beta = 0.5 \), \( b = 0.125 \), \( \gamma = 0.125 \).

**Theorem 3.5.** Suppose \( \mu \in \mathcal{P}([0,1]) \) is a weak* limit point of the invariant measures \( \{ \mu^N \}_{N \in \mathbb{N}^*} \).

(i) If \( \alpha = 0 \), then \( \mu \) is a Dirac measure supported at 0, i.e., \( \mu = \delta_0 \).
(ii) If \( \alpha > 0 \), then (10) has a unique equilibrium \( p \), which is positive and globally asymptotically stable, and \( \mu \) is a Dirac measure supported at \( p \), i.e., \( \mu = \delta_p \).

Proof. (i) follows from the Markov chain theory. (ii) follows from Theorem 3.4. \( \square \)

The following assertion gives precise description about the connection between the deterministic and stochastic models.

**Theorem 3.6.** Let \( U \subset [0, 1] \) be any neighborhood of \( p \). Then a.s.

\[
\lim_{N \to \infty} \lim_{n \to \infty} \tau_n^N(U) = 1,
\]

where \( \tau_n^N(\cdot) = \frac{1}{n+1} \sum_{k=0}^{n} \delta_{X_k^N}(\cdot) \).

Proof. The idea of the proof is borrowed from [3]. In order to prove the theorem, it suffices to show that there exists a full measure set \( \Omega_0 \), for every \( \omega \in \Omega_0 \),

\[
\lim_{N \to \infty} \limsup_{n \to \infty} \tau_n^N(U^c) = 0,
\]

where \( U^c \) denotes the complementary set of \( U \) in \([0, 1] \).

Given \( N \in \mathbb{N}^* \), for \( \omega \in \Omega \), there exists a subsequence \( n_i \to +\infty \) (depending on \( \omega, N \)) such that

\[
\limsup_{n \to \infty} \tau_{n_i}^N(U^c) = \lim_{n \to \infty} \tau_{n_i}^N(U^c).
\]

By the tightness of \( \{\tau_{n_i}^N\}_{i \in \mathbb{N}} \), we can suppose that \( \{\tau_{n_i}^N\}_{i \in \mathbb{N}} \) weakly converges toward some probability measure \( \mu^N \). On the other hand, the process \( \{X_k^N\}_{k \in \mathbb{N}} \) is a finite Markov chain and the assumption \( \alpha > 0 \) implies that this chain is irreducible and all states are positively recurrent. Therefore the sample-path ergodic theorem implies that \( \mu^N \) is almost surely (\( \Omega_N \)) an invariant probability measure of \( \{X_k^N\}_{k \in \mathbb{N}} \) (see, e.g., [10, Proposition 3.3.2, p. 45]). Let \( g : [0, 1] \to [0, 1] \) be a smooth function

\[
g(x) = \begin{cases} 
1, & x \in [0, 1] \setminus U; \\
0, & x \in W,
\end{cases}
\]

where \( W \subset U \) is a neighborhood of \( p \). Then

\[
\lim_{n \to \infty} \tau_n^N(U^c) \leq \lim_{n \to \infty} \int_{[0,1]} g(x) \tau_n^N(dx) = \int_{[0,1]} g(x) \mu^N(dx).
\]

This implies that \( \limsup_{n \to \infty} \tau_n^N(U^c) \leq \int_{[0,1]} g(x) \mu^N(dx) \) a.s. \( \Omega_0 = \bigcap_{N=1}^{\infty} \Omega_N \). Then \( \Omega_0 \) is a full measure set and for every \( \omega \in \Omega_0 \), we get

\[
\int_{[0,1]} g(x) \mu^N(dx), \forall N \geq 1.
\]

We claim that

\[
\lim_{N \to \infty} \int_{[0,1]} g(x) \mu^N(dx) = 0.
\]

Assume that the claim is not true. Then there exists a sequence \( N_i \to \infty \) such that

\[
\lim_{N_i \to \infty} \int_{[0,1]} g(x) \mu^{N_i}(dx) > 0.
\]
By the tightness of $\{\mu^N_i\}_{N_i \in \mathbb{N}}$, we can suppose $\mu^N_i \overset{w}{\to} \mu = \delta_p$, and Theorem 3.5 (ii) implies that $\mu([0, 1] \setminus W) = 0$. Thus we have
\[
\lim_{N_i \to \infty} \int_{[0,1]} g(x)\mu^N_i(dx) = \int_{[0,1]} g(x)\mu(dx) = \int_W g(x)\mu(dx) = 0,
\]
contradicting to (11). Therefore we obtain that $\lim\limsup_{n \to \infty} \tau^N_n(U^c) = 0$ for every $\omega \in \Omega_0$.

Theorem 3.6 shows that as $N \to \infty$, the process $\{X^N_k\}_{k \in \mathbb{N}}$ almost surely spends most of the time in the neighborhood of the unique equilibrium $p$.

**Corollary 3.7.** Let $\eta > 0$. Then with probability 1, for every $\delta > 0$ there exists $N_* \geq 1$ having the following property: for every $N \geq N_*$, there exists $n_* = n_*(N) \geq 1$ such that if $n \geq n_*$, then
\[
\sharp\{k \in J_n : \|X^N_k - p\| > \eta\} < \delta,
\]
where $J_n = \{0, 1, 2, \cdots, n\}$, and $\sharp$ denotes cardinality.

4. **Application to SIR epidemic models with stochastic incidence.** In this section we will build the stochastic SIR model. Then we shall apply our previous theoretic results to analyze the long-run behavior of this stochastic model.

In SIR epidemic model, individuals in the population are classified by three status: susceptible, infected or removed which are denoted by $S$, $I$ and $R$, respectively. The classic SIR model with a standard incidence is given by the following form:
\[
\begin{align*}
\frac{dS}{dt} &= b(I + R) - \beta SI, \\
\frac{dI}{dt} &= \beta SI - (b + \gamma)I, \\
\frac{dR}{dt} &= \gamma I - bR,
\end{align*}
\]
where $b > 0$ is the birth rate, $\gamma > 0$ is the recovery rate, $\beta > 0$ is the contact rate and $N = S(t) + I(t) + R(t)$ is the total population size. As explained in SIS model, we have still assumed that birth rate equals the death rate, so that the total population size $N$ is constant. For details we refer to Hethcote [9] and the references therein.

Let $s(t) = \frac{S(t)}{N}$, $i(t) = \frac{I(t)}{N}$ and $r(t) = \frac{R(t)}{N}$ be the susceptible, infectious and removed fractions, respectively. Dividing the equations (12) by the constant total population size $N$ yields
\[
\begin{align*}
\frac{ds}{dt} &= b(i + r) - \beta si, \\
\frac{di}{dt} &= \beta si - (b + \gamma)i, \\
\frac{dr}{dt} &= \gamma i - br.
\end{align*}
\]
Note that $r(t) = 1 - s(t) - i(t)$, we can consider two independent variables $s(t)$ and $i(t)$. Therefore, we merely consider the following form:
\[
\begin{align*}
\frac{ds}{dt} &= -\beta si + b(1 - s), \\
\frac{di}{dt} &= \beta si - (b + \gamma)i,
\end{align*}
\]
where $(s, i) \in D_1 = \{(s, i) : s \geq 0, i \geq 0, s + i \leq 1\}.
Let $R_0 = \frac{\beta}{\gamma}$ denote the basic reproduction number, then the asymptotic behavior of model (14) is summed as follows (see, e.g., [9]).

\textbf{Theorem 4.1.} Let $(s(t), i(t))$ be a solution of (14).

(i) If $R_0 \leq 1$, then there is a unique equilibrium $p_0 = (1, 0)$ in $D_1$ and it is globally asymptotically stable.

(ii) If $R_0 > 1$, then $p^* = \left(\frac{1}{R_0}, \frac{1}{R_0}(1 - \frac{1}{R_0})\right)$ is a unique positive equilibrium in $D_1 \setminus \{0\}$ and it is globally asymptotically stable for the positive initial conditions.

We now use the language of probability theory to build stochastic SIR models which are similar to the previous stochastic SIS models. Let the total population size $N$ be fixed. At any discrete time $k \in \mathbb{N}$, let $S_k$, $I_k$ and $R_k$ denote random variables for the number of susceptible, infected and removed individuals at time $k$, respectively. Let $R_k = N - S_k - I_k$, there are two independent variables $S_k$ and $R_k$. The process $\{(S_k, I_k)\}_{k \in \mathbb{N}}$ has the Markov property. Let $X_k = \frac{S_k}{N}$ and $Y_k = \frac{I_k}{N}$ denote the susceptible and infectious fractions at time $k$, respectively. The process $\{(X_k, Y_k)\}_{k \in \mathbb{N}}$ also enjoys Markov property from the assumption of $\{(S_k, I_k)\}_{k \in \mathbb{N}}$.

For $x, y \in [0,1]$ and suppose $X_k = x$, $Y_k = y$. We assume that, at any time $k \in \mathbb{N}$, there is at most one individual changing her status (susceptible, infected or removed) in the population during time interval $k \rightarrow k + 1$. This implies that the possible values $X_{k+1} = x - \frac{1}{N}$, $Y_{k+1} = y + \frac{1}{N}$ occurs if and only if the probability that she (called $Q_k$) is chosen from susceptible class, which is $x$, times the conditional probability that $Q_k$ becomes infected at next time $k + 1$, which is $\beta y$. Thus we have $P(X_{k+1} = x - \frac{1}{N}, Y_{k+1} = y - \frac{1}{N}|X_k = x, Y_k = y) = \beta xy$.

The conditional probability of the event that $X_{k+1} = x$, $Y_{k+1} = y - \frac{1}{N}$ occurs if and only if the probability that $Q_k$ is chosen from infective class, which is $y$, times the conditional probability that $Q_k$ is recovery at time $k + 1$, which is $\gamma y$. That is, $P(X_{k+1} = x, Y_{k+1} = y - \frac{1}{N}|X_k = x, Y_k = y) = \gamma y$.

The conditional probability of the event that $X_{k+1} = x + \frac{1}{N}$, $Y_{k+1} = y$ occurs if and only if the probability that $Q_k$ is chosen from removed class, which is $1 - x - y$, times the conditional probability that $Q_k$ is death at next time $k + 1$ (since the total population size $N$ is constant), which is $b$. We obtain $P(X_{k+1} = x + \frac{1}{N}, Y_{k+1} = y - \frac{1}{N}|X_k = x, Y_k = y) = b(1 - x - y)$.

Similarly, $X_{k+1} = x + \frac{1}{N}$, $Y_{k+1} = y - \frac{1}{N}$ if and only if $Q_k$ is chosen from infective class, and at time $k + 1$, she is death, i.e., $P(X_{k+1} = x + \frac{1}{N}, Y_{k+1} = y - \frac{1}{N}|X_k = x, Y_k = y) = by$.

Finally, the conditional probability of no change (the only other possibility), $X_{k+1} = x$, $Y_{k+1} = y$ is $1 - \beta xy - \gamma y - by - b(1 - x - y)$. That is, $P(X_{k+1} = x, Y_{k+1} = 0|X_k = x, Y_k = y) = 1 - \beta xy - \gamma y - by - b(1 - x - y) = 1 - \beta xy - \gamma y - b(1 - x)$.

For simplicity, we make additional assumption that $b$, $\beta$ and $\gamma$ are all positive constants such that $1 - \beta xy - \gamma y - b(1 - x) \geq 0$ for all $(x, y) \in D_1$.

For each $N \in \mathbb{N}^*$, the sequence of two-dimensional random variable $(X^N, Y^N) = \{(X_k^N, Y_k^N) : k \in \mathbb{N}\}$ is a discrete-time Markov chain taking value in a finite lattice $\{0, \frac{1}{N}, \cdots, \frac{N-1}{N}, 1\} \times \{0, \frac{1}{N}, \cdots, \frac{N-1}{N}, 1\} \cap D_1 \subset [0,1]^2$. It is easy to show that the state $\{(1,0)\}$ is absorbing and that all other states are transient. Thus it has a
unique invariant probability measure $\delta_{\{(1,0)\}}$ (see the discussion of the SIS stochastic model). And it implies that the disease eventually dies out.

Compare this asymptotic result to Theorem 4.1, there is a big difference between the deterministic and stochastic models.

We adjust the model by introducing a stochastic incidence. The only useful difference is that

$$P(X_{k+1} - X_k = -\frac{1}{N}, Y_{k+1} - Y_k = \frac{1}{N}|X_k = x, Y_k = y) = \beta xy + \alpha x,$$

where $\alpha$ is for simplicity a small positive probability, because the disease can occur sporadically in the susceptible class by the environmental variability. Then it is easy to see that Markov chain $\{(X_k, Y_k)\}_{k \in \mathbb{N}}$ is irreducible and every state is positive recurrent. Therefore, $\{(X_k, Y_k)\}_{k \in \mathbb{N}}$ has a unique invariant probability measure $\mu^N$, where each component of $\mu^N$ is positive. In order to show the average equation, we compute the conditional probability that $P\left(X_{k+1} - X_k = -\frac{1}{N}|X_k = x, Y_k = y\right) = b(1 - x), P\left(X_{k+1} - X_k = -\frac{1}{N}|X_k = x, Y_k = y\right) = \beta xy + \alpha x, P(Y_{k+1} - Y_k = \frac{1}{N}|X_k = x, Y_k = y) = \gamma y + by$. Then the expected values of $X_{k+1} - X_k$ and $Y_{k+1} - Y_k$ are easily calculated,

$$\mathbb{E}[X_{k+1} - X_k|X_k = x, Y_k = y] = \frac{1}{N}\mathbb{P}(X_{k+1} - X_k = \frac{1}{N}|X_k = x, Y_k = y) - \frac{1}{N}\mathbb{P}(X_{k+1} - X_k = -\frac{1}{N}|X_k = x, Y_k = y) = \frac{1}{N}b(1 - x) - \beta xy - \alpha x$$

and

$$\mathbb{E}[Y_{k+1} - Y_k|X_k = x, Y_k = y] = \frac{1}{N}\mathbb{P}(Y_{k+1} - Y_k = \frac{1}{N}|X_k = x, Y_k = y) - \frac{1}{N}\mathbb{P}(Y_{k+1} - Y_k = -\frac{1}{N}|X_k = x, Y_k = y) = \frac{1}{N}[\beta xy + \alpha x - (b + \gamma)y] = \frac{1}{N}Q(x,y),$$

where $P(x,y) = b(1 - x) - \beta xy - \alpha x$ and $Q(x,y) = \beta xy - (b + \gamma)y + \alpha x$.

Let $F(x,y) = (P(x,y), Q(x,y))$ and $Z_k = (X_k, Y_k)$, it is easy to verify that

$$Z_{k+1} - Z_k = \frac{1}{N}F(Z_k) + Z_{k+1} - \mathbb{E}[Z_{k+1}|Z_k] = \frac{1}{N}[F(Z_k) + U_{k+1}],$$

where $U_k = N(Z_{k+1} - \mathbb{E}[Z_{k+1}|Z_k])$.

Let $\sigma$-algebra $\mathcal{F}_k$ denote by $\sigma\{Z_0, \cdots, Z_k\}$. It is clear that $\{U_k\}_{k \in \mathbb{N}^*}$ is a martingale difference sequence, i.e., $U_k$ is measurable with respect to $\mathcal{F}_k$ and satisfies

$$\mathbb{E}[U_{k+1}|\mathcal{F}_k] = 0.$$

Together with above assumptions, we also have

$$\|U_k\| \leq \max_{(x,y) \in [0,1]^2} \|F(x,y)\| + 2.$$  \hfill (15)

Let $\epsilon = \frac{1}{N}$, then the process $\{Z_k\}_{k \in \mathbb{N}}$, parameterized by $\epsilon$, also satisfies the recursion form (6). We now consider the average equation:

$$\begin{cases}
\frac{dx}{dt} = -\beta xy + b(1 - x) - \alpha x = P(x,y), \\
\frac{dy}{dt} = \beta xy - (b + \gamma)y + \alpha x = Q(x,y).
\end{cases}$$  \hfill (16)

Here $\beta, b, \gamma$ and $\alpha$ are positive numbers and $(x,y) \in D_1$.

The asymptotic behavior of solution paths is described in the following theorem.
Theorem 4.2. Let $F = (P(x,y), Q(x,y))$ be the vector field defined by (16), then there exists a unique positive equilibrium $p$ and it is globally asymptotically stable.

Proof. The proof of this theorem follows the line of argument in Hethcote [9] using standard phase plane methods. The set $D_1$ is positively invariant since no direction vectors at the boundary of $D_1$ are outward (see, e.g., [13] and the references therein). It is easy to see that the system (16) has a unique positive equilibrium $p = (x^*, y^*)$ in $D_1$. It is not difficult to verify that $\det \begin{pmatrix} \frac{\partial P}{\partial x} & \frac{\partial P}{\partial y} \\ \frac{\partial Q}{\partial x} & \frac{\partial Q}{\partial y} \end{pmatrix}_p > 0$ and $\left( \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right)_p < 0$, i.e., the characteristic roots of the linearization around $p$ have negative real parts, this implies that it is locally asymptotically stable. Choose $B(x,y) = \frac{1}{y}$. Then from routine computation, $\frac{\partial (BP)}{\partial x} + \frac{\partial (BQ)}{\partial y} = \frac{1}{y} (b - \alpha) - \beta - \frac{\alpha y^2}{y^2} < 0$. Thus, by Dulac’s criterion, the system (16) has no periodic orbit in $D_1$. Consequently, the unique equilibrium $p$ is globally asymptotically stable.

Using Theorem 4.2 it is easy to derive from Theorem 3.4 the following assertion.

Theorem 4.3. Suppose $\mu \in \mathcal{P}([0,1]^2)$ is a weak* limit point of the invariant measures $\{\mu^N\}_{N \in \mathbb{N}^*}$.

(i) If $\alpha = 0$, then $\mu$ is a Dirac measure supported at $\{(1,0)\}$, i.e., $\mu = \delta_{\{(1,0)\}}$.

(ii) If $\alpha > 0$, then (16) has a unique equilibrium $p$, which is positive and globally asymptotically stable, and $\mu$ is a Dirac measure supported at $p$, i.e., $\mu = \delta_p$.

As in the stochastic SIS case we can also have the same result as Corollary 3.7.

In SIS and SIR models with stochastic incidence, the results of Theorem 3.5 (ii) and Theorem 4.3 (ii) imply that, the deterministic models have a unique globally stable equilibrium $p$. And for the discrete time stochastic models, it follows from Corollary 3.7 that sample paths tend to cluster near $p$ as $N \to \infty$. This implies that the results of deterministic and stochastic models are reconciled.
If $R_0 \leq 1$, pictures (a) of Figures 2 and 3 show that the fluctuation of equilibrium $p$ is in a sufficiently small neighborhood of the disease-free equilibrium, granted the parameter $\alpha$ is small enough. Biologically, it means that the disease occurs sporadically but not becomes endemic disease. But for $R_0 > 1$, pictures (b) of Figures 2 and 3 show that the unique equilibrium $p$ is bounded away from the disease-free equilibrium, which means that the disease will still become the endemic disease with prevalence $p$ and the influence of stochastic incidence can be neglected. Hence, from the practical viewpoint, the biologically models with stochastic incidence are more realistic.

Acknowledgments. We are grateful to the editors and referees for their careful reading and helpful comments which led to an improvement of our original manuscript. The second author is greatly appreciate Prof. M. Y. Li for sending the preprint of the paper [11] and for his invaluable communication on this topic.

REFERENCES

[1] M. Benaïm, Recursive algorithms, urn processes and chaining number of chain recurrent sets, Ergodic Theory Dynam. Systems, 18 (1998), 53–87.
[2] M. Benaïm, Dynamics of stochastic approximation algorithms, Séminaire de Probabilités, XXXIII, 1–68, Lecture Notes in Math., 1709, Springer, Berlin, 1999.
[3] M. Benaïm and M. W. Hirsch, Stochastic approximation algorithms with constant step size whose average is cooperative, Ann. Appl. Probab., 9 (1999), 216–241.
[4] M. Benaïm and M. W. Hirsch, Differential and stochastic epidemic models, Fields Inst. Commun., 21 (1999), 31–44.
[5] R. Bowen, $\omega$-limit sets of Axiom A diffeomorphisms, J. Differential Equations, 18 (1975), 333–339.
[6] D. L. Burkholder, B. J. Davis and R. F. Gundy, Integral inequalities for convex functions of operators on martingales, Proc. 6th Berkeley Symp. Math. Statist. and Probab., 2 (1972), 223–240.
[7] C. C. Conley, Isolated Invariant Sets and the Morse Index, CBMS Regional Conference Series in Mathematics. American Mathematical Society, Providence, 1978.
[8] J. C. Fort and G. Pagès, Asymptotic behavior of a Markovian stochastic algorithm with constant step, SIAM J. Control Optim., 37 (1999), 1456–1482.
[9] H. W. Hethcote, Qualitative analyses of communicable disease models, Math. Biosci., 28 (1976), 335–356.
[10] O. Hernández-Lerma and J. B. Lasserre, Markov Chains and Invariant Probabilities, Birkhäuser Verlag, Basel, 2003.
[11] M. Y. Li and C. Xu, Limit threshold theorems for stochastic epidemic models, preprint, 2014.
[12] R. Mañé, Ergodic Theory and Differentiable Dynamics, Springer-Verlag, Berlin, 1987.
[13] A. Lajmanovich and J. Yorke, A deterministic model for gonorrhea in a nonhomogeneous population, Math. Biosci., 28 (1976), 221–236.

Received December 2014; revised July 2015.

E-mail address: 1000360929@smai1.shnu.edu.cn
E-mail address: jiangjf@shnu.edu.cn