DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES B
Volume **24**, Number **8**, August **2019**

doi:10.3934/dcdsb.2019125

pp. **4417–4443**

ASYMPTOTIC BEHAVIOR OF GENE EXPRESSION WITH COMPLETE MEMORY AND TWO-TIME SCALES BASED ON THE CHEMICAL LANGEVIN EQUATIONS

Yun Li and Fuke Wu*

School of Mathematics and Statistics
Huazhong University of Science and Technology
Wuhan, Hubei, 430074, China

George Yin

Department of Mathematics
Wayne State University
Detroit, MI 48202 USA

(Communicated by Xiaoying Han)

Abstract. Gene regulatory networks, which are complex high-dimensional stochastic dynamical systems, are often subject to evident intrinsic fluctuations. It is deemed reasonable to model the systems by the chemical Langevin equations. Since the mRNA dynamics are faster than the protein dynamics, we have a two-time scales system. In general, the process of protein degradation involves time delays. In this paper, we take the system memory into consideration in which we consider a model with a complete memory represented by an integral delay from 0 to t. Based on the averaging principle and perturbed test function method, this work examines the weak convergence of the slow-varying process. By treating the fast-varying process as a random noise, under appropriate conditions, it is shown that the slow-varying process converges weakly to the solution of a stochastic differential delay equation whose coefficients are the average of those of the original slow-varying process with respect to the invariant measure of the fast-varying process.

1. **Introduction.** Gene expression is a complex process involving many biochemical reactions with proteins as the final products. Most reactions are not instantaneous, there exists natural time delays in the evolution of cell states [3, 18]. For example, the process of degradation of both mRNA and protein [5] often consists of several steps and can naturally be modeled by using time delays. Delayed degradation of JAK2 protein in signaling pathways was considered in [5] and delayed protein

2010 *Mathematics Subject Classification.* Primary: 60H20, 65C30, 60H30, 34K50; Secondary: 62L20, 92C45, 93C70, 34E10.

Key words and phrases. Weak convergence, two-time scales, time delay, gene expression, chemical Langevin equation.

Yun Li is supported by in part by the National Natural Science Foundations of China (Grant Nos. 61873320 and 11761130072). Fuke Wu is supported by the National Natural Science Foundations of China (Grant Nos. 61873320 and 11761130072) and the Royal Society-Newton Advanced Fellowship. George Yin is supported by the National Science Foundation under grant DMS-1710827.

* Corresponding author: Fuke Wu.

degradation was studied in [4]. [3] also considered the stability of the Hes1 gene expression consisting of a cascade of reactions with discrete as well as distributed delays.

In many biochemical reactions occurring in living cells, the number of various molecules might be low with significant stochastic fluctuations. For the biochemical reaction systems subject to the intrinsic noise that originates from the inner stochasticity of the systems and is generated by intermolecular collisions affecting the timing of individual reaction [2], the reaction processes can be modeled as a discrete Markov process with jumps from one discrete state to another representing chemical reactions. The stochastic simulation algorithm (SSA), originally proposed by Gillespie [9], is an exact simulation for the systems subject to intrinsic noise, but it is often computationally expensive. This is especially true in case of highly reactive biochemical systems comprising a large number of molecular species. To reduce the computational load, one of the ways is to use the chemical Langevin equation (CLE) [10, 17].

In a classical model of gene expression [21], molecules of mRNA are produced from DNA in the process of transcription and then give rise to the production of protein molecules in the process of translation. Both types of molecules may degrade. Since the mRNA dynamics are faster than the protein dynamics, we have a two-time scales system; see [22, 23]. Denote the intensities of the biochemical reactions by k_r/ε, k_p, γ_r/ε and γ_p, respectively,

$$\begin{aligned} &\text{DNA} \xrightarrow{k_r/\varepsilon} \text{mRNA},\ \text{mRNA} \xrightarrow{\gamma_r/\varepsilon} \emptyset,\\ &\text{mRNA} \xrightarrow{k_p} \text{Protein},\ \text{Protein} \xrightarrow{\gamma_p} \emptyset, \end{aligned} \tag{1}$$

where the small parameter ε shows that the mRNA dynamics are faster than protein. Denote the concentrations of mRNA and protein by r^ε and q^ε, respectively. Then the standard equations of chemical kinetics read

$$\begin{cases} \dot{r}^\varepsilon(t) = \dfrac{1}{\varepsilon}(k_r - \gamma_r r^\varepsilon(t)),\\ \dot{q}^\varepsilon(t) = k_p r^\varepsilon(t) - \gamma_p q^\varepsilon(t). \end{cases} \tag{2}$$

Following the work [4], [18] took into account the process of protein degradation with time delays. In [4] and [18], to simplify the mathematical models, only fixed time delay is considered, whereas distributed delays treated as memory were considered in [3]. When the complete memory is considered, integral delay from 0 to t is more suitable. Then system (1) can be rewritten as

$$\begin{aligned} &\text{DNA} \xrightarrow{k_r/\varepsilon} \text{mRNA},\ \text{mRNA} \xrightarrow{\gamma_r/\varepsilon} \emptyset,\\ &\text{mRNA} \xrightarrow{k_p} \text{Protein},\ \text{Protein} \xrightarrow{\gamma_p} \emptyset,\ \text{Protein} \xrightarrow{\gamma_d(s)} \emptyset, \end{aligned} \tag{3}$$

and the second equation in (2) can be rewritten as

$$\dot{q}^\varepsilon(t) = k_p r^\varepsilon(t) - \gamma_p q^\varepsilon(t) - \int_0^t \gamma_d(s) q^\varepsilon(s) ds,$$

where $\gamma_d(s)$ can be seen as the degradation intensity of the proteins produced at time $s \in [0, t]$. When the intrinsic fluctuations are considered, the corresponding

chemical Langevin equation (see [23]) is given by

$$\begin{cases} dr^{\varepsilon}(t) = \dfrac{1}{\varepsilon}(k_r - \gamma_r r^{\varepsilon}(t))dt + \dfrac{1}{\sqrt{\varepsilon}}\sqrt{k_r + \gamma_r r^{\varepsilon}(t)}dw_1(t), \\ dq^{\varepsilon}(t) = \Big(k_p r^{\varepsilon}(t) - \gamma_p q^{\varepsilon}(t) - \displaystyle\int_0^t \gamma_d(s)q^{\varepsilon}(s)ds\Big)dt \\ \quad + \sqrt{k_p r^{\varepsilon}(t) + \gamma_p q^{\varepsilon}(t) + \displaystyle\int_0^t \gamma_d(s)q^{\varepsilon}(s)ds}dw_2(t), \end{cases} \tag{4}$$

where $w_1(t)$ and $w_2(t)$ are two independent Brownian motions.

Since the information of DNA is from protein, the synthesis of mRNA can be generalized as a function of the concentrations of protein. In general, gene regulatory networks are complex high-dimensional stochastic dynamical systems. System (4) can be generalized as

$$\begin{cases} dr_i^{\varepsilon}(t) = \dfrac{1}{\varepsilon}\big(f_i(q^{\varepsilon}(t)) - \gamma_i r_i^{\varepsilon}(t)\big)dt + \dfrac{1}{\sqrt{\varepsilon}}\sqrt{f_i(q^{\varepsilon}(t)) + \gamma_i r_i^{\varepsilon}(t)}dw_{i1}(t), \\ dq_i^{\varepsilon}(t) = \Big(g_i(r_i^{\varepsilon}(t)) - \delta_i q_i^{\varepsilon}(t) - q_{i,t}^{\varepsilon}\Big)dt + \sqrt{g_i(r_i^{\varepsilon}(t)) + \delta_i q_i^{\varepsilon}(t) + q_{i,t}^{\varepsilon}}dw_{i2}(t), \end{cases} \tag{5}$$

for $i = 1, 2, ..., n$ with deterministic initial data $r(0) \in \mathbb{R}^n$ and $q(0) \in \mathbb{R}^n$, where $r^{\varepsilon}(t) = (r_1^{\varepsilon}(t), r_2^{\varepsilon}(t), \ldots, r_n^{\varepsilon}(t))'$ and $q^{\varepsilon}(t) = (q_1^{\varepsilon}(t), q_2^{\varepsilon}(t), \ldots, q_n^{\varepsilon}(t))'$ represent the concentrations of mRNA and protein at time t, respectively, $f_i : \mathbb{R}^n \to \mathbb{R}_+$ and $g_i : \mathbb{R} \to \mathbb{R}_+$ represent synthesis of mRNA and protein in the gene i, respectively, $q_{i,t}^{\varepsilon} = \int_0^t \pi_i(s)q_i^{\varepsilon}(s)ds$ represents that the complete memory is considered, $\pi_i(s)$ can be seen as the degradation intensity of the proteins produced at time $s \in [0, t]$, $\gamma_i, \delta_i > 0$ are degradation rates of mRNA and protein, respectively, $w_{i1}(t)$ and $w_{i2}(t)$ $(i = 1, 2, \ldots, n)$ are independent standard Brownian motions.

[23] considered the asymptotic behavior of gene regulatory networks with two-time scales by virtue of the averaging principle and the Fokker-Planck equation. This paper considers not only two-time scales, but also the complete memory represented by an integral from 0 to t under the special function g_i. Since the second equation in (5) involves the delay, its solution is non-Markov. Thus the techniques in the literature which only treating Markov processes are not applicable. It is certainly important to establish a complexity reduction method for the delay system with two-time scales since no existing results are available to date. By treating the fast-varying process as a random noise, this paper will overcome the difficulties so as to achieve the complexity reduction.

The rest of the paper is arranged as follows. Section 2 provides necessary notation, assumptions and some preliminaries. Section 3 examines the transition probability density and invariant measure of the fast-varying process $r_i^{\varepsilon}(t)$ for $i = 1, 2, ..., n$. By using the perturbed test function, martingale method and weak convergence techniques, Section 4 shows that the slow-varying process $q^{\varepsilon}(\cdot)$ converges weakly to the solution of a stochastic differential delay equation whose coefficients are the average of those of the original slow-varying process with respect to the invariant measure of the fast-varying process as $\varepsilon \to 0$. Based on the established results, this section also examines the stochastic differential delay equation (4) and gives its asymptotic properties. The final section gives some concluding remarks.

2. **Notation, assumptions and preliminaries.** Throughout this paper, unless otherwise specified, we use the following notation. Let $\mathbb{R}^n$ denote the n-dimensional Euclidean space with the Euclidean norm $|\cdot|$, and $\mathcal{B}(\mathbb{R}^n)$ be the Borel σ-algebra of $\mathbb{R}^n$. For each $N>0$, let $S_N=\{x:|x|\le N\}$ be a ball with radius N centered at the origin. For a vector or matrix A, denote its transpose by A'; for a matrix A, denote its trace norm by $|A|=\sqrt{Tr(A'A)}$. Denote by $D([0,T];\mathbb{R}^n)$ the family of functions on $[0,T]$ with values in $\mathbb{R}^n$ that are right continuous with left limits endowed with the Skorohod topology. Denote by $C^m(\mathbb{R}^n;\mathbb{R})$ the family of functions on $\mathbb{R}^n$ with values in $\mathbb{R}$ that have continuous partial derivatives up to the mth-order, $C_0^m(\mathbb{R}^n;\mathbb{R})$ the family of $C^m(\mathbb{R}^n;\mathbb{R})$ functions with compact support, and $B_b(\mathbb{R}^n;\mathbb{R})$ the family of bounded and measurable functions on $\mathbb{R}^n$ with values in $\mathbb{R}$. Denote by $C^{k,m}([0,T]\times\mathbb{R}^n;\mathbb{R})$ the family of functions $V(t,x)$ on $[0,T]\times\mathbb{R}^n$ that are kth-order continuously differentiable with respect to t and mth-order continuously differentiable with respect to x, and $C_0^{k,m}([0,T]\times\mathbb{R}^n;\mathbb{R})$ the family of $C^{k,m}([0,T]\times\mathbb{R}^n;\mathbb{R})$ functions with compact support. $\mathbf{L}^2([0,T]\times\Omega;\mathbb{R}^{n\times l})$ denotes the family of all $\mathbb{R}^{n\times l}$-valued measurable $\mathcal{F}_t$-adapted processes $\Phi(t)$ such that $\mathbb{E}\int_0^T|\Phi(t)|^2dt<\infty$. Throughout this paper, K denotes a generic positive constant, whose value may change for different usage. Thus, $K+K=K$ and $KK=K$ are understood in an appropriate sense.

In this paper, if $q(t)$ is a stochastic process, denote by $\mathcal{F}_t^q$ the σ-algebra generated by $\{q(s):s\le t\}$ and $\mathbb{E}_t^q$ the corresponding conditional expectation. For the stochastic process $q^\varepsilon(t)$ and $r^\varepsilon(t)$ depending on ε, we denote by $\mathcal{F}_t^\varepsilon$ the σ-algebra generated by $\{q^\varepsilon(s),r^\varepsilon(s):s\le t\}$ and $\mathbb{E}_t^\varepsilon$ the corresponding conditional expectation.

Let $\mathscr{M}$ denote the set of real-valued progressively measurable processes that are non-zero only on a bounded t-interval and

$$\bar{\mathscr{M}}^\varepsilon=\Big\{f\in\mathscr{M}:\sup_t\mathbb{E}|f(t)|<\infty \text{ and } f(t) \text{ is } \mathcal{F}_t^\varepsilon-\text{measurable}\Big\}. \tag{6}$$

Following [12, 14], let us recall the definitions of the p-lim and the infinitesimal operator $\hat{\mathcal{L}}^\varepsilon$ as follows.

Definition 2.1. Let $f,\ f^\delta\in\bar{\mathscr{M}}^\varepsilon$ for each $\delta>0$. We say $f=\text{p-}\lim\limits_{\delta\to0_+}f^\delta$ if and only if

$$\begin{cases}\sup\limits_{t,\delta}\mathbb{E}|f^\delta(t)|<\infty,\\ \lim\limits_{\delta\to0_+}\mathbb{E}|f^\delta(t)-f(t)|=0,\ \forall\ 0\le t\le T.\end{cases}$$

Definition 2.2. Let $f,\ g\in\bar{\mathscr{M}}^\varepsilon$. If

$$\text{p-}\lim_{\delta\to0_+}\Big(\frac{\mathbb{E}_t^\varepsilon f(t+\delta)-f(t)}{\delta}-g(t)\Big)=0,$$

we say that $f(\cdot)\in\mathscr{D}(\hat{\mathcal{L}}^\varepsilon)$ and $\hat{\mathcal{L}}^\varepsilon f=g$, where $\mathscr{D}(\hat{\mathcal{L}}^\varepsilon)$ denotes the domain of the operator $\hat{\mathcal{L}}^\varepsilon$.

Thus $\hat{\mathcal{L}}^\varepsilon$ is a type of infinitesimal operator. The following lemma was proved in Kurtz [12].

Lemma 2.3. *If $f\in\mathscr{D}(\hat{\mathcal{L}}^\varepsilon)$, then*

$$M_f(t)=f(t)-\int_0^t\hat{\mathcal{L}}^\varepsilon f(u)du$$

is a martingale, and

$$\mathbb{E}_t^\varepsilon f(t+s) - f(t) = \mathbb{E}_t^\varepsilon \int_t^{t+s} \hat{\mathcal{L}}^\varepsilon f(u)du, \ \ w.p.1.$$

3. **Invariant measure of the fast-varying process.** To obtain the weak convergence of the slow-varying process $q^\varepsilon(t)$ as $\varepsilon \to 0$, the properties of the fast-varying process $r^\varepsilon(t)$ are crucial. In (5), $r^\varepsilon(t)$ is rapidly varying in contrast to $q^\varepsilon(t)$. To proceed, we first consider asymptotic properties of $r^\varepsilon(t)$. Let us define the process $R_i^\varepsilon(t) = r_i^\varepsilon(\varepsilon t)$ for any $i = 1, 2, ..., n$. Then $R_i^\varepsilon(t)$ satisfies

$$dR_i^\varepsilon(t) = \big(f_i(q^\varepsilon(\varepsilon t)) - \gamma_i R_i^\varepsilon(t)\big)dt + \sqrt{f_i(q^\varepsilon(\varepsilon t)) + \gamma_i R_i^\varepsilon(t)}d\tilde{w}_{i1}(t),$$

where $\tilde{w}_{i1}(t) = w_{i1}(\varepsilon t)/\sqrt{\varepsilon}$ is a standard Brownian motion. When the fast-varying process is analyzed and ε is small, the slow-varying process q^ε varies slowly. Let us define a fixed-Q process $R_i(\cdot) = R_i(\cdot|Q)$ by

$$dR_i(t) = \big(f_i(Q) - \gamma_i R_i(t)\big)dt + \sqrt{f_i(Q) + \gamma_i R_i(t)}d\tilde{w}_{i1}(t), \tag{7}$$

where Q is treated as a parameter. Hence, although $R(t) = (R_1(t), R_2(t), \ldots, R_n(t))'$ is a n-dimensional process, it can be seen as n scalar processes and satisfies the following theorem.

Theorem 3.1. *There exists a unique global solution $R_i(t)$ for (7) for all $t \geq 0$ and this solution holds the following properties:*

(i) $f_i(Q) + \gamma_i R_i(t) \geq 0$;

(ii) $R_i(t)$ *is a homogeneous Markov process and there exists transition probability density*

$$p_i(x, s; x_0, t) = \begin{cases} \gamma_i \hat{p}_i(\beta_i + \gamma_i x, s; \beta_i + \gamma_i x_0, t), & x \geq -\varphi_i, \\ 0, & x < -\varphi_i, \end{cases} \tag{8}$$

where $\hat{p}_i(y, s; y_0, t) = ce^{-u-v}(v/u)^{\frac{q}{2}} I_q(2(uv)^{\frac{1}{2}})$, $y \geq 0$, $c = 2/[\gamma_i(1-e^{-\gamma_i(s-t)})]$, $u = cy_0 e^{-\gamma_i(s-t)}$, $v = cy$, $q = \frac{4\beta_i}{\gamma_i} - 1$ *and* I_q *is the modified Bessel function of the first kind of order* q *given by*

$$I_q(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+q+1)} \left(\frac{x}{2}\right)^{2k+q},$$

in which $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$ *is a gamma function. Moreover, there exists an invariant measure with the density*

$$\mu_i(x) = \begin{cases} \dfrac{2^{4\varphi_i}}{\Gamma(4\varphi_i)} (x + \varphi_i)^{4\varphi_i - 1} e^{-2(x+\varphi_i)}, & x \geq -\varphi_i, \\ 0, & x < -\varphi_i, \end{cases} \tag{9}$$

where $\varphi_i = f_i(Q)/\gamma_i$;

(iii) *the* m*th-order moment* $\mathbb{E}[R_i(t)]^m$ *is uniformly bounded for any integer* $m > 0$.

Proof. Define $\beta_i = f_i(Q)$ and $\hat{R}_i(t) = \beta_i + \gamma_i R_i(t)$. Then, according to (7), $\hat{R}_i(t)$ satisfies the following stochastic differential equation

$$d\hat{R}_i(t) = \gamma_i(2\beta_i - \hat{R}_i(t))dt + \gamma_i\sqrt{\hat{R}_i(t)}d\tilde{w}_{i1}(t), \tag{10}$$

which is a mean-reverting square root process. This implies that (10) has a unique nonnegative solution $\hat{R}_i(t)$ on $t \geq 0$, see [6,16], which implies that (7) has a unique

solution $R_i(t)$ on t$\geq$ 0 with $f_i(Q)+\gamma_i R_i(t)\geq 0$, that is, $R_i(t)\geq -\varphi_i$. Hence, nonnegativity of the solution of (7) cannot be guaranteed. Moreover, the solution $\hat{R}_i(t)$ of (10) is a homogeneous Markov process and its transition probability density is given by $\hat{p}_i(y,s;y_0,t)$; see [6,23]. As $s\to\infty$, it is easily observed that $\hat{R}_i(\cdot)$ tends to a gamma distribution with the density

$$\hat{\mu}_i(y)=\begin{cases}\dfrac{\left(\frac{2}{\gamma_i}\right)^{4\varphi_i}}{\Gamma(4\varphi_i)}y^{4\varphi_i-1}e^{-\frac{2y}{\gamma_i}}, & y\geq 0,\\ 0, & y<0,\end{cases}$$

according to $\hat{p}_i(y,s;y_0,t)$.

Note that $\hat{R}_i(\cdot)=\beta_i+\gamma_i R_i(\cdot)$. For $x\geq -\varphi_i$, the transition probability density of the solution $R_i(\cdot)$ of (7) is given by

$$\begin{aligned}p_i(x,s;x_0,t) &= \gamma_i\hat{p}_i(\beta_i+\gamma_i x,s;R_i(t)=x_0,t)\\ &= \gamma_i\hat{p}_i(\beta_i+\gamma_i x,s;\beta_i+\gamma_i x_0,t).\end{aligned}\tag{11}$$

For $x<-\varphi_i$, the transition probability density $p_i(x,s;x_0,t)=0$. The stationary density of $R_i(\cdot)$ can be expressed as follows

$$\mu_i(x)=\gamma_i\hat{\mu}_i(\beta_i+\gamma_i x)=\begin{cases}\dfrac{2^{4\varphi_i}}{\Gamma(4\varphi_i)}(x+\varphi_i)^{4\varphi_i-1}e^{-2(x+\varphi_i)}, & x\geq-\varphi_i,\\ 0, & x<-\varphi_i.\end{cases}\tag{12}$$

Moreover, one can compute the moment generating function with respect to the invariant measure

$$\begin{aligned}\mathbb{E}_{\mu_i}(e^{\alpha R_i}) &= \int_{-\varphi_i}^{+\infty}e^{\alpha x}\frac{2^{4\varphi_i}}{\Gamma(4\varphi_i)}(x+\varphi_i)^{4\varphi_i-1}e^{-2(x+\varphi_i)}dx\\ &= \frac{2^{4\varphi_i}}{\Gamma(4\varphi_i)}e^{-\alpha\varphi_i}\int_0^{+\infty}y^{4\varphi_i-1}e^{-(2-\alpha)y}dy\\ &= \frac{2^{4\varphi_i}}{\Gamma(4\varphi_i)}e^{-\alpha\varphi_i}\frac{1}{(2-\alpha)^{4\varphi_i}}\int_0^{+\infty}y^{4\varphi_i-1}e^{-y}dy\\ &= \left(1-\frac{\alpha}{2}\right)^{-4\varphi_i}e^{-\alpha\varphi_i},\end{aligned}\tag{13}$$

where $0<\alpha<2$. This implies that the mth-order moment with respect to the invariant measure for any integer $m>0$ is given by

$$\begin{aligned}\mathbb{E}_{\mu_i}(R_i)^m &= (-1)^m\varphi_i^m+2m(-1)^{m-1}\varphi_i^m+4m(m-1)(-1)^{m-2}\varphi_i^m\\ &\quad +m(m-1)(-1)^{m-2}\varphi_i^{m-1}+\cdots+\frac{\varphi_i(4\varphi_i+1)\cdots(4\varphi_i+m-1)}{2^{m-2}}.\end{aligned}\tag{14}$$

Since initial value $R_i(0)=r_i(0)$ is a constant, we can obtain the probability density function of $R_i(t)$

$$F_{R_i(t)}(x)=\begin{cases}\gamma_i\bar{c}e^{-\bar{u}-\bar{v}}\left(\dfrac{\bar{v}}{\bar{u}}\right)^{\frac{q}{2}}I_q(2(\bar{u}\bar{v})^{\frac{1}{2}}), & x\geq-\varphi_i,\\ 0, & x<-\varphi_i,\end{cases}\tag{15}$$

where $\bar{c}=2/[\gamma_i(1-e^{-\gamma_i t})]$, $\bar{u}=\bar{c}(\beta_i+\gamma_i r_i(0))e^{-\gamma_i t}$, $\bar{v}=\bar{c}(\beta_i+\gamma_i x)$. Consequently, the moment generating function of $R_i(t)$ is given by

$$\begin{aligned}
\mathbb{E}e^{\alpha R_i(t)} &= \int_{-\varphi_i}^{+\infty} e^{\alpha x}\gamma_i\bar{c}e^{-\bar{u}-\bar{v}}\Big(\frac{\bar{v}}{\bar{u}}\Big)^{\frac{q}{2}} I_q(2(\bar{u}\bar{v})^{\frac{1}{2}})dx \\
&= \gamma_i e^{-\bar{u}}\sum_{k=0}^{\infty}\frac{\bar{u}^k\bar{c}^{k+q+1}}{k!\Gamma(k+q+1)}\int_{-\varphi_i}^{+\infty}(\beta_i+\gamma_i x)^{k+q}e^{\alpha x}e^{-\bar{c}(\beta_i+\gamma_i x)}dx \\
&= e^{-\bar{u}}e^{-\alpha\varphi_i}\sum_{k=0}^{\infty}\frac{\bar{u}^k\bar{c}^{k+q+1}}{k!\Gamma(k+q+1)}\int_0^{+\infty} y^{k+q}e^{-(\bar{c}-\frac{\alpha}{\gamma_i})y}dy \\
&= e^{-\bar{u}}e^{-\alpha\varphi_i}\sum_{k=0}^{\infty}\frac{\bar{u}^k\bar{c}^{k+q+1}}{k!\Gamma(k+q+1)}\Big(\bar{c}-\frac{\alpha}{\gamma_i}\Big)^{-(k+q+1)}\int_0^{+\infty} y^{k+q}e^{-y}dy \\
&= e^{-\bar{u}}e^{-\alpha\varphi_i}\sum_{k=0}^{\infty}\frac{\bar{u}^k}{k!}\Big(1-\frac{\alpha}{\bar{c}\gamma_i}\Big)^{-(k+q+1)}, \qquad (16)
\end{aligned}$$

in which $\bar{c}-\alpha/\gamma_i>0$. Note that $\bar{u}\to 0$ as $t\to\infty$. According to (13) and (16), it is easy to verify that

$$\lim_{t\to\infty}\mathbb{E}e^{\alpha R_i(t)} = \mathbb{E}_{\mu_i}(e^{\alpha R_i}).$$

Hence,

$$\lim_{t\to\infty}\mathbb{E}[R_i(t)]^m = \mathbb{E}_{\mu_i}(R_i)^m,\ \forall m>0.$$

Thus, the mth-order moment of $R_i(t)$ is uniformly bounded for any integer $m>0$. This completes the proof. □

4. **Weak convergence and averaged system.** In this section, we show that the sequence $q^\varepsilon(t)$ converges weakly to a stochastic process that is the solution of an appropriate stochastic differential equation. In order to obtain the desired weak convergence, we first need to prove tightness.

To address this issue, we need to verify

$$\lim_{N_0\to\infty}\limsup_{\varepsilon\to 0}\mathbb{P}\Big(\Big\{\sup_{0\le t\le T}|q^\varepsilon(t)|\ge N_0\Big\}\Big)=0,\ \text{for each } T<\infty, \qquad (17)$$

where $\mathbb{P}(A)$ denotes the probability of A. The verification of (17) is usually quite involved, and requires complicated calculations. To circumvent the difficulties, we use the truncation technique as follows. For any $i=1,2,...,n$, and any $(x,y,r)\in\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n$, define

$$b_i(x,y,r)=g_i(r_i)-\delta_i x_i-y_i \ \text{ and } \ \psi_i(x,y,r)=\sqrt{g_i(r_i)+\delta_i x_i+y_i}.$$

The second equation of (5) can therefore be rewritten as

$$dq_i^\varepsilon(t)=b_i(q^\varepsilon(t),q_t^\varepsilon,r^\varepsilon(t))dt+\psi_i(q^\varepsilon(t),q_t^\varepsilon,r^\varepsilon(t))dw_{i2}(t),$$

where $r^\varepsilon(t)=(r_1^\varepsilon(t),r_2^\varepsilon(t),\dots,r_n^\varepsilon(t))'=(R_1(\frac{t}{\varepsilon}),R_2(\frac{t}{\varepsilon}),\dots,R_n(\frac{t}{\varepsilon}))'$ is the solution of the equation (7) under fixed q^ε. For each $N>0$ sufficient large such that $|q(0)|\le N$, define

$$dq_i^{\varepsilon,N}(t)=b_i^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))dt+\psi_i^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))dw_{i2}(t), \qquad (18)$$

where $b_i^N(x,y,r)=b_i(x,y,r)h^N(x)$, $\psi_i^N(x,y,r)=\psi_i(x,y,r)h^N(x)$ and

$$h^N(x)=\begin{cases}1, & \text{if } x\in S_N,\\ 0, & \text{if } x\in\mathbb{R}^n-S_{N+1},\\ \text{smooth}, & \text{otherwise.}\end{cases}$$

Hence, according to the definitions of $b_i^N(x,y,r)$ and $\psi_i^N(x,y,r)$, (18) can be rewritten as

$$\begin{aligned} dq_i^{\varepsilon,N}(t) &= \Big(g_i(r_i^\varepsilon(t)) - \delta_i q_i^{\varepsilon,N}(t) - q_{i,t}^{\varepsilon,N}\Big)h^N(q^{\varepsilon,N}(t))dt \\ &\quad + \sqrt{g_i(r_i^\varepsilon(t)) + \delta_i q_i^{\varepsilon,N}(t) + q_{i,t}^{\varepsilon,N}}\, h^N(q^{\varepsilon,N}(t))dw_{i2}(t). \end{aligned} \tag{19}$$

From the definition, it can be seen that $q^{\varepsilon,N}(t) = q^\varepsilon(t)$ up until the first exit from $S_N = \{x : |x| \le N\}$. Then $q^{\varepsilon,N}(t)$ is said to be the N-truncation of $q^\varepsilon(t)$. Let $\mathcal{F}_t^{\varepsilon,N} = \sigma(q^{\varepsilon,N}(s), r^\varepsilon(s) : s \le t)$. We can also give the corresponding definitions for $\mathscr{M}^{\varepsilon,N}$ and $\hat{\mathcal{L}}^{\varepsilon,N}$. It follows that $\mathcal{F}_t^{\varepsilon,N} \subset \sigma(\mathcal{F}_t^{W_2} \vee \mathcal{F}_t^{r^\varepsilon})$, where $\mathcal{F}_t^{W_2}$ and $\mathcal{F}_t^{r^\varepsilon}$ are generated by the n-dimensional Brownian motion $W_2(t) = (w_{12}(t), w_{22}(t), \ldots, w_{n2}(t))'$ and the stochastic process $r^\varepsilon(t)$, respectively. To proceed, the following assumptions are needed.

(**A1**) The second equation of (5) has a unique strong solution $q_i^\varepsilon(t)$ on $[0,T]$ for each $r_i^\varepsilon(t)$ with each deterministic initial value $q_i(0)$, for any $i = 1, 2, ..., n$.

(**A2**) $f_i(x)$, $\frac{\partial}{\partial x_j} f_i(x)$ and $\frac{\partial^2}{\partial x_j^2} f_i(x)$ are uniformly bounded with respect to $x \in G$ for any $i, j = 1, 2, ..., n$, where $G \subset \mathbb{R}^n$ is a compact set. For $i = 1, 2, ..., n$, $g_i(r_i) = a_i r_i^2 + b_i r_i + c_i$ for any $a_i, b_i, c_i \in \mathbb{R}$ and $\pi_i(t)$ is uniformly bounded with respect to $t \in [0,T]$. Denote $b(x,y,r) = (b_1(x,y,r), b_2(x,y,r), \ldots, b_n(x,y,r))'$, $\psi(x,y,r) = \mathrm{diag}(\psi_1(x,y,r), \psi_2(x,y,r), \ldots, \psi_n(x,y,r))$, $a(x,y,r) = \mathrm{diag}(a_1(x,y,r), a_2(x,y,r), \ldots, a_n(x,y,r))$, where $a_i(x,y,r) = \psi_i^2(x,y,r)$.

Remark 1. Here we only consider $g_i(r_i) = a_i r_i^2 + b_i r_i + c_i$. In the concluding remarks of Section 5, we show that for any sufficiently smooth function g_i, the results still hold. With respect to the invariant measure in the Theorem 3.1, the expectation of $b_i(x,y,r) = g_i(r_i) - \delta_i x_i - y_i$ and $a_i(x,y,r) = g_i(r_i) + \delta_i x_i + y_i$ is given by

$$\begin{aligned} \int_{\mathbb{R}} b_i(x,y,r)\mu_i(r_i)dr_i &= \int_{\mathbb{R}} g_i(r_i)\mu_i(r_i)dr_i - \delta_i x_i - y_i \\ &= \frac{a_i\beta_i^2}{\gamma_i^2} + \frac{a_i\beta_i}{\gamma_i} + \frac{b_i\beta_i}{\gamma_i} + c_i - \delta_i x_i - y_i \\ &=: \bar{b}_i(x,y) \end{aligned}$$

and

$$\begin{aligned} \int_{\mathbb{R}} a_i(x,y,r)\mu_i(r_i)dr_i &= \int_{\mathbb{R}} g_i(r_i)\mu_i(r_i)dr_i + \delta_i x_i + y_i \\ &= \frac{a_i\beta_i^2}{\gamma_i^2} + \frac{a_i\beta_i}{\gamma_i} + \frac{b_i\beta_i}{\gamma_i} + c_i + \delta_i x_i + y_i \\ &=: \bar{a}_i(x,y), \end{aligned}$$

for any $i = 1, 2, ..., n$.

Remark 2. If $g_i(r_i^\varepsilon(t)) + \delta_i x_i + y_i > 0$ for any $t \in [0,T]$, $\omega \in \Omega$, $x_i \in \mathbb{R}$ and $y_i \in \mathbb{R}$, then the second equation of (5) has a unique strong solution. Under Assumption (A1), since the second equation of (5) has a unique strong solution $q_i^\varepsilon(t)$ on $[0,T]$ for each $r_i^\varepsilon(t)$, it can be observed that $\bar{a}_i(x,y) \ge 0$ since $\int_{\mathbb{R}} g_i(r_i)\mu_i(r_i)dr_i$ is the mean of $g_i(\cdot)$ with respect to the invariant measure.

(**A3**) The following stochastic differential equation

$$dq(t) = \bar{b}(q(t), q_t)dt + \bar{\psi}(q(t), q_t)dB(t) \tag{20}$$

has a unique weak solution (i.e., uniqueness in the sense of the distribution) on $[0,T]$ for each deterministic initial value $q(0)$, where $\bar{b}(x,y)=(\bar{b}_1(x,y),\bar{b}_2(x,y),\ldots,\bar{b}_n(x,y))'$, $\bar{\psi}(x,y)=\mathrm{diag}(\bar{\psi}_1(x,y),\bar{\psi}_2(x,y),\ldots,\bar{\psi}_n(x,y))$, where $\bar{\psi}_i^2(x,y)=\bar{a}_i(x,y)$, $B(t)$ is an n-dimensional standard Brownian motion.

Next, we state the main theorem in this paper. Its proof will be divided into several parts.

Theorem 4.1. *If* (A1)-(A3) *hold, then* $\{q^\varepsilon(\cdot)\}$ *is tight in* $D([0,T];\mathbb{R}^n)$*, and the limit of any weakly convergent subsequence satisfies equation* (20) *with the same initial value as* $q^\varepsilon(0)=q(0)$ *which is deterministic and independent of* ε*.*

We say that $q(t)$ of (20), is a solution of the martingale problem with operator $\bar{L}$, in that for any function $f\in C_0^{1,2}([0,T]\times\mathbb{R}^n;\mathbb{R})$,

$$M_f(t)=f(t,q(t))-f(0,q(0))-\int_0^t \bar{L}(q(s),q_s)f(s,q(s))ds \tag{21}$$

is a martingale, where for any $x,y\in\mathbb{R}^n$,

$$\bar{L}(x,y)\cdot=\frac{\partial\cdot}{\partial t}+\sum_{i=1}^n \bar{b}_i(x,y)\frac{\partial\cdot}{\partial x_i}+\frac{1}{2}\sum_{i=1}^n \bar{a}_i(x,y)\frac{\partial^2\cdot}{\partial x_i^2}. \tag{22}$$

As mentioned, it is difficult to verify (17). We thus begin the proof of Theorem 4.1 by working with the N-truncated process. Corresponding to this truncation, we have the operators $\hat{\mathcal{L}}^{\varepsilon,N}$ and $\bar{L}^N$, which are operators $\hat{\mathcal{L}}^\varepsilon$ and $\bar{L}$ with x, y, $\bar{b}$ and $\bar{\psi}$ replaced by x^N, y^N, $\bar{b}^N$ and $\bar{\psi}^N$, respectively. Not only can assumption (A1) guarantee the existence and uniqueness of the strong solution of the truncated stochastic differential equation (19), but also the tightness. We proceed with the following theorem.

Theorem 4.2. *Under assumption* (A1)*, there exists a unique strong solution* $q^{\varepsilon,N}(t)$ *for the truncated stochastic differential equation* (19) *for any initial value* $q^{\varepsilon,N}(0)=q(0)\in S_N$ *that is deterministic and independent of* ε*. Moreover, this solution is continuous,* $\mathcal{F}_t^{\varepsilon,N}$*-adapted and tight in* $D([0,T];\mathbb{R}^n)$*.*

To prove this theorem, we need the following Lemma 4.3 (see [16, Theorem 7.1, p.39] for a proof) and Lemma 4.4 (see [14, Theorem 5, p.32]).

Lemma 4.3. *Let* $p\geq 2$ *and* $\Phi\in\mathbf{L}^2([0,T]\times\Omega;\mathbb{R}^{n\times l})$ *such that*

$$\mathbb{E}\int_0^T|\Phi(t)|^p dt<\infty.$$

Then

$$\mathbb{E}\Big|\int_0^T\Phi(t)dW(t)\Big|^p\leq\Big(\frac{p(p-1)}{2}\Big)^{\frac{p}{2}}T^{\frac{p-2}{2}}\mathbb{E}\int_0^T|\Phi(t)|^p dt, \tag{23}$$

where $W(t)$ *is an* l*-dimensional standard Brownian motion.*

Lemma 4.4. *Let* $\{Q^\varepsilon(\cdot)\}$ *be a sequence of* $\mathcal{F}_t^\varepsilon$*-adapted process with paths in* $D([0,T];\mathbb{R}^n)$*. If this sequence satisfies*

$$\lim_{N_0\to\infty}\limsup_{\varepsilon\to 0}\mathbb{P}\Big(\Big\{\sup_{0\leq t\leq T}|Q^\varepsilon(t)|\geq N_0\Big\}\Big)=0 \tag{24}$$

and there are nondecreasing continuous function $F(\cdot)$ *and* $a>1$*,* $\gamma>0$ *such that*

$$\mathbb{E}|Q^\varepsilon(t)-Q^\varepsilon(t_1)|^\gamma|Q^\varepsilon(t_2)-Q^\varepsilon(t)|^\gamma\leq[F(t_2)-F(t_1)]^a, \tag{25}$$

where $0\leq t_1<t<t_2\leq T$*. Then* $Q^\varepsilon(\cdot)$ *is tight in* $D([0,T];\mathbb{R}^n)$*.*

With these two lemmas in hand, we can give a proof of Theorem 4.2.

Proof of Theorem 4.2. According to assumption (A1), the truncated stochastic differential equation (19) has a unique continuous and $\mathcal{F}_t^{\varepsilon,N}$-adapted strong solution.

To prove the tightness of $\{q^{\varepsilon,N}(\cdot)\}$, we need only to verify that the conditions in Lemma 4.4 are satisfied. In fact, under the truncation technique, (24) holds. Hence, we need only to show that (25) holds for the truncated process $q^{\varepsilon,N}(t)$.

From (18),

$$dq^{\varepsilon,N}(t) = b^N(q^{\varepsilon,N}(t), q_t^{\varepsilon,N}, r^\varepsilon(t))dt + \psi^N(q^{\varepsilon,N}(t), q_t^{\varepsilon,N}, r^\varepsilon(t))dW_2(t), \tag{26}$$

where $b^N(x,y,r) = b(x,y,r)h^N(x)$, $\psi^N(x,y,r) = \psi(x,y,r)h^N(x)$, $W_2(t) = (w_{12}(t), w_{22}(t), \ldots, w_{n2}(t))'$ is an n-dimensional standard Brownian motion. For any $0 \le t_1 < t < t_2 \le T$, (26) and the elementary inequality $(a+b)^2 \le 2(a^2+b^2)$ yield

$$\begin{aligned}\left|q^{\varepsilon,N}(t) - q^{\varepsilon,N}(t_1)\right|^2 &\le 2\Big|\int_{t_1}^t b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2 \\ &\quad +2\Big|\int_{t_1}^t \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2\end{aligned}$$

and

$$\begin{aligned}\left|q^{\varepsilon,N}(t_2) - q^{\varepsilon,N}(t)\right|^2 &\le 2\Big|\int_{t}^{t_2} b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2 \\ &\quad +2\Big|\int_{t}^{t_2} \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2.\end{aligned}$$

Applying properties of conditional expectation gives

$$\begin{aligned}&\mathbb{E}\left|q^{\varepsilon,N}(t) - q^{\varepsilon,N}(t_1)\right|^2\left|q^{\varepsilon,N}(t_2) - q^{\varepsilon,N}(t)\right|^2 \\ \le\ & 4\mathbb{E}\Big|\int_{t_1}^t b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2\Big|\int_{t}^{t_2} b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2 \\ &+4\mathbb{E}\Big|\int_{t_1}^t b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2\Big|\int_{t}^{t_2} \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2 \\ &+4\mathbb{E}\Big|\int_{t_1}^t \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2\Big|\int_{t}^{t_2} b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2 \\ &+4\mathbb{E}\Big|\int_{t_1}^t \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2\Big|\int_{t}^{t_2} \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2 \\ \le\ & 4\mathbb{E}\Big|\int_{t_1}^t b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2\Big|\int_{t}^{t_2} b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2 \\ &+4\mathbb{E}\Big|\int_{t_1}^t b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2\mathbb{E}\Big[\int_{t}^{t_2} |\psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))|^2 ds\big|\mathcal{F}_t\Big] \\ &+4\mathbb{E}\Big|\int_{t_1}^t \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2\mathbb{E}\Big[\Big|\int_{t}^{t_2} b^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))ds\Big|^2\big|\mathcal{F}_t\Big] \\ &+4\mathbb{E}\Big|\int_{t_1}^t \psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))dW_2(s)\Big|^2\mathbb{E}\Big[\int_{t}^{t_2} |\psi^N(q^{\varepsilon,N}(s), q_s^{\varepsilon,N}, r^\varepsilon(s))|^2 ds\big|\mathcal{F}_t\Big].\end{aligned}$$

According to Theorem 3.1, the mth-order moment of $R_i(t)$ is uniformly bounded for any integer $m > 0$. Applying the Young inequality, the Hölder inequality and

Lemma 4.3, gives

$$
\begin{aligned}
&\mathbb{E}\big|q^{\varepsilon,N}(t)-q^{\varepsilon,N}(t_1)\big|^2\big|q^{\varepsilon,N}(t_2)-q^{\varepsilon,N}(t)\big|^2\\
\le\ & 4\mathbb{E}\Big|\int_{t_1}^{t} b^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))ds\Big|^4+4\mathbb{E}\Big|\int_{t}^{t_2} b^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))ds\Big|^4\\
&+4\mathbb{E}\Big|\int_{t}^{t_2}|\psi^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))|^2ds\Big|^2\\
&+4\mathbb{E}\Big|\int_{t_1}^{t}\psi^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))dW_2(s)\Big|^4\\
\le\ & 4(t-t_1)^3\mathbb{E}\int_{t_1}^{t}\big|b^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))\big|^4ds\\
&+4(t_2-t)^3\mathbb{E}\int_{t}^{t_2}\big|b^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))\big|^4ds\\
&+4(t_2-t)\mathbb{E}\int_{t}^{t_2}\big|\psi^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))\big|^4ds\\
&+144(t-t_1)\mathbb{E}\int_{t_1}^{t}\big|\psi^N(q^{\varepsilon,N}(s),q_s^{\varepsilon,N},r^{\varepsilon}(s))\big|^4ds\\
\le\ & 8T^2(t_2-t_1)n\sum_{i=1}^{n}\int_{t_1}^{t_2}\mathbb{E}\Big|g_i\Big(R_i\Big(\frac{s}{\varepsilon}\Big)\Big)-\delta_i q_i^{\varepsilon,N}(s)-q_{i,s}^{\varepsilon,N}\Big|^4ds\\
&+148(t_2-t_1)n\sum_{i=1}^{n}\int_{t_1}^{t_2}\mathbb{E}\Big|g_i\Big(R_i\Big(\frac{s}{\varepsilon}\Big)\Big)+\delta_i q_i^{\varepsilon,N}(s)+q_{i,s}^{\varepsilon,N}\Big|^2ds\\
\le\ & K(t_2-t_1)^2. \qquad (27)
\end{aligned}
$$

Thus, (25) holds for the truncated process $q^{\varepsilon,N}(t)$. Lemma 4.4, implies that $\{q^{\varepsilon,N}(\cdot)\}$ is tight in $D([0,T];\mathbb{R}^n)$. The proof is completed. □

Since $q^{\varepsilon,N}(\cdot)$ is tight, by Prohorov's theorem, it is sequentially compact. Thus, we can extract a weakly convergent subsequence and we still lable it by ε. Moreover, the limit is defined as $q^N(\cdot)$. By the Skorohod representation, without changing notation, we may assume that $q^{\varepsilon,N}(\cdot)$ converges to $q^N(\cdot)$ in the sense of w.p.1. We proceed to characterize the limit process $q^N(\cdot)$ by using the averaged system. In what follows, we characterize the weak limit by applying the following lemma [19, 24].

Lemma 4.5. *Let $Q^{\varepsilon}(\cdot)$ be an $\mathbb{R}^n$-valued process defined on $[0,T]$, with $Q^{\varepsilon}(0)=Q(0)$ being deterministic and independent of ε. Let $\{Q^{\varepsilon}(\cdot)\}$ be tight in $D([0,T];\mathbb{R}^n)$. Suppose* (A3) *holds and $\bar{L}$ is the corresponding operator defined by* (22). *For each $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$(or any dense subset of it) and each $T<\infty$, there exists $f^{\varepsilon}(\cdot)\in\mathscr{D}(\hat{\mathcal{L}}^{\varepsilon})$ such that*

$$\text{p-}\lim_{\varepsilon\to 0_+}[f^{\varepsilon}(\cdot)-f(Q^{\varepsilon}(\cdot))]=0 \qquad (28)$$

and

$$\text{p-}\lim_{\varepsilon\to 0_+}[\hat{\mathcal{L}}^{\varepsilon}f^{\varepsilon}(\cdot)-\bar{L}(Q^{\varepsilon}(\cdot),Q_{\cdot}^{\varepsilon})f(Q^{\varepsilon}(\cdot))]=0. \qquad (29)$$

Then, $Q^{\varepsilon}(\cdot)\Rightarrow q(\cdot)$, where $q(\cdot)$ is the solution of the stochastic differential equation (20).

Remark 3. In the process of the averaging, the fast-varying process $r^\varepsilon(t)$ is treated as noise and is averaged out. We use the perturbed test function method to examine the weak convergence. Introducing the perturbed test functions allows us to eliminate the noise terms $r^\varepsilon(t)$ through averaging, and obtain the desired results in the limit.

With these results in hand, we next give a proof of Theorem 4.1.

Proof of Theorem 4.1. According to the definition of p-lim, to prove (28) for $q^{\varepsilon,N}(t)$ and for any $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$, we need to find $f^{\varepsilon,N}(\cdot)\in\mathscr{D}(\hat{\mathcal{L}}^{\varepsilon,N})$ and verify

$$\begin{cases} \sup\limits_{0\le t\le T,\varepsilon}\mathbb{E}|f^{\varepsilon,N}(t)-f(q^{\varepsilon,N}(t))|<\infty,\\ \lim\limits_{\varepsilon\to 0_+}\mathbb{E}|f^{\varepsilon,N}(t)-f(q^{\varepsilon,N}(t))|=0,\ \forall 0\le t\le T. \end{cases}\tag{30}$$

Similarly, to prove (29) for the above $q^{\varepsilon,N}(t)$ and $f(\cdot)$, we need to verify

$$\begin{cases} \sup\limits_{0\le t\le T,\varepsilon}\mathbb{E}|\hat{\mathcal{L}}^{\varepsilon,N}f^{\varepsilon,N}(t)-\bar{L}^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N})f(q^{\varepsilon,N}(t))|<\infty,\\ \lim\limits_{\varepsilon\to 0_+}\mathbb{E}|\hat{\mathcal{L}}^{\varepsilon,N}f^{\varepsilon,N}(t)-\bar{L}^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N})f(q^{\varepsilon,N}(t))|=0,\ \forall 0\le t\le T. \end{cases}\tag{31}$$

Step 1. Constructing the function $f^{\varepsilon,N}(\cdot)$ by virtue of the perturbed test function method.

For any $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$, to use the perturbed test function method, for $t<T$, define

$$\begin{aligned} V_1(t,x)&:=\int_t^T f_x(x)\mathbb{E}_t^{r^\varepsilon}[b^N(x,q_t^{\varepsilon,N},r^\varepsilon(s))-\bar{b}^N(x,q_t^{\varepsilon,N})]ds,\\ V_2(t,x)&:=\sum_{i=1}^n\int_t^T f_{x_ix_i}(x)\mathbb{E}_t^{r^\varepsilon}[a_i^N(x,q_t^{\varepsilon,N},r^\varepsilon(s))-\bar{a}_i^N(x,q_t^{\varepsilon,N})]ds, \end{aligned}$$

$f_1^{\varepsilon,N}(t)=V_1(t,q^{\varepsilon,N}(t))$ and $f_2^{\varepsilon,N}(t)=V_2(t,q^{\varepsilon,N}(t))$. In the process of building the perturbed test functions, the slow-varying process $q_i^{\varepsilon,N}(t)$ and $q_{i,t}^{\varepsilon,N}=\int_0^t\pi_i(s)q_i^{\varepsilon,N}(s)ds$ are considered as parameters for any $i=1,2,...,n$. Making change of variable s/ε to s yields that

$$\begin{aligned} f_1^{\varepsilon,N}(t)&=\varepsilon\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_x(q^{\varepsilon,N}(t))\mathbb{E}_{\frac{t}{\varepsilon}}^R[b^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},R(s))-\bar{b}^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N})]ds\\ &=\varepsilon\sum_{i=1}^n\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(q^{\varepsilon,N}(t))\mathbb{E}_{\frac{t}{\varepsilon}}^R[b_i^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},R(s))-\bar{b}_i^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N})]ds\\ &=\varepsilon\sum_{i=1}^n\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds \end{aligned}\tag{32}$$

and

$$\begin{aligned} &f_2^{\varepsilon,N}(t)\\ &=\varepsilon\sum_{i=1}^n\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(q^{\varepsilon,N}(t))\mathbb{E}_{\frac{t}{\varepsilon}}^R[a_i^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},R(s))-\bar{a}_i^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N})]ds \end{aligned}$$

$$= \varepsilon \sum_{i=1}^{n} \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i x_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2 \Big\{ \mathbb{E}\Big[g_i(R_i(s)) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i)) \Big\} ds, \tag{33}$$

where $\mu_i(g_i(r_i)) = \int_{\mathbb{R}} g_i(r_i)\mu_i(r_i)dr_i$. Define

$$f^{\varepsilon,N}(t) = f(q^{\varepsilon,N}(t)) + f_1^{\varepsilon,N}(t) + \frac{1}{2} f_2^{\varepsilon,N}(t). \tag{34}$$

Step 2. Verifying (30).

According to (30) and the form of $f^{\varepsilon,N}(t)$, we need only to estimate $f_1^{\varepsilon,N}(t)$ and $f_2^{\varepsilon,N}(t)$. Let us first estimate $f_1^{\varepsilon,N}(t)$. Since $f(\cdot) \in C_0^4(\mathbb{R}^n;\mathbb{R})$, $f_{x_i}(\cdot)$ is uniformly bounded. Note that $g_i(r_i) = a_i r_i^2 + b_i r_i + c_i$. According to the definition of the truncation function $h^N(\cdot)$, $h^N(\cdot) \in C_0^\infty(\mathbb{R}^n;\mathbb{R})$. Thus,

$$\begin{aligned}
&\sup_{0\le t\le T} \mathbb{E}|f_1^{\varepsilon,N}(t)| \\
=\ & \varepsilon \sup_{0\le t\le T} \mathbb{E}\Big| \sum_{i=1}^{n} \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(q^{\varepsilon,N}(t)) h^N(q^{\varepsilon,N}(t)) \Big\{ \mathbb{E}\Big[g_i(R_i(s)) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] \\
&\qquad - \mu_i(g_i(r_i)) \Big\} ds \Big| \\
\le\ & \varepsilon \sum_{i=1}^{n} \sup_{0\le t\le T} \mathbb{E}\Big| \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(q^{\varepsilon,N}(t)) h^N(q^{\varepsilon,N}(t)) \Big\{ a_i \mathbb{E}\Big[R_i^2(s) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] \\
&+ b_i \mathbb{E}\Big[R_i(s) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - a_i \int_{\mathbb{R}} r_i^2 \mu_i(r_i) dr_i - b_i \int_{\mathbb{R}} r_i \mu_i(r_i) dr_i \Big\} ds \Big| \\
\le\ & \varepsilon K \sum_{i=1}^{n} \sup_{0\le t\le T} \mathbb{E} \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big| a_i \mathbb{E}\Big[R_i^2(s) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] + b_i \mathbb{E}\Big[R_i(s) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] \\
&- a_i \int_{\mathbb{R}} r_i^2 \mu_i(r_i) dr_i - b_i \int_{\mathbb{R}} r_i \mu_i(r_i) dr_i \Big| ds.
\end{aligned} \tag{35}$$

According to Theorem 3.1, for any $1 \le i \le n$, by virtue of the transition probability density $p_i(x,s;x_0,t)$ and the invariant measure of the solution $R_i(t)$ of (7), we obtain

$$\begin{aligned}
&\mathbb{E}\Big[g_i(R_i(s)) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i)) \\
=\ & a_i \mathbb{E}\Big[R_i^2(s) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] + b_i \mathbb{E}\Big[R_i(s) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - a_i \int_{\mathbb{R}} r_i^2 \mu_i(r_i) dr_i - b_i \int_{\mathbb{R}} r_i \mu_i(r_i) dr_i \\
=\ & a_i \int_{-\tilde{\varphi}_i}^{+\infty} x^2 p_i\Big(x, s; R_i\Big(\frac{t}{\varepsilon}\Big), \frac{t}{\varepsilon}\Big) dx + b_i \int_{-\tilde{\varphi}_i}^{+\infty} x p_i\Big(x, s; R_i\Big(\frac{t}{\varepsilon}\Big), \frac{t}{\varepsilon}\Big) dx \\
&- a_i \int_{\mathbb{R}} r_i^2 \mu_i(r_i) dr_i - b_i \int_{\mathbb{R}} r_i \mu_i(r_i) dr_i \\
=\ & \frac{a_i \tilde{u}^2}{\gamma_i^2 \tilde{c}^2} + \frac{2a_i \tilde{u}}{\gamma_i^2 \tilde{c}^2} + \frac{2a_i \tilde{u}(\tilde{q}+1)}{\gamma_i^2 \tilde{c}^2} + \frac{a_i(\tilde{q}+1)(\tilde{q}+2)}{\gamma_i^2 \tilde{c}^2} - \frac{2a_i \tilde{u}\tilde{\beta}_i}{\gamma_i^2 \tilde{c}} - \frac{2a_i(\tilde{q}+1)\tilde{\beta}_i}{\gamma_i^2 \tilde{c}} \\
&- \frac{a_i \tilde{\beta}_i}{\gamma_i} + \frac{b_i \tilde{u}}{\gamma_i \tilde{c}} + \frac{b_i(\tilde{q}+1)}{\gamma_i \tilde{c}} - \frac{2b_i \tilde{\beta}_i}{\gamma_i} \\
=\ & \frac{a_i \hat{R}_i^2(\frac{t}{\varepsilon})}{\gamma_i^2} e^{-2\gamma_i(s-\frac{t}{\varepsilon})} + \frac{a_i \hat{R}_i(\frac{t}{\varepsilon})}{\gamma_i} e^{-\gamma_i(s-\frac{t}{\varepsilon})} \big(1 - e^{-\gamma_i(s-\frac{t}{\varepsilon})}\big)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{2a_i\hat{R}_i(\frac{t}{\varepsilon})\tilde{\beta}_i}{\gamma_i^2}e^{-\gamma_i(s-\frac{t}{\varepsilon})}-\frac{4a_i\hat{R}_i(\frac{t}{\varepsilon})\tilde{\beta}_i}{\gamma_i^2}e^{-2\gamma_i(s-\frac{t}{\varepsilon})}+\frac{4a_i\tilde{\beta}_i^2}{\gamma_i^2}\big(1-e^{-\gamma_i(s-\frac{t}{\varepsilon})}\big)^2\\
&+\frac{a_i\tilde{\beta}_i}{\gamma_i}\big(1-e^{-\gamma_i(s-\frac{t}{\varepsilon})}\big)^2-\frac{4a_i\tilde{\beta}_i^2}{\gamma_i^2}\big(1-e^{-\gamma_i(s-\frac{t}{\varepsilon})}\big)\\
&-\frac{a_i\tilde{\beta}_i}{\gamma_i}+\frac{b_i\hat{R}_i(\frac{t}{\varepsilon})}{\gamma_i}e^{-\gamma_i(s-\frac{t}{\varepsilon})}-\frac{2b_i\tilde{\beta}_i}{\gamma_i}e^{-\gamma_i(s-\frac{t}{\varepsilon})}\\
=\;&\frac{a_i}{\gamma_i^2}\tilde{\beta}_i^2e^{-2\gamma_i(s-\frac{t}{\varepsilon})}-\frac{2a_iR_i(\frac{t}{\varepsilon})}{\gamma_i}\tilde{\beta}_ie^{-2\gamma_i(s-\frac{t}{\varepsilon})}-\frac{a_i}{\gamma_i}\tilde{\beta}_ie^{-\gamma_i(s-\frac{t}{\varepsilon})}-\frac{2a_i}{\gamma_i^2}\tilde{\beta}_i^2e^{-\gamma_i(s-\frac{t}{\varepsilon})}\\
&+\frac{2a_iR_i(\frac{t}{\varepsilon})}{\gamma_i}\tilde{\beta}_ie^{-\gamma_i(s-\frac{t}{\varepsilon})}+a_iR_i^2(\frac{t}{\varepsilon})e^{-2\gamma_i(s-\frac{t}{\varepsilon})}+a_iR_i(\frac{t}{\varepsilon})e^{-\gamma_i(s-\frac{t}{\varepsilon})}\\
&-a_iR_i(\frac{t}{\varepsilon})e^{-2\gamma_i(s-\frac{t}{\varepsilon})}-\frac{b_i}{\gamma_i}\tilde{\beta}_ie^{-\gamma_i(s-\frac{t}{\varepsilon})}+b_iR_i(\frac{t}{\varepsilon})e^{-\gamma_i(s-\frac{t}{\varepsilon})}\\
=\;&\frac{a_i}{\gamma_i^2}f_i^2(q^{\varepsilon,N}(t))e^{-2\gamma_i(s-\frac{t}{\varepsilon})}-\frac{2a_iR_i(\frac{t}{\varepsilon})}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-2\gamma_i(s-\frac{t}{\varepsilon})}\\
&-\frac{a_i}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}-\frac{2a_i}{\gamma_i^2}f_i^2(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}\\
&+\frac{2a_iR_i(\frac{t}{\varepsilon})}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}+a_iR_i^2\Big(\frac{t}{\varepsilon}\Big)e^{-2\gamma_i(s-\frac{t}{\varepsilon})}\\
&+a_iR_i\Big(\frac{t}{\varepsilon}\Big)e^{-\gamma_i(s-\frac{t}{\varepsilon})}-a_iR_i\Big(\frac{t}{\varepsilon}\Big)e^{-2\gamma_i(s-\frac{t}{\varepsilon})}-\frac{b_i}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}\\
&+b_iR_i\Big(\frac{t}{\varepsilon}\Big)e^{-\gamma_i(s-\frac{t}{\varepsilon})},
\end{aligned}
\tag{36}
$$

where

$$
\tilde{c}=\frac{2}{\gamma_i\big(1-e^{-\gamma_i(s-\frac{t}{\varepsilon})}\big)},\tilde{u}=\tilde{c}\hat{R}_i\Big(\frac{t}{\varepsilon}\Big)e^{-\gamma_i(s-\frac{t}{\varepsilon})},\tilde{\beta}_i=f_i(q^{\varepsilon,N}(t)),\tilde{q}=\frac{4\tilde{\beta}_i}{\gamma_i}-1,\tilde{\varphi}_i=\frac{\tilde{\beta}_i}{\gamma_i}.
$$

Assumption (A2) shows that $f_i(x)$ is uniformly bounded with respect to $x\in G$, which implies that $f_i(x)$ is bounded for any $x\in S_{N+1}$. According to Theorem 3.1, the mth-order moment of $R_i(t)$ is uniformly bounded for any integer $m>0$. Denote $\gamma:=\max\{\gamma_i:1\le i\le n\}$. Consequently, for any $1\le i\le n$,

$$
\begin{aligned}
&\varepsilon\sup_{0\le t\le T}\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|a_i\mathbb{E}\Big[R_i^2(s)\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]+b_i\mathbb{E}\Big[R_i(s)\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-a_i\int_{\mathbb{R}}r_i^2\mu_i(r_i)dr_i\\
&-b_i\int_{\mathbb{R}}r_i\mu_i(r_i)dr_i\Big|ds\\
\le\;&\varepsilon\sup_{0\le t\le T}\mathbb{E}\Big[\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\frac{a_i}{\gamma_i^2}f_i^2(q^{\varepsilon,N}(t))e^{-2\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds\\
&+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\frac{2a_iR_i(\frac{t}{\varepsilon})}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-2\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds\\
&+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\frac{a_i}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\frac{2a_i}{\gamma_i^2}f_i^2(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds\\
&+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\frac{2a_iR_i(\frac{t}{\varepsilon})}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|a_iR_i^2\Big(\frac{t}{\varepsilon}\Big)e^{-2\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|a_iR_i\Big(\frac{t}{\varepsilon}\Big)e^{-\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|a_iR_i\Big(\frac{t}{\varepsilon}\Big)e^{-2\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds\\
&+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\frac{b_i}{\gamma_i}f_i(q^{\varepsilon,N}(t))e^{-\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds+\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|b_iR_i\Big(\frac{t}{\varepsilon}\Big)e^{-\gamma_i(s-\frac{t}{\varepsilon})}\Big|ds\Big]\\
\le\ & \varepsilon K\sup_{0\le t\le T}\Big(\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|^2+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|+1\Big)\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds\\
\le\ & \varepsilon K\sup_{0\le t\le T}\big(1-e^{-\frac{\gamma_i(T-t)}{\varepsilon}}\big)\le\varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon). \qquad (37)
\end{aligned}
$$

According to (35) and (37), we obtain

$$
\sup_{0\le t\le T}\mathbb{E}|f_1^{\varepsilon,N}(t)|\le\varepsilon K\sum_{i=1}^{n}\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)\le\varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon), \qquad (38)
$$

which implies $\sup_{0\le t\le T}\mathbb{E}|f_1^{\varepsilon,N}(t)|\to 0$ as $\varepsilon\to 0$.

To proceed, let us estimate $f_2^{\varepsilon,N}(t)$. Since $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot)\in C_0^\infty(\mathbb{R}^n;\mathbb{R})$, $f_{x_ix_i}(\cdot)$ and $h^N(\cdot)$ are uniformly bounded. Note that $g_i(r_i)=a_ir_i^2+b_ir_i+c_i$. According to (36), applying the technique similar to (37), gives

$$
\begin{aligned}
&\sup_{0\le t\le T}\mathbb{E}|f_2^{\varepsilon,N}(t)|\\
=\ & \varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{i=1}^{n}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}f_{x_ix_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]\\
&\qquad -\mu_i(g_i(r_i))\Big\}ds\Big|\\
\le\ & \varepsilon\sum_{i=1}^{n}\sup_{0\le t\le T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}f_{x_ix_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2\Big\{a_i\mathbb{E}\Big[R_i^2(s)\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]\\
&+b_i\mathbb{E}\Big[R_i(s)\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-a_i\int_{\mathbb{R}}r_i^2\mu_i(r_i)dr_i-b_i\int_{\mathbb{R}}r_i\mu_i(r_i)dr_i\Big\}ds\Big|\\
\le\ & \varepsilon K\sum_{i=1}^{n}\sup_{0\le t\le T}\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|a_i\mathbb{E}\Big[R_i^2(s)\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]+b_i\mathbb{E}\Big[R_i(s)\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]\\
&-a_i\int_{\mathbb{R}}r_i^2\mu_i(r_i)dr_i-b_i\int_{\mathbb{R}}r_i\mu_i(r_i)dr_i\Big|ds\\
\le\ & \varepsilon K\sum_{i=1}^{n}\sup_{0\le t\le T}\Big(\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|^2+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|+1\Big)\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds\\
\le\ & \varepsilon K\sum_{i=1}^{n}\sup_{0\le t\le T}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds\le\varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon), \qquad (39)
\end{aligned}
$$

which implies $\sup_{0\le t\le T}\mathbb{E}|f_2^{\varepsilon,N}(t)|\to 0$ as $\varepsilon\to 0$. Note that

$$
\mathbb{E}|f^{\varepsilon,N}(t)-f(q^{\varepsilon,N}(t))|\le\mathbb{E}|f_1^{\varepsilon,N}(t)|+\frac{1}{2}\mathbb{E}|f_2^{\varepsilon,N}(t)|.
$$

(38) and (39) imply

$$
\sup_{0\le t\le T}\mathbb{E}|f^{\varepsilon,N}(t)-f(q^{\varepsilon,N}(t))|\to 0,\ as\ \varepsilon\to 0. \qquad (40)
$$

Thus, (30) holds.

Step 3. Verifying (31).

For any $(t,x)\in[0,T]\times\mathbb{R}^n$, let us define

$$V(t,x)=V_1(t,x)+\frac{1}{2}V_2(t,x).$$

According to the definitions of $f^{\varepsilon,N}(\cdot)$, $\hat{\mathcal{L}}^{\varepsilon,N}$ and L^N, gives

$$\begin{aligned}
&\hat{\mathcal{L}}^{\varepsilon,N}f^{\varepsilon,N}(t)\\
&=\text{p-}\lim_{\delta\to0}\frac{\mathbb{E}_t^{\varepsilon,N}f^{\varepsilon,N}(t+\delta)-f^{\varepsilon,N}(t)}{\delta}\\
&=\text{p-}\lim_{\delta\to0}\frac{\mathbb{E}_t^{\varepsilon,N}f(q^{\varepsilon,N}(t+\delta))-f(q^{\varepsilon,N}(t))}{\delta}\\
&+\text{p-}\lim_{\delta\to0}\frac{\mathbb{E}_t^{\varepsilon,N}V(t+\delta,q^{\varepsilon,N}(t+\delta))-V(t,q^{\varepsilon,N}(t))}{\delta}\qquad(41)\\
&=L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))f(q^{\varepsilon,N}(t))+L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V(t,q^{\varepsilon,N}(t)),
\end{aligned}$$

where $\mathbb{E}_t^{\varepsilon,N}$ is the conditional expectation with respect to $\mathcal{F}_t^{\varepsilon,N}$ and

$$\begin{aligned}
L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V(t,q^{\varepsilon,N}(t)) &= L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_1(t,q^{\varepsilon,N}(t))\\
&\quad+\frac{1}{2}L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_2(t,q^{\varepsilon,N}(t)).
\end{aligned}$$

According to the form of $\hat{\mathcal{L}}^{\varepsilon,N}f^{\varepsilon,N}(t)$, we first need to estimate $L^N(q^{\varepsilon,N}(t),\,q_t^{\varepsilon,N},\,r^\varepsilon(t))V_1(t,\,q^{\varepsilon,N}(t))$ and $L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_2(t,q^{\varepsilon,N}(t))$.

Let us first estimate $L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_1(t,q^{\varepsilon,N}(t))$. Note that

$$\begin{aligned}
&L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_1(t,q^{\varepsilon,N}(t))\\
&=\varepsilon\sum_{i=1}^n L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))\Big\{\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}f_{x_i}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]\\
&\quad-\mu_i(g_i(r_i))\Big\}ds\Big\}.\qquad(42)
\end{aligned}$$

For any $1\le i\le n$, applying the Itô formula gives

$$\begin{aligned}
&\varepsilon L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))\Big\{\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}f_{x_i}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]\\
&\quad-\mu_i(g_i(r_i))\Big\}ds\Big\}\\
&=-f_{x_i}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\Big[g_i\Big(R_i\Big(\frac{t}{\varepsilon}\Big)\Big)-\mu_i(g_i(r_i))\Big]+\varepsilon\sum_{j=1}^n I_{1j}^\varepsilon(t,q^{\varepsilon,N}(t))\\
&+\varepsilon\sum_{j=1}^n I_{2j}^\varepsilon(t,q^{\varepsilon,N}(t))+\varepsilon\sum_{j=1}^n I_{3j}^\varepsilon(t,q^{\varepsilon,N}(t))+\frac{\varepsilon}{2}\sum_{j=1}^n I_{4j}^\varepsilon(t,q^{\varepsilon,N}(t))\\
&+\varepsilon\sum_{j=1}^n I_{5j}^\varepsilon(t,q^{\varepsilon,N}(t))+\varepsilon\sum_{j=1}^n I_{6j}^\varepsilon(t,q^{\varepsilon,N}(t))+\frac{\varepsilon}{2}\sum_{j=1}^n I_{7j}^\varepsilon(t,q^{\varepsilon,N}(t))\\
&+\varepsilon\sum_{j=1}^n I_{8j}^\varepsilon(t,q^{\varepsilon,N}(t))+\frac{\varepsilon}{2}\sum_{j=1}^n I_{9j}^\varepsilon(t,q^{\varepsilon,N}(t)),\qquad(43)
\end{aligned}$$

where

$$
\begin{aligned}
I_{1j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_j}(x)h^N(x)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times b_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
I_{2j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(x)\frac{\partial}{\partial x_j}h^N(x)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times b_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
I_{3j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(x)h^N(x)\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times b_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
I_{4j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_jx_j}(x)h^N(x)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
I_{5j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_j}(x)\frac{\partial}{\partial x_j}h^N(x)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
I_{6j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_j}(x)h^N(x)\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t))\\
I_{7j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(x)\frac{\partial^2}{\partial x_j^2}h^N(x)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
I_{8j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(x)\frac{\partial}{\partial x_j}h^N(x)\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
I_{9j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(x)h^N(x)\frac{\partial^2}{\partial x_j^2}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad\times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)).
\end{aligned}
$$

Since $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot)\in C_0^{\infty}(\mathbb{R}^n;\mathbb{R})$, $f_{x_ix_j}(\cdot)$ and $h^N(\cdot)$ are uniformly bounded. Consequently,

$$
\begin{aligned}
&\varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^{n}I_{1j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|\\
\le\;&\varepsilon K\sum_{j=1}^{n}\sup_{0\le t\le T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\times b_j^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))\Big|.
\end{aligned}
$$

For any $1 \le j \le n$, note that $b_j^N(x,y,r) = h^N(x)\big(g_j(r_j) - \delta_j x_j - y_j\big)$ and $g_j(r_j) = a_j r_j^2 + b_j r_j + c_j$. Assumption (A2) shows that $\pi_i(t)$ is uniformly bounded with respect to $t \in [0,T]$, which implies that $q_{i,t}^{\varepsilon,N} = \int_0^t \pi_i(s) q_i^{\varepsilon,N}(s)ds$ is uniformly bounded. Recall that $h^N(\cdot) \in C_0^\infty(\mathbb{R}^n;\mathbb{R})$ and the mth-order moment of $R_i(t)$ is uniformly bounded for any integer $m > 0$. $R_i(t)$ and $R_j(t)$ are independent for any $i \ne j$ and $t \ge 0$. According to (36), we obtain

$$
\begin{aligned}
&\mathbb{E}\Big| \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i))\Big\}ds \cdot b_j^N(q^{\varepsilon,N}(t), q_t^{\varepsilon,N}, r^\varepsilon(t))\Big| \\
\le\ & \mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big|g_j\Big(R_j\Big(\frac{t}{\varepsilon}\Big)\Big)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i))\Big\}\Big|ds \\
&+\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big|\big(\delta_j q_j^{\varepsilon,N}(t) + q_{j,t}^{\varepsilon,N}\big)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i))\Big\}\Big|ds \\
\le\ & K\Big(\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|^2\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2 + \mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|^2\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big| + \mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2 \\
&+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|^2 + \mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big| + \mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2 + \mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big| \\
&+\mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big| + 1\Big)\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds \\
\le\ & K\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds \\
\le\ & K\big(1 - e^{-\gamma_i(\frac{T}{\varepsilon}-\frac{t}{\varepsilon})}\big), \qquad (44)
\end{aligned}
$$

which implies

$$
\begin{aligned}
\varepsilon \sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n I_{1j}^\varepsilon(t, q^{\varepsilon,N}(t))\Big| &\le \varepsilon K\sum_{j=1}^n \sup_{0\le t\le T}\big(1 - e^{-\gamma_i(\frac{T}{\varepsilon}-\frac{t}{\varepsilon})}\big) \\
&\le \varepsilon K\big(1 - e^{-\frac{\gamma T}{\varepsilon}}\big) = O(\varepsilon).
\end{aligned}
$$

Since $f(\cdot) \in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot) \in C_0^\infty(\mathbb{R}^n;\mathbb{R})$, $f_{x_i}(\cdot)$, $h^N(\cdot)$ and $\frac{\partial}{\partial x_j}h^N(\cdot)$ are uniformly bounded. Note that $b_j^N(x,y,r) = h^N(x)\big(g_j(r_j) - \delta_j x_j - y_j\big)$ for any $j = 1, 2, ..., n$. Applying the technique similar to (44), gives

$$
\begin{aligned}
&\varepsilon \sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n I_{2j}^\varepsilon(t, q^{\varepsilon,N}(t))\Big| \\
\le\ & \varepsilon K\sum_{j=1}^n \sup_{0\le t\le T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i))\Big\}ds \\
&\times b_j^N(q^{\varepsilon,N}(t), q_t^{\varepsilon,N}, r^\varepsilon(t))\Big| \\
\le\ & \varepsilon K\sum_{j=1}^n \sup_{0\le t\le T}\Big\{\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big|g_j\Big(R_j\Big(\frac{t}{\varepsilon}\Big)\Big)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i))\Big\}\Big|ds \\
&+\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big|\big(\delta_j q_j^{\varepsilon,N}(t) + q_{j,t}^{\varepsilon,N}\big)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i))\Big\}\Big|ds\Big\}
\end{aligned}
$$

$$\leq \quad \varepsilon K\sum_{j=1}^{n}\sup_{0\leq t\leq T}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds \leq \varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon).$$

Since $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot)\in C_0^\infty(\mathbb{R}^n;\mathbb{R})$, $f_{x_ix_jx_j}(\cdot)$ and $h^N(\cdot)$ are uniformly bounded. Note that $a_j^N(x,y,r)=(h^N(x))^2\big(g_j(r_j)+\delta_j x_j+y_j\big)$ for any $j=1,2,...,n$. Applying the technique similar to (44), gives

$$\begin{aligned}
&\frac{\varepsilon}{2}\sup_{0\leq t\leq T}\mathbb{E}\Big|\sum_{j=1}^{n} I_{4j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|\\
\leq\ & \frac{\varepsilon}{2}K\sum_{j=1}^{n}\sup_{0\leq t\leq T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\times a_j^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))\Big|\\
\leq\ & \varepsilon K\sum_{j=1}^{n}\sup_{0\leq t\leq T}\Big\{\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|g_j\Big(R_j\Big(\frac{t}{\varepsilon}\Big)\Big)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds\\
&+\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\big(\delta_j q_j^{\varepsilon,N}(t)+q_{j,t}^{\varepsilon,N}\big)\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds\Big\}\\
\leq\ & \varepsilon K\sum_{j=1}^{n}\sup_{0\leq t\leq T}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds \leq \varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon).
\end{aligned}$$

Note that $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$, $h^N(\cdot)\in C_0^\infty(\mathbb{R}^n;\mathbb{R})$ and $a_j^N(x,y,r)=(h^N(x))^2\big(g_j(r_j)+\delta_j x_j+y_j\big)$ for any $j=1,2,...,n$. Similarly,

$$\begin{aligned}
&\varepsilon\sup_{0\leq t\leq T}\mathbb{E}\Big|\sum_{j=1}^{n} I_{5j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|\\
\leq\ & \varepsilon K\sum_{j=1}^{n}\sup_{0\leq t\leq T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\times a_j^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))\Big|\\
\leq\ & \varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\varepsilon}{2}\sup_{0\leq t\leq T}\mathbb{E}\Big|\sum_{j=1}^{n} I_{7j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|\\
\leq\ & \frac{\varepsilon}{2}K\sum_{j=1}^{n}\sup_{0\leq t\leq T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\times a_j^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))\Big|\\
=\ & O(\varepsilon).
\end{aligned}$$

In order to estimate all remaining terms, we first need to calculate the partial derivatives of the conditional expectation. For any $1\leq j\leq n$,

$$\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}$$

$$
\begin{aligned}
&= \frac{2a_i}{\gamma_i^2} f_i(q^{\varepsilon,N}(t)) \frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) e^{-2\gamma_i(s-\frac{t}{\varepsilon})} - \frac{2a_i R_i(\frac{t}{\varepsilon})}{\gamma_i} \frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) e^{-2\gamma_i(s-\frac{t}{\varepsilon})} \\
&\quad - \frac{a_i}{\gamma_i} \frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})} - \frac{4a_i}{\gamma_i^2} f_i(q^{\varepsilon,N}(t)) \frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})} \\
&\quad + \frac{2a_i R_i(\frac{t}{\varepsilon})}{\gamma_i} \frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})} - \frac{b_i}{\gamma_i} \frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})} \qquad (45)
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{\partial^2}{\partial x_j^2} \Big\{ \mathbb{E}\Big[g_i(R_i(s)) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i)) \Big\} \\
&= \frac{2a_i}{\gamma_i^2} \Big[\frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) \Big]^2 e^{-2\gamma_i(s-\frac{t}{\varepsilon})} + \frac{2a_i}{\gamma_i^2} f_i(q^{\varepsilon,N}(t)) \frac{\partial^2}{\partial x_j^2} f_i(q^{\varepsilon,N}(t)) e^{-2\gamma_i(s-\frac{t}{\varepsilon})} \\
&\quad - \frac{2a_i R_i(\frac{t}{\varepsilon})}{\gamma_i} \frac{\partial^2}{\partial x_j^2} f_i(q^{\varepsilon,N}(t)) e^{-2\gamma_i(s-\frac{t}{\varepsilon})} - \frac{a_i}{\gamma_i} \frac{\partial^2}{\partial x_j^2} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})} \\
&\quad - \frac{4a_i}{\gamma_i^2} \Big[\frac{\partial}{\partial x_j} f_i(q^{\varepsilon,N}(t)) \Big]^2 e^{-\gamma_i(s-\frac{t}{\varepsilon})} - \frac{4a_i}{\gamma_i^2} f_i(q^{\varepsilon,N}(t)) \frac{\partial^2}{\partial x_j^2} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})} \\
&\quad + \frac{2a_i R_i(\frac{t}{\varepsilon})}{\gamma_i} \frac{\partial^2}{\partial x_j^2} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})} - \frac{b_i}{\gamma_i} \frac{\partial^2}{\partial x_j^2} f_i(q^{\varepsilon,N}(t)) e^{-\gamma_i(s-\frac{t}{\varepsilon})}. \qquad (46)
\end{aligned}
$$

Assumption (A2) shows that $\frac{\partial f_i}{\partial x_j}$ is uniformly bounded with respect to $x \in G$, which implies that $\frac{\partial f_i}{\partial x_j}$ is bounded for any $x \in S_{N+1}$. Since $f(\cdot) \in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot) \in C_0^\infty(\mathbb{R}^n;\mathbb{R})$, $f_{x_i}(\cdot)$ and $h^N(\cdot)$ are uniformly bounded. For any $1 \le j \le n$, note that $b_j^N(x,y,r) = h^N(x)\big(g_j(r_j) - \delta_j x_j - y_j\big)$ and $g_j(r_j) = a_j r_j^2 + b_j r_j + c_j$. Therefore, according to (45) and applying the same technique as the estimate of (44), we obtain

$$
\begin{aligned}
&\varepsilon \sup_{0\le t\le T} \mathbb{E}\Big| \sum_{j=1}^n I_{3j}^\varepsilon(t, q^{\varepsilon,N}(t)) \Big| \\
&\le \varepsilon \sum_{j=1}^n \sup_{0\le t\le T} \mathbb{E}\Big| \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(q^{\varepsilon,N}(t)) h^N(q^{\varepsilon,N}(t)) \frac{\partial}{\partial x_j} \Big\{ \mathbb{E}\Big[g_i(R_i(s)) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] \\
&\quad - \mu_i(g_i(r_i)) \Big\} ds b_j^N(q^{\varepsilon,N}(t), q_t^{\varepsilon,N}, r^\varepsilon(t)) \Big| \\
&\le \varepsilon K \sum_{j=1}^n \sup_{0\le t\le T} \Big\{ \mathbb{E} \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big| g_j\Big(R_j\Big(\frac{t}{\varepsilon}\Big)\Big) \frac{\partial}{\partial x_j} \Big\{ \mathbb{E}\Big[g_i(R_i(s)) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i)) \Big\} \Big| ds \\
&\quad + \mathbb{E} \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} \Big| \big(\delta_j q_j^{\varepsilon,N}(t) + q_{j,t}^{\varepsilon,N}\big) \frac{\partial}{\partial x_j} \Big\{ \mathbb{E}\Big[g_i(R_i(s)) \Big| R_i\Big(\frac{t}{\varepsilon}\Big)\Big] - \mu_i(g_i(r_i)) \Big\} \Big| ds \Big\} \\
&\le \varepsilon K \sum_{j=1}^n \sup_{0\le t\le T} \Big(\mathbb{E}\Big| R_i\Big(\frac{t}{\varepsilon}\Big) \Big| \Big| R_j\Big(\frac{t}{\varepsilon}\Big) \Big|^2 + \mathbb{E}\Big| R_i\Big(\frac{t}{\varepsilon}\Big) \Big| \Big| R_j\Big(\frac{t}{\varepsilon}\Big) \Big| + \mathbb{E}\Big| R_j\Big(\frac{t}{\varepsilon}\Big) \Big|^2 \\
&\quad + \mathbb{E}\Big| R_i\Big(\frac{t}{\varepsilon}\Big) \Big| + \mathbb{E}\Big| R_j\Big(\frac{t}{\varepsilon}\Big) \Big| + 1 \Big) \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})} ds \\
&\le \varepsilon K \sup_{0\le t\le T} \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})} ds
\end{aligned}
$$

$$\leq \quad \varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon).$$

Since $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot)\in C_0^\infty(\mathbb{R}^n;\mathbb{R})$, $f_{x_i}(\cdot)$, $f_{x_ix_j}(\cdot)$, $h^N(\cdot)$ and $\frac{\partial}{\partial x_j}h^N(\cdot)$ are uniformly bounded. Note that $a_j^N(x,y,r)=(h^N(x))^2\big(g_j(r_j)+\delta_j x_j+y_j\big)$ for any $j=1,2,...,n$. Similarly,

$$\begin{aligned}
&\varepsilon\sup_{0\leq t\leq T}\mathbb{E}\Big|\sum_{j=1}^n I_{6j}^\varepsilon(t,q^{\varepsilon,N}(t))\Big|\\
\leq\ &\varepsilon\sum_{j=1}^n\sup_{0\leq t\leq T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_j}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]\\
&-\mu_i(g_i(r_i))\Big\}ds\times a_j^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))\Big|\\
\leq\ &\varepsilon K\sum_{j=1}^n\sup_{0\leq t\leq T}\Big\{\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|g_j\Big(R_j\Big(\frac{t}{\varepsilon}\Big)\Big)\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds\\
&+\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\big(\delta_j q_j^{\varepsilon,N}(t)+q_{j,t}^{\varepsilon,N}\big)\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds\Big\}\\
\leq\ &\varepsilon K\sum_{j=1}^n\sup_{0\leq t\leq T}\Big(\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|+\mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2\\
&+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|+\mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|+1\Big)\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds\\
\leq\ &\varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon)
\end{aligned}$$

and

$$\begin{aligned}
&\varepsilon\sup_{0\leq t\leq T}\mathbb{E}\Big|\sum_{j=1}^n I_{8j}^\varepsilon(t,q^{\varepsilon,N}(t))\Big|\\
\leq\ &\varepsilon\sum_{j=1}^n\sup_{0\leq t\leq T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(q^{\varepsilon,N}(t))\frac{\partial}{\partial x_j}h^N(q^{\varepsilon,N}(t))\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]\\
&-\mu_i(g_i(r_i))\Big\}ds a_j^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))\Big|\\
\leq\ &\varepsilon K\sum_{j=1}^n\sup_{0\leq t\leq T}\Big\{\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|g_j\Big(R_j\Big(\frac{t}{\varepsilon}\Big)\Big)\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds\\
&+\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\big(\delta_j q_j^{\varepsilon,N}(t)+q_{j,t}^{\varepsilon,N}\big)\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds\Big\}\\
\leq\ &\varepsilon K\sum_{j=1}^n\sup_{0\leq t\leq T}\Big(\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|+\mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2\\
&+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|+\mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|+1\Big)\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds\\
\leq\ &\varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big)=O(\varepsilon).
\end{aligned}$$

Assumption (A2) shows that $\frac{\partial^2 f_i}{\partial x_j^2}$ is uniformly bounded with respect to $x \in G$, which implies that $\frac{\partial^2 f_i}{\partial x_j^2}$ is bounded for any $x \in S_{N+1}$. Since $f(\cdot) \in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot) \in C_0^\infty(\mathbb{R}^n;\mathbb{R})$, $f_{x_i}(\cdot)$ and $h^N(\cdot)$ are uniformly bounded. Note that $a_j^N(x,y,r) = (h^N(x))^2\big(g_j(r_j) + \delta_j x_j + y_j\big)$ for any $j = 1,2,...,n$. Therefore, according to (46) and applying the same technique as the estimate of (44), we obtain

$$
\begin{aligned}
&\frac{\varepsilon}{2}\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n I_{9j}^\varepsilon(t,q^{\varepsilon,N}(t))\Big| \\
\le\ & \frac{\varepsilon}{2}\sum_{j=1}^n \sup_{0\le t\le T}\mathbb{E}\Big|\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_i}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\frac{\partial^2}{\partial x_j^2}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] \\
&-\mu_i(g_i(r_i))\Big\}ds a_j^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))\Big| \\
\le\ & \varepsilon K\sum_{j=1}^n \sup_{0\le t\le T}\Big\{\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|g_j\Big(R_j\Big(\frac{t}{\varepsilon}\Big)\Big)\frac{\partial^2}{\partial x_j^2}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds \\
&+\mathbb{E}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}}\Big|\big(\delta_j q_j^{\varepsilon,N}(t)+q_{j,t}^{\varepsilon,N}\big)\frac{\partial^2}{\partial x_j^2}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}\Big|ds\Big\} \\
\le\ & \varepsilon K\sum_{j=1}^n \sup_{0\le t\le T}\Big(\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|+\mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|^2 \\
&+\mathbb{E}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big|+\mathbb{E}\Big|R_j\Big(\frac{t}{\varepsilon}\Big)\Big|+1\Big)\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds \\
&\le \varepsilon K\sup_{0\le t\le T}\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} e^{-\gamma_i(s-\frac{t}{\varepsilon})}ds \le \varepsilon K\big(1-e^{-\frac{\gamma T}{\varepsilon}}\big) = O(\varepsilon).
\end{aligned}
$$

According to (42) and (43), these estimates lead to

$$
\begin{aligned}
&\sup_{0\le t\le T}\mathbb{E}\Big|L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_1(t,q^{\varepsilon,N}(t)) \\
&+\sum_{i=1}^n f_{x_i}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\Big[g_i\Big(R_i\Big(\frac{t}{\varepsilon}\Big)\Big)-\mu_i(g_i(r_i))\Big]\Big| \\
=\ & O(\varepsilon). \qquad (47)
\end{aligned}
$$

To proceed, let us estimate $L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_2(t,q^{\varepsilon,N}(t))$. Note that

$$
\begin{aligned}
&L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_2(t,q^{\varepsilon,N}(t)) \\
=\ & \varepsilon\sum_{i=1}^n L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))\Big\{\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2 \\
&\times\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\Big\}. \qquad (48)
\end{aligned}
$$

For any $1 \le i \le n$, applying the Itô formula, gives

$$
\begin{aligned}
&\varepsilon L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))\Big\{\int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2 \\
&\qquad\times\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\Big\}
\end{aligned}
$$

$$
\begin{aligned}
= \quad & -f_{x_ix_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2\Big[g_i\Big(R_i\Big(\frac{t}{\varepsilon}\Big)\Big)-\mu_i(g_i(r_i))\Big]+\varepsilon\sum_{j=1}^{n}J_{1j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\\
& +\varepsilon\sum_{j=1}^{n}J_{2j}^{\varepsilon}(t,q^{\varepsilon,N}(t))+\varepsilon\sum_{j=1}^{n}J_{3j}^{\varepsilon}(t,q^{\varepsilon,N}(t))+\frac{\varepsilon}{2}\sum_{j=1}^{n}J_{4j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\\
& +\varepsilon\sum_{j=1}^{n}J_{5j}^{\varepsilon}(t,q^{\varepsilon,N}(t))+\varepsilon\sum_{j=1}^{n}J_{6j}^{\varepsilon}(t,q^{\varepsilon,N}(t))+\frac{\varepsilon}{2}\sum_{j=1}^{n}J_{7j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\\
& +\varepsilon\sum_{j=1}^{n}J_{8j}^{\varepsilon}(t,q^{\varepsilon,N}(t))+\frac{\varepsilon}{2}\sum_{j=1}^{n}J_{9j}^{\varepsilon}(t,q^{\varepsilon,N}(t)),
\end{aligned}
\tag{49}
$$

where

$$
\begin{aligned}
J_{1j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_ix_j}(x)(h^N(x))^2\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times b_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
J_{2j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(x)\frac{\partial}{\partial x_j}(h^N(x))^2\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times b_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
J_{3j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(x)(h^N(x))^2\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times b_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
J_{4j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_ix_jx_j}(x)(h^N(x))^2\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
J_{5j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_ix_j}(x)\frac{\partial}{\partial x_j}(h^N(x))^2\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
J_{6j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_ix_j}(x)(h^N(x))^2\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
J_{7j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(x)\frac{\partial^2}{\partial x_j^2}(h^N(x))^2\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),\\
J_{8j}^{\varepsilon}(t,x) &= \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(x)\frac{\partial}{\partial x_j}(h^N(x))^2\frac{\partial}{\partial x_j}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds\\
&\quad \times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)),
\end{aligned}
$$

$$J_{9j}^{\varepsilon}(t,x) = \int_{\frac{t}{\varepsilon}}^{\frac{T}{\varepsilon}} f_{x_ix_i}(x)(h^N(x))^2\frac{\partial^2}{\partial x_j^2}\Big\{\mathbb{E}\Big[g_i(R_i(s))\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big]-\mu_i(g_i(r_i))\Big\}ds \times a_j^N(x,q_t^{\varepsilon,N},r^{\varepsilon}(t)).$$

Recall that $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$ and $h^N(\cdot)\in C_0^{\infty}(\mathbb{R}^n;\mathbb{R})$. Applying the same technique as the estimates of $I_{1j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$, $I_{2j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$, $I_{4j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$, $I_{5j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$ and $I_{7j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$, we obtain

$$\varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{1j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon),\ \varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{2j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon),$$
$$\frac{\varepsilon}{2}\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{4j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon),\ \varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{5j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon)$$

and

$$\frac{\varepsilon}{2}\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{7j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon).$$

According to the partial derivatives (45) and (46) of the conditional expectation. Similarly, applying the same technique as the estimates of $I_{3j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$, $I_{6j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$, $I_{8j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$ and $I_{9j}^{\varepsilon}(t,q^{\varepsilon,N}(t))$, we obtain

$$\varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{3j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon),\ \varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{6j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon),$$
$$\varepsilon\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{8j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon),\ \frac{\varepsilon}{2}\sup_{0\le t\le T}\mathbb{E}\Big|\sum_{j=1}^n J_{9j}^{\varepsilon}(t,q^{\varepsilon,N}(t))\Big|=O(\varepsilon).$$

According to (48) and (49), these estimates lead to

$$\sup_{0\le t\le T}\mathbb{E}\Big|L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))V_2(t,q^{\varepsilon,N}(t)) + \sum_{i=1}^n f_{x_ix_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2\Big[g_i\Big(R_i\Big(\frac{t}{\varepsilon}\Big)\Big)-\mu_i(g_i(r_i))\Big]\Big|=O(\varepsilon). \quad (50)$$

The estimates of $L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))V_1(t,q^{\varepsilon,N}(t))$ and $L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))\times V_2(t,q^{\varepsilon,N}(t))$, together with (41), yield

$$\begin{aligned}
&\sup_{0\le t\le T}\mathbb{E}\big|\hat{\mathcal{L}}^{\varepsilon,N}f^{\varepsilon,N}(t)-\bar{L}^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N})f(q^{\varepsilon,N}(t))\big|\\
=\ &\sup_{0\le t\le T}\mathbb{E}\big|L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))f(q^{\varepsilon,N}(t))\\
&+L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))V_1(t,q^{\varepsilon,N}(t))\\
&+\frac{1}{2}L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))V_2(t,q^{\varepsilon,N}(t))-\bar{L}^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N})f(q^{\varepsilon,N}(t))\big|\\
\le\ &\sup_{0\le t\le T}\mathbb{E}\Big|L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^{\varepsilon}(t))V_1(t,q^{\varepsilon,N}(t))\\
&+\sum_{i=1}^n f_{x_i}(q^{\varepsilon,N}(t))h^N(q^{\varepsilon,N}(t))\Big[g_i\Big(R_i\Big(\frac{t}{\varepsilon}\Big)\Big)-\mu_i(g_i(r_i))\Big]\Big|
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\sup_{0\le t\le T}\mathbb{E}\Big|L^N(q^{\varepsilon,N}(t),q_t^{\varepsilon,N},r^\varepsilon(t))V_2(t,q^{\varepsilon,N}(t))\\
&+\sum_{i=1}^{n}f_{x_ix_i}(q^{\varepsilon,N}(t))[h^N(q^{\varepsilon,N}(t))]^2\Big[g_i\Big(R_i\Big(\frac{t}{\varepsilon}\Big)\Big)-\mu_i(g_i(r_i))\Big]\Big|\\
=\;& O(\varepsilon), \qquad (51)
\end{aligned}
$$

which implies that (31) holds. This, together with (30) yield $q^{\varepsilon,N}(\cdot)\Rightarrow q^N(\cdot)$ as $\varepsilon\to 0$ by virtue of Lemma 4.5, where $q^N(\cdot)$ satisfies the stochastic differential equation (20).

Step 4. Removing truncation for the weak convergence.

The argument is similar to that of [14, p.46]. For any deterministic initial value $q(0)$, let $\mathbb{P}(\cdot)$ and $\mathbb{P}^N(\cdot)$ denote the probabilities induced by $q(\cdot)$ and $q^N(\cdot)$, respectively, on the Borel sets of $D([0,T];\mathbb{R}^n)$. By (A3), the martingale problem has a unique solution for each $q(0)$, so $\mathbb{P}(\cdot)$ is unique. For each $T<\infty$, the uniqueness implies that $\mathbb{P}(\cdot)$ agrees with $\mathbb{P}^N(\cdot)$ on all Borel sets of the set of paths in $D([0,T];S_N)$ for each $t\le T$. However, $\mathbb{P}\{\sup_{0\le t\le T}|q(t)|\le N\}\to 1$ as $N\to\infty$. This together with the weak convergence of $q^{\varepsilon,N}(t)$ imply that $q^\varepsilon(\cdot)\Rightarrow q(\cdot)$. Moreover, the uniqueness implies that the limit does not depend on the chosen subsequences. The proof is thus completed. □

At the end of this section, let us examine equation (4) by using this theorem. For the first equation of (4), letting $R(t)=r^\varepsilon(\varepsilon t)$ yields

$$dR(t)=(k_r-\gamma_rR(t))dt+\sqrt{k_r+\gamma_rR(t)}d\tilde{w}_1(t).$$

Moreover, define $\hat{R}(t)=k_r+\gamma_rR(t)$. Then,

$$d\hat{R}(t)=\gamma_r(2k_r-\hat{R}(t))dt+\gamma_r\sqrt{\hat{R}(t)}d\tilde{w}_1(t),$$

which has a unique global solution, known as the mean-reverting square root process. Moreover, this solution has a unique invariant measure μ. According to section 3, it is easy to obtain that

$$\lim_{t\to\infty}\mathbb{E}R(t)=\mathbb{E}_\mu R=\frac{k_r}{\gamma_r},\ \lim_{t\to\infty}\mathbb{E}[R(t)]^2=\mathbb{E}_\mu R^2=\frac{k_r^2}{\gamma_r^2}+\frac{k_r}{\gamma_r}.$$

To obtain the desired asymptotic results, we assume that (a) $\gamma_d(t)$ is uniformly bounded with respect to $t\in[0,T]$ and (b) the second equation of (4) has a unique strong solution $q^\varepsilon(t)$ on $[0,T]$ for each $r^\varepsilon(t)=R(t/\varepsilon)$ with each deterministic initial value $q(0)$. Applying Theorem 4.1 yields that there exists a standard Brownian motion $B(t)$ such that $q^\varepsilon(t)$ in (4) converges weakly to $q(t)$ satisfying the following stochastic differential delay equation

$$
\begin{aligned}
dq(t)=&\Big(\frac{k_pk_r}{\gamma_r}-\gamma_pq(t)-\int_0^t\gamma_d(s)q(s)ds\Big)dt\\
&+\sqrt{\frac{k_pk_r}{\gamma_r}+\gamma_pq(t)+\int_0^t\gamma_d(s)q(s)ds}dB(t).
\end{aligned}
\qquad (52)
$$

5. **Concluding remarks.** The results of Theorem 4.1 also hold when $g_i(\cdot):\mathbb{R}\to\mathbb{R}_+$ is an arbitrary order polynomial and satisfies (A1)-(A3). In fact, for any $0<\alpha<2$,

$$\begin{aligned}\mathbb{E}\Big[e^{\alpha R_i(s)}\Big|R_i\Big(\frac{t}{\varepsilon}\Big)\Big] &= \int_{-\tilde{\varphi}_i}^{+\infty} e^{\alpha x} p_i\Big(x,s;R_i\Big(\frac{t}{\varepsilon}\Big),\frac{t}{\varepsilon}\Big)dx\\ &= \int_0^{+\infty} e^{\alpha\left(\frac{y-\tilde{\beta}_i}{\gamma_i}\right)}\hat{p}_i\Big(y,s;\hat{R}_i\Big(\frac{t}{\varepsilon}\Big),\frac{t}{\varepsilon}\Big)dy\\ &= e^{-\alpha\tilde{\varphi}_i}e^{-\tilde{u}}\sum_{m=0}^{\infty}\frac{\tilde{c}^{m+\tilde{q}+1}\tilde{u}^m}{m!\Gamma(m+\tilde{q}+1)}\int_0^{+\infty}e^{-\left(\tilde{c}-\frac{\alpha}{\gamma_i}\right)y}y^{m+\tilde{q}}dy\\ &= e^{-\alpha\tilde{\varphi}_i}e^{-\tilde{u}}\sum_{m=0}^{\infty}\frac{\tilde{u}^m}{m!}\Big(1-\frac{\alpha}{\tilde{c}\gamma_i}\Big)^{-(m+\tilde{q}+1)} \qquad (53)\end{aligned}$$

and

$$\begin{aligned}\mu_i(e^{\alpha r_i}) &= \int_{-\tilde{\varphi}_i}^{+\infty} e^{\alpha r_i}\frac{2^{4\tilde{\varphi}_i}}{\Gamma(4\tilde{\varphi}_i)}(r_i+\tilde{\varphi}_i)^{4\tilde{\varphi}_i-1}e^{-2(r_i+\tilde{\varphi}_i)}dr_i\\ &= \frac{2^{4\tilde{\varphi}_i}}{\Gamma(4\tilde{\varphi}_i)}e^{-\alpha\tilde{\varphi}_i}\int_0^{+\infty}y^{4\tilde{\varphi}_i-1}e^{-(2-\alpha)y}dy\\ &= \frac{2^{4\tilde{\varphi}_i}}{\Gamma(4\tilde{\varphi}_i)}e^{-\alpha\tilde{\varphi}_i}\frac{1}{(2-\alpha)^{4\tilde{\varphi}_i}}\int_0^{+\infty}y^{4\tilde{\varphi}_i-1}e^{-y}dy\\ &= \Big(1-\frac{\alpha}{2}\Big)^{-4\tilde{\varphi}_i}e^{-\alpha\tilde{\varphi}_i}. \qquad (54)\end{aligned}$$

Since a sufficiently smooth function can be approximated by polynomial functions, the results of Theorem 4.1 also hold when $g_i(\cdot):\mathbb{R}\to\mathbb{R}_+$ is a sufficiently smooth function and satisfies (A1)-(A3). However, for more general function $g_i(\cdot):\mathbb{R}\to\mathbb{R}_+$, it will be our future work how to establish the weak convergence.

REFERENCES

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter *Molecular Biology of the Cell*, 4^{nd} edition, New York: Garland, 2002.

[2] G. Balázsi, A. van Oudenaarden and J. J. Collins, Cellular decision making and biological noise: From microbes to mammals, *Cell*, **144** (2011), 910–925.

[3] M. Bodnar, General model of a cascade of reactions with times: Global stability analysis, *J. Differential Eqs.*, **259** (2015), 777–795.

[4] D. Bratsun, D. Volfson, L. S. Tsimring and J. Hasty, Delay-induced stochastic oscillations in gene regulation, *Proc. National Academy of Sci. USA*, **102** (2005), 14593–14598.

[5] C. Clayton and M. Shapira, Post-transcriptional regulation of gene expression in trypanosomes and leishmanias, *Molecular and Biochemical Parasitology*, **156** (2007), 93–101.

[6] J. C. Cox, J. E. Ingersoll and S. A. Ross, A theory of the term structure of interest rates, *Econometrica*, **53** (1985), 385–407.

[7] D. Denault, J. Loros and J. C. Dunlap, WC-2 mediates WC-1-FRQ interaction within the PAS protein-linked circadian feedback loop of Neurospora, *EMBO J.*, **20** (2001), 109–117.

[8] T. C. Gard, *Introduction to Stochastic Differential Equation*, Marcel Dekker Inc, New York, 1988.

[9] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *Journal of computational physics*, **22** (1976), 403–434.

[10] D. T. Gillespie, The chemical Langevin equation, *The Journal of Chemical Physics*, **113** (2000), 297–306.

[11] R. Z. Khasminskii, On stochastic processes defined by differential equations with a small parameter, (Russian) *Teor. Verojatnost. i Primenen*, **11** (1966), 240–259.

[12] T. Kurtz, Semigroups of conditioned shifts and approximation of Markov processes, *Ann. Probab.*, **3** (1975), 618–642.
[13] T. Kurtz, *Approximation of Population Processes,* CBMS-NSF Regional Conference Series in Applied Mathematics, 36, SIAM, Philadelphia, Pa., 1981.
[14] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.
[15] H. J. Kushner, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, MA, 1990.
[16] X. Mao, *Stochastic Differential Equations and Applications*, 2^{nd} edition, Horwood, Chichester, 2008.
[17] B. Mélykúti, K. Burrage and K. C. Zygalakis, Fast stochastic simulation of biochemical reaction systems by alternative formulations of the chemical Langevin equation, *The Journal of chemical physics*, **132** (2010), 164109.
[18] J. Miekisz, J. Poleszczuk and M. Bodnar, Stochastic models of gene expression with delayed degradation, *Bull. Math. Bio.*, **73** (2011), 2231–2247.
[19] K. M. Ramachandran, A singularly perturbed stochastic delay system with small parameter, *Stochastic Anal. Appl.*, **11** (1993), 209–230.
[20] K. M. Ramachandran, Stability of stochastic delay differential equation with a small parameter, *Stochastic Anal. Appl.*, **26** (2008), 710–723.
[21] M. Thattai and A. van Oudenaarden, Intrinsic noise in gene regulatory networks, *Proc. Natl. Acad. Sci. USA*, **98** (2001), 8614–8619.
[22] M. Turcotte, J. Garcia-Ojalvo and G. M. Süel, A genetic timer through noise-induced stabilization of an unstable state, *Proc. Natl. Acad. Sci. USA*, **105** (2008), 15732–15737.
[23] F. Wu, T. Tian, J. B. Rawlings and G. Yin, Approximate method for stochastic chemical kinetics with two-time scales by chemical Langevin equations, *J. Chemical Phy.*, **144** (2016), 174112.
[24] G. Yin and K. M. Ramachandran, A differential delay equation with wideband noise perturbations, *Stochastic Process. Appl.*, **35** (1990), 231–249.
[25] G. Yin and H. Q. Zhang, Singularly perturbed Markov chains: Limit results and applications, *Ann. Appl. Probab.*, **17** (2007), 207–229.
[26] G. Yin and Q. Zhang, *Continuous-time Markov Chains and Applications: A Two-time-scale Approach*, 2^{nd} edition, Springer, Now York, 2013.

Received July 2018; revised November 2018.

E-mail address: li_yun@hust.edu.cn
E-mail address: wufuke@hust.edu.cn
E-mail address: gyin@math.wayne.edu