CONVERGENCE OF NONLINEAR FILTERINGS FOR MULTISCALE SYSTEMS WITH CORRELATED LÉVY NOISES*

HUIJIE QIAO

School of Mathematics, Southeast University
Nanjing, Jiangsu 211189, China

Department of Mathematics, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

hjqiaoge@seu.edu.cn

Abstract. In the paper, we consider nonlinear filtering problems of multiscale systems in two cases—correlated sensor Lévy noises and correlated Lévy noises. First of all, we prove that the slow part of the origin system converges to the homogenized system in the uniform mean square sense. And then based on the convergence result, in the case of correlated sensor Lévy noises, the nonlinear filtering of the slow part is shown to approximate that of the homogenized system in $L^1$ sense. However, in the case of correlated Lévy noises, we prove that the nonlinear filtering of the slow part converges weakly to that of the homogenized system.

1. INTRODUCTION

Nowadays, more and more high dimensional and complex mathematical models are used in engineering and science (c.f. [1, 3, 8, 11, 12, 17, 19, 20]). For example, in some climate models, it is common to simulate the dynamics of the atmosphere and ocean on varying spatial grids with distinct time scale separations.

Simultaneously, controlling, estimating and forecasting these models become more and more interesting (c.f. [1, 3, 8, 11, 17, 20]). However, different time scales make much trouble. Therefore, how to treat these scales is the first important task. A kind of usual methods is to reduce the dimension of these high dimensional mathematical models and study low dimensional ones with the similar dynamical structure. Thus, by estimating the low dimensional ones, we can master the origin high dimensional ones. And nonlinear filtering problems are just right to estimate unobservable and complicated phenomena by observing some simple objects. So, by solving some suitable filtering problems, high dimensional and complex models can be controlled and estimated.

In the paper, we are mainly interested in the nonlinear filtering problems of the following two multiscale systems. For a fixed time $T > 0$, given a completed filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})$. Consider the following slow-fast system on $\mathbb{R}^n \times \mathbb{R}^m$ and the

\*This work was partly supported by NSF of China (No. 11001051, 11371352) and China Scholarship Council under Grant No. 201906095034.
observation process on $\mathbb{R}^d$, for $0 \leq t \leq T$,
\begin{align}
\begin{cases}
    dX_t^e = b_1(X_t^e, Z_t^e)dt + \sigma_1(X_t^e)dB_t + \int_{U_1} f_1(X_{t-}^e, u)\tilde{N}_{p_1}(dt, du), \\
    X_0^e = x_0, \\
    dZ_t^e = b_2(X_t^e, Z_t^e)dt + \frac{1}{\sqrt{\epsilon}}\sigma_2(X_t^e, Z_t^e)dB_t + \int_{U_2} f_2(X_{t-}^e, Z_{t-}^e, u)\tilde{N}_{p_2}(dt, du), \\
    Z_0^e = z_0, \\
    dY_t^e = h(X_t^e)dt + \sigma_3 dV_t + \sigma_4 dB_t, \\
    Y_0^e = 0,
\end{cases}
\end{align}
(1)

where $V, W, B$ are $l$-dimensional, $m$-dimensional and $j$-dimensional standard Brownian motions, respectively, and $p_1, p_2$ are two stationary Poisson point processes of the class (quasi left-continuous) defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})$ with values in $\mathbb{U}$ and the characteristic measure $\nu_1, \nu_2$, respectively. Here $\nu_1, \nu_2$ are two $\sigma$-finite measures defined on a measurable space $(\mathbb{U}, \mathcal{U})$. Fix $U_1, U_2 \in \mathcal{U}$ with $\nu_1(U \setminus U_1) < \infty$ and $\nu_2(U \setminus U_2) < \infty$. Let $N_{p_1}((0,t], du)$ be the counting measure of $p_1(t)$, a Poisson random measure and then $EN_{p_1}((0,t], A) = t \nu_1(A)$ for $A \in \mathcal{U}$. Denote

\[ \tilde{N}_{p_1}((0,t], du) := N_{p_1}((0,t], du) - t \nu_1(du), \quad A \in \mathcal{U}|_{U_1}, \]

the compensated measure of $N_{p_1}((0,t], du)$. By the same way, we could define $N_{p_2}((0,t], du)$, $\tilde{N}_{p_2}((0,t], du)$. And $N_{p_2}((0,t], du)$ is another Poisson random measure on $(\mathbb{U}, \mathcal{U})$ such that $EN_{p_2}((0,t], A) = \frac{1}{t} \nu_2(A)$ for $A \in \mathcal{U}$. Moreover, $V_t, W_t, B_t, N_{p_1}, N_{p_2}$ are mutually independent. The mappings $b_1 : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$, $b_2 : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m$, $\sigma_1 : \mathbb{R}^n \mapsto \mathbb{R}^{n \times l}$, $\sigma_2 : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^{m \times m}$, $f_1 : \mathbb{R}^n \times U_1 \mapsto \mathbb{R}^n$, $f_2 : \mathbb{R}^n \times \mathbb{R}^m \times U_2 \mapsto \mathbb{R}^m$, and $h : \mathbb{R}^n \mapsto \mathbb{R}^d$ are all Borel measurable. The matrices $\sigma_3, \sigma_4$ are $d \times l, d \times j$, respectively. The system (1) is usually called as a correlated noise model.

We also consider the following slow-fast system on $\mathbb{R}^n \times \mathbb{R}^m$ and the observation process on $\mathbb{R}^d$, for $0 \leq t \leq T$, $l = d$,
\begin{align}
\begin{cases}
    d\tilde{X}_t^e = \tilde{b}_1(\tilde{X}_t^e, \tilde{Z}_t^e)dt + \tilde{\sigma}_0(\tilde{X}_t^e)dB_t + \tilde{\sigma}_1(\tilde{X}_t^e)dB_t + \int_{U_1} \tilde{f}_1(\tilde{X}_{t-}^e, u)\tilde{N}_{p_1}(dt, du), \\
    \tilde{X}_0^e = \tilde{x}_0, \\
    d\tilde{Z}_t^e = \frac{1}{\sqrt{\epsilon}}\tilde{b}_2(\tilde{X}_t^e, \tilde{Z}_t^e)dt + \frac{1}{\sqrt{\epsilon}}\tilde{\sigma}_2(\tilde{X}_t^e, \tilde{Z}_t^e)dB_t + \int_{U_2} \tilde{f}_2(\tilde{X}_{t-}^e, \tilde{Z}_{t-}^e, u)\tilde{N}_{p_2}(dt, du), \\
    \tilde{Z}_0^e = \tilde{z}_0, \\
    d\tilde{Y}_t^e = h(\tilde{X}_t^e)dt + \tilde{\sigma}_3 dV_t + \tilde{\sigma}_4 dB_t, \\
    \tilde{Y}_0^e = 0,
\end{cases}
\end{align}
(2)

where $N_{\lambda}(dt, du)$ is a random measure with a predictable compensator $\lambda(t, \tilde{X}_t^e, u)dt\nu_3(du)$. Here the function $\lambda : [0,T] \times \mathbb{R}^n \times \mathbb{U} \mapsto (0,1)$ is Borel measurable and $\nu_3$ is a $\sigma$-finite measure defined on $\mathbb{U}$ with $\nu_3(U \setminus U_3) < \infty$ and $\int_{U_3} \|u\| \nu_3(du) < \infty$ for a fixed $U_3 \in \mathcal{U}$. Concretely speaking, set

\[ \tilde{N}_{\lambda}((0,t], A) := N_{\lambda}((0,t], A) - \int_{0}^{t} \int_{A} \lambda(s, X_s^e, u)ds\nu_3(du), \quad t \in [0,T], A \in \mathcal{U}|_{U_3}, \]

and then $\tilde{N}_{\lambda}((0,t], du)$ is the compensated martingale measure of $N_{\lambda}((0,t], du)$. Moreover, $V_t, W_t, B_t, N_{p_1}, N_{p_2}, \tilde{N}_{p_2}, N_{\lambda}$ are mutually independent. These mappings $\tilde{b}_1 : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$, $\tilde{b}_2 : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m$, $\tilde{\sigma}_0 : \mathbb{R}^n \mapsto \mathbb{R}^{n \times l}$, $\tilde{\sigma}_1 : \mathbb{R}^n \mapsto \mathbb{R}^{n \times d}$, $\tilde{\sigma}_2 : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^{m \times m}$, $\tilde{f}_1 : \mathbb{R}^n \times U_1 \mapsto \mathbb{R}^n$, $\tilde{f}_2 : \mathbb{R}^n \times \mathbb{R}^m \times U_2 \mapsto \mathbb{R}^m$, $\tilde{h} : \mathbb{R}^n \mapsto \mathbb{R}^d$, $\tilde{f}_3 : [0,T] \times U_3 \mapsto \mathbb{R}^d$ and $\tilde{g}_3 : [0,T] \times (U \setminus U_3) \mapsto \mathbb{R}^d$ are all Borel measurable. As usual, the system (2) is called as a correlated noise model.
Note that in the systems (1) (2), the unobservable processes and the observable ones have correlated parts. The type of the multiscale correlated filtering problems usually stems from atmospheric and climatology problems. For example, coupled atmosphere-ocean models provide a multiscale model with fast atmospheric and slow ocean dynamics. In the case of climate prediction, the ocean memory, due to its heat capacity, holds important information. Hence, the improved estimate of the ocean state, which is often the slow component, is of greater interest.

In the paper, we firstly prove that the slow part of a fast-slow system converges to the homogenized system in the uniform mean square sense. And then based on the convergence result, for the system (1), the nonlinear filtering of the slow part is shown to approximate that of the homogenized system in $L^1$ sense. But for the system (2), we prove that the nonlinear filtering of the slow part converges weakly to that of the homogenized system.

It is worthwhile to mentioning our method and results. Firstly, for the system (1), since the driving processes of the fast-slow system are correlated with that of the observation, we can not obtain the Zakai equation of the homogenized system (cf. [15]). Thus, those methods by means of the Zakai equation do not work (cf. [6, 7, 12]). Therefore, we make use of the exponential martingale to prove the convergence for the filtering of the slow part to that of the homogenized system. However, for the system (2), we can deduce the Zakai equations of the slow system and the homogenized system, and then show their filtering convergence. Secondly, here we prove the uniform mean square convergence stronger than weak convergence in [1, 8, 12] and convergence in probability in [11]. Thirdly, when $f_1 = f_2 = 0$ in (1) and $\tilde{f}_1 = \tilde{f}_2 = \tilde{f}_3 = \tilde{g}_3 = 0$ in (2), two types of multiscale correlated filtering problems have appeared in [1] and [9], respectively. In [1], when the slow part of the origin system converges to the homogenized system in distribution, Beeson and Namachchivaya only stated that the filtering of the slow part also converges to the filtering of the homogenized system in $L_p$ sense. Regretfully, they didn’t prove the result. Here we show the convergence in $L^1$ sense when $f_1 \neq 0, f_2 \neq 0$. Therefore, our result generalizes the result in a manner. In [9], Lucic and Heunis proved that the slow part converges weakly to the homogenized system, and the filtering of the slow part also converges weakly to that of the homogenized system. Here, we establish that the slow part converges to the homogenized system in the uniform mean square sense, and the filtering of the slow part converges to that of the homogenized system in $L^1$ sense when $f_1 \neq 0, f_2 \neq 0$. Thus, our result is better. Finally, in [12], we considered the nonlinear filtering problem of the system (2) with $\tilde{\sigma}_1 = 0$. Here, we permit $\tilde{\sigma}_1 \neq 0$. Therefore, our result is more general in some sense.

The paper is arranged as follows. In next section, we consider strong convergence for the fast-slow system. In Section 3, we define nonlinear filtering problem and then show that the filtering of the slow part for the system (1) converges to that of the homogenize system. In Section 4, the filtering of the slow part for the system (2) is proved to converge weakly to that of the homogenize system. We summarize all the results in Section 5.

The following convention will be used throughout the paper: $C$ with or without indices will denote different positive constants whose values may change from one place to another.

2. Convergence of some processes

In the section, we study strong convergence for the fast-slow system (1) when $\varepsilon \to 0$. 

3
2.1. **A slow-fast system.** In the subsection, we introduce slow-fast systems and the existence and uniqueness of their solutions.

Let us consider the system (I). First of all, we give out our assumptions and state some related results.

**Assumption 1.**

\( (H^1_{b_1, \sigma_1, f_1}) \) For \( x_1, x_2 \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^m \), there exist \( L_{b_1}, L_{\sigma_1}, L_{f_1} > 0 \) such that

\[
\begin{align*}
|b_1(x_1, z_1) - b_1(x_2, z_2)|^2 &\leq L_{b_1}(|x_1 - x_2|^2 + |z_1 - z_2|^2), \\
\|\sigma_1(x_1) - \sigma_1(x_2)\|^2 &\leq L_{\sigma_1}|x_1 - x_2|^2, \\
\int_{U_1} |f_1(x_1, u) - f_1(x_2, u)|^2 \nu_1(du) &\leq L_{f_1}|x_1 - x_2|^2,
\end{align*}
\]

where \(|\cdot|\) and \(\|\cdot\|\) denote the length of a vector and the Hilbert-Schmidt norm of a matrix, respectively.

\( (H^2_{b_1, \sigma_1, f_1}) \) For \( x \in \mathbb{R}^n, z \in \mathbb{R}^m \), there exists a \( L_{b_1, \sigma_1, f_1} > 0 \) such that

\[
|b_1(x, z)|^2 + \|\sigma_1(x)\|^2 + \int_{U_1} |f_1(x, u)|^2 \nu_1(du) \leq L_{b_1, \sigma_1, f_1}.
\]

\( (H^1_{b_2}) \) (i) \( b_2 \) is bi-continuous in \((x, z)\),

(ii) There exist \( L_{b_2} \geq 0, \bar{L}_{b_2} > 0 \) such that

\[
\begin{align*}
|b_2(x_1, z) - b_2(x_2, z)| &\leq L_{b_2}|x_1 - x_2|, & x_1, x_2 \in \mathbb{R}^n, z \in \mathbb{R}^m, \\
\langle z_1 - z_2, b_2(x_1, z_1) - b_2(x_2, z_2) \rangle &\leq -\bar{L}_{b_2}|z_1 - z_2|^2, & x \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^m,
\end{align*}
\]

(iii) For \( x \in \mathbb{R}^n, z \in \mathbb{R}^m \), there exists a constant \( \bar{L}_{b_2} > 0 \) such that

\[
|b_2(x, z)| \leq \bar{L}_{b_2}(1 + |x| + |z|).
\]

\( (H^1_{\sigma_2}) \) For \( x_1, x_2 \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^m \), there exists a constant \( L_{\sigma_2} > 0 \) such that

\[
\|\sigma_2(x_1, z_1) - \sigma_2(x_2, z_2)\| \leq L_{\sigma_2}(|x_1 - x_2| + |z_1 - z_2|).
\]

\( (H^1_{f_2}) \) There exists a positive function \( L(u) \) satisfying

\[
\sup_{u \in U_2} L(u) \leq \gamma < 1 \quad \text{and} \quad \int_{U_2} L(u)^2 \nu_2(du) < +\infty,
\]

such that for any \( x_1, x_2 \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^m \) and \( u \in U_2 \)

\[
|f_2(x_1, z_1, u) - f_2(x_2, z_2, u)| \leq L(u)(|x_1 - x_2| + |z_1 - z_2|),
\]

and

\[
|f_2(0, 0, u)| \leq L(u).
\]

Under **Assumption 1.**, by Theorem 1.2 in [14], we know that the system (I) has a unique strong solution denoted by \((X_t^x, Z_t^x)\).
2.2. The fast equation. In the subsection, we mainly study the second part of the system (1).

First, take any \( x \in \mathbb{R}^n \) and fix it. And consider the following SDE in \( \mathbb{R}^m \):
\[
\begin{aligned}
  \text{d}Z_t^x &= b_2(x, Z_t^x)\text{d}t + \sigma_2(x, Z_t^x)\text{d}W_t + \int_{U_2} f_2(x, Z_t^x, u)\tilde{N}_{p_2}(\text{d}t, \text{d}u),
  
  Z_0^x &= z_0, \quad t \geq 0.
\end{aligned}
\]
Under the assumption (H\(_{b_2}^1\)) (H\(_{b_2}^3\)) (H\(_{\sigma_2}^1\)), the above equation has a unique solution \( Z_t^x \).

In addition, it is a Markov process and its transition probability is denoted by \( p(x; z_0, t, A) \) for \( t \geq 0 \) and \( A \in \mathcal{B}(\mathbb{R}^m) \). We assume:

Assumption 2.
(H\(_{\sigma_2}^2\)) There exists a function \( \alpha_1(x) > 0 \) such that
\[
\langle \sigma_2(x, z)h, h \rangle \geq \sqrt{\alpha_1(x)}|h|^2, \quad z, h \in \mathbb{R}^m,
\]
and
\[
\|\sigma_{\alpha_1}(x, z_1) - \sigma_{\alpha_1}(x, z_2)\|^2 \leq L_{\alpha_1}|z_1 - z_2|^2, \quad z_1, z_2 \in \mathbb{R}^m,
\]
where \( \sigma_{\alpha_1}(x, z) \) is the unique symmetric nonnegative definite matrix such that \( \sigma_{\alpha_1}(x, z)\sigma_{\alpha_1}(x, \cdot) = \sigma_2(x, z)\sigma_2^T(x, \cdot) - \alpha_1(x)I \) for the unit matrix \( I \).

(H\(_{b_2, \alpha_1, f_2}^1\)) There exist a \( r > 2 \) and two functions \( \alpha_2(x) > 0, \alpha_3(x) \geq 0 \) such that for all \( z \in \mathbb{R}^m \)
\[
2\langle z, b_2(x, z) \rangle + \|\sigma_2(x, z)\|^2 + \int_{U_2} |f_2(x, z, u)|^2\nu_2(\text{d}u) \leq -\alpha_2(x)|z|^r + \alpha_3(x).
\]

(H\(_{b_2, \alpha_2, f_2}^2\))
\[
M := 2\tilde{L}_{b_2} - L_{b_2} - 2L_{\alpha_2}^2 - 2\int_{U_2} L^2(\text{d}u)\nu_2(\text{d}u) > 0.
\]

Under the assumptions (H\(_{b_2}^1\)) (H\(_{b_2}^3\)) (H\(_{b_2}^1\)) (H\(_{\sigma_2}^1\)) (H\(_{b_2, \alpha_2, f_2}^1\)), by Theorem 1.3 in [13] it holds that there exists a unique invariant probability measure \( \bar{p}(x, \cdot) \) for \( Z_t^x \) such that
\[
\|p(x; z_0, t, \cdot) - \bar{p}(x, \cdot)\|_{\text{var}} \leq Ce^{-\alpha t}, \quad t > 0,
\]
where \( \| \cdot \|_{\text{var}} \) is the total variance norm and \( C, \alpha > 0 \) are two constants independent of \( z_0, t \).

2.3. The homogenized equation. In the subsection, we construct a homogenized equation and study the relationship between the origin equation and the homogenized one.

Next, set
\[
\bar{b}_1(x) := \int_{\mathbb{R}^m} b_1(x, z)\bar{p}(x, \text{d}z),
\]
and by [12] Lemma 3.1 we know that \( \bar{b}_1 \) is Lipschitz continuous. So, we construct a SDE on the probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})\) as follows:
\[
\begin{aligned}
  \text{d}X^0_t &= \bar{b}_1(X^0_t)\text{d}t + \sigma_1(X^0_t)\text{d}V_t + \int_{U_1} f_1(X^0_t, u)\tilde{N}_{p_1}(\text{d}t, \text{d}u),
  
  X^0_0 &= x_0, \quad 0 \leq t \leq T.
\end{aligned}
\]
Based on the assumptions (H\(_{b_1, \alpha_1, f_1}^1\)) (H\(_{b_1, \alpha_1, f_1}^2\)), it holds that Eq. (1) has a unique strong solution denoted as \( X^\varepsilon \). And then we study the relation between \( X^\varepsilon \) and \( X^0 \). To do this,
we realize a partition of \([0, T]\) into intervals of size \(\delta \varepsilon > 0\), and introduce an auxiliary processes:

\[
d\hat{Z}_t^\varepsilon = \frac{1}{\varepsilon} b_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_t^\varepsilon)dt + \frac{1}{\sqrt{\varepsilon}} \sigma_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_t^\varepsilon)dW_t + \int_{U_2} f_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_{t-}^\varepsilon, u)\tilde{N}_{p_2}^\varepsilon(dt, du),
\]

\[
t \in [k\delta \varepsilon, (k+1)\delta \varepsilon],
\]

\[
\hat{Z}_{k\delta \varepsilon}^\varepsilon = Z_{k\delta \varepsilon}^\varepsilon,
\]

for \(k = 0, \cdots, \lceil \frac{T}{\delta \varepsilon} \rceil\), where \(\lceil \frac{T}{\delta \varepsilon} \rceil\) denotes the integer part of \(\frac{T}{\delta \varepsilon}\). Moreover, we mention the fact that \(\lceil \frac{T}{\delta \varepsilon} \rceil = k\) for \(t \in [k\delta \varepsilon, (k+1)\delta \varepsilon)\). The following lemma gives the relationship between \(Z^\varepsilon\) and \(\hat{Z}^\varepsilon\).

**Lemma 2.1.** Under Assumption 1.-2., it holds that

\[
\sup_{0 \leq s \leq T} \mathbb{E}|Z_s^\varepsilon - \hat{Z}_s^\varepsilon|^2 \leq \frac{L_{b_2} + 2L_{\sigma_2}^2 + 2\int_{U_2} L^2(u)\nu_2(du)}{\varepsilon} (\delta \varepsilon + 1)L_{b_1, \sigma_1, f_1} \delta \varepsilon^2.
\]

**Proof.** By the equations (1)-(5), it holds that for \(s \in [k\delta \varepsilon, (k+1)\delta \varepsilon)\)

\[
Z_s^\varepsilon - \hat{Z}_s^\varepsilon = \frac{1}{\varepsilon} \int_{k\delta \varepsilon}^s \left( b_2(X_r^\varepsilon, Z_r^\varepsilon) - b_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_r^\varepsilon) \right)dr
\]

\[
+ \frac{1}{\sqrt{\varepsilon}} \int_{k\delta \varepsilon}^s \left( \sigma_2(X_r^\varepsilon, Z_r^\varepsilon) - \sigma_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_r^\varepsilon) \right)dW_r
\]

\[
+ \int_{k\delta \varepsilon}^s \int_{U_2} \left( f_2(X_r^\varepsilon, Z_r^\varepsilon, u) - f_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_r^\varepsilon, u) \right)\tilde{N}_{p_2}^\varepsilon(dr, du).
\]

Applying the Itô formula to \(Z_s^\varepsilon - \hat{Z}_s^\varepsilon\) and taking the expectation on two sides, we have that

\[
\mathbb{E}|Z_s^\varepsilon - \hat{Z}_s^\varepsilon|^2 = \frac{2}{\varepsilon} \mathbb{E} \int_{k\delta \varepsilon}^s \langle Z_r^\varepsilon - \hat{Z}_r^\varepsilon, b_2(X_r^\varepsilon, Z_r^\varepsilon) - b_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_r^\varepsilon) \rangle dr
\]

\[
+ \frac{1}{\varepsilon} \mathbb{E} \int_{k\delta \varepsilon}^s \|\sigma_2(X_r^\varepsilon, Z_r^\varepsilon) - \sigma_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_r^\varepsilon)\|^2 dr
\]

\[
+ \frac{1}{\varepsilon} \mathbb{E} \int_{k\delta \varepsilon}^s \int_{U_2} |f_2(X_r^\varepsilon, Z_r^\varepsilon, u) - f_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_r^\varepsilon, u)|^2\nu_2(du)dr
\]

\[
\leq \frac{2}{\varepsilon} \mathbb{E} \int_{k\delta \varepsilon}^s \langle Z_r^\varepsilon - \hat{Z}_r^\varepsilon, b_2(X_r^\varepsilon, Z_r^\varepsilon) - b_2(X_{k\delta \varepsilon}^\varepsilon, Z_{k\delta \varepsilon}^\varepsilon) \rangle dr
\]

\[
+ \frac{2}{\varepsilon} \mathbb{E} \int_{k\delta \varepsilon}^s \langle Z_r^\varepsilon - \hat{Z}_r^\varepsilon, b_2(X_{k\delta \varepsilon}^\varepsilon, Z_{k\delta \varepsilon}^\varepsilon) - b_2(X_{k\delta \varepsilon}^\varepsilon, \hat{Z}_{k\delta \varepsilon}^\varepsilon) \rangle dr
\]

\[
+ \frac{2L_{\sigma_2}^2}{\varepsilon} \int_{k\delta \varepsilon}^s \left( |X_r^\varepsilon - X_{k\delta \varepsilon}^\varepsilon|^2 + |Z_r^\varepsilon - \hat{Z}_r^\varepsilon|^2 \right)dr
\]

\[
+ \frac{2}{\varepsilon} \int_{U_2} L^2(u)\nu_2(du)\mathbb{E} \int_{k\delta \varepsilon}^s \left( |X_r^\varepsilon - X_{k\delta \varepsilon}^\varepsilon|^2 + |Z_r^\varepsilon - \hat{Z}_r^\varepsilon|^2 \right)dr
\]

\[
\leq \frac{2}{\varepsilon} \mathbb{E} \int_{k\delta \varepsilon}^s \left| Z_r^\varepsilon - \hat{Z}_r^\varepsilon \right|^2 dr
\]

\[
- \frac{2L_{b_2}}{\varepsilon} \int_{k\delta \varepsilon}^s |Z_r^\varepsilon - \hat{Z}_r^\varepsilon|^2 dr
\]
Thus, by \((H_0)\), \((H_1)\), \((H_2)\) it holds that

\[
\mathbb{E}|Z_\varepsilon^r - \hat{Z}_\varepsilon^r|^2 \leq \frac{L_{b_2} + 2L_{\sigma_2}^2 + 2 \int_{\mathcal{U}_2} L^2(u) \nu_2(du) \mathbb{E} \int_{k\delta_\varepsilon}^{k\delta_\varepsilon} |X_\varepsilon^r - X_{k\delta_\varepsilon}^r|^2 dr}{\varepsilon}.
\]

Thus, by \((H_{b_2,\sigma_2,f_2})\) it holds that

\[
\mathbb{E}|Z_\varepsilon^r - \hat{Z}_\varepsilon^r|^2 \leq \frac{L_{b_2} + 2L_{\sigma_2}^2 + 2 \int_{\mathcal{U}_2} L^2(u) \nu_2(du) \mathbb{E} \int_{k\delta_\varepsilon}^{k\delta_\varepsilon} |X_\varepsilon^r - X_{k\delta_\varepsilon}^r|^2 dr}{\varepsilon}.
\]  \hfill (7)

To obtain \((6)\), we only need to estimate \(\mathbb{E}|X_\varepsilon^r - X_{k\delta_\varepsilon}^r|^2\) for \(r \in [k\delta_\varepsilon, (k + 1)\delta_\varepsilon)\). Note that

\[
X_\varepsilon^r - X_{k\delta_\varepsilon}^r = \int_{k\delta_\varepsilon}^{r} b_1(X_\varepsilon^r, Z_\varepsilon^r) dv + \int_{k\delta_\varepsilon}^{r} \sigma_1(X_\varepsilon^r) dV_v + \int_{k\delta_\varepsilon}^{r} \int_{\mathcal{U}_1} f_1(X_\varepsilon^r, u) \tilde{N}_{p_1}(dv, du).
\]

So, by the Hölder inequality and \((H_{b_1,\sigma_1,f_1})\) we obtain that

\[
\mathbb{E}|X_\varepsilon^r - X_{k\delta_\varepsilon}^r|^2 \leq 3 \left| \int_{k\delta_\varepsilon}^{r} b_1(X_\varepsilon^r, Z_\varepsilon^r) dv \right|^2 + 3 \left| \int_{k\delta_\varepsilon}^{r} \sigma_1(X_\varepsilon^r) dV_v \right|^2 + 3 \left| \int_{k\delta_\varepsilon}^{r} \int_{\mathcal{U}_1} f_1(X_\varepsilon^r, u) \tilde{N}_{p_1}(dv, du) \right|^2
\]

\[
\leq 3(r - k\delta_\varepsilon) \mathbb{E} \int_{k\delta_\varepsilon}^{r} |b_1(X_\varepsilon^r, Z_\varepsilon^r)|^2 dv + 3 \mathbb{E} \int_{k\delta_\varepsilon}^{r} \|\sigma_1(X_\varepsilon^r)\|^2 dv + 3 \int_{k\delta_\varepsilon}^{r} \int_{\mathcal{U}_1} |f_1(X_\varepsilon^r, u)|^2 \nu_1(du) dv
\]

\[
\leq 3(\delta_\varepsilon + 1) L_{b_1,\sigma_1,f_1,\delta_\varepsilon}.
\]  \hfill (8)

By inserting \((8)\) in \((7)\), it holds that

\[
\mathbb{E}|Z_\varepsilon^r - \hat{Z}_\varepsilon^r|^2 \leq \frac{L_{b_2} + 2L_{\sigma_2}^2 + 2 \int_{\mathcal{U}_2} L^2(u) \nu_2(du) \mathbb{E} \int_{k\delta_\varepsilon}^{k\delta_\varepsilon} |X_\varepsilon^r - X_{k\delta_\varepsilon}^r|^2 dr}{\varepsilon}.
\]

This is just right \((6)\). Thus, the proof is complete. \hfill \square

Next, we apply \((6)\) to estimate \(|X_\varepsilon^r - X_\varepsilon^0|^2\). The main result in the section is the following theorem.
Theorem 2.2. Suppose that Assumption 1.2. hold. Then there exists a constant $C > 0$ independent of $\varepsilon, \delta \varepsilon$ such that

$$
\mathbb{E}\left( \sup_{0 \leq t \leq T} |X^\varepsilon_t - X^0_t|^2 \right) \leq \left( C \frac{\varepsilon}{\delta \varepsilon} + C(\delta \varepsilon + 1)\delta \varepsilon + C(\delta \varepsilon + 1) \frac{\delta^2}{\varepsilon} \right) e^{CT}.
$$

(9)

Proof. By the equations (1)-(4), we know that

$$
X^\varepsilon_t - X^0_t = \int_0^t \left( b_1(X^\varepsilon_s, Z^\varepsilon_s) - \bar{b}_1(X^0_s) \right) ds + \int_0^t \left( \sigma_1(X^\varepsilon_s) - \sigma_1(X^0_s) \right) dV_s
$$

$$
+ \int_0^t \int_{U_1} \left( f_1(X^\varepsilon_{s-}, u) - f_1(X^0_{s-}, u) \right) \tilde{N}_{p_1}(ds, du), \quad t \in [0, T].
$$

And then by the Burkholder-Davis-Gundy inequality and the Hölder inequality, it holds that

$$
\mathbb{E}\left( \sup_{0 \leq t \leq T} |X^\varepsilon_t - X^0_t|^2 \right) \leq 3 \mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_0^t \left( b_1(X^\varepsilon_s, Z^\varepsilon_s) - \bar{b}_1(X^0_s) \right) ds \right|^2 \right)
$$

$$
+ 3 \mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_0^t \left( \sigma_1(X^\varepsilon_s) - \sigma_1(X^0_s) \right) dV_s \right|^2 \right)
$$

$$
+ 3 \mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_0^t \int_{U_1} \left( f_1(X^\varepsilon_{s-}, u) - f_1(X^0_{s-}, u) \right) \tilde{N}_{p_1}(ds, du) \right|^2 \right)
$$

$$
\leq 12 \mathbb{E}\left( \left| \int_0^t \left( b_1(X^\varepsilon_s, Z^\varepsilon_s) - b_1(X^\varepsilon_s, \hat{Z}^\varepsilon_s) \right) ds \right|^2 \right)
$$

$$
+ 12 \mathbb{E}\left( \left| \int_0^t \left( \hat{b}_1(X^\varepsilon_s, \hat{Z}^\varepsilon_s) - \bar{b}_1(X^\varepsilon_s) \right) ds \right|^2 \right)
$$

$$
+ 12 \mathbb{E}\left( \left| \int_0^t \left( \bar{b}_1(X^\varepsilon_s) - \bar{b}_1(X^0_s) \right) ds \right|^2 \right)
$$

$$
+ 12 \mathbb{E}\left( \left| \int_0^T \left( \sigma_1(X^\varepsilon_s) - \sigma_1(X^0_s) \right) ds \right|^2 \right)
$$

$$
+ 12 \mathbb{E}\left( \left| \int_0^T \int_{U_1} \left( f_1(X^\varepsilon_{s-}, u) - f_1(X^0_{s-}, u) \right) \tilde{N}_{p_1}(ds, du) \right|^2 \right)
$$

$$
\leq 12TLb_1 \int_0^T \left( \mathbb{E}|X^\varepsilon_s - X^0_{s-}|^2 + \mathbb{E}|Z^\varepsilon_s - \hat{Z}^\varepsilon_s|^2 \right) ds
$$

$$
+ 12 \mathbb{E}\left( \left| \int_0^t \left( b_1(X^\varepsilon_s, \hat{Z}^\varepsilon_s) - \bar{b}_1(X^\varepsilon_s) \right) ds \right|^2 \right)
$$

$$
+ 12TC \int_0^T \mathbb{E}|X^\varepsilon_{s-} - X^\varepsilon_s|^2 ds
$$

$$
+ \left( 12TC + 12L_{\sigma_1} + 12L_{f_1} \right) \int_0^T \mathbb{E}|X^\varepsilon_s - X^0_s|^2 ds
$$

8
\[
\begin{align*}
\leq & \ 12\mathbb{E}\left(\sup_{0 \leq t \leq T} \left| \int_0^t \left( b_1(X_{k_0}^\varepsilon, \hat{Z}_{s}^\varepsilon) - \bar{b}_1(X_{k_0}^\varepsilon) \right) \, ds \right|^2 \right) \\
& + \left( 12TL_b + 12TC \right) \int_0^T \mathbb{E}|X_{k\delta}^\varepsilon - X_s^\varepsilon|^2 \, ds \\
& + 12TL_b \int_0^T \mathbb{E}|Z_s^\varepsilon - \hat{Z}_s^\varepsilon|^2 \, ds \\
& + \left( 12TC + 12L\sigma_1 + 12L_{f_1} \right) \int_0^T \mathbb{E}\left( \sup_{0 \leq r \leq s} |X_r^\varepsilon - X_r^0|^2 \right) \, ds \\
= & \ I_1 + I_2 + I_3 + I_4,
\end{align*}
\]
By combining (14) with (11), it holds that

\[ E \left( b_1(X_{kδ, s}^ε, \tilde{Z}_{s}^ε) - \tilde{b}_1(X_{kδ, s}^ε) \right) dsdr = 2ε^2 \int_{0}^{δ_ε/ε} \int_{r}^{δ_ε/ε} E \left( b_1(X_{kδ, r}^ε, \tilde{Z}_{r}^ε) - \tilde{b}_1(X_{kδ, r}^ε) \right) dsdr. \quad (12) \]

And then we investigate the integrand of the above integration. By the Hölder inequality it holds that

\[ E \left( b_1(X_{kδ, r}^ε, \tilde{Z}_{r}^ε) - \tilde{b}_1(X_{kδ, r}^ε) \right) \left( b_1(X_{kδ, r}^ε, \tilde{Z}_{r}^ε) - \tilde{b}_1(X_{kδ, r}^ε) \right) \leq \left( E \left( b_1(X_{kδ, r}^ε, \tilde{Z}_{r}^ε) - \tilde{b}_1(X_{kδ, r}^ε) \right)^2 \right)^{1/2} \left( E \left( b_1(X_{kδ, r}^ε, \tilde{Z}_{r}^ε) - \tilde{b}_1(X_{kδ, r}^ε) \right)^2 \right)^{1/2} \leq C e^{-α(s-ε)}, \quad (13) \]

where the last inequality is based on \((H_{b_1}^ε, σ, δ_ε, \tilde{Z}_r^ε) \leq \sigma(\tilde{Z}_r^ε, 0 \leq v \leq r) \lor \mathcal{N}\) and \(\mathcal{N}\) is the collection of all \(P\)-measure zero sets. Inserting (13) in (12), we furthermore obtain that

\[ E \left| \int_{0}^{δ_ε/ε} \left( b_1(X_{kδ, r}^ε, \tilde{Z}_{kδ, r+s}^ε) - \tilde{b}_1(X_{kδ, r}^ε) \right) ds \right|^2 \leq 2ε^2 \int_{0}^{δ_ε/ε} \int_{r}^{δ_ε/ε} C e^{-α(s-ε)} dsdr \leq Cε^2 δ_ε/ε. \quad (14) \]

By combining (14) with (11), it holds that

\[ I_1 \leq \frac{C}{δ_ε/ε}. \quad (15) \]

Finally, applying (15) (8) (6) to (10), we have that

\[ E \left( \sup_{0 ≤ t ≤ T} |X_t^ε - X_t^0|^2 \right) \leq \frac{C}{δ_ε/ε} + C(δ_ε + 1)δ_ε + C(δ_ε + 1)δ_ε^2/ε + C \left( \sup_{0 ≤ t ≤ s} |X_r^ε - X_r^0|^2 \right) ds. \]

The Gronwall inequality admits us to obtain that

\[ E \left( \sup_{0 ≤ t ≤ T} |X_t^ε - X_t^0|^2 \right) \leq \left( \frac{C}{δ_ε} + C(δ_ε + 1)δ_ε + C(δ_ε + 1)δ_ε^2/ε \right) e^{CT}. \]

The proof is complete. □

**Remark 2.3.** Based on Theorem 2.2, it holds that \(X_t^ε\) converges to \(X_t^0\) in the mean square sense if \(\frac{ε}{δ_ε} \to 0\) and \(\frac{δ_ε}{ε} \to 0\) as \(ε \to 0\). For example, we take \(δ_ε = ε^{2/3}\), and have that \(\frac{ε}{δ_ε} = ε^{1/3} \to 0\), and \(\frac{δ_ε}{ε} = ε^{1/3} \to 0\) when \(ε \to 0\).

3. **Convergence of nonlinear filterings with correlated sensor noises**

In the section, we introduce the nonlinear filtering problems for \(X_t^ε\) and \(X_t^0\) and their relationship.
3.1. **Nonlinear filtering problems with the system (1).** In the subsection, we introduce nonlinear filtering problems of $X^\varepsilon_t$ and $X^0_t$.

For

$$Y^\varepsilon_t = \int_0^t h(X^\varepsilon_s)ds + \sigma_3 V_t + \sigma_4 B_t,$$

we make the following hypotheses:

**Assumption 3.**

$(H_h)$ $h$ is bounded.

$(H_{\sigma_3, \sigma_4})$ $\sigma_3^T \sigma_3 + \sigma_4 \sigma_4^T = I$, where $\sigma_3^T$ stands for the transpose of the matrix $\sigma_3$ and $I$ is the $d$ order unit matrix.

By $(H_{\sigma_3, \sigma_4})$, we know that $U_t := \sigma_3 V_t + \sigma_4 B_t$ is a $d$ dimensional Brownian motion. Denote

$$(\gamma^\varepsilon_t)^{-1} := \exp \left\{ -\int_0^t h^i(X^\varepsilon_s)dU^i_s - \frac{1}{2} \int_0^t |h(X^\varepsilon_s)|^2 ds \right\}.$$ 

Here and hereafter, we use the convention that repeated indices imply summation. And then by $(H_h)$ we know that $(\gamma^\varepsilon_t)^{-1}$ is an exponential martingale. Define a measure $\mathbb{P}^\varepsilon$ via

$$\frac{d\mathbb{P}^\varepsilon}{d\mathbb{P}} = (\gamma^\varepsilon_T)^{-1}.$$ 

By the Girsanov theorem for Brownian motions, one can obtain that

$$Y^\varepsilon_t = U_t + \int_0^t h(X^\varepsilon_s)ds$$

is a $\mathcal{F}_t$-Brownian motion under the probability measure $\mathbb{P}^\varepsilon$.

Next, we rewrite $\gamma^\varepsilon_t$ as

$$\gamma^\varepsilon_t = \exp \left\{ \int_0^t h^i(X^\varepsilon_s)dY^\varepsilon^i_s - \frac{1}{2} \int_0^t |h(X^\varepsilon_s)|^2 ds \right\}.$$ 

Define

$$\rho^\varepsilon_t(\psi) := \mathbb{E}^{\mathbb{P}^\varepsilon}[\psi(X^\varepsilon_t)\gamma^\varepsilon_t | \mathcal{F}_t],$$

$$\pi^\varepsilon_t(\psi) := \mathbb{E}[\psi(X^\varepsilon_t) | \mathcal{F}_t^{Y^\varepsilon}], \quad \psi \in \mathcal{B}(\mathbb{R}^n),$$

where $\mathbb{E}^{\mathbb{P}^\varepsilon}$ denotes the expectation under the measure $\mathbb{P}^\varepsilon$, $\mathcal{F}_t^{Y^\varepsilon} \triangleq \sigma(Y^\varepsilon_s : 0 \leq s \leq t) \vee \mathcal{N}$, $\mathcal{N}$ is the collection of all $\mathbb{P}$-measure zero sets and $\mathcal{B}(\mathbb{R}^n)$ denotes the collection of all bounded and Borel measurable functions on $\mathbb{R}^n$. $\rho^\varepsilon_t$ and $\pi^\varepsilon_t$ are called the nonnormalized filtering and the normalized filtering of $X^\varepsilon_t$ with respect to $\mathcal{F}_t^{Y^\varepsilon}$, respectively. And then by the Kallianpur-Striebel formula it holds that

$$\pi^\varepsilon_t(\psi) = \frac{\rho^\varepsilon_t(\psi)}{\rho^\varepsilon_t(1)}.$$ 

Set

$$\gamma^0_t := \exp \left\{ \int_0^t h^i(X^0_s)dY^\varepsilon^i_s - \frac{1}{2} \int_0^t |h(X^0_s)|^2 ds \right\},$$

and furthermore

$$\rho^0_t(\psi) := \mathbb{E}^{\mathbb{P}^\varepsilon}[\psi(X^0_t)\gamma^0_t | \mathcal{F}_t],$$
\[ \pi^0_t(\psi) := \frac{\rho^0_t(\psi)}{\rho^0_t(1)}. \]

And then we will prove that \( \pi^0 \) could be understood as the nonlinear filtering problem for \( X^0_t \) with respect to \( \mathcal{F}^Y_t \).

3.2. The relation between \( \pi^0_\varepsilon \) and \( \pi^0 \). In this subsection we will show that \( \pi^0_\varepsilon \) converges to \( \pi^0_t \) as \( \varepsilon \to 0 \) in a suitable sense. Let us start with a key lemma.

**Lemma 3.1.** Under (H\(_n\)), there exists a constant \( C > 0 \) such that

\[ \mathbb{E} |\rho^0_t(1)|^{-p} \leq \exp \{ (2p^2 + p + 1)CT/2 \}, \quad t \in [0, T], \quad p > 1. \]

**Proof.** Although the proof is similar to Lemma 4.1 in [17], we prove it to the readers convenience. For \( \mathbb{E} |\rho^0_t(1)|^{-p} \), we compute

\[ \mathbb{E} |\rho^0_t(1)|^{-p} = \mathbb{E}^\varepsilon |\rho^0_t(1)|^{-p} \gamma^\varepsilon \leq (\mathbb{E}^\varepsilon |\rho^0_t(1)|^{-2p})^{1/2}(\mathbb{E}^\varepsilon (\gamma^\varepsilon T)^2)^{1/2}, \]

where the last inequality is based on the Hölder inequality. For \( \mathbb{E}^\varepsilon |\rho^0_t(1)|^{-2p} \), note that \( \rho^0_t(1) = \mathbb{E}^\varepsilon [\gamma^0_t|\mathcal{F}^Y_t] \). And then it follows from the Jensen inequality that

\[ \mathbb{E}^\varepsilon |\rho^0_t(1)|^{-2p} = \mathbb{E}^\varepsilon [\gamma^0_t|\mathcal{F}^Y_t]|^{-2p} \leq \mathbb{E}^\varepsilon [\mathbb{E}^\varepsilon [\gamma^0_t|\mathcal{F}^Y_t]] = \mathbb{E}^\varepsilon [\gamma^0_t|^{-2p}]. \]

Thus, the definition of \( \gamma^0_t \) allows us to obtain that

\[ \mathbb{E}^\varepsilon [\gamma^0_t|^{-2p}] = \mathbb{E}^\varepsilon \left[ \exp \left\{ -2p \int_0^t h(X^0_s)dY^\varepsilon_s + \frac{2p}{2} \int_0^t |h(X^0_s)|^2 ds \right\} \right] \]

\[ = \mathbb{E}^\varepsilon \left[ \exp \left\{ -2p \int_0^t h(X^0_s)dY^\varepsilon_s - \frac{4p^2}{2} \int_0^t |h(X^0_s)|^2 ds \right\} \right] \]

\[ \cdot \exp \left\{ \left( \frac{4p^2}{2} + \frac{2p}{2} \right) \int_0^t |h(X^0_s)|^2 ds \right\} \]

\[ \leq \exp \{ (2p^2 + p)CT \} \mathbb{E}^\varepsilon \left[ \exp \left\{ -2p \int_0^t h(X^0_s)dY^\varepsilon_s - \frac{4p^2}{2} \int_0^t |h(X^0_s)|^2 ds \right\} \right] \]

\[ = \exp \{ (2p^2 + p)CT \}, \]

where the last step is based on the fact that \( \exp \left\{ -2p \int_0^t h(X^0_s)dY^\varepsilon_s - \frac{4p^2}{2} \int_0^t |h(X^0_s)|^2 ds \right\} \)

is an exponential martingale under \( \mathbb{P}^\varepsilon \).

Similarly, we know that \( \mathbb{E}^\varepsilon (\gamma^\varepsilon T)^2 \leq \exp \{ CT \} \). So, by simple calculation, it holds that \( \mathbb{E} |\rho^0_t(1)|^{-p} \leq \exp \{ (2p^2 + p + 1)CT/2 \} \). The proof is complete. \( \square \)

**Theorem 3.2.** Suppose that **Assume 1.-3.** hold. Then it holds that for \( \phi \in C^1_b(\mathbb{R}^n) \)

\[ \lim_{\varepsilon \to 0} \mathbb{E} |\pi^\varepsilon_t(\phi) - \pi^0_t(\phi)| = 0, \quad \text{(17)} \]

where \( C^1_b(\mathbb{R}^n) \) denotes the collection of all the functions which themselves and their first order partial derivatives are bounded and Borel measurable.

**Proof.** For \( \phi \in C^1_b(\mathbb{R}^n) \), it follows from the Hölder inequality and Lemma 3.1 that

\[ \mathbb{E} |\pi^\varepsilon_t(\phi) - \pi^0_t(\phi)| = \mathbb{E} \left| \frac{\rho^\varepsilon_t(\phi) - \rho^0_t(\phi)}{\rho^0_t(1)} - \pi^\varepsilon_t(\phi) \frac{\rho^\varepsilon_t(1) - \rho^0_t(1)}{\rho^0_t(1)} \right| \]
\[
\begin{align*}
\leq & \mathbb{E} \left| \frac{\rho_t^\varepsilon(\phi) - \rho_t^0(\phi)}{\rho_t^0(1)} \right| + \mathbb{E} \left| \pi_t^\varepsilon(\phi) \frac{\rho_t^\varepsilon(1) - \rho_t^0(1)}{\rho_t^0(1)} \right| \\
\leq & \left( \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \right)^{1/r_1} \left( \mathbb{E} \left| \rho_t^0(1) \right|^{-r_2} \right)^{1/r_2} \\
& + \|\phi\|_{C^1_{\mathbb{R}^n}} \left( \mathbb{E} \left| \rho_t^\varepsilon(1) - \rho_t^0(1) \right|^{r_1} \right)^{1/r_1} \left( \mathbb{E} \left| \rho_t^0(1) \right|^{-r_2} \right)^{1/r_2} \\
\leq & C \left( \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \right)^{1/r_1} + C \|\phi\|_{C^1_{\mathbb{R}^n}} \left( \mathbb{E} \left| \rho_t^\varepsilon(1) - \rho_t^0(1) \right|^{r_1} \right)^{1/r_1},
\end{align*}
\]

where \(1 < r_1 < 2, r_2 > 1\) and \(1/r_1 + 1/r_2 = 1\).

Next, we estimate \(\mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1}\). Note that
\[
\begin{align*}
\mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} &= \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \gamma_{T}^{r_1} \\
&\leq \left( \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \right)^{1/p_1} \left( \mathbb{E} \left| \gamma_t^{r_1} \right|^{p_2} \right)^{1/p_2} \\
&\leq \exp \{CT\} \left( \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \right)^{1/p_1},
\end{align*}
\]

where \(1 < p_1 < 2, 1 < r_1 p_1 < 2, p_2 > 1\) and \(1/p_1 + 1/p_2 = 1\). And then we only need to observe \(\mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1 p_1}\). Based on the definitions of \(\rho_t^\varepsilon(\phi), \rho_t^0(\phi)\) and the Jensen inequality, it holds that
\[
\begin{align*}
\mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1 p_1} &= \mathbb{E} \left[ \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \gamma_t^{r_1} \right]^{p_1} \\
&= \mathbb{E} \left[ \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \gamma_t^{r_1} \right]^{p_1} \\
&\leq \mathbb{E} \left[ \mathbb{E} \left[ \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right]^{r_1 p_1} \right] \\
&\leq 2^{r_1 p_1 - 1} \mathbb{E} \left[ \left( \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right)^{r_1 p_1} \right] \\
&\quad + 2^{r_1 p_1 - 1} \mathbb{E} \left[ \left( \rho_t^0(\phi) - \rho_t^\varepsilon(\phi) \right)^{r_1 p_1} \right] \\
&=: I_1 + I_2.
\end{align*}
\]

First, we deal with \(I_1\). By the Hölder inequality, it holds that
\[
\begin{align*}
I_1 &\leq 2^{r_1 p_1 - 1} \left( \mathbb{E} \left[ \left( \phi(X_t^\varepsilon) - \phi(X_t^0) \right)^{r_1 p_1 q_1} \right] \right)^{1/q_1} \left( \mathbb{E} \left[ \left| \gamma_t^{r_1} \right|^{r_1 p_1 q_2} \right] \right)^{1/q_2} \\
&\leq 2^{r_1 p_1 - 1} \|\phi\|_{C^1_{\mathbb{R}^n}} \left( \mathbb{E} \left[ \left| X_t^\varepsilon - X_t^0 \right|^{r_1 p_1 q_1} \right] \right)^{1/q_1} \\
&\quad \cdot \left( \mathbb{E} \exp \left\{ r_1 p_1 q_2 \int_0^t h(X_s^\varepsilon) dY_s \right\} - \frac{(r_1 p_1 q_2)^2}{2} \int_0^t |h(X_s^\varepsilon)|^2 ds \right) \\
&\quad \cdot \exp \left\{ \frac{(r_1 p_1 q_2)^2}{2} \int_0^t |h(X_s^\varepsilon)|^2 ds - \frac{r_1 p_1 q_2}{2} \int_0^t |h(X_s^\varepsilon)|^2 ds \right\}^{1/q_2} \\
&\leq 2^{r_1 p_1 - 1} \|\phi\|_{C^1_{\mathbb{R}^n}} \left( \mathbb{E} \left[ \left| X_t^\varepsilon - X_t^0 \right|^{r_1 p_1 q_1} \right] \right)^{1/q_1} e^{\frac{r_1 p_1 q_2}{2} (r_1 p_1 q_2 - 1) CT},
\end{align*}
\]

where \(1 < q_1 < 2, 1 < r_1 p_1 q_1 < 2, q_2 > 1\) and \(1/q_1 + 1/q_2 = 1\), and the last step is based on the fact that the process \(\exp \left\{ r_1 p_1 q_2 \int_0^t h(X_s^\varepsilon) dY_s - \frac{(r_1 p_1 q_2)^2}{2} \int_0^t |h(X_s^\varepsilon)|^2 ds \right\} \) is an exponential martingale under \(\mathbb{P}^\varepsilon\). Note that
\[
\begin{align*}
\mathbb{E} \left| X_t^\varepsilon - X_t^0 \right|^{r_1 p_1 q_1} &= \mathbb{E} \left| X_t^\varepsilon - X_t^0 \right|^{r_1 p_1 q_1} \gamma_t^{-1} \\
&\leq \left( \mathbb{E} \left| X_t^\varepsilon - X_t^0 \right|^2 \right)^{r_1 p_1 q_1 / 2} \left( \mathbb{E} \gamma_t^{-2} \right)^{(2 - r_1 p_1 q_1) / 2 - (2 - r_1 p_1 q_1) / 2}
\end{align*}
\]
\[ \leq CR(\varepsilon)^{r_1 p_1 q_1 / 2}, \]  

where \( R(\varepsilon) := (C_{\delta_\varepsilon} + C(\delta_\varepsilon + 1)\delta_\varepsilon + C(\delta_\varepsilon + 1)\varepsilon^2) e^{CT} \) and the last step is based on Theorem 2.2. Thus, by inserting (20) in (19), we have that

\[ I_1 \leq C\|\phi\|_{C^1_{\delta_\varepsilon}(\mathbb{R}^n)} R(\varepsilon)^{r_1 p_1 / 2}. \]

We choose \( \delta_\varepsilon \) as that in Remark 2.3 and obtain that \( \lim_{\varepsilon \to 0} R(\varepsilon) = 0 \) and

\[ \lim_{\varepsilon \to 0} I_1 = 0. \]  

Next, for \( I_2 \), we know that

\[ I_2 \leq 2^{r_1 p_1 - 1} \|\phi\|_{C^1_{\delta_\varepsilon}(\mathbb{R}^n)} \mathbb{E} \left[ |\gamma_t^\varepsilon - \gamma_t^0|^{r_1 p_1} \right] = 2^{r_1 p_1 - 1} \|\phi\|_{C^1_{\delta_\varepsilon}(\mathbb{R}^n)} \mathbb{E} \left[ |\gamma_t^\varepsilon - \gamma_t^0|^{r_1 p_1} \right] (\varepsilon_T)^{-1} \]

\[ \leq 2^{r_1 p_1 - 1} \|\phi\|_{C^1_{\delta_\varepsilon}(\mathbb{R}^n)} \left( \mathbb{E} |\gamma_t^\varepsilon - \gamma_t^0|^2 \right)^{r_1 p_1 / 2} \left( \mathbb{E} (\gamma_t^\varepsilon)^{2/(2-r_1 p_1)} \right)^{2-r_1 p_1 / 2} \]

\[ \leq C \left( \mathbb{E} |\gamma_t^\varepsilon - \gamma_t^0|^2 \right)^{r_1 p_1 / 2}. \]

Note that \( \gamma_t^\varepsilon, \gamma_t^0 \) have the following expressions

\[ \gamma_t^\varepsilon = \exp \left\{ \int_0^t h(X_s^\varepsilon)^i dU_s^j + \frac{1}{2} \int_0^t |h(X_s^\varepsilon)|^2 ds \right\}, \]

\[ \gamma_t^0 = \exp \left\{ \int_0^t h(X_s^0)^i dU_s^j + \int_0^t h(X_s^0)^i h(X_s^\varepsilon)^i ds - \frac{1}{2} \int_0^t |h(X_s^0)|^2 ds \right\}. \]

So, by Theorem 2.2 and simple calculation, it holds that

\[ \lim_{\varepsilon \to 0} |\gamma_t^\varepsilon - \gamma_t^0| = 0. \]

Moreover, (H) admits us to get that

\[ |\gamma_t^\varepsilon|^2 \leq \exp \left\{ \int_0^t 2h(X_s^\varepsilon)^i dU_s^j - \frac{1}{2} \int_0^t |2h(X_s^\varepsilon)|^2 ds \right\} e^{CT}, \]

\[ |\gamma_t^0|^2 \leq \exp \left\{ \int_0^t 2h(X_s^0)^i dU_s^j - \frac{1}{2} \int_0^t |2h(X_s^0)|^2 ds \right\} e^{CT}, \]

and then \( \mathbb{E} |\gamma_t^\varepsilon|^2 \leq e^{CT}, \mathbb{E} |\gamma_t^0|^2 \leq e^{CT} \). Thus, the dominated convergence theorem yields that

\[ \lim_{\varepsilon \to 0} I_2 = 0. \]  

Finally, combining (21) (22) with (18), we obtain that

\[ \lim_{\varepsilon \to 0} \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1 p_1} = 0, \]

and furthermore

\[ \lim_{\varepsilon \to 0} \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} = 0. \]

We recall that

\[ \mathbb{E} |\pi_t^\varepsilon(\phi) - \pi_t^0(\phi)| \leq C \left( \mathbb{E} \left| \rho_t^\varepsilon(\phi) - \rho_t^0(\phi) \right|^{r_1} \right)^{1/r_1} + C \|\phi\|_{C^1_{\delta_\varepsilon}(\mathbb{R}^n)} \mathbb{E} \left| \rho_t^\varepsilon(1) - \rho_t^0(1) \right|^{r_1} \]  

Thus, taking the limit on two sides as \( \varepsilon \to 0 \), one can get (17). The proof is complete. \( \square \)
**Remark 3.3.** Here we can not give out the convergence rate of \(\pi_\varepsilon^x\) to \(\pi_0^x\). That is because the convergence of the slow part to the homogenized system is in \(L^2\) sense and is not in \(L^p\) sense for any \(p > 1\).

4. Convergence of nonlinear filterings with correlated noises

In the section, we study the nonlinear filtering problem of the system (2). First of all, we give out our assumption.

**Assumption 4.**

(i) \(\tilde{b}_1, \tilde{\sigma}_0, \tilde{\sigma}_1, \tilde{f}_1\) satisfy \((H_{b_1,\sigma_1,f_1})-(H_{b_1,\sigma_1,f_1})\), where \(\tilde{b}_1, (\sigma_0, \sigma_1), \tilde{f}_1\) replace \(b_1, \sigma_1, f_1\);

(ii) \(\tilde{b}_2, \tilde{\sigma}_2, \tilde{f}_2\) satisfy \((H_{b_2}, (H_{\sigma_2}), \tilde{f}_2\), respectively;

(iii) \(\tilde{b}_2, \tilde{\sigma}_2, \tilde{f}_2\) satisfy \((H_{\sigma_2}^2), (H_{b_2,\sigma_2,f_2}^1)-(H_{b_2,\sigma_2,f_2}^2), \) where \(\tilde{b}_2, \tilde{\sigma}_2, \tilde{f}_2\) replace \(b_2, \sigma_2, f_2\).

Under **Assumption 4. (i)-(ii)**, by Theorem 1.2 in [14], the system (2) has a unique strong solution denoted by \((\hat{X}_t^\varepsilon, \hat{Z}_t^\varepsilon)\). And then take any \(x \in \mathbb{R}^n\) and fix it. And consider the following SDE in \(\mathbb{R}^n\):

\[
\begin{align*}
\begin{cases}
\ d\hat{Z}_t^\varepsilon &= \hat{b}_2(x, \hat{Z}_t^\varepsilon)dt + \hat{\sigma}_2(x, \hat{Z}_t^\varepsilon)dW_t + \int_{U_2} \hat{f}_2(x, \hat{Z}_t^\varepsilon, u)\tilde{N}_{p_2}(dt, du), \\
\ Z_0^\varepsilon &= \tilde{z}_0,
\end{cases}
\end{align*}
\]

Based on **Assumption 4. (ii)-(iii)**, it holds that the above equation has a unique invariant probability measure denoted as \(\tilde{p}(x, d\varepsilon)\). So, set

\[
\tilde{b}_1(x) := \int_{\mathbb{R}^m} \tilde{b}_1(x, z)\tilde{p}(x, dz),
\]

and by [12, Lemma 3.1], we know that \(\tilde{b}_1\) is Lipschitz continuous. So, we construct a SDE on the probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})\) as follows:

\[
\begin{align*}
\begin{cases}
\ d\hat{X}_t^0 &= \tilde{b}_1(X_t^0)dt + \tilde{\sigma}_0(X_t^0)d\hat{B}_t + \tilde{\sigma}_1(X_t^0)d\hat{V}_t + \int_{U_1} \tilde{f}_1(X_t^0, u)\tilde{N}_{p_1}(dt, du), \\
\ X_0^0 &= \hat{x}_0,
\end{cases}
\text{for } 0 \leq t \leq T.
\end{align*}
\]

The solution of Eq. (23) is denoted as \(\hat{X}_t^0\). By the same deduction to that in Theorem 2.2 we can obtain the following theorem.

**Theorem 4.1.** There exists a constant \(C \geq 0\) independent of \(\varepsilon, \delta_\varepsilon\) such that

\[
\mathbb{E} \left( \sup_{0 \leq t \leq T} |\hat{X}_t^\varepsilon - \hat{X}_t^0|^2 \right) \leq \left( C \frac{\varepsilon}{\delta_\varepsilon} + C(\delta_\varepsilon + 1)\delta_\varepsilon + C(\delta_\varepsilon + 1)\frac{\delta_\varepsilon^2}{\varepsilon} \right) e^{CT}.
\]

4.1 Nonlinear filtering problems with the system (2). Next, for the observation process \(\hat{Y}_t^\varepsilon\) defined in (2), i.e.

\[
\hat{Y}_t^\varepsilon = V_t + \int_0^t \hat{h}(\hat{X}_s^\varepsilon)ds + \int_0^t \int_{U_3} \hat{f}_3(s, u)\tilde{N}_\lambda(ds, du) + \int_0^t \int_{U_3} \tilde{g}_3(s, u)N_\lambda(ds, du),
\]

we assume:

**Assumption 5.**

(i) \(\hat{h}\) is bounded and

\[
\int_0^T \int_{U_3} |\hat{f}_3(s, u)|^2\nu_3(du)ds < \infty.
\]
(ii) There exists a positive function $\tilde{L}(u)$ satisfying
\[ \int_{U_3} \frac{(1 - \tilde{L}(u))^2}{\tilde{L}(u)} \nu_3(du) < \infty \]
such that $0 < \tilde{l} \leq \tilde{L}(u) < \lambda(t, x, u) < 1$ for $u \in U_3$, where $\tilde{l}$ is a constant.

Now, denote
\[ (\lambda_\varepsilon^{-1}) = \exp \left\{ - \int_0^t \hat{h}(X_s^\varepsilon) dV_s^\varepsilon - \frac{1}{2} \int_0^t |\hat{h}(X_s^\varepsilon)|^2 ds - \int_0^t \int_{U_3} \log \lambda(s, X_s^\varepsilon, u) \nu_3(ds, du) \right\}. \]

Thus, by Assumption 5. we know that $(\lambda_\varepsilon^{-1})$ is an exponential martingale. Define a measure $\tilde{P}_\varepsilon$ via
\[ \frac{d\tilde{P}_\varepsilon}{dP} = (\lambda_\varepsilon^{-1}). \]

Under the probability measure $\tilde{P}_\varepsilon$, it follows from the Girsanov theorem that $\tilde{V}_t := V_t + \int_0^t \hat{h}(X_s^\varepsilon) ds$ is a Brownian motion and $N_\varepsilon(dt, du)$ is a Poisson random measure with the predictable compensator $d\nu_3(du)$. Moreover, by the same deduction to that in [16 Lemma 3.1], we know that $\lambda_\varepsilon$ satisfies the following equation
\[ \lambda_\varepsilon = 1 + \int_0^t \lambda_\varepsilon \hat{h}(X_s^\varepsilon) dV_s^\varepsilon + \int_0^t \int_{U_3} \lambda_\varepsilon - (\lambda(s, X_s^\varepsilon, u) - 1) \tilde{N}(ds, du), \tag{24} \]

where $\tilde{N}(ds, du) := N_\varepsilon(dt, du) - d\nu_3(du)$. Set
\[ \tilde{\rho}_\varepsilon(\psi) := \mathbb{E}_{\tilde{P}_\varepsilon}[\psi(X_t^\varepsilon) \lambda_\varepsilon | \mathcal{F}_t^{\tilde{Y}_t^\varepsilon}], \]
\[ \tilde{\pi}_\varepsilon(\psi) := \mathbb{E}[\psi(X_t^\varepsilon) | \mathcal{F}_t^{\tilde{Y}_t^\varepsilon}], \quad \psi \in \mathcal{B}(\mathbb{R}^n), \]

where $\mathbb{E}_{\tilde{P}_\varepsilon}$ stands for the expectation under the probability measure $\tilde{P}_\varepsilon$. And then by the Kallianpur-Striebel formula it holds that
\[ \tilde{\pi}_\varepsilon(\psi) = \frac{\tilde{\rho}_\varepsilon(\psi)}{\tilde{\rho}_\varepsilon(1)}. \]

In addition, we have the following result.

**Theorem 4.2.** (The Zakai equation) For $\psi \in \mathcal{C}_b^2(\mathbb{R}^n)$, $\tilde{\rho}_\varepsilon(\psi)$ satisfies the following Zakai equation
\[ \tilde{\rho}_\varepsilon(\psi) = \tilde{\rho}_0(\psi) + \int_0^t \tilde{\rho}_s \left( (\mathcal{L}^{X_s^\varepsilon} \psi)(\cdot, \tilde{Z}_s^\varepsilon) \right) ds + \int_0^t \left( \tilde{\rho}_s(\psi \hat{h}^i) + \tilde{\rho}_s((\partial_j \psi) \tilde{\sigma}_1^{ij}) \right) dV_s^i \]
\[ + \int_0^t \int_{U_3} \tilde{\rho}_s \left( \psi(\lambda(s, \cdot, u) - 1) \right) \tilde{N}(ds, du), \tag{25} \]

where
\[ (\mathcal{L}^{X_s^\varepsilon} \psi)(x, z) := \frac{\partial \psi(x)}{\partial x_1} b_1^i(x, z) + \frac{1}{2} \frac{\partial^2 \psi(x)}{\partial x_i \partial x_j} (\tilde{\sigma}_0 \tilde{T}^i_j)(x) + \frac{1}{2} \frac{\partial^2 \psi(x)}{\partial x_i \partial x_j} (\tilde{\sigma}_1 \tilde{T}^i_j)(x) \]
\[ + \int_{U_1} \left[ \psi(x + \tilde{f}_1(x, u)) - \psi(x) - \frac{\partial \psi(x)}{\partial x_i} \tilde{f}_1^i(x, u) \right] \nu_1(du). \]
Proof. Applying the Itô formula to $\psi(\bar{X}_t^\varepsilon)$, one can have that

$$
\psi(\bar{X}_t^\varepsilon) = \psi(\bar{x}_0) + \int_0^t (L\bar{X}_s^\varepsilon \psi)(\bar{X}_s^\varepsilon, \bar{Z}_s^\varepsilon)ds + \int_0^t (\nabla \psi)(\bar{X}_s^\varepsilon)\tilde{\sigma}_0(\bar{X}_s^\varepsilon)dB_s
+ \int_0^t (\nabla \psi)(\bar{X}_s^\varepsilon)\bar{\sigma}_1(\bar{X}_s^\varepsilon)d(\bar{V}_s - \int_0^s \tilde{h}(\bar{X}_r^\varepsilon)dr)
+ \int_0^t \int_{U_1} [\psi(\bar{X}_s^\varepsilon + \bar{f}_1(\bar{X}_s^\varepsilon, u)) - \psi(\bar{X}_s^\varepsilon)]\bar{N}_{p_1}(ds, du).
$$

Note that $\lambda_s^\varepsilon$ satisfies (24). So, it follows from the Itô formula that

$$
\psi(\bar{X}_t^\varepsilon)\lambda_t^\varepsilon = \psi(\bar{x}_0) + \int_0^t \psi(\bar{X}_s^\varepsilon)\lambda_s^\varepsilon\tilde{h}(\bar{X}_s^\varepsilon)i d\bar{V}_s^i
+ \int_0^t \int_{U_3} \psi(\bar{X}_s^\varepsilon)\lambda_s^\varepsilon_- (\lambda(s, \bar{X}_s^\varepsilon, u) - 1)\bar{N}(ds, du)
+ \int_0^t \lambda_s^\varepsilon(L\bar{X}_s^\varepsilon \psi)(\bar{X}_s^\varepsilon, \bar{Z}_s^\varepsilon)ds + \int_0^t \lambda_s^\varepsilon(\nabla \psi)(\bar{X}_s^\varepsilon)\tilde{\sigma}_0(\bar{X}_s^\varepsilon)dB_s
+ \int_0^t \lambda_s^\varepsilon(\nabla \psi)(\bar{X}_s^\varepsilon)\bar{\sigma}_1(\bar{X}_s^\varepsilon)d\left(\bar{V}_s - \int_0^s \tilde{h}(\bar{X}_r^\varepsilon)dr\right)
+ \int_0^t \int_{U_1} \lambda_s^\varepsilon_- [\psi(\bar{X}_s^\varepsilon + \bar{f}_1(\bar{X}_s^\varepsilon, u)) - \psi(\bar{X}_s^\varepsilon)]\bar{N}_{p_1}(ds, du)
+ \int_0^t (\partial_j \psi)(\bar{X}_s^\varepsilon)\bar{\sigma}_1^j(\bar{X}_s^\varepsilon)\lambda_s^\varepsilon\tilde{h}(\bar{X}_s^\varepsilon)i ds
= \psi(\bar{x}_0) + \int_0^t \left(\psi(\bar{X}_s^\varepsilon)\lambda_s^\varepsilon\tilde{h}(\bar{X}_s^\varepsilon)i + \lambda_s^\varepsilon(\partial_j \psi)(\bar{X}_s^\varepsilon)\bar{\sigma}_1^j(\bar{X}_s^\varepsilon)\right) d\bar{V}_s^i
+ \int_0^t \int_{U_3} \psi(\bar{X}_s^\varepsilon)\lambda_s^\varepsilon_- (\lambda(s, \bar{X}_s^\varepsilon, u) - 1)\bar{N}(ds, du)
+ \int_0^t \lambda_s^\varepsilon(L\bar{X}_s^\varepsilon \psi)(\bar{X}_s^\varepsilon, \bar{Z}_s^\varepsilon)ds + \int_0^t \lambda_s^\varepsilon(\nabla \psi)(\bar{X}_s^\varepsilon)\tilde{\sigma}_0(\bar{X}_s^\varepsilon)dB_s
+ \int_0^t \int_{U_1} \lambda_s^\varepsilon_- [\psi(\bar{X}_s^\varepsilon + \bar{f}_1(\bar{X}_s^\varepsilon, u)) - \psi(\bar{X}_s^\varepsilon)]\bar{N}_{p_1}(ds, du).
$$

Taking the conditional expectation with respect to $\mathcal{F}_t^{\bar{Y}_s^\varepsilon}$ under $\mathbb{P}^\varepsilon$ on two hand sides of the above equality, one could obtain that

$$
\mathbb{E}^{\mathbb{P}^\varepsilon}[\psi(\bar{X}_t^\varepsilon)\lambda_t^\varepsilon|\mathcal{F}_t^{\bar{Y}_s^\varepsilon}] = \mathbb{E}^{\mathbb{P}^\varepsilon}[\psi(\bar{x}_0)|\mathcal{F}_t^{\bar{Y}_s^\varepsilon}]
+ \int_0^t \mathbb{E}^{\mathbb{P}^\varepsilon}[\left(\psi(\bar{X}_s^\varepsilon)\lambda_s^\varepsilon\tilde{h}(\bar{X}_s^\varepsilon)i + \lambda_s^\varepsilon(\partial_j \psi)(\bar{X}_s^\varepsilon)\bar{\sigma}_1^j(\bar{X}_s^\varepsilon)\right)|\mathcal{F}_t^{\bar{Y}_s^\varepsilon}]d\bar{V}_s^i
+ \int_0^t \int_{U_3} \mathbb{E}^{\mathbb{P}^\varepsilon}[\psi(\bar{X}_s^\varepsilon)\lambda_s^\varepsilon_- (\lambda(s, \bar{X}_s^\varepsilon, u) - 1)|\mathcal{F}_t^{\bar{Y}_s^\varepsilon}]\bar{N}(ds, du)
+ \int_0^t \mathbb{E}^{\mathbb{P}^\varepsilon}[\lambda_s^\varepsilon(L\bar{X}_s^\varepsilon \psi)(\bar{X}_s^\varepsilon, \bar{Z}_s^\varepsilon)|\mathcal{F}_t^{\bar{Y}_s^\varepsilon}]ds,
$$

17
where [18, Theorem 1.4.7] is used. That is, it holds that

\[ \bar{\rho}_t(\psi) = \bar{\rho}_0(\psi) + \int_0^t \bar{\rho}_s \left( (\mathcal{L}^{\psi}X^\varepsilon) \cdot \bar{\mu}^\varepsilon_s \right) ds + \int_0^t \left( \bar{\rho}_s(\psi \bar{h}^i) + \bar{\rho}_s((\partial_j \psi) \bar{\sigma}_1^i) \right) d\tilde{V}_s^i \\
+ \int_0^t \int_{U_3} \bar{\rho}_s(\psi(\lambda(s, \cdot, u) - 1)) \tilde{N}(ds, du). \]

The proof is complete. \(\Box\)

In the following, we define the nonlinear filtering of \(\bar{X}_t^0\) with respect to \(\mathcal{F}_t^\varepsilon\). Set

\[ \lambda_t^0 := \exp \left\{ \int_0^t \tilde{h}^i(\bar{X}_s^0) d\tilde{V}_s^i - \frac{1}{2} \int_0^t |\tilde{h}(\bar{X}_s^0)|^2 ds + \int_0^t \int_{U_3} \log \lambda(s, \bar{X}_s^0, u) N_\lambda(ds, du) + \int_0^t \int_{U_3} (1 - \lambda(s, \bar{X}_s^0, u)) \nu_3(du) ds \right\}, \]

and furthermore

\[ \bar{\rho}_t^0(\psi) := \mathbb{E}^{\mathbb{P}_0}[\psi(\bar{X}_t^0) \lambda_t^0 | \mathcal{F}_t^\varepsilon], \]
\[ \bar{\pi}_t^0(\psi) := \frac{\bar{\rho}_t^0(\psi)}{\bar{\rho}_t^0(1)}. \]

And then by the similar deduction to that in Theorem 4.2, it holds that \(\bar{\rho}_t^0\) satisfies the following equation

\[ \bar{\rho}_t^0(\psi) = \bar{\rho}_0^0(\psi) + \int_0^t \bar{\rho}_s^0 \left( (\mathcal{L}^{\psi}X^0) \cdot \bar{\mu}_s^0 \right) ds + \int_0^t \bar{\rho}_s^0 \left( ((\partial_j \psi) \bar{\sigma}_1^i) (\bar{h}^i(\cdot) - \bar{h}^i(\bar{X}_s^0)) \right) ds \\
+ \int_0^t \left( \bar{\rho}_s^0(\psi \bar{h}^i) + \bar{\rho}_s^0((\partial_j \psi) \bar{\sigma}_1^i) \right) d\tilde{V}_s^i \\
+ \int_0^t \int_{U_3} \bar{\rho}_s^0(\psi(\lambda(s, \cdot, u) - 1)) \tilde{N}(ds, du), \quad (26) \]

where

\[ (\mathcal{L}^{\psi}X^0)(x) := \frac{\partial \psi(x)}{\partial x_1} f_1^0(x) + \frac{1}{2} \frac{\partial^2 \psi(x)}{\partial x_i \partial x_j} (\bar{\sigma}_0 \bar{\sigma}_0^T)^ij(x) + \frac{1}{2} \frac{\partial^2 \psi(x)}{\partial x_i \partial x_j} (\bar{\sigma}_1 \bar{\sigma}_1^T)^ij(x) \\
+ \int_{U_1} \left[ \psi(x + \tilde{f}_1(x, u)) - \psi(x) - \frac{\partial \psi(x)}{\partial x_1} \tilde{f}_1(x, u) \right] \nu_1(du). \]

4.2. The relationship of \(\bar{\pi}^\varepsilon\) and \(\bar{\pi}^0\).

4.2.1. The case of \(\tilde{f}_3 = \bar{g}_3 = 0\). In the case of \(\tilde{f}_3 = \bar{g}_3 = 0\), by the similar deduction to that in Theorem 3.2, one can obtain the following result.

**Theorem 4.3.** Suppose that Assume 4.-5. hold. Then it holds that for \(\phi \in C_b^1(\mathbb{R}^n)\),

\[ \lim_{\varepsilon \to 0} \mathbb{E} |\bar{\pi}^\varepsilon_t(\phi) - \bar{\pi}^0_t(\phi)| = 0. \]
4.2.2. The case of $\hat{f}_3 \neq \hat{g}_3 \neq 0$. In the case of $\hat{f}_3 \neq \hat{g}_3 \neq 0$, we first prepare two important lemmas. Since their proofs are similar to that of [12, Lemma 5.1, 5.2], we omit them.

**Lemma 4.4.** Under Assumption 4.-5., it holds that for any $t \in [0, T]$,
\[
(\tilde{\rho}_t^0(1))^{-1} < \infty, \quad \mathbb{P} \text{ a.s.}
\]  
(27)

**Lemma 4.5.** Under Assumption 4.-5., \{\tilde{\rho}_t^0, t \in [0, T]\} is relatively weakly compact in $D([0, T], \mathcal{M}(\mathbb{R}^n))$, where $\mathcal{M}(\mathbb{R}^n)$ denotes the set of bounded Borel measures on $\mathbb{R}^n$.

In the following, we assume more:

**Assumption 6.**
\{\hat{Z}_{\xi t}, t \in [0, T]\} is tight.

Now, we state and prove the main theorem in the section.

**Theorem 4.6.** Suppose that Assumption 4.-6. hold. Then $\hat{\pi}^\varepsilon_t$ converges weakly to $\hat{\pi}^0_t$ as $\varepsilon \to 0$ for any $t \in [0, T]$.

**Proof.** By the definition of $\hat{\pi}^\varepsilon_t$, $\hat{\pi}^0_t$, it holds that for $\phi \in C^2_b(\mathbb{R}^n)$,
\[
\hat{\pi}^\varepsilon_t(\phi) - \hat{\pi}^0_t(\phi) = \hat{\rho}_t^\varepsilon(\phi) - \hat{\rho}_t^0(\phi) - \hat{\pi}^\varepsilon_t(\phi) - \hat{\pi}^0_t(\phi).
\]
So, in order to prove $\hat{\pi}^\varepsilon_t(\phi) - \hat{\pi}^0_t(\phi)$ converges weakly to 0, in terms of Lemma 4.4 we only need to show that $\hat{\rho}_t^\varepsilon(\phi) - \hat{\rho}_t^0(\phi)$ converges weakly to 0 as $\varepsilon \to 0$. Besides, it follows from Lemma 4.5 that there exist a weak convergence subsequence $\{\hat{\rho}_t^{\varepsilon_k}, k \in \mathbb{N}\}$ and a measure-valued process $\tilde{\rho}_t$ such that $\hat{\rho}_t^{\varepsilon_k}$ converges weakly to $\tilde{\rho}_t$ as $k \to \infty$. Therefore, we just need to prove that for $t \in [0, T]$, $\tilde{\pi}_t(\phi) - \tilde{\pi}_t^0(\phi)$ converges weakly to 0 as $\varepsilon \to 0$.

Next, we search for the equations which $\tilde{\rho}_t(\phi)$ solves. By Theorem 4.1 and (25), we follow up the line of [12, Theorem 5.3] and obtain that $\tilde{\rho}_t(\phi)$ satisfies the following equation
\[
\tilde{\rho}_t(\phi) = \tilde{\rho}_0(\phi) + \int_t^0 \tilde{\rho}_s(\mathcal{L}_s \phi) ds + \int_t^0 \left( \tilde{\rho}_s((\partial_j \phi) \tilde{\sigma}_t^j) \right) d\tilde{V}_s^i + \int_t^0 \left( \tilde{\rho}_s((\partial_j \phi) \tilde{\sigma}_t^j) \right) d\tilde{N}(ds, du).
\]
(28)

Besides, by (26) and Theorem 4.1, we know that, there exists a measure-valued process $\tilde{\rho}_t^0$ such that $\tilde{\rho}_t^0(\phi)$ converges to $\tilde{\rho}_t^0$ a.s. and $\tilde{\rho}_t^0$ satisfies the following equation
\[
\tilde{\rho}_t^0(\phi) = \tilde{\rho}_0^0(\phi) + \int_t^0 \tilde{\rho}_s^0(\mathcal{L}_s \phi) ds + \int_t^0 \left( \tilde{\rho}_s^0((\partial_j \phi) \tilde{\sigma}_t^j) \right) d\tilde{V}_s^i + \int_t^0 \left( \tilde{\rho}_s^0((\partial_j \phi) \tilde{\sigma}_t^j) \right) d\tilde{N}(ds, du).
\]
(29)

Note that Eq.(28) and Eq.(29) are the same. Thus, it follows from [12, Theorem 3.9] that for any $t \in [0, T]$,
\[
\tilde{\rho}_t = \tilde{\rho}_t^0, \quad \mathbb{P} \text{ a.s.}
\]
That is, $\tilde{\rho}_t(\phi) - \tilde{\rho}_t^0(\phi)$ converges weakly to 0 as $\varepsilon \to 0$. The proof is complete. \hfill \Box
5. Conclusion

In the paper, we consider nonlinear filtering problems of multiscale systems in two cases—correlated sensor Lévy noises and correlated Lévy noises. First of all, we prove that the slow part of the origin system converges to the homogenized system in the uniform mean square sense. Next, in the case of correlated sensor Lévy noises, the nonlinear filtering of the slow part is shown to approximate that of the homogenized system in $L^1$ sense. However, in the case of correlated Lévy noises, we prove that the nonlinear filtering of the slow part converges weakly to that of the homogenized system.

Acknowledgements:
The author would like to thank Professor Xicheng Zhang for his valuable discussions. The author also thanks Professor Renming Song for providing her an excellent environment to work in the University of Illinois at Urbana-Champaign.

References

[1] R. Beeson and N. S. Namachchivaya: Nudged Particle Filters in Multiscale Chaotic Systems with Correlated Sensor Noise, G. Yin and Q. Zhang (eds.), Modeling, Stochastic Control, Optimization, and Applications, The IMA Volumes in Mathematics and its Applications 164.
[2] D. Crisan: Particle approximations for a class of stochastic partial differential equations, *Applied Mathematics and Optimization*, 54(2006)293-314.
[3] P. D. Ditlevsen: Observation of α-stable noise induced millennial climate changes from an ice record, *Geophysical Research Letters*, 26(1999)1441-1444.
[4] D. Givon: Strong convergence rate for two-time-scale jumpdiffusion stochastic differential systems, *Multiscale Model. Simul.*, 6(2007)577-594.
[5] N. Ikeda and S. Watanabe: *Stochastic differential equations and diffusion processes*, 2nd ed., North-Holland/Kodanska, Amsterdam/Tokyo, 1989.
[6] P. Imkeller, N. S. Namachchivaya, N. Perkowski and H. C. Yeong: Dimensional reduction in nonlinear filtering: a homogenization approach, *The Annals of Applied Probability*, 23(2013)2290-2326.
[7] H. J. Kushner: *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Systems & Control: Foundations & Applications 3. Birkhäuser, Boston, 1990.
[8] H. J. Kushner: Robustness and convergence of approximations to nonlinear filters for jump-diffusions, *Computational and Applied Math.*, 16(1996)153-183.
[9] V. M. Lucic and A. J. Heunis: Convergence of nonlinear filters for randomly perturbed dynamical systems, *Appl. Math. Optim.*, 48(2003)93-128.
[10] T. Meyer-Brandis and F. Proske: Explicit solution of a non-linear filtering problems for Lévy Processes with application to finance, *Applied Mathematics and Optimization*, 50(2004)119-134.
[11] A. Papunicalaou and K. Spiliopoulos: Dimension reduction in statistical estimation of partially observed multiscale processes, *SIAM J. on Uncertainty Quantification*, 5(2017)1220-1247.
[12] H. J. Qiao: Convergence of nonlinear filtering for stochastic dynamical systems with Lévy noises, https://arxiv.org/abs/1707.07824.
[13] H. J. Qiao: Exponential ergodicity for SDEs with jumps and non-Lipschitz coefficients, *Journal of Theoretical Probability*, 27(2014)137-152.
[14] H. J. Qiao: Euler-Maruyama approximation for SDEs with jumps and non-Lipschitz coefficients, *Osaka Journal of Mathematics*, 51(2014)47-66.
[15] H. J. Qiao: Nonlinear filtering of stochastic differential equations driven by correlated Lévy noises, [http://arxiv.org/abs/1907.06779](http://arxiv.org/abs/1907.06779).
[16] H. J. Qiao and J. Q. Duan: Nonlinear Filtering of Stochastic Dynamical Systems with Lévy Noises, *Advances in Applied Probability*, 47(2015)902-918.
[17] H. J. Qiao, Y. J. Zhang and J. Q. Duan: Effective filtering on a random slow manifold, *Nonlinearity*, 31(2018) 4649-4666.
[18] B. L. Rozovskii: *Stochastic Evolution Systems, Linear Theory and Applications to Nonlinear Filtering*, Math. Appl. (Sov. Ser.), vol. 35, Kluwer Academic Publishers, 1990.

[19] G. Yin and Q. Zhang: *Continuous Time Markov Chains and Applications: A Two Time-Scale Approach*, Springer-Verlag, New York, 2013.

[20] Y. J. Zhang, H. J. Qiao and J. Q. Duan: Effective filtering analysis for non-Gaussian dynamic systems. Appear in *Applied Mathematics and Optimization*. 