CONVERGENCE OF NONLINEAR FILTERING FOR STOCHASTIC DYNAMICAL SYSTEMS WITH LÉVY NOISES*

HUIJIE QIAO

Department of Mathematics, Southeast University
Nanjing, Jiangsu 211189, China
hjqiaogean@seu.edu.cn

Abstract. We consider the nonlinear filtering problem of multiscale non-Gaussian signal processes and observation processes with jumps. Firstly, we prove that the dimension for the signal system could be reduced. Secondly, convergence of the corresponding nonlinear filtering to the homogenized filtering is shown by weak convergence approach.

1. Introduction

For a fixed time $T > 0$, given a completed filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, P).$ Consider the following slow-fast system on $\mathbb{R}^n \times \mathbb{R}^m$: for $0 \leq t \leq T,$

\[
\begin{cases}
    dX^\varepsilon_t = b_1(X^\varepsilon_t, Z^\varepsilon_t)dt + \sigma_1(X^\varepsilon_t, Z^\varepsilon_t)dV_t + \int_{U_1} f_1(X^\varepsilon_{t^-}, u)\tilde{N}_{p_1}(dt, du), \\
    X^\varepsilon_0 = x_0, \\
    dZ^\varepsilon_t = \frac{1}{\varepsilon}b_2(X^\varepsilon_t, Z^\varepsilon_t)dt + \frac{1}{\sqrt{\varepsilon}}\sigma_2(X^\varepsilon_t, Z^\varepsilon_t)dW_t + \int_{U_2} f_2(X^\varepsilon_{t^-}, Z^\varepsilon_{t^-}, u)\tilde{N}_{p_2}(dt, du), \\
    Z^\varepsilon_0 = z_0,
\end{cases}
\]

where $V, W$ are $\varepsilon$-dimensional and $m$-dimensional standard Brownian motion, respectively, and $p_1, p_2$ are two stationary Poisson point processes of the class (quasi left-continuous) defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, P)$ with values in $U$ and the characteristic measure $\nu_1, \nu_2$, respectively. Here $\nu_1, \nu_2$ are two $\sigma$-finite measures defined on a measurable space $(U, \mathcal{U})$. 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 $E N_{p_1}((0, t], A) = t \nu_1(A)$ for $A \in \mathcal{U}$. Denote

\[
\tilde{N}_{p_1}((0, t], du) := N_{p_1}((0, t], du) - t \nu_1(du),
\]

the compensated measure of $p_1(t)$. 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 $(U, \mathcal{U})$ such that $E N_{p_2}((0, t], A) = \frac{1}{\varepsilon} t \nu_2(A)$ for $A \in \mathcal{U}$. Moreover, $V_t, W_t, N_{p_1, p_2}, N_{p_2}^\varepsilon$ are mutually independent. The mappings $b_1 : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n, b_2 : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m, \sigma_1 : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^{n \times l}, \sigma_2 : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times m}, f_1 : \mathbb{R}^n \times U_1 \rightarrow \mathbb{R}^n$ and $f_2 : \mathbb{R}^n \times \mathbb{R}^m \times U_2 \rightarrow \mathbb{R}^m$ are all Borel measurable.

The slow-fast dynamical system (1) is usually called multiscale processes, where the rates of change of different variables differ by orders of magnitude. And multiple time scales models are widely applied in the science and engineering fields. For example, fast

AMS Subject Classification(2010): 60G35, 60G51, 60H10.
Keywords: Nonlinear filtering, dimensional reduction, homogenization, weak convergence.

*This work was partly supported by NSF of China (No. 11001051, 11371352).
atmospheric and slow oceanic dynamics describe the climate evolution and state dynamic in electric power systems consists of fast- and slowly-varying elements.

Next, define an observation process \( Y^\varepsilon \) by

\[
Y^\varepsilon_t = \int_0^t h(X^\varepsilon_s, Z^\varepsilon_s)ds + B_t + \int_0^t \int_{U_3} f_3(s, u)\tilde{N}_\lambda(ds, du) + \int_0^t \int_{U \cup U_3} g_3(s, u)N_\lambda(ds, du),
\]

where \( B \) is a \( d \)-dimensional standard Brownian motion and \( N_\lambda((0, t], du) \) is a Poisson random measure with a predictable compensator \( \lambda(t, X^\varepsilon_t, u)tv_3(du) \). Here the function \( \lambda(t, x, u) \in (0, 1) \) and \( v_3 \) is another \( \sigma \)-finite measure defined on \( U \) with \( v_3(U \setminus U_3) < \infty \) and \( \int_{U_3} ||u||^2_2 v_3(du) < \infty \) for a fixed \( U_3 \in \mathcal{U} \), where \( ||\cdot||_U \) denotes the norm on \((U, \mathcal{U})\). Set \( \tilde{N}_\lambda((0, t], du) := N_\lambda((0, t], du) - \lambda(t, X^\varepsilon_t, u)tv_3(du) \), and then \( \tilde{N}_\lambda((0, t], du) \) is the compensated martingale measure of \( N_\lambda((0, t], du) \). Moreover, \( V_t, W_t, B_t, N_{p_1}, N_{p_2}, N^\sigma_{p_2}, N_\lambda \) are mutually independent. \( h : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^d, f_3 : [0, T] \times U_3 \mapsto \mathbb{R}^d \) and \( g_3 : [0, T] \times (U \setminus U_3) \mapsto \mathbb{R}^d \) are all Borel measurable. For a Borel measurable function \( F \), the nonlinear filtering problem for the slow component \( X^\varepsilon_t \) with respect to \( \{Y^\varepsilon_s, 0 \leq s \leq t\} \) leads to evaluating the ‘filter’ \( \mathbb{E}[F(X^\varepsilon_t) | \mathcal{F}^Y_t] \), where \( \mathcal{F}^Y_t \) is the \( \sigma \)-algebra generated by \( \{Y^\varepsilon_s, 0 \leq s \leq t\} \) and \( \mathbb{E}|F(X^\varepsilon_t)| < \infty \) for \( t \in [0, T] \).

When \( f_1 = f_2 = f_3 = g_3 = 0 \), this problem has been studied alternatively. Let us recall some works. In [3], Park-Sowers-Namachchivaya considered the filtering problem with a two-dimensional plant and a one-dimensional observation process. There they used the time change and decomposition methods. And for the high dimension case, Park-Namachchivaya-Yeong [7] presented a numerical algorithm method. Later, Imkeller-Namachchivaya-Perkowski-Yeong [2] showed that for the high dimension slow-fast dynamical system [1], the filter \( \mathbb{E}[F(X^\varepsilon_t) | \mathcal{F}^Y_t] \) converges to the homogenized filter (See Section 4) by double backward stochastic differential equations and asymptotic techniques.

When \( f_1 \neq 0, f_2 \neq 0, f_3 = g_3 = 0 \), Kushner [4] studied this problem by a weak convergence method.

In the paper, we observe this problem with \( f_1 \neq 0, f_2 \neq 0, f_3 \neq 0, g_3 \neq 0 \), i.e. multiscale non-Gaussian signal processes and observation processes with jumps. Firstly, the dimension for the slow-fast system is proved to be reduced. Secondly, convergence of the corresponding nonlinear filtering to the homogenized filtering is shown.

It is worthwhile to mention our methods. By a martingale problem method we reduce the dimension of the slow-fast system. For the filtering problem for the slow component \( X^\varepsilon_t \) with respect to \( \{Y^\varepsilon_s, 0 \leq s \leq t\} \), since the time change is only useful for a one-dimensional process, and the theory for double backward stochastic differential equations with jumps is short, these techniques are not applied to the present case. Here we compute the difference between \( \mathbb{E}[F(X^\varepsilon_t) | \mathcal{F}^Y_t] \) and the homogenized filter and then convert it to the difference between two unnormalized filterings. With the help of the weak convergence method in [4], we know that the difference between two unnormalized filterings converges to zero. Thus, we prove that \( \mathbb{E}[F(X^\varepsilon_t) | \mathcal{F}^Y_t] \) converges weakly to the homogenized filter.

The paper is arranged as follows. In the next section, we introduce some notation, terminology and concepts used in the sequel. The dimension reducing for the slow-fast system is placed in Section 3. In Section 4, nonlinear filtering problems are introduced. And convergence of the corresponding nonlinear filtering to the homogenized filtering is proved in Section 5.
The following convention will be used throughout the paper: $C$ with or without indices will denote different positive constants (depending on the indices) whose values may change from one place to another.

2. Preliminary

In the section, we introduce some notation, terminology and concepts used in the sequel. Firstly, introduce the following notation and terminology:

(i) For a separable metric space $E$, let $\mathcal{B}(E)$ denote the Borel $\sigma$-algebra on $E$ and $B(E)$ denote the set of all real-valued uniformly bounded Borel-measurable mappings on $E$. Also let $C(E)$ be the set of all real-valued continuous functions on $E$, put $\bar{C}(E) := B(E) \cap C(E)$, and let $C_c(E)$ be the set of all members of $\bar{C}(E)$ which have compact support. When $E$ is locally compact, let $\hat{C}(E)$ be the collection of all members of $C(E)$ which vanish at infinity.

(ii) For a positive integer $r$, let $C^r(\mathbb{R}^q)$ denote the collection of all members of $C(\mathbb{R}^q)$ with continuous derivatives of each order, up to and including $r$. Let $C_c^\infty(\mathbb{R}^q)$ denote the collection of all members of $C(\mathbb{R}^q)$ with continuous derivatives of all orders and compact support. For $E$ a metric space and $r$ some positive integer, write $C^{r,0}(\mathbb{R}^q \times E)$ for the collection of all mappings $f \in C(\mathbb{R}^q \times E)$ whose partial derivatives of every order up to and including $r$, with respect to its first $q$ real-valued arguments, exist and are members of $C(\mathbb{R}^q \times E)$, and put $C^{r,0}_c(\mathbb{R}^q \times E) := C^{r,0}(\mathbb{R}^q \times E) \cap C_c(\mathbb{R}^q \times E)$.

(iii) When $E$ is a complete separable metric space, let $\mathcal{P}(E)$ denote the collection of all probability measures on the measurable space $(E, \mathcal{B}(E))$ with the usual topology of weak (or narrow) convergence; and if $X : (\Omega, \mathcal{F}, \mathbb{P}) \mapsto E$ is $\mathcal{F}/\mathcal{B}(E)$-measurable, then let $L(X)$ be the distribution of $X$ on $(E, \mathcal{B}(E))$. Also, for a $\mathcal{B}(E)$-measurable mapping $f : E \mapsto \mathbb{R}$ which is integrable with respect to $\mu \in \mathcal{P}(E)$, we put $\mu f := \int_E f \mu$.

Secondly, we introduce some concepts. Suppose that $E$ is a separable metric space.

**Definition 2.1.** Let $\mathcal{A} \subset B(E) \times B(E)$ be a relation with domain $\mathcal{D}(\mathcal{A})$, and let $\mu \in \mathcal{P}(E)$. Then a progressively measurable solution of the martingale problem for $\mathcal{A}$ (for $(\mathcal{A}, \mu)$) is some pair $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}), (\tilde{X}_t)\}$, in which $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ is a complete filtered probability space and $(\tilde{X}_t)$ is an $E$-valued $\tilde{\mathcal{F}}_t$-progressively measurable process such that $f(\tilde{X}_t) - \int_0^t \mathcal{A} f(\tilde{X}_s) ds$ is an $\tilde{\mathcal{F}}_t$-martingale for each $f \in \mathcal{D}(\mathcal{A})$ (and $L(\tilde{X}_0) = \mu$). The martingale problem for $(\mathcal{A}, \mu)$ has the property of existence when there exists some progressively measurable solution of the martingale problem for $(\mathcal{A}, \mu)$, and has the property of uniqueness when, given any two progressively measurable solutions $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}), (\tilde{X}_t)\}$ and $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \hat{\mathbb{P}}), (\hat{X}_t)\}$ of the martingale problem for $(\mathcal{A}, \mu)$, the $E$-valued processes $\tilde{X}$ and $\hat{X}$ necessarily have identical finite-dimensional distributions. The martingale problem for $(\mathcal{A}, \mu)$ is called well-posed when it has the properties of both existence and uniqueness. Finally, the martingale problem for $\mathcal{A}$ is well-posed when the martingale problem for $(\mathcal{A}, \mu)$ is well-posed for each $\mu \in \mathcal{P}(E)$. 

3. Convergence of some processes

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

We make the following assumptions, in order to guarantee existence and uniqueness of the solution for the system (1).

**Assumption 1.**
For \( x_1, x_2 \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^m \), there exists a \( L_1 > 0 \) such that

\[
|b_1(x_1, z_1) - b_1(x_2, z_2)| \leq L_1(|x_1 - x_2| + |z_1 - z_2|),
\]

\[
\|\sigma_1(x_1, z_1) - \sigma_1(x_2, z_2)\| \leq L_1(|x_1 - x_2|^2 + |z_1 - z_2|^2),
\]

\[
\int_{U_1} |f_1(x_1, u) - f_1(x_2, u)|^2 \nu_1(du) \leq L_1|x_1 - x_2|^2,
\]

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

For \( x \in \mathbb{R}^n, z \in \mathbb{R}^m \), there exists a \( L_2 > 0 \) such that

\[
|b_1(x, z)|^2 + \|\sigma_1(x, z)\|^2 + \int_{U_1} |f_1(x, u)|^2 \nu_1(du) \leq L_2.
\]

For \( x_1, x_2 \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^m \), there exists a \( L_3 > 0 \) such that

\[
|b_2(x_1, z_1) - b_2(x_2, z_2)| \leq L_3(|x_1 - x_2| + |z_1 - z_2|),
\]

\[
\|\sigma_2(x_1, z_1) - \sigma_2(x_2, z_2)\| \leq L_3(|x_1 - x_2|^2 + |z_1 - z_2|^2),
\]

\[
\int_{U_2} |f_2(x_1, z_1, u) - f_2(x_2, z_2, u)|^2 \nu_2(du) \leq L_3(|x_1 - x_2|^2 + |z_1 - z_2|^2).
\]

Under Assumption 1, by Theorem 1.2 in [10], the system (1) has a unique strong solution denoted by \((X_t^x, Z_t^z)\). Moreover, the infinitesimal generator of the system (1) is given by

\[
(L^c H)(x, z) = (L^{X^c} H)(x, z) + (L^{Z^c} H)(x, z), \quad H \in D(L^c),
\]

where

\[
(L^{X^c} H)(x, z) := \frac{\partial H(x, z)}{\partial x_i} b_1^i(x, z) + \frac{1}{2} \frac{\partial^2 H(x, z)}{\partial x_i \partial x_j} (\sigma_1 \sigma_1^T)^{ij}(x, z)
\]

\[
+ \int_{U_1} \left[ H(x + f_1(x, u), z) - H(x, z) - \frac{\partial H(x, z)}{\partial x_i} f_1^i(x, u) \right] \nu_1(du),
\]

and

\[
(L^{Z^c} H)(x, z) := \int_0^1 \frac{\partial H(x, z)}{\partial z_i} b_2^i(x, z) + \frac{1}{2} \frac{\partial^2 H(x, z)}{\partial z_i \partial z_j} (\sigma_2 \sigma_2^T)^{ij}(x, z)
\]

\[
+ \int_{U_2} \left[ H(x, z + f_2(x, z, u)) - H(x, z) - \frac{\partial H(x, z)}{\partial z_i} f_2^i(x, z, u) \right] \nu_2(du).
\]

Here and hereafter, we use the convention that repeated indices imply summation.

Next take any \( x \in \mathbb{R}^n \) and fix it. And consider the following SDE in \( \mathbb{R}^m \):

\[
\begin{cases}
    dZ_t^x = b_2(x, Z_t^x)dt + \sigma_2(x, Z_t^x)dW_t + \int_{U_2} f_2(x, Z_t^x, u)\tilde{N}_{p_2}(dt, du), \\
    Z_0^x = z_0,
\end{cases}
\quad t \geq 0.
\]

Under the assumption (\( H_{b_2, \sigma_2, f_2} \)), 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 \( t > 0 \) and \( A \in \mathcal{B}(\mathbb{R}^m) \). Set \( (T_t \varphi)(z_0) := \int_{\mathbb{R}^m} \varphi(z') p(x; z_0, t, dz') \) for any \( \varphi \in C(\mathbb{R}^m) \), and then \( \{ T_t, t \geq 0 \} \) is its transition semigroup and \( \varepsilon L^{Z^c} \) is its infinitesimal generator. For \( Z_t^x \), we assume:
Assumption 2. There exists a unique invariant probability measure $\bar{\mu}(x; dz)$ for $Z^x_t$ and
\[ \int_0^\infty \left| \int_{\mathbb{R}^n} \varphi(z') p(x; z, s, dz') - \int_{\mathbb{R}^n} \varphi(z') \bar{\mu}(x; dz') \right| ds < \infty \]
for any $\varphi \in C(\mathbb{R}^m)$.

About conditions for existence of a unique invariant probability measure for $Z^x_t$, please refer to \cite{9}. Define an operator $\bar{L}$ as follows:
\[ \mathcal{D}(\bar{L}) := C_c^\infty(\mathbb{R}^n), \]
\[ (\bar{L}g)(x) := \int_{\mathbb{R}^m} (\mathcal{L}^x g)(x, z) \bar{\mu}(x; dz) \]
\[ = \frac{\partial g(x)}{\partial x_i} \bar{b}_1(x) + \frac{1}{2} \frac{\partial^2 g(x)}{\partial x_i \partial x_j} (\sigma_1 \sigma_1^T)^{ij}(x) \]
\[ + \int_{\mathcal{U}_1} \left[ g(x + f_1(x, u)) - g(x) - \frac{\partial g(x)}{\partial x_i} f_1^i(x, u) \right] \nu_1(du), \quad g \in \mathcal{D}(\bar{L}), \]
where
\[ \bar{b}_1(x) := \int_{\mathbb{R}^m} b_1(x, z) \bar{\mu}(x, dz), \quad (\sigma_1 \sigma_1^T)(x) := \int_{\mathbb{R}^m} (\sigma_1 \sigma_1^T)(x, z) \bar{\mu}(x, dz). \]

It is clear that $\bar{L}$ is a diffusion operator. So, we could construct a SDE generated by $\bar{L}$ on the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \bar{\mathbb{P}})$ as follows:
\[ \begin{cases} 
\mathrm{d}X^0_t = \bar{b}_1(X^0_t) dt + \bar{\sigma}_1(X^0_t) d\bar{V}_t + \int_{\mathcal{U}_1} f_1(X^0_t, u) \bar{N}(dt, du), \\
X^0_0 = x_0,
\end{cases} \quad 0 \leq t \leq T, \tag{2} \]
where $\bar{V}$ is a $l$-dimensional standard Brownian motion, and $\bar{N}(dt, du)$ is a Poisson random measure with the characteristic measure $\nu_1$ and $\bar{N}(dt, du) = N(dt, du) - \nu_1(du)dt$. For the operator $\bar{L}$, we make the following requirement.

Assumption 3. The martingale problem for $(\bar{L}, \delta_{x_0})$ is well-posed.

Theorem 3.1. Under all the above hypotheses $\{X^\varepsilon_t, t \in [0, T]\}$ converges weakly to $\{X^0_t, t \in [0, T]\}$ in $D([0, T], \mathbb{R}^n)$.

Proof. Step 1. We prove that $\{X^\varepsilon_t, t \in [0, T]\}$ is relatively weakly compact in $D([0, T], \mathbb{R}^n)$.

Firstly of all, consider the martingale problem associated with $\mathcal{L}^\varepsilon$. For $H \in \mathcal{D}(\mathcal{L}^\varepsilon)$,
\[ M_H(t) := H(X^\varepsilon_t, Z^\varepsilon_t) - H(x_0, z_0) - \int_0^t (\mathcal{L}^\varepsilon H)(X^\varepsilon_s, Z^\varepsilon_s) ds \tag{3} \]
is a square integrable martingale and
\[ \langle M_H(\cdot), M_H(\cdot) \rangle_t = \int_0^t \frac{\partial H(X^\varepsilon_s, Z^\varepsilon_s)}{\partial x_i} \frac{\partial H(X^\varepsilon_s, Z^\varepsilon_s)}{\partial z_j} (\sigma_1 \sigma_1^T)^{ij}(X^\varepsilon_s, Z^\varepsilon_s) ds \]
\[ + \int_0^t \int_{\mathcal{U}_1} [H(X^\varepsilon_s + f_1(X^\varepsilon_s, u), Z^\varepsilon_s) - H(X^\varepsilon_s, Z^\varepsilon_s)] \nu_1(du) ds \]
\[ + \frac{1}{\varepsilon} \int_0^t \frac{\partial H(X^\varepsilon_s, Z^\varepsilon_s)}{\partial z_i} \frac{\partial H(X^\varepsilon_s, Z^\varepsilon_s)}{\partial z_j} (\sigma_2 \sigma_2^T)^{ij}(X^\varepsilon_s, Z^\varepsilon_s) ds \]
where the last inequality is based on the condition (H and the Itô isometry, it holds that for any $E \in \mathcal{F}_t$ independent of $\epsilon$ and the stopping time $\tau\leq T$. Let $\tau$ be any $(\mathcal{F}_t)_{t\geq 0}$-stopping time no more than $T$. And then the Prohorov theorem admits us to obtain that

$$\sup_{t \leq T} \mathbb{E}[|X_t - X_t^\epsilon|^2] < \infty.$$ 

By the similar deduction to above, one could furthermore get

$$\mathbb{E}[|X_{\tau+\delta} - X_{\tau}^\epsilon|^2] = 0.$$

Thus, it follows from Theorem 2.7 in [3] that $\{X_t^\epsilon, t \in [0, T]\}$ is tight in $D([0, T], \mathbb{R}^n)$. And then the Prohorov theorem admits us to obtain that $\{X_t^\epsilon, t \in [0, T]\}$ is relatively weakly compact in $D([0, T], \mathbb{R}^n)$.

**Step 2.** We prove that the weak limit of $\{X_t^\epsilon, t \in [0, T]\}$ is $\{X_t^0, t \in [0, T]\}$.

Taking $H(x, z) = g(x)$ in (3), where $g$ is a smooth and bounded function, we have that

$$g(X_t^\epsilon) - g(X_s^\epsilon) - \int_s^t (\mathcal{L}X^\epsilon_g)(X_r^\epsilon, Z_r^\epsilon)dr = M_g(t) - M_g(s),$$

and then

$$g(X_t^\epsilon) - g(X_s^\epsilon) - \int_s^t (\mathcal{L}g)(X_r^\epsilon)dr = \int_s^t [(\mathcal{L}X^\epsilon_g)(X_r^\epsilon, Z_r^\epsilon) - (\mathcal{L}g)(X_r^\epsilon)] dr + M_g(t) - M_g(s).$$
Moreover, multiplying a bounded $\mathcal{F}_s$-measurable functional $\chi_s$ of the process $\{X_t^\epsilon, t \in [0, T]\}$ and taking expectation on the two hand sides of the above equality, we know
\[
\mathbb{E} \left[ \chi_s \left( g(X_t^\epsilon) - g(X_s^\epsilon) - \int_s^t (\tilde{L}g)(X_r^\epsilon) dr \right) \right] = \mathbb{E} \left[ \chi_s \int_s^t \left[ (\mathcal{L}^X g)(X_r^\epsilon, Z_r^\epsilon) - (\tilde{L}g)(X_r^\epsilon) \right] dr \right].
\]  
(4)

Next we compute $\lim_{\epsilon \downarrow 0} \mathbb{E} \left[ \chi_s \int_s^t \left[ (\mathcal{L}^X g)(X_r^\epsilon, Z_r^\epsilon) - (\tilde{L}g)(X_r^\epsilon) \right] dr \right]$. On one hand, set
\[
\Psi(x, z, A) := \int_0^\infty [p(x, z, t, A) - \tilde{p}(x; A)] dt,
\]
and then
\[
\Psi_g(x, z) := \int_{\mathbb{R}^m} \left[ (\mathcal{L}^X g)(x, z') - (\tilde{L}g)(x) \right] [p(x, z, t, d\nu') - \tilde{p}(x; d\nu')] dt
\]
\[
= \int_0^\infty \left( \int_{\mathbb{R}^m} \left[ (\mathcal{L}^X g)(x, z') - (\tilde{L}g)(x) \right] [p(x, z, t, d\nu') - \tilde{p}(x; d\nu')] \right) dt
\]
\[
= \int_0^\infty T_t[(\mathcal{L}^X g) - (\tilde{L}g)](x, z) dt.
\]
Furthermore, it holds that
\[
\epsilon(\mathcal{Z}^\epsilon \Psi_g)(x, z) = \int_0^\infty (\epsilon(\mathcal{Z}^\epsilon T_t)[(\mathcal{L}^X g) - (\tilde{L}g)](x, z) dt
\]
\[
= \int_0^\infty \frac{dT_t[(\mathcal{L}^X g) - (\tilde{L}g)](x, z)}{dt} dt
\]
\[
= \lim_{t \to \infty} T_t[(\mathcal{L}^X g) - (\tilde{L}g)](x, z) - [(\mathcal{L}^X g)(x, z) - (\tilde{L}g)(x)]
\]
\[
= \lim_{t \to \infty} \int_{\mathbb{R}^m} \left[ (\mathcal{L}^X g)(x, z') - (\tilde{L}g)(x) \right] [p(x, z, t, d\nu') - \tilde{p}(x; d\nu')] dt - [(\mathcal{L}^X g)(x, z) - (\tilde{L}g)(x)]
\]
\[
= -[(\mathcal{L}^X g)(x, z) - (\tilde{L}g)(x)],
\]
(5)

where the last equality is based on Assumption 2. On the other hand, taking $H(x, z) = \epsilon(\Psi_g)(x, z)$ again in (3), we get that
\[
\epsilon(\Psi_g)(X_t^\epsilon, Z_t^\epsilon) - \epsilon(\Psi_g)(X_s^\epsilon, Z_s^\epsilon) - \epsilon \int_s^t (\mathcal{L}^X \Psi_g)(X_r^\epsilon, Z_r^\epsilon) dr
\]
\[
= \int_s^t \epsilon(\mathcal{Z}^\epsilon \Psi_g)(X_r^\epsilon, Z_r^\epsilon) dr + M_{\epsilon\Psi_g}(t) - M_{\epsilon\Psi_g}(s).
\]
So, by multiplying $\chi_s$ and taking expectation on the two hand sides of the above equality, it holds that
\[
\epsilon \mathbb{E} \left[ \chi_s \left( \Psi_g(X_t^\epsilon, Z_t^\epsilon) - \Psi_g(X_s^\epsilon, Z_s^\epsilon) - \int_s^t (\mathcal{L}^X \Psi_g)(X_r^\epsilon, Z_r^\epsilon) dr \right) \right] = \mathbb{E} \left[ \chi_s \int_s^t \epsilon(\mathcal{Z}^\epsilon \Psi_g)(X_r^\epsilon, Z_r^\epsilon) dr \right].
\]
\[
\begin{align*}
&\quad = -\mathbb{E}\left[\chi_s \int_s^t \left[ (\mathcal{L}X^\varepsilon)(X^\varepsilon_s, Z^\varepsilon_s) - (\bar{\mathcal{L}}g)(X^\varepsilon_t) \right] \, dr \right],
\end{align*}
\]
where the last equality is based on (5). As \( \varepsilon \to 0 \), it is easy to see that
\[
\lim_{\varepsilon \downarrow 0} \mathbb{E}\left[\chi_s \int_s^t \left[ (\mathcal{L}X^\varepsilon)(X^\varepsilon_s, Z^\varepsilon_s) - (\bar{\mathcal{L}}g)(X^\varepsilon_t) \right] \, dr \right] = 0.
\]
The above limit, together with (4), yields that
\[
\lim_{\varepsilon \downarrow 0} \mathbb{E}\left[\chi_s \left( g(X^\varepsilon_t) - g(X^\varepsilon_s) - \int_s^t (\bar{\mathcal{L}}g)(X^\varepsilon_r) \, dr \right) \right] = 0,
\]
which means that the weak limit of \( \{X^\varepsilon_t, t \in [0, T]\} \) is a solution of the martingale problem
for \( (\bar{\mathcal{L}}, \delta_{x_0}) \). By Assumption 3., the weak limit of \( \{X^\varepsilon_t, t \in [0, T]\} \) is
\( \{X^0_t, t \in [0, T]\} \). □

4. NONLINEAR FILTERING PROBLEMS

In the section, we study nonlinear filtering problems for \( X^\varepsilon \) and \( X^0 \).
For \( Y^\varepsilon \), we assume:

**Assumption 4.** \( h \) is bounded and
\[
\int_0^T \int_{U_3} |f_3(s, u)|^2 \nu_3(du) ds < \infty.
\]

Under Assumption 4., \( Y^\varepsilon \) is well defined. Set
\[
(\Lambda^\varepsilon_t)^{-1} : = \exp \left\{ - \int_0^t h(X^\varepsilon_s, Z^\varepsilon_s) dB^\varepsilon_s - \frac{1}{2} \int_0^t |h(X^\varepsilon_s, Z^\varepsilon_s)|^2 ds 
\right.
\]
\[
- \int_0^t \int_{U_3} \log \lambda(s, X^\varepsilon_s, u) N_\lambda(ds, du) - \int_0^t \int_{U_3} (1 - \lambda(s, X^\varepsilon_s, u)) \nu_3(du) ds \right\}.
\]

**Assumption 5.** There exists a positive function \( L(u) \) satisfying
\[
\int_{U_3} \frac{(1 - L(u))^2}{L(u)} \nu_3(du) < \infty
\]
such that \( 0 < l \leq L(u) < \lambda(t, x, u) < 1 \) for \( u \in U_3 \), where \( l \) is a constant.

Under Assumption 5., it holds that
\[
\mathbb{E} \left[ \exp \left\{ \int_0^T \int_{U_3} \frac{(1 - \lambda(s, X^\varepsilon_s, u))^2}{\lambda(s, X^\varepsilon_s, u)} \nu_3(du) ds \right\} \right]
\]
\[
< \exp \left\{ \int_0^T \int_{U_3} \frac{(1 - L(u))^2}{L(u)} \nu_3(du) ds \right\}
\]
\[
< \infty.
\]
Thus, by the same deduction to that in [11], we know that \( (\Lambda^\varepsilon_T)^{-1} \) is an exponential martingale. By use of \( (\Lambda^\varepsilon_T)^{-1} \), one could define a measure \( \mathbb{P}^\varepsilon \) via
\[
\frac{d\mathbb{P}^\varepsilon}{d\mathbb{P}} = (\Lambda^\varepsilon_T)^{-1}.
\]
By the Girsanov theorem for Brownian motions and random measures, we can obtain that under the measure $\mathbb{P}^\varepsilon$, $\tilde{B}_t := B_t + \int_0^t h(X_s^\varepsilon, Z_s^\varepsilon)ds$ is a Brownian motion and $N_\lambda((0, t], du)$ is a Poisson random measure with the predictable compensator $\pi$.

Next, rewrite

$$\Lambda_t^\varepsilon = \exp \left\{ \int_0^t h(X_s^\varepsilon, Z_s^\varepsilon) i \, dB_s^i - \frac{1}{2} \int_0^t |h(X_s^\varepsilon, Z_s^\varepsilon)|^2 \, ds \right. \left. + \int_0^t \int_{\mathbb{U}_3} \log \lambda(s, X_{s-}^\varepsilon, u) N_\lambda(ds, du) + \int_0^t \int_{\mathbb{U}_3} (1 - \lambda(s, X_s^\varepsilon, u)) \nu_3(du)ds \right\},$$

and define

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

where $\mathbb{E}^{\mathbb{P}^\varepsilon}$ denotes the expectation under the measure $\mathbb{P}^\varepsilon$ and $\mathcal{F}_t^{Y^\varepsilon}$ stands for the $\sigma$-algebra generated by $\{Y_s^\varepsilon, 0 \leq s \leq t \}$. Again set

$$\pi_t^\varepsilon(\psi) := \mathbb{E}[\psi(X_t^\varepsilon) | \mathcal{F}_t^{Y^\varepsilon}],$$

and by the Kallianpur-Striebel formula it holds that

$$\pi_t^\varepsilon(\psi) = \frac{\rho_t^\varepsilon(\psi)}{\rho_t^\varepsilon(1)}.$$

Set

$$\tilde{h}(x) := \int_{\mathbb{R}^m} h(x, z) \bar{p}(x, dz),$$

and then $\tilde{h}$ is an averaged version of $h$. So, we make use of $\tilde{h}$ to define

$$\bar{\Lambda}_t := \exp \left\{ \int_0^t \tilde{h}(X_s^0)^i dB_s^i - \frac{1}{2} \int_0^t |\tilde{h}(X_s^0)|^2 \, ds \right. \left. + \int_0^t \int_{\mathbb{U}_3} \log \lambda(s, X_{s-}^0, u) N_\lambda(ds, du) + \int_0^t \int_{\mathbb{U}_3} (1 - \lambda(s, X_s^0, u)) \nu_3(du)ds \right\},$$

$$\rho_t^0(\psi) := \mathbb{E}^{\mathbb{P}^0}[\psi(X_t^0) | \mathcal{F}_t^{Y^\varepsilon}],$$

where $X_t^0$ is the limit process in Section 3. Put

$$\pi_t^0(\psi) := \frac{\rho_t^0(\psi)}{\rho_t^0(1)},$$

and then we study the relation between $\pi_t^0$ and $\pi_t^\varepsilon$ as $\varepsilon \to 0$ in the next section.

At the first look, it is more reasonable to define the limit observable process

$$Y_t^0 := \int_0^t \tilde{h}(X_s^0) ds + B_t + \int_0^t \int_{\mathbb{U}_3} f_3(s, u) \bar{N}_\lambda(ds, du) + \int_0^t \int_{\mathbb{U}_3} g_3(s, u) N_\lambda(ds, du),$$

where $\bar{N}_\lambda((0, t], du)$ is a Poisson random measure with a predictable compensator $\lambda(t, X_t^0, u) tv_3(du)$, and the corresponding nonlinear filtering

$$\mathbb{P}_t^0(\psi) := \mathbb{E}[\psi(X_t^0) | \mathcal{F}_t^{Y^0}],$$

and discuss the relation between $\mathbb{P}_t^0$ and $\pi_t^\varepsilon$ as $\varepsilon \to 0$. In fact, since $X_t^0$ couldn’t be obtained genuinely, $Y_t^0$ is not observable. However, should such homogenized observation be available, using it would lead to loss of information for estimating the signal compared to using the actual observation. Therefore, we only consider $X_t^0$ under $\mathcal{F}_t^{Y^\varepsilon}$. 
5. CONVERGENCE OF NONLINEAR FILTERINGS

In the section, we prove that $\pi_t^\varepsilon$ converges weakly to $\pi_t^0$ as $\varepsilon \to 0$ for any $t \in [0,T]$. Firstly, let us prove two key lemmas.

**Lemma 5.1.** Suppose that $h, \lambda$ satisfy Assumption 4-5.. Then $(\rho_t^0(1))^{-1} < \infty \ P \text{ a.s. for any } t \in [0,T]$.

**Proof.** By the H"older inequality, it holds that

$$\mathbb{E}(\rho_t^0(1))^{-1} = \mathbb{E}_{\rho_t^0} (\rho_t^0(1))^{-1} \Lambda_T \leq (\mathbb{E}_{\rho_t^0} (\rho_t^0(1))^{-2})^{1/2} (\mathbb{E}_{\rho_t^0} (\Lambda_T^2))^{1/2}.$$

Let us firstly estimate $\mathbb{E}_{\rho_t^0} (\rho_t^0(1))^{-2}$. Note that $\rho_t^0(1) = \mathbb{E}_{\rho_t^0} [\Lambda_t | \mathcal{F}_t^Y]$ and $x \mapsto x^{-2}$ is convex. Thus, we know by the Jensen inequality that

$$\mathbb{E}_{\rho_t^0} (\rho_t^0(1))^{-2} = \mathbb{E}_{\rho_t^0} (\mathbb{E}_{\rho_t^0} [\Lambda_t | \mathcal{F}_t^Y])^{-2} \leq \mathbb{E}_{\rho_t^0} (\mathbb{E}_{\rho_t^0} [(\Lambda_t)^{-2} | \mathcal{F}_t^Y]) = \mathbb{E}_{\rho_t^0} (\Lambda_t)^{-2}.$$

So, we estimate $\mathbb{E}_{\rho_t^0} (\Lambda_t)^{-2}$. Applying the Itô formula to $(\Lambda_t)^{-1}$, one could obtain that

$$(\Lambda_t)^{-1} = 1 + \int_0^t (\Lambda_s)^{-1} |\dot h(X_s^0)|^2 ds + \int_0^t \int_{\mathcal{U}_3} (\Lambda_s)^{-1} \frac{(1 - \lambda(s, X_s^0, u))^2}{\lambda(s, X_s^0, u)} \nu_3(du)ds$$

$$- \int_0^t (\Lambda_s)^{-1} h(X_s^0) \dot B_s^i + \int_0^t (\Lambda_s)^{-1} h(X_s^0) \lambda(s, X_s^0, u) \tilde{N}_\lambda(ds, du).$$

Furthermore, it follows from the H"older inequality and the Itô isometry that

$$\mathbb{E}_{\rho_t^0} (\Lambda_t)^{-2} \leq 5 + 5\mathbb{E}_{\rho_t^0} \left| \int_0^t (\Lambda_s)^{-1} |\dot h(X_s^0)|^2 ds + \int_0^t (\Lambda_s)^{-1} h(X_s^0) \dot B_s^i \right|^2$$

$$+ 5\mathbb{E}_{\rho_t^0} \left| \int_0^t \int_{\mathcal{U}_3} (\Lambda_s)^{-1} \frac{(1 - \lambda(s, X_s^0, u))^2}{\lambda(s, X_s^0, u)} \nu_3(du)ds \right|^2$$

$$+ 5\mathbb{E}_{\rho_t^0} \left| \int_0^t \int_{\mathcal{U}_3} (\Lambda_s)^{-1} \frac{1 - \lambda(s, X_s^0, u)}{\lambda(s, X_s^0, u)} \tilde{N}_\lambda(ds, du) \right|^2$$

$$\leq 5 + 5T \mathbb{E}_{\rho_t^0} \int_0^t (\Lambda_s)^{-2} |\dot h(X_s^0)|^4 ds + 5\mathbb{E}_{\rho_t^0} \int_0^t (\Lambda_s)^{-2} |\dot h(X_s^0)|^2 ds$$

$$+ 5T \mathbb{E}_{\rho_t^0} \int_0^t (\Lambda_s)^{-2} \left| \int_{\mathcal{U}_3} \frac{(1 - \lambda(s, X_s^0, u))^2}{\lambda(s, X_s^0, u)} \nu_3(du)ds \right|^2 ds$$

$$+ 5\mathbb{E}_{\rho_t^0} \int_0^t \int_{\mathcal{U}_3} (\Lambda_s)^{-2} \frac{1 - \lambda(s, X_s^0, u)}{\lambda(s, X_s^0, u)} \tilde{N}_\lambda(ds, du) ds$$

$$\leq 5 + C \int_0^t \mathbb{E}_{\rho_t^0} (\Lambda_t)^{-2} ds,$$

where the last step is based on Assumption 4-5.. The Gronwall inequality admits us to have $\mathbb{E}_{\rho_t^0} (\Lambda_t)^{-2} < \infty$.

Next, deal with $\mathbb{E}_{\rho_t^0} (\Lambda_T^2)$. Applying the Itô formula to $\Lambda_T^\varepsilon$, we obtain that

$$\Lambda_T^\varepsilon = 1 + \int_0^t \Lambda_s^\varepsilon h(X_s^\varepsilon, Z_s^\varepsilon) \dot B_s^i + \int_0^t \int_{\mathcal{U}_3} \Lambda_s^\varepsilon (\lambda(s, X_s^\varepsilon, u) - 1) \tilde{N}_\lambda(ds, du).$$

(6)

Thus, by the similar deduction to $\mathbb{E}_{\rho_t^0} (\Lambda_t)^{-2}$ it holds that $\mathbb{E}_{\rho_t^0} (\Lambda_T^2)^2 < \infty$.

In conclusion, $\mathbb{E}(\rho_t^0(1))^{-1} < \infty$. The proof is completed. \qed
Lemma 5.2. Under Assumption 4-5., \{\rho^\varepsilon_t, t \in [0, T]\} is relatively weakly compact in \(D([0, T], \mathcal{M}(\mathbb{R}^n))\).

Proof. First of all, we explain \(\rho^\varepsilon_t \in \mathcal{M}(\mathbb{R}^n)\). Note that \(\rho^\varepsilon_t(\mathbb{R}^n) = \rho^\varepsilon_t(1_{\mathbb{R}^n}) = \rho^\varepsilon_t(1) = \mathbb{E}^\varepsilon[\Lambda_t^\varepsilon | \mathcal{F}^\varepsilon_t]\). And then by the Hölder inequality, it holds that
\[
\mathbb{E}^\varepsilon(\rho^\varepsilon_t(1) = \mathbb{E}^\varepsilon[\rho^\varepsilon_t(1)\Lambda_t^\varepsilon] \leq (\mathbb{E}^\varepsilon((\rho^\varepsilon_t(1))^2)^{1/2} (\mathbb{E}^\varepsilon(\Lambda_t^\varepsilon)^2)^{1/2}.
\]

On one hand, the Jensen inequality admits us to obtain that
\[
\mathbb{E}^\varepsilon((\rho^\varepsilon_t(1))^2 = \mathbb{E}^\varepsilon(\mathbb{E}^\varepsilon[\Lambda_t^\varepsilon | \mathcal{F}^\varepsilon_t]^2) \leq \mathbb{E}^\varepsilon[\mathbb{E}^\varepsilon[(\Lambda_t^\varepsilon)^2 | \mathcal{F}^\varepsilon_t]] = \mathbb{E}^\varepsilon(\Lambda_t^\varepsilon)^2.
\]

By the proof of Lemma 5.1 one could get \(\mathbb{E}^\varepsilon((\rho^\varepsilon_t(1))^2 < \infty\). On the other hand, it follows from the proof of Lemma 5.1 that \(\mathbb{E}^\varepsilon(\Lambda_t^\varepsilon)^2 < \infty\). Thus, \(\rho^\varepsilon_t(\mathbb{R}^n) < \infty \text{ a.s. } \mathbb{P}\). Other measure properties of \(\rho^\varepsilon_t\) are easy to justify by means of properties of conditional expectations.

Next, we deduce the equation for \(\rho^\varepsilon_t\). For \(\psi \in C^2_b(\mathbb{R}^n)\), applying the Itô formula to \(\psi(X^\varepsilon_t)\), we have that
\[
\psi(X^\varepsilon_t) = \psi(X^\varepsilon_0) + \int_0^t (\mathcal{L}X^\varepsilon_t)\psi(X^\varepsilon_s, Z^\varepsilon_s)ds + \int_0^t \nabla \psi(X^\varepsilon_s, Z^\varepsilon_s)\sigma_1(X^\varepsilon_s, Z^\varepsilon_s)dV_s
\]
\[
+ \int_0^t \int_{\mathcal{U}_1} \psi(X^\varepsilon_{s-} + f_1(X^\varepsilon_{s-}, u)) - \psi(X^\varepsilon_{s-})]dN_{p_1}(ds, du).
\]

Note that \(\Lambda_t^\varepsilon\) satisfies Eq. (6). So, it follows from the Itô formula that
\[
\psi(X^\varepsilon_t)\Lambda_t^\varepsilon = \psi(X^\varepsilon_0) + \int_0^t \psi(X^\varepsilon_s)\Lambda_t^\varepsilon h(X^\varepsilon_s, Z^\varepsilon_s)d\bar{B}_s^i
\]
\[
+ \int_0^t \int_{\mathcal{U}_3} \psi(X^\varepsilon_{s-})\Lambda_t^\varepsilon (\lambda(s, X^\varepsilon_{s-}, u) - 1)\bar{N}_{\lambda}(ds, du)
\]
\[
+ \int_0^t \int_{\mathcal{U}_1} \Lambda_t^\varepsilon (\mathcal{L}X^\varepsilon_t)\psi(X^\varepsilon_s, Z^\varepsilon_s)ds + \int_0^t \Lambda_t^\varepsilon \nabla \psi(X^\varepsilon_s, Z^\varepsilon_s)dV_s
\]
\[
+ \int_0^t \int_{\mathcal{U}_1} \Lambda_t^\varepsilon [\psi(X^\varepsilon_{s-} + f_1(X^\varepsilon_{s-}, u)) - \psi(X^\varepsilon_{s-})]dN_{p_1}(ds, du).
\]

Taking the conditional expectation with respect to \(\mathcal{F}^\varepsilon_t\) under \(\mathbb{P}\) on two hand sides of the above equality, one could obtain that
\[
\mathbb{E}^\varepsilon[\psi(X^\varepsilon_t)\Lambda_t^\varepsilon | \mathcal{F}^\varepsilon_t] = \mathbb{E}^\varepsilon[\psi(X^\varepsilon_0) | \mathcal{F}^\varepsilon_0] + \int_0^t \mathbb{E}^\varepsilon[\psi(X^\varepsilon_s)\Lambda_t^\varepsilon h(X^\varepsilon_s, Z^\varepsilon_s)d\bar{B}_s^i
\]
\[
+ \int_0^t \int_{\mathcal{U}_3} \mathbb{E}^\varepsilon[\psi(X^\varepsilon_{s-})\Lambda_t^\varepsilon (\lambda(s, X^\varepsilon_{s-}, u) - 1) | \mathcal{F}^\varepsilon_s]dN_{\lambda}(ds, du)
\]
\[
+ \int_0^t \mathbb{E}^\varepsilon[\Lambda_t^\varepsilon (\mathcal{L}X^\varepsilon_t)\psi(X^\varepsilon_s, Z^\varepsilon_s) | \mathcal{F}^\varepsilon_s]ds,
\]

i.e.
\[
\rho^\varepsilon_t(\psi) = \rho^\varepsilon_0(\psi) + \int_0^t \rho^\varepsilon_s \left( (\mathcal{L}X^\varepsilon_s)(\cdot, Z^\varepsilon_s) \right) ds + \int_0^t \rho^\varepsilon_s (\psi h(\cdot, Z^\varepsilon_s) ds + \int_0^t \int_{\mathcal{U}_3} \rho^\varepsilon_s (\psi(\lambda(s, \cdot, u) - 1)) dN_{\lambda}(ds, du).
\]

(7)
For the detailed deduction of the above equation, please refer to the proof of Theorem 3.3 in [11].

Let $\tau$ be any $(\mathcal{F}_t)_{t \geq 0}$—stopping time no more than $T$. For any $\delta > 0$, we compute $\mathbb{E}[\rho^\varepsilon_{\tau+\delta}(\psi) - \rho^\varepsilon_{\tau}(\psi)]$. It follows from the Hölder inequality that

$$\mathbb{E}[\rho^\varepsilon_{\tau+\delta}(\psi) - \rho^\varepsilon_{\tau}(\psi)] = \mathbb{E}^{\mathbb{P}_{\varepsilon}}[\rho^\varepsilon_{\tau+\delta}(\psi) - \rho^\varepsilon_{\tau}(\psi)] \leq \left( \mathbb{E}^{\mathbb{P}_{\varepsilon}}[\rho^\varepsilon_{\tau+\delta}(\psi) - \rho^\varepsilon_{\tau}(\psi)]^2 \right)^{1/2} \left( \mathbb{E}^{\mathbb{P}_{\varepsilon}}(\Lambda^\varepsilon_T)^2 \right)^{1/2}.$$

Since $\mathbb{E}^{\mathbb{P}_{\varepsilon}}(\Lambda^\varepsilon_T)^2 < C$, which has been proved in Lemma [11], we only consider $\mathbb{E}^{\mathbb{P}_{\varepsilon}}[\rho^\varepsilon_{\tau+\delta}(\psi) - \rho^\varepsilon_{\tau}(\psi)]^2$. The Hölder inequality and the Itô isometry admit us to get

$$\mathbb{E}^{\mathbb{P}_{\varepsilon}}[\rho^\varepsilon_{\tau+\delta}(\psi) - \rho^\varepsilon_{\tau}(\psi)]^2 \leq 3\mathbb{E}^{\mathbb{P}_{\varepsilon}} \left| \int_\tau^{\tau+\delta} \rho^\varepsilon_s \left( \mathcal{L}^\varepsilon \psi \right) (\cdot, Z^\varepsilon_s) \, ds \right|^2$$

$$+ 3\mathbb{E}^{\mathbb{P}_{\varepsilon}} \left| \int_\tau^{\tau+\delta} \rho^\varepsilon_s \left( \psi h(\cdot, Z^\varepsilon_s) \right) d\tilde{B}^i_s \right|^2$$

$$+ 3\mathbb{E}^{\mathbb{P}_{\varepsilon}} \left| \int_\tau^{\tau+\delta} \int_{U_3} \rho^\varepsilon_s \left( \psi(\lambda(s, \cdot, u) - 1) \right) \hat{N}_\lambda(du, ds) \right|^2$$

$$\leq 3\varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_\tau^{\tau+\delta} \left| \rho^\varepsilon_s \left( \mathcal{L}^\varepsilon \psi \right) (\cdot, Z^\varepsilon_s) \right|^2 \, ds$$

$$+ 3\mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_\tau^{\tau+\delta} \left| \rho^\varepsilon_s \left( \psi h(\cdot, Z^\varepsilon_s) \right) \right|^2 \, ds$$

$$+ 3\mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_\tau^{\tau+\delta} \int_{U_3} \left| \rho^\varepsilon_s \left( \psi(\lambda(s, \cdot, u) - 1) \right) \right|^2 \nu_3(du) \, ds$$

$$=: I_1 + I_2 + I_3.$$

Firstly, deal with $I_1$. By the Jensen inequality and $(H^2_{b_1, \sigma_1, f_1})$, it holds that

$$I_1 = 3\varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_\tau^{\tau+\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ \mathcal{E}_s^\varepsilon (\mathcal{L}^\varepsilon \psi)(X^\varepsilon_s, Z^\varepsilon_s) \right]^2 \, ds$$

$$\leq 3\varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_\tau^{\tau+\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ (\mathcal{E}_s^\varepsilon)^2 \right] \left[ (\mathcal{L}^\varepsilon \psi)(X^\varepsilon_s, Z^\varepsilon_s) \right]^2 \, ds$$

$$\leq 3C \varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_\tau^{\tau+\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ (\mathcal{E}_s^\varepsilon)^2 \right] \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ \mathcal{F}_s^\varepsilon \right] \, ds$$

$$\leq 3C \varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_0^{\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ (\mathcal{E}_{\tau+s}^\varepsilon)^2 \right] \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ \mathcal{F}_{\tau+s}^\varepsilon \right] \, ds$$

$$\leq 3C \varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_0^{\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ (\mathcal{E}_{\tau+s}^\varepsilon)^2 \right] \, ds$$

$$\leq 3C \varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_0^{\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ (\mathcal{E}_{\tau+s}^\varepsilon)^2 \right] \, ds$$

$$\leq 3C \varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_0^{\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ (\mathcal{E}_{\tau+s}^\varepsilon)^2 \right] \, ds$$

$$\leq 3C \varepsilon \mathbb{E}^{\mathbb{P}_{\varepsilon}} \int_0^{\delta} \mathbb{E}^{\mathbb{P}_{\varepsilon}} \left[ (\mathcal{E}_{\tau+s}^\varepsilon)^2 \right] \, ds$$

where $C$ is independent of $\varepsilon, \delta$. By the same deduction to $I_1$, we get that $I_2 + I_3 \leq C\delta$. Thus,

$$\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \sup_{\tau \leq T} \mathbb{E}[\rho^\varepsilon_{\tau+\delta}(\psi) - \rho^\varepsilon_{\tau}(\psi)] = 0.$$  

(8)
Based on the similar calculation to above, it holds that
\[
\sup_{\varepsilon,t \in [0,T]} \mathbb{E} |\rho_t^\varepsilon(\psi)| < \infty. \tag{9}
\]
So, combining (9) with (8), we know from Theorem 5.1 in [4] that \(\{\rho_t^\varepsilon(\psi), t \in [0,T]\}\) is relatively weakly compact in \(D([0,T], \mathbb{R})\). Moreover, Theorem 6.2 in [4] admits us to obtain that \(\{\rho_t^\varepsilon, t \in [0,T]\}\) is relatively weakly compact in \(D([0,T], \mathcal{M}(\mathbb{R}^n))\). \(\square\)

To attain the convergence of \(\pi_t^\varepsilon\) to \(\pi_t^0\) as \(\varepsilon \to 0\), we assume more:

**Assumption 6.** \(\{Z_t^\varepsilon, t \in [0,T]\}\) is tight.

Now, it is the position to state the main result in the section.

**Theorem 5.3.** Under Assumption 1.-6., \(\pi_t^\varepsilon\) converges weakly to \(\pi_t^0\) as \(\varepsilon \to 0\) for any \(t \in [0,T]\).

**Proof.** For \(\psi \in C^2_b(\mathbb{R}^n)\), it holds that
\[
\pi_t^\varepsilon(\psi) - \pi_t^0(\psi) = \frac{\rho_t^\varepsilon(\psi) - \rho_t^0(\psi)}{\rho_t^0(1)} - \pi_t^\varepsilon(\psi) \frac{\rho_t^0(1) - \rho_t^0(\psi)}{\rho_t^0(1)}.
\]
Thus, in order to prove \(\pi_t^\varepsilon(\psi) - \pi_t^0(\psi)\) converges weakly to 0, by Lemma 5.1 and the conditional expectation property of \(\pi_t^\varepsilon(\psi)\), we only need to show that \(\rho_t^\varepsilon(\psi)\) converges weakly to \(\rho_t^0(\psi)\) as \(\varepsilon \to 0\).

On one hand, we compute the weak limit of \(\rho_t^\varepsilon(\psi)\) as \(\varepsilon \to 0\). By Lemma 5.2 there exist a weakly convergence subsequence \(\{\rho_t^{\varepsilon_k}, k \in \mathbb{N}\}\) and a measure-valued process \(\bar{\rho}_t\) such that \(\rho_t^{\varepsilon_k}(\psi)\) converges weakly to \(\bar{\rho}_t(\psi)\) as \(k \to \infty\). To compare \(\bar{\rho}_t(\psi)\) with \(\rho_t^0(\psi)\), we deduce the equation which \(\bar{\rho}_t(\psi)\) satisfies. Note that \(\rho_t^0(\psi)\) solves Eq. (7). And then we consider the weak limits of three integrals in Eq. (7). For the first integral, it holds that
\[
\rho_s^{\varepsilon_k}\left((\mathcal{L}^{X^{\varepsilon_k}}\psi)(\cdot, Z_s^{\varepsilon_k})\right) - \bar{\rho}_s\left(\bar{\mathcal{L}}\psi\right) = \rho_s^{\varepsilon_k}\left((\mathcal{L}^{X^{\varepsilon_k}}\psi)(\cdot, Z_s^{\varepsilon_k})\right) - \rho_s^{\varepsilon_k}\left(\bar{\mathcal{L}}\psi\right) + \rho_s^{\varepsilon_k}\left(\bar{\mathcal{L}}\psi\right) - \bar{\rho}_s\left(\bar{\mathcal{L}}\psi\right) = I_1 + I_2.
\]

For \(I_1\), one know that
\[
I_1 = \mathbb{E}^{\mathbb{P}\varepsilon_k} \left[ \Lambda_{s_{k}}^{\varepsilon_k} \left((\mathcal{L}^{X^{\varepsilon_k}}\psi)(X^{\varepsilon_k}_s, Z_s^{\varepsilon_k}) - (\bar{\mathcal{L}}\psi)(X^{\varepsilon_k}_s)\right) \bigg| \mathcal{F}_{Y^{\varepsilon_k}}^s \right]
\]
\[
= \mathbb{E}^{\mathbb{P}\varepsilon_k} \left[ \Lambda_{s_{k}}^{\varepsilon_k} \frac{\partial \psi}{\partial x_i}(X^{\varepsilon_k}_s) \left[b_1^{ij}(X^{\varepsilon_k}_s, Z_s^{\varepsilon_k}) - \bar{b}_1^{ij}(X^{\varepsilon_k}_s)\right] \bigg| \mathcal{F}_{Y^{\varepsilon_k}}^s \right]
\]
\[
+ \frac{1}{2} \mathbb{E}^{\mathbb{P}\varepsilon_k} \left[ \Lambda_{s_{k}}^{\varepsilon_k} \frac{\partial^2 \psi}{\partial x_i \partial x_j}(X^{\varepsilon_k}_s) \left[(\sigma_1 \sigma_1^T)^{ij}(X^{\varepsilon_k}_s, Z_s^{\varepsilon_k}) - (\bar{\sigma_1} \bar{\sigma_1}^T)^{ij}(X^{\varepsilon_k}_s)\right] \bigg| \mathcal{F}_{Y^{\varepsilon_k}}^s \right]
\]
\[
=: I_{11} + I_{12}.
\]

Let us deal with \(I_{11}\). Since
\[
\lim_{n \to \infty} \sum_{j=0}^{n-1} \mathbb{E}^{\mathbb{P}\varepsilon_k} \left[ \Lambda_{(j+1)t/n}^{\varepsilon_k} \frac{\partial \psi}{\partial x_i}(X_{(j+1)t/n}^{\varepsilon_k}, Z_{(j+1)t/n}^{\varepsilon_k}) \left[b_1^{ij}(X_{(j+1)t/n}^{\varepsilon_k}, Z_{(j+1)t/n}^{\varepsilon_k}) - \bar{b}_1^{ij}(X_{(j+1)t/n}^{\varepsilon_k})\right] \bigg| \mathcal{F}_{Y^{\varepsilon_k}}^{(j+1)t/n} \right] I_{(jt/n,(j+1)t/n)}(s)
\]
\[ = \mathbb{E}^{\varepsilon_k} \left[ \Lambda^\varepsilon_k \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] \mathcal{F}^{\varepsilon_k}_s \right], \ a.s., \mathbb{P},
\]
we only consider \( \mathbb{E}^{\varepsilon_k} \left[ \Lambda^\varepsilon_k \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] \mathcal{F}^{\varepsilon_k}_s \right)
for \( s \in (jt/n, (j + 1)t/n] \). Based on independence of \( X^\varepsilon_k, Z^\varepsilon_k \) and \( Y^\varepsilon_k \) under \( \mathbb{P}^{\varepsilon_k} \), it holds that
\[ = \mathbb{E}^{\varepsilon_k} \left[ \Lambda^\varepsilon_k \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] \mathcal{F}^{\varepsilon_k}_s \right).
\]
where \( l \) is a positive integer such that \( s - \varepsilon_k l > 0 \). And then we compute
\[ \mathbb{E}^{\varepsilon_k} \left[ \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] X^\varepsilon_k_{s-\varepsilon_k l}, Z^\varepsilon_k_{s-\varepsilon_k l} \right).
\]
On one side, it is easy to see that
\[ = \mathbb{E}^{\varepsilon_k} \left[ \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] X^\varepsilon_k_{s-\varepsilon_k l}, Z^\varepsilon_k_{s-\varepsilon_k l} \right].
\]
\[ = \mathbb{E}^{\varepsilon_k} \left[ \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] X^\varepsilon_k_{s-\varepsilon_k l}, Z^\varepsilon_k_{s-\varepsilon_k l} \right].
\]
\[ =: I_{111} + I_{112}.\]
Based on tightness of \( \{(X_t, Z_t^l), t \in [0, T]\} \) and \((H^1_{b_1, \sigma_1, f_1})\), it holds that \( \lim_{k \to \infty} I_{111} = 0 \). By the definition of \( p(X^\varepsilon_k Z^\varepsilon_k, l, dz) \) and \( \bar{p}(X^\varepsilon_k, dz) \), we know that \( \lim_{l \to \infty} I_{112} = 0 \). On the other side, it follows from the dominated convergence theorem that
\[ \lim_{n \to \infty} \int \mathbb{E}^{\varepsilon_k} \left[ \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] = \int \mathbb{E}^{\varepsilon_k} \left[ \frac{\partial \psi}{\partial x_i}(X^\varepsilon_k) \left[ b^i_1(X^\varepsilon_k, Z^\varepsilon_k) - \bar{b}^i_1(X^\varepsilon_k) \right] \right] d\mu \]
Thus, the dominated convergence theorem admits us to obtain \( \lim_{k \to \infty} I_{111} = 0 \).
By the same deduction to that for $I_{11}$, it holds that $I_{12}$ goes to zero a.s. as $k \to \infty$. Thus, $I_1$ converges to zero as $k \to \infty$, which together with weak convergence of $I_2$ to zero as $k \to \infty$ yields that $\rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right)$ converges weakly to $\bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right)$ as $k \to \infty$. Besides, set

$$
\left( \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) \right)^{(n)} := \sum_{j=0}^{n-1} \rho_{(j+1)t/n}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{(j+1)t/n}^\varepsilon) \right) I(t/n,(j+1)t/n)(s),
$$

and then

$$
\lim_{n \to \infty} \left( \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) \right)^{(n)} = \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right), \quad \text{a.s.} \mathbb{P},
$$

$$
\lim_{n \to \infty} \left( \bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right) \right) = \bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right), \quad \text{a.s.} \mathbb{P}.
$$

Moreover, by the dominated convergence theorem, it holds that

$$
\lim_{n \to \infty} \mathbb{E}^{\mathbb{P}_{s}^k} \left( \int_{0}^{t} \left| \left( \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) \right)^{(n)} - \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) \right|^2 ds \right) = 0,
$$

$$
\lim_{n \to \infty} \mathbb{E}^{\mathbb{P}_{s}^k} \left( \int_{0}^{t} \left| \left( \bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right) \right)^{(n)} - \bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right) \right|^2 ds \right) = 0.
$$

So, the H"{o}lder inequality admits us to obtain that

$$
\lim_{n \to \infty} \mathbb{E}^{\mathbb{P}_{s}^k} \left| \int_{0}^{t} \left( \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) \right)^{(n)} ds - \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) ds \right|^2 = 0,
$$

$$
\lim_{n \to \infty} \mathbb{E}^{\mathbb{P}_{s}^k} \left| \int_{0}^{t} \left( \bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right) \right)^{(n)} ds - \int_{0}^{t} \bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right) ds \right|^2 = 0.
$$

From this, it follows that

$$
\int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) ds - \int_{0}^{t} \bar{\rho}_s \left( \bar{\mathcal{L}} \psi \right) ds
$$

$$
= \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) ds - \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right)^{(n)} ds + \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right)^{(n)} ds - \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) ds
$$

$$
+ \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right)^{(n)} ds - \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) ds
$$

$$
= \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right) ds - \int_{0}^{t} \rho_{s}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{s}^\varepsilon) \right)^{(n)} ds
$$

$$
+ \sum_{j=0}^{n-1} \left( \rho_{(j+1)t/n}^\varepsilon \left( (\mathcal{L}^X_{s} \psi)(\cdot, Z_{(j+1)t/n}^\varepsilon) \right) - \bar{\rho}_{(j+1)t/n} \left( \bar{\mathcal{L}} \psi \right) \right) \left( (j + 1)t/n - jt/n \right)
\[ + \int_0^t (\rho_s(\mathcal{L}\psi))(n) \, ds - \int_0^t \bar{\rho}_s(\mathcal{L}\psi) \, ds \xrightarrow{w} 0, \]

i.e.
\[
\int_0^t \rho_s^{\varepsilon_k}((\mathcal{L}X_s^{\varepsilon_k})(\cdot, Z_s^{\varepsilon_k})) \, ds \xrightarrow{w} \int_0^t \bar{\rho}_s(\mathcal{L}\psi) \, ds. \tag{10}
\]

In the following, we treat the second integral in Eq. (10). By the similar deduction to above one could have that \( \rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i) \) converges weakly to \( \bar{\rho}_s(\psi \bar{h}^i) \) as \( k \to \infty \). Besides, define
\[
(\rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i))(n) := \sum_{j=0}^{n-1} \rho_{(j+1)t/n}^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i) I_{(jt/n,(j+1)t/n]}(s),
\]

\[
(\bar{\rho}_s(\psi \bar{h}^i))(n) := \sum_{j=0}^{n-1} \rho_{(j+1)t/n}(\psi \bar{h}^i) I_{(jt/n,(j+1)t/n]}(s),
\]

and then
\[
\lim_{n \to \infty} (\rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i))(n) = \rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i), \quad a.s. \mathbb{P},
\]
\[
\lim_{n \to \infty} (\bar{\rho}_s(\psi \bar{h}^i))(n) = \bar{\rho}_s(\psi \bar{h}^i), \quad a.s. \mathbb{P}.
\]

Furthermore it follows from the dominated convergence theorem that
\[
\lim_{k \to \infty} \mathbb{E}^{\mathbb{P}^{\varepsilon_k}} \left( \int_0^t \left| (\rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i))^{(n)} - \rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i) \right|^2 \, ds \right) = 0,
\]
\[
\lim_{k \to \infty} \mathbb{E}^{\mathbb{P}^{\varepsilon_k}} \left( \int_0^t \left| (\bar{\rho}_s(\psi \bar{h}^i))^{(n)} - \bar{\rho}_s(\psi \bar{h}^i) \right|^2 \, ds \right) = 0.
\]

Based on the Itô isometry, it holds that
\[
\mathbb{E}^{\mathbb{P}^{\varepsilon_k}} \left| \int_0^t \left( \rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i) \right)^{(n)} \, dB_s^i - \int_0^t \rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i) \, dB_s^i \right|^2 = \sum_{i=1}^m \mathbb{E}^{\mathbb{P}^{\varepsilon_k}} \left( \int_0^t \left| (\rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i))^{(n)} - \rho_s^{\varepsilon_k}(\psi h(\cdot, Z_s^{\varepsilon_k})^i) \right|^2 \, ds \right),
\]
\[
\mathbb{E}^{\mathbb{P}^{\varepsilon_k}} \left| \int_0^t \left( \bar{\rho}_s(\psi \bar{h}^i) \right)^{(n)} \, dB_s^i - \int_0^t \bar{\rho}_s(\psi \bar{h}^i) \, dB_s^i \right|^2 = \sum_{i=1}^m \mathbb{E}^{\mathbb{P}^{\varepsilon_k}} \left( \int_0^t \left| (\bar{\rho}_s(\psi \bar{h}^i))^{(n)} - \bar{\rho}_s(\psi \bar{h}^i) \right|^2 \, ds \right).
Thus, \( f_0^t \left( \rho_s^k \left( \psi h(\cdot, Z_{s}^{j k}) \right) \right) \) and \( f_0^t \left( \bar{\rho}_s \left( \psi \tilde{h} \right) \right) \) converge in mean square to:

\[
\int_0^t \rho_s^k \left( \psi h(\cdot, Z_{s}^{j k}) \right) d\bar{B}_s^j \quad \text{and} \quad \int_0^t \bar{\rho}_s \left( \psi \tilde{h} \right) d\bar{B}_s^j,
\]

respectively. Let us compute

\[
\int_0^t \rho_s^k \left( \psi h(\cdot, Z_{s}^{j k}) \right) d\bar{B}_s^j - \int_0^t \bar{\rho}_s \left( \psi \tilde{h} \right) d\bar{B}_s^j \quad \rightarrow \quad 0,
\]

that is,

\[
\int_0^t \rho_s^k \left( \psi h(\cdot, Z_{s}^{j k}) \right) d\bar{B}_s^j \quad \rightarrow \quad \int_0^t \bar{\rho}_s \left( \psi \tilde{h} \right) d\bar{B}_s^j. \tag{11}
\]

For the third integral in Eq. (7), by the similar deduction to the second integral it holds that

\[
\int_0^t \int_{U_3} \rho_s^k \left( \psi(\lambda(s, \cdot, u) - 1) \right) \tilde{N}_\lambda(ds, du) \quad \rightarrow \quad \int_0^t \int_{U_3} \bar{\rho}_s \left( \psi(\lambda(s, \cdot, u) - 1) \right) \tilde{N}_\lambda(ds, du). \tag{12}
\]

Combining (12) with (10) (11) and taking weak limits on two hand sides of (7) as \( k \rightarrow \infty \), we obtain that

\[
\bar{\rho}_t(\psi) = \rho_0(\psi) + \int_0^t \bar{\rho}_s(\tilde{\mathcal{L}} \psi) ds + \int_0^t \bar{\rho}_s \left( \psi \tilde{h} \right) d\bar{B}_s^j 
+ \int_0^t \int_{U_3} \bar{\rho}_s \left( \psi(\lambda(s, \cdot, u) - 1) \right) \tilde{N}_\lambda(ds, du).
\]

On the other hand, we consider \( \rho_0^t(\psi) \). By the similar deduction to \( \rho_t^k(\psi) \), it holds that

\[
\rho_0^t(\psi) = \rho_0^0(\psi) + \int_0^t \rho_0^s(\tilde{\mathcal{L}} \psi) ds + \int_0^t \rho_0^s \left( \psi \tilde{h} \right) d\bar{B}_s^j 
+ \int_0^t \int_{U_3} \rho_0^s \left( \psi(\lambda(s, \cdot, u) - 1) \right) \tilde{N}_\lambda(ds, du).
\]
Thus, $\bar{\rho}$ and $\rho^0$ solve the same equation

$$
\rho_t(\psi) = \rho_0(\psi) + \int_0^t \rho_s\left(\bar{\mathcal{L}}\psi\right)ds + \int_0^t \rho_s\left(\psi\bar{h}^i\right)d\bar{B}^i_s \\
+ \int_0^t \int_{U^3} \rho_s\left(\psi(\lambda(s, \cdot, u) - 1)\right)\tilde{N}_\lambda(ds, du).
$$

(13)

Besides, based on Theorem 4.2 in [11], Eq. (13) has a unique solution. So, for $t \in [0, T]$

$$\bar{\rho}_t(\psi) = \rho^0_t(\psi), \quad \text{a.s.}\ \mathbb{P}.$$ 

That is, as $k \to \infty$, $\rho^k_t(\psi)$ converges weakly to $\rho^0_t(\psi)$. The proof is completed. $\square$

REFERENCES

[1] N. Ikeda and S. Watanabe: *Stochastic differential equations and diffusion processes*, 2nd ed., North-Holland/Kodanska, Amsterdam/Tokyo, 1989.

[2] 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.

[3] T. G. Kurtz: *Approximation of Population Processes*, Vol. 36 of CBMS-NSF Regional Conf. Series in Appl. Math., SIAM, Philadelphia, 1981.

[4] H. J. Kushner: *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Systems & Control: Foundations & Applications 3. Birkhäuser, Boston, 1990.

[5] V. M. Lucic and A. J. Heunis: Convergence of nonlinear filters for randomly perturbed dynamical systems, *Appl. Math. Optim.*, 48(2003)93-128.

[6] A. Pazy: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag Berlin Heidelberg New York, 1983.

[7] J. H. Park, N. S. Namachchivaya and H. C. Yeong: Particle filters in a multiscale environment: Homogenized hybrid particle filter, *J. Appl. Mech.*, 78(2011)1-10.

[8] J. H. Park, R. B. Sowers and N. S. Namachchivaya: Dimensional reduction in nonlinear filtering, *Nonlinearity*, 23(2010)305-324.

[9] H. J. Qiao: Exponential ergodicity for SDEs with jumps and non-Lipschitz coefficients, *Journal of Theoretical Probability*, 27(2014)137-152.

[10] H. J. Qiao: Euler approximation for SDEs with jumps and non-Lipschitz coefficients, *Osaka Journal of Mathematics*, 51(2014)47-66.

[11] H. J. Qiao and J. Q. Duan: Nonlinear Filtering of Stochastic Dynamical Systems with Lévy Noises, *Advances in Applied Probability*, 47(2015)902-918.

[12] D. W. Stroock and S. R. S. Varadhan: Diffusion processes with boundary conditions, *Comm. Pure Appl. Math.*, 24(1971)147-225.