Weak Convergence of a Sequence of Semimartingales to a Process of Diffusion Type with Discontinuous Coefficients

Yi-Ju Chao
127 Vincent Hall, 206 Church St. S.E.,
Math Department, University of Minnesota,
Minneapolis, MN 55455,
E-mail: chao@math.umn.edu

Abstract

1 Introduction

Problems regarding conditions for processes to converge wealy to a process of diffusion type under various assumptions have attracted much work, for example [1]-[18]. The goal of this article is to explore sufficient conditions for weak convergence in the case when the pre-limit processes are semimartingales.

R. Sh. Liptser and A. V. Shiryaev's article in 1984 [18] discusses some sufficient conditions for the weak convergence of a sequence of semimartigales to a process of diffusion type. One of the sufficient conditions is requiring the drift coefficients and diffusion coefficients of limit processes are both continuous in the Skorokhod-Lindvall topology. This article considers discontinuous coefficients of limit processes. Besides some sufficient conditions discussed in [18], requiring the sets of discontinuity of the coefficients of limit processes are of Lebesque measure zero, and coercivity of the diffusion coefficients can make weak convergence achieved.

The problem discussed in this article is applicable from a practical point of view. The results can apply to considerable queueing models, such as least-load-balancing models [21] [23], and models of congestion control [22]. In these models, the drift coefficients of the limit processes are discontinuous with respect to state-space. By formulating queue-length processes as semimartingales and applying the results in this article, diffusion approximation can be readily obtained.

This article is organized as follows. The introduction is in section 1. The notations and main results are presented in section 2. Section 3 gives preliminary for proving the main results, they are theorems in control diffusion processes and supermatingale characterization of a class of stochastic integrals. Section 4 proves the main results. Section 5 proposes futher study.

2 Notations and main results

Let $(\Omega, \mathcal{F}, P)$ be a given complete probability space, and on it let $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$, be right-continuous nondecreasing families of σ-algebras, where the σ-algebras $\mathcal{F}_0$ is completed by the sets in $\mathcal{F}$ having zero probability.

Let D be the space of right-continuous functions with left limits, $X = (X_t)_{t\geq 0}$ from $R^+ = [0, \infty)$ to a d-dimensional Euclidean space R^d, equipped with the Skorokhod-Lindvall topology. Let $\mathcal{D} = (D_t)_{t\geq 0}$ be the family of σ-algebras

$$D_t = \bigcap_{\epsilon>0} \sigma\{X : X_s, 0 \leq s \leq t + \epsilon\}$$

and

$$\mathcal{D}_\infty = \sigma(\bigcup_{t\geq 0} D_t).$$

Let $X = (X_t, \mathcal{F}_t)_{t\geq 0}$ denote a R^d-valued semimartingale with sample paths in D. Let Q be the probability distribution of X, i.e. the probability measure such that for any $\Gamma \in \mathcal{D}_\infty$

$$Q(\Gamma) = P(X \in \Gamma).$$

Each semimartingale X admits a canonical representation (see [17]).

$$X_t = X_0 + B_t + X_t^c + \int_0^t \int_{|x|\leq a} xd(\mu - \nu) + \int_0^t \int_{|x|>a} xd\mu, \tag{2.1}$$

where $a \in (0, 1]$, $B_t = (B_t, \mathcal{F}_t)_{t\geq 0}$ in R^d is a predictable process of locally bounded variation, with $B_0 = 0$, and $X_t^c = (X_t^c, \mathcal{F}_t)_{t\geq 0}$ is a local martingale, $X_0^c = 0$, with sample paths in the space C, the space of continuous functions, and with quadratic characteristic A_t, μ is the random integer-valued jump measure of the process X, $(\mu((0, t] \times B)) = \sum_{0<s\leq t} I(\Delta X_s \in B))$ with Borel sets $B \in \mathcal{B}(R_0^d)$, where $R_0^d = R^d \setminus \{0\})$, $\Delta X_s = X_s - X_{s-}$, where $X_{s-} = \lim_{u\to s^-} X_u$, and ν is the compensator of the jump measure μ. We call that the martingale X_t has triplet of predictable characteristics $T = (B_t, A_t, \nu)$, which uniquely determines the semimartingale X_t.

We also recall that the compensator ν of a jump measure μ is a positive random measure on $R^+ \times R^d$ having the following properties (for details, see [17]):

(A) $\nu([0, \infty) \times \{0\}) = \nu(\{0\} \times R^d) = 0$,

(B)

$$\int_0 \int_{R^d} (x^2 \wedge 1) d\nu < \infty (P - a.s.), t > 0, \tag{2.2}$$

(C) For each Borel set $\Gamma \in \mathcal{B}(R_0^d)$, the process $(\nu((0,t],\Gamma), \mathcal{F}_t)$ is a predictable and such that the process

$$(\mu((0,t],\Gamma) - \nu((0,t],\Gamma), \mathcal{F}_t)_{t \geq 0}$$

is a local martingale,

(D) $\int_{|x|\leq a} x\nu(t,dx) = \Delta B_t, t > 0.$

We let

$$M_t^a = X_t^c + \int_0^t \int_{|x|\leq a} xd(\mu - \nu).$$

$M_t^a = (M_t^a, \mathcal{F}_t^a)$ is thus a locally square-integrable martingale with predictable quadratic characteristic

$$< M^a >_t = A_t + \int_0^t \int_{|x|\leq a} xx^* d\nu - \sum_{0<s\leq t} \hat{x}_s^a (\hat{x}_s^a)^*,$$

where

$$\hat{x}_t^a = \int_{|x|\leq a} x\nu(t,dx),$$

and $*$ if the transposition sign.

We consider a sequence of complete probability spaces $(\Omega^n, \mathcal{F}^n, P^n)$, and a sequence of semimartingales $X^n = (X_t^n, \mathcal{F}_t^n)$ with triplets of predictable characteristics $T^n = (B^n, A^n, \nu^n)$. The canonical representations of X^n are

$$X_t^n = X_0^n + B_t^n + M_t^{a,n} + \int_0^t \int_{|x|>a} xd\mu^n, \tag{2.3}$$

where $B_t^n = (B_{1,t}^n, \cdots, B_{d,t}^n)$ in R^d, and $< M^{a,n} >_t = (< M_i^{a,n}, M_j^{a,n} >_t)_{i,j=1,\cdots,d}$ are $d \times d$-matrices. We also denote Q^n the probability distribution of X^n, i.e. the probability measure such that for any $\Gamma \in \mathcal{D}_\infty$

$$Q^n(\Gamma) = P^n(X^n \in \Gamma).$$

We also assume a R^d-valued semimartingale $(X_t, \mathcal{D}_t)$ with the triplet $T = (B(X), A(X), 0)$, with

$$B_t(X) = \int_0^t b(s,X) ds$$

and

$$A_t(X) = \int_0^t a(s,X) ds,$$

where the function $b(t,X)=(b^1(t,X),\cdots,b^d(t,X))$, taking values in R^d, is predictable for each $t\in R^+$ and $X\in D$, and a $d\times d$-matrix $a(t,X)=(a^{ij}(t,X))_{i,j=1}^d$ is symmetric and nonegative definite. The semartingale X_t is assumed to admit the canonical representation (possibly on an enlargement probability space)

$$X_t = X_0 + \int_0^t b(s,X)ds + \int_0^t \sqrt{a(s,X)}dw_s, \tag{2.4}$$

with R^d-valued Wiener process $(w_t)_{t\geq 0}$ consisting of independent components.

We present some conditions appearing in the statements of our main results. For each $T>0$, $\epsilon>0$,

Condition (N): for all $a\in(0,1]$

$$\lim_n P^n(\int_0^T\int_{|x|>a} d\nu^n \geq \epsilon) = 0.$$

Condition (sup B):

$$\lim_n P^n(\sup_{t\leq T}|B_t^n - \int_0^t b(s,X^n)ds| \geq \epsilon) = 0.$$

Condition (sup A):

$$\lim_n P^n(\sup_{t\leq T}| < M^{a,n} >_t - \int_0^t a(s,X^n)ds| \geq \epsilon) = 0.$$

Condition (Linear growth (I)): *There exists a function $L(t)$ satisfying*

$$\int_0^t L(s)ds < \infty,$$

such that

$$|b(t,X)| \leq L(t)(1+\sup_{s\leq t}|X_s|), \tag{2.5}$$

and

$$|a(t,X)| \leq L(t)(1+\sup_{s\leq t}|X_s|^2). \tag{2.6}$$

It's well-known that if conditions **(sup A) (sup B) (N)** and **(Linear growth (I))** are satisfied, the functions $b(t,X)$ and $a(t,X)$ are continuous with respect to $X\in C$ in the Skorokhod-Lindvall topology, and X_0^n converges to some random variable X_0 in distribution, then the semimartingales $X^n(t)$ (2.3) converge to (2.4) weakly. The reader can find the theorems in articles [18],[17].

If we consider "state-dependent" drfit and diffusion coefficients, i.e.

Condition (S): $a(t,X) \equiv a(t,x)$, *and* $b(t,X) \equiv b(t,x)$ *with* $t \in R^{+}$, and $x \in R^{d}$.

We also allow the coefficients $b(t,x)$ and $a(t,x)$ to be discontinuous with respect to $x \in R^d$. We discover a set of sufficient conditions on the coefficients (**(Coercive), (Mb), (Ma)**, which will be described later) for weak convergence of the semimartingales (2.3) to the diffusion process

$$X_t = X_0 + \int_0^t b(s,X_s)ds + \int_0^t \sqrt{a(s,X_s)}dw_s, \tag{2.7}$$

where w_s is a d-dimensional Wiener process with independent components. The main idea of our approach is that the limit diffusion process can not spend too much time in the sets of discontinuity of the coefficients $a(t,x)$ and $b(t,x)$ with respect to x. This property can be achieved while the diffusion coefficient $a(t,x)$ is coercive and the sets of discontinuity of the coefficients are of Lebesgue measure zero.

We now present the sufficient conditions in our main results. Under condition **(S)**, we also consider the following conditions:

Condition (Coercive): *There exists a constant* $\gamma > 0$ *such that*

$$\gamma \le (a(t,x)\lambda, \lambda),$$

for all $t \in R^+$, $x \in R^d$, *for all d-dimensional unit vectors* λ, *where* $(\cdot,\cdot)$ *denotes the inner product.*

For all $t \in R^+$, we define the set of discontinuity of the coefficient $a(t,x)$

$$M_{a,t} := \{(t,x) : a(t,.) \text{ is discontinuous in } x \in R^d\}$$

and a subset in $R^+ \times R^d$

$$M_a := \bigcup_{t \le T} M_{a,t}.$$

We also define the set of discontinuity of the coefficient $b(t,x)$ in the same way

$$M_{b,t} := \{(t,x) : b(t,.) \text{ is discontinuous in } x \in R^d\}$$

and a subset in $R^+ \times R^d$

$$M_b := \bigcup_{t \le T} M_{b,t}.$$

Condition (Ma): *For all* $\epsilon > 0$ *there exists an open set* O *in* $R^+ \times R^d$ *such that*

$$M_a \subset O,$$

and

$$\int_0^T \int_{R^d} I_O(t,x)dxdt < \epsilon.$$

Condition (Mb): *For all $\epsilon > 0$ there exists an open set O in $R^+ \times R^d$ such that*

$$M_b \subset O,$$

and

$$\int_0^T \int_{R^d} I_O(t,x)dxdt < \epsilon.$$

We now state the main results.

Theorem 2.1 *(A) Assume that conditions* **(Linear growth(I))**, **(N)**, **(sup A)**, **(sup B)** *are fulfilled, and also assume*

$$\lim_{l\to\infty} \limsup_n P^n(|X_0^n| \geq l) = 0. \tag{2.8}$$

Then the family of measure $(Q^n)_{n\geq 1}$ is relatively compact and the limit Q' of any weakly convergent subsequence of $(Q^n)_{n\geq 1}$ possesses the property $Q'(C) = 1$.

(B) In addition to the conditions in part (A), we also assume $a(s,X)$ and $b(s,X)$ are continuous in $X \in C$ in the Skorokhod topology, and X_0^n converges to X_0 in distribution instead of (2.8). Then the semimartingales (2.3) converge weakly to the solution of (2.7).

(C) We assume the conditions in part (A) are atill fulfilled. We also assume that condition **(S)** *is satisfied. Under assumption* **(S)**, *we moreover assume that conditions* **(Ma)**, **(Mb)** *and* **(Coercive)** *are fulfilled, and X_0^n satisfying*

$$E|X_0^n|^2 < N,$$

for some constant N not depending on n, and converging to X_0 in distribution instead of (2.8). Moreover, we assume that the coefficients $a(t,x)$ and $b(t,x)$ uniquely determine the distribution Q of the weak solution of the diffusion process defined by Itô integral equation

$$X_t = X_0 + \int_0^t b(s,X_s)ds + \int_0^t \sqrt{a(s,X_s)}dw_s, \tag{2.9}$$

where w_s is a Wiener process.

Then the semimartingales (2.3) converge weakly to the solution of (2.9).

We further investigate the sufficient conditions on the coefficients $a(t,x)$ and $b(t,x)$ for the weak convergence.

Theorem 2.2 *Besides the sufficient conditions mentioned in part (C) in theorem 2.1, we also assume that the coefficient $a(t,x)$ is continuous in $x \in R^d$ for all $t \in R^+$, then the semimartingale (2.3) converges weakly to the solution of (2.9).*

Remark 2.1 *Conditions* **(Linear growth (I))**, **(Coercive)**, **(Mb)** *and* **(Ma)** *are also only applied to $a(t,x)$ and $b(t,x)$. This weak convergence also requires that the coefficients $a(t,x)$ and $b(t,x)$ ensure the weak existence and uniqueness of the solution of the limit process. But there is no constraint on the drift coefficient and diffusion coefficient of the semimartingales except conditions* **(sup A) (sup B)**.

3 Preliminaries

3.1 Control diffusion process

The following theory is stated on page 51 in article [28]. Let A be a set of pairs (a,b), where a is a $d \times d$-dimensional nonnegative matrix and b is a d-dimensional vector. Assume on a probability space $(\Omega, \mathcal{F} = (\mathcal{F}_t)_{t\geq 0}, P)$, with a random process $(a_t, b_t) \in A$ for all (ω, t) the process

$$x_t = x_0 + \int_0^t b_s ds + \int_0^t \sqrt{a_s} dw_s, \tag{3.10}$$

is defined, where $(w_t, \mathcal{F}_t)$ is a d-dimensional Wiener process.

Let D be a bounded region in the Euclidean space R^d, and x_0 be a fixed point of D. Assume that a_t, b_t are progressively measurable with respect to $(\mathcal{F}_t)_{t\geq 0}$ and are bounded for all $t \in [0,T]$ and for all ω. We also denote the first exit time of x_t from the region D by τ_D, and define a set $H = [0,T] \times D$ in $R^+ \times R^d$.

We shall see further in the proof of our main results in section 4 that the estimate of the form plays an essential role

$$E \int_0^{\tau_D} |g(t, x_t)| dt \leq N \|g\|_{d+1,H}, \tag{3.11}$$

where $g : R^+ \times R^d \mapsto R$ is an arbitrary real-valued Borel function, and $\|g\|_{d+1,H}$ is the $L_{d+1}(H)$-norm of g

$$\|g\|_{d+1,H} := (\int_H |g(t,x)|^{d+1} dx dt)^{\frac{1}{d+1}}.$$

A crucial fact is that the constant N in (3.11) does not depend on a specified process with coefficients (a_t, b_t), but is given instead by the set A.

Definition 3.1 *A nonnegative function $F(a)$ defined on the set of all nonnegative definite symmetric matrices a of dimension $d \times d$ is said to be* **regular** *if for each $\epsilon > 0$ there is a constant $k(\epsilon)$ such that for all a and for all d-dimensional unit vectors λ*

$$F(a) \leq \epsilon tr(a) + k(\epsilon)(a\lambda, \lambda),$$

where $tr(a)$ denotes the trace of the matrix a, and $(\cdot,\cdot)$ denotes the inner product.

Theorem 3.1 *Assume that there is a regular function F such that the coefficients (a_t, b_t) in (3.10) satisfying $|b_t| \le F(a_t)$ for all (t, ω). Then there exist constants N depending only on the dimesion d, the function $F(a)$, and the diameter of the region D such that for all $s \ge 0$, and for all Borel functions $g(t,x)$ on $R^+ \times R^d$, we have*

$$E\{\int_s^{\tau_D} (det(a_t))^{\frac{1}{d+1}} |g(t,x_t)| dt | \mathcal{F}_s\} \le N \|g\|_{d+1,H},$$

Proof. The proof of this theorem is stated in Chapter 2 in article [29].

3.2 A supermartingale characterization of a class of stochastic integrals

For fixed $T > 0$, for $t \in [0, T]$, and stochastic processes $y. \in C = C(R^+, R^d)$, we denote the value of $y.$ at the point t by y_t, and denote the smallest σ-field in C by N_t such that all functions y_s as the functions on C are N_t-measurable for $s \le t$.

For all $t \in R^+$, $y. \in C$, let a set $A_t(y.)$ be defined with elements of the type (a, b), where $a = (a^{ij})$ is a symmetric nonnegative $d \times d$ matrix, and $b = (b^i) \in R^d$. Assume that for every $t, y.$ the set $A_t(y.)$ as a subset of R^{d_1}, with $d_1 = d^2 + d$, is a closed and convex. We furthermore assume that $(a, b) \in A_t(y.)$ satisfies the following condition.

Condition (Linear growth (II)): There exists a function $L(t)$ satisfying

$$\int_0^t L(s) ds < \infty$$

such that

$$|a^{ij}| \le L(t)(1 + \sup_{s \le t} |y_s|^2); \tag{3.12}$$

and

$$|b^i| \le L(t)(1 + \sup_{s \le t} |y_s|). \tag{3.13}$$

We also define

$$G_t(u_{ij}, u_i, y.) = \sup\{\sum_{i,j=1}^d a^{ij} u_{ij} + \sum_{i=1}^d b^i u_i : (a, b) \in A_t(y.)\}.$$

We suppose that $G_t(u_{ij}, u_i, y.)$ is N_t-measurable with respect to $y.$ for every $t \in [0, T]$, $(u_{ij}, u_i) \in R^{d_1}$, and is measurable with respect to $(t, y.)$ for every $(u_{ij}, u_i) \in R^{d_1}$. We also denote by $C_0^\infty(R^d)$ the set of all infinitely differentiable (real) functions each of which have compact support in R^d.

Theorem 3.2 *Assume on a given complete probability space* $(\Omega, \mathcal{F}, P)$, $x_t = x_t(\omega)$, $t \in [0,T]$, $\omega \in \Omega$, *be a continuous* R^d*-valued process satisfying*

$$E \sup_{t \leq T} |x_t|^2 < \infty. \tag{3.14}$$

If for every $m = 1, 2, \cdots$,

$$0 \leq t_1 \leq \cdots \leq t_m \leq s \leq t \leq T,$$

and

$$u, f_1, \cdots, f_m \in C_0^\infty(R^d), f_i \geq 0$$

we have

$$Ef_1(x_{t_1}) \cdots f_m(x_{tm})(u(x_t) - u(x_s)) \leq$$

$$Ef_1(x_{t_1}) \cdots f_m(x_{tm}) \int_s^t G_r(u_{x^i x^j}(x_r), u_{x^i}(x_r), x.)dr, \tag{3.15}$$

then there exists a function $(\tilde{a}_t(y.), \tilde{b}_t(y.))$ N_t*-measurable with respect to* $y.$ *for every* $t \in R^+$, *and satifying* $(\tilde{a}_t(y.), \tilde{b}_t(y.)) \in A_t(y.)$ *for all* $t \in [0,T]$, *and on some enlargement of* $(\Omega, \mathcal{F}, P)$ *there exists a d-dimensional Wiener process* w_t *such that for every* $s \in [0,T]$ *the process* x_t *regarded on* $[0,s]$ *and* $w_t - w_s$ *regarded on* $[s,T]$ *are independent and*

$$x_t = x_0 + \int_0^t \sqrt{\tilde{a}_r(x.)}dw_r + \int_0^t \tilde{b}_r(x.)dr \qquad \forall t \in [0,T] \qquad (a.s.). \tag{3.16}$$

Remark 3.1 *The proof of theorem 3.2 in this article is pretty similar the theorem proved in N. V. Krylov's article [28] except for some adjustments. In article [28], the author assumes that* $A_t(y.)$ *is uniformly bounded in* $t \in [0,T]$ *and for almost all* ω *without assuming that* x_t *is continuous and (3.14) and prove that the process* x_t *is continuous. We here consider* $(a, b) \in A_t(y.)$ *satisfying* **(Linear growth (II))** *condition, and furthermore assume that the process* x_t *is continuous, and (3.14) to get a similar result.*

Proof. Step 1. Consider the following spaces of real functions on $[0,T] \times C \times R^d$. Let $C^{1,2}$ be the set of all bounded functions $u(t, y., x)$, N_t-measurable in $y.$ for every t, x, with first derivatives with respect to (t, x), and second derivatives with respect to x continuous in (t, x) for every $y.$, and the second derivatives are bounded with respect to $(t, y., x)$. Let L_2 be the set of all N_t-adapted functions $u(t, y., x)$ which are measurable in $(t, y., x)$ and

$$E \int_0^T \int_{R^d} |u(t, y.(\omega), x)|^2 dxdt < \infty. \tag{3.17}$$

We view the space L_2 as a Hilbert space endowed with the norm of element u being the square root of what is given by (3.17).

We want to prove that for $u \in C^{1,2}$, we have

$$E[u(T, x., x_T) - u(0, x., x_0)] \leq$$

$$\int_0^T E[\frac{\partial u}{\partial t}(t, x_t) + G_t(u_{x^i x^j}(t, x., x_t), u_{x^i}(t, x., x_t))]dt. \tag{3.18}$$

Notice that according to (3.15) we have the inequality

$$E[u(t, x., x_t) - u(s, x., x_s)] = E[u(t, x., x_t) - u(s, x., x_t)] + E[u(s, x., x_t) - u(s, x., x_s)]$$

$$\leq \int_s^t E\frac{\partial u}{\partial t}(r, x., x_t)dr + E\int_s^t G_r(u_{x^i x^j}(s, x., x_r), u_{x^i}(s, x., x_r), x.)dr.$$

Also notice that the second term of the right-hand side of the above inequality is absolutely continuous in t because of

$$\int_s^t EG_r(u_{x^i x^j}(s, x., x_r), u_{x^i}(s, x., x_r), x.)dr =$$

$$\int_s^t E \sup_{(a,b)\in A_r(x.)} (\sum_{i,j=1}^d a_{ij} u_{x^i x^j}(s, x., x_r) + \sum_{i=1}^d b_i u_{x^i}(s, x., x_r))dr \leq$$

$$\int_s^t E[\sum_{i,j=1}^d |u_{x^i x^j}| L(r)(1 + \sup_{u\leq r} |x_u|^2) + \sum_{i=1}^d |u_{x^i}| L(r)(1 + \sup_{u\leq r} |x_u|)]dr \leq$$

$$M(t-s)(1 + E\sup_{t\leq T} |x_t|^2),$$

for some constant $M > 0$. In the above, the first inequality follows from **(Linear growth II)**, and the last term is bounded by (3.14). One can also take $-u$ instead of u in the above inequality and then see that $Eu(t, x., x_t)$ is absolutely continuous with respect to t and its derivative is no greater than the integrand in (3.18). So we can get (3.18).

Step 2. Take a nonnegative test function $\zeta \in C_0^\infty(R)$ so that $\zeta(t) = 0$ for $t \leq 0$, $\int \zeta dt = 1$, and for $\gamma > 0$ define

$$\zeta_\gamma(t, x) = \gamma^{-d-1}\zeta(\frac{t}{\gamma})\prod_{i=1}^d \zeta(\frac{x^i}{\gamma}).$$

Considering all the functions u in L_2 (which is defined in step 1) defined for $t > 0$ and equal to zero for $t \leq 0$, we define an operator T_γ on L_2 by

$$(T_\gamma u)(t, y., x) := \int u(t-s, y., x-z)\zeta_\gamma(s, z)dsdz, \qquad t \leq T.$$

The well-known properties of convolutions imply that T_γ is bounded as an operator from L_2 to L_2 and its norm is less than 1. Moreover, the function $T_\gamma u$ is infinitely differentiable with respect to t, x for every $y. \in C$ and

$$\sup_{t\le T} E[\sum_{i,j=1}^{d} |(T_\gamma u)_{x^i x^j}| + \sum_{i=1}^{d} |(T_\gamma u)_{x^i}| + |T_\gamma u| + |\frac{\partial}{\partial t} T_\gamma u|](t, x., x_t) \le N||u||_{L_2}, \tag{3.19}$$

where the constant N is independent of u. Finally, inequality (3.18) remains true if u is replaced by $T_\gamma v$ for every $v \in L_2$. This fact follows from (3.19) by approximating the function v by bounded functions.

Step 3. For all $\gamma > 0$ we define K_γ as the set of all linear functions l on L_2 of the type

$$l(u) := ET_\gamma u(T, x., x_T) - E\int_0^T LT_\gamma u(t, x., x_t)dt,$$

where

$$Lv(t, x., x_t) := \frac{\partial v}{\partial t}(t, x., x_t) + \sum_{i,j=1}^{d} a^{ij}(t, x.)v_{x^i x^j}(t, x., x_t) + \sum_{i=1}^{d} b^i(t, x.)v_{x^i}(t, x., x_t),$$

and $(a, b)(t, y.)$ is measurable in $(t, y.)$, N_t-measurable in $y.$ for every t, and $(a, b)(t, y.) \in A_t(y.)$ for almost all (t, ω) for all $y.$. The set K_γ is clearly convex and bounded according to step 2.

Next we show that it is closed in weak topology. Let $l_n \in K_\gamma$ and $l_n \to l$ on L_2 as $n \to \infty$. Define L_n as the operator corresponding to l_n and (a_n, b_n) as the coefficients of L_n. According to **(Linear growth II)** on $A_t(x.)$ and the assumption (3.14), the norms of $(a_n, b_n)(t, x.)$ in $L_2([0, T] \times \Omega)$ are uniformly bounded. Banach-Saks theorem implies that there exits a subsequence $n_k \to \infty$ such that the arithemetic means of $(a_{n_k}, b_{n_k})(t, x.)$ converge in $L_2([0, T] \times \Omega)$. Without loss of generality we can suppose that $n_k = k$. For some subsequence $m_k \to \infty$ these means converge for almost all (t, ω). Now, we define

$$(a, b)(t, y.) := \lim_{k\to\infty} \frac{1}{m_k} \sum_{i=1}^{m_k} (a_i, b_i)(t, y.) \tag{3.20}$$

for those $(t, y.)$ for which the right-hand side of (3.20) is well-defined and for remaining $(t, y.)$ let $(a, b)(t, y.)$ be equal to a fixed point in $A_t(y.)$(??). Indeed, the (a, b) is measurable and N_t-measurable with respect to $y.$ for every t.

Moreover, (3.20) holds for almost all (t, ω) if $y.$ is replaced by $x.(\omega)$. It also follows that $(a, b)(t, x.) \in A_t(x.)$ $(a.e.(t, \omega))$, since we assume that the set $A_t(y.)$ is convex and closed for t and $y.$. This implies that $l = \lim_n l_n$ is of the given form and $l \in K_\gamma$.

Step 4. Prove that for all $\gamma > 0$, $u \in L_2$

$$\min_{l\in K_\gamma} l(u) = ET_\gamma u(T, x., x_T) - E\int_0^T [\frac{\partial T_\gamma u}{\partial t} + G_t(T_\gamma u_{x^i x^j}, T_\gamma u_{x^i})(t, x., x_t)]dt \le 0. \tag{3.21}$$

Indeed, the inequality is discusses in step 2. To prove the equality, we note that the minimum in (3.21) is attained in view of weak compactness of K_γ and we take in $R^{d_1} = \{(a, b)\}$ a dense countable subset $\{(a_r, b_r) : r = 1, 2, \cdots\}$. For $r = 1, 2, \cdots$, we define $(a_r, b_r)(t, y.)$ as the closest point in $A_t(y.)$ to (a_r, b_r). The sets $A_t(y.)$ are closed convex and measurable and this implies (??) that $(a_r, b_r)(t, y.)$ is well-defined and is measurable in an appropriate sense.

It is readily seen that for all $t \in [0, T]$, $y. \in C$,

$$\lim_{n\to\infty} \max_{r\le n}[\sum_{i,j=1}^d a_r^{i,j}(t, y.)(T_\gamma u)_{x^i x^j}(t, y., y_t) + \sum_{i=1}^d b_r^i(t, y.)(T_\gamma u)_{x^i}(t, y., y_t)]$$

$$= \sup_r[\sum_{i,j=1}^d a_r^{i,j}(t, y.)(T_\gamma u)_{x^i x^j}(t, y., y_t) + \sum_{i=1}^d b_r^i(t, y.)(T_\gamma u)_{x^i}(t, y., y_t)]$$

$$= G_t((T_\gamma u)_{x^i x^j}(t, y., y_t), (T_\gamma u)_{x^i}(t, y., y_t), y.) \tag{3.22}$$

By the measurable choice of their ordering number, we can construct measurable $(\tilde{a}_n, \tilde{b}_n)(t, y.) \in A_t(y.)$ such that

$$\sum_{i,j=1}^d \tilde{a}_n^{i,j}(t, y.)(T_\gamma u)_{x^i x^j}(t, y., y_t) + \sum_{i=1}^d \tilde{b}_n^i(t, y.)(T_\gamma u)_{x^i}(t, y., y_t)$$

$$= \max_{r\le n}[\sum_{i,j=1}^d a_r^{i,j}(t, y.)(T_\gamma u)_{x^i x^j}(t, y., y_t) + \sum_{i=1}^d b_r^i(t, y.)(T_\gamma u)_{x^i}(t, y., y_t)]$$

converges to the last expression in (3.22) for all $t, y.$. Hence

$$\lim_{n\to\infty} E[\int_0^T [\sum_{i,j=1}^d \tilde{a}_n^{i,j}(t, x.)(T_\gamma u)_{x^i x^j}(t, x., x_t) + \sum_{i=1}^d \tilde{b}_n^i(t, x.)(T_\gamma u)_{x^i}(t, x., x_t)]dt$$

$$= E\int_0^T G_t(T_\gamma u_{x^i x^j}, T_\gamma u_{x^i})(t, x., x_t)dt.$$

Thus the first expression in (3.21) is not greater than the second one. The opposite inequality is obvious and (3.21) is proved.

Step 5. In the reflexive space L_2 the unit sphere is weakly compact, K_γ is weakly compact, the function $l(u)$ is bilinear and weakly continuous in l and in u and by the Ky Fan theorem and by (3.21)

$$0 \ge \max_{||u||\le 1} \min_{l\in K_\gamma} l(u) = \min_{l\in K_\gamma} \max_{||u||\le 1} l(u) = \max_{||u||\le 1} l_\gamma(u)$$

for some $l_\gamma \in K_\gamma$. We see that $l_\gamma(u) \le 0$ for $||u|| \le 1$. The linearlity of l implies that $l_\gamma(u) = 0$ for all $u \in L_2$.

Step 6. Let L_γ be the operator corresponding to l_γ. Repeating in the third step and using the weak limits of the coefficients of L_γ as $\gamma \to 0$ it is not difficult to construct an measurable function $(a,b)(t,y.) \in A_t(y.)$ N_t-measurable in y. for every t such that

$$Eu(T,x.,x_T) = E\int_0^T [\frac{\partial u}{\partial t} + \sum_{i,j=1}^d a^{ij}u_{x^ix^j} + \sum_{i=1}^d b^i u_{x^i}](t,x.,x_t)dt + Eu(0,x.,x_0) \tag{3.23}$$

for every $u \in C^{1,2}$ which equals zero for $t=0$. Substitute $u(t,y.,x)\zeta_n(t)$ instead of u, where $\zeta_n(t) = \zeta(nt)$, $\zeta(t)$ is infinitely differentiable, $\zeta(t) = 0$ for $t \le 0$, $\zeta(t) = 1$ for $t \le 1$, $\zeta' \ge 0$. Note that $\zeta_n(t) \to 1$ as $n \to \infty$, for $t > 0$

$$\lim_{n\to\infty} E\int_0^T \zeta_n'(t)u(t,x.,x_t)dt = \lim_{n\to\infty} E\int_0^{\frac{1}{n}} \zeta_n'(t)u(t,x.,x_t)dt$$

$$= \lim_{n\to\infty} E\int_0^{\frac{1}{n}} \zeta_n'(t)dt \times u(0,x.,x_0) = Eu(0,x.,x_0)$$

if $u \in C^{1,2}$. Then we see that (3.23) holds true for all $u \in C^{1,2}$.

Step 7. Let $\tau(y.)$ be a N_t-Markovian time, $\tau \le T$, $u(x)$ be a function on R^d twice continuously differentiable in x and bounded with derivatives up to the second order. Take ζ_n from step 6 and define

$$u_n(t,y.,x) = u(x)(1 - \zeta_n(t - \tau(y.))).$$

The set $\{y. : \zeta_n(t - \tau(y.)) < c\}$ for every c, if it is nonempty, can be represented as $\{y. : t - \tau(y.) < h\}$ with $h \ge 0$. Hence u_n is N_t-measurable and by (3.23) we obtain

$$Eu(x_0) = \lim_{n\to\infty}\{u(x_T)(1-\zeta_n(t-\tau(y.))) - \int_0^T (1-\zeta_n(t-\tau))[\sum_{i,j=1}^d a^{ij}(t,x.)u_{x^ix^j}(x_t) +$$

$$\sum_{i=1}^d b^i(t,x.)u_{x^i}(x_t)]dt + \int_0^T \zeta_n'(t-\tau)u(x_t)dt\}$$

$$= E\{u(x_\tau) - \int_0^\tau [\sum_{i,j=1}^d a^{ij}(t,x.)u_{x^ix^j}(x_t) + \sum_{i=1}^d b^i(t,x.)u_{x^i}(x_t)]dt\}. \tag{3.24}$$

Step 8. The equality between extreme terms in (3.24) for every stopping time $\tau \le T$ implies that the process

$$u(x_t) - \int_0^t [\sum_{i,j=1}^d a^{ij}(s,x.)u_{x^ix^j}(x_s) + \sum_{i=1}^d b_i(s,x.)u_{x^i}(x_s)]ds, \ t \le T,$$

is martingale and the assertion of our theorem now follows from the Doob theorem [16] and the Strook-Varadhan theorem [32]. The theorem is thus proved.

4 Proof of the main theorems

In the defintion in section 2, for all n, the semimartingales X^n are defined on different complete probability spaces $(\Omega^n, \mathcal{F}^n, P^n)$. By Skorokhod representation theorem, we can find a complete probability space $(\Omega, \mathcal{F} = (\mathcal{F}_t)_{t\geq 0}, P)$, and a sequence of semimartingales Y^n on $(\Omega, \mathcal{F}, P)$ such that for all n, X^n and Y^n are the same in law. For convenience, we consider a sequence of semimartingales in the same probability space, and conditions **(N)**, **(sup B)** and **(sup A)** imply the following conditions, respectively, if assuming that X^n, for all n, are in the same probability space : for each $T > 0$, for all $\epsilon > 0$,
Condition (N'): *For all $a \in (0, 1]$,*

$$\lim_n P(\int_0^T \int_{|x|>a} d\nu^n \geq \epsilon) = 0.$$

Condition (sup B'):

$$\lim_n P(\sup_{t\leq T} |B_t^n - \int_0^t b(s, X^n) ds| \geq \epsilon) = 0.$$

Condition (sup A'):

$$\lim_n P(\sup_{t\leq T} | < M^{a,n} >_t - \int_0^t a(s, X^n) ds| \geq \epsilon) = 0.$$

Before proving theorem 2.1, and theorem 2.2, we prove some lemmas.

Lemma 4.1 *Assume on a complete probability space $(\Omega, \mathcal{F}, P)$, given a sequence of semimartingales*

$$X_t^n = X_0^n + B_t^n + M_t^{a,n} + \int_0^t \int_{|x|>a} x d\mu^n, \tag{4.25}$$

where ν^n (which is the compensator of μ^n), B_t^n, and $M_t^{a,n}$, satisfy conditions **(N')**, **(sup $B^{'}$)**, *and* **(sup $A^{'}$)**, *respectively. Futhermore, assum that*

$$E|X_0^n|^2 < N,$$

for some constant N not depending on n. Then for fixed $T > 0$, we have

$$E \sup_{t\leq T} |X_t^n|^2 \leq K,$$

for some constant $K > 0$ not depending on n.

Remark 4.1 *In this case, for each n, we say that X_t^n is locally square integrable. By theorem 11.31 in [26], we thus have: (1) for all $t > 0$, $x^2 * \nu_t < \infty$, (2) $X_t^n = X_0^n + m_t^n + \alpha_t^n$, where α_t^n is a predictable process with finite variation, and m_t^n is a locally square-integrable martingale.*

Proof. For brevity we will use the notation

$$(Z_t)^* := \sup_{s \leq t} |Z_s|.$$

We first consider the case when X_t^n is a special semimartingale. From (2.3) we get

$$((X_t^n)^*)^2 \leq 3[(X_0^n)^2 + ((B_t^n)^* + \int_0^t \int_{|x|>a} |x| d\mu^n)^2 + (M_t^{a,n})^*)^2] \tag{4.26}$$

Let us evaluate the right-hand side of (4.26). From **(Linear growth (I))**, it follows that

$$(B_t^n)^* + \int_0^t \int_{|x|>a} |x| d\mu^n$$

$$\leq (\int_0^t b(s, X^n) ds)^* + (B_t^n - \int_0^t b(s, X^n) ds)^* + \int_0^t \int_{|x|>a} |x| d\mu^n$$

$$\leq \int_0^t L(s)(1 + (X_s^n)^*) ds + (B_t^n - \int_0^t b(s, X^n) ds)^* + \int_0^t \int_{|x|>a} |x| d\mu^n. \tag{4.27}$$

The expectation of the second term of the right-hand side of (4.27) is uniformly bounded according to **(sup B')** . The third term of the right-hand side of (4.27)

$$\int_0^t \int_{|x|>a} |x| d\mu^n \tag{4.28}$$

converges to zero in probability according to **(N')** (see formula (50) in [20]). Thus the expectation of the (4.28) is uniformly bounded.

According to Doob inequality (see Theorem 1.9.2 in [17])

$$E((M_t^{a,n})^*)^2 \leq 4E < M^{a,n} >_t .$$

To estimate $(M_t^{a.n})^*$, we can consider

$$< M^{a,n} >_t \leq (\int_0^t a(s, X^n) ds)^* + (< M^{a,n} >_t - \int_0^t a(s, X^n) ds)^*.$$

$$\leq (\int_0^t L(s)(1 + |(X_s^n)^*|^2) ds + +(< M^{a,n} >_t - \int_0^t a(s, X^n) ds)^*, \tag{4.29}$$

where the second inequality follows from **(Linear growth (I))**. We also notice that the expectation of the second term of the right-hand side of (4.29) is uniformly bounded according to **(sup** A'**)**.

The inequalities (4.27) (4.29) imply that

$$E((X_t^n)^*)^2 \leq K_1 + K_2 \int_0^t E((X_s^n)^*)^2 ds, \qquad \text{for } t \leq T,$$

with some constants $K_1 > 0$ and $K_2 > 0$ not depending on n. If $\int_0^t E((X_s^n)^*)^2 ds < \infty$, $t \leq T$, then by the Gronwall-Bellman inequality

$$E((X_T^n)^*)^2 \leq K_1 \exp(K_2 T). \tag{4.30}$$

If X_t^n is a semimartingale, we define stopping times

$$\tau_k^n := \inf_t \{|X_t^n| > k\},$$

and consider stopping processes $X_{t\wedge\tau_k^n}^n$, which are special semimartingales. By (4.30) we thus have

$$E((X_{T\wedge\tau_k^n}^n)^*)^2 \leq K_1 \exp(K_2 T).$$

By Fatou's lemma we have

$$E((X_T^n)^*)^2 = E \liminf_k ((X_{T\wedge\tau_k^n}^n)^*)^2$$

$$\leq \liminf_k E((X_{T\wedge\tau_k^n}^n)^*)^2 \leq K_1 \exp(K_2 T).$$

The lemma is thus proved.

Lemma 4.2 *Assume the conditions of lemma 4.1 are satisfied. We also assume that the processes* $(X_t^n)_{n\geq 1}$ *satisfy*

$$\lim_{n\to\infty} X_t^n = X_t \text{ uniformly on } [0,T] \qquad P - a.s. \tag{4.31}$$

for some stochastic process X_t*, which is continuous in* t*,* $P - a.s.$*. We also assume the function* $(a(t,y.), b(t,y.))$ *in conditions* **(sup** B'**)** *and* **(sup** A'**)** *satisfying* $(a(t,y.), b(t,y.)) \in A_t(y.)$ *for* $t \in [0,T]$*, and* $y. \in C$*, where* $A_t(y.)$ *is defined in section 3.2.*

Then there exist N_t*-adapted function* $(\tilde{a}_t(y.), \tilde{b}_t(y.)) \in A_t(y.)$ *for* $t \in [0,T]$*, Furthermore, there exists an enlargement of probability space* $(\Omega, \mathcal{F}, P)$*, and a d-dimensional Wiener process* w_t *on the enlargment space such that for every* $s \in [0,T]$ *the process* w_t *regarded on* $[0,s]$ *and* $w_t - w_s$ *regarded on* $[s,T]$ *are independent, and*

$$X_t = X_0 + \int_0^t \tilde{b}_r(X.)dr + \int_0^t \sqrt{\tilde{a}_r(X.)}dw_r \qquad \forall t \in [0,T] \qquad (a.s.). \tag{4.32}$$

Proof. This lemma is an application of theorem 3.2. Since lemma 4.1 and (4.31) imply that (3.14) is satisfied, we only need to check (3.15).

For $u \in C_0^\infty(R^d)$, for $0 \le s \le t \le T$, applying Itô's formula (see corollary 11.27 in [26] or theorem 27.1 in [31]) to equation (4.25), along with (N'), we get

$$u(X_t^n) - u(X_s^n)$$

$$= \int_s^t \sum_{i=1}^d \frac{\partial u}{\partial x_i}(X_{r-}^n) dB_{i,r}^n + \frac{1}{2}\int_s^t \sum_{i,j=1}^d \frac{\partial^2 u}{\partial x_i x_j}(X_{r-}^n) d[M_i^{a,n}, M_j^{a,n}]_r$$

$$+[u(X_u^n) - u(X_{u-}^n) - \sum_{i=1}^d \frac{\partial u}{\partial x^i}(X_{u-}^n)x_i - \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2 u}{\partial x_i x_j}(X_{u-}^n)x_i x_j]I_{|x|<a})I_{s<u\le t} * \nu^n$$

$$+m^n(t) - m^n(s),$$

where

$$m^n(t) = \int_0^t \sum_{i=1}^d \frac{\partial u}{\partial x_i}(X_{u-}^n) dM_i^{a,n} \tag{4.33}$$

is a locally martingale.

For $0 \le t_1 \le, \cdots, \le t_m \le s \le t \le T$, and $u, f_1, \cdots, f_m \in C_0^\infty(R^d)$, and $f_i \ge 0$, $i = 1, \cdots, d$, we consider

$$f_1(X_{t_1}^n)\cdots f_m(X_{t_m}^n)(u(X_t^n) - u(X_s^n)) - f_1(X_{t_1}^n)\cdots f_m(X_{t_m}^n)(m^n(t) - m^n(s))$$

$$= f_1(X_{t_1}^n)\cdots f_m(X_{t_m}^n)[\int_s^t \sum_{i=1}^d \frac{\partial u}{\partial x_i}(X_{r-}^n) dB_{i,r}^n]$$

$$+f_1(X_{t_1}^n)\cdots f_m(X_{t_m}^n)[\frac{1}{2}\int_s^t \sum_{ij=1}^d \frac{\partial^2 u}{\partial x_i x_j}(X_{r-}^n) d[M_i^{a,n}, M_j^{a,n}]_r)$$

$$+f_1(X_{t_1}^n)\cdots f_m(X_{t_m}^n)[(u(X_u^n) - u(X_{u-}^n) - \sum_{i\le d} \frac{\partial u}{\partial x^i}(X_{u-}^n)x_i$$

$$-\frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2 u}{\partial x_i x_j}(X_{u-}^n)x_i x_j)I_{|x|<a}I_{s<u\le t} * \nu^n]. \tag{4.34}$$

We denote the left-hand side of the above equality by (A), and the first, the second, and the third term of the right-hand side by (B), (C), and (D), respectively. We take expectation then take limits on terms (A)-(D), and want to estimate them seperately.

For (A), we have

$$\lim_n E(A) = \lim_n E[f_1(X_{t_1}^n)\cdots f_m(X_{t_m}^n)(u(X_t^n) - u(X_s^n))]$$

$$+\lim_n E[f_1(X^n_{t_1})\cdots f_m(X^n_{t_m})(m^n(s)-m^n(t))]. \tag{4.35}$$

By bounded convergence theorem, the first term of the right-hand side of (4.35) converges to

$$E[f_1(X_{t_1})\cdots f_m(X_{t_m})(u(X_t)-u(X_s))]$$

since X^n_t converge to X_t uniformly in $[0,T]$ as $n\to\infty$, and X_t is continuous in t (a.s.). The second term of the right-hand side of (4.35) can be estimated by localization. Let

$$\tau^n_k=\inf_t\{|m^n_t|>k\},$$

then for all n and k, we thus have

$$E[f_1(X^n_{t_1\wedge\tau^n_k})\cdots f_m(X^n_{t_m\wedge\tau^n_k})(m^n(s\wedge\tau^n_k)-m^n(t\wedge\tau^n_k)]=0,$$

according to lemma 8.5 (a) in Chapter 2 in [30]. Notice that (4.33), **(Linear growth(I))**, **(sup A')**, and Doob's inequality imply

$$E\sup_{t\le T}\sup_{t\le T}|m^n_t|\le KE\sup_{t\le T}<M^{a,n}>_t$$

$$\le K_1E\sup_{t\le T}\int_0^t(1+\sup_{r\le t}|X^n_r|^2)dr<N$$

for constants $K,K_1>0$, for some constant N not depending on n, and the last inequality follows from lemma 4.1. By Fatou's lemma we thus have

$$\lim_n E[f_1(X^n_{t_1})\cdots f_m(X^n_{t_m})(m^n(s)-m^n(t)]$$

$$=\lim_n E\liminf_k[f_1(X^n_{t_1\wedge\tau^n_k})\cdots f_m(X^n_{t_m\wedge\tau^n_k})(m^n(s\wedge\tau^n_k)-m^n(t\wedge\tau^n_k)]$$

$$\le\lim_n\liminf_k E[f_1(X^n_{t_1\wedge\tau^n_k})\cdots f_m(X^n_{t_m\wedge\tau^n_k})(m^n(s\wedge\tau^n_k)-m^n(t\wedge\tau^n_k)]=0.$$

We now estimate $\lim_n E(B)$. Since B^n_t is with finite variation and $\frac{\partial u}{x_i}(X^n_t)$ is a semimartingale, by corollary 9.35 in [26] we have, for $i=1,\cdots,d$,

$$f_1(X^n_{t_1}),\cdots,f_m(X^n_{t_m})[\int_s^t\frac{\partial u}{\partial x_i}(X^n_{r-})dB^n_{i,r}-\int_s^t\frac{\partial u}{\partial x_i}(X^n_{r-})b_i(r,X^n)dr]$$

$$=f_1(X^n_{t_1}),\cdots,f_m(X^n_{t_m})[\frac{\partial u}{\partial x_i}(X^n_{r-})(B^n_{i,r}-\int_s^t b_i(u,X^n)du)]|_s^t$$

$$-\int_s^t(B^n_{i,r}-\int_s^t b_i(u,X^n)du)d(\frac{\partial u}{\partial x_i}(X^n_r)).$$

$$\le B_1[\sup_{t\le T}|B^n_t-\int_0^t b(r,X^n)dr|],\qquad (a.s.) \tag{4.36}$$

for some constant $B_1 > 0$, the last inequality is because $\frac{\partial u}{\partial x_i}$, $f_1, \cdots, f_m$ are bounded. By **(sup B')**, we thus have

$$\lim_n E(B) = \lim_n E[\int_s^t \sum_{i=1}^d \frac{\partial u}{\partial x_i}(X_{r-}^n) b_i(r, X^n) dr].$$

(Linear growth (I)) and lemma 4.1 imply

$$E[\int_s^t \frac{\partial u}{\partial x_i}(X_{r-}^n) b_i(r, X^n) dr] \le \int_s^t L(r)(1 + E \sup_{u \le r} |X_u^n|) dr < N,$$

with N not depending on n. We thus can apply dominated convergence theorem to the above formula to get

$$\lim_n E(B) = E \lim_n [\int_s^t \frac{\partial u}{\partial x_i}(X_{r-}^n) b_i(r, X^n) dr].$$

For $\lim_n E(C)$, using the same argument as in estimating $\lim_n E(B)$, we have the following estimate, for $i, j = 1, \cdots, d$,

$$f_1(X_{t_1}^n) \cdots f_m(X_{t_m}^n)[\int_s^t \frac{\partial^2 u}{\partial x_i x_j}(X_{r-}^n) d[M_i^{a,n}, M_j^{a,n}]_r - \int_s^t \frac{\partial^2 u}{\partial x_i x_j}(X_{r-}^n) a_{ij}(r, X^n) dr$$

$$\le B_2 [\sup_{t \le T} |[M^{a,n}]_t - \int_0^t a(r, X^n) dr|] \qquad (a.s.), \tag{4.37}$$

for some constant $B_2 > 0$. We also notice (N') implies that for $\epsilon > 0$,

$$\lim_n P(\sup_{t \le T} |[M^{a,n}]_t - < M_t^{a,n} | > \epsilon) = 0,$$

according to formula (4.11) in article [18]. By **(sup A')** **(Linear growth (I))** (3.14) and dominated convergence theorem, we thus have

$$\lim_n E(C) = E \lim_n [\int_s^t \sum_{i,j=1}^d \frac{\partial^2 u}{\partial x_i x_j}(X_{r-}^n) a_{ij}(r, X^n) dr].$$

For $\lim_n E(D)$, we first notice that by Taylor expansion

$$(u(X_u^n) - u(X_{u-}^n) - \sum_{i=1}^d \frac{\partial u}{\partial x^i}(X_{u-}^n) x_i - \frac{\partial^2 u}{\partial x^i x^j}(X_{u-}^n) x_i x_j) I_{|x|<a} I_{s<u\le t} * \nu^n$$

$$\le B_3 |x|^3 I_{|x|<a} I_{s<u\le t} * \nu^n,$$

for some constant $B_3 > 0$. We also notice for all $\epsilon \in (0, 1]$,

$$\int_s^t \int_{|x|<\epsilon} |x|^3 \nu^n(du, dx) \le \epsilon \int_s^t \int_{R^d} (|x|^2 \wedge 1) \nu^n(du, dx) \le B_4 \epsilon,$$

for some constant B_4 not depending on ϵ. By (2.2), it follows

$$(D) \leq B_3 \int_s^t \int_{|x|<a} |x|^3 \nu^n(du, dx) = B_5 (\int_s^t \int_{\epsilon<|x|<a} |x|^3 \nu^n(du, dx) + \epsilon).$$

Applying **(N')** to the above term implies that (D) converges to zero in probability. We thus prove $\lim_n E(D)$ is equal to zero.

Applying the above estimates on (A)-(D), we have

$$E[f_1(X_{t_1}) \cdots f_m(X_{t_m})(u(X_t) - u(X_s))]$$

$$\leq E \limsup_n f_1(X_{t_1}^n) \cdots f_m(X_{t_m}^n)[\int_s^t \sum_{i,j=1}^d \frac{\partial^2 u}{\partial x_i x_j}(X_{r-}^n) a_{ij}(r, X^n)$$

$$+ \sum_{i=1}^d \frac{\partial u}{\partial x_i}(X_{r-}^n) b_i(r, X^n) dr].$$

$$= E f_1(X_{t_1}) \cdots f_m(X_{t_m}) \int_s^t \limsup_n [\sum_{i,j=1}^d \frac{\partial^2 u}{\partial x_i x_j}(X_{r-}^n) a_{ij}(r, X^n)$$

$$+ \sum_{i=1}^d \frac{\partial u}{\partial x_i}(X_{r-}^n) b_i(r, X^n) dr]. \tag{4.38}$$

We choose $A_t(X.)$ such that, for all $t \in [0, T]$,

$$(\limsup_n |a_{ij}(t, X^n)|, \limsup_n |b_i(t, X^n)|) \in A_t(X.).$$

Notice that **(Linear growth (II)** on $A_t(X.)$ still can be achieved, since $a_{ij}(t, X.)$ and $b(t, X.)$ satisfy **(Linear growth (I))**. According to the definition of G and notice that it is a upper-continuous function, we thus have

$$E[f_1(X_{t_1}) \cdots f_m(X_{t_m})(u(X_t) - u(X_s))]$$

$$\leq E[f_1(X_{t_1}) \cdots f_m(X_{t_m}) \int_s^t G_r(u_{ij}(X_r), u_i(X_r), X.) dr]$$

The inequality (3.15) is therefore proved. Applying theorem 3.2, we thus prove this lemma.

Corollary 4.1 *In addition to conditions in lemma 4.2, if we further assume that the function $a(t, X)$ in* **(sup A')** *satisfies* **(Coercive)**, *then the coefficient $\tilde{a}$ in (4.32) also satisfies* **(Coercive)**.

Proof. For $x \in A$, the first d^2 elements represent a nonnegative $d \times d$ matrix, which is denoted by a_x. We define a set

$$B := \{x = (x_1, \cdots, x_{d^2}, \cdots, x_{d_1}) : \gamma \leq (a_x \lambda, \lambda) \ \forall \lambda : R^d\text{-valued unit vector}\},$$

where γ is taken from **(Coercive)**. For all t, and $y. \in C$, we define sets in R^{d_1}

$$\hat{A}_t(y.) = A_t(y.) \cap B,$$

where $A_t(y.)$ is as in lemma 4.2. The sets $\hat{A}_t(y.)$ is still closed, since $A_t(y.)$ closed, and B is closed. The set B is convex since for every $0 \leq \beta \leq 1$ we have the inequality

$$\gamma|\lambda|^2 \leq \beta(a_1\lambda, \lambda) + (1-\beta)(a_2\lambda, \lambda) = ((\beta a_1 + (1-\beta)a_2)\lambda, \lambda),$$

for all R^d-valued unit vector λ. Therefore, $\hat{A}_t(y.)$ are also convex. Theorem 4.2 is thus fulfilled with policy sets $\hat{A}_t(y.)$, so the pair of coefficients $(\tilde{a}_t(y.), \tilde{b}_t(y.))$ in (4.32) is belonging to $\hat{A}_t(y.)$. Therefore, the coefficient $\tilde{a}_t$ satisfies **(Coercive)**. The corollary is thus proved.

Lemma 4.3 *Let $T > 0$, assume that on a complete probability space $(\Omega, \mathcal{F}, P)$, there exists a diffusion process X_t which satisfy*

$$X_t = X_0 + \int_0^t \tilde{b}_s ds + \int_0^t \sqrt{\tilde{a}_s} dw_s \ \forall t \leq T, \ P-a.s. \tag{4.39}$$

where $\tilde{a}_s$ satisfies **(Coercive)** *and $\tilde{b}_s$ is bounded. We also assume a bounded domain D in R^d, and $X_0 \in D$. The first exit time of the process (4.39) from the domain D is denoted by τ_D. For a measurable subset Γ in $[0,T] \times D$, the random subsets in $D_\Gamma \in (0,T]$ is defined to be*

$$D_\Gamma := \{t : t \in [0, \tau_D \wedge T), (t, X_t) \in \Gamma\},$$

which is the time that the process (t, X_t) stays in the set Γ before exiting from the domain $[0,T] \times D$.

Then for every $C > 0$, there is a constant $N = N(C, D, d, \gamma)$ such that

$$P(|D_\Gamma| > C) \leq N|\Gamma|, \tag{4.40}$$

where $|\Gamma|$ denotes Lebesgue measure of the set Γ.

Proof. We use theorem 3.1 to prove this lemma. We can choose a function $F \equiv K(1 \vee \gamma)$, where K is an upper bound of $|\tilde{b}_t|$. Since $\tilde{a}$ is coercive, F is a regular function (see Definition 3.1). We choose $g(t,x)$ in theorem 3.1 to be $I_\Gamma(t,x)$, the indicator function on the set Γ. Therefore, we can get

$$E[\int_0^{\tau_D} \det(\tilde{a}_s)^{\frac{1}{d+1}} I_\Gamma(s, X_s) ds] \leq N|\Gamma|.$$

(Coercive) implies $det(\tilde{a}_s)$ is strictly larger than zero for all $s \in [0,T]$. We thus can get from the above inequality

$$E[|D_\Gamma|] \leq N_1|\Gamma|,$$

for some constant $N_1 > 0$. Applying Chebychev's inequality on the above inequality, we can get for all $C > 0$

$$P(|D_\Gamma| > C) \leq \frac{E[|D_\Gamma|]}{C} \leq N_1|\Gamma|$$

for some constant $N_1 = N_1(C, D, d, \gamma) > 0$. The lemma is thus proved.

We now prove the main results.

Proof of theorem 2.1. The proof of part (A), and (B) can be found in Theorem 3.1 in Chapter 8 in [17]. We only prove part (C). According to part (A), there exists a subsequence of $(X_t^n)_{n\geq 1}$, say $(X_t^{n'})_{n'\geq 1}$, with distribution $Q^{n'}$ converge to a probability measure Q'. By using Skorokhod representation theorem (see [27] or Theorem 15.42 in [26]), we can find a probability space, say $(\Omega', \mathcal{F}', P')$, and D^d- valued semimartingales $\hat{X}^{n'}$ and $\hat{X}$ on this probability space Ω', where $\hat{X}^{n'}$ has distributions $Q^{n'}$ and $\hat{X}$ have distribution Q'. Moreover,

$$\lim_{n'} \hat{X}^{n'} = \hat{X} \qquad P' - a.s. \tag{4.41}$$

in the Skorokhod-Lindvall topology. From part (A), we also have $Q'(C) = P'(\hat{X} \in C) = 1$. Thus, we have

$$\lim_{n'} \hat{X}^{n'}(t) = \hat{X}(t) \tag{4.42}$$

uniformly in $[0,T]$ P'-a.s. (see Ch.6 Sec.1 Problem 5 in [17]).

For fixed $a \in (0,1]$, the process $(\hat{X}_{n'})_{n'\geq 1}$ also sdmit a uniqe canonical representation with the triplet of predictable characteristics $T = (\hat{B}^{n'}, \hat{A}^{n'}, \hat{\nu}^{n'})$. The process $\hat{X}^{n'}$ can be written as

$$\hat{X}_t^{n'} = \hat{X}_0^{n'} + \hat{B}_t^{n'} + \hat{M}_t^{a,n'} + \int_0^t \int_{|x|>a} x d\hat{\nu}^{n'}.$$

The process $X^{n'}$ and $\hat{X}^{n'}$ are the same in law. So **(N), (sup A), (sup B)** imply **(*N'*), (sup *A'*), (sup *B'*)**, respectively (see Ch.4 Sec.3 Theorem 2 in [17]).

Notice that the coefficients $a(t,x)$ and $b(t,x)$ determine the distribution of the weak solution of (2.9). Therefore, by Theorem 1 (3) in [19], to prove that $X^n(t)$ converges weakly to the solution of (2.9), it suffices to prove that for fixed $T > 0$ and for all $\epsilon > 0$

$$\lim_{n'} P'[\sup_{t\leq T} |\hat{B}_t^{n'} - \int_0^t b(s,\hat{X})ds| > \epsilon] = 0; \tag{4.43}$$

and

$$\lim_{n'} P'[\sup_{t\le T} | < M^{a,n'} >_t - \int_0^t a(s,\hat{X})ds| > \epsilon] = 0. \tag{4.44}$$

We observe that the integrand in (4.43) satisfies

$$|\hat{B}_t^{n'} - \int_0^t b(s,\hat{X})ds| \le |\hat{B}_t^{n'} - \int_0^t b(s,\hat{X}^{n'})ds| + |\int_0^t b(s,\hat{X}^{n'}) - b(s,\hat{X})ds|.$$

We also notice the function $b(t,X)$ satifying **(S)**. Along with **(sup B')**, to prove (4.43), we only need to prove that for all $T>0$ and for all $\epsilon>0$,

$$\lim_{n'} P'(\sup_{t\le T} \int_0^t |b(s,\hat{X}_s^{n'}) - b(s,\hat{X}_s)|ds > \epsilon) = 0. \tag{4.45}$$

Notice that first of all, according to corollary 4.1 we know there is an enlargment of the probability space $(\Omega',\mathcal{F}',P')$, say $(\Omega'',\mathcal{F}'',P'')$, and $\tilde{a}$ and $\tilde{b}$, satisfying **(Linear growth (I))**, such that

$$\hat{X}_t = \hat{X}_0 + \int_0^t \sqrt{\tilde{a}}dw_r + \int_0^t \tilde{b}dr, \qquad \forall t\in[0,T] \qquad (P''-a.s.),$$

moreover, the diffusion term $\tilde{a}$ satisfies **(Coercive)**.

Secondly, we notice that for $\epsilon>0$, there is a constant $K_\epsilon>0$ such that

$$\sup_n P''(\sup_{t\le T}|\hat{X}_t| > K_\epsilon) < \epsilon.$$

We define a subset D of R^d

$$D := \{x\in R^d : |x| \le K_\epsilon\}.$$

and notice according to lemma 4.3, for every measurable subset Γ of $R^+\times R^d$, the set

$$D_\Gamma = \{t : t\in[0,\tau_D\wedge T], (t,\hat{X}_t)\in\Gamma\},$$

satisfies

$$P''(|D_\Gamma| > C) \le N|\Gamma|, \tag{4.46}$$

where $N = N(C,D,d,\gamma)$.

We now prove (4.45). We choose Γ as the set O in **(Mb)**, and consider the inequality

$$\int_0^T |b(s,\hat{X}_s^{n'}) - b(s,\hat{X}_s)|ds \le$$

$$\int_{D_O} |b(s,\hat{X}_s^{n'}) - b(s,\hat{X}_s)|ds + \int_{[0,T]\setminus D_O} |b(s,\hat{X}_s^{n'}) - b(s,\hat{X}_s)|ds, \tag{4.47}$$

where $D_O = \{t : t \in [0, \tau_D \wedge O], (t, \hat{X}_t) \in O\}$. The second term of (4.47) converges to zero since $b(s, \cdot)$ is uniformly continuous in D_O^c, the completment of the set D_O. The first term of (4.47) is less than $2B|D_O|$, for some constant B, a bound of function $|\tilde{b}$, since (4.41) implies

$$\tilde{b}(s, \hat{X}^{n'}| \leq K(1 + \sup_{t \leq T} |\hat{X}_t|) \leq K(1 + K_\epsilon).$$

According to (4.46), for $\epsilon' > 0$,

$$P'(B|D_O| > \epsilon') \leq N|O|,$$

with a constant N only depending on ϵ', not depending on $|O|$.

According to **(Mb)**, we can choose an open set O in $R^+ \times R^d$ as small as possible such that

$$P'(B|D_O| > \epsilon') \leq N|O| \leq \epsilon. \tag{4.48}$$

The equality (4.45) is thus proved. To prove (4.44), we can use the same argument as the one in proving (4.43). The theorem is thus proved.

Proof of theorem 2.2 According to theorem 2.1, we only need to check that the coefficients $a(t, x)$ and $b(t, x)$ uniquely determine the distribution of Q of the diffusion process (2.9). This is true according to theorem 7.2.1 and theorem 10.2.2 in [32]. The theorem is thus proved.

Acknowledgement

References

[1] A. V. Skorokhod, *Studies in the theory of random processes*, Izdat. Kiev. Univ. Kiev. 1961; English transl. Addison-Wesley, 1965.

[2] I. I. Gikhman and A. V. Skorokhod, *Introduction to the Theory of Random Processes*, "Nauka", Moscow 1965; English transl., Sauders, Philadelphia, Pa., 1969.

[3] A. A. Borovkov, *Theorems on the Convergence to Markov Diffusion Processes*, Z. Wahrsch. View. Gebie. 16 (1970), 47-76.

[4] H. J. Kushner, *On the Weak Convergence of Interpolated Markov Chain to a Diffusion*, Ann. Probab. 2 (1974), 40-50.

[5] S. Ya. Makhno, *On the Question of Weak Solutions of Stochastic Differential Equations*, Proc. School-Series Theory of Random

Processes (Druskininkai, 1974), Part 1, Inst. Fiz. i Mat. Akad. Nauk. Litovsk. SSR Vilnius, 1975, pp.237-255. (Russian).

[6] R. Markvenas, *The Weak Convergence of Random Process to the Solution of a Martingale Problem*, Litoves Math Sb. 15 (1975) no. 2, 67-75; English transl. in Lithuanian Math J. 15 (1975).

[7] Harry A. Gauss and John H. Gillespie, *Diffusion Approximations to Linear Stochastic Differential Equation with Stationary Coefficients*, J. Appl. Prob. 14 (1997), 58-74.

[8] M. Frank Norman, *Diffusion Approximation of non-Markovian Processes* Ann. Prob. 3 (1975), 358-364.

[9] Walter A. Rosenkrantz, *Limit Theorems for Solution to a Class of Stochastic Differential Equations*, Indian. Univ. Math. J. 24 (1974/75), 613-625.

[10] A. A. Borovkov, *Asymptotic Methods in Queueing Theory*, "Nauka", Moscow, 1980 (Russian).

[11] Inge. S. Helland, *Mimimal Condotions for Weak Convergence to a Diffusion Process on the Line*, Ann. Probab. 9 (1981), 429-452.

[12] H. J. Kushner and H. Huang, *On the Weak Convergence of a Sequence of General Stochastic Differential Equation to a Diffusion*, SIAM J. Appl. Math. 40 (1981), 528-541.

[13] Harry A. Guess, *An invariance Principle for Solution to Stochastic Differential Equations*, J. Appl. Probab. 18 (1981), 548-553.

[14] Ingl. S. Helland, *Convergence to Diffusions with Regular Boundaries*, Stochastic Processes Appl. 13 (1982), 27-58.

[15] B. Grigelionis and R. Mikulyavichyus, *On the Weak Convergence of Semimartingales*, Litovsk. Math. Sb. 21 (1981), no. 3, 9-24; English transl. in Lithuanian Math J. 21 (1981).

[16] J. L. Doob, *Stochastic Processes*, New York, Wiley, 1953.

[17] R. Sh. Liptser and A. N. Shiryayev, *Theory of Martingales*, (Kluwer Acacemic Publishers, 1989).

[18] R. Sh. Liptser, and A. N. Shiryaev, *Weak Convergence of a Sequence of Martingales to a Process of Diffusion Type*, Math. USSR Sbornik Vol. 49 (1984), No. 1.

[19] R. Sh. Liptser and A. N. Shiryayev, *On Weak Convergence of Semimartingales to Stochastically Continuous Processes with Independent and Conditionally Independent Increments* Math USSR. Sbornik, Vol 44 (1983) No.3.

[20] R. SH. Liptser and N. Shiryayev, *A Functional Central Limit Theorem for Semimartingales*, Theory Probab. Appl. 25 (1980) 667-688.

[21] Yi-Ju Chao, *Diffusion Approximations for a System of Infinite Servers with Load Balancing*, in preparation.

[22] A. Das and R. Srikant, *Diffusion Approximations for Models of Congestion Control in High-Speed Networks*, in preparation.

[23] P. J. Fleming and B. Simon, *Heavy Traffic Approximation for a System of Infinite Servers with Load Balancing*, Probability in the Engineering and Informational Sciences, preprint.

[24] Jean Jacod, *Calcul Stochastique et problemes de martingales*, Lecture Notes in Math., Vol 714, (Spring-Verlag, 1979).

[25] Yu. M. Kabanov, R. Sh. Liptser, and A. V. Shiryaev, *Absolute continuity and singularity of locally absolutely continuous probability measures I*, Math. Sb., 107 (149) (1978), English transl. in Math USSR Sb. 35 (1979), No. 5.

[26] Sheng-wu He, Jia-gang Wang and Jia-an Yan, *Semimartingale Theory and Stochastic Calculus*, (Science Press and CRC Press Inc., 1992).

[27] A. N. Ethier, and T. G. Kurtz, *Markov Process Characterization and Convergence*, (John Wiley and Sons, 1985).

[28] N. V. Krylov, *A Supermartingale Characterization of a Class of Stochastic Integrals*, Statistics and Control of Stochastic Processes, Vol.2. Optimalizatio Software, New York , (1989).

[29] N. V. Krylov, *Control Diffusion Process*, (Springer-Verlag, 1980).

[30] N. V. Krylov, *Introduction to the Theory of Diffusion Processes* (1994).

[31] M. Metivier, *Semimartingales: a Course on Stochastic Process*, (Walter de Gruyter, 1982).

[32] D. W. Strook and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, (Springer-Verlag, 1979).