Markov marginal problems and their applications to Markov optimal control

Toshio Mikami

Series #399. December 1997

HOKKAIDO UNIVERSITY
PREPRINT SERIES IN MATHEMATICS

#374 J. Inoue and S.-E. Takahashi, On characterizations of the image of Gelfand transform of commutative Banach algebras, 30 pages. 1997.

#375 L. Solomon and H. Terao, The double coxeter arrangement, 21 pages. 1997.

#376 G. Ishikawa and T. Morimoto, Solution surfaces of Monge-Ampère equations, 15 pages. 1997.

#377 G. Ishikawa, A relative transversality theorem and its applications, 16 pages. 1997.

#378 J. Inoue and T. Nakazi, On the zeroes of solutions of an extremal problem in H^1, 14 pages. 1997.

#379 Y. Giga and K. Ito, On pinching of curves moved by surface diffusion, 12 pages. 1997.

#380 F. Hiroshima, Weak coupling limit removing an ultraviolet cut-off for a Hamiltonian of particles interacting with a scalar field, 39 pages. 1997.

#381 Y. Giga, S. Matsui and Y. Shimizu, On estimates in Hardy spaces for the Stokes flow in a half space, 13 pages. 1997.

#382 A. Arai, A new estimate for the ground state energy of Schrödinger operators, 12 pages. 1997.

#383 M. Nakamura and T. Ozawa, The Cauchy problem for nonlinear wave equations in the Sobolev space of critical order, 24 pages. 1997.

#384 K. Ito, Asymptotic stability of planar rarefaction wave for scalar viscous conservation law, 8 pages. 1997.

#385 A. Arai, Representation-theoretic aspects of two-dimensional quantum systems in singular vector potentials: canonical commutation relations, quantum algebras, and reduction to lattice quantum systems, 32 pages. 1997.

#386 P. Aviles and Y. Giga, On lower semicontinuity of a defect energy obtained by a singular limit of the Ginzburg-Landau type energy for gradient fields, 21 pages. 1997.

#387 T. Nakazi and T. Yamamoto, Norms of some singular integral operators and their inverse operators, 28 pages. 1997.

#388 M.-H. Giga and Y. Giga, Remarks on convergence of evolving graphs by nonlocal curvature, 18 pages. 1997.

#389 T. Tsukada, Reticular Lagrangian singularities, 41 pages. 1997.

#390 M. Nakamura and T. Ozawa, The Cauchy problem for nonlinear wave equations in the homogeneous Sobolev space, 12 pages. 1997.

#391 Y. Giga, M. Ohnuma and M.-H. Sato, On strong maximum principle and large time behaviour of generalized mean curvature flow with the Neumann boundary condition, 24 pages. 1997.

#392 T. Tsujishita and H. Watanabe, Monoidal closedness of the category of simulations, 24 pages. 1997.

#393 T. Arase, A remark on the quantale structure of multisets, 10 pages. 1997.

#394 N. H. Bingham and A. Inoue, Extension of the Drasin-Shea-Jordan theorem, 16 pages. 1997.

#395 N. H. Bingham and A. Inoue, Ratio Mercerian theorems with applications to Hankel and Fourier transforms, 30 pages. 1997.

#396 Y. Nishiura and D. Ueyama, A skeleton structure of self-replicating dynamics, 27 pages. 1997.

#397 K. Hirata and K. Sugano, On semisimple extensions of serial rings, 6 pages. 1997.

#398 D. Pei, Singularities of $\mathbb{R}P^2$-valued Gauss maps of surfaces in Minkowski 3-space, 15 pages. 1997.

Markov marginal problems and their applications to Markov optimal control

Toshio Mikami
Department of Mathematics
Hokkaido University
Sapporo 060
JAPAN
email;mikami@math.sci.hokudai.ac.jp

ABSTRACT.- In this paper we discuss a class of Markov marginal problems (MMP). By MMP, we mean the problem to construct a Markov process with given (marginal) constraints on the path space. As an application we consider Markov optimal control problems.

Key words; Markov marginal problem, Markov optimal control, copula, covariance kernel

1. Introduction.

Let $\mathbf{M}$ and $\mathbf{X}$ be a topological space, and $\{\mu_t\}_{t\in\mathbf{M}}$ be a family of Borel probability measures on $(\mathbf{X}, \mathbf{B}(\mathbf{X}))$. As a classical marginal problem, the following is known.
(MP). Find a probability measure $\mathbf{Q}$ on $(\mathbf{X}^{\mathbf{M}}, \mathbf{B}(\mathbf{X}^{\mathbf{M}}))$ such that

$$\mathbf{Q}(f \in \mathbf{X}^{\mathbf{M}} | f(t) \in dx) = \mu_t(dx) \quad \text{for all } t \in \mathbf{M}. \tag{1.1}$$

Let $\mathbf{P}(\mathbf{X})$ denote the space of Borel probability measures on $(\mathbf{X}, \mathbf{B}(\mathbf{X}))$, equipped with the weak topology, and let $C(\mathbf{M}, \mathbf{X})$ denote the space of continuous functions from $\mathbf{M}$ to $\mathbf{X}$, equipped with the topology of uniform convergence. With respect to MP, the following is known (see [2, Theorem 2.1] and also [1]).

Theorem 1.1. *Let $\mathbf{M}$ and $\mathbf{X}$ be compact metric spaces and let $\mathbf{X}$ be connected and locally connected. Suppose that $\mu : \mathbf{M} \mapsto \mathbf{P}(\mathbf{X})$ is continuous and that $supp(\mu_t) = \mathbf{X}$ for all $t \in \mathbf{M}$. Then there exists $\mathbf{Q} \in P(C(\mathbf{M}, \mathbf{X}))$ such that*

$$\mathbf{Q}(f \in C(\mathbf{M}, \mathbf{X}) | f(t) \in dx) = \mu_t(dx) \quad \text{for all } t \in \mathbf{M}. \tag{1.2}$$

Instead of giving the topological structure to **M** and **X**, giving the order structure to them, similar problem has been considered in [10, 14, 15] (see also the references therein).

In this paper we consider the case **M** is an interval on R, such as $[0,1]$ and $[0,\infty)$, and we would like to construct a Markov process with given constraints.

As a typical problem in this direction, one can mention Nelson's stochastic quantization. Let us briefly introduce the problem. Let $\psi(t,x)$ $(t \geq 0,\ x \in R^d)$ be the solution of the following Schrödinger equation:

$$\begin{aligned} &(-1)^{1/2}\partial\psi(t,x)/\partial t = -\Delta_x\psi(t,x)/2 + V(x)\psi(t,x) \quad (t>0, x\in R^d), \\ &\int_{R^d} |\psi(0,x)|^2 dx = 1, \end{aligned} \qquad (1.3).$$

for some function $V(\cdot): R^d \mapsto R$. Here we put $\Delta_x \equiv \sum_{i=1}^d \partial^2/\partial x_i^2$.

One of basic problems in Nelson's stochastic quantization (see [22], [23]) is to construct a Markov process $\{X(t)\}_{t\geq 0}$ on a probability space $(\Omega, \mathbf{B}, P)$ such that

$$P(X(t)\in dx) = |\psi(t,x)|^2 dx \quad \text{for all } t\geq 0. \qquad (1.4).$$

From (1.3), for any infinitely differentiable function $f: R^d \mapsto R$ with a compact support, the following holds: for $t > 0$,

$$d[\int_{R^d} f(x)|\psi(t,x)|^2 dx]/dt = \int_{R^d} [\Delta_x f(x)/2 + < b(t,x;\psi), \nabla_x f(x) >] |\psi(t,x)|^2 dx, \quad (1.5).$$

where we put

$$b(t,x;\psi) \equiv \begin{cases} Re(\nabla_x\psi(t,x)/\psi(t,x)) + Im(\nabla_x\psi(t,x)/\psi(t,x)) & \text{if } \psi(t,x)\neq 0, \\ 0 & \text{if } \psi(t,x) = 0. \end{cases} \qquad (1.6).$$

Here we put $\nabla_x \equiv (\partial/\partial x_i)_{i=1}^d$, and $< \cdot,\cdot >$ denotes the inner product in R^d. As an answer to the above problem, the following is known (see [6] and also [7, 21, 30]).

Theorem 1.2. *Suppose that the following condition holds: for all* $t > 0$,

$$\int_0^t ds \int_{R^d} |b(s,x;\psi)|^2 |\psi(s,x)|^2 dx < \infty.$$

Then there exists a Markov process $\{X(t)\}_{t\geq 0}$ such that (1.4) holds and that

$$dX(t) = b(t, X(t); \psi)dt + dW(t) \quad (t > 0), \tag{1.7}$$

where $\{W(t)\}_{t\geq 0}$ is a d-dimensional Wiener process (see [16, 28]).

When Δ_x is replaced, in (1.5), by $\sum_{i,j=1}^{d} a(t,x)^{ij}\partial^2/\partial x_i \partial x_j$ for a symmetric, uniformly positive definite $d \times d$-matrix $(a(t,x)^{ij})_{i,j=1}^{d}$, the problem was considered in [18, 24].

As we discussed in [17], this problem is related to Markov optimal control problem. Let us briefly introduce it.

Let $(\Omega, \mathbf{B}, P)$ be a probability space and $\{\mathbf{B}_t\}_{t\geq 0}$ be a right continuous, increasing family of sub σ- fields of $\mathbf{B}$, and let $\{W(t)\}_{t\geq 0}$ denote a d-dimensional $(\mathbf{B}_t)$-Wiener process on $(\Omega, \mathbf{B}, P)$, and $\{\sigma(t,x)\}_{t\geq 0, x\in R^d}$ be a Borel measurable $d\times d$-matrix, and $\{u(t)\}_{t\geq 0}$ be a $(\mathbf{B}_t)$-progressively measurable vector in R^d. Consider a semimartingale $\{X^u(t)\}_{t\geq 0}$ which satisfies the follows: for $t > 0$,

$$dX^u(t) = u(t)dt + \sigma(t, X^u(t))dW(t). \tag{1.8}$$

For Borel measurable functions $L(t,x;u) : [0,\infty) \times R^d \times R^d \mapsto R$ and $\Phi(x) : R^d \mapsto R$, study the following problem (Markov optimal control problem).

(MOCP). For $T > 0$, consider if the following is true:

$$\begin{aligned}
&\inf\{\int_0^T E[L(t, X^u(t); u(t))]dt + \Phi(X^u(T)); \{u(t)\}_{t\in[0,T]}\} \qquad (1.9)\\
&= \inf\{\int_0^T E[L(t, X^u(t); u(t))]dt + \Phi(X^u(T)); u(t) = b(t, X^u(t)) \\
&\qquad \text{for some } b(t,x)((t,x) \in [0,T] \times R^d)\}.
\end{aligned}$$

About Markov optimal control, we refer the reader to [11]. With respect to this problem, the following is known from Theorem 1.2, by Jensen's inequality (see [17]).

Theorem 1.3. *Suppose that σ is an identity matrix or $d = 1$ and that $L(t,x;\cdot)$ is convex for all $(t,x) \in [0,T] \times R^d$. Then the following holds:*

$$\inf\{\int_0^T E[L(t,X^u(t);u(t))]dt+\Phi(X^u(T)); \int_0^T E[|u(t)|^2]dt<\infty\} \qquad (1.10).$$
$$=\inf\{\int_0^T E[L(t,X^u(t);b(t,X^u(t)))]dt+\Phi(X^u(T)); \int_0^T E[|u(t)|^2]dt<\infty,$$
$$u(t)=b(t,X^u(t)) \text{ for some } b(t,x)((t,x)\in[0,T]\times R^d)\},$$

provided that l.h.s. is finite.

Remark 1.1. As a similar problem to MOCP, one can mention the Monge problem that is still open (see [8, 13, 25, 29] and references therein). In a simple case, it can be stated as follows. Let d_1, $d_2 \in \mathbf{N}$. For a measurable function $c : R^{d_1}\times R^{d_2} \mapsto R$ and Borel probability measures Q_1 and Q_2 on R^{d_1} and R^{d_2}, respectively, study if the following holds:

$$\inf\{\int_{R^{d_1}\times R^{d_2}} c(x,y)\mu(dxdy); \mu(dxR^{d_2})=Q_1(dx), \mu(R^{d_1}dy)=Q_2(dy)\} \qquad (1.11).$$
$$=\inf\{\int_{R^{d_1}} c(x,\phi(x))Q_1(dx); \phi : R^{d_1}\mapsto R^{d_2} \text{ satisfies } Q_2(dy)=(Q_1)^{\phi^{-1}}(dy)\}.$$

Minimizing ϕ in (1.11) is called Monge function.

In this paper we give a new approach to MOCP, by studying the marginal problem that we call Markov marginal problem (MMP). Roughly speaking, our approach is as follows: for $(X^u(t),u(t))_{t\in[0,T]}$ in (1.8), find $b(t,x)((t,x)\in[0,T]\times R^d)$ for which the following has a weak solution

$$X(t)=X^u(0)+\int_0^t b(s,X(s))ds+\int_0^t \sigma(s,X(s))dW(s), \quad (t\in[0,T]), \qquad (1.12).$$

and

$$P((X^u(t),u(t))\in dx)=P((X(t),b(t,X(t)))\in dx) \quad \text{for a.e. } t\in[0,T]. \qquad (1.13).$$

If this can be done for the minimizing X^u, then (1.9) is true. This MMP is a different kinds of marginal problem from Theorem 1.1 in that we construct "Markov process". The problem in Nelson's stochastic quantization can be considered as MMP. In section 2, we give a partial answer to the above problem and give the application to MOCP in section 3. The reader will notice that our approach was inspired by the idea of covariance kernels on central limit theorems (see Remark 2.2 and [3-5, 20]) and that we used the relation between 2-copulas and t-norm (see Remark 2.3 and [26, 27]).

2. Marginal problems.

In this section we state and prove theorems on Markov marginal problems (MMP) introduced in section 1. From now on, we fix $T > 0$.

The following result is known (see [19, Theorem 2.1]) and will be used in the proof of Theorem 2.2.

Theorem 2.1. *For any family of distribution functions $\{F_t(\cdot)\}_{t\in[0,T]}$ on R^d for which $F_t(\cdot)$ is continuous for $0 \le t \le T$, there exists a R^d-valued Markov process $\{X(t)\}_{0\le t\le T}$ on a probability space $(\Omega, \mathbf{B}, P)$ such that*

$$P(X(t) \in dx) = dF_t(x) \quad \text{for all } 0 \le t \le T. \tag{2.1}$$

Put $P_0(\mathbf{X}) \equiv \{\mu \in P(\mathbf{X}); supp(\mu) = \mathbf{X}\}$. The next theorem can be proved from Theorem 2.1.

Theorem 2.2. *Suppose that $\mathbf{X}$ is a connected and locally connected compact metric space, and that $\mu_t \in P_0(\mathbf{X})$ for all $t \in [0, T]$. Then there exists a measurable $\overline{\varphi}_t : R \mapsto \mathbf{X}$ $(t \in [0, T])$ such that*

$$P(\overline{\varphi}_t(W(t+1)) \in dx) = \mu_t(dx) \quad \text{for all } t \in [0, T], \tag{2.2}$$

where $\{W(t)\}_{t\ge0}$ is a one-dimensional Wiener process on a probability space $(\Omega, \mathbf{B}, P)$.

The next theorem played a crucial in [2, Theorem 2.2] and will be used in the proof of Theorem 2.2.

Theorem 2.3. *Suppose that* $\mathbf{X}$ *is a connected and locally connected compact metric space. Then there exists a continuous function* $\tilde{\varphi} : P_0(\mathbf{X}) \mapsto P_0([0,1])$ *and a continuous function* $\varphi : [0,1] \mapsto \mathbf{X}$ *such that*

$$(\tilde{\varphi}(\mu))^{\varphi^{-1}} = \mu \quad \text{for all } \mu \in P_0(\mathbf{X}). \tag{2.3}$$

Let us prove Theorem 2.2.

(Proof of Theorem 2.2). Put for $x \in R$ and $t \in [0,T]$,

$$\begin{aligned} F_{1,t}(x) &= P(W(t+1) \leq x), \\ F_{2,t}(x) &= \tilde{\varphi}(\mu_t)((-\infty, x]), \end{aligned} \tag{2.4}$$

where $\tilde{\varphi}$ is a function stated in Theorem 2.3. For a distribution function F on R, put

$$F^{-1}(u) \equiv \begin{cases} \sup\{x \in R; F(x) < u\} & \text{for } u \in (0,1], \\ \sup\{x \in R; F(x) = 0\} & \text{for } u = 0. \end{cases} \tag{2.5}$$

If the set where the supremum is taken is empty, then we put the sup $= -\infty$.

In the same way as in [19],

$$P(F_{2,t}^{-1}(F_{1,t}(W(t+1))) \in dx) = \tilde{\varphi}(\mu_t)(dx) \quad \text{for all } 0 \leq t \leq T. \tag{2.6}$$

In fact, for $x \in R$ and $t \in [0,T]$,

$$P(F_{2,t}^{-1}(F_{1,t}(W(t+1))) \leq x) = P(F_{1,t}(W(t+1)) \leq F_{2,t}(x)) = F_{2,t}(x), \tag{2.7}$$

since $F_{1,t}(W(t+1))$ is uniformly distributed on $[0,1]$, and since for a distribution function F on R and $u \in (0,1]$,

$$F^{-1}(u) \leq x \quad \text{iff} \quad u \leq F(x). \tag{2.8}$$

From Theorem 2.3 and (2.6), putting $\overline{\varphi}_t(x) = \varphi(F_{2,t}^{-1}(F_{1,t}(x)))$ $(x \in R,\ t \in [0,T])$, one get

$$P(\overline{\varphi}_t(W(t+1)) \in dx) = \tilde{\varphi}(\mu_t)^{\varphi^{-1}}(dx) = \mu_t(dx) \quad \text{for all } 0 \leq t \leq T, \tag{2.9}$$

which completes the proof.

Q.E.D.

To give theorems on MMP discussed below Remark 1.1, let us give the assumption which will be used later.

(A.1). d=1. $\sigma(t,x) : [0,T] \times R \mapsto R$ is bounded and is continuous in x, uniformly in t. There exists $\nu > 0$ such that

$$\sigma(t,x) \geq \nu^{1/2}$$

for all $(t,x) \in [0,T] \times R$. X_o is a $(\mathbf{B}_0)$-adapted random variable on $(\Omega, \mathbf{B}, P)$ such that $P(|X_o| < \infty) = 1$.

(A.2). (1.8) has a solution such that $X^u(0) = X_o$.

(A.3).

$$E[\int_0^T |u(t)|dt] < \infty, \quad P(\int_0^T |u(t)|^2 dt < \infty) = 1.$$

(A.4).

$$\limsup_{r \to \infty} P(|X_o| \geq r) r^4 < \infty.$$

Remark 2.1. Under (A.1), (1.8) has a unique weak solution X^0 which is strong Markov if $u \equiv 0$ (see [28]). Under (A.1)-(A.3), there exist a jointly measurable $u'(t,x) : [0,T] \times C([0,T];R) \mapsto R$ for which $u'(t,\cdot)$ is $\mathbf{B}(C([0,t];R))_+$-measurable for $t \in [0,T]$, and $\sigma[X^u(s); 0 \leq s \leq t]$-adapted Wiener process $\overline{W}(t)$ such that for $t \in [0,T]$,

$$X^u(t) = X_o + \int_0^t u'(s, X^u) ds + \int_0^t \sigma(s, X^u(s)) d\overline{W}(s). \qquad (2.10).$$

Moreover $P^{(X^u)^{-1}}$ is absolutely continuous with respect to $P^{(X^0)^{-1}}$ (see [16, section 7.6]), and $P(X^0(t) \in dx)$ is absolutely continuous with respect to dx for $t \in (0,T]$ (see [28]).

Let us state a technical result which can be obtained exactly in the same way as in [12, p. 371, Theorem 3.1].

Theorem 2.4. *Suppose that (A.1) holds and that $B(t,v) : [0,T] \times R \mapsto R$ is bounded, measurable, and globally Lipschitz continuous in v, uniformly in $t \in [0,T]$. Then the following integral equation has a unique solution; for $(t,x) \in [0,T] \times R$,*

$$v(t,x) = E[\exp(\int_0^t \sigma(s,X^0(s))^{-1}B(s,v(s,X^0(s)))dW(s) \qquad (2.11).$$
$$- \int_0^t \sigma(s,X^0(s))^{-2}B(s,v(s,X^0(s)))^2 ds/2); X^0(t) \le x],$$

where X^0 is a weak solution to (1.8) with $u(t) \equiv 0$ and with $X^0(0) = X_o$.

Let X^B be the solution of the following: for $t \in [0,T]$,

$$X^B(t) = X_o + \int_0^t B(s,v(s,X^B(s)))ds + \int_0^t \sigma(s,X^B(s))dW(s), \qquad (2.12).$$

which has a unique weak solution under the assumption in Theorem 2.4 (see [28]). Then the following result can be obtained and be considered as MMP.

Theorem 2.5. *Under the same assumption as in Theorem 2.4, for $v(t,x)$ constructed in Theorem 2.4, (2.12) has a unique weak solution, and the following holds:*

$$P(X^B(t) \le x) = v(t,x) \qquad (2.13).$$

for $(t,x) \in [0,T] \times R$. In particular, for any $\varphi \in C_o^\infty(R;R)$,

$$d(\int_R \varphi(x)v(t,dx))/dt \qquad (2.14).$$
$$= \int_R (2^{-1}\sigma(t,x)^2\partial^2\varphi(x)/\partial x^2 + B(t,v(t,x))\partial\varphi(x)/\partial x)v(t,dx), \quad dt-a.e. \text{ on } (0,T).$$

(Proof of Theorem 2.5). The first part of the theorem can be easily proved by Cameron-Martin-Maruyama-Girsanov formula (see [16, 28]). The second part is an easy consequence of the Ito formula.

Q. E. D.

In Theorem 2.5, we fixed a diffusion coefficient. If we allow that it can be changed, then we get the following result.

Theorem 2.6. *Suppose that (A.1)-(A.2) and (A.4) hold, and that $\{|u(t)|\}_{0\le t\le T}$ is dominated by a positive constant $C_u > 0$ a.s.. Then there exist measurable $a^n(\cdot,\cdot) : [0,T]\times R \mapsto [\nu,\infty)$ and $b^n(\cdot,\cdot) : [0,T]\times R \mapsto R$ such that the following has a weak solution: for $t \in [0,T]$*

$$dX^n(t) = b^n(t, X^n(t))dt + a^n(t, X^n(t))^{1/2}dW(t) \qquad (2.15).$$

$(n \ge 1)$, and the following holds:

$$\begin{aligned} \lim_{n\to\infty} P(X^n(t) \in dx) &= P(X^u(t) \in dx) \quad (t \in [0,T]), \\ \lim_{n\to\infty} P(b^n(t, X^n(t)) \in dx) &= P(u(t) \in dx) \quad (t \in (0,T]), \end{aligned} \qquad (2.16).$$

weakly.

Let us prove Theorem 2.6 with the aid of lemmas given later.

(Proof of Theorem 2.6). For $n \ge 1$, $t \in [0,T]$ and $x \in R$, put

$$\begin{gathered} F_t(x) = P(X^u(t) \le x), \quad c^n(x) = Cn(1 + (nx)^2)^{-3/2}, \quad C = (\int_R (1+x^2)^{-3/2}dx)^{-1}, \\ F_t^n(x) = \int_R (\int_{-\infty}^{x-y} c^n(z)dz)F_t(dy), \quad f_t^n(x) = \int_R c^n(x-y)F_t(dy). \end{gathered} \qquad (2.17).$$

Then the following holds (see Lemma 2.7): for $n \ge 1$ and $x \in R$,

$$\partial f_t^n(x)/\partial t = 2^{-1}\partial^2(a^n(t,x)f_t^n(x))/\partial x^2 - \partial(b^n(t,x)f_t^n(x))/\partial x \quad dt\text{-a.e. on } (0,T), \quad (2.18).$$

where we put for $n \ge 1$, $x \in R$ and $t \in [0,T]$,

$$\begin{gathered} G_t(x) = P(u(t) \le x), \quad b(t,x) = G_t^{-1}(1 - F_t(x)) \quad \text{(see (2.5) for notation)}, \\ b^n(t,x) = \int_R b(t,y)c^n(x-y)F_t(dy)/f_t^n(x), \\ \tilde{a}^n(t,x) = \int_R \sigma(t,y)^2 c^n(x-y)F_t(dy)/f_t^n(x), \\ \tilde{b}(t,x) = E[u(t)|X^u(t) = x], \quad \tilde{b}^n(t,x) = \int_R \tilde{b}(t,y)c^n(x-y)F_t(dy)/f_t^n(x), \\ a^n(t,x) = \tilde{a}^n(t,x) + 1_{(0,T]}(t)2f_t^n(x)^{-1}\int_{-\infty}^x [b^n(t,y) - \tilde{b}^n(t,y)]f_t^n(y)dy. \end{gathered} \qquad (2.19).$$

From the assumption, one can show that there exists a unique strong solution to (2.15) with $P(X^n(0) \in dx) = F_0^n(dx)$, and that $P(X^n(t) \in dx) = F_t^n(dx)$ $(0 \le t \le T)$, and that (2.16) holds for a^n and b^n defined in (2.19). To prove the uniqueness and the existence part, one only has to show that the following holds (see [28]):

$$a^n(t,x) \ge \nu \quad (0 \le t \le T, x \in R), \qquad (2.20).$$

$$\sup_{0 \le t \le T, x \in R} (|\partial a^n(t,x)/\partial x| + |\partial b^n(t,x)/\partial x| + |a^n(t,0)| + |b^n(t,0)|) < \infty, \qquad (2.21).$$

since

$$\int_R |y| f_0^n(y) dy < \infty.$$

(2.20)-(2.21) and (2.16) will be proved in Lemmas 2.8-2.9 and in Lemmas 2.10-2.11 which will be given later.

Q. E. D.

Remark 2.2. In (2.18), we determined $a^n(t,x)$ from a given function $b^n(t,x)$ to derive PDE for $f_t^n(x)$. This idea is that of covariance kernel (see [3-5, 20]). Let us introduce it. Let f be a probability density function on R such that $\int_R y f(y) dy = \mu$ and $\int_R (y-\mu)^2 f(y) dy = \sigma^2 < \infty$. The following function is called a covariance kernel or ω-function of f:

$$\omega(x) \equiv \int_{-\infty}^{x} (\mu - y) f(y) dy / [\sigma^2 f(x)] \qquad (2.22).$$

on the set $\{y \in R : f(y) > 0\}$. One can see that for any $\phi \in C_0^\infty(R:R)$,

$$\int_R [\omega(x) \partial^2 \phi(x)/\partial x^2 + \sigma^{-2}(\mu - x) \partial \phi(x)/\partial x] f(x) dx = 0. \qquad (2.23).$$

In the rest part of this section, we prove (2.16), (2.18) and (2.20)-(2.21) as Lemmas 2.7-2.11.

Let us first prove (2.18).

Lemma 2.7. *Under the assumption in Theorem 2.6, for $n \geq 1$ and $x \in R$,*

$$\partial f_t^n(x)/\partial t = 2^{-1}\partial^2(a^n(t,x)f_t^n(x))/\partial x^2 - \partial(b^n(t,x)f_t^n(x))/\partial x \quad dt-a.e. \text{ on } (0,T). \quad (2.24).$$

(Proof). We only have to prove that the following holds: for $n \geq 1$ and $x \in R$,

$$\partial f_t^n(x)/\partial t = 2^{-1}\partial^2(\tilde{a}^n(t,x)f_t^n(x))/\partial x^2 - \partial(\tilde{b}^n(t,x)f_t^n(x))/\partial x \quad dt - a.e. \text{ on } (0,T). \quad (2.25).$$

This is true, since $f_t^n(x) = E[c^n(x - X^u(t))]$ and since, by the Ito formula

$$\begin{aligned}
&\partial E[c^n(x - X^u(t))]/\partial t \\
&= E[2^{-1}\sigma(t, X^u(t))^2\partial^2 c^n(x - X^u(t))/\partial x^2 - u(t)\partial c^n(x - X^u(t))/\partial x] \\
&= E[2^{-1}\sigma(t, X^u(t))^2\partial^2 c^n(x - X^u(t))/\partial x^2 - \tilde{b}(t, X^u(t))\partial c^n(x - X^u(t))/\partial x] \\
&= \int_R [2^{-1}\sigma(t,y)^2\partial^2 c^n(x-y)/\partial x^2 - \tilde{b}(t,y)\partial c^n(x-y)/\partial x]F_t(dy) \\
&= 2^{-1}\partial^2[\int_R \sigma(t,y)^2 c^n(x-y)F_t(dy)]/\partial x^2 - \partial[\int_R \tilde{b}(t,y)c^n(x-y)F_t(dy)]/\partial x.
\end{aligned}$$

Q. E. D.

We proceed to the proof of (2.20).

Lemma 2.8. *Under the assumption in Theorem 2.6, for $x \in R$ and $t \in (0,T]$,*

$$\int_{-\infty}^{x} [b^n(t,y) - \tilde{b}^n(t,y)]f_t^n(y)dy \geq 0. \quad (2.26).$$

(Proof). For $x \in R$ and $t \in (0,T]$,

$$\begin{aligned}
\int_{-\infty}^{x} [b^n(t,y) - \tilde{b}^n(t,y)]f_t^n(y)dy &= \int_{-\infty}^{x} dy \int_R [b(t,z) - \tilde{b}(t,z)]c^n(y-z)F_t(dz) \quad (2.27). \\
&= \int_R [b(t,z) - \tilde{b}(t,z)][\int_{-\infty}^{x-z} c^n(y)dy]F_t(dz) \\
&= \int_R c^n(y)dy \int_{-\infty}^{x-y} [b(t,z) - \tilde{b}(t,z)]F_t(dz),
\end{aligned}$$

and

$$\int_{-\infty}^{x} b(t,z)F_t(dz) = E[G_t^{-1}(1-F_t(X^u(t))); X^u(t) \le x],$$
$$\int_{-\infty}^{x} \tilde{b}(t,z)F_t(dz) = E[E[u(t)|X^u(t)]; X^u(t) \le x] = E[u(t); X^u(t) \le x]. \qquad (2.28).$$

For any distribution function Φ on R^2 such that $\int_{R^2} |y|\Phi(dxdy) < \infty$,

$$\int_{(-\infty,x]\times R} y\Phi(dxdy) = \int_0^{\infty} [\Phi(x,\infty) - \Phi(x,y)]dy - \int_{-\infty}^{0} \Phi(x,y)dy. \qquad (2.29).$$

This and (2.27)-(2.28) together with the following completes the proof: for any distribution function Φ on R^2,

$$\Phi(x,y) \ge max(\Phi(x,\infty) + \Phi(\infty,y) - 1, 0) \quad (x, y \in R) \qquad (2.30).$$

(see [26, 27]). In fact, from (2.8),

$$\begin{aligned} &P(X^u(t) \le x, b(t, X^u(t)) \le y) \qquad (2.31).\\ &= P(X^u(t) \le x, G_t^{-1}(1-F_t(X^u(t))) \le y)\\ &= P(1-G_t(y) \le F_t(X^u(t)) \le F_t(x))\\ &= max(F_t(x) + G_t(y) - 1, 0) \le P(X^u(t) \le x, u(t) \le y) \quad \text{(from (2.30))}. \end{aligned}$$

Here we used the fact that $F_t(X^u(t))$ is uniformly distributed on $[0,1]$ for $t \in (0,T]$.

Q. E. D.

Remark 2.3. (2.31) implies that our approach is effective only in case $d = 1$. In fact, for any distribution function $\Phi(x_1, \cdots, , x_d)$ on R^d, there exists a distribution function C_Φ on $[0,1]^d$ which is uniquely determined on $Range(F_1) \times \cdots Range(F_d)$ (see [26, 27]) such that the following holds: for $(x_i)_{i=1}^d \in R^d$,

$$\min(F_1(x_1), \cdots, F_d(x_d)) \geq \Phi(x_1, \cdots, , x_d) = C_\Phi(F_1(x_1), \cdots, , F_d(x_d)) \qquad (2.32).$$
$$\geq max(F_1(x_1) + \cdots + F_d(x_d) + d - 1, 0).$$

Here we denote by $F_i(x)$ the value of Φ when $x_k = \infty$ $(k \neq i)$ and $x_i = x$. t-norm $W(u_1, \cdots, u_d) = max(u_1 + \cdots + u_d + d - 1, 0)$ is a distribution function on $[0,1]^d$ iff $d = 2$.

Let us prove (2.21) which will also be used in the proof of Lemma 2.10.

Lemma 2.9. *Under the assumption in Theorem 2.6, for any $n \geq 1$, there exists $C_n > 0$ such that*

$$\sup_{0 \leq t \leq T, x \in R} (|\partial a^n(t,x)/\partial x| + |\partial b^n(t,x)/\partial x| + |a^n(t,0)| + |b^n(t,0)|) \leq C_n. \qquad (2.33).$$

(Proof). We only have to show that the following holds:

$$|\partial c^n(x)/\partial x| \leq (3/2)nc^n(x), \qquad (2.34).$$
$$\sup_{0 \leq t \leq T, x \in R} f_t^n(x)^{-1}(1 + |x|)^{-3} < \infty, \qquad (2.35).$$
$$\sup_{0 \leq t \leq T, x \in R} |\partial f_t^n(x)/\partial x|(1 + |x|)^4 < \infty, \qquad (2.36).$$
$$\sup_{0 \leq t \leq T} \{\sup_{x<0} F_t^n(x)(1 + |x|)^2 + \sup_{x>0}(1 - F_t^n(x))(1 + |x|)^2\} < \infty. \qquad (2.37).$$

This is true, since

$$\begin{aligned}\partial a^n(t,x)/\partial x = &\int_R \sigma(t,y)^2(\partial c^n(x-y)/\partial x)F_t(dy)/f_t^n(x) - a^n(t,x)[\partial f_t^n(x)/\partial x]/f_t^n(x)\\ &+ 1_{(0,T]}(t)2[b^n(t,x) - \tilde{b}^n(t,x)],\\ \partial b^n(t,x)/\partial x = &\int_R b(t,y)(\partial c^n(x-y)/\partial x)F_t(dy)/f_t^n(x) - b^n(t,x)[\partial f_t^n(x)/\partial x]/f_t^n(x),\end{aligned}$$

and since

$$|b^n(t,x)| + |\tilde{b}^n(t,x)| \le \sup_{0\le s\le T, y\in R} \{|b(s,y)| + |\tilde{b}(s,y)|\} \le 2C_u,$$

$$|a^n(t,x)| \le \sup_{0\le s\le T, y\in R} |\sigma(s,y)|^2 + 1_{(0,T]}(t) 4C_u \min(F_t^n(x), 1 - F_t^n(x))/f_t^n(x).$$

Here we used the fact that $u(t)$ and $G_t^{-1}(1 - F_t(X^u(t)))$ has the same probability law for $t \in (0,T]$ and that

$$\int_R (b^n(t,x) - \tilde{b}^n(t,x)) f_t^n(x)dx = E[u(t) - G_t^{-1}(1 - F_t(X^u(t)))] = 0$$

(see (2.27)-(2.28) and (2.31)).

(2.34) can be easily shown as follows:

$$\partial c^n(x)/\partial x = c^n(x)(3/2)n(-2nx)/[1 + n^2x^2]. \tag{2.38}$$

(2.35) can be shown as follows: for $r > 0$,

$$\begin{aligned} f_t^n(x) &= \int_R nC(1 + n^2(x-y)^2)^{-3/2} F_t(dy) \qquad (2.39). \\ &\ge \int_{|y|\le r} nC(1 + 2n^2(x^2+y^2))^{-3/2} F_t(dy) \ge nCP(|X^u(t)| \le r)(1 + 2n^2(x^2+r^2))^{-3/2}, \end{aligned}$$

and

$$\begin{aligned} &P(|X^u(t)| \ge r) \qquad (2.40). \\ &\le P(|X^u(0)| \ge r/3) + P(|\int_0^t u(s)ds| \ge r/3) + P(|\int_0^t \sigma(s, X^u(s))dW(s)| \ge r/3) \\ &\le P(|X^u(0)| \ge r/3) + (3tC_u/r)^4 + 3t^2 \sup_{0\le s\le T, y\in R} |\sigma(s,y)|^4 (3/r)^4 \end{aligned}$$

by Chebychev's inequality.

Next we prove (2.36). For $x \in R$,

$$|\partial f_t^n(x)/\partial x| \le \int_R (3/2)n^2 C 2n|x-y|/(1+n^2(x-y)^2)^{5/2} F_t(dy) \le (3/2)n^2 C. \qquad (2.41).$$

For x for which $|x| > 2$,

$$\begin{aligned} &|\partial f_t^n(x)/\partial x| \qquad (2.42). \\ &\le \Big(\int_{|x-y|<|x|/2} + \int_{|x-y|\ge|x|/2}\Big)(3/2)n^2 C 2n|x-y|/(1+n^2(x-y)^2)^{5/2} F_t(dy) \\ &\le (3/2)n^2 C P(|X^u(t)| \ge |x|/2) + (3/2)n^2 C|x|/(1+(|x|/2)^2)^{5/2}, \end{aligned}$$

since $2r/(1+r^2)^{5/2}$ is decreasing on $[1,\infty)$. This together with (2.40) completes the proof from (A.4).

(2.37) can be proved from (2.36). In fact

$$F_t^n(x) = \int_{-\infty}^x dy \int_{-\infty}^y (\partial f_t^n(z)/\partial z) dz = 1 + \int_x^\infty dy \int_y^\infty (\partial f_t^n(z)/\partial z) dz, \qquad (2.43).$$

since

$$\lim_{x\to-\infty} F_t^n(x) = 0, \quad \lim_{|x|\to\infty} f_t^n(x) = 0.$$

Q. E. D.

Next we prove the first part of (2.16).

Lemma 2.10. *Under the assumption of Theorem 2.6, for $n \ge 1$, $t \in [0,T]$ and $x \in R$,*

$$P(X^n(t) \le x) = F_t^n(x). \qquad (2.44).$$

(Proof). Fix $n \ge 1$ and put, for $k \ge 1$, $t \in [0,T]$ and $x \in R$,

$$\begin{aligned} &g^k(x) = k/(2\pi)^{1/2}\exp(-k^2x^2/2), \quad G^k(x) = \int_{-\infty}^x g^k(y)dy, \\ &h_t(x) = P(X^n(t) \le x) - F_t^n(x), \quad u^k(t,x) = \int_R G^k(x-y)h_t(dy). \end{aligned} \qquad (2.45).$$

Then one can show that the following holds: for $k \geq 1$ and $t \in [0, T]$

$$\int_R |u(t,x)|^2 dx \leq \lim_{k\to\infty} \int_R |u^k(t,x)|^2 dx = 0, \tag{2.46}$$

which completes the proof.

Since by Lemma 2.9 and (2.36),

$$\sup_{0\leq t\leq T} \{E[|X^n(t)|] + \int_R |y| f_t^n(y) dy\} < \infty, \tag{2.47}$$

letting $r \to \infty$ and then $k \to \infty$ in (2.48) below, we get (2.46), again from Lemma 2.9: for $r > 0$, by (2.18) and the Ito formula,

$$\begin{aligned}
&\int_{-r}^{r} |u^k(t,x)|^2 dx \qquad\qquad (2.48).\\
&\leq (\int_0^t ds |[\int_R a^n(s,y) g^k(x-y) h_s(dy) u^k(s,x)]_{x=-r}^{r}| \\
&\quad + \int_0^t ds \int_{R^3} C_n |y_3/k| g^k(y_1 + y_3/k - y_2) |dh_s(y_1)/dy| |dh_s(y_2)/dy| g^1(y_3) dy_1 dy_2 dy_3 \\
&\quad + 2\int_0^t ds \int_{R^2} C_n |y_1/k| g^1(y_1) |dh_s(y_2)/dy| dy_1 dy_2) \exp(tC_u^2/\nu).
\end{aligned}$$

(2.48) is true, since

$$\begin{aligned}
&\int_{-r}^{r} |u^k(t,x)|^2/2 dx \\
&= \int_0^t ds \int_{-r}^{r} \{\int_R [2^{-1} a^n(s,y)(dg^k(x-y)/dx) - b^n(s,y) g^k(x-y)] h_s(dy)\} u^k(s,x) dx \\
&= \int_0^t ds [\int_R 2^{-1} a^n(s,y) g^k(x-y) h_s(dy) u^k(s,x)]_{x=-r}^{r} \\
&\quad - \int_0^t ds \int_{-r}^{r} \{\int_R 2^{-1} a^n(s,y) g^k(x-y) h_s(dy)\} [\partial u^k(s,x)/\partial x] dx \\
&\quad - \int_0^t ds \int_{-r}^{r} \{\int_R b^n(s,y) g^k(x-y) h_s(dy)\} u^k(s,x) dx
\end{aligned}$$

by the integration by parts; and

$$
\begin{aligned}
&-\int_0^t ds \int_{-r}^{r} \{\int_R 2^{-1} a^n(s,y) g^k(x-y) h_s(dy)\} [\partial u^k(s,x)/\partial x] dx \\
&= -\int_0^t ds \int_{-r}^{r} 2^{-1} a^n(s,x) |\partial u^k(s,x)/\partial x|^2 dx \\
&\quad + \int_0^t ds \int_{-r}^{r} \{\int_R 2^{-1} [a^n(s,x) - a^n(s,y)] g^k(x-y) h_s(dy)\} [\partial u^k(s,x)/\partial x] dx \\
&\le -(\nu/2) \int_0^t ds \int_{-r}^{r} |\partial u^k(s,x)/\partial x|^2 dx \\
&\quad + \int_0^t ds \int_{-r}^{r} \{\int_R 2^{-1} [a^n(s,x) - a^n(s,y_1)] g^k(x-y_1) h_s(dy_1)\} \{\int_R g^k(x-y_2) h_s(dy_2)\} dx,
\end{aligned}
$$

and for $s \in [0,t]$ and y_1, $y_2 \in R$,

$$
\int_R |a^n(s,x) - a^n(s,y_1)| g^k(x-y_1) g^k(x-y_2) dx \le C_n \int_R |y_3/k| g^1(y_3) g^k(y_1 + y_3/k - y_2) dy_3,
$$

by the mean value theorem, from Lemma 2.9; and

$$
\begin{aligned}
&-\int_0^t ds \int_{-r}^{r} \{\int_R b^n(s,y) g^k(x-y) h_s(dy)\} u^k(s,x) dx \\
&= -\int_0^t ds \int_{-r}^{r} b^n(s,x) [\partial u^k(s,x)/\partial x] u^k(s,x) dx \\
&\quad + \int_0^t ds \int_{-r}^{r} \{\int_R [b^n(s,x) - b^n(s,y)] g^k(x-y) h_s(dy)\} u^k(s,x) dx \\
&\le \int_0^t ds \int_{-r}^{r} [(\nu/2) |\partial u^k(s,x)/\partial x|^2 + 1/(2\nu) |b^n(s,x) u^k(s,x)|^2] dx \\
&\quad + \int_0^t ds \int_{R^2} C_n |x-y| g^k(x-y) |dh_s(y)/dy| dx dy,
\end{aligned}
$$

by the mean value theorem, from Lemma 2.9.

Q. E. D.

Finally we prove the second part of (2.16).

Lemma 2.11. *Under the assumption of Theorem 2.6, for $t \in (0, T]$ and the continuity point x of $G_t(\cdot)$,*

$$P(b^n(t, X^n(t)) \le x) \to G_t(x) \quad \text{as } n \to \infty. \qquad (2.49).$$

(Proof). For $t \in (0, T]$, $\delta > 0$ and the continuity point x of $G_t(\cdot)$,

$$\begin{aligned} &P(G_t^{-1}(1 - F_t(X^n(t) - \delta)) + 2C_uC(n^2\delta^3)^{-1} f_t^n(X^n(t))^{-1} \le x) \qquad (2.50). \\ &\le P(b^n(t, X^n(t)) \le x) \\ &\le P(G_t^{-1}(1 - F_t(X^n(t) + \delta)) - 2C_uC(n^2\delta^3)^{-1} f_t^n(X^n(t))^{-1} \le x). \end{aligned}$$

This is true, since for $x \in R$,

$$\begin{aligned} b(t, x + \delta) \int_{|x-y|<\delta} c^n(x - y) F_t(dy)/f_t^n(x) &\le \int_{|x-y|<\delta} b(t, y) c^n(x - y) F_t(dy)/f_t^n(x) \\ &\le b(t, x - \delta) \int_{|x-y|<\delta} c^n(x - y) F_t(dy)/f_t^n(x) \end{aligned}$$

(notice that $b(t, \cdot)$ is nonincreasing), and since

$$\int_{|x-y|\ge\delta} c^n(x - y) F_t(dy)/f_t^n(x) \le nC(1 + n^2\delta^2)^{-3/2}/f_t^n(x) < C(n^2\delta^3)^{-1}/f_t^n(x).$$

Let us show that the probability in both ends of (2.50) converge to $G_t(x)$ as $n \to \infty$ and then $\delta \to 0$. The left end can be considered as follows:

$$\begin{aligned} &P(G_t^{-1}(1 - F_t(X^n(t) - \delta)) + 2C_uC(n^2\delta^3)^{-1} f_t^n(X^n(t))^{-1} \le x) \qquad (2.51). \\ &\ge P(G_t^{-1}(1 - F_t(X^n(t) - \delta)) \le x - \delta) - P(2C_uC(n^2\delta^3)^{-1} f_t^n(X^n(t))^{-1} \ge \delta), \end{aligned}$$

and

$$P(G_t^{-1}(1-F_t(X^n(t)-\delta)) \le x-\delta) \tag{2.52}$$
$$= P(1-F_t(X^n(t)-\delta) \le G_t(x-\delta)) \quad \text{(see (2.8))}$$
$$\to P(1-G_t(x-\delta) \le F_t(X^u(t)-\delta)) \quad \text{as } n \to \infty \text{ (by Lemma 2.10)}$$
$$\to P(1-G_t(x) \le F_t(X^u(t))) = G_t(x) \quad \text{as } \delta \to 0.$$

Take $k \ge 1$ and let $n \ge k$. Then by Lemma 2.10,

$$P(2C_uC(n^2\delta^3)^{-1}f_t^n(X^n(t))^{-1} \ge \delta) \tag{2.53}$$
$$= P(2C_u\delta^{-3}(\int_R (n^{-2}+(X^n(t)-y)^2)^{-3/2}F_t(dy))^{-1} \ge 1)$$
$$\le P(2C_u\delta^{-3}(\int_R (k^{-2}+(X^n(t)-y)^2)^{-3/2}F_t(dy))^{-1} \ge 1)$$
$$\to P(2C_u\delta^{-3}(\int_R (k^{-2}+(X^u(t)-y)^2)^{-3/2}F_t(dy))^{-1} \ge 1) \quad \text{as } n \to \infty$$
$$\to P(2C_u\delta^{-3}(\int_R |X^u(t)-y|^{-3}F_t(dy))^{-1} \ge 1) = 0 \quad \text{as } k \to \infty,$$

since $P(X^u(t) \in dx)$ is absolutely continuous with respect to dx for all $t \in (0,T]$ (see Lemma 2.12 given below).

The right end of (2.50) can be shown to converge to $G_t(x)$, as $n \to \infty$ and then $\delta \to 0$, in the same way as above.

Q. E. D.

Let us prove the last equality of (2.53).

Lemma 2.12. *Suppose that (A.1)-(A.3) hold. Then for $t \in (0,T]$,*

$$P(\int_R |X^u(t)-y|^{-3}F_t(dy) < \infty) = 0. \tag{2.54}$$

(Proof). Put

$$S_t \equiv \{x \in R; f_t(x) \equiv \lim_{h\to 0}(F_t(x+h)-F_t(x))/h > 0\}. \tag{2.55}$$

Then we only have to show that for $x \in S_t$,

$$\int_R |x-y|^{-3} F_t(dy) = \infty, \tag{2.56}$$

since for $t \in (0, T]$

$$P(X^u(t) \in S_t) = 1.$$

This is true, since $P(X^u(t) \in dx)$ is absolutely continuous with respect to dx for $t \in (0, T]$ (see Remark 2.1).

Let us prove (2.56). For $x \in S_t$ and $\delta > 0$,

$$\begin{aligned}
&\int_R |x-y|^{-3} F_t(dy) \\
&> \int_{|x-y|<\delta} |x-y|^{-3} F_t(dy) \\
&> \delta^{-3}(F_t(x+\delta) - F_t(x-\delta)) \to \infty \quad \text{as } \delta \to 0.
\end{aligned} \tag{2.57}$$

Q. E. D.

3. Applications to Markov optimal control problems.

In this section we consider the applications of Markov marginal problems discussed in section 2 to Markov optimal control problems.

Let $k_1(t,x)$, $k_2(t,x) : [0,T] \times R^d \mapsto R$ and $K_1(x)$, $K_2(x) : R^d \mapsto R$ be bounded from below and continuous. For a R^d-valued stochastic process $\{X(t)\}_{0 \le t \le T}$ on a probability space $(\Omega, \mathbf{B}, P)$, put

$$\begin{aligned} F_t^X(x) &\equiv P(X^i(t) \le x^i (i = 1, \cdots, d)) \quad (x = (x^i)_{i=1}^d \in R^d), \\ J_1(X) &\equiv E[\int_0^T k_1(t, X(t))dt + K_1(X(0)) + K_2(X(T))]. \end{aligned} \tag{3.1}$$

For the solution to (1.8), put

$$\begin{aligned} J_2(X^u) &\equiv E[\int_0^T k_2(t, u(t))dt + K_1(X^u(0))], \\ J_3(X^u) &\equiv E[\int_0^T [k_1(t, X^u(t)) + k_2(t, u(t))]dt + K_1(X^u(0)) + K_2(X^u(T))]. \end{aligned} \tag{3.2}$$

The following theorem can be found in [19].

Theorem 3.1.

$$\begin{aligned} &\inf\{J_1(X); F_t^X(\cdot) \text{ is continuous for } t \in [0,T]\} \\ &= \inf\{J_1(X); F_t^X(\cdot) \text{ is continuous for } t \in [0,T], \{X(t)\}_{0 \le t \le T} \text{ is a Markov process }\}, \end{aligned} \tag{3.3}$$

provided that the left hand side is finite.

The following theorem can be proved from Theorem 2.5 and will be proved later.

Theorem 3.2. *Suppose that (A.1)-(A.2) hold. Then for any $\varepsilon > 0$, there exist bounded measurable $b^\varepsilon(\cdot,\cdot) : [0,T] \times R \mapsto R$ and a unique weak solution to the following : for $t \in [0,T]$*

$$X^\varepsilon(t) = X_o + \int_0^t b^\varepsilon(s, X^\varepsilon(s))ds + \int_0^t \sigma(s, X^\varepsilon(s))dW(s) \tag{3.4}$$

such that the following holds: for $t \in (0,T]$

$$\lim_{\varepsilon\to 0} P(b^\varepsilon(t, X^\varepsilon(t)) \in dx) = P(u(t) \in dx) \quad \text{weakly.} \tag{3.5}$$

As an application of Theorem 3.2, the following can be proved easily and the proof is omitted.

Corollary 3.3. *Suppose that (A.1) holds. Then the following holds:*

$$\begin{aligned} &\inf\{J_2(X^u); \text{(A.2) holds for } u(t)\} \\ &= \inf\{J_2(X^u); \text{(A.2) holds for } u(t) = b(t, X(t)) \text{ for some } b(t,x)\}, \end{aligned} \tag{3.6}$$

provided that the left hand side is finite.

The following theorem can be proved from Theorem 2.6 and will be proved later.

Theorem 3.4. *Suppose that (A.1)-(A.4) hold. Then for any $\varepsilon > 0$ and $n \geq 1$, there exist bounded, measurable $\tilde{b}^{\varepsilon,n}(\cdot,\cdot) : [0,T] \times R \mapsto R$, and measurable $\tilde{a}^{\varepsilon,n}(\cdot,\cdot) : [0,T] \times R \mapsto [\nu, \infty)$ such that there exists a unique strong solution to the following: for $t \in [0,T]$*

$$\tilde{X}^{\varepsilon,n}(t) = X^{\varepsilon,n}(0) + \int_0^t \tilde{b}^{\varepsilon,n}(s, \tilde{X}^{\varepsilon,n}(s))ds + \int_0^t \tilde{a}^{\varepsilon,n}(s, \tilde{X}^{\varepsilon,n}(s))^{1/2} dW(s), \tag{3.7}$$

and such that the following holds: for $t \in (0,T]$

$$\begin{aligned} \lim_{\varepsilon\to 0}\lim_{n\to 0} P(\tilde{X}^{\varepsilon,n}(t) \in dx) &= P(X^u(t) \in dx) \quad (0 \leq t \leq T) \text{ weakly}, \\ \lim_{\varepsilon\to 0}\lim_{n\to 0} P(\tilde{b}^{\varepsilon,n}(t, \tilde{X}^{\varepsilon,n}(t)) \in dx) &= P(u(t) \in dx) \quad (0 < t \leq T) \text{ weakly}. \end{aligned} \tag{3.8}$$

When we clarify the dependence of X^u on σ, we write $X^{u,\sigma}$. As an application of Theorem 3.4, the following can be proved easily and the proof is omitted.

Corollary 3.5. *Suppose that (A.1) and (A.4) hold. Then the following holds:*

$$\begin{aligned} &\inf\{J_3(X^{u,\sigma});\ \text{(A.2)-(A.3) hold for } u(t)\} \\ &= \inf\{J_3(X^{u,\sigma'});\ \text{(A.2)-(A.3) hold for } u(t) = b(t, X^{u,\sigma'}(t)) \text{ for some } b(t,x), \sigma'(t,x)^2 \geq \nu\}, \end{aligned} \tag{3.9}$$

provided that the left hand side is finite.

Let us prove Theorem 3.2 first.

(Proof of Theorem 3.2.) For any $\varepsilon > 0$, put

$$G_{t,\varepsilon}^{-1}(u) = \begin{cases} \max(G_t^{-1}(G_t(-1/\varepsilon)), -1/\varepsilon) & \text{if } u \leq G_t(-1/\varepsilon), \\ G_t^{-1}(u) & \text{if } G_t(-1/\varepsilon) < u < G_t(1/\varepsilon), \\ G_t^{-1}(G_t(1/\varepsilon)) & \text{if } u \geq G_t(1/\varepsilon) \end{cases} \tag{3.10.}$$

(see (2.5) for the notation of G_t^{-1}). For $\varphi \in C_o^\infty(R; [0,\infty))$ which is not identically zero, put

$$B^\varepsilon(t,u) = \int_R \varphi((u-y)/\varepsilon) G_{t,\varepsilon}^{-1}(y) dy / C_\varepsilon, \tag{3.11.}$$

where we put $C_\varepsilon = \varepsilon \int_R \varphi(y) dy$. Then for $\varepsilon > 0$, $B^\varepsilon(t,u)$ is bounded and globally Lipschitz continuous in u, uniformly in $t \in [0,T]$, since

$$|G_{t,\varepsilon}^{-1}(u)| \leq 1/\varepsilon.$$

Let $v^\varepsilon(t,x)$ be the solution to (2.11) with $B = B^\varepsilon$. Then the following has a unique weak solution from Theorem 2.5: for $t \in [0,T]$

$$X^\varepsilon(t) = X_o + \int_0^t B^\varepsilon(s, v^\varepsilon(s, X^\varepsilon(s)))ds + \int_0^t \sigma(s, X^\varepsilon(s))dW(s), \tag{3.12.}$$

and letting $\varepsilon \to 0$, for $t \in (0,T]$ and the continuity point x of G_t,

$$P(B^\varepsilon(t, u^\varepsilon(t, X^\varepsilon(t))) \leq x) \to G_t(x). \tag{3.13.}$$

Let us show that (3.13) is true to complete the proof. Take $r > 0$ so that $supp(\varphi) \subset [-r, r]$. For $x \in R$,

$$\begin{aligned} P(G_{t,\varepsilon}^{-1}(v^\varepsilon(t, X^\varepsilon(t)) + r\varepsilon) \leq x) &\leq P(B^\varepsilon(t, v^\varepsilon(t, X^\varepsilon(t))) \leq x) \\ &\leq P(G_{t,\varepsilon}^{-1}(v^\varepsilon(t, X^\varepsilon(t)) - r\varepsilon) \leq x), \end{aligned} \tag{3.14.}$$

since $G_{t,\varepsilon}^{-1}$ is nondecreasing.

Suppose that $G_t(x) \in (0,1]$. Then for sufficiently small ε, $G_t(-1/\varepsilon) < G_t(x) \leq G_t(1/\varepsilon)$, and

$$
\begin{aligned}
&P(B^\varepsilon(t, v^\varepsilon(t, X^\varepsilon(t))) \leq x) \geq P(G_{t,\varepsilon}^{-1}(v^\varepsilon(t, X^\varepsilon(t)) + r\varepsilon) \leq x) \quad (\text{from } (3.14)) \quad (3.15). \\
&\geq P(G_{t,\varepsilon}^{-1}(v^\varepsilon(t, X^\varepsilon(t)) + r\varepsilon) \leq G_t^{-1}(G_t(x))) \quad (\text{since } G_t^{-1}(G_t(x)) \leq x) \\
&\geq P(G_{t,\varepsilon}^{-1}(v^\varepsilon(t, X^\varepsilon(t)) + r\varepsilon) \leq G_{t,\varepsilon}^{-1}(G_t(x))) \quad (\text{see } (3.10)) \\
&\geq P(v^\varepsilon(t, X^\varepsilon(t)) + r\varepsilon \leq G_t(x)) = G_t(x) - r\varepsilon \to G_t(x),
\end{aligned}
$$

as $\varepsilon \to 0$, since $G_{t,\varepsilon}^{-1}$ is nondecreasing (see (3.10)) and since $v^\varepsilon(t, X^\varepsilon(t))$ is uniformly distributed on $[0,1]$ for $t \in (0,T]$ (see (2.13)).

Suppose that $G_t(x) \in [0,1)$. Then for sufficiently small ε, $G_t(-1/\varepsilon) \leq G_t(x) < G_t(1/\varepsilon)$. Suppose also that $G_t(x) = G_t(x-)$. Then from (3.14),

$$
\begin{aligned}
&P(B^\varepsilon(t, v^\varepsilon(t, X^\varepsilon(t))) \leq x) \leq P(G_{t,\varepsilon}^{-1}(v^\varepsilon(t, X^\varepsilon(t)) - r\varepsilon) \leq x) \qquad (3.16). \\
&\leq P(v^\varepsilon(t, X^\varepsilon(t)) - r\varepsilon \leq G_t(x)) = G_t(x) + r\varepsilon \to G_t(x),
\end{aligned}
$$

as $\varepsilon \to 0$, in the same way as in (3.15). Here we used the follows: if $u > G_t(x)$, then

$$
G_{t,\varepsilon}^{-1}(u) = \begin{cases} G_t^{-1}(u) > x & \text{if } G_t(1/\varepsilon) > u > G_t(x), \\ G_t^{-1}(G_t(1/\varepsilon)) > x & \text{if } u \geq G_t(1/\varepsilon). \end{cases} \qquad (3.17).
$$

Q. E. D.

Finally we prove Theorem 3.4.

(Proof of Theorem 3.4). For $\varepsilon > 0$, let $\{X^{u,\varepsilon}(t)\}_{0 \leq t \leq T}$ satisfy the following; for $t \in [0,T]$,

$$
X^{u,\varepsilon}(t) = X_o + \int_0^t \min(\max(u'(s, X^{u,\varepsilon}), -1/\varepsilon), 1/\varepsilon) ds + \int_0^t \sigma(s, X^{u,\varepsilon}(s)) d\overline{W}(s). \quad (3.18).
$$

The existence and uniqueness of solution to (3.18) can be easily shown by Cameron-Martin-Maruyama-Girsanov formula (see Remark 2.1 and [16]).

From Theorem 2.6, there exist measurable $\tilde{a}^{\varepsilon,n}(\cdot,\cdot) : [0,T] \times R \mapsto [\nu,\infty)$ and $\tilde{b}^{\varepsilon,n}(\cdot,\cdot) : [0,T] \times R \mapsto R$ such that the following holds: for $t \in [0,T]$,

$$\tilde{X}^{\varepsilon,n}(t) = \tilde{X}^{\varepsilon,n}(0) + \int_0^t \tilde{b}^{\varepsilon,n}(s,\tilde{X}^{\varepsilon,n}(s))ds + \int_0^t \tilde{a}^{\varepsilon,n}(s,\tilde{X}^{\varepsilon,n}(s))^{1/2}d\tilde{W}(s), \qquad (3.19).$$

and

$$P(\tilde{X}^{\varepsilon,n}(t) \in dx) \to P(X^{u,\varepsilon}(t) \in dx) \to P(X^u(t) \in dx) \quad (0 \le t \le T), (3.20).$$

$$\begin{aligned} P(\tilde{b}^{\varepsilon,n}(t,\tilde{X}^{\varepsilon,n}(t)) \in dx) &\to P(\min(\max(u'(t,X^{u,\varepsilon}),-1/\varepsilon),1/\varepsilon) \in dx) \\ &\to P(u(t) \in dx) \quad (0 < t \le T), \qquad (3.21). \end{aligned}$$

weakly as $n \to \infty$ and then $\varepsilon \to 0$.

Q. E. D.

4. Discussion.

Our approach to MOCP is, roughly speaking, as follows. For X^u in (1.8), find nonlinear PDE:

$$\begin{aligned} \partial v(t,x)/\partial t &= 2^{-1}\partial(\Phi_2(t,x,v(\cdot,\cdot),\partial v(\cdot,\cdot)/\partial x,\cdots)\partial v(t,x)/\partial x)/\partial x \\ &\quad - \Phi_1(t,x,v(\cdot,\cdot),\partial v(\cdot,\cdot)/\partial x,\cdots)\partial v(t,x)/\partial x \quad ((t,x) \in (0,T) \times R), \quad (4.1). \\ v(0,x) &= P(X^u(0) \le x) \quad x \in R \end{aligned}$$

such that the following SDE has a weak solution:

$$\begin{aligned} dX(t) &= \Phi_2(t,X(t),v(\cdot,\cdot),\partial v(\cdot,\cdot)/\partial x,\cdots)^{1/2}dW(t) \qquad (4.2). \\ &\quad + \Phi_1(t,X(t),v(\cdot,\cdot),\partial v(\cdot,\cdot)/\partial x,\cdots)dt \quad (t > 0), \end{aligned}$$

and that

$$P((X^u(t), u(t)) \in dx) \tag{4.3.}$$
$$= P((X(t), \Phi_1(t, X(t), v(\cdot,\cdot), \partial v(\cdot,\cdot)/\partial x, \cdots)) \in dx) \quad (t > 0).$$

If this can be done for the minimizing X^u in (1.9), then MOCP can be solved completely.

In Theorems 3.2 and 3.4, we put heuristically,

$$\Phi_2 = \sigma^2(t,x) \quad \text{and } \Phi_1 = G_t^{-1}(v(t,x)), \tag{4.4.}$$

where $G_t(x) = P(u(t) \le x)$; and

$$\Phi_1 = G_t^{-1}(1 - v(t,x)),$$
$$\Phi_2 = \sigma(t,x)^2 + 1_{(0,T]}(t) 2[\partial v(t,x)/\partial x]^{-1} \int_{-\infty}^{x} [G_t^{-1}(1 - v(t,y)) - b(t,y)][\partial v(t,y)/\partial y] dy \tag{4.5.}$$

where $b(t,y) = E[u(t)|X^u(t) = y]$, respectively.

ACKNOWLEDGEMENT. MOCP was introduced to me when I was a student of Professor Wendell H. Fleming's during 1988-1989. I would like to thank him for his constant encouragement and useful suggestion.

REFERENCES.

[1] Blumenthal, R. M. and Corson, H. H., On continuous collections of measures, Ann. Inst. Fourier. **20**, 2 (1970) 193-199.

[2] Blumenthal, R. M. and Corson, H. H., On continuous collections of measures, Proc. 6th Berkeley Sympos. Math. Statist. Probab. **2** (1972) 33-40.

[3] Cacoullos, T. and Papathanasiou, V., Characterizations of distributions by variance bounds, Statist. Probab. Lett. **7** (1989) 351-356.

[4] Cacoullos, T., Papathanasiou, V. and Utev, S. A., Another characterization of the normal law and a proof of the central limit theorem connected with it, Theory Probab. Appl. **37** (1992) 581-588.

[5] Cacoullos, T., Papathanasiou, V. and Utev, S. A., Variational inequalities with examples and an application to the central limit theorem, Ann. Probab. **22** (1994) 1607-1618.

[6] Carlen E. A., Conservative diffusions, Commun. Math. Phys. **94** (1984) 293-315.

[7] Carmona R., Probabilistic constraction of Nelson processes, In: Itô, K., Ikeda, N. (eds.) Probabilistic Methods in Mathematical Physics. Proceedings, Katata 1985, Kinokuniya, Tokyo, 1987, pp. 55-81.

[8] Cuesta-Albertos, J. A. and Tuero-Diaz, A., A characterization for the solution of the Monge-Kantorovich mass transference problem, Statist. Probab. Lett. **16** (1993) 147-152.

[9] Darsow W. F., Nguyen B. and Olsen E. T., Copulas and Markov processes, Illinois. J. Math. **36** (1992) 600-642.

[10] Edwards, D. A., On the existence of probability measures with given marginals, Ann. Inst. Fourier. **28** 4 (1978) 53-78.

[11] Fleming W. H. and Soner H. M., Controlled Markov processes and Viscosity solutions, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1993.

[12] Freidlin, M. I., Functional integration and partial differential equations, Princeton University Press, Princeton, 1985.

[13] Gangbo, W. and McCann, R. J., The geometry of optimal transpotation, preprint.

[14] Kamae, T. and Krengel, U., Stochastic partial ordering, Ann. Probab. **6** (1978) 1044-1049.

[15] Kellerer, H. G., Representation of Markov kernels by random mappings under order conditions, In: Benes, V., Stepan, J.(eds.) Distributions with given Marginals and Moment Problems, Proceedings, Prague 1996, Kluwer Academic Publishers, Dordrecht, Boston, London, 1997, pp. 143-160.

[16] Liptser, R. S. and Shiryaev, A. N., Statistics of random processes I, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1977.

[17] Mikami, T., Small random perturbations of dynamical systems and Markov processes from marginal distributions, Thesis, Brown University 1989.

[18] Mikami, T., Variational processes from the weak forward equation, Commun.. Math.

Phys. **135** (1990) 19-40.

[19] Mikami, T., Copula fields and their applications, Proc. Japan Acad. Ser. A **71** (1995) 221-224.

[20] Mikami, T., Equivalent conditions on the central limit theorem for a sequence of probability measures on R, to appear in Statist. Probab. Lett..

[21] Nagasawa M., Transformations of diffusion and Schrödinger process, Probab. Th. Rel. Fields **82** (1989) 109-136.

[22] Nelson E., Dynamical theories of Brownian motion, Princeton University Press, Princeton, 1967.

[23] Nelson E., Quantum fluctuations, Princeton University Press, Princeton, 1984.

[24] Quastel, J. and Varadhan, S. R. S., Diffusion semigroups and diffusion processes corresponding to degenerate divergence form operators, Comm. Pure Appl. Math. **50** (1997) 667-706.

[25] Rüschendorf, L. and Uckelmann, L., On optimal multivariate couplings, In: Benes, V., Stepan, J.(eds.) Distributions with given Marginals and Moment Problems, Proceedings, Prague 1996, Kluwer Academic Publishers, Dordrecht, Boston, London, 1997, pp. 143-160.

[26] Schweizer B., Thirty years of copulas, In: Dall'Aglio, G., Kotz, S., Salinetti, G. (eds.) Advances in Probability Distributions with Given Marginals, Proceedings, Rome 1990, Kluwer Academic Publishers, Dordrecht, Boston, London, 1991, pp. 13-50.

[27] Schweizer B. and Sklar A., Probabilistic metric space, North - Holland , New York, Amsterdam, 1983.

[28] Stroock, D. W. and Varadhan, S. R. S., Multidimensional diffusion processes, Springer - Verlag, Berlin, Heidelberg, New York, Tokyo, 1979.

[29] Sudakov, V. N., Geometric problems in the theory of infinite dimensional probability distributions, Proc. Steklov Inst. Math. **141** (1979) 1-178.

[30] Zheng W. A., Tightness results for laws of diffusion processes application to stochastic mechanics, Ann. Inst. Henri Poincaré, Sect. B **21** (1985) 103-124.