Electron. Commun. Probab. **27** (2022), article no. 26, 1–12.
https://doi.org/10.1214/22-ECP466
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Order preservation and positive correlation for nonlinear Fokker-Planck equation*

Panpan Ren[†] [‡]

Abstract

By investigating McKean-Vlasov SDEs, the order preservation and positive correlation are characterized for nonlinear Fokker-Planck equations. The main results recover the corresponding criteria on these properties established in [3, 5] for diffusion processes or linear Fokker-Planck equations.

Keywords: nonlinear Fokker-Planck equation; order preservation; positive correlation.
MSC2020 subject classifications: Primary 60J60, Secondary 60H30.
Submitted to ECP on October 24, 2021, final version accepted on April 16, 2022.

1 Introduction

Based on [5], complete criteria have been established in [3] for the order preservation and positive correlation for diffusion processes corresponding to linear Fokker-Planck equations, where the order preservation links to comparison theorem in the literature of SDEs, and the positive correlation arises from statistics is known as Fortuin–Kasteleyn–Ginibre (FKG) inequality due to [4]. In the present paper we aim to extend these criteria to nonlinear Fokker-Planck equations associated with McKean-Vlasov SDEs.

In this paper, we aim to extend the above results to nonlinear Fokker-Planck equations on the Wasserstein space $\mathcal{P}_2 = \{\mu \in \mathcal{P} : \mu(|\cdot|^2) < \infty\}$, a space of probability measures on $\mathbb{R}^d$ with second moment. Under the Wasserstein distance

$$\mathbb{W}_2(\mu,\nu) := \inf_{\pi\in\mathcal{C}(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \pi(\mathrm{d}x,\mathrm{d}y)\right)^{\frac{1}{2}},$$

$\mathcal{P}_2$ is a Polish space., where $\mathcal{C}(\mu,\nu)$ is the space of all couplings of μ and ν. Consider the following time-distribution dependent second order differential operators

$$L_{t,\mu} := \mathrm{tr}\{a(t,\cdot,\mu)\nabla^2\} + b(t,\cdot,\mu)\cdot\nabla, \quad \bar{L}_{t,\mu} := \mathrm{tr}\{\bar{a}(t,\cdot,\mu)\nabla^2\} + \bar{b}(t,\cdot,\mu)\cdot\nabla, \tag{1.1}$$

where $a = (a_{ij})_{1\le i,j\le d},\ \bar{a} = (\bar{a}_{ij})_{1\le i,j\le d},\ b = (b_i)_{1\le i\le d},\ \bar{b} = (\bar{b}_i)_{1\le i\le d}$ are continuous on $[0,\infty)\times\mathbb{R}^d\times\mathcal{P}_2$. The nonlinear Fokker-Planck equations for L and $\bar{L}$ are formulated as

$$\partial_t\mu_t = L^*_{t,\mu_t}\mu_t, \quad \partial_t\bar{\mu}_t = \bar{L}^*_{t,\bar{\mu}_t}\bar{\mu}_t, \quad t \ge s. \tag{1.2}$$

*The author greatefully acknowledge support by European Union through the ERC-AdG 694405"Ricci-Bounds" and Alexander von Humboldt Foundation.

†Department of Mathematics, University of Bonn, Bonn 53115, Germany.
‡Department of Mathematics, Technical University of Kaiserslautern, Kaiserslautern, 67663, Germany. E-mail: RPPZOE@GMAIL.COM.

We call $(\mu_t, \bar{\mu}_t)_{t\geq s} \in C([s,\infty);\mathcal{P}_2)\times C([s,\infty);\mathcal{P}_2)$ a solution to (1.2), if for any $f \in C_0^\infty(\mathbb{R}^d)$,

$$\mu_t(f) = \mu_s(f) + \int_s^t \mu_r(L_{r,\mu_r} f)\mathrm{d}r, \;\; \bar{\mu}_t(f) = \mu_s(f) + \int_s^t \bar{\mu}_r(\bar{L}_{r,\bar{\mu}_r} f)\mathrm{d}r, \;\; t \geq s.$$

To characterize (1.2) using McKean-Vlasov SDEs, we make the following assumption.

(A) $b, \bar{b}, a, \bar{a}$ are continuous on $[0,\infty)\times\mathbb{R}^d\times\mathcal{P}_2$, a and $\bar{a}$ are positive definite,

$$\mu\big(|b(t,\cdot,\mu)| + \|a(t,\cdot,\mu)\|\big) := \int_{\mathbb{R}^d} \big(|b(t,\cdot,\mu)| + \|a(t,\cdot,\mu)\|\big)\mathrm{d}\mu$$

is locally bounded in $(t,\mu) \in [0,\infty)\times\mathcal{P}_2$, and there exits an increasing function $K : [0,\infty) \to [0,\infty)$ such that $b, \bar{b}, \sigma := \sqrt{2a}$ and $\bar{\sigma} := \sqrt{2\bar{a}}$ satisfy

$$\begin{aligned} \max\Big\{&2\langle b(t,x,\mu) - b(t,y,\nu), x-y\rangle + \|\sigma(t,x,\mu) - \sigma(t,x,\nu)\|_{HS}^2, \\ &2\langle \bar{b}(t,x,\mu) - \bar{b}(t,y,\nu), x-y\rangle + \|\bar{\sigma}(t,x,\mu) - \bar{\sigma}(t,x,\nu)\|_{HS}^2\Big\} \\ &\leq K(t)(|x-y|^2 + \mathbb{W}_2(\mu,\nu)^2), \quad t\geq 0, \;\; x,y\in\mathbb{R}^d, \;\; \mu,\nu\in\mathcal{P}_2, \end{aligned} \tag{1.3}$$

where $\|\cdot\|_{HS}$ stands for the Hilbert-Schmidt norm of a matrix $\cdot$.

Consider the distribution dependent SDEs

$$\begin{aligned} \mathrm{d}X_t &= b(t, X_t, \mathcal{L}_{X_t})\mathrm{d}t + \sqrt{2a(t,X_t,\mathcal{L}_{X_t})}\mathrm{d}W_t, \\ \mathrm{d}\bar{X}_t &= \bar{b}(t, \bar{X}_t, \mathcal{L}_{\bar{X}_t})\mathrm{d}t + \sqrt{2\bar{a}(t,\bar{X}_t,\mathcal{L}_{\bar{X}_t})}\mathrm{d}W_t, \end{aligned} \tag{1.4}$$

where W_t is a d-dimensional Brownian motion on a complete filtration probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, and $\mathcal{L}_\xi$ denotes the distribution of a random variable ξ.

According to [13], **(A)** implies the strong and weak well-posedness of (1.4) for initial distributions in $\mathcal{P}_2$; that is, for any $s \geq 0$ and any $\mathcal{F}_s$-measurable $(X_s, \bar{X}_s)$ with $\mu := \mathcal{L}_{X_s}, \bar{\mu} := \mathcal{L}_{\bar{X}_s} \in \mathcal{P}_2$, the (1.4) for $t\geq s$ has a unique strong solution $(X_t, \bar{X}_t)_{t\geq s}$ as well as a unique weak solution with initial distributions $(\mu,\bar{\mu})$ at times s, such that

$$P_{s,t}^*\mu := \mathcal{L}_{X_t}, \;\; \bar{P}_{s,t}^*\bar{\mu} := \mathcal{L}_{\bar{X}_t} \;\text{ for } \mu = \mathcal{L}_{X_s}, \bar{\mu} = \mathcal{L}_{\bar{X}_s} \tag{1.5}$$

are continuous in $\mathcal{P}_2$ with respect to $t \geq s$. Therefore, by the superposition principle in [1], **(A)** implies that $(\mu_t, \bar{\mu}_t)_{t\geq s} := (P_{s,t}^*\mu, \bar{P}_{s,t}^*\bar{\mu})_{t\geq s}$ is the unique solution of (1.2) with $(\mu_s, \bar{\mu}_s) = (\mu, \bar{\mu})$.

We will also consider the order preservation and positive correlations for

$$\Lambda_s\mu := \mathcal{L}_{(X_t)_{t\geq s}}, \quad \bar{\Lambda}_s\bar{\mu} = \mathcal{L}_{(\bar{X}_t)_{t\geq s}}, \quad s\geq 0, \mu, \bar{\mu} \in \mathcal{P}_2.$$

Unlike in the setting of standard Markov processes, due to the non-linearity these properties for $(\Lambda_s\mu, \bar{\Lambda}_s\bar{\mu})$ do not imply by their time-marginals $(P_{s,t}^*\mu, \bar{P}_{s,t}^*\bar{\mu})_{t\geq s}$.

In Section 2, we state our main results on the order preservation for $(P_{s,t}^*, \bar{P}_{s,t}^*)$ and $(\Lambda_s, \bar{\Lambda}_s)$, as well as on the positive correlations for Λ_s. To prove these results, in Section 3 we extend the main results of [3] to the time inhomogeneous setting which are also new in the literature. Finally, the main results are proved in Section 4.

2 Main results

We first define the order preservation and positive correlations in the present setting. We denote $x \leq y$ for $x := (x_i)_{1\leq i\leq d}, \; y := (y_i)_{1\leq i\leq d} \in \mathbb{R}^d$ if $x_i \leq y_i$ for any $1 \leq i \leq d$. Let $\mathcal{B}_b(\mathbb{R}^d)$ be the set of all bounded measurable functions on $\mathbb{R}^d$. Consider the class of bounded measurable increasing functions: $\mathcal{U}_b := \big\{f \in \mathcal{B}_b(\mathbb{R}^d)\big| f(x) \le f(y)$ for $x, y \in \mathbb{R}^d$ with $x \le y\big\}$, and the family of probability measures of positive correlations: $\mathcal{P}_+ := \big\{\mu \in \mathcal{P}|\, \mu(fg) \ge \mu(f)\mu(g) \quad f, g \in \mathcal{U}_b\big\}$. If $\mu \in \mathcal{P}_+$, then μ is said to satisfy the FKG inequality. Moreover, we write $\mu \preceq \nu$ for any two probability measures $\mu, \nu \in \mathcal{P}$, if $\mu(f) \le \nu(f)$ holds for any $f \in \mathcal{U}_b(\mathbb{R}^d)$. Note that definitions of $\mathcal{P}_+$ and $\mu \preceq \nu$ do not change if we replace $\mathcal{B}_b(\mathbb{R}^d)$ by $C_b^k(\mathbb{R}^d)$ for $k \in \mathbb{Z}_+ \cup \{\infty\}$, where $C_b^0(\mathbb{R}^d) = C_b(\mathbb{R}^d)$ denotes the set of all bounded continuous functions on $\mathbb{R}^d$, while when $k \ge 1$ the class $C_b^k(\mathbb{R}^d)$ consists of bounded functions on $\mathbb{R}^d$ having bounded derivatives up to order k. For any $\xi, \eta \in C_s := C([s,\infty);\mathbb{R}^d)$, we denote $\xi \le \eta$ if $\xi_t \le \eta_t$ for all $t \ge s$. For any two probability measures Φ_1, Φ_2 on the path space C_s, we denote $\Phi_1 \preceq \Phi_2$ if $\Phi_1(F) \le \Phi_2(F)$ holds for any bounded increasing function F on C_s. Similarly, let $\mathcal{P}_+^s$ denote the set of probability measures on C_s satisfying the FKG inequality for bounded increasing functions on C_s.

Definition 2.1. Let $t \ge s \ge 0$ and let $(P_{s,t}, \bar{P}_{s,t}^*)$ be in (1.5).

(1) We write $\bar{P}_{s,t}^* \preceq P_{s,t}^*$, if $\bar{P}_{s,t}^*\mu \preceq P_{s,t}^*\nu$ holds for any $\mu, \nu \in \mathcal{P}_2$ with $\mu \preceq \nu$.
(2) We write $\bar{\Lambda}_s \preceq \Lambda_s$, if $\bar{\Lambda}_s\mu \preceq \Lambda_s\nu$ holds for any $\mu, \nu \in \mathcal{P}_2$ with $\mu \preceq \nu$.
(3) We write $P_{s,t}^* \in \mathcal{P}_+$ if $P_{s,t}^*\mathcal{P}_+ \subset \mathcal{P}_+$; and $\Lambda_s \in \mathcal{P}_+^s$ if $\Lambda_s\mu \in \mathcal{P}_+^s$ holds for all $\mu \in \mathcal{P}_+$.

Obviously, $\bar{\Lambda}_s \preceq \Lambda_s$ for all $s \ge 0$ implies $\bar{P}_{s,t}^* \preceq P_{s,t}^*$ for all $t \ge s \ge 0$, but the inverse may not be true in the nonlinear setting. Similarly, $\Lambda_s \in \mathcal{P}_+^s$ implies $P_{s,t}^* \in \mathcal{P}_+$ for any $t \ge s$ but the inverse may not be true.

2.1 Order preservation

The following result provides sufficient conditions for the order preservation.

Theorem 2.1. Assume **(A)** and the following two conditions:

(1) For any $1 \le i \le d$ and $s \ge 0$, $\bar{b}_i(s,x,\nu) \le b_i(s,y,\mu)$ holds for $x \le y$ with $x_i = y_i$ and $\nu \preceq \mu$;
(2) $a = \bar{a}$, and for any $1 \le i, j \le d, s \ge 0$ and $\mu \in \mathcal{P}_2$, $a_{ij}(s,x,\mu)$ depends only on x_i and x_j.

Then $\bar{\Lambda}_s \preceq \Lambda_s$ for all $s \ge 0$. Consequently, $\bar{P}_{s,t}^* \preceq P_{s,t}^*$ for $t \ge s$.

The next two results include necessary conditions for the order preservation, which are weaker than the sufficient ones given in Theorem 2.1. However, they coincide with the sufficient conditions and hence become sufficient and necessary conditions when $b(t,x,\mu)$ and $a(t,x,\mu)$ do not depend on μ. Indeed, when the coefficients do not depend on the distribution, we may take $\nu = \delta_x$ and $\mu = \delta_y$ for $x \le y$ with $x_i = y_i$ in Theorem 2.2 (i), and with $x_i = y_i, x_j = y_j$ in Theorem 2.2(ii), such that these necessary conditions coincide with the sufficient ones in Theorem 2.1.

For any $\mu \in \mathcal{P}$ and $I \subset \{1,\cdots,d\}$, let $\mu_I(A) := \mu(\{x \in \mathbb{R}^d : x_I \in A\}),\ \ A \in \mathcal{B}(\mathbb{R}^{\#I})$ be the marginal distribution of μ with respect to components indexed by I, where $\#I$ denotes the number of elements in I. In particular, we simply denote $\mu_i = \mu_{\{i\}}$

Theorem 2.2. Assume **(A)** and if $\bar{\Lambda}_s \preceq \Lambda_s$ for all $s \ge 0$, then the following conditions hold:

(i) for any $\nu \preceq \mu$ with $\nu_i = \mu_i$, $1 \le i \le d$, there exists a coupling $\pi \in \mathcal{C}(\nu,\mu)$ with $\pi\big(\{x \le y\}\big) = 1$ such that

$$\bar{b}_i(s,x,\nu) \le b_i(s,y,\mu), \quad s \ge 0,\ (x,y) \in \mathrm{supp}\pi.$$

Consequently, $\bar{b}_i(s,x,\mu) \le b_i(s,x,\mu)$ for $s \ge 0, x \in \mathbb{R}^d, \mu \in \mathcal{P}_2$.

(ii) for any $\nu \preceq \mu$ with $\nu_{ij} = \mu_{ij}, 1 \le i, j \le d$, there exists a coupling $\pi \in \mathcal{C}(\nu, \mu)$ with $\pi\big(\{x \le y\}\big) = 1$ such that

$$\bar{a}_{ij}(s, x, \nu) = a_{ij}(s, y, \mu),\ s \ge 0,\ (x, y) \in \text{supp}\pi.$$

Consequently, $a(s, x, \mu) = \bar{a}(s, x, \mu)$ for any $s \ge 0, x \in \mathbb{R}^d, \mu \in \mathcal{P}_2$.

Since $\bar{\Lambda}_s \preceq \Lambda_s$ implies $\bar{P}^*_{s,t} \preceq P^*_{s,t}$, conditions in the following result are also necessary for $\bar{\Lambda}_s \preceq \Lambda_s$.

Theorem 2.3. Assume **(A)**. If $\bar{P}^*_{s,t} \preceq P^*_{s,t}$ for $t \ge s \ge 0$, then

(i) For any $s \ge 0$ and $1 \le i \le d$, $\nu(\bar{b}_i(s, \cdot, \nu)) \le \mu(b_i(s, \cdot, \mu))$ holds for $\nu \preceq \mu$ with $\nu_i = \mu_i$.

(ii) For any $s \ge 0$ and $1 \le i, j \le d$, $\bar{a}_{ij}(s, x, \delta_x) = a_{ij}(s, x, \delta_x)$ holds and $a_{ij}(s, x, \delta_x)$ depends only on x_i and x_j.

2.2 Positive correlations

We first present sufficient conditions for the positive correlations.

Theorem 2.4. Assume **(A)**. If

(1) For any $s \ge 0$ and $1 \le i \le d$, $b_i(s, x, \nu) \le b_i(s, y, \mu),$ for $\nu \preceq \mu, x \le y$ with $x_i = y_i$,

(2) For any $1 \le i, j \le d$, $a_{ij} \ge 0$, and for any $\mu \in \mathcal{P}_+$, $a_{ij}(s, x, \mu)$ depends only on x_i and x_j,

then $\Lambda_s \in \mathcal{P}^s_+$, and consequently, $P^*_{s,t} \in \mathcal{P}_+$ for any $t \ge s \ge 0$.

Theorem 2.5. If $\Lambda^\mu_s \in \mathcal{P}^s_+$ for $s \ge 0$ and $\mu = \mu_{\{ij\}} \times \mu_{\{ij\}^c} \in \mathcal{P}_+$, then

(1) For any $s \ge 0$, $1 \le i, j \le d$, $a_{ij}(s, x, \delta_x) \ge 0$ and $a_{ij}(s, x, \mu)$ depends only on x_i and x_j.

(2) For any $s \ge 0$, $1 \le i \le d$ and $f \in \mathcal{U}_b$ independent on x_i,

$$\mu(b_i(s, \cdot, \mu)f) \ge \mu(f)\mu(b_i(s, \cdot, \mu)).$$

When σ and b do not depend on the distribution, by taking μ and ν being Dirac measures we see that conditions in Theorems 2.4 and 2.5 coincide each other so that they become sufficient and necessary for the positive correlations.

3 Time-inhomogeneous diffusion processes

Consider the time-dependent second order diffusion operators: for $t \ge 0$ and $x \in \mathbb{R}^d$,

$$L_t := \frac{1}{2}\text{tr}\{a(t, x)\nabla^2\} + b(t, x) \cdot \nabla,\ \ \bar{L}_t := \frac{1}{2}\text{tr}\{\bar{a}(t, x)\nabla^2\} + \bar{b}(t, x) \cdot \nabla. \tag{3.1}$$

where $a = (a_{ij}), \bar{a} = (\bar{a}_{ij}), b = (b_i), \bar{b} = (\bar{b}_i)$ are continuous in $[0, \infty) \times \mathbb{R}^d$. Assume that the martingale problems associated with $(L_t)_{t\ge 0}$ and $(\bar{L}_t)_{t\ge 0}$ are well-posed so that there exist unique time-inhomogeneous diffusion processes $(X_{s,t})_{t\ge s\ge 0}$ and $(\bar{X}_{s,t})_{t\ge s\ge 0}$ corresponding to $(L_t)_{t\ge 0}$ and $(\bar{L}_t)_{t\ge 0}$, respectively. Let $(P_{s,t})_{t\ge s\ge 0}$ and $(\bar{P}_{s,t})_{t\ge s\ge 0}$ be the Markov semigroups generated by $(X^x_{s,t})_{\{ij\}} \times \mu_{\{ij\}^c}$ and $(\bar{X}^x_{s,t})_{\{ij\}} \times \mu_{\{ij\}^c}$ with the initial value $X_{s,s} = \bar{X}_{s,s} = x$, respectively, i.e.,

$$P_{s,t}f(x) = \mathbb{E}f(X^x_{s,t}),\ \bar{P}_{s,t}f(x) = \mathbb{E}f(\bar{X}^x_{s,t}),\quad f \in \mathcal{B}_b(\mathbb{R}^d). \tag{3.2}$$

It is well known that for any $f \in C^\infty_0(\mathbb{R}^d)$

$$\frac{\text{d}}{\text{d}s}P_{s,t}f(x) = -P_{s,t}L_sf,\qquad \frac{\text{d}}{\text{d}t}P_{s,t}f(x) = L_tP_{s,t}f,\quad t \ge s \ge 0. \tag{3.3}$$

For any $x, y \in \mathbb{R}^d$ with $x \le y$, $f \in \mathcal{U}_b$ and $t \ge s \ge 0$, if $\bar{P}_{s,t}f(x) \le P_{s,t}f(y)$, we call $P_{s,t}$ preserving order, written as $\bar{P}^*_{s,t} \preceq P^*_{s,t}$, where for any $\mu \in \mathcal{P}$, $P^*_{s,t}\mu, \bar{P}^*_{s,t}\mu \in \mathcal{P}$ is given by

$$(P^*_{s,t}\mu)(f) := \mu(P_{s,t}f), \quad (\bar{P}^*_{s,t}\mu)(f) := \mu(\bar{P}_{s,t}f), \quad f \in \mathcal{B}_b(\mathbb{R}^d).$$

Moreover, we denote $P^*_{s,t} \in \mathcal{P}_+$ if $P^*_{s,t}\mathcal{P}_+ \subset \mathcal{P}_+$.

For any $\mu \in \mathcal{P}$, let $\Lambda_s\mu$ and $\bar{\Lambda}_s\mu$ be the distributions of the processes starting at μ from time s generated by L and $\bar{L}$ respectively. By the standard Markov property we see that $\bar{P}^*_{s,t} \preceq P^*_{s,t}$ for $t \ge s \ge 0$ if and only if $\bar{\Lambda}_s \preceq \Lambda_s$ for $s \ge 0$, while $P^*_{s,t} \in \mathcal{P}_+$ for $t \ge s$ is equivalent to $\Lambda_s \in \mathcal{P}^s_+$.

Theorem 3.1. $\bar{P}^*_{s,t} \preceq P^*_{s,t}$ for $t \ge s \ge 0$, equivalently $\bar{\Lambda}_s \preceq \Lambda_s$ for $s \ge 0$, if and only if the following conditions hold:

(**1**) For any $s \ge 0$ and $1 \le i \le d$, $\bar{b}_i(s,x) \le b_i(s,y)$ for $x \le y$ and $x_i = y_i$.
(**2**) For any $s \ge 0$ and $1 \le i, j \le d$, $\bar{a}_{ij} = a_{ij}$ and $a_{ij}(s,x)$ only depends on x_i and x_j.

Theorem 3.2. $P^*_{s,t} \in \mathcal{P}_+$ for $t \ge s$, equivalently $\Lambda_s \in \mathcal{P}^s_+$ for $s \ge 0$, if and only if the following conditions hold:

(1) For any $s \ge 0$ and $1 \le i \le d$, $b_i(s,x) \le b_i(s,y)$ for $x \le y$ and $x_i = y_i$;
(2) For any $s \ge 0$ and $1 \le i \le d$, $a_{ij} \ge 0$ and $a_{ij}(s,x) \ge 0$ depends only on x_i and x_j.

Proof of Theorem 3.1. (a) We first prove the necessity. For any $t \ge s \ge 0$ and $x \in \mathbb{R}^d$, let Λ^x_s (resp. $\bar{\Lambda}^x_s$) be the distribution of the L_t-diffusion (resp. $\bar{L}_t$-diffusion) process on the path space $C_s := C([s,\infty);\mathbb{R}^d)$ starting from x at time s.

For $x \in \mathbb{R}^d$ and $0 \le s_0 \le s_1 < s_2 < \cdots < s_n$, let $\Lambda^x_{s_0,s_1,\cdots,s_n}$ be the marginal distribution of $\Lambda^x_{s_0}$ at the time sequence $(s_1,\cdots,s_n)$, which can be expressed via the Markov property as below

$$\Lambda^x_{s_0,s_1,\cdots,s_n}(\mathrm{d}y_1,\mathrm{d}y_2,\cdots,\mathrm{d}y_n) = P_{s_0,s_1}(x,\mathrm{d}y_1)P_{s_1,s_2}(y_1,\mathrm{d}y_2)\cdots P_{s_{n-1},s_n}(y_{n-1},\mathrm{d}y_n).$$

Then, by an inductive argument, together with the Markov property of the associated Markov process, $\bar{P}^*_{s,t} \preceq P^*_{s,t}$ implies $\bar{\Lambda}^x_s \le \Lambda^y_s$ (i.e., $\bar{\Lambda}^x_s(f) \le \Lambda^y_s(f)$ for any $f \in \mathcal{U}_b \cap C_s$). Therefore, there exists a coupling $\mathbb{P}^{x,y}_s \in \mathcal{C}(\bar{\Lambda}^x_s, \Lambda^y_s)$ such that

$$\mathbb{P}^{x,y}_s\big((\xi,\eta) \in C_s \times C_s : \eta \preceq \xi\big) = 1. \tag{3.4}$$

Let $(\Omega,\mathcal{F},\mathbb{P}) = (C_s \times C_s, \mathcal{B}(C_s \times C_s), \mathbb{P}^{x,y}_s)$ with the natural filtration $(\mathcal{F}_t)_{t\ge s}$ induced by the coordinate process $(\xi_t,\eta_t)_{t\ge s}$ solving

$$\begin{cases} \mathrm{d}\xi_t = b(t,\xi_t)\mathrm{d}t + \sigma(t,\xi_t)\mathrm{d}B^1_t, & \xi_s = y \\ d\eta_t = \bar{b}(t,\eta_t)\mathrm{d}t + \bar{\sigma}(t,\eta_t)\mathrm{d}B^2_t, & \eta_s = x \end{cases} \tag{3.5}$$

for some d-dimensional Brownian motions $(B^1_t)_{t\ge s}$ and $(B^2_t)_{t\ge s}$, and some measurable mappings $\sigma, \bar{\sigma} : [s,\infty) \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ with $a = \sigma\sigma^*$, $\bar{a} = \bar{\sigma}\,\bar{\sigma}^*$. Then, from (3.4), we have $\xi_t \ge \eta_t$, $\mathbb{P}^{x,y}_s$-a.s., for all $t \ge s$.

Let $x \le y$ with $x_i = y_i$. Since $\xi_t \ge \eta_t$, $\mathbb{P}^{x,y}_s$-a.s., and $(\xi_s)_i = (\eta_s)_i$ due to $x_i = y_i$, we derive from (3.5) that

$$\int_s^t \big(b_i(r,\xi_r) - \bar{b}_i(r,\eta_r)\big)\mathrm{d}r \ge \int_s^t \big\langle \bar{\sigma}_{i\cdot}(r,\xi_r), \mathrm{d}B^2_r\big\rangle - \int_s^t \big\langle \sigma_{i\cdot}(r,\xi_r), \mathrm{d}B^1_r\big\rangle,$$

where $\sigma_{i\cdot}$ means the i-th row of σ. Taking conditional expectation $\mathbb{P}^{x,y}_{s_0}(\cdot|\mathcal{F}_{s_0})$ on both sides yields

$$\int_s^t \mathbb{E}\big(\big(b_i(r,\xi_r) - \bar{b}_i(r,\eta_r)\big)|\mathcal{F}_s\big)\mathrm{d}r \ge 0, \quad t \ge s.$$

This implies the assertion (**1**) by taking the continuity of $b_i,\ \bar{b}_i$ and $(\xi_\cdot,\ \eta_\cdot)$ into account.

Let $x \le y$ with $(x_i, x_j) = (y_i, y_j)$. Then, by using $\xi_t \ge \eta_t, \mathbb{P}^{x,y}_{s_0}$-a.s., again, we have

$$\int_{s_0}^t b_k(s,\xi_s)\mathrm{d}s + \int_{s_0}^t \left\langle \sigma_{k\cdot}(s,\xi_s), \mathrm{d}B^1_s\right\rangle \ge \int_{s_0}^t \bar{b}_k(s,\eta_s)\mathrm{d}s + \int_{s_0}^t \left\langle \bar{\sigma}_{k\cdot}(s,\eta_s), \mathrm{d}B^2_s\right\rangle,\ k = i,j. \quad (3.6)$$

Note that as $t \downarrow s_0$,

$$\frac{1}{\sqrt{t-s_0}}\Big(\int_{s_0}^t \left\langle \sigma_{i\cdot}(s,\xi_s), \mathrm{d}B^1_s\right\rangle, \int_{s_0}^t \left\langle \sigma_{j\cdot}(s,\xi_s), \mathrm{d}B^1_s\right\rangle\Big)\overrightarrow{weakly}\ N\left(0, \begin{pmatrix} a_{ii}(s,y) & a_{ij}(s,y) \\ a_{ji}(s,y) & a_{jj}(s,y)\end{pmatrix}\right) =: \mu,$$

$$\frac{1}{\sqrt{t-s_0}}\Big(\int_{s_0}^t \left\langle \bar{\sigma}_{i\cdot}(s,\xi_s), \mathrm{d}B^2_s\right\rangle, \int_{s_0}^t \left\langle \bar{\sigma}_{j\cdot}(s,\xi_s), \mathrm{d}B^1_s\right\rangle\Big)\overrightarrow{weakly}\ N\left(0, \begin{pmatrix} \bar{a}_{ii}(s,y) & \bar{a}_{ij}(s,y) \\ \bar{a}_{ji}(s,y) & \bar{a}_{jj}(s,y)\end{pmatrix}\right) =: \bar{\mu}.$$

Then (3.6) implies $\bar{\mu} \preceq \mu$. Similarly, $\mu \preceq \bar{\mu}$. Therefore, we have $\mu = \bar{\mu}$ so that $a = \bar{a}$. For the assertion that a_{ij} depends only on x_i and x_j of (**2**), it can be proven by following exactly the arguments of [3, Lemmas 2.1 & Lemma 2.3].

(b) Following exactly the arguments of [3, Lemma 2.4, 2.5 & Theorem 1.3] with replacing time homogeneous semi-group P_t by time inhomogeneous semi-group $P_{s,t}$, we prove the sufficiency by the following Theorem 3.3 on the monotonicity. □

Theorem 3.3. $P^*_{s,t}$ is monotone, i.e., $P^*_{s,t}\mu \preceq P^*_{s,t}\nu$ with $\mu \preceq \nu$ for $t \ge s \ge 0$, provided the following two conditions hold:

($\mathbf{1}'$) $b_i(s,\cdot)$ is smooth, $b_i(s,x) \le b_i(s,y)$ with $x \le y$ and $x_i = y_i$;
($\mathbf{2}'$) $a_{ij}(s,\cdot)$ is smooth, $a_{ij}(s,x)$ depends only on x_i and x_j.

Proof. As in [3], by an approximation argument we assume that $(a,b) \in C^\infty_b([0,T]\times\mathbb{R}^d)$ for any $T > 0$. To prove $P_{s,t}f \in \mathcal{U}_b$ for $t \ge s$ and $f \in \mathcal{U}_b$, it suffices to show

$$\nabla P_{s,t}f(x) \ge 0, \quad t \ge s,\ f \in \mathcal{U}_b \cap C^\infty_b(\mathbb{R}^d)$$

since $\mathcal{U}_b \cap C^\infty_b(\mathbb{R}^d)$ is dense in $\mathcal{U}_b$. Below, we assume $f \in \mathcal{U}_b \cap C^\infty_b(\mathbb{R}^d)$. Let $u_{s,t} = P_{s,t}f,\ t \ge s$. Then by (3.3), we have

$$\partial_t u_{s,t} = L_t u_{s,t}, \quad t \ge s, \qquad u_{s,s} = f.$$

Taking the partial derivative w.r.t. the k-th component (i.e., ∂_k) on both sides yields

$$\partial_t(\partial_k u_{s,t}) = \partial_k\partial_t u_{s,t} = L^k_t(\partial_k u_{s,t}) + \sum_{j=1}^d \alpha_{kj}(t,\cdot)\partial_j u_{s,t}, \quad (3.7)$$

where $L^k_t := L_t + \sum_{j=1}^d \left[(1-\frac{\delta_{jk}}{2})\partial_k a_{jk}(t,\cdot)\right]\partial_j + \partial_k b_k(t,\cdot),\ \alpha_{kj}(t,\cdot) := (\partial_k b_j(t,\cdot))I_{\{k\neq j\}}$. Since L^k_t is a time-inhomogeneous Schrödinger operator, it generates a positivity-preserving semigroup $\left(T^k_{s,t}\right)_{t\ge s}$. So, the operator $\tilde{L}_t := \left(L^k_t\right)_{1\le k\le d}$ defined on $C^2(\mathbb{R}^d;\mathbb{R}^d)$ by $\tilde{L}_t V := \left(L^k_t V_k\right)_{1\le k\le d}$ generates a positivity preserving semigroup $T_{s,t} := \left(T^k_{s,t}\right)_{1\le k\le d}$, $t \ge s$. Let $D_r = \left(\alpha_{kj}(r,\cdot)_{1\le k,j\le d}\right)$ and $V_{s,t} = \nabla P_{s,t}f = \nabla u_{s,t}$. Then (3.7) implies

$$\partial_t V_{s,t} = \tilde{L}_t V_{s,t} + D_t V_{s,t}, \quad t \ge s, \quad V_{s,s} = \nabla f.$$

This, together with Duhamel's formula, gives

$$V_{s,t} = T_{s,t}V_{s,s} + \int_s^t T_{r,t}D_r V_{s,r}\mathrm{d}r, \quad t \ge s.$$

Since $V_{s,s} = \nabla f \ge 0$ and $T_{s,t},\ D_r$ are positivity preserving, this implies $V_{s,t} = \nabla P_{s,t}f \ge 0$. □

Proof of Theorem 3.2. Theorem 3.2 can be proved using the same arguments in [5, Proposition 4.1] by combining Theorem 3.1 and 3.3. So, we omit the details to save space. □

4 Proofs of Theorems 2.1-2.5

Proof of Theorem 2.1. Since $\bar{P}^*_{s,t}\nu$ and $P^*_{s,t}\mu$ are marginal distributions at time t of $\bar{\Lambda}_s\nu$ and $\Lambda_s\mu$, $\bar{P}^*_{s,t} \preceq P^*_{s,t}$ for $t \geq s$ follows from $\bar{\Lambda}_s \preceq \Lambda_s$. Therefore, to obtain the desired assertion, it is sufficient to show $\bar{\Lambda}_s\nu \preceq \Lambda_s\mu$. Below, we set $\mu, \nu \in \mathcal{P}_2$ with $\nu \preceq \mu$. For any $T > s$, set

$$\mathcal{P}^{\nu,\mu}_{s,T} := \big\{(\mu^{(1)}, \mu^{(2)}) \in C([s,T]; \mathcal{P}_2 \times \mathcal{P}_2) : \mu^{(1)}_t \preceq \mu^{(2)}_t,\ t \in [s,T],\ \mu^{(1)}_s = \nu, \mu^{(2)}_s = \mu\big\},$$

which is a complete metric space under the metric for $\lambda > 0$,

$$\rho_\lambda\big((\mu^{(1)}, \mu^{(2)}), (\tilde{\mu}^{(1)}, \tilde{\mu}^{(2)})\big) := \sup_{t\in[s,T]} \mathrm{e}^{-\lambda t}\big\{\mathbb{W}_2(\mu^{(1)}_t, \tilde{\mu}^{(1)}_t) + \mathbb{W}_2(\mu^{(2)}_t, \tilde{\mu}^{(2)}_t)\big\}.$$

For any $(\mu^{(1)}, \mu^{(2)}) \in \mathcal{P}^{\nu,\mu}_{s,T}$, consider the following time-dependent SDEs:

$$\begin{cases} \mathrm{d}X^{(1),\mu^{(1)}}_t = \bar{b}(t, X^{(1),\mu^{(1)}}_t, \mu^{(1)}_t)\mathrm{d}t + \bar{\sigma}(t, X^{(1),\mu^{(1)}}_t, \mu^{(1)}_t)\mathrm{d}W_t \quad t \geq s, \quad X^{(1),\mu^{(1)}}_s = \xi \sim \nu, \\ \mathrm{d}X^{(2),\mu^{(2)}}_t = b(t, X^{(2),\mu^{(2)}}_t, \mu^{(2)}_t)\mathrm{d}t + \sigma(t, X^{(2),\mu^{(2)}}_t, \mu^{(2)}_t)\mathrm{d}W_t \quad t \geq s, \quad X^{(2),\mu^{(2)}}_s = \eta \sim \mu, \end{cases} \tag{4.1}$$

where $\sigma = \sqrt{2a}$ and $\bar{\sigma} = \sqrt{2\bar{a}}$, and $\xi \sim \nu$ means $\mathcal{L}_\xi = \nu$. Define the mapping on $\mathcal{P}^{\nu,\mu}_{s,T}$ by

$$H\big((\mu^{(1)}, \mu^{(2)})\big)(t) = \Big(\mathcal{L}_{X^{(1),\mu^{(1)}}_t}, \mathcal{L}_{X^{(2),\mu^{(2)}}_t}\Big), \quad t \geq s. \tag{4.2}$$

Since $\mu^{(1)}_t \preceq \mu^{(2)}_t$, by Theorem 3.1 the conditions in Theorem 2.1 imply $\mathcal{L}_{X^{(1),\mu^{(1)}}_{[s,T]}} \preceq \mathcal{L}_{X^{(2),\mu^{(2)}}_{[s,T]}}$, so that $H : \mathcal{P}^{\nu,\mu}_{s,T} \to \mathcal{P}^{\nu,\mu}_{s,T}$. By Itô's formula and the assumption $(\mathbf{A})$, it is easy to see that H is contractive under the metric ρ_λ for large enough $\lambda > 0$, so that it has a unique fixed point and hence the proof is finished. □

Proof of Theorem 2.2. Let $s \geq 0$ and $\nu \preceq \mu$ with $\nu_i = \mu_i$. By $\bar{\Lambda}_s \preceq \Lambda_s$, we have $\bar{\Lambda}_s\nu \preceq \Lambda_s\mu$. According to [8, Theorem 5], there exists $\mathbb{P}_s \in \mathcal{C}(\bar{\Lambda}_s\nu, \Lambda_s\mu)$ such that

$$\mathbb{P}_s\big(\{(\xi,\eta) \in C_s \times C_s : \xi_t \geq \eta_t, t \geq s\}\big) = 1. \tag{4.3}$$

Since $\bar{\Lambda}_s\nu$ and $\Lambda_s\mu$ are solutions to the martingale problems associated with the operators $\bar{L}$ and L in (1.1), respectively, according to the superposition principle (see [12]), we have $\mathcal{L}_{(\xi,\eta)} = \mathbb{P}_s$, where (ξ_t, η_t) solves

$$\begin{cases} \mathrm{d}\xi_t = \bar{b}(t, \eta_t, \mathcal{L}_{\eta_t})\mathrm{d}t + \bar{\sigma}(t, \eta_t, \mathcal{L}_{\eta_t})\mathrm{d}B^1_t, \ \ t \geq s, \\ \mathrm{d}\eta_t = b(t, \xi_t, \mathcal{L}_{\xi_t})\mathrm{d}t + \sigma(t, \xi_t, \mathcal{L}_{\xi_t})\mathrm{d}B^2_t, \ \ t \geq s, \end{cases} \tag{4.4}$$

for some $2d$-dimensional Brownian motions $(B^1_t, B^2_t)_{t\geq s}$ on the probability space $(C_s \times C_s, \mathcal{B}(C_s \times C_s), \{\mathcal{F}_t\}_{t\geq s}, \mathbb{P}_s)$, where $\{\mathcal{F}_t\}_{t\geq s}$ is induced by $(\xi_t, \eta_t)_{t\geq s}$.

Since $\mathcal{L}_{(\xi,\eta)} = \mathbb{P}_s$ satisfying (4.3), we have $\xi_t \geq \eta_t$ for all $t \geq s$. Moreover, note that $\mathcal{L}_{(\xi_s,\eta_s)} \in \mathcal{C}(\nu,\mu)$ and $\nu_i = \mu_i$ imply $\xi^i_s = \eta^i_s$. Thus, we find $\mathbb{P}_s$-a.s.

$$\int_s^t \bar{b}_i(r, \xi_r, \mu^{(1)}_r)\mathrm{d}r + \int_s^t \bar{\sigma}_{i\cdot}(r, \xi_r, \mu^{(1)}_r)\mathrm{d}B^1_t \leq \int_s^t b_i(r, \eta_r, \mu^{(2)}_r) + \int_s^t \sigma_{i\cdot}(r, \eta_r, \mu^{(2)}_r)\mathrm{d}B^1_t,\ t \geq s. \tag{4.5}$$

Taking conditional expectation with respect to $\mathcal{F}_s$, we drive

$$\int_s^t \mathbb{E}\big(\bar{b}_i(r, \xi_r, \mu^{(1)}_r)|\mathcal{F}_s\big)\mathrm{d}r \leq \int_s^t \mathbb{E}\big(b_i(r, \eta_r, \mu^{(2)}_r)|\mathcal{F}_s\big)\mathrm{d}r, \quad t \geq s.$$

By the continuity of $\bar{b}$ and b and $\mu_r^{(1)} \to \nu, \mu_r^{(2)} \to \mu$ weakly as $r \downarrow s$, we obtain

$$\bar{b}_i(s, \xi_s, \nu) \le \bar{b}_i(s, \eta_s, \nu), \quad t \ge s, \ \mathbb{P}_s - a.s.$$

Consequently, letting $\pi := \mathcal{L}_{(\xi_s, \eta_s)} \in \mathcal{C}(\nu, \mu)$, we have $\pi\big(\{x \le y\}\big) = 1$ and

$$\bar{b}_i(s, x, \nu) \le b_i(s, y, \mu), \ (x, y) \in \mathrm{supp}\pi.$$

Thus, the first assertion of (i) holds true. Hence, for $\nu = \mu$, $\pi(\{x \le y\}) = 1$ implies $x = y, \pi - a.s.$ Whence, we have

$$\bar{b}_i(s, x, \mu) \le b_i(s, x, \mu), \ x \in \mathrm{supp}\mu.$$

In general, for any $x \in \mathbb{R}^d$, let $\mu_\varepsilon = (1-\varepsilon)\mu + \varepsilon\delta_x$. It is easy to see that $x \in \mathrm{supp}\mu_\varepsilon$. Thus applying (4) with μ_ε replaced by μ yields

$$\bar{b}_i(s, x, \mu_\varepsilon) \le b_i(s, x, \mu_\varepsilon), \ s \ge 0, \ \varepsilon > 0.$$

Consequently, the second assertion in (i) follows by taking $\varepsilon \downarrow 0$.

Below we assume $\nu \preceq \mu$ with $\nu_{ij} = \mu_{ij}$ so that $\nu_i = \mu_i, \nu_j = \mu_j$. Thus, we deduce from (4.5) that for any $\varepsilon \in [0, 1]$,

$$\begin{aligned}
&\int_s^t \big[\varepsilon\bar{b}_i(r, \xi_r, \mu_r^{(1)}) + (1-\varepsilon)\bar{b}_j(r, \xi_r, \mu_r^{(1)})\big]\mathrm{d}r + \int_s^t \big[\varepsilon\bar{\sigma}_{i\cdot}(r, \xi_r, \mu_r^{(1)}) + (1-\varepsilon)\bar{\sigma}_{j\cdot}(r, \xi_r, \mu_r^{(1)})\big]\mathrm{d}B_r^1 \\
&\quad \le \int_s^t \big[\varepsilon\bar{b}_i(r, \eta_r, \mu_r^{(2)}) + (1-\varepsilon)\bar{b}_j(r, \eta_r, \mu_r^{(2)})\big]\mathrm{d}r \\
&\qquad + \int_s^t \big[\varepsilon\bar{\sigma}_{i\cdot}(r, \eta_r, \mu_r^{(2)}) + (1-\varepsilon)\bar{\sigma}_{j\cdot}(r, \eta_r, \mu_r^{(2)})\big]\mathrm{d}B_r^2
\end{aligned}$$

Dividing both side by $\frac{1}{\sqrt{t-s}}$ and letting $t \downarrow s$, we find

$$\begin{aligned}
&N\left(0, \varepsilon^2\bar{a}_{ii}(s, \xi_s, \nu) + 2\varepsilon(1-\varepsilon)\bar{a}_{ij}(s, \xi_s, \nu) + (1-\varepsilon)^2\bar{a}_{jj}(s, \xi_s, \nu)\right) \\
&\le N\left(0, \varepsilon^2 a_{ii}(s, \eta_s, \mu) + 2\varepsilon(1-\varepsilon)a_{ij}(s, \eta_s, \mu) + (1-\varepsilon)^2 a_{jj}(s, \eta_s, \mu)\right).
\end{aligned}$$

By the symmetry of centred normal distribution, this further implies

$$\begin{aligned}
&\varepsilon^2\bar{a}_{ii}(s, \xi_s, \nu) + 2\varepsilon(1-\varepsilon)\bar{a}_{ij}(s, \xi_s, \nu) + (1-\varepsilon)^2\bar{a}_{jj}(s, \xi_s, \nu) \\
&= \varepsilon^2 a_{ii}(s, \eta_s, \mu) + 2\varepsilon(1-\varepsilon)a_{ij}(s, \eta_s, \mu) + (1-\varepsilon)^2 a_{jj}(s, \eta_s, \mu), \varepsilon \in [0, 1].
\end{aligned}$$

Consequently, dividing by ε^2 on both sides yields

$$\bar{a}_{ij}(s, \xi_s, \nu) = a_{ij}(s, \eta_s, \mu), \ \mathbb{P}_s - a.s., \tag{4.6}$$

which gives for $\pi = \mathcal{L}_{(\xi_s, \eta_s)} \in \mathcal{C}(\nu, \mu)$,

$$\bar{a}_{ij}(s, x, \nu) = a_{ij}(s, y, \mu), \quad (x, y) \in \mathrm{supp}\pi, \ s \ge 0.$$

Thus, by the approximation trick above, we can obtain the second assertion in (ii). □

Proof of Theorem 2.3. Due to $\bar{P}_{s,t}^* \preceq P_{s,t}^*$, we have $\mathcal{L}_{\bar{X}_{s,t}} = \bar{P}_{s,t}^*\nu \preceq P_{s,t}^*\mu = \mathcal{L}_{X_{s,t}}$ for $\nu \preceq \mu$. Therefore, in particular for $f(x) = x_i \in \mathcal{U}$, we obtain

$$\mathbb{E}(\bar{X}_{s,t})_i \le \mathbb{E}(X_{s,t})_i.$$

Since $\nu_i(f) = \mu_i(f)$, we then deduce from (4.4) with ξ_t and η_t replaced by $X_{s,t}$ and $\bar{X}_{s,t}$ in (4.1) that

$$\int_s^t \mathbb{E}(\bar{b}_i(s, \bar{X}_{s,r}, \mathcal{L}_{\bar{X}_{s,r}}))\mathrm{d}r \le \int_s^t \mathbb{E}(b_i(s, X_{s,r}, \mathcal{L}_{X_{s,r}}))\mathrm{d}r.$$

Dividing by $t-s$ on both side followed by $t \downarrow s$, we get (i).

Since $\bar{P}^*_{s,t}\mu \le P^*_{s,t}\mu$, for $f \in \mathcal{U}_b \cap C_b^\infty(\mathbb{R}^d)$, we have $\mu(\bar{L}_{s,\mu}f) \le \mu(L_{s,\mu}f)$. In particular, taking $\mu = \delta_x$ yields $\bar{L}_{s,\delta_x}f(x) \le L_{s,\delta_x}f(x)$. With this at hand, we can get the assertion (ii) by following exactly the argument of [3, Lemma 3.4]. □

To prove Theorems 2.4 and 2.5, we first present some lemmas.

Lemma 4.1. Let $\mu = \frac{1}{2}(\mu^{(1)} + \mu^{(2)})$, where $\mu^{(1)}, \mu^{(2)} \in \mathcal{P}_+$ such that $\mu^{(1)} \preceq \mu^{(2)}$. Then, $\mu \in \mathcal{P}_+$.

Proof. Let $f, g \in \mathcal{U}_b$. By $\mu^{(1)} \preceq \mu^{(2)}$, we have $\big(\mu^{(1)}(f) - \mu^{(2)}(f)\big)\big(\mu^{(1)}(g) - \mu^{(2)}(g)\big) \ge 0$. That is,

$$2\big(\mu^{(1)}(f)\mu^{(1)}(g) + \mu^{(2)}(f)\mu^{(2)}(g)\big) \ge \big(\mu^{(1)}(f) + \mu^{(2)}(f)\big)\big(\mu^{(1)}(g) + \mu^{(2)}(g)\big).$$

Combining this with $\mu^{(i)}(fg) \ge \mu^{(i)}(f)\mu^{(i)}(g)$ due to $\mu^{(i)} \in \mathcal{P}_+$ for $i = 1, 2$, we obtain

$$\mu(fg) = \frac{1}{2}\big(\mu^{(1)}(fg) + \mu^{(2)}(fg)\big) \ge \frac{1}{4}\big(\mu^{(1)}(f) + \mu^{(2)}(f)\big)\big(\mu^{(1)}(g) + \mu^{(2)}(g)\big) = \mu(f)\mu(g). \quad \square$$

Lemma 4.2. Suppose that $(P^*_{s,t})_{t\ge s}$ preserves positive correlations and let μ be in Lemma 4.1. Then, for $f, g \in \mathcal{U}_b \cap C_b^\infty(\mathbb{R}^d)$ with $\mu(fg) = \mu(f)\mu(g)$, we have

$$\begin{aligned} 2\big(\mu^{(1)}(L_{s,\mu}(fg)) + \mu^{(2)}(L_{s,\mu}(fg))\big) &\ge \big(\mu^{(1)}(L_{s,\mu}(f)) + \mu^{(2)}(L_{s,\mu}(f))\big)\big(\mu^{(1)}(g) + \mu^{(2)}(g)\big) \\ &\quad + \big(\mu^{(1)}(L_{s,\mu}(g)) + \mu^{(2)}(L_{s,\mu}(g))\big)\big(\mu^{(1)}(f) + \mu^{(2)}(f)\big). \end{aligned} \tag{4.7}$$

$$\begin{aligned} 2\big(\mu^{(1)}(\Gamma_1(f,g)) + \mu^{(2)}(\Gamma_1(f,g))\big) &\ge \big(\mu^{(2)}(L_{s,\mu}(g)) - \mu^{(1)}(L_{s,\mu}(g))\big)\big(\mu^{(1)}(f) - \mu^{(2)}(f)\big) \\ &\quad + \big(\mu^{(2)}(L_{s,\mu}(f)) + \mu^{(1)}(L_{s,\mu}(f))\big)\big(\mu^{(1)}(g) - \mu^{(2)}(g)\big), \end{aligned} \tag{4.8}$$

where $\Gamma_1(f,g) := L_{s,\mu}(fg) - fL_{s,\mu}g - gL_{s,\mu}f = \langle a(t,\cdot,\mu)\nabla f, \nabla g\rangle$.

Proof. By a direct calculation, we see that (4.7) is equivalent to (4.8). So, it suffices to prove (4.7). From Lemma 4.1, we have $\mu \in \mathcal{P}_+$. Since $(P^*_{s,t})_{t\ge s}$ preserves positive correlations, we have

$$(P^*_{s,t}\mu)(fg) \ge (P^*_{s,t}\mu)(f)(P^*_{s,t}\mu)(g), \quad t \ge s,\ f, g \in \mathcal{U}_b \cap C_b^\infty(\mathbb{R}^d).$$

This, together with $\mu = \frac{1}{2}(\mu^{(1)} + \mu^{(2)})$, yields

$$2\big((P^*_{s,t}\mu^{(1)})(fg) + (P^*_{s,t}\mu^{(2)})(fg)\big) \ge \big((P^*_{s,t}\mu^{(1)})(f) + (P^*_{s,t}\mu^{(2)})(f)\big)\big((P^*_{s,t}\mu^{(1)})(g) + (P^*_{s,t}\mu^{(2)})(g)\big).$$

Combining this with $\mu(fg) = \mu(f)\mu(g)$ we derive

$$\begin{aligned} &\frac{2}{t-s}\big((P^*_{s,t}\mu^{(1)})(fg) + (P^*_{s,t}\mu^{(2)})(fg) - (\mu^{(1)})(fg) + (\mu^{(2)})(fg)\big) \\ &\ge \frac{1}{t-s}\big\{\big((P^*_{s,t}\mu^{(1)})(f) + (P^*_{s,t}\mu^{(2)})(f)\big)\big((P^*_{s,t}\mu^{(1)})(g) + (P^*_{s,t}\mu^{(2)})(g)\big) \\ &\quad - \big(\mu^{(1)}(f) + \mu^{(2)}(f)\big)\big(\mu^{(1)}(g) + \mu^{(2)}(g)\big)\big\}. \end{aligned}$$

Consequently, the assertion (4.7) follows by taking $t \downarrow s$. □

Lemma 4.3. If $(P^*_{s,t})_{t\ge s}$ preserves positive correlation, then $a_{ij}(s,x,\delta_x) \ge 0$ for any $s \ge 0, x \in \mathbb{R}^d$.

Proof. By Lemma 4.2 with $\mu^{(1)} = \mu^{(2)} = \delta_x$, for any $f, g \in \mathcal{U}_b \cap C_b^\infty(\mathbb{R}^d)$,

$$L_{s,\delta_x}(fg)(x) \geq \{fL_{s,\delta_x}g + gL_{s,\delta_x}f\}(x), \quad s \geq 0.$$

By choosing $f, g \in \mathcal{U}_b \cap C_b^\infty(\mathbb{R}^d)$ such that $f(z) = z_i - x_i$ and $g(z) = z_j - x_j$ holds in a neighbourhood of x, we obtain $a_{ij}(t, x, \delta_x u) \geq 0$. □

Lemma 4.4. Let $\mu^{(1)} := \mu_i^{(1)} \times \mu_{\{i\}^c}^{(1)}$ and $\mu^{(2)} := \mu_i^{(2)} \times \mu_{\{i\}^c}^{(2)}$ with $\mu_i^{(1)} = \mu_i^{(2)}$, where $\mu_i^{(1)}$, $\mu_i^{(2)}$, $\mu_{\{i\}^c}^{(1)}$, $\mu_{\{i\}^c}^{(2)} \in \mathcal{P}_+$. If $(P_{s,t}^*)_{t\geq s}$ preserves positive correlation, then $\mu^{(1)}(b_i(t,\cdot,\mu^{(1)})) \leq \mu^{(2)}(b_i(t,\cdot,\mu^{(2)}))$.

Proof. Since $\mu_i^{(1)}$, $\mu_i^{(2)}$, $\mu_{\{i\}^c}^{(1)}$, $\mu_{\{i\}^c}^{(2)} \in \mathcal{P}_+$, we deduce $\mu^{(1)}, \mu^{(2)} \in \mathcal{P}_+$. For given i and $k \neq i$, take $f, g \in \mathcal{U}_b \cap C_b^\infty(\mathbb{R}^d)$ such that in a neighbourhood of x,

$$f(z) = z_i - \int_{\mathbb{R}} r\mu_i^{(1)}(\mathrm{d}r), \quad g(z) = \frac{\frac{h(z_k)}{1+h(z_k)} - \mu^{(1)}(\frac{h}{1+h})}{\mu^{(2)}(\frac{h}{1+h}) - \mu^{(1)}(\frac{h}{1+h})},$$

where $h \in C^\infty(\mathbb{R}; \mathbb{R}_+)$ is an increasing function. Since $\mu(fg) = \mu(f) = 0$ and $\mu = (\mu^{(1)} + \mu^{(2)})/2$, Lemma 4.2 with the above defined f and g implies $\mu^{(1)}\big(b_i(t,\cdot,\mu^{(1)})\big) \leq \mu^{(2)}\big(b_i(t,\cdot,\mu^{(2)}\big)$. □

Proof of Theorem 2.4. Since $\bar{P}_{s,t}^*\mu$ is the marginal distributions of $\Lambda_s\mu$ at time t, $\Lambda_s \in \mathcal{P}_+$ implies $P_{s,t}^* \in \mathcal{P}_+$ for $t \geq s$. So, it suffices to prove $\Lambda_s \in \mathcal{P}_+^s$. We only need to prove that for any $\mu_0 \in \mathcal{P}_+$ and $T > s \geq 0$, the marginal distribution $\Lambda_{s,T}\mu_0$ of $\Lambda_s\mu_0$ on $C_{s,T} := C([s,T]; \mathbb{R}^d)$ satisfies

$$(\Lambda_{s,T}\mu_0)(FG) \geq (\Lambda_{s,T}\mu_0)(F)\Lambda_{s,T}(G) \tag{4.9}$$

for any bounded increasing functions F, G on $C_{s,T}$. To achieve this, let

$$\mathcal{D}_+ = \big\{\nu \in C([s,T]; \mathcal{P}_2(\mathbb{R}^d)) : \nu_s = \mu_0,\ \nu_t \in \mathcal{P}_+,\ t \in [s,T]\big\},$$

which is a Polish space under the metric for $\lambda > 0$:

$$\mathbb{W}_{2,\lambda}(\mu,\nu) := \sup_{t\in[s,T]} (\mathrm{e}^{-\lambda t}\mathbb{W}_2(\mu_t,\nu_t)). \tag{4.10}$$

For $\nu \in \mathcal{D}_+$ and $b_t^\nu(x) := b_t(x,\nu_t), \sigma_t^\nu(x) := \sqrt{2a_t(x,\nu_t)}$, consider the time-dependent SDE

$$\mathrm{d}X_t^\nu = b_t^\nu(X_t^\nu)\mathrm{d}t + \sigma_t^\nu(X_t^\nu)\mathrm{d}W_t, \quad t \in [s,T], \quad X_s^\nu = X_s \sim \mu_0, \tag{4.11}$$

By $\mu_0 \in \mathcal{P}_+$ and conditions (1)-(2), Theorem 3.2 implies

$$\mathcal{D}_+ \ni \nu \mapsto \Phi(\nu) \in \mathcal{D}_+; \quad (\Phi(\nu))_t := \mathcal{L}_{X_t^\nu}, \quad t \in [s,T].$$

By Itô's formula and (1.3), it is easy to see that Φ is contractive under the complete metric

$$\mathbb{W}_{2,\lambda}(\nu^1\nu^2) := \sup_{t\in[s,T]} \big(\mathrm{e}^{-\lambda t}\mathbb{W}_2\big(\nu_t^1,\nu_t^2\big)\big)$$

on $\mathcal{D}_+$, so that it has a unique fixed point $\nu \in \mathcal{D}_+$. Thus, $\Lambda_{s,T}\mu_0 := \mathcal{L}_{(X_t^\nu)_{t\in[s,T]}}$ with $\nu_t = \mathcal{L}_{X_t^\nu} \in \mathcal{P}_+,\ t \in [s,T]$. Therefore, by applying Theorem 3.2 to the diffusion process generated by L_t with coefficients (b^ν, a^ν), we conclude that the present conditions (1) and (2) imply (4.9) as desired.

□

Proof of Theorem 2.5. Consider the decoupled SDE

$$\mathrm{d}X_{s,t}^{x,\mu} = b(t, X_{s,t}^{x,\mu}, P_{s,t}^*\mu)\mathrm{d}t + \sigma(t, X_{s,t}^{x,\mu}, P_{s,t}^*\mu)\mathrm{d}W_t, \quad t \geq s, \;\; X_{s,s}^{x,\mu} = x, \tag{4.12}$$

where $P_{s,t}^*\mu$ is the marginal distribution of Λ_s at the time t. For $x \in \mathbb{R}^d$, let $\Lambda_s^{x,\mu} = \mathcal{L}_{X_{[s,\infty)}^{x,\mu}}$. Then for any $\nu \in \mathcal{P}$, $\Lambda_s^{\nu,\mu} := \int_{\mathbb{R}^d} \Lambda_s^{x,\mu}\nu(dx)$ is the law of $X_{[s,\infty)}^{\nu,\mu}$ with initial distribution ν. Noting that $\Lambda_s^{\mu} = \Lambda_s^{\mu,\mu} = \mathcal{L}_{X_{[s,\infty)}^{\mu}}$, by $\Lambda_s^\mu \in \mathcal{P}_+^s$ we obtain

$$\Lambda_s^\mu(FG) \geq \Lambda_s^\mu(F)\Lambda_s^\mu(G), \;\; F, G \in \mathcal{U}(C_s), \mu \in \mathcal{P}_+. \tag{4.13}$$

For $\gamma \in C_s$, let $F(\gamma) = f(\gamma_s)$ with $0 \leq f \in \mathcal{U}(\mathbb{R}^d)$. Then (4.13) becomes

$$\Lambda_s^{\nu,\mu}(G) \geq \Lambda_s^{\mu,\mu}(G), \; G \in \mathcal{U}(C_s),$$

where $\nu(dx) := \frac{f(x)\mu(dx)}{\mu(f)}$. That is, $\Lambda_s^{\nu,\mu} \geq \Lambda_s^{\mu,\mu}$. Then there exit $\pi_s \in \mathcal{C}(\Lambda_s^{\nu,\mu}, \Lambda_s^{\mu,\mu})$ and Brownian motions B_t^1 and B_t^2 on $(\Omega, \mathcal{F}_t, \mathbb{P}) := (C_s, \sigma(\gamma_r : r \in [s,t]), \pi_s)$ such that

$$\begin{cases} \mathrm{d}\xi_t = b(t, \xi_t, P_{s,t}^*\mu)\mathrm{d}t + \sigma(t, \xi_t, P_{s,t}^*\mu)\mathrm{d}B_t^1, \;\; \mathcal{L}_{\xi_s} = \nu, \;\; t \geq s \\ \mathrm{d}\eta_t = b(t, \eta_t, P_{s,t}^*\mu)\mathrm{d}t + \sigma(t, \eta_t, P_{s,t}^*\mu)\mathrm{d}B_t^2, \;\; \mathcal{L}_{\eta_s} = \mu, \;\; t \geq s. \end{cases} \tag{4.14}$$

satisfy $\pi_s(\xi \geq \eta) = 1$.

For any increasing $0 \leq f \in \mathcal{U}_b$, which does not depend on x_i, x_j, we have $\nu := \frac{fd\mu}{\mu(f)}$ with $\mu = \mu_{\{ij\}} \times \mu_{\{ij\}^c}$ such that $\nu_{ij} = \mu_{ij}$. Thus, $\xi_s^i = \eta_s^i, \; \xi_s^j = \eta_s^j$. So,

$$\begin{aligned} \xi_t^k - \eta_t^k &= \int_s^t \big(b_k(r, \xi_r, P_{s,r}^*\mu) - b_k(r, \eta_r, P_{s,r}^*\mu)\big)\mathrm{d}r \\ &+ \int_s^t \big\langle \sigma_{k\cdot}(r, \eta_r, P_{s,r}^*\mu), \mathrm{d}B_t^1 \big\rangle - \int_s^t \big\langle \sigma_{k\cdot}(r, \eta_r, P_{s,r}^*\mu), \mathrm{d}B_t^2 \big\rangle \geq 0, \quad k = i, j. \end{aligned} \tag{4.15}$$

Thus, by following the argument to derive (4.6), we have

$$a_{ij}(s, \xi_s, \mu) = \sigma_{i\cdot}(s, \xi_s, \mu)\sigma_{i\cdot}(s, \xi_s, \mu) = \sigma_{i\cdot}(s, \eta_s, \mu)\sigma_{j\cdot}(s, \eta_s, \mu) = a_{ij}(s, \eta_s, \mu).$$

This, together with $\nu = \frac{fd\mu}{\mu(f)}$, $\mathcal{L}_{\xi_s} = \nu$ and $\mathcal{L}_{\eta_s} = \mu$, leads to

$$\begin{aligned} \int f(x)a_{ij}(s, x, \mu)\mu(dx) &= \mu(f)\mathbb{E}a_{ij}(s, \xi_s, \mu) \\ &= \mu(f)\mathbb{E}a_{ij}(s, \eta_s, \mu) = \mu(f)\int a_{ij}(s, x, \mu)\mu(dx). \end{aligned} \tag{4.16}$$

Let g be a function such that $\mu(fg) = \mu(f)\mu(g)$. Then for $f(x) = I_A(x_k : k \neq i, j)$ with $A \in \mathcal{B}(\mathbb{R}^{(d-2)})$, we obtain

$$\mathbb{E}^\mu(I_A g) = \int_{\mathbb{R}^d} I_A(x)g(x)\mu(dx) = \mu(A)\mu(g).$$

Now, by the definition of conditional expectation we get $\mathbb{E}^\mu\big(g | x_k : k \neq i, j\big) = \mu(g)$, which obviously implies that g depends only on x_i, x_j. Thus, (4.16) yields the first assertion.

Dividing by $t - s$ on both side of (4.15) and taking $t \to s$, we get $\mathbb{E}b_i(s, \xi_s, \mu) \geq \mathbb{E}b_i(s, \xi_s, \mu)$. This, together with $\nu = \frac{fd\mu}{\mu(f)}$, $\mathcal{L}_{\xi_s} = \nu$ and $\mathcal{L}_{\eta_s} = \mu$, leads to the second assertion. □

References

[1] Barbu, V., Röckner, M.: From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. *Ann. Probab.* **48**, (2020), 1902–1920. MR4124528

[2] Øksendal, B.: Stochastic Differential Equations: An Introduction with Applications. *Springer-Verlag* (2003). MR2001996

[3] Chen, M-F., Wang, F-Y.: On order-preservation and positive correlations for multidimensional diffusion processes. *Probab Theory Relat. Fields* **95**, (1993), 421–28. MR1213199

[4] Forruin, C.M., Kasteleyn, P.W., Ginibre, J.: Correlation inequality on some partially ordered sets. *Comm. Math. Phys.* **22** (1971), 89–109. MR0309498

[5] Herbst, I., Pitt, L.: Diffusion equation techniques in stochastic monotonicity and positive correlations. *Probab Theory Relat. Fields* **87**, (1991), 275–312. MR1084331

[6] Huang, X., Wang F-Y.: Distribution dependent SDEs with singular coefficients. *Stoch. Proc. Appl.* **129**, (2019), 4747–4770. MR4013879

[7] Huang, X., Liu, C., Wang F-Y.: Order Preservation for Path-Distribution Dependent SDEs. *Commun. Pure Appl. Anal.* **17**, (2018), 2125–2133. MR3809144

[8] Kamae, T., Krengel, U., O'Brien, G. L.: Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5** (1977), 899–912. MR0494447

[9] Trevisan. D.: Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Elect. J. Probab.* **21**, (2016),1–44. MR3485364

[10] Ren, P., Wang, F.-Y.: Donsker-Varadhan Large Deviations for Path-Distribution Dependent SPDEs. *J. Math. Anal. Appl.* **499**, (2021), 125000. MR4207324

[11] Mishura, Yu. S.,Veretennikov, A. Yu.: Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations. *arXiv:1805.10841*.

[12] Strook, D., Varadhan, S.R.S: Multidimensional Diffusion Processes. *Springer*, 2007. MR2190038

[13] Wang, F-Y.: Distribution dependent SDEs for Landau type equations. *Stoch. Proc. Appl.* **128**, (2018), 595–621. MR3739509

Acknowledgement. I would like to thank the referee and Professor Feng-Yu Wang and Jianhai Bao for helpful comments and corrections.