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.

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 $P^2 = \{\mu \in P : \mu(|\cdot|^2) < \infty\}$, a space of probability measures on $\mathbb{R}^d$ with second moment. Under the Wasserstein distance

$$W_2(\mu, \nu) := \inf_{\pi \in \mathcal{C}(\mu, \nu)} \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dx, dy) \right)^{\frac{1}{2}},$$

$P_2$ is a Polish space, where $\mathcal{C}(\mu, \nu)$ is the space of all couplings of $\mu$ and $\nu$. Consider the following time-distribution dependent second order differential operators

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

where $a = (a_{ij})_{1 \leq i, j \leq d}$, $\bar{a} = (\bar{a}_{ij})_{1 \leq i, j \leq d}$, $b = (b_i)_{1 \leq i \leq d}$, $\bar{b} = (\bar{b}_i)_{1 \leq i \leq d}$ are continuous on $[0, \infty) \times \mathbb{R}^d \times P_2$. The nonlinear Fokker-Planck equations for $L$ and $\bar{L}$ are formulated as

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

*The author gratefully 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: RPP20E@GMAIL.COM.
Order preservation and positive correlations

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 \(f \in C^\infty_0(\mathbb{R}^d)\),

\[
\mu_t(f) = \mu_s(f) + \int_s^t \mu_r(L_{r, \mu}, f)dr, \quad \bar{\mu}_t(f) = \bar{\mu}_s(f) + \int_s^t \bar{\mu}_r(L_{r, \bar{\mu}}, f)dr, \quad t \geq s.
\]

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

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

\[
\mu \left([b(t, s, \mu) + \|a(t, s, \mu)\|] := \int_{\mathbb{R}^d} \left([b(t, s, \mu) + \|a(t, s, \mu)\|] \right) d\mu
\]

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

\[
\begin{align*}
\max \left\{ & b(t, x, \mu) - b(t, y, \nu), x - y \right\} + \|\sigma(t, x, \mu) - \sigma(t, x, \nu)\|_{HS}^2, \\
2\bar{b}(t, x, \mu) - \bar{b}(t, y, \nu), x - y \right\} + \|\bar{\sigma}(t, x, \mu) - \bar{\sigma}(t, x, \nu)\|_{HS}^2 \}
\end{align*}
\]

\[
\leq K(t)(|x - y|^2 + W_2(\mu, \nu)^2), \quad t \geq 0, \quad x, y \in \mathbb{R}^d, \quad \mu, \nu \in \mathcal{P}_2,
\]

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

Consider the distribution dependent SDEs

\[
\begin{align*}
dX_t &= b(t, X_t, \mathcal{L}_{X_t})dt + \sqrt{2a(t, X_t, \mathcal{L}_{X_t})}dW_t, \\
d\bar{X}_t &= \bar{b}(t, \bar{X}_t, \mathcal{L}_{\bar{X}_t})dt + \sqrt{2\bar{a}(t, \bar{X}_t, \mathcal{L}_{\bar{X}_t})}dW_t,
\end{align*}
\]

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 \(\xi\).

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 \(\mathcal{F}_t\)-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

\[
P^s_{t, \mu} := \mathcal{L}_{X_t}, \quad \bar{P}^s_{t, \bar{\mu}} := \mathcal{L}_{\bar{X}_t}, \quad \text{for } \mu = \mathcal{L}_{X_s}, \bar{\mu} = \mathcal{L}_{\bar{X}_s}
\]

are continuous in \(\mathcal{P}_2\) with respect to \(t \geq s\). Therefore, by the superposition principle in [1], (A) implies that \((\mu_t, \bar{\mu}_t)_{t \geq s} := \left(\left.P^s_{t, \mu}, \bar{P}^s_{t, \bar{\mu}}\right)_{t \geq s}\right)\) 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 \((P^s_{t, \mu}, \bar{P}^s_{t, \bar{\mu}})\) and \((\Lambda_s, \bar{\Lambda}_s)\), as well as on the positive correlations for \(\Lambda_s\). 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 \(\mathcal{B}(\mathbb{R}^d)\) be the set of all bounded measurable functions on \(\mathbb{R}^d\). Consider the class
of bounded measurable increasing functions: \( U_b := \{ f \in B_b(\mathbb{R}^d) \mid f(x) \leq f(y) \text{ for } x, y \in \mathbb{R}^d \text{ with } x \leq y \} \), and the family of probability measures of positive correlations: \( P_+ := \{ \mu \in P \mid \mu(fg) \geq \mu(f)\mu(g) \text{ for } f, g \in U_b \} \). If \( \mu \in P_+ \), then \( \mu \) is said to satisfy the FKG inequality. Moreover, we write \( \mu \preceq \nu \) for any two probability measures \( \mu, \nu \in P \), if \( \mu(f) \leq \nu(f) \) holds for any \( f \in U_b(\mathbb{R}^d) \). Note that definitions of \( P_+ \) and \( \mu \preceq \nu \) do not change if we replace \( B_0(\mathbb{R}^d) \) by \( C_0^k(\mathbb{R}^d) \) for \( k \in \mathbb{Z}_+ \cup \{ \infty \} \), where \( C_0^k(\mathbb{R}^d) = C_b(\mathbb{R}^d) \) denotes the set of all bounded continuous functions on \( \mathbb{R}^d \), while when \( k \geq 1 \) the class \( C_0^k(\mathbb{R}^d) \) consists of bounded functions on \( \mathbb{R}^d \) having bounded derivatives up to order \( k \).

For any \( t, \eta \in C_s := C([s, \infty); \mathbb{R}^d) \), we denote \( \xi \leq \eta \) if \( \xi_t \leq \eta_t \) for all \( t \geq s \). For any two probability measures \( \Phi_1, \Phi_2 \) on the path space \( C_s \), we denote \( \Phi_1 \preceq \Phi_2 \) if \( \Phi_1(F) \leq \Phi_2(F) \) holds for any bounded increasing function \( F \) on \( C_s \). Similarly, let \( P^s_+ \) denote the set of probability measures on \( C_s \) satisfying the FKG inequality for bounded increasing functions on \( C_s \).

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

1. We write \( \bar{P}^s_{s,t} \preceq P^s_{s,t} \mu \preceq P^s_{s,t} \nu \) holds for any \( \mu, \nu \in P_2 \) with \( \mu \preceq \nu \).
2. We write \( \Lambda_s \preceq \Lambda_s, \bar{\Lambda}_s \preceq \Lambda_s, \mu \preceq \nu \) holds for any \( \mu, \nu \in P_2 \) with \( \mu \preceq \nu \).
3. We write \( P^s_{s,t} \in P_+ \) if \( P^s_{s,t} \in P_+ \); and \( \Lambda_s \in P^s_+ \) if \( \Lambda_s \in P^s_+ \) for all \( \mu \in P_+ \).

Obviously, \( \bar{\Lambda}_s \preceq \Lambda_s \) for all \( s \geq 0 \) implies \( \bar{P}^s_{s,t} \preceq P^s_{s,t} \) for all \( t \geq s \geq 0 \), but the inverse may not be true in the nonlinear setting. Similarly, \( \Lambda_s \in P^s_+ \) implies \( P^s_{s,t} \in P_+ \) for any \( t \geq s \) 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 \( 1 \leq i \leq d \) and \( s \geq 0 \), \( \bar{b}_i(s, x, \nu) \leq b_i(s, y, \mu) \) holds for \( x \leq y \) with \( x_i = y_i \) and \( \nu \preceq \mu \);
2. \( a = \bar{a} \), and for any \( 1 \leq i, j \leq d, s \geq 0 \) and \( \mu \in P_2, a_{ij}(s, x, \mu) \) depends only on \( x_i \) and \( x_j \).

Then \( \bar{\Lambda}_s \preceq \Lambda_s \) for all \( s \geq 0 \). Consequently, \( \bar{P}^s_{s,t} \preceq P^s_{s,t} \), for \( t \geq s \).

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 necessary and sufficient conditions when \( b(t, x, \mu) \) and \( a(t, x, \mu) \) do not depend on \( \mu \). Indeed, when the coefficients do not depend on the distribution, we may take \( \nu = \delta_x \) and \( \mu = \delta_y \) for \( x \leq y \) 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 \( \mu \in P \) and \( I \subset \{1, \ldots, d\} \), let \( \mu_I(A) := \mu(\{ x \in \mathbb{R}^d : x_I \in A \}) \), \( A \in B(\mathbb{R}^{#I}) \) be the marginal distribution of \( \mu \) 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 \( s \geq 0 \), then the following conditions hold:

1. For any \( \nu \preceq \mu \) with \( \nu_i = \mu_i, 1 \leq i \leq d \), there exists a coupling \( \pi \in C(\nu, \mu) \) with \( \pi(\{ x \leq y \}) = 1 \) such that
   \[ b_i(s, x, \nu) \leq b_i(s, y, \mu), \quad s \geq 0, \quad (x, y) \in \text{supp} \pi. \]

Consequently, \( b_i(s, x, \mu) \leq b_i(s, x, \mu) \) for \( s \geq 0, x \in \mathbb{R}^d, \mu \in P_2 \).
Order preservation and positive correlations

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

$$a_{ij}(x, y) + \nu_{ij} \in \sup \mathcal{P}.$$ 

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

Since $\Lambda_s \preceq \Lambda_s$ implies $P^{s,t} \succeq P^{s,t}$, conditions in the following result are also necessary for $\Lambda_s \leq \Lambda_s$.

**Theorem 2.3.** Assume (A). If $P^{s,t} \succeq P^{s,t}$ for $t \geq s \geq 0$, then

(i) For any $s \geq 0$ and $1 \leq i \leq d$, $b_i(s, x, \nu) \preceq b_i(s, y, \mu)$ holds for $\nu \preceq \mu$ with $\nu_i = \mu_i$.
(ii) For any $s \geq 0$ and $1 \leq i,j \leq d$, $a_{ij}(s, x, \delta_x) = a_{ij}(s, x, \delta_x)$ 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 $s \geq 0$ and $1 \leq i \leq d$, $b_i(s, x, \nu) \preceq b_i(s, y, \mu)$, for $\nu \preceq \mu, x \preceq y$ with $x_i = y_i$.
2. For any $1 \leq i,j \leq d$, $a_{ij} \geq 0$, and for any $\mu \in \mathcal{P}_+$, $a_{ij}(s, x, \mu)$ depends only on $x_i$ and $x_j$.

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

**Theorem 2.5.** If $\Lambda_s \in \mathcal{P}_+$ for $s \geq 0$ and $\mu = \mu_{(ij)} \times \mu_{(ij)^c} \in \mathcal{P}_+$, then

1. For any $s \geq 0, 1 \leq i,j \leq d$, $a_{ij}(s, x, \delta_x) \geq 0$ and $a_{ij}(s, x, \mu)$ depends only on $x_i$ and $x_j$.
2. For any $s \geq 0, 1 \leq i,j \leq d$ and $f \in \mathcal{B}_b$ independent on $x_i$,

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

When $\sigma$ and $b$ do not depend on the distribution, by taking $\mu$ and $\nu$ 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 $t \geq 0$ and $x \in \mathbb{R}^d$,

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

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 $(L_t)_{t \geq 0}$ and $(\bar{L}_t)_{t \geq 0}$ are well-posed so that there exist unique time-inhomogeneous diffusion processes $(X^x_{s,t})_{t \geq s \geq 0}$ and $(\bar{X}^x_{s,t})_{t \geq s \geq 0}$ corresponding to $(L_t)_{t \geq 0}$ and $(\bar{L}_t)_{t \geq 0}$, respectively. Let $(P^x_{s,t})_{t \geq s \geq 0}$ and $(\bar{P}^x_{s,t})_{t \geq s \geq 0}$ be the Markov semigroups generated by $(X^x_{s,t})_{t \geq s \geq 0}$ and $(\bar{X}^x_{s,t})_{t \geq s \geq 0}$ with the initial value $X^x_{s,s} = X^x_{s,s} = x$, respectively, i.e.,

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

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

$$\frac{d}{ds} P^x_{s,t} f(x) = -L_t P^x_{s,t} f, \quad \frac{d}{dt} \bar{P}^x_{s,t} f(x) = \bar{L}_t \bar{P}^x_{s,t} f, \quad t \geq s \geq 0. \quad (3.3)$$
Order preservation and positive correlations

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

\[
(\bar{P}_{s,t}^* \mu)(f) := \mu(\bar{P}_{s,t} f), \quad (\bar{P}_{s,t}^* \mu)(f) := \mu(\bar{P}_{s,t} f), \quad f \in B_b(\mathbb{R}^d).
\]

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 \( \Lambda_{s,t} \mu \) and \( \bar{\Lambda}_{s,t} \mu \) be the distributions of the processes starting at \( \mu \) from time \( s \) generated by \( L \) and \( \bar{L} \) respectively. By the standard Markov property we see that \( \bar{P}_{s,t}^* \leq P_{s,t}^* \) for \( t \geq s \geq 0 \) if and only if \( \bar{\Lambda}_{s,t} \preceq \Lambda_{s,t} \) for \( s \geq 0 \), while \( P_{s,t}^* \in \mathcal{P}_+ \) for \( t \geq s \) is equivalent to \( \bar{\Lambda}_{s,t} \preceq \Lambda_{s,t} \) for \( s \geq 0 \).

**Theorem 3.1.** \( \bar{P}_{s,t}^* \leq P_{s,t}^* \) for \( t \geq s \geq 0 \), equivalently \( \bar{\Lambda}_{s,t} \preceq \Lambda_{s,t} \) for \( s \geq 0 \), if and only if the following conditions hold:

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

**Theorem 3.2.** \( P_{s,t}^* \in \mathcal{P}_+ \) for \( t \geq s \), equivalently \( \bar{\Lambda}_{s,t} \preceq \Lambda_{s,t} \) for \( s \geq 0 \), if and only if the following conditions hold:

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

**Proof of Theorem 3.1.** (a) We first prove the necessity. For any \( t \geq s \geq 0 \) and \( x \in \mathbb{R}^d \), let \( \Lambda^x_t \) (resp. \( \bar{\Lambda}^x_t \)) 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 \( 0 \leq s_0 \leq s_1 \leq \cdots s_n \), let \( \Lambda^x_{s_0, s_1, \cdots, s_n} \) be the marginal distribution of \( \Lambda^x_t \) at the time sequence \( (s_1, \cdots, s_n) \), which can be expressed via the Markov property as above

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

Then, by an inductive argument, together with the Markov property of the associated Markov process, \( \bar{P}_{s,t}^* \leq P_{s,t}^* \) implies \( \bar{\Lambda}^x_t \leq \Lambda^x_t \) (i.e., \( \bar{\Lambda}^x_t(f) \leq \Lambda^x_t(f) \) for any \( f \in U_b \cap C_s \)). Therefore, there exists a coupling \( \bar{P}^x_{s,y} \in C(\bar{\Lambda}^x_t, \Lambda^x_t) \) such that

\[
\bar{P}^x_{s,y}((\xi, \eta) \in C_s \times C_s : \eta \preceq \xi) = 1. \tag{3.4}
\]

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

\[
\begin{align*}
\frac{d\xi_t}{dt} &= b(t, \xi_t)dt + \sigma(t, \xi_t)d B^1_t, \quad \xi_s = y \\
\frac{d\eta_t}{dt} &= \bar{b}(t, \eta_t)dt + \bar{\sigma}(t, \eta_t)d B^2_t, \quad \eta_s = x
\end{align*} \tag{3.5}
\]

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

Let \( x \leq y \) with \( x_i = y_i \). Since \( \xi_t \geq \eta_t, \bar{P}^x_{s,y}\text{-a.s.}, \) and \( (\xi_t)_t = (\eta_t)_t \) due to \( x = y \), we derive from (3.5) that

\[
\int_s^t (b_i(r, \xi_r) - b_i(r, \eta_r))dr \geq \int_s^t \langle \sigma_i(r, \xi_r), dB^1_r \rangle - \int_s^t \langle \sigma_i(r, \xi_r), dB^1_r \rangle,
\]

where \( \sigma_i \) means the \( i \)-th row of \( \sigma \). Taking conditional expectation \( \bar{P}^x_{s_0}(\cdot | \mathcal{F}_{s_0}) \) on both sides yields

\[
\int_s^t E((b_i(r, \xi_r) - b_i(r, \eta_r)) | \mathcal{F}_{s_0})dr \geq 0, \quad t \geq s.
\]
Order preservation and positive correlations

This implies the assertion (1) by taking the continuity of \( b_i, \tilde{b}_i \) and \((\xi, \eta)\) into account.

Let \( x \leq y \) with \((x_i, x_j) = (y_i, y_j)\). Then, by using \( \tilde{\xi}_t \geq \eta_t, P_{\tilde{\eta}}^{t,s} \)-a.s., again, we have

\[
\int_0^t b_k(s, \xi_s) \, ds + \int_0^t \langle \sigma_k(s, \xi_s), dB^1_k \rangle \geq \int_0^t \tilde{b}_k(s, \eta_s) \, ds + \int_0^t \langle \tilde{\sigma}_k(s, \eta_s), dB^1_k \rangle, \quad k = i, j. \tag{3.6}
\]

Note that as \( t \downarrow s_0, \)

\[
\frac{1}{\sqrt{t - s_0}} \left( \int_s^t \langle \sigma_i(s, \xi_s), dB^1_k \rangle, \int_s^t \langle \sigma_j(s, \xi_s), dB^1_k \rangle \right) \text{ weakly} \rightarrow N \left( 0, \begin{pmatrix} a_{ij}(s, y) & a_{ij}(s, y) \\ a_{ij}(s, y) & a_{jj}(s, y) \end{pmatrix} \right) =: \mu_
\]

\[
\frac{1}{\sqrt{t - s_0}} \left( \int_s^t \langle \sigma_i(s, \xi_s), dB^1_k \rangle, \int_s^t \langle \sigma_j(s, \xi_s), dB^1_k \rangle \right) \text{ weakly} \rightarrow N \left( 0, \begin{pmatrix} \tilde{a}_{ij}(s, y) & \tilde{a}_{ij}(s, y) \\ \tilde{a}_{ij}(s, y) & \tilde{a}_{jj}(s, y) \end{pmatrix} \right) =: \tilde{\mu}.
\]

Then (3.6) implies \( \tilde{\mu} \preceq \mu \). Similarly, \( \mu \preceq \tilde{\mu} \). Therefore, we have \( \mu = \tilde{\mu} \) so that \( a = \tilde{a} \). 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., \( P_{s,t}^* \mu \leq P_{s,t}^* \nu \) with \( \mu \leq \nu \) for \( t \geq s \geq 0 \), provided the following two conditions hold:

1. \( b_i(s, \cdot) \) is smooth, \( b_i(s, x) \leq b_i(s, y) \) with \( x \leq y \) and \( x_i = y_i; \)
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 \( t \geq s \) and \( f \in \mathcal{U}_b \), it suffices to show

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

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 \( u_{s,t} = P_{s,t} f, t \geq s \). Then by (3.3), we have

\[
\partial_t u_{s,t} = L_t u_{s,t}, \quad t \geq s, \quad u_{s,s} = f.
\]

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

\[
\partial_k(\partial_k u_{s,t}) = \partial_k(\partial_k u_{s,t}) = L_k^k(\partial_k u_{s,t}) + \sum_{j=1}^d \alpha_{kj}(t, \cdot) \partial_j u_{s,t}, \tag{3.7}
\]

where \( L_k^k := L_t + \sum_{j=1}^d \left( 1 - \frac{1}{b_k(t, \cdot)} \partial_k a_{jk}(t, \cdot) \right) \partial_j + \partial_k b_k(t, \cdot), \quad \alpha_{kj}(t, \cdot) := (\partial_k b_k(t, \cdot)I_{k \neq j}). \)

Since \( L_k^k \) is a time-inhomogeneous Schrödinger operator, it generates a positivity-preserving semigroup \( (T_{s,t}^k)_{t \geq s} \). So, the operator \( \hat{L}_t := (L_k^k)_{1 \leq k \leq d} \) defined on \( C^2(\mathbb{R}^d; \mathbb{R}^d) \) by \( \hat{L}_t V := (L_k^k V_k)_{1 \leq k \leq d} \) generates a positivity preserving semigroup \( T_{s,t} := (T_{s,t}^k)_{1 \leq k \leq d} \) \( t \geq s \). Let \( D_r := (\alpha_{kj}(r, \cdot) \chi_{k \leq j \leq d}) \) and \( V_{s,t} = \nabla P_{s,t} f = \nabla u_{s,t} \). Then (3.7) implies

\[
\partial_t V_{s,t} = \hat{L}_t V_{s,t} + D_r V_{s,t}, \quad t \geq s, \quad V_{s,s} = \nabla f.
\]

This, together with Duhamel’s formula, gives

\[
V_{s,t} = T_{s,t} V_{s,s} + \int_s^t T_{r,t} D_r V_{s,r} \, dr, \quad t \geq s.
\]

Since \( V_{s,s} = \nabla f \geq 0 \) and \( T_{s,t} \), \( D_r \) are positivity preserving, this implies \( V_{s,t} = \nabla P_{s,t} f \geq 0 \).

**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. \( \square \)
Order preservation and positive correlations

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 \( \Lambda_s \mu \), \( \bar{P}_{s,t} \preceq P_{s,t}^\nu \) 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
\[
\mathcal{P}_{s,T}^{\nu,\mu} := \{ (\mu^{(1)}, \mu^{(2)}) \in \mathcal{C}([s,T]; \mathcal{P}_2 \times \mathcal{P}_2) : \mu^{(1)} \ll \mu^{(2)}, \, t \in [s,T], \, \mu_s^{(1)} = \nu, \mu_s^{(2)} = \mu \},
\]
which is a complete metric space under the metric for \( \lambda > 0 \),
\[
\rho_\lambda((\mu^{(1)}, \mu^{(2)}), (\mu^{(1)}_t, \mu^{(2)}_t)) := \sup_{t \in [s,T]} \mathbb{e}^{-\lambda t} \{ W_2(\mu^{(1)}_t, \mu^{(1)}_t) + W_2(\mu^{(2)}_t, \mu^{(2)}_t) \}.
\]
For any \( (\mu^{(1)}, \mu^{(2)}) \in \mathcal{P}_{s,T}^{\nu,\mu} \), consider the following time-dependent SDEs:
\[
\begin{align*}
\begin{cases}
\mathrm{d}X^{(1)}_{t}(\nu,\mu) & = b(t, X^{(1)}_{t}(\nu,\mu), \mu^{(1)}_t) \, \mathrm{d}t + \sigma(t, X^{(1)}_{t}(\nu,\mu), \mu^{(1)}_t) \, \mathrm{d}W_t, & t \geq s, \\
\mathrm{d}X^{(2,\nu,\mu)}_{t} & = b(t, X^{(2,\nu,\mu)}_{t}, \mu^{(2)}_t) \, \mathrm{d}t + \sigma(t, X^{(2,\nu,\mu)}_{t}, \mu^{(2)}_t) \, \mathrm{d}W_t, & t \geq s,
\end{cases}
\end{align*}
\]
where \( \sigma = \sqrt{2\alpha} \) and \( \bar{\sigma} = \sqrt{2\alpha} \), and \( \xi \sim \nu \) means \( \mathcal{L}_\xi = \nu \). Define the mapping on \( \mathcal{P}_{s,T}^{\nu,\mu} \) by
\[
H((\mu^{(1)}, \mu^{(2)}))(t) = \left( X^{(1),(\nu,\mu)}_{t}, X^{(2,\nu,\mu)}_{t} \right), \quad t \geq s. \tag{4.2}
\]
Since \( \mu^{(1)} \ll \mu^{(2)} \), by Theorem 3.1 the conditions in Theorem 2.1 imply \( \mathcal{L}_{X^{(1),(\nu,\mu)}} \preceq \mathcal{L}_{X^{(2,\nu,\mu)}} \), so that \( H : \mathcal{P}_{s,T}^{\nu,\mu} \to \mathcal{P}_{s,T}^{\nu,\mu} \). By Itô’s formula and the assumption (A), it is easy to see that \( H \) is contractive under the metric \( \rho_\lambda \) for large enough \( \lambda > 0 \), so that it has a unique fixed point and hence the proof is finished. \( \square \)

Proof of Theorem 2.2. Let \( s \geq 0 \) and \( \nu \preceq \mu \) with \( \nu_t = \mu_t \). 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
\[
\mathbb{P}_s \{ \{ (\xi, \eta) \in C_s : \xi_t \geq \eta_t, t \geq s \} \} = 1. \tag{4.3}
\]
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 \( \mathcal{L}_{(\xi, \eta)} \in \mathbb{P}_s \), where \( (\xi_t, \eta_t) \) solves
\[
\begin{align*}
\begin{cases}
\mathrm{d}\xi_t & = b(t, \xi_t, \mathcal{L}_{\eta_t}) \, \mathrm{d}t + \bar{\sigma}(t, \xi_t, \mathcal{L}_{\eta_t}) \, \mathrm{d}B_t^1, & t \geq s, \\
\mathrm{d}\eta_t & = b(t, \xi_t, \mathcal{L}_{\xi_t}) \, \mathrm{d}t + \bar{\sigma}(t, \xi_t, \mathcal{L}_{\xi_t}) \, \mathrm{d}B_t^2, & t \geq s,
\end{cases}
\end{align*}
\]
for some 2d-dimensional Brownian motions \( (B_t^1, B_t^2)_{t \geq s} \) 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 \( \mathcal{L}_{(\xi_t, \eta_t)} \in \mathcal{C}(\nu, \mu) \) and \( \nu_t = \mu_t \) imply \( \xi_s^t = \eta_s^t \). Thus, we find \( \mathbb{P}_s \)-a.s.
\[
\int_s^t b_s(r, \xi_r, \mu^{(1)}_r) \, \mathrm{d}r + \int_s^t \bar{\sigma}_s(r, \xi_r, \mu^{(1)}_r) \, \mathrm{d}B_r^1 \leq \int_s^t b_s(r, \eta_r, \mu^{(2)}_r) + \int_s^t \bar{\sigma}_s(r, \eta_r, \mu^{(2)}_r) \, \mathrm{d}B_r^1, \quad t \geq s. \tag{4.5}
\]
Taking conditional expectation with respect to \( \mathcal{F}_s \), we drive
\[
\int_s^t \mathbb{E}(b_s(r, \xi_r, \mu^{(1)}_r) | \mathcal{F}_s) \, \mathrm{d}r \leq \int_s^t \mathbb{E}(b_s(r, \eta_r, \mu^{(2)}_r) | \mathcal{F}_s) \, \mathrm{d}r, \quad t \geq s.
\]


Page 7/12
Order preservation and positive correlations

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

\[
\bar{b}_t(s, \xi, \nu) \leq \bar{b}_t(s, \eta, \nu), \quad t \geq s, \quad \mathbb{P}_s - a.s.
\]

Consequently, letting \( \pi := L(\xi, \eta) \in \mathcal{C}(\nu, \mu) \), we have \( \pi(\{x \leq y\}) = 1 \) and

\[
\bar{b}_t(s, x, \nu) \leq b_t(s, y, \mu), \quad (x, y) \in \text{supp}\pi.
\]

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

\[
b_t(s, x, \mu) \leq b_t(s, x, \mu), \quad x \in \text{supp}\mu.
\]

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

\[
\bar{b}_t(s, x, \mu_\varepsilon) \leq b_t(s, x, \mu), \quad s \geq 0, \quad \varepsilon > 0.
\]

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 \( \varepsilon \in [0, 1] \),

\[
\int_s^t \left[ \varepsilon \bar{b}_t(r, \xi, \mu^{(1)}_i) + (1 - \varepsilon)\bar{b}_t(r, \xi, \mu^{(1)}_j) \right] \, dr + \int_s^t \left[ \varepsilon \bar{a}_t(r, \delta_{\xi}, \mu^{(1)}_i) + (1 - \varepsilon)\bar{a}_t(r, \delta_{\xi}, \mu^{(1)}_j) \right] \, dB_r^1
\]

\[
\leq \int_s^t \left[ \varepsilon \bar{b}_t(r, \eta, \mu^{(2)}_i) + (1 - \varepsilon)\bar{b}_t(r, \eta, \mu^{(2)}_j) \right] \, dr
\]

\[
+ \int_s^t \left[ \varepsilon \bar{a}_t(r, \delta_{\eta}, \mu^{(2)}_i) + (1 - \varepsilon)\bar{a}_t(r, \delta_{\eta}, \mu^{(2)}_j) \right] \, dB_r^2
\]

Dividing both sides by \( \frac{1}{\sqrt{t - s}} \) and letting \( t \downarrow s \), we find

\[
N(0, \varepsilon^2 \bar{a}_{ij}(s, \xi, \nu) + 2\varepsilon(1 - \varepsilon)\bar{a}_{jj}(s, \xi, \nu) + (1 - \varepsilon)^2 \bar{a}_{jj}(s, \xi, \nu))
\]

\[
\leq N(0, \varepsilon^2 a_{ij}(s, \eta, \mu) + 2\varepsilon(1 - \varepsilon)a_{ij}(s, \eta, \mu) + (1 - \varepsilon)^2 a_{jj}(s, \eta, \mu)).
\]

By the symmetry of centred normal distribution, this further implies

\[
\varepsilon^2 \bar{a}_{ij}(s, \xi, \nu) + 2\varepsilon(1 - \varepsilon)\bar{a}_{jj}(s, \xi, \nu) + (1 - \varepsilon)^2 \bar{a}_{jj}(s, \xi, \nu)
\]

\[
= \varepsilon^2 a_{ij}(s, \eta, \mu) + 2\varepsilon(1 - \varepsilon)a_{ij}(s, \eta, \mu) + (1 - \varepsilon)^2 a_{jj}(s, \eta, \mu), \varepsilon \in [0, 1].
\]

Consequently, dividing by \( \varepsilon^2 \) on both sides yields

\[
\bar{a}_{ij}(s, \xi, \nu) = a_{ij}(s, \eta, \mu), \quad \mathbb{P}_s - a.s., \quad (4.6)
\]

which gives for \( \pi = L(\xi, \eta) \in \mathcal{C}(\nu, \mu) \),

\[
\bar{a}_{ij}(s, x, \nu) = a_{ij}(s, y, \mu), \quad (x, y) \in \text{supp}\pi, \quad s \geq 0.
\]

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

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

\[
\mathbb{E}(X_{s,t})_{i} \leq \mathbb{E}(X_{s,t}).
\]

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

\[
\int_s^t \mathbb{E}(b_t(s, x, \mathcal{L}_{X_{s,t}})) \, dr \leq \int_s^t \mathbb{E}(b_t(s, x, \mathcal{L}_{\bar{X}_{s,t}})) \, dr.
\]
Dividing by $t - s$ on both side followed by $t \downarrow s$, we get (i).

Since $P^*_s \leq P^*_{t,s}$, for $f \in \mathcal{C}^\infty_b(\mathbb{R}^d)$, we have $\mu(L_{s}\mu f) \leq \mu(L_{t}\mu f)$. In particular, taking $\mu = \delta_x$ yields $L_{s}\delta_x f(x) \leq L_{t}\delta_x f(x)$. 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 $\mu^{(1)} \leq \mu^{(2)}$. Then, $\mu \in \mathcal{P}_+$.

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

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

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

$$\mu(fg) = \frac{1}{2}(\mu^{(1)}(fg) + \mu^{(2)}(fg)) \geq \frac{1}{4}(\mu^{(1)}(f) + \mu^{(2)}(f))(\mu^{(1)}(g) + \mu^{(2)}(g)) = \mu(f)\mu(g).$$

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

$$2(\mu^{(1)}(L_{s}\mu f) + \mu^{(2)}(L_{s}\mu g)) \geq (\mu^{(1)}(L_{s}\mu f) + \mu^{(2)}(L_{s}\mu f))(\mu^{(1)}(g) + \mu^{(2)}(g))$$

$$+ (\mu^{(1)}(L_{s}\mu g) + \mu^{(2)}(L_{s}\mu g))(\mu^{(1)}(f) + \mu^{(2)}(f)).$$

(4.7)

$$2(\mu^{(1)}(\Gamma_1 f) + \mu^{(2)}(\Gamma_1 f)) \geq (\mu^{(2)}(L_{s}\mu g) - \mu^{(1)}(L_{s}\mu g))(\mu^{(1)}(f) - \mu^{(2)}(f))$$

$$+ (\mu^{(2)}(L_{s}\mu f) + \mu^{(1)}(L_{s}\mu f))(\mu^{(1)}(g) - \mu^{(2)}(g)).$$

(4.8)

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

**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 $(P^*_{s,t})_{s \geq t}$ preserves positive correlations, we have

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

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

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

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

$$\frac{2}{t - s}((P^*_{s,t}\mu^{(1)}(f) + (P^*_{s,t}\mu^{(2)}(f))(fg) - (\mu^{(1)}(f) + (\mu^{(2)}(fg))$$

$$\geq \frac{1}{t - s} \left\{ ((P^*_{s,t}\mu^{(1)}(f) + (P^*_{s,t}\mu^{(2)}(f))(P^*_{s,t}\mu^{(1)}(g) + (P^*_{s,t}\mu^{(2)}(g)) - (\mu^{(1)}(f) + (\mu^{(2)}(f))(\mu^{(1)}(g) + \mu^{(2)}(g)).$$

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

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

Proof. By Lemma 4.2 with $\mu^{(1)} = \mu^{(2)} = \delta_x$, for any $f, g \in U_0 \cap C^\infty_b (\mathbb{R}^d)$,

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

By choosing $f, g \in U_0 \cap C^\infty_b (\mathbb{R}^d)$ such that $f(z) = z_i - x_i$ and $g(z) = z_j - x_j$ holds in a neighbourhood of $x$, we obtain $a_{ij}(t, x, \delta_s u) \geq 0$.

**Lemma 4.4.** Let $\mu^{(1)} := \mu_1^{(1)} \times \mu_i^{(1)}$ and $\mu^{(2)} := \mu_1^{(2)} \times \mu_i^{(2)}$ with $\mu_1^{(1)} = \mu_i^{(2)}$, where $\mu_1^{(1)}$, $\mu_i^{(2)}$, $\mu_1^{(1)}$, $\mu_i^{(2)} \in \mathcal{P}_+$. If $(P_{s,t})_{t \geq s}$ preserves positive correlation, then $\mu^{(1)} \left( (b_1(t, \cdot, \mu^{(1)})) \right) \leq \mu^{(2)} \left( (b_1(t, \cdot, \mu^{(2)})) \right)$.

Proof. Since $\mu_1^{(1)}$, $\mu_i^{(2)}$, $\mu_i^{(1)}$, $\mu_i^{(2)} \in \mathcal{P}_+$, we deduce $\mu^{(1)}$, $\mu^{(2)} \in \mathcal{P}_+$. For given $i$ and $k \neq i$, take $f, g \in U_0 \cap C^\infty_b (\mathbb{R}^d)$ such that in a neighbourhood of $x$,

$$f(z) = z_i - \int_{\mathbb{R}} r \mu_1^{(1)}(dr), \quad g(z) = \frac{h(z_i)}{\mu^{(1)}(h(1 + T \| z \|_1))} - \frac{h(z_i)}{\mu^{(2)}(h(1 + T \| z \|_1))} \mu^{(1)}(1 + T \| z \|_1),$$

where $h \in C^\infty(\mathbb{R}; \mathbb{R}^d)$ 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 $\mu^{(1)} \left( (b_1(t, \cdot, \mu^{(1)})) \right) \leq \mu^{(2)} \left( (b_1(t, \cdot, \mu^{(2)})) \right)$.

**Proof of Theorem 2.4.** Since $P_{s,t}^{\ast, \mu}$ is the marginal distributions of $\Lambda_s \mu$ at time $t$, $\Lambda_s \in \mathcal{P}_+$ implies $P_{s,t}^{\ast, \mu} \in \mathcal{P}_+$ for $t \geq s$. So, it suffices to prove $\Lambda_s \in \mathcal{P}_+$. 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

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

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

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

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

$$W_{2,\lambda}(\mu, \nu) := \sup_{t \in [s, T]} \left( e^{-\lambda t} W_2(\mu_t, \nu_t) \right). \quad (4.10)$$

For $\nu \in D_{\ast}$ and $b_\nu'(x) := b_1(x, \nu_t), \sigma_\nu'(x) := \sqrt{2a_\nu(x, \nu_t)}$, consider the time-dependent SDE

$$dX_\nu^t = b_\nu'(X_\nu^t) dt + \sigma_\nu'(X_\nu^t) dW_t, \quad t \in [s, T], \quad X_\nu^s = X_s \sim \mu_0, \quad (4.11)$$

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

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

By Itô’s formula and (1.3), it is easy to see that $\Phi$ is contractive under the complete metric

$$W_{2,\lambda}(\nu_1, \nu_2^2) := \sup_{t \in [s, T]} \left( e^{-\lambda t} W_2(\nu_1^t, \nu_2^t) \right)$$

on $D_{\ast}$, so that it has a unique fixed point $\nu \in D_{\ast}$. Thus, $\Lambda_{s,T} \mu_0 := \mathcal{L}_{\{X_\nu^t \mid t \in [s, T]\}}$ with $\nu_t = \mathcal{L}_{X_t} \in \mathcal{P}_+$, $t \in [s, T]$. Therefore, by applying Theorem 3.2 to the diffusion process generated by $L_\nu$ with coefficients $(b_\nu', \sigma_\nu')$, we conclude that the present conditions (1) and (2) imply (4.9) as desired.
Order preservation and positive correlations

**Proof of Theorem 2.5.** Consider the decoupled SDE

\[ dX^{x,\mu}_{s,t} = b(t, X^{x,\mu}_{s,t}, P^{*}_{s,t,\mu})dt + \sigma(t, X^{x,\mu}_{s,t}, P^{*}_{s,t,\mu})dW_t, \quad t \geq s, \quad X^{x,\mu}_{s,s} = x, \]  

(4.12)

where \( P^{*}_{s,t,\mu} \) is the marginal distribution of \( \Lambda^\mu \), at the time \( t \). For \( x \in \mathbb{R}^d \), let \( \Lambda^\mu_{s,t} = \mathbb{L}X^{x,\mu}_{[s,t]} \).

Then for any \( \nu \in \mathcal{P}, \Lambda^\mu_{s,t} := \int_{\mathbb{R}^d} \Lambda^\mu_{s,t}(dx) \) is the law of \( X^{x,\mu}_{[s,t]} \) with initial distribution \( \nu \).

Noting that \( \Lambda^\mu_s = \Lambda^\mu_{s,s} = \mathbb{L}X^{x,\mu}_{[s,s]} \), by \( \Lambda^\mu \in \mathcal{P}^1 \) we obtain

\[ \Lambda^\mu_{s,t}(FG) \geq \Lambda^\mu_{s,t}(F) \Lambda^\mu_{s,t}(G), \quad F,G \in \mathcal{U}(C_s), \mu \in \mathcal{P}_+. \]  

(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

\[ \Lambda^\mu_{s,t}(\gamma) \geq \Lambda^\mu_{s,t}(F) \Lambda^\mu_{s,t}(G), \quad G \in \mathcal{U}(C_s), \]  

where \( \nu(dx) := \frac{f(x)\mu(dx)}{\mu(f)} \). That is, \( \Lambda^\nu_{s,t} \geq \Lambda^\mu_{s,t} \).

Thus, there exist \( \pi \in \mathcal{C}(\Lambda^\nu_{s,t}, \Lambda^\mu_{s,t}) \) and Brownian motions \( B^1_t \) and \( B^2_t \) on \( (\Omega, \mathcal{F}, \mathbb{P}) := (C_s, \sigma(\gamma_r : r \in [s,t]), \pi_s) \) such that

\[ \begin{align*}
\frac{d\xi_t}{dt} &= b(t, \xi_t, P^*_{s,t,\mu})dt + \sigma(t, \xi_t, P^*_{s,t,\mu})dB^1_t, \quad \mathcal{L}_{\xi_t} = \nu, \quad t \geq s, \\
\frac{d\eta_t}{dt} &= b(t, \eta_t, P^*_{s,t,\mu})dt + \sigma(t, \eta_t, P^*_{s,t,\mu})dB^2_t, \quad \mathcal{L}_{\eta_t} = \mu, \quad t \geq s.
\end{align*} \]  

(4.14)

satisfy \( \pi_s(\xi_t \geq \eta_t) = 1 \).

For any increasing \( 0 \leq f \in \mathcal{U}_b \) which does not depend on \( x_i, x_j \), we have \( \nu := \int fdu \) with \( \mu = \mu(x_i) \times \mu(x_j) \) such that \( \nu_{ij} = \mu_{ij} \). Thus, \( \xi^i_s = \eta^i_s, \xi^j_s = \eta^j_s \).

So,

\[ \xi^k_t - \eta^k_t = \int_s^t \left( b_k(r, \xi^r, P^*_{s,r,\mu}) - b_k(r, \eta^r, P^*_{s,r,\mu}) \right)dr \]

\[ + \int_s^t \langle \sigma_k(r, \eta^r, P^*_{s,r,\mu}), dB^1_t \rangle - \int_s^t \langle \sigma_k(r, \eta^r, P^*_{s,r,\mu}), dB^2_t \rangle \geq 0, \quad k = i, j. \]  

(4.15)

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

\[ a_{ij}(s, \xi, \mu) = \sigma_i(s, \xi, \mu)\sigma_i(s, \xi, \mu) = \sigma_i(s, \eta, \mu)\sigma_j(s, \eta, \mu) = a_{ij}(s, \eta, \mu). \]

This, together with \( \nu = \int fdu \), \( \mathcal{L}_{\xi_t} = \nu \) and \( \mathcal{L}_{\eta_t} = \mu \), leads to

\[ \int f(x)a_{ij}(s, x, \mu)\mu(dx) = \mu(f)\mathbb{E}a_{ij}(s, \xi, \mu) \]

\[ = \mu(f)\mathbb{E}a_{ij}(s, \eta, \mu) = \mu(f)\int a_{ij}(s, x, \mu)\mu(dx). \]  

(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

\[ \mathbb{E}^\mu(I_{Ag}) = \int_{\mathbb{R}^d} I_A(x)g(x)\mu(dx) = \mu(A)\mu(g). \]

Now, by the definition of conditional expectation we get \( \mathbb{E}^\mu(g|x_k: k \neq i, j) = \mu(g) \), 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 \( \mathbb{E}b_i(s, \xi, \mu) \geq \mathbb{E}b_i(s, \xi, \mu) \). This, together with \( \nu = \int fdu \), \( \mathcal{L}_{\xi_t} = \nu \) and \( \mathcal{L}_{\eta_t} = \mu \), leads to the second assertion. \( \square \)
Order preservation and positive correlations

References


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