Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 105, 1–52.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP357

Hölder continuity of the solutions to a class of SPDE's arising from branching particle systems in a random environment

Yaozhong Hu*† David Nualart‡§ Panqiu Xia¶

Abstract

We consider a d-dimensional branching particle system in a random environment. Suppose that the initial measures converge weakly to a measure with bounded density. Under the Mytnik-Sturm branching mechanism, we prove that the corresponding empirical measure X_t^n converges weakly in the Skorohod space $D([0,T];M_F(\mathbb{R}^d))$ and the limit has a density $u_t(x)$, where $M_F(\mathbb{R}^d)$ is the space of finite measures on $\mathbb{R}^d$. We also derive a stochastic partial differential equation $u_t(x)$ satisfies. By using the techniques of Malliavin calculus, we prove that $u_t(x)$ is jointly Hölder continuous in time with exponent $\frac{1}{2}-\epsilon$ and in space with exponent $1-\epsilon$ for any $\epsilon>0$.

Keywords: Branching particle system; random environment; stochastic partial differential equations; Malliavin calculus; Hölder continuity.
AMS MSC 2010: 60H07; 60H15; 60J68.
Submitted to EJP on October 18, 2018, final version accepted on September 8, 2019.

1 Introduction

Consider a d-dimensional branching particle system in a random environment. For any integer $n\geq 1$, the branching events happen at time $\frac{k}{n}$, $k=1,2,\dots$. The dynamics of each particle, labelled by a multi-index α, is described by the stochastic differential equation (SDE):

$$dx_t^{\alpha,n}=dB_t^{\alpha}+\int_{\mathbb{R}^d}h(y-x_t^{\alpha,n})W(dt,dy), \tag{1.1}$$

where h is a $d\times d$ matrix-valued function on $\mathbb{R}^d$, whose entries $h^{ij}\in L^2(\mathbb{R}^d)$, B^α are d-dimensional independent Brownian motions, and W is a d-dimensional space-time white

*Department of Mathematical and Statistical Science, University of Alberta at Edmonton, Edmonton, Canada. E-mail: yaozhong@ualberta.ca
†Supported by an NSERC discovery grant.
‡Department of Mathematics, University of Kansas, Lawrence, KS, USA. E-mail: nualart@ku.edu
§Supported by the NSF grant DMS 1811181.
¶Department of Mathematics, University of Kansas, Lawrence, KS, USA. E-mail: pqxia@ku.edu

Gaussian random field on $\mathbb{R}_+ \times \mathbb{R}^d$ independent of the family $\{B^\alpha\}$. The random field W can be regarded as the random environment for the particle system. The existence and uniqueness of the Feller process $x^{\alpha,n}$ that solves the SDE (1.1) will be proved in Section 2.

At any branching time each particle dies and it randomly generates offspring. The new particles are born at the death position of their parents, and inherit the branching-dynamics mechanism. The branching mechanism we use in this paper follows the one introduced by Mytnik [23], and studied further by Sturm [30]. Let $X^n = \{X^n_t, t \geq 0\}$ denote the empirical measure of the particle system. One of the main results of this work is to prove that the empirical measure-valued processes converge weakly to a process $X = \{X_t, t \geq 0\}$, such that for almost every $t \geq 0$, X_t has a density $u_t(x)$ almost surely. By using the techniques of Malliavin calculus, we also establish the almost surely joint Hölder continuity of u with exponent $\frac{1}{2} - \epsilon$ in time and $1 - \epsilon$ in space for any $\epsilon > 0$.

To compare our results with the classical ones. Let us recall briefly some existing work in the literature. The one-dimensional model was initially introduced and studied by Wang ([32, 33]). In these papers, he proved that under the classical Dawson-Watanabe branching mechanism, the empirical measure X^n converges weakly to a process $X = \{X_t, t \geq 0\}$, which is the unique solution to a martingale problem.

For the above one dimensional model Dawson et al. [8] proved that for almost every $t > 0$, the limit measure-value process X has a density $u_t(x)$ a.s. and u is the weak solution to the following stochastic partial differential equation (SPDE):

$$\begin{aligned} u_t(x) =& \mu(x) + \int_0^t \frac{1}{2}(1 + \|h\|_2^2)\Delta u_s(x) ds - \int_0^t \int_{\mathbb{R}} \nabla_x [h(y-x)u_s(x)] W(ds, dy) \\ &+ \int_0^t \sqrt{u_s(x)} \frac{V(ds, dx)}{dx}, \end{aligned} \tag{1.2}$$

where $\|h\|_2$ is the L^2-norm of h, and V is a space-time white Gaussian random field on $\mathbb{R}_+ \times \mathbb{R}$ independent of W.

Suppose further that h is in the Sobolev space $H_2^2(\mathbb{R})$ and the initial measure has a density $\mu \in H_2^1(\mathbb{R})$. Then Li et al. [20] proved that $u_t(x)$ is almost surely jointly Hölder continuous. By using the techniques of Malliavin calculus, Hu et al. [14] improved their result to obtain the sharp Hölder continuity: they proved that the Hölder exponents are $\frac{1}{4} - \epsilon$ in time and $\frac{1}{2} - \epsilon$ in space, for any $\epsilon > 0$.

Our paper is concerned with higher dimensions ($d > 1$). However in this case, the super Brownian motion (a special case when $h = 0$) does not have a density (see e.g. Corollary 2.4 of Dawson and Hochberg [6]). Thus in higher dimensional case we have to abandon the classical Dawson-Watanabe branching mechanism and adopt the Mytnik-Sturm one. As a consequence, the difficult term $\sqrt{u_s(x)}$ in the SPDE (1.2) becomes $u_s(x)$ (see equation (3.1) in Section 3 for the exact form of the equation).

We follow the approach introduced in [14] to study the Hölder continuity of the conditional density of a particle motion using Malliavin calculus. However, because of the multidimensional setting considered here, new difficulties arise. On one hand, the integration by parts formulas require higher order Malliavin derivatives which make computations more complex. To lower the order of Malliavin differentiability in our framework, we use the combination of Riesz transform and Malliavin calculus, previously studied in depth by Bally and Caramellino [1] (see Appendix A for the density formula that we are using). Another difficulty is the fact that in the one-dimensional case considered in [14], the Malliavin derivative can be expressed explicitly and this type of formula for the Malliavin derivative is no longer available here. We have to use another approach to obtain appropriate sharp estimates. More details are given in Appendix A.

This paper is organized as follows. In Section 2 we shall briefly describe the branching mechanism used in this paper. In Section 3 we state the main results obtained in this paper. These include three theorems. The first one (Theorem 3.3) is about the existence and uniqueness of a (linear) stochastic partial differential equation (equation (3.1)), which is proved (Theorem 3.2) to be satisfied by the density of the limiting empirical measure process X^n of the particle system (see (2.12)). The core result of this paper is Theorem 3.4 which intends to give sharp Hölder continuity of the solution $u_t(x)$ to (3.1).

Section 4 presents the proofs for Theorems 3.2 and 3.3. The proof of Theorem 3.4 is the objective of the remaining sections. First, in Section 5, we focus on the one-particle motion with no branching. By using the techniques from Malliavin calculus, we obtain a Gaussian type estimates for the transition probability density of the particle motion conditional on W. This estimate plays a crucial role in the proof of Theorem 3.4. In Section 6, we derive a conditional convolution representation of the weak solution to the SPDE (3.1), which is used to establish the Hölder continuity. In Section 7, we show that the solution u to (3.1) is Hölder continuous.

Lastly, the martingale problem (4.4)–(4.5) is introduced in Section 4 to prove Theorems 3.2 and 3.3. The well-posedness of the martingale problem can be proved under the assumption that the initial measure has a bounded density. We conjecture that it also holds for an arbitrary finite initial measure. We will not pursue this in this paper (see Remark 4.12 (ii)).

2 Branching particle system

We split this section into two parts. In Section 2.1, we consider a finite branching-free particle system, and prove the existence and uniqueness of this system. In Section 2.2, we give a brief induction to the Mytnik-Sturm branching mechanism.

2.1 Finite branching-free particle system

In this section, we will show the existence and uniqueness of the finite branching-free particle system that is determined by (1.1). The one-dimensional analogue is given by Lemma 1.3 of Wang [32].

Fix a time interval $[0,T]$. Let $W=\{W(t,x),(t,x)\in[0,T]\times\mathbb{R}^d\}$ be a d-dimensional space-time white Gaussian random field. For any positive integer n, let $\{B^i\}_{i\in\{1,\dots,n\}}$ be a family of independent d-dimensional Brownian motions that is independent of W. Consider an n-particle system, where the motion of each particle is described by the following stochastic differential equation in a random environment W:

$$dx_t^i = dB_t^i + \int_{\mathbb{R}^d} h(y-x_t^i)W(dt,dy), \tag{2.1}$$

with initial condition $x_0^i\in\mathbb{R}^d$ for all $i=1,\dots n$. In the case $n=1$, we omit all upper indexes in equation (2.1) without confusion.

The following hypothesis for h will be used throughout this paper:

[H0] $h=(h^{ij})_{1\le i,j\le d}\in H_2^3(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$. That is, the entries h^{ij} of h belongs to the Sobolev space $H_2^3(\mathbb{R}^d)$.

For $k=0,1,2,3$, denote by $\|\cdot\|_{k,2}$ the Sobolev norm on $H_2^k(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$, that is

$$\|h\|_{k,2}^2 := \sum_{i,j=1}^d \|h^{ij}\|_{k,2}^2 = \sum_{i,j=1}^d \Big(\int_{\mathbb{R}^d}|h^{ij}(x)|^2dx\Big)^{\frac12} + \sum_{i,j=1}^d\sum_{l=1}^k\sum_{i_1,\dots,i_l=1}^d\Big(\int_{\mathbb{R}^d}\Big|\frac{\partial^l}{\partial x_1\cdots\partial x_l}h^{ij}(x)\Big|^2dx\Big)^{\frac12}.$$

Let $\rho : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be given by

$$\rho(x) = \int_{\mathbb{R}^d} h(z-x)h^*(z)dz, \tag{2.2}$$

where $h^* = (h^{ji})_{1\le i,j\le d}$ denotes the transpose of h. Then, for any $1 \le i, j \le d$, and $x \in \mathbb{R}^d$, by Cauchy-Schwarz's inequality, we have

$$|\rho^{ij}(x)| \le \sum_{k=1}^{d} \|h^{ik}\|_2 \|h^{kj}\|_2.$$

We denote by $\|\cdot\|_2$ the Hilbert Schmidt norm for matrices. Then, by Cauchy-Schwarz's inequality again, we have

$$\begin{aligned}\|\rho\|_\infty := \sup_{x\in\mathbb{R}^d} \|\rho(x)\|_2 = \sup_{x\in\mathbb{R}^d} \Big(\sum_{i,j=1}^{d} |\rho^{ij}(x)|^2 \Big)^{\frac{1}{2}} \\ \le \Big(\sum_{i,j=1}^{d} \Big| \sum_{k=1}^{d} \|h^{ik}\|_2 \|h^{kj}\|_2 \Big|^2 \Big)^{\frac{1}{2}} \\ \le \Big(\sum_{i,k=1}^{d} \|h^{ik}\|_2^2 \sum_{j,k=1}^{d} \|h^{kj}\|_2^2 \Big)^{\frac{1}{2}} \le \|h\|_2^2 \le \|h\|_{3,2}^2.\end{aligned}$$

Similarly, we can show that the first, second, and third partial derivatives of ρ are bounded in $\mathbb{R}^d$. We make use of the following notations:

$$\|\rho\|_{k,\infty} := \sup_{x\in\mathbb{R}} \Big(\sum_{i,j=1}^{d} \sum_{i_1,\dots,i_k=1}^{d} \Big| \frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}} \rho^{ij}(x) \Big|^2 \Big)^{\frac{1}{2}},$$

for $k = 1, 2, 3$. Now let us study the SDE's (2.1). These equations are not coupled and we solve them for each i separately. For this reason in the next theorem, which provides the existence and uniqueness of the equation (for each fixed i), we suppress the superscript index i.

Theorem 2.1. *Assume the hypothesis **[H0]**. Then, there exists a d-dimensional stochastic process $x = \{x_t, 0 \le t \le T\}$ that is the unique strong solution to the SDE (2.1) (for each fixed i) with initial condition $x_0 = x \in \mathbb{R}^d$.*

Proof. We prove this theorem by Picard iteration. Let

$$x_t^{(0)} = B_t + \int_0^t \int_{\mathbb{R}^d} h(y-x)W(ds,dy),$$

and let

$$x_t^{(m)} = B_t + \int_0^t \int_{\mathbb{R}^d} h(y - x_s^{(m-1)})W(ds,dy),$$

for all $m \ge 1$. Denote by $d_t^{(m)} = x_t^{(m)} - x_t^{(m-1)}$ for all $t \in [0,T]$. Then $d_t^{(m)}$ satisfies the following equation

$$d_t^{(m)} = \int_0^t \int_{\mathbb{R}^d} \Big[h(y - x_s^{(m)}) - h(y - x_s^{(m-1)}) \Big] W(ds,dy). \tag{2.3}$$

An application of the Itô isometry yields that

$$\begin{aligned}
\|d_t^{(m)}\|_2^2 =& \left\| \int_0^t \int_{\mathbb{R}^d} \left[h(y - x_s^{(m)}) - h(y - x_s^{(m-1)}) \right] W(ds, dy) \right\|_2^2 \\
=& \mathbb{E} \int_0^t \int_{\mathbb{R}^d} \left\| h(y - x_s^{(m)}) - h(y - x_s^{(m-1)}) \right\|_2^2 dy ds \\
=& 2\|h\|_2^2 t - 2 \sum_{i,j=1}^d \int_0^t \rho^{ij}(d_s^{(m-1)}) ds \\
=& 2 \sum_{i,j=1}^d \int_0^t \left[\rho^{ij}(d_s^{(m-1)}) - \rho^{ij}(0) \right] ds \,, \qquad (2.4)
\end{aligned}$$

since $\sum_{i,j=1}^d \rho^{ij}(0) = \|h\|_2^2$. Noticing that ρ^{ij} has bounded first partial derivatives, we have

$$\|d_t^{(m)}\|_2^2 \le C \int_0^t \|d_s^{(m-1)}\|_2^2 ds,$$

for some constant C independent of m. On the other hand, we can show that

$$\begin{aligned}
\|x_t^{(0)} - x\|_2^2 =& \left\| B_t + \int_0^t \int_{\mathbb{R}^d} h(y - x) W(ds, dy) - x \right\|_2^2 \\
\le& t + t\|h\|_2^2 + |x|^2.
\end{aligned}$$

By iteration, we can conclude that

$$\|d_t^{(m)}\|_2^2 \le \frac{1}{(m+1)!}(1 + \|h\|_2^2) t^{m+1} + \frac{1}{m!}|x|^2 t^m, \qquad (2.5)$$

which is summable in m. In other words, for any $t \in [0, T]$, the sequence $x_t^{(m)}$ is convergent in $L^2(\Omega)$. Denote by x_t the limit of this sequence.

We claim that $x = \{x_t, 0 \le t \le T\}$ is a strong solution to (2.1) (recall we suppress the superscript). It suffices to show that as $m \to \infty$,

$$\int_0^t \int_{\mathbb{R}^d} h(y - x_s^{(m)}) W(ds, dy) \to \int_0^t \int_{\mathbb{R}^d} h(y - x_s) W(ds, dy),$$

in $L^2(\Omega)$ for all $t \in [0, T]$. We can easily check this convergence by arguments similar to those in (2.3)–(2.5).

Suppose that there are two solutions x and $\widetilde{x}$ to the SDE (2.1). Let $d = x - \widetilde{x}$. Again, by a similar argument as in (2.3)–(2.5), we have the following inequality

$$\|d_t\|_2^2 \le C \int_0^t \|d_s\|_2^2 ds.$$

Notice that

$$\|d_t\|_2^2 \le 2\|x_t\|_2^2 + 2\|\widetilde{x}_t\|_2^2 \le 4(t + t\|h\|_2^2) < \infty.$$

An application of Gronwall's inequality yields $\|d_t\|_2^2 \equiv 0$. □

While equations (2.1) can be solved separately for each fixed i the solutions $x^1, \dots, x^n$ are not independent since all of them depend on the common random environment W. It is easy to see that $(x^1, \dots, x^n)$ is an nd-dimensional Feller process governed by the generator

$$A^{(n)} f(y_1, \dots, y_n) = \frac{1}{2}(\Delta^{(n)} + B^{(n)}) f(y_1, \dots, y_n), \qquad (2.6)$$

where $\Delta^{(n)}$ is the Laplace operator in $\mathbb{R}^{nd}$,

$$B^{(n)}f(y_1,\dots,y_n) = \sum_{k_1,k_2=1}^{n}\sum_{i,j=1}^{d}\rho^{ij}(y_{k_1}-y_{k_2})\frac{\partial^2 f}{\partial y^i_{k_1}\partial y^j_{k_2}(y^1,\dots,y^n)}, \tag{2.7}$$

and $y_k = (y^1_k,\dots,y^d_k)\in\mathbb{R}^d$ for all $k=1,\dots,n$. This is similar to (1.19) of Wang [32] for the one-dimensional case.

2.2 The Mytnik-Sturm branching mechanism

In this section, we briefly construct the branching particle system. For further study of this branching mechanism, we refer the readers to Mytnik's and Sturm's papers (see [23, 30]).

We start this section by introducing some notation. For any integer $k\geq 0$, we denote by $C^k_b(\mathbb{R}^d)$ the space of k times continuously differentiable functions on $\mathbb{R}^d$ which are bounded together with their derivatives up to the order k. Also, $H^k_2(\mathbb{R}^d)$ is the Sobolev space of square integrable functions on $\mathbb{R}^d$ which have square integrable derivatives up to the order k. For any differentiable function ϕ on $\mathbb{R}^d$, we make use of the notation $\partial_{i_1\cdots i_m}\phi(x) = \frac{\partial^m}{\partial x_{i_1}\cdots\partial x_{i_m}}\phi(x)$.

We write $M_F(\mathbb{R}^d)$ for the space of finite measures on $\mathbb{R}^d$ equipped with the weak topology. We denote by $D([0,T],M_F(\mathbb{R}^d))$ the Skorokhod space of càdlàg functions on time interval $[0,T]$, taking values in $M_F(\mathbb{R}^d)$. For any $\phi\in C_b(\mathbb{R}^d)$ and $\mu\in M_F(\mathbb{R}^d)$, we write

$$\langle\mu,\phi\rangle = \mu(\phi) := \int_{\mathbb{R}^d}\phi(x)\mu(dx). \tag{2.8}$$

Let $\mathcal{I} := \{\alpha = (\alpha_0,\alpha_1,\dots,\alpha_N), \alpha_0\in\{1,2,3\dots\}, \alpha_i\in\{1,2\},\ \text{for}\ 1\leq i\leq N\}$ be a set of multi-indexes. In our model $\mathcal{I}$ is the index set of all possible particles. In other words, initially there are a finite number of particles and each particle generates at most 2 offspring. For any particle $\alpha = (\alpha_0,\alpha_1,\dots,\alpha_N)\in\mathcal{I}$, let $\alpha-1 = (\alpha_0\dots,\alpha_{N-1}), \alpha-2 = (\alpha_0,\dots,\alpha_{N-2}),\dots,\alpha-N = (\alpha_0)$ be the ancestors of α. Then, $|\alpha| = N$ is the number of the ancestors of the particle α. It is easy to see that the ancestors of any particle α are uniquely determined.

Fix a time interval $[0,T]$. Let $(\Omega,\mathcal{F},P)$ be a complete probability space, on which $\{B^\alpha_t, t\in[0,T]\}_{\alpha\in\mathcal{I}}$ are independent d-dimensional standard Brownian motions, and W is a d-dimensional space-time white Gaussian random field on $[0,T]\times\mathbb{R}^d$ independent of the family $\{B^\alpha\}$.

Let $x_t = x(x_0,B^\alpha,r,t)$, where $0\leq r\leq t\leq T$, be the unique solution to the following SDE:

$$x_t = x_0 + B^\alpha_t - B^\alpha_r + \int_r^t\int_{\mathbb{R}^d} h(y-x_s)W(ds,dy), \tag{2.9}$$

where $x_0\in\mathbb{R}^d$, $r\in[0,t)$ and h is a $d\times d$ matrix-valued function. We assume that h satisfies hypothesis **[H0]**.

For any $t\in[0,T]$, let $t_n = \frac{\lfloor nt\rfloor}{n}$ be the last branching time before t. For any $\alpha = (\alpha_0,\alpha_1,\dots,\alpha_N)$, if $nt_n = \lfloor nt\rfloor\leq N$, let $\alpha_t = (\alpha_0,\dots,\alpha_{\lfloor nt\rfloor})$ be the ancestor of α at time t. Suppose that each particle, which starts from the death place of its parent, moves in $\mathbb{R}^d$ following the motion described by the SDE (2.9) during its lifetime. Then, the path of any particle α and all its ancestors, denoted by $x^{\alpha,n}_t$, is given by

$$x^{\alpha,n}_t = x^{\alpha_t,n}_t = \begin{cases} x\left(x^n_{\alpha_0}, B^{(\alpha_0)}, 0, t\right), & 0\leq t<\frac{1}{n},\\ x\big(x^{\alpha_t-1,n}_{t_n^-}, B^{\alpha_t}, t_n, t\big), & \frac{1}{n}\leq t<\frac{N+1}{n},\\ \partial, & \text{otherwise.}\end{cases}$$

Here $x^n_{\alpha_0} \in \mathbb{R}^d$ is the initial position of particle (α_0), $x^{\alpha_t-1,n}_{t_n^-} := \lim_{s\uparrow t_n} x^{\alpha_t-1,n}_s$, and ∂ denotes the "cemetery"-state, that can be understood as a point at infinity.

Let $\xi = \{\xi(x), x \in \mathbb{R}^d\}$ be a real-valued random field on $\mathbb{R}^d$ with covariance

$$\mathbb{E}\big(\xi(x)\xi(y)\big) = \kappa(x,y), \tag{2.10}$$

for all $x, y \in \mathbb{R}^d$. Assume that ξ satisfies the following conditions:

[H1] (i) ξ is symmetric, that is $\mathbb{P}(\xi(x) > z) = \mathbb{P}(\xi(x) < -z)$ for all $x \in \mathbb{R}^d$ and $z \in \mathbb{R}$.

(ii) $\sup_{x\in\mathbb{R}^d} \mathbb{E}\big(|\xi(x)|^p\big) < \infty$ for some $p > 2$.

(iii) κ is continuous and bounded on $\mathbb{R}^d \times \mathbb{R}^d$.

For any $n \geq 1$, the random field ξ is used to define the offspring distribution after a scaling $\frac{1}{\sqrt{n}}$. In order to make the offspring distribution a probability measure, we introduce the truncation of the random field ξ, denoted by ξ^n, as follows:

$$\xi^n(x) = \begin{cases} \sqrt{n}, & \text{if } \xi(x) > \sqrt{n}, \\ -\sqrt{n}, & \text{if } \xi(x) < -\sqrt{n}, \\ \xi(x), & \text{otherwise.} \end{cases} \tag{2.11}$$

The correlation function of the truncated random field is then given by

$$\kappa_n(x,y) = \mathbb{E}\big(\xi^n(x)\xi^n(y)\big).$$

Let $(\xi^n_i)_{i\geq 0}$ be independent copies of ξ^n. Denote by ξ^{n+}_i and ξ^{n-}_i the positive and negative part of ξ^n_i respectively. Let $N^{\alpha,n} \in \{0,1,2\}$ be the offspring number of the particle α at the branching time $\frac{\lfloor\alpha\rfloor+1}{n}$. Assume that $\{N^{\alpha,n}, |\alpha| = i\}$ are conditionally independent given ξ^n_i and the position of α at its branching time, with a distribution given by

$$\begin{aligned} P\Big(N^{\alpha,n} = 2 | \, \xi^n_i, x^{\alpha,n}_{\frac{i+1}{n}-}\Big) &= \frac{1}{\sqrt{n}} \xi^{n+}_i\Big(x^{\alpha,n}_{\frac{i+1}{n}-}\Big), \\ P\Big(N^{\alpha,n} = 0 | \, \xi^n_i, x^{\alpha,n}_{\frac{i+1}{n}-}\Big) &= \frac{1}{\sqrt{n}} \xi^{n-}_i\Big(x^{\alpha,n}_{\frac{i+1}{n}-}\Big), \\ P\Big(N^{\alpha,n} = 1 | \, \xi^n_i, x^{\alpha,n}_{\frac{i+1}{n}-}\Big) &= 1 - \frac{1}{\sqrt{n}} |\xi^n_i| \Big(x^{\alpha,n}_{\frac{i+1}{n}-}\Big). \end{aligned}$$

For any particle $\alpha = (\alpha_0, \dots, \alpha_N)$, α is called to be alive at time t, denoted by $\alpha \sim_n t$, if the following conditions are satisfied:

(i) There are exactly N branching before or at t: $\lfloor nt \rfloor = N$.

(ii) α has an unbroken ancestors line: $\alpha_{N-i+1} \leq N^{\alpha-i,n}$, for all $i = 1, 2, \dots, N$.

[Introduction of $N^{\alpha,n}$ allows the particle α produce one more generation, namely, produce new particle $(\alpha, N^{\alpha,n})$. However, $(\alpha, 0)$ is considered no longer alive and will not produce offspring any more.] For any n, denote by $X^n = \{X^n_t, t \in [0,T]\}$ the empirical measure-valued process of the particle system. Then, X^n is a discrete measure-valued process, given by

$$X^n_t = \frac{1}{n} \sum_{\alpha\sim_n t} \delta_{x^{\alpha,n}_t}, \tag{2.12}$$

where δ_x is the Dirac measure at $x \in \mathbb{R}^d$, and the sum is among all alive particles at time $t \in [0,T]$. Then, for any $\phi \in C^2_b(\mathbb{R}^d)$, with the notation (2.8), we have

$$X^n_t(\phi) = \int_{\mathbb{R}^d} \phi(x) X^n_t(dx) = \frac{1}{n} \sum_{\alpha\sim n} \phi(x^{\alpha,n}_t).$$

3 Main results

Let $(\Omega, \mathcal{F}, \{F_t\}_{t\in[0,T]}, P)$ be a complete filtered probability space that satisfies the usual conditions. Suppose that W is a d-dimensional space-time white Gaussian random field on $[0,T]\times\mathbb{R}^d$, and V is a one-dimensional Gaussian random field on $[0,T]\times\mathbb{R}^d$ independent of W, that is time white and spatially colored with correlation κ defined in (2.10):

$$\mathbb{E}\big(V(t,x)V(s,y)\big) = (t\wedge s)\kappa(x,y),$$

for all $s,t\in[0,T]$ and $x,y\in\mathbb{R}^d$. Assume that $\{W(t,x), x\in\mathbb{R}^d\}$, $\{V(t,x), x\in\mathbb{R}^d\}$ are $\mathcal{F}_t$-measurable for all $t\in[0,T]$, and $\{W(t,x)-W(s,x), x\in\mathbb{R}^d\}$, $\{V(t,x)-V(s,x), x\in\mathbb{R}^d\}$ are independent of $\mathcal{F}_s$ for all $0\le s<t\le T$.

Denote by A^* the adjoint of A, where $A=A^{(1)}$ is the generator defined in (2.6). Consider the following SPDE:

$$\begin{aligned} u_t(x) =& \mu(x) + \int_0^t A^* u_s(x)ds - \sum_{i,j=1}^d \int_0^t \int_{\mathbb{R}^d} \frac{\partial}{\partial x_i}\left[h^{ij}(y-x)u_s(x)\right] W^j(ds,dy) \\ &+ \int_0^t u_s(x)\frac{V(ds,dx)}{dx}. \end{aligned} \tag{3.1}$$

Definition 3.1. *Let $u=\{u_t(x), t\in[0,T], x\in\mathbb{R}^d\}$ be a random field. Then,*

(i) u is said to be a strong solution to the SPDE (3.1), if u is jointly measurable on $[0,T]\times\mathbb{R}^d\times\Omega$, adapted to $\{\mathcal{F}_t\}_{t\in[0,T]}$ and for any $\phi\in C_b^2(\mathbb{R}^d)$, the following equation holds for every $t\in[0,T]$:

$$\begin{aligned} \int_{\mathbb{R}^d}\phi(x)u_t(x)dx =& \int_{\mathbb{R}^d}\phi(x)\mu(x)dx + \int_0^t\int_{\mathbb{R}^d} A\phi(x)u_s(x)dxds \\ &+ \int_0^t\int_{\mathbb{R}^d}\Big[\int_{\mathbb{R}^d}\nabla\phi(x)^* h(y-x)u_s(x)dx\Big]W(ds,dy) \\ &+ \int_0^t\int_{\mathbb{R}^d}\phi(x)u_s(x)V(ds,dx),\ a.s. \end{aligned} \tag{3.2}$$

where the last two stochastic integrals are Walsh's integral (see e.g. Walsh [31]).

The solution to (3.1) is said to be pathwise unique, if whenever u and $\widetilde{u}$ are two solutions to (3.1), then there exists a set $G\in\mathcal{F}$ of probability one, such that $u_t(\omega)=\widetilde{u}_t(\omega)$ for all $t\in[0,T]$ and $\omega\in G$.

(ii) u is said to be a weak solution to the SPDE (3.1), if there exists a filtered probability space, on which W and V are independent random fields that satisfy the above properties, such that u is a strong solution with this probability space.

Let $X^n=\{X_t^n, 0\le t\le T\}$ be defined by (2.12). In order to show the convergence of X^n in $D([0,T];M_F(\mathbb{R}^d))$, we make use of the following hypotheses on the initial measures X_0^n:

[H2] (i) $\sup_{n\ge1}|X_0^{(n)}(1)|<\infty$.

(ii) $X_0^n\Rightarrow X_0$ in $M_F(\mathbb{R}^d)$ as $n\to\infty$.

(iii) X_0 has a bounded density μ.

In Section 4 we prove the following two theorems.

Theorem 3.2. *Let X^n be defined in (2.12). Then, under hypotheses* **[H1]** *and* **[H2]***, we have the following results:*

(i) $X^n \Rightarrow X$ *in* $D([0,T], M_F(\mathbb{R}^d))$ *as* $n \to \infty$.

(ii) The limit $X = \{X_t, t \in [0,T]\}$ *is a continuous* $M_F(\mathbb{R}^d)$*-valued process. In addition, for almost all* $\omega \in \Omega$ *and every* $t \in [0,T]$*, as a finite measure on* $\mathbb{R}^d$*,* $X_t(\omega)$ *has a density* $u_t(x,\omega)$*.*

(iii) $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ *is a weak solution to the SPDE (3.1) in the sense of Definition 3.1.*

Theorem 3.3. *Assume the hypotheses* **[H1]** *and* **[H2]** *(iii). The SPDE (3.1) has a jointly continuous strong solution, which is pathwise unique in the space of jointly continuous solutions in the sense of Definition 3.1.*

The last main result in this paper is the following theorem concerning the Hölder continuity of the solution to the SPDE (3.1).

Theorem 3.4. *Let* $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ *be the strong solution to the SPDE (3.1) in the sense of Definition 3.1. Then, for any* $\beta_1, \beta_2 \in (0,1)$ *and* $p > 1$*, there exists a constant* C *that depends on* T*,* d*,* h*,* p*,* β_1*, and* β_2*, such that for all* $x, y \in \mathbb{R}^d$ *and* $0 < s < t \leq T$

$$\|u_t(x) - u_s(y)\|_{2p} \leq C s^{-\frac{1}{2}} \left(|x-y|^{\beta_1} + |t-s|^{\frac{1}{2}\beta_2}\right).$$

Hence by Kolmogorov's criteria, $u_t(x)$ *is almost surely jointly Hölder continuous on* $(0,T] \times \mathbb{R}^d$*, with exponent* $\beta_1 \in (0,1)$ *in space and* $\beta_2 \in (0, \frac{1}{2})$ *in time.*

4 Proof of Theorems 3.2 and 3.3

We prove Theorems 3.2 and 3.3 in the following steps:

(i) In Section 4.1, we show that $\{X^n\}_{n\geq 1}$ is a tight sequence in $D([0,T]; M_F(\mathbb{R}^d))$, and the limit of any convergent subsequence in law solves a martingale problem.

(ii) In Section 4.2, we show that any solution to the martingale problem has a density almost surely.

(iii) In Section 4.3, we show the equivalence between martingale problem (see e.g. (4.4)–(4.5) below) and the SPDE (3.1). Finally, we prove Theorems 3.2 and 3.3.

4.1 Tightness and martingale problem

Recall the empirical measure-valued process $X^n = \{X_t^n, t \in [0,T]\}$ given by (2.12). Let $A = A^{(1)}$ be the generator of one particle motion defined in (2.6). For any $\phi \in C_b^2(\mathbb{R}^d)$, similar to equality (49) of Sturm [30], we can decompose X_t^n as follows:

$$X_t^n(\phi) = X_0^n(\phi) + Z_t^n(\phi) + M_t^{b,n}(\phi) + B_t^n(\phi) + U_t^n(\phi), \tag{4.1}$$

where

$$Z_t^n(\phi) = \int_0^t X_u^n(A\phi)du,$$

$$M_t^{b,n}(\phi) = M_{t_n}^{b,n}(\phi) = \frac{1}{n} \sum_{s_n < t_n} \sum_{\alpha \sim_n s_n} \phi\big(x_{s_n+\frac{1}{n}}^{\alpha,n}\big)(N^{\alpha,n} - 1),$$

$$B_t^n(\phi) = \frac{1}{n}\Big(\sum_{s_n < t_n} \sum_{\alpha \sim_n s_n} \int_{s_n}^{s_n+\frac{1}{n}} \nabla\phi(x_u^{\alpha,n})^* dB_u^\alpha + \sum_{\alpha \sim_n t} \int_{t_n}^t \nabla\phi(x_u^{\alpha,n})^* dB_u^\alpha \Big),$$

and

$$U_t^n(\phi) = \frac{1}{n}\Big(\sum_{s_n<t_n} \sum_{\alpha\sim_n s_n} \int_{s_n}^{s_n+\frac{1}{n}} \int_{\mathbb{R}^d} \nabla\phi(x_u^{\alpha,n})^* h(y - x_u^{\alpha,n}) W(du,dy) \\ + \sum_{\alpha\sim_n t} \int_{t_n}^{t} \int_{\mathbb{R}^d} \nabla\phi(x_u^{\alpha,n})^* h(y - x_u^{\alpha,n}) W(du,dy)\Big).$$

Notice that

$$\begin{aligned} &\mathbb{E} \sum_{\alpha\sim_n s_n} \int_{s_n}^{s_n+\frac{1}{n}} \int_{\mathbb{R}^d} \Big|\nabla\phi(x_u^{\alpha,n})^* h(y - x_u^{\alpha,n})\Big|^2 dydu \\ \le& \sum_{|\alpha|=\lfloor sn\rfloor} \mathbb{E} \int_{s_n}^{s_n+\frac{1}{n}} \int_{\mathbb{R}^d} \Big|\nabla\phi(x_u^{\alpha,n})^* h(y - x_u^{\alpha,n})\Big|^d ydu \\ \le& 2^{nT} n^{-1} N_0 \|\phi\|_{1,\infty}\|h\|_2 < \infty, \end{aligned}$$

where N_0 denotes the number of initial particles, that is a finite integer. Therefore, by the stochastic Fubini theorem (see, e.g., Lemma 4.1 on page 116 of Ikeda and Watanabe [15]), we can write:

$$U_t^n(\phi) = \int_0^t \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \nabla\phi(x)^* h(y-x) X_u(dx)\Big) W(du,dy).$$

As in Sturm [30], consider the natural filtration, generated by the process X^n

$$\mathcal{F}_t^n = \sigma\left(\{x^{\alpha,n}, N^{\alpha,n} \big| |\alpha| < \lfloor nt\rfloor\} \cup \{x_s^{\alpha,n}, s\le t, |\alpha| = \lfloor nt\rfloor\}\right),$$

and a discrete filtration at branching times

$$\widetilde{\mathcal{F}}_{t_n}^n = \sigma\big(\mathcal{F}_{t_n}^n \cup \{x^{\alpha,n} \big| |\alpha| = nt_n\}\big) = \mathcal{F}_{(t_n+n^{-1})-}^n.$$

Then, $B_t^n(\phi)$ and $U_t^n(\phi)$ are continuous $\mathcal{F}_t^n$-martingales, while $M_t^{b,n}(\phi)$ is a discrete $\widetilde{\mathcal{F}}_{t_n}^n$-martingale.

Lemma 4.1. *Assume hypotheses **[H0]**, **[H1]**, **[H2]** (i) and (ii). Let $p > 2$ be given in hypothesis **[H1]**. Then, for all $\phi \in C_b^2(\mathbb{R}^d)$,*

(i) $\mathbb{E}\Big(\sup_{0\le t\le T} |X_t^n(\phi)|^p\Big)$, $\mathbb{E}\Big(\sup_{0\le t\le T} |M_t^{b,n}(\phi)|^p\Big)$ *and* $\mathbb{E}\Big(\sup_{0\le t\le T} |U_t^n(\phi)|^p\Big)$ *are bounded uniformly in* $n\ge 1$.

(ii) $\mathbb{E}\Big(\sup_{0\le t\le T} |B_t^n(\phi)|^p\Big) \to 0$, *as* $n\to\infty$.

Proof. (i) By the same argument as that for Lemma 3.1 of Sturm [30], we can show that

$$\mathbb{E}\Big(\sup_{0\le t\le T} |M_t^{b,n}(1)|^p\Big) \le C \int_0^T \mathbb{E}\Big(\sup_{0\le s\le t} |X_s^n(1)|^p\Big) dt$$

where the constant $C > 0$ does not depend on n. Again similarly as Strum did for (58) of [30], we can also deduce the following inequality

$$\begin{aligned} \mathbb{E}\Big(\sup_{0\le t\le T} |X_t^n(1)|^p\Big) \le& C\Big(1 + \mathbb{E}\Big(\sup_{0\le t\le T} |M_{t_n}^{b,n}(1)|^p\Big)\Big) \\ \le& C_1 + C_2 \int_0^T \mathbb{E}\Big(\sup_{0\le s\le t} |X_s^n(1)|^p\Big) dt, \end{aligned}$$

where C_1, C_2 are constants independent of n. Notice that

$$\sup_{0\le t\le T} |X_t^n(1)| \le 2^{nT}\frac{N_0^n}{n},$$

that is bounded for fixed n. Then, it follows from Gronwall's inequality that the sequence $\Big\{\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(1)|^p\Big)\Big\}_{n\ge1}$ is uniformly bounded in n.

The uniform boundedness of $\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(\phi)|^p\Big)$ and $\mathbb{E}\Big(\sup_{0\le t\le T}|M_t^{b,n}(\phi)|^p\Big)$ follows immediately.

We estimate $U_t^n(\phi)$ as follows:

$$\begin{aligned}
\langle U^n(\phi)\rangle_t =& \Big\langle \int_0^{\cdot}\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}\nabla\phi(x)^* h(y-x)X_u^n(dx)\Big)W(du,dy)\Big\rangle_t \\
=& \sum_{j=1}^d \int_0^t\int_{\mathbb{R}^d}\Big(\sum_{i=1}^d\int_{\mathbb{R}^d}\partial_i\phi(x)h^{ij}(y-x)X_u^n(dx)\Big)^2 dydu \\
=& \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}\nabla\phi(x)^*\rho(x-z)\nabla\phi(z)X_u^n(dx)X_u^n(dz)du \\
\le& \|\rho\|_\infty\|\phi\|_{1,\infty}^2\int_0^T |X_u^n(1)|^2\,du. \qquad (4.2)
\end{aligned}$$

Thus by (4.2), Burkholder-Davis-Gundy's and Jensen's inequalities, we have

$$\begin{aligned}
\mathbb{E}\Big(\sup_{0\le t\le T}|U_t^n(\phi)|^p\Big) \le& c_p\mathbb{E}\big(\langle U^n(\phi)\rangle_T^{\frac{p}{2}}\big) \le c_p\|\rho\|_\infty^{\frac{p}{2}}\|\phi\|_{1,\infty}^p T^{\frac{p}{2}-1}\mathbb{E}\Big(\int_0^T|X_u^n(1)|^p du\Big) \\
\le& c_p\|\rho\|_\infty^{\frac{p}{2}}\|\phi\|_{1,\infty}^p T^{\frac{p}{2}}\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(1)|^p\Big), \qquad (4.3)
\end{aligned}$$

that is also uniformly bounded in n.

(ii) Note that $\{B^\alpha\}$ are independent Brownian motions. Then, by Burkholder-Davis-Gundy's inequality, we have

$$\begin{aligned}
\mathbb{E}\Big(\sup_{0\le t\le T}|B_t^n(\phi)|^2\Big) \le& \frac{c_2}{n^2}\Big[\sum_{s_n<T_n}\sum_{\alpha\sim_n s_n}\mathbb{E}\Big(\int_{s_n}^{s_n+\frac{1}{n}}|\nabla\phi(x_u^{\alpha,n})|^2du\Big) \\
&+\sum_{\alpha\sim_n t}\mathbb{E}\Big(\int_{T_n}^T|\nabla\phi(x_u^{\alpha,n})|^2du\Big)\Big] \\
=& \frac{c_2}{n}\mathbb{E}\Big(\int_0^T\int_{\mathbb{R}^d}|\nabla\phi(x)|^2X_u(dx)du\Big) \le \frac{c_2}{n}\|\phi\|_{1,\infty}^2 T\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(1)|^p\Big)\to 0,
\end{aligned}$$

because $\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(1)|^p\Big)$ is uniformly bounded in n. □

As a consequence of Lemma 4.1, the collection

$$\Big\{\sup_{0\le t\le T}|X_t^n(\phi)|^2,\ \sup_{0\le t\le T}|M_t^{b,n}(\phi)|^2,\ \sup_{0\le t\le T}|U_t^n(\phi)|^2\Big\}_{n\ge1}$$

is uniformly integrable.

Definition 4.2. *Let $\{X^\alpha\}$ be a collection of real-valued stochastic processes. A family of stochastic processes $\{X^\alpha\}$ is said to be C-tight, if it is tight, and the limit of any subsequence is continuous.*

Lemma 4.3. *Assume hypotheses **[H0]**, **[H1]**, **[H2]** (i) and (ii). For all $\phi \in C_b^2(\mathbb{R}^d)$, $M^{b,n}(\phi)$, $Z^n(\phi)$, and $X^n(\phi)$ and $U^n(\phi)$ are C-tight sequences in $D([0,T],\mathbb{R})$.*

Proof. By an argument that used by Sturm in the proof of Lemma 3.6 in [30], we can deduce the C-tightness of $M^{b,n}(\phi)$ and $Z^n(\phi)$.

We prove the tightness of $X_t^n(\phi)$ by checking Aldous's conditions (see e.g. Theorem 4.5.4 of Dawson [5]). By Chebyshev's inequality, for any fixed $t \in [0,T]$, and $N > 0$, we have

$$\mathbb{P}\left(|X_t^n(\phi)| > N\right) \leq \frac{1}{N^p}\mathbb{E}\big(|X_t^n(\phi)|^p\big) \leq \frac{1}{N^p}\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^n(\phi)|^p\Big) \to 0,$$

uniformly in n as $N \to \infty$ by Lemma 4.1 (i).

On the other hand, for any $n \geq 1$, we extend X^n to the time interval $[0, T_n + \frac{1}{n}]$ in such a way that X^n performs the same diffusion/branching mechanism as before on $[T, T_n + \frac{1}{n}]$. Denote by $\widetilde{X}^n = \{\widetilde{X}^n(t), t \in [0, T_n + \frac{1}{n}]\}$ the extended process. Then, by Theorem 10.13 of Dynkin [10], we know that $\widetilde{X}^n$ is a strong Markov process on $[0, T_n + \frac{1}{n}]$.

Let $\{\tau_n\}_{n\geq 1}$ be any collection of stopping times bounded by T and let $\{\delta_n\}_{n\geq1}$ be any positive sequence that decreases to 0, such that $\tau_n + \delta_n \leq T$. Then, due to the uniform boundedness of $\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^n(\phi)|^p\Big)$ and the strong Markov property of $\widetilde{X}^n$, we have

$$\begin{aligned}\mathbb{P}\Big(\big|X^n_{\tau_n+\delta_n}(\phi) - X^n_{\tau_n}(\phi)\big| > \epsilon\Big) =&\mathbb{P}\left(\Big|\widetilde{X}^n_{\tau_n+\delta_n}(\phi) - \widetilde{X}^n_{\tau_n}(\phi)\Big| > \epsilon\right)\\ =&\mathbb{P}\left(\Big|\widetilde{X}^n_{\delta_n}(\phi) - \widetilde{X}^n_{0}(\phi)\Big| > \epsilon\right) \leq \frac{1}{\epsilon^p}\mathbb{E}\Big(\Big|\widetilde{X}^n_{\delta_n}(\phi) - \widetilde{X}^n_{0}(\phi)\Big|^p\Big)\\ \leq&\frac{\delta_n^{\frac{p}{2}}}{\epsilon^p}c_p\|\rho\|_\infty^{\frac{p}{2}}\|\phi\|_\infty^p\Big[\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^n(1)|^p\Big) + \mathbb{E}\Big(|X_0^n(1)|^p\Big)\Big]\\ \to&0,\end{aligned}$$

as $n \to 0$. Thus both of Aldous's conditions are satisfied, and then it follows that $X_t^n(\phi)$ is tight in $D([0,T],\mathbb{R})$.

Recall the decomposition formula (4.1):

$$X_t^n(\phi) = X_0^n(\phi) + Z_t^n(\phi) + M_t^{b,n}(\phi) + B_t^n(\phi) + U_t^n(\phi).$$

Notice that $X^n(\phi)$, $Z^n(\phi)$, $M^{b,n}(\phi)$ are tight sequences as proved just above, $X_0^n(\phi)$ converges weakly by assumption, and $B_t^n(\phi)$ converges 0 in $L^2(\Omega)$ uniformly for all $t \in [0,T]$ by Lemma 4.1 (ii). As a consequence, $U^n(\phi)$ is tight in $D([0,T],\mathbb{R})$. In addition, by Proposition VI.3.26 of Jacod and Shiryaev [16], every limit of a tight sequence of continuous process $U^n(\phi)$ is continuous. It follows that $U^n(\phi)$ and $X^n(\phi)$ are C-tight sequences in $D([0,T];\mathbb{R})$. □

Let $\mathscr{S} = \mathscr{S}(\mathbb{R}^d)$ be the Schwartz space on $\mathbb{R}^d$, and let $\mathscr{S}'$ be the dual of $\mathscr{S}$. Then, we have the following lemma.

Lemma 4.4. *Assume hypotheses **[H0]**, **[H1]** and **[H2]** (i), (ii). Then,*

(i) $\{X^n\}_{n\geq1}$ *is a C-tight sequence in* $D([0,T];M_F(\mathbb{R}^d))$.

(ii) $\{B^n\}_{n\geq1}$, $\{M^{b,n}\}_{n\geq1}$, *and* $\{U^n\}_{n\geq1}$ *are C-tight in* $D([0,T];\mathscr{S}')$.

Proof. Let $\widehat{\mathbb{R}}^d = \mathbb{R}^d \cup \{\partial\}$ be the one point compactification of $\mathbb{R}^d$. Then, by Theorem 4.6.1 of Dawson [5] and Lemma 4.3, $\{X^n\}_{n\geq1}$ is a tight sequence in $D([0,T];M_F(\widehat{\mathbb{R}}^d))$.

On the other hand, by the same argument as in Lemma 3.9 of Sturm [30], we can show that any limit of a weakly convergent subsequence X^{n_k} in $D([0,T];M_F(\widehat{\mathbb{R}}^d))$ belongs to

$C([0,T];M_F(\mathbb{R}^d))$, the space of continuous $M_F(\mathbb{R}^d)$-valued functions on $[0,T]$. Therefore, $\{X^n\}_{n\geq 1}$ is a C-tight sequence in $D([0,T];M_F(\mathbb{R}^d))$.

To show property (ii), notice that $\mathscr{S} \subset C_b^2(\mathbb{R}^d)$. Then, by Theorem 4.1 of Mitoma [22], $\{B^n\}_{n\geq 1}$, $\{M^{b,n}\}_{n\geq 1}$, and $\{U^n\}_{n\geq 1}$ are C-tight in $D([0,T];\mathscr{S}')$. □

Proposition 4.5. *Assume hypotheses* ***[H0]****,* ***[H1]****,* ***[H2]*** *(i) and (ii). Let X be the limit of a weakly convergent subsequence $\{X^{n_k}\}_{k\geq 1}$ in $D([0,T];M_F(\mathbb{R}^d))$. Then, X is a solution to the following martingale problem: for any $\phi \in C_b^2(\mathbb{R}^d)$, the process $M(\phi) = \{M_t(\phi) : 0 \leq t \leq T\}$, given by*

$$M_t(\phi) := X_t(\phi) - X_0(\phi) - \int_0^t X_s(A\phi)ds, \tag{4.4}$$

is a continuous and square integrable $\mathcal{F}_t^X$-adapted martingale with quadratic variation:

$$\begin{aligned}\langle M(\phi)\rangle_t = &\int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \nabla\phi(x)^*\rho(x-y)\nabla\phi(y)X_s(dx)X_s(dy)ds\\ &+\int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \kappa(x,y)\phi(x)\phi(y)X_s(dx)X_s(dy)ds.\end{aligned} \tag{4.5}$$

Proof. Let $\{X^{n_k}\}_{k\geq 1}$ be a weakly convergent subsequence in $D([0,T];M_F(\mathbb{R}^d))$. By taking further subsequences, we can assume, in view of Lemma 4.4 (ii), that $\{B^{n_k}\}_{k\geq 1}$, $\{M^{b,n_k}\}_{k\geq 1}$, and $\{U^{n_k}\}_{k\geq 1}$ are weakly convergent subsequences in $D([0,T];\mathscr{S}')$.

Therefore, by Skorokhod's representation theorem, there exists a probability space $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}})$, on which $(\widetilde{X}^{n_k},\widetilde{M}^{b,n_k},\widetilde{B}^{n_k},\widetilde{U}^{n_k})$ has the same joint distribution as $(X^{n_k},M^{b,n_k},B^{n_k},U^{n_k})$ for all $k\geq 1$, and converge a.s. to $(\widetilde{X},\widetilde{M}^b,\widetilde{B},\widetilde{U})$ in the product space $D([0,T],M_F(\widehat{\mathbb{R}}^d))\times D([0,T],\mathscr{S}')^3$.

Then, for any $\phi \in \mathscr{S}'$, $(\widetilde{X}^{n_k}(\phi),\widetilde{M}^{b,n_k}(\phi),\widetilde{B}^{n_k}(\phi),\widetilde{U}^{n_k}(\phi))$ converges a.s. in $D([0,T],\mathbb{R})^4$. Since $\left\{\sup_{0\leq t\leq T}|X_t^n(\phi)|^2, \sup_{0\leq t\leq T}|M_t^{b,n}(\phi)|^2, \sup_{0\leq t\leq T}|U_t^n(\phi)|^2\right\}_{n\geq 1}$ is uniformly integrable, the convergence is also in $L^2([0,T]\times\Omega)$.

For any $t\in[0,T]$, let

$$\widetilde{M}_t^{n_k}(\phi) := \widetilde{X}_t^{n_k}(\phi) - \widetilde{X}_0^{n_k}(\phi) - \int_0^t \widetilde{X}_s^{n_k}(A\phi)ds = \widetilde{M}_t^{b,n_k}(\phi) + \widetilde{B}_t^{n_k}(\phi) + \widetilde{U}_t^{n_k}(\phi).$$

Then, it converges to a continuous and square integrable martingale $\widetilde{M}(\phi) = \widetilde{M}^b(\phi) + \widetilde{U}(\phi)$ in $L^2(\widetilde{\Omega})$ with respect to its natural filtration.

The next step is to compute the quadratic variation of $\widetilde{M}(\phi)$. Notice that W and $\{B^\alpha\}$ are independent, then U^n and B^n are orthogonal. As a consequence, $\widetilde{U}^{n_k}$ and $\widetilde{B}^{n_k}$ are also orthogonal. On the other hand, $\widetilde{M}^{b,n}(\phi)$ is a pure jump martingale, while $\widetilde{U}^{n_k}(\phi)$ and $\widetilde{B}^{n_k}(\phi)$ are continuous martingales. Due to Theorem 43 on page 353 of Dellacherie and Meyer [9], $\widetilde{M}^{b,n}(\phi)$, $\widetilde{B}^{n_k}(\phi)$ and $\widetilde{U}^{n_k}(\phi)$ are mutually orthogonal. By the same argument as in Lemma 4.1, we can show that $\langle\widetilde{M}^{b,n_k}(\phi)+\widetilde{B}^{b,n_k}(\phi)+\widetilde{U}^{n_k}(\phi)\rangle_t = \langle\widetilde{M}^{b,n_k}(\phi)\rangle_t + \langle\widetilde{B}^{b,n_k}(\phi)\rangle_t + \langle\widetilde{U}^{n_k}(\phi)\rangle_t$ are uniformly integrable. Then, by Theorem II.4.5 of Perkins [26], we have

$$\begin{aligned}\langle\widetilde{M}^{b,n_k}(\phi)+\widetilde{B}^{b,n_k}(\phi)+\widetilde{U}^{n_k}(\phi)\rangle_t &= \langle\widetilde{M}^{b,n_k}(\phi)\rangle_t + \langle\widetilde{B}^{b,n_k}(\phi)\rangle_t + \langle\widetilde{U}^{n_k}(\phi)\rangle_t\\ &\to \langle\widetilde{M}^b(\phi)\rangle_t + \langle\widetilde{U}(\phi)\rangle_t = \langle\widetilde{M}(\phi)\rangle_t\end{aligned}$$

as $k\to\infty$ in $D([0,T],\mathbb{R})$ in probability.

On the other hand, by the same argument of Lemma 3.8 of Sturm [30], we have

$$\langle\widetilde{M}^b(\phi)\rangle_t = \lim_{k\to\infty}\langle\widetilde{M}^{b,n_k}(\phi)\rangle_t = \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}\kappa(x,y)\phi(x)\phi(y)\widetilde{X}_s(dx)\widetilde{X}_s(dy)ds, \ a.s.$$

For $\langle \widetilde{U}(\phi)\rangle_t$, by (4.2), since $\widetilde{X}^{n_k}(\phi) \to \widetilde{X}(\phi)$ in $L^2([0,T]\times\Omega)$, it follows that

$$\lim_{k\to\infty} \langle \widetilde{U}^{n_k}(\phi)\rangle_t = \lim_{k\to\infty} \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \nabla\phi(x)^*\rho(x-z)\nabla\phi(z)\widetilde{X}^n_u(dx)\widetilde{X}^n_u(dz)du$$
$$= \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \nabla\phi(x)^*\rho(x-z)\nabla\phi(z)\widetilde{X}_u(dx)\widetilde{X}_u(dz)du.$$

As a consequence, $\widetilde{M} = \{\widetilde{M}_t, t\in[0,T]\}$, where

$$\widetilde{M}_t(\phi) = \widetilde{X}_t(\phi) - \widetilde{X}_0(\phi) - \int_0^t \widetilde{X}_s(A\phi)ds = \widetilde{M}^b_t(\phi) + \widetilde{B}_t(\phi) + \widetilde{U}_t(\phi),$$

is a continuously square integrable martingale with the quadratic variation given by the expression (4.5).

Finally, by the same argument as in Theorem II in Section 4.2 of Perkins [26], we can show $\widetilde{M}(\phi)$ is an $\mathcal{F}^{\widetilde{X}}$-adapted martingale. □

4.2 Absolute continuity

Assume hypotheses **[H0]** and **[H1]**. Let X_t be a solution to the martingale problem (4.4)–(4.5). In this section, we show that for almost every $t\in[0,T]$, as an $M_F(\mathbb{R}^d)$-valued random variable, X_t has a density almost surely.

For any $n\geq 1$, $f\in C_b^2(\mathbb{R}^{nd})$, and $\mu\in M_F(\mathbb{R}^d)$, we define

$$\mu^{\otimes n}(f) := \int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d} f(x_1,\dots,x_n)\mu(dx_1)\cdots\mu(dx_n).$$

We derive the moment formula $\mathbb{E}(X_t^{\otimes n}(f))$ of the process X. In the one-dimensional Dawson-Watanabe branching case, Skoulakis and Adler [27] obtained the formula by computing the limit of particle approximations. An alternative approach by Xiong [34] consists in differentiating a conditional stochastic log-Laplace equation. In the present paper we use the techniques of moment duality to derive the moment formula. It can be also formulated by computing the limit of particle approximations.

For any integers $n\geq 2$ and $k\leq n$, we make use of the notation $x_k = (x_k^1,\dots,x_k^d)\in\mathbb{R}^d$ and $x=(x_1,\dots,x_n)\in\mathbb{R}^{nd}$. Let $\Phi_{ij}^{(n)} : C_b^2(\mathbb{R}^{nd}) \to C_b^2(\mathbb{R}^{nd})$, and $F^{(n)}, G^{(n)} : C_b^2(\mathbb{R}^{nd})\times M_F(\mathbb{R}^d)\to\mathbb{R}$ be given by

$$(\Phi_{ij}^{(n)}f)(x_1,\dots,x_n) := \kappa(x_i,x_j)f(x_1,\dots,x_n), \quad i,j\in\{1,2,\dots,n\},$$
$$F^{(n)}(f,\mu) := \mu^{\otimes n}(f),$$

and

$$G^{(n)}(f,\mu) := \mu^{\otimes n}(A^{(n)}f) + \frac{1}{2}\sum_{\substack{1\,\leq\, i,\,j\,\leq\, n\\ i\,\neq\, j}} \mu^{\otimes n}(\Phi_{ij}^{(n)}f),$$

where $\kappa\in C_b^2(\mathbb{R}^{2d})$ is the correlation of the random field ξ given by (2.10), and $A^{(n)}$ is the generator of n-particle motion defined in (2.6).

Lemma 4.6. *Let X_t be a solution to the martingale problem (4.4)–(4.5). Then, for any $n\geq 2$ and $f\in C_b^2(\mathbb{R}^{nd})$, the following process*

$$F^{(n)}(f,X_t) - \int_0^t G^{(n)}(f,X_s)ds$$

is a martingale.

Proof. See Lemma 1.3.2 of Xiong [35]. □

Let $\{T_t^{(n)}\}_{t\geq 0}$ be the semigroup generated by $A^{(n)}$, that is, $T_t^{(n)} : C_b^2(\mathbb{R}^{nd}) \to C_b^2(\mathbb{R}^{nd})$, given by

$$T_t^{(n)} f(x_1, \dots, x_n) = \int_{\mathbb{R}^{nd}} p(t, (x_1, \dots, x_n), (y_1, \dots, y_n)) f(y_1, \dots, y_n) dy_1 \dots dy_n,$$

where p is the transition density of n-particle-motion.

Let $\{S_k^{(n)}\}_{k\geq 1}$ be i.i.d. uniformly distributed random variables taking values in the set $\{\Phi_{ij}, 1 \leq i, j \leq n, i \neq j\}$. Let $\{\tau_k\}_{k\geq 1}$ be i.i.d exponential random variables independent of $\{S_k^{(n)}\}_{k\geq 1}$, with rate $\lambda_n = \frac{1}{2}n(n-1)$. Let $\eta_0 \equiv 0$, and $\eta_j = \sum_{i=1}^{j} \tau_i$ for all $j \geq 1$. For any $f \in C_b^2(\mathbb{R}^{nd})$, we define a $C_b^2(\mathbb{R}^{nd})$-valued random process $Y^{(n)} = \{Y_t^{(n)}, 0 \leq t \leq T\}$ as follows: for any $j \geq 0$ and $t \in [\eta_j, \eta_{j+1})$,

$$Y_t^{(n)} := T_{t-\eta_j}^{(n)} S_j^{(n)} T_{\tau_j}^{(n)} \cdots S_2^{(n)} T_{\tau_2}^{(n)} S_1^{(n)} T_{\tau_1}^{(n)} f. \tag{4.6}$$

Then, $Y^{(n)}$ is a Markov process with $Y_0^{(n)} = f$. It involves countable many i.i.d. jumps $S_k^{(n)}$, controlled by i.i.d. exponential clocks τ_k. In between jumps, the process evolves deterministically by the continuous semigroup $T_t^{(n)}$. Notice that the exponential clock is memoryless, and the semigroup $T_t^{(n)}$ is generated by a time homogeneous Markov process. Therefore, $Y^{(n)}$ is also time homogeneous.

Lemma 4.7. *For any $n \geq 2$ and $f \in C_b^2(\mathbb{R}^{nd})$, let $Y_t^{(n)}$ be defined in (4.6). Then*

$$\mathbb{E}\Big(\sup_{x\in\mathbb{R}^{nd}} \big|Y_t^{(n)}(x)\big|\Big) \leq \|f\|_\infty \exp\left(\|\kappa\|_\infty \lambda_n t\right). \tag{4.7}$$

Proof. Since $T_t^{(n)}$ is the semigroup generated by a Markov process, for any $t > 0$ and $f \in C_b^2(\mathbb{R}^{nd})$, $\|T_t^{(n)} f\|_\infty \leq \|f\|_\infty$. By definition of jump operators $\{S_j^{(n)}\}_{j\geq 1}$, it is easy to see that $\|S_j^{(n)} f\|_\infty \leq \|\kappa\|_\infty \|f\|_\infty$. Thus we have

$$\mathbb{E}\Big(\sup_{x\in\mathbb{R}^{nd}} \big|Y_t^{(n)}(x)\big|\Big) \leq \|f\|_\infty \sum_{j=0}^{\infty} \big[\|\kappa\|_\infty^j \mathbb{P}(\eta_j < t)\big]. \tag{4.8}$$

Notice that η_j is the sum of i.i.d. exponential random variables. Then, we have

$$\mathbb{P}(\eta_j < t) = 1 - \exp\left(-\lambda_n t\right) \sum_{k=0}^{j-1} \frac{(\lambda_n t)^k}{k!} = \exp(\lambda_n(t'-t)) \frac{(\lambda_n t)^j}{j!}, \tag{4.9}$$

for some $t' \in (0, t)$. Therefore, (4.7) follows from (4.8) and (4.9). □

Let $H^{(n)} : C_b^2(\mathbb{R}^{nd}) \times M_F(R^d) \to \mathbb{R}$ be given by

$$H^{(n)}(f, \mu) := G^{(n)}(f, \mu) - \lambda_n F^{(n)}(f, \mu).$$

Lemma 4.8. *Let $n \geq 2$ and $\mu \in M_F(\mathbb{R}^d)$. Then, the process*

$$F^{(n)}(Y_t^{(n)}, \mu) - \int_0^t H^{(n)}(Y_s^{(n)}, \mu) ds \tag{4.10}$$

is a martingale.

Proof. Let $\mu^{(n)}$ be any finite measure on $\mathbb{R}^{nd}$. Then, we have

$$\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\big) = \mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\mathbf{1}_{\{\tau_1>t\}}\big) + \mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\mathbf{1}_{\{\eta_1\le t<\eta_2\}}\big) + o(t). \tag{4.11}$$

For the first term, we have

$$\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\mathbf{1}_{\{\tau_1>t\}}\big) =\mu^{(n)}(T_t^{(n)}f)\mathbb{P}(\tau_1>t) = \mu^{(n)}(T_t^{(n)}f)\exp(-\lambda_n t)). \tag{4.12}$$

For the second term, since τ_1, τ_2 are independent, then for any $0\le s\le t$, we have

$$\mathbb{P}(\tau_1+\tau_2>t, \tau_1\le s) = \int_0^s\int_{t-s_1}^\infty \lambda_n^2\exp(-\lambda_n(s_1+s_2))ds_2ds_1 = \lambda_n s e^{-\lambda_n t}. \tag{4.13}$$

Note that by Lemma 4.7, $|Y_\cdot^{(n)}|$ is integrable on $[0,T]\times\mathbb{R}^{nd}\times\Omega$ with respect to the product measure $dt\times\mu^{(n)}(dx)\times P(d\omega)$. Then, by (4.13), Fubini's theorem, and the mean value theorem, we have

$$\begin{aligned}
&\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\mathbf{1}_{\{\eta_1\le t<\eta_2\}}\big)\\
=&\frac{1}{2}\sum_{\substack{1\le i,j\le n\\ i\ne j}}\int_0^t\int_{\mathbb{R}^{nd}}\big(T_{t-s}^{(n)}\Phi_{ij}^{(n)}T_s^{(n)}f\big)(x)\exp\left(-\lambda_n t\right)\mu^{(n)}(dx)ds\\
=&\frac{t}{2}\exp\left(-\lambda_n t\right)\sum_{\substack{1\le i,j\le n\\ i\ne j}}\int_{\mathbb{R}^{nd}}\big(T_{t-t'}^{(n)}\Phi_{ij}^{(n)}T_{t'}^{(n)}f\big)(x)\mu^{(n)}(dx),
\end{aligned} \tag{4.14}$$

for some $t'\in(0,t)$. Combining (4.11), (4.12), and (4.14), we have

$$\lim_{t\downarrow 0}\frac{\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\big)-\mu^{(n)}(f)}{t} = \mu^{(n)}(A^{(n)}f) + \frac{1}{2}\sum_{\substack{1\le i,j\le n\\ i\ne j}}\mu^{(n)}\big(\Phi_{ij}^{(n)}f-f\big).$$

By Proposition 4.1.7 of Ethier and Kurtz [11], the following process:

$$\mu^{(n)}(Y_t^{(n)}) - \int_0^t\Big[\mu^{(n)}(A^{(n)}Y_s^{(n)}) + \frac{1}{2}\sum_{\substack{1\le i,j\le n\\ i\ne j}}\mu^{(n)}(\Phi_{ij}^{(n)}Y_s^{(n)}-Y_s^{(n)})\Big]ds, \tag{4.15}$$

is a martingale. Then, the lemma follows by choosing $\mu^{(n)}=\mu^{\otimes n}$. □

By Lemma 4.6, 4.8 and Corollary 3.2 of Dawson and Kurtz [7], we have the following moment equality:

$$\mathbb{E}\big(X_t^{\otimes n}(f)\big) = \mathbb{E}\Big[X_0^{\otimes n}(Y_t^{(n)})\exp\Big(\int_0^t\lambda_n ds\Big)\Big] = \exp\Big(\frac{1}{2}n(n-1)t\Big)\mathbb{E}\big(X_0^{\otimes n}(Y_t^{(n)})\big). \tag{4.16}$$

Lemma 4.9. *Let $n\ge 2$, and let $f\in C_b^2(\mathbb{R}^{nd})$.*

(i) The following PDE

$$\partial_t v_t(x) = A^{(n)}v_t(x) + \frac{1}{2}\sum_{\substack{1\le i,j\le n\\ i\ne j}}\kappa(x_i,x_j)v(t,x), \tag{4.17}$$

with initial value $v_0(x)=f(x)$, has a unique solution.

(ii) Let $X=\{X_t, t\in[0,T]\}$ be a solution to the martingale problem (4.4)–(4.5). Then,

$$\mathbb{E}\big(X_t^{\otimes n}(f)\big) = X_0^{\otimes n}(v_t). \tag{4.18}$$

Proof. Firstly, we claim that the operator $A^{(n)} = \frac{1}{2}(\Delta + B^{(n)})$ is uniformly parabolic in the sense of Friedman (see Section 1.1 of [12]). Because Δ is uniformly parabolic, then it suffices to analyse the properties of $B^{(n)}$. For any $k = 1, \dots, n$, $i = 1, \dots, d$, and $\xi_k^i \in \mathbb{R}$, let $\xi_k = (\xi_k^1, \dots, \xi_k^d)$. Then, we have

$$\sum_{k_1,k_2=1}^{n} \sum_{i,j=1}^{d} \rho^{ij}(x_{k_1} - x_{k_2})\xi_{k_1}^i \xi_{k_2}^j = \int_{\mathbb{R}^d} \Big| \sum_{k=1}^{n} h^*(z - x_k)\xi_k \Big|^2 dz \geq 0.$$

Thus $B^{(n)}$ is parabolic. On the other hand, by Jensen's inequality, we have

$$\sum_{k_1,k_2=1}^{n} \sum_{i,j=1}^{d} \rho^{ij}(x_{k_1} - x_{k_2})\xi_{k_1}^i \xi_{k_2}^j = \int_{\mathbb{R}^d} \Big| \sum_{k=1}^{n} h^*(z - x_k)\xi_k \Big|^2 dz \leq n\|\rho\|_\infty \sum_{k=1}^{n} |\xi_k|^2.$$

It follows that $A^{(n)} = \frac{1}{2}(\Delta + B^{(n)})$ is uniformly parabolic.

Since $h \in H_2^3(\mathbb{R}^d; \mathbb{R}^d \otimes \mathbb{R}^d)$, $\rho(x-y) = \int_{\mathbb{R}^d} h(z-x)h^*(z-y)dz$ has bounded derivatives up to order three, then by Theorem 1.12 and 1.16 of Friedman [12], the PDE (4.17) has a unique solution.

In order to show (ii), let

$$\widetilde{v}_t(x) = \mathbb{E}\big(Y_t^{(n)}(x)\big),$$

where $Y^{(n)}$ is defined by (4.6). By the same argument as we did in the proof of Lemma 4.7, we can show that for any $t \in [0,T]$ and $x \in \mathbb{R}^{nd}$

$$\mathbb{E}\Big(\sup_{x \in \mathbb{R}^d} \big|A^{(n)} Y_t^{(n)}(x)\big| \Big) < \infty.$$

Then, by the dominated convergence theorem, we have

$$\mathbb{E}\big(A^{(n)} Y_t^{(n)}(x)\big) = A^{(n)} \mathbb{E}(Y_t^{(n)}(x)).$$

Let $\mu^{(n)}$ be any finite measure on $\mathbb{R}^{nd}$. Recall that the process defined by (4.15) is a martingale, then the following equality follows from Fubini's theorem:

$$\begin{aligned}
\mu^{(n)}(\widetilde{v}_t) =& \mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\big) = \mu^{(n)}(f) + \int_0^t \big\langle \mu^{(n)}, \mathbb{E}\big(A^{(n)} Y_s^{(n)}\big) \big\rangle ds \\
&+ \frac{1}{2} \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} \int_0^t \big\langle \mu^{(n)}, [k(\cdot_i, \cdot_j) - 1]\mathbb{E}(Y_s^{(n)}) \big\rangle ds \\
=& \mu^{(n)}(f) + \int_0^t \big\langle \mu^{(n)}, A^{(n)} \widetilde{v}_s \big\rangle ds + \frac{1}{2} \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} \int_0^t \big\langle \mu^{(n)}, [k(\cdot_i, \cdot_j) - 1]\widetilde{v}_s \big\rangle ds.
\end{aligned}$$

In other words,

$$\Big\langle \mu^{(n)}, \widetilde{v}_t - f - \int_0^t \Big[A^{(n)} \widetilde{v}_s - \frac{1}{2} \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} (k(\cdot_i, \cdot_j) - 1)\widetilde{v}_s \Big] ds \Big\rangle = 0,$$

for all $\mu^{(n)} \in M_F(\mathbb{R}^{nd})$. It follows that $\widetilde{v} = \{\widetilde{v}_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ solves the following PDE

$$\partial_t \widetilde{v}_t(x) = A^{(n)} \widetilde{v}_t(x) + \frac{1}{2} \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} [\kappa(x_i, x_j) - 1]\widetilde{v}_t(x), \tag{4.19}$$

with the initial value $\widetilde{v}_0(x) = f(x)$. This solution is unique by the same argument as in part (i). Observe that

$$v_t(x) := \widetilde{v}_t(x) \exp\Big(\frac{1}{2}n(n-1)t\Big) \tag{4.20}$$

solves equation (4.17). Therefore, (4.18) follows from (4.20) and the moment duality (4.16). □

In Lemma 4.9, we derived the moment formula for $\mathbb{E}\big(X_t^{(n)}(f)\big)$ in the case when $n \geq 2$. If $n = 1$, the dual process only involves a deterministic evolution driven by the semigroup of one particle motion, which makes things much simpler. We write the formula below and skip the proof. Let $p(t, x, y)$ be the transition density of the one particle motion, then for any $\phi \in C_b^2(\mathbb{R}^d)$,

$$\mathbb{E}(X_t(\phi)) = X_0(T_t^{(1)}\phi) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p(t, x, y)\phi(y)dyX_0(dx).$$

The existence of the density of X_t will be derived following Wang's idea (see Theorem 2.1 of [32]). For any $\epsilon > 0$, $x \in \mathbb{R}^d$, let p_ϵ be the heat kernel on $\mathbb{R}^d$, that is

$$p_\epsilon(x) = (2\pi\epsilon)^{-\frac{d}{2}} \exp\Big(-\frac{|x|^2}{2\epsilon}\Big).$$

Lemma 4.10. *Let $X = \{X_t, t \in [0, T]\}$ be a solution to the martingale problem (4.4)–(4.5). Assume that the initial measure $X_0 \in M_F(\mathbb{R}^d)$ has a bounded density μ. Then,*

$$\int_0^T \int_{\mathbb{R}^d} \mathbb{E}\big(\big|X_t(p_\epsilon(x-\cdot))\big|^2\big)dxdt < \infty, \tag{4.21}$$

and

$$\lim_{\epsilon_1,\epsilon_2\downarrow 0} \int_0^T \int_{\mathbb{R}^d} \mathbb{E}\big(\big|X_t(p_{\epsilon_1}(x-\cdot)) - X_t(p_{\epsilon_2}(x-\cdot))\big|^2\big)dxdt = 0. \tag{4.22}$$

Proof. Let $\Gamma(t, (y_1, y_2); r, (z_1, z_2))$ be the fundamental solution to the PDE (4.17) when $n = 2$ (see Chapter 1 of Friedman [12] for a detailed account on existence and properties of fundamental solutions to parabolic PDEs). We write $y = (y_1, y_2)$ and $z = (z_1, z_2) \in \mathbb{R}^{2d}$. Then, for $f \in C_b^2(\mathbb{R}^{2d})$,

$$v(t, y) = \int_{\mathbb{R}^{2d}} \Gamma(t, y; 0, z)f(z)dz,$$

is the unique solution to the PDE (4.17) with initial condition $v_0 = f$. Thus by Lemma 4.9, we have

$$\begin{aligned}&\mathbb{E}\big[X_t(p_{\epsilon_1}(x-\cdot))X_t(p_{\epsilon_2}(x-\cdot))\big]\\ &= \int_{\mathbb{R}^{2d}}\int_{\mathbb{R}^{2d}} \Gamma(t, y; 0, z)p_{\epsilon_1}(x-z_1)p_{\epsilon_2}(x-z_2)dzX_0^{\otimes 2}(dy).\end{aligned} \tag{4.23}$$

By inequality (6.12) of Friedman [12] on page 24, we know that there exists $C_\Gamma, \lambda > 0$, such that for any $0 \leq r < t \leq T$,

$$|\Gamma(t, y; r; z)| \leq C_\Gamma p_{\frac{t-r}{\lambda}}(y_1 - z_1)p_{\frac{t-r}{\lambda}}(y_2 - z_2). \tag{4.24}$$

Therefore, by the semigroup property of heat kernels and Fubini's theorem, we have

$$\begin{aligned}&\int_0^T \int_{\mathbb{R}^d} \mathbb{E}\big[X_t(p_{\epsilon_1}(x-\cdot))X_t(p_{\epsilon_2}(x-\cdot))\big]dxdt\\ &= \int_0^T \int_{\mathbb{R}^{2d}}\int_{\mathbb{R}^{2d}} \Gamma(t, y; 0, z)p_{\epsilon_1+\epsilon_2}(z_1 - z_2)dzX_0^{\otimes 2}(dy)dt.\end{aligned} \tag{4.25}$$

From (4.24), (4.25) and the fact that $X_0 \in M_F(\mathbb{R}^d)$ has a bounded density μ, it follows that (4.21) is true.

Let $\mathcal{M}$ be the function on $\mathbb{R}^{2d}$, given by

$$\mathcal{M}(z) = \int_0^T \int_{\mathbb{R}^{2d}} \Gamma(t, y; 0, z) X_0^{\otimes 2}(dy) dt.$$

By (6.13) of Friedman [12] on page 24, we know that $\Gamma(t, y; r, x)$ is uniformly continuous in the spatial argument for any fixed r and t such that $0 < r < t < T$. As a consequence $\mathcal{M}$ is continuous. Therefore, by (4.24) and the continuity of $\mathcal{M}$, the function $\mathcal{N}$ on $\mathbb{R}^d$ given by

$$\mathcal{N}(x) := \int_{\mathbb{R}^d} \mathcal{M}(z_1, z_1 - x) dz_1,$$

is integrable and continuous everywhere. It follows that

$$\begin{aligned} &\lim_{\epsilon_1, \epsilon_2 \to 0} \int_0^t \int_{\mathbb{R}^d} \mathbb{E}\big[X_t(p_{\epsilon_1}(x - \cdot)) X_t(p_{\epsilon_2}(x - \cdot))\big] dx dt \\ = &\lim_{\epsilon_1, \epsilon_2 \to 0} \int_{\mathbb{R}^{2d}} \mathcal{M}(z) p_{\epsilon_1 + \epsilon_2}(z_1 - z_2) dz \\ = &\lim_{\epsilon_1, \epsilon_2 \to 0} \int_{\mathbb{R}^d} \mathcal{N}(y) p_{\epsilon_1 + \epsilon_2}(y) dy \\ = &\mathcal{N}(0) = \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^{2d}} \Gamma(t, y; 0, (x, x)) X_0^{\otimes 2}(dy) dx dt. \end{aligned} \tag{4.26}$$

Therefore, (4.22) is a consequence of (4.26). □

Proposition 4.11. *Let $X = \{X_t, t \in [0, T]\}$ be a solution to the martingale problem (4.4)–(4.5). Assume that the initial measure $X_0 \in M_F(\mathbb{R}^d)$ has a bounded density μ. Then, for almost every $t \in (0, T]$, X_t is absolutely continuous with respect to the Lebesgue measure almost surely.*

Proof. As proved in Lemma 4.10, for any $x \in \mathbb{R}^d$ and $\epsilon_n \downarrow 0$, the sequence $\{X_t(p_{\epsilon_n}^x)\}_{n \geq 1}$ is Cauchy in $L^2(\Omega \times \mathbb{R}^d \times [0, T])$. Then, it converges to some square integrable random field. By the same argument as in Theorem 2.1 of Wang [32], we can show that the limit random field is the density of X_t almost surely. □

Remark 4.12.

(i) The assumption in Proposition 4.11, that the initial measure has a bounded density, cannot be removed. Actually, if we choose $X_0 = \delta_0$, the Dirac delta mass at 0, then $\int_0^T \int_{\mathbb{R}^d} \Gamma(t, 0; 0, (x, x)) dx dt$ behaves like $\int_0^T t^{-\frac{d}{2}} dt$, which is finite only if $d = 1$. This is another difference from the one dimensional situation, in which case $X_0(1) < \infty$ is enough to imply the existence of the density (see Theorem 2.1 Wang [32] for the Dawson-Watanabe branching model).

(ii) The method of duality is conventionally used to prove the well-posedness of martingale problems arisen from branching mechanisms. In the one-dimensional Dawson-Watanabe model, Wang proved the well-posedness by solving a moment problem (see Section 4 of [33]). This requires a moment bound of the form $\sum_{n=1}^{\infty} r^n \mathbb{E}(|X_t(1)|^n)/n! < \infty$ for some positive r. However, this method does not work in our model and here is the explanation. In the next section, we will prove that the density u is a solution to equation (3.1) and when $h \equiv 0$, we have that $\mathbb{E}(\langle u_t, 1 \rangle^n)$ behaves like $c_1 e^{c_2 n^{1+\epsilon}}$ for some $\epsilon > 0$ (see e.g. Theorem 4.4 of Chen et al. [3] and Theorem 4.3 of Hu et al. [13] for some sharp bounds of similar SPDE's).

Therefore, the condition $\sum_{n=1}^{\infty} r^n \mathbb{E}(|\langle u_t, 1\rangle|^n)/n! < \infty$ for some positive r cannot be satisfied in our model. In the next section, we prove the well-posedness of the martingale problem (4.4)–(4.5) by the Yamada-Watanabe argument assuming the existence of the density. Without the existence of the density, we are currently not able to use the moment duality to show the well-posedness of the martingale problem. We are not pursue this in the current paper.

4.3 Proof of Theorems 3.2 and 3.3

The proof of Theorems 3.2 and 3.3 is based on the equivalence of the martingale problem (4.4)–(4.5) and the SPDE (3.1).

The equivalence between martingale problems and SDE's in finite dimensions was observed in the 1970s (see Stroock and Varachan [29]). An alternative proof given by Kurtz [19] consists of the "Markov mapping theorem". In a recent paper [2] Biswas et al. generalized this result to the infinite dimensional cases with one noise following Kurtz's idea. Here in the present paper, we establish a similar result with two noises by using the martingale representation theorem.

Proposition 4.13. *Let $\mu \in C_b(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ be a nonnegative function on $\mathbb{R}^d$. Then, $u = \{u_t, t \in [0,T]\}$ is the density of a solution of the martingale problem (4.4)–(4.5) with initial density μ, if and only if u is a weak solution to the SPDE (3.1).*

Proof. If u is a weak solution to (3.1), then, as a consequence of Itô's formula, u is the density of a measure-valued process that solves the martingale problem (4.4)–(4.5). It suffices to show the converse statement.

Let $X = \{X_t, t \in [0,T]\}$ be a solution to the martingale problem (4.4)–(4.5) with initial density μ. Then, by Proposition 4.11, for almost every $t \in [0,T]$, X_t has a density almost surely. We denote by u_t the density of X_t.

Consider $M = \{M_t, t \in [0,T]\}$ defined by (4.4) as an $\mathscr{S}'$-martingale (see Definition 2.1.2 of Kallianpur and Xiong [17]). Then, by Theorem 3.1.4 of [17], there exists a Hilbert space $\mathcal{H}^* \supset L^2(\mathbb{R}^d)$, such that M is an $\mathcal{H}^*$-valued martingale. Denote by $\mathcal{H}$ the dual space of $\mathcal{H}^*$.

Let $\mathfrak{H}_1 = L^2(\mathbb{R}^d; \mathbb{R}^d)$, and let $\mathfrak{H}_2$ be the completion of $\mathscr{S}$ with the inner product

$$\langle \phi, \varphi \rangle_{\mathfrak{H}_2} := \int_{\mathbb{R}^d \times \mathbb{R}^d} \kappa(x,y)\phi(x)\varphi(y)dxdy.$$

Consider the product space $\mathfrak{H} = \mathfrak{H}_1 \times \mathfrak{H}_2$. Then, $\mathfrak{H}$ is a Hilbert space equipped with the inner product

$$\big\langle (\phi_1, \phi_2), (\varphi_1, \varphi_2) \big\rangle_{\mathfrak{H}} := \langle \phi_1, \varphi_1 \rangle_{\mathfrak{H}_1} + \langle \phi_2, \varphi_2 \rangle_{\mathfrak{H}_2}.$$

For any $t \in [0,T]$, let $\Psi_t : \mathcal{H} \to \mathfrak{H}$ be given by $\Psi_t(\phi)(x,y) = \big(\Psi_t^1(\phi)(x), \Psi_t^2(\phi)(y)\big)$, where

$$\Psi_t^1(\phi)(x) := \int_{\mathbb{R}^d} \nabla\phi(y)^* h(x-y) u_t(y) dy,$$

and

$$\Psi_t^2(\phi)(x) := \phi(x) u_t(x).$$

Then, for any $\phi, \varphi \in \mathcal{H}$, we have

$$\begin{aligned}
\langle M(\phi), M(\varphi) \rangle_t &= \int_0^t \nabla\phi(x)^* \rho(x-y) \nabla\varphi(y) X_s(dx) X_s(dy) ds \\
&\quad + \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} \kappa(x,y)\phi(x)\phi(y) X_s(dx) X_s(dy) ds \\
&= \int_0^t \langle \Phi_s(\phi), \Phi_s(\varphi) \rangle_{\mathfrak{H}} ds.
\end{aligned}$$

Therefore, by the martingale representation theorem (see e.g. Theorem 3.3.5 of Kallianpur and Xiong [17]), there exists a $\mathfrak{H}$-cylindrical Brownian motion $\mathfrak{B} = \{\mathfrak{B}_t, 0 \le t \le T\}$, such that

$$M_t(\phi) = \int_0^t \big\langle \Psi_s(\phi), d\mathfrak{B}_s \big\rangle.$$

Let $\mathfrak{B}^1 = \{\mathfrak{B}_t^1(\phi), t \in [0,T], \phi \in \mathfrak{H}_1\}$ and $\mathfrak{B}^2 = \{\mathfrak{B}_t^2(\varphi), t \in [0,T], \varphi \in \mathfrak{H}_2\}$ be given by

$$\mathfrak{B}_t^1(\phi) = \mathfrak{B}_t(\phi, 0) \text{ and } \mathfrak{B}_t^2(\varphi) = \mathfrak{B}_t(0, \varphi).$$

Then, $\mathfrak{B}^1$ and $\mathfrak{B}^2$ are $\mathfrak{H}^1$- and $\mathfrak{H}^2$-cylindrical Brownian motion respectively, and they are independent. As a consequence, we have

$$M_t(\phi) = \int_0^t \Big\langle \int_{\mathbb{R}^d} \nabla\phi(z)^* h(\cdot - z) X_s(dz), d\mathfrak{B}_s^1 \Big\rangle + \int_0^t \big\langle \phi u_s, d\mathfrak{B}_s^2 \big\rangle. \tag{4.27}$$

Let $\{e_j\}_{j\ge1}$ be a complete orthonormal basis of $\mathfrak{H}_2$. Then, by Theorem 3.2.5 of [17], $V = \{V_t, t \in [0,T]\}$, defined by

$$V_t := \sum_{j=1}^{\infty} \mathfrak{B}_t^2(e_j) e_j,$$

is a $\mathscr{S}'$-valued Wiener process with covariance

$$\mathbb{E}\big[V_s(\phi) V_t(\varphi)\big] = s \wedge t \int_{\mathbb{R}^d \times \mathbb{R}^d} \kappa(x,y) \phi(x) \varphi(y) dx dy,$$

for any $\phi, \varphi \in \mathscr{S}$. Therefore, by (4.27) and the equivalence of stochastic integrals with respect to Hilbert space valued Brownian motion and Walsh's integrals (see e.g. Proposition 2.6 of Dalang and Quer-Sardanyons [4] for spatial homogeneous noises), u is a weak solution to the SPDE (3.1). □

Proof of Theorems 3.2 and 3.3. By Propositions 4.5 and 4.13, the SPDE (3.1) has a weak solution, that can be obtained by the branching particle approximation. We do not prove the continuity here, because later in Section 7, we will show that the solution is not only continuous, but also Hölder continuous. The continuity of u yields an improved version of Proposition 4.11. Namely, if X_t is a continuous measure-valued process (e.g. the limit of the particle approximation), then it has a density for all $t \in [0,T]$ almost surely.

In the next step, we prove the pathwise uniqueness of equation (3.1). Assume that u and $\widetilde{u}$ are two continuous strong solutions to (3.1) with initial condition μ. Let $d = u - \widetilde{u}$. Then, $d = \{d_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is a solution to (3.1), with initial condition $\mu \equiv 0$, that is continuous in two parameters. Thus d is also the density of a solution to the martingale problem (4.4)–(4.5), with initial measure $X_0 \equiv 0$.[1] By the moment duality (4.16), for any $\phi \in C_b^2(\mathbb{R}^d)$, we have

$$\mathbb{E}\langle d_t, \phi\rangle^2 = \exp(t)\mathbb{E}\big(X_0(Y_t^{(2)})\big) \equiv 0,$$

where $Y^{(2)}$ is the dual process defined by (4.6) in the case $n = 2$. Since d is continuous in t, it follows that $u = \widetilde{u}$ almost surely. Therefore, by the Yamada-Watanabe argument (see Yamada and Watanabe [36] and Kurtz [18]), we obtain the strong existence and weak uniqueness of equation (3.1). This proves Theorem 3.3. Recall Propositions 4.5 and 4.13. The weak uniqueness of equation (3.1) also implies that every limit of the convergent

[1] $d_t(x)$ may be negative for some $(t,x) \in [0,T] \times \mathbb{R}^d$. In this case d_t is considered as the density of a signed measure ν, where $|\nu|(1) \le |u_t(1)| + |\widetilde{u}_t(1)| < \infty$ a.s.. The moment duality still holds.

subsequence of $\{X^n\}_{n\geq 1}$ constructed in Section 4.1 is continuous (see Lemma 4.4) and unique in law. In other words, $\{X^n\}_{n\geq 1}$ is convergent in $D([0,T];M_F)$ to a continuous $M_F(\mathbb{R}^d)$-valued process in law. The limit has a density almost surely, that is a weak solution the SPDE (3.1). □

The following corollary is a direct consequence of Theorem 3.3 and Proposition 4.13.

Corollary 4.14. *Assume Hypotheses* ***[H0]***, ***[H1]***, *and assume that* $X_0 \in M_F(\mathbb{R}^d)$ *has a bounded density. Then, the martingale problem (4.4)–(4.5) is well-posed.*

5 Moment estimates for one-particle motion

In this section, we focus on the one-particle motion without branching. By using the techniques of Malliavin calculus, we will obtain moment estimates for the transition probability density of the particle motion conditional on the environment W. A brief introduction and several theorems on Malliavin calculus are stated in Appendix A. For a detailed account on this topic, we refer the readers to the book of Nualart [24].

Fix a time interval $[0,T]$. Let $B=\{B_t, 0\leq t\leq T\}$ be a standard d-dimensional Brownian motion and let W be a d-dimensional space-time white Gaussian random field on $[0,T]\times\mathbb{R}^d$ that is independent of B. Assume that $h\in H_2^3(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$. For any $0\leq r<t\leq T$, we denote by $\xi_t=\xi_t^{r,x}$, the path of one-particle motion, with initial position $\xi_r = x$. It satisfies the SDE

$$\xi_t = x + B_t - B_r + \int_r^t \int_{\mathbb{R}^d} h(y-\xi_u) W(du,dy). \tag{5.1}$$

We will apply the Malliavin calculus on ξ_t with respect to the Brownian motion B. Let $H = L^2([0,T];\mathbb{R}^d)$ be the associated Hilbert space. By the Picard iteration scheme (see e.g. Theorem 2.2.1 of Nualart [24]), we can prove that for any $t\in(r,T]$, $\xi_t\in\cap_{p\geq 1}\mathbb{D}^{3,p}(\mathbb{R}^d)$. Particularly, $D\xi_t$ satisfies the following system of SDE's

$$D_\theta^{(k)}\xi_t^i = \delta_{ik} - \sum_{j_1,j_2=1}^d \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_1} h^{ij_2}(y-\xi_s) D_\theta^{(k)}\xi_s^{j_1} W^{j_2}(ds,dy), \quad 1\leq i,k\leq d, \tag{5.2}$$

for any $\theta\in[r,t]$, and $D_\theta^{(k)}\xi_t^i = 0$ for all $\theta > t$.

In order to simplify the expressions, we rewrite the stochastic integrals in (5.2) as integrals with respect to martingales. To this end, let $M=\{M_t, r\leq t\leq T\}$ be the $d\times d$ matrix-valued process given by

$$M_t = \sum_{k=1}^d \int_r^t \int_{\mathbb{R}^d} g_k(s,y) W^k(ds,dy),$$

where $g_k: \Omega\times[r,T]\times\mathbb{R}^d \to \mathbb{R}^d\otimes\mathbb{R}^d$ is given by

$$g_k^{ij}(t,y) = \partial_i h^{jk}(y-\xi_t), \quad 1\leq i,j,k\leq d.$$

Notice that M_t is the sum of stochastic integrals, so it is a matrix-valued martingale. The quadratic covariations of $\{M^{ij}\}_{i,j=1}^d$ are bounded and deterministic:

$$\begin{aligned}\langle M^{i_1j_1}, M^{i_2j_2}\rangle_t =& \sum_{k=1}^d \int_r^t \int_{\mathbb{R}^d} \partial_{i_1} h^{j_1k}(y-\xi_s)\partial_{i_2} h^{j_2k}(y-\xi_s) dyds \\ =&(t-r)\sum_{k=1}^d \int_{\mathbb{R}^d} \partial_{i_1} h^{j_1k}(y)\partial_{i_2} h^{j_2k}(y) dy := Q_{i_2,j_2}^{i_1,j_1}(t-r) \leq \|h\|_{3,2}(t-r).\end{aligned} \tag{5.3}$$

Now equation (5.2) can be rewritten as follows:

$$D_\theta^{(k)}\xi_t^i = \delta_{ik} - \sum_{j=1}^d \int_\theta^t \int_{\mathbb{R}^d} D_\theta^{(k)}\xi_s^j dM_s^{ji}, \quad 1 \le i,k \le d. \tag{5.4}$$

Lemma 5.1. *For any $0 \le r < t \le T$, $x \in \mathbb{R}^d$, let $\gamma_t = \gamma_{\xi_t}$ be the Malliavin matrix of $\xi_t = \xi_t^{r,x}$, then γ_t is nondegenerate almost surely.*

Proof. We prove the lemma following Stroock's idea (see Chapter 8 of Stroock [28]). Let $\lambda_\theta(t)$ be the $d \times d$ symmetric random matrix given by

$$\lambda_\theta^{ij}(t) = \sum_{k=1}^d D_\theta^{(k)}\xi_t^i D_\theta^{(k)}\xi_t^j.$$

Then, the Malliavin matrix of ξ_t is the integral of $\lambda_\theta(t)$:

$$\gamma_t = \int_r^t \lambda_\theta(t) d\theta.$$

By (5.2), (5.3) and Itô's formula, we have

$$\begin{aligned} D_\theta^{(k)}\xi_t^i D_\theta^{(k)}\xi_t^j =& \delta_{ik}\delta_{kj} - \sum_{k_1=1}^d \int_\theta^t D_\theta^{(k)}\xi_s^i D_\theta^{(k)}\xi_s^{k_1} dM_s^{k_1 j} - \sum_{k_2=1}^d \int_\theta^t D_\theta^{(k)}\xi_s^j D_\theta^{(k)}\xi_s^{k_2} dM_s^{k_2 i} \\ &+ \sum_{k_1,k_2=1}^d Q_{k_2,i}^{k_1,j} \int_\theta^t D_\theta^{(k)}\xi_s^{k_1} D_\theta^{(k)}\xi_s^{k_2} ds. \end{aligned}$$

Therefore,

$$\begin{aligned} \lambda_\theta(t) =& I - \int_\theta^t \lambda_\theta(s) dM_s - \int_\theta^t dM_s^* \cdot \lambda_\theta(s) \\ &+ \sum_{k=1}^d \int_\theta^t \int_{\mathbb{R}^d} g_k^*(s,y)\lambda_\theta(s) g_k(s,y) dy ds. \end{aligned} \tag{5.5}$$

For any $\theta \in [r,t]$, we claim that $\lambda_\theta(t)$ is invertible almost surely, and its inverse $\beta_\theta(t)$ satisfies the following SDE:

$$\begin{aligned} \beta_\theta(t) =& I + \int_\theta^t \beta_\theta(s) dM_s^* + \int_\theta^t dM_s \cdot \beta_\theta(s) \\ &+ \sum_{k=1}^d \int_\theta^t \int_{\mathbb{R}^n} \left(g_k(s,y)^2 \beta_\theta(s) + g_k(s,y)\beta_\theta(s) g_k^*(s,y) + \beta_\theta(s) g_k^*(s,y)^2 \right) dy ds. \end{aligned} \tag{5.6}$$

Indeed, by Itô's formula, we have

$$\begin{aligned} d[\lambda_\theta(t)\beta_\theta(t)] =& - dM_t^* \cdot [\lambda_\theta(t)\beta_\theta(t)] + [\lambda_\theta(t)\beta_\theta(t)] dM_t^* \\ &+ \sum_{k=1}^d \Big(\int_{\mathbb{R}^d} \left([\lambda_\theta(t)\beta_\theta(t)] g_k^*(t,y)^2 - g_k^*(t,y)[\lambda_\theta(t)\beta_\theta(t)] g_k^*(t,y) \right) dy \Big) dt. \end{aligned} \tag{5.7}$$

Note that $\lambda_\theta(t)\beta_\theta(t) \equiv I$ solves the SDE (5.7) with initial value $\lambda_\theta(\theta)\beta_\theta(\theta) = I$. Therefore, the strong uniqueness of the linear SDE (5.7) implies that $\lambda_\theta^{-1}(t) = \beta_\theta(t)$ almost surely.

Denote by $\|\cdot\|_2$ the Hilbert-Schmidt norm of matrices. By Jensen's inequality (see Lemma 8.14 of Stroock [28]), the following inequality holds almost surely

$$\big\|\gamma_t^{-1}\big\|_2 = \Big\|\Big(\int_r^t \lambda_\theta(t)d\theta\Big)^{-1}\Big\|_2 \leq \frac{1}{(t-r)^2}\Big\|\int_r^t \beta_\theta(t)d\theta\Big\|_2. \tag{5.8}$$

It is easy to show that $\sup_{\theta\in[r,t]}\big\|\|\beta_\theta(t)\|_2\big\|_{2p} < \infty$ for all $p \geq 1$. Therefore, the right-hand side of (5.8) is finite a.s., and thus γ_t is nondegenerate almost surely. □

We denote by $\sigma_t = \gamma_t^{-1}$ the inverse of the Malliavin matrix of ξ_t. In the following lemma, we obtain some moment estimates for the derivatives of ξ_t and σ_t. Before estimates, we introduce the following generalized Cauchy-Schwarz's inequality.

Lemma 5.2. *Let n_1, n_2 be nonnegative integers, let $u_1 \in L^{2p}(\Omega; (H^{\otimes n_1}))$, and let $u_2 \in L^{2p}(\Omega, (H^{\otimes n_2}))$. Then, $u_1 \otimes u_2 \in L^p(\Omega; (H^{\otimes(n_1+n_2)}))$, and*

$$\big\|\|u_1\otimes u_2\|_{H^{\otimes(n_1+n_2)}}\big\|_p \leq \big\|\|u_1\|_{H^{\otimes n_1}}\big\|_{2p}\big\|\|u_2\|_{H^{\otimes n_2}}\big\|_{2p}. \tag{5.9}$$

Proof. The lemma can be obtained by the classical Cauchy-Schwarz inequality. □

Lemma 5.3. *For any $p \geq 1$ and $0 \leq r < t \leq T$, there exists a constant $C > 0$, that depends on T, d, $\|h\|_{3,2}$, p, such that*

$$\max_{1\leq i\leq d}\big\|\|D\xi_t^i\|_H\big\|_{2p} \leq C(t-r)^{\frac{1}{2}}. \tag{5.10}$$

$$\max_{1\leq i,j\leq d}\big\|\sigma_t^{ij}\big\|_{2p} \leq C(t-r)^{-1}, \tag{5.11}$$

$$\max_{1\leq i,j\leq d}\big\|\|D\sigma_t^{ij}\|_H\big\|_{2p} \leq C, \tag{5.12}$$

$$\max_{1\leq i\leq d}\big\|\|D^2\xi_t^i\|_{H^{\otimes 2}}\big\|_{2p} \leq C(t-r)^{\frac{3}{2}}. \tag{5.13}$$

$$\max_{1\leq i,j\leq d}\big\|\|D^2\sigma_t^{ij}\|_{H^{\otimes 2}}\big\|_{2p} \leq C(t-r)^{\frac{1}{2}}, \tag{5.14}$$

$$\max_{1\leq i\leq d}\big\|\|D^3\xi_t^i\|_{H^{\otimes 3}}\big\|_{2p} \leq C(t-r)^2. \tag{5.15}$$

Proof of (5.10). By (5.3), (5.4), Jensen's, Burkholder-Davis-Gundy's, and Minkowski's inequalities, we have

$$\begin{aligned}\sum_{i,k=1}^d \|D_\theta^{(k)}\xi_t^i\|_{2p}^2 \leq& \sum_{i,k=1}^d\Big(\delta_{ik} + \sum_{j=1}^d\Big\|\int_\theta^t\int_{\mathbb{R}^d} D_\theta^{(k)}\xi_s^j dM_s^{ji}\Big\|_{2p}\Big)^2\\ \leq&(d+1)\sum_{i,k=1}^d\Big(\delta_{ik} + \sum_{j=1}^d\Big\|\int_\theta^t\int_{\mathbb{R}^d} D_\theta^{(k)}\xi_s^j dM_s^{ji}\Big\|_{2p}^2\Big)\\ \leq& d(d+1) + (d+1)c_p\sum_{i,j,k=1}^d Q_{ji}^{ji}\Big\|\int_\theta^t \big|D_\theta^{(k)}\xi_s^j\big|^2 ds\Big\|_p\\ \leq& d(d+1) + 2c_p d(d+1)\|h\|_{3,2}^2\sum_{j,k=1}^d\int_\theta^t \|D_\theta^{(k)}\xi_s^j\|_{2p}^2 ds. \end{aligned}\tag{5.16}$$

Thus by Gronwall's lemma, we have

$$\sum_{i,j=1}^d \|D_\theta^{(k)}\xi_t^j\|_{2p}^2 \leq d(d+1)\exp\big(2c_p d(d+1)\|h\|_{3,2}^2 T\big) := C. \tag{5.17}$$

Therefore, by (5.17) and Minkowski's inequality, we have

$$\| \| D\xi_t^i \|_H \|_{2p}^2 = \left\| \sum_{k=1}^d \int_r^t |D_\theta^{(k)} \xi_t^i|^2 d\theta \right\|_p \leq \sum_{k=1}^d \int_r^t \| D_\theta^{(k)} \xi_t^i \|_{2p}^2 d\theta \leq C(t-r).$$

This completes the proof of (5.10). □

Proof of (5.11). In order to prove (5.11), we rewrite the SDE (5.6) in the following way:

$$\begin{aligned} \beta_\theta^{ij}(t) =& \delta_{ij} + \sum_{k_1=1}^d \int_\theta^t \beta_\theta^{ik_1}(s) dM_s^{jk_1} + \sum_{k_2=1}^d \int_\theta^t \beta_\theta^{k_2 j}(s) dM_s^{ik_2} \\ &+ \sum_{k_1,k_2=1}^d \left(Q_{k_1,k_2}^{i,k_1} \int_\theta^t \beta_\theta^{k_2 j}(s) ds \right) + \sum_{k_1,k_2=1}^d \left(Q_{j,k_2}^{i,k_1} \int_\theta^t \beta_\theta^{k_1 k_2}(s) ds \right) \\ &+ \sum_{k_1,k_2=1}^d \left(Q_{j,k_2}^{k_2,k_1} \int_\theta^t \beta_\theta^{ik_1}(s) ds \right). \end{aligned} \tag{5.18}$$

Similarly as we did in step **(i)**, by Burkholder-Davis-Gundy's, and Minkowski's inequalities, we can show that the martingale terms satisfies the following inequality

$$\left\| \int_\theta^t \beta_\theta^{ik_1}(s) dM_s^{jk_1} \right\|_{2p}^2 \leq 2c_p \|h\|_{3,2}^2 \int_\theta^t \left\| \beta_\theta^{ik_1}(s) \right\|_{2p}^2 ds. \tag{5.19}$$

For the drift terms, by Minkowski's and Jensen's inequality, we have

$$\left\| \int_\theta^t \beta_\theta^{k_1 k_2}(s) ds \right\|_{2p}^2 \leq (t-\theta) \int_\theta^t \left\| \beta_\theta^{k_1 k_2}(s) \right\|_{2p}^2 ds. \tag{5.20}$$

Then, by (5.18)–(5.20), and Gronwall's lemma, we have

$$\sum_{i,j=1}^d \left\| \beta_\theta^{ij}(t) \right\|_{2p}^2 \leq C.$$

Thus by Minkowski's and Jensen's inequalities, we have

$$\left\| \left\| \int_r^t \beta_\theta(t) d\theta \right\|_2 \right\|_{2p} \leq c_d \sum_{i,j=1}^d \int_r^t \left\| \beta_\theta^{ij}(t) \right\|_{2p} d\theta \leq C(t-r). \tag{5.21}$$

Therefore, (5.11) follows from (5.8), (5.21), Minkowski's and Jensen's inequalities. □

Proof of (5.12). By integrating equation (5.5) on both sides with respect to θ, and applying the stochastic Fubini theorem (see e.g. Lemma 4.1 on page 116 of Ikeda and Watanabe [15]), we have

$$\begin{aligned} \gamma_t = \int_r^t \lambda_\theta(t) d\theta =& I(t-r) - \int_r^t \gamma_s dM_s - \int_r^t dM_s^* \cdot \gamma_s \\ &+ \sum_{m=1}^d \int_r^t \int_{\mathbb{R}^d} g_m^*(y,s) \gamma_s g_m(y,s) dy ds. \end{aligned} \tag{5.22}$$

Taking the Malliavin derivative on both sides of (5.22), we have the following SDE:

$$
\begin{aligned}
D_\theta^{(k)}\gamma_t^{ij} =& -\sum_{k_1=1}^d \int_\theta^t D_\theta^{(k)}\gamma_s^{ik_1} dM_s^{k_1 j} - \sum_{k_1=1}^d \int_\theta^t \gamma_s^{ik_1} d\left(D_\theta^{(k)} M_s^{k_1 j}\right) \\
& - \sum_{k_2=1}^d \int_\theta^t D_\theta^{(k)}\gamma_s^{k_2 j} dM_s^{k_2 i} - \sum_{k_2=1}^d \int_\theta^t \gamma_s^{k_2 j} d\left(D_\theta^{(k)} M_s^{k_2 i}\right) \\
& + \sum_{k_1,k_2=1}^d \left(Q_{k_2,j}^{k_1,i} \int_\theta^t D_\theta^{(k)}\gamma_s^{k_1 k_2} ds\right),
\end{aligned} \tag{5.23}
$$

where

$$
D_\theta^{(k)} M_s^{ij} = -\sum_{i_1,i_2=1}^d \int_\theta^s \int_{\mathbb{R}^d} \partial_{i,i_2} h^{ji_1}(y-\xi_r) D_\theta^{(k)} \xi_r^{i_2} W^{i_1}(dr,dy). \tag{5.24}
$$

For the first and the third term, by similar arguments as in (5.16), we can show that

$$
\left\| \int_\theta^t D_\theta^{(k)}\gamma_s^{ik_1} dM_s^{k_1 j} \right\|_{2p}^2 \le c_{d,p} \|h\|_{3,2}^2 \int_\theta^t \left\| D_\theta^{(k)}\gamma_s^{ik_1}\right\|_{2p}^2 ds. \tag{5.25}
$$

To estimate the second and the fourth term, notice that by (5.10), we have

$$
\begin{aligned}
\max_{1\le i,j\le d} \left\|\gamma_t^{ij}\right\|_{2p} =& \max_{1\le i,j\le d} \left\|\langle D\xi_t^i, D\xi_t^j\rangle_H\right\|_{2p} \\
\le& \max_{1\le i\le d} \left\| \|D\xi_t^i\|_H \right\|_{4p} \max_{1\le j\le d} \left\| \|D\xi_t^j\|_H \right\|_{4p} \le C(t-r).
\end{aligned} \tag{5.26}
$$

Therefore, by (5.17), (5.24), (5.26), Jensen's, Burkholder-Davis-Gundy's, Minkowski's, and Cauchy-Schwarz's inequalities, we have

$$
\begin{aligned}
\left\| \int_\theta^t \gamma_s^{ik_1} d\big(D_\theta^{(k)} M_s^{k_1 j}\big) \right\|_{2p}^2 \le& c_{d,p} \|h\|_{3,2}^2 \sum_{k_2=1}^d \int_\theta^t \left\|\gamma_s^{ik_1}\right\|_{4p}^2 \left\| D_\theta^{(k)} \xi_s^{k_2} \right\|_{4p}^2 ds \\
\le& C(t-r)^3.
\end{aligned} \tag{5.27}
$$

For the last term, by Minkowski's and Jensen's inequalities, we have

$$
\left\| \int_\theta^t D_\theta^{(k)}\gamma_s^{k_1 k_2} ds \right\|_{2p}^2 \le (t-\theta) \int_\theta^t \left\|D_\theta^{(k)}\gamma_s^{k_1 k_2}\right\|_{2p}^2 ds \le T \int_\theta^t \left\|D_\theta^{(k)}\gamma_s^{k_1 k_2}\right\|_{2p}^2 ds. \tag{5.28}
$$

Combining (5.23)–(5.28), we obtain the following inequality

$$
\sum_{i,j=1}^d \left\|D_\theta^{(k)}\gamma_t^{ij}\right\|_{2p}^2 \le c_1 \int_\theta^t \sum_{i,j=1}^d \left\|D_\theta^{(k)}\gamma_s^{ij}\right\|_{2p}^2 ds + c_2(t-r)^3, \tag{5.29}
$$

where c_1, c_2 depends on T, d, $\|h\|_{3,2}^2$, and p. Thus by Gronwall's lemma, we have

$$
\sum_{i,j=1}^d \left\|D_\theta^{(k)}\gamma_t^{ij}\right\|_{2p}^2 \le C(t-r)^3. \tag{5.30}
$$

It follows that

$$
\left\| \|D\gamma_t^{ij}\|_H \right\|_{2p} \le C(t-r)^2. \tag{5.31}
$$

Notice that $\gamma_t\sigma_t = I$, a.s., as a consequence, $D(\gamma_t\sigma_t) = DI \equiv 0$. That implies

$$
D\sigma_t^{ij} = -\sum_{i_1,i_2=1}^d \sigma_t^{ii_1} D\gamma_t^{i_1 i_2} \sigma_t^{i_2 j}. \tag{5.32}
$$

Then, (5.12) follows from (5.9), (5.11), (5.31) and (5.32). □

Proof of (5.13). Fix $0 \le r < t \le T$. For any $\theta_1, \theta_2 \in [r,t]$, let $\theta = \theta_1 \vee \theta_2$. Taking the Malliavin derivative on both sides of (5.4), we have the following SDE:

$$D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^i_t = -\sum_{j_1=1}^d \int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_1}_s dM^{j_1 i}_s$$
$$+ \sum_{j_1,j_2,j_3=1}^d \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D^{(k_1)}_{\theta_1}\xi^{j_2}_s D^{(k_2)}_{\theta_2}\xi^{j_3}_s W^{j_1}(ds,dy). \tag{5.33}$$

Similarly as in (5.16), we can show the following inequalities

$$\Big\| \int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_1}_s dM^{j_1 i}_s \Big\|^2_{2p} \le c_{d,p}\|h\|^2_{3,2} \int_\theta^t \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_1}_s\|^2_{2p} ds, \tag{5.34}$$

and

$$\Big\| \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D^{(k_1)}_{\theta_1}\xi^{j_2}_s D^{(k_2)}_{\theta_2}\xi^{j_3}_s W^{j_1}(ds,dy) \Big\|^2_{2p}$$
$$\le c_p\|h\|^2_{3,2} \int_\theta^t \|D^{(k_1)}_{\theta_1}\xi^{j_2}_s\|^2_{4p} \|D^{(k_2)}_{\theta_2}\xi^{j_3}_s\|^2_{4p} ds \le C(t-r). \tag{5.35}$$

Thus combining (5.33)–(5.35), we have

$$\sum_{i=1}^d \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^i_t\|^2_{2p} \le c_1 \sum_{i=1}^d \int_\theta^t \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^i_s\|^2_{2p} ds + c_2(t-r).$$

Then, it follows from Gronwall's lemma that

$$\sum_{i=1}^d \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^i_t\|^2_{2p} \le C(t-r). \tag{5.36}$$

Inequality (5.13) is a consequence of (5.36), Jensen's and Minkowski's inequalities. □

Proof of (5.14). For any $\theta_1, \theta_2 \in [r,t]$ and $\theta = \theta_1 \vee \theta_2$, by taking the Malliavin derivative on both sides of (5.23), we have

$$D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{ij}_t = -\sum_{i_1=1}^d \Big(\int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{ii_1}_s dM^{i_1 j}_s + \int_\theta^t D^{(k_1)}_{\theta_2}\gamma^{ii_1}_s d\big(D^{(k_2)}_{\theta_1} M^{i_1 j}_s\big)\Big)$$
$$-\sum_{i_1=1}^d \Big(\int_\theta^t D^{(k_2)}_{\theta_2}\gamma^{ii_1}_s d\big(D^{(k_1)}_{\theta_1} M^{i_1 j}_s\big) + \int_\theta^t \gamma^{ii_1}_s d\big(D^{(k_1,k_2)}_{\theta_1,\theta_2} M^{i_1 j}_s\big)\Big)$$
$$-\sum_{i_2=1}^d \Big(\int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{i_2 j}_s dM^{i_2 i}_s + \int_\theta^t D^{(k_1)}_{\theta}\gamma^{i_2 j}_s d\big(D^{(k_2)} M^{i_2 i}_s\big)\Big)$$
$$-\sum_{i_2=1}^d \Big(\int_\theta^t D^{(k_2)}_{\theta_1}\gamma^{i_2 j}_s d\big(D^{(k_1)}_{\theta_2} M^{i_2 i}_s\big) + \int_\theta^t \gamma^{i_2 j}_s d\big(D^{(k_1,k_2)}_{\theta_1,\theta_2} M^{i_2 i}_s\big)\Big)$$
$$+\sum_{i_1,i_2=1}^d \Big(Q^{i_1,i}_{i_2,j} \int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{i_1 i_2}_s ds\Big), \tag{5.37}$$

where

$$D^{(k_1,k_2)}_{\theta_1,\theta_2} M^{ij}_s = -\sum_{j_1,j_2,j_3=1}^d \int_\theta^s \int_{\mathbb{R}^d} \partial_{i,j_2,j_3} h^{jj_1}(y-\xi_r) D^{(k_1)}_{\theta_1}\xi^{j_2}_r D^{(k_2)}_{\theta_2}\xi^{j_3}_r W^{j_1}(dr,dy)$$
$$+ \sum_{j_1,j_2=1}^d \int_\theta^s \int_{\mathbb{R}^d} \partial_{i,j_2} h^{jj_1}(y-\xi_r) D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_2}_r W^{j_1}(dr,dy).$$

By (5.17), (5.26), (5.30), (5.36), Burkholder-Davis-Gundy's, Minkowski's and Hölder's inequalities, we have the following inequalities

$$\Big\| \int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma_s^{ii_1} dM_s^{i_1 j}\Big\|_{2p}^2 \le c_{d,p}\|h\|_{3,2}^2 \int_\theta^t \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma_s^{ii_1}\|_{2p}^2 ds, \tag{5.38}$$

$$\begin{aligned}\Big\| \int_\theta^t D^{(k_1)}_{\theta_2}\gamma_s^{ii_1} d\big(D^{(k_2)}_{\theta_1} M_s^{i_1 j}\big)\Big\|_{2p}^2 \le & c_{d,p}\|h\|_{3,2}^2 \sum_{i_2=1}^d \int_\theta^t \|D^{(k_1)}_{\theta_2}\gamma_s^{ii_1} D^{(k_2)}_{\theta_2}\xi_s^{i_2}\|_{2p}^2 ds \\ \le & c_{d,p}\|h\|_{3,2}^2 \sum_{i_2=1}^d \int_\theta^t \|D^{(k_1)}_{\theta_2}\gamma_s^{ii_1}\|_{4p}^2 \|D^{(k_2)}_{\theta_2}\xi_s^{i_2}\|_{4p}^2 ds \\ \le & C(t-r)^4, \end{aligned}\tag{5.39}$$

and

$$\begin{aligned}&\Big\| \int_\theta^t \gamma_t^{ii_1} d\big(D^{(k_1,k_2)}_{\theta_1,\theta_2} M_s^{i_1 j}\big)\Big\|_{2p}^2 \\ \le & c_d\Big(\sum_{j_1,j_2,j_3=1}^d \Big\| \int_\theta^t \int_{\mathbb{R}^d} \gamma_s^{ii_1}\partial_{i_1,j_2,j_3} h^{jj_1}(y-\xi_r)\, D^{(k_1)}_{\theta_1}\xi_s^{j_2} D^{(k_2)}_{\theta_2}\xi_s^{j_3} W^{j_1}(ds,dy)\Big\|_{2p}^2 \\ & + \sum_{j_1,j_2=1}^d \Big\| \int_\theta^t \int_{\mathbb{R}^d} \gamma_s^{ii_1}\partial_{i_1,j_2} h^{jj_1}(y-\xi_s)\, D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi_s^{j_2} W^{j_1}(ds,dy)\Big\|_{2p}^2 \Big) := c_d\,(I_1+I_2)\,. \end{aligned}$$

We estimate I_1, I_2 as follows:

$$I_1 \le d\|h\|_{3,2}^2 \sum_{j_2,j_3=1}^d \int_\theta^t \|\gamma_s^{ii_1}\|_{6p}^2 \|D^{(k_1)}_{\theta_1}\xi_s^{j_2}\|_{6p}^2 \|D^{(k_2)}_{\theta_2}\xi_s^{j_3}\|_{6p}^2 ds \le C(t-r)^3,$$

and

$$I_2 \le d\|h\|_{3,2}^2 \sum_{j_2=1}^d \int_\theta^t \|\gamma_s^{ii_1}\|_{4p}^2 \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi_s^{j_2}\|_{4p}^2 ds \le C(t-r)^4 \le CT(t-r)^3.$$

Thus we have

$$\Big\| \int_\theta^t \gamma_t^{ii_1} d\big(D^{(k_1,k_2)}_{\theta_1,\theta_2} M_s^{i_1 j}\big)\Big\|_{2p}^2 \le C(t-r)^3. \tag{5.40}$$

Therefore, combine (5.37)–(5.40), we have

$$\sum_{i,j=1}^d \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma_t^{ij}\|_{2p}^2 \le c_1(t-r)^3 + c_2 \sum_{i,j=1}^d \int_\theta^t \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma_s^{ij}\|_{2p}^2 ds.$$

By Gronwall's lemma, we have

$$\sum_{i,j=1}^d \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma_t^{ij}\|_{2p}^2 \le C(t-r)^3, \tag{5.41}$$

which implies

$$\big\| \|D^2\gamma_t^{ij}\|_{H^{\otimes 2}}\big\|_{2p} \le C(t-r)^{\frac{5}{2}}.$$

By taking the second Malliavin derivative of $\gamma_t\sigma_t \equiv I$, we have

$$D^2\sigma_t^{ij} = -\sum_{i_1,i_2=1}^d \sigma_t^{ii_1}\big(D^2\gamma_t^{i_1i_2}\sigma_t^{i_2j} + D\gamma_t^{i_1i_2}\otimes D\sigma_t^{i_2j} + D\sigma_t^{i_2j}\otimes D\gamma_t^{i_1i_2}\big). \tag{5.42}$$

Then, (5.14) can be deduced by (5.9), (5.11), (5.12), (5.31) and (5.42). □

Proof of (5.15). For any $\theta_1, \theta_2, \theta_3 \in [r, t]$, let $\theta = \theta_1 \vee \theta_2 \vee \theta_3$. Taking the Malliavin derivative on both sides of (5.33), we have

$$
\begin{aligned}
D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi^i_t = & \sum_{j_1,j_2,j_3=1}^{d} \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_2}_s D^{(k_3)}_{\theta_3}\xi^{j_3}_s W^{j_1}(ds,dy) \\
& - \sum_{j_1,j_2=1}^{d} \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2} h^{ij_1}(y-\xi_s) D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi^{j_2}_s W^{j_1}(ds,dy) \\
& - \sum_{j_1,\dots,j_4=1}^{d} \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3,j_4} h^{ij_1}(y-\xi_s) D^{(k_1)}_{\theta_1}\xi^{j_2}_s D^{(k_2)}_{\theta_2}\xi^{j_3}_s D^{(k_3)}_{\theta_3}\xi^{j_4}_s W^{j_1}(ds,dy) \\
& + \sum_{j_1,j_2,j_3=1}^{d} \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D^{(k_1,k_3)}_{\theta_1,\theta_3}\xi^{j_2}_s D^{(k_2)}_{\theta_2}\xi^{j_3}_s W^{j_1}(ds,dy) \\
& + \sum_{j_1,j_2,j_3=1}^{d} \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D^{(k_1)}_{\theta_1}\xi^{j_2}_s D^{(k_2,k_3)}_{\theta_2,\theta_3}\xi^{j_3}_s W^{j_1}(ds,dy). \quad (5.43)
\end{aligned}
$$

By (5.17), (5.36), Burkholder-Davis-Gundy's, Minkowski's, and Hölder's inequalities, we have the following inequalities:

$$
\begin{aligned}
& \Big\| \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_2}_s D^{(k_3)}_{\theta_3}\xi^{j_3}_s W^{k_1}(ds,dy) \Big\|^2_{2p} \\
\le & c_p \|h\|^2_{3,2} \int_\theta^t \big\| D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_2}_s \big\|^2_{4p} \big\| D^{(k_3)}_{\theta_3}\xi^{j_3}_s \big\|^2_{4p} ds \le C(t-r)^2, \quad (5.44)
\end{aligned}
$$

$$
\Big\| \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2} h^{ij_1}(y-\xi_s) D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi^{j_2}_s W^{j_1}(ds,dy) \Big\|^2_{2p} \le c_p \|h\|^2_{3,2} \int_\theta^t \big\| D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi^{j_2}_s \big\|^2_{2p} ds, \quad (5.45)
$$

and

$$
\begin{aligned}
& \Big\| \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3,j_4} h^{ij_1}(y-\xi_s) D^{(k_1)}_{\theta_1}\xi^{j_2}_s D^{(k_2)}_{\theta_2}\xi^{j_3}_s D^{(k_3)}_{\theta_3}\xi^{j_3}_s W^{j_1}(ds,dy) \Big\|^2_{2p} \\
\le & c_p \|h\|^2_{3,2} \int_\theta^t \big\| D^{(k_1)}_{\theta_1}\xi^{j_2}_s \big\|^2_{6p} \big\| D^{(k_2)}_{\theta_2}\xi^{j_3}_s \big\|^2_{6p} \big\| D^{(k_3)}_{\theta_3}\xi^{j_3}_s \big\|^2_{6p} ds \le C(t-r). \quad (5.46)
\end{aligned}
$$

Thus combining (5.43)–(5.46), by Jensen's inequality, we have

$$
\sum_{i=1}^d \big\| D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi^i_t \big\|^2_{2p} \le c_1 \sum_{i=1}^d \int_\theta^t \big\| D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi^i_t \big\|^2_{2p} ds + c_2(t-r).
$$

Then, the following inequality follows from Gronwall's lemma

$$
\sum_{i=1}^d \big\| D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi^i_t \big\|^2_{2p} \le C(t-r). \quad (5.47)
$$

Therefore, (5.15) is a consequence of (5.47). □

In the next lemma, we derive estimates for the moments of increments of the derivatives of ξ_t and σ_t.

Lemma 5.4. *For any $p \geq 1$, $0 \leq r < s < t \leq T$, and $1 \leq i, j \leq d$, there exists a constant $C > 0$ depends on T, d, p, and $\|h\|_{3,2}$, such that*

$$\max_{1\leq i\leq d} \left\| \|D\xi_t^i - D\xi_s^i\|_H \right\|_{2p} \leq C(t-s)^{\frac{1}{2}}, \tag{5.48}$$

$$\max_{1\leq i,j\leq d} \left\| \sigma_t^{ij} - \sigma_s^{ij} \right\|_{2p} \leq C(t-r)^{-\frac{1}{2}}(s-r)^{-1}(t-s)^{\frac{1}{2}}, \tag{5.49}$$

$$\max_{1\leq i,j\leq d} \left\| \|D\sigma_t^{ij} - D\sigma_s^{ij}\|_H \right\|_{2p} \leq C(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}, \tag{5.50}$$

$$\max_{1\leq i\leq d} \left\| \|D\xi_t^i - D^2\xi_s^i\|_{H^{\otimes 2}} \right\|_{2p} \leq C(t-r)(t-s)^{\frac{1}{2}}. \tag{5.51}$$

Proof of (5.48). By (5.4), we have

$$D_\theta^{(k)}\xi_t^i - D_\theta^{(k)}\xi_s^i = \delta_{ik}\mathbf{1}_{[s,t]}(\theta) - \sum_{j=1}^d \int_{\theta\vee s}^t D_\theta^{(k)}\xi_u^j dM_u^{ji}.$$

Thus by (5.17), Burkholder-Davis-Gundy's, Jensen's, and Minkowski's inequalities, we have

$$\left\|D_\theta^{(k)}\xi_t^i - D_\theta^{(k)}\xi_s^i\right\|_{2p}^2 \leq C\left[\delta_{ik}\mathbf{1}_{[s,t]}(\theta) + (t-s)\right].$$

Thus we can show (5.48) by Minkowski's inequality:

$$\begin{aligned}\left\| \|D\xi_t^i - D\xi_s^i\|_H \right\|_{2p}^2 &\leq \sum_{k=1}^d \int_r^t \left\|D_\theta^{(k)}\xi_t^i - D_\theta^{(k)}\xi_s^i\right\|_{2p}^2 d\theta \\ &\leq \sum_{k=1}^d C\Big(\int_s^t \delta_{ik} d\theta + \int_r^t (t-s)d\theta\Big) \leq C(t-s).\end{aligned}$$ □

Proof of (5.49). Note that $\sigma_t - \sigma_s = \sigma_t(\gamma_s - \gamma_t)\sigma_s$. Then, by (5.11) and Hölder's inequality, it suffices to estimate the moment of $\gamma_t - \gamma_s$. By (5.22), we have

$$\begin{aligned}\gamma_t^{ij} - \gamma_s^{ij} =& \delta_{ij}(t-s) - \sum_{k_1=1}^d \int_s^t \gamma_u^{ik_1} dM_u^{k_1 j} - \sum_{k_2=1}^d \int_s^t \gamma_u^{jk_2} dM_u^{k_2 i} \\ &+ \sum_{k_1,k_2=1}^d Q_{k_2,j}^{i,k_1} \int_s^t \gamma_u^{k_1k_2} du.\end{aligned}$$

Then, by (5.26), Minkowski's, Jensen's, and Burkholder-Davis-Gundy's inequalities, for all $1 \leq i, j \leq d$, we have

$$\begin{aligned}\left\|\gamma_t^{ij} - \gamma_s^{ij}\right\|_{2p}^2 \leq& C\left((t-s)^2 + (t-r)^2(t-s) + (t-r)^2(t-s)^2\right) \\ \leq& C(1+T)^2(t-r)(t-s). \end{aligned} \tag{5.52}$$

Then, (5.49) is a consequence of (5.11) and (5.52). □

Proof of (5.50). By (5.23), we have the following equation:

$$\begin{aligned}D_\theta^{(k)}\gamma_t^{ij} - D_\theta^{(k)}\gamma_s^{ij} =& -\sum_{k_1=1}^d \int_{\theta\vee s}^t D_\theta^{(k)}\gamma_u^{ik_1} dM_u^{k_1 j} - \sum_{k_1=1}^d \int_{\theta\vee s}^t \gamma_u^{ik_1} d\big(D_\theta^{(k)} M_u^{k_2 j}\big) \\ &- \sum_{k_2=1}^d \int_{\theta\vee s}^t D_\theta^{(k)}\gamma_u^{k_2 j} dM_u^{k_2 i} - \sum_{k_2=1}^d \int_{\theta\vee s}^t \gamma_u^{k_2 j} d\big(D_\theta^{(k)} M_u^{k_2 i}\big) \\ &+ \sum_{k_1,k_2=1}^d \Big(Q_{k_2,j}^{k_1,i} \int_{\theta\vee s}^t D_\theta^{(k)}\gamma_u^{k_1k_2} du\Big).\end{aligned}$$

Then, by (5.17), (5.26), and (5.30), Burkholder-Davis-Gundy's, Jensen's, Minkowski's, and Cauchy-Schwarz's inequalities, we have

$$\begin{aligned}\big\|D_\theta^{(k)}\gamma_t^{ij}-D_\theta^{(k)}\gamma_s^{ij}\big\|_{2p}^2 &\le c_{d,p}\|h\|_{3,2}^2\Big[\sum_{k_1=1}^d\int_{\theta\vee s}^t\big\|D_\theta^{(k)}\gamma_u^{ik_1}\big\|_{2p}^2du\\ &\quad+\sum_{k_2=1}^d\int_{\theta\vee s}^t\big\|\gamma_u^{ik_1}\big\|_{4p}^2\big\|D_\theta^{(k)}\xi_u^{k_2}\big\|_{4p}^2du+(t-s)\int_{\theta\vee s}^t\big\|D_\theta^{(k)}\gamma_u^{k_1k_2}\big\|_{2p}^2du\Big]\\ &\le C(t-r)^2(t-s).\end{aligned}$$

This implies

$$\big\|\|D\gamma_t^{ij}-D\gamma_s^{ij}\|_H\big\|_{2p}\le C(t-r)^{\frac32}(t-s)^{\frac12}. \tag{5.53}$$

By (5.32), we have

$$\begin{aligned}D\sigma_t^{ij}-D\sigma_s^{ij} &=\sum_{i_1,i_2=1}^d\left(\sigma_t^{ii_1}D\gamma_t^{i_1i_2}\sigma_t^{i_2j}-\sigma_s^{ii_1}D\gamma_s^{i_1i_2}\sigma_s^{i_2j}\right)\\ &=\sum_{i_1,i_2=1}^d\sigma_t^{ii_1}\left(D\gamma_t^{i_1i_2}-D\gamma_s^{i_1i_2}\right)\sigma_t^{i_2j}+\sum_{i_1,i_2=1}^d\left(\sigma_t^{ii_1}-\sigma_s^{ii_1}\right)D\gamma_s^{i_1i_2}\sigma_t^{i_2j}\\ &\quad+\sum_{i_1,i_2=1}^d\sigma_s^{ii_1}D\gamma_s^{i_1i_2}\left(\sigma_t^{i_2j}-\sigma_s^{i_2j}\right).\end{aligned}$$

Thus (5.50) follows from (5.9), (5.11), (5.31), (5.49) and (5.53). □

Proof of (5.51). Let $\theta=\theta_1\vee\theta_2$, by (5.33), we have the following equation:

$$\begin{aligned}D_{\theta_1,\theta_2}^{(k_1,k_2)}\xi_t^i-D_{\theta_1,\theta_2}^{(k_1,k_2)}\xi_s^i &=-\sum_{j_1,j_2=1}^d\int_{\theta\vee s}^t\int_{\mathbb{R}^d}\partial_{j_2}h^{ij_1}(y-\xi_u)D_{\theta_1,\theta_2}^{(k_1,k_2)}\xi_u^{j_2}W^{j_1}(du,dy)\\ &\quad+\sum_{j_1,j_2,j_3=1}^d\int_{\theta\vee s}^t\int_{\mathbb{R}^d}\partial_{j_2,j_3}h^{ij_1}(y-\xi_u)D_{\theta_1}^{(k_1)}\xi_u^{j_2}D_{\theta_2}^{(k_2)}\xi_u^{j_3}W^{j_1}(du,dy).\end{aligned}$$

As a consequence, by (5.17), (5.36), Burkholder-Davis-Gundy's, Minkowski's, and Cauchy-Schwarz's inequalities, we have

$$\begin{aligned}\big\|D_{\theta_1,\theta_2}^{(i,j)}\xi_t^k-D_{\theta_1,\theta_2}^{(i,j)}\xi_s^k\big\|_{2p}^2 \le& c_p\Big[\sum_{j_1=1}^d\|h\|_{3,2}^2\int_{\theta\vee s}^t\big\|D_{\theta_1,\theta_2}^{(i,j)}\xi_u^{j_1}\big\|_{2p}^2du\\ &+\sum_{j_1,j_2}^d\|h\|_{3,2}^2\int_{\theta\vee s}^t\big\|D_{\theta_1}^{(i)}\xi_u^{j_1}\big\|_{4p}^2\big\|D_{\theta_2}^{(j)}\xi_u^{j_2}\big\|_{4p}^2du\Big]\\ \le& C(t-s).\end{aligned} \tag{5.54}$$

Therefore, we obtain (5.51) by integrating (5.54) and Minkowski's inequality. □

We define the following functionals of ξ_t

$$H_{(i)}(\xi_t,1)=-\sum_{j=1}^d\delta\big(\sigma_t^{ji}D\xi_t^j\big),\quad 1\le i\le d, \tag{5.55}$$

and

$$H_{(i,j)}(\xi_t,1) = -\sum_{k=1}^{d} \sigma\big(H_{(i)}(\xi_t,1)\sigma_t^{kj} D\xi_t^k\big), \quad 1 \le i,j \le d. \tag{5.56}$$

A more detailed description of these functionals can be seen in Appendix A. In the next lemma, we establish moment estimates for the functionals $H_{(i)}(\xi_t,1)$ and $H_{(i,j)}(\xi_t,1)$.

Lemma 5.5. *Suppose that $h \in H_2^3(\mathbb{R}^d;\mathbb{R}^d \otimes \mathbb{R}^d)$, then the following inequalities are satisfied:*

$$\max_{1\le i\le d} \big\|H_{(i)}(\xi_t,1)\big\|_{2p} \le C(t-r)^{-\frac{1}{2}}, \tag{5.57}$$

$$\max_{1\le i,j\le d} \big\|H_{(i,j)}(\xi_t,1)\big\|_{2p} \le C(t-r)^{-1}. \tag{5.58}$$

Proof. Due to Meyer's inequality (see e.g. Proposition 1.5.4 and 2.1.4 of Nualart [24]), it suffices to estimate

$$\big\|\|\sigma_t^{ji} D\xi_t^j\|_H\big\|_{2p},\ \big\|\|D\big(\sigma_t^{ji} D\xi_t^j\big)\|_{H^{\otimes 2}}\big\|_{2p} \text{ and } \big\|\|D^2\big(\sigma_t^{ji} D\xi_t^j\big)\|_{H^{\otimes 3}}\big\|_{2p}.$$

By (5.10) and Lemma 5.2–5.3, we have

$$\big\|\|\sigma_t^{ji} D\xi_t^j\|_H\big\|_{2p} \le \big\|\sigma_t^{ji}\big\|_{4p} \big\|\|D\xi_t^j\|_H\big\|_{4p} \le C(t-r)^{-\frac{1}{2}},$$

$$\begin{aligned}
\big\|\|D\big(\sigma_t^{ji} D\xi_t^j\big)\|_{H^{\otimes 2}}\big\|_{2p} \le& \big\|\|D\sigma_t^{ji} \otimes D\xi_t^j\|_{H^{\otimes 2}}\big\|_{2p} + \big\|\|\sigma_t^{ji} D^2\xi_t^j\|_{H^{\otimes 2}}\big\|_{2p} \\
\le& \big\|\|D\sigma_t^{ji}\|_H\big\|_{4p} \big\|\|D\xi_t^j\|_H\big\|_{4p} + \big\|\sigma_t^{ji}\big\|_{4p} \big\|\|D^2\xi_t^j\|_{H^{\otimes 2}}\big\|_{4p} \\
\le& C(t-r)^{\frac{1}{2}},
\end{aligned}$$

and

$$\begin{aligned}
\big\|\|D^2\big(\sigma_t^{ji} D\xi_t^j\big)\|_{H^{\otimes 3}}\big\|_{2p} \le& \big\|\|D^2\sigma_t^{ji} \otimes D\xi_t^j\|_{H^{\otimes 2}}\big\|_{2p} \\
&+ \big\|\|D\sigma_t^{ji} \otimes D^2\xi_t^j\|_{H^{\otimes 2}}\big\|_{2p} + \big\|\|\sigma_t^{ji} D^3\xi_t^j\|_{H^{\otimes 2}}\big\|_{2p} \\
\le& C(t-r).
\end{aligned}$$

The above inequalities hold for all $1 \le i,j \le d$. Then, (5.57) and (5.58) follows. $\square$

The next lemma provides the moment estimate for the increment of $H_{(i)}(\xi_t,1)$.

Lemma 5.6. *Suppose that $h \in H_2^3(\mathbb{R}^d;\mathbb{R}^d \otimes \mathbb{R}^d)$. Then,*

$$\max_{1\le i\le d} \big\|H_{(i)}(\xi_t,1) - H_{(i)}(\xi_s,1)\big\|_{2p} \le C(s-r)^{-\frac{1}{2}}(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{5.59}$$

Proof. Notice that, by definition, we have

$$\begin{aligned}
H_{(i)}(\xi_t,1) - H_{(i)}(\xi_s,1) =& -\sum_{j=1}^{d} \delta\big(\sigma_t^{ji} D\xi_t^j\big) + \sum_{j=1}^{d} \delta\big(\sigma_s^{ji} D\xi_s^j\big) \\
=& -\sum_{j=1}^{d} \delta\big(\sigma_t^{ji} D\xi_t^j - \sigma_s^{ji} D\xi_s^j\big).
\end{aligned}$$

Thus by Meyer's inequality again, it suffices to estimate

$$I_1 := \big\|\|\sigma_t^{ji} D\xi_t^j - \sigma_s^{ji} D\xi_s^j\|_H\big\|_{2p} \text{ and } I_2 := \big\|\|D\big(\sigma_t^{ji} D\xi_t^j - \sigma_s^{ji} D\xi_s^j\big)\|_{H^{\otimes 2}}\big\|_{2p}.$$

For I_1, we have

$$I_1 \leq \big\|\big\|(\sigma_t^{ji} - \sigma_s^{ji}) D\xi_s^k\big\|_H\big\|_{2p} + \big\|\|\sigma_t^{ji}(D\xi_t^j - D\xi_s^j)\|_H\big\|_{2p}.$$

Notice that by Lemmas 5.2–5.4, we can write

$$\begin{aligned}\big\|\big\|(\sigma_t^{ji} - \sigma_s^{ji}) D\xi_s^j\big\|_H\big\|_{2p} \leq & \|\sigma_t^{ji} - \sigma_s^{ji}\|_{4p}\big\|\|D\xi_s^j\|_H\big\|_{4p} \\ \leq & C(t-r)^{-\frac{1}{2}}(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}\end{aligned}$$

and

$$\begin{aligned}&\big\|\big\|\sigma_t^{ji}(D\xi_t^j - D\xi_s^j)\big\|_H\big\|_{2p} \leq \|\sigma_t^{ji}\|_{4p}\big\|\big\|D\xi_t^j - D\xi_s^j\big\|_H\big\|_{4p} \\ \leq & C(t-r)^{-1}(t-s)^{\frac{1}{2}} \leq C(t-r)^{-\frac{1}{2}}(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}.\end{aligned}$$

Thus combining the above inequalities, we have the following estimate for I_1:

$$I_1 \leq C(t-r)^{-\frac{1}{2}}(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{5.60}$$

By Lemmas 5.2–5.4, we have the following estimate for I_2:

$$\begin{aligned}I_2 \leq & \big\|D\sigma_t^{ji} \otimes D\xi_t^j - D\sigma_t^{ji} \otimes D\xi_s^j\big\|_{2p,H^{\otimes 2}} + \big\|\sigma_t^{ji} D^2\xi_s^j - \sigma_s^{ji} D^2\xi_s^j\big\|_{2p,H^{\otimes 2}} \\ \leq & \big\|\|D\sigma_t^{ji}\|_H\big\|_{4p}\big\|\big\|(D\xi_t^j - D\xi_s^j)\big\|_H\big\|_{4p} + \big\|\big\|(D\sigma_t^{ji} - D\sigma_s^{ji})\big\|_H\big\|_{4p}\big\|\big\|D\xi_s^j\big\|_H\big\|_{4p} \\ & + \big\|\sigma_t^{ji}\big\|_{4p}\big\|\|D^2\xi_t^j - D^2\xi_s^j\|_{H^{\otimes 2}}\big\|_{4p} + \big\|\sigma_t^{ji} - \sigma_s^{ji}\big\|_{4p}\big\|\|D^2\xi_s^j\|_{H^{\otimes 2}}\big\|_{2p} \\ \leq & C(t-s)^{\frac{1}{2}}.\end{aligned} \tag{5.61}$$

Therefore, (5.59) follows from (5.60), (5.61) and Meyer's inequality. □

The next lemma shows that ξ is a d-dimensional Gaussian process in the whole probability space. Notice that, however, conditionally on W, the process ξ is no longer Gaussian, because it is the solution to a nonlinear SDE.

Lemma 5.7. *The process ξ given by equation (5.1) is a d-dimensional Gaussian process, with mean x and covariance matrix*

$$\Sigma_{s,t} = (t \wedge s - r)(I + \rho(0)), \tag{5.62}$$

where $\rho(0)$ is defined in (2.2). Moreover, the probability density of ξ_t, denoted by $p_{\xi_t}(y)$, is bounded by a Gaussian density:

$$p_{\xi_t}(y) \leq (2\pi(t-r))^{-\frac{d}{2}} \exp\Big(-\frac{k|x-y|^2}{t-r}\Big), \tag{5.63}$$

where

$$k = [2(d\|h\|_{2,3}^2 + 1)]^{-1}. \tag{5.64}$$

Proof. Since B is a d-dimensional Brownian motion and W is a d-dimensional space-time white Gaussian random field independent of B, then $\xi = \{\xi_t, r \leq t \leq T\}$ is a square integrable d-dimensional martingale. The quadratic covariation of ξ is given by

$$\begin{aligned}\langle \xi^i, \xi^j \rangle_t = & \delta_{ij}(t-r) + \sum_{k=1}^d \int_r^t \int_{\mathbb{R}^d} h^{ik}(\xi_s - y) h^{jk}(\xi_s - y) dy ds \\ = & \big(\delta_{ij} + \rho^{ij}(0)\big)(t-r).\end{aligned} \tag{5.65}$$

Note that $\rho(0)$ is a symmetric nonnegative definite matrix. As a consequence, $I + \rho(0)$ is strictly positive definite, and thus nondegenerate. Therefore, we can find a nondegenerate matrix M, such that $M^*(I+\rho(0))M = I$. Let $\eta = M\xi$, then $\eta = \{\eta_t, t \in [0,T]\}$ is a martingale with quadratic covariation

$$\langle \eta^i, \eta^j \rangle_t = (t-r) \sum_{k_1,k_2=1}^{d} M^{ik_1} M^{jk_2} \langle \xi^{k_1}, \xi^{k_2} \rangle_t = \delta_{ij}(t-r).$$

By Levy's martingale characterization, η is a d-dimensional Brownian motion. Then, $\xi = M^{-1}\eta$ is a Gaussian process, with covariance matrix (5.62).

Since for any $t > r$, $\Sigma_t := \Sigma_{t,t} = (t-r)(I + \rho(0))$ is symmetric and positive definite, the probability density of the Gaussian random vector ξ_t is given by

$$p_{\xi_t}(y) = \frac{1}{\sqrt{(2\pi)^d |\Sigma_t|}} \exp\Big(-\frac{1}{2}(y-x)^* \Sigma_t^{-1} (y-x) \Big). \tag{5.66}$$

Recall that $\rho(0)$ is symmetric and nonnegative definite. Then it has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0$. Let λ be the diagonal matrix with diagonal elements $\lambda_1, \ldots, \lambda_d$. There is an orthogonal matrix U, such that $\rho(0) = U^* \lambda U$. Let k be defined in (5.64). It follows that

$$\lambda_1 + 1 \leq \sum_{i,j=1}^{d} |\rho^{ij}(0)| + 1 \leq \|\rho\|_\infty + 1 \leq d\|h\|_{3,2}^2 + 1 = \frac{1}{2k}.$$

Thus for any nonzero $x \in \mathbb{R}^d$, we have

$$\begin{aligned} \frac{1}{2} x^* \Sigma_t^{-1} x - \frac{k}{t-r} x^* x =& \frac{1}{2} x^* \Big(\Sigma_t^{-1} - \frac{2k}{t-r} I \Big) x \\ =& \frac{1}{2(t-r)} x^* U^* \big((I+\lambda)^{-1} - 2kI \big) U x \geq 0, \end{aligned}$$

because $(I+\lambda)^{-1} - 2kI$ is a nonnegative diagonal matrix. Thus for any $x, y \in \mathbb{R}^d$, $t > r$, we have

$$\exp\Big(-\frac{1}{2}(y-x)^* \Sigma_t^{-1} (y-x) \Big) \leq \exp\Big(-\frac{k|x-y|^2}{t-r} \Big). \tag{5.67}$$

On the other hand, we have

$$|\Sigma_t| = |U^* \left(I + \lambda \right) U (t-r)| \geq (t-r)^d. \tag{5.68}$$

Therefore, we obtain (5.63) by plugging (5.67)–(5.68) into (5.66). □

Denote by $\mathbb{P}^W$, $\mathbb{E}^W$, and $\|\cdot\|_p^W$ the probability, expectation and L^p-norm conditional on W. The following two propositions are estimates for the conditional distribution of ξ.

Proposition 5.8. *Fix $0 \leq r < t \leq T$ and recall that $\xi_r = \xi_r^{r,x} = x$. Let $c > 0$, choose $\rho \in (0, c\sqrt{t-r}]$. Then, for any $p_1, p_2 \geq 1$ and $y \in \mathbb{R}^d$, there exists $C > 0$, depending on p_1, p_2, c, $\|h\|_2$, and d, such that*

$$\big\| \mathbb{P}^W (|\xi_t - y| \leq \rho)^{\frac{1}{p_1}} \big\|_{p_2} \leq C \exp\Big(-\frac{k|x-y|^2}{p(t-r)} \Big), \tag{5.69}$$

where k is defined in (5.64) and $p = p_1 \vee p_2$.

Proof. Let $p = p_1 \vee p_2$. Then, by Jensen's inequality, we have

$$\big\| \mathbb{P}^W (|\xi_t - y| \leq \rho)^{\frac{1}{p_1}} \big\|_{p_2} = \big\| \big\| \mathbf{1}_{\{\xi_t - y| \leq \rho\}} \big\|_{p_1}^W \big\|_{p_2} \leq \big\| \mathbf{1}_{\{\xi_t - y| \leq \rho\}} \big\|_p.$$

We consider two different cases.

(i) Suppose that $2\rho \le |x-y|$. If $|\xi_t - y| \le \rho \le c\sqrt{t-r}$, then

$$|\xi_t - x| \ge |x-y| - |\xi_t - y| \ge |x-y| - \rho \ge \frac{|x-y|}{2},$$

and equivalently $\{|\xi_t - y| < \rho\} \subset \{|\xi_t - x| \ge \frac{|x-y|}{2}\}$. Then, by Lemma 5.7, we have

$$\begin{aligned}\left\|\mathbb{P}^W(|\xi_t - y| \le \rho)^{\frac{1}{p_1}}\right\|_{p_2} =& \left\|\mathbf{1}_{\{|\xi_t - x| \ge \frac{|x-y|}{2}\} \cap \{|\xi_t - y| < \rho\}}\right\|_p \le C\Big[V_d \rho^d \sup_{|z-x| \ge \frac{|x-y|}{2}} p_{\xi_t}(z)\Big]^{\frac{1}{p}} \\ \le& C\Big[V_d c^d (2\pi)^{-\frac{d}{2}} \exp\Big(-\frac{k|x-y|^2}{t-r}\Big)\Big]^{\frac{1}{p}}, \end{aligned} \tag{5.70}$$

where $V_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(1+\frac{d}{2})}$ is the volume of the unit sphere in $\mathbb{R}^d$.

(ii) On the other hand, suppose that $2\rho > |x-y|$. Then $|x-y| \le 2\rho \le 2c\sqrt{t-r}$. Thus by Lemma 5.7 again, we have

$$\begin{aligned}\left\|\mathbb{P}^W(|\xi_t - y| \le \rho)^{\frac{1}{p_1}}\right\|_{p_2} \le& C\big(V_d \rho^d (2\pi(t-r))^{-\frac{d}{2}}\big)^{\frac{1}{p}} \\ \le& C\big(V_d c^d (2\pi)^{-\frac{d}{2}}\big)^{\frac{1}{p}} \exp\Big(\frac{4kc^2}{p} - \frac{4kc^2}{p}\Big) \\ \le& C\big(V_d c^d (2\pi)^{-\frac{d}{2}}\big)^{\frac{1}{p}} e^{\frac{4kc^2}{p}} \exp\Big(-\frac{k|x-y|^2}{p(t-r)}\Big). \end{aligned} \tag{5.71}$$

Therefore, (5.69) follows from (5.70)–(5.71). □

Denote by $p^W(r,x;t,y)$ the transition probability density of ξ conditional on W. In other words, $p^W(r,x;t,y)$ is the conditional probability density of $\xi_t = \xi_t^{r,x}$. The existence of $p^W(r,x;t,y)$ is guaranteed by Theorem A.3. By applying Theorem A.4, we can further obtain the following estimate:

Proposition 5.9. *For any $0 \le r < t \le T$, $p \ge 1$, and $y \in \mathbb{R}^d$, there exist $C > 0$, depending on T, d, $\|h\|_{3,2}$, p, and q, such that*

$$\left\|p^W(r,x;t,y)\right\|_{2p} \le C \exp\Big(-\frac{k|x-y|^2}{6pd(t-r)}\Big)(t-r)^{-\frac{d}{2}}, \tag{5.72}$$

where k is defined in (5.64).

Proof. Choose $p_1 \in (d, 3pd]$, let $p_2 = 2p_1$, and $p_3 = \frac{p_1 p_2}{p_2 - p_1} = p_2$. Then, by (A.12) and Hölder's inequality, we have

$$\begin{aligned}\left\|p^W_{\xi_t}(y)\right\|_{2p} \le C \max_{1\le i\le d}\Big\{&\left\|\mathbb{P}^W(|\xi_t - y| < 2\rho)^{\frac{1}{p_2}}\right\|_{6p} \left\|\big(\|H_{(i)}(\xi_t,1)\|^W_{p_1}\big)^{d-1}\right\|_{6p} \\ &\times \Big[\frac{1}{\rho} + \left\|\|H_{(i)}(\xi_t,1)\|^W_{p_2}\right\|_{6p}\Big]\Big\}. \end{aligned} \tag{5.73}$$

By Jensen's inequality, we have for any $1 \le i \le d$

$$\left\|\big(\|H_{(i)}(\xi_t,1)\|^W_{p_1}\big)^{d-1}\right\|_{6p} \le \left\|H_{(i)}(\xi_t,1)\right\|^{d-1}_{6p\vee p_1} \le \left\|H_{(i)}(\xi_t,1)\right\|^{d-1}_{6pd}, \tag{5.74}$$

and

$$\left\|\|H_{(i)}(\xi_t,1)\|^W_{p_2}\right\|_{6p} \le \left\|H_{(i)}(\xi_t,1)\right\|_{6pd}. \tag{5.75}$$

$\rho = \frac{\sqrt{t-r}}{4}$. (5.72) is a consequence of (5.73)–(5.75), Lemma 5.5, and Proposition 5.8. □

6 A conditional convolution representation

In this section, we follow the idea of Li et al. (see Section 3 of [20]) to obtain a conditional convolution formulation of the SPDE (3.1). Consider the following SPDE:

$$u_t(x) = \int_{\mathbb{R}^d} \mu(z) p^W(0,z;t,x)dz + \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) u_r(z) V(dr,dz), \tag{6.1}$$

where W and V are the same random fields as in (3.1), p^W is the transition density of ξ_t given by (5.1) conditional on W.

In order to define the stochastic integral on the right-hand side of (6.1), we introduce the following filtrations. First, for any $t \in [0,T]$, we set

$$\mathcal{F}_t := \sigma\{W(s,x),(s,x) \in [0,T] \times \mathbb{R}^d\} \vee \sigma\{V(s,x),(s,x) \in [0,t] \times \mathbb{R}^d\}. \tag{6.2}$$

The stochastic integral in (6.1) is defined for all $\mathcal{F}_t$-adapted processes. But later we will see that the solution u, as a limit of Picard iteration, is in fact adapted to a smaller filtration defined as follows: for any $t \in [0,T]$,

$$\mathcal{G}_t := \sigma\{W(s,x),(s,x) \in [0,t] \times \mathbb{R}^d\} \vee \sigma\{V(s,x),(s,x) \in [0,t] \times \mathbb{R}^d\}. \tag{6.3}$$

Definition 6.1. *A random field $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is said to be a strong solution to the SPDE (6.1), if the following properties are satisfied:*

(i) u is $\mathcal{G}_t$-adapted.

(ii) u is square integrable in the following sense:

$$\mathbb{E}\Big(\int_0^T \int_{\mathbb{R}^d} |u_t(x)|^2 dxdt\Big) < \infty. \tag{6.4}$$

(iii) The stochastic integral in (6.1) is defined as Walsh's integral and the equality holds almost surely for all $t \in [0,T]$ and almost every $x \in \mathbb{R}^d$.

Lemma 6.2. *Assume that κ and μ are bounded. Then the SPDE (6.1) has a unique strong solution (in the sense of Definition 6.1). Denote the solution by $u = \{u_t(x), 0 \le t \le T, x \in \mathbb{R}^d\}$. Then, for any $p \ge 1$, the following inequality holds:*

$$\sup_{0\le t\le T} \sup_{x\in\mathbb{R}^d} \|u_t(x)\|_{2p} < \infty. \tag{6.5}$$

Proof. We prove the lemma by the Picard iteration. Let $u_0(t,x) \equiv \mu(x)$ and let

$$u_n(t,x) = \int_{\mathbb{R}^d} \mu(z) p^W(0,z;t,x)dz + \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) u_{n-1}(r,z) V(dr,dz), \tag{6.6}$$

for all $n \ge 1$ and $0 \le t \le T$. Since W and V are independent, then V is a martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$. Notice that for any $r \in [0,T]$, $\mathcal{F}_t$ includes all the information of W, and p^W depends only on W. Then, $p^W(r,z;t,x)$ is $\mathcal{F}_r$-measurable, and by induction $u_{n-1}(r,z)$ is $\mathcal{F}_r$-measurable for all $[r,t] \subset [0,T]$ and $x,z \in \mathbb{R}^d$. Thus the stochastic integral is well-defined, and u_n is an $\mathcal{F}_t$-adapted random field. In addition, we know that $p^W(r,z;t,x)$ is $\mathcal{G}_t$-measurable, and by induction we can assume that $u_{n-1}(t)$ is $\mathcal{G}_t$-measurable as well. Thus the stochastic integral in (6.6) is $\mathcal{G}_t$-measurable. Therefore, the limit of $u_n(t,x)$ in $L^2(\Omega)$, if exists, is also $\mathcal{G}_t$-measurable.

Let $d_n(t,x) := u_{n+1}(t,x) - u_n(t,x)$. Then

$$d_n(t,x) := \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) d_{n-1}(r,z) V(dr,dz).$$

For any $p \geq 1$, let

$$d_n^*(t) := \int_{\mathbb{R}^d} \|d_n(t,x)\|_{2p}^2 dx. \tag{6.7}$$

We aim to prove the existence and convergence of $\{u^n\}_{n\geq 1}$ in $L^{2p}(\Omega; L^2(\mathbb{R}^d))$ by showing that $\sqrt{d_n^*(t)}$ is summable in n. Then, we will show that the limit is a solution to (6.1).

By the definition of $u_n(t)$, Burkholder-Davis-Gundy, Minkowski's and Cauchy-Schwarz's inequalities, we have

$$d_n^*(t) \leq c_p \|\kappa\|_\infty \int_{\mathbb{R}^d} \int_0^t \Big(\int_{\mathbb{R}^d} \big\| p^W(r,z;t,x) d_{n-1}(r,z) \big\|_{2p} dz \Big)^2 dr dx. \tag{6.8}$$

By the Markov property, $p^W(r,z;t,x)$ depends only on $\{W(s,z) - W(r,z), s \in (r,t], z \in \mathbb{R}^d\}$. On the other hand, $d_{n-1}(r,z)$ depends on V and $\{W(s,z), s \in [0,r], z \in \mathbb{R}^d\}$. Thus, $p^W(r,z;t,x)$ and $d_{n-1}(r,z)$ are independent. That implies

$$\mathbb{E}\big(|p^W(r,z;t,x) d_{n-1}(r,z)|^{2p}\big) = \mathbb{E}\big(|p^W(r,z;t,x)|^{2p}\big) \mathbb{E}\big(|d_{n-1}(r,z)|^{2p}\big). \tag{6.9}$$

Then, by (6.8), (6.9), Young's convolution inequality, Fubini's theorem and Proposition 5.9, we have

$$\begin{aligned} d_n^*(t) \leq & c_p \|\kappa\|_\infty \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\mathbb{R}^d} \|p^W(r,z_1;t,x)\|_{2p} \|p^W(r,z_2;t,x)\|_{2p} dx \\ & \qquad \times \|d_{n-1}(r,z_1)\|_{2p} \|d_{n-1}(r,z_2)\|_{2p} dz_1 dz_2 dr \\ \leq & C \int_0^t (t-r)^{-\frac{d}{2}} \exp\Big(- \frac{k|z_1 - z_2|^2}{12pd(t-r)} \Big) \|d_{n-1}(r,z_1)\|_{2p} \|d_{n-1}(r,z_2)\|_{2p} dz_1 dz_2 dr \\ \leq & C \int_0^t d_{n-1}^*(r) dr, \end{aligned} \tag{6.10}$$

where $C > 0$ depends on p, T, d, h, and $\|\kappa\|_\infty$.

Thus by iteration, we have

$$d_n^*(t) \leq C^n \int_0^t \int_0^{r_n} \cdots \int_0^{r_2} d_0^*(r_1) dr_1 \cdots dr_n. \tag{6.11}$$

To estimate d_0^*, we observe that

$$\begin{aligned} d_0^*(t) = & \int_{\mathbb{R}^d} \Big\| \int_{\mathbb{R}^d} (\mu(z) - \mu(x)) \, p^W(0,z;t,x) dz \\ & \qquad + \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) \mu(z) V(dr,dz) \Big\|_{2p}^2 dx \\ \leq & 3 \int_{\mathbb{R}^d} \Big\| \int_{\mathbb{R}^d} \mu(z) p^W(0,z;t,x) dz \Big\|_{2p}^2 dx + 3 \int_{\mathbb{R}^d} \Big\| \int_{\mathbb{R}^d} \mu(x) p^W(0,z;t,x) dz \Big\|_{2p}^2 dx \\ & + 3 \int_{\mathbb{R}^d} \Big\| \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) \mu(z) V(dr,dz) \Big\|_{2p}^2 dx. \end{aligned} \tag{6.12}$$

By an argument similar to the proof of (6.10), we can show that $d_0^*(t) < C$. Therefore, we have

$$d_n^*(t) \leq C \int_0^t \int_0^{r_n} \cdots \int_0^{r_2} 1 dr_1 \dots dr_n = C \frac{t^n}{n!}. \tag{6.13}$$

Notice that $\sqrt{d^n(t)}$ is summable in n and the corresponding series is bounded on [0,T]. Therefore, for any fixed $t \in [0,T]$, $\{u_n(t,\cdot)\}_{n\geq 0}$ is convergent in $L^{2p}(\Omega; L^2(\mathbb{R}^d))$. Denote by $u_t(x)$ the limit of this sequence. We claim that $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is a strong solution to (6.1). Clearly u satisfies (6.4) and is $\mathcal{G}_t$-adapted. Therefore, it suffices to show that as $n \to \infty$,

$$\int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,\cdot) u_n(r,z) V(dr,dz) \to \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,\cdot) u(r,z) V(dr,dz), \tag{6.14}$$

in $L^{2p}(\Omega)$ for all $t \in [0,T]$. Actually, by Burkholder-Davis-Gundy's, Minkowski's, Young's convolution inequalities, and the fact that $\{p^W(r,z;t,x), x,z \in \mathbb{R}^d\}$ and $\{u_n(r,z) - u(r,z), z \in \mathbb{R}^d\}$ are independent, we can write

$$\Big\| \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x)\left(u_n(r,z) - u(r,z)\right) V(dr,dz) \Big\|_{2p}^2 \le C \int_0^t \int_{\mathbb{R}^d} \|u_n(r,z) - u(r,z)\|_{2p}^2 dz dr.$$

This implies that (6.14) is true. As we discussed before, the limit $u(t,x)$ is $\mathcal{G}_t$-measurable, it follows that $u(t,x)$ is a strong solution to (6.1).

In order to show the uniqueness, we assume that $v = \{v_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is another strong solution to (6.1). Let $d_t(x) = u_t(x) - v_t(x)$ for any $t \in [0,T]$ and $x \in \mathbb{R}^d$. Then,

$$d_t(x) = \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) d_r(z) V(dr,dz).$$

By the Ito isometry, Minkowski's and Young's convolution inequalities and the fact that the families $\{d_r(x), x \in \mathbb{R}^d\}$ and $\{p^W(r,z;t,x), x,z \in \mathbb{R}^d\}$ are independent, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \|d_t(x)\|_2^2 dx \le& \int_0^t \sup_{x \in \mathbb{R}^d} \|d_r(x)\|_2^2 \Big(\int_{\mathbb{R}^d} \left\|p^W(r,z;t,x)\right\|_2 dz \Big)^2 dr \\ \le& C \int_0^t \int_{\mathbb{R}^d} \|d_r(x)\|_2^2 dx dr. \end{aligned} \tag{6.15}$$

Notice that by definition,

$$\int_{\mathbb{R}^d} \|d_t(x)\|_2^2 dx \le \int_{\mathbb{R}^d} \mathbb{E}|u_t(x)|^2 dx + \int_{\mathbb{R}^d} \mathbb{E}|v_t(x)|^2 dx < \infty,$$

for almost every $t \in [0,T]$. As a consequence of Gronwall's lemma and the fact that $d_0 \equiv 0$, inequality (6.15) implies $d(t,x) \equiv 0$, a.s for almost every $(t,x) \in [0,T] \times \mathbb{R}^d$. It follows that the solution to (6.1) in the sense of Definition 6.1 is unique.

In order to obtain the uniform boundedness (6.5), we need to estimate the following expression when applying the Picard iteration:

$$\widetilde{d_n^*}(t) := \sup_{x \in \mathbb{R}^d} \|d_n(t,x)\|_{2p}^2,$$

instead of $d_n^*(t)$ defined in (6.7). By a similar argument as we did before, the following inequality can be proved:

$$\widetilde{d_n^*}(t) \le C \frac{T^n}{n!},$$

where $C > 0$ is independent of n. Then, inequality (6.5) follows immediately. □

Proposition 6.3. *Assume that κ and μ are bounded. Let $u = \{u_t(x), 0 < t \le T, x \in \mathbb{R}^d\}$ be the unique strong solution to (6.1) in the sense of Definition 6.1. Then, u is the strong solution to (3.1) in the sense of Definition 3.1.*

Proof. Let $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ be the unique solution to the SPDE (6.1), and write $Z(dt,dx) = u_t(x)V(dt,dx)$ for all $t \in [0,T]$ and $x \in \mathbb{R}^d$. Then, it suffices to show that u satisfies the following equation:

$$\begin{aligned}\langle u_t, \phi\rangle =& \langle \mu, \phi\rangle + \int_0^t \langle u_s, A\phi\rangle ds + \int_0^t \int_{\mathbb{R}^d} \langle u_s, \nabla\phi^* h(y-\cdot)\rangle W(ds,dy) \\ &+ \int_0^t \int_{\mathbb{R}^d} \phi(x) Z(ds,dx),\end{aligned} \tag{6.16}$$

for any $\phi \in C_b^2\left(\mathbb{R}^d\right)$.

Denote by

$$\mathbb{E}^W_{s,x}(\phi(\xi_t)) := \mathbb{E}\big(\phi(\xi_t)|W, \xi_s = x\big) = \int_{\mathbb{R}^d} \phi(z) p^W(s,x;t,z)dz.$$

As u is the strong solution to (6.1), the following equations are satisfied

$$\langle u_t, \phi\rangle = \big\langle \mu, \mathbb{E}^W_{0,\cdot}(\phi(\xi_t))\big\rangle + \int_0^t \int_{\mathbb{R}^d} \mathbb{E}^W_{s,z}(\phi(\xi_t)) Z(ds,dz),$$

$$\int_0^t \langle u_s, A\phi\rangle ds = \int_0^t \big\langle \mu, \mathbb{E}^W_{0,\cdot}(A\phi(\xi_s))\big\rangle ds + \int_0^t \int_0^s \int_{\mathbb{R}^d} \mathbb{E}^W_{r,z}(A\phi(\xi_s)) Z(dr,dz) ds,$$

and

$$\begin{aligned}\int_0^t \int_{\mathbb{R}^d} \langle u_s, \nabla\phi^* h(y-\cdot)\rangle W(ds,dy) =& \int_0^t \int_{\mathbb{R}^d} \big\langle \mu, \mathbb{E}^W_{0,\cdot}\big(\nabla\phi(\xi_s)^* h(y-\xi_s)\big)\big\rangle W(ds,dy) \\ &+ \int_0^t \int_{\mathbb{R}^d} \int_0^s \int_{\mathbb{R}^d} \mathbb{E}^W_{r,z}\big((\nabla\phi(\xi_s)^* h(y-\xi_s)\big) Z(dr,dz) W(ds,dy).\end{aligned}$$

Notice that $\phi \in C_b^2(\mathbb{R}^d)$, $h \in H_2^3(\mathbb{R}^d; \mathbb{R}^d \otimes \mathbb{R}^d)$, and $\|u_t(x)\|_2^2$ is integrable on $[0,T] \times \mathbb{R}^d$. These properties allow us to write

$$\mathbb{E}\Big(\int_0^T \int_{\mathbb{R}^d} |\nabla\phi(\xi_s)^* h(y-\xi_s)|^2 dy ds\Big) \le T\|\phi\|_{1,\infty}\|h\|_2^2 < \infty,$$

$$\begin{aligned}&\mathbb{E}\int_0^T \int_0^T \int_{\mathbb{R}^d\times\mathbb{R}^d} |A\phi(\xi_s)||\kappa(z_1,z_2) u_r(z_1) u_r(z_2)| dz_1 dz_2 ds dr \\ \le& \|\phi\|_{2,\infty}\|\kappa\|_\infty \int_0^T \int_{\mathbb{R}^d} \|u_r(x)\|_2^2 dx dr < \infty,\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\Big(\int_0^T \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\nabla\phi(\xi_s)^* h(y-\xi_s)|^2 |\kappa(z_1,z_2) u_r(z_1) u_r(z_2)| dy dz_1 dz_2 ds dr\Big) \\ \le& \|\phi\|_{1,\infty}\|h\|_2\|\kappa\|_\infty \int_0^T \int_{\mathbb{R}^d} \|u_r(x)\|_2^2 dx dr < \infty.\end{aligned}$$

Thus by the stochastic Fubini theorem (see e.g. Lemma 4.1 on page 116 of Ikeda and Watanabe [15]), we have

$$\begin{aligned}&\langle u_t, \phi\rangle - \langle \mu, \phi\rangle - \int_0^t \langle u_s, A\phi\rangle ds - \int_0^t \int_{\mathbb{R}^d} \langle u_s, \nabla\phi^* h(y-\cdot)\rangle W(ds,dy) \\ =& \Big\langle \mu, \mathbb{E}^W_{0,\cdot}\Big(\phi(\xi_t) - \phi(\xi_0) - \int_0^t A\phi(\xi_s) ds - \int_0^t \int_{\mathbb{R}^d} \nabla\phi(\xi_s)^* h(y-\xi_s) W(ds,dy)\Big)\Big\rangle \\ &+ \int_0^t \int_{\mathbb{R}^d} \mathbb{E}^W_{s,z}\Big(\phi(\xi_t) - \int_s^t A\phi(\xi_r) dr - \int_s^t \int_{\mathbb{R}^d} \nabla\phi(\xi_r)^* h(y-\xi_r) W(dr,dy)\Big) Z(ds,dz).\end{aligned} \tag{6.17}$$

The last stochastic integral in (6.17) is well-defined, because the integrand is an $\mathcal{F}_s$-adapted process, where $\mathcal{F}_s$ is defined in (6.2). Notice that by Itô's formula, we have

$$\begin{aligned}\phi(\xi_t^{s,x}) =&\phi(x) + \int_s^t A\phi(\xi_r^{s,x})dr + \int_s^r \nabla\phi(\xi_r^{s,x})^* dB_r \\ &+ \int_s^t \int_{\mathbb{R}^d} \nabla\phi(\xi_r^{s,x})^* h(y-\xi_r^{s,x}) W(dr,dy). \end{aligned} \tag{6.18}$$

Then, (6.16) follows from (6.17) and (6.18). □

7 Proof of Theorem 3.4

In this section, we prove Theorem 3.4 by showing the the Hölder continuity of $u_t(x)$ in spatial and time variables separately:

Proposition 7.1. *Suppose that $h \in H_3^2\left(\mathbb{R}^d\right)$, $\|\kappa\|_\infty < \infty$, and $\mu \in L^1\left(\mathbb{R}^d\right)$ is bounded. Then, for any $0 < s < t \le T$, $x, y \in \mathbb{R}^d$ $\beta \in (0,1)$ and $p > 1$, there exists a constant C depending on T, d, $\|h\|_{3,2}$, $\|\mu\|_\infty$, $\|\kappa\|_\infty$, p, and β, such that the following inequalities are satisfied:*

$$\|u_t(y) - u_t(x)\|_{2p} \le Ct^{-\frac{1}{2}}(y-x)^\beta, \tag{7.1}$$

$$\|u_t(x) - u_s(x)\|_{2p} \le Cs^{-\frac{1}{2}}(t-s)^{\frac{1}{2}\beta}. \tag{7.2}$$

Then, Theorem 3.4 is simply a corollary of Proposition 7.1. In order to prove Proposition 7.1, we need the following Hölder continuity results for the conditional transition density $p^W(r,z;t,x)$:

Lemma 7.2. *Suppose that $h \in H_2^3(\mathbb{R}^d)$, $0 \le r < s < t \le T$, $x, y \in \mathbb{R}^d$, and $\beta \in (0,1)$. Then, there exists $C > 0$, depending on T, d, $\|h\|_{3,2}$, p and β, such that the following inequalities are satisfied:*

$$\int_{\mathbb{R}^d} \left\|p^W(r,z;t,y) - p^W(r,z;t,x)\right\|_{2p} dz \le C(t-r)^{-\frac{1}{2}\beta} |y-x|^\beta, \tag{7.3}$$

$$\int_{\mathbb{R}^d} \left\|p^W(r,z;t,x) - p^W(r,z;s,x)\right\|_{2p} dz \le C(s-r)^{-\frac{1}{2}\beta}(t-s)^{\frac{1}{2}\beta}. \tag{7.4}$$

Before showing the proof, let us firstly derive a variant of the density formula (A.11). It will be used in the proof of (7.4). Choose $\phi \in C_b^2\left(\mathbb{R}^n\right)$, such that $\mathbf{1}_{B(0,1)} \le \phi \le \mathbf{1}_{B(0,4)}$, and its first and second partial derivatives are all bounded by 1. For any $x \in \mathbb{R}^d$ and $\rho > 0$, we set $\phi_\rho^x := \phi(\frac{\cdot - x}{\rho})$. Assume that F satisfies all the properties in Theorem A.3. Let Q_n be the n-dimensional Poisson kernel (see (A.10)). Then, the density of F can be represented as follows:

$$\begin{aligned} p_F(x) =& \sum_{i,j_1,j_2=1}^n \mathbb{E}\left[\partial_{j_1} Q_n(F-x) \left\langle DF^{j_1}, DF^{j_2}\right\rangle_H \sigma^{j_2 i} H_{(i)}(F, \phi_\rho^x(F))\right] \\ =&\mathbb{E}\Big[\Big\langle DQ_n(F-x), \sum_{i,j_2=1}^m H_{(i)}(F,\phi_\rho^x(F))\sigma^{j_2 i} DF^{j_2}\Big\rangle_H\Big] \\ =&\sum_{i=1}^m \mathbb{E}\Big[Q_n(F-x) \sum_{j_2=1}^m \delta\left[H_{(i)}(F,\phi_\rho^x(F))\sigma^{j_2 i} DF^{j_2}\right]\Big] \\ =&-\sum_{i=1}^m \mathbb{E}\big[Q_n(F-x) H_{(i,i)}(F,\phi_\rho^x(F))\big]. \end{aligned} \tag{7.5}$$

Let $\xi_t = \xi_t^{r,z}$ be defined in (5.1).

Proof of (7.3). Choose $p_1 \in (d, 3pd]$, let $p_2 = 2p_1$, and $p_3 = \frac{p_1 p_2}{p_2 - p_1} = p_2$. Then, by (A.13) and Hölder's inequality, for any fixed $z, x, y \in \mathbb{R}^d$ and $\rho > 0$, we can show that

$$\begin{aligned} I(z) :=& \|p^W(r, z; t, x) - p^W(r, z; t, y)\|_{2p} \\ \leq& C|y - x| \left\| \mathbb{P}^W(\xi_t - \tau \leq 4\rho)^{\frac{1}{p_2}} \right\|_{6p} \max_{1 \leq i \leq d} \Big\{ \left\| \left(\|H_{(i)}(\xi_t; 1)\|^W_{p_2} \right)^{d-1} \right\|_{6p} \\ & \times \Big(\frac{1}{\rho^2} + \frac{2}{\rho} \left\| \|H_{(i)(\xi_t;1)}\|^W_{p_2} \right\|_{6p} + \left\| \|H_{(i,j)}(\xi_t; 1)\|^W_{p_2} \right\|_{6p} \Big) \Big\}, \end{aligned}$$

where $\tau = cx + (1 - c)y$, for some $c \in (0, 1)$ that depends on z, x, y.

Let $\rho = \frac{\sqrt{t-r}}{8}$. Similarly as proved in Proposition 5.9, we can show that

$$\begin{aligned} I(z) \leq& C|y - x|(t - r)^{-\frac{d+1}{2}} \exp\Big(- \frac{k|\tau - z|^2}{(6p \vee p_2)(t - r)} \Big) \\ \leq& C|y - x|(t - r)^{-\frac{d+1}{2}} \exp\Big(- \frac{k|\tau - z|^2}{6pd(t - r)} \Big), \end{aligned} \tag{7.6}$$

where k is defined in (5.64) and $C > 0$ depends on T, d, p, and $\|h\|_{3,2}$.

Notice that even if we fix $x, y \in \mathbb{R}^d$, τ is still a function of z that does not have an explicit formulation. Thus it is not easy to calculate the integral of I directly. Without losing generality, assume that $x = 0$, and $y = (y_1, 0, \ldots, 0)$, where $y_1 \geq 0$. Then $\tau = ((1 - c)y_1, 0, \ldots, 0)$, where $c = c(z) \in (0, 1)$. Let $\widehat{k} = \frac{k}{6pd}$. For any $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$, we consider the following cases.

(a) If $z_1 \leq 0$, then

$$\exp\Big(- \frac{k|\tau - z|^2}{6pd(t - r)} \Big) \leq \exp\Big(- \frac{\widehat{k}|z|^2}{t - r} \Big). \tag{7.7}$$

(b) If $z_1 \geq y_1$, then

$$\exp\Big(- \frac{k|\tau - z|^2}{6pd(t - r)} \Big) \leq \exp\Big(- \frac{\widehat{k}|y - z|^2}{t - r} \Big). \tag{7.8}$$

(c) If $0 < z_1 < y_1$, then

$$\exp\Big(- \frac{k|\tau - z|^2}{6pd(t - r)} \Big) \leq \exp\Big(- \frac{\widehat{k}|\tau_0 - z|^2}{t - r} \Big), \tag{7.9}$$

where $\tau_0 = (z_1, 0, \ldots, 0)$.

Therefore, combining (7.6)–(7.9), we have

$$\int_{\mathbb{R}^d} I(z)dz \leq C|y - x|(t - r)^{-\frac{d+1}{2}} (I_1 + I_2 + I_3), \tag{7.10}$$

where

$$\begin{aligned} I_1 &= \int_{-\infty}^0 dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(- \frac{\widehat{k}|z|^2}{t - r} \Big) dz_d \ldots dz_2, \\ I_2 &= \int_{|y|}^\infty dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(- \frac{\widehat{k}|y - z|^2}{t - r} \Big) dz_d \ldots dz_2, \\ I_3 &= \int_0^{|y|} dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(- \frac{\widehat{k}|\tau_0 - z|^2}{t - r} \Big) dz_d \ldots dz_2. \end{aligned}$$

By a changing of variables, it is easy to show that

$$I_1 + I_2 = \int_{\mathbb{R}^d} \exp\Big(-\frac{\widehat{k}|z|^2}{t-r}\Big)dz = \widehat{k}^{-\frac{d}{2}}(t-r)^{\frac{d}{2}}. \tag{7.11}$$

For I_3, we compute the integral as follows:

$$\begin{aligned} I_3 =& \int_0^{|y|} dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(-\frac{\widehat{k}\,(z_2^2+\dots z_d^2)}{t-r}\Big)dz_d\dots dz_2 \\ =& \big(2\pi\widehat{k}^{-1}\big)^{\frac{d-1}{2}}(t-r)^{\frac{d-1}{2}}|y|. \end{aligned} \tag{7.12}$$

Thus combining (7.10)–(7.12), we have

$$\begin{aligned} \int_{\mathbb{R}^d} I(z)dz \le& C\big[(t-r)^{-\frac{1}{2}}|y| + (t-r)^{-1}|y|^2\big] \\ =& C\big[(t-r)^{-\frac{1}{2}}|y-x| + (t-r)^{-1}|y-x|^2\big]. \end{aligned} \tag{7.13}$$

It is easy to see that inequality (7.13) holds for all $x, y \in \mathbb{R}^d$.

On the other hand, by Proposition 5.9, we have

$$\int_{\mathbb{R}^d} I(z)dz \le \int_{\mathbb{R}^d} \|p^W(r,z;t,y)\|_{2p} + \|p^W(r,z;t,x)\|_{2p}dz \le C. \tag{7.14}$$

Therefore by (7.13) and (7.14), for any $\beta_1, \beta_2 \in (0,1)$, we have

$$\int_{\mathbb{R}^d} I(z)dz \le C\big[(t-r)^{-\frac{1}{2}\beta_1}\,|y-x|^{\beta_1} + (t-r)^{-\beta_2}\,|y-x|^{2\beta_2}\,\big].$$

Then, (7.3) follows by choosing $\beta = \beta_1 = 2\beta_2$. □

Proof of (7.4). Let $\rho_1 = \sqrt{t-r}$ and $\rho_2 = \sqrt{s-r}$. By density formula (7.5), we have

$$\begin{aligned} &\big|p^W(r,z;t,x) - p^W(r,z;s,x)\big| \\ \le& \sum_{i=1}^d \big|\mathbb{E}^W\,\{[Q_d(\xi_t - x) - Q_d(\xi_s - x)]\,H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\}\,\big| \\ &+ \sum_{i=1}^d \big|\mathbb{E}^W\,\{Q_d(\xi_t - x)\,\big[H_{(i,i)}(\xi_t, \phi^x_{\rho_1}(\xi_t)) - H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\big]\}\,\big| \\ =& I_1 + I_2. \end{aligned} \tag{7.15}$$

Estimation for I_1: Note that by the local property of δ (see Proposition 1.3.15 of Nulart [24]), $H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))$ vanishes except if $\xi_s \in B(x, 4\rho_2)$. Choose $p_1 \in (d, 2pd]$. Let $p_2 = 3p_1$ and $p_3 = \frac{3p_1}{3p_1-2}$. Then, $\frac{2}{p_2} + \frac{1}{p_3} = 1$. Thus, by Hölder's inequality, we have

$$\begin{aligned} \|I_1\|_{2p} \le& d\big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|^W_{p_2}\big\|_{6p}\big\|\|Q_d(\xi_t - x) - Q_d(\xi_s - x)\|^W_{p_3}\big\|_{6p} \\ &\times \max_{1\le i\le d}\big\|\|H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\|^W_{p_2}\big\|_{6p}. \end{aligned} \tag{7.16}$$

By Proposition 5.8, and the fact that $p_2 = 3p_1 \le 6pd$, the first factor satisfies the following inequality

$$\big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|^W_{p_2}\big\|_{6p} = \big\|\mathbb{P}^W(|\xi_s - x| < 4\rho_2)^{\frac{1}{p_2}}\big\|_{6p} \le C\exp\Big(-\frac{k|z-x|}{6pd(s-r)}\Big). \tag{7.17}$$

By Lemmas 5.5 and A.2, for all $1 \le i \le d$, the last factor can be estimated as follows:

$$\begin{aligned} \big\| \|H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\|^W_{p_2} \big\|_{6p} \le& \frac{1}{\rho_2^2} + \frac{2}{\rho_2} \big\| \|H_{(i)}(\xi_s,1)\|^W_{p_2}\big\|_{6p} + \big\|\|H_{(i,i)}(\xi_s,1)\|^W_{p_2}\big\|_{6p} \\ \le& C(s-r)^{-1}. \end{aligned} \tag{7.18}$$

We estimate the second factor by the mean value theorem. Let $\eta_1 = |\xi_t - x|$ and $\eta_2 = |\xi_s - x|$. Then, we can write

$$Q_d(\xi_t - x) - Q_d(\xi_s - x) = \begin{cases} A_2^{-1}\left(\log \eta_1 - \log \eta_2\right), & \text{if } d = 2, \\ -A_d^{-1}\left[\eta_1^{-(d-2)} - \eta_2^{-(d-2)}\right], & \text{if } d \ge 3. \end{cases}$$

Thus, by the mean value theorem, it follows that

$$|Q_d(\xi_t - x) - Q_d(\xi_s - x)| = \frac{c_d|\eta_1 - \eta_2|}{|\zeta\eta_1 + (1-\zeta)\eta_2|^{d-1}},$$

where c_d is a constant coming from the Poisson kernel, and $\zeta \in (0,1)$ is a random number that depends on η_1 and η_2. Notice that $f(x) = x^{-(d-1)}$ is a convex function on $(0,\infty)$, and $\mathbb{P}(\eta_1 > 0) = \mathbb{P}(\eta_2 > 0) = 1$, then we have

$$|\zeta\eta_1 + (1-\zeta)\eta_2|^{-(d-1)} \le |\zeta\eta_1|^{-(d-1)} + |(1-\zeta)\eta_2|^{-(d-1)}, \ a.s.$$

Let $q = \frac{p_1}{p_1 - 1}$, then $\frac{1}{q} + \frac{1}{p_2} = \frac{1}{p_3}$. As a consequence of Hölder's inequality, we have

$$\begin{aligned} \big\|\|Q_d(\xi_t - x) - Q_d(\xi_s - x)\|^W_{p_3}\big\|_{6p} \le& c_d \Big\| \Big\| \frac{|\eta_1 - \eta_2|}{|\zeta\eta_1 + (1-\zeta)\eta_2|^{d-1}} \Big\|^W_{p_3} \Big\|_{6p} \\ \le& C \big\|\|\eta_1 - \eta_2\|^W_{p_2}\big\|_{12p} \big\|\||\zeta\eta_1 + (1-\zeta)\eta_2|^{-(d-1)}\|^W_q\big\|_{12p} \\ \le& C\|\eta_1 - \eta_2\|_{12pd}\left[\big\|\||\zeta\eta_1^{-(d-1)}\|^W_q\big\|_{12p} + \big\|\|(1-\zeta)\eta_2^{-(d-1)}\|^W_q\big\|_{12p}\right] \\ \le& C\|\xi_t - \xi_s\|_{12pd}\left[\big\|\||\xi_t - y|^{-(d-1)}\|^W_q\big\|_{12p} + \big\|\||\xi_s - y|^{-(d-1)}\|^W_q\big\|_{12p}\right]. \end{aligned} \tag{7.19}$$

The negative moments of $\xi_t - y$ can be estimated by (5.57), Jensen's inequality, and Lemma A.6:

$$\begin{aligned} \big\|\||\xi_t - x|^{-(d-1)}\|^W_q\big\|_{12p} \le& C \max_{1\le i\le d} \big\|\left(\|H_i(\xi_t,1)\|^W_{p_1}\right)^{d-1}\big\|_{12p} \\ \le& C \max_{1\le i\le d} \big\|H_{(i)}(\xi_t,1)\big\|^{d-1}_{12pd} \le C(t-r)^{-\frac{d-1}{2}}. \end{aligned} \tag{7.20}$$

Then, by (7.19)–(7.20), we have

$$\big\|\|Q_d(\xi_t - x) - Q_d(\xi_s - x)\|^W_{p_3}\big\|_{6p} \le C(t-s)^{\frac{1}{2}}(s-r)^{-\frac{d-1}{2}}. \tag{7.21}$$

Thus combining (7.16), (7.17), (7.18) and (7.21), we have

$$\|I_1\|_{2p} \le C \exp\Big(-\frac{k|z-x|}{6pd(s-r)}\Big)(s-r)^{-\frac{d+1}{2}}(t-s)^{\frac{1}{2}}.$$

This implies

$$\int_{\mathbb{R}^d} \|I_1\|_{2p} dz \le C(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{7.22}$$

Estimates for I_2: Recall that $\gamma_t = (\langle D\xi^i, D\xi^j\rangle_H)_{i,j=1}^d = \sigma_t^{-1}$. By computation analogue to (7.5) going backward, we can show that

$$\begin{aligned}
&\mathbb{E}^W\big[Q_d(\xi_t - x)\big(H_{(i,i)}(\xi_t, \phi^x_{\rho_1}(\xi_t)) - H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\big)\big] \\
=&-\sum_{j_1,j_2=1}^d \mathbb{E}^W\big[\partial_{j_2} Q_d(\xi_t - x)\langle D\xi_t^{j_2}, D\xi_t^{j_1}\rangle_H H_{(i)}\big(\xi_t, \phi^x_{\rho_1}(\xi_t)\big)\sigma_t^{j_1 i}\big] \\
&+\sum_{j_1,j_2=1}^d \mathbb{E}^W\big[\partial_{j_2} Q_d(\xi_t - x)\langle D\xi_t^{j_2}, D\xi_s^{j_1}\rangle_H H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s^{r,z})\right)\sigma_s^{j_1 i}\big] \\
=&-\mathbb{E}^W\big[\partial_i Q_d(\xi_t - x)\left(H_{(i)}\left(\xi_t, \phi^x_{\rho_1}(\xi_t)\right) - H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s)\right)\right)\big] \\
&+\sum_{j_1,j_2=1}^d \mathbb{E}^W\big[\partial_{j_2} Q_d(\xi_t - x)\langle D\xi_t^{j_2} - D\xi_s^{j_2}, D\xi_s^{j_1}\rangle_H H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s)\right)\sigma_s^{j_1 i}\big] \\
:=&J_1 + J_2.
\end{aligned} \tag{7.23}$$

By Lemma A.2, we have

$$\begin{aligned}
\left|H_{(i)}\left(\xi_t, \phi^x_{\rho_1}(\xi_t)\right) - H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s)\right)\right| &\le \left|\partial_i\phi^x_{\rho_1}(\xi_t) - \partial_i\phi^x_{\rho_2}(\xi_s)\right| \\
&+ |\phi^x_{\rho_2}(\xi_s)|\left|H_{(i)}(\xi_t,1) - H_{(i)}(\xi_s,1)\right| + \left|H_{(i)}(\xi_t,1)\right|\left|\phi^x_{\rho_1}(\xi_t) - \phi^x_{\rho_2}(\xi_s)\right|.
\end{aligned} \tag{7.24}$$

By the mean value theorem, for some random numbers $c_1, c_2 \in (0,1)$, we have

$$\begin{aligned}
\left|\phi^x_{\rho_1}(\xi_t) - \phi^x_{\rho_2}(\xi_s)\right| =&\left|\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right|\left|\phi\Big(\frac{\xi_t - x}{\rho_1}\Big) - \phi\Big(\frac{\xi_s - x}{\rho_2}\Big)\right| \\
=&\left|\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right| \\
&\times\left|\nabla\phi\Big(c_1\frac{\xi_t - x}{\rho_1} + (1-c_1)\frac{\xi_s - x}{\rho_2}\Big)^* \cdot \Big(\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\Big)\right| \\
\le&\left|\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right|\left|\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\right|,
\end{aligned} \tag{7.25}$$

and

$$\begin{aligned}
\left|\partial_i\phi^x_{\rho_1}(\xi_t) - \partial_i\phi^x_{\rho_2}(\xi_s)\right| =&\left|\rho_1^{-1}\partial_i\phi\Big(\frac{\xi_t - x}{\rho_1}\Big) - \rho_2^{-1}\partial_i\phi\Big(\frac{\xi_s - x}{\rho_2}\Big)\right| \\
\le&\frac{1}{\rho_1}\left|\nabla\partial_i\phi\Big(c_2\frac{\xi_t - x}{\rho_1} + (1-c_2)\frac{\xi_s - x}{\rho_2}\Big)^* \cdot \Big(\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\Big)\right| \\
&+\left|\partial_i\phi^x_{\rho_2}(\xi_s)\right|\left|\frac{1}{\rho_1} - \frac{1}{\rho_2}\right| \\
\le&\frac{1}{\rho_1}\left(\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right)\left|\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\right| \\
&+\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\left|\frac{1}{\rho_1} - \frac{1}{\rho_2}\right|.
\end{aligned} \tag{7.26}$$

Choose $q \in (d, 3pd]$, let $p_1 = \frac{q}{q-1}$, $p_2 = 2q$, $p_3 = 4q$. Then,

$$\frac{1}{p_1} + \frac{2}{p_2} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{2}{p_3} = 1.$$

Then, by (7.24)–(7.26), and Hölder's inequality, we have

$$\begin{aligned}\|J_1\|_{2p} \leq & \rho_1^{-1}\Big\|\|\partial_i Q_d(\xi_t - x)\|_{p_1}^W\Big\|_{6p}\Big\|\|\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W\Big\|_{6p} \\ & \times \Big\|\Big\|\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\Big\|_{p_2}^W\Big\|_{6p} \\ & + \Big\|\|\partial_i Q_d(\xi_t - x)\|_{p_1}^W\Big\|_{6p}\Big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W\Big\|_{6p}\Big\|\|\rho_1^{-1} - \rho_2^{-1}\|_{p_2}\Big\|_{6p} \\ & + \Big\|\|\partial_i Q_d(\xi_t - x)\|_{p_1}^W\Big\|_{6p}\Big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W\Big\|_{6p}\Big\|\|H_{(i)}(\xi_t,1) - H_{(i)}(\xi_s,1)\|_{p_2}^W\Big\|_{6p} \\ & + \Big\|\|\partial_i Q_d(\xi_t - x)\|_{p_1}^W\Big\|_{6p}\Big\|\|\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W\Big\|_{6p} \\ & \times \Big\|\Big\|\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\Big\|_{p_3}^W\Big\|_{12p} \quad \Big\|\|H_{(i)}(\xi_t,1)\|_{p_3}^W\Big\|_{12p} \\ := & L_1 + L_2 + L_3 + L_4. \end{aligned} \tag{7.27}$$

In order to estimate the moments of $\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}$, we rewrite this random vector in the following way:

$$\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2} = \frac{\xi_t - \xi_s}{\rho_1} + (\xi_s - z)\left(\frac{1}{\rho_1} - \frac{1}{\rho_2}\right) + (z - x)\left(\frac{1}{\rho_1} - \frac{1}{\rho_2}\right).$$

It follows that

$$\begin{aligned}\Big\|\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\Big\|_{12p \vee p_3} \leq & (t-r)^{-\frac{1}{2}}\big\|\xi_t - \xi_s\big\|_{12pd} \\ & + \frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}}\big\|\xi_s - z\big\|_{12pd} + |z - x|\frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}}.\end{aligned}$$

According to Lemma 5.7, $\xi_t - \xi_s$ and $\xi_s - z$ are Gaussian random vectors with mean 0, and covariance matrix $(t-s)(I + \rho(0))$ and $(s-r)(I + \rho(0))$ respectively. Therefore, we have

$$\begin{aligned}\Big\|\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}\Big\|_{12pd} \leq & c_{p,d}(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}} + c_{p,d}\frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}}(s-r)^{\frac{1}{2}} \\ & + |z - x|\frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}} \\ \leq & C\big(|z-x|(s-r)^{-\frac{1}{2}} + 1\big)(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}.\end{aligned} \tag{7.28}$$

Therefore, by (7.28), Proposition 5.8 and Lemma A.6, we have

$$\begin{aligned}L_1 + L_4 \leq & C(t-r)^{-\frac{d}{2}}\Big[\exp\Big(-\frac{k|z-x|^2}{6pd(t-r)}\Big) + \exp\Big(-\frac{k|z-x|^2}{6pd(s-r)}\Big)\Big] \\ & \times \big(1 + |z-x|(s-r)^{-\frac{1}{2}}\big)(t-s)^{\frac{1}{2}},\end{aligned} \tag{7.29}$$

and

$$L_2 + L_3 \leq C(t-r)^{-\frac{d}{2}}\exp\Big(-\frac{k|z-x|^2}{6pd(s-r)}\Big)(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{7.30}$$

Plugging (7.29) and (7.30) into (7.27), we have

$$\int_{\mathbb{R}^d}\|J_1\|_{2p}\,dz \leq C(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{7.31}$$

For J_2, notice that, by definition,

$$\langle D\xi_t^{j_2} - D\xi_s^{j_2}, D\xi_s^{j_1}\rangle_H = \sum_{k=1}^{d}\int_r^s \big(D_\theta^{(k)}\xi_t^{j_2} - D_\theta^{(k)}\xi_s^{j_2}\big)D_\theta^{(k)}\xi_s^{j_1}\,d\theta.$$

By (5.2), we have

$$D_\theta^{(k)}\xi_t^{j_2} - D_\theta^{(k)}\xi_s^{j_2} = \mathbf{1}_{[s,t]}(\theta)\delta_{j_2k} - \sum_{i=1}^{d}\mathbf{1}_{[r,t]}(\theta)\int_s^t D_\theta^{(k)}\xi_r^i dM_r^{ij_2}.$$

By an argument similar to the one used in the proof of Lemma 5.3, we can show that

$$\big\|\mathbf{1}_{[r,s]}(\theta)\big(D_\theta^{(k)}\xi_t^{j_2} - D_\theta^{(k)}\xi_s^{j_2}\big)\big\|_{2p}^2 \le C\mathbf{1}_{[r,s]}(\theta)(t-s).$$

Therefore, by Hölder's and Minkowski's inequalities, we have

$$\begin{aligned}\big\|\langle D\xi_t^{j_2} - D\xi_s^{j_2}, D\xi_s^{j_1}\rangle_H\big\|_{2p} \le& \sum_{k=1}^{d}\int_r^s \big\|\mathbf{1}_{[r,s]}(\theta)\big(D_\theta^{(k)}\xi_t^{j_2} - D_\theta^{(k)}\xi_s^{j_2}\big)\big\|_{4p}\big\|D_\theta^{(k)}\xi_s^{j_1}\big\|_{4p}d\theta\\ \le& C(s-r)(t-s)^{\frac12}.\end{aligned}\tag{7.32}$$

Choose $q\in(d,3pd]$. Let $p_1=\frac{q}{q-1}$, $p_2=2q$ and $p_3=6q$. Then $\frac{1}{p_1}+\frac{1}{p_2}+\frac{3}{p_3}=1$. Thus, by (7.32), Hölder's inequality, Lemmas 5.3, 5.5, A.6, and Proposition 5.8, we have

$$\begin{aligned}\|J_2\|_{2p} \le& \sum_{j_1,j_2=1}^{d}\big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W\big\|_{6p}\big\|\|\partial_{j_2}Q_d(\xi_t-x)\|_{p_1}^W\big\|_{6p}\\ &\times\big\|\|\langle D\xi_t^{j_2} - D\xi_s^{j_2}, D\xi_s^{j_1}\rangle_H\|_{p_3}^W\big\|_{18p}\big\|\|H_{(i)}(\xi_s,\phi_{\rho_2}^y(\xi_s))\|_{p_3}^W\big\|_{18p}\big\|\|\sigma_s^{j_1i}\|_{p_3}^W\big\|_{18p}\\ \le& C\exp\Big(-\frac{k|z-x|^2}{6pd(s-r)}\Big)(t-r)^{-\frac{d-1}{2}}(t-s)^{\frac12}(s-r)^{-\frac12}.\end{aligned}$$

As a consequence, we have

$$\int_{\mathbb{R}^d}\|J_2\|_{2p}\,dz \le C(t-s)^{\frac12}.\tag{7.33}$$

Finally, combining (7.22), (7.31) and (7.33), we have

$$\int_{\mathbb{R}^d}\big\|p^W(r,z;t,x) - p^W(r,z;s,x)\big\|_{2p}\,dz \le C(s-r)^{-\frac12}(t-s)^{\frac12}.\tag{7.34}$$

On the other hand, by (5.72), we have

$$\int_{\mathbb{R}^d}\|I_2\|_{2p}dz \le \int_{\mathbb{R}^d}\|p^W(r,z;t,y)\|_{2p} + \|p^W(r,z;s,y)\|_{2p} \le C.\tag{7.35}$$

Thus (7.4) follows from (7.34) and (7.35). □

Proof of Proposition 7.1. By the convolution representation (6.1), Burkholder-Davis-Gundy's, and Minkowski's inequalities, we have

$$\begin{aligned}\|u_t(y)-u_t(x)\|_{2p} \le& \Big\|\int_{\mathbb{R}^d}\mu(z)\left(p^W(0,z;t,y) - p^W(0,z;t,x)\right)dz\Big\|_{2p}\\ &+\Big\|\int_0^t\int_{\mathbb{R}^d}u_r(z)\left(p^W(r,z;t,y) - p^W(r,z;t,x)\right)V(dz,dr)\Big\|_{2p}\\ \le& \|\mu\|_\infty\int_{\mathbb{R}^d}\big\|p^W(0,z;t,y) - p^W(0,z;t,x)\big\|_{2p}\,dz\\ &+\|\kappa\|_\infty^{\frac12}\Big(\int_0^t\Big(\int_{\mathbb{R}^d}\big\|u_r(z)\left(p^W(r,z;t,y) - p^W(r,z;t,x)\right)\big\|_{2p}\,dz\Big)^2dr\Big)^{\frac12}\\ :=& I_1 + \|\kappa\|_\infty^{\frac12}I_2.\end{aligned}\tag{7.36}$$

Note that I_1 can be estimated by Lemma 7.2. For I_2, recall that $u(r,z)$ is independent of $p^W(r,z;t,y)^2$. Then, by Lemma 6.2 and 7.2, we have

$$\begin{aligned}
I_2 \leq & \Big(\int_0^t \sup_{z\in\mathbb{R}^d} \|u_r(z)\|_{2p}^2 \Big(\int_{\mathbb{R}^d} \big\|p^W(r,z;t,y) - p^W(r,z;t,x)\big\|_{2p} dz \Big)^2 dr \Big)^{\frac12} \\
\leq & C|y-x|^\beta \Big(\int_0^t (t-r)^{-\beta} dr \Big)^{\frac12} \leq \frac{Ct^{\frac{1-\beta}{2}}}{\sqrt{1-\beta}} |y-x|^\beta. \qquad (7.37)
\end{aligned}$$

Therefore (7.1) follows from (7.3), (7.36) and (7.37).

The proof of (7.2) is quite similar. As in (7.36), we can show that

$$\begin{aligned}
&\|u_t(x) - u_s(x)\|_{2p} \leq \|\mu\|_\infty \int_{\mathbb{R}^d} \big\|p^W(0,z;t,x) - p^W(0,z;s,x)\big\|_{2p} dz \\
&\quad + C\|\kappa\|_\infty^{\frac12} \Big[\int_s^t \sup_{z\in\mathbb{R}^d} \|u_r(z)\|_{2p}^2 \Big(\int_{\mathbb{R}^d} \big\|p^W(r,z;t,x)\big\|_{2p} dz \Big)^2 dr \Big]^{\frac12} \\
&\quad + C\|\kappa\|_\infty^{\frac12} \Big[\int_0^s \sup_{z\in\mathbb{R}^d} \|u_r(z)\|_{2p}^2 \Big(\int_{\mathbb{R}^d} \big\|\big(p^W(r,z;t,x) - p^W(r,z;s,x)\big)\big\|_{2p} dz \Big)^2 dr \Big]^{\frac12}.
\end{aligned}$$

Then, the estimate (7.2) follows from (7.4), Proposition 5.9 and Lemma 6.2. □

A Basic introduction on Malliavin calculus

In this section, we present some preliminaries on the Malliavin calculus. We refer the readers to book of Nualart [24] for a detailed account on this topic.

Fix a time interval $[0,T]$. Let $B = \{B_t^1, \dots, B_t^d, 0 \leq t \leq T\}$ be a standard d-dimensional Brownian motion on $[0,T]$. Denote by $\mathcal{S}$ the class of smooth random variables of the form

$$G = g(B_{t_1}, \dots, B_{t_m}) = g\left(B_{t_1}^1, \dots, B_{t_1}^d, \dots, B_{t_m}^1, \dots, B_{t_m}^d\right), \qquad (A.1)$$

where m is any positive integer, $0 \leq t_1 < \cdots < t_m \leq T$, and $g: \mathbb{R}^{md} \to \mathbb{R}$ is a smooth function that has all partial derivatives with at most polynomial growth. We make use of the notation $x = \left(x_i^k\right)_{1\leq i\leq m, 1\leq k\leq d}$ for any element $x \in \mathbb{R}^{md}$. The basic Hilbert space associated with B is $H = L^2\left([0,T];\mathbb{R}^d\right)$.

Definition A.1. *For any $G \in \mathcal{S}$ given by (A.1), the Malliavin derivative, is the H-valued random variable DG given by*

$$D_\theta^{(k)} G = \sum_{i=1}^m \frac{\partial g}{\partial x_i^k}(B_{t_1}, \dots, B_{t_m}) \mathbf{1}_{[0,t_i]}(\theta), \quad 1 \leq k \leq d, \ \theta \in [0,T].$$

In the same way, for any $n \geq 1$, the iterated derivative $D^n G$ of a random variable of the form (A.1) is a random variable with values in $H^{\otimes n} = L^2\left([0,T]^n; \mathbb{R}^{d^n}\right)$. For each $p \geq 1$, the iterated derivative D^n is a closable and unbounded operator on $L^p(\Omega)$ taking values in $L^p(\Omega; H^{\otimes n})$. For any $n \geq 1$, $p \geq 1$ and any Hilbert space V, we can introduce the Sobolev space $\mathbb{D}^{n,p}(V)$ of V-valued random variables as the closure of $\mathcal{S}$ with respect to the norm

$$\begin{aligned}
\|G\|_{n,p,V}^2 =& \|G\|_{L^p(\Omega;V)}^2 + \sum_{k=1}^n \|D^k G\|_{L^p(\Omega; H^{\otimes k}\otimes V)}^2 \\
=& \left[\mathbb{E}\left(\|G\|_V^p\right)\right]^{\frac2p} + \sum_{k=1}^n \left[\mathbb{E}\left(\|D^k G\|_{H^{\otimes k}\otimes V}^p\right)\right]^{\frac2p}.
\end{aligned}$$

[2] The same idea has been used in the proof of Lemma 6.2.

By definition, the divergence operator δ is the adjoint operator of D in $L^2(\Omega)$. More precisely, δ is an unbounded operator on $L^2(\Omega; H)$, taking values in $L^2(\Omega)$. We denote by $\mathrm{Dom}(\delta)$ the domain of δ. Then, for any $u = (u^1, \dots, u^d) \in \mathrm{Dom}(\delta)$, $\delta(u)$ is characterized by the duality relationship: for all for all $G \in \mathbb{D}^{1,2} = \mathbb{D}^{1,2}(\mathbb{R})$.

$$\mathbb{E}\left(\delta(u)G\right) = \mathbb{E}\left(\langle DG, u\rangle_H\right). \tag{A.2}$$

Let F be an n-dimensional random vector, with components $F^i \in \mathbb{D}^{1,1}, 1 \le i \le n$. We associate to F an $n \times n$ random symmetric nonnegative definite matrix, called the Malliavin matrix of F, denoted by γ_F. The entries of γ_F are defined by

$$\gamma_F^{ij} = \left\langle DF^i, DF^j \right\rangle_H = \sum_{k=1}^d \int_0^T D_\theta^{(k)} F^i D_\theta^{(k)} F^j d\theta. \tag{A.3}$$

Suppose that $F \in \cap_{p\ge1}\mathbb{D}^{2,p}(\mathbb{R}^n)$, and its Malliavin matrix γ_F is invertible. Denote by σ_F the inverse of γ_F. Assume that $\sigma_F^{ij} \in \cap_{p\ge1}\mathbb{D}^{1,p}$ for all $1 \le i,j \le n$. Let $G \in \cap_{p\ge1}\mathbb{D}^{1,2}$. Then $G\sigma_F^{ij}DF^k \in \mathrm{Dom}(\delta)$ for all $1 \le i,j,k \le n$. For such F and G, we define

$$H_{(i)}(F,G) = -\sum_{j=1}^n \delta\big(G\sigma_F^{ji}DF^j\big), \quad 1 \le i \le n. \tag{A.4}$$

If furthermore $H_{(i)}(F,G) \in \cap_{p\ge1}\mathbb{D}^{1,p}$ for all $1 \le i \le n$, then we define

$$H_{(i,j)}(F,G) = H_{(j)}\left(F, H_{(i)}(F,G)\right), \quad 1 \le i,j \le n. \tag{A.5}$$

The following lemma is a Wiener functional version of Lemma 9 of Bally and Caramellino [1].

Lemma A.2. *Suppose that $F \in \cap_{p\ge1}\mathbb{D}^{2,p}(\mathbb{R}^n)$, $(\gamma_F^{-1})^{ij} = \sigma_F^{ij} \in \cap_{p\ge1}\mathbb{D}^{2,p}$ for all $1 \le i,j \le n$, and $\phi \in C_b^1(\mathbb{R}^n)$. Then, for any $1 \le i \le n$, we have*

$$H_{(i)}\left(F, \phi(F)\right) = \partial_i\phi(F) + \phi(F)H_{(i)}(F,1). \tag{A.6}$$

Suppose that $F \in \cap_{p\ge1}\mathbb{D}^{3,p}(\mathbb{R}^n)$ and $\phi \in C_b^2(\mathbb{R}^n)$. Then, for any $1 \le i,j \le n$, we have

$$\begin{aligned} H_{(i,j)}\left(F, \phi(F)\right) =& \partial_{ij}\phi(F) + \partial_i\phi(F)H_{(j)}(F,1) \\ &+ \partial_j\phi(F)H_{(i)}(F,1) + \phi(F)H_{(i,j)}(F,1). \end{aligned} \tag{A.7}$$

Proof. For any $F \in \cap_{p\ge1}\mathbb{D}^{2,p}(\mathbb{R}^n)$ and $\phi \in C_b^1(\mathbb{R}^n)$, it is easy to check that $\phi(F) \in \cap_{p\ge1}\mathbb{D}^{1,p}$. Then, $H_{(i)}(F,\phi(F))$ is well defined. For any $G \in \mathbb{D}^{1,2}$, by the duality of D and δ, we have

$$\begin{aligned} \mathbb{E}\left(H_{(i)}\left(F,\phi(F)\right)G\right) &= -\sum_{j=1}^n \mathbb{E}\big(\delta\big(\phi(F)\sigma_F^{ji}DF^j\big)G\big) \\ &= -\sum_{j=1}^n \mathbb{E}\big(\phi(F)\sigma_F^{ji}\langle DF^j, DG\rangle_H\big). \end{aligned} \tag{A.8}$$

On the other hand, by the product rule for the operator D, we have

$$\mathbb{E}\left(\phi(F)H_{(i)}(F,1)G\right) = -\sum_{j=1}^m \mathbb{E}\big(\langle\sigma_F^{ji}DF^j, D\big(\phi(F)G\big)\rangle_H\big)$$

$$= -\sum_{j=1}^m \mathbb{E}\big(\phi(F)\sigma_F^{ji}\langle DF^j, DG\rangle_H\big) - \sum_{j_1,j_2=1}^m \mathbb{E}\big(G\partial_{j_2}\phi(F)\sigma_F^{j_1 i}\langle DF^{j_1}, DF^{j_2}\rangle_H\big).$$

Note that σ_F is the inverse of $\gamma_F = \big(\langle DF^i, DF^j\rangle_H\big)_{i,j=1}^n$, then

$$\sum_{j_1,j_2=1}^{m} \mathbb{E}\big(G\partial_{j_2}\phi(F)\sigma_F^{j_1 i}\big\langle DF^{j_1}, DF^{j_2}\big\rangle_H\big) = \mathbb{E}\big(G\partial_i\phi(F)\big). \tag{A.9}$$

Then, (A.6) follows from (A.8)–(A.9). Equality (A.7) can be proved similarly. □

The next theorem is a density formula using the Riesz transformation. The formula was first introduced by Malliavin and Thalmaier (see Theorem Section 4.23 of [21]), then further studied by Bally and Caramenillo [1].

For any integer $n \geq 2$, let Q_n be the n-dimensional Poisson kernel. That is,

$$Q_n(x) = \begin{cases} A_2^{-1}\log|x|, & n = 1, \\ -A_n^{-1}|x|^{2-n}, & n > 2, \end{cases} \tag{A.10}$$

where A_n is the area of the unit sphere in $\mathbb{R}^n$. Then, $\partial_i Q_n(x) = c_n x_i\,|x|^{-n}$, where $c_2 = A_2^{-1}$ and $c_n = (\frac{n}{2}-1)A_n^{-1}$ for $n > 2$.

The theorem below is the density formula for a class of differentiable random variables.

Theorem A.3. *(Proposition 10 of Bally and Caramenillo [1]) Let $F \in \cap_{p\geq 1}\mathbb{D}^{2,p}(\mathbb{R}^n)$. Assume that $(\gamma_F^{-1})^{ij} = \sigma_F^{ij} \in \cap_{p\geq 1}\mathbb{D}^{1,p}$ for all $1 \leq i,j \leq n$. Then, the law of F has a density p_F.*

More precisely, for any $x \in \mathbb{R}^n$ and $r > 0$, let $B(x,r)$ be the sphere on $\mathbb{R}^n$ centered at x with radius r. Suppose that $\phi \in C_b^1(\mathbb{R}^d)$, such that $\mathbf{1}_{B(0,1)} \leq \phi \leq \mathbf{1}_{B(0,2)}$, and $|\nabla\phi| \leq 1$. Define $\phi_\rho^x := \phi(\frac{\cdot - x}{\rho})$ for any $\rho > 0$ and $x \in \mathbb{R}^n$. Then,

$$\begin{aligned} p_F(x) &= \sum_{i=1}^n \mathbb{E}\big(\partial_i Q_n(F-x)H_{(i)}(F,1)\big) \\ &= \sum_{i=1}^n \mathbb{E}\big(\partial_i Q_n(F-x)H_{(i)}(F,\phi_\rho^x(F))\big) \\ &= \sum_{i=1}^n \mathbb{E}\big(\mathbf{1}_{B(x,2\rho)}(F)\partial_i Q_n(F-x)H_{(i)}(F,\phi_\rho^x(F))\big). \end{aligned} \tag{A.11}$$

The next theorem provides the estimates for the density and its increment.

Theorem A.4. *Suppose that F satisfies the conditions in Theorem A.3. Then, for any $p_2 > p_1 > n$, let $p_3 = \frac{p_1 p_2}{p_2 - p_1}$, there exists a constant C that depends on p_1, p_2 and n, such that*

$$p_F(x) \leq C\mathbb{P}(|F-x| < 2\rho)^{\frac{1}{p_3}} \max_{1\leq i\leq n}\left[\big\|H_{(i)}(F,1)\big\|_{p_1}^{n-1}\left(\frac{1}{\rho} + \big\|H_{(i)}(F,1)\big\|_{p_2}\right)\right]. \tag{A.12}$$

If furthermore, $F \in \cap_{p\geq 1}\mathbb{D}^{3,p}(\mathbb{R}^n)$, there exists a constant C that depends on p_1, p_2, and m, such that for all $x_1, x_2 \in \mathbb{R}^n$,

$$\begin{aligned} |p_F(x_1) - p_F(x_2)| \leq{}& C|x_1 - x_2|\mathbb{P}(|F-y| < 4\rho)^{\frac{1}{p_3}} \\ &\times \max_{1\leq i,j\leq n}\left[\big\|H_{(i)}(F,1)\big\|_{p_1}^{n-1}\left(\frac{1}{\rho^2} + \frac{2}{\rho}\big\|H_{(i)}(F,1)\big\|_{p_2} + \big\|H_{(i,j)}(F,1)\big\|_{q_2}\right)\right], \end{aligned} \tag{A.13}$$

where $y = cx_1 + (1-c)x_2$ for some $c \in (0,1)$ which may depend on x_1 and x_2.

Remark A.5. Inequalities stated in Theorem A.4 are an improved version of those estimates by Bally and Caramillino (see Theorem 8 of [1]). We refer to Nualart and Nualart (see Lemma 7.3.2 of [25]) for a related statement. For the sake of completeness, we present below a proof of Theorem A.4. The proof follows the same idea as in Theorem 8 of [1]. The only difference occurs when choosing the radius of the ball in the estimate for the Poisson kernel. If we optimize the radius, then the exponent of $\|H_{(i)}(F,1)\|_p$ is $n-1$, instead of $\frac{q_1(n-1)}{q_1-n} > n-1$ in [1].

In order to prove Theorem A.4, we first give the estimate for the Poisson kernel:

Lemma A.6. *Suppose that F satisfy the conditions in Theorem A.3. For any $p>n$, let $q=\frac{p}{p-1}$. Then, there exists a constant $C>0$ depends on m and p, such that*

$$\sup_{x\in\mathbb{R}^n}\|\partial_i Q_n(F-x)\|_q \le \sup_{x\in\mathbb{R}^n}\left\||F-x|^{-(n-1)}\right\|_q \le C\max_{1\le i\le n}\left\|H_{(i)}(F,1)\right\|_p^{n-1}. \tag{A.14}$$

Proof. Assume that

$$\|p_F\|_\infty := \sup_{x\in\mathbb{R}^d} p_F(x) < \infty.$$

Denote by $M=\sup_{1\le i\le n}\|H_{(i)}(F,1)\|_p$. Then by Hölder's inequality, for all $x\in\mathbb{R}^d$, we have

$$\begin{aligned}
p_F(x) =& \sum_{i=1}^n \mathbb{E}\big(\partial_i Q_n(F-x)H_{(i)}(F,1)\big) \le \sum_{i=1}^m \|\partial_i Q_n(F-x)\|_q\|H_{(i)}(F,1)\|_p \\
\le& n\sup_{x\in\mathbb{R}^n}\left\||F-x|^{-(n-1)}\right\|_q M,
\end{aligned}$$

which implies

$$\|p_F\|_\infty \le n\sup_{x\in\mathbb{R}^n}\left\||F-x|^{-(n-1)}\right\|_q M. \tag{A.15}$$

In order to estimate $\||F-x|^{-(n-1)}\|_q$, choose any $\rho>0$. Then for all $x\in\mathbb{R}^n$,

$$\begin{aligned}
\mathbb{E}(|F-x|^{-(n-1)q}) =& \int_{\mathbb{R}^d}|y-x|^{-(n-1)q}p_F(y)dy \\
=& \int_{|y-x|\le\rho}|y-x|^{-(n-1)q}p_F(y)dy + \int_{|y-x|>\rho}|y-x|^{-(n-1)q}p_F(y)dy \\
\le& \|p_F\|_\infty\int_0^\rho r^{-(n-1)q}r^{n-1}dr + \rho^{-(n-1)q} \\
=& k_{n,q}\|p_F\|_\infty\rho^{1-(n-1)(q-1)} + \rho^{-(n-1)q},
\end{aligned} \tag{A.16}$$

where $k_{n,q}=[1-(n-1)(q-1)]^{-1}$. The last equality is due to the fact that $1-(n-1)(q-1)>0$.

Combining (A.15) and (A.16), we have

$$\|p_F\|_\infty \le \left[nk_{n,q}^{\frac{1}{q}}\|p_F\|_\infty^{\frac{1}{q}}\rho^{\frac{1-(n-1)(q-1)}{q}} + \rho^{-(n-1)}\right]M. \tag{A.17}$$

By optimizing the right-hand side of (A.17), we choose

$$\rho=\rho^* := \left[\frac{(n-1)q}{n}\right]^{\frac{q}{n}}\|p_F\|_\infty^{-\frac{1}{n}}.$$

Plugging ρ^* into (A.17), we obtain

$$\|p_F\|_\infty \le \left(nk_{n,q}^{\frac{1}{q}}\left[\frac{(n-1)q}{n}\right]^{\frac{1-(n-1)(q-1)}{n}} + \left[\frac{(n-1)qM}{n}\right]^{-\frac{q(n-1)}{n}}\right)M\|p_F\|_\infty^{\frac{n-1}{n}}.$$

Then, it follows that

$$\|p_F\|_\infty \le CM^n = C \max_{1\le i\le n} \|H_{(i)}(F,1)\|_p^n, \tag{A.18}$$

where C is a constant that depends on p and n. Thus (A.14) follows from (A.17) and (A.18).

The result can be generalized to the case without the assumption $\|p_F\|_\infty < \infty$ by the same argument as in Theorem 5 of [1]. □

Proof of Theorem A.4. Choose $p_2 > p_1 > n$, let $p_3 = \frac{p_1 p_2}{p_2 - p_1}$ and $q = \frac{p_1}{p_1 - 1}$. Then $\frac{1}{q} + \frac{1}{p_2} + \frac{1}{p_3} = 1$. Thus by density formula (A.11) and Hölder's inequality, we have

$$p_F(x) \le \sum_{i=1}^n \|\mathbf{1}_{B_{(x,2\rho)}}(F)\|_{p_3} \|\partial_i Q_n(F-x)\|_q \|H_{(i)}(F, \phi_\rho^x(F))\|_{p_2}. \tag{A.19}$$

Then, (A.12) is a consequence of (A.19), Lemma A.2 and A.6. Inequality (A.13) can be proved similarly. □

References

[1] Bally, V., and Caramellino, L.: Riesz transform and integration by parts formulas for random variables. *Stochastic Process. Appl.* **121**, no. 6, (2011), 1332–1355. MR-2794979

[2] Biswas, N., Etheridge, A., and Klimek, A.: The spatial Lambda-Fleming-Viot process with fluctuating selection. *arXiv preprint* arXiv:1802.08188.

[3] Chen, X., Hu, Y., Nualart, D, and Tindel, S.: Spatial asymptotics for the parabolic Anderson model driven by a Gaussian rough noise. *Electron. J. Probab.*, **22**, no. 65, (2017), 1–38. MR-3690290

[4] Dalang, R. C., and Quer-Sardanyons, L.: Stochastic integrals for SPDE's: a comparison. *Expositiones Mathematicae*, **29**, no. 1, (2011), 67–109. MR-2785545

[5] Dawson, D. A.: Infinitely divisible random measures and superprocesses. In *Stochastic Analysis and Related Topics*, vol. 31 of *Progress in Probability*. Springer, 1992, pp. 1–129. MR-1203373

[6] Dawson, D. A., and Hochberg, K. J.: The carrying dimension of a stochastic measure diffusion. *Ann. Probab.*, **7**, no. 4, (1979), 693–703. MR-0537215

[7] Dawson, D. A., and Kurtz, T. G.: Applications of duality to measure-valued diffusion processes. In *Advances in filtering and optimal stochastic control*. Springer, Berlin, 1982, pp. 91–105. MR-0794506

[8] Dawson, D. A., Vaillancourt, J., and Wang, H.: Stochastic partial differential equations for a class of interacting measure-valued diffusions. *Ann. Inst. Henri. Poincaré Probab. Stat.*, **36**, no. 2, (2000), 167–180. MR-1751657

[9] Dellacherie, C., and Meyer, P.-A.: *Probabilities and potential. B, volume 72 of North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. MR-0745449

[10] Dynkin, E. B.: *Markov processes*. Springer, Berlin, 1965. MR-0193670

[11] Ethier, S. N., and Kurtz, T. G.: *Markov processes: characterization and convergence*. J. Wiley & Sons, Hoboken, 1986. MR-0838085

[12] Friedman, A.: *Partial differential equations of parabolic type*. Courier Dover Publications, Mineola, 2008. MR-0181836

[13] Hu Y., Huang J., Lê K., Nualart D., and Tindel S.: Parabolic Anderson Model with Rough Dependence in Space. In *Abel symposium 2016: Computation and Combinatorics in Dynamics, Stochastics and Control.*, vol. 13 of *Abel Symposia*. Springer, Cham, 2016, pp. 477–498.

[14] Hu, Y., Lu, F., and Nualart, D.: Hölder continuity of the solutions for a class of nonlinear SPDE's arising from one dimensional superprocesses. *Probab. Theory and Related Fields*, **156**, no. 1-2, (2013), 27–49. MR-3055251

[15] Ikeda, N. and Watanabe, S.: *Stochastic differential equations and diffusion processes*. North-Holland Publishing Co, Amsterdam, 1989. MR-0637061

[16] Jacod, J., and Shiryaev, A. N.: *Limit theorems for stochastic processes*, vol. 288 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2013. MR-0959133

[17] Kallianpur, G., and Xiong, J.: *Stochastic differential equations in infinite dimensional spaces*, vol. 26 of *IMS Lecture Notes-Monograph Series*. Institute of Mathematical Statistics, Hayward, 1995. MR-1465436

[18] Kurtz, T. G.: The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.*, **12**, no. 33, (2007), 951–965. MR-2336594

[19] Kurtz, T. G.: Equivalence of stochastic equations and martingale problems. In *Stochastic analysis 2010*. Springer, Berlin, 2011, pp. 113–130. MR-2789081

[20] Li, Z., Wang, H., Xiong, J., and Zhou, X.: Joint continuity of the solutions to a class of nonlinear SPDEs. *Probab. Theory and Related Fields*, **153**, no. 3-4, (2012), 441–469. MR-2948683

[21] Malliavin, P., and Thalmaier, A.: *Stochastic calculus of variations in mathematical finance*. Springer, Berlin 2006. MR-2189710

[22] Mitoma, I.: Tightness of probabilities on $C([0,1];\mathscr{S}')$ and $D([0,1];\mathscr{S}')$. *Ann. Probab.*, **11**, no. 4, (1983), 989–999. MR-0714961

[23] Mytnik, L.: Superprocesses in random environments. *Ann. Probab.*, **24**, no. 4, (1996), 1953–1978. MR-1415235

[24] Nualart, D.: *The Malliavin calculus and related topics*. Springer, Berlin, 2006. MR-2200233

[25] Nualart, D., and Nualart, E.: *Introduction to Malliavin calculus*. Cambridge University Press, Cambridge, 2018. MR-3838464

[26] Perkins, E.: Dawson-Watanabe superprocesses and measure-valued diffusions. In *Lectures on probability theory and statistics*, vol. 1781 of *Lecture notes in Mathematics*. Springer, Berlin, 2002, pp. 125–329. MR-1915445

[27] Skoulakis, G., and Adler, R. J.: Superprocesses over a stochastic flow. *Ann. Appl. Probab.*, **11**, no. 2, (2001), 488–543. MR-1843056

[28] Stroock, D.: Some applications of stochastic calculus to partial differential equations. In *Ecole d'Eté de Probabilités de Saint-Flour XI-1981*, vol. 976 of *Lecture notes in Mathematics*. Springer, Berlin, 1983, pp. 267–382. MR-0722984

[29] Stroock, D. W., and Varadhan, S. R.: On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971)*, vol. 3. 1972, pp. 333–359. MR-0400425

[30] Sturm, A.: On convergence of population processes in random environments to the stochastic heat equations with colored noise. *Electron. J. Probab.*, **8**, no. 6, (2003), 1–39. MR-1986838

[31] Walsh, J. B.: An introduction to stochastic partial differential equations. In *École d'Été de Probabilités de Saint Flour XIV-1984*. Springer, Berlin, 1986, pp. 265–439. MR-0876085

[32] Wang, H.: State classification for a class of measure-valued branching diffusions in a brownian medium. *Probab. Theory and Related Fields*, **109**, no. 1, (1997), 39–55. MR-1469919

[33] Wang, H.: A class of measure-valued branching diffusions in a random medium. *Stochastic Anal. Appl.*, **16**, no. 4, (1998), 753–786. MR-1632574

[34] Xiong, J.: A stochastic log-Laplace equation. *Ann. Probab.*, **32**, no. 3B, (2004), 2362–2388. MR-2078543

[35] Xiong, J.: *Three classes of nonlinear stochastic partial differential equations*. World Scientific, Singapore, 2013. MR-3235846

[36] Yamada, T., Watanabe, S.: On the uniqueness of solutions of stochastic differential equations. *J. of Math. of Kyoto Univ.*, **11**, no. 1, (1971), 155–167. MR-0278420

Acknowledgments. We would like to thank two anonymous referees for their constructive and helpful comments. The third author would like to thank Alison Etheridge and Jie Xiong for helpful discussions on Proposition 4.13.