Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2017, Vol. 53, No. 1, 259–286
DOI: 10.1214/15-AIHP716
© Association des Publications de l'Institut Henri Poincaré, 2017

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Strong existence and uniqueness for degenerate SDE with Hölder drift

P. E. Chaudru de Raynal

Université Nice Sophia Antipolis, CNRS, LJAD, UMR 7351, 06100 Nice, France. E-mail: pe.deraynal@univ-savoie.fr

Received 5 June 2012; revised 14 September 2015; accepted 14 September 2015

Abstract. In this paper, we prove pathwise uniqueness for stochastic degenerate systems with a Hölder drift, for a Hölder exponent larger than the critical value 2/3. This work extends to the degenerate setting the earlier results obtained by Zvonkin (*Mat. Sb. (N.S.)* **93(135)** (1974) 129–149, 152), Veretennikov (*Mat. Sb. (N.S.)* **111(153)** (1980) 434–452, 480), Krylov and Röckner (*Probab. Theory Related Fields* **131**(2) (2005) 154–196) from non-degenerate to degenerate cases. The existence of a threshold for the Hölder exponent in the degenerate case may be understood as the price to pay to balance the degeneracy of the noise. Our proof relies on regularization properties of the associated PDE, which is degenerate in the current framework and is based on a parametrix method.

Résumé. Dans ce travail, on montre qu'un système hypoelliptique, composé d'une composante diffusive et d'une composante totalement dégénérée, est fortement résoluble lorsque l'exposant de la régularité Hölder de la dérive par rapport à la composante dégénérée est strictement supérieur à 2/3. Ce travail étend au cadre dégénéré les travaux antérieurs de Zvonkin (*Mat. Sb. (N.S.)* **93(135)** (1974) 129–149, 152), Veretennikov (*Mat. Sb. (N.S.)* **111(153)** (1980) 434–452, 480), Krylov et Röckner (*Probab. Theory Related Fields* **131**(2) (2005) 154–196). L'apparition d'un seuil critique pour l'exposant peut-être vue comme le prix à payer pour la dégénérescence. La preuve repose sur des résultats de régularité de la solution de l'EDP associée, qui est dégénérée, et est basée sur une méthode parametrix.

MSC: 60H10; 60H30; 34F05

Keywords: Strong uniqueness; Degeneracy; Hölder drift; Parametrix; Stochastic differential equation

1. Introduction

Let T be a positive number and d be a positive integer, we consider the following $\mathbb{R}^d \times \mathbb{R}^d$ system for any t in $[0, T]$:

$$\begin{cases} \mathrm{d}X_t^1 = F_1(t, X_t^1, X_t^2)\,\mathrm{d}t + \sigma(t, X_t^1, X_t^2)\,\mathrm{d}W_t, & X_0^1 = x_1, \\ \mathrm{d}X_t^2 = F_2(t, X_t^1, X_t^2)\,\mathrm{d}t, & X_0^2 = x_2, \end{cases} \tag{1.1}$$

where x_1, x_2 belong to $\mathbb{R}^d$, $(W_t, t \geq 0)$ is a standard d-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ and $F_1, F_2, \sigma : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{M}_d(\mathbb{R})$ (the set of real $d \times d$ matrices) are measurable functions. The diffusion matrix $a := \sigma\sigma^*$ is supposed to be uniformly elliptic. The notation "*" stands for the transpose.

In this paper, we investigate the well posedness of (1.1) outside the Cauchy–Lipschitz framework. Notably, we are interested in the strong posedness, i.e. strong existence and uniqueness of a solution. Strong existence means that there exists a process $(X_t^1, X_t^2, 0 \leq t \leq T)$ adapted to the filtration generated by the Brownian motion $(W_t, 0 \leq t \leq T)$ which satisfies (1.1). Strong uniqueness means that if two processes satisfy this equation with the same initial conditions,

their trajectories are almost surely indistinguishable. Here, we show that under a suitable Hölder assumption on the drift coefficients and Lipschitz condition on the diffusion matrix, the strong well-posedness holds for (1.1).

It may be a real challenge to prove the existence of a unique solution for a differential system without Lipschitz conditions on the coefficients. For example, in [18] Zhang showed that SDE with Sobolev coefficients admits a unique generalized solution (as Lebesgue almost everywhere stochastic flow, see Definition 2.1 in [19]) under integrability conditions on the drift, the divergence of the drift and the diffusion coefficient. The case of drift with divergence of polynomial growth being also handled.

Concerning strong solutions, the first result in this direction is due to Zvonkin. In [20], the author showed that the strong well-posedness holds for the one-dimensional system

$$Y_t = \int_0^t b(s, Y_s)\,\mathrm{d}s + W_t, \qquad Y_0 = y, t \in [0, T], \tag{1.2}$$

for a measurable function b in $\mathbb{L}^\infty$. Then, Veretennikov [16] generalized this result to the multi-dimensional case. Krylov and Röckner showed in [12] the strong well-posedness for b in $\mathbb{L}^p_{\mathrm{loc}}$, $p > d$ and Zhang in [19] handled the case of non-constant, Sobolev and non-degenerate diffusion coefficient. Finally, when b is a measurable and bounded function, Davie showed in [2] that for almost every Brownian path, there exists a unique solution for the system (1.2). We emphasize that this result implies the strong uniqueness, but the converse is not true. Indeed, in such a case, there exists an a priori set $\Omega' \subset \Omega$ with $\mathbb{P}(\Omega') = 1$ such that for all ω in Ω' the solution of (1.2) is unique.

All these results rely on the regularization of differential systems by adding a non-degenerate noise, and we refer to [6] for a partial review on this subject. The proofs of such results rely on the deep connection between SDEs and PDEs (see [1] or [9] for a partial review in the elliptic and parabolic cases). The generator associated to the Markov process Y is a linear partial differential operator of second order (usually denoted by $\mathcal{L}$) with the transition density of Y as fundamental solution. As explained by Fedrizzi and Flandoli in [5]: "if we have a good theory for the PDE:

$$\frac{\partial}{\partial t} u + \mathcal{L} u = \Phi, \quad \text{on } [0, T], u_T = \mathbf{0}, \tag{1.3}$$

where the source term Φ has the same regularity as the drift, then, we have the main tools to prove strong uniqueness".

In (1.1), the noise added is completely degenerate w.r.t. the component X^2. This sort of system has also been studied by Veretennikov in [17] but without considering any regularization in the degenerate direction. Indeed, the author showed that strong well-posedness holds when the drift is measurable and bounded and the diffusion matrix is Lipschitz w.r.t. the non-degenerate component X^1 and when both the drift and the diffusion matrix are twice continuously differentiable functions with bounded derivatives w.r.t. the degenerate component.

In this paper, we show that the noise regularizes, even in the degenerate direction, by means of the random drift. Unfortunately, there is a price to pay to balance the degeneracy of the noise. First, the drift must be at least 2/3-Hölder continuous w.r.t. the degenerate component. We do not know how sharp is this critical value, but it is consistent with our approach. Secondly, the drift F_2 of the second component must be Lipschitz continuous w.r.t. the first component and its derivative in this direction has to be uniformly non-degenerate: this allows the noise to propagate through the system and then the drift to regularize.

Our proof also relies on regularization properties of the associated PDE, and the aforementioned "good theory" is here a "strong theory": a Lipschitz bound on the solution of (1.3) and on its derivative w.r.t. the first component. We emphasize that, in our case, the generator $\mathcal{L}$ is given by: for all ψ in $C^{1,2,1}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$[1]

$$\begin{aligned}\mathcal{L}\psi(t, x_1, x_2) = \frac{1}{2}\operatorname{Tr}\big(a(t, x_1, x_2) D^2_{x_1}\psi(t, x_1, x_2)\big) + \big[F_1(t, x_1, x_2)\big] \cdot \big[D_{x_1}\psi(t, x_1, x_2)\big] \\ + \big[F_2(t, x_1, x_2)\big] \cdot \big[D_{x_2}\psi(t, x_1, x_2)\big],\end{aligned} \tag{1.4}$$

where $\operatorname{Tr}(a)$ stands for the trace of the matrix a, "$\cdot$" denotes the standard Euclidean inner product on $\mathbb{R}^d$ and where for any z in $\mathbb{R}^d$, the notation D_z means the derivative w.r.t. the variable z. Here, the crucial point is that the operator is not uniformly parabolic.

[1]That is, continuously differentiable w.r.t. the first variable, twice continuously differentiable w.r.t. the second variable and once continuously differentiable w.r.t. the third variable.

When the coefficients are smooth and when the Lie algebra generated by the vector fields spans the whole space, it is known that such an operator admits a smooth fundamental solution (see [10]): it is said to be hypoelliptic and the coefficients are said to satisfy a Hörmander condition. The assumption on the uniform non-degeneracy of the derivative of the drift F_2 w.r.t. x_1 together with the uniform ellipticity of the matrix σ can be understood as a sort of weak Hörmander condition.

This particular form of degeneracy is a non-linear generalization of Kolmogorov's degeneracy, in reference to the first work [11] of Kolmogorov in this direction. Degenerate operators of this form have been studied by many authors see e.g. the work of Di Francesco and Polidoro [4], and Delarue and Menozzi [3]. We also emphasize that, in [14], Menozzi deduced from the regularization property exhibited in ([3]) the well weak posedness of a generalization of (1.1). Nevertheless, to the best of our knowledge, there does not exist a strong theory, in the sense defined above, for the PDE (1.3) when $\mathcal{L}$ is defined by (1.4). We investigate it by using the so called parametrix approach (see [8] for partial review in the elliptic setting).

To conclude this introduction, we just come back to the regularity assumed on the drift of the non-degenerate component. In comparison with the works of Veretennikov [16,17], Krylov and Röckner [12], and Flandoli and Fedrizzi [5], asking for F_1 to be only in $\mathbb{L}^p$, $p > d$ might appear as the right framework. Since the parametrix is a perturbation method and since we are interested in $\mathbb{L}^\infty$ estimates, we suppose the drift F_1 to be Hölder continuous w.r.t. x_1.

1.1. *Organization of this paper*

Section 1.2 states useful notations. In Section 1.3 we give the detailed assumptions and the main result of this paper: strong existence and uniqueness for (1.1). In Section 1.4, we expose the strategy to prove this result. It is based on the regularization properties of the associated PDE which are, in fact, the main contribution of this work. These properties are given in Section 1.5. Finally, our main result is proved in Section 1.6.

The remainder of this paper is dedicated to the proof of the regularization properties of the associated PDE.

We present in Section 2 the linear and Brownian heuristic. It explains how the proof of the regularization properties in a simple case works and allows to understand our assumptions and how the proof in the general case can be achieved. Then, we give in Section 3 the mathematical tools, and the proof of the regularization properties of the PDE is given in Section 4. This is the technical part of this paper.

1.2. *Notations*

In order to simplify the notations, we adopt the following convention: x, y, z, ξ, etc. denote the $2d$-dimensional real variables $(x_1, x_2), (y_1, y_2), (z_1, z_2), (\xi_1, \xi_2)$, etc. Consequently, each component of the d-dimensional variables $x_k, k = 1, 2$ is denoted by $x_{kl}, l = 1, \dots, d$. We denote by $g(t, X_t)$ any function $g(t, X_t^1, X_t^2)$ from $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^N$, $N \in \mathbb{N}$. Here, $X_t = (X_t^1, X_t^2)$ and then $F(t, X_t)$ is the $\mathbb{R}^{2d}$ valued function $(F_1(t, X_t^1, X_t^2), F_2(t, X_t^1, X_t^2))^*$. We rewrite the system (1.1) in a shortened form:

$$\mathrm{d}X_t = F(t, X_t)\,\mathrm{d}t + B\sigma(t, X_t)\,\mathrm{d}W_t, \tag{1.5}$$

where B is the $2d \times d$ matrix: $B = (\mathrm{Id}, 0_{\mathbb{R}^d \times \mathbb{R}^d})^*$. "Id" stands for the identity matrix of $\mathcal{M}_d(\mathbb{R})$, the set of real $d \times d$ matrices. When necessary, we write $(X_s^{t,x})_{t \le s \le T}$ the process in (1.1) which starts from x at time t, i.e. such that $X_t^{t,x} = x$.

We recall that the canonical Euclidean inner product on $\mathbb{R}^d$ is denoted by "$\cdot$". We denote by $\mathrm{GL}_d(\mathbb{R})$ the set of $d \times d$ invertible matrices with real coefficients and by ϕ a measurable function from $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^2$. Each one-dimensional component of this function is denoted by ϕ_i, $i = 1, 2$ and plays the role of one coordinate of F_i. Hence, ϕ_i satisfies the same regularity as F_i given latter. We recall that a denotes the square of the diffusion matrix σ, $a := \sigma\sigma^*$. Subsequently, we denote by c, C, c', C', c'' etc. a positive constant, depending only on known parameters in (H), given just below, that may change from line to line and from an equation to another.

We denote by $C^{1,2,1}$ the space of functions that are continuously differentiable w.r.t. the first variable, twice continuously differentiable w.r.t. the second variable and once continuously differentiable w.r.t. the third variable.

The notation D stands for the total space derivative. For any function from $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ we denote by D_1 (resp. D_2) the derivative with respect to the first (resp. second) d-dimensional space component. In the same spirit,

the notation D_z means the derivative w.r.t. the variable z. Hence, for all integer n, D_z^n is the nth derivative w.r.t. z and for all integer m the $n \times m$ cross differentiations w.r.t. z, y are denoted by $D_z^n D_y^m$. Furthermore, the partial derivative ∂/∂_t is denoted by ∂_t.

1.3. *Main result*

Assumptions (H). *We say that assumptions (H) hold if the following assumptions are satisfied*:

(H1) Regularity of the coefficients. *There exist* $0 < \beta_i^j < 1$, $1 \le i, j \le 2$ *and three positive constants* C_1, C_2, C_σ *such that for all* t *in* $[0, T]$ *and all* (x_1, x_2) *and* (y_1, y_2) *in* $\mathbb{R}^d \times \mathbb{R}^d$

$$\begin{aligned}
&\big|F_1(t, x_1, x_2) - F_1(t, y_1, y_2)\big| \le C_1\big(|x_1 - y_1|^{\beta_1^1} + |x_2 - y_2|^{\beta_1^2}\big),\\
&\big|F_2(t, x_1, x_2) - F_2(t, y_1, y_2)\big| \le C_2\big(|x_1 - y_1| + |x_2 - y_2|^{\beta_2^2}\big),\\
&\big|\sigma(t, x_1, x_2) - \sigma(t, y_1, y_2)\big| \le C_\sigma\big(|x_1 - y_1| + |x_2 - y_2|\big).
\end{aligned}$$

Moreover, the coefficients are supposed to be continuous w.r.t. the time and the exponents $\beta_i^2, i = 1, 2$ *are supposed to be strictly greater than* $2/3$. *Thereafter, we set* $\beta_2^1 = 1$ *for notational convenience.*

(H2) Uniform ellipticity of $\sigma\sigma^*$. *The function* $\sigma\sigma^*$ *satisfies the uniform ellipticity hypothesis*:

$$\exists \Lambda > 1, \forall \zeta \in \mathbb{R}^{2d}, \quad \Lambda^{-1}|\zeta|^2 \le \big[\sigma\sigma^*(t, x_1, x_2)\zeta\big] \cdot \zeta \le \Lambda|\zeta|^2,$$

for all $(t, x_1, x_2) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$.

(H3-a) Differentiability and regularity of $x_1 \mapsto F_2(\cdot, x_1, \cdot)$. *For all* $(t, x_2) \in [0, T] \times \mathbb{R}^d$, *the function* $F_2(t, \cdot, x_2) : x_1 \mapsto F_2(t, x_1, x_2)$ *is continuously differentiable and there exist* $0 < \alpha^1 < 1$ *and a positive constant* $\bar{C}_2$ *such that, for all* (t, x_2) *in* $[0, T] \times \mathbb{R}^d$ *and* x_1, y_1 *in* $\mathbb{R}^d$

$$\big|D_1 F_2(t, x_1, x_2) - D_1 F_2(t, y_1, x_2)\big| \le \bar{C}_2|x_1 - y_1|^{\alpha^1}.$$

(H3-b) Non-degeneracy of $(D_1 F_2)(D_1 F_2)^*$. *There exists a closed convex subset* $\mathcal{E} \subset \mathrm{GL}_d(\mathbb{R})$ (*the set of* $d \times d$ *invertible matrices with real coefficients*) *such that for all* t *in* $[0, T]$ *and* (x_1, x_2) *in* $\mathbb{R}^{2d}$ *the matrix* $D_1 F_2(t, x_1, x_2)$ *belongs to* $\mathcal{E}$. *We emphasize that this implies that*

$$\exists \bar{\Lambda} > 1, \forall \zeta \in \mathbb{R}^{2d}, \quad \bar{\Lambda}^{-1}|\zeta|^2 \le \big[(D_1 F_2)(D_1 F_2)^*(t, x_1, x_2)\zeta\big] \cdot \zeta \le \bar{\Lambda}|\zeta|^2,$$

for all $(t, x_1, x_2) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$.

Remark 1. *The reason for the existence of the critical value* $2/3$ *for the Hölder regularity of the drift in* (H1) *and the particular "convexity" assumption* (H3-b) *are discussed in Section 2. In the following, the sentence "known parameters in (H)" refers to the parameters belonging to these assumptions.*

The following Theorem is the main result of this paper and regards the strong well-posedness of the system (1.1).

Theorem 1.1. *Under (H), strong existence and uniqueness hold for (1.1) for any positif* T.

1.4. *Strategy of proof*

Let us expose the basic arguments to prove Theorem 1.1. Existence of a weak solution follows from a compactness argument. Then, if the strong uniqueness holds, the strong existence follows. The main issue consists in proving the strong uniqueness.

This works as follows: suppose that there exists a unique $C^{1,2,1}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ solution $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)^*$ of the linear system of PDEs:

$$\begin{cases} \partial_t \mathbf{u}_i(t, x) + \mathcal{L}\mathbf{u}_i(t, x) = F_i(t, x), & \text{for } (t, x) \in [0, T) \times \mathbb{R}^{2d},\\ \mathbf{u}_i(T, x) = 0_{\mathbb{R}^d}, & i = 1, 2, \end{cases} \tag{1.6}$$

then, thanks to Itô's formula, for all t in $[0, T]$ we have that

$$\int_0^t F\big(s, X_s^x\big)\,\mathrm{d}s = \mathbf{u}\big(t, X_t^x\big) - \mathbf{u}(0, x) - \int_0^t D\mathbf{u}\big(s, X_s^x\big) B\sigma\big(s, X_s^x\big)\,\mathrm{d}W_s.$$

Thus, if the functions $\mathbf{u}$ and $D\mathbf{u}B = (D_1\mathbf{u}, 0_{\mathbb{R}^d\times\mathbb{R}^d})^*$ are C_T Lipschitz continuous, the drift $\int_0^t F(s, X_s^x)\,\mathrm{d}s$ is Lipschitz continuous w.r.t. the argument X^x, with Lipschitz constant C_T.

Now suppose that C_T tends to 0 when T goes to 0. Uniqueness then follows by mean of classical and circular arguments for T small enough. Since the strategy can be iterated, one can deduce strong uniqueness on any positive interval.

The main issue here is then to obtain a strong theory for the PDE (1.6). The problem is that, to the best of our knowledge, it is not known that this PDE admits a Lipschitz $C^{1,2,1}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ solution under our weak Hörmander assumptions. Nevertheless, in our analysis we do not need to obtain the existence of a regular solution, but only the existence of the Lipschitz bounds for u and D_1u depending only on known parameters in (H).

Therefore, we investigate these bounds in a smooth setting. Thanks to assumptions (H), there exists a sequence of smooth (say infinitely differentiable with bounded derivatives of all order greater than 1) mollified coefficients $(a^n, F_1^n, F_2^n)_{n\geq 1}$ satisfying (H) uniformly (in n), such that this sequence converges in supremum norm to (a, F_1, F_2). More details on the regularization procedure are given in Section 1.5 below. In this regularized setting, we know that the PDE (1.6) admits a unique smooth solution.

Then, thanks to a first order parametrix expansion of the solution $\mathbf{u}^n$ of the regularized PDE (1.6) (see e.g. [8]), we show that for T small enough there exists a positive constant C_T, which is independent of the regularization procedure, such that the supremum norm of $D_1\mathbf{u}^n$, $D_2\mathbf{u}^n$, $D_1^2\mathbf{u}^n$ and $D_2D_1\mathbf{u}^n$ are bounded by C_T. We then recover the Lipschitz regularity of the drift in small time and we obtain strong uniqueness by letting the regularization procedure tends to the infinity.

1.5. *PDE's results*

This section summarizes the PDE's results used for proving Theorem 1.1.

The mollifying procedure. Let us first detail how the smooth approximation of the coefficients a, F_1, F_2 works. For all positive integer n, we set:

$$F_2^n(t, x) = \int F_2(t - s, x - y)\varphi_1^n(y)\varphi_2^n(s)\,\mathrm{d}y\,\mathrm{d}s,$$

where $\varphi_1^n(\cdot) = c_1 n^{2d}\varphi(n|\cdot|)$ and $\varphi_2^n(\cdot) = c_2 n\varphi(n|\cdot|)$ for c_1, c_2 two constants of normalization and for a smooth function φ with support in the unit ball. For example $\varphi : z \in \mathbb{R} \mapsto \exp(-\frac{1}{1-z^2})\mathbf{1}_{]-1;1[}(z)$.

By defining $(F_1^n)_{n\geq 1}$ and $(a^n)_{n\geq 1}$ with the same procedure, it is then clear that for every n the mollified coefficients a^n, F_1^n, F_2^n are infinitely differentiable with bounded derivatives of all order greater than 1 and such that

$$\big(a^n, F_1^n, F_2^n\big) \underset{n\to+\infty}{\longrightarrow} (a, F_1, F_2), \tag{1.7}$$

uniformly on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$. Moreover, it is well-seen that they satisfy the same assumptions as (a, F_1, F_2) uniformly in n.

Let us just check the non-degeneracy assumption (H3-b) on $D_1F_2^n$. For all positive δ, one can find a positive integer $N(\delta)$ and a sequence of rectangles $(R_k)_{1\leq k\leq N(\delta)}$ having sides of length less than δ and a family of points $\{(s_k, y_k) \in R_k, 1 \leq k \leq N(\delta)\}$ such that, for all (t, x) in $[0, T] \times \mathbb{R}^{2d}$:

$$D_1F_2^n(t, x) = \lim_{\delta\to 0}\sum_{k=1}^{N(\delta)} D_1F_2(t - s_k, x - y_k)\int_{R_k}\varphi_1^n(y)\varphi_2^n(s)\,\mathrm{d}y\,\mathrm{d}s.$$

Since D_1F_2 belongs to the closed convex subset $\mathcal{E}$, it is clear that $D_1F_2^n$ belongs to $\mathcal{E}$.

The regularized PDE. As we said, we do not solve the limit PDE problem (1.6). The investigations are done with the mollified coefficients $(a^n, F_1^n, F_2^n)_{n\geq 1}$ defined above. Let us denote by $\mathcal{L}^n$ the regularized version of $\mathcal{L}$ (that is the version of $\mathcal{L}$ with mollified coefficients). We have from Section 2.1, Chapter II of [7] (note that the time dependence here is not a problem to do so):

Lemma 1.2. *Let n be a positive integer. The PDE,*

$$\partial_t \mathbf{u}_i^n(t,x) + \mathcal{L}^n \mathbf{u}_i^n(t,x) = F_i^n(t,x), \quad \text{for } (t,x) \in [0,T] \times \mathbb{R}^{2d}, \mathbf{u}_i^n(T,x) = 0_{\mathbb{R}^d}, i = 1,2, \tag{1.8}$$

where $\mathcal{L}^n$ is the regularized version of the operator $\mathcal{L}$ defined by (1.4), admits a unique solution $\mathbf{u}^n = (\mathbf{u}_1^n, \mathbf{u}_2^n)^*$, which is infinitely differentiable.

Besides, the terminal condition $\mathbf{u}^n(T,\cdot) = 0_{\mathbb{R}^{2d}}$ is very important: it guarantees that the solution and its derivatives vanish at time T. Hence, it allows to control the Lipschitz constant of $\mathbf{u}^n$ by a constant small as T is small. Indeed, we show in Section 4 that the solutions $\mathbf{u}^n, n \geq 1$ satisfy:

Proposition 1.3. *There exist a positive $\mathcal{T}$, a positive $\delta_{1.3}$ and a positive constant C, depending only on known parameters in (H) and not on n, such that, for all positive T less than $\mathcal{T}$:*

$$\left\| D_1 \mathbf{u}^n \right\|_\infty + \left\| D_2 \mathbf{u}^n \right\|_\infty + \left\| D_1^2 \mathbf{u}^n \right\|_\infty + \left\| D_1 D_2 \mathbf{u}^n \right\|_\infty \leq C T^{\delta_{1.3}}.$$

In order to prove these results, we emphasize that each coordinate of the vectorial solution $\mathbf{u}_i^n$ of the decoupled linear PDE (1.8) can be described by the PDE

$$\partial_t u_i^n(t,x) + \mathcal{L}^n u_i(t,x) = \phi_i^n(t,x), \quad \text{for } (t,x) \in [0,T] \times \mathbb{R}^{2d}, u_i^n(T,x) = 0, i = 1,2, \tag{1.9}$$

where $\phi_i^n : \mathbb{R}^{2d} \to \mathbb{R}$ denotes the mollified (by the procedure described above) coefficient ϕ_i: a function that satisfies the same regularity assumptions as F_i given in (H1) (this function plays the role of one of the coordinates of F_i). Therefore, we only have to prove Lemma 1.2 and Proposition 1.3 for (1.9) instead of (1.8).

Since the estimates on the solutions $\mathbf{u}^n, n \geq 1$ are obtained uniformly in n (that is independently of the procedure of regularization), when we investigate the properties of the solution of the PDE (1.9) in the following sections, we forget the superscript "n" which arises from the mollifying procedure, and we further assume that the following assumptions hold:

Assumptions (HR). *We say that assumptions (HR) hold if assumptions (H) hold true and F_1, F_2, ϕ_1, ϕ_2 and a are infinitely differentiable functions with bounded derivatives of all order greater than* 1.

1.6. *Proof of Theorem 1.1*

We know from Theorem 6.1.7 of [15] that the system (1.1) admits a weak solution (we emphasize that this result remains valid under the linear growth conditions assumed on the coefficients). Hence, we only have to prove the strong uniqueness. Thereafter, we denote by "$\mathbf{1}$" the $2d \times 2d$ matrix:

$$\begin{pmatrix} \mathrm{Id} & 0_{\mathbb{R}^d \times \mathbb{R}^d} \\ 0_{\mathbb{R}^d \times \mathbb{R}^d} & 0_{\mathbb{R}^d \times \mathbb{R}^d} \end{pmatrix}. \tag{1.10}$$

Let $(X_t, t \geq 0)$ and $(Y_t, t \geq 0)$ be two solutions of (1.1) with the same initial condition x in $\mathbb{R}^{2d}$. Let $\mathbf{u}^n$ be the solution of the linear system of PDEs (1.8). Thanks to Lemma 1.2, we can apply Itô's formula on $\mathbf{u}^n(t, X_t) - X_t$ and we obtain

$$\begin{aligned} \mathbf{u}^n(t, X_t) - X_t = &\int_0^t \left[\partial_t \mathbf{u}^n + \mathcal{L}\mathbf{u}^n\right](s, X_s)\,\mathrm{d}s - \int_0^t F(s, X_s)\,\mathrm{d}s + \mathbf{u}^n(0,x) - x \\ &+ \int_0^t \left[D\mathbf{u}^n - \mathbf{1}\right] B\sigma(s, X_s)\,\mathrm{d}W_s. \end{aligned}$$

In order to use the fact that $\partial_t \mathbf{u}^n + \mathcal{L}^n \mathbf{u}^n = F^n$, we rewrite

$$\mathbf{u}^n(t, X_t) - X_t = \int_0^t \big[\partial_t \mathbf{u}^n + \mathcal{L}^n \mathbf{u}^n\big](s, X_s)\,\mathrm{d}s + \int_0^t \big(\mathcal{L} - \mathcal{L}^n\big)\mathbf{u}^n(s, X_s)\,\mathrm{d}s - \int_0^t F(s, X_s)\,\mathrm{d}s + \mathbf{u}^n(0, x) - x$$
$$+ \int_0^t \big[D\mathbf{u}^n - \mathbf{1}\big] B\sigma(s, X_s)\,\mathrm{d}W_s,$$

and then,

$$\mathbf{u}^n(t, X_t) - X_t = \int_0^t \big(\mathcal{L} - \mathcal{L}^n\big)\mathbf{u}^n(s, X_s)\,\mathrm{d}s + \int_0^t \big(F^n(s, X_s) - F(s, X_s)\big)\,\mathrm{d}s + \mathbf{u}^n(0, x) - x$$
$$+ \int_0^t \big[D\mathbf{u}^n - \mathbf{1}\big] B\sigma(s, X_s)\,\mathrm{d}W_s.$$

By the same arguments, we have:

$$\mathbf{u}^n(t, Y_t) - Y_t = \int_0^t \big(\mathcal{L} - \mathcal{L}^n\big)\mathbf{u}^n(s, Y_s)\,\mathrm{d}s + \int_0^t \big(F^n(s, Y_s) - F(s, Y_s)\big)\,\mathrm{d}s + \mathbf{u}^n(0, x) - x$$
$$+ \int_0^t \big[D\mathbf{u}^n - \mathbf{1}\big] B\sigma(s, Y_s)\,\mathrm{d}W_s.$$

By taking the expectation of the supremum over t of the square norm of the difference of the two equalities above, we get from Doob's inequality that

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t - Y_t|^2\Big] \le C\bigg\{\mathbb{E}\Big[\sup_{t\in[0,T]} \big|\mathbf{u}^n(t, X_t) - \mathbf{u}^n(t, Y_t)\big|^2\Big]$$
$$+ \mathbb{E}\bigg[\int_0^T \big|\big[D\mathbf{u}^n B\big](s, X_s) - \big[D\mathbf{u}^n B\big](s, Y_s)\big|^2 \|\sigma\|_\infty^2\,\mathrm{d}s\bigg]$$
$$+ \mathbb{E}\bigg[\int_0^T \big(\big\|D\mathbf{u}^n B\big\|_\infty + \mathbf{1}\big)\big|\big[\sigma(s, Y_s) - \sigma(s, X_s)\big]\big|^2\,\mathrm{d}s\bigg] + \mathcal{R}(n, T)\bigg\},$$

where

$$\mathcal{R}(n, T) = T\Big(\mathbb{E}\Big[\sup_{t\in[0,T]} \big|F^{(n)}(t, Y_t) - F(t, Y_t)\big|^2\Big] + \mathbb{E}\Big[\sup_{t\in[0,T]} \big|\big(\mathcal{L}^n - \mathcal{L}\big)\mathbf{u}^n(t, Y_t)\big|^2\Big]$$
$$+ \mathbb{E}\Big[\sup_{t\in[0,T]} \big|F^{(n)}(t, X_t) - F(t, X_t)\big|^2\Big] + \mathbb{E}\Big[\sup_{t\in[0,T]} \big|\big(\mathcal{L}^n - \mathcal{L}\big)\mathbf{u}^n(t, X_t)\big|^2\Big]\Big).$$

First, note that from (1.7), for both Y_t and X_t, we have:

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \big|F^n(t, X_t) - F(t, X_t)\big|^2\Big] + \mathbb{E}\Big[\sup_{t\in[0,T]} \big|\big(\mathcal{L}^n - \mathcal{L}\big)\mathbf{u}^n(t, X_t)\big|^2\Big] \to 0, \quad \text{as } n \to \infty,$$

so that $\mathcal{R}(n, T)$ tends to 0 when n tends to $+\infty$. Secondly, we know from Proposition 1.3, that there exists a positive $\mathcal{T}$ such that for all T less than $\mathcal{T}$ and for all t in $[0, T]$, the functions $\mathbf{u}^n$ and $D_1\mathbf{u}^n$ are Lipschitz continuous in space, with a Lipschitz constant independent of n. Since $D\mathbf{u}^n B = (D_1\mathbf{u}^n, 0_{\mathbb{R}^d\times\mathbb{R}^d})$, by letting n tends to $+\infty$ and using the two arguments above, we deduce that for all T less than $\mathcal{T}$:

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t - Y_t|^2\Big] \le C(T)\bigg\{\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t - Y_t|^2\Big] + \mathbb{E}\bigg[\int_0^T |X_s - Y_s|^2\,\mathrm{d}s\bigg]\bigg\},$$

where $C(T)$ tends to 0 when T tends to 0. Hence, one can now find another positive $\mathcal{T}'$, depending only on known parameters in (H), such that strong uniqueness holds for all T less than $\mathcal{T}'$. By iterating this computation, the same result holds on any finite intervals and so on $[0,\infty)$. □

2. The linear and Brownian heuristic

This section introduces the main issue when proving Proposition 1.3 in a simple case. It allows to understand our strategy and the role of some of the assumptions in (H). Furthermore, this presents in a simple form the effects of the degeneracy. By "simple", we mean that the assumptions (HL) below hold true.

Assumptions (HL). *We say that hypotheses (HL) hold if (H) and (HR) hold with*: $F_1 \equiv 0_{\mathbb{R}^d}$, $\sigma \equiv \mathrm{Id}$, *for all* (t,x_1,x_2) *in* $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$, $F_2(t,x_1,x_2)=\bar{F}_2(x_2)+\Gamma_t x_1$. *This implies in particular that for all* t *in* $[0,T]$, Γ_t *belongs to the convex subset* $\mathcal{E}$ *of* $\mathrm{GL}_d(\mathbb{R})$.

Under (HL), the SDE (1.1) becomes:

$$\begin{cases} \mathrm{d}X_s^{1,t,x}=\mathrm{d}W_s, & X_t^{1,t,x}=x_1,\\ \mathrm{d}X_s^{2,t,x}=(\bar{F}_2(X_s^{2,t,x})+\Gamma_s X_s^{1,t,x})\,\mathrm{d}s, & X_t^{2,t,x}=x_2,\end{cases} \tag{2.1}$$

for all $t<s$ in $[0,T]^2$, x in $\mathbb{R}^{2d}$ and admits a unique strong solution X. We recall that the associated PDE is

$$\begin{cases} \partial_t u_i(t,x)+\mathcal{L}u_i(t,x)=\phi_i(t,x), & \text{for } (t,x)\in[0,T]\times\mathbb{R}^{2d},\\ u_i(T,x)=0, & i=1,2,\end{cases} \tag{2.2}$$

where $\mathcal{L}$ is the generator of (2.1).

Our strategy to study the solution of the PDE (2.2) rests upon parametrix. This method is based on the following observation: in small time, the generator of the solution of an SDE with smooth and variable coefficients and the generator of the solution of the same SDE with fixed (frozen at some point) coefficients are "closed". The variable generator is then seen as a perturbation of the frozen generator, which usually enjoys well known properties.

Here, we know the explicit form of the fundamental solution of the frozen generator (which is the transition density of the solution of the frozen SDE). Especially, we can prove that this fundamental solution and its derivatives admit Gaussian type bounds. Hence, thanks to the parametrix, we write the solution of the PDE (2.2) as a time–space integral of some perturbed kernel against this fundamental solution and we can study it.

We emphasize that the choice of the freezing point for the coefficients plays a central role: the perturbation done in the parametrix has to be of the order of the typical trajectories of the process associated to the frozen operator.

2.1. *The frozen system*

Kolmogorov's example. To understand how the frozen system could be, we go back to the work of Kolmogorov [11] where the author studied the prototype system (1.1). When $d=1$, Kolmogorov showed that the solution of $\mathrm{d}Y_s^1=\mathrm{d}W_s$, $\mathrm{d}Y_s^2=\alpha Y_s^1\,\mathrm{d}s$ $(\alpha\neq 0)$, with initial condition (x_1,x_2) in $\mathbb{R}^2$, admits a density. Notably, this density is Gaussian and given by, for all s in $(0,T]$ and (y_1,y_2) in $\mathbb{R}^2$

$$p(0,x_1,x_2;s,y_1,y_2)=\frac{\sqrt{3}}{\alpha\pi s^2}\exp\left(-\frac{1}{2}\left|K_s^{-1/2}(y_1-x_1,y_2-x_2-s\alpha x_1)^*\right|^2\right), \tag{2.3}$$

with the following covariance matrix

$$K_s:=\begin{pmatrix} s & (1/2)\alpha s^2\\ (1/2)\alpha s^2 & (1/3)\alpha^2 s^3\end{pmatrix}. \tag{2.4}$$

This example illustrates the behaviour of the system in small time: it is not diffusive. The first coordinate oscillates with fluctuations of order $1/2$, while the second one oscillates with fluctuations of order $3/2$. As a direct consequence,

the transport of the initial condition of the first coordinate has a key role in the second one. This observation is crucial when freezing the coefficients.

The frozen system. As Kolmogorov's example suggests, we have to keep track of the transport of the initial condition when we freeze the coefficients. This allows us to reproduce a perturbation of the order of the typical trajectories of the frozen process. Then, we freeze the system (2.1) along the forward transport $\theta_{\tau,s} = (\theta^1_{\tau,s}, \theta^2_{\tau,s})^*$, s in $[\tau, T]$ that solves the ODE:

$$\frac{\mathrm{d}}{\mathrm{d}s}\theta_{\tau,s} = \big(0_{\mathbb{R}^d}, \bar{F}_2\big(\theta^2_{\tau,s}(\xi)\big) + \Gamma_s\theta^1_{\tau,s}(\xi)\big)^*, \quad \theta_{\tau,\tau}(\xi) = \xi, \tag{2.5}$$

for a given (τ, ξ) in $[0, T] \times \mathbb{R}^{2d}$ (we emphasize that in the regularized setting this ODE is well-posed). Hence, we obtain the following frozen system

$$\begin{cases} \mathrm{d}\bar{X}^{1,t,x}_s = \mathrm{d}W_s, & \bar{X}^{1,t,x}_t = x_1, \\ \mathrm{d}\bar{X}^{2,t,x}_s = (\bar{F}_2(\theta^2_{\tau,s}(\xi)) + \Gamma_s\bar{X}^{1,t,x}_s)\,\mathrm{d}s, & \bar{X}^{2,t,x}_t = x_2, \end{cases} \tag{2.6}$$

for all s in $(t, T]$. This is our candidate to approximate (2.1).[2] Obviously, in order to reproduce the typical trajectories of the frozen process, the couple of variables (τ, ξ) in (2.5) will be chosen as the initial data (t, x) of the solution of the SDE (2.6).

2.2. *Existence and Gaussian bound of the density of the frozen system*

In this case, the crucial point is the specific form of the covariance matrix $\bar{\Sigma}_{t,\cdot}$ of $\bar{X}^{t,x}_\cdot$ defined in (2.6). For any s in $(t, T]$, standard computations show that

$$\bar{\Sigma}_{t,s} = \begin{pmatrix} (s-t) & \int_t^s\int_t^r \Gamma_u\,\mathrm{d}u\,\mathrm{d}r \\ \int_t^s\int_t^r \Gamma_u\,\mathrm{d}u\,\mathrm{d}r & \int_t^s(\int_t^r \Gamma_u\,\mathrm{d}u)(\int_t^r \Gamma_u\,\mathrm{d}u)^*\,\mathrm{d}r \end{pmatrix}. \tag{2.7}$$

Therefore, the existence and the Gaussian estimates of the transition density of $\bar{X}^{t,x}_s$ stem from the control of the spectrum of $\bar{\Sigma}_{t,s}$. Such an investigation has been already done by Delarue and Menozzi in [3]. The two following Lemmas shortly describe some of their results that are useful for us. The proofs are not given. For further details, we refer to Section 3, pp. 18–24 of their paper. They prove that

Lemma 2.1. *Suppose that assumptions (HL) hold, then, a sufficient condition for the non-degeneracy of the variance matrix $\bar{\Sigma}_{t,s}$, s in $(t, T]$ is given by*

$$\det[\Gamma_r] > 0 \quad \textit{for a.e. } r \in [t, s].$$

In that case, the solution of (2.6) admits a transition density $\bar{q}$ given by, for all s in $(t, T]$ and y_1, y_2 in $\mathbb{R}^d$:

$$\bar{q}(t, x_1, x_2; s, y_1, y_2) = \frac{1}{(2\pi)^{d/2}}\det(\bar{\Sigma}_{t,s})^{-1/2}\exp\bigg(-\frac{1}{2}\big|\bar{\Sigma}^{-1/2}_{t,s}\big(y_1 - x_1, y_2 - m^{2,\tau,\xi}_{t,s}(x)\big)^*\big|^2\bigg), \tag{2.8}$$

where

$$m^{2,\tau,\xi}_{t,s}(x) = x_2 + \int_t^s \Gamma_r x_1\,\mathrm{d}r + \int_t^s \bar{F}_2\big(\theta^2_{\tau,r}(\xi)\big)\,\mathrm{d}r,$$

and where $\bar{\Sigma}_{t,s}$ is the uniformly non-degenerate matrix given by (2.7).

From this expression, we can give the following Gaussian type estimate on the transition density of the solution of the SDE (2.6) and on its derivatives (the Gaussian bounds on the derivatives are not proven in the aforementioned work, but are proven in a more general case in the proof of Proposition 3.1 below):

[2]Note that if $\tau > s$, we can extend the definition of θ and suppose that $\forall(v > r) \in [0, T]^2, \forall\xi \in \mathbb{R}^{2d}, \theta_{v,r}(\xi) = 0$.

Lemma 2.2. *Suppose that assumptions (HL) hold, then, the transition kernel $\bar{q}$ and its derivatives admit Gaussian-type bounds: there exists a positive constant c depending only on known parameters in (H) such that for all ξ in $\mathbb{R}^{2d}$:*

$$\begin{aligned}&\left|D_{x_1}^{N^{x_1}} D_{x_2}^{N^{x_2}} D_{y_1}^{N^{y_1}} \bar{q}(t, x_1, x_2; s, y_1, y_2)\right| \\ &\quad \le (s-t)^{-[3N^{x_2}+N^{x_1}+N^{y_1}]/2} \frac{c}{(s-t)^{2d}} \exp\left(-c\left(\frac{|y_1-x_1|^2}{s-t} + \frac{|y_2 - m_{t,s}^{2,\tau,\xi}(x)|^2}{(s-t)^3}\right)\right),\end{aligned} \tag{2.9}$$

for all s in $(t, T]$, y_1, y_2 in $\mathbb{R}^d$ and any N^{x_1}, N^{x_2}, N^{y_1} less than 2.

We emphasize that the constant c that appears in the exponential in estimate (2.9) does not depend on Γ, as suggested by Lemma 2.1. This uniform control is not obvious and is related to the "closed convex" assumption (H3-b).

If the control is not uniform, Delarue and Menozzi show in [3] (see Example 3.5) that one can find a sequence of matrix $(\Gamma_1^n)_{n\ge 0}$ with positive constant determinant such that $\det[\Sigma_{0,1}^n]$ converges towards 0 as n tends to the infinity. The crucial point in their example is that the sequence of functions $(t \in [0,1] \mapsto \Gamma_t^n)_{n\ge 0}$ weakly converges towards 0. Hence, to overcome this problem, the authors need some closure for the weak topology. The closed convex assumption (H3-b) allows them to obtain compactness for the weak topology.

Note that for all s in $[t, T]$, the mean $(x_1, m_{t,s}^{2,t,x}(x))$ of $\bar{X}_s^{t,x}$ satisfies the ODE (2.5) with initial data $(\tau, \xi) = (t, x)$. Since this equation admits a unique solution under (HL), we deduce that for all s in $[t, T]$, the forward transport function defined by (2.5) with the starting point x as initial condition is equal to the mean: $\theta_{t,s}(x) = m_{t,s}^{t,x}(x)$.

Finally, as we will show in the proof of Proposition 3.1 in Section 3, the transition density $\bar{q}$ enjoys the following symmetry property:

$$\forall (t < s, x, y) \in [0, T]^2 \times \mathbb{R}^d \times \mathbb{R}^d, \quad D_{x_2}\bar{q}(t, x; s, y) = -D_{y_2}\bar{q}(t, x; s, y). \tag{2.10}$$

This plays a crucial role in the proof of the Lipschitz estimates of the solution u of the PDE (2.2) and of its derivative $D_1 u$.

2.3. *Representation of the solution by parametrix*

Let $\bar{\mathcal{L}}^{\tau,\xi} := (1/2)\Delta_{x_1} + [\bar{F}_2(\theta_{\tau,t}^2(\xi)) + \Gamma_t x_1] \cdot D_{x_2}$ be the generator of the frozen process $\bar{X}$. We can write the PDE (2.2) as

$$\begin{cases} \partial_t u_i(t, x) + \bar{\mathcal{L}}^{\tau,\xi} u_i(t, x) = \phi_i(t, x) + (\bar{\mathcal{L}}^{\tau,\xi} - \mathcal{L}) u_i(t, x), & \text{for } (t, x) \in [0, T) \times \mathbb{R}^{2d}, \\ u_i(T, x) = 0, & i = 1, 2. \end{cases}$$

Since $\bar{q}$ is the fundamental solution of $\bar{\mathcal{L}}^{\tau,\xi}$ we have, for all (t, x) in $[0, T] \times \mathbb{R}^{2d}$, that the unique solution of this PDE reads:

$$u_i(t, x) = \int_t^T \int_{\mathbb{R}^{2d}} \left\{\phi_i(s, y) - \left[\bar{F}_2(y_2) - \bar{F}_2\left(\theta_{\tau,s}^2(\xi)\right)\right] \cdot D_{y_2} u_i(s, y)\right\} \bar{q}(t, x; s, y)\, \mathrm{d}y\, \mathrm{d}s,$$

for $i = 1, 2$ and for all (τ, ξ) in $[0, T] \times \mathbb{R}^{2d}$. We emphasize that u is the unique solution of (2.2) and does not depend on the choice of the freezing point "(τ, ξ)". For every (t, x) in $[0, T] \times \mathbb{R}^{2d}$, we can then choose τ as the current evaluation time and then write

$$u_i(t, x) = \int_t^T \int_{\mathbb{R}^{2d}} \left\{\phi_i(s, y) - \left[\bar{F}_2(y_2) - \bar{F}_2\left(\theta_{t,s}^2(\xi)\right)\right] \cdot D_{y_2} u_i(s, y)\right\} \bar{q}(t, x; s, y)\, \mathrm{d}y\, \mathrm{d}s. \tag{2.11}$$

From now on, for every given time t in $[0, T]$, we let $\tau = t$.

2.4. *A priori estimates*

In order to prove the bounds of Proposition 1.3, we need to obtain estimates of the supremum norm of the first and second order derivatives of the u_i, $i = 1, 2$. Having in mind to invert the differentiation and integral operators, we see any differentiation of u_i as an integral of a certain function against the derivative of the degenerate Gaussian kernel $\bar{q}$.

As shown in Lemma 2.2, the differentiation of this kernel generates a time-singularity. Each differentiation of the transition kernel w.r.t. the first component gives a time-singularity of order $1/2$ while the differentiation w.r.t. the second component gives a time-singularity of order $3/2$.

The main issue consists in smoothing this singularity by using the regularity of the coefficients assumed in (H) together with Gaussian decay in $\bar{q}$ by letting the freezing point ξ be the starting point of the process.

Let us illustrate the computations with the worst case in Proposition 1.3, that is, the cross derivative $D_1 D_2 u_i$ (which gives a time singularity of order 2, see Lemma 2.2). In order to invert the integral and the differentiation operators we have to show that the derivative of the time integrand in (2.11) is suitably bounded. When it is evaluated at point $\xi = x$ we show in the sequel that this is indeed the case. Let (t, x) in $[0, T] \times \mathbb{R}^{2d}$, for i in $\{1, 2\}$, we denote by $\mathcal{I}_i(s, x)$, s in $(t, T]$, the time integrand in (2.11). By switching the differentiation and the (space) integral operator we have:

$$\begin{aligned} &D_{x_1} D_{x_2} \mathcal{I}_i(s, x) \\ &\quad = \int_{\mathbb{R}^{2d}} \big\{\phi_i(s, y) - \big[\bar{F}_2(y_2) - \bar{F}_2\big(\theta^2_{t,s}(\xi)\big)\big] \cdot D_{y_2} u_i(s, y)\big\} D_{x_1} D_{x_2} \bar{q}(t, x; s, y)\, \mathrm{d}y. \end{aligned} \tag{2.12}$$

We deal with the right hand side of (2.12). Let us first deal with the first term in the time-integrand. For all s in $(t, T]$ we have:

$$\begin{aligned} &\int_{\mathbb{R}^{2d}} \phi_i(s, y_1, y_2) D_{x_1} D_{x_2} \bar{q}(t, x_1, x_2; s, y_1, y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2 \\ &\quad = \int_{\mathbb{R}^{2d}} \big(\phi_i(s, y_1, y_2) - \phi_i\big(s, y_1, \theta^2_{t,s}(\xi)\big)\big) D_{x_1} D_{x_2} \bar{q}(t, x_1, x_2; s, y_1, y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2 \\ &\qquad + \int_{\mathbb{R}^{2d}} \phi_i\big(s, y_1, \theta^2_{t,s}(\xi)\big) D_{x_1} D_{x_2} \bar{q}(t, x_1, x_2; s, y_1, y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2, \end{aligned}$$

where, thanks to the symmetry (2.10) and an integration by parts argument, the last term in the right hand side is equal to 0. In the sequel, we refer to this argument as *the centering argument*. Combining this argument and the estimate of $D_{x_1} D_{x_2} \bar{q}$ in Lemma 2.2, we have, for all s in $(t, T]$:

$$\begin{aligned} &\bigg|\int_{\mathbb{R}^{2d}} \big(\phi_i(s, y_1, y_2) - \phi_i\big(s, y_1, \theta^2_{t,s}(\xi)\big)\big) D_{x_1} D_{x_2} \bar{q}(t, x_1, x_2; s, y_1, y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2\bigg| \\ &\quad \le \int_{\mathbb{R}^{2d}} \bigg\{(s-t)^{-2} \big|\phi_i(s, y_1, y_2) - \phi_i\big(s, y_1, \theta^2_{t,s}(\xi)\big)\big| \\ &\qquad \times \frac{c}{(s-t)^{2d}} \exp\bigg(-c\bigg(\frac{|y_1 - x_1|^2}{s-t} + \frac{|y_2 - m^{2,t,\xi}_{t,s}(x)|^2}{(s-t)^3}\bigg)\bigg)\bigg\}\, \mathrm{d}y_1\, \mathrm{d}y_2, \end{aligned}$$

where c depends only on known parameters in (H). By using the Hölder regularity of ϕ_i assumed in (H1) we have

$$\begin{aligned} &\bigg|\int_{\mathbb{R}^{2d}} (s-t)^{-2} \big|\phi_i(s, y_1, y_2) - \phi_i\big(s, y_1, \theta^2_{t,s}(\xi)\big)\big| \bar{q}(t, x_1, x_2; s, y_1, y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2\bigg| \\ &\quad \le C \int_{\mathbb{R}^{2d}} \bigg\{(s-t)^{-2} \big|y_2 - \theta^2_{t,s}(\xi)\big|^{\beta_i^2} \\ &\qquad \times \frac{c}{(s-t)^{2d}} \exp\bigg(-c\bigg(\frac{|y_1 - x_1|^2}{s-t} + \frac{|y_2 - m^{2,t,\xi}_{t,s}(x)|^2}{(s-t)^3}\bigg)\bigg)\bigg\}\, \mathrm{d}y_1\, \mathrm{d}y_2. \end{aligned} \tag{2.13}$$

Now, we use the Gaussian off-diagonal decay of $\bar{q}$ to smooth the time-singularity: by letting $\xi = x$ (and then $\theta^2_{t,s}(x) = m^{2,t,x}_{t,s}(x)$), for all positive η, there exists a positive constant $\bar{C}$ such that[3]

$$\left(\frac{|y_2 - m^{2,t,x}_{t,s}(x)|}{(s-t)^{3/2}}\right)^{\beta_i^2} \times \exp\left(-\eta\left(\frac{|y_1 - x_1|^2}{s-t} + \frac{|y_2 - m^{2,t,x}_{t,s}(x)|^2}{(s-t)^3}\right)\right) \le \bar{C},$$

where $\bar{C}$ depends on η and β_i^2 only. Thus, by damaging the constant c in the exponential in (2.13), we obtain the following estimate:

$$\begin{aligned}&\left|\int_{\mathbb{R}^{2d}} (s-t)^{-2}\left|\phi_i(s, y_1, y_2) - \phi_i\left(s, y_1, \theta^2_{t,s}(x)\right)\right|\bar{q}(t, x_1, x_2; s, y_1, y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2\right| \\ &\quad \le C' \int_{\mathbb{R}^{2d}} (s-t)^{-2+3\beta_i^2/2} \frac{c}{(s-t)^{2d}} \exp\left(-c'\left(\frac{|y_1 - x_1|^2}{s-t} + \frac{|y_2 - m^{2,t,x}_{t,s}(x)|^2}{(s-t)^3}\right)\right)\mathrm{d}y_1\,\mathrm{d}y_2. \end{aligned} \tag{2.14}$$

Therefore, by choosing the value of β_i^2 strictly greater than $2/3$, the singularity $(s-t)^{-2+3\beta_i^2/2}$ becomes integrable. By applying the same procedure (without centering) with the term

$$\int_{\mathbb{R}^{2d}} \left\{\left[\bar{F}_2(y_2) - \bar{F}_2\left(\theta^2_{t,s}(\xi)\right)\right] \cdot D_{y_2} u_i(s, y)\right\}\bar{q}(t, x; s, y)\,\mathrm{d}y,$$

we can deduce that

$$\|D_1 D_2 u_i\|_\infty \le C''\left(T^{-1+3\beta_i^2/2} + T^{-1+3\beta_2^2/2}\|D_2 u_i\|_\infty\right). \tag{2.15}$$

The main problem here is that the supremum norm of $D_2 u_i$ appears in the bound, so that we also have to estimate this quantity in term on known parameters in (H) to overcome the problem. It is well seen that the same arguments lead to

$$\|D_2 u_i\|_\infty \le C'''\left(T^{(-1+3\beta_i^2)/2} + T^{(-1+3\beta_2^2)/2}\|D_2 u_i\|_\infty\right).$$

By choosing T sufficiently small (e.g. such that $C''' T^{(-1+3\beta_2^2)/2}$ is less than $1/2$) we obtain that

$$\|D_2 u_i\|_\infty \le 2C''' T^{(-1+3\beta_i^2)/2}. \tag{2.16}$$

We refer to this argument as the *circular argument* in the following. By plugging this bound in (2.15) and by applying the same strategy with $D_1 u_i$ and $D_1^2 u_i$, Proposition 1.3 under (HL) follows for T less than $\mathcal{T} := (1/(2C'''))^{2/(-1+3\beta_2^2)}$.

From this discussion, one can also see the specific choice of the freezing curve as the one that matches the off-diagonal decay of the exponential in $\bar{q}$ when $\xi = x$.

3. Mathematical tools

In this section, we introduce the ingredients for the proof of Proposition 1.3.

3.1. *The frozen system*

Consider the frozen system:

$$\begin{aligned}\mathrm{d}\tilde{X}^{1,t,x}_s &= F_1\left(s, \theta_{\tau,s}(\xi)\right)\mathrm{d}s + \sigma\left(s, \theta_{\tau,s}(\xi)\right)\mathrm{d}W_s, \\ \mathrm{d}\tilde{X}^{2,t,x}_s &= \left[F_2\left(s, \theta_{\tau,s}(\xi)\right) + D_1 F_2\left(s, \theta_{\tau,s}(\xi)\right)\left(\tilde{X}^{1,t,x}_s - \theta^1_{\tau,s}(\xi)\right)\right]\mathrm{d}s\end{aligned} \tag{3.1}$$

[3] By using the inequality: $\forall \eta > 0, \forall q > 0, \exists \bar{C} > 0$ s.t. $\forall \sigma > 0, \sigma^q e^{-\eta\sigma} \le \bar{C}$.

for all s in $(t,T]$, any t in $[0,T]$, and for any initial condition x in $\mathbb{R}^{2d}$ at time t and any $(\tau,\xi)\in[0,T]\times\mathbb{R}^{2d}$, linearized around the transport $(\theta_{\tau,s}(\xi))_{\tau\le s\le T}$ defined[4] by

$$\frac{\mathrm{d}}{\mathrm{d}s}\theta_{\tau,s}(\xi)=F\big(s,\theta_{\tau,s}(\xi)\big),\qquad \theta_{\tau,\tau}(\xi)=\xi. \tag{3.2}$$

The following Proposition holds:

Proposition 3.1. *Suppose that assumptions (HR) hold, then*:

(i) *There exists a unique (strong) solution of (3.1) with mean*

$$\big(m^{\tau,\xi}_{t,s}\big)_{t\le s\le T}=\big(m^{1,\tau,\xi}_{t,s},m^{2,\tau,\xi}_{t,s}\big)_{t\le s\le T},$$

where

$$\begin{aligned}
m^{1,\tau,\xi}_{t,s}(x)&=x_1+\int_t^s F_1\big(r,\theta_{\tau,r}(\xi)\big)\,\mathrm{d}r,\\
m^{2,\tau,\xi}_{t,s}(x)&=x_2+\int_t^s\Big[F_2\big(r,\theta_{\tau,r}(\xi)\big)+D_1F_2\big(r,\theta_{\tau,r}(\xi)\big)\big(x_1-\theta^1_{\tau,r}(\xi)\big)\\
&\qquad+D_1F_2\big(r,\theta_{\tau,r}(\xi)\big)\int_t^r F_1\big(v,\theta_{\tau,v}(\xi)\big)\,\mathrm{d}v\Big]\,\mathrm{d}r,
\end{aligned} \tag{3.3}$$

and uniformly non-degenerate covariance matrix $(\tilde\Sigma_{t,s})_{t\le s\le T}$:

$$\tilde\Sigma_{t,s}=\begin{pmatrix}\int_t^s\sigma\sigma^*(r,\theta_{\tau,r}(\xi))\,\mathrm{d}r & \int_t^s R_{r,s}(\tau,\xi)\sigma\sigma^*(r,\theta_{\tau,r}(\xi))\,\mathrm{d}r\\ \int_t^s\sigma\sigma^*(r,\theta_{\tau,r}(\xi))R^*_{r,s}(\tau,\xi)\,\mathrm{d}r & \int_t^s R_{t,r}(\tau,\xi)\sigma\sigma^*(r,\theta_{\tau,r}(\xi))R^*_{t,r}(\tau,\xi)\,\mathrm{d}r\end{pmatrix}, \tag{3.4}$$

where:

$$R_{t,r}(\tau,\xi)=\Big[\int_t^r D_1F_2\big(v,\theta_{\tau,v}(\xi)\big)\,\mathrm{d}v\Big],\quad t\le r\le s\le T.$$

(ii) *This solution is a Gaussian process with transition density*:

$$\begin{aligned}
&\tilde q(t,x_1,x_2;s,y_1,y_2)\\
&\quad=\frac{3^{d/2}}{(2\pi)^{d/2}}\big(\det[\tilde\Sigma_{t,s}]\big)^{-1/2}\exp\Big(-\frac12\big|\tilde\Sigma^{-1/2}_{t,s}\big(y_1-m^{1,\tau,\xi}_{t,s}(x),y_2-m^{2,\tau,\xi}_{t,s}(x)\big)^*\big|^2\Big),
\end{aligned} \tag{3.5}$$

for all s in $(t,T]$.

(iii) *This transition density $\tilde q$ is the fundamental solution of the PDE driven by $\tilde{\mathcal{L}}^{\tau,\xi}$ and given by*:

$$\begin{aligned}
\tilde{\mathcal{L}}^{\tau,\xi}:=\frac12\operatorname{Tr}\big[a\big(t,\theta_{\tau,t}(\xi)\big)D^2_{x_1}\big]&+\big[F_1\big(t,\theta_{\tau,t}(\xi)\big)\big]\cdot D_{x_1}\\
&+\big[F_2\big(t,\theta_{\tau,t}(\xi)\big)+D_1F_2\big(t,\theta_{\tau,t}(\xi)\big)\big(x_1-\theta^1_{\tau,t}(\xi)\big)\big]\cdot D_{x_2}.
\end{aligned} \tag{3.6}$$

(iv) *There exist two positive constants c and C, depending only on known parameters in (H), such that*

$$\tilde q(t,x_1,x_2;s,y_1,y_2)\le C\hat q_c(t,x_1,x_2;s,y_1,y_2), \tag{3.7}$$

[4]Note again that if $\tau>s$, we can extend the definition of θ and suppose that $\forall(v>r)\in[0,T]^2,\forall\xi\in\mathbb{R}^{2d},\theta_{v,r}(\xi)=0$.

where

$$\hat{q}_c(t,x_1,x_2;s,y_1,y_2)=\frac{c}{(s-t)^{2d}}\exp\biggl(-c\biggl(\frac{|y_1-m_{t,s}^{1,\tau,\xi}(x)|^2}{s-t}+\frac{|y_2-m_{t,s}^{2,\tau,\xi}(x)|^2}{(s-t)^3}\biggr)\biggr),$$

and

$$\bigl|D_{x_1}^{N^{x_1}}D_{x_2}^{N^{x_2}}D_{y_1}^{N^{y_1}}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\bigr|\le C(s-t)^{-[3N^{x_2}+N^{x_1}+N^{y_1}]/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2), \tag{3.8}$$

for all s in $(t,T]$ and any integers N^{x_1},N^{x_2},N^{y_1} less than 2.

Proof. (i) First of all, note that, under (HR), the ODE: $[\mathrm{d}/\mathrm{d}s]\theta_{\tau,s}(\xi)=F(s,\theta_{\tau,s}(\xi))$, $\theta_{\tau,\tau}(\xi)=\xi$ admits a unique solution and that (3.1) admits a unique strong solution $\tilde{X}$. By rewriting (3.1) in integral form and by plugging the obtained representation of $\tilde{X}^1$ in $\tilde{X}^2$, it is easily seen that the expressions of the mean (3.3) and the variance (3.4) follow from the stochastic Fubini Theorem and standard computations. The uniform non-degeneracy of $(\tilde{\Sigma}_{t,s})_{t<s\le T}$ arises from assumptions (H) and Proposition 3.1 in [3].

(ii)–(iii) These assertions result from standard computations.

(iv) For all s in $(t,T]$, we know from Proposition 3.1 in [3] that the matrix $\tilde{\Sigma}_{t,s}$ is symmetric and uniformly non-degenerate. Besides, from Proposition 3.4 in [3] there exists a constant C depending only on known parameters in (H) such that for all s in $(t,T]$, for all (x,y,ξ) in $\mathbb{R}^{2d}\times\mathbb{R}^{2d}\times\mathbb{R}^{2d}$,

$$\begin{aligned}&-\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigl(y_1-m_{t,s}^{1,\tau,\xi}(x),y_2-m_{t,s}^{2,\tau,\xi}(x)\bigr)^*\bigr]\cdot\bigl[\bigl(y_1-m_{t,s}^{1,\tau,\xi}(x),y_2-m_{t,s}^{2,\tau,\xi}(x)\bigr)^*\bigr]\\&\quad\le -C\biggl[\biggl(\frac{y_1-m_{t,s}^{1,\tau,\xi}(x)}{(s-t)^{1/2}},\frac{y_2-m_{t,s}^{2,\tau,\xi}(x)}{(s-t)^{3/2}}\biggr)^*\biggr]\cdot\biggl[\biggl(\frac{y_1-m_{t,s}^{1,\tau,\xi}(x)}{(s-t)^{1/2}},\frac{y_2-m_{t,s}^{2,\tau,\xi}(x)}{(s-t)^{3/2}}\biggr)^*\biggr].\end{aligned}$$

For $i,j=1,2$, let $[\tilde{\Sigma}_{t,s}^{-1}]_{i,j}$ denote the block of size $d\times d$ of the matrix $\tilde{\Sigma}_{t,s}^{-1}$ at the $(i-1)d+1$, $(j-1)d+1$ rank. We can deduce from (3.4) that there exists a positive constant C depending only on known parameters in (H) such that (we also refer the reader to Lemma 3.6 and to the proof of Lemma 5.5 in [3] for more details), for all s in $(t,T]$, for all ζ in $\mathbb{R}^d$:

$$\begin{aligned}&\bigl|\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{1,1}\zeta\bigr|\le C(s-t)^{-1}|\zeta|,\\&\bigl|\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{1,2}\zeta\bigr|+\bigl|\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{2,1}\zeta\bigr|\le C(s-t)^{-2}|\zeta|,\\&\bigl|\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{2,2}\zeta\bigr|\le C(s-t)^{-3}|\zeta|,\end{aligned}\tag{3.9}$$

hence, $\tilde{\Sigma}_{t,\cdot}^{-1}$ has the same structure as $K_{\cdot-t}^{-1}$ in (2.4).

Now, we compute the derivatives w.r.t. each component and estimate it with the help of (3.9). Let $(t<s,x,y)$ in $[0,T]^2\times\mathbb{R}^{2d}\times\mathbb{R}^{2d}$, we have:

$$\begin{aligned}&\bigl|D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\bigr|\\&\quad=\bigl|\bigl(-2\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{2,1}\bigl(y_1-m_{t,s}^{1,\tau,\xi}(x)\bigr)-2\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{2,2}\bigl(y_2-m_{t,s}^{2,\tau,\xi}(x)\bigr)\bigr)\tilde{q}(t,x_1,x_2;s,y_1,y_2)\bigr|\\&\quad\le C(s-t)^{-3/2}\biggl(\biggl|\frac{(y_1-m_{t,s}^{2,\tau,\xi}(x))}{(s-t)^{1/2}}\biggr|+\biggl|\frac{(y_2-m_{t,s}^{2,\tau,\xi}(x))}{(s-t)^{3/2}}\biggr|\biggr)\tilde{q}(t,x_1,x_2;s,y_1,y_2)\\&\quad\le C'(s-t)^{-3/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2).\end{aligned}$$

Note that the symmetry $D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)=-D_{y_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)$ holds. Now, we have

$$\begin{aligned}&\bigl|D_{y_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\bigr|\\&\quad=\bigl|\bigl(2\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{1,1}\bigl(y_1-m_{t,s}^{1,\tau,\xi}(x)\bigr)+2\bigl[\tilde{\Sigma}_{t,s}^{-1}\bigr]_{1,2}\bigl(y_2-m_{t,s}^{2,\tau,\xi}(x)\bigr)\bigr)\tilde{q}(t,x_1,x_2;s,y_1,y_2)\bigr|\\&\quad\le C(s-t)^{-1/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2).\end{aligned}$$

Unfortunately, the transport of the initial condition of the diffusive component in the degenerate component breaks the symmetry and $D_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2) \neq -D_{y_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)$. Indeed

$$\begin{aligned} D_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2) = \big(&-2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,1}\big(y_1 - m_{t,s}^{1,\tau,\xi}(x)\big) - 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,2}\big(y_2 - m_{t,s}^{2,\tau,\xi}(x)\big) \\ &- 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,2}\big[\big(R_{t,s}(\tau,\xi)\big)\big(y_1 - m_{t,s}^{1,\tau,\xi}(x)\big)\big] \\ &- 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{2,2}\big[\big(R_{t,s}(\tau,\xi)\big)\big(y_2 - m_{t,s}^{2,\tau,\xi}(x)\big)\big]\big)\tilde{q}(t,x_1,x_2;s,y_1,y_2). \end{aligned}$$

Since the term $R_{t,s}(\xi)$ is of order $(s-t)$ (this is the transport of the initial condition from time t to s), we deduce that

$$\begin{aligned} \big|D_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\big| &\leq C(s-t)^{-1/2}\bigg\{\bigg|\frac{(y_1 - m_{t,s}^{2,\tau,\xi}(x))}{(s-t)^{1/2}}\bigg| + \bigg|\frac{(y_2 - m_{t,s}^{2,\tau,\xi}(x))}{(s-t)^{3/2}}\bigg| \\ &\quad + \bigg|\frac{(y_1 - m_{t,s}^{2,\tau,\xi}(x))}{(s-t)^{1/2}}\bigg| + \bigg|\frac{(y_2 - m_{t,s}^{2,\tau,\xi}(x))}{(s-t)^{3/2}}\bigg|\bigg\}\tilde{q}(t,x_1,x_2;s,y_1,y_2) \\ &\leq C'(s-t)^{-1/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2). \end{aligned}$$

Finally,

$$\begin{aligned} D_{x_1}^2\tilde{q}(t,x_1,x_2;s,y_1,y_2) = \big(&-2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,1}D_{x_1}m_{t,s}^{1,\tau,\xi}(x) - 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,2}D_{x_1}m_{t,s}^{2,\tau,\xi}(x) \\ &- 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,2}\big[\big(R_{t,s}(\tau,\xi)\big)D_{x_1}m_{t,s}^{1,\tau,\xi}(x)\big] \\ &- 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{2,2}\big[\big(R_{t,s}(\tau,\xi)\big)D_{x_1}m_{t,s}^{2,\tau,\xi}(x)\big]\big)\tilde{q}(t,x_1,x_2;s,y_1,y_2) \\ &+ \big(-2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,1}\big(y_1 - m_{t,s}^{1,\tau,\xi}(x)\big) - 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,2}\big(y_2 - m_{t,s}^{2,\tau,\xi}(x)\big) \\ &- 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{1,2}\big[\big(R_{t,s}(\tau,\xi)\big)\big(y_1 - m_{t,s}^{1,\tau,\xi}(x)\big)\big] \\ &- 2\big[\tilde{\Sigma}_{t,s}^{-1}\big]_{2,2}\big[\big(R_{t,s}(\tau,\xi)\big)\big(y_2 - m_{t,s}^{2,\tau,\xi}(x)\big)\big]\big)^2\tilde{q}(t,x_1,x_2;s,y_1,y_2). \end{aligned}$$

Note that, from (3.3) we have $D_{x_1}m_{t,s}^{\tau,\xi}(x) = (\mathrm{Id}, R_{t,s}(\tau,\xi))^*$, so that,

$$\big|D_{x_1}^2\tilde{q}(t,x_1,x_2;s,y_1,y_2)\big| \leq C(s-t)^{-1}\hat{q}_c(t,x_1,x_2;s,y_1,y_2).$$

The other derivatives can be deduced from these computations and estimate (3.8) follows. □

Remark 2. *From this proof, one can deduce that the symmetry $D_{x_2}\tilde{q} = -D_{y_2}\tilde{q}$ holds. Therefore, by an integration by parts argument, for all t in $[0,T]$, all s in $[t,T]$ and y_1, x_1, x_2 in $\mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\,\mathrm{d}y_2 = 0. \tag{3.10}$$

This argument is very useful in the sequel.

3.2. *Definitions and rules of calculus*

We introduce some definitions and rules of computations that will be useful in the following section. Let us begin by the following definition:

Definition 3.2. *For all ζ in $\mathbb{R}^{2d}$ we denote by $\Delta(\zeta)$ the perturbation operator around ζ acting on any function f from $[0,T]\times\mathbb{R}^{2d}$ as follows:*

$$\forall (s,y) \in [0,T]\times\mathbb{R}^{2d}, \quad \Delta(\zeta)f(s,y) = f(s,y) - f(s,\zeta),$$

and for $j = 1, 2$, we denote by $\Delta^j(\zeta)$ the perturbation operator around ζ_j acting on any function f from $[0, T] \times \mathbb{R}^{2d}$ as follows:

$$\forall (s, y) \in [0, T] \times \mathbb{R}^{2d}, \quad \Delta^1(\zeta) f(s, y_1, y_2) = f(s, y_1, \zeta_2) - f(s, \zeta_1, \zeta_2),$$

and

$$\forall (s, y) \in [0, T] \times \mathbb{R}^{2d}, \quad \Delta^2(\zeta) f(s, y_1, y_2) = f(s, y_1, y_2) - f(s, y_1, \zeta_2).$$

Especially, the notation $\Delta^j(\zeta) y_j$ stands for $y_j - \zeta_j$.

Given this definition we can give a generic *centering argument*, as introduced in Section 2.4 in the linear and Brownian heuristic:

Claim 3.3. *Let $\tilde{q}$ be the function defined by (3.5) in Proposition 3.1 and let f and g be two continuous functions defined on $[0, T] \times \mathbb{R}^{2d}$. For all N^1, N^2 in $\mathbb{N}$, for all $t < s$ in $[0, T]^2$, x in $\mathbb{R}^{2d}$ and ζ in $\mathbb{R}^{2d}$ we have that*:

(a)

$$D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} f(s, y) \tilde{q}(t, x; s, y) \, \mathrm{d}y = D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \Delta(\zeta) f(s, y) \tilde{q}(t, x; s, y) \, \mathrm{d}y, \quad \text{if } N_1 + N_2 > 0,$$

(b)

$$D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} f(s, y) \tilde{q}(t, x; s, y) \, \mathrm{d}y = D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \Delta^2(\zeta) f(s, y) \tilde{q}(t, x; s, y) \, \mathrm{d}y, \quad \text{if } N_2 > 0,$$

(c)

$$\begin{aligned} &D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \Delta^1(\zeta) f(s, y) g(s, y) \tilde{q}(t, x; s, y) \, \mathrm{d}y \\ &\quad = D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \big(\Delta^1(\zeta) f(s, y)\big)\big(\Delta^2(\zeta) g(s, y)\big) \tilde{q}(t, x; s, y) \, \mathrm{d}y, \quad \text{if } N_2 > 0. \end{aligned}$$

Proof. Let f and g be defined as in Claim 3.3 and let $t < s$ in $[0, T]^2$, x in $\mathbb{R}^{2d}$. We have, by Definition 3.2:

$$\begin{aligned} D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} f(s, y) \tilde{q}(t, x; s, y) \, \mathrm{d}y &= D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \Delta(\zeta) f(s, y) \tilde{q}(t, x; s, y) \, \mathrm{d}y \\ &\quad + D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} f(s, \zeta) \tilde{q}(t, x; s, y) \, \mathrm{d}y, \end{aligned}$$

for all ζ in $\mathbb{R}^{2d}$. The last term in the right hand side is equal to 0 since it does not depend on x after integrating. This concludes the proof of (a). Now, we prove (c). For all ζ in $\mathbb{R}^{2d}$ we have

$$\begin{aligned} &D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \big(\Delta^1(\zeta) f\big)(s, y_1) g(s, y_1, y_2) \tilde{q}(t, x_1, x_2; s, y_1, y_2) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \\ &\quad = D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \big(\Delta^1(\zeta) f\big)(s, y_1) \big(\Delta^2(\zeta) g\big)(s, y_1, y_2) \tilde{q}(t, x_1, x_2; s, y_1, y_2) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \\ &\qquad + D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} \big(\Delta^1(\zeta) f\big)(s, y_1) g(s, y_1, \zeta_2) \tilde{q}(t, x_1, x_2; s, y_1, y_2) \, \mathrm{d}y_1 \, \mathrm{d}y_2. \end{aligned}$$

By using differentiation under the integral sign Theorem, (3.10) (since N^2 is positive) together with an integration by parts, the last term in the right hand side is equal to 0. Finally, assertion (b) follows from the same arguments. This concludes the proof of the Claim. □

3.3. *Representation and differentiation of the solution of the PDE* (1.9)

Lemma 3.4. *Suppose that assumptions* (HR) *hold, then, for all x in $\mathbb{R}^{2d}$ and t in $[0,T]$ the solution $u=(u_1,u_2)^*$ of the PDE* (1.9) *can be written as*

$$
\begin{aligned}
u_i(t,x) &= E\Big[\int_t^T \phi_i\big(s,\tilde{X}_s^{t,x}\big) - \big(\mathcal{L}-\tilde{\mathcal{L}}^{t,\xi}\big)u_i\big(s,\tilde{X}_s^{t,x}\big)\,\mathrm{d}s\Big] \qquad (3.11)\\
&= \int_t^T\int_{\mathbb{R}^{2d}} \phi_i(s,y)\tilde{q}(t,x;s,y)\,\mathrm{d}y\,\mathrm{d}s\\
&\quad - \int_t^T\int_{\mathbb{R}^{2d}} \frac{1}{2}\operatorname{Tr}\big[\Delta\big(\theta_{t,s}(\xi)\big)a(s,y)D^2_{y_1}u_i(s,y)\big]\tilde{q}(t,x;s,y)\,\mathrm{d}y\,\mathrm{d}s\\
&\quad - \int_t^T\int_{\mathbb{R}^{2d}} \big\{\big[\Delta\big(\theta_{t,s}(\xi)\big)F_1(s,y)\big]\cdot D_{y_1}u_i(s,y)\big\}\tilde{q}(t,x;s,y)\,\mathrm{d}y\,\mathrm{d}s\\
&\quad - \int_t^T\int_{\mathbb{R}^{2d}} \big\{\big[\Delta\big(\theta_{t,s}(\xi)\big)F_2(s,y) - D_1F_2\big(s,\theta_{t,s}(\xi)\big)\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\big]\cdot D_{y_2}u_i(s,y)\big\}\tilde{q}(t,x;s,y)\,\mathrm{d}y\,\mathrm{d}s\\
&=: \sum_{j=1}^4\int_t^T\int_{\mathbb{R}^{2d}} \mathcal{H}_i^j\big(s,y,\theta_{t,s}(\xi)\big)\tilde{q}(t,x;s,y)\,\mathrm{d}y\,\mathrm{d}s \qquad (3.12)\\
&=: \sum_{j=1}^4\int_t^T \mathcal{I}_i^j(s,x)\,\mathrm{d}s \qquad (3.13)
\end{aligned}
$$

for all ξ in $\mathbb{R}^{2d}$. It is infinitely differentiable on $[0,T]\times\mathbb{R}^{2d}$.

Remark 3. *We emphasize that the solution "u" does not depend on the choice of the freezing data "(τ,ξ)". Above we chose to set $\tau=t$ when the solution u is evaluated at time t. We keep this choice from now on.*

Proof of Lemma 3.4. Thanks to Lemma 1.2 the PDE (1.9) is well posed and can be rewritten as

$$
\begin{cases}
\partial_t u_i(t,x) + \tilde{\mathcal{L}}^{\tau,\xi}u_i(t,x) = -(\mathcal{L}-\tilde{\mathcal{L}}^{\tau,\xi})u_i(t,x) + \phi_i(t,x), & (t,x)\in[0,T)\times\mathbb{R}^{2d},\\
u_i(T,x)=0, & i=1,2,
\end{cases}
\qquad (3.14)
$$

for all (τ,ξ) in $[0,T]\times\mathbb{R}^{2d}$, so that (3.11) follows from Feynman–Kac representation. Equality (3.13) follows from the definition of $\tilde{q}$ in Proposition 3.1. Next, given a positive ε, we have for all (t,x) in $[0,T]\times\mathbb{R}^{2d}$,

$$
\begin{aligned}
u_i(t,x) &= \int_{t+\varepsilon}^T \mathbb{E}\big[\phi_i\big(s,\tilde{X}_s^{t,x}\big) - \big(\mathcal{L}-\tilde{\mathcal{L}}^{\tau,\xi}\big)u_i\big(s,\tilde{X}_s^{t,x}\big)\big]\,\mathrm{d}s\\
&\quad + \int_t^{t+\varepsilon} \mathbb{E}\big[\phi_i\big(s,\tilde{X}_s^{t,x}\big) - \big(\mathcal{L}-\tilde{\mathcal{L}}^{\tau,\xi}\big)u_i\big(s,\tilde{X}_s^{t,x}\big)\big]\,\mathrm{d}s.
\end{aligned}
$$

Under (HR), the coefficients of $\mathcal{L}$, $\tilde{\mathcal{L}}^{\tau,\xi}$ and the functions ϕ_i, u_i are smooth (Lemma 1.2). From classical regularity results on the solution of the SDE (2.6) (see e.g. [13]) we can deduce that the solution is infinitely differentiable under (HR). □

We now derive a representation formula for the derivatives of $\mathcal{I}_i^j$, $i=1,2$, $j=1,\dots,4$, that involve a differentiation w.r.t. the degenerate variable. This allows us to handle the singularity of the derivative of the kernel $\tilde{q}$ (see assertion (iv) in Proposition 3.1) as done in Section 2.4 (*centering argument*).

Lemma 3.5. *Let $t < T$ in $\mathbb{R}^+$. For all i in $\{1,2\}$, all (s,x) in $(t,T]\times\mathbb{R}^{2d}$ and all integer n the terms $D^n_{x_1}D_{x_2}\mathcal{I}^j_i(s,x)$, $j=1,\dots,4$, can be written as*:

$$D^n_{x_1}D_{x_2}\mathcal{I}^1_i(s,x) = -\int_{\mathbb{R}^{2d}} \Delta^2\big(\theta_{t,s}(\xi)\big)\phi_i(s,y)D^n_{x_1}D_{x_2}\tilde{q}(t,x;s,y)\,\mathrm{d}y \tag{3.15}$$

and

$$\begin{aligned}
&D^n_{x_1}D_{x_2}\mathcal{I}^2_i(s,x)\\
&\quad= -\frac{1}{2}\int_{\mathbb{R}^{2d}}\Bigg\{\mathrm{Tr}\big[\big[\Delta^2\big(\theta_{t,s}(\xi)\big)a(s,y)\big]D^2_{y_1}u_i(s,y)\big]\\
&\qquad-\sum_{l=1}^d\Big[\frac{\partial}{\partial y_{1l}}a_{l.}(s,y)\Big]\cdot\big[\Delta^2\big(\theta_{t,s}(\xi)\big)D_{y_1}u_i(s,y)\big]\Bigg\}D^n_{x_1}D_{x_2}\tilde{q}(t,x;s,y)\,\mathrm{d}y\\
&\qquad+\frac{1}{2}\sum_{l=1}^d\int_{\mathbb{R}^{2d}}\big\{\big[\Delta^1\big(\theta_{t,s}(\xi)\big)a_{l.}(s,y)\big]\cdot\big[\Delta^2\big(\theta_{t,s}(\xi)\big)D_{y_1}u_i(s,y)\big]\big\}D^n_{x_1}D_{x_2}\Big(\frac{\partial}{\partial y_{1l}}\tilde{q}(t,x;s,y)\Big)\,\mathrm{d}y,
\end{aligned}\tag{3.16}$$

where "$a_{l.}$" denotes the lth line of the matrix a, and

$$\begin{aligned}
D^n_{x_1}D_{x_2}\mathcal{I}^3_i(s,x) = -\int_{\mathbb{R}^{2d}}&\big\{\big[\Delta^2\big(\theta_{t,s}(\xi)\big)F_1(s,y)\big]\cdot D_{y_1}u_i(s,y)\\
&+\big[\Delta^1\big(\theta_{t,s}(\xi)\big)F_1(s,y)\big]\cdot\big[\Delta^2\big(\theta_{t,s}(\xi)\big)D_{y_1}u_i(s,y)\big]\big\}D^n_{x_1}D_{x_2}\tilde{q}(t,x;s,y)\,\mathrm{d}y,
\end{aligned}\tag{3.17}$$

and finally:

$$\begin{aligned}
D^n_{x_1}D_{x_2}\mathcal{I}^4_i(s,x) = -\int_{\mathbb{R}^{2d}}&\big\{\big[\Delta^1\big(\theta_{t,s}(\xi)\big)F_2(s,y) - D_1F_2\big(s,\theta_{t,s}(\xi)\big)\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\big]\cdot\big[\Delta^2\big(\theta_{t,s}(\xi)\big)D_{y_2}u_i(s,y)\big]\\
&+\big[\Delta^2\big(\theta_{t,s}(\xi)\big)F_2(s,y)\big]\cdot D_{y_2}u_i(s,y)\big\}D^n_{x_1}D_{x_2}\tilde{q}(t,x;s,y)\,\mathrm{d}y.
\end{aligned}\tag{3.18}$$

Proof. Representation (3.15) is a direct consequence of assertion (b) in Claim 3.3. Next, we deal with (3.16). By using first the decomposition $\Delta = \Delta^1 + \Delta^2$ and then by integrating by parts, we have

$$\begin{aligned}
\mathcal{I}^2_i(s,x) = -\frac{1}{2}\int_{\mathbb{R}^{2d}}&\Bigg\{\mathrm{Tr}\big[\big[\Delta^2\big(\theta_{t,s}(\xi)\big)a(s,y)\big]D^2_{y_1}u_i(s,y)\big]\\
&-\sum_{l=1}^d\Big[\frac{\partial}{\partial y_{1l}}\Delta^1\big(\theta_{t,s}(\xi)\big)a_{l.}(s,y)\Big]\cdot D_{y_1}u_i(s,y)\Bigg\}\tilde{q}(t,x;s,y)\,\mathrm{d}y\\
&+\frac{1}{2}\sum_{l=1}^d\int_{\mathbb{R}^{2d}}\big\{\big[\Delta^1\big(\theta_{t,s}(\xi)\big)a_{l.}(s,y)\big]\cdot D_{y_1}u_i(s,y)\big\}\frac{\partial}{\partial y_{1l}}\tilde{q}(t,x;s,y)\,\mathrm{d}y.
\end{aligned}$$

Note that, for all l in $\{1,\dots,d\}$, $[\partial/\partial y_{1l}]\Delta^1(\theta_{t,s}(\xi))a_{l.}(s,y_1,\theta^2_{t,s}(\xi)) = [\partial/\partial y_{1l}]a_{l.}(s,y_1,\theta^2_{t,s}(\xi))$. We conclude by differentiating and then by applying assertion (c) in Claim 3.3 with $f=a$ and $g=D_1u_i$. This gives (3.16).

By using again first the decomposition $\Delta = \Delta^1 + \Delta^2$, assertions (3.17) and (3.18) are immediate consequences of assertion (c) in Claim 3.3. This concludes the proof of Lemma 3.5. □

We conclude this section with the following definition:

Definition 3.6 (The centered integrands). *When differentiating at least once the terms $\mathcal{I}^j_i$, $j=1,\dots,4$, w.r.t. the degenerate variable "x_2", Lemma 3.5 allows us to identify the $\mathcal{H}^j_i$, $j=1,3,4$, defined in (3.12) with the integrand of*

(3.15), (3.17), *and* (3.18) *respectively. In the same spirit we use in that case the notation* $\mathcal{H}_i^2$ *to denote both the sum of the two integrands appearing in the representation* (3.16) *of* $D_{x_1}^n D_{x_2}\mathcal{I}_i^2$ *and the integrand defined in* (3.12).

When we identify the $\mathcal{H}_i^j$, $j = 1, \dots, 4$, *with the integrand of* (3.15), (3.16), (3.17) *and* (3.18) *respectively, we will call them the* centered integrands.

4. Proof of Proposition 1.3

4.1. *From intermediate gradient estimates to Proposition* 1.3

We here give the intermediate estimates that allow to prove Proposition 1.3. In the following, u_i denotes the ith component of the solution $u = (u_1, u_2)^*$ of the linear system of PDE (1.9). The following arguments and Lemmas hold for $i = 1, 2$.

Since the representation (3.13) of each u_i involves its derivatives, we prove Proposition 1.3 by using a *circular argument* (see (2.16)). We recall that for any function from $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ we denote by D_1 (resp. D_2) the derivative with respect to the first (resp. second) d-dimensional space component. We first show that

Lemma 4.1. *Suppose that assumptions* (HR) *hold. There exist a positive* $\mathcal{T}_{4.1}$, *two positive numbers* $\delta_{4.1}$ *and* $\bar{\delta}_{4.1}$ *and a positive constant* C, *depending only on known parameters in* (H), *such that*

(i) $\|D_1^2 u_i\|_\infty \le T^{\delta_{4.1}} C(1 + \|D_2 u_i\|_\infty)$,

(ii) $\|D_1 u_i\|_\infty \le T^{\bar{\delta}_{4.1}} C(1 + \|D_2 u_i\|_\infty)$,

for all T *less than* $\mathcal{T}_{4.1}$. *We recall that* D_1 *denotes the derivative with respect to the first* d*-dimensional space component.*

Then, we estimate the gradients that involve the derivatives w.r.t. the degenerate variable "D_2". To this aim, we differentiate the representation (3.13) and we estimate it. Thanks to Proposition 3.1, we know that this differentiation generates the worst singularity in the time space integrals. As shown in Section 2.4, we can use the regularity of the coefficients assumed in (H) in order to smooth this singularity. But in this more general case, we also have to use the regularity of the solution itself. Notably, we need to estimate the Hölder regularity of $D_2 u_i$. Hence, we prove the following sort of Hölder estimate on $D_2 u_i$:

Lemma 4.2. *Suppose that assumptions* (HR) *hold and let*

$$M(D_2 u_i, T) := \sup_{w_1, w_2 \neq w_2' \in \mathbb{R}^d, t \in [0,T]} \frac{|D_2 u_i(t, w_1, w_2) - D_2 u_i(t, w_1, w_2')|}{|w_2 - w_2'|^{\gamma/3} + |w_2 - w_2'|^{\beta_2^2} + |w_2 - w_2'|^{\beta_1^2} + |w_2 - w_2'|}, \tag{4.1}$$

for some positive number γ. *There exist a positive* $\mathcal{T}_{4.2}$, *a positive constant* C *and a positive number* $\delta_{4.2}$, *depending only on known parameters in* (H), *such that for all positive* γ *strictly less than* $3\inf\{\beta_1^2, \beta_2^2\} - 1$,

$$M(D_2 u_i, T) \le C T^{\delta_{4.2}} \big(1 + \|D_2 u_i\|_\infty + \|D_1 D_2 u_i\|_\infty\big),$$

for all T *less than* $\mathcal{T}_{4.2}$.

This allows us to obtain the following estimates

Lemma 4.3. *Suppose that assumptions* (HR) *hold and let* n *in* $\{0, 1\}$. *There exist a positive* $\mathcal{T}_{4.3}(n)$, *a positive number* $\delta_{4.3}(n)$ *and a positive constant* $C(n)$, *depending only on known parameters in* (H) *and* n, *such that*:

$$\big\|D_1^n D_2 u_i\big\|_\infty \le C(n) T^{\delta_{4.3}(n)},$$

for all T *less than* $\mathcal{T}_{4.3}(n)$.

Finally, we can deduce Proposition 1.3 from Lemmas 4.1 and 4.3. □

4.2. *Intermediate gradient estimates*: *Proofs*

4.2.1. *Proof of Lemma* 4.1

We first show assertion (i). Let (t,x) belong to $[0,T]\times\mathbb{R}^{2d}$, from the representation (3.13) in Lemma 3.4 we have:

$$D^2_{x_1}u_i(t,x)=\sum_{j=1}^{4}D^2_{x_1}\int_t^T\mathcal{I}_i^j(s,x)\,\mathrm{d}s. \tag{4.2}$$

As done in the Section 2.4 we have to obtain a suitable bound on the $D_1^2\mathcal{I}_i^j$ in order to estimate $D_1^2u_i$. Thanks to assertion (a) in Claim 3.3 applied on the integrand of $D^2_{x_1}\mathcal{I}_i^1$ we have for all s in $[t,T]$:

$$\begin{aligned}\sum_{j=1}^{4}D^2_{x_1}\mathcal{I}_i^j(s,x)=&\int_{\mathbb{R}^{2d}}\Delta\big(\theta_{t,s}(\xi)\big)\phi_i(s,y)\big[D^2_{x_1}\tilde{q}(t,x;s,y)\big]\,\mathrm{d}y\\&-\frac{1}{2}\int_{\mathbb{R}^{2d}}\mathrm{Tr}\big[\Delta\big(\theta_{t,s}(\xi)\big)a(s,y)D^2_{y_1}u_i(s,y)\big]\big[D^2_{x_1}\tilde{q}(t,x;s,y)\big]\,\mathrm{d}y\\&-\int_{\mathbb{R}^{2d}}\big\{\big[\Delta\big(\theta_{t,s}(\xi)\big)F_1(s,y)\big]\cdot D_{y_1}u_i(s,y)\big\}\big[D^2_{x_1}\tilde{q}(t,x;s,y)\big]\,\mathrm{d}y\\&-\int_{\mathbb{R}^{2d}}\big\{\big[\Delta\big(\theta_{t,s}(\xi)\big)F_2(s,y)-D_1F_2\big(s,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi)\big)\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\big]\cdot\big[D_{y_2}u_i(s,y)\big]\big\}\\&\times\big[D^2_{x_1}\tilde{q}(t,x;s,y)\big]\,\mathrm{d}y.\end{aligned} \tag{4.3}$$

Note that by using the fact that $\Delta=\Delta^1+\Delta^2$, a Taylor expansion of order 0 with integrable remainder of the mapping $\theta^1_{t,s}(\xi)\mapsto F_2(\cdot,\theta^1_{t,s}(\xi),\cdot)$ around y_1 in $\mathbb{R}^d$ and (H3-a), we have that:

$$\begin{aligned}\forall(s,y_2)\in(t,T]\times\mathbb{R}^d,\quad&\big|\big[\Delta\big(\theta_{t,s}(\xi)\big)F_2(s,y)-D_1F_2\big(s,\theta_{t,s}(\xi)\big)\big(\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\big)\big]\cdot D_{y_2}u_i(s,y)\big|\\&\le C\|D_2u_i\|_\infty\big(\big|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\big|^{\beta_2^2}+\big|\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\big|^{1+\alpha^1}\big).\end{aligned} \tag{4.4}$$

From Proposition 3.1, we know that for all s in $(t,T]$ and y in $\mathbb{R}^{2d}$, $|D^2_{x_1}\tilde{q}(t,x;s,y)|\le C'(s-t)^{-1}\hat{q}_c(t,x;s,y)$. By plugging this estimate in (4.3), together with the regularity of the coefficients assumed in (H) and (4.4), we obtain that

$$\begin{aligned}\left|\sum_{j=1}^{4}D^2_{x_1}\mathcal{I}_i^j(s,x)\right|\le C''(s-t)^{-1}\int_{\mathbb{R}^{2d}}\Bigg\{&\sum_{j=1}^{2}\Bigg\{(s-t)^{(j-1/2)\beta_i^j}\left|\frac{\Delta^j(\theta_{t,s}(\xi))y_j}{(s-t)^{(j-1/2)}}\right|^{\beta_i^j}\\&+\|D_1^2u_i\|_\infty(s-t)^{(j-1/2)}\left|\frac{\Delta^j(\theta_{t,s}(\xi))y_j}{(s-t)^{(j-1/2)}}\right|\\&+\|D_1u_i\|_\infty(s-t)^{(j-1/2)\beta_1^j}\left|\frac{\Delta^j(\theta_{t,s}(\xi))y_j}{(s-t)^{(j-1/2)}}\right|^{\beta_1^j}\Bigg\}\\&+\|D_2u_i\|_\infty\left[(s-t)^{3\beta_2^2/2}\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|^{\beta_2^2}+(s-t)^{(1+\alpha^1)/2}\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|^{1+\alpha^1}\right]\Bigg\}\\&\times\hat{q}_c(t,x;s,y)\,\mathrm{d}y.\end{aligned}$$

Set $\xi=x$, we now use the Gaussian off-diagonal decay in $\hat{q}_c$ (see the computations in Section 2.4 for more details):

$$\begin{aligned}\forall\kappa>0,\exists\underline{\mathrm{C}},\underline{\mathrm{c}}>0\quad\text{s.t.}\quad\forall(s,y)\in(t,T]\times\mathbb{R}^{2d}:\quad&\big|\Delta^l\big(\theta_{t,s}(x)\big)y_l\big|^\kappa\hat{q}_c(t,x;s,y)\\&\le\underline{\mathrm{C}}(s-t)^{(l-1/2)\kappa}\hat{q}_{\underline{\mathrm{c}}}(t,x;s,y)\end{aligned} \tag{4.5}$$

for $l = 1, 2$ and where $\underline{C}$ and $\underline{c}$ depend only on known parameters in (H) and on κ. Hence, by integrating w.r.t. the space variable we have:

$$\begin{aligned}
&\left|\sum_{j=1}^{4} D_{x_1}^2 \mathcal{I}_i^j(s,x)\right| \\
&\quad \le C'''\Bigg\{\sum_{j=1}^{2}\big\{(s-t)^{-1+(j-1/2)\beta_i^j} + \big((s-t)^{-1/2}+(s-t)^{1/2}\big)\big\|D_1^2 u_i\big\|_\infty + (s-t)^{(j-1/2)\beta_2^j-1}\|D_1 u_i\|_\infty\big\} \\
&\qquad + \big((s-t)^{-1+3\beta_1^2/2} + (s-t)^{(\alpha^1-1)/2}\big)\|D_2 u_i\|_\infty\Bigg\}.
\end{aligned}$$

So that we can inverse the differentiation and integral operators in (4.2) and then deduce that

$$\begin{aligned}
\big\|D_1^2 u_i\big\|_\infty \le C'''\big\{&\big(T^{3\beta_2^2/2} + T^{(1+\alpha^1)/2}\big)\|D_2 u_i\|_\infty + \big(T^{\beta_1^1/2} + T^{3\beta_1^2/2}\big)\|D_1 u_i\|_\infty \\
&+ \big(T^{1/2} + T^{3/2}\big)\big\|D_1^2 u_i\big\|_\infty + T^{\beta_i^1/2} + T^{3\beta_i^2/2}\big\}.
\end{aligned}$$

Then, by setting $\mathcal{T}_{4.1}^{(i)} = \sup\{T > 0,\ \text{such that } C'''(T^{1/2} + T^{3/2}) \le 1/2\}$ we deduce the assertion (i) for all T less than $\mathcal{T}_{4.1}^{(i)}$ from a *circular argument* (see (2.16)). The proof of the second statement (ii) can be done by the same arguments. This concludes the proof of Lemma 4.1. □

4.2.2. *Proof of Lemma 4.3*
Let (t, x) belongs to $[0, T] \times \mathbb{R}^{2d}$. We have from representation (3.13) in Lemma 3.4 that, for all n in $\{0, 1\}$ and all positive ε

$$D_{x_1}^n D_{x_2} u_i(t,x) = \sum_{j=1}^{4} D_{x_1}^n D_{x_2} \int_t^T \mathcal{I}_i^j(s,x)\,\mathrm{d}s. \tag{4.6}$$

Again, we look for a suitable bound on the $D_{x_1}^n D_{x_2}\mathcal{I}_i^j$ in order to estimate $D_{x_1}^n D_{x_2} u_i$. Thanks to representation (3.12), we know that the derivatives $D_{x_1}^n D_{x_2}\mathcal{I}_i^j$ can be written as the integral of some function against the derivative of the degenerate Gaussian kernel $\tilde{q}$. Since from Proposition 3.1 we have that for all s in $(t, T]$ and y in $\mathbb{R}^{2d}$

$$\big|D_{x_1}^n D_{x_2}\tilde{q}(t,x;s,y)\big| \le C(s-t)^{-(3+n)/2}\hat{q}_c(t,x;s,y), \tag{4.7}$$

we use the regularity of the coefficients together with the regularity of the solution u_i and its derivatives to smooth the time singularity appearing in (4.7), as done in the previous subsection and in Section 2.4.

Hence, we first give a bound on each *centered integrand* $\mathcal{H}_i^j$ of $D_{x_1}^n D_{x_2}\mathcal{I}_i^j$, $j = 1, \dots, 4$, given by Definition 3.6. For all s in $(t, T]$, y in $\mathbb{R}^{2d}$ we have, from the regularity of the coefficients assumed in (H), that:

$$\big|\mathcal{H}_i^1\big(s,y,\theta_{t,s}(\xi)\big)\big| \le C\big|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\big|^{\beta_i^2}. \tag{4.8}$$

Then, we recall that from Mean Value Theorem (MVT) we have

$$\big|\Delta^2\big(\theta_{t,s}(\xi)\big)D_{y_1}u_i(s,y)\big| \le \|D_1 D_2 u_i\|_\infty\big|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\big|, \tag{4.9}$$

so that

$$\begin{aligned}
\big|\mathcal{H}_i^2\big(s,y,\theta_{t,s}(\xi)\big)\big| \le C'\big\{&\big\|D_1^2 u_i\big\|_\infty\big|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\big| \\
&+ \|D_1 D_2 u_i\|_\infty\big(\big|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\big| + \big|\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\big|\big|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\big|\big)\big\}.
\end{aligned} \tag{4.10}$$

By the same way, we get that

$$\begin{aligned}\left|\mathcal{H}_i^3\big(s,y,\theta_{t,s}(\xi)\big)\right| &\le C''\big\{\|D_1u_i\|_\infty\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\beta_1^2}\\ &\quad+\|D_1D_2u_i\|_\infty\left|\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\right|^{\beta_1^1}\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|\big\}.\end{aligned}\tag{4.11}$$

And finally, since we have from Lemma 4.2

$$\begin{aligned}&\left|\Delta^2\big(\theta_{t,s}(\xi)\big)D_{y_2}u_i(s,y)\right|\\ &\quad\le M(D_2u_i,T)\big(\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\gamma/3}+\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\beta_1^2}+\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\beta_2^2}+\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|\big),\end{aligned}\tag{4.12}$$

where $M(D_2u_i,T)$ is defined by (4.1) we have from the regularity of the coefficients assumed in (H) and (4.4) that

$$\begin{aligned}&\left|\mathcal{H}_i^4\big(s,y,\theta_{t,s}(\xi)\big)\right|\\ &\quad\le C'''\|D_2u_i\|_\infty\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\beta_2^2}\\ &\qquad+M(D_2u_i,T)\left|\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\right|^{1+\alpha^1}\\ &\qquad\times\big(\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\gamma/3}+\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\beta_1^2}+\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|^{\beta_2^2}+\left|\Delta^2\big(\theta_{t,s}(\xi)\big)y_2\right|\big).\end{aligned}\tag{4.13}$$

Finally, let us recall from Proposition 3.1 that for all (s,y) in $(t,T]\times\mathbb{R}^{2d}$,

$$\left|D_{x_1}^nD_{x_2}\frac{\partial}{\partial y_{1l}}\tilde{q}(t,x;s,y)\right|\le C''''(s-t)^{-(4+n)/2}\hat{q}_{c'}(t,x;s,y),\tag{4.14}$$

for all l in $\{1,\dots,d\}$.

Now, we can plug together some of the above estimates in the corresponding $D_{x_1}^nD_{x_2}\mathcal{I}_i^j$, $j=1,\dots,4$, defined in Lemma 3.5. By using (4.7) with (4.8) (resp. (4.11) and (4.13)) in (3.15) (resp. (3.17) and (3.18)), estimates (4.7) and (4.14) with (4.10) in (3.16), by letting next $\xi=x$ and by using (4.5) in all these terms, we can deduce that for all s in $(t,T]$

$$\begin{aligned}&\left|\sum_{j=1}^4D_{x_1}^nD_{x_2}\mathcal{I}_i^j(s,x)\right|\\ &\quad\le c'\int_{\mathbb{R}^{2d}}\big\{(s-t)^{(3(\beta_i^2-1)-n)/2}+(s-t)^{-n/2}\big(\left\|D_1^2u_i\right\|_\infty+\|D_1D_2u_i\|_\infty\big)\\ &\qquad+\|D_1u_i\|_\infty(s-t)^{(3(\beta_1^2-1)-n)/2}+\|D_1D_2u_i\|_\infty(s-t)^{(\beta_1^1-n)/2}\\ &\qquad+M(D_2u_i,T)(s-t)^{-1+(\alpha^1-n)/2}\big((s-t)^{\gamma/2}+(s-t)^{\beta_1^2/2}+(s-t)^{\beta_2^2/2}+(s-t)\big)\\ &\qquad+\|D_2u_i\|_\infty(s-t)^{(3(\beta_2^2-1)-n)/2}\big\}\hat{q}_{c''}(t,x;s,y)\,\mathrm{d}y.\end{aligned}\tag{4.15}$$

Then, we have from Lemma 4.2 that for all T less than $\mathcal{T}_{4.2}$:

$$\forall\gamma\quad\text{s.t.}\quad 0<\gamma<3\inf\big\{\beta_1^2,\beta_2^2\big\}-1,\qquad M(D_2u_i,T)\le c'''\mathcal{T}_{4.2}\big(1+\|D_2u_i\|_\infty+\|D_1D_2u_i\|_\infty\big).\tag{4.16}$$

So that we can inverse the differentiation and integral operators in (4.6). Hence, by plugging (4.16) in (4.15) and by choosing n equal to 0, after integrating in time, we deduce from a *circular argument* (as described in (2.16)) that there exist two positive numbers $\mathcal{T}'_{4.3}(0)$, $\delta_{4.3}(0)$ and a positive constant $\bar{C}$, depending only on known parameters in (H), such that

$$\|D_2u_i\|_\infty\le\bar{C}T^{\delta_{4.3}(0)}\big(1+\|D_1D_2u_i\|_\infty\big),\tag{4.17}$$

for all T less than $\mathcal{T}'_{4.3}(0)$. Next, by letting n be equal to 1 in (4.15) (recall that Lemma 4.2 hold for all $0 < \gamma < 3\inf\{\beta_1^2, \beta_2^2\} - 1$ and that $3\inf\{\beta_1^2, \beta_2^2\} > 1$ so that all the time-singularity are integrable), by using the same arguments as in the case $n = 0$ together with (4.17), we can show that there exist two positive numbers $\mathcal{T}_{4.3}(1)$, $\delta_{4.3}(1)$ and a positive constant $\bar{C}'$, depending only on known parameters in (H), such that

$$\|D_1 D_2 u_i\|_\infty \le \bar{C}' T^{\delta_{4.3}(1)},$$

for all T less than $\mathcal{T}_{4.3}(1)$. This conclude the proof for $n = 1$. The case $n = 0$ follows from plugging this estimate in (4.17). □

Remark 4. *Note that the estimate on supremum norm of $D_2 u_i$ (i.e. when $n = 0$) could be obtain without Lemma 4.2, by bounding the left hand side of (4.12) by the supremum norm of $D_2 u_i$.*

4.2.3. *Proof of Lemma 4.2*
From (3.13) in Lemma 3.4, Lemma 4.3 and Remark 4, for all (t, x_1) in $[0, T] \times \mathbb{R}^d$ and $x_2 \neq z_2$ in $\mathbb{R}^d$ we have:

$$\begin{aligned}
&\left|D_2 u_i(t, x_1, x_2) - D_2 u_i(t, x_1, z_2)\right| \\
&\quad \le \left|\sum_{j=1}^{4} \int_t^T D_{x_2}\mathcal{I}_i^j(s, x_1, x_2) - D_{z_2}\mathcal{I}_i^j(s, x_1, z_2)\,\mathrm{d}s\right|.
\end{aligned} \tag{4.18}$$

We recall that the $\mathcal{I}_i^j$, $j = 1, \dots, 4$, depend on the freezing point $\xi = (\xi_1, \xi_2)^*$ of the process which started from (x_1, x_2) and (x_1, z_2) at time t. Here, we choose the same freezing point "ξ" for the two processes (with different initial conditions).

Then, we split the time interval w.r.t. the characteristic scale of the second component of the system (3.1) in order to study the perturbation on each interval. Hence, we set $\mathcal{S} = \{s \in (t, T] \text{ s.t. } |x_2 - z_2| < (s - t)^{3/2}\}$ and $\mathcal{S}^c = \{s \in (t, T] \text{ s.t. } |x_2 - z_2| \ge (s - t)^{3/2}\}$. We have for all s in $(t, T]$:

$$\begin{aligned}
&\left|\sum_{j=1}^{4} D_{x_2}\mathcal{I}_i^j(s, x_1, x_2) - D_{z_2}\mathcal{I}_i^j(s, x_1, z_2)\right| \\
&\quad = \sum_{j=1}^{4} \mathbf{1}_{\mathcal{S}} \int_{\mathbb{R}^{2d}} \Big\{\mathcal{H}_i^j\big(s, y_1, y_2, \theta_{t,s}^1(\xi), \theta_{t,s}^2(\xi)\big) \\
&\qquad \times \big(D_{x_2}\tilde{q}(t, x_1, x_2; s, y_1, y_2) - D_{z_2}\tilde{q}(t, x_1, z_2; s, y_1, y_2)\big)\Big\}\,\mathrm{d}y_1\,\mathrm{d}y_2 \\
&\qquad + \sum_{j=1}^{4} \mathbf{1}_{\mathcal{S}^c} \int_{\mathbb{R}^{2d}} \Big\{\mathcal{H}_i^j\big(s, y_1, y_2, \theta_{t,s}^1(\xi), \theta_{t,s}^2(\xi)\big) \\
&\qquad \times \big(D_{x_2}\tilde{q}(t, x_1, x_2; s, y_1, y_2) - D_{z_2}\tilde{q}(t, x_1, z_2; s, y_1, y_2)\big)\Big\}\,\mathrm{d}y_1\,\mathrm{d}y_2 \\
&\quad := \mathcal{P}_s(\mathcal{S}) + \mathcal{P}_s\big(\mathcal{S}^c\big),
\end{aligned} \tag{4.19}$$

where the $\mathcal{H}_i^j$, $j = 1, \dots, 4$, are given by (3.12). We now bound the first and the second sum in the right hand side of the last equality above.

Estimation of $\mathcal{P}_s(\mathcal{S})$ on $(t, T]$. As a first step, we bound the first sum in the right hand side of (4.19). At the end of this part, it is proven that:

Claim 4.4. *For all s in $\mathcal{S}$, y_1, y_2 in $\mathbb{R}^d$, the following inequality holds*:

$$\begin{aligned}
&\left|D_{x_2}\tilde{q}(t, x_1, x_2; s, y_1, y_2) - D_{z_2}\tilde{q}(t, x_1, z_2; s, y_1, y_2)\right| \\
&\quad \le C(s - t)^{-(3+\gamma)/2}\hat{q}_c(t, x_1, x_2; s, y_1, y_2)|x_2 - z_2|^{\gamma/3},
\end{aligned} \tag{4.20}$$

where c and C depend only on known parameters in (H) and where $0<\gamma<3$. Moreover

$$\left|\frac{\partial}{\partial y_{1l}}D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)-\frac{\partial}{\partial y_{1l}}D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\right| \\ \leq C'(s-t)^{-(3+\gamma+1)/2}\hat{q}_{c'}(t,x_1,x_2;s,y_1,y_2)|x_2-z_2|^{\gamma/3}, \tag{4.21}$$

for any $l=1,\dots,d$, where c' and C' depend only on known parameters in (H) and where $0<\gamma<3$.

On the other hand, Lemma 3.5 allows to choose the $\mathcal{H}_i^j$, $j=1,\dots,4$, appearing in $\mathcal{P}_\varepsilon(\mathcal{S})$ as the *centered integrands* (see Definition 3.6). So that this term is similar to the term studied in the proof of Lemma 4.3. In fact, the only difference with the terms in the proof of Lemma 4.3 (when $n=0$) is that the integration is now against the perturbed differentiated kernel defined by the left hand side of (4.20) or (4.21):[5] in comparison with (4.7) and (4.14), we "loose" a $\gamma/2$ in the time-singularity in the estimation.

Hence, we can use the same arguments as the ones of Lemma 4.3 to bound $\mathcal{P}_s(\mathcal{S})$: by using estimates (4.8), (4.10), (4.11), (4.13) together with (4.20) or (4.21) in $s(\mathcal{S})$ and next by letting $\xi=x$ together with (4.5) we deduce that

$$\begin{aligned}\left|\mathcal{P}_s(\mathcal{S})\right| \leq C'(s-t)^{-\gamma/2}\big\{&(s-t)^{-3(1-\beta_i^2)/2}+\left\|D_1^2u_i\right\|_\infty+\|D_1D_2u_i\|_\infty\\ &+\|D_1u_i\|_\infty(s-t)^{-3(1-\beta_1^2)/2}+\|D_1D_2u_i\|_\infty(s-t)^{\beta_1^1/2}\\ &+\|D_2u_i\|_\infty(s-t)^{-3(1-\beta_2^2)/2}\\ &+M(D_2u_i,T)(s-t)^{-1+\alpha^1/2}\big((s-t)^{\gamma/2}+(s-t)^{3\beta_1^2/2}+(s-t)^{3\beta_2^2/2}+(s-t)^{3/2}\big)\big\}|x_2-z_2|^{\gamma/3},\end{aligned}$$

for all $\gamma<3\inf(\beta_1^2,\beta_2^2)-1$. By integrating in space and by using Lemma 4.1, we obtain that there exists a positive number $\delta_{4.2}^1$, depending only on known parameters in (H), such that:

$$\begin{aligned}\int_t^T\left|\mathcal{P}_s(\mathcal{S})\right|\mathrm{d}s \leq C''T^{\delta_{4.2}^1}&\big(M(D_2u_i,T)+\|D_2u_i\|_\infty+\|D_1D_2u_i\|_\infty+1\big)\\ &\times\big(|x_2-z_2|^{\gamma/3}+|x_2-z_2|^{\beta_2^2}+|x_2-z_2|^{\beta_1^2}+|x_2-z_2|\big),\end{aligned} \tag{4.22}$$

for all $\gamma<3\inf(\beta_1^2,\beta_2^2)-1$.

Estimation of $\mathcal{P}_\varepsilon(\mathcal{S}^c)$. As a second step, we bound the sum $\mathcal{P}_s(\mathcal{S}^c)$ on $(t,T]$ in (4.19). Note that

$$\mathcal{P}_s\big(\mathcal{S}^c\big)=\sum_{j=1}^4\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\mathcal{H}_i^j\big(s,y_1,y_2,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 \tag{4.23}$$

$$-\sum_{j=1}^4\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\mathcal{H}_i^j\big(s,y_1,y_2,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 \tag{4.24}$$

and that for all s in $\mathcal{S}^c$ we have:

$$1\leq(s-t)^{-\gamma/2}|x_2-z_2|^{\gamma/3}. \tag{4.25}$$

On the one hand we bound the right hand side of (4.23). Again, thanks to Lemma 3.5, we can identify the $\mathcal{H}_i^j$, $j=1,\dots,4$, appearing in this term as the *centered integrands* (see Definition 3.6). And again, by proceeding exactly as in the proof of Lemma 4.3 when $n=0$ (the restriction of the time integration on the set $\mathcal{S}^c$ is not a problem to do

[5] And that the time integration is done on the set $\mathcal{S}$.

so), we obtain that this term is bounded by the right hand side of (4.15). By using (4.25), we deduce that

$$\begin{aligned}
&\left|\sum_{j=1}^{4}\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\mathcal{H}_i^j\big(s,y_1,y_2,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2\right|\\
&\quad\le C(s-t)^{-\gamma/2}\big\{(s-t)^{-3(1-\beta_1^2)/2}+\|D_1^2u_i\|_\infty+\|D_1D_2u_i\|_\infty\\
&\qquad+\|D_1u_i\|_\infty(s-t)^{-3(1-\beta_1^2)/2}+\|D_1D_2u_i\|_\infty(s-t)^{-\beta_1^1/2}\\
&\qquad+\|D_2u_i\|_\infty(s-t)^{3(\beta_2^2-1)/2}\\
&\qquad+M(D_2u_i,T)(s-t)^{-1+\alpha^1/2}\big((s-t)^{\gamma/2}+(s-t)^{3\beta_1^2/2}+(s-t)^{3\beta_2^2/2}+(s-t)^{3/2}\big)\big\}|x_2-z_2|^{\gamma/3}.
\end{aligned}\tag{4.26}$$

On the other hand, we have to deal with the term (4.24). The crucial point here is that the frozen transition density is evaluated at point z_2 but frozen along the transport of x_2 (since we already chose $\xi=x$). This is because we took the same freezing point for the two solutions with different initial conditions. Hence, we have to re-center carefully each *integrand* $\mathcal{H}_i^j$, $j=1,\dots,4$, in order to use the Gaussian off-diagonal decay in $\tilde{q}$ for smoothing the time singularity. Indeed, in this case, for all s in $[t,T]$ and y_2 in $\mathbb{R}^d$, this off-diagonal decay is $|y_2-m_{t,s}^{2,\xi}(x_1,z_2)|$.

Hence, as we did in Lemma 3.5, we have to re-center the integrands around $(m_{t,s}^{2,\xi}(x_1,\tilde{\xi}_2))_{t<s\le T}$, for some $\tilde{\xi}_2$ in $\mathbb{R}^{2d}$ with the help of Claim 3.3 and then choose $\tilde{\xi}_2=z_2$. For notational convenience we write z_2 instead of $\tilde{\xi}_2$ as from now. Let

$$\begin{aligned}
&\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\mathcal{H}_i^j\big(s,y_1,y_2,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2\\
&\quad:=\tilde{\mathcal{I}}_i^j(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c},
\end{aligned}\tag{4.27}$$

for $j=1,\dots,4$. Below, we re-center each integrand of $\tilde{\mathcal{I}}_i^j(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}$, $j=1,\dots,4$, and we estimate it.

Bound of $\tilde{\mathcal{I}}_i^1(c,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}$. We deduce from assertion (b) in Claim 3.3 and (4.25) that:

$$\big|\tilde{\mathcal{I}}_i^1(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\big|\le(s-t)^{-\gamma/2}\int_{\mathbb{R}^{2d}}\Delta^2\big(\theta_{t,s}(z_2)\big)\phi_i(s,y_1,y_2)D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2|x_2-z_2|^{\gamma/3}.$$

Then, by using estimate on $D_{x_2}\tilde{q}$ from Proposition 3.1, regularity of ϕ_i under (H) a

$$\big|\tilde{\mathcal{I}}_i^1(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\big|\le C\mathbf{1}_{\mathcal{S}^c}(s-t)^{-3(1-\beta_i^2+\gamma/3)/2}|x_2-z_2|^{\gamma/3},\tag{4.28}$$

for all positive γ strictly less than $3\beta_i^2-1$.

Bound of $\tilde{\mathcal{I}}_i^2(s;t,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}$. We first split this term as:

$$\begin{aligned}
&\tilde{\mathcal{I}}_i^2(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\\
&\quad=-\frac{1}{2}\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\mathrm{Tr}\big[\big(\Delta^2\big(m_{t,s}^{t,\xi}(x_1,z_2)\big)a(s,y)\big)D_{y_1}^2u_i(s,y)\big]D_{z_2}\tilde{q}(t,x_1,z_2;s,y)\,\mathrm{d}y\\
&\qquad-\frac{1}{2}\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\mathrm{Tr}\big[\big(a\big(s,y_1,m_{t,s}^{2,t,\xi}(x_1,z_2)\big)-a\big(s,\theta_{t,s}(\xi)\big)\big)D_{y_1}^2u_i(s,y)\big]D_{z_2}\tilde{q}(t,x_1,z_2;s,y)\,\mathrm{d}y.
\end{aligned}$$

By applying assertion (b) in Claim 3.3 on the first term in the right hand side, by integrating by parts (see the proof of Lemma 3.5) the second term in the right hand side and then by applying assertion (c) in Claim 3.3 we get:

$$\begin{aligned}
&\tilde{\mathcal{I}}_i^2(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\\
&\quad=-\frac{1}{2}\int_{t+\varepsilon}^{T}\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\Big\{\mathrm{Tr}\big[\big(\Delta^2\big(m_{t,s}^{t,\xi}(x_1,z_2)\big)a(s,y_1,y_2)\big)D_{y_1}^2u_i(s,y_1,y_2)\big]
\end{aligned}$$

$$
\begin{aligned}
&-\sum_{l=1}^{d}\Bigg(\Bigg[\frac{\partial}{\partial y_{1l}}a_{l.}\big(s,y_1,m_{t,s}^{t,\xi}(x_1,z_2)\big)\Bigg]\\
&\cdot\big[\Delta^2\big(m_{t,s}^{t,\xi}(x_1,z_2)\big)D_{y_1}u_i(s,y_1,y_2)\big]\Bigg)\Bigg\}\times\big[D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\big]\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\\
&+\frac{1}{2}\sum_{l=1}^{d}\int_{t+\varepsilon}^{T}\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\Big\{\big[a_{l.}\big(s,y_1,m_{t,s}^{2,t,\xi}(x_1,z_2)\big)-a_{l.}\big(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)\big]\\
&\cdot\big[\Delta^2\big(m_{t,s}^{t,\xi}(x_1,z_2)\big)D_{y_1}u_i(s,y_1,y_2)\big]\times\Bigg[\frac{\partial}{\partial y_{1l}}D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\Bigg]\Big\}\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s.
\end{aligned}
$$

Therefore, from the estimates on the derivatives of $\tilde{q}$ from Proposition 3.1, by using MVT (4.9) and the regularity of a from (H) together with estimate (4.5) and by integrating next in space we obtain that

$$
\begin{aligned}
&\big|\tilde{\mathcal{I}}_i^2(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\big|\\
&\quad\le C'\mathbf{1}_{\mathcal{S}^c}\big\{\big(\big\|D_1^2u_i\big\|_\infty(s-t)^{-\gamma/2}+\|D_1D_2u_i\|_\infty(s-t)^{-\gamma/2}\big)|x_2-z_2|^{\gamma/3}\\
&\qquad+\|D_1D_2u_i\|_\infty(s-t)^{-1/2}\big|m_{t,s}^{2,t,\xi}(x_1,z_2)-\theta_{t,s}^2(\xi)\big|\big\},
\end{aligned}
\tag{4.29}
$$

by using next the estimate (4.25) in the first term on the right hand side. This holds for all $\gamma<2$.

Bound of $\tilde{\mathcal{I}}_i^3(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}$. By using the definition of Δ^2 and then assertion (c) in Claim 3.3, this term can be centered as follows

$$
\begin{aligned}
\tilde{\mathcal{I}}_i^3(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}=-\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}&\big\{\big[\Delta^2\big(m_{t,s}^{t,\xi}(x_1,z_2)\big)F_1(s,y_1,y_2)\big]\cdot\big[D_{y_1}u_i(s,y_1,y_2)\big]\\
&+\big[F_1\big(s,y_1,m_{t,s}^{2,t,\xi}(x_1,z_2)\big)-F_1\big(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)\big]\\
&\cdot\big[\Delta^2\big(m_{t,s}^{2,t,\xi}(x_1,z_2)\big)D_{y_1}u_i(s,y_1,y_2)\big]\big\}D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2.
\end{aligned}
$$

Now, thanks to the regularity of F_1 assumed in (H), if we apply the estimate on the derivative of $\tilde{q}$ from Proposition 3.1, MVT (4.9) and estimate (4.25), we obtain that

$$
\begin{aligned}
&\big|\tilde{\mathcal{I}}_i^3(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\big|\\
&\quad\le C''\mathbf{1}_{\mathcal{S}^c}\big\{\big(\|D_1u_i\|_\infty(s-t)^{-3(1-\beta_1^2+\gamma/3)/2}+\|D_1D_2u_i\|_\infty(s-t)^{(\beta_1^1-\gamma)/2}\big)|x_2-z_2|^{\gamma/3}\\
&\qquad+\|D_1D_2u_i\|_\infty\big|m_{t,s}^{2,t,\xi}(x_1,z_2)-\theta_{t,s}^2(\xi)\big|^{\beta_1^2}\big\}
\end{aligned}
\tag{4.30}
$$

for all $\gamma<3\beta_1^2-1$.

Bound of $\tilde{\mathcal{I}}_i^4(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}$. From representation (3.13), by using the definition of Δ^2 and then by centering the term D_2u_i thanks to assertion (c) in Claim 3.3 we can write

$$
\begin{aligned}
&\tilde{\mathcal{I}}_i^4(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\\
&\quad=-\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\big\{\big[\Delta^2\big(m_{t,s}^{t,\xi}(x_1,z_2)\big)F_2(s,y_1,y_2)\big]\cdot\big[D_{y_2}u_i(s,y_1,y_2)\big]\\
&\qquad+\big[F_2\big(s,y_1,m_{t,s}^{2,t,\xi}(x_1,z_2)\big)-F_2\big(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)-D_1F_2\big(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)\big)\Delta^1\big(\theta_{t,s}(\xi)\big)y_1\big]\\
&\qquad\cdot\big[\Delta^2\big(m_{t,s}^{2,t,\xi}(x_1,z_2)\big)D_{y_2}u_i(s,y_1,y_2)\big]\big\}D_{z_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2.
\end{aligned}
$$

By using the regularity of the coefficients from (H), (4.4) and estimate on $D_2\tilde{q}$ from Proposition 3.1 and (4.25) we have:

$$\begin{aligned}
&\big|\tilde{\mathcal{I}}_i^4(s,x_1,z_2)\mathbf{1}_{\mathcal{S}^c}\big| \\
&\quad\le C'''\mathbf{1}_{\mathcal{S}^c}\big\{\|D_2u_i\|_\infty(s-t)^{-3(1-\beta_2^2+\gamma/3)/2}|x_2-z_2|^{\gamma/3}+M(D_2u_i,T)\big[\big|m_{t,s}^{2,t,\xi}(x_1,z_2)-\theta_{t,s}^2(\xi)\big| \\
&\qquad+(s-t)^{-1+\alpha^1/2}\big(|x_2-z_2|^{\gamma/3}+|x_2-z_2|^{\beta_1^2}+|x_2-z_2|^{\beta_2^2}+|x_2-z_2|\big)\big]\big\},
\end{aligned} \tag{4.31}$$

for all $\gamma<3\beta_1^2/2-1$.

Now, note that from the definition (3.3) of m,

$$\forall s\in[t,T],\quad m_{t,s}^{2,t,x}(x_1,z_2)-\theta_{t,s}^2(x)=z_2-x_2.$$

Hence, by letting $\xi=x$ in (4.28), (4.29), (4.30) and (4.31) and combining the resulting estimates with (4.26), we deduce that there exist a positive constant C'''' and a positive number $\delta_{4.2}^2$, depending only on known parameters in (H), such that:

$$\begin{aligned}
\int_t^T\big|\mathcal{P}_s\big(\mathcal{S}^c\big)\big|\,\mathrm{d}s &\le C''''T^{\delta_{4.2}^2}\big(M(D_2u_i,T)+\|D_2u_i\|_\infty+\|D_1D_2u_i\|_\infty+1\big) \\
&\quad\times\big(|x_2-z_2|^{\gamma/3}+|x_2-z_2|^{\beta_2^2}+|x_2-z_2|^{\beta_1^2}+|x_2-z_2|\big),
\end{aligned} \tag{4.32}$$

for all $\gamma<3\inf(\beta_1^2,\beta_2^2)-1$.

"Hölder estimate" on D_2u_i. Finally, by plugging estimates (4.22) and (4.32) in (4.18), we deduce that there exist a positive constant C and a positive number $\delta_{4.2}$, depending only on known parameters in (H), such that:

$$\begin{aligned}
\big|D_2u_i(t,x_1,x_2)-D_2u_i(t,x_1,z_2)\big| &\le CT^{\delta_{4.2}}\big(\|D_2u_i\|_\infty+\|D_1D_2u_i\|_\infty+M(D_2u_i,T)+1\big) \\
&\quad\times\big(|x_2-z_2|^{\gamma/3}+|x_2-z_2|^{\beta_2^2}+|x_2-z_2|^{\beta_1^2}+|x_2-z_2|\big).
\end{aligned}$$

A circular argument concludes the proof of Lemma 4.2. □

Proof of Claim 4.4. Let $(t<s,x,y)$ in $[0,T]^2\times\mathbb{R}^{2d}\times\mathbb{R}^{2d}$, by using MVT and the Gaussian estimate of $D_2^2\tilde{q}$ from Proposition 3.1 we have:

$$\begin{aligned}
&\big|(D_2\tilde{q})(t,x_1,x_2;s,y_1,y_2)-(D_2\tilde{q})(t,x_1,z_2;s,y_1,y_2)\big| \\
&\quad\le\sup_{\rho\in(0,1)}\big|\big(D_2^2\tilde{q}\big)\big(t,x_1,x_2+\rho(x_2-z_2);s,y_1,y_2\big)\big||x_2-z_2| \\
&\quad\le C'(s-t)^{-3}\sup_{\rho\in(0,1)}\hat{q}_{\bar{c}}\big(t,x_1,x_2+\rho(x_2-z_2);s,y_1,y_2\big)|x_2-z_2|,
\end{aligned} \tag{4.33}$$

where $\bar{c}$ is a positive constant depending only on known parameters in (H). Note that on $\mathcal{S}$:

$$\sup_{\rho\in(0,1)}\hat{q}_{\bar{c}}\big(t,x_1,x_2+\rho(x_2-z_2);s,y_1,y_2\big)\le C''\hat{q}_c(t,x_1,x_2;s,y_1,y_2). \tag{4.34}$$

Combining (4.33) and (4.34), we obtain:

$$\begin{aligned}
&\big|(D_2\tilde{q})(t,x_1,x_2;s,y_1,y_2)-(D_2\tilde{q})(t,x_1,z_2;s,y_1,y_2)\big| \\
&\quad\le C'''(s-t)^{-3}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)|x_2-z_2|.
\end{aligned}$$

Rewrite $|x_2-z_2|=|x_2-z_2|^{1-\gamma/3}|x_2-z_2|^{\gamma/3}$. Since $|x_2-z_2|<(s-t)^{3/2}$ we have $|x_2-z_2|<(s-t)^{3/2-\gamma/2}\times|x_2-z_2|^{\gamma/3}$ and (4.20) follows.

The second assertion follows from the same arguments. □

Acknowledgements

I thank François Delarue for his suggestions, large comments and careful reading of the paper as well as the anonymous referees for providing valuable inputs that improved the paper.

References

[1] R. F. Bass. *Diffusions and Elliptic Operators*. *Probability and Its Applications (New York)*. Springer-Verlag, New York, 1998. MR1483890
[2] A. M. Davie. Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not. IMRN* **24** (2007) Article ID rnm124. MR2377011
[3] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.* **259** (6) (2010) 1577–1630. MR2659772
[4] M. Di Francesco and S. Polidoro. Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. *Adv. Differential Equations* **11** (11) (2006) 1261–1320. MR2277064
[5] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics* **83** (3) (2011) 241–257. MR2810591
[6] F. Flandoli. *Random Perturbation of PDEs and Fluid Dynamic Models*. *Lecture Notes in Mathematics* **2015**. Springer, Heidelberg, 2011. MR2796837
[7] M. I. Freidlin. *Functional Integration and Partial Differential Equations*. Princeton University Press, Princeton, NJ, 1985. MR0833742
[8] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, NJ, 1964. MR0181836
[9] A. Friedman. *Stochastic Differential Equations and Applications*. Dover, Mineola, NY, 2006. MR2295424
[10] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.* **119** (1967) 147–171. MR0222474
[11] A. Kolmogorov. Zufällige bewegungen (zur theorie der brownschen bewegung). *Ann. of Math. (2)* **35** (1934) 116–117. MR1503147
[12] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** (2) (2005) 154–196. MR2117951
[13] H. Kunita. Stochastic differential equation and stochastic flows of diffeomorphisms. In *École d'été de probabilités de Saint-Flour, XII – 1982* 143–303. Springer-Verlag, Berlin, 1982. MR0876080
[14] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab.* **16** (2011) 234–250. MR2802040
[15] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **233**. Springer-Verlag, Berlin, 1979. MR0532498
[16] A. J. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S.)* **111(153)** (1980) 434–452, 480. MR0568986
[17] A. Y. Veretennikov. Stochastic equations with diffusion that degenerates with respect to part of the variables. *Izv. Ross. Akad. Nauk Ser. Mat.* **47** (1) (1983) 189–196. MR0688921
[18] X. Zhang. Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Rev. Mat. Iberoamericana* **29** (1) (2010) 25–52. MR3010120
[19] X. Zhang. Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** (2011) 1096–1116. MR2820071
[20] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)* **93(135)** (1974) 129–149. 152. MR0336813