Accepted Manuscript



Martingale problems for some degenerate Kolmogorov equations

Stéphane Menozzi

PII: S0304-4149(17)30156-4
DOI: http://dx.doi.org/10.1016/j.spa.2017.06.001
Reference: SPA 3140

To appear in: *Stochastic Processes and their Applications*

Received date : 21 May 2016
Revised date : 6 February 2017
Accepted date : 7 June 2017

Please cite this article as: S. Menozzi, Martingale problems for some degenerate Kolmogorov equations, *Stochastic Processes and their Applications* (2017), http://dx.doi.org/10.1016/j.spa.2017.06.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

*Manuscript

MARTINGALE PROBLEMS FOR SOME DEGENERATE KOLMOGOROV EQUATIONS

STÉPHANE MENOZZI

Abstract. We obtain Calderón-Zygmund estimates for some degenerate equations of Kolmogorov type with inhomogeneous nonlinear coefficients. We then derive the well-posedness of the martingale problem associated with related degenerate operators, and therefore uniqueness in law for the corresponding stochastic differential equations. Some density estimates are established as well.

1. Introduction

1.1. **Statement of the problem.** Consider the following system of Stochastic Differential Equations (SDEs in short)

$$\begin{aligned} &dX_t^1 = F_1(t, X_t^1, \dots, X_t^n)dt + \sigma(t, X_t^1, \dots, X_t^n)dW_t, \\ &dX_t^2 = F_2(t, X_t^1, \dots, X_t^n)dt, \\ &dX_t^3 = F_3(t, X_t^2, \dots, X_t^n)dt, \qquad\qquad t \geq 0, \\ &\dots \\ &dX_t^n = F_n(t, X_t^{n-1}, X_t^n)dt, \end{aligned} \tag{1.1}$$

$(W_t)_{t\geq 0}$ standing for a d-dimensional Brownian motion, and each $(X_t^i)_{t\geq 0}$, $i \in [\![1, n]\!]$, being $\mathbb{R}^d$-valued as well.

From the applicative viewpoint, systems of type (1.1) appear in many fields. Let us for instance mention for $n = 2$ stochastic Hamiltonian systems (see e.g. Soize [Soi94] for a general overview or Talay [Tal02] and Hérau and Nier [HN04] for convergence to equilibrium). Again for $n = 2$, the above dynamics is used in mathematical finance to price Asian options (see for example [BPV01]). For $n \geq 2$, it appears in heat conduction models (see e.g. Eckmann et al. [EPRB99] and Rey-Bellet and Thomas [RBT00] when the chain of differential equations is forced by two heat baths).

Assume first that the coefficients $(F_i)_{i\in[\![1,n]\!]}$ are Lipschitz continuous in space and that the diffusion matrix $a(t,.) := \sigma\sigma^*(t,.)$ is bounded. If we additionally suppose that $a(t,.)$ and $(D_{x_{i-1}}F_i(t,.))_{i\in[\![2,n]\!]}$ are non-degenerate (weak Hörmander condition) and Hölder continuous in space, with respective Hölder exponents in $(1/2, 1]$ and $(0, 1]$, some multi-scale Gaussian Aronson like estimates have been proved in [DM10] for the density of (1.1) uniformly on the time set $(0, T]$, for fixed $T > 0$ (see Example 2 and Theorem 1.1 of that reference). Those results extend to the case of an arbitrary Hölder exponent in $(0, 1]$ for $a(t,.)$ thanks to uniqueness in law arguments that have been investigated in [Men11] through the well posedness

Date: February 6, 2017.

2000 *Mathematics Subject Classification.* Primary 60H10, 60G46; Secondary 60H30, 35K65.

Key words and phrases. Degenerate SDEs, martingale problem, Calderón-Zygmund estimates.

ACCEPTED MANUSCRIPT

of the martingale problem. This result is established exploiting specifically some regularizing effect of an underlying *parametrix* kernel.

Anyhow, when studying the martingale problem, the natural framework is to consider non-degenerate continuous coefficients[1]. In the special case $n = 1$ if $a(t,.)$ is bounded and uniformly elliptic, i.e. (1.1) corresponds to a non-degenerate SDE, it is well known that the martingale problem associated with the generator $(L_t)_{t\geq 0}$ of (1.1) is well posed as soon as the coefficient F_1 is bounded measurable and that $a(t,.)$ is continuous in space, see e.g. Stroock and Varadhan [SV79].

The key ingredient consists in proving Calderón-Zygmund estimates. In that framework, those estimates write as controls in L^p norms of suitable singular integrals related to the Gaussian density of (1.1) when the diffusion coefficient is constant in space and $F_1 = 0$. These controls then allow, through an operator inversion, to derive the well posedness of the martingale problem, or from a PDE viewpoint of the Cauchy problem with L^p source term, when the diffusion coefficient does not vary much. The case of a bounded drift can then be handled through a Girsanov transform. Eventually, the well posedness of the martingale problem is established under the sole continuity assumption on a and boundedness on b thanks to a localization procedure.

We refer to the monographs of Stein [Ste70] or Gilbarg and Trudinger [GT83] for a presentation of the Calderón-Zygmund theory for non-degenerate elliptic equations. In that framework a more probabilistic approach is proposed in Bass [Bas95]. We also mention the monograph of Coifman and Weiss [CW71], from which some Calderón-Zygmund estimates can be derived in some degenerate frameworks when there is an underlying *homogeneous space*.

In this work, for $n > 1$, under the previous assumptions of non-degeneracy and continuity on a and a weak Hörmander condition of the $(D_{x_{i-1}}F_i(t,.)_{i\in[\![2,n]\!]})$, we are interested in proving the well-posedness of the martingale problem for the generator $(L_t)_{t\geq 0}$ of (1.1).

To achieve this goal, we will establish Calderón-Zygmund estimates for a singular Gaussian kernel derived from a suitable linearization of the degenerate system (1.1), already used in [DM10], [Men11]. The linearization is here crucial since it provides the *proxy* density on which some good controls can be established. Observe indeed that when σ is constant in space and the coefficients F are linear and s.t. $(D_{x_{i-1}}F_i)_{i\in[\![2,n]\!]}$ satisfy a weak Hörmander condition, then the SDE has a multi-scale Gaussian density (see the seminal paper of Kolmogorov [Kol34], Di Francesco and Polidoro [FP06], [DM10]). Roughly speaking, the non degeneracy of a and the Hörmander assumption on the drift allow to say that the i^{th} component of the SDE *feels* a noise whose typical scale corresponds to the one of the $(i-1)^{\text{th}}$ iterated integral of the Brownian motion which is $t^{1/2+(i-1)} = t^{(2i-1)/2}$ at time t. On the other hand, the Gaussian density exhibits deviations w.r.t. the transport of the initial condition by the deterministic differential system[2] having unbounded coefficients. The multi-scale Gaussian densities of Kolmogorov type will play here the same role as the standard Gaussian one in the non-degenerate setting of [SV79].

[1]This assumption yields even in the non-degenerate case, estimates in L^q spaces for the density (see e.g. Chapter 9 in [SV79]), whereas the Hölder continuity gives, still in a weak solution framework, pointwise controls (see e.g. Sheu [She91]).

[2]corresponding to (1.1) when $\sigma = 0$

ACCEPTED MANUSCRIPT

Let us now mention that in the linear case, Calderón-Zygmund estimates have been obtained by Bramanti *et al.* in [BCLP10], [BCLP13]. Precisely, they consider an operator

$$\mathcal{A} := \sum_{i,j=1}^{p_0} a_{ij}\partial_{x_i,x_j} + \sum_{i,j=1}^{N} b_{ij}x_i\partial_{x_j},$$

where the matrices $(a_{ij})_{(i,j)\in[\![1,p_0]\!]^2}$ are symmetric positive definite, constant in [BCLP10], with continuous variable homogeneous coefficients which do not vary much in [BCLP13], and the $(b_{ij})_{(i,j)\in[\![1,N]\!]^2}$ are s.t. $\mathcal{A}$ is hypoelliptic. The authors then establish global L^p estimates $p \in (1,+\infty)$ of the following type: $\exists c := c(a,b,p_0,N,p),\ \forall u \in C_0^2(\mathbb{R}^N)$,

$$\|\partial_{x_ix_j}u\|_{L^p(\mathbb{R}^N)} \leq c\{\|\mathcal{A}u\|_{L^p(\mathbb{R}^N)} + \|u\|_{L^p(\mathbb{R}^N)}\},\ (i,j) \in [\![1,p_0]\!]^2. \tag{1.2}$$

Weak (1-1) estimates are also obtained. It has been proved in [LP94] that for some suitable basis, the matrices b have the same form as in (1.1), i.e. in our setting the coefficient F would write $F(x) = Bx$ where $B_{i,j} = 0_{d\times d}$ for $j < i-1,\ (i,j) \in [\![1,n]\!]^2$. Hence, the operator $\mathcal{A}$ can be seen as a particular case of generator associated with (1.1).

The strategy in those works still consists in estimating suitable singular integrals[3] related to the Gaussian fundamental solution of $\mathscr{L} = \mathcal{A} - \partial_t$ that enjoys the previously described properties (multi-scale and unbounded transport). Even in the linear framework, when the matrix B has strictly upper diagonal entries it is not possible to enter the Coifman and Weiss [CW71] framework of homogeneous spaces. Here, the underlying homogeneous norm would be the one derived from the various time scales of the components, corresponding once again to those of the iterated integrals of the Brownian motion. Namely,

$$\forall (s,x_1,\cdots,x_n) \in \mathbb{R}\times\mathbb{R}^{nd},\ \bar{\rho}(s,x_1,\cdots,x_n) := |s|^{1/2} + \sum_{i=1}^{n}|x_i|^{1/(2i-1)}. \tag{1.3}$$

Now, for a given $i \in [\![1,n-1]\!]$, the entries $B_{i,j} \in \mathbb{R}^d \otimes \mathbb{R}^d,\ j \in [\![i+1,n]\!]$ are associated with components that have negligible time scale, namely $t^{(2j-1)/2}$, w.r.t. to the current one of order $t^{(2i-1)/2}$ in small time. This property has been exploited thoroughly in [DM10], [Men11] and Section 3 to derive pointwise estimates, but in the current framework it breaks the global homogeneity in (1.3) when considering, for $(s,x),(t,y) \in \mathbb{R}\times\mathbb{R}^{nd}$, $\bar{\mathbf{d}}((s,x),(t,y)) = \bar{\rho}(t-s, x-\exp(-B(t-s))y)$ which appears as a natural candidate to be a quasi-distance taking into account the transport. It can be shown that $\bar{\mathbf{d}}((s,x),(t,y))$ and $\bar{\mathbf{d}}((t,y),(s,x))$ are equivalent on quasi metric balls only (whereas they are actually globally equivalent when there are no strictly upper-diagonal contributions in B, see also Section 3). This observation could lead to consider the associated metric balls as homogeneous spaces to be in the Coifman-Weiss setting. The problem with this choice is that it is not clear anymore that the (sub)-balls enjoy the doubling property. This is why the authors in [BCLP10], [BCLP13] rely on some specific estimates established by Bramanti [Bra10] on possibly non doubling spaces (see Section 3 for details). Also, their analysis strongly relies on some underlying Lie group structure.

Let us mention that for variable homogeneous coefficients (namely in VMO_{loc} (resp. C^α) w.r.t. the distance induced by the vector fields), local L^p (resp.

[3]precisely second order derivatives w.r.t. to the non degenerate components

ACCEPTED MANUSCRIPT

Schauder) estimates have been obtained by Bramanti and Zhu [BZ13] following the same lines. Concerning the link between the L^p estimates of [BCLP13] and the well-posedness of the martingale problem, we can refer to the recent work of Priola [Pri15] who introduces a rather general localization procedure that allows to extend the well posedness of the martingale problem from the case of *almost constant* coefficients to the natural one of continuous coefficients. The contribution of [Pri15], w.r.t. to the classic localization results of Stroock and Varadhan (see e.g. Chapter 6.6 in [SV79]), being the handling of rather general unbounded coefficients.

The main novelty in our approach consists in considering inhomogeneous coefficients and rather general non-linear drifts for the degenerate part of equation (1.1). To this end we introduce a suitable kernel, for which we establish specific Calderón-Zygmund estimates, and do not exploit some underlying Lie group properties appearing in the quoted works and which fail in our setting since they are heavily related to a time homogeneous and linear drift. Importantly, even in the time-homogeneous setting, the non-linear structure and the unboundedness of the drift in (1.1) does not allow to directly use the results of [BCLP10], [BCLP13] through a linearization and localization procedure. The key idea consists in viewing (1.1) as a non-linear ODE *perturbed* by a noise. This naturally yields to consider balls that are build around the characteristic lines of the ODE and reflect the multiscale behavior of the process, where the various scales are once again those of the Brownian motion and its iterated integrals. This approach also allows, through a suitable localization procedure, to establish density estimates in L^q (see equation (2.2) in Theorem 2.1), which is the natural framework for diffusion coefficients that are just continuous. This is to our best knowledge the first result of this kind in the weak Hörmander setting, even for a homogeneous linear drift.

Let us also mention that, as a byproduct of our Caldéron-Zygmund estimates, it should be possible to get the well posedness of a decoupled degenerate BSDE having a Hölder continuous in space driver (or from the analytical viewpoint, to develop a strong theory for semi-linear degenerate PDEs with Hölder in space source term) following the lines of Delarue and Guatteri [DG06].

A challenging open problem would consist in extending the density estimates of equation (2.2) to degenerate Itô processes of the form (1.1) where the diffusion coefficient would simply be measurable, bounded from above and from below. This would indeed give degenerate Krylov like estimates (see Sections 2 and 3 of Chapter 2 in [Kry87] in the uniformly elliptic setting) which would be the crux to get existence and uniqueness results for fully coupled degenerate Backward SDEs (or again to get a strong theory of quasilinear degenerate PDEs) with the previous type of drift (see Delarue [Del02], [Del03] for an exposition of the strategy in the non-degenerate case).

The article is organized as follows. We state our assumptions and main results in Section 2. We then introduce in Section 3 the degenerate Gaussian kernel for which we establish Calderón-Zygmund estimates, recalling formally how uniqueness can be derived from these controls when the coefficients do not vary much. In Section 4 we specify the various steps that lead to the Calderón-Zygmund estimates of Theorem 3.1. We then perform in Section 5 a localization procedure and give some local and global controls on the density from the previous estimates. This requires some careful extensions of the arguments of the non-degenerate framework, see e.g. Sections 7.1, 7.2 and 9.1 in [SV79]), exploiting again the characteristic lines of the

ACCEPTED MANUSCRIPT

underlying ODE. Section 6 is the technical core of the paper and is devoted to the proof of the technical results of Section 4.

2. Assumptions and Main Results

2.1. **Notations and Assumptions.** In what follows, we denote a quantity in $\mathbb{R}^{nd}$ by a bold letter: i.e. $\mathbf{0}$, stands for zero in $\mathbb{R}^{nd}$ and the solution $(X_t^1,\dots,X_t^n)_{t\geq 0}$ to (1.1) is denoted by $(\mathbf{X}_t)_{t\geq 0}$. Introducing the embedding matrix B from $\mathbb{R}^d$ into $\mathbb{R}^{nd}$, i.e. $B=(I_d,0,\dots,0)^*$, where "$*$" stands for the transpose, we rewrite (1.1) in the shortened form

$$d\mathbf{X}_t=\mathbf{F}(t,\mathbf{X}_t)dt+B\sigma(t,\mathbf{X}_t)dW_t,$$

where $\mathbf{F}=(F_1,\dots,F_n)$ is an $\mathbb{R}^{nd}$-valued function.

With these notations the generator of (1.1) writes for all $t\geq 0$:

(2.1)
$$\forall\varphi\in C_0^2(\mathbb{R}^{nd}),\ \forall\mathbf{x}\in\mathbb{R}^{nd},\ L_t\varphi(\mathbf{x})=\langle\mathbf{F}(t,\mathbf{x}),\mathbf{D}_{\mathbf{x}}\varphi(\mathbf{x})\rangle+\frac{1}{2}\mathrm{tr}(a(t,\mathbf{x})D^2_{\mathbf{x}_1}\varphi(\mathbf{x})).$$

Also, for a point $\mathbf{x}:=(\mathbf{x}_1,\cdots,\mathbf{x}_n)\in\mathbb{R}^{nd}$, we will often denote for all $i\in[\![1,n-1]\!]$, $\mathbf{x}^{i,n}:=(\mathbf{x}_i,\cdots,\mathbf{x}_n)$. The notation $|\cdot|$ stands for the (Euclidean) norm on $\mathbb{R}^m$ or $\mathbb{R}^m\otimes\mathbb{R}^m$, $m\in\{d,nd\}$.

Let us now introduce some assumptions concerning the coefficients of (1.1).

(C) The diffusion coefficient $(a(t,.))_{t\geq 0}$ is bounded measurable and continuous in space, i.e.

$$\lim_{\mathbf{y}\to\mathbf{x}}\sup_{0\leq s\leq T}|a(s,\mathbf{y})-a(s,\mathbf{x})|=0$$

for all $T>0$ and $\mathbf{x}\in\mathbb{R}^{nd}$.

(UE) There exists $\Lambda\geq 1$, $\forall t\geq 0,\mathbf{x}\in\mathbb{R}^{nd},\ \xi\in\mathbb{R}^d,\ \Lambda^{-1}|\xi|^2\leq\langle a(t,\mathbf{x})\xi,\xi\rangle\leq\Lambda|\xi|^2$.

(S) The $(F_i)_{i\in[\![1,n]\!]}$ are bounded measurable in time, globally Lipschitz continuous in space. Also, the functions $(F_i)_{i\in[\![2,n]\!]}$ are once continuously differentiable in the variable $\mathbf{x}_{i-1}$ and the $(D_{\mathbf{x}_{i-1}}F_i)_{i\in[\![2,n]\!]}$ are η-Hölder continuous in space, $\eta\in(0,1]$.

(ND) There exists a closed convex subset $\mathcal{E}_{i-1}\subset GL_{\mathbf{d}}(\mathbb{R})$ (set of invertible $d\times d$ matrices) s.t., for all $t\geq 0$ and $(\mathbf{x}_{i-1},\dots,\mathbf{x}_n)\in\mathbb{R}^{(n-i+2)d}$, $D_{\mathbf{x}_{i-1}}F_i(t,\mathbf{x}_{i-1},\dots,\mathbf{x}_n)\in\mathcal{E}_{i-1}$. For example, $\mathcal{E}_i$, $i\in[\![1,n-1]\!]$, may be a closed ball included in $GL_{\mathbf{d}}(\mathbb{R})$, which is an open set.

Assumptions **(UE)**, **(ND)** can be seen as a kind of (weak) Hörmander condition. They allow to transmit the non degenerate noise of the first component to the other ones. Let us also recall that the last part of Assumption **(ND)** and the particular structure of $\mathbf{F}(t,.)=(F_1(t,.),\cdots,F_n(t,.))$ yield that the i^{th} component of the system (1.1) has intrinsic time scale $(2i-1)/2, i\in[\![1,n]\!]$. This fact will be thoroughly used in our analysis (see Section 3 for details). We notice that the coefficients may be irregular in time, see **(S)**. We say that assumption **(A)** is in force when **(C)**, **(UE)**, **(S)**, **(ND)** hold.

2.2. **Main Results.** Our main result is the following theorem.

Theorem 2.1. Under **(A)** the martingale problem associated with $(L_t)_{t\geq 0}$ in (2.1) is well-posed. That is, for every $\mathbf{x}\in\mathbb{R}^{nd}$, there exists a unique probability measure $\mathbb{P}$ on $C(\mathbb{R}^+,\mathbb{R}^{nd})$ s.t. denoting by $(\mathbf{X}_t)_{t\geq 0}$ the canonical process, $\mathbb{P}[\mathbf{X}_0=\mathbf{x}]=1$ and for all $\varphi\in C_0^{1,2}(\mathbb{R}^+\times\mathbb{R}^{nd},\mathbb{R})$, $\varphi(t,\mathbf{X}_t)-\varphi(0,\mathbf{x})-\int_0^t(\partial_s+L_s)\varphi(s,\mathbf{X}_s)ds$ is a $\mathbb{P}$-martingale. In particular, weak uniqueness in law holds for the SDE (1.1).

ACCEPTED MANUSCRIPT

Also, if the diffusion coefficient a is uniformly continuous, the unique weak solution of (1.1) admits a density in the following sense. Letting $P(s,t,\mathbf{x},.)$ be the transition probability determined by $(L_t)_{t\geq 0}$, then for a given $T>0$, almost all $t\in(s,T]$ and all $\Gamma\in\mathcal{B}(\mathbb{R}^{nd})$, $P(s,t,\mathbf{x},\Gamma)=\int_\Gamma p(s,t,\mathbf{x},\mathbf{y})d\mathbf{y}$.

More specifically, for any $f\in L^p([0,T]\times\mathbb{R}^{nd})$, $p>\frac{(n^2d+2)}{2}$, there exists $C_{2.2}:=C_{2.2}(T,p,(\mathbf{A}))$ s.t. for all $(s,\mathbf{x})\in[0,T)\times\mathbb{R}^{nd}$:

$$|\mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}[\int_s^T f(t,\mathbf{X}_t)dt]|\leq C_{2.2}(1+|\mathbf{x}|)\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})}, \tag{2.2}$$

where $\mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}$ denotes the expectation w.r.t. $\mathbb{P}_{s,\mathbf{x}}[\cdot]:=\mathbb{P}[\cdot|\mathbf{X}_s=\mathbf{x}]$.

Remark 2.1. Let us first emphasize that by duality, the previous control gives a bound for the density in $L^q([0,T]\times\mathbb{R}^{nd})$ where $q^{-1}+p^{-1}=1$. Also, the contribution in $\mathbf{x}$ in the r.h.s. of (2.2) is specifically linked to the unboundedness of the drift term in (1.1). It derives from the localization procedure needed for the analysis. Namely, we are led to consider a suitable partition of $[0,T]\times\mathbb{R}^{nd}$ on which the coefficients of (1.1) satisfy a same given continuity constraint. Consider for instance a given point $(s,\mathbf{x})\in[0,T]\times\mathbb{R}^{nd}$, and a given threshold $\varepsilon>0$. It is then clear that, for the deterministic differential system deriving from (1.1) with dynamics

$$\dot{\boldsymbol{\theta}}_{t,s}(\mathbf{x})=\mathbf{F}(t,\boldsymbol{\theta}_{t,s}(\mathbf{x})),t\geq s,\boldsymbol{\theta}_{s,s}(\mathbf{x})=\mathbf{x}, \tag{2.3}$$

which can somehow be seen as the *mean* of the system, one has $|\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{x}|\leq\int_s^t(\kappa|\boldsymbol{\theta}_{u,s}(\mathbf{x})|+c)du\leq(t-s)(c+\kappa|\mathbf{x}|)\exp(\kappa(t-s))$, $t\in[s,T]$, where $c:=\sup_{u\in[0,T]}|F(u,\mathbf{0})|$ and κ stands for the Lipschitz constant of $\mathbf{F}$. Hence, one has $|\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{x}|\leq\varepsilon$ for $|t-s|\leq C/(1+|\mathbf{x}|), C:=C(c,\kappa,\varepsilon)$, which corresponds to the time step of the partition for a given $|\mathbf{x}|$. For a fixed $T>0$, $\lceil(1+|\mathbf{x}|)C^{-1}T\rceil$ is then an upper bound for the total number of time-steps.

3. "Frozen" Kernel and Formal derivation of uniqueness from Calderón-Zygmund estimates

Assume **(A)** is in force. One of the main differences between the uniform Hölder continuity assumed in [BP09] in the non degenerate case or in [Men11] for the current framework and the continuity statement of **(C)** is that in the first two cases no localization is needed. Indeed, the global Hölder continuity allows to remove globally the time singularities coming from the second order spatial derivatives of suitable Gaussian kernels arising in a parametrix like expansion of the density. In the current framework we first focus on the "local case". As in the non-degenerate case, we assume the diffusion coefficient $a(t,.):=\sigma\sigma^*(t,.)$ of (1.1) "does not vary much" (see e.g. Chapter 7 of [SV79]). Precisely, we first assume that there exists a measurable function $\varsigma:[0,T]\to\mathcal{S}_d$ (symmetric matrices of dimension d) satisfying **(UE)** and such that

$$\varepsilon_a:=\sup_{0\leq t\leq T}\sup_{\mathbf{x}\in\mathbb{R}^{nd}}|a(t,\mathbf{x})-\varsigma(t)|, \tag{3.1}$$

is small. In particular, we do not assume any *a priori* continuity of a. The continuity assumption **(C)** will actually allow, through a suitable localization procedure described in Section 5, to have (3.1) for all $\mathbf{x}_0\in\mathbb{R}^{nd}$ with $\varsigma(t)=a(t,\mathbf{x}_0)$ on some neighborhood of $\mathbf{x}_0$.

ACCEPTED MANUSCRIPT

To define the Gaussian kernel needed for the analysis we first introduce the backward deterministic differential system associated with (1.1). For fixed $T > 0$, $\mathbf{y} \in \mathbb{R}^{nd}$ and $t \in [0,T]$, we define:

$$\dot{\boldsymbol{\theta}}_{t,T}(\mathbf{y}) = \mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y})),\ \boldsymbol{\theta}_{T,T}(\mathbf{y}) = \mathbf{y}. \tag{3.2}$$

Remark 3.1. Observe from (3.2), (2.3) that $\boldsymbol{\theta}_{s,t}(\mathbf{x})$ is well defined for all $s,t \in [0,T], \mathbf{x} \in \mathbb{R}^{nd}$. The associated differential dynamics in s runs forward in time if $s \geq t$ and backward otherwise.

Consider now the deterministic ODE

$$\frac{d}{dt}\tilde{\boldsymbol{\phi}}_t = \mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y})) + D\mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y}))[\tilde{\boldsymbol{\phi}}_t - \boldsymbol{\theta}_{t,T}(\mathbf{y})], \quad t \geq 0, \tag{3.3}$$

where for all $\mathbf{x} \in \mathbb{R}^{nd}$,

$$D\mathbf{F}(t,\mathbf{x}) = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ D_{\mathbf{x}_1}\mathbf{F}_2(t,\mathbf{x}) & 0 & \cdots & \cdots & 0 \\ 0 & D_{\mathbf{x}_2}\mathbf{F}_3(t,\mathbf{x}) & 0 & 0 & \vdots \\ \vdots & 0 & \ddots & \vdots & \\ 0 & \cdots & 0 & D_{\mathbf{x}_{n-1}}\mathbf{F}_n(t,\mathbf{x}) & 0 \end{pmatrix}$$

denotes the subdiagonal of the Jacobian matrix $\mathbf{D}_{\mathbf{x}}\mathbf{F}$ at point $\mathbf{x}$.

Introduce now for a given $(T,\mathbf{y}) \in \mathbb{R}^{+*} \times \mathbb{R}^{nd}$, the resolvent $(\tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s))_{s,t\geq 0}$ associated with the partial gradients $(D\mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y})))_{t\geq 0}$ which satisfies for $(s,t) \in (\mathbb{R}^+)^2$:

$$\begin{aligned} &\partial_t \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s) = D\mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y}))\tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s),\ \tilde{\mathbf{R}}^{T,\mathbf{y}}(s,s) = I_{nd\times nd}, \\ &\partial_s \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s) = -\tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)D\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})),\ \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,t) = I_{nd\times nd}. \end{aligned} \tag{3.4}$$

Note in particular that since the partial gradients are subdiagonal $\det(\tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)) = 1$.

Setting as well for $0 \leq s,t \leq T$,

$$\tilde{\mathbf{m}}^{T,\mathbf{y}}(s,t) := \int_s^t \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,u)\big(\mathbf{F}(u, \boldsymbol{\theta}_{u,T}(\mathbf{y})) - D\mathbf{F}(u, \boldsymbol{\theta}_{u,T}(\mathbf{y}))\boldsymbol{\theta}_{u,T}(\mathbf{y})\big)du,$$

and denoting by $(\tilde{\boldsymbol{\theta}}^{T,\mathbf{y}}_{t,s})_{t,s\geq 0}$ the flow associated with (3.3), i.e. $\tilde{\boldsymbol{\theta}}^{T,\mathbf{y}}_{t,s}(\mathbf{x})$ is the value of $\tilde{\boldsymbol{\phi}}_t$ when $\tilde{\boldsymbol{\phi}}_s = \mathbf{x}$, we thus derive:

$$\begin{aligned} \tilde{\boldsymbol{\theta}}^{T,\mathbf{y}}_{t,s}(\mathbf{x}) &= \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)\mathbf{x} \\ &\quad + \int_s^t \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,u)\big(\mathbf{F}(u, \boldsymbol{\theta}_{u,T}(\mathbf{y})) - D\mathbf{F}(u, \boldsymbol{\theta}_{u,T}(\mathbf{y}))\boldsymbol{\theta}_{u,T}(\mathbf{y})\big)du \\ &= \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)\mathbf{x} + \tilde{\mathbf{m}}^{T,\mathbf{y}}(s,t). \end{aligned} \tag{3.5}$$

Note that the flow is affine.

We now introduce for all $0 \leq s < t,\ (\mathbf{x},\mathbf{y}) \in (\mathbb{R}^{nd})^2$ the kernel:

$$\begin{aligned} \tilde{q}(s,t,\mathbf{x},\mathbf{y}) &:= \frac{1}{(2\pi)^{nd/2}\det(\tilde{\mathbf{K}}^{\mathbf{y}}(s,t))^{1/2}} \\ &\times \exp\left(-\frac{1}{2}\langle \tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}) - \mathbf{y}), \tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}) - \mathbf{y}\rangle\right), \end{aligned} \tag{3.6}$$

ACCEPTED MANUSCRIPT

where $\tilde{\mathbf{K}}^{\mathbf{y}}(s,t) := \int_s^t \tilde{\mathbf{R}}^{t,\mathbf{y}}(t,u)B\varsigma(u)B^*\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,u)^*du$. In other words, under **(A)**, denoting by $\varsigma(u)^{1/2}$ the only subdiagonal matrix s.t. $\varsigma(u)^{1/2}[\varsigma(u)^{1/2}]^* = \varsigma(u)$, $\tilde{q}(s,t,\mathbf{x},\mathbf{y})$ is the density at time t and point $\mathbf{y}$ of the diffusion $(\tilde{\mathbf{X}}_u^{t,\mathbf{y}})_{u\in[s,t]}$ with dynamics:

$$d\tilde{\mathbf{X}}_u^{t,\mathbf{y}} = [\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y})) + D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}))(\tilde{\mathbf{X}}_u^{t,\mathbf{y}} - \boldsymbol{\theta}_{u,t}(\mathbf{y}))]du + B\varsigma(u)^{1/2}dW_u,$$
$$\forall u\in[s,t],\ \tilde{\mathbf{X}}_s^{t,\mathbf{y}} = \mathbf{x}. \tag{3.7}$$

Remark 3.2. We importantly point out that the kernel $\tilde{q}(s,t,\mathbf{x},\mathbf{y})$ introduced in (3.6) emphasizes a strong non-linear dependence in the $\mathbf{y}$ variable. Indeed, if the drift $F(t,\mathbf{x}) = A_t\mathbf{x}$ is linear in (1.1), the kernel $\tilde{q}$ in (3.6) rewrites:

$$\tilde{q}(s,t,\mathbf{x},\mathbf{y}) := \frac{1}{(2\pi)^{nd/2}\det(\mathbf{K}(s,t))^{1/2}}$$
$$\times\exp\left(-\frac{1}{2}\langle\mathbf{K}(s,t)^{-1}(\mathbf{R}(t,s)\mathbf{x}-\mathbf{y}),\mathbf{R}(t,s)\mathbf{x}-\mathbf{y}\rangle\right),$$

where $(\mathbf{R}(t,s))_{0\le s,t\le T}$ stands for the resolvent associated with $(A_t)_{t\in[0,T]}$ and $\mathbf{K}(s,t) := \int_s^t \mathbf{R}(t,u)B\varsigma(u)B^*\mathbf{R}(t,u)^*du$. In that case, denoting by $\mathbf{K}(s,t)^{-1/2}$ a square root of $\mathbf{K}(s,t)^{-1}$, the changes of variables $\mathbf{y}\mapsto\mathbf{K}(s,t)^{-1/2}(\mathbf{R}(t,s)\mathbf{x}-\mathbf{y})$ and $\mathbf{x}\mapsto\mathbf{K}(s,t)^{-1/2}(\mathbf{R}(t,s)\mathbf{x}-\mathbf{y})$ are linear in both variables $\mathbf{y}$ and $\mathbf{x}$. Such changes of variables turn out to be very useful in the study of singular integrals associated with the kernel $\tilde{q}$ (see [BCLP10] and [BCLP13]). In the general setting we consider, the mapping $\mathbf{x}\mapsto\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1/2}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})$ is still linear w.r.t. $\mathbf{x}$ but it is definitely not the case for $\mathbf{y}\mapsto\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1/2}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})$ for which the strong non-linear dependence in $\mathbf{y}$ appears in both $\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)$ and $\tilde{\boldsymbol{\theta}}_{s,t}^{t,\mathbf{y}}$. This fact will lead us to perform a specific analysis for the singular integrals that will be associated with this kernel. In particular, we cannot directly apply the results for the degenerate Gaussian kernels with linear drift established in [BCLP10], [BCLP13]. We still have degenerate linear/affine Gaussian kernels, but the linearization of the drift in (1.1) will depend on the integration variable $\mathbf{y}$, see the Green kernel $\tilde{G}f$ introduced in (3.12) below and Section 4.2 for details.

Assumption **(A)** also guarantees that the covariance matrix $(\tilde{\mathbf{K}}^{\mathbf{y}}(s,t))_{0\le s<t}$ satisfies *uniformly* in $\mathbf{y}\in\mathbb{R}^{nd}$ a *good scaling property* in the sense of Definition 3.2 in [DM10] (see also Proposition 3.4 of that reference). That is: for all fixed $T>0$, there exists $C_{3.8} := C_{3.8}(T,(\mathbf{A}))\ge 1$ s.t. for all $0\le s<t\le T$, for all $\mathbf{y}\in\mathbb{R}^{nd}$:

$$\forall\boldsymbol{\xi}\in\mathbb{R}^{nd},\ C_{3.8}^{-1}(t-s)^{-1}|\mathbb{T}_{t-s}\boldsymbol{\xi}|^2 \le \langle\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)\boldsymbol{\xi},\boldsymbol{\xi}\rangle \le C_{3.8}(t-s)^{-1}|\mathbb{T}_{t-s}\boldsymbol{\xi}|^2, \tag{3.8}$$

where for all $t>0$, $\mathbb{T}_t = \mathrm{diag}((t^iI_d)_{i\in[\![1,n]\!]})$ is a scale matrix. As pointed out in the introduction, equation (3.8) indicates that the i^{th} component of (3.7) has characteristic time scale of order $(2i-1)/2$.

From (3.6) and (3.8), we directly derive that for all $T>0$ there exists $C_{3.9} := C_{3.9}(T,(\mathbf{A}))\ge 1$ s.t. for all $0\le s<t\le T$, $(\mathbf{x},\mathbf{y})\in(\mathbb{R}^{nd})^2$:

$$C_{3.9}^{-1}(t-s)^{-n^2d/2}\exp(-C_{3.9}(t-s)|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|^2) \le \tilde{q}(s,t,\mathbf{x},\mathbf{y})$$
$$\le C_{3.9}(t-s)^{-n^2d/2}\exp(-C_{3.9}^{-1}(t-s)|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|^2). \tag{3.9}$$

Now, Lemma 5.3 and Equation (5.11) from the proof of Lemma 5.5 in [DM10] (see also Sections 4.2 and 6.2 below for details) give that there exists $C_{3.10} :=$

ACCEPTED MANUSCRIPT

$C_{3.10}(T,(\mathbf{A}))\geq 1$ s.t.:

$$\begin{aligned}
&C_{3.10}^{-1}|\mathbb{T}_{t-s}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|\leq|\mathbb{T}_{t-s}^{-1}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|\leq C_{3.10}|\mathbb{T}_{t-s}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|,\\
&C_{3.10}^{-1}|\mathbb{T}_{t-s}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|\leq|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|\leq C_{3.10}|\mathbb{T}_{t-s}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|,\\
&|D_{\mathbf{x}_j}\tilde{q}(s,t,\mathbf{x},\mathbf{y})|\leq C_{3.10}(t-s)^{-j+1}|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|\tilde{q}(s,t,\mathbf{x},\mathbf{y}),\ j\in[\![1,n]\!].
\end{aligned}\tag{3.10}$$

On the other hand, it is crucial to observe that $\tilde{q}(s,t,\mathbf{x},\mathbf{y})$ satisfies the following Backward Kolmogorov equation for all $(t,\mathbf{y})\in\mathbb{R}^{+*}\times\mathbb{R}^{nd}$:

$$\left(\partial_s+\tilde{L}_s^{t,\mathbf{y}}\right)\tilde{q}(s,t,\mathbf{x},\mathbf{y})=0,\ (s,\mathbf{x})\in[0,t)\times\mathbb{R}^{nd},\ \tilde{q}(s,t,\mathbf{x},\cdot)\underset{t\downarrow s}{\longrightarrow}\delta_{\mathbf{x}}(.).\tag{3.11}$$

In the above equation we wrote:

$$\begin{aligned}
\tilde{L}_s^{t,\mathbf{y}}\tilde{q}(s,t,\mathbf{x},\mathbf{y}) &:= \langle\mathbf{F}(s,\boldsymbol{\theta}_{s,t}(\mathbf{y}))+D\mathbf{F}(s,\boldsymbol{\theta}_{s,t}(\mathbf{y}))(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y})),\mathbf{D}_{\mathbf{x}}\tilde{q}(s,t,\mathbf{x},\mathbf{y})\rangle\\
&\quad+\frac{1}{2}\mathrm{tr}(\varsigma(s)D_{\mathbf{x}_1}^2\tilde{q}(s,t,\mathbf{x},\mathbf{y})),
\end{aligned}$$

i.e. $\tilde{L}_s^{t,\mathbf{y}}$ is the generator of $(\tilde{\mathbf{X}}_u^{t,\mathbf{y}})_{u\in[s,t]}$ defined in (3.7) at time s. Also, the convergence towards the Dirac mass in (3.11) follows from Lemma 2.4 in [Men11].

For the rest of the section we assume w.l.o.g. that $T\leq 1$. For $0\leq s<T$ and a function $f\in C_0^\infty([0,T)\times\mathbb{R}^{nd})$ we now define for all $\mathbf{x}\in\mathbb{R}^{nd}$:

$$\tilde{G}f(s,\mathbf{x}):=\int_s^T dt\int_{\mathbb{R}^{nd}}\tilde{q}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})d\mathbf{y}.\tag{3.12}$$

From (3.11) one easily gets that

$$\partial_s\tilde{G}f(s,\mathbf{x})+\tilde{M}f(s,\mathbf{x})=-f(s,\mathbf{x}),\ (s,\mathbf{x})\in[0,T)\times\mathbb{R}^{nd},\ \tilde{G}f(s,\cdot)\underset{s\uparrow T}{\longrightarrow}0,$$

with $\tilde{M}f(s,\mathbf{x}):=\int_s^T dt\int_{\mathbb{R}^{nd}}d\mathbf{y}\tilde{L}_s^{t,\mathbf{y}}\tilde{q}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})$. Observe that, because of the strong non-linear dependence of $\tilde{q}(s,t,\mathbf{x},\mathbf{y})$ on the variable $\mathbf{y}$ (see Remark 3.2), $\tilde{M}f(s,\mathbf{x})$ is not simply a differential operator (in the $\mathbf{x}$ variable) applied to the Green kernel (3.12). Again, the linearization procedure does depend on the integration variable.

We now write,

$$\partial_s\tilde{G}f(s,\mathbf{x})+L_s\tilde{G}f(s,\mathbf{x})=(-f+Rf)(s,\mathbf{x}),\ (s,\mathbf{x})\in[0,T)\times\mathbb{R}^{nd},$$

where $Rf(s,\mathbf{x}):=(L_s\tilde{G}f-\tilde{M}f)(s,\mathbf{x})=\int_s^T dt\int_{\mathbb{R}^{nd}}d\mathbf{y}(L_s-\tilde{L}_s^{t,\mathbf{y}})\tilde{q}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})$.

Now, the local condition (3.1) yields:

$$|Rf(s,\mathbf{x})|\leq|Nf(s,\mathbf{x})|+\sum_{i=2}^n|D_{\mathbf{x}_i}R_if(s,\mathbf{x})|+\frac{\varepsilon_a}{2}|D_{\mathbf{x}_1}^2\tilde{G}f(s,\mathbf{x})|,\tag{3.13}$$

where setting

$$\begin{aligned}
\mathbf{F}^{t,\mathbf{y}}(s,\mathbf{x}) &:= \big(\mathbf{F}_1(s,\boldsymbol{\theta}_{s,t}(\mathbf{y})),\mathbf{F}_2(s,\mathbf{x}_1,(\boldsymbol{\theta}_{s,t}(\mathbf{y}))^{2,n}),\mathbf{F}_3(s,\mathbf{x}_2,(\boldsymbol{\theta}_{s,t}(\mathbf{y}))^{3,n}),\cdots,\\
&\quad\mathbf{F}_n(s,\mathbf{x}_{n-1},(\boldsymbol{\theta}_{s,t}(\mathbf{y}))_n)\big),\\
Nf(s,\mathbf{x}) &:= \sum_{i=1}^n\int_s^T dt\int_{\mathbb{R}^{nd}}d\mathbf{y}\langle\big((\mathbf{F}-\mathbf{F}^{t,\mathbf{y}})(s,\mathbf{x})\big)_i,\mathbf{D}_{\mathbf{x}_i}\tilde{q}(s,t,\mathbf{x},\mathbf{y})\rangle f(t,\mathbf{y}),
\end{aligned}\tag{3.14}$$

ACCEPTED MANUSCRIPT

and for all $i \in [\![2,n]\!]$,

$$
\begin{aligned}
R_i f(s,\mathbf{x}) \quad := \quad & \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{q}(s,t,\mathbf{x},\mathbf{y}) f(t,\mathbf{y}) \big\{ \mathbf{F}_i^{t,\mathbf{y}}(s,\mathbf{x}) \\
& - \big[\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{y})) + D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{y}))(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_{i-1} \big] \big\}. \qquad (3.15)
\end{aligned}
$$

Remark 3.3. Observe from the above equation that if, for all $i \in [\![2,n]\!]$, the function $\mathbf{F}_i$ is linear w.r.t. to the $(i-1)^{\text{th}}$ variable (component that transmits the noise), then for all $(s,\mathbf{x}) \in [0,T) \times \mathbb{R}^{nd}$, $R_i f(s,\mathbf{x}) = 0$.

The terms Nf in (3.14) and $(D_{\mathbf{x}_i} R_i f)_{i \in [\![2,n]\!]}$ in (3.15) do not have time singularities. Let us justify this point.

Using (3.10), (3.9), we derive from (3.14) that:

$$
\begin{aligned}
|Nf(s,\mathbf{x})| &\leq C \sum_{i=1}^n \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} |(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))^{i,n}| |D_{\mathbf{x}_i} \tilde{q}(s,t,\mathbf{x},\mathbf{y})| |f(t,\mathbf{y})| \\
&\leq C \sum_{i=1}^n \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\{ \frac{|(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))^{i,n}|}{(t-s)^{(2i-1)/2}} \right\} (t-s)^{1/2} |\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x}) - \mathbf{y})| \\
&\qquad \times \tilde{q}(s,t,\mathbf{x},\mathbf{y}) |f(t,\mathbf{y})| \\
&\leq C \int_s^T dt \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t-s)^{n^2 d/2}} \exp(-C^{-1}(t-s) |\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x}) - \mathbf{y})|^2) |f(t,\mathbf{y})|,
\end{aligned}
$$

for $C := C((\mathbf{A}))$. Now, as a consequence of Hölder's inequality we derive that for all $p > 1$, $p^{-1} + q^{-1} = 1$:

$$
\begin{aligned}
|Nf(s,\mathbf{x})|^p \leq C(p,(\mathbf{A})) T^{p/q} \int_s^T dt \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t-s)^{n^2 d/2}} \\
\exp(-C^{-1}(t-s) |\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x}) - \mathbf{y})|^2) |f(t,\mathbf{y})|^p.
\end{aligned}
$$

The Fubini Theorem and (3.10) then yields:

$$
\|Nf\|_{L^p([0,T)\times\mathbb{R}^{nd})} \leq C(p,(\mathbf{A})) T \|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}. \qquad (3.16)
$$

From (3.9), (3.10) we also derive from (3.15) that for all $i \in [\![2,n]\!]$:

$$
\begin{aligned}
|D_{\mathbf{x}_i} R_i f(s,\mathbf{x})| \quad \leq \quad & C \int_s^T dt \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t-s)^{n^2 d/2}} (t-s)^{-i+1/2} |(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_{i-1}|^{1+\eta} \\
& \times \exp\left(-C^{-1}(t-s) |\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x}) - \mathbf{y})|^2 \right) |f(t,\mathbf{y})| \\
\leq \quad & C \int_s^T dt (t-s)^{-1+(i-\frac{3}{2})\eta} \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t-s)^{n^2 d/2}} |f(t,\mathbf{y})| \\
& \times \exp\left(-C^{-1}(t-s) |\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x}) - \mathbf{y})|^2 \right),
\end{aligned}
$$

ACCEPTED MANUSCRIPT

where $\eta \in (0,1]$ stands for the Hölder continuity index of $D_{\mathbf{x}_{i-1}}\mathbf{F}_i$, see Assumption **(S)**. Hence, by the Hölder inequality and for $p>2,\ p^{-1}+q^{-1}=1$:

$$|D_{\mathbf{x}_i}R_if(s,\mathbf{x})|^p \le C(p,(\mathbf{A}))\left(\int_s^T \frac{dt}{(t-s)^{1-\frac{\eta}{2}}}\right)^{p/q}\left(\int_s^T \frac{dt}{(t-s)^{1-\frac{\eta}{2}}}\right.$$
$$\left.\times \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t-s)^{n^2d/2}} \exp\left(-C^{-1}(t-s)|\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})|^2\right)|f(t,\mathbf{y})|^p\right),$$

$$\|D_{\mathbf{x}_i}R_if\|_{L^p([0,T)\times\mathbb{R}^{nd})} \le C(p,(\mathbf{A}))T^{\eta/2}\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}, \tag{3.17}$$

where the last control again follows from Fubini's theorem.

Now, the key tool to prove uniqueness for the martingale problem derives from the following Calderón and Zygmund type estimate for the Green function $\tilde{G}f$. Namely, we have the following theorem which is proved in Section 4.3.

Theorem 3.1. Assume that Assumption **(A)** is in force. Suppose also that $T \in (0,T_0]$, $T_0 := T_0((\mathbf{A})) \le 1$. Then, for all $p \in (1,+\infty)$ there exists $C_{3.18} := C_{3.18}(T_0,p,(\mathbf{A}))$ s.t. for all $f \in L^p([0,T)\times\mathbb{R}^{nd})$,

$$\|D^2_{\mathbf{x}_1}\tilde{G}f\|_{L^p([0,T)\times\mathbb{R}^{nd})} \le C_{3.18}\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}, \tag{3.18}$$

where the Green function $\tilde{G}f$ is defined in (3.12) with the kernel $\tilde{q}$ introduced in (3.6).

Remark 3.4. Let us specify that the *small time* condition appearing here is due to the fact that we are led to compare the flow $\boldsymbol{\theta}$ and its *linearization* appearing in the considered kernel $\tilde{q}$ of the Green function $\tilde{G}f$ (see equations (3.6) and (3.12)). It is clear that this procedure can be a good approximation in small time only.

Hence, from (3.18), (3.17) and (3.16), integrating in (3.13) we derive that under **(A)**, with the notations of (3.1), for $p>2$,

$$\|Rf\|_{L^p([0,T)\times\mathbb{R}^{nd})} \le (\frac{\varepsilon_a}{2}C_{3.18}+C(p,(\mathbf{A}))T^{\frac{\eta}{2}})\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}. \tag{3.19}$$

Thus, for $\varepsilon_a < C^{-1}_{3.18}$ and $T < (4C(p,(\mathbf{A})))^{-2/\eta}$, the operator $I-R$ admits a bounded inverse on $L^p([0,T)\times\mathbb{R}^{nd})$, and formally $Gf(s,\mathbf{x}) := \tilde{G}\circ(I-R)^{-1}f(s,\mathbf{x})$, $(s,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$ solves the Cauchy problem:

$$\begin{cases}(\partial_t+L_t)u(t,\mathbf{x}) = -f(t,\mathbf{x}), (t,\mathbf{x}) \in [0,T)\times\mathbb{R}^{nd},\\ u(T,\mathbf{x})=0,\end{cases}$$

for $f \in L^p([0,T)\times\mathbb{R}^{nd})$, $p > (n^2d+2)/2$. This last condition on p is needed to give a pointwise sense to Gf. Observe indeed from Hölder's inequality and the upper-bound in (3.9) that, for all $p > (n^2d+2)/2$, there exists $C_{3.20} := C_{3.20}(p,(\mathbf{A}))$, s.t. for all $f \in L^p([0,T)\times\mathbb{R}^{nd})$, $(s,\mathbf{x}) \in [0,T)\times\mathbb{R}^{nd}$,

$$|\tilde{G}f(s,\mathbf{x})| \le C_{3.20}T^{1-(2+n^2d)/(2p)}\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})}. \tag{3.20}$$

The above control is an extension of Lemma 7.1.1. in [SV79] in the non-degenerate case. From the probabilistic viewpoint we will prove that there is only one probability $\mathbb{P}$ on $C([0,T],\mathbb{R}^{nd})$ solving the martingale problem and therefore derive $Gf(s,\mathbf{x}) = \mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}[\int_s^T f(t,\mathbf{X}_t)dt]$, $(\mathbf{X}_t)_{t\in[0,T]}$ being the canonical process. A localization argument similar to the one in Priola [Pri15] then allows to extend the well

ACCEPTED MANUSCRIPT

posedness of the martingale problem under the sole continuity assumption **(C)** in **(A)**, i.e. without the local condition (3.1), see Section 5.

Remark 3.5. Let us mention that we considered the condition (3.1) to derive our main estimate in Theorem 2.1 through probabilistic localization arguments, see Section 5. However, from a more analytical viewpoint, we could obtain from the proof of Theorem 3.1 estimates similar to those in (1.2), replacing therein the operator $\mathcal{A}$ by the L defined in (2.1), provided the diffusion coefficient is uniformly continuous. This would extend the results in [BCLP13] to the case of a non-homogeneous and non-linear first order vector field.

Remark 3.6 (Some Points about the Drift). One can wonder if the assumptions on the drift $\mathbf{F}$ can be weakened in order to conserve the global well-posedness of the martingale problem. To answer this query one needs to consider separately $\mathbf{F}_1$, associated with the non degenerate component, and the $(\mathbf{F}_i)_{i\in[\![2,N]\!]}$, associated with the degenerate ones.

- For $\mathbf{F}_1$ the assumptions in **(S)** can be considerably weakened. Indeed if $\mathbf{F}_1 \in L^p([0,T]\times\mathbb{R}^{nd}), p > n^2d+2$ then the martingale problem is still well posed provided the $(\mathbf{F}_i)_{i\in[\![2,n]\!]}$ satisfy **(S)**. To see this, instead of (3.2), (3.3), we consider the following dynamics to define the underlying Gaussian kernel:

$$\dot{\bar{\boldsymbol{\theta}}}_{t,T}(\mathbf{y}) = \bar{\mathbf{F}}(t,\bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})),\ \bar{\boldsymbol{\theta}}_{T,T}(\mathbf{y}) = \mathbf{y},$$

where for all $\mathbf{x}\in\mathbb{R}^{nd}$, $\bar{\mathbf{F}}(t,\mathbf{x}) := (\mathbf{0},\mathbf{F}_2(t,\mathbf{x}),\cdots,\mathbf{F}_n(t,\mathbf{x}))$ and

$$\frac{d}{dt}\tilde{\bar{\boldsymbol{\phi}}}_t = \bar{\mathbf{F}}(t,\bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})) + D\mathbf{F}(t,\bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y}))[\tilde{\bar{\boldsymbol{\phi}}}_t - \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})],\quad t\geq 0,$$

i.e. we put the non degenerate drift to 0 in our proxy model. This would yield in (3.13), (3.15) the additional contribution

$$\begin{aligned}|D_{\mathbf{x}_1}R_1f(s,\mathbf{x})| &:= |\int_s^T dt\int_{\mathbb{R}^{nd}} d\mathbf{y}\langle\mathbf{F}_1(s,\mathbf{x}),D_{\mathbf{x}_1}\tilde{q}(s,t,\mathbf{x},\mathbf{y})\rangle f(t,\mathbf{y})|\\ &\leq C|\mathbf{F}_1(s,\mathbf{x})|\int_s^T dt\int_{\mathbb{R}^{nd}}\frac{d\mathbf{y}}{(t-s)^{(n^2d+1)/2}}\\ &\times\exp\left(-C^{-1}(t-s)|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|^2\right)|f(t,\mathbf{y})|,\end{aligned}\tag{3.21}$$

exploiting (3.10) for the last inequality. We get the same time-singularity as in the non-degenerate case. Thus, if $f\in L^p([0,T)\times\mathbb{R}^{nd})$:

$$\begin{aligned}&\|D_{\mathbf{x}_1}R_1f\|^p_{L^p([0,T)\times\mathbb{R}^{nd})}\\ \leq\int_0^T ds\int_{\mathbb{R}^{nd}} d\mathbf{x}|\mathbf{F}_1(s,\mathbf{x})|^p\|f\|^p_{L^p([0,T)\times\mathbb{R}^{nd})}&(\int_s^T dt(t-s)^{-(n^2d/2(q-1)+q/2)})^{p/q},\\ &\|D_{\mathbf{x}_1}R_1f\|_{L^p([0,T)\times\mathbb{R}^{nd})}\leq\|\mathbf{F}_1\|_{L^p([0,T)\times\mathbb{R}^{nd})}\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}T^\beta,\end{aligned}$$

where $\beta := \frac{1}{2}(1-\frac{2+n^2d}{p}) > 0$ under the previous condition on p. Adding this contribution in (3.19), we derive that the operator inversion can still be performed provided T is small enough. In the non-degenerate case, we refer to the work of Krylov and Röckner [KR05], or Fedrizzi and Flandoli [FF11] for an alternative proof, for additional results concerning strong solvability for drifts in L^p-spaces. Let us also mention that, as in the non-degenerate case, a bounded measurable drift

ACCEPTED MANUSCRIPT

$\mathbf{F}_1$ does not alter the well posedness of the martingale problem. This can be seen from (3.21) similarly to the previous computations for the terms $(D_{\mathbf{x}_{i-1}} R_i f)_{i\in[\![2,n]\!]}$.

- For $(\mathbf{F}_i)_{i\in[\![2,n]\!]}$ the situation is more complicated. To have as "proxy" a Gaussian process that satisfies in the whole space the *good scaling property* (3.8), it seems rather natural to impose that $(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i\in[\![2,n]\!]}$ are pointwise defined and non degenerate (Hörmander like assumption). In the current framework, a natural question consists in relaxing the Hölder continuity of the $(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i\in[\![2,n]\!]}$. Assuming simply continuity on those functions would again lead to consider a singular integral operator. Observe indeed from the previous computations that for all $i \in [\![2,n]\!]$,

$$|\mathbf{F}_i^{t,\mathbf{y}}(s,\mathbf{x}) - \{\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{y})) + D_{\mathbf{x}_{i-1}}\mathbf{F}_i(s,\mathbf{x})(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))_{i-1}\}||D_{\mathbf{x}_i}\tilde{q}(s,t,\mathbf{x},\mathbf{y})| \leq \frac{C}{(t-s)^{n^2d/2+1}}\exp\left(C^{-1}(t-s)|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|^2\right),$$

which is the expected time singularity for the convolution kernel of a singular integral operator. Anyhow, it would be in this case rather delicate to establish the *cancellation* property needed to complete the analysis, see also Proposition 4.2 for the properties required on a Calderón-Zygmund kernel. The η-Hölder continuity of the $(\mathbf{F}_i)_{i\in[\![2,n]\!]}$ is a sufficient condition to *globally* get rid of the time singularity (see again (3.17)).

Let us eventually mention that for $n = 2$, under **(UE)**, **(ND)**, when σ is Lipschitz continuous, $\mathbf{F}$ is Lipschitz in $\mathbf{x}_1$ and $D_{\mathbf{x}_1}\mathbf{F}_2$ is Hölder continuous, strong uniqueness has been established for (1.1) by Chaudru de Raynal [Ray14] provided $\mathbf{F}_1, \mathbf{F}_2$ are η-Hölder continuous in $\mathbf{x}_2$ with $\eta > 2/3$. Such results have been extended to the case of Dini-continuous drifts with the same critical exponents and non-Lipschitz diffusion coefficients by Wang and Zhang [WZ16]. It has also been recently emphasized by Chaudru de Raynal [Ray16] that, in his previous framework of Hölder continuous coefficients, the weak well posedness holds for $\eta > 1/3$. A counter-example is also provided emphasizing that weak uniqueness might fail if $\eta < 1/3$. Eventually, recent works by Fedrizzi *et al.* [FFPV16] and Zhang [Zha16] address the problem of strong/weak well posedness when $n = 2$, in the linear case for the degenerate variable, but for a large class of "rough" drifts in the non-degenerate dynamics.

4. Derivation of the Calderón-Zygmund estimates

We assume **(A)** is in force and that $T \leq T_0((\mathbf{A})) \leq 1$.

4.1. **Quasi Metric Structure and Covering.** To derive Theorem 3.1, a crucial step consists in considering a "good" parabolic metric and in taking into account the unbounded transport term in (1.1). In order to take into consideration our various time-scales, associated with the propagation of the noise into the system, we introduce the following metric:

$$\forall (t,\mathbf{x}) \in \mathbb{R}\times\mathbb{R}^{nd},\ \rho(t,\mathbf{x}) := |t|^{1/2} + \sum_{i=1}^{n}|\mathbf{x}_i|^{1/(2i-1)}. \tag{4.1}$$

Remark 4.1. Recalling the definition of the scale matrix $\mathbb{T}_t := \mathrm{diag}((t^i I_d)_{i\in[\![1,n]\!]})$, $t \geq 0$, we can now observe that $\mathbf{x} \in \mathbb{R}^{nd} \mapsto \rho(t, t^{-1/2}\mathbb{T}_t\mathbf{x})$ is $1/2$ homogeneous in the time variable, i.e. $\rho(t, t^{-1/2}\mathbb{T}_t\mathbf{x}) = t^{1/2}\rho(1,\mathbf{x})$.

ACCEPTED MANUSCRIPT

The metric introduced in (4.1) is similar to the one appearing in [BCLP10], [BCLP13] for L^p regularity.

Introducing now the strip $S := [-T,T]\times\mathbb{R}^{nd}$, we then define for $\big((s,\mathbf{x}),(t,\mathbf{y})\big) \in S^2$ the quasi-distances:

$$\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) := \rho(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}),$$
$$\mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y})) := \mathbf{d}((t,\mathbf{y}),(s,\mathbf{x})) = \rho(t-s,\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y})), \tag{4.2}$$

with ρ as in (4.1). We now define the "balls" associated with the quasi-metric $\mathbf{d}$ ($\mathbf{d}$-balls) in the following way:

$$\forall (s,\mathbf{x})\in S,\ \forall \delta>0,$$
$$B((s,\mathbf{x}),\delta) := \{(t,\mathbf{y})\in S : \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))\le\delta\}. \tag{4.3}$$

We mention that the natural extension of the balls considered in [BCLP10], [BCLP13] would have been to consider $\mathbf{d}^*$ in the above definition. For this choice, in the linear, homogeneous case $\boldsymbol{\theta}_{s,t}(\mathbf{y}) := \mathbf{R}_{s-t}\mathbf{y}$, $\mathbf{R}$ standing for the resolvent of the linear differential system deriving from (1.1), which can indeed be seen as a group action. We choose here to follow the characteristic associated with the center of the ball, considering a *metric tube* around it. Anyhow those choices are very close and locally equivalent, see Proposition 4.1.

There is now, as in the previously mentioned works, a double difficulty. First, the quasi-distance used to define the balls satisfies the quasi-triangle inequality only locally. A natural choice would then consist in considering singular integrals for the "homogeneous space" associated with the balls of the above form. But here comes the second difficulty. In that case, it is not clear that such balls, seen as homogeneous spaces, enjoy the doubling property, which is however satisfied on the whole strip $S = [-T,T]\times\mathbb{R}^{nd}$.

The first key-point is the following result.

Proposition 4.1. Let $S := [-T,T]\times\mathbb{R}^{nd}$. The space $(S,\mathbf{d},dtd\mathbf{x})$ is a locally invariant quasi-metric space in the following sense: for a given $\Lambda\in(0,1]$ there exists a constant $C_{4.1} := C_{4.1}((\mathbf{A}),T,\Lambda) > 0$ s.t.

a) For all $(s,\mathbf{x}),(t,\mathbf{y})\in S$, if $\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))\le\Lambda$ then

$$\mathbf{d}((t,\mathbf{y}),(s,\mathbf{x})) \le C_{4.1}\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) = \rho(|t-s|,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})),$$

and for $(\sigma,\boldsymbol{\xi})\in S$ s.t. $\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))\le\Lambda$ and $\mathbf{d}((t,\mathbf{y}),(\sigma,\boldsymbol{\xi}))\le\Lambda$ then

$$\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) \le C_{4.1}(\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})) + \mathbf{d}((\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))).$$

b) Every $\mathbf{d}$-ball in the sense of (4.3) has positive and finite measure and every non-empty intersection of two balls has positive measure.

c) There exists $R>0$ s.t. for $0<R_1<R_2\le R$ there exists $C := C(R_1,R_2)$ s.t. for all $(s,\mathbf{x})\in S$,

$$|B((s,\mathbf{x}),R_2)| \le C|B((s,\mathbf{x}),R_1)|,$$

where $|.|$ stands here for the Lebesgue measure of the balls.

Remark 4.2 (General and Subdiagonal structure). Let us stress, as it will appear from the proof of Proposition 4.1 in Section 6.1, that for the general form of $\mathbf{F}$ in the dynamics of $\boldsymbol{\theta}$, the constant $C_{4.1}$ appearing in Proposition 4.1 depends on the specific radius, here $\Lambda\le 1$, chosen for the balls. In the following, we assume Λ "small enough" and refer to Section 6.3 for a specific discussion on the choice of $\Lambda := \Lambda((\mathbf{A}))$.

ACCEPTED MANUSCRIPT

However, the proof also emphasizes that when the function $\mathbf{F}$ has the following structure, $\mathbf{F}_1(t,\mathbf{x}) = \mathbf{F}_1(t,\mathbf{x}_1)$, $\forall i \in [\![2,n]\!]$, $\mathbf{F}_i(t,\mathbf{x}^{i-1,n}) = \mathbf{F}_i(t,\mathbf{x}_{i-1},\mathbf{x}_i)$ (subdiagonal case), then the constant $C_{4.1}$ does not depend on the radius (see Remark 6.1). Hence, in this latter case, point a) of the proposition gives that the quasi-distances $\mathbf{d}$ and $\mathbf{d}^*$ involving respectively the forward and backward transport are actually equivalent. In such a case $\mathbf{d}$ is a usual quasi-distance in the sense of Coifman and Weiss [CW71] and the strip S can be seen as a homogeneous space.

From Proposition 4.1 we can use Theorem 25 in [BCLP10] that we now state in our specific case. It gives that the strip S can be covered by metric balls.

Theorem 4.1 (Covering Theorem). For every $\delta_0 > 0$ and $K > 1$ there exists $\delta \in (0,\delta_0)$, a positive integer M and a countable set $\big((s_i,\mathbf{x}_i)\big)_{i\in A} \subset S$ s.t.

1. $S = \bigcup_{i\in A} B((s_i,\mathbf{x}_i),\delta)$.
2. $\sum_{i\in A} \mathbb{I}_{B((s_i,\mathbf{x}_i),K\delta)} \le M^2$.

4.2. **Singular kernel and associated estimates.** Let us first define for $s>0$, $\varsigma(-s) := \varsigma(s)$, i.e. we symmetrize the diffusion coefficient. Fix now $T>0$ and introduce:

$$\forall (s,t,\mathbf{x},\mathbf{y}) \in \mathbb{R}^2 \times (\mathbb{R}^{nd})^2, k\big(s,t,\mathbf{x},\mathbf{y}\big) := \mathbb{I}_{t>s} D^2_{\mathbf{x}_1} \tilde{q}(s,t,\mathbf{x},\mathbf{y}),$$

with $\tilde{q}$ defined in (3.6). We recall from Remark 3.2 that $\tilde{q}$ has strong a non-linear dependence in the parameter $\mathbf{y}$ that serves as integration variable in the Green kernel 3.12.

From (3.5) and (3.6) a direct computation yields (see also the proof of Lemma 5.5 in [DM10]):

$$\begin{aligned} k(s,t,\mathbf{x},\mathbf{y}) &= \mathbb{I}_{t>s}\Big(-[\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)^*\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)]_{11} \\ &\quad +[\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)^*\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})]_1^{\otimes 2}\Big)\,\tilde{q}(s,t,\mathbf{x},\mathbf{y}). \end{aligned} \tag{4.4}$$

In the above equation, for a matrix $\mathbf{M} \in \mathbb{R}^{nd}\otimes\mathbb{R}^{nd}$ (resp. a vector $\mathbf{z}\in\mathbb{R}^{nd}$), the notation $[\mathbf{M}]_{11}$ stands for the $d\times d$ submatrix $(\mathbf{M}_{ij})_{(i,j)\in[\![1,d]\!]}$ (resp. $[\mathbf{z}]_1$ stands for the subvector of $\mathbb{R}^d$, $(\mathbf{z}_i)_{i\in[\![1,d]\!]}$).

From (3.6), (3.8) and the scaling Lemma 6.2 below (see also equations (5.10), (5.11) in [DM10]), we have that there exists $C := C(T,(\mathbf{A}))$ s.t.:

$$|[\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)^*\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})]_i| \le C\left((t-s)^{-i+1}|\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})|\right),$$

$$\begin{aligned} &|[\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)^*\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)]_{11} + [\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)^*\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})]_1^{\otimes 2}| \\ &\qquad\qquad \le C((t-s)^{-1}|I_d| + |\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})|^2), \end{aligned}$$

so that (3.8), (4.4) yield that $\exists (c_{4.5},C_{4.5}) := (c_{4.5},C_{4.5})(T,(\mathbf{A}))$ s.t.

$$|k(s,t,\mathbf{x},\mathbf{y})| \le C_{4.5}\mathbb{I}_{t>s}(t-s)^{-1}q_{c_{4.5}}(s,t,\mathbf{x},\mathbf{y}), \tag{4.5}$$

where for all $c>0$,

$$q_c(s,t,\mathbf{x},\mathbf{y}) := \frac{c^{nd/2}}{(2\pi)^{nd/2}(t-s)^{n^2d/2}}\exp\left(-\frac{c}{2}(t-s)|\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})|^2\right).$$

Observe that this is the same order of singularity than in the non-degenerate case. This is anyhow expectable since we are considering the derivatives w.r.t. the non-degenerate variables. From equation (4.5), we get that for all $\epsilon\in(0,1)$, $\forall (i,j)\in$

ACCEPTED MANUSCRIPT

$[\![1,d]\!]^2,\ f\in L^\infty(\mathbb{R}\times\mathbb{R}^{nd}),\ (s,\mathbf{x})\in\mathbb{R}^{1+nd}$,

$$K_{ij}^\epsilon f(s,\mathbf{x}):=\int_{S\cap\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon}dtd\mathbf{y}k_{ij}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y}),\ \forall(i,j)\in[\![1,d]\!]^2, \tag{4.6}$$

is well defined.

Remark 4.3. To establish L^p estimates for $K_{ij}^\epsilon f$, which in turn prove Theorem 3.1, there are two main differences with the articles [BCLP10], [BCLP13]. The first difference is that the distance $d\big((s,\mathbf{x}),(t,\mathbf{y}))\big)$ defining the integration domain in (4.6) involves the non-linear flow associated with the drift in equation (1.1). This choice is rather natural to have *good* properties on the underlying quasi-distance (see Proposition 4.1 and its proof in Section 6). Secondly, the singular kernel of K_{ij}^ϵ corresponds to the second order spatial derivatives w.r.t. non-degenerate variables of the linearized Gaussian kernel introduced in (3.6). Again, the dependence of the linearized kernel in the integration variable $\mathbf{y}$ is strongly non-linear (see Remark 3.2). In [BCLP10], [BCLP13], the flow associated with the drift in (1.1) does not depend on the integration variable $\mathbf{y}$. Hence, to use directly the results of those works, it would be necessary to linearize w.r.t. an auxiliary fixed given spatial point which is not the current integration variable and to freeze as well the coefficients in time, since the above works deal with time-homogeneous coefficients.

A possibility would be to consider a kernel deriving from a *forward* parametrix like expansion (as opposed to the current approach deriving from the *backward* parametrix developed in [DM10]) linearizing around the starting point. This strategy would lead to reinvestigate the controls concerning the drift, estimates for the terms $Nf,(D_{\mathbf{x}_i}R_if)_{i\in[\![2,n]\!]}$ introduced in Section 3 (see equations (3.14), (3.15) for the definition of those terms and (3.16), (3.17) for the corresponding controls). It would as well require some additional smoothness of the drift in the time variable, as it is already the case in the non-degenerate setting (see [IKO62], [Fri64]). Such a choice would also be likely to make the localization procedure even trickier than the current one.

Hence, it does not seem to be direct to adapt the approach of [BCLP10], [BCLP13] to the current framework. With our approach, the linearization of the drift around the "backward" deterministic flow allows to control the drift terms almost as in the previous work [DM10], [Men11]. The counterpart is that we have to precisely handle the linearization procedure in the related singular integral estimates. It must be carefully checked that, in short time, the linearized flow allows to obtain the required estimates in [Bra10] for the kernel we consider.

4.3. **Proof of Theorem 3.1.** The first thing to do consists in splitting the kernel into a singular and a non singular part observing that the singularity is *diagonal*. Specifically, for a given fixed $\delta>0$ there exist $(c,C):=(c,C)(T,(\mathbf{A}),\delta)$ s.t. if for $\big((s,\mathbf{x}),(t,\mathbf{y})\big)\in S^2,\ \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))=\rho(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})\geq\delta$, then from (4.5) we have:

$$|k(s,t,\mathbf{x},\mathbf{y})|\mathbb{I}_{\rho(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})\geq\delta}\leq Cq_c(s,t,\mathbf{x},\mathbf{y}). \tag{4.7}$$

Indeed, from the definitions in (4.2), we have either $|t-s|^{1/2}\geq\delta/(n+1)$ or that there exists $i\in[\![1,n]\!]$ s.t. $|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_i|^{1/(2i-1)}\geq\delta/(n+1)$. Thus:

- If $|t-s|^{1/2}\geq\delta/(n+1)$, there is no singularity in (4.5) and (4.7) holds.

ACCEPTED MANUSCRIPT

- If there exists $i \in [\![1,n]\!],\ |(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_i|^{1/(2i-1)} \geq \delta/(n+1)$ we have from (4.5):

$$|k(s,t,\mathbf{x},\mathbf{y})|\mathbb{I}_{\rho(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})\geq\delta} \leq \frac{C_{4.5}(n+1)^2}{\delta^2}\frac{|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_i|^{2/(2i-1)}}{t-s}q_{c_{4.5}}(s,t,\mathbf{x},\mathbf{y})$$

$$\leq \frac{C_{4.5}(n+1)^2}{\delta^2}((t-s)|\mathbb{T}_{t-s}^{-1}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|^2)^{1/(2i-1)}q_{c_{4.5}}(s,t,\mathbf{x},\mathbf{y})$$

$$\overset{(3.10)}{\leq} \frac{C_{4.5}C_{3.10}(n+1)^2}{\delta^2}((t-s)|\mathbb{T}_{t-s}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|^2)^{1/(2i-1)}q_{c_{4.5}}(s,t,\mathbf{x},\mathbf{y})$$

$$\overset{(3.10)}{\leq} \frac{C_{4.5}C_{3.10}^2(n+1)^2}{\delta^2}((t-s)|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|^2)^{1/(2i-1)}q_{c_{4.5}}(s,t,\mathbf{x},\mathbf{y}),$$

which again yields (4.7) from the definition of $q_{c_{4.5}}$ after (4.5).

Let us now define:

$$K_{ij}^{\epsilon,d}f(s,\mathbf{x}) := \int_{S\cap\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} dtd\mathbf{y}k_{ij}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})\eta_\delta(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}),$$

$$K_{ij}^{\epsilon,\infty}f(s,\mathbf{x}) := \int_{S\cap\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} dtd\mathbf{y}k_{ij}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})(1-\eta_\delta)(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}),$$

where η_δ is a smooth non-negative cut-off function s.t. for all $(u,z)\in\mathbb{R}\times\mathbb{R}^{nd},\ \eta_\delta(u,z)=1$ if $\rho(u,z)\leq\delta$ and $\eta_\delta(u,z)=0$ if $\rho(u,z)\geq 2\delta$. We thus have:

$$K_{ij}^{\epsilon}f(s,\mathbf{x}) = K_{ij}^{\epsilon,d}f(s,\mathbf{x}) + K_{ij}^{\epsilon,\infty}f(s,\mathbf{x}), \tag{4.8}$$

It is then easily seen from (4.7), that for all $f\in L^p(S),\ p\in[1,+\infty]$,

$$\|K_{ij}^{\epsilon,\infty}f\|_{L^p(S)} \leq C_{4.9}\|f\|_{L^p(S)},\ C_{4.9} := C_{4.9}(T,(\mathbf{A}),\delta,p). \tag{4.9}$$

The singular part of the kernel requires a much more subtle handling. Setting $k_{ij}^d(s,t,\mathbf{x},\mathbf{y}) = \eta_\delta(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})k_{ij}(s,t,\mathbf{x},\mathbf{y})$ we will prove the following proposition.

Proposition 4.2 (Calderón-Zygmund Kernel).

i) $\exists C_{4.2} := C_{4.2}(T,(\mathbf{A}),\delta),\ \forall\big((s,\mathbf{x}),(t,\mathbf{y})\big)\in S^2,\ |k_{ij}^d(s,t,\mathbf{x},\mathbf{y})| \leq \frac{C_{4.2}}{\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))^{n^2d+2}}$.

ii) There exists a constant $c_{4.2}$ s.t.

$$|k_{ij}^d(s,t,\mathbf{x},\mathbf{y})-k_{ij}^d(\sigma,t,\boldsymbol{\xi},\mathbf{y})| \leq C_{4.2}\{\frac{\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))^\eta}{\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))^{n^2d+2+\eta}} + \frac{1}{\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))^{n^2d+2-\eta}}\},$$

$\forall\big((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})\big)\in S^2,\ c_{4.2}\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})) \leq \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) \leq \Lambda$ for some $\Lambda\leq 1$ that will be specified later on.

iii) The two previous "standard estimates" hold for the adjoint kernel

$$k_{ij}^{d,*}(s,t,\mathbf{x},\mathbf{y}) := k_{ij}^d(t,s,\mathbf{y},\mathbf{x}).$$

ACCEPTED MANUSCRIPT

iv) Cancellation Property:

$$\sup_{\epsilon>0}\Big|\int_{\mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y}\Big| + \sup_{\epsilon>0}\Big|\int_{\mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d,*}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y}\Big| < +\infty.$$

Also the limits:

$$h(s,\mathbf{x}) := \lim_{\epsilon\to 0}\int_{\mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y},$$
$$h^*(s,\mathbf{x}) := \lim_{\epsilon\to 0}\int_{\mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d,*}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y}$$

exist, are finite for almost every $(s,\mathbf{x})\in S$ and s.t. $h-h^*$ is in $C^\gamma(S)$ for some $\gamma>0$.

Remark 4.4. Let us emphasize that the constant Λ in point *ii)* is rather arbitrary since it is only needed to split the singular and non-singular part of the kernel. In practice, we will choose $\Lambda = 2\delta$ small enough to suitably control the linearization (3.5) of the initial deterministic differential system (2.3). Let us as well mention that it is precisely this linearization error that also yields the second term in the r.h.s. of point *ii)*, which is not *usual* in the estimates for a singular kernel, but gives an integrable singularity, w.r.t. the singular control of point *i)*. This term could have been avoided by modifying the definition of the singular kernel at hand which would anyhow have seemed more complicated and less *natural*. We refer to Lemma 6.3 and the proof of Proposition 4.2 in Section 6.3 for details.

The strategy is now to exploit those estimates to derive L^p controls on the covering of S with the $\mathbf{d}$-balls introduced in Theorem 4.1. But to do so, we have to carefully check that the cancellation property appearing in Proposition 4.2 for the whole space still holds on the metric balls. This property can be conserved thanks to a Hölder continuous cut-off as in Proposition 18 from [BCLP10]. Namely, from Proposition 4.2 it can be derived similarly to the previous reference that:

Proposition 4.3 (Localized Cancellation). There exists a constant $R_0>0$ s.t. for $(s_0,\mathbf{x}_0)\in S$, $R\le R_0$, if a,b stand for two cut-off functions belonging to $C^\alpha(\mathbb{R}^{n+1},\mathbb{R})$, $\alpha>0$ and with support in $B((s_0,\mathbf{x}_0),R)$, then defining

$$k_{ij}^{d,\mathrm{loc}}(s,t,\mathbf{x},\mathbf{y}) := a(s,\mathbf{x})k_{ij}^{d}(s,t,\mathbf{x},\mathbf{y})b(t,\mathbf{y}),$$
$$k_{ij}^{d,*,\mathrm{loc}}(s,t,\mathbf{x},\mathbf{y}) := a(s,\mathbf{x})k_{ij}^{d,*}(s,t,\mathbf{x},\mathbf{y})b(t,\mathbf{y}),$$

we have that:

- $k_{ij}^{d,\mathrm{loc}}$, $k_{ij}^{d,*,\mathrm{loc}}$ satisfy the first three points of Proposition (4.2) and for all $(s,\mathbf{x})\in B((s_0,\mathbf{x}_0),R)$:

$$\sup_{\epsilon>0}\Big|\int_{(t,\mathbf{y})\in B((s_0,\mathbf{x}_0),R),\ \mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d,\mathrm{loc}}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y}\Big| + \sup_{\epsilon>0}\Big|\int_{(t,\mathbf{y})\in B((s_0,\mathbf{x}_0),R),\ \mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d,*,\mathrm{loc}}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y}\Big| < +\infty.$$

ACCEPTED MANUSCRIPT

- For almost all $(s,\mathbf{x}) \in B((s_0,\mathbf{x}_0),R)$ the limits

$$h^{\mathrm{loc}}(s,\mathbf{x}) := \lim_{\epsilon\to 0} \int_{(t,\mathbf{y})\in B((s_0,\mathbf{x}_0),R),\ \mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d,\mathrm{loc}}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y},$$

$$h^{*,\mathrm{loc}}(s,\mathbf{x}) := \lim_{\epsilon\to 0} \int_{(t,\mathbf{y})\in B((s_0,\mathbf{x}_0),R),\ \mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d,*,\mathrm{loc}}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y}$$

exist, are finite and s.t. $h^{\mathrm{loc}} - h^{*,\mathrm{loc}}$ is in $C^\gamma(B((s_0,\mathbf{x}_0),R))$ for some $\gamma > 0$. .

Now from Propositions 4.2 and 4.3 we derive from Theorem 3 in [Bra10], recalling from Proposition 4.1 that $\mathbf{d}$ and $\mathbf{d}^*$ are equivalent on "quasi" metric balls, that for every $(s_0,\mathbf{x}_0) \in S,\ R \le R_0,\ p \in (1,+\infty)$, there exists a constant $C_{p,T,(\mathbf{A})}$ independent of $(s_0,\mathbf{x}_0)$ s.t. setting

$$T_{ij}f(s,\mathbf{x}) := \lim_{\epsilon\to 0} \int_{(t,\mathbf{y})\in B((s_0,\mathbf{x}_0),R),\ \mathbf{d}^*((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon} k_{ij}^{d,\mathrm{loc}}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})dtd\mathbf{y},$$

$$\|T_{ij}f\|_{L^p(B((s_0,\mathbf{x}_0),R))} \le C_{p,T,(\mathbf{A})}\|f\|_{L^p(B((s_0,\mathbf{x}_0),R))}.$$

The covering Theorem 4.1 then gives, similarly to the proof of Theorem 22 in [BCLP10], that setting $K_{ij}^d f(s,\mathbf{x}) := \lim_{\epsilon\to 0} K_{ij}^{\epsilon,d} f(s,\mathbf{x})$, for every $p \in (1,+\infty)$, there exists a constant $C_{p,T,(\mathbf{A})}$ s.t. $\|K_{ij}^d f\|_{L^p(S)} \le C_{p,T,(\mathbf{A})}\|f\|_{L^p(S)}$. Combining this control with equations (4.8) and (4.9) eventually yields that

$$\|D^2_{\mathbf{x}_1^i,\mathbf{x}_1^j}\tilde{G}f\|_{L^p(S)} \le C_{p,T,(\mathbf{A})}\|f\|_{L^p(S)},$$

up to a modification of $C_{p,T,(\mathbf{A})}$. This concludes the proof of Theorem 3.1 under **(A)**.

5. Derivation of Theorem 2.1 from the Calderón-Zygmund estimates

5.1. **Well posedness of the martingale problem.** Existence can be obtained by usual compactness arguments, see e.g. Theorem 6.1.7 in [SV79] that can be adapted to the current framework. We will therefore focus on uniqueness.

The strategy is the following. We first prove the well-posedness of the martingale problem under the local condition (3.1) for $T > 0$ small enough. Still under (3.1), we then get rid of the small time constraint thanks to a chaining/gluing argument (see e.g. Chapter 6 in [SV79], and Chapter 4.6 in Ethier Kurtz [EK97] or Chapter 4.11 in Kolokoltsov [Kol11] in the more general framework of càdlàg processes). We eventually derive the well-posedness on the whole space thanks to the localization results in Priola [Pri15].

5.1.1. *Well Posedness with Local Condition.* In this section we must adapt carefully the arguments in Chapter 7 of [SV79], who consider a zero or bounded drift term. In our model the drift is simply crucial. In the linear case, we can refer to the work of Priola [Pri15], who derived through resolvents the well posedness under (3.1) for an arbitrary time. The nonlinear drift $\mathbf{F}$ yields, for the linearization to be efficient, additional small time-constraints.

There are two key steps to derive uniqueness. The first one is the following Lemma.

ACCEPTED MANUSCRIPT

Lemma 5.1. For $T > 0$ small enough and under (3.1), whenever $\mathbb{P}$ solves the martingale problem associated with $(L_t)_{t\in[0,T]}$, for every $p > (n^2d+2)/2$, there exists $C_{5.1} := C_{5.1}(p,(\mathbf{A}))$ s.t. for all $(s,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$, $f \in C_0^\infty([s,T]\times\mathbb{R}^{nd})$:

$$\left|\mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}\left[\int_s^T f(t,\mathbf{X}_t)dt\right]\right| \leq C_{5.1}\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})}. \tag{5.1}$$

Observe that this means that, for every solution of the martingale problem, the associated canonical process has a density whose L^q norm, where $q^{-1}+p^{-1} = 1, p, q > 1$, is uniformly controlled.

Proposition 5.1. Under **(A)**, if $T > 0$ is small enough and the local condition (3.1) is fulfilled, then the martingale problem is well posed on $[0,T]$ and if $\mathbb{P}_{s,\mathbf{x}}$ stands for the associated family of solutions, using the notations of Section 3, we have for all $(s,\mathbf{x}) \in [0,T)\times\mathbb{R}^{nd}, f \in C_0^\infty([s,T]\times\mathbb{R}^{nd})$:

$$\mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}\left[\int_s^T f(t,\mathbf{X}_t)dt\right] = \tilde{G}\circ(I-R)^{-1}f(s,\mathbf{x}).$$

Also, if there exists $\bar{\mathbf{x}} \in \mathbb{R}^{nd}$ s.t.

$$\varepsilon_{a,\infty} := \sup_{t\geq 0}\sup_{\mathbf{x}\in\mathbb{R}^{nd}}|a(t,\mathbf{x}) - a(t,\bar{\mathbf{x}})|,$$

is small, then the martingale problem is well posed on $\mathbb{R}^+$.

The proofs of Lemma 5.1 and Proposition 5.1 are postponed to Section 5.1.3.

5.1.2. *Derivation of the Global Well-Posedness.* Now, without the local condition (3.1), the continuity assumed in **(A)** allows to localize. Precisely, it is possible to consider a countable covering of $E := [0,+\infty)\times\mathbb{R}^{nd} = \cup_{i\in\mathbb{N}}\mathcal{G}_i^\delta$ where

$$\mathcal{G}_i^\delta := [(s_i-\delta)\vee 0, s_i+\delta]\times B^{Eucl}(\mathbf{x}_i,\delta),$$

for $(s_i,\mathbf{x}_i) \in \mathbb{R}^+\times\mathbb{R}^{nd}$, $B^{Eucl}(\mathbf{x}_i,\delta)$ stands for the usual *Euclidean* ball of $\mathbb{R}^{nd}$ with center $\mathbf{x}_i$ and radius $\delta > 0$. The parameter δ is chosen s.t. $2\delta \leq T$, for T as in the previous paragraph, and

$$\sup_{(t,\mathbf{y})\in\mathcal{G}_i^\delta}|a(t,\mathbf{y}) - a(t,\mathbf{x}_i)| \leq \varepsilon_a,$$

for a sufficiently small ε_a (condition (3.1)). This statement can be proved by compactness arguments. In the homogeneous case, we refer to the proof of Theorem 6 p. 263 in [Pri15]. Under **(C)**, those arguments extend to the current framework. Define then $\forall(t,\mathbf{y}) \in \mathbb{R}^+\times\mathbb{R}^{nd}$, $\tilde{a}_i(t,\mathbf{y}) = a(t,\mathbf{y})\mathbb{I}_{(t,\mathbf{y})\in\mathcal{G}_i^\delta} + (1-\mathbb{I}_{(t,\mathbf{y})\in\mathcal{G}_i^\delta})a(s_i,\mathbf{x}_i)$. Denoting by $\tilde{L}_i$ the generator of (1.1) associated with diffusion coefficient $\tilde{a}_i$, since $\varepsilon_{\tilde{a}_i,\infty} \leq \varepsilon_a$ which can be chosen small enough thanks to the continuity assumption, we have from Proposition 5.1 that the martingale problem is well posed for $\tilde{L}_i$. Also, $\tilde{L}_i = L$ on $\mathcal{G}_i^\delta$. The same holds if we add the time derivative in the operators (in order to take into account the inhomogeneity) or equivalently if we consider the time space processes. We can then conclude to global uniqueness from Theorem 26 in [Pri15] taking as state space E.

ACCEPTED MANUSCRIPT

5.1.3. *Proof of Lemma 5.1 and Proposition 5.1.*

Proof of Lemma 5.1

- **Step 1.** Let us introduce for given $(s,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$ and a measurable function $\varsigma : [0,T] \to \mathcal{S}^d$ satisfying **(UE)** the process,

$$\bar{\mathbf{X}}_t^{s,\mathbf{x},\varsigma} := \mathbf{x} + \int_s^t \mathbf{F}(u, \bar{\mathbf{X}}_u^{s,\mathbf{x},\varsigma})du + \int_s^t B\varsigma^{1/2}(u)d\beta_u,$$

defined on some filtered probability space $(\mathcal{E},\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mu)$ on which $(\beta_u)_{u\geq 0}$ is a Brownian motion.

The previous dynamics corresponds to a modification of (1.1) where we consider a deterministic non-degenerate non-homogeneous diffusion coefficient $\varsigma^{1/2}$. Observe that under **(A)**, we can derive from Theorem 1.1 in [DM10] that for all $t > s$, $\bar{\mathbf{X}}_t^{s,\mathbf{x},\varsigma}$ has a multiscale Gaussian density $\bar{p}^{\varsigma}(s,t,\mathbf{x},\cdot)$. Precisely, $\bar{p}^{\varsigma}(s,t,\mathbf{x},\mathbf{y})$ satisfies (3.9) with $\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})$ replaced by $\boldsymbol{\theta}_{t,s}(\mathbf{x})$ solving (2.3).

- **Step 2**: Let $\pi^N([0,T])$ be the partition of the interval $[0,T]$ with time-step $h := T/N,\ N \in \mathbb{N}^*$. Define as well for $u \in [0,T]$, $\phi(u) := \{t_i := ih, t_i \leq u < t_{i+1}\}$ (i.e. $\phi(u)$ is the largest discretization time lower or equal than u). Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$ be a filtered probability space on which β is a d-dimensional Brownian motion. For given $(s,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$, let $(\boldsymbol{\xi}_t)_{t\in[s,T]}$ solve:

$$\boldsymbol{\xi}_t = \mathbf{x} + \int_s^t \mathbf{F}(u,\boldsymbol{\xi}_u)du + \int_s^t B\Sigma_u d\beta_u,$$

where the coefficient Σ_u is $\mathcal{F}_{\phi(u)}$- measurable. Conditioning iteratively w.r.t. the $(\mathcal{F}_{t_i})_{i\in[\![h^{-1}\phi(s),h^{-1}T]\!]}$, it can be easily deduced that $(\boldsymbol{\xi}_t)_{t\in(s,T]}$ has a density. It can indeed be written as a convolution of the densities introduced in Step 1. If we additionally assume that, setting $A_u = \Sigma_u\Sigma_u^*$, the following local condition holds: there exists a measurable function $\varsigma : [0,T] \to \mathcal{S}^d$ satisfying **(UE)** s.t.

$$\sup_{(u,\omega)\in[0,T]\times\Omega} \|\varsigma(u) - A_u(\omega)\| \leq \varepsilon_a, \tag{5.2}$$

for a small enough ε_a, the point is now to establish that for all $(s,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$, $f \in C_0^\infty([0,T)\times\mathbb{R}^{nd})$, $p > (n^2d+2)/2$, there exists $C_p := C_p((\mathbf{A}))$ s.t:

$$|\mathbb{E}^{\mathbb{P}}[\int_s^T f(t,\boldsymbol{\xi}_t)dt]| \leq C_p\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})}.$$

Recall from (3.12) that for $\tilde{G}f(s,\mathbf{x}) = \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} f(t,\mathbf{y})\tilde{q}(s,t,\mathbf{x},\mathbf{y})$, where $\tilde{q}(s,t,\mathbf{x},\mathbf{y})$ stands for the density of $(\tilde{\mathbf{X}}_u^{t,\mathbf{y}})_{u\geq s}$ introduced in (3.7), we have:

$$\partial_s\tilde{G}f(s,\mathbf{x}) + \tilde{M}f(s,x) = -f(s,\mathbf{x}),\ (s,\mathbf{x}) \in [0,T)\times\mathbb{R}^{nd},$$

ACCEPTED MANUSCRIPT

for $\tilde{M}f(s,\mathbf{x}) = \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{L}_s^{t,\mathbf{y}} \tilde{q}(s,t,\mathbf{x},\mathbf{y}) f(t,\mathbf{y})$. Denoting by $L_t^{\boldsymbol{\xi}}$ the generator of $\boldsymbol{\xi}$ at time t, let us then write:

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}[\tilde{G}f(s,\boldsymbol{\xi}_s)] &= -\mathbb{E}^{\mathbb{P}}[\int_s^T \{\partial_t \tilde{G}f(t,\boldsymbol{\xi}_t) + \tilde{M}f(t,\boldsymbol{\xi}_t)\}dt] \\ &\quad -\mathbb{E}^{\mathbb{P}}[\int_s^T \{L_t^{\boldsymbol{\xi}}\tilde{G}f(t,\boldsymbol{\xi}_t) - \tilde{M}f(t,\boldsymbol{\xi}_t)\}dt] \\ &=: \mathbb{E}^{\mathbb{P}}[\int_s^T f(t,\boldsymbol{\xi}_t)dt] - \mathbb{E}^{\mathbb{P}}[\int_s^T \bar{R}_\Sigma f(t,\boldsymbol{\xi}_t)dt]. \qquad (5.3)\end{aligned}$$

For $(t,\mathbf{z}) \in [0,T)\times\mathbb{R}^{nd}$, the term $\bar{R}_\Sigma f(t,\mathbf{z}) = L_t^{\boldsymbol{\xi}}\tilde{G}f(t,\mathbf{z}) - \tilde{M}f(t,\mathbf{z})$ can be controlled, thanks to (5.2), exactly as the contribution $Rf(t,\mathbf{z})$ in Section 3 (see equation (3.19)).

From equation (5.3), we derive from (3.19) and (3.20) that:

$$\begin{aligned}&|\mathbb{E}^{\mathbb{P}}[\int_s^T f(t,\boldsymbol{\xi}_t)dt]| \\ &\leq [C_{3.20}T^{1-(n^2d+2)/(2p)} + \frac{\varepsilon_a}{2}C_{3.18} + C(p,(\mathbf{A}))T^{\eta/2}]\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})}.\end{aligned}$$

We have thus proved the estimate (5.1) of Lemma 5.1 for processes of the form $(\boldsymbol{\xi}_t)_{t\geq 0}$.

- **Step 3.** It now remains to extend the previous control to an arbitrary solution $\mathbb{P}$ of the martingale problem on $[0,T]$ under the local condition (3.1). To this end, for a progressively measurable $A : [0,T]\times\Omega \to \mathcal{S}^d$ satisfying (5.2), we set for a family of non-negative mollifiers $(\zeta_\varepsilon)_{\varepsilon>0}$ with compact support on $\mathbb{R}$, i.e. $\int_{\mathbb{R}}\zeta_\varepsilon(t)dt = 1,\ \zeta_\varepsilon \in C_0^\infty(\mathbb{R})$:

$$A_\varepsilon(t) := \int_0^t \zeta_\varepsilon(t-s)A(s)ds,\ \varsigma_\varepsilon(t) := \int_0^t \zeta_\varepsilon(t-s)\varsigma(s)ds,\ t\geq 0.$$

For $N \geq 1$, define now $A_{\varepsilon,N}(t) := A_\varepsilon(\phi(t)), \varsigma_{\varepsilon,N}(t) := \varsigma_\varepsilon(\phi(t))$ where $\phi(t)$ denotes as in **Step 2** the largest discretization time lower or equal to T for the time-step $h = T/N$. It is clear that the pair $(A_{\varepsilon,N},\varsigma_{\varepsilon,N})$ satisfies (5.2). Also, $\lim_{\varepsilon\downarrow 0}\lim_N \mathbb{E}^{\mathbb{P}}[\int_s^T |A_{\varepsilon,N}(t) - A(t)|^2 dt] = 0$. Let now $\mathbf{X}$ be the canonical process associated with $\mathbb{P}$ and set $A_u := a(u,\mathbf{X}_u)$. Introduce then for $t \geq s$, $\beta_t := \int_s^t A^{-1/2}(u)d\mathbf{X}_u^1 - \int_s^t A^{-1/2}(u)\mathbf{F}_1(u,\mathbf{X}_u)du$. Then β is a Brownian motion after time s (see e.g. Theorem 4.5.1 in [SV79] that extends to the current framework) and:

$$\mathbf{X}_t = \mathbf{X}_s + \int_s^t \mathbf{F}(u,\mathbf{X}_u)du + \int_s^t BA^{1/2}(u)d\beta_u,\ t\geq s.$$

Set

$$\boldsymbol{\xi}_t^{\varepsilon,N} = \mathbf{X}_s + \int_s^t \mathbf{F}(u,\boldsymbol{\xi}_u^{\varepsilon,N})du + \int_s^t BA_{\varepsilon,N}^{1/2}(u)d\beta_u,\ t\geq s.$$

From **Step 2**, we derive that for all $f \in C_0^\infty([0,T)\times\mathbb{R}^{nd})$:

$$|\mathbb{E}^{\mathbb{P}}[\int_s^T f(t,\boldsymbol{\xi}_t^{\varepsilon,N})dt]| \leq C_p\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})}.$$

ACCEPTED MANUSCRIPT

On the other hand, since $\mathbf{F}$ is globally Lipschitz, we get from Gronwall's Lemma and Doob's inequality

$$\mathbb{E}^{\mathbb{P}}[\sup_{t\in[s,T]}|\boldsymbol{\xi}_t^{\varepsilon,N}-\mathbf{X}_t|^2]\leq C_p\mathbb{E}^{\mathbb{P}}[\int_s^T|A_{\varepsilon,N}(t)-A(t)|dt].$$

Thus,

$$\left|\mathbb{E}^{\mathbb{P}}[\int_s^T f(t,\mathbf{X}_t)dt]\right|=\lim_{\varepsilon\downarrow 0}\lim_N\left|\mathbb{E}[\int_s^T f(t,\boldsymbol{\xi}_t^{\varepsilon,N})dt]\right|\leq C_p\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})},$$

which completes the proof.

Proof of Proposition 5.1.

Let now $\mathbb{P}$ be an arbitrary solution to the martingale problem. From Lemma 5.1 we get that for all $(s,\mathbf{x})\in[0,T)\times\mathbb{R}^{nd}$, there exists $\varphi\in L^q([s,T)\times\mathbb{R}^{nd}), q\in(1,1+\frac{2}{n^2d})$ s.t.:

$$\mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}[\int_s^T f(t,\mathbf{X}_t)dt]=\int_s^T dt\int_{\mathbb{R}^{nd}}f(t,\mathbf{y})\varphi(t,\mathbf{y})d\mathbf{y},$$

for all $f\in C_0^\infty([0,T]\times\mathbb{R}^{nd})$. On the other hand, similarly to (5.3):

$$\begin{aligned}\tilde{G}f(s,\mathbf{x}) &= \mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}[\int_s^T f(t,\mathbf{X}_t)dt]-\mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}[\int_s^T Rf(t,\mathbf{X}_t)dt]\\ &= \int_s^T dt\int_{\mathbb{R}^{nd}}(I-R)f(t,\mathbf{y})\varphi(t,\mathbf{y})d\mathbf{y}.\end{aligned}$$

Since both sides are continuous w.r.t. the L^p convergence and that $I-R$ is invertible on L^p (see eq. (3.19)) we conclude that:

$$\mathbb{E}^{\mathbb{P}_{s,\mathbf{x}}}[\int_s^T f(t,\mathbf{X}_t)dt]=\int_s^T dt\int_{\mathbb{R}^{nd}}f(t,\mathbf{y})\varphi(t,\mathbf{y})d\mathbf{y}=\tilde{G}\circ(I-R)^{-1}f(s,\mathbf{x}).$$

This gives uniqueness on $[0,T]$. Set now $\mathbb{P}^1=\mathbb{P}$, the unique solution of the martingale problem on $[0,T]$. If $\varepsilon_{a,\infty}$ is small, we have as well that, for all $i\geq 2$, there is a unique solution $\mathbb{P}^i$ to the martingale problem on $[(i-1)T,iT]=:\mathcal{I}_i$, i.e. given $(s,\mathbf{x})\in\mathcal{I}_i\times\mathbb{R}^{nd}$ there is a unique $\mathbb{P}^i_{s,\mathbf{x}}$ on $C([0,\infty),\mathbb{R}^{nd}),\mathcal{B}(C([0,\infty),\mathbb{R}^{nd})))$ s.t. $\mathbb{P}^i_{s,\mathbf{x}}[\mathbf{X}_t=\mathbf{x},0\leq t\leq s]$ and $f(\mathbf{X}_{t\wedge iT})-\int_s^{t\wedge iT}L_uf(\mathbf{X}_u)du$ is a $\mathbb{P}^i_{s,\mathbf{x}}$ martingale after time s for all $f\in C_0^\infty(\mathbb{R}^{nd})$. For fixed $(s,\mathbf{x})\in[0,T]\times\mathbb{R}^{nd}$, define inductively $\mathbb{Q}^1_{s,\mathbf{x}}:=\mathbb{P}^1_{s,\mathbf{x}}$, $\mathbb{Q}^N_{s,\mathbf{x}}:=\mathbb{Q}^{N-1}_{s,\mathbf{x}}\otimes_{(N-1)T}\mathbb{P}^N_{(N-1)T,\mathbf{X}_{(N-1)T}}$, $N\geq 2$. It can be shown from the well posedness of the martingale problem on each time interval $(\mathcal{I}_i)_{i\in\mathbb{N}^*}$ and Theorem 6.1.2 in [SV79], considering the $(iT)_{i\geq 1}$ as stopping times, that the martingale problem is, under the local condition (3.1) well posed on $[0,+\infty)$. Roughly speaking, the $\mathbb{P}^i$ can be "glued" together yielding global well posedness.

5.2. **Existence of the Density and Associated Estimates.** The goal of this section is to prove the statement (2.2) of Theorem 2.1. To this end, we will need the following result which extends to our current degenerate setting Theorem 9.1.9 in [SV79] in *small* time.

Theorem 5.1 (Local existence of the density and associated estimates). Assume $T\leq T_0((\mathbf{A}))\leq 1$ as in Theorem 3.1.

ACCEPTED MANUSCRIPT

- If the diffusion coefficient a is uniformly continuous, then, for $0 \leq s < T$, the unique weak solution of (1.1) admits a density in the following sense. Letting $P(s,t,\mathbf{x},.)$ be the transition probability determined by $(L_t)_{t\geq 0}$, then for a given $T \in (0, T_0]$, almost all $t \in (s,T]$ and all $\Gamma \in \mathcal{B}(\mathbb{R}^{nd})$, $P(s,t,\mathbf{x},\Gamma) = \int_\Gamma p(s,t,\mathbf{x},\mathbf{y})d\mathbf{y}$. Also, for $q \in [1,2)$, the density p satisfies:
$$\left(\int_s^T dt(t-s)^\alpha \int_{\mathbb{R}^{nd}} d\mathbf{y}|p(s,t,\mathbf{x},\mathbf{y})|^q\right)^{1/q} \leq C^1_{5.1}(1+|\mathbf{x}|),$$
where $\alpha = \left(\frac{(n^2d+2)}{2}\right)(q-1)$ and $C^1_{5.1} := C^1_{5.1}(T,q,(\mathbf{A}),\delta_T)$, denoting for all $\varepsilon > 0$, $\delta_T(\varepsilon) := \text{argmax}_{\zeta\in\mathbb{R}^+}\{\sup_{s\in[0,T],|\mathbf{x}-\mathbf{y}|\leq\zeta}|a(s,\mathbf{x}) - a(s,\mathbf{y})| < \varepsilon\}$ the modulus of continuity of a.
- For all $0 \leq s < T$ and $q \in [1,2)$, $\delta > 0$,
$$\left(\int_s^T dt \int_{\mathbb{R}^{nd}\setminus B^{Eucl}(\boldsymbol{\theta}_{t,s}(\mathbf{x}),\delta)} d\mathbf{y}|p(s,t,\mathbf{x},\mathbf{y})|^q\right)^{1/q} \leq C^2_{5.1}(1+|\mathbf{x}|),$$
with $C^2_{5.1} := C^2_{5.1}(T,q,(\mathbf{A}),\delta_T,\delta)$, $B^{Eucl}(\boldsymbol{\theta}_{t,s}(\mathbf{x}),\delta)$ standing again for the Euclidean ball of $\mathbb{R}^{nd}$ with radius δ and center $\boldsymbol{\theta}_{t,s}(\mathbf{x})$, recalling $\boldsymbol{\theta}_{t,s}(\mathbf{x}) = \mathbf{x} + \int_s^t \mathbf{F}(u,\boldsymbol{\theta}_{u,s}(\mathbf{x}))du$ (i.e. $\boldsymbol{\theta}_{t,s}(\mathbf{x})$ is the solution at time t of the deterministic differential system associated with (1.1) starting from $\mathbf{x}$ at time s).

Remark 5.1. There are three differences w.r.t. to Theorem 9.1.9 in [SV79]. First, the norm of the initial point in the r.h.s. of the above controls is due to the transport by unbounded coefficients. Second, the small time constraint follows from our linearization strategy employed to derive the Calderón-Zygmund estimates of Theorem 3.1. At last, the upper bound on q comes from the control on the remainders in (3.17) (which imposes the constraint $p > 2$ with $p^{-1} + q^{-1} = 1$). When the system is linear w.r.t. the components that transmit the noise (see Remark 3.3) this constraint disappears and the result of Theorem 5.1 hold for $q \in [1, +\infty)$.

We provide below the principal lines needed to adapt the proof of Theorem 9.1.9 in [SV79], stressing which specific modifications are needed in the degenerate case and mainly concern the *localization* arguments. Once again the key idea is to localize along the characteristic lines associated with the deterministic differential system instead of using spatial balls only as in [SV79]. Observe anyhow that, when the drift is bounded, the product of the time interval and the spatial ball can be seen as a tube along a *characteristic line*. Indeed, if the drift is 0 then the deterministic differential system does not leave its initial condition; if it is bounded, the image of a spatial ball by the deterministic system will stay uniformly in time in a ball whose radius only depend on the bound of the drift, the final time and the initial radius, **but not** on the points of the initial ball.

5.2.1. *Controls for Slowly Varying Coefficients.* We use here freely the notations of Section 3 for the operators $\tilde{G}, R$ (see equations (3.12)-(3.15)). Also, in order to keep notations close to those in [SV79], we introduce for $r > 2$, the class $\mathcal{A}(r,T)$ of measurable coefficients $a : \mathbb{R}^+ \times \mathbb{R}^{nd}$ satisfying **(UE)** and $\mathbf{F}$ satisfying **(ND)**, **(S)**, for which there exists $\mathbf{x}_0 \in \mathbb{R}^{nd}$ s.t. with the notations of (3.1), for all $\rho \in [r, \frac{n^2d+4}{2} \vee r]$, $(\frac{\varepsilon_a}{2}C_{3.18}(T,\rho,(\mathbf{A})) + C(\rho,(\mathbf{A}))T^{\frac{\eta}{2}}) < 3/4$ so that from (3.19), we have that $(I-R)^{-1}$ is consistent as bounded operator from $L^\rho([0,T]\times\mathbb{R}^{nd})$ into

ACCEPTED MANUSCRIPT

itself. In particular this imposes that $T \leq 1$ is sufficiently small. Setting then $K := \tilde{G} \circ (I - R)^{-1}$, we thus derive from the Young inequality that it is consistent as bounded operator from $L^\rho([0,T] \times \mathbb{R}^{nd})$ into $L^\sigma([0,T] \times \mathbb{R}^{nd})$ provided that

$$0 \leq \frac{1}{\rho} - \frac{1}{\sigma} < \frac{2}{n^2 d + 2}.$$

Here comes the first Lemma emphasizing some regularizing effects of K which can be derived similarly to Lemma 9.1.2 in [SV79].

Lemma 5.2. Let $a, \mathbf{F} \in \mathcal{A}(r,T)$. Then, for $N = \lceil \frac{n^2 d+2}{2} \frac{1}{r} \rceil$, K^{N+1} maps $L^r([0,T] \times \mathbb{R}^{nd})$ into $C_b([0,T] \times \mathbb{R}^{nd})$ (space of real valued bounded continuous functions), i.e. it is L^r-strong Feller. Precisely, for all $(s, \mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$:

$$|K^{N+1} f(s,\mathbf{x})| \leq C_{5.2} \|f\|_{L^r([0,T]\times\mathbb{R}^{nd})},$$

where $C_{5.2} := C_{5.2}(T, r, (\mathbf{A}))$.

From Lemma 5.2, Lemma 9.1.3 in [SV79] and the well posedness of the martingale problem, denoting by $P(s,t,\mathbf{x},.)$ the associated transition function, one then gets: $Pf(s,\mathbf{x}) := \int_s^T dt \int_{\mathbb{R}^{nd}} P(s,t,\mathbf{x},d\mathbf{y}) f(t,\mathbf{y}) = Kf(s,\mathbf{x})$ if $f \in L^\rho([0,T] \times \mathbb{R}^{nd}) \cap L^\infty([0,T] \times \mathbb{R}^{nd})$. Therefore, for $N = \lceil (n^2 d+2)/(2r) \rceil$, $f \in C_0([0,T) \times \mathbb{R}^{nd})$ (continuous functions with compact support),

$$|P^{N+1} f(s,\mathbf{x})| \leq C_{5.2} \|f\|_{L^r([0,T]\times\mathbb{R}^{nd})}. \tag{5.4}$$

This observation then yields the following result.

Lemma 5.3. If $a, \mathbf{F} \in \mathcal{A}(r,T)$ denoting by P the transition function associated with $(L_t)_{t\in[0,T]}$ then for $r \leq \rho \leq +\infty$,

$$\left| \int_s^T dt (t-s)^N \int_{\mathbb{R}^{nd}} P(s,t,\mathbf{x},d\mathbf{y}) f(t,\mathbf{y}) \right| \leq C^1_{5.3} \|f\|_{L^\rho([0,T]\times\mathbb{R}^{nd})},$$

with $C^1_{5.3} := C^1_{5.3}(T, r, (\mathbf{A}))$. Also, for each $\delta > 0$, $r < \rho \leq \infty$,

$$\left| \int_s^T dt \int_{\mathbb{R}^{nd} \setminus B^{Eucl}(\boldsymbol{\theta}_{t,s}(\mathbf{x}),\delta)} P(s,t,\mathbf{x},d\mathbf{y}) f(t,\mathbf{y}) \right| \leq C^2_{5.3} \|f\|_{L^\rho([0,T]\times\mathbb{R}^{nd})},$$

where $C^2_{5.3} := C^2_{5.3}(T, r, \rho, (\mathbf{A}), \delta)$.

Remark 5.2. This is the first statement that differs from [SV79]. Indeed, the unbounded transport contribution appears here for the first time. To fully justify this aspect we give below the full proof of this result.

Proof. The first statement of the Lemma still follows from Lemma 9.1.3 in [SV79] and (5.4) from an interpolation argument. For the second one, we can assume w.l.o.g. that $T \geq s + c,\ c > 0$. In that case:

$$\begin{aligned} &\left| \int_s^T dt \int_{\mathbb{R}^{nd} \setminus B^{Eucl}(\boldsymbol{\theta}_{t,s}(\mathbf{x}),\delta)} P(s,t,\mathbf{x},d\mathbf{y}) f(t,\mathbf{y}) \right| \\ &\leq \sum_{k\geq 1} \int_{s+c/(k+1)}^{s+c/k} dt \int_{\mathbb{R}^{nd} \setminus B^{Eucl}(\boldsymbol{\theta}_{t,s}(\mathbf{x}),\delta)} P(s,t,\mathbf{x},d\mathbf{y}) |f(t,\mathbf{y})| \\ &\quad + c^{-N} \int_{s+c}^T dt (t-s)^N \int_{\mathbb{R}^{nd}} P(s,t,\mathbf{x},d\mathbf{y}) |f(t,\mathbf{y})|. \end{aligned}$$

ACCEPTED MANUSCRIPT

The last contribution can be bounded directly by the first statement of the Lemma. To control the sum, we see that introducing

$$\Lambda^k_{s,\mathbf{x}}\varphi := \int_{s+c/(k+1)}^{s+c/k} dt \int_{\mathbb{R}^{nd}\setminus B^{Eucl}(\boldsymbol{\theta}_{t,s}(\mathbf{x}),\delta)} P(s,t,\mathbf{x},d\mathbf{y})\varphi(t,\mathbf{y}),$$

we indeed get, from Lemma 9.1.3 in [SV79] and the first part of the lemma, that as a linear operator on $L^r([0,T]\times\mathbb{R}^{nd})$, $\Lambda^k_{s,\mathbf{x}}$ is bounded by $N!((k+1)/c)^N$. Now for $\varphi\in L^\infty([0,T]\times\mathbb{R}^d)$,

$$|\Lambda^k_{s,\mathbf{x}}\varphi| \leq \frac{c}{k}|\varphi|_\infty \mathbb{P}_{s,\mathbf{x}}[\sup_{t\in[s,s+c/k]}|\mathbf{X}_t-\boldsymbol{\theta}_{t,s}(\mathbf{x})|\geq\delta].$$

Let us emphasize that it is precisely because we consider the deviations of the process from the deterministic differential system, that we can control the previous term with Bernstein like inequalities. Precisely, from Gronwall's lemma:

$$|\mathbf{X}_t-\boldsymbol{\theta}_{t,s}(\mathbf{x})| \leq \exp(CT)|\int_s^t \sigma(u,\mathbf{X}_u)dW_u| \leq \exp(C)|\int_s^t \sigma(u,\mathbf{X}_u)dW_u|,$$

with $C:=C((\mathbf{A}))$, so that, from Bernstein's inequality:

$$\mathbb{P}_{s,\mathbf{x}}[\sup_{t\in[s,s+c/k]}|\mathbf{X}_t-\boldsymbol{\theta}_{t,s}(\mathbf{x})|\geq\delta] \leq C\exp(-C^{-1}k\delta^2/c),\ C:=C((\mathbf{A})),$$

up to a modification of C. The result then once again follows from standard interpolation. □

5.2.2. *Localization arguments.* Now we adapt more significantly the arguments in [SV79] to get our results. The leading idea is the same as in the proof of Lemma 5.3: to exploit the Bernstein-like deviations of the process from the deterministic system. We now want to localize carefully to get rid off the *quasi-constant* diffusion coefficient of the previous section. We have the following *tubular* localization estimate.

Lemma 5.4 (Tubular estimate). For $s_0\in[0,T)$, $\mathbf{x}_1\in\mathbb{R}^{nd}$ let $\mathbb{P}_{s_0,\mathbf{x}_1}$ denote the solution to the martingale problem associated with $(L_t)_{t\in[s_0,T]}$. For $0<R_1<R_2$, $\mathbf{x}_0\in\mathbb{R}^{nd}$ defining $\tau_{-1}=s_0$ and for all $k\in\mathbb{N}$,

$$\tau_{2k} := \inf\{t\geq\tau_{2k-1} : |\mathbf{X}_t-\boldsymbol{\theta}_{t,s_0}(\mathbf{x}_0)|=R_2\},$$
$$\tau_{2k+1} := \inf\{t\geq\tau_{2k} : |\mathbf{X}_t-\boldsymbol{\theta}_{t,s_0}(\mathbf{x}_0)|=R_1\},$$

then

$$\mathbb{E}^{\mathbb{P}_{s_0,\mathbf{x}_1}}[\sum_{k\geq0}\mathbb{I}_{\tau_{2k}\in[0,T]}] \leq C_{5.4} := C_{5.4}(T,(\mathbf{A}),R_2-R_1).$$

The proof can be performed as in Lemma 9.1.6 in [SV79]. The previous definitions of the stopping times allows to apply the required Bernstein like arguments similarly to the proof of Lemma 5.3.

The following result differs once again in the localization argument from Lemma 9.1.7 in [SV79], even though it can be proved rather similarly from Lemma 5.4. We emphasize here that the localization has to be performed in time and space. Roughly speaking this is needed in order to *partition* in time the characteristic tubes in subtubes for which the local condition (3.1) is valid. This is the key of the proof.

ACCEPTED MANUSCRIPT

Lemma 5.5 (First Localization Lemma). Let $\mathbb{P}_{s_0,\mathbf{x}_1}$ solve the martingale problem for $(L_t)_{t\in[0,T]}$ starting from $(s_0,\mathbf{x}_1)\in[0,T]\times\mathbb{R}^{nd}$. Suppose now that the martingale problem associated with the operator

$$\tilde{L}_t = \mathbf{F}(t,\cdot)\cdot\mathbf{D}_{\mathbf{x}} + \frac{1}{2}\mathrm{Tr}(\tilde{a}(t,\cdot)D^2_{\mathbf{x}_1})$$

is well posed and that $\tilde{a}=a$ on

$$\mathcal{C}_{\underline{t},\bar{t},R}(s_0,\mathbf{x}_0) := \{(t,\mathbf{y})\in[0,T]\times\mathbb{R}^{nd} : t\in[\underline{t},\bar{t}], \boldsymbol{\theta}_{s_0,t}(\mathbf{y})\in B^{Eucl}(\mathbf{x}_0,R)\}, \tag{5.5}$$

for some $s_0\le\underline{t}<\bar{t}\le T,\ R>0$. Let us denote by $\tilde{\mathbb{P}}_{s_0,\mathbf{x}_1}$ the solution to the martingale problem for $(\tilde{L}_t)_{t\in[0,T]}$ starting from $(s_0,\mathbf{x}_1)\in[0,T]\times\mathbb{R}^{nd}$. Then for each $\delta\in(0,R)$:

$$|\mathbb{E}^{\mathbb{P}_{s_0,\mathbf{x}_1}}[\int_{s_0}^T f(t,\mathbf{X}_t)dt]| \le \mathbb{E}^{\tilde{\mathbb{P}}_{s_0,\mathbf{x}_1}}[\int_{s_0}^T |f(t,\mathbf{X}_t)|dt]$$
$$+C_{5.5}\sup_{(s,\mathbf{x})\in\partial\mathcal{C}_{\underline{t},\bar{t},R-\delta}(s_0,\mathbf{x}_0)}\mathbb{E}^{\tilde{\mathbb{P}}_{s,\mathbf{x}}}[\int_s^T |f(t,\mathbf{X}_t)|dt],$$

for all $f\in C_0(\mathcal{C}_{\underline{t},\bar{t},R-\delta}(s_0,\mathbf{x}_0))$ and $C_{5.5}:=C_{5.5}(T,(\mathbf{A}),R,\delta)$.

We now specify how this Lemma needs to be used. As a direct corollary of Lemmas 5.5 and 5.3 we derive:

Lemma 5.6 (Second Localization Lemma). Let $\tilde{a}(s,\mathbf{x}):=a(s,\mathbf{x})$, in $\mathcal{C}_{\underline{t},\bar{t},R}(s_0,\mathbf{x}_0)$ and $\tilde{a}(s,\mathbf{x}):=a(s_0,\mathbf{x}_0)$, elsewhere. Assume that the functions $\tilde{a},\mathbf{F}\in\mathcal{A}(r,T)$ for some $r\in(2,+\infty)$. Let $\mathbb{P}_{s_0,\mathbf{x}_1}$ solve the martingale problem for L starting at $s_0,\mathbf{x}_1\in[0,T]\times\mathbb{R}^{nd}$. Then for each $0<\alpha<R$ and $r<\rho\le+\infty$.

$$|\mathbb{E}^{\mathbb{P}_{s_0,\mathbf{x}_1}}[\int_{s_0}^T (t-s_0)^N f(t,\mathbf{X}_t)dt]| \le C^1_{5.6}\|f\|_{L^\rho([s_0,T]\times\mathbb{R}^{nd})},$$

for all function $f\in C_0(\mathcal{C}_{\underline{t},\bar{t},\alpha}(s_0,\mathbf{x}_0))$, where $N=\lceil(n^2d+2)/(2r)\rceil$ and $C^1_{5.6}:=C^1_{5.6}(T,(\mathbf{A}),r,\rho,R-\alpha)$. If additionally, $|\mathbf{x}_1-\mathbf{x}_0|>\alpha$, then

$$|\mathbb{E}^{\mathbb{P}_{s_0,\mathbf{x}_1}}[\int_{s_0}^T f(t,\mathbf{X}_t)dt]| \le C^2_{5.6}\|f\|_{L^\rho([s_0,T]\times\mathbb{R}^{nd})},$$

where $C^2_{5.6}:=C^2_{5.6}(T,(\mathbf{A}),r,\rho,R-\alpha,|\mathbf{x}_1-\mathbf{x}_0|-\alpha)$.

5.2.3. *Proof of Theorem 5.1.* From the previous localization Lemmas, the idea is now to specifically *partition* the space in order to have crowns of the previous type, $\mathcal{C}_{\underline{t},\bar{t},R}(s_0,\mathbf{x}_0)$, introduced in (5.5), on which the local condition (3.1) holds. Let q' denote the conjugate of $q\in(1,2)$ and choose $r\in(2,q')$ s.t. $N:=\lceil\frac{n^2d+2}{2}\frac{1}{r}\rceil=\lceil\frac{n^2d+2}{2}\frac{1}{q'}\rceil$. Choose $T\le(4(\sup_{\rho\in[r,(\frac{n^2d+2}{2}\vee r)+1]}C(\rho,(\mathbf{A}))))^{-2}$ in (3.19) and set $\varepsilon^{-1}:=C_{3.18}(T,r,(\mathbf{A}))\vee C_{3.18}(T,(\frac{n^2d+2}{2}\vee r)+1,(\mathbf{A}))\vee\varepsilon_a^{-1}$. Let us introduce for a fixed starting point $(s,\mathbf{x})$ of the martingale problem, the spatial balls

$$Q_{\mathbf{k}}:=\{\mathbf{x}+\mathbf{y} : |\mathbf{y}_j-\mathbf{k}_j\gamma/(nd)^{1/2}|\le\gamma/(nd)^{1/2},\ j\in[\![1,nd]\!]\},\ \mathbf{k}\in\mathbb{Z}^{nd},$$

where $\gamma:=\frac{\delta_T(\varepsilon)}{C_1}$, and recalling that δ_T stands for the modulus of continuity of a, the constant C_1 is then chosen large enough so that for all $\mathbf{k}\in\mathbb{Z}^{nd},\ \mathbf{y}_0,\mathbf{y}_1\in Q_{\mathbf{k}},\ t\in[s,T]$,

$$|a(t,\boldsymbol{\theta}_{t,s}(\mathbf{y}_0))-a(t,\boldsymbol{\theta}_{t,s}(\mathbf{y}_1))|\le\varepsilon/2. \tag{5.6}$$

ACCEPTED MANUSCRIPT

This means that the local condition is satisfied on the time section of the transport of $Q_{\mathbf{k}}$ by the flow. In order to apply the previous results, we also need to handle the time contribution. Define now $h^{\mathbf{k}} := \frac{T-s}{\lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil}$ where the constant $C_2 := C_2(\varepsilon)$ is chosen large enough, so that for all $i \in [\![0, \lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil]\!]$, setting $t_i^{\mathbf{k}} := s+ih^{\mathbf{k}}$, the coefficient a restricted to $\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\beta\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})$, $\beta > 1$, coincides with some $\tilde{a}$ s.t. $\tilde{a}, \mathbf{F}$ belong to the class $\mathcal{A}(r,T)$.

This choice simply means that the length of the time intervals for which we partition the set $\boldsymbol{\theta}(T,s,Q_{\mathbf{k}}) := \{(t,\mathbf{z}) \in [s,T]\times\mathbb{R}^{nd} : \boldsymbol{\theta}_{s,t}(\mathbf{z}) \in Q_{\mathbf{k}}\}$ (image of $Q_{\mathbf{k}}$ by the flow between times s and T) highly depends on the norm of the starting point. This is specifically due to the unbounded drift. Precisely we can write:

$$\boldsymbol{\theta}(T,s,Q_{\mathbf{k}}) := \bigcup_{i=0}^{\lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil-1} \mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2}). \tag{5.7}$$

Now from Lemma 5.6 we get that for $\rho = (r+q')/2$, for all $i \in [\![0, \lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil - 1]\!]$

$$\begin{aligned} &|\int_s^T dt(t-s)^N \int \mathbb{I}_{(t,\mathbf{y})\in\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})} P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})| \\ &= |\int_{t_i^{\mathbf{k}}}^{t_{i+1}^{\mathbf{k}}} dt(t-s)^N \int_{\boldsymbol{\theta}_{t,s}(Q_{\mathbf{k}})} P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})| \le C_{5.8}\|f\|_{L^\rho([0,T]\times\mathbb{R}^{nd})}, \end{aligned} \tag{5.8}$$

where $C_{5.8} := C_{5.8}(T,(\mathbf{A}),r,\rho)$. On the other hand, comparing deviations along the characteristics allows once again to use Bernstein inequalities, similarly to the proof of Lemma 5.3. Namely,

$$\begin{aligned} &P(s,t,\mathbf{x},\boldsymbol{\theta}_{t,s}(Q_{\mathbf{k}})) \\ &\le \mathbb{P}_{s,\mathbf{x}}[\exists u\in[s,t], |\mathbf{X}_u - \boldsymbol{\theta}_{u,s}(\mathbf{x})| > |\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \boldsymbol{\theta}_{u,s}(\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})|/2] \\ &\le \mathbb{P}_{s,\mathbf{x}}[\sup_{u\in[s,T]} |\int_s^u \sigma(v,\mathbf{X}_v)dW_v| \ge C|\mathbf{k}\gamma|] \le 2d\exp(-\bar{C}^{-1}\frac{|\mathbf{k}|^2\gamma^2}{T\mathbf{d}^2}),\ \bar{C} := \bar{C}((\mathbf{A})) \ge 1, \end{aligned}$$

using Gronwall's Lemma (see proof of Lemma 5.3) and the bi-Lipschitz property of the flow for the last but one inequality. We therefore obtain:

$$\begin{aligned} &|\int_s^T dt(t-s)^N \int \mathbb{I}_{(t,\mathbf{y})\in\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})} P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})| \\ &\qquad\qquad \le C_{5.9}\exp(-\bar{C}^{-1}\frac{|\mathbf{k}|^2\gamma^2}{T\mathbf{d}^2})|f|_\infty, \end{aligned} \tag{5.9}$$

where $C_{5.9} := C_{5.9}(T,n,d,N)$. We thus get by interpolation that for $\vartheta = 1-\frac{\rho}{q'} \in (0,1)$:

$$\begin{aligned} &|\int_s^T dt(t-s)^N \int \mathbb{I}_{(t,\mathbf{y})\in\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})} P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})| \\ &\qquad\qquad \le C_{5.8}^{1-\vartheta}C_{5.9}^{\vartheta}\exp(-\bar{C}^{-1}\frac{\vartheta|\mathbf{k}|^2\gamma^2}{T\mathbf{d}^2})\|f\|_{L^{q'}([0,T]\times\mathbb{R}^{nd})}. \end{aligned}$$

ACCEPTED MANUSCRIPT

Summing first, for a given $\mathbf{k}\in\mathbb{Z}^d$, over $i\in[\![0,\lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil-1]\!]$ (that is, according to (5.7), integrating on $\boldsymbol{\theta}(T,s,Q_{\mathbf{k}})$) we obtain

$$
\begin{aligned}
&|\int_s^T dt(t-s)^N\int \mathbb{I}_{(t,\mathbf{y})\in\boldsymbol{\theta}(T,s,Q_{\mathbf{k}})}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})|\\
=&|\int_s^T dt(t-s)^N\int_{\boldsymbol{\theta}_{t,s}(Q_{\mathbf{k}})}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})|\\
\leq&\tilde{C}(T^{1/2}+|\mathbf{x}|)C_{5.8}^{1-\vartheta}C_{5.9}^{\vartheta}\exp(-\frac{\bar{C}^{-1}}{2}\frac{\vartheta|\mathbf{k}|^2\gamma^2}{T\mathbf{d}^2})\|f\|_{L^{q'}([0,T]\times\mathbb{R}^{nd})},\tilde{C}:=\tilde{C}((\mathbf{A}),\vartheta).
\end{aligned}
$$

Summing now over $\mathbf{k}\in\mathbb{Z}^{nd}$ yields:

$$
\begin{aligned}
&|\int_s^T dt(t-s)^N\int_{\mathbb{R}^{nd}}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})|\\
&\leq C_{5.10}(1+|\mathbf{x}|)\|f\|_{L^{q'}([0,T]\times\mathbb{R}^{nd})},C_{5.10}:=C_{5.10}(T,q,(\mathbf{A}),\gamma).
\end{aligned}
\tag{5.10}
$$

This contribution already emphasizes the main difference w.r.t. the non degenerate case: the estimate depends on the initial point. The proof of Theorem 5.1 can then be completed similarly to the one of Theorem 9.1.9 in [SV79]. This achieves the proof of the existence of the density and the associated estimates when T is small enough.

The existence of the density in Theorem 2.1 then follows from a chaining argument.

5.2.4. *Derivation of equation* (2.2) *in Theorem 2.1.* To prove (2.2) we will inductively apply the results of the previous section along a time grid whose time-steps are lower than T_0 in Theorem 5.1. We can assume w.l.o.g. that $T:=T_0N,\ N\in\mathbb{N}$. Setting now $t_i:=iT_0,\ i\in[\![0,N]\!]$, write from the strong Markov property:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}_{0,\mathbf{x}}}[\int_0^T f(t,\mathbf{X}_t)dt]=\sum_{i=0}^{N-1}\mathbb{E}^{\mathbb{P}_{0,\mathbf{x}}}[\mathbb{E}^{\mathbb{P}_{t_i,\mathbf{X}_{t_i}}}[\int_{t_i}^{t_{i+1}}f(t,\mathbf{X}_t)dt]]\\
&=\Big\{\sum_{i=0}^{N-1}\mathbb{E}^{\mathbb{P}_{0,\mathbf{x}}}[\mathbb{E}^{\mathbb{P}_{t_i,\mathbf{X}_{t_i}}}[\int_{t_i}^{t_{i+1}}f(t,\mathbf{X}_t)\mathbb{I}_{\mathbf{X}_t\in B^{Eucl}(\boldsymbol{\theta}_{t,t_i}(\mathbf{X}_{t_i}),\delta)}dt]]\Big\}\\
&+\Big\{\sum_{i=0}^{N-1}\mathbb{E}^{\mathbb{P}_{0,\mathbf{x}}}[\mathbb{E}^{\mathbb{P}_{t_i,\mathbf{X}_{t_i}}}[\int_{t_i}^{t_{i+1}}f(t,\mathbf{X}_t)\mathbb{I}_{\mathbf{X}_t\notin B^{Eucl}(\boldsymbol{\theta}_{t,t_i}(\mathbf{X}_{t_i}),\delta)}dt]]\Big\}:=T_C+T_F,
\end{aligned}
\tag{5.11}
$$

for some $\delta>0$ to be specified later on. From the second part of Theorem 5.1 we obtain that for $p>n^2d/2+1$,

$$
|T_F|\leq C_{5.1}^2\sum_{i=0}^{N-1}\mathbb{E}^{\mathbb{P}_{0,\mathbf{x}}}[(1+|\mathbf{X}_{t_i}|)]\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})}\leq NC_{5.1}^2C_F(1+|\mathbf{x}|)\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})},
\tag{5.12}
$$

where $C_F:=C_F(T_0,(\mathbf{A}))$. On the other hand, we can follow the localization procedure of the previous proof (see equation (5.6)), and find $\delta>0$ s.t. setting $h_i:=\frac{t_{i+1}-t_i}{\lceil C_2(1+|\mathbf{X}_{t_i}|)\rceil}$ and $t_i^j:=t_i+h_ij,\ j\in[\![0,\lceil C_2(1+|\mathbf{X}_{t_i}|)\rceil$ then the coefficient $\tilde{a}$ is equal to a on $\mathcal{C}_{t_i^j,t_i^{j+1},2\delta}(t_i,\mathbf{X}_{t_i})$ and to $a(t_i,\mathbf{X}_{t_i})$ elsewhere, belongs to the class

ACCEPTED MANUSCRIPT

$\mathcal{A}(r,T)$ for some $r \in (2, n^2d/2+1)$. We then derive from Lemma 5.5:

$$|T_C| \leq \sum_{i=0}^{N-1} \mathbb{E}^{\mathbb{P}_{0,\mathbf{x}}}\Big[\sum_{j=0}^{\lceil C_2(1+|\mathbf{X}_{t_i}|)\rceil-1} \mathbb{E}^{\mathbb{P}_{t_i,\mathbf{X}_{t_i}}}\big[\mathbb{E}^{\tilde{\mathbb{P}}_{t_i^j,\mathbf{X}_{t_i^j}}}\big[\int_{t_i^j}^{t_i^{j+1}} |f(t,\mathbf{X}_t)|\mathbb{I}_{\mathbf{X}_t\in B^{Eucl}(\boldsymbol{\theta}_{t,t_i}(\mathbf{X}_{t_i}),\delta)}dt\big] + \sup_{(s,\mathbf{y})\in\partial\mathcal{C}_{t_i^j,t_i^{j+1},\delta}(t_i,\mathbf{X}_{t_i})} \mathbb{E}^{\tilde{\mathbb{P}}_{s,\mathbf{y}}}\big[\int_s^{t_i^{j+1}} |f(t,\mathbf{X}_t)|\mathbb{I}_{\mathbf{X}_t\in B^{Eucl}(\boldsymbol{\theta}_{t,t_i}(\mathbf{X}_{t_i}),\delta)}dt\big]\Big].$$

Now, from equation (3.20), we derive

(5.13)

$$|T_C| \leq C\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})} \sum_{i=0}^{N-1} \mathbb{E}^{\mathbb{P}_{0,\mathbf{x}}}[\lceil C_2(1+|\mathbf{X}_{t_i}|)\rceil] \leq CN(1+|\mathbf{x}|)\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})},$$

up to a modification of C. The result follows from (5.13), (5.12), (5.11).

6. Proofs of the technical results

6.1. **Proofs concerning the quasi-metric structure (Proposition 4.1).** Let us first observe from the definition of the $\mathbf{d}$-balls, see equations (4.1), (4.2), that there exists $C_1 := C_1((\mathbf{A})) > 0$ s.t. for all $\delta > 0, (s,\mathbf{x}) \in S$,

$$|B((s,\mathbf{x}),\delta)| \leq C_1\delta^{2+d\sum_{i=1}^n(2i-1)} = C_1\delta^{2+n^2d}.$$

On the other hand, introducing

$$\bar{B}((s,\mathbf{x}),\delta) := \big\{(t,\mathbf{y}) \in S : |t-s| \leq \frac{\delta^2}{4},\ \rho_{\mathrm{Sp}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}) \leq \frac{\delta}{2}\big\},$$

$$\forall \mathbf{z} \in \mathbb{R}^{nd},\ \rho_{\mathrm{Sp}}(\mathbf{z}) := \sum_{i=1}^n |\mathbf{z}_i|^{1/(2i-1)},$$

i.e. ρ_{Sp} corresponds to the spatial contribution in the metric (4.1), we have that $\bar{B}((s,\mathbf{x}),\delta) \subset B((s,\mathbf{x}),\delta)$. Indeed, for all $(t,\mathbf{y}) \in \bar{B}((s,\mathbf{x}),\delta)$, $\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) := \rho(|t-s|,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}) \leq |t-s|^{1/2} + \rho_{\mathrm{Sp}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}) \leq \delta$. Since we also have, up to a modification of C_1 that for all $(s,\mathbf{x}) \in S,\ \delta > 0,\ |\bar{B}((s,\mathbf{x}),\delta)| \geq C_1^{-1}\delta^{2+n^2d}$ we therefore derive:

$$C_1^{-1}\delta^{2+n^2d} \leq |B((s,\mathbf{x}),\delta)| \leq C_1\delta^{2+n^2d},$$

which gives b) and c). To derive a), we need to exploit the specific structure of the dynamics. Let us first recall how to relate the forward and backward dynamics. Precisely, one has for all $v \in I(t,s) := \begin{cases}[s,t], \text{ if } s<t\\ [t,s], \text{ if } s\geq t\end{cases}$,

$$\boldsymbol{\theta}_{v,s}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{y}) = \boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y} - \int_v^t (\mathbf{F}(u,\boldsymbol{\theta}_{u,s}(\mathbf{x})) - \mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y})))du, \tag{6.1}$$

which for $v = s$ yields:

$$\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}) = \boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y} - \int_s^t (\mathbf{F}(u,\boldsymbol{\theta}_{u,s}(\mathbf{x})) - \mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y})))du.$$

ACCEPTED MANUSCRIPT

Starting from the last components, and assuming w.l.o.g. that $t > s$, we have:

$$|(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_n| \leq |(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_n| + C_2 \int_s^t \left(|(\boldsymbol{\theta}_{v,s}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{y}))_{n-1}| + |(\boldsymbol{\theta}_{v,s}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{y}))_n|\right) dv$$
$$\leq \exp(C_2(t-s)) \left(|(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_n| + C_2 \int_s^t |(\boldsymbol{\theta}_{v,s}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{y}))_{n-1}| dv \right),$$

where $C_2 := C_2((\mathbf{A}))$ and using Gronwall's Lemma for the last inequality. Exploiting iteratively (6.1) and Gronwall's Lemma we derive that there exists $C_3 := C_3(T, (\mathbf{A}))$ s.t.

$$|(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_n| \leq C_3 \sum_{j=1}^{n} |(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j| |t-s|^{n-j}. \tag{6.2}$$

Using Young's inequality with $p_j = \frac{2n-1}{2j-1}$, $q_j = \frac{2n-1}{2(n-j)}$ in order to make the homogeneous exponent of coordinate $j \in [\![1, n-1]\!]$ appear, we get:

$$|(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_n|^{1/(2n-1)} \leq C_3^{1/(2n-1)} \Bigg[|(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_n|^{1/(2n-1)} +$$
$$\sum_{j=1}^{n-1} \left(\frac{|(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j)|^{1/(2j-1)}}{p_j} + \frac{|t-s|^{1/2}}{q_j} \right) \Bigg] \leq C_4 \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})), \tag{6.3}$$

for $C_4 := C_4(T, (\mathbf{A}))$.

The above estimate does not exploit the fact that $\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) \leq \Lambda \leq 1$. This last assumption is actually needed for the components $i \in [\![1, n-1]\!]$ whose differential dynamics potentially involve coordinates $j > i$ with higher characteristic time-scales in small times but that are not negligible in the "homogeneous" norm we consider. Namely, similarly to (6.2) we derive for all $i \in [\![1, n-1]\!]$ up to a modification of C_3:

$$|(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_i| \leq C_3 \left(\sum_{j=1}^{i} |(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j| |t-s|^{i-j} + \sum_{j=i+1}^{n} |(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j| |t-s| \right). \tag{6.4}$$

Thus,

$$|(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_i|^{1/(2i-1)} \leq C_3^{1/(2i-1)} \left(\sum_{j=1}^{i} \left(|(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j| |t-s|^{i-j} \right)^{1/(2i-1)} \right.$$
$$\left. + \sum_{j=i+1}^{n} \left(|(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j| |t-s| \right)^{1/(2i-1)} \right).$$

For the first contribution of the r.h.s. we can use again Young's inequality with $p_j = \frac{2i-1}{2j-1}$, $q_j = \frac{2i-1}{2(i-j)}$. For the second contribution we exploit that since $\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) \leq \Lambda \leq 1$ then for all $j \in [\![1, n]\!]$, $|(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j| \leq 1$ which for $j \geq i+1$ yields $|(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j|^{1/(2i-1)} \leq |(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_j|^{1/(2j-1)}$. We therefore get up to a modification of C_4 that:

$$|(\mathbf{x} - \boldsymbol{\theta}_{s,t}(\mathbf{y}))_i|^{1/(2i-1)} \leq C_4 \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})),$$

ACCEPTED MANUSCRIPT

which together with (6.3) indeed gives that there exists $C_{4.1} := C_{4.1}((\mathbf{A}), T))$ s.t. $\mathbf{d}((t,\mathbf{y}),(s,\mathbf{x})) \leq C_{4.1}\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))$ for $(s,\mathbf{x}),(t,\mathbf{y}) \in S$, $\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) \leq \Lambda \leq 1$ which is the first part of a). It remains to prove the quasi-triangle inequality. Recalling that ρ defined in (4.1) satisfies the quasi-triangle inequality, let us write:

$$\begin{aligned}\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) &= \rho(t-\sigma+\sigma-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})+\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})-\mathbf{y})\\ &\leq K(\rho(\sigma-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}))+\rho(t-\sigma,\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})-\mathbf{y}))\\ &:= K(\rho(\sigma-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}))+\mathbf{d}((\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))).\end{aligned}$$

(6.5)

On the other hand, using the specific form of $\mathbf{F}$ in the dynamics of $\boldsymbol{\theta}$, we can derive similarly to (6.2), (6.4) using the direct forward dynamics that for all $i \in [\![1,n]\!]$:

$$\begin{aligned}&|(\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}))_i|\\ &\leq C_3\left(\sum_{j=1}^{i}|(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})_j||t-\sigma|^{i-j}+\sum_{j=i+1}^{n}|(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})_j||t-\sigma|\right).\end{aligned}$$

Thus, using as above Young inequalities and the fact that $\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})) \leq \Lambda \leq 1$ we get for all $i \in [\![1,n]\!]$,

$$\begin{aligned}&|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}))_i|^{1/(2i-1)}\\ \leq\ & C_3^{1/(2i-1)}\Bigg(\sum_{j=1}^{i}\left[\frac{|(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})_j|)^{1/(2j-1)}}{p_j}+\frac{|t-\sigma|^{1/2}+|\sigma-s|^{1/2}}{q_j}\right]\\ &+\sum_{j=i+1}^{n}|(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})_j|^{1/(2j-1)}|t-\sigma|^{1/(2i-1)}\Bigg)\\ \leq\ & C_4(\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))+\mathbf{d}((\sigma,\boldsymbol{\xi}),(t,\mathbf{y})),\end{aligned}$$

with $p_j = \frac{2i-1}{2j-1}$, $q_j = \frac{2i-1}{2(i-j)}$ in the last but one equality. Hence $\rho(\sigma-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})) \leq C_5(\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))+\mathbf{d}((\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))$, $C_5 := C_5((\mathbf{A}),T)$, which plugged into (6.5) concludes the proof up to a modification of $C_{4.1}$. □

Remark 6.1 (Subdiagonal structure). Observe from the previous proof that when the function $\mathbf{F}$ in the dynamics of $\boldsymbol{\theta}$ has the following structure, $\mathbf{F}_1(t,\mathbf{x}) = \mathbf{F}_1(t,\mathbf{x}_1)$, and for all $i \in [\![2,n]\!]$, $\mathbf{F}_i(t,\mathbf{x}^{i-1,n}) = \mathbf{F}_i(t,\mathbf{x}_{i-1},\mathbf{x}_i)$, then the terms in $j \in [\![i+1,n]\!]$ do not appear in equation (6.4). Hence, the distances respectively associated with the forward and backward transport are actually equivalent.

6.2. **Controls on the flows and the frozen kernel.** We first state a Lemma that gives some controls and equivalences for the scaled forward, backward and linearized flows, specifying the estimates given in (3.10).

Lemma 6.1 (Controls and Equivalences of the scaled flows). There exists a constant $C := C(T,(\mathbf{A}))$ s.t. for all $0 \leq s \leq u < v \leq t \leq T$, $(\mathbf{x},\mathbf{y}) \in (\mathbb{R}^{nd})^2$, $w \geq v-u$,

$$C^{-1}|\mathbb{T}_w^{-1}(\mathbf{x}-\boldsymbol{\theta}_{u,v}(\mathbf{y}))| \leq |\mathbb{T}_w^{-1}(\boldsymbol{\theta}_{v,u}(\mathbf{x})-\mathbf{y})| \leq C|\mathbb{T}_w^{-1}(\mathbf{x}-\boldsymbol{\theta}_{u,v}(\mathbf{y}))|.$$

We also have:

$$C^{-1}|\mathbb{T}_{t-s}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))| \leq |\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})| \leq C|\mathbb{T}_{t-s}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|.$$

ACCEPTED MANUSCRIPT

Proof. The first control can be derived from the structure of the drift in (1.1) using Gronwall's Lemma. Indeed, similarly to (6.1):

$$\begin{aligned}|\mathbb{T}_w^{-1}(\boldsymbol{\theta}_{v,u}(\mathbf{x})-\mathbf{y})| &= |\mathbb{T}_w^{-1}(\mathbf{x}-\boldsymbol{\theta}_{u,v}(\mathbf{y}))+\int_u^v \mathbb{T}_w^{-1}(\mathbf{F}(r,\boldsymbol{\theta}_{r,u}(\mathbf{x}))-\mathbf{F}(r,\boldsymbol{\theta}_{r,v}(\mathbf{y})))dr| \\ &\leq |\mathbb{T}_w^{-1}(\mathbf{x}-\boldsymbol{\theta}_{u,v}(\mathbf{y}))|+Cw^{-1}\int_u^v |\mathbb{T}_w^{-1}(\boldsymbol{\theta}_{r,u}(\mathbf{x})-\boldsymbol{\theta}_{r,v}(\mathbf{y}))|dr \\ &\leq \exp(Cw^{-1}\int_u^v dr)|\mathbb{T}_w^{-1}(\mathbf{x}-\boldsymbol{\theta}_{u,v}(\mathbf{y}))|,\end{aligned}$$

where $C := C(T,(\mathbf{A}))$. Since $v-u \leq w$, this gives the r.h.s. The l.h.s. is proved similarly.

To prove the second control, we need the following auxiliary yet important Scaling Lemma.

Lemma 6.2 (Scaling Lemma). Let $0 \leq s < t \leq T$ and $(s_0,\mathbf{y}_0) \in [0,T]\times\mathbb{R}^{nd}$ be given. The resolvent matrices in (3.4) can be written in the following way: for all $u \in [s,t]$,

$$\begin{aligned}\tilde{\mathbf{R}}^{s_0,\mathbf{y}_0}(u,s) &= \mathbb{T}_{t-s}\bar{\mathbf{R}}^{s,t,(s_0,\mathbf{y}_0)}_{\frac{u-s}{t-s}}\mathbb{T}^{-1}_{t-s}, \\ \tilde{\mathbf{R}}^{s_0,\mathbf{y}_0}(s,u) &= \mathbb{T}_{t-s}\bar{\mathbf{R}}^{s,t,(s_0,\mathbf{y}_0)}_{\frac{s-u}{t-s}}\mathbb{T}^{-1}_{t-s},\end{aligned}$$

where there exists $C_{6.2} := C_{6.2}(T,(\mathbf{A})) > 1$ s.t. for all $(u,v) \in [s,t]^2$, $|\bar{\mathbf{R}}^{s,t,(s_0,\mathbf{y}_0)}_{\frac{u-v}{t-s}}| \leq C_{6.2}$.

Proof. Let us define for $w \in [0,1]$, $\bar{\mathbf{R}}^{s,t,(s_0,\mathbf{y}_0)}_w := \mathbb{T}^{-1}_{t-s}\tilde{\mathbf{R}}^{s_0,\mathbf{y}_0}(s+w(t-s),s)\mathbb{T}_{t-s}$. Observe from the differential dynamics in (3.4) that:

$$\begin{aligned}\partial_w\bar{\mathbf{R}}^{s,t,(s_0,\mathbf{y}_0)}_w &= \mathbb{T}^{-1}_{t-s}(t-s)D\mathbf{F}(s+w(t-s),\boldsymbol{\theta}_{s+w(t-s),s_0}(\mathbf{y}_0))\mathbb{T}_{t-s} \\ &\quad\times\left[\mathbb{T}^{-1}_{t-s}\tilde{\mathbf{R}}^{s_0,\mathbf{y}_0}(s+w(t-s),s)\mathbb{T}_{t-s}\right] \\ &= \left\{\mathbb{T}^{-1}_{t-s}(t-s)D\mathbf{F}(s+w(t-s),\boldsymbol{\theta}_{s+w(t-s),s_0}(\mathbf{y}_0))\mathbb{T}_{t-s}\right\}\bar{\mathbf{R}}^{s,t,(s_0,\mathbf{y}_0)}_w.\end{aligned}$$

Setting $\alpha^{s,t,(s_0,\mathbf{y}_0)}_w := \left\{\mathbb{T}^{-1}_{t-s}(t-s)D\mathbf{F}(s+w(t-s),\boldsymbol{\theta}_{s+w(t-s),s_0}(\mathbf{y}_0))\mathbb{T}_{t-s}\right\}$, we derive from the subdiagonal structure of the partial gradient $D\mathbf{F}$ that $|\alpha^{s,t,(s_0,\mathbf{y}_0)}_w| \leq C := C((\mathbf{A}))$. This gives the first statement taking $w = \frac{u-s}{t-s}$. The second follows by symmetry setting, for $w \in [-1,0]$, $\bar{\mathbf{R}}^{s,t,(s_0,\mathbf{y}_0)}_w := \mathbb{T}^{-1}_{t-s}\tilde{\mathbf{R}}^{s_0,\mathbf{y}_0}(s,s-w(t-s))\mathbb{T}_{t-s}$, differentiating in w as above and taking eventually $w = -\frac{u-s}{t-s}$. □

Observe now from equation (3.5) that

$$\tilde{\mathbf{R}}^{t,\mathbf{y}}(s,t)(\mathbf{y}-\tilde{\mathbf{m}}^{t,\mathbf{y}}(s,t)) = \boldsymbol{\theta}_{s,t}(\mathbf{y}), \tag{6.6}$$

i.e. we get the pull-back by the deterministic system of the final point $\mathbf{y}$ from t to s. Hence, using Lemma 6.2, we obtain:

$$\begin{aligned}|\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})| = |\mathbb{T}^{-1}_{t-s}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))| &= |\bar{\mathbf{R}}^{s,t,(t,\mathbf{y})}_1\mathbb{T}^{-1}_{t-s}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))| \\ &\leq C_{6.2}|\mathbb{T}^{-1}_{t-s}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|.\end{aligned}$$

This gives the r.h.s of the second statement in Lemma 6.1. Once again, the l.h.s. can be proved similarly. □

ACCEPTED MANUSCRIPT

As a consequence of Lemma 6.2, we obtain the following controls for the derivatives of the *frozen* density (3.6) (see also the arguments in Section 5 of [DM10])

Proposition 6.1 (Derivatives of the Frozen Density). There exist constants $C_{6.1} := C_{6.1}(T,(\mathbf{A}))$, $c_{6.1} := c_{6.1}(T,(\mathbf{A}))$ s.t. for all multi index $\alpha = (\alpha_1,\cdots,\alpha_n) \in \mathbb{N}^n$, $|\alpha| := \sum_{i=1}^n \alpha_i \leq 3$ we have $\forall 0 \leq s < t \leq T, \forall (\mathbf{x},\mathbf{y}) \in (\mathbb{R}^{nd})^2$,

$$|\partial_{\mathbf{x}}^{\alpha} \tilde{q}(s,t,\mathbf{x},\mathbf{y})| \leq \frac{C_{6.1}}{(t-s)^{\sum_{i=1}^n \frac{(2i-1)\alpha_i}{2}}} q_{c_{6.1}}(s,t,\mathbf{x},\mathbf{y}),$$

$$\forall c>0,\ q_c(s,t,\mathbf{x},\mathbf{y}) := \frac{c^{nd/2}}{(2\pi)^{nd/2}(t-s)^{n^2d/2}} \exp\left(-\frac{c}{2}(t-s)|\mathbb{T}_{t-s}^{-1}(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\mathbf{y})|^2\right).$$

6.3. **Proof of Proposition 4.2.** This section is devoted to the proof of Proposition 4.2 which provides the key estimates to derive Theorem 3.1 under assumption **(A)**.

6.3.1. *Some Preliminary Notations and Control of the Linearization Error.* Introduce first $\Sigma_1 := \{(u,\mathbf{z}) \in \mathbb{R}\times\mathbb{R}^{nd} : \rho(u,\mathbf{z}) = 1\}$ with ρ defined in (4.1), i.e. Σ_1 is the level curve at 1 of the parabolic metric. With this definition we can introduce, for given points $(s,\mathbf{x}),(t,\mathbf{y}),(\sigma,\boldsymbol{\xi}) \in [-T,T]\times\mathbb{R}^{nd}$, the mappings:

$$J_1 : (t,\mathbf{y}) \in [-T,T]\times\mathbb{R}^{nd} \mapsto \left(s+\rho^2\tilde{s}, \boldsymbol{\theta}_{t,s}(\mathbf{x}) + \rho^{-1}\mathbb{T}_{\rho^2}\tilde{\mathbf{x}}\right),\ \rho := \rho(t-s,\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x}))$$
$$(\tilde{s},\tilde{\mathbf{x}}) := \left((t-s)\rho^{-2}, \rho\mathbb{T}_{\rho^{-2}}\left(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})\right)\right) \in \Sigma_1. \tag{6.7}$$

$$J_2 : (\sigma,\boldsymbol{\xi}) \in [-T,T]\times\mathbb{R}^{nd} \mapsto \left(s+\alpha^2\bar{s}, \boldsymbol{\theta}_{\sigma,s}(\mathbf{x}) + \alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}\right),\ \alpha := \rho(\sigma-s,\boldsymbol{\xi}-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))$$
$$(\bar{s},\bar{\mathbf{x}}) := \left((\sigma-s)\alpha^{-2}, \alpha\mathbb{T}_{\alpha^{-2}}\left(\boldsymbol{\xi}-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})\right)\right) \in \Sigma_1. \tag{6.8}$$

Define now

$$\mathcal{R}_{t,s}^{\rho}(\mathbf{x},\mathbf{y}) := \rho\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})). \tag{6.9}$$

From a stability analysis similar to the one of equations (A-8), (A-10) in [Men11], we derive the following lemma, which allows to control the linearization error.

Lemma 6.3. Assume that **(A)** holds and that $0 < \rho \leq \Lambda$ for some $\Lambda \leq 1$. Then, there exists $C_{6.3} := C_{6.3}((\mathbf{A}))$ s.t. with the notation of (6.9):

$$|\mathcal{R}_{t,s}^{\rho}(\mathbf{x},\mathbf{y})| \leq C_{6.3}(\rho^{\eta}+(t-s))|\tilde{\mathbf{x}}|. \tag{6.10}$$

Also, if $\rho = \rho(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}) \geq c_{\infty}\alpha = c_{\infty}\rho(\sigma-s,\boldsymbol{\xi}-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))$, then

$$|\mathcal{R}_{t,\sigma}^{\rho}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})| \leq C_{6.3}(\rho^{\eta}+(t-\sigma))|\tilde{\mathbf{x}}|. \tag{6.11}$$

This implies that taking $j_0 \in [\![1,n]\!]$ s.t. $|\tilde{\mathbf{x}}_{j_0}| := \sup_{i\in[\![1,n]\!]}|\tilde{\mathbf{x}}_i|$ one has for all $j \in [\![1,n]\!]$:

$$|\{\mathcal{R}_{t,s}^{\rho}(\mathbf{x},\mathbf{y})\}_j| \leq C_{6.3}n(\Lambda^{\eta}+(t-s))|\tilde{\mathbf{x}}_{j_0}|,$$
$$|\{\mathcal{R}_{t,\sigma}^{\rho}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})\}_j| \leq C_{6.3}n(\Lambda^{\eta}+(t-\sigma))|\tilde{\mathbf{x}}_{j_0}|. \tag{6.12}$$

Eventually, if $|\sigma-s| \leq K|t-\sigma|$, we also have

$$(t-\sigma)^{1/2}|\mathbb{T}_{t-\sigma}^{-1}(\tilde{\boldsymbol{\theta}}_{\sigma,s}^{t,\mathbf{y}}(\mathbf{x})-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))| \leq C\{\rho^{\eta}+|\sigma-s|\}(t-\sigma)^{1/2}|\mathbb{T}_{t-\sigma}^{-1}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|. \tag{6.13}$$

ACCEPTED MANUSCRIPT

Proof. Let us prove (6.10). Recalling equations (3.2), (3.3) and (3.14), (3.15), we write:

$$\begin{aligned}\mathcal{R}^{\rho}_{t,s}(\mathbf{x},\mathbf{y}) &:= \rho\mathbb{T}_{\rho^{-2}}\left\{\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})\right\} = \\ &\rho\mathbb{T}_{\rho^{-2}}\left\{\int_s^t du\Big[\Big(\mathbf{F}(u,\boldsymbol{\theta}_{u,s}(\mathbf{x})) - \mathbf{F}^{t,\mathbf{y}}(u,\boldsymbol{\theta}_{u,s}(\mathbf{x}))\Big)\right. \\ &\qquad + \Big(D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}))(\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{u,s}(\mathbf{x}))\Big) \\ &\quad + \Big(\int_0^1 d\delta\left(D\mathbf{F}^{t,\mathbf{y}}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}) + \delta(\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}(\mathbf{y}))\right) \\ &\qquad \left.\left. - D\mathbf{F}^{t,\mathbf{y}}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}))\right)(\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}(\mathbf{y}))\Big)\Big]\right\} \\ &:= (\mathcal{R}^{\rho,1}_{t,s} + \mathcal{R}^{\rho,2}_{t,s} + \mathcal{R}^{\rho,3}_{t,s})(\mathbf{x},\mathbf{y}), \end{aligned} \tag{6.14}$$

where, accordingly with the notations of (3.14), for $(u,\mathbf{z}) \in [s,t]\times\mathbb{R}^{nd}$, $D\mathbf{F}^{t,\mathbf{y}}(u,\mathbf{z})$ is the $(nd)\times(nd)$ matrix with only non zero $d\times d$ matrix entries $(D\mathbf{F}^{t,\mathbf{y}}(u,\mathbf{z}))_{j,j-1} := D_{\mathbf{x}_{j-1}}\mathbf{F}_j(u,\mathbf{z}_{j-1},\boldsymbol{\theta}_{u,t}(\mathbf{y})^{j,n})$, $j \in [\![2,n]\!]$, so that in particular $D\mathbf{F}^{t,\mathbf{y}}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y})) = D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}))$.

The structure of the "partial gradient" $D\mathbf{F}^{t,\mathbf{y}}$ and its Hölder property yield that there exists $C_3 := C_3(T,(\mathbf{A}))$ s.t. for all $j \in [\![2,n]\!]$:

$$\begin{aligned}|(\mathcal{R}^{\rho,3}_{t,s}(\mathbf{x},\mathbf{y}))_j| &\le C_3\rho^{1-2j}\int_s^t du|(\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}(\mathbf{y}))_{j-1}||\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}(\mathbf{y})|^{\eta} \\ &\le C_3|\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y}|^{\eta}\rho^{-2}\int_s^t du(\sum_{k=2}^n \rho^{1-2(k-1)}|(\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}(\mathbf{y}))_{k-1}|).\end{aligned}$$

Since $\rho = \rho(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y}) = d\big((s,\mathbf{x}),(t,\mathbf{y})\big) \le \Lambda \le 1$ we derive:

$$|\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y}| \le C\sum_{i=1}^n |(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})_i|^{1/(2i-1)} \le C\rho,\ C := C(n).$$

Hence, up to modifications of C_3,

$$\begin{aligned}|(\mathcal{R}^{\rho,3}_{t,s}(\mathbf{x},\mathbf{y}))_j| &\le C_3\rho^{\eta}\times\rho^{-2}\int_s^t du(\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{u,s}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}(\mathbf{y}))|) \\ &\le C_3\rho^{-2+\eta}\int_s^t du(\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})|) \\ &\le C_3\rho^{1+\eta}|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y})|, \end{aligned} \tag{6.15}$$

using Lemma 6.1 for the last but one inequality and recalling from (6.7) that $|t-s|/\rho^2 \le 1$ for the last one.

ACCEPTED MANUSCRIPT

On the other hand, the term $\mathcal{R}^{\rho,1}_{t,s}(\mathbf{x},\mathbf{y})$ can be seen as a remainder w.r.t. the characteristic time scales. Precisely, there exists $C_1 := C_1(T,(\mathbf{A}))$ (possibly changing from line to line) s.t. for all $j \in [\![1,n]\!]$:

$$\begin{aligned}|(\mathcal{R}^{\rho,1}_{t,s}(\mathbf{x},\mathbf{y}))_j| &\leq C_1\rho^{1-2j}\int_s^t du\sum_{k=j}^n|(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\boldsymbol{\theta}_{u,t}(\mathbf{y}))_k|\\ &\leq C_1\int_s^t du\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\boldsymbol{\theta}_{u,t}(\mathbf{y}))|\\ &\leq C_1(t-s)\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})| \qquad (6.16)\end{aligned}$$

using once again Lemma 6.1 for the last inequality.

Recall now that $\mathcal{R}^{\rho,2}_{t,s}(\mathbf{x},\mathbf{y})$ is the linear part of equation (6.14). Setting

$$\begin{aligned}\forall u\in[s,t],\ \alpha^\rho_{t,u}(\mathbf{y}) &:= \left\{\rho\mathbb{T}_{\rho^{-2}}D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}))\rho^{-1}\mathbb{T}_{\rho^2}\right\},\\ \tilde{\mathcal{R}}^{\rho,t}_{u,s}(\mathbf{x},\mathbf{y}) &:= \rho\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{u,s}(\mathbf{x})),\end{aligned}$$

it can be rewritten

$$\begin{aligned}\mathcal{R}^{\rho,2}_{t,s}(\mathbf{x},\mathbf{y}) &= \int_s^t du\left\{\rho\mathbb{T}_{\rho^{-2}}D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}))\rho^{-1}\mathbb{T}_{\rho^2}\right\}\left(\rho\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{u,s}(\mathbf{x}))\right)\\ &= \int_s^t du\alpha^\rho_{t,u}(\mathbf{y})\left(\rho\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{u,s}(\mathbf{x}))\right)\\ &= \int_s^t du\alpha^\rho_{t,u}(\mathbf{y})\tilde{\mathcal{R}}^{\rho,t}_{u,s}(\mathbf{x},\mathbf{y}),\end{aligned}$$

where there exists a constant $C_2 := C_2(T,(\mathbf{A}))$ s.t. $\int_s^t du|\alpha^\rho_{t,u}(\mathbf{y})| \leq C_2$. From (6.16), (6.15), (6.14), observing as well that $\mathcal{R}^\rho_{t,s}(\mathbf{x},\mathbf{y}) = \tilde{\mathcal{R}}^{\rho,t}_{t,s}(\mathbf{x},\mathbf{y})$, we derive from Gronwall's Lemma:

$$\begin{aligned}\exists C_4 := C_4(T,(\mathbf{A})),\ |\mathcal{R}^\rho_{t,s}(\mathbf{y})| &\leq C_4(\rho^\eta+(t-s))\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|\\ &\leq C_4(\rho^\eta+(t-s))|\tilde{\mathbf{x}}|\end{aligned}$$

recalling (6.7) for the last inequality. This gives equation (6.10) of the Lemma. For (6.11), the previous proof can be adapted with obvious modifications using thoroughly that $\boldsymbol{\theta}_{u,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})) := \boldsymbol{\theta}_{u,s}(\mathbf{x})$ and Lemma 6.1. The main differences are that the time integrals are taken between σ and t. Following the computations leading to (6.15), the contribution $\mathcal{R}^{\rho,3}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})$ would be bounded by $C_3\rho^{-2+\eta}|t-\sigma|\{\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|\} \leq C_3\rho^{-2+\eta}(|t-s|+|s-\sigma|)\{\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|\}$. Recalling also that $|t-s|+|\sigma-s| \leq \rho^2+\alpha^2 \leq (1+c_\infty)\rho^2$ on the considered set, we get that (6.15) still holds in that case. We would similarly have $|\mathcal{R}^{\rho,1}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})| \leq C_1(t-\sigma)\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|$ giving (6.16) in that case. Eventually, the same previous triangle inequality would give that, on the considered set $\int_\sigma^t du|\alpha^\rho_{t,y}(\mathbf{y})| \leq C$, so that $\mathcal{R}^{\rho,2}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})$ can still be viewed as the well controlled linear part of the inequality. The proof then again follows from Gronwall's lemma.

Eventually, (6.13) is established similarly exploiting again Lemma 6.1 and the condition $|\sigma-s| \leq K|t-\sigma|$ which is needed to control the linear part through Gronwall's lemma. □

ACCEPTED MANUSCRIPT

6.3.2. *Main Proof of Proposition 4.2.* Point *i)* can be derived for both kernels k_{ij}^d, $k_{ij}^{d,*}$ recalling from (4.5) and Lemma 6.1 that there exists $C := C((\mathbf{A}), T) \geq 1$ s.t. $\forall((s,\mathbf{x}),(t,\mathbf{y})) \in S^2$,

$$\begin{aligned} &|k_{ij}^d(s,t,\mathbf{x},\mathbf{y})| + |k_{ij}^{d,*}(s,t,\mathbf{x},\mathbf{y})| \\ &\leq \frac{C}{(t-s)^{1+n^2d/2}} \exp\left(-C^{-1}(t-s)|\mathbb{T}_{t-s}^{-1}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})^2|\right). \end{aligned} \tag{6.17}$$

Now for a given $c_1 > 2$, if $c_1|t-s|^{1/2} > \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))$ then the r.h.s of (6.17) can directly be upper bounded by $Cc_1^{2+n^2d}/\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))^{2+n^2d}$. On the other hand, if $c_1|t-s|^{1/2} \leq \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))$ then, by definition of d in (4.2) we derive that $\exists i_0 \in [\![1,n]\!]$ s.t.

$$|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_{i_0}|^{1/(2i_0-1)} \geq \frac{1}{n}(1-\frac{1}{c_1})\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})).$$

This property yields:

$$\begin{aligned} |k_{ij}^d(s,t,\mathbf{x},\mathbf{y})| + |k_{ij}^{d,*}(s,t,\mathbf{x},\mathbf{y})| &\leq \frac{C}{|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_{i_0}|^{\frac{n^2d+2}{2i_0-1}}} \left(\frac{|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_{i_0}|}{|t-s|^{1/2(2i_0-1)}}\right)^{\frac{n^2d+2}{2i_0-1}} \\ &\quad \times \exp\left(-C^{-1}(t-s)|\mathbb{T}_{t-s}^{-1}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})^2|\right) \\ &\leq \frac{\tilde{C}}{\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))^{n^2d+2}} \exp\left(-\bar{C}^{-1}(t-s)|\mathbb{T}_{t-s}^{-1}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})^2|\right), \end{aligned}$$

where $\tilde{C} := \tilde{C}((\mathbf{A}), T, c_1)$, $\bar{C} := \bar{C}((\mathbf{A}), T)$. This gives the first claim.

Let us now establish point *ii)* for the kernel k_{ij}^d, recalling that below

$$c_\infty \mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})) \leq \mathbf{d}((s,\mathbf{x}),(t,\mathbf{y})) \leq \Lambda \leq 1, \tag{6.18}$$

for $c_\infty := c_\infty((\mathbf{A}))$ large enough and Λ small enough to be specified later on. We can w.l.o.g. assume that $|\sigma - s| \leq K|t-\sigma|$, for some $K := K((\mathbf{A}), T) > 1$ and write:

$$\begin{aligned} |k_{ij}(s,t,\mathbf{x},\mathbf{y}) - k_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y})| &\leq |k_{ij}(s,t,\mathbf{x},\mathbf{y}) - k_{ij}(\sigma,t,\tilde{\boldsymbol{\theta}}_{\sigma,s}^{t,\mathbf{y}}(\mathbf{x}),\mathbf{y})| \\ &\quad + |k_{ij}(\sigma,t,\tilde{\boldsymbol{\theta}}_{\sigma,s}^{t,\mathbf{y}}(\mathbf{x}),\mathbf{y}) - k_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y})| \\ &:= I_1(s,\sigma,t,\mathbf{x},\mathbf{y}) + I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y}). \end{aligned} \tag{6.19}$$

Remark 6.2. The previous splitting of $|k_{ij}(s,t,\mathbf{x},\mathbf{y}) - k_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y})|$ has been done to separate the time and space sensitivities. In I_1 the space variable is frozen and from (4.4) and the flow property of $\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}$ its value is equal to $\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x}) - \mathbf{y}$. In I_2 the time variables are equal to $t-\sigma$. Also, the intermediate spatial point $\tilde{\boldsymbol{\theta}}_{\sigma,s}^{t,\mathbf{y}}(\mathbf{x})$ yields from I_2 a difference of the form $|\tilde{\boldsymbol{\theta}}_{\sigma,s}^{t,\mathbf{y}}(\mathbf{x}) - \boldsymbol{\xi}|$ which up to a linearization error has the same order as $|\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}) - \boldsymbol{\xi}|$, norm of the spatial point appearing in (6.18). The condition $|\sigma - s| \leq K|t-\sigma|$ is here needed to use properties on the rescaled flows for the spatial sensitivity (see Lemma 6.3 and equation (6.25)) which allow to control the linearization error. We emphasize that if $|\sigma - s| \geq K|t-\sigma|$ (and therefore $|t-s| \geq (1-1/K)|\sigma - s|$) then the integrand has to be split differently,

ACCEPTED MANUSCRIPT

writing

$$|k_{ij}(s,t,\mathbf{x},\mathbf{y})-k_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y})|\leq |k_{ij}(s,t,\mathbf{x},\mathbf{y})-k_{ij}(s,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{s,\sigma}(\boldsymbol{\xi}),\mathbf{y})|$$
$$+|k_{ij}(s,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{s,\sigma}(\boldsymbol{\xi}),\mathbf{y})-k_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y})|.$$

The above terms could be analyzed similarly to those appearing in (6.19) following the procedure below.

Setting for all $-T\leq s<t\leq T,\ (\mathbf{z},\mathbf{y})\in(\mathbb{R}^{nd})^2$:

$$\begin{aligned}\bar{k}(s,t,\mathbf{z},\mathbf{y}) &:= \mathbb{I}_{-T\leq s<t\leq T}\Big(-[\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)^{*}\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)]_{11}\\ &\quad+[\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,s)^{*}\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}(\mathbf{z}-\mathbf{y})]_1^{\otimes 2}\Big)\times\Big(\frac{1}{(2\pi)^{nd}\det(\tilde{\mathbf{K}}^{\mathbf{y}}(s,t))^{1/2}}\\ &\quad\times\exp(-\frac{1}{2}\langle\tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}(\mathbf{z}-\mathbf{y}),\mathbf{z}-\mathbf{y}\rangle)\Big),\end{aligned}$$

(4.4) and (6.19) yield $I_1(s,\sigma,t,\mathbf{x},\mathbf{y})=|\bar{k}_{ij}(s,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}),\mathbf{y})-\bar{k}_{ij}(\sigma,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}),\mathbf{y})|$. Thus, from (4.4), we derive similarly to (4.5) (see also the proof of Proposition 3.7 in [DM10] for a thorough discussion on the time sensitivities of the covariance matrix) that $\exists(c,C):=(c,C)(T,(\mathbf{A}))>0$ s.t.

$$\begin{aligned}I_1(s,\sigma,t,\mathbf{x},\mathbf{y}) &\leq |s-\sigma|\sup_{\tau\in[s\wedge\sigma,(s\vee\sigma)\wedge t[}|\partial_\tau\bar{k}_{ij}(\tau,t,\mathbf{z},\mathbf{y})|_{\mathbf{z}=\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})}\\ &\leq C|s-\sigma|\times\sup_{\tau\in[s\wedge\sigma,(s\vee\sigma)\wedge t[}\frac{\bar{q}_c(\tau,t,\mathbf{z},\mathbf{y})}{(t-\tau)^2}|_{\mathbf{z}=\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})},\end{aligned}\tag{6.20}$$

where $\bar{q}_c(\tau,t,\mathbf{z},\mathbf{y}):=\frac{c^{nd/2}}{(2\pi)^{nd/2}(t-\tau)^{n^2d/2}}\exp(-\frac{c}{2}(t-\tau)|\mathbb{T}^{-1}_{t-\tau}(\mathbf{z}-\mathbf{y})|^2)$.

We have to consider the terms I_1,I_2 under the condition (6.18) that rewrites $\{(t,\mathbf{y})\in[0,T]\times\mathbb{R}^{nd}:\rho=\rho(t-s,\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x}))\geq c_\infty\alpha=c_\infty\rho(\sigma-s,\boldsymbol{\xi}-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))\}$.

From (6.7) we get that for all $\tau\in[s\wedge\sigma,(s\vee\sigma)\wedge t[$:

$$\begin{aligned}\frac{\bar{q}_c(\tau,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}),\mathbf{y})}{(t-\tau)^2} &\leq \frac{C}{\rho^{4+n^2d}}\frac{1}{\Big(\tilde{s}-\frac{\tau-s}{\rho^2}\Big)^{2+n^2d/2}}\\ &\quad\times\exp\left(-c(t-\tau)|\mathbb{T}^{-1}_{t-\tau}(\rho^{-1}\mathbb{T}_{\rho^2}\{\tilde{\mathbf{x}}+\mathcal{R}^{\rho}_{t,s}(\mathbf{x},\mathbf{y})\})|^2\right),\end{aligned}\tag{6.21}$$

using the notation introduced in (6.9), i.e. the term $\mathcal{R}^{\rho}_{t,s}(\mathbf{x},\mathbf{y})$ measures the difference associated with the approximation of the non-linear flow by the affine one.

Now, if $|\tilde{s}-\frac{\tau-s}{\rho^2}|^{1/2}\geq\frac{1}{c_\infty}=:\tilde{c}$, for $c_\infty\geq 1$ to be specified later on, we have from (6.21) and (6.20) that $I_1(s,\sigma,t,\mathbf{x},\mathbf{y})\leq\frac{C|s-\sigma|}{\rho^{4+n^2d}}\leq\frac{C\alpha^2}{\rho^{4+n^2d}}$ using (6.8) for the last inequality. On the other hand, from (6.18) $\rho\geq c_\infty\alpha$. Hence, $|\tau-s|\leq|\sigma-s|\leq\alpha^2\leq\frac{\rho^2}{c_\infty^2}$. Since $(\tilde{s},\tilde{\mathbf{x}})\in\Sigma_1$, we thus derive:

$$\sum_{j=1}^{n}|\tilde{\mathbf{x}}_j|^{1/(2j-1)}+\left|\tilde{s}-\frac{\tau-s}{\rho^2}\right|^{1/2}\geq 1-\left|\frac{(\tau-s)}{\rho^2}\right|^{1/2}\geq 1-\frac{1}{c_\infty}=1-\tilde{c}.$$

ACCEPTED MANUSCRIPT

Hence, for $|\tilde{s}-\frac{\tau-s}{\rho^2}|^{1/2}\le\tilde{c}$, we obtain

$$\sum_{j=1}^{n}|\tilde{\mathbf{x}}_j|^{1/(2j-1)}\ge 1-2\tilde{c} \text{ and } \exists j_0\in[\![1,n]\!],\ |\tilde{\mathbf{x}}_{j_0}|^{1/(2j_0-1)}\ge\frac{1-2\tilde{c}}{n}>0, \text{ for } c_\infty>2. \tag{6.22}$$

Write now:

$$(t-\tau)|\mathbb{T}^{-1}_{t-\tau}(\rho^{-1}\mathbb{T}_{\rho^2}\{\tilde{\mathbf{x}}+\mathcal{R}^\rho_{t,s}(\mathbf{x},\mathbf{y})\})|^2=\sum_{j=1}^{n}\left(\tilde{s}-\frac{\tau-s}{\rho^2}\right)^{-(2j-1)}|\{\tilde{\mathbf{x}}+\mathcal{R}^\rho_{t,s}(\mathbf{x},\mathbf{y})\}_j|^2.$$

Thus, we get from Lemma 6.3, equation (6.12), that for T and Λ s.t. $C_{6.3}n(\Lambda^\eta+(t-s))\le 1/2$:

$$\begin{aligned}&(t-\tau)|\mathbb{T}^{-1}_{t-\tau}(\rho^{-1}\mathbb{T}_{\rho^2}\{\tilde{\mathbf{x}}+\mathcal{R}^\rho_{t,s}(\mathbf{x},\mathbf{y})\})|^2\\&\ge\left(\tilde{s}-\frac{\tau-s}{\rho^2}\right)^{-(2j_0-1)}|\tilde{\mathbf{x}}_{j_0}|^2\left(\frac{1}{2}-(C_{6.3}n(\Lambda^\eta+(t-s)))^2\right)\\&\ge\frac{1}{4}\left(\tilde{s}-\frac{\tau-s}{\rho^2}\right)^{-(2j_0-1)}\left(\frac{1-2\tilde{c}}{n}\right)^{2(2j_0-1)},\end{aligned}$$

using (6.22) for the last inequality. Plugging the above control into (6.21) yields:

$$\begin{aligned}&\frac{\bar{q}_c(\tau,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}),\mathbf{y})}{(t-\tau)^2}\le\frac{C}{\rho^{4+n^2d}}\frac{1}{\left(\tilde{s}-\frac{\tau-s}{\rho^2}\right)^{2+n^2d/2}}\\&\times\exp\left(-\frac{c}{4}\left(\tilde{s}-\frac{\tau-s}{\rho^2}\right)^{-(2j_0-1)}\left(\frac{1-2\tilde{c}}{n}\right)^{2(2j_0-1)}\right)\le\frac{C}{\rho^{4+n^2d}}.\end{aligned}\tag{6.23}$$

From (6.23) and (6.20) we finally get the global bound:

$$\exists C_1:=C_1(T,(\mathbf{A}))>0,\ I_1(s,\sigma,t,\mathbf{x},\mathbf{y})\le\frac{C_1\alpha^2}{\rho^{4+n^2d}}. \tag{6.24}$$

Let us now turn to $I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y})$. From Proposition 6.1 and (4.4), we get similarly to (6.20) that $\exists(c,C):=(c,C)(T,(\mathbf{A}))>0$ s.t.:

$$\begin{aligned}I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y})\ &\le\ C(t-\sigma)^{1/2}|\mathbb{T}^{-1}_{t-\sigma}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}))|\\&\quad\times\frac{1}{(t-\sigma)}\sup_{\gamma\in[0,1]}\bar{q}_c(\sigma,t,\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}),\mathbf{y})\\&\le\ C(t-\sigma)^{1/2}|\mathbb{T}^{-1}_{t-\sigma}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})|\\&\quad\times\frac{1}{(t-\sigma)}\sup_{\gamma\in[0,1]}\bar{q}_c(\sigma,t,\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}),\mathbf{y})\\&\le\ C(t-\sigma)^{1/2}\{|\mathbb{T}^{-1}_{t-\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})|+|\mathbb{T}^{-1}_{t-\sigma}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{\sigma,s}(\mathbf{x})-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))|\}\\&\quad\times\frac{1}{(t-\sigma)}\sup_{\gamma\in[0,1]}\bar{q}_c(\sigma,t,\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}),\mathbf{y}),\end{aligned}\tag{6.25}$$

where the last but one inequality is derived similarly to the first statement in Lemma 6.1 using the dynamics (3.3) associated with (3.5). Plugging equation (6.13) from

ACCEPTED MANUSCRIPT

Lemma 6.3 into (6.25) now yields:

$$I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y})$$
$$\leq C\Big\{\sum_{k=1}^{n}\frac{|(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})_k|+(\rho^\eta+|s-\sigma|)|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_k|}{\rho^{2+(2k-1)+n^2d}}\frac{1}{\Big|\tilde{s}-\frac{\sigma-s}{\rho^2}\Big|^{\frac{2+(2k-1)+n^2d}{2}}}\Big\}$$
$$\times\sup_{\gamma\in[0,1]}\exp\left(-c(t-\sigma)\left|\mathbb{T}_{t-\sigma}^{-1}\Big(\gamma\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}_{t,\sigma}^{t,\mathbf{y}}(\boldsymbol{\xi})-\mathbf{y}\Big)\right|^2\right). \tag{6.26}$$

Thus, if $|\tilde{s}-\frac{\sigma-s}{\rho^2}|^{1/2}\geq\tilde{c}$, we get recalling (6.18), (6.7), (6.8):

$$I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y})\leq C\Big\{\sum_{k=1}^{n}\frac{|(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})_k|}{\rho^{2+(2k-1)+n^2d}}+\frac{1}{\rho^{2-\eta+n^2d}}\Big\}$$
$$\leq C\Big\{\sum_{k=1}^{n}\frac{\alpha^{2k-1}}{\rho^{2+(2k-1)+n^2d}}+\frac{1}{\rho^{2-\eta+n^2d}}\Big\},$$

using (6.8) for the last inequality. Recall now from (6.7), (6.8) and (6.9) that

$$\begin{aligned}
&\big(\gamma\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}_{t,\sigma}^{t,\mathbf{y}}(\boldsymbol{\xi})\big)-\mathbf{y}\\
&=\gamma\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}_{t,\sigma}^{t,\mathbf{y}}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})+\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}})-\mathbf{y}\\
&\overset{(3.5)}{=}\gamma\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})+(1-\gamma)\Big\{\tilde{\boldsymbol{\theta}}_{t,\sigma}^{t,\mathbf{y}}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))+\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}\Big\}-\mathbf{y}\\
&=\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}+\gamma(\tilde{\boldsymbol{\theta}}_{t,s}^{t,\mathbf{y}}(\mathbf{x})-\boldsymbol{\theta}_{t,s}(\mathbf{x}))+(1-\gamma)(\tilde{\boldsymbol{\theta}}_{t,\sigma}^{t,\mathbf{y}}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))-\boldsymbol{\theta}_{t,s}(\mathbf{x}))\\
&\quad+(1-\gamma)\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}\\
&:=-\rho^{-1}\mathbb{T}_{\rho^2}\{\tilde{\mathbf{x}}+\gamma\mathcal{R}_{t,s}^{\rho}(\mathbf{x},\mathbf{y})+(1-\gamma)\mathcal{R}_{t,\sigma}^{\rho}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})\}+(1-\gamma)\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}},
\end{aligned}\tag{6.27}$$

where $(\bar{s},\bar{\mathbf{x}})\in\Sigma_1$. Observe that, from (6.12) in Lemma 6.3 we have for all $j\in[\![1,n]\!]$,

$$|(\mathcal{R}_{t,s}^{\rho}(\mathbf{x},\mathbf{y}))_j|+|(\mathcal{R}_{t,\sigma}^{\rho}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y}))_j|\leq C_{6.3}n(2\Lambda^\eta+(t-\sigma)+(t-s))|\tilde{\mathbf{x}}_{j_0}|. \tag{6.28}$$

On the other hand, from the scaling Lemma 6.2 we obtain that

$$(t-\sigma)^{1/2}\mathbb{T}_{t-\sigma}^{-1}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}=\bar{\mathbf{R}}_1^{\sigma,t,(t,\mathbf{y})}(t-\sigma)^{1/2}\mathbb{T}_{t-\sigma}^{-1}\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}$$

with $|\bar{\mathbf{R}}_1^{\sigma,t,(t,\mathbf{y})}|\leq\hat{C}:=\hat{C}(T,(\mathbf{A}))$. Thus, recalling that from the structure of the linearized system the resolvent is subdiagonal (see (3.3), (3.5)), we derive for all $j\in[\![1,n]\!]$:

$$\begin{aligned}
(t-\sigma)^{1/2}|(\mathbb{T}_{t-\sigma}^{-1}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}})_j|&=(t-\sigma)^{1/2}|(\bar{\mathbf{R}}_1^{\sigma,t,(t,\mathbf{y})}\mathbb{T}_{t-\sigma}^{-1}\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}})_j|\\
&\leq\hat{C}\sum_{i=1}^{j}\left(\frac{\alpha^2}{t-\sigma}\right)^{i-1/2}|\bar{\mathbf{x}}_i|\leq\hat{C}\sum_{i=1}^{j}\left(\frac{\alpha^2}{\rho^2|\tilde{s}-\frac{\sigma-s}{\rho^2}|}\right)^{i-1/2}|\bar{\mathbf{x}}|\\
&\leq\hat{C}\left|\tilde{s}-\frac{\sigma-s}{\rho^2}\right|^{-(j-1/2)}\sum_{i=1}^{j}\left(\frac{\alpha^2}{\rho^2}\right)^{i-1/2}|\bar{\mathbf{x}}|
\end{aligned}\tag{6.29}$$

ACCEPTED MANUSCRIPT

as soon as $\tilde{c} \leq 1$ and $|\tilde{s} - \frac{\sigma - s}{\rho^2}|^{1/2} \leq \tilde{c}$. In that case, using (6.22), (6.27), (6.28), (6.29) we then derive that:

$$
\begin{aligned}
&(t-\sigma)|\mathbb{T}^{-1}_{t-\sigma}(\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}) + (1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}) - \mathbf{y})|^2 \\
&\geq \left|\tilde{s} - \frac{\sigma - s}{\rho^2}\right|^{-(2j_0-1)} \left(\frac{1}{2}|\tilde{\mathbf{x}}_{j_0}|^2 - 2\Big\{(|(\mathcal{R}^{\rho}_{t,s}(\mathbf{x},\mathbf{y}))_{j_0}| + |(\mathcal{R}^{\rho}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y}))_{j_0}|)^2 \right. \\
&\qquad \left. + \hat{C}^2\left\{\sum_{i=1}^{j_0}\left(\frac{\alpha^2}{\rho^2}\right)^{i-1/2}\right\}^2 |\bar{\mathbf{x}}|^2\Big\}\right) \\
&\geq \left|\tilde{s} - \frac{\sigma - s}{\rho^2}\right|^{-(2j_0-1)} \left(\frac{1}{2}|\tilde{\mathbf{x}}_{j_0}|^2 - 2n^2[C^2_{6.3}(2\Lambda^{\eta} + 2(t-\sigma\wedge s))^2 + \hat{C}^2 c_{\infty}^{-2}]\right) \\
&\geq \left|\tilde{s} - \frac{\sigma - s}{\rho^2}\right|^{-(2j_0-1)} \bar{c},
\end{aligned}
$$

where $\bar{c} > 0$ for T, Λ small enough and a sufficiently large c_{∞}. Plugging this last inequality in (6.26), we thus obtain the global bound:

(6.30)
$$
\exists C_2 := C_2(T,(\mathbf{A})) > 0,\ I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y}) \leq C_2\left(\sum_{k=1}^{n}\frac{\alpha^{2k-1}}{\rho^{2+(2k-1)+n^2d}} + \frac{1}{\rho^{2-\eta+n^2d}}\right).
$$

Plugging (6.24), (6.30) into (6.19), and recalling as well that (6.18) holds, gives the point.

Let us now turn to the estimates concerning the *adjoint* kernel $k^{d,*}_{ij}$. Some additional contributions need to be taken into account. Namely, when investigating the difference

$$
\begin{aligned}
D^d_{ij}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y})) &:= k^{d,*}_{ij}(s,t,\mathbf{x},\mathbf{y}) - k^{d,*}_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y}) \\
&= k^{d}_{ij}(t,s,\mathbf{y},\mathbf{x}) - k^{d}_{ij}(t,\sigma,\mathbf{y},\boldsymbol{\xi}),
\end{aligned} \tag{6.31}
$$

we are led to consider the linearized systems $\tilde{\boldsymbol{\theta}}^{s,\mathbf{x}}_{s,t}(\mathbf{y}), \tilde{\boldsymbol{\theta}}^{\sigma,\boldsymbol{\xi}}_{\sigma,t}(\mathbf{y})$. Define now for all $(s,t,u,\mathbf{x}) \in [-T,T]^3 \times \mathbb{R}^{nd}$:

$$
\begin{aligned}
\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u) &:= \tilde{\mathbf{R}}^{s,\mathbf{x}}(t,s)\tilde{\mathbf{K}}^{s,\mathbf{x}}(t,u)\tilde{\mathbf{R}}^{s,\mathbf{x}}(t,s)^*, \\
\tilde{\mathbf{K}}^{s,\mathbf{x}}(t,u) &:= \textstyle\int_t^u \tilde{\mathbf{R}}^{s,\mathbf{x}}(s,v)B\varsigma(v)B^*[\tilde{\mathbf{R}}^{s,\mathbf{x}}(s,v)]^* dv.
\end{aligned} \tag{6.32}
$$

Remark 6.3. Let us note that for $u = s$ we have $\tilde{\mathbf{K}}^{s,\mathbf{x}}(t,s) = \tilde{\mathbf{K}}^{\mathbf{x}}(t,s)$ introduced after (3.6). Observe also, from the above definition and the specific structure of the resolvent (see equations (3.3)-(3.5)) that we actually have $\det(\tilde{\mathbf{R}}^{s,\mathbf{x}}(t,s)) = 1$ and therefore $\det(\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)) = \det(\tilde{\mathbf{K}}^{s,\mathbf{x}}(t,u))$.

From the definition in (6.32) and rewriting (6.6) in the current variables, the exponential bounds write:

$$
\begin{aligned}
&\langle\tilde{\mathbf{K}}^{\mathbf{x}}(t,s)^{-1}(\tilde{\boldsymbol{\theta}}^{s,\mathbf{x}}_{s,t}(\mathbf{y}) - \mathbf{x}), \tilde{\boldsymbol{\theta}}^{s,\mathbf{x}}_{s,t}(\mathbf{y}) - \mathbf{x}\rangle \\
&= \langle\tilde{\mathbf{R}}^{s,\mathbf{x}}(s,t)^*\tilde{\mathbf{K}}^{\mathbf{x}}(t,s)^{-1}\tilde{\mathbf{R}}^{s,\mathbf{x}}(s,t)(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})), \mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})\rangle \\
&= \langle\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,s)^{-1}(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})), \mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})\rangle,
\end{aligned}
$$

$$
\langle\tilde{\mathbf{K}}^{\boldsymbol{\xi}}(t,\sigma)^{-1}(\tilde{\boldsymbol{\theta}}^{\sigma,\boldsymbol{\xi}}_{\sigma,t}(\mathbf{y}) - \boldsymbol{\xi}), \tilde{\boldsymbol{\theta}}^{\sigma,\boldsymbol{\xi}}_{\sigma,t}(\mathbf{y}) - \boldsymbol{\xi}\rangle = \langle\tilde{\mathbf{H}}^{\sigma,\boldsymbol{\xi}}(t,\sigma)^{-1}(\mathbf{y} - \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})), \mathbf{y} - \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})\rangle.
$$

ACCEPTED MANUSCRIPT

Introducing, for all $(s,t,u,\mathbf{x}) \in [-T,T]^3 \times \mathbb{R}^{nd}$, $\mathbf{z} \in \mathbb{R}^{nd}$,

$$\check{k}^{s,\mathbf{x},u,t}(\mathbf{z}) := \mathbb{I}_{s-t>0}\left\{-[\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)^{-1}]_{1,1} + [\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)^{-1}\mathbf{z}]_1^{\otimes 2}\right\}$$

$$\frac{1}{(2\pi)^{nd/2}\det(\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u))^{1/2}} \times \exp\left(-\frac{1}{2}\langle \tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)^{-1}\mathbf{z},\mathbf{z}\rangle\right), \tag{6.33}$$

we can rewrite from (6.31) and (4.4):

$$\begin{aligned}
|D_{ij}^d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))| &= |\check{k}_{ij}^{s,\mathbf{x},s,t}(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})) - \check{k}_{ij}^{\sigma,\boldsymbol{\xi},\sigma,t}(\mathbf{y}-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}))| \\
&\leq |\check{k}_{ij}^{s,\mathbf{x},s,t}(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})) - \check{k}_{ij}^{s,\mathbf{x},\sigma,t}(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x}))| \\
&\quad + |\check{k}_{ij}^{s,\mathbf{x},\sigma,t}(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})) - \check{k}_{ij}^{\sigma,\boldsymbol{\xi},\sigma,t}(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x}))| \\
+|\check{k}_{ij}^{\sigma,\boldsymbol{\xi},\sigma,t}(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})) - \check{k}_{ij}^{\sigma,\boldsymbol{\xi},\sigma,t}(\mathbf{y}-\boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}))| &=: \sum_{l=1}^3 |D_{ij}^{d,l}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))|.
\end{aligned}$$

Now the terms $|D_{ij}^{d,1}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))|$ and $|D_{ij}^{d,3}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))|$ respectively involve time and space sensitivities when the freezing parameters in the covariance matrix are fixed. Those contributions can therefore be investigated as terms I_1 and I_2 in (6.24), (6.30). Once again, the previous splitting is associated w.l.o.g. to the case $|\sigma - s| \leq K|t-\sigma|$, see also Remark 6.2. The term $|D_{ij}^{d,2}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))|$ involves two different covariance matrices observed at the same time but that are respectively associated with the freezing points $(s,\mathbf{x})$ and $(\sigma,\boldsymbol{\xi})$ in the linearization of (3.2). To analyze this difference we proceed as in the proof of Lemma 2.4 in [Men11]. Namely, using (6.32) and the Scaling Lemma 6.2, we rewrite:

$$\begin{aligned}
\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,\sigma) &= \int_t^\sigma \tilde{\mathbf{R}}^{s,\mathbf{x}}(t,u)B_\varsigma(u)B^*[\tilde{\mathbf{R}}^{s,\mathbf{x}}(t,u)]^* du \\
&= \mathbb{T}_{\sigma-t}\int_t^\sigma \bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}}\mathbb{T}_{\sigma-t}^{-1}B_\varsigma(u)B^*\mathbb{T}_{\sigma-t}^{-1}[\bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}}]^* du\mathbb{T}_{\sigma-t} \\
&= (\sigma-t)^{-1}\mathbb{T}_{\sigma-t}\left[\frac{1}{\sigma-t}\int_t^\sigma \bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}}B_\varsigma(u)B^*[\bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}}]^* du\right]\mathbb{T}_{\sigma-t}.
\end{aligned}$$

Defining,

$$\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})} := \left[\frac{1}{\sigma-t}\int_t^\sigma \bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}}B_\varsigma(u)B^*[\bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}}]^* du\right], \tag{6.34}$$

yields:

$$\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,\sigma) = (\sigma-t)^{-1}\mathbb{T}_{\sigma-t}\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})}\mathbb{T}_{\sigma-t}.$$

Observe now from Lemma 6.2 and the non degeneracy assumption on c in **(A)**, that $\bar{\mathbf{H}}^{t,\sigma,(s,\mathbf{x})}$ is a bounded uniformly elliptic matrix of $\mathbb{R}^{nd}\otimes\mathbb{R}^{nd}$. Similarly,

$$\begin{aligned}
\tilde{\mathbf{H}}^{\sigma,\boldsymbol{\xi}}(t,\sigma) &= (\sigma-t)^{-1}\mathbb{T}_{\sigma-t}\left[\frac{1}{\sigma-t}\int_t^\sigma \bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-u}{\sigma-t}}B_\varsigma(u)B^*[\bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-u}{\sigma-t}}]^* du\right]\mathbb{T}_{\sigma-t} \\
&=: (\sigma-t)^{-1}\mathbb{T}_{\sigma-t}\bar{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})}\mathbb{T}_{\sigma-t}, \qquad (6.35)
\end{aligned}$$

where $\bar{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})}$ is again a uniformly elliptic bounded matrix on $\mathbb{R}^{nd}\otimes\mathbb{R}^{nd}$. Thus,

$$\begin{aligned}
&\langle(\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,\sigma) - \tilde{\mathbf{H}}^{\sigma,\boldsymbol{\xi}}(t,\sigma))(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})),\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})\rangle \\
&= \langle(\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})} - \bar{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})})\mathbf{z},\mathbf{z}\rangle\Big|_{\mathbf{z}=(\sigma-t)^{-1/2}\mathbb{T}_{\sigma-t}(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x}))}. \qquad (6.36)
\end{aligned}$$

ACCEPTED MANUSCRIPT

We now want to control the difference $(\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})} - \bar{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})})$ in (6.36). From the definitions in (6.34) and (6.35):

$$\begin{aligned}\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})} - \tilde{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})} &= (\sigma-t)^{-1}\int_t^\sigma \Big\{ \bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}} B\varsigma(u) B^* [\bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}}]^* \\ &\quad - \bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-u}{\sigma-t}} B\varsigma(u) B^* [\bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-u}{\sigma-t}}]^* \Big\} du, \\ |\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})} - \bar{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})}| &\le C(\sigma-t)^{-1}\int_t^\sigma \Big| \bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}} - \bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-u}{\sigma-t}} \Big| du. \end{aligned} \tag{6.37}$$

Let us now write, still from Lemma 6.2 and (3.4):

$$\begin{aligned} |\bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-u}{\sigma-t}} - \bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-u}{\sigma-t}}| &= |\mathbb{T}^{-1}_{\sigma-t}(\tilde{\mathbf{R}}^{s,\mathbf{x}}(t,u) - \tilde{\mathbf{R}}^{\sigma,\boldsymbol{\xi}}(t,u))\mathbb{T}_{\sigma-t}| = \\ \Big| \mathbb{T}^{-1}_{\sigma-t}\int_t^u \Big\{ \tilde{\mathbf{R}}^{s,\mathbf{x}}(t,v) D\mathbf{F}(v,\boldsymbol{\theta}_{v,s}(\mathbf{x})) - \tilde{\mathbf{R}}^{\sigma,\boldsymbol{\xi}}(t,v) D\mathbf{F}(v,\boldsymbol{\theta}_{v,\sigma}(\boldsymbol{\xi})) \Big\} dv \mathbb{T}_{\sigma-t} \Big| \\ \le \Big| \int_t^u (\bar{\mathbf{R}}^{t,\sigma,(s,\mathbf{x})}_{\frac{t-v}{\sigma-t}} - \bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-v}{\sigma-t}}) \left\{ \mathbb{T}^{-1}_{\sigma-t} D\mathbf{F}(v,\boldsymbol{\theta}_{v,s}(\mathbf{x}))\mathbb{T}_{\sigma-t} \right\} dv \\ + \int_t^u \bar{\mathbf{R}}^{t,\sigma,(\sigma,\boldsymbol{\xi})}_{\frac{t-v}{\sigma-t}} \mathbb{T}^{-1}_{\sigma-t}(D\mathbf{F}(v,\boldsymbol{\theta}_{v,s}(\mathbf{x})) - D\mathbf{F}(v,\boldsymbol{\theta}_{v,\sigma}(\boldsymbol{\xi})))\mathbb{T}_{\sigma-t} dv \Big| \le C|\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}) - \boldsymbol{\xi}|^\eta, \end{aligned}$$

where $C := C((\mathbf{A}))$, using the smoothness conditions assumed in **(S)**, the sub-diagonal structure of $D\mathbf{F}$ (see eq. (3.3)), the Liscphitz property of the flow and Gronwall's Lemma for the last inequality. From the above equation and (6.37), (6.36), we thus derive:

$$\begin{aligned} &\langle (\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,\sigma) - \tilde{\mathbf{H}}^{\sigma,\boldsymbol{\xi}}(t,\sigma))(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})), \mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})\rangle \\ &\le C|\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}) - \boldsymbol{\xi}|^\eta (\sigma-t)^{-1/2} |\mathbb{T}_{\sigma-t}(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x}))|^2,\ C := C((\mathbf{A})). \end{aligned} \tag{6.38}$$

Because of the non-degeneracy of c, the inverse matrices $(\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})})^{-1}$, $(\bar{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})})^{-1}$ have the same spatial Hölder regularity. Indeed, up to a change of coordinates one can assume that one of the two matrices is diagonal at the considered point and that the other has dominant diagonal if $|\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}) - \boldsymbol{\xi}|$ is small enough (depending on the ellipticity bounds in **(A)** and the dimension). This reduces to the scalar case. Hence,

$$\begin{aligned} &\langle ((\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,\sigma))^{-1} - (\tilde{\mathbf{H}}^{\sigma,\boldsymbol{\xi}}(t,\sigma))^{-1})(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})), \mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})\rangle \\ &= \langle ((\bar{\mathbf{H}}_1^{t,\sigma,(s,\mathbf{x})})^{-1} - (\bar{\mathbf{H}}_1^{t,\sigma,(\sigma,\boldsymbol{\xi})})^{-1})\mathbf{z}, \mathbf{z}\rangle \Big|_{\mathbf{z} = (\sigma-t)^{1/2}\mathbb{T}^{-1}_{\sigma-t}(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x}))} \\ &\le C|\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}) - \boldsymbol{\xi}|^\eta (\sigma-t)^{1/2} \mathbb{T}^{-1}_{\sigma-t}(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x}))|^2. \end{aligned} \tag{6.39}$$

The difference of the determinants can be investigated similarly. With the notations following (6.33), we therefore derive:

$$|D^{d,2}_{ij}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))| \le C\frac{\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))^\eta}{|t-\sigma|}\bar{q}_c(\sigma,t,\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y}),$$

ACCEPTED MANUSCRIPT

which can be analyzed similarly to I_2 (see equation (6.19) and page 40) in the previous proof and yields:

(6.40)
$$|D_{ij}^{d,2}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))| \leq C\frac{\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))^{\eta}}{\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))^{2+n^2d}} \leq C\frac{\mathbf{d}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))^{\eta}}{\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}))^{2+\eta+n^2d}},$$

recalling that $\mathbf{d}((s,\mathbf{x}),(t,\mathbf{y}) \leq \Lambda \leq 1$. This gives points *ii)* and *iii)*.

The cancellation property *iv)* is the more subtle to derive. Let us first prove:

$$\sup_{\epsilon>0}|\int_{\mathbf{d}((t,\mathbf{y}),(s,\mathbf{x}))>\epsilon} k_{ij}^{d}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y}| < +\infty.$$

Write with the notation of (6.33):

$$\int_{\mathbf{d}((t,\mathbf{y}),(s,\mathbf{x}))>\epsilon} k_{ij}^{d}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y} = \int_{\rho(t-s,\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))\in(\epsilon,\delta)} \check{k}_{ij}^{t,\mathbf{y},t,s}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))dtd\mathbf{y}$$
$$+\int_{\rho(t-s,\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))\in(\delta,2\delta)} \check{k}_{ij}^{t,\mathbf{y},t,s}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))dtd\mathbf{y} =: O_1^{\epsilon}+O_2^{\epsilon}.$$

Recall now from Section 4.3 that the kernel involves a cut-off that localizes the singularities. Hence, it is easily seen from (4.7) and the computations following that equation that $|O_2^{\epsilon}| \leq C := C(T,(\mathbf{A}),\delta)$. Let us now focus on O_1^{ϵ}. Set $\mathbf{z} := \mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y})$ that yields $d\mathbf{z} = \det(\mathrm{Jac}_{\boldsymbol{\theta}_{s,t}(\mathbf{y})})d\mathbf{y}$ where for T small enough $\det(\mathrm{Jac}_{\boldsymbol{\theta}_{s,t}(\mathbf{y})}) = 1 + O(|t-s|)$. We thus derive:

(6.41)
$$O_1^{\epsilon} = \int_{\rho(t-s,\mathbf{z})\in(\epsilon,\delta)} \exp(-\frac{1}{2}\langle(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1}\mathbf{z},\mathbf{z}\rangle)$$
$$\times P_{ij}((\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1},\mathbf{z})\frac{1}{(2\pi)^{nd/2}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{1/2}}dtd\mathbf{z}+O(1),$$

denoting for $(\mathbf{A},\mathbf{z}) \in \mathbb{R}^{nd}\otimes\mathbb{R}^{nd}\times\mathbb{R}^{nd}$, $P_{ij}(\mathbf{A},\mathbf{z}) := \left\{[\mathbf{A}]_{1,1}+[\mathbf{A}\mathbf{z}]_1^{\otimes 2}\right\}_{ij}$.

To conclude the analysis we need an additional regularization of the drift. To this end, we now introduce for given $((s,\mathbf{x}),(t,\mathbf{z})) \in S$:

(6.42)
$$q^{s,t,\mathbf{x},(t-s)*}(\mathbf{z}) := \left\{-[\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t)^{-1}]_{1,1}+[\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t)^{-1}\mathbf{z}]_1^{\otimes 2}\right\}$$
$$\frac{1}{(2\pi)^{nd/2}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t))^{1/2}}\times\exp\left(-\frac{1}{2}\langle\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t)^{-1}\mathbf{z},\mathbf{z}\rangle\right),$$

with

$$\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t) := \int_s^t \tilde{\mathbf{R}}_{s,u}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}B\varsigma(u)B^*[\tilde{\mathbf{R}}_{s,u}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}]^*du,$$
$$\partial_u\tilde{\mathbf{R}}_{s,u}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*} = -\tilde{\mathbf{R}}_{s,u}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}D\mathbf{F}^{(t-s)*}(u,\boldsymbol{\theta}_{u,s}(\mathbf{x}-\mathbf{z})),\ u\in[s,t],$$
$$\tilde{\mathbf{R}}_{s,s}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*} = I_{nd\times nd},$$
$$D\mathbf{F}^{(t-s)*}(u,\boldsymbol{\theta}_{u,s}(\mathbf{x}-\mathbf{z})) := D\mathbf{F}(u,\boldsymbol{\theta}_{u,s}(\cdot))*\zeta_{t-s}(\mathbf{x}-\mathbf{z}),$$

where the last $*$ stands for the spatial convolution and $\zeta_{t-s} : \mathbb{R}^{nd} \to [0,1]$ is a smooth mollifyer s.t. for $\mathbf{y}\in\mathbb{R}^{nd}$, $\zeta_{t-s}(\mathbf{y}) = 1$ if $|\mathbf{y}| \leq |t-s|^{1/8}$ and 0 if $|\mathbf{y}| \geq 2|t-s|^{1/8}$. Hence, there exists $C>0$ s.t. $|\mathbf{D}_{\mathbf{y}_i}\zeta(\mathbf{y})| \leq C|t-s|^{-1/8}, \forall i \in [\![1,d]\!]$, $|\mathbf{D}^2_{\mathbf{y}_i\mathbf{y}_j}\zeta(\mathbf{y})| \leq C|t-s|^{-1/4}$, $\forall(i,j)\in[\![1,d]\!]^2$.

ACCEPTED MANUSCRIPT

Under **(A)**, one easily gets that the *mollified* matrix $\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t)$ satisfies the good scaling property (3.8). Computations similar to those leading from (6.36) to (6.38), (6.39) also yield for all $\mathbf{z}\in\mathbb{R}^{nd}$:

$$|\langle[(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t))^{-1}-(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1}]\mathbf{z},\mathbf{z}\rangle| \leq C|t-s|^{\eta/8}|(t-s)^{1/2}\mathbb{T}_{t-s}^{-1}\mathbf{z}|^2, \tag{6.43}$$

$$\langle[\mathbf{D}_{\mathbf{z}_1^i}(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t))^{-1}+\mathbf{D}^2_{\mathbf{z}_1^i\mathbf{z}_1^j}(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t))^{-1}]\mathbf{z},\mathbf{z}\rangle \leq C|t-s|^{-1/4}|(t-s)^{1/2}\mathbb{T}_{t-s}^{-1}\mathbf{z}|^2,\ \forall(i,j)\in[\![1,d]\!]^2. \tag{6.44}$$

Similar controls would hold for the difference and sensitivities of the determinants. Precisely, we observe that the sensitivities of the *mollified* covariance matrices w.r.t. the freezing parameters induce additional integrable singularities. Write then from (6.41), (6.42):

$$O_1^\epsilon=\int_{\rho(t-s,\mathbf{z})\in(\epsilon,\delta)}[q^{s,t,\mathbf{x},0}(\mathbf{z})-q^{s,t,\mathbf{x},(t-s)*}(\mathbf{z})]dtd\mathbf{z} + \int_{\rho(t-s,\mathbf{z})\in(\epsilon,\delta)} q^{s,t,\mathbf{x},(t-s)*}(\mathbf{z})+O(1)=:O_{11}^\epsilon+O_{12}^\epsilon+O(1),$$

denoting with a slight abuse of notation by $q^{s,t,\mathbf{x},0}(\mathbf{z})$ the term in (6.42) when there is no convolution. From the definition of $q^{s,t,\mathbf{x},\cdot}$ in (6.42), the good scaling property property (3.8) satisfied by both $\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})},\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),*(t-s)}$ and equation (6.43), we get that the contribution $O_{11}^\epsilon=O(1)$. Indeed, the singularities are integrable for that term. On the other hand

$$q^{s,t,\mathbf{x},(t-s)*}(\mathbf{z}) = D^2_{\mathbf{z}_1^i\mathbf{z}_1^j}\left\{\frac{\exp(-\frac{1}{2}\langle(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),*(t-s)}(s,t))^{-1}\mathbf{z},\mathbf{z}\rangle)}{(2\pi)^{nd}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),*(t-s)}(s,t))^{1/2}}\right\}+\mathcal{R}_{i,j}(s,t,\mathbf{x},\mathbf{z}).$$

The remainder term $\mathcal{R}_{i,j}$ gathers the contributions deriving from the sensitivities of $(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t))^{-1}$ w.r.t. $\mathbf{z}_1^i,\mathbf{z}_1^j$. Actually, the regularization of the coefficient $D\mathbf{F}$ is just required here to differentiate the dynamics of the resolvent. From (6.44) it can be checked that:

$$|\mathcal{R}_{i,j}(s,t,\mathbf{x},\mathbf{z})|\leq\frac{C}{(t-s)^{n^2d/2+3/4}}\exp(-(t-s)|\mathbb{T}_{t-s}^{-1}\mathbf{z}|^2),\ C:=C((\mathbf{A}),T).$$

We thus write

$$O_{12}^\epsilon = \int_{\rho(t-s,\mathbf{z})\in(\epsilon,\delta)} D^2_{\mathbf{z}_1^i\mathbf{z}_1^j}\left\{q_{\mathbf{H},*}(s,t,\mathbf{z})\right\}dtd\mathbf{z}+O(1),$$

$$q_{\mathbf{H},*}(s,t,\mathbf{z}) := \frac{\exp(-\frac{1}{2}\langle(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t))^{-1}\mathbf{z},\mathbf{z}\rangle)}{(2\pi)^{nd/2}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}(s,t))^{1/2}}.$$

By the divergence theorem:

$$|O_{12}^\epsilon|\leq\sum_{\beta\in\{\epsilon,\delta\}}\left|\int_{\rho(t-s,\mathbf{z})=\beta}\partial_{\mathbf{z}_1^i}q_{\mathbf{H},*}(s,t,\mathbf{z})n_j d\nu((t-s),\mathbf{z})\right|.$$

Now, from the metric homogeneity (see Remark 4.1), the good scaling property (3.8) that is valid for $\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z}),(t-s)*}$, it can be shown, changing variables similarly to

ACCEPTED MANUSCRIPT

(6.7), that for ϵ small enough:

$$|O^{\epsilon}_{12}| \leq C \int_{\rho(|\tilde{s}|,\bar{\mathbf{z}})=1} \frac{d\nu(\tilde{s},\bar{\mathbf{z}})}{|\tilde{s}|^{n^2 d/2+1/2}} \exp\left(-C|\tilde{s}||\mathbb{T}^{-1}_{|\tilde{s}|}\bar{\mathbf{z}}|^2\right) < +\infty, \tag{6.45}$$

and that O^{ϵ}_{12} admits a limit when $\epsilon \to 0$. From (6.45) we thus get point *iv)* for the kernel k^d_{ij}. The proof of the cancellation property for the adjoint kernel, as well as the γ-Hölder continuity for $h - h^*$ for some $\gamma > 0$, can be proved similarly exploiting the equivalence of the "forward" and "backward" distance on compact sets (see Proposition 4.1). □

Acknowledgments

The article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

[Bas95] R. F. Bass. *Probabilistic techniques in Analysis*. Springer, 1995.

[BCLP10] Marco Bramanti, Giovanni Cupini, Ermanno Lanconelli, and Enrico Priola. Global L^p estimates for degenerate Ornstein-Uhlenbeck operators. *Math. Z.*, 266(4):789–816, 2010.

[BCLP13] Marco Bramanti, Giovanni Cupini, Ermanno Lanconelli, and Enrico Priola. Global L^p estimates for degenerate Ornstein-Uhlenbeck operators with variable coefficients. *Math. Nachr.*, 286(11-12):1087–1101, 2013.

[BP09] R.F. Bass and E.A. Perkins. A new technique for proving uniqueness for martingale problems. *From Probability to Geometry (I): Volume in Honor of the 60th Birthday of Jean-Michel Bismut*, pages 47–53, 2009.

[BPV01] E. Barucci, S. Polidoro, and V. Vespri. Some results on partial differential equations and asian options. *Math. Models Methods Appl. Sci*, 3:475–497, 2001.

[Bra10] M. Bramanti. Singular integrals in nonhomogeneous spaces: L^2 and L^p continuity from Hölder estimates. *Revista Matematica Iberoamericana*, 26–1:347–366, 2010.

[BZ13] Marco Bramanti and Maochun Zhu. L^p and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift. *Anal. PDE*, 6(8):1793–1855, 2013.

[CW71] R. Coifman and G. Weiss. *Analyse Harmonique non-commutative sur certains espaces homogènes*, volume 242. Springer, Lecture Notes in Math., 1971.

[Del02] F. Delarue. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stoch. Proc. and App.*, 99:209–286, 2002.

[Del03] F. Delarue. Estimates of the solutions of a system of quasilinear PDEs. A probabilistic scheme. *Séminaire de Probabilités*, XXXVII:290–332, 2003.

[DG06] F. Delarue and G. Guatteri. Weak Solvability Theorem for FBSDEs. *Stoc. Proc. Appl.*, 12:1712–1742, 2006.

[DM10] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *Journal of Functional Analysis*, 259–6:1577–1630, 2010.

[EK97] E. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. Wiley, 1997.

[EPRB99] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Comm. Math. Phys.*, 201–3:657–697, 1999.

[FF11] E. Fedrizzi and F. Flandoli. Pathwise Uniqueness and Continuous Dependence for SDEs with Nonregular Drift. *Stochastics*, 83-3:241–257, 2011.

[FFPV16] E. Fedrizzi, F. Flandoli, E. Priola, and J. Vovelle. Regularity of Stochastic Kinetic Equations. *arXiv:1606.01088*, 2016.

[FP06] 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*, 11:1261–1320, 2006.

ACCEPTED MANUSCRIPT

[Fri64] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, 1964.

[GT83] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Springer Verlag, 1983.

[HN04] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker-planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, 171–2:151–218, 2004.

[IKO62] A. M. Il'in, A. S. Kalashnikov, and O. A. Oleinik. Second-order linear equations of parabolic type. *Uspehi Mat. Nauk*, 17–3(105):3–146, 1962.

[Kol34] A. N. Kolmogorov. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math.*, 2-35:116–117, 1934.

[Kol11] V. Kolokoltsov. *Markov Processes, Semigroups and Generators*. De Gruyter, 2011.

[KR05] N. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Prob. Theory Rel. Fields*, 131:154–196, 2005.

[Kry87] N.V. Krylov. *Controlled diffusion processes*. Springer, 1987.

[LP94] E. Lanconelli and S. Polidoro. On a class of hypoelliptic evolution operators. *Rend. Sem. Mat. Univ. Politec. Torino*, 52–1:29–63, 1994.

[Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electronic Communications in Probability*, 17:234–250, 2011.

[Pri15] Enrico Priola. On weak uniqueness for some degenerate SDEs by global L^p estimates. *Potential Anal.*, 42(1):247–281, 2015.

[Ray14] P.E. Chaudru De Raynal. Strong existence and uniqueness for stochastic differential equation with Hölder drift and degenerate noise. *hal-00702532v3, To Appear in Ann. Intitut H. Poincaré*, 2014.

[Ray16] P.E. Chaudru De Raynal. Weak well posedness for hypoelliptic stochastic differential equation with singular drift: A sharp result. *arXiv:1606.05458*, 2016.

[RBT00] L. Rey-Bellet and L. Thomas. Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. *Comm. Math. Phys.*, 215–1:1–24, 2000.

[She91] S. J. Sheu. Some estimates of the transition density of a nondegenerate diffusion Markov process. *Ann. Probab.*, 19–2:538–561, 1991.

[Soi94] C. Soize. *The Fokker-Planck equation for stochastic dynamical systems and its explicit steady state solutions*. Series on Advances in Mathematics for Applied Sciences, 17. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[Ste70] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton university press, 1970.

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.

[Tal02] D. Talay. Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8–2:163–198, 2002.

[WZ16] F. Y. Wang and X. Zhang. Degenerate SDE with Hölder-Dini drift and non-Lipschitz noise coefficient. *SIAM J. Math. Anal.*, 48(3):2189–2226, 2016.

[Zha16] X. Zhang. Stochastic Hamiltonian flows with singular cefficients. *arXiv:1606.04360*, 2016.

Laboratoire de Modélisation Mathématique d'Evry (LaMME), Université d'Evry Val d'Essonne, 23 Boulevard de France 91037 Evry, France and Laboratory of Stochastic Analysis, HSE, Shabolovka 31, Moscow, Russian Federation.

E-mail address: stephane.menozzi@univ-evry.fr