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)

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

$(W_t)_{t\geq 0}$ 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).

Under *mild* assumptions, namely spatial Lipschitz continuity of the coefficients $(F_i)_{i\in[\![1,n]\!]}, \sigma$, non degeneracy of the bounded diffusion matrix $a(t,.) := \sigma\sigma^*(t,.)$ and the gradients $(D_{x_{i-1}}F_i(t,.))_{i\in[\![2,n]\!]}$ (weak Hörmander condition), 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 Theorem 1.1 of that reference). Those results extend to the case of Hölder continuous diffusion coefficient $a(t,.)$ and gradients $(D_{x_{i-1}}F_i(t,.))_{i\in[\![2,n]\!]}$ thanks to uniqueness in law arguments that have been investigated in [Men11] through the well posedness of the martingale problem.

Date: August 17, 2014.

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

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

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 to 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].

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

[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 $\sigma = 0$

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 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: $\exists c := c(a,b,p_0,N,p),\ \forall u \in C_0^2(\mathbb{R}^N)$,

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

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 $j < i-1,\ (i,j) \in [\![1,n]\!]$. Hence, the operator $\mathcal{A}$ can be seen as a particular case of generator associated to (1.1).

The strategy in those works still consists in estimating suitable singular integrals[3] related to the Gaussian fundamental solution of $L = \mathcal{A} - \partial_t$ 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,

$$\forall (s,x_1,\cdots,x_n) \in \mathbb{R}\times\mathbb{R}^{nd},\ \bar{\rho}(s,x_1,\cdots,x_n) := |s|^{1/2} + \sum_{i=1}^{n} |x_i|^{1/(2i-1)}. \tag{1.2}$$

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 (1.2) when considering, for $(s,x),(t,y) \in \mathbb{R}\times\mathbb{R}^{nd}$, $\bar{d}((s,x),(t,y)) = \bar{\rho}(t-s, x-\exp(-B(t-s))y)$ which appears as a natural candidate to be a quasi-distance taking into account the transport. It can be shown that $\bar{d}((s,x),(t,y))$ and $\bar{d}((t,y),(s,x))$ are equivalent on 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. Schauder) estimates have been obtained by Bramanti and Zhu [BZ11] 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

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

[Pri13] 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 [Pri13], 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 and non-linear coefficients for the degenerate part of equation (1.1). To this end we introduce a suitable kernel and do not exploit some underlying Lie group properties appearing in the quoted works. The key idea consists in viewing (1.1) as an ODE *perturbed* by a noise. This naturally yields to consider balls that are build around the characteristic lines of the ODE and reflect the multi-scale 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 spaces.

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 adaptations of the arguments of the non-degenerate framework, see e.g. Sections 7.1, 7.2 and 9.1 in [SV79]). 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 $i\in[\![1,n-1]\!]$, $\mathbf{x}^{i,n}:=(\mathbf{x}_i,\cdots,\mathbf{x}_n)$.

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 $(F_i)_{i\in[\![1,n]\!]}$ and the $(D_{\mathbf{x}_{i-1}}F_i)_{i\in[\![1,n]\!]}$ are twice continuously differentiable with bounded derivatives in space, bounded measurable in time.

(ND) There exists a closed convex subset $\mathcal{E}_{i-1}\subset GL_d(\mathbb{R})$ (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 $GL_d(\mathbb{R})$, 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)**. Also, the spatial smoothness on F is required to adapt the techniques of [BCLP10], which involve the adjoint kernel. 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).

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 $(s,\mathbf{x})\in[0,T)\times\mathbb{R}^{nd}$:

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

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 $|\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{x}|\leq\kappa\int_s^t|\boldsymbol{\theta}_{u,s}(\mathbf{x})|du\leq\kappa(t-s)|\mathbf{x}|\exp(\kappa(t-s))$, $t\geq s$, where κ stands for the Lipschitz constant of $\mathbf{F}$. Hence, one has $|\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{x}|\leq\varepsilon$ for $|t-s|\leq C/|\mathbf{x}|, C:=C(\kappa,\varepsilon)$, which corresponds to the time step of the partition for a given $|\mathbf{x}|$. For a fixed $T>0$, $\lceil|\mathbf{x}|C^{-1}T\rceil$ 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]). For technical reasons (see Section 6.3 and the proof of Lemma 6.3), we suppose that the gradients $(D_{\mathbf{x}_{i-1}}F_i(t,.))_{i\in[\![2,n]\!]}$ do not vary much as well, that is the degenerate components are "almost affine" w.r.t the components that transmit the noise. Precisely, we first assume that there exists $\mathbf{x}_0 \in \mathbb{R}^{nd}$ such that:

$$\begin{aligned}\varepsilon_a &:= \sup_{0\le t\le T}\sup_{\mathbf{x}\in\mathbb{R}^{nd}}|a(t,\mathbf{x})-a(t,\mathbf{x}_0)|,\\ \varepsilon_{\mathbf{F}} &:= \sup_{i\in[\![2,n]\!]}\sup_{0\le t\le T}\sup_{(\mathbf{x},\mathbf{z})\in(\mathbb{R}^{nd})^2}|D_{\mathbf{x}_{i-1}}F_i(t,\mathbf{x}_{i-1},\mathbf{z}^{i,n})-D_{\mathbf{x}_{i-1}}F_i(t,(\mathbf{x}_0)_{i-1},\mathbf{z}^{i,n})|\end{aligned} \tag{3.1}$$

are small, without supposing *a priori* continuity of a.

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

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

Consider now the deterministic ODE

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

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 $\mathbf{x}$.

We denote by $(\tilde{\boldsymbol{\theta}}^{T,\mathbf{y}}_{t,s})_{s,t\ge 0}$ the associated flow, i.e. $\tilde{\boldsymbol{\theta}}^{T,\mathbf{y}}_{t,s}(\mathbf{x})$ is the value of $\tilde{\boldsymbol{\phi}}_t$ when $\tilde{\boldsymbol{\phi}}_s = \mathbf{x}$. It is affine:

$$\begin{aligned}\tilde{\boldsymbol{\theta}}^{T,\mathbf{y}}_{t,s}(\mathbf{x}) &= \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)\mathbf{x}\\ &\quad+\int_s^t \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,u)\big(\mathbf{F}(u,\boldsymbol{\theta}_{u,T}(\mathbf{y})) - D\mathbf{F}(u,\boldsymbol{\theta}_{u,T}(\mathbf{y}))\boldsymbol{\theta}_{u,T}(\mathbf{y})\big)du\\ &:= \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)\mathbf{x} + \mathbf{m}^{T,\mathbf{y}}(s,t).\end{aligned} \tag{3.4}$$

Above, $(\tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s))_{s,t\ge 0}$ stands for the resolvent associated with the subdiagonal matrices $(D\mathbf{F}(t,\boldsymbol{\theta}_{t,T}(\mathbf{y})))_{t\ge 0}$. Note in particular that $\det(\tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)) = 1$.

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

$$\tilde{q}(s,t,\mathbf{x},\mathbf{y}) := \frac{1}{(2\pi)^{nd/2}(\tilde{\mathbf{K}}^{\mathbf{y}}(s,t))^{1/2}}$$
$$\times \exp\left(-\frac{1}{2}\langle \tilde{\mathbf{K}}^{\mathbf{y}}(s,t)^{-1}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y}), \tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y}\rangle\right), \tag{3.5}$$

where $\tilde{\mathbf{K}}^{\mathbf{y}}(s,t) := \int_s^t \tilde{\mathbf{R}}^{t,\mathbf{y}}(t,u)Ba(u,\mathbf{x}_0)B^*\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,u)^*du$. In other words, under **(A)**, $\tilde{q}(s,t,\mathbf{x},\mathbf{y})$ is the density at time t and point $\mathbf{y}$ of the diffusion $(\tilde{\mathbf{X}}^{t,\mathbf{y}}_u)_{u\in[s,t]}$ with dynamics:

$$d\tilde{\mathbf{X}}^{t,\mathbf{y}}_u = [\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y})) + D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{y}))(\tilde{\mathbf{X}}^{t,\mathbf{y}}_u - \boldsymbol{\theta}_{u,t}(\mathbf{y}))]du + B\sigma(u,\mathbf{x}_0)dW_u,$$
$$\forall u \in [s,t],\ \tilde{\mathbf{X}}^{t,\mathbf{y}}_s = \mathbf{x}. \tag{3.6}$$

Assumption **(A)** also guarantees that the covariance matrix $(\tilde{\mathbf{K}}^{\mathbf{y}}(s,t))_{0\leq s<t}$ 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 $C_{3.7} := C_{3.7}(T,(\mathbf{A})) \geq 1$ s.t. for all $0\leq s<t\leq T$, for all $\mathbf{y}\in\mathbb{R}^{nd}$:

$$\forall \boldsymbol{\xi}\in\mathbb{R}^{nd},\ C_{3.7}^{-1}(t-s)^{-1}|\mathbb{T}_{t-s}\boldsymbol{\xi}|^2 \leq \langle \tilde{\mathbf{K}}^{\mathbf{y}}(s,t)\boldsymbol{\xi},\boldsymbol{\xi}\rangle \leq C_{3.7}(t-s)^{-1}|\mathbb{T}_{t-s}\boldsymbol{\xi}|^2, \tag{3.7}$$

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 (3.7) indicates that the i^{th} component of (3.6) has characteristic time scale of order $(2i-1)/2$.

From (3.5) and (3.7), we directly derive that for all $T>0$ there exists $C_{3.8} := C_{3.8}(T,(\mathbf{A})) \geq 1$ s.t. for all $0\leq s<t\leq T,\ (\mathbf{x},\mathbf{y})\in(\mathbb{R}^{nd})^2$:

$$C_{3.8}^{-1}(t-s)^{-n^2d/2}\exp(-C_{3.8}(t-s)|\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})|^2) \leq \tilde{q}(s,t,\mathbf{x},\mathbf{y})$$
$$\leq C_{3.8}(t-s)^{-n^2d/2}\exp(-C_{3.8}^{-1}(t-s)|\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})|^2). \tag{3.8}$$

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 for details) give that:

$$C^{-1}|\mathbb{T}^{-1}_{t-s}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))| \leq |\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})| \leq C|\mathbb{T}^{-1}_{t-s}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))|,$$
$$|D_{\mathbf{x}_j}\tilde{q}(s,t,\mathbf{x},\mathbf{y})| \leq C(t-s)^{-j+1}|\mathbb{T}^{-1}_{t-s}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\mathbf{y})|\tilde{q}(s,t,\mathbf{x},\mathbf{y}),\ j\in[\![1,n]\!]. \tag{3.9}$$

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 $(t,\mathbf{y})\in\mathbb{R}^{+*}\times\mathbb{R}^{nd}$:

$$\left(\partial_s + \tilde{L}^{t,\mathbf{y}}_s\right)\tilde{q}(s,t,\mathbf{x},\mathbf{y}) = 0,\ (s,\mathbf{x})\in[0,t)\times\mathbb{R}^{nd},\ \tilde{q}(s,t,.,\mathbf{y}) \underset{s\uparrow t}{\longrightarrow} \delta_{\mathbf{y}}(.). \tag{3.10}$$

In the above equation we wrote:

$$\tilde{L}^{t,\mathbf{y}}_s\tilde{q}(s,t,\mathbf{x},\mathbf{y}) := \langle \mathbf{F}(s,\boldsymbol{\theta}_{s,t}(\mathbf{y})) + D\mathbf{F}(s,\boldsymbol{\theta}_{s,t}(\mathbf{y}))(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y})), \mathbf{D}_{\mathbf{x}}\tilde{q}(s,t,\mathbf{x},\mathbf{y})\rangle$$
$$+\frac{1}{2}\mathrm{tr}(a(s,\mathbf{x}_0)D^2_{\mathbf{x}_1}\tilde{q}(s,t,\mathbf{x},\mathbf{y})).$$

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 $\mathbf{x}\in\mathbb{R}^{nd}$:

$$\tilde{G}f(s,\mathbf{x}) := \int_s^T dt\int_{\mathbb{R}^{nd}} \tilde{q}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})d\mathbf{y}. \tag{3.11}$$

From (3.10) 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 $\tilde{M}f(s,\mathbf{x}) := \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{L}_s^{t,\mathbf{y}} \tilde{q}(s,t,\mathbf{x},\mathbf{y}) f(t,\mathbf{y})$.

Hence,

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

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:

$$|Rf(s,\mathbf{x})| \leq |Nf(s,\mathbf{x})| + \sum_{i=2}^n |D_{\mathbf{x}_i} R_i f(s,\mathbf{x})| + \frac{\varepsilon_a}{2} |D^2_{\mathbf{x}_1} \tilde{G}f(s,\mathbf{x})|, \tag{3.12}$$

where setting

$$\begin{aligned} \mathbf{F}^{t,\mathbf{y}}(s,\mathbf{x}) &:= \big(\mathbf{F}_1(s,\boldsymbol{\theta}_{s,t}(\mathbf{y})), \mathbf{F}_2(s,\mathbf{x}_1,(\boldsymbol{\theta}_{s,t}(\mathbf{y}))^{2,n}), \mathbf{F}_3(s,\mathbf{x}_2,(\boldsymbol{\theta}_{s,t}(\mathbf{y}))^{3,n}),\cdots, \\ &\quad \mathbf{F}_n(s,\mathbf{x}_{n-1},(\boldsymbol{\theta}_{s,t}(\mathbf{y}))_n)\big), \\ Nf(s,\mathbf{x}) &:= \sum_{i=1}^n \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} \langle \big((\mathbf{F}-\mathbf{F}^{t,\mathbf{y}})(s,\mathbf{x})\big)_i, \mathbf{D}_{\mathbf{x}_i} \tilde{q}(s,t,\mathbf{x},\mathbf{y})\rangle f(t,\mathbf{y}), \end{aligned} \tag{3.13}$$

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

$$\begin{aligned} R_i f(s,\mathbf{x}) &:= \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{q}(s,t,\mathbf{x},\mathbf{y}) f(t,\mathbf{y}) \big\{ \mathbf{F}_i^{t,\mathbf{y}}(s,\mathbf{x}) \\ &\quad - \big[\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{y})) + D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{y}))(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))_{i-1}\big]\big\}. \end{aligned} \tag{3.14}$$

Remark 3.1. 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 (3.13) and $(D_{\mathbf{x}_i} R_i f)_{i\in[\![2,n]\!]}$ in (3.14) do not have time singularities. Let us justify this point.

Using (3.8), (3.9), we derive from (3.13) that:

$$|Nf(s,\mathbf{x})| \leq C \sum_{i=1}^n \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} |(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))^{i,n}| |D_{\mathbf{x}_i} \tilde{q}(s,t,\mathbf{x},\mathbf{y})| |f(t,\mathbf{y})|$$

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

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

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 (3.9) then yields:

$$\|Nf\|_{L^p([0,T)\times\mathbb{R}^{nd})} \leq C(p,(\mathbf{A}))T\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}. \tag{3.15}$$

From (3.8), (3.9) we also derive for all $i \in [\![2,n]\!]$:

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

Hence, by Hölder's inequality and for $p>2,\ p^{-1}+q^{-1}=1$:

$$|D_{\mathbf{x}_i}R_if(s,\mathbf{x})|^p \leq C(p,(\mathbf{A}))\left(\int_s^T dt(t-s)^{(-1+(i-\frac{3}{2}))q}\right)^{p/q}$$

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

$$\|D_{\mathbf{x}_i}R_if\|_{L^p([0,T)\times\mathbb{R}^{nd})} \leq C(p,(\mathbf{A}))T^{1/2}\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}, \tag{3.16}$$

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.

Theorem 3.1. Assume that Assumption **(A)** is in force. Suppose also that the local condition (3.1) holds for $\mathbf{F}$, i.e. $\varepsilon_{\mathbf{F}} \leq \varepsilon_0(n,d)$ for a small enough $\varepsilon_0(n,d)$, and that $T\in(0,T_0],\ T_0 := T_0((\mathbf{A})) \leq 1$. Then, for all $p\in(1,+\infty)$ there exists $C_{3.17} := C_{3.17}(T_0,p,(\mathbf{A}))$ s.t. for all $f\in L^p([0,T)\times\mathbb{R}^{nd})$,

$$\|D^2_{\mathbf{x}_1}\tilde{G}f\|_{L^p([0,T)\times\mathbb{R}^{nd})} \leq C_{3.17}\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}, \tag{3.17}$$

where the Green function $\tilde{G}f$ is defined in (3.11) with the kernel $\tilde{q}$ introduced in (3.5).

Remark 3.2. 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*. It is clear that this procedure can be a good approximation in small time only.

Hence, plugging (3.17), (3.16) and (3.15) into (3.12) we derive that under **(A)**, for $p>2$,

$$\|Rf\|_{L^p([0,T)\times\mathbb{R}^{nd})} \leq (\frac{\varepsilon_a}{2}C_{3.17}+C(p,(\mathbf{A}))T^{\frac{1}{2}})\|f\|_{L^p([0,T)\times\mathbb{R}^{nd})}. \tag{3.18}$$

Thus, for $\varepsilon_a < C^{-1}_{3.17}$ and $T<(4C(p,(\mathbf{A})))^{-2}$, 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 (3.8) that, for all $p > (n^2d+2)/2$, $f \in L^p([0,T)\times\mathbb{R}^{nd})$,

$$|\tilde{G}f(s,\mathbf{x})| \leq CT^{1-(2+n^2d)/(2p)}\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})}. \tag{3.19}$$

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 $Gf(s,\mathbf{x}) = \mathbb{E}^{\mathbb{P}}[\int_s^T f(t,\mathbf{X}_t)dt]$, $(\mathbf{X}_t)_{t\in[0,T]}$ being the canonical process. A localization argument similar to the one in Priola [Pri13] then allows to extend the well posedness of the martingale problem under the sole continuity assumption **(C)** in **(A)**, i.e. without the local condition (3.1), see Section 5.

4. Derivation of the Calderón-Zygmund estimates

We assume **(A)** is in force and that $T \leq T_0((\mathbf{A})) \leq 1$. We also suppose for the whole section that the local condition (3.1) on $\mathbf{F}$ holds.

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 to 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. $\rho(t, t^{-1/2}\mathbb{T}_t\mathbf{x}) = t^{1/2}\rho(1,\mathbf{x})$.

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 $(s,\mathbf{x}),(t,\mathbf{y}) \in S^2$ the distances:

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

with ρ as in (4.1). We now define the metric balls in the following way:

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

We mention that the natural extension of the balls considered in [BCLP10], [BCLP13] would have been to consider 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 "distance" used to define the balls satisfies the triangle inequality only locally. A

natural choice would then consist in considering singular integrals for the "homogeneous space" associated to the balls of the above form, but in such case it is not clear that such balls 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 $(S, d, dtd\mathbf{x})$ is a locally invariant quasi-metric space in the following sense: there exists a constant $C_{4.1} := C_{4.1}((\mathbf{A}), T) > 0$ s.t.

a) For all $(s, \mathbf{x}), (t, \mathbf{y}) \in S$, if $d((s, \mathbf{x}), (t, \mathbf{y})) \leq 1$ then

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

and for $(\sigma, \boldsymbol{\xi}) \in S$ s.t. $d((s, \mathbf{x}), (\sigma, \boldsymbol{\xi})) \leq 1$ and $d((t, \mathbf{y}), (\sigma, \boldsymbol{\xi})) \leq 1$ then

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

b) Every 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 $0 < R_1 < R_2 \leq R$ there exists $C := C(R_1, R_2)$ s.t. for all $(s, \mathbf{x}) \in S$,

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

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 here depends on the specific radius, here 1, chosen for the balls.

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 distances d and d^* involving respectively the forward and backward transport are actually equivalent. In such a case 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.

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

4.2. **Singular kernel and associated estimates.** Let us first define for $s > 0, \mathbf{x} \in \mathbb{R}^{nd}$, $a(-s, \mathbf{x}) := a(s, \mathbf{x})$, i.e. we symmetrize the diffusion coefficient. Fix now $T > 0$ and introduce:

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

From (3.4) and (3.5) 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 (3.5), (3.7) and the scaling Lemma 6.2 (see also equations (5.10), (5.11) in [DM10]), we have that there exists $C:=C(T,(\mathbf{A}))$ s.t.:

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

so that (3.7), (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 $c>0$,

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

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 $\epsilon\in(0,1)$, $\forall(i,j)\in[\![1,d]\!]^2$, $f\in L^\infty(\mathbb{R}\times\mathbb{R}^{nd})$, $(s,\mathbf{x})\in\mathbb{R}^{1+nd}$,

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

is well defined.

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 $(s,\mathbf{x}),(t,\mathbf{y})\in S^2$, $d((s,\mathbf{x}),(t,\mathbf{y}))=\rho(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})\ge\delta$, then from (4.5) we have:

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

Let us now write:

$$\begin{aligned} K^\epsilon_{ij}f(s,\mathbf{x}) &= \int_{S\cap d((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon}dtd\mathbf{y}k_{ij}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})\eta_\delta(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}) \\ &\quad+\int_{S\cap d((s,\mathbf{x}),(t,\mathbf{y}))>\epsilon}dtd\mathbf{y}k_{ij}(s,t,\mathbf{x},\mathbf{y})f(t,\mathbf{y})(1-\eta_\delta)(t-s,\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}) \\ &:= K^{\epsilon,d}_{ij}f(s,\mathbf{x})+K^{\epsilon,\infty}_{ij}f(s,\mathbf{x}), \end{aligned}$$

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 $\rho(u,z)\le\delta$ and $\eta_\delta(u,z)=0$ if $\rho(u,z)\ge2\delta$. It is then

easily seen from (4.7) (see also the proof of of point *i)* in Proposition 4.2) that for all $f \in L^p(S),\ p \in [1, +\infty]$,

$$\|K_{ij}^{\epsilon,\infty} f\|_{L^p(S)} \leq C_{4.8} \|f\|_{L^p(S)},\ C_{4.8} := C_{4.8}(T, (\mathbf{A}), \delta, p). \tag{4.8}$$

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

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

$$|k_{ij}^d(s,t,\mathbf{x},\mathbf{y}) - k_{ij}^d(\sigma,t,\boldsymbol{\xi},\mathbf{y})| \leq C_{4.2} \frac{d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))}{d((s,\mathbf{x}),(t,\mathbf{y}))^{n^2 d+3}},$$

$\forall (s,\mathbf{x}), (\sigma,\boldsymbol{\xi}) \in S^2,\ c_{4.2} d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})) \leq d((s,\mathbf{x}),(t,\mathbf{y})) \leq 1.$

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

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

iv) Cancellation Property:

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

Also the limits:

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

exist and are finite for almost every $\mathbf{x} \in S$.

The strategy is now to exploit those estimates to derive L^p controls on the covering of S with our metric balls introduced in the 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 $(s_0, \mathbf{x}_0) \in S,\ R \leq R_0$, 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

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

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 $(s,\mathbf{x}) \in B((s_0,\mathbf{x}_0), R)$:

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

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

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

exist and are finite.

Now from Propositions 4.2 and 4.3 we derive from Theorem 3 in [Bra10], recalling from Proposition 4.1 that d and d^* are equivalent on 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

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

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

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 equation (4.8) 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. On the other hand, from the definition in (3.11) and (3.8), observe that for $f\in C_0^\infty([0,T]\times\mathbb{R}^{nd})$ and $p>(n^2d+2)/2$,

$$\exists C_{5.1} := C_{5.1}(p,T,(\mathbf{A})),\ \forall (s,\mathbf{x})\in[0,T)\times\mathbb{R}^{nd},$$
$$|\tilde{G}f(s,\mathbf{x})| \le C_{5.1}T^{1-(n^2d+2)/(2p)}\|f\|_{L^p([s,T)\times\mathbb{R}^{nd})}. \tag{5.1}$$

Assume now that $\varepsilon_{\mathbf{F}}\le\varepsilon_0(n,d)$ small enough and $\varepsilon_a\le C_{3.17}^{-1}$ in the local condition (3.1). Choose also $T<(4C(p,(\mathbf{A})))^{-2}$ in (3.18). Exploiting (5.1), (3.18), the uniqueness to the martingale problem can be derived following the steps of Section 7.1 in [SV79].

Now, without the local condition, the continuity assumed in **(A)** allows to localize. Precisely, for all $(s,\mathbf{x})\in[0,T]\times\mathbb{R}^{nd}$ we can find $\delta>0$ s.t. if $\mathcal{G} := [(s-\delta)\vee 0, s+\delta\wedge T]\times B(\mathbf{x},\delta)$ then

$$\sup_{(t,\mathbf{y})\in\mathcal{G}}|a(t,\mathbf{y})-a(t,\mathbf{x})|\le\varepsilon_a,$$
$$\sup_{i\in[\![2,n]\!]}\sup_{(t,\mathbf{y})\in\mathcal{G},\ \mathbf{z}\in\mathbb{R}^{nd}}|D_{\mathbf{x}_{i-1}}F_i(t,\mathbf{y}_{i-1},\mathbf{z}^{i,n})-D_{\mathbf{x}_{i-1}}F_i(t,\mathbf{x}_{i-1},\mathbf{z}^{i,n})|\le\varepsilon_{\mathbf{F}}.$$

Define then $\forall(t,\mathbf{y})\in[0,T]\times\mathbb{R}^{nd},\ \tilde{a}(t,\mathbf{y}) = a(t,\mathbf{y})\mathbb{I}_{(t,\mathbf{y})\in\mathcal{G}} + (1-\mathbb{I}_{(t,\mathbf{y})\in\mathcal{G}})a(t,\mathbf{x})$ and $\tilde{F}$ s.t. $D_{\mathbf{x}_{i-1}}\tilde{F}_i(t,\mathbf{y}_{i-1},\mathbf{z}^{i,n}) := D_{\mathbf{x}_{i-1}}F_i(t,(\Pi_{\mathbf{x}_{i-1},\delta}(\mathbf{y}_{i-1}),\mathbf{z}^{i,n})$ where:

$$\Pi_{\mathbf{x}_{i-1},\delta}(\mathbf{y}_{i-1}) = \mathbf{y}_{i-1}\mathbb{I}_{|\mathbf{y}_{i-1}-\mathbf{x}_{i-1}|\le\delta} + \left(\mathbf{x}_{i-1}+\frac{\mathbf{y}_{i-1}-\mathbf{x}_{i-1}}{|\mathbf{y}_{i-1}-\mathbf{x}_{i-1}|}\delta\right)\mathbb{I}_{|\mathbf{y}_{i-1}-\mathbf{x}_{i-1}|>\delta}.$$

Then the coefficients $\tilde{a}, \tilde{\mathbf{F}}$ satisfy **(A)** and the local condition (3.1). We can then invoke Theorem 27 in Priola [Pri13] to derive global uniqueness in *small time*. Indeed, it is not hard to directly construct a covering of S with sets of the previous type. An alternative consists in exploiting the covering result of Theorem 4.1. Let us emphasize that we cannot here directly apply Theorem 6.6.1 in [SV79], which being valid for bounded coefficients would have required a truncation of the drift. It is not clear to derive Calderón-Zygmund estimates of the previous type that are uniform w.r.t. to the truncation.

The well posedness of the martingale for an arbitrary given $T > 0$ then follows from the Markov property (see Chapter 6.2 of [SV79] for details).

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.

- If the diffusion coefficient a is uniformly continuous, then, for $0 \leq s < T$, and $q \in [1, 2)$, 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 $\alpha = \left(\frac{(n^2 d+2)}{2} \right)(q-1)$ 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 $q \in [1, 2)$, $\delta > 0$,

$$\left(\int_s^T dt \int_{\mathbb{R}^{nd} \setminus B^E(\boldsymbol{\theta}_{t,s}(\mathbf{x}), \delta)} d\mathbf{y} |p(s,t,\mathbf{x},\mathbf{y})|^q \right)^{1/q} \leq C^2_{5.1}(1 + |\mathbf{x}|),$$

with $C^2_{5.1} := C^2_{5.1}(T, q, (\mathbf{A}), \delta_T, \delta)$, $B^E(\boldsymbol{\theta}_{t,s}(\mathbf{x}), \delta)$ standing 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 to (1.1) starting from $\mathbf{x}$ at time s).

Remark 5.1. There are three differences w.r.t. to the indicated theorem. 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 Calderon-Zygmund estimates of Theorem 3.1. At last, the upper bound on q comes from the control on the remainders in (3.16). When the system is linear w.r.t. the components that transmit the noise (see Remark 3.1) 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 that theorem, 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 to 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 (3.11)-(3.14)). 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), $\varepsilon_{\mathbf{F}} \leq \varepsilon_0$ and for all $\rho \in [r, \frac{n^2d+4}{2} \vee r]$, $(\frac{\varepsilon_a}{2} C_{3.17}(T,\rho,(\mathbf{A})) + C(\rho,(\mathbf{A}))T^{\frac{1}{2}}) < 3/4$ so that from (3.18), we have that $(I-R)^{-1}$ is consistent as bounded operator from $L^\rho([0,T]\times\mathbb{R}^{nd})$ into itself. In particular this imposes that $T \leq 1$ is sufficiently small. Setting then $K := \tilde{G} \circ (I-R)^{-1}$, we thus derive 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

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

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

Lemma 5.1. Let $a, \mathbf{F} \in \mathcal{A}(r,T)$. Then, for $N = \lceil \frac{n^2d}{2} \frac{1}{r} \rceil$, 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). Precisely, for all $(s,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$:

$$|K^{N+1} f(s,\mathbf{x})| \leq C_{5.1} \|f\|_{L^r([0,T]\times\mathbb{R}^{nd})},$$

where $C_{5.1} := C_{5.1}(T,r,(\mathbf{A}))$.

From Lemma 5.1, Lemma 9.1.3 in [SV79] and the well posedness of the martingale problem, denoting by $P(s,x,t,.)$ the associated transition function, one then gets: $Pf(s,x) = \int_s^T dt \int_{\mathbb{R}^{nd}} P(s,t,\mathbf{x},d\mathbf{y}) f(t,\mathbf{y}) = Kf(s,x)$ if $f \in L^\rho([0,T]\times\mathbb{R}^{nd}) \cap L^\infty([0,T]\times\mathbb{R}^{nd})$. Therefore, for $N = \lceil (n^2d+2)/2r \rceil$, $f \in C_0([0,T]\times\mathbb{R}^{nd})$ (functions with compact support),

$$|P^{N+1} f(s,\mathbf{x})| \leq C_{5.1} \|f\|_{L^r([0,T]\times\mathbb{R}^{nd})}. \tag{5.2}$$

This observation then yields the following result.

Lemma 5.2. If $a, \mathbf{F} \in \mathcal{A}(r,T)$ denoting by P the transition function associated to $(L_t)_{t\in[0,T]}$ then for $r \leq \rho \leq +\infty$,

$$\int_s^T dt (t-s)^N \int_{\mathbb{R}^{nd}} P(s,t,\mathbf{x},d\mathbf{y}) f(t,\mathbf{y}) \leq C^1_{5.2} \|f\|_{L^\rho([0,T]\times\mathbb{R}^{nd})},$$

with $C^1_{5.2} := C^1_{5.2}(T,r,(\mathbf{A}))$. Also, for each $\delta > 0$, $r < \rho \leq \infty$,

$$\int_s^T dt \int_{\mathbb{R}^{nd} \setminus B^E(\boldsymbol{\theta}_{t,s}(\mathbf{x}),\delta)} P(s,t,\mathbf{x},d\mathbf{y}) f(t,\mathbf{y}) \leq C^2_{5.2} \|f\|_{L^\rho([0,T]\times\mathbb{R}^{nd})},$$

where $C^2_{5.2} := C^2_{5.2}(T,r,\rho,(\mathbf{A}))$.

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 (5.2) 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:

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

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

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

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

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

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:

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

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 5.2: 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* coefficients of the previous section. We have the following *tubular* localization.

Lemma 5.3 (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 to $(L_t)_{t\in[s_0,T]}$. For $0 < R_1 < R_2$, $\mathbf{x}_0 \in \mathbb{R}^{nd}$ defining $\tau_{-1} = s_0$ and for all $n \in \mathbb{N}$,

$$\begin{aligned}
\tau_{2n} &:= \inf\{t \geq \tau_{2n-1} : |\mathbf{X}_t - \boldsymbol{\theta}_{t,s}(\mathbf{x}_0)| = R_2\}, \\
\tau_{2n+1} &:= \inf\{t \geq \tau_{2n} : |\mathbf{X}_t - \boldsymbol{\theta}_{t,s}(\mathbf{x}_0)| = R_1\},
\end{aligned}$$

then

$$\mathbb{E}^{\mathbb{P}_{s_0,\mathbf{x}_1}}[\sum_{n\geq 0} \mathbb{I}_{\tau_{2n}\in[0,T]}] \leq C_{5.3} := C_{5.3}(T,(\mathbf{A}),R_2 - R_1).$$

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 5.2.

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 5.3. 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.

Lemma 5.4 (First Localization Lemma). Let P 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 to the operator

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

is well posed and that $\tilde{\mathbf{F}}=\mathbf{F},\tilde{a}=a$ on $\mathcal{C}_{\underline{t},\bar{t},R}(s_0,\mathbf{x}_0):=\{(t,\mathbf{y})\in[0,T]\times\mathbb{R}^{nd}:t\in[\underline{t},\bar{t}],\boldsymbol{\theta}_{s_0,t}(\mathbf{y})\in B^E(\mathbf{x}_0,R)\}$ for some $0\leq\underline{t}<\bar{t}\leq T,\ R>0$. Then for each $\delta\in(0,R)$:

$$|\mathbb{E}^{\mathbb{P}}[\int_{s_0}^T f(t,\mathbf{X}_t)dt]|\leq\mathbb{E}^{\tilde{\mathbb{P}}_{s_0,\mathbf{x}_1}}[\int_{s_0}^T|f(t,\mathbf{X}_t)|dt]$$
$$+C_{5.4}\sup_{(s,\mathbf{x})\in\partial\mathcal{C}_{\underline{t},\bar{t},R-\delta}(s_0,\mathbf{x}_0)}\mathbb{E}^{\tilde{\mathbb{P}}_{s,\mathbf{x}}}[\int_s^T|f(t,\mathbf{X}_t)|dt],$$

for all $f\in C_0(\mathcal{C}_{\underline{t},\bar{t},R-\delta}(s_0,\mathbf{x}_0))$ and $C_{5.4}:=C_{5.4}(T,(\mathbf{A}),R,\delta)$.

We now specify how this Lemma needs to be used. As a direct corollary of Lemmas 5.4 and 5.2 we derive:

Lemma 5.5 (Second Localization Lemma). Let $\tilde{a}(s,\mathbf{x}):=a(s,\mathbf{x}),\tilde{\mathbf{F}}(s,\mathbf{x}):=\mathbf{F}(s,\mathbf{x})$ in $\mathcal{C}_{\underline{t},\bar{t},R}(s_0,\mathbf{x}_0)$ and $\tilde{a}(s,\mathbf{x}):=a(s,\mathbf{x}_0),\ \tilde{\mathbf{F}}(s,\mathbf{x}):=\mathbf{F}(s,\mathbf{x}_0)$ elsewhere. Assume that the functions $\tilde{a},\tilde{\mathbf{F}}\in\mathcal{A}(r,T)$ for some $r\in(2,+\infty)$. Let P 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 $r<\rho\leq+\infty$.

$$|\mathbb{E}^{\mathbb{P}}[\int_{s_0}^T(t-s_0)^N f(t,\mathbf{X}_t)dt]|\leq C^1_{5.5}\|f\|_{L^\rho([s_0,T]\times\mathbb{R}^{nd})},$$

for all $f\in C_0(\mathcal{C}_{\underline{t},\bar{t},\alpha}(s_0,\mathbf{x}_0))$, where $N=\lceil(n^2d+2)/2r\rceil$ and $C^1_{5.5}:=C^1_{5.5}(T,(\mathbf{A}),r,\rho,R-\alpha)$. If additionally, $|\mathbf{x}_1-\mathbf{x}_0|>\alpha$, then

$$|\mathbb{E}^{\mathbb{P}}[\int_{s_0}^T f(t,\mathbf{X}_t)dt]|\leq C^2_{5.5}\|f\|_{L^\rho([s_0,T]\times\mathbb{R}^{nd})},$$

where $C^2_{5.5}:=C^2_{5.5}(T,(\mathbf{A}),r,\rho,R-\alpha,|\mathbf{x}_1-\mathbf{x}_0|-\alpha)$.

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, $\mathcal{C}_{\underline{t},\bar{t},R}(s_0,\mathbf{x}_0)$ 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 $T\leq(4(\sup_{\rho\in[r,(\frac{n^2d+2}{2}\vee r)+1]}C(\rho,(\mathbf{A}))))^{-2}$ in (3.18) and set $\varepsilon^{-1}=C_{3.17}(T,r,(\mathbf{A}))\vee$

$C_{3.17}(T,(\frac{n^2d+2}{2}\vee r)+1,(\mathbf{A}))\vee\varepsilon_0^{-1}$. Let us introduce for a fixed starting point $(s,\mathbf{x})$ of the martingale problem, the spatial balls

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

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 $\mathbf{k}\in\mathbb{Z}^{nd},\ \mathbf{y}_0,\mathbf{y}_1\in Q_{\mathbf{k}},\ t\in[s,T]$,

$$\begin{aligned}&|a(t,\boldsymbol{\theta}_{t,s}(\mathbf{y}_0))-a(t,\boldsymbol{\theta}_{t,s}(\mathbf{y}_1))|\\&+\sup_{i\in[\![2,n]\!],0\leq t\leq T,(\mathbf{x},\mathbf{z})\in(\mathbb{R}^{nd})^2}|D_{\mathbf{x}_{i-1}}F_i(t,(\boldsymbol{\theta}_{t,s}(\mathbf{y}_0))_{i-1},\mathbf{z}^{i,n})\\&-D_{\mathbf{x}_{i-1}}F_i(t,((\boldsymbol{\theta}_{t,s}(\mathbf{y}_1))_{i-1})_{i-1},\mathbf{z}^{i,n})|\leq\varepsilon/2.\end{aligned}\tag{5.3}$$

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 coefficients $a,\mathbf{F}$ restricted to $\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\beta\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2}),\ \beta>1$, coincide with some $\tilde{a},\tilde{\mathbf{F}}$ belonging 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:

$$\boldsymbol{\theta}(T,s,Q_{\mathbf{k}}):=\bigcup_{i=0}^{\lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil-1}\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2}).\tag{5.4}$$

Now from Lemma 5.5 we get that for $\rho=(r+q')/2$, for all $i\in[\![0,\lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil-1]\!]$

$$\begin{aligned}&|\int_s^T dt(t-s)^N\int\mathbb{I}_{(t,\mathbf{y})\in\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})|\\&=|\int_{t_i^{\mathbf{k}}}^{t_{i+1}^{\mathbf{k}}}dt(t-s)^N\int_{\boldsymbol{\theta}_{t,s}(Q_{\mathbf{k}})}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})|\leq C_{5.5}\|f\|_{L^\rho([0,T]\times\mathbb{R}^{nd})},\end{aligned}\tag{5.5}$$

where $C_{5.5}:=C_{5.5}(T,(\mathbf{A}),r,\rho)$. On the other hand, comparing deviations along the characteristics allows once again to use Bernstein inequalities, similarly to the proof of Lemma 5.2. Namely,

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

using Gronwall's Lemma (see proof of Lemma 5.2) and the bi-Lipschitz property of the flow for the last but one inequality. We therefore obtain:

$$|\int_s^T dt(t-s)^N \int \mathbb{I}_{(t,\mathbf{y})\in\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})| \leq C_{5.6}\exp(-\bar{C}^{-1}\frac{|\mathbf{k}|^2\gamma^2}{Td^2})|f|_\infty, \tag{5.6}$$

where $C_{5.6} := C_{5.6}(T,n,d,N)$. We thus get by interpolation that for $\vartheta = 1-\frac{\rho}{q'} \in (0,1)$:

$$|\int_s^T dt(t-s)^N \int \mathbb{I}_{(t,\mathbf{y})\in\mathcal{C}_{t_i^{\mathbf{k}},t_{i+1}^{\mathbf{k}},\gamma}(s,\mathbf{x}+\mathbf{k}\gamma/(nd)^{1/2})}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})| \leq C_{5.5}^{1-\vartheta}C_{5.6}^{\vartheta}\exp(-\bar{C}^{-1}\frac{\vartheta|\mathbf{k}|^2\gamma^2}{Td^2})\|f\|_{L^{q'}([0,T]\times\mathbb{R}^{nd})}.$$

Summing for a given $\mathbf{k}\in\mathbb{Z}^d$ first over $i\in[\![0,\lceil C_2(1+|\mathbf{k}\gamma/(nd)^{1/2}+\mathbf{x}|)\rceil-1]\!]$ (that is according to (5.4) on $\boldsymbol{\theta}(T,s,Q_{\mathbf{k}})$) we obtain

$$\begin{aligned}
&|\int_s^T dt(t-s)^N \int \mathbb{I}_{(t,\mathbf{y})\in\boldsymbol{\theta}(T,s,Q_{\mathbf{k}})}P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})|\\
=&|\int_s^T dt(t-s)^N \int_{\boldsymbol{\theta}_{t,s}(Q_{\mathbf{k}})} P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})|\\
\leq&\tilde{C}(T^{1/2}+|\mathbf{x}|)C_{5.5}^{1-\vartheta}C_{5.6}^{\vartheta}\exp(-\bar{C}^{-1}\frac{\vartheta|\mathbf{k}|^2\gamma^2}{Td^2})\|f\|_{L^{q'}([0,T]\times\mathbb{R}^{nd})},\tilde{C}:=\tilde{C}((\mathbf{A}),\vartheta).
\end{aligned}$$

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

$$|\int_s^T dt(t-s)^N \int_{\mathbb{R}^{nd}} P(s,t,\mathbf{x},d\mathbf{y})f(t,\mathbf{y})| \leq C_{5.7}(1+|\mathbf{x}|)\|f\|_{L^{q'}([0,T]\times\mathbb{R}^{nd})}, C_{5.7}:=C_{5.7}(T,q,(\mathbf{A}),\gamma). \tag{5.7}$$

This contribution already emphasizes the main difference w.r.t. the non degenerate case: the estimate depends on the initial point. The proof can then be completed similarly to 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 which 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:

$$\begin{aligned}
\mathbb{E}_{\mathbf{x}}[\int_0^T f(t,\mathbf{X}_t)dt] &= \sum_{i=0}^{N-1}\mathbb{E}_{\mathbf{x}}[\mathbb{E}_{t_i,\mathbf{X}_{t_i}}[\int_{t_i}^{t_{i+1}} f(t,\mathbf{X}_t)dt]]\\
&= \Big\{\sum_{i=0}^{N-1}\mathbb{E}_{\mathbf{x}}[\mathbb{E}_{t_i,\mathbf{X}_{t_i}}[\int_{t_i}^{t_{i+1}} f(t,\mathbf{X}_t)\mathbb{I}_{\mathbf{X}_t\in B^E(\boldsymbol{\theta}_{t,t_i}(\mathbf{X}_{t_i}),\delta)}dt]]\Big\}\\
&+ \Big\{\sum_{i=0}^{N-1}\mathbb{E}_{\mathbf{x}}[\mathbb{E}_{t_i,\mathbf{X}_{t_i}}[\int_{t_i}^{t_{i+1}} f(t,\mathbf{X}_t)\mathbb{I}_{\mathbf{X}_t\notin B^E(\boldsymbol{\theta}_{t,t_i}(\mathbf{X}_{t_i}),\delta)}dt]]\Big\} := T_C+T_F,
\end{aligned} \tag{5.8}$$

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

(5.9)
$$|T_F| \leq C_{5.1}^2 \sum_{i=0}^{N-1} \mathbb{E}_{\mathbf{x}}[(1+|\mathbf{X}_{t_i}|)]\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})} \leq NC_{5.1}^2 C_F(1+|\mathbf{x}|)\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})},$$

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 (5.3)), 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 $t_i^j := t_i + h_i j,\ j \in [\![0, \lceil C_2(1+|\mathbf{X}_{t_i}|)\rceil$ then the coefficients $\tilde{a}, \tilde{\mathbf{F}}$ equal to $a, \mathbf{F}$ on $\mathcal{C}_{t_i^j, t_i^{j+1}, 2\delta}(t_i, \mathbf{X}_{t_i})$ and to $a(t_i, \mathbf{X}_{t_i}), \mathbf{F}(t_i, \mathbf{X}_{t_i})$ elsewhere, belong to the class $\mathcal{A}(r, T)$ for some $r \in (2, n^2 d/2 + 1)$. We then derive from Lemma 5.4:

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

Now, from equation (3.19), we derive

(5.10)
$$|T_C| \leq C\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})} \sum_{i=0}^{N-1} \mathbb{E}[\lceil C_2(1+|\mathbf{X}_{t_i}|)\rceil] \leq CN(1+|\mathbf{x}|)\|f\|_{L^p([0,T]\times\mathbb{R}^{nd})},$$

up to a modification of C. The result follows from (5.10), (5.9), (5.8).

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 balls, see equations (4.1), (4.2), that there exists $C_1 := C_1((\mathbf{A})) > 0$ s.t. for all $\delta > 0, (s, \mathbf{x}) \in S$,

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

On the other hand, introducing

$$\begin{aligned}
\bar{B}((s,\mathbf{x}), \delta) &:= \big\{(t,\mathbf{y}) \in S : |t-s| \leq \frac{\delta^2}{4},\ \rho_{\mathrm{Sp}}(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y}) \leq \frac{\delta}{2}\big\} \\
&= \bigcup_{|t-s| \leq \frac{\delta^2}{4},\ s \in [-T,T]} \{\mathbf{y} \in \mathbb{R}^{nd} : \rho_{\mathrm{Sp}}(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y}) \leq \frac{\delta}{2}\}, \\
\forall \mathbf{z} \in \mathbb{R}^{nd},\ \rho_{\mathrm{Sp}}(\mathbf{z}) &:= \sum_{i=1}^n |\mathbf{z}_i|^{1/(2i-1)},
\end{aligned}$$

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

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

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 $v \in I(t,s) := ([s,t]\mathbb{I}_{s<t}) \cup ([t,s]\mathbb{I}_{s\geq t})$,

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

which for $v = s$ yields:

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

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

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

where $C_2 := C_2((\mathbf{A}))$ and using Gronwall's Lemma for the last inequality. Using 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:

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

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

The above estimate does not exploit the fact that $d((s,\mathbf{x}),(t,\mathbf{y})) \leq 1$. This last assumption is actually needed for the components $i \in [\![1, n-1]\!]$ whose differential dynamics potentially involves 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,

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

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 $d((s,\mathbf{x}),(t,\mathbf{y})) \le 1$ then for all $j \in [\![1,n]\!]$, $|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_j| \le 1$ which for $j \ge i+1$ yields $|(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_j|^{1/(2i-1)} \le |(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})_j|^{1/(2j-1)}$. We therefore get up to a modification of C_4 that:

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

which together with (6.3) indeed gives that there exists $C_{4.1} := C_{4.1}((\mathbf{A}),T))$ s.t. $d((t,\mathbf{y}),(s,\mathbf{x})) \le C_{4.1} d((s,\mathbf{x}),(t,\mathbf{y}))$ for $(s,\mathbf{x}),(t,\mathbf{y}) \in S,\ d((s,\mathbf{x}),(t,\mathbf{y})) \le 1$ 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:

$$\begin{aligned} d((s,\mathbf{x}),(t,\mathbf{y})) &= \rho(|t-\sigma+\sigma-s|, \boldsymbol{\theta}_{t,s}(\mathbf{x}) - \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}) + \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}) - \mathbf{y}) \\ &\le K(\rho(|\sigma-s|, \boldsymbol{\theta}_{t,s}(\mathbf{x}) - \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})) + \rho(|t-\sigma|, \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi}) - \mathbf{y})) \\ &:= K(\rho(|\sigma-s|, \boldsymbol{\theta}_{t,s}(\mathbf{x}) - \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})) + d((\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))). \end{aligned} \tag{6.5}$$

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 $i \in [\![1,n]\!]$:

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

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

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

with $p_j = \frac{2i-1}{2j-1}$, $q_j = \frac{2i-1}{2(i-j)}$ in the last but one equality. Hence $\rho(\sigma-s, \boldsymbol{\theta}_{t,s}(\mathbf{x}) - \boldsymbol{\theta}_{t,\sigma}(\boldsymbol{\xi})) \le C_5 (d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})) + d((\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))$, $C_5 := C_5((\mathbf{A}),T)$, 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 to 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 controls given in (3.9).

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

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

We also have:

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

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

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

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

To prove the second control, we need an auxiliary important scaling Lemma whose proof can be derived from Lemma 3.6 in [DM10].

Lemma 6.2 (Scaling Lemma). Let $0 \leq s < t \leq T$ and $\mathbf{y} \in \mathbb{R}^{nd}$ be given. The resolvent and covariance matrices respectively associated to (3.4) and (3.6) can then be written in the following way:

$$\forall u \in [s,t],\ \tilde{\mathbf{R}}^{t,\mathbf{y}}(u,s) = \mathbb{T}_{t-s}\bar{\mathbf{R}}^{t,\mathbf{y}}_{\frac{u-s}{t-s}}\mathbb{T}_{t-s}^{-1},\ \tilde{\mathbf{K}}^{t,\mathbf{y}}(s,u) = (t-s)^{-1}\mathbb{T}_{t-s}\bar{\mathbf{K}}^{t,\mathbf{y}}_{\frac{u-s}{t-s}}\mathbb{T}_{t-s},$$

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

Observe now from equation (3.4) that $\tilde{\mathbf{R}}^{t,\mathbf{y}}(s,t)(\mathbf{y}-\tilde{\mathbf{m}}^{t,\mathbf{y}}(s,t)) = \boldsymbol{\theta}_{s,t}(\mathbf{y})$ (pull-back by the deterministic system of the final point $\mathbf{y}$ from t to s). Hence:

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

giving the r.h.s. Once again, the l.h.s. can be proved similarly. □

As a consequence of Lemma 6.2, we derive the following controls for the derivatives of the *frozen* density (3.5) (see also the arguments in Section 5 of [DM10]).

Proposition 6.1. 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

$\forall 0 \le s < t \le T, \forall (\mathbf{x},\mathbf{y}) \in (\mathbb{R}^{nd})^2$,

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

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

6.3. **Proof of Propositions 4.2 and 4.3.** This section is devoted to the proof of Propositions 4.2 and 4.3 which provide the key estimates to derive Theorem 3.1 under assumption **(A)**.

6.3.1. *Proof of Proposition 4.2.* Point *i)* can be derived for both kernels k^d_{ij}, $k^{d,*}_{ij}$ recalling from (4.5) and (3.9) that there exists $C := C((\mathbf{A}),T) \ge 1$ s.t. $\forall ((s,\mathbf{x}),(t,\mathbf{y})) \in S^2$,

$$\begin{aligned} &|k^d_{i,j}(s,t,\mathbf{x},\mathbf{y})| + |k^{d,*}_{i,j}(s,t,\mathbf{x},\mathbf{y})| \\ &\le \frac{C}{(t-s)^{1+n^2d/2}} \exp\left(-C^{-1}(t-s)|\mathbb{T}^{-1}_{t-s}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})^2|\right). \end{aligned} \tag{6.6}$$

Now for a given $c_1 > 2$, if $c_1|t-s|^{1/2} \ge d((s,\mathbf{x}),(t,\mathbf{y}))$ then the r.h.s of (6.6) can directly be upper bounded by $Cc_1^{2+n^2d}/d((s,\mathbf{x}),(t,\mathbf{y}))^{2+n^2d}$. On the other hand, if $c_1|t-s|^{1/2} \le d((s,\mathbf{x}),(t,\mathbf{y}))$ then, by definition of d in (4.2) we derive that $\exists i \in [\![1,n]\!]$ s.t.

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

This property yields:

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

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 $k^d_{i,j}$, recalling that below

$$c_\infty d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi})) \le d((s,\mathbf{x}),(t,\mathbf{y})) \le 1, \tag{6.7}$$

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

$$\begin{aligned} |k_{ij}(s,t,\mathbf{x},\mathbf{y}) - k_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y})| &\le |k_{ij}(s,t,\mathbf{x},\mathbf{y}) - k_{ij}(\sigma,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{\sigma,s}(\mathbf{x}),\mathbf{y})| \\ &\quad + |k_{ij}(\sigma,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{\sigma,s}(\mathbf{x}),\mathbf{y}) - k_{ij}(\sigma,t,\boldsymbol{\xi},\mathbf{y})| \\ &:= I_1(s,\sigma,t,\mathbf{x},\mathbf{y}) + I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y}). \end{aligned} \tag{6.8}$$

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 $\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}) - \mathbf{y}$. In I_2 the time variables are equal to $t-\sigma$. Also, the intermediate spatial point $\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{\sigma,s}(\mathbf{x})$

yields from I_2 a difference of the form $|\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi}|$ 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 (6.7). The condition $|\sigma-s|\leq K|t-\sigma|$ is here needed to use properties on the rescaled flows for the spatial sensitivity (see Lemma 6.1 and equation (6.16)) 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, writing

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

The above terms could be analyzed similarly to those appearing in (6.8) following the procedure below.

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

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

we can rewrite $I_1(s,\sigma,t,\mathbf{x},\mathbf{y})=|\bar{k}_{ij}(s,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}),\mathbf{y})-\bar{k}_{ij}(\sigma,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}),\mathbf{y})|$. 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.

$$\begin{aligned}I_1(s,\sigma,t,\mathbf{x},\mathbf{y})\quad&\leq\quad|s-\sigma|\sup_{\tau\in[s\wedge\sigma,(s\vee\sigma)\wedge t[}|\partial_\tau\bar{k}_{ij}(\tau,t,\mathbf{z},\mathbf{y})|_{\mathbf{z}=\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})}\\&\leq\quad C|s-\sigma|\times\sup_{\tau\in[s\wedge\sigma,(s\vee\sigma)\wedge t[}\frac{\bar{q}_c(\tau,t,\mathbf{z},\mathbf{y})}{(t-\tau)^2}|_{\mathbf{z}=\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})},\end{aligned}\tag{6.9}$$

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

Introduce now $\Sigma_1:=\{(u,\mathbf{z})\in\mathbb{R}\times\mathbb{R}^{nd}:\rho(u,\mathbf{z})=1\}$ 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 the mappings:

$$\begin{aligned}&J_1:(t,\mathbf{y})\in[-T,T]\times\mathbb{R}^{nd}\mapsto\left(s+\rho^2\tilde{s},\boldsymbol{\theta}_{t,s}(\mathbf{x})+\rho^{-1}\mathbb{T}_{\rho^2}\tilde{\mathbf{x}}\right),\ \rho:=\rho(t-s,\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x}))\\&\qquad(\tilde{s},\tilde{\mathbf{x}}):=\left((t-s)\rho^{-2},\rho\mathbb{T}_{\rho^{-2}}\left(\mathbf{y}-\boldsymbol{\theta}_{t,s}(\mathbf{x})\right)\right)\in\Sigma_1.\end{aligned}\tag{6.10}$$

$$\begin{aligned}&J_2:(\sigma,\boldsymbol{\xi})\in[-T,T]\times\mathbb{R}^{nd}\mapsto\left(s+\alpha^2\bar{s},\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})+\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}\right),\ \alpha:=\rho(\sigma-s,\boldsymbol{\xi}-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))\\&\qquad(\bar{s},\bar{\mathbf{x}}):=\left((\sigma-s)\alpha^{-2},\alpha\mathbb{T}_{\alpha^{-2}}\left(\boldsymbol{\xi}-\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})\right)\right)\in\Sigma_1.\end{aligned}\tag{6.11}$$

We have to consider the terms I_1,I_2 under the condition (6.7) 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 (6.10) we get that for all $\tau\in[s\wedge\sigma,(s\vee\sigma)\wedge t[$:

$$\begin{aligned}\frac{\bar{q}_c(\tau,t,\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}),\mathbf{y})}{(t-\tau)^2}\quad&\leq\quad\frac{C}{\rho^{4+n^2d}}\frac{1}{\left(\tilde{s}-\frac{\tau-s}{\rho^2}\right)^{2+n^2d/2}}\\&\times\exp\left(-c(t-\tau)|\mathbb{T}^{-1}_{t-\tau}(\rho^{-1}\mathbb{T}_{\rho^2}\{\tilde{\mathbf{x}}+\mathcal{R}^\rho_{t,s}(\mathbf{x},\mathbf{y})\})|^2\right),\end{aligned}\tag{6.12}$$

where $\mathcal{R}^{\rho}_{t,s}(\mathbf{x},\mathbf{y}) := \rho\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x}))$ is a term that takes into account the approximation of the non-linear flow by the linear 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 the above expression and (6.9) 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 (6.11) for the last inequality. On the other hand, from (6.7) $\rho \geq c_\infty\alpha$. 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:

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

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

$$\sum_{j=1}^{n} |\tilde{\mathbf{x}}_j|^{1/(2j-1)} \geq 1 - 2\tilde{c} \text{ and } \exists j_0 \in [\![1,n]\!],\ |\tilde{\mathbf{x}}_{j_0}|^{1/(2j_0-1)} \geq \frac{1-2\tilde{c}}{n} > 0,\ \text{for } c_\infty > 2. \tag{6.13}$$

Write now:

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

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, whose proof is for the sake of completeness proposed in Appendix A.

Lemma 6.3. Assume that **(A)** holds and that $0 < \rho \leq \Lambda$. Then, under the local condition (3.1), there exists $C_{6.3} := C_{6.3}((\mathbf{A}),\Lambda)$ s.t.:

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

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

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

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 $j \in [\![1,n]\!]$:

$$\begin{aligned}|\{\mathcal{R}^{\rho}_{t,s}(\mathbf{x},\mathbf{y})\}_j| &\leq C_{6.3}n(\varepsilon_{\mathbf{F}} + (t-s))|\tilde{\mathbf{x}}_{j_0}|,\\ |\{\mathcal{R}^{\rho}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})\}_j| &\leq C_{6.3}n(\varepsilon_{\mathbf{F}} + (t-\sigma))|\tilde{\mathbf{x}}_{j_0}|.\end{aligned} \tag{6.14}$$

Thus, for T and $\varepsilon_{\mathbf{F}}$ s.t. $C_{6.3}n(\varepsilon_{\mathbf{F}} + (t-s)) \leq 1/2$:

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

using (6.13) for the last inequality. Plugging the above control into (6.12) yields:

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

From the above computations, we finally get the global bound:

$$\exists C_1:=C_1(T,(\mathbf{A}))>0,\ I_1(s,\sigma,t,\mathbf{x},\mathbf{y})\leq\frac{C_1\alpha^2}{\rho^{4+n^2d}}. \tag{6.15}$$

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 (6.9) that $\exists(c,C):=(c,C)(T,(\mathbf{A}))>0$ s.t.:

$$\begin{aligned}I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y})&\leq C(t-\sigma)^{1/2}|\mathbb{T}^{-1}_{t-\sigma}(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}))|\\&\quad\times\frac{1}{(t-\sigma)}\sup_{\gamma\in[0,1]}\bar{q}_c(\sigma,t,\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}),\mathbf{y})\\&\leq C(t-\sigma)^{1/2}|\mathbb{T}^{-1}_{t-\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})-\boldsymbol{\xi})|\\&\quad\times\frac{1}{(t-\sigma)}\sup_{\gamma\in[0,1]}\bar{q}_c(\sigma,t,\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{\sigma,s}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi}),\mathbf{y}),\end{aligned} \tag{6.16}$$

using also Lemma 6.1 recalling that we have assumed $|\sigma-s|\leq K|t-\sigma|$ for the last inequality. Hence:

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

Thus, if $|\tilde{s}-\frac{\sigma-s}{\rho^2}|^{1/2}\geq\tilde{c}$, we get

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

using (6.11) for the last inequality. Recall now from (6.10), (6.11) that

$$\begin{aligned}&\left(\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\xi})\right)-\mathbf{y}\\&=\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})+(1-\gamma)\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x})+\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}})-\mathbf{y}\\&\overset{(3.4)}{=}\gamma\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})+(1-\gamma)\left\{\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))+\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}\right\}-\mathbf{y}\\&=\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y}+\gamma(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,s}(\mathbf{x})-\boldsymbol{\theta}_{t,s}(\mathbf{x}))+(1-\gamma)(\tilde{\boldsymbol{\theta}}^{t,\mathbf{y}}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}))-\boldsymbol{\theta}_{t,s}(\mathbf{x}))\\&\quad+(1-\gamma)\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}\\&:=-\rho^{-1}\mathbb{T}_{\rho^2}\{\tilde{\mathbf{x}}+\gamma\mathcal{R}^\rho_{t,s}(\mathbf{x},\mathbf{y})+(1-\gamma)\mathcal{R}^\rho_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})\}+(1-\gamma)\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}},\end{aligned} \tag{6.17}$$

where $(\bar{s},\bar{\mathbf{x}})\in\Sigma_1$. Observe that, from (6.14) we have for all $j\in[\![1,n]\!]$,

$$|(\mathcal{R}^{\rho}_{t,s}(\mathbf{x},\mathbf{y}))_j|+|(\mathcal{R}^{\rho}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y}))_j|\leq C_{6.3}n(2\varepsilon_{\mathbf{F}}+(t-\sigma)+(t-s))|\tilde{\mathbf{x}}_{j_0}|. \tag{6.18}$$

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

$$(t-\sigma)^{1/2}\mathbb{T}^{-1}_{t-\sigma}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}=\hat{\tilde{\mathbf{R}}}^{t,\mathbf{y},t-\sigma}(1,0)(t-\sigma)^{1/2}\mathbb{T}^{-1}_{t-\sigma}\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}}$$

with $|\hat{\tilde{\mathbf{R}}}^{t,\mathbf{y},t-\sigma}(1,0)|\leq\hat{C}:=\hat{C}(T,(\mathbf{A}))$. Thus, recalling that from the structure of the linearized system the resolvent is subdiagonal (see (3.3), (3.4)), we derive for all $j\in[\![1,n]\!]$:

$$\begin{aligned}(t-\sigma)^{1/2}|(\mathbb{T}^{-1}_{t-\sigma}\tilde{\mathbf{R}}^{t,\mathbf{y}}(t,\sigma)\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}})_j|&=(t-\sigma)^{1/2}|(\hat{\tilde{\mathbf{R}}}^{t,\mathbf{y},t-\sigma}(1,0)\mathbb{T}^{-1}_{t-\sigma}\alpha^{-1}\mathbb{T}_{\alpha^2}\bar{\mathbf{x}})_j|\\&\leq\hat{C}\sum_{i=1}^{j}\left(\frac{\alpha^2}{t-\sigma}\right)^{i-1/2}|\bar{\mathbf{x}}_i|\leq\hat{C}\sum_{i=1}^{j}\left(\frac{\alpha^2}{\rho^2\{\tilde{s}-\frac{\sigma-s}{\rho^2}\}}\right)^{i-1/2}|\bar{\mathbf{x}}|\\&\leq\hat{C}\left(\tilde{s}-\frac{\sigma-s}{\rho^2}\right)^{-(j-1/2)}\sum_{i=1}^{j}\left(\frac{\alpha^2}{\rho^2}\right)^{i-1/2}|\bar{\mathbf{x}}|\end{aligned} \tag{6.19}$$

as soon as $\tilde{c}\leq 1$ and $|\tilde{s}-\frac{\sigma-s}{\rho^2}|^{1/2}\leq\tilde{c}$. In that case, using (6.13), (6.17), (6.18), (6.19) we then derive that:

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

where $\bar{c}>0$ for $T,\varepsilon_{\mathbf{F}}$ small enough and a sufficiently large c_∞. We thus obtain the global bound:

$$\exists C_2:=C_2(T,(\mathbf{A}))>0,\ I_2(\sigma,t,\mathbf{x},\boldsymbol{\xi},\mathbf{y})\leq C_2\left(\sum_{k=1}^{n}\frac{\alpha^{2k-1}}{\rho^{2+(2k-1)+n^2d}}\right). \tag{6.20}$$

Plugging (6.15), (6.20) into (6.8) gives the point.

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

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

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

$$\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u) := \tilde{\mathbf{R}}^{s,\mathbf{x}}(t,s)\tilde{\mathbf{K}}^{\mathbf{x}}(t,u)\tilde{\mathbf{R}}^{s,\mathbf{x}}(t,s)^*. \tag{6.21}$$

Remark 6.3. Observe that from the above definition and the specific structure of the resolvent (see equations (3.3), (3.4)) we actually have $\det(\tilde{\mathbf{R}}^{s,\mathbf{x}}(t,s)) = 1$ and therefore $\det(\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)) = \det(\tilde{\mathbf{K}}^{\mathbf{x}}(t,u))$.

From the definition in (6.21) and (3.5) the exponential bounds write:

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

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

$$\begin{aligned}
&\check{k}^{s,\mathbf{x},u,t}(\mathbf{z}) := \mathbb{I}_{s-t>0}\left\{-[\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)^{-1}]_{1,1} + [\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)^{-1}\mathbf{z}]_1^{\otimes 2}\right\} \\
&\frac{1}{(2\pi)^{nd/2}\det(\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u))^{1/2}} \times \exp\left(-\frac{1}{2}\langle \tilde{\mathbf{H}}^{s,\mathbf{x}}(t,u)^{-1}\mathbf{z}, \mathbf{z}\rangle\right),
\end{aligned} \tag{6.22}$$

we can rewrite:

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

Now the terms $|\{D^d_{i,j}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))\}_1|$ and $|\{D^d_{i,j}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))\}_3|$ 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 (6.15), (6.20). 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 $|\{D^d_{i,j}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))\}_2|$ involves two different covariance matrices observed at the same time but that are respectively associated to 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 the scaling lemma 6.2, we rewrite:

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

where $\check{\tilde{\mathbf{H}}}^{s,\mathbf{x},\sigma,t}_1$, $\check{\tilde{\mathbf{H}}}^{\sigma,\boldsymbol{\xi},\sigma,t}_1$ are uniformly elliptic bounded matrices on $\mathbb{R}^{nd}\otimes\mathbb{R}^{nd}$. Thus,

$$\begin{aligned}
&\langle (\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,\sigma) - \tilde{\mathbf{H}}^{\sigma,\boldsymbol{\xi}}(t,\sigma))(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})), \mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})\rangle \\
&= \langle (\check{\tilde{\mathbf{H}}}^{s,\mathbf{x},\sigma,t}_1 - \check{\tilde{\mathbf{H}}}^{\sigma,\boldsymbol{\xi},\sigma,t}_1)((\sigma - t)^{-1/2}\mathbb{T}_{\sigma-t}(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x}))), (\sigma - t)^{-1/2}\mathbb{T}_{\sigma-t}(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x}))\rangle.
\end{aligned} \tag{6.23}$$

We now want to control the difference $(\check{\tilde{\mathbf{H}}}_1^{s,\mathbf{x},\sigma,t} - \check{\tilde{\mathbf{H}}}_1^{\sigma,\boldsymbol{\xi},\sigma,t})$ in (6.23). From (6.21) and the scaling Lemma 6.2:

$$\begin{aligned}
\check{\tilde{\mathbf{H}}}_1^{s,\mathbf{x},\sigma,t} - \check{\tilde{\mathbf{H}}}_1^{\sigma,\boldsymbol{\xi},\sigma,t} &= (\sigma-t)\int_t^\sigma \Big\{ \bar{\mathbf{R}}^{s,\mathbf{x}}_{\frac{t-u}{\sigma-t}} \mathbb{T}^{-1}_{\sigma-t} Ba(u,x_0) B^* \{\bar{\mathbf{R}}^{s,\mathbf{x}}_{\frac{t-u}{\sigma-t}} \mathbb{T}^{-1}_{\sigma-t}\}^* \\
&\quad - \bar{\mathbf{R}}^{\sigma,\boldsymbol{\xi}}_{\frac{t-u}{\sigma-t}} \mathbb{T}^{-1}_{\sigma-t} Ba(u,x_0) B^* \{\bar{\mathbf{R}}^{\sigma,\boldsymbol{\xi}}_{\frac{t-u}{\sigma-t}} \mathbb{T}^{-1}_{\sigma-t}\}^* \Big\} du \\
&= (\sigma-t)^{-1}\int_t^\sigma \Big\{ \bar{\mathbf{R}}^{s,\mathbf{x}}_{\frac{t-u}{\sigma-t}} Ba(u,x_0) B^* (\bar{\mathbf{R}}^{s,\mathbf{x}}_{\frac{t-u}{\sigma-t}})^* \\
&\quad - \bar{\mathbf{R}}^{\sigma,\boldsymbol{\xi}}_{\frac{t-u}{\sigma-t}} Ba(u,x_0) B^* (\bar{\mathbf{R}}^{\sigma,\boldsymbol{\xi}}_{\frac{t-u}{\sigma-t}})^* \Big\} du, \\
|\check{\tilde{\mathbf{H}}}_1^{s,\mathbf{x},\sigma,t} - \check{\tilde{\mathbf{H}}}_1^{\sigma,\boldsymbol{\xi},\sigma,t}| &= C(\sigma-t)^{-1}\int_t^\sigma \Big| \bar{\mathbf{R}}^{s,\mathbf{x}}_{\frac{t-u}{\sigma-t}} - \bar{\mathbf{R}}^{\sigma,\boldsymbol{\xi}}_{\frac{t-u}{\sigma-t}} \Big| du.
\end{aligned} \tag{6.24}$$

Let us now write:

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

where $C := C((\mathbf{A}))$, using the smoothness conditions assumed in **(S)**, the scaling Lemma 6.2 and Gronwall's Lemma for the last inequality. From the above equation and (6.24), (6.23), we thus derive:

$$\begin{aligned}
&\langle (\tilde{\mathbf{H}}^{s,\mathbf{x}}(t,\sigma) - \tilde{\mathbf{H}}^{\sigma,\boldsymbol{\xi}}(t,\sigma))(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})), \mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x})\rangle \\
&\leq C|\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}) - \boldsymbol{\xi}||(\sigma-t)^{-1/2}\mathbb{T}_{\sigma-t}(\mathbf{y} - \boldsymbol{\theta}_{t,s}(\mathbf{x}))|^2,\ C := C((\mathbf{A})).
\end{aligned} \tag{6.25}$$

Because of the non-degeneracy of a, the inverse matrices $(\check{\tilde{\mathbf{H}}}_1^{s,\mathbf{x},\sigma,t})^{-1}, (\check{\tilde{\mathbf{H}}}_1^{\sigma,\boldsymbol{\xi},\sigma,t})^{-1}$ have the same spatial Lipschitz regularity. Indeed, up to a change of coordinates on 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,

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

The difference of the determinants can be investigated similarly. We therefore derive:

$$|\{D^d_{i,j}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))\}_2| \leq C\frac{d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))}{|t-\sigma|}\bar{q}_c(\sigma,t,\boldsymbol{\theta}_{t,s}(\mathbf{x}) - \mathbf{y}),$$

which can be analyzed similarly to I_2 (see equation (6.8) and page 28) in the previous proof and yields:

(6.26)

$$|\{D^d_{i,j}((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}),(t,\mathbf{y}))\}_2| \leq C\frac{d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))}{d((s,\mathbf{x}),(t,\mathbf{y}))^{2+n^2d}} \leq C\frac{d((s,\mathbf{x}),(\sigma,\boldsymbol{\xi}))}{d((s,\mathbf{x}),(t,\mathbf{y}))^{3+n^2d}},$$

recalling that $d((s,\mathbf{x}),(t,\mathbf{y}) \leq 1$. This gives points *ii)* and *iii)*. The cancellation property *iv)* is the more subtle to derive. Let us first prove:

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

Write:

$$\int_{d((t,\mathbf{y}),(s,\mathbf{x}))>\epsilon} k^d_{i,j}(s,t,\mathbf{x},\mathbf{y})dtd\mathbf{y} = \int_{\rho(t-s,\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))\in(\epsilon,\delta)} \check{k}^{t,\mathbf{y},s,t}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))dtd\mathbf{y}$$
$$+\int_{\rho(t-s,\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))\in(\delta,2\delta)} \check{k}^{t,\mathbf{y},s,t}(\mathbf{x}-\boldsymbol{\theta}_{s,t}(\mathbf{y}))dtd\mathbf{y} := O^\epsilon_1 + O^\epsilon_2.$$

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) that $|O^\epsilon_2| \leq C := C(T,(\mathbf{A}),\delta)$. Let us now focus on O^ϵ_1. 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:

$$O^\epsilon_1 = \int_{\rho(t-s,\mathbf{z})\in(\epsilon,\delta)} \exp(-\frac{1}{2}\langle(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1}\mathbf{z},\mathbf{z}\rangle)$$
$$\times P_{i,j}(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1},\mathbf{z})\frac{1}{(2\pi)^{nd}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{1/2}}dtd\mathbf{z} + O(1),$$

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

Now:

$$\exp(-\frac{1}{2}\langle(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1}\mathbf{z},\mathbf{z}\rangle)$$
$$\times P_{i,j}(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1},\mathbf{z})\frac{1}{(2\pi)^{nd}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{1/2}}$$
$$= \partial_{\mathbf{z}^i_1\mathbf{z}^j_1}\left\{\exp(-\frac{1}{2}\langle(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1}\mathbf{z},\mathbf{z}\rangle)\frac{1}{(2\pi)^{nd}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{1/2}}\right\}$$
$$+\mathcal{R}_{i,j}(s,t,\mathbf{x},\mathbf{z}).$$

Reproducing the arguments that lead to (6.25) from (6.23), we observe that the sensitivities of the covariance matrices w.r.t. the freezing parameters do not induce additional singularities. Hence, it can be checked that:

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

The smoothness on the coefficient $\mathbf{F}$, which is assumed under **(A)** to be C^2 in space and s.t. the $D_{\mathbf{x}_{i-1}}\mathbf{F}_i$ is C^2 in space as well, is actually just required here to differentiate twice the dynamics of the resolvent which already corresponds to a first order linearization of the initial system. Such terms appear in $\mathcal{R}_{i,j}$ and derive

from the sensitivities of $(\mathbf{H}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1}$ w.r.t. $\mathbf{z}_1^i, \mathbf{z}_1^j$. We thus write

$$\begin{aligned} O_1^\epsilon &= \int_{\rho(t-s,\mathbf{z})\in(\epsilon,\delta)} \partial_{\mathbf{z}_1^i\mathbf{z}_1^j}\Big\{q_{\mathbf{H}}(s,t,\mathbf{z})\Big\}dtd\mathbf{z} + O(1), \\ q_{\mathbf{H}}(s,t,\mathbf{z}) &:= \frac{\exp(-\frac{1}{2}\langle(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{-1}\mathbf{z},\mathbf{z}\rangle)}{(2\pi)^{nd}\det(\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}(s,t))^{1/2}}. \end{aligned}$$

By the divergence theorem:

$$|O_1^\epsilon| \le \sum_{\beta\in\{\epsilon,\delta\}} \left| \int_{\rho(t-s,\mathbf{z})=\beta} \partial_{\mathbf{z}_1^i} q_{\mathbf{H}}(s,t,\mathbf{z}) n_j d\nu((t-s),\mathbf{z}) \right|,$$

Now, from the metric homogeneity (see Remark 4.1), the good scaling property (3.7) that is valid for $\tilde{\mathbf{H}}^{t,\boldsymbol{\theta}_{t,s}(\mathbf{x}-\mathbf{z})}$, it can be shown, changing variables similarly to (6.10), that for ϵ small enough:

$$|O_1^\epsilon| \le C \int_{\rho(|\tilde{s}|,\bar{\mathbf{z}})=1} \frac{d\nu_1(\tilde{s},\bar{\mathbf{z}})}{\tilde{s}^{n^2d/2+1/2}} \exp\left(-C|\tilde{s}||\mathbb{T}_{|\tilde{s}|}^{-1}\bar{\mathbf{z}}|^2\right) < +\infty, \tag{6.27}$$

and that O_1^ϵ admits a limit when $\epsilon \to 0$. From (6.27) we thus get point *iv)* for the kernel $k_{i,j}^d$. The proof of the cancellation property for the adjoint kernel can be proved similarly exploiting the equivalence of the "forward" and "backward" distance on compact sets (see Proposition 4.1). □

Appendix A. Proof of Lemma 6.3

Let us prove the first claim of the Lemma. Recalling equations (3.2), (3.3) and (3.13), (3.14), we write:

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

where 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 the small variation condition (3.1) yield that there exists $C_3 := C_3(T,(\mathbf{A}))$ s.t. for all $j \in [\![2,d]\!]$:

$$\begin{aligned}
|(\mathcal{R}^{\rho,3}_{t,s}(\mathbf{x},\mathbf{y}))_j| &\leq C_3\varepsilon_{\mathbf{F}}\rho^{1-2j}\int_s^t du|(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\boldsymbol{\theta}_{u,t}(\mathbf{y}))_{j-1}| \\
&\leq C_3\varepsilon_{\mathbf{F}}\rho^{-2}\int_s^t du(\sum_{k=2}^n \rho^{1-2(k-1)}|(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\boldsymbol{\theta}_{u,t}(\mathbf{y}))_{k-1}|) \\
&\leq C_3\varepsilon_{\mathbf{F}}\rho^{-2}\int_s^t du(\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\boldsymbol{\theta}_{u,t}(\mathbf{y}))|) \\
&\leq C_3\varepsilon_{\mathbf{F}}\rho^{-2}\int_s^t du(\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|) \\
&\leq C_3\varepsilon_{\mathbf{F}}\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|, \qquad \text{(A.2)}
\end{aligned}$$

up to a modification of C_3, using the bi-Lipschitz property of the flow $\boldsymbol{\theta}$ for the last but one inequality (see the end of the proof of Proposition 5.1 in [DM10] for details) and recalling from (6.10) that $|t-s|/\rho^2 \leq 1$ for the last one.

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 $j \in [\![1,n]\!]$:

$$\begin{aligned}
|(\mathcal{R}^{\rho,1}_{t,s}(\mathbf{x},\mathbf{y}))_j| &\leq C_1\rho^{1-2j}\int_s^t du\sum_{k=j}^n|(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\boldsymbol{\theta}_{u,t}(\mathbf{y}))_k| \\
&\leq C_1\int_s^t du\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{u,s}(\mathbf{x})-\boldsymbol{\theta}_{u,t}(\mathbf{y}))| \\
&\leq C_1(t-s)\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})| \qquad \text{(A.3)}
\end{aligned}$$

using once again the bi-Lipschitz property of the flow $\boldsymbol{\theta}$ for the last inequality.

Remark A.1. Let us observe carefully that the second inequality is valid in whole generality when ρ is bounded. On the other hand, under $(\mathbf{A}_{SD})$ there is only one entry corresponding to $k=j$ in the first inequality and the second one is always true. These are exactly the assumptions required in the statement of the Lemma.

Recall now that $\mathcal{R}^{\rho,2}_{t,s}(\mathbf{x},\mathbf{y})$ is the linear part of equation (A.1), i.e. it can be rewritten

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

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 (A.3), (A.2), (A.1) and Gronwall's Lemma we derive

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

recalling (6.10) for the last inequality. This gives the first part of the Lemma. For the second one, 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 the bi-Lipschitz property of the flow. The main differences are that the time integrals are taken between σ and t. Following the computations leading to (A.2), the contribution $\mathcal{R}^{\rho;3}_{t,\sigma}(\boldsymbol{\theta}_{\sigma,s}(\mathbf{x}),\mathbf{y})$ would be bounded by $C_3\varepsilon_{\mathbf{F}}\rho^{-2}|t-\sigma|\{\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|\} \leq C_3\varepsilon_{\mathbf{F}}\rho^{-2}(|t-s|+|s-\sigma|)\{\rho|\mathbb{T}_{\rho^{-2}}(\boldsymbol{\theta}_{t,s}(\mathbf{x})-\mathbf{y})|\}$. 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 (A.2) 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 (A.3) 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. □

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.

[BZ11] M. Bramanti and M. Zhu. L^p and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift. *Preprint*, 2011.

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

[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.

[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.

[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.

[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.

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

[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.

[Pri09] Enrico Priola. Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.*, 194(2):117–153, 2009.

[Pri13] Enrico Priola. On weak uniqueness for some degenerate SDEs by global L^p estimates. 2013.

[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.

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, Moscow.

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