Sharp Schauder Estimates for some Degenerate Kolmogorov Equations

Paul-Éric Chaudru de Raynal, Igor Honoré, Stéphane Menozzi

▶ **To cite this version:**

Paul-Éric Chaudru de Raynal, Igor Honoré, Stéphane Menozzi. Sharp Schauder Estimates for some Degenerate Kolmogorov Equations. Annali della Scuola Normale Superiore di Pisa, Classe di Scienze, 2021, pp.989-1089. 10.2422/2036-2145.201811_011 . hal-01906573v2

HAL Id: hal-01906573
https://hal.science/hal-01906573v2

Submitted on 7 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SHARP SCHAUDER ESTIMATES FOR SOME DEGENERATE KOLMOGOROV EQUATIONS

PAUL-ÉRIC CHAUDRU DE RAYNAL, IGOR HONORÉ AND STÉPHANE MENOZZI

Abstract. We provide here some sharp Schauder estimates for degenerate PDEs of Kolmogorov type when the coefficients lie in some suitable anisotropic Hölder spaces and the first order term is non-linear and unbounded. We proceed through a perturbative approach based on forward parametrix expansions.

Due to the low regularizing properties of the degenerate variables, for the procedure to work, we heavily exploit duality results between appropriate Besov spaces.

Our method can be seen as *constructive* and provides, even in the non-degenerate case, an alternative approach to Schauder estimates.

Keywords: Schauder estimates, Kolmogorov degenerate PDEs, parametrix, Besov spaces.
MSC: Primary: 34F05, 60H10 ; Secondary: 60H30.

1. Introduction and Main Results

1.1. **Statement of the problem.** For a fixed time horizon $T > 0$ and given integers $n, d \in \mathbb{N}$, we aim at proving Schauder estimates for degenerate scalar valued Kolmogorov PDEs of the form:

$$\begin{cases} \partial_t u(t,\mathbf{x}) + \langle \mathbf{F}(t,\mathbf{x}), \mathbf{D}u(t,\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1} u(t,\mathbf{x}) a(t,\mathbf{x})\big) = -f(t,\mathbf{x}),\ (t,\mathbf{x}) \in [0,T)\times\mathbb{R}^{nd}, \\ u(T,\mathbf{x}) = g(\mathbf{x}),\ \mathbf{x}\in\mathbb{R}^{nd}, \end{cases} \tag{1.1}$$

where $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_n) \in \mathbb{R}^{nd}$ and for each $i \in [\![1, n]\!]$, $\mathbf{x}_i \in \mathbb{R}^d$. Above, the source f and the terminal condition g are bounded and scalar valued mappings and the bounded diffusion matrix a is $\mathbb{R}^d \otimes \mathbb{R}^d$-valued. Also, $\mathbf{F}(t,\mathbf{x}) := (\mathbf{F}_1(t,\mathbf{x}), \cdots, \mathbf{F}_n(t,\mathbf{x}))$ is a vector of $\mathbb{R}^d$-valued unbounded mappings $\mathbf{F}_i$ which have, for $i \in [\![2, n]\!]$, the following structure:

$$\forall (t,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd},\ \mathbf{F}_i(t,\mathbf{x}) := \mathbf{F}_i(t,\mathbf{x}^{i-1:n}),\ \mathbf{x}^{i-1:n} := (\mathbf{x}_{i-1}, \cdots, \mathbf{x}_n). \tag{1.2}$$

The notation $\mathbf{D} = (D_{\mathbf{x}_1}, \cdots, D_{\mathbf{x}_n})$ stands for the full spatial gradient and $D_{\mathbf{x}_i}$ denotes the partial gradient w.r.t. to $\mathbf{x}_i$. For notational convenience we will denote the spatial operator in (1.1) by $(L_t)_{t\in[0,T]}$, i.e. for any $\varphi \in C^2_0(\mathbb{R}^{nd}, \mathbb{R})$ (space of twice continuously differentiable functions with compact support):

$$L_t\varphi(\mathbf{x}) = \langle \mathbf{F}(t,\mathbf{x}), \mathbf{D}\varphi(\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1}\varphi(\mathbf{x}) a(t,\mathbf{x})\big). \tag{1.3}$$

In this work, this operator is supposed to be of weak Hörmander type, i.e. we suppose that the diffusion matrix a is uniformly elliptic and that the matrices $\big(D_{\mathbf{x}_{i-1}}\mathbf{F}_i(t,\cdot)\big)_{i\in[\![2,n]\!]}$ have full rank.

Acknowledgments: For the first Author, this work has been partially supported by the ANR project ANR-15-IDEX-02. For the third author, the study has been funded by the Grant RSF 17-11-01098 from the Russian Science Foundation.

1.2. **Preliminaries.** Under suitable regularity assumptions on $a, \mathbf{F}$ it can be shown that the martingale problem associated with (1.3) is well posed, see e.g. [27], [28], [4]. In that case, there exists a unique weak solution to the stochastic differential equation

$$\begin{aligned} d\mathbf{X}^1_t &= \mathbf{F}_1(t,\mathbf{X}^1_t,\ldots,\mathbf{X}^n_t)dt + \sigma(t,\mathbf{X}^1_t,\ldots,\mathbf{X}^n_t)dW_t, \\ d\mathbf{X}^2_t &= \mathbf{F}_2(t,\mathbf{X}^1_t,\ldots,\mathbf{X}^n_t)dt, \\ d\mathbf{X}^3_t &= \mathbf{F}_3(t,\mathbf{X}^2_t,\ldots,\mathbf{X}^n_t)dt, \qquad\qquad t \geq 0, \\ &\vdots \\ d\mathbf{X}^n_t &= \mathbf{F}_n(t,\mathbf{X}^{n-1}_t,\mathbf{X}^n_t)dt, \end{aligned} \tag{1.4}$$

where $(W_t)_{t\geq 0}$ is a Brownian motion on some filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. The operator $(L_t)_{t\geq 0}$ then corresponds to the generator of the process in (1.4) where σ is a square root of a. The well posedness of the martingale problem in particular implies that (1.1) admits a solution in the mild sense on a suitable

function space (see e.g. Kolokoltsov [14]).

Starting from this framework and assuming that the coefficients $a, \mathbf{F}$, the source f as well as the terminal condition g lie in appropriate Hölder spaces, we here prove that the Cauchy problem (1.1) is well posed in the weak sense and that its unique weak solution satisfies some appropriate Schauder estimates.

These Schauder estimates allow in particular to quantify precisely the *parabolic bootstrap* associated with $(L_t)_{t\geq 0}$, hence emphasizing its intrinsic regularization properties. In [4], the authors show how in a Hölder framework for the coefficients $\mathbf{F}, \sigma$, some minimal thresholds allow to guarantee the well posedness of the martingale problem associated with $(L_t)_{t\geq 0}$ and how such thresholds depend on the level $i \in [\![1, n]\!]$ of the chain and the variable $j \in [\![(i-1) \vee 1, n]\!]$. Adding to these thresholds some suitable regularity (depending also on the level and on the variable) allows to derive Schauder estimates. They provide accurate quantitative bounds associated with the global smoothing effect of $(L_t)_{t\geq 0}$ and reflect how the additional regularity propagates through the operator. From another point of view, which relies on the associated stochastic system (1.4), such estimates quantify the *regularizing* properties of the Brownian motion when propagating through the aforementioned system. In any case, they underline the degenerate structure of the operator $(L_t)_{t\geq 0}$ (or equivalently of the system (1.4)) which leads to more tricky smoothing effects than those exhibited in the non-degenerate setting.

It is indeed known from the seminal work of Lunardi [24] on the topic that Schauder estimates for degenerate Kolmogorov equations differ from those in the usual non-degenerate setting (see e.g. [16] or [10]). They reflect in some sense the multiple scales in the systems (1.1) and (1.4) (see Section 1.3 below) and are stated in terms of anisotropic Hölder spaces. In particular, those spaces emphasize that the higher is the index of the considered variable in $[\![1, n]\!]$, the weaker is the associated regularity gain.

Mathematical background. Let us shortly describe particular cases of dynamics of type (1.1) for which some Schauder estimates have already been proved. We again mention the article by Lunardi [24], who considered the special case of a homogeneous linear drift $\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ satisfying the structure condition (1.2). Precisely, the matrix $\mathbf{A}$ writes in this case:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_{1,1} & \cdots & \cdots & \cdots & \mathbf{a}_{1,n} \\ \mathbf{a}_{2,1} & \cdots & \cdots & \cdots & \mathbf{a}_{2,n} \\ \mathbf{0}_{d,d} & \mathbf{a}_{3,2} & \cdots & \cdots & \mathbf{a}_{3,n} \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & \mathbf{a}_{n,n-1} & \mathbf{a}_{n,n} \end{pmatrix},$$

where the entries $(\mathbf{a}_{i,j})_{ij\in[\![1,n]\!]^2}$ are in $\mathbb{R}^d \otimes \mathbb{R}^d$ s.t. $\left(\mathbf{a}_{i,i-1}\right)_{i\in[\![2,n]\!]}$ are non-degenerate elements of $\mathbb{R}^d \otimes \mathbb{R}^d$ (which expresses the weak Hörmander condition)*. Also, the homogeneous diffusion coefficient a belongs to an appropriate anisotropic Hölder space and asymptotically converges when $|\mathbf{x}| \to \infty$ to a non-degenerate constant matrix of $\mathbb{R}^d \otimes \mathbb{R}^d$. The assumptions on the asymptotic behavior on the diffusion coefficient have then been relaxed by Lorenzi [22], in the kinetic framework, i.e. $n = 2$ with the notations of (1.4), up to additional regularity assumptions on a which could also be unbounded.

Priola established later in [29] Schauder estimates, without dimensional constraints for time homogeneous drifts of the form

$$\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \begin{pmatrix} \tilde{\mathbf{F}}_1(\mathbf{x}) \\ \mathbf{0}_{(n-1)d,d} \end{pmatrix}, \tag{1.5}$$

for a non-linear drift $\tilde{\mathbf{F}}_1$ acting on the non-degenerate variable in the expected anisotropic Hölder space. The underlying technique consisted in establishing bounds on the derivatives of the semi-group of the perturbed degenerate Ornstein-Uhlenbeck process (i.e. with $\tilde{\mathbf{F}}_1$) from the usual unperturbed one (with $(\mathbf{A}\mathbf{x})_1$ only) through the Girsanov theorem assuming first that $\tilde{\mathbf{F}}_1$ is smooth. This initial smoothness of $\tilde{\mathbf{F}}_1$ is required in order to compute the associated tangent flows. Through the continuity approach, the author then managed to obtain the estimates for a bounded variable diffusion coefficient lying in the natural *expected* Hölder space similar to the one of [24] with the same asymptotic conditions. The smoothness of $\tilde{\mathbf{F}}_1$ is relaxed through an approximation procedure viewing the difference between the Hölder drift and its mollification as a source term and exploiting the estimates established for the *smooth* drift.

Finally, we mention the work of Di Francesco and Polidoro [8] who derived Schauder estimates for a linear drift of the previous type using an alternative notion of continuity regarding the diffusion coefficient a, which

*Actually the non-zero entries $(\mathbf{a}_{i,j})$ of $\mathbf{A}$ can be non-square and simply have full rank in [24]. We restrict here to square matrices for the sake of simplicity.

somehow involves the unbounded transport associated with the drift.

Hence, in the current framework of degenerate Kolmogorov equations, focusing on the drift, the Schauder estimates hold, to the best of our knowledge, for either linear drifts or Hölder perturbations on the non-degenerate variable of a linear drift.

Mathematical outline. There will be two main difficulties to overcome in order to prove Schauder estimates in our framework: the degeneracy as well as the non-linearity and unboundedness of the drift. Concerning this second issue let us also mention that, in the non-degenerate setting, Schauder estimates for unbounded non-linear drift coefficients were obtained under mild smoothness assumptions by Krylov and Priola [17] who heavily used the flow associated with the first order vector field in L_t, i.e. $\dot{\boldsymbol{\theta}}_t(\mathbf{x}) = \mathbf{F}(t, \boldsymbol{\theta}_t(\mathbf{x}))$, to precisely get rid of the unbounded terms.

In this work we will prove, in the framework of Hölder spaces for the source f, the terminal condition g and the coefficients $a, \mathbf{F}$, Schauder estimates similar to those of the previously quoted works ([24], [22], [29]). The diffusion coefficient a and the source term f will have, as in the non-degenerate case, the same regularity. We mention that, in contrast with the non-degenerate case, this will not be the case for the drift $\mathbf{F}$ for which some additional smoothness on the degenerate entries $(\mathbf{F}_i)_{i\in[\![2,n]\!]}$ is needed to guarantee the well posedness of (1.4). In particular, the Hölder indexes of $\mathbf{F}$ will be above the minimal thresholds appearing in [4]. The flow associated with the drift term will again play a key role in our setting. Eventually, we do not impose any particular spatial asymptotic condition on the diffusion a.

To prove our result, we will here proceed through a perturbative approach. The idea is to perform a first order *parametrix* expansion (or Duhamel expansion) of a solution of (1.1) with mollified coefficients around a suitable linearized Ornstein-Uhlenbeck type semi-group. The main idea behind consists in exploiting this easier framework in order to subsequently obtain a tractable control on the error expansion. When applying such a strategy, the common way to proceed consists in adopting the so-called *backward parametrix approach*, as successfully considered by Il'in *et al.* [13], Friedman [10] or McKean and Singer [26] in the non degenerate setting. This technique has been extended to the current degenerate setting, which involves unbounded coefficients, and successfully exploited for handling the corresponding martingale problem or density estimates of the fundamental solution of (1.1) in [4] and [7].

Unfortunately, this approach does not seem very adapted to derive Schauder estimates since it does not allow to easily deal with gradient and sensitivity estimates associated with the degenerate directions.

We therefore adopt here an alternative *forward parametrix approach* which is better designed to deal with gradient estimates. It has indeed been successfully applied e.g. in [2], [3] to derive strong uniqueness for degenerate kinetic SDEs of type (1.4) (i.e. $n = 2$ with the previous notations). Especially, this approach is better tailored to exploit cancellation techniques which are crucial when derivatives come in, as opposed to the *backward* one.

The perturbative approach is not usual to establish Schauder type estimates. The standard way is to proceed through *a priori* estimates to establish for a given solution of the PDE in a suitable function space, the expected bound. Existence and uniqueness issues, in the considered function space, for the solution of the equation are addressed in a second time. We can refer to [16] for a clear presentation of this approach and to [17] for an extension of this method to non-degenerate operators with unbounded drift coefficients. We will here obtain that the solutions of (1.1) with mollified coefficients satisfy, uniformly w.r.t. the mollification parameter, a Schauder type estimate (see Sections 3 to 5 below). From the well posedness of the martingale problem established in [4] under our current assumptions, we will then derive that the martingale solution to (1.1) actually itself satisfies the Schauder controls. Since we want to be in the *sharpest* possible Hölder setting for the coefficients, source and terminal functions, we will need to establish some subtle controls (in particular we have no *true* derivatives of the coefficients) which will heavily rely on duality results for Besov spaces (see Section 4.2 below and e.g. Chapter 3 in Lemarié-Rieusset [20]).

Let us emphasize that the perturbative approach developed here provides, even in the non-degenerate case, a new alternative to establish Schauder estimates. It can be seen as a constructive one in the sense that, from a sequence of smooth solutions, that uniformly satisfy the expected control, we will extract through convergence in law arguments a limit solution which also satisfies the bound. Uniqueness of the solution in the considered class then again follows from uniqueness in law of the underlying limit process.

The drawback of our approach is that, for the parabolic problem (1.1), we first have to establish our estimates in small time. This is intuitively clear since the perturbative methods (expansions along an *ad hoc proxy*)

are precisely designed for small times. To obtain the result for an arbitrary given time, we then have to iterate the estimate, which is precisely natural since Schauder estimates provide a kind of stability in the considered function space. We are therefore *far* from the *optimal* constants for the Schauder estimates established in the non-degenerate setting for time dependent coefficients by Krylov and Priola [18]. However, to the best of our knowledge, only two approaches allow to derive Schauder estimates for parabolic degenerate Kolmogorov equations with fully non-linear drift in Hölder space: ours and the one of Hao *al.* [11] developed in the kinetic non-local case and based on Littlewood-Paley decompositions. Also, we think the strategy developed in the current work should apply for the elliptic degenerate Kolmogorov equation with *good* potential term, i.e. the *negative sign* of the potential would allow to integrate on an infinite time horizon (getting therefore rid of the small time constraint) or for unbounded sources f that would need in that cases to be somehow controlled by an associated potential as in [17].

To conclude with this outline, we come back to the stochastic counterpart of (1.1): system (1.4). In connection with the *strong* regularizing properties of Brownian motion, the perturbative approach we develop here allows as well to address the problem of strong well posedness for the SDE (1.4) in a Hölder framework for the coefficients. This was done by Chaudru de Raynal in [3] for $n = 2$ (let us mention as well in this *kinetic case* the works of Fedrizzi *et al.* [9] and Zhang [32] who derived strong uniqueness for L^p drifts on the non-degenerate component and a linear degenerate dynamics in (1.4)). Namely, in the companion paper [6], we establish through a similar approach strong uniqueness for the full chain for some suitable related Hölder thresholds for the drift. Indeed, from a PDE viewpoint, strong uniqueness for the associated SDE is heavily related to pointwise controls of the gradient of the solution of the PDE in all the directions (including the degenerate ones). These controls hence require some additional regularity and are then obtained under some slightly stronger Hölder regularity assumptions on the coefficients (which for strong uniqueness issues then in turn become the source term in the PDE with the Zvonkin approach). Some related issues were also considered under additional smoothness conditions by Lorenzi [21] in the case of a linear drift.

Before stating our main results, we recall some properties associated with the system (1.1). We first describe in Section 1.3 how the intrinsic multi-scales of the degenerate Kolmogorov like equations appear. We then introduce the appropriate setting of Hölder spaces to consider in Section 1.4. We eventually conclude the introduction stating in Section 1.5 our main results concerning Schauder estimates associated with (1.1).

1.3. **Intrinsic scales of the system and associated distance.** Let us now briefly expose how the system *typically* behaves. To do so, consider the following operator:

$$\mathscr{L}_0 := \partial_t + \langle \mathbf{A}_0 \mathbf{x}, \mathbf{D} \rangle + \frac{1}{2}\Delta_{\mathbf{x}_1}, \ \mathbf{A}_0 = \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{I}_{d,d} & \ddots & \cdots & \cdots & \vdots \\ \mathbf{0}_{d,d} & \mathbf{I}_{d,d} & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & \mathbf{I}_{d,d} & \mathbf{0}_{d,d} \end{pmatrix}, \tag{1.6}$$

which can be viewed as a typical model for the operator in (1.1). Introducing now, for $\lambda > 0$, the dilation operator $\delta_\lambda : (t, \mathbf{x}) \in \mathbb{R}^+ \times \mathbb{R}^{nd} \mapsto \delta_\lambda(t, \mathbf{x}) = \big(\lambda^2 t, \lambda \mathbf{x}_1, \lambda^3 \mathbf{x}_2, \cdots, \lambda^{2n-1} \mathbf{x}_n\big) \in \mathbb{R}^+ \times \mathbb{R}^{nd}$, i.e. with a slight abuse of notation, $\big(\delta_\lambda(t, \mathbf{x})\big)_0 := \lambda^2 t$ and for each $i \in [\![1, n]\!]$, $\big(\delta_\lambda(t, \mathbf{x})\big)_i := \lambda^{2i-1} \mathbf{x}_i$, we have that

$$\mathscr{L}_0 v = 0 \implies \mathscr{L}_0(v \circ \delta_\lambda) = 0. \tag{1.7}$$

This hence leads us to introduce the homogeneous quasi-distance[†] corresponding to the dilation operator δ_λ. Precisely, setting for all $0 \le t \le s < +\infty$, $(\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^{nd})^2$:

$$\mathbf{d}_P\big((t, \mathbf{x}), (s, \mathbf{y})\big) = (s - t)^{\frac{1}{2}} + \sum_{i=1}^{d} |\mathbf{y}_i - \mathbf{x}_i|^{\frac{1}{2i-1}}, \tag{1.8}$$

we indeed have $\mathbf{d}_P\big(\delta_\lambda\big((t, \mathbf{x})\big), \delta_\lambda\big((s, \mathbf{y})\big)\big) = \lambda \mathbf{d}_P\big((t, \mathbf{x}), (s, \mathbf{y})\big)$. In our current setting we will mainly use the spatial part deriving from the parabolic homogeneous quasi-distance $\mathbf{d}_P$ in (1.8). We set accordingly,

$$\mathbf{d}\big(\mathbf{x}, \mathbf{y}\big) = \sum_{i=1}^{d} |\mathbf{y}_i - \mathbf{x}_i|^{\frac{1}{2i-1}}. \tag{1.9}$$

[†]the triangle inequality holds up to some multiplicative constant.

From a technical point of view, these (quasi-)distances express the spatial homogeneity associated with the intrinsic time scales of the variances for the Gaussian process with generator $\mathscr{L}_0$. From a more probabilistic viewpoint the exponents in (1.9) can be related to the characteristic time-scales of the iterated integrals of the Brownian motion (see e.g. [7], [4]).

The non-degeneracy and boundedness assumption on a as well as the Hörmander condition on the $(\mathbf{F}_i)_{i\in[\![2,n]\!]}$ that we assume for (1.1)-(1.4) (see assumptions **(UE)** and **(H)** below) will allow us to consider for our analysis the previous quasi-distances associated with the simplest yet typical equation in the class described by (1.1)-(1.4).

1.4. **Associated Hölder spaces.** We first recall some useful notations and spaces. We denote for $k\in\mathbb{N},\beta\in(0,1)$ by $\|\cdot\|_{C^{k+\beta}(\mathbb{R}^m,\mathbb{R}^\ell)}$, $m\in\{1,d,nd\}$, $\ell\in\{1,d,d^2,nd\}$ the usual homogeneous Hölder semi-norm, see e.g. Krylov [16]. Precisely, for $\psi\in C^{k+\beta}(\mathbb{R}^m,\mathbb{R}^\ell)$, denoting by $\vartheta=(\vartheta_1,\cdots,\vartheta_m)\in\mathbb{N}^m$ a generic multi-index and $|\vartheta|=\sum_{i=1}^m\vartheta_i$, we define the semi-norm:

$$\begin{aligned}\|\psi\|_{C^{k+\beta}(\mathbb{R}^m,\mathbb{R}^\ell)} &:= \sum_{i=1}^k\sup_{|\vartheta|=i}\|D^\vartheta\psi\|_{L^\infty(\mathbb{R}^m,\mathbb{R}^\ell)}+\sup_{|\vartheta|=k}[D^\vartheta\psi]_\beta,\\ [D^\vartheta\psi]_\beta &:= \sup_{(x,y)\in(\mathbb{R}^m)^2,x\neq y}\frac{|D^\vartheta\psi(x)-D^\vartheta\psi(y)|}{|x-y|^\beta},\end{aligned}\tag{1.10}$$

where $|\cdot|$ denotes the Euclidean norm on the considered space. We will also need to consider the associated subspace with bounded elements. Namely, we set:

$$C_b^{k+\beta}(\mathbb{R}^m,\mathbb{R}^\ell):=\{\psi\in C^{k+\beta}(\mathbb{R}^m,\mathbb{R}^\ell):\|\psi\|_{L^\infty(\mathbb{R}^m,\mathbb{R}^\ell)}<+\infty\}.$$

We define correspondingly the Hölder norm:

$$\|\psi\|_{C_b^{k+\beta}(\mathbb{R}^m,\mathbb{R}^\ell)}:=\|\psi\|_{C^{k+\beta}(\mathbb{R}^m,\mathbb{R}^\ell)}+\|\psi\|_{L^\infty(\mathbb{R}^m,\mathbb{R}^\ell)}.\tag{1.11}$$

We are now in position to define our anisotropic Hölder spaces with multi-index of regularity. Let $\psi:\mathbb{R}^{nd}\to\mathbb{R}^\ell$ be a smooth function. We first introduce, for $i\in[\![1,n]\!]$, $x\in\mathbb{R}^d$ the perturbation operator that writes:

$$\forall\mathbf{z}\in\mathbb{R}^{nd},\Pi_i^x(\psi)(\mathbf{z}):=\psi(\mathbf{z}_1,\cdots,\mathbf{z}_i+x,\cdots,\mathbf{z}_n).\tag{1.12}$$

We then define for each $i\in[\![1,n]\!]$, the mapping

$$(\mathbf{z},x)\in\mathbb{R}^{nd}\times\mathbb{R}^d\longmapsto\psi_i(\mathbf{z},x):=\Pi_i^x(\psi)(\mathbf{z}).\tag{1.13}$$

Let us introduce the following anisotropic Hölder space in $\mathbf{d}$-quasi-metric: given a parameter $\gamma\in(0,1)$, and $k\in\mathbb{N}$, we say that ψ is in $C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd},\mathbb{R}^\ell)$, if

$$\|\psi\|_{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd},\mathbb{R}^\ell)}:=\sum_{i=1}^n\sup_{\mathbf{z}\in\mathbb{R}^{nd}}\|\psi_i(\mathbf{z},\cdot)\|_{C^{\frac{k+\gamma}{2i-1}}(\mathbb{R}^d,\mathbb{R}^\ell)}<+\infty.\tag{1.14}$$

For the sake of simplicity, we will write:

$$\|\psi\|_{L^\infty}:=\|\psi\|_{L^\infty(\mathbb{R}^{nd},\mathbb{R}^\ell)},\text{ and }\|\psi\|_{C_{\mathbf{d}}^{k+\gamma}}:=\|\psi\|_{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd},\mathbb{R}^\ell)}.$$

The subscript $\mathbf{d}$ stands here to indicate the dependence of the Hölder exponents appearing in the r.h.s. on the underlying quasi-distance $\mathbf{d}$ reflecting the scale invariance of the system (see equation (1.9) and the comments above for details). Note in particular that, for $k=0$, there exists $C:=C(n,d)\geq1$ s.t.:

$$\begin{aligned}C^{-1}[\psi]_{\gamma,\mathbf{d}}\leq\|\psi\|_{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd},\mathbb{R}^\ell)}\leq C[\psi]_{\gamma,\mathbf{d}},\\ [\psi]_{\gamma,\mathbf{d}}:=\sup_{\mathbf{x}\neq\mathbf{x}',(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2}\frac{|\psi(\mathbf{x})-\psi(\mathbf{x}')|}{\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')},\end{aligned}\tag{1.15}$$

see also e.g. Lunardi [24].

From (1.11) and (1.14), we write that $\psi\in C_{b,\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd},\mathbb{R}^\ell)$ if

$$\|\psi\|_{C_{b,\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd},\mathbb{R}^\ell)}:=\sum_{i=1}^n\sup_{\mathbf{z}\in\mathbb{R}^{nd}}\|\psi_i(\mathbf{z},\cdot)\|_{C_b^{\frac{k+\gamma}{2i-1}}(\mathbb{R}^d,\mathbb{R}^\ell)}<+\infty.$$

From now on, we will denote:

$$\|\psi\|_{C_{b,\mathbf{d}}^\gamma}:=\|\psi\|_{C_{b,\mathbf{d}}^\gamma(\mathbb{R}^{nd},\mathbb{R}^\ell)}.$$

Finally, through the article, we use the following notation for all $\varphi_1 \in L^\infty\big([0,T], C^{k+\gamma}_{b,\mathbf{d}}(\mathbb{R}^m,\mathbb{R}^\ell)\big)$ and $\varphi_2 \in L^\infty\big([0,T], C^{k+\gamma}_{\mathbf{d}}(\mathbb{R}^m,\mathbb{R}^\ell)\big)$:

$$\|\varphi_1\|_{L^\infty(C^{k+\gamma}_{b,\mathbf{d}})} := \sup_{t\in[0,T]} \|\varphi_1(t,\cdot)\|_{C^{k+\gamma}_{b,\mathbf{d}}(\mathbb{R}^m,\mathbb{R}^\ell)}, \text{ and } \|\varphi_2\|_{L^\infty(C^{k+\gamma}_{\mathbf{d}})} := \sup_{t\in[0,T]} \|\varphi_2(t,\cdot)\|_{C^{k+\gamma}_{\mathbf{d}}(\mathbb{R}^m,\mathbb{R}^\ell)}.$$

1.5. **Assumptions and main result.** With these notations at hand we can now state our assumptions and main results. In the following, we will assume:

(UE) **Uniform Ellipticity of the diffusion Coefficient.** There exists $\kappa \geq 1$ s.t. for all $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^{nd}$, $z \in \mathbb{R}^d$,

$$\kappa^{-1}|z|^2 \leq \langle a(t,\mathbf{x})z, z\rangle \leq \kappa|z|^2,$$

where $|\cdot|$ again denotes the Euclidean norm and $\langle\cdot,\cdot\rangle$ is the inner product.

(H) **Weak Hörmander like condition.** For each $i \in [\![2,n]\!]$, there exists a closed convex subset $\mathcal{E}_{i-1} \subset GL_d(\mathbb{R})$ (set of invertible $d\times d$ matrices) such that, 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, $\mathcal{E}_{i-1}$ may be a closed ball included in $GL_d(\mathbb{R})$, which is an open set.

(S) **Smoothness of the Coefficients.** Fix $\gamma \in (0,1)$. We suppose the following conditions hold.

(i) **Smoothness of the diffusion coefficient.** We assume that a is measurable in time and that $a \in L^\infty\big([0,T], C^\gamma_{b,\mathbf{d}}(\mathbb{R}^{nd}, \mathbb{R}^d\otimes\mathbb{R}^d)\big)$.

(ii) **Smoothness of the drift in time.** We only assume here that the drift is measurable in time and bounded at the origin, i.e. the measurable mapping $t \mapsto \mathbf{F}(t,\mathbf{0})$ is bounded.

(iii) **Smoothness of the drift in space.** We now state, for each level $i \in [\![1,n]\!]$ the smoothness assumptions on the drift component $\mathbf{F}_i$ (see the remark below for more explanations):

$$\mathbf{F}_i \in L^\infty\Big([0,T], C^{(2i-3)\vee 0+\gamma}_{\mathbf{d}}(\mathbb{R}^{((n-i+2)\wedge n)d}, \mathbb{R}^d\Big). \tag{1.16}$$

For a fixed parameter $\gamma \in (0,1)$, we will say that **(A)** is in force as soon as **(UE)**, **(H)**, **(S)** hold.

Remark 1. *Let us come back to assumption* **(S)***-(iii), which may seem difficult to understand at first sight. Namely, we here explain a little bit how the particular thresholds appearing in this assumption come from as well as the precise regularity imposed on each component of the drift* $\mathbf{F}$ *w.r.t. any space variable.*

• *Note first that for* $i = 1$ *assumption* **(S)***-(iii) readily says, with the previous notations for Hölder spaces, that* $\mathbf{F}_1 \in L^\infty([0,T], C^\gamma_{\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R}^d))$.

• *For each level* $i \in [\![2,n]\!]$ *of the chain in* (1.2)*, we shall consider different types of assumptions on* $\mathbf{F}_i$ *depending on the variables* $\mathbf{x}_{i-1}$ *and* $\mathbf{x}^{i:n} = (\mathbf{x}_i,\cdots,\mathbf{x}_n)$ *respectively. Let us now fix* $i \in [\![2,n]\!]$.

The component $\mathbf{x}_{i-1}$ *is hence the one which transmits the noise. Coherently with the usual Hörmander setting, we need some differentiability of* $\mathbf{F}_i$ *w.r.t.* $\mathbf{x}_{i-1}$*. In order to have a global coherence, in terms of time-space homogeneity, for all the considered variables, the specific smoothness to be considered for that variable is that* $\mathbf{F}_i(t,\cdot,\mathbf{x}^{i:n})$ *is in* $C^{1+\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d,\mathbb{R}^d)$*. Recalling now the previous definition of* $\mathbf{d}$ *and of the associated Hölder spaces, we have* $C^{1+\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d,\mathbb{R}^d) = C^{\frac{2i-3+\gamma}{2(i-1)-1}}(\mathbb{R}^d,\mathbb{R}^d)$.

Now, at level i*, the components* $\mathbf{x}^{i:n}$ *are above the current characteristic time-scale, i.e. the vector of their associated time rescaling, which writes according to the homogenous quasi-metric* $\mathbf{d}_P$ *in* (1.8) *as* $(t^{i-\frac12}, t^{(i+1)-\frac12}, \cdots, t^{n-\frac12})$*, has in small time entries that are actually smaller or equal than the time rescaling of the current variable* $\mathbf{x}_i$ *in* $t^{i-\frac12}$*. We recall as well that, in order to have the well posedness of the martingale problem associated with the operator* $(L_t)_{t\geq 0}$ *in* (1.3)*, some natural minimal thresholds of Hölder continuity appear for these variables. Precisely, at level* i*,* $\mathbf{F}_i$ *must be Hölder continuous in* $\mathbf{x}_j$*,* $j \in [\![i,n]\!]$*, with index strictly greater than* $\frac{2i-3}{2j-1}$ *(see* [4] *for details). Here, still to have a global coherence, in terms of time-space homogeneity, for all the considered variables, we assume that* $\mathbf{F}_i$ *is* $\frac{2i-3}{2j-1} + \frac{\gamma}{2j-1}$ *Hölder continuous in its* j^{th} *variable. This precisely corresponds to the minimal threshold required to which we add the intrinsic* γ*-Hölder regularity w.r.t to the associated scale appearing in* $\mathbf{d}$ *for the considered entry. Thus, with a slight abuse of notations,* $z \mapsto \mathbf{F}_i(t,\mathbf{x}_{i-1},\mathbf{x}_i,\cdots,\mathbf{x}_{j-1},z,\mathbf{x}_{j+1},\cdots,\mathbf{x}_n)$ *is supposed to be in* $C^{\frac{2i-3}{2j-1}+\frac{\gamma}{2j-1}}(\mathbb{R}^d,\mathbb{R}^d)$.

• *"Gathering" the regularity conditions assumed on each variable for each component of* $\mathbf{F}$ *hence gives assumption* **(S)***-(iii).*

Remark 2. *Concerning the time regularity in the previous assumptions we can refer to the earlier work of Kruzhkov et al.* [15] *who first consider this type of regularity to establish Schauder estimates in the classical non*

degenerate framework with bounded coefficients. We can also mention Lorenzi [23] *for extensions to unbounded coefficients.*

We are now in position to state our main result.

Theorem 1 (Schauder Estimates for degenerate Kolmogorov Equations with general drifts.)**.** *Let $\gamma \in (0,1)$ be given. Suppose that* **(A)** *is in force and that the terminal condition g and source term f of the Cauchy problem* (1.1) *satisfy: $g \in C^{2+\gamma}_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})$ and $f \in L^\infty\big([0,T], C^\gamma_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})\big)$.*

Then, there exists a unique mild and weak solution u in $C^{2+\gamma}_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})$ to (1.1)*. Furthermore, there exists a constant $C_1 := C_1\big((\mathbf{A}),T\big)$ s.t.*

$$\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \le C_1\big(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\big). \tag{1.17}$$

Note that some Schauder estimates results were recently established for a non-local operator associated with a stable process. Let us mention [5] for the non-degenerate case ($n = 1$) with super-critical drift (stability index strictly less than 1), [11] for a kinetic case ($n = 2$) and [25] for the complete degenerate chain. If the second order term in (1.1) is replaced by an α-stable operator, $\alpha \in (0,2)$, then the parabolic regularity gain associated with the parabolic distance $\mathbf{d}$, established therein, is α instead of 2 in Theorem 1.

Section 2 below is dedicated to the presentation and description of the various steps that we perform to obtain Theorem 1. From now on we will denote by C a generic constant that may change from line to line but only depends on known parameters in **(A)** and the considered fixed final time T, i.e. $C := C((\mathbf{A}),T)$. We reserve the notation c for generic constants that may also change from line to line, depend on **(A)** but are also independent of T, i.e. $c := c((\mathbf{A}))$.

2. Detailed Guide to the proof

The various steps of our procedure could be roughly summed up as follows: we first mollify the coefficients in equation (1.1) in order to work with well defined objects. In the following, we will call, with a slight terminology abuse, by *regularized* or *mollified* solution of (1.1) the solution of (1.1) associated with the *regularized* or *mollified* coefficients.

We then derive the estimate of Theorem 1 in this framework but uniformly in the mollification parameter. To do so, we will expand the regularized solution of (1.1) around a well chosen proxy. This expansion will allow us to obtain an explicit representation of the mollified solution of (1.1) for which we will derive the desired estimates in small time. A key point is that such a representation is of implicit form, so that, when applying our strategy, the upper-bound of the Hölder estimate will involve the Hölder norm of the smoothed solution itself. To overcome this problem, the main idea consists in using a circular argument, bringing together the Hölder norms of the solutions on the same side of the inequality. This strategy then requires to obtain constants in front of the bounds depending of the solution as small as needed. This property will be fulfilled when working with an appropriately rescaled version of the smoothed solution. We can then transfer estimates on the (regularized) rescaled version of the solution to the original (regularized) one and then extend it to arbitrary time length intervals by using a chaining argument. We then conclude the proof of the estimates in Theorem 1 through a compactness argument, allowing us to get rid of the regularization parameters, and eventually show that the mild solution of (1.1) is a weak solution thanks to suitable controls deriving from our analysis.

The main objective of this section is to introduce the approach shortly described above. Especially the derivation of the estimates in Theorem 1 for the regularized solution in small time, involving norms of the solution itself. This part is actually the core of the paper (we refer to Section 6 for the last steps of our procedure: scaling arguments, finite time issues and weak solution property). The main point is thus to emphasize the various difficulties arising when applying our strategy and to introduce the adapted tools that can be used to circumvent them.

2.1. **The mollifying procedure.**

The first step of our strategy is to mollify equation (1.1) in order to get a well-posed Cauchy problem in the classical sense. Precisely, for $\varphi \in C^2_0(\mathbb{R}^{nd},\mathbb{R})$, $m \in \mathbb{N}$ and $t \in [0,T]$ we define the operator:

$$L^m_t\varphi(x) := \langle \mathbf{F}_m(t,\mathbf{x}), \mathbf{D}\varphi(\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1}\varphi(\mathbf{x})a_m(t,\mathbf{x})\big), \tag{2.1}$$

where $\mathbf{F}_m, a_m$ are mollified versions in space of the initial coefficients $\mathbf{F}, a$ in (1.3), i.e. $\mathbf{F}_m(t,\mathbf{x}) = \mathbf{F}(t,\cdot)\star\phi_m(\mathbf{x}), a_m(t,\mathbf{x}) = a(t,\cdot)\star\phi_m(\mathbf{x})$, where for any $\mathbf{z}\in\mathbb{R}^{nd}$, $\phi_m(\mathbf{z}) := m^{nd}\phi(\mathbf{z}m)$ for a smooth, i.e. C^∞, non-negative function $\phi : \mathbb{R}^{nd} \to \mathbb{R}^+$ s.t. $\int_{\mathbb{R}^{nd}}\phi(\mathbf{z})d\mathbf{z} = 1$ and the previous convolutions are to be understood

componentwise. We introduce correspondingly the stochastic differential equation with generator $(L_t^m)_{t\geq 0}$. Namely, for fixed $(t,\mathbf{x})\in[0,T]\times\mathbb{R}^{nd}$ and $s\geq t$,

$$\mathbf{X}_s^{m,t,\mathbf{x}} = \mathbf{x} + \int_t^s \mathbf{F}_m(u,\mathbf{X}_u^{m,t,\mathbf{x}})du + \int_t^s B\sigma_m(u,\mathbf{X}_u^{m,t,\mathbf{x}})dW_u, \tag{2.2}$$

where σ_m is a square root of a_m and $B=(\mathbf{I}_{d,d},\mathbf{0}_{d,d},\dots,\mathbf{0}_{d,d})^*$ is the embedding matrix from $\mathbb{R}^{nd}$ to $\mathbb{R}^d$. The dynamics in (2.2) is similar to the one in (1.4) up to the mollification of the coefficients. It can be deduced from the well-posedness of the martingale problem, which holds under our current assumptions from [4], that $(\mathbf{X}_t^m)_{t\in[0,T]}\Rightarrow_m(\mathbf{X}_t)_{t\in[0,T]}$ (convergence in law on the path space) where $(\mathbf{X}_t)_{t\in[0,T]}$ is the unique weak solution of (1.4) (see also [30]).

Consider now mollified versions f_m, g_m of the source f and the final condition g in (1.1). It is then rather direct to derive through stochastic flows techniques, see e.g. Kunita [19], that

$$u_m(t,\mathbf{x}) := \mathbb{E}[g_m(\mathbf{X}_T^{m,t,\mathbf{x}})] + \int_t^T \mathbb{E}[f_m(s,\mathbf{X}_s^{m,t,\mathbf{x}})]ds, \tag{2.3}$$

belongs for any given m to $C_b^\infty(\mathbb{R}^{nd},\mathbb{R})$ (space of infinitely differentiable functions with bounded derivatives) and precisely solves:

$$\begin{cases}\partial_t u_m(t,\mathbf{x}) + \langle \mathbf{F}_m(t,\mathbf{x}),\mathbf{D}u_m(t,\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1}u_m(t,\mathbf{x})a_m(t,\mathbf{x})\big) = -f_m(t,\mathbf{x}),\ (t,\mathbf{x})\in[0,T)\times\mathbb{R}^{nd},\\ u_m(T,\mathbf{x}) = g_m(\mathbf{x}),\ \mathbf{x}\in\mathbb{R}^{nd}.\end{cases} \tag{2.4}$$

2.2. **Proxy and explicit representation of u_m.** The idea is now to obtain controls of the norms $\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}$ which are uniform w.r.t. the mollifying parameter m. To this end, we will use a perturbative method by expanding u_m around a suitable Ornstein-Uhlenbeck like Gaussian proxy corresponding to an appropriate linearization of the dynamics in (2.2). Consider first the deterministic dynamics deriving from (2.2) obtained setting σ_m to $\mathbf{0}_{d,d}$, i.e.

$$\dot{\boldsymbol{\theta}}^m_{v,\tau}(\boldsymbol{\xi}) = \mathbf{F}_m(v,\boldsymbol{\theta}^m_{v,\tau}(\boldsymbol{\xi})),\ v\in[0,T],\ \boldsymbol{\theta}^m_{\tau,\tau}(\boldsymbol{\xi}) = \boldsymbol{\xi}, \tag{2.5}$$

where $(\tau,\boldsymbol{\xi})\in[0,T]\times\mathbb{R}^{nd}$ are *freezing parameters*, respectively in time and space to be specified.

Fix $0\leq t<s\leq T$ and $\mathbf{x}\in\mathbb{R}^{nd}$. The typical linearization of (2.2) on the time interval $[t,s]$ around $(\boldsymbol{\theta}^m_{v,\tau}(\boldsymbol{\xi}))_{v\in[t,s]}$ writes:

$$\tilde{\mathbf{X}}_v^{m,(\tau,\boldsymbol{\xi})} = \mathbf{x} + \int_t^v [\mathbf{F}_m(r,\boldsymbol{\theta}^m_{r,\tau}(\boldsymbol{\xi})) + D\mathbf{F}_m(r,\boldsymbol{\theta}_{r,\tau}(\boldsymbol{\xi}))(\tilde{\mathbf{X}}_r^{m,(\tau,\boldsymbol{\xi})} - \boldsymbol{\theta}^m_{r,\tau}(\boldsymbol{\xi}))]dr + \int_t^v B\sigma_m(r,\boldsymbol{\theta}^m_{r,\tau}(\boldsymbol{\xi}))dW_r, \tag{2.6}$$

where for any $\mathbf{z}\in\mathbb{R}^{nd}$,

$$D\mathbf{F}_m(v,\mathbf{z}) := \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d}\\ D_{\mathbf{z}_1}\mathbf{F}_{m,2}(v,\mathbf{z}) & \mathbf{0}_{d,d} & \cdots & \cdots & \mathbf{0}_{d,d}\\ \mathbf{0}_{d,d} & D_{\mathbf{z}_2}\mathbf{F}_{m,3}(v,\mathbf{z}_{2:n}) & \mathbf{0}_{d,d} & \mathbf{0}_{d,d} & \vdots\\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \vdots\\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & D_{\mathbf{z}_{n-1}}\mathbf{F}_{m,n}(v,\mathbf{z}_{n-1},\mathbf{z}_n) & \mathbf{0}_{d,d}\end{pmatrix} \tag{2.7}$$

denotes the subdiagonal of the Jacobian matrix $D_{\mathbf{z}}\mathbf{F}_m(v,\cdot)$ at point $\mathbf{z}$. From our previous assumptions (non-degeneracy of σ and Hörmander like condition), the Gaussian process with dynamics (2.6) admits a well controlled multi-scale density $\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\cdot)$ (see e.g. Section 3.1 below and for instance [7], [4]). Namely, there exists $C:=C((\mathbf{A}),T)\geq 1$ s.t. for $j\in\{0,1,2\}$, for all $k\in[\![1,n]\!]^2$, $\ell\in\{0,1\}$, and for all $0\leq t<s\leq T$, $(\mathbf{x},\mathbf{y})\in(\mathbb{R}^{nd})^2$:

$$\begin{aligned}|D^\ell_{\mathbf{x}_k}D^j_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})| &\leq \frac{C}{(s-t)^{\ell(k-\frac{1}{2})+\frac{j}{2}+\frac{n^2d}{2}}}\exp\Big(-C^{-1}(s-t)|\mathbb{T}^{-1}_{s-t}(\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t}(\mathbf{x})-\mathbf{y})|^2\Big)\\ &=: \frac{C}{(s-t)^{\ell(k-\frac{1}{2})+\frac{j}{2}}}\bar{p}^{(\tau,\boldsymbol{\xi})}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y}),\end{aligned} \tag{2.8}$$

where $\mathbf{m}_{s,t}^{m,(\tau,\boldsymbol{\xi})}(\mathbf{x})$ stands for the mean of $\tilde{\mathbf{X}}_s^{m,(\tau,\boldsymbol{\xi})}$ and for any $u>0$, $\mathbb{T}_u$ is the *intrinsic* scale matrix:

$$\mathbb{T}_u = \begin{pmatrix} u\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & u^2\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & u^n\mathbf{I}_{d,d} \end{pmatrix}, \tag{2.9}$$

that is, the i^{th} diagonal entry of $u^{-\frac{1}{2}}\mathbb{T}_u$ reflects the time order of the variances of the $(i-1)^{\text{th}}$ iterated integral of the standard Brownian motion at time u. Observe as well that the time singularities in (2.8) precisely reflect the typical scale of the associated variable, i.e. differentiating in $\mathbf{x}_k$ yields an additional time singularity in $(s-t)^{-k+\frac{1}{2}}$ where $(s-t)^{k-\frac{1}{2}}$ is exactly the order of the standard deviation of the $(k-1)^{\text{th}}$ iterated integral of the Brownian motion.

Denoting by $(\tilde{L}_v^{m,(\tau,\boldsymbol{\xi})})_{v\in[t,T]}$ the generator of (2.6), it also holds that:

$$\begin{aligned} &\big(\partial_s - (\tilde{L}_s^{m,(\tau,\boldsymbol{\xi})})^*\big)\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) = 0,\ \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\cdot) \to_{s\downarrow t} \delta_{\mathbf{x}}(\cdot), \\ &\big(\partial_t + \tilde{L}_t^{m,(\tau,\boldsymbol{\xi})}\big)\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) = 0,\ \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\cdot,\mathbf{y}) \to_{t\uparrow s} \delta_{\mathbf{y}}(\cdot). \end{aligned}$$

The above equations are respectively the *forward* and *backward* Kolmogorov equations. In the first one, the operator $(\tilde{L}_s^{m,(\tau,\boldsymbol{\xi})})^*$ acts on the *forward* variable $\mathbf{y}$ whereas in the second one, $\tilde{L}_t^{m,(\tau,\boldsymbol{\xi})}$ acts in the *backward* variable $\mathbf{x}$. We will use the notation $\tilde{P}_{T,t}^{m,(\tau,\boldsymbol{\xi})}$ for the corresponding semi-group, i.e.

$$\tilde{P}_{T,t}^{m,(\tau,\boldsymbol{\xi})} g_m(\mathbf{x}) := \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,T,\mathbf{x},\mathbf{y}) g_m(\mathbf{y}) d\mathbf{y},$$

as well as

$$\tilde{G}^{m,(\tau,\boldsymbol{\xi})} f_m(t,\mathbf{x}) \ := \ \int_t^T ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) f_m(s,\mathbf{y}) d\mathbf{y}, \tag{2.10}$$

for the associated Green kernel (with fixed final time $T>0$).

For fixed $(t,\mathbf{x})\in[0,T]\times\mathbb{R}^{nd}$ and the above Gaussian proxy, for which $(\tau,\boldsymbol{\xi})$ still remain to be specified, we recall that Duhamel's formula (first order parametrix expansion) yields that:

$$u_m(t,\mathbf{x}) = \tilde{P}_{T,t}^{m,(\tau,\boldsymbol{\xi})} g_m(\mathbf{x}) + \tilde{G}^{m,(\tau,\boldsymbol{\xi})} f_m(t,\mathbf{x}) + \int_t^T ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\boldsymbol{\xi})}) u_m(s,\mathbf{y}) d\mathbf{y}. \tag{2.11}$$

Note for instance that the superscript $(\tau,\boldsymbol{\xi})$, which stands for the freezing parameters, does not appear in the regularized solution u_m. This is because the smoothed solution does not depend on the freezing parameters. Hence, the above representation is valid for **any** choice of $(\tau,\boldsymbol{\xi})$.

2.3. **Estimates of the supremum norm of the second order derivative w.r.t. the non degenerate variables: introduction of the Besov duality argument.** Recall from the statement of our main Theorem 1 that we have to give bounds on $\|u_m\|_{L^\infty(C_{b,\mathbf{d}}^{2+\gamma})}$. For this introduction to the proof, we will focus on the contribution $D_{\mathbf{x}_1}^2 u_m$, that already exhibits almost all the difficulties and for which we want to establish a control in time-space supremum norm and for the γ-Hölder modulus associated with the distance $\mathbf{d}$.

Differentiating in $D_{\mathbf{x}_1}^2$ equation (2.11) gives:

$$\begin{aligned} D_{\mathbf{x}_1}^2 u_m(t,\mathbf{x}) \ &= \ D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{m,(\tau,\boldsymbol{\xi})} g_m(\mathbf{x}) + D_{\mathbf{x}_1}^2 \tilde{G}^{m,(\tau,\boldsymbol{\xi})} f_m(t,\mathbf{x}) \\ &\quad + \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\boldsymbol{\xi})}) u_m(s,\mathbf{y}) d\mathbf{y}. \end{aligned} \tag{2.12}$$

Concentrating on the last term, which turns out to be the most delicate, we see that the choice of $(\tau,\boldsymbol{\xi})$ must be made in order to balance the time singularities coming from $D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})$. Let us first consider the non-degenerate part coming from the difference $(L_s^m - \tilde{L}_s^{m,(\tau,\boldsymbol{\xi})}) u_m(s,\mathbf{y})$ which explicitly writes from (2.1) and (2.6):

$$\begin{aligned} &\langle \mathbf{F}_{m,1}(s,\mathbf{y}) - \mathbf{F}_{m,1}(s,\boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi})), D_{\mathbf{y}_1} u_m(s,\mathbf{y})\rangle + \frac{1}{2}\mathrm{Tr}\Big((a_m(s,\mathbf{y}) - a_m(s,\boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi})))D_{\mathbf{y}_1}^2 u_m(s,\mathbf{y})\Big) \\ =: \ &\Delta_{1,\mathbf{F}_m,\sigma_m}(\tau,s,\mathbf{y},\boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}),u_m), \end{aligned} \tag{2.13}$$

and can be upper-bounded from the Hölder continuity assumption (w.r.t. the underlying homogeneous metric $\mathbf{d}$) on $\mathbf{F}_1$ and a as:

$$\begin{aligned}&|\Delta_{1,\mathbf{F}_m,\sigma_m}(\tau,s,\mathbf{y},\boldsymbol{\theta}^m_{s,\tau}(\boldsymbol{\xi}),u_m)|\\ &\leq \Big([\mathbf{F}_1(s,\cdot)]_{\mathbf{d},\gamma}\|D_{\mathbf{y}_1}u_m(s,\cdot)\|_{L^\infty}+\frac{1}{2}[a(s,\cdot)]_{\mathbf{d},\gamma}\|D^2_{\mathbf{y}_1}u_m(s,\cdot)\|_{L^\infty}\Big)\mathbf{d}^\gamma(\mathbf{y},\boldsymbol{\theta}^m_{s,\tau}(\boldsymbol{\xi})).\end{aligned}\tag{2.14}$$

The contribution $\mathbf{d}^\gamma(\mathbf{y},\boldsymbol{\theta}^m_{s,\tau}(\boldsymbol{\xi}))$ in the above r.h.s. must then equilibrate the time singularity in $(s-t)^{-1}$ coming from $D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})$ (see (2.8) and Proposition 3 below). This is possible if $\mathbf{d}^\gamma(\mathbf{y},\boldsymbol{\theta}^m_{s,\tau}(\boldsymbol{\xi}))$ is *compatible* with the off-diagonal bound $(s-t)|\mathbb{T}^{-1}_{s-t}(\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t}(\mathbf{x})-\mathbf{y})|^2$ in (2.8). This is precisely the case considering $(\tau,\boldsymbol{\xi})=(t,\mathbf{x})$ which gives

$$\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t}(\mathbf{x})|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})}=\boldsymbol{\theta}^m_{s,t}(\mathbf{x}),\tag{2.15}$$

as it can readily be checked from (2.5), (2.6) (taking the expectation) and the Grönwall's lemma. Therefore, observing precisely from the metric homogeneity (see equations (1.8) and (1.9)) that:

$$\begin{aligned}\mathbf{d}^\gamma(\mathbf{y},\boldsymbol{\theta}^m_{s,t}(\mathbf{x}))&=(s-t)^{\frac{\gamma}{2}}\mathbf{d}^\gamma((s-t)^{\frac12}\mathbb{T}^{-1}_{s-t}\mathbf{y},(s-t)^{\frac12}\mathbb{T}^{-1}_{s-t}\boldsymbol{\theta}^m_{s,t}(\mathbf{x}))\\ &\leq C(s-t)^{\frac{\gamma}{2}}\Big(\sum_{i=1}^n|(s-t)^{-\frac{2i-1}{2}}(\mathbf{y}-\boldsymbol{\theta}^m_{s,t}(\mathbf{x}))_i|^{\frac{\gamma}{2i-1}}\Big),\end{aligned}$$

we get that the terms of last contribution in the above r.h.s. can precisely be absorbed by the exponential off-diagonal bound in (2.8).

We therefore eventually derive for the non-degenerate contribution with the notation of (2.8):

$$\begin{aligned}&\Big|\int_t^T ds\int_{\mathbb{R}^d}D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\Delta_{1,\mathbf{F}_m,\sigma_m}(\tau,s,\boldsymbol{\theta}^m_{s,t}(\boldsymbol{\xi}),\mathbf{y},u_m)d\mathbf{y}\Big|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})}\\ \leq&\int_t^T\frac{ds}{(s-t)^{1-\frac{\gamma}{2}}}\int_{\mathbb{R}^{nd}}C\big(\|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}+\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})}\big)\bar{p}^{(\tau,\boldsymbol{\xi})}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})\\ &\times\Big(\|D_{\mathbf{y}_1}u_m(s,\cdot)\|_{L^\infty}+\|D^2_{\mathbf{y}_1}u_m(s,\cdot)\|_{L^\infty}\Big)d\mathbf{y}\\ \leq&\frac{2\Lambda}{\gamma}(T-t)^{\frac{\gamma}{2}}\big(\|D_{\mathbf{y}_1}u_m\|_{L^\infty}+\|D^2_{\mathbf{y}_1}u_m\|_{L^\infty}\big)\\ \leq&\frac{2\Lambda}{\gamma}(T-t)^{\frac{\gamma}{2}}\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}.\end{aligned}\tag{2.16}$$

Remark 3 (Constants depending on the Hölder moduli of the coefficients)**.** *In equation* (2.16)*, we denoted* $\Lambda:=C\big(\|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}+\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})}\big)$*, i.e.* Λ *explicitly depends on the Hölder moduli of the coefficients* a *and* $\mathbf{F}_1$ *on the time interval* $[0,T]$*. Importantly, in the following, we will keep the generic notation* Λ *for any constant depending on the Hölder moduli of* $a,\mathbf{F}$*, but* ***not*** *on the supremum norms of* a *and* $(D\mathbf{F}_i)_{i\in[\![2,n]\!]}$ *and such that* $\Lambda\to 0$ *when the Hölder moduli of* a*,* $\mathbf{F}_1$ *and for any* $i\in[\![2,n]\!]$*,* $\mathbf{F}_i$ *w.r.t. to the variables* i *to* n*, themselves tend to 0. In other words,* Λ *is meant to tend to 0 when the coefficients do not vary much. In the computations below* Λ *may change from line to line but will always enjoy the previous property.*

Equation (2.16) thus precisely yields a time smoothing effect corresponding exactly to the Hölder continuity exponent γ of the coefficients. The previous choice of $(\tau,\boldsymbol{\xi})$ is known as the *forward* parametrix and seems adapted as soon as one is led to estimate derivatives of the solution.

Let us mention that, as far as one is concerned with density estimates, which formally amounts to replace $u_m(s,\mathbf{x})$, $u_m(s,\mathbf{y})$ in (2.11) with $p^m(t,T,\mathbf{x},\mathbf{z})$, $p^m(s,T,\mathbf{y},\mathbf{z})$ (density at some fixed point $\mathbf{z}\in\mathbb{R}^{nd}$ of $\mathbf{X}^m_T$ starting from $\mathbf{x}$ at time t), or with the well-posedness of the martingale problem, another choice, consisting in freezing in $(\tau,\boldsymbol{\xi})=(s,\mathbf{y})$ in the above equation, could also be considered. Note that the freezing parameters would here depend on the time and spatial integration variables. This *backward* approach was first introduced by Il'in *et al* [13] (see also Friedman [10] or McKean and Singer [26]) and led successfully to density estimates and well-posedness of the martingale problem for the current model (1.4) in the respective works [7], [28], [4].

However, when dealing with derivatives, the forward perturbative approach appears more flexible since it allows to exploit cancellation techniques whereas this is much trickier in the backward case for which $\tilde{p}^{m,(s,\mathbf{y})}(t,s,\mathbf{x},\mathbf{y})$ is not a density w.r.t. $\mathbf{y}$. Some associated errors associated with this approach are thoroughly discussed in [4].

Let us emphasize that, in view of the above calculations, from now on, the choice of the freezing parameters $(\tau,\boldsymbol{\xi})$ will *by default* be $(\tau,\boldsymbol{\xi})=(t,\mathbf{x})$. We will hence sometimes forget the superscript on quantities that depend on these parameters for the sake of clarity and assume implicitly this choice. It may happen in the

following that another choice for the freezing space point $\boldsymbol{\xi}$ will be done. If so, we will specify it. In any case, when the regularized solution u_m will be evaluated at point t in $[0,T]$, we will choose $\tau = t$ so that this choice shall be assumed in the following.

Let us now turn to the contributions associated with the degenerate variables in the difference $(L_s^m - \tilde{L}_s^{m,(\tau,\boldsymbol{\xi})})u_m(s,\mathbf{y})$. They precisely write:

$$\begin{aligned}&\sum_{i=2}^n \Big\langle \Big(\mathbf{F}_{m,i}(s,\mathbf{y}) - \mathbf{F}_{m,i}(s,\boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}}\mathbf{F}_{m,i}(s,\boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}))(\mathbf{y}-\boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}))_{i-1}\Big), D_{\mathbf{y}_i}u_m(s,\mathbf{y})\Big\rangle \\ =:\ &\sum_{i=2}^n \big\langle \Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y}), D_{\mathbf{y}_i}u_m(s,\mathbf{y})\big\rangle. \end{aligned} \tag{2.17}$$

Under the current assumptions, we do not expect to have uniform controls w.r.t. to the smoothing parameter m for the derivatives $(D_{\mathbf{y}_i}u_m)_{i\in[\![2,n]\!]}$ in the degenerate directions. Our strategy will first consist for those terms in performing an integration by parts leading to:

$$\begin{aligned}&\Big|\sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\big\langle \Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y}), D_{\mathbf{y}_i}u_m(s,\mathbf{y})\big\rangle\Big|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\ \leq\ &\sum_{i=2}^n \int_t^T ds \Big| \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{y}_i}\cdot \Big(\big(D_{\mathbf{x}_1}^2\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\otimes \Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y})\big)\Big) u_m(s,\mathbf{y})\Big|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})}, \end{aligned} \tag{2.18}$$

where the notation "$\otimes$" stands for the usual tensor product. In particular, the term $\big(D_{\mathbf{x}_1}^2\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\otimes \Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y})\big)$ is a tensor lying in $(\mathbb{R}^d)^{\otimes 3}$. Furthermore, $D_{\mathbf{y}_i}\cdot$ refers to an extended form of the divergence over the i^{th} variable ($\mathbf{y}_i\in\mathbb{R}^d$). Precisely, from (2.18), we rewrite for all $i\in[\![2,n]\!]$, $(s,\mathbf{y})\in[t,T]\times\mathbb{R}^{nd}$:

$$\begin{aligned}&D_{\mathbf{y}_i}\cdot\Big(\big(D_{\mathbf{x}_1}^2\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\otimes \Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y})\big)\Big) \\ =\ &\sum_{j=1}^d \partial_{\mathbf{y}_i^j}\Big(D_{\mathbf{x}_1}^2\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\big(\Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y})\big)_j\Big), \end{aligned}$$

with $\mathbf{y}_i = (\mathbf{y}_i^1,\cdots,\mathbf{y}_i^n)$. In other words, this "enhanced" divergence form decreases by one the order of the input tensor. As a particular case, if $d=1$, $\Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y})$ is a scalar and the divergence form corresponds to the standard differentiation, i.e. $D_{\mathbf{y}_i}\cdot = \partial_{\mathbf{y}_i}$.

Introduce now for notational convenience, for a multi-index $\vartheta\in\mathbb{N}^n$ the quantity:

$$\Theta_{i,(t,\mathbf{x})}^{m,\vartheta}(s,\mathbf{y}) := D_{\mathbf{x}}^{\vartheta}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\otimes\Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y}).$$

With this notation at hand equation (2.18) rewrites:

$$\begin{aligned}&\Big|\sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\big\langle \Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}),\mathbf{y}), D_{\mathbf{y}_i}u_m(s,\mathbf{y})\big\rangle\Big|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\ =\ &\sum_{i=2}^n \int_t^T ds \Big| \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{y}_i}\cdot\big(\Theta_{i,(t,\mathbf{x})}^{m,\vartheta}(s,\mathbf{y})\big)u_m(s,\mathbf{y})\Big|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})}, \end{aligned} \tag{2.19}$$

with $\vartheta = (2,0,\cdots,0)$, i.e. the multi-index here involves the second order derivatives of the frozen heat-kernel w.r.t. to its non-degenerate components.

In view of our main estimates in Theorem 1, we will use the duality between suitable Besov spaces to derive bounds for the spatial integrals in (2.19). Introduce, for each fixed $i\in[\![2,n]\!]$ and any spatial point $(\mathbf{y}_1,\cdots,\mathbf{y}_{i-1},\mathbf{y}_{i+1},\cdots,\mathbf{y}_n) =: (\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})\in\mathbb{R}^{(n-1)d}$ the mappings

$$\begin{aligned} u_m^{i,(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}:\mathbf{y}_i &\mapsto u_m(s,\mathbf{y}_{1:i-1},\mathbf{y}_i,\mathbf{y}_{i+1:n}), \\ \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^{m,\vartheta}:\mathbf{y}_i &\mapsto D_{\mathbf{y}_i}\cdot\big(\Theta_{i,(t,\mathbf{x})}^{m,\vartheta}(s,\mathbf{y})\big). \end{aligned} \tag{2.20}$$

The underlying idea is that we actually want, for the i^{th} variable, to control uniformly in m the Hölder modulus $[u_m^{i,(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}]_{\frac{2+\gamma}{2i-1}}$ uniformly in $(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})\in[t,T]\times\mathbb{R}^{(n-1)d}$. To prove this property we recall that, setting $\tilde{\alpha}_i := \frac{2+\gamma}{2i-1}$, $C_b^{\tilde{\alpha}_i}(\mathbb{R}^d,\mathbb{R}) = B_{\infty,\infty}^{\tilde{\alpha}_i}(\mathbb{R}^d,\mathbb{R})$ with the usual notations for Besov spaces (see e.g. Triebel [31]).

Let us now recall some definitions/characterizations from Section 2.6.4 of Triebel [31]. For $\tilde{\alpha} \in \mathbb{R}, q \in (0,+\infty], p \in (0,\infty]$, $B^{\tilde{\alpha}}_{p,q}(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{\mathcal{H}^{\tilde{\alpha}}_{p,q}} < +\infty\}$ where $\mathcal{S}(\mathbb{R}^d)$ stands for the Schwarz class and

$$\|f\|_{\mathcal{H}^{\tilde{\alpha}}_{p,q}} := \|\varphi(D)f\|_{L^p(\mathbb{R}^d)} + \Big(\int_0^1 \frac{dv}{v} v^{(m-\frac{\tilde{\alpha}}{2})q} \|\partial_v^m h_v \star f\|^q_{L^p(\mathbb{R}^d)}\Big)^{\frac{1}{q}}, \tag{2.21}$$

with $\varphi \in C^\infty_0(\mathbb{R}^d)$ (smooth function with compact support) is s.t. $\varphi(0) \neq 0$, $\varphi(D)f := (\varphi \hat{f})^\vee$ where $\hat{f}$ and $(\varphi \hat{f})^\vee$ respectively denote the Fourier transform of f and the inverse Fourier transform of $\varphi \hat{f}$. When $\tilde{\alpha} > d(1/p-1) \vee 0$ then $\|\varphi(D)f\|_{L^p(\mathbb{R}^d)}$ in (2.21) can be replaced by $\|f\|_{L^p(\mathbb{R}^d)}$. The parameter m is an integer s.t. $m > \frac{\tilde{\alpha}}{2}$ and for $v > 0$, $z \in \mathbb{R}^d$, $h_v(z) := \frac{1}{(2\pi v)^{\frac{d}{2}}} \exp\big(-\frac{|z|^2}{2v}\big)$ is the usual heat kernel of $\mathbb{R}^d$. We point out that the quantities in (2.21) are well defined for $q < \infty$. The modifications for $q = +\infty$ are obvious and can be written passing to the limit.

Observe that the quantity $\|f\|_{\mathcal{H}^{\tilde{\alpha}}_{p,q}}$, where the subscript $\mathcal{H}$ stands to indicate the dependence on the heat-kernel, depends on the considered function φ and the chosen $m \in \mathbb{N}$. It also defines a quasi-norm on $B^s_{p,q}(\mathbb{R}^d)$. The previous definition of $B^{\tilde{\alpha}}_{p,q}(\mathbb{R}^d)$ is known as the thermic characterization of Besov spaces and is particularly well adapted to our current framework. By abuse of notation we will write as soon as this quantity is finite $\|f\|_{\mathcal{H}^{\tilde{\alpha}}_{p,q}} =: \|f\|_{B^{\tilde{\alpha}}_{p,q}}$.

As indicated above, it is easily seen from (2.21) that $C^{\tilde{\alpha}_i}_b(\mathbb{R}^d,\mathbb{R}) = B^{\tilde{\alpha}_i}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R})$. It is also well known that $B^{\tilde{\alpha}_i}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R})$ and $B^{-\tilde{\alpha}_i}_{1,1}(\mathbb{R}^d,\mathbb{R})$ are in duality (see e.g. Proposition 3.6 in [20]). Namely $B^{\tilde{\alpha}_i}_{\infty,\infty}$ is the dual of the closure of the Schwartz class $\mathcal{S}$ in $B^{-\tilde{\alpha}_i}_{1,1}$. But $\mathcal{S}$ is dense in $B^{-\tilde{\alpha}_i}_{1,1}$ (see for instance Theorem 4.1.3 in [1]). We will therefore write from (2.18) and with the notations of (2.20)

$$\begin{aligned}
&\sum_{i=2}^n \int_t^T ds \Big| \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{y}_i} \cdot \big(\Theta^{m,\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y})\big) u_m(s,\mathbf{y})\Big|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\
\leq\ & \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \|\Psi^{m,\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{B^{-\tilde{\alpha}_i}_{1,1}} \|u_m^{i,(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{B^{\tilde{\alpha}_i}_{\infty,\infty}} \\
\leq\ & C\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \|\Psi^{m,\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{B^{-\tilde{\alpha}_i}_{1,1}}.
\end{aligned} \tag{2.22}$$

Exploiting the thermic characterization of Besov spaces (see again (2.21) and Section 4.2), it will be shown in Lemma 11 below that there exists a constant $\Lambda := \Lambda((\mathbf{A}),T)$ as in Remark 3 s.t. for all $i \in [\![2,n]\!]$ and $m \in \mathbb{N}$:

$$\int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \|\Psi^{m,\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{B^{-\tilde{\alpha}_i}_{1,1}} \leq \frac{\Lambda}{(s-t)^{1-\frac{\gamma}{2}}}. \tag{2.23}$$

Therefore:

$$\sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \|\Psi^{m,\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{B^{-\tilde{\alpha}_i}_{1,1}} \leq \Lambda (T-t)^{\frac{\gamma}{2}}, \tag{2.24}$$

which plugged into (2.22) eventually gives the following global smoothing effect for the degenerate contributions. Namely,

$$\begin{aligned}
&\Big|\sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D^2_{\mathbf{x}_1} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) \big\langle \Delta_{i,\mathbf{F}_m}(\tau,s,\boldsymbol{\theta}^m_{s,t}(\boldsymbol{\xi}),\mathbf{y}) D_{\mathbf{y}_i} u_m(s,\mathbf{y})\big\rangle\Big|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\
\leq\ & \Lambda (T-t)^{\frac{\gamma}{2}} \|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})},
\end{aligned} \tag{2.25}$$

which is precisely homogeneous to the bound obtained for the non-degenerate variables in (2.16). In both cases, the contribution $(T-t)^{\frac{\gamma}{2}}$ derives from the assumed smoothness of the coefficients $a, \mathbf{F}$ w.r.t. $\mathbf{d}$ which exactly leads to the same global control for the *a priori* most singular part of expansion (2.12).

From the previous bounds and (2.12) we thus obtain:

$$|D^2_{\mathbf{x}_1} u_m(t,\mathbf{x})| \leq \Big(|D^2_{\mathbf{x}_1} \tilde{P}^{m,(\tau,\boldsymbol{\xi})}_{T,t} g_m(\mathbf{x})| + |D^2_{\mathbf{x}_1} \tilde{G}^{m,(\tau,\boldsymbol{\xi})} f_m(t,\mathbf{x})|\Big)\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} + \Lambda (T-t)^{\frac{\gamma}{2}} \|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}. \tag{2.26}$$

Since $\tilde{P}^{m,(\tau,\xi)}_{T,t}$ is a *true* semi-group, and $\tilde{G}^{m,(\tau,\boldsymbol{\xi})}$ the associated Green kernel (precisely because we used a *forward* perturbative expansion), it will be derived in Lemma 12 (thanks to cancellation techniques) that there

exists $C := C((\mathbf{A}))$ s.t. for any $(t,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$:

$$\begin{aligned}\Big(|D^2_{\mathbf{x}_1}\tilde{P}^{m,(\tau,\boldsymbol{\xi})}_{T,t}g_m(\mathbf{x})| + |D^2_{\mathbf{x}_1}\tilde{G}^{m,(\tau,\boldsymbol{\xi})}f_m(t,\mathbf{x})|\Big)\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} &\leq C(\|g_m\|_{C^{2+\gamma}_{b,\mathbf{d}}} + (T-t)^{\frac{\gamma}{2}}\|f_m\|_{L^\infty(C^\gamma_{b,\mathbf{d}})})\\ &\leq C(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + (T-t)^{\frac{\gamma}{2}}\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}).\end{aligned} \tag{2.27}$$

Equation (2.27) eventually leads to the following estimate on $|D^2_{\mathbf{x}_1}u_m(t,\mathbf{x})|$:

$$|D^2_{\mathbf{x}_1}u_m(t,\mathbf{x})| \leq C\Big(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + (T-t)^{\frac{\gamma}{2}}(\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\Big) + \Lambda(T-t)^{\frac{\gamma}{2}}\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}. \tag{2.28}$$

For T small enough, this equation would be compatible with the estimates of Theorem 1. Equation (2.28) might even seem *too strong* since it also exhibits, additionally to the control of the term associated with the perturbation, a small contribution (in $(T-t)^{\frac{\gamma}{2}}$ for a small enough T) w.r.t. to the source f_m. This is precisely because $\|D^2_{\mathbf{x}_1}u_m(t,\cdot)\|_{L^\infty}$ is not one of the *critical* terms in the Hölder norm $\|u_m(t,\cdot)\|_{C^{2+\gamma}_{b,\mathbf{d}}}$, i.e. the regularity of the coefficients still gives that it can be viewed as a remainder at first sight.

2.4. **Estimates on the Hölder modulus of the second order derivative w.r.t. the non degenerate variables: introduction of the various change of regime of the system.** Now, a typical critical term of the Hölder norm, for which we precisely exploit totally the spatial regularity of the coefficients, is $[D^2_{\mathbf{x}_1}u_m(t,\cdot)]^\gamma_{\mathbf{d}}$ (see assumption **(S)** and (1.15)). Let us now detail how we can handle it and in which sense it can be viewed as critical.

Of course, if $g=0$, for $t\in[0,T]$ and given spatial points $(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2$ we can assume w.l.o.g. that, for some constant c_0 to be specified later on and meant to be *small*, $c_0^{\frac{1}{2}}\mathbf{d}(\mathbf{x},\mathbf{x}') \leq (T-t)^{\frac{1}{2}}$, i.e. the spatial points are *close* w.r.t. the characteristic time scale $(T-t)^{\frac{1}{2}}$ for the homogeneous metric $\mathbf{d}$. Indeed, if $c_0^{\frac{1}{2}}\mathbf{d}(\mathbf{x},\mathbf{x}') > (T-t)^{\frac{1}{2}}$, equation (2.28) readily gives:

$$\begin{aligned}|D^2_{\mathbf{x}_1}u_m(t,\mathbf{x}) - D^2_{\mathbf{x}_1}u_m(t,\mathbf{x}')| &\leq |D^2_{\mathbf{x}_1}u_m(t,\mathbf{x})| + |D^2_{\mathbf{x}_1}u_m(t,\mathbf{x}')|\\ &\leq 2(T-t)^{\frac{\gamma}{2}}(C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} + \Lambda\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})})\\ &\leq 2c_0^{\frac{\gamma}{2}}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')(C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} + \Lambda\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}).\end{aligned} \tag{2.29}$$

Note that the above bound is now critical in the sense discussed above. Let us now focus, as before, on the Hölder control associated with the perturbative contribution in (2.12) when $c_0^{\frac{1}{2}}\mathbf{d}(\mathbf{x},\mathbf{x}') \leq (T-t)^{\frac{1}{2}}$. Namely,

$$\begin{aligned}D^2_{\mathbf{x}_1}\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_m(t,T,\mathbf{x},\mathbf{x}') &:= \int_t^T ds\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(L^m_s - \tilde{L}^{m,(\tau,\boldsymbol{\xi})}_s)u_m(s,\mathbf{y})d\mathbf{y}\\ &\quad - \int_t^T ds\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s - \tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_s)u_m(s,\mathbf{y})d\mathbf{y},\end{aligned} \tag{2.30}$$

where we recall that *a priori* the spatial freezing points $(\boldsymbol{\xi},\boldsymbol{\xi}')$ in (2.30) (see also (2.12)) should be different for $\mathbf{x}$ and $\mathbf{x}'$ and depend on the position of $\mathbf{d}(\mathbf{x},\mathbf{x}')$ w.r.t. the current characteristic time scale in the time integral. Following the terminology of heat kernels, we will say that at time $s\in[t,T]$ the points $\mathbf{x},\mathbf{x}'$ are in the *diagonal regime* if $c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}') \leq s-t$, i.e. their homogeneous distance is small w.r.t. the characteristic time for a parameter c_0 to be specified later on.

We insist again that we have the usual equivalence between time and space, i.e. time has to be compared with the square of the spatial metric $\mathbf{d}$. Similarly, we will say that the *off-diagonal* regime holds when $c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}') > (s-t)$. Observing that in the diagonal case $s \geq t + c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$ (and in the off-diagonal one $s < t + c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$) we split the time integral in (2.30) as:

$$\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_m(t,T,\mathbf{x},\mathbf{x}') := \Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}') + \Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}'),$$

with

$$\begin{aligned}\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{off-diag}}(t,\mathbf{x},\mathbf{x}') &:= \int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds\int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi})}_s)u_m(s,\mathbf{y})d\mathbf{y}\\ &\quad -\int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds\int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_s)u_m(s,\mathbf{y})d\mathbf{y},\\ \Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{diag}}(t,T,\mathbf{x},\mathbf{x}') &:= \int_{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi})}_s)u_m(s,\mathbf{y})d\mathbf{y}\\ &\quad -\int_{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_s)u_m(s,\mathbf{y})d\mathbf{y}.\end{aligned} \tag{2.31}$$

Intuitively, for the term $D^2_{\mathbf{x}_1}\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{off-diag}}(t,\mathbf{x},\mathbf{x}')$, since $\mathbf{x},\mathbf{x}'$ are *far* at the characteristic time scale $(s-t)^{\frac12}$, there is no expectable gain in expanding $D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})-D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})$. One therefore writes:

$$\begin{aligned}&|D^2_{\mathbf{x}_1}\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{off-diag}}(t,\mathbf{x},\mathbf{x}')|\\ \le\ & |\int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi})}_s)u_m(s,\mathbf{y})d\mathbf{y}|\\ &+|\int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_s)u_m(s,\mathbf{y})d\mathbf{y}|.\end{aligned}$$

Now, provided $\boldsymbol{\xi}=\mathbf{x}$, $\boldsymbol{\xi}'=\mathbf{x}'$ one derives from the previous equation, similarly to (2.22), (2.23), that

$$|D^2_{\mathbf{x}_1}\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{off-diag}}(t,\mathbf{x},\mathbf{x}')|\le\Lambda\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}\frac{ds}{(s-t)^{1-\frac{\gamma}{2}}}\le\Lambda\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}c_0^{\frac{\gamma}{2}}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'). \tag{2.32}$$

For c_0 small enough, we obtain again an estimate that would be compatible with the global bound on $\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}$ stated in Theorem 1.

Turning now to $D^2_{\mathbf{x}_1}\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{diag}}(t,T,\mathbf{x},\mathbf{x}')$ one would therefore be tempted to carry on the analysis with the previous freezing points $\boldsymbol{\xi}=\mathbf{x}$, $\boldsymbol{\xi}'=\mathbf{x}'$. Intuitively, in the diagonal regime this should not have too much impact. This is only partly true, since if we proceed so we will be led to investigate the difference of operators at different freezing spatial points and this leads to compare quantities like $\boldsymbol{\theta}^m_{s,t}(\mathbf{x})-\boldsymbol{\theta}^m_{s,t}(\mathbf{x}')$ for which we want a uniform control w.r.t. m. Since the initial (unmollified) coefficients $a,\mathbf{F}$ are only Hölder continuous in space, this quantity is typically controlled (see Lemma 7) as:

$$\mathbf{d}\big(\boldsymbol{\theta}^m_{s,t}(\mathbf{x}),\boldsymbol{\theta}^m_{s,t}(\mathbf{x}')\big)\le C\big(\mathbf{d}(\mathbf{x},\mathbf{x}')+(s-t)^{\frac12}\big), \tag{2.33}$$

where the time contribution is precisely due to the quasi-distance $\mathbf{d}$ (see the proof of Lemma 7 in Appendix A.1.1).

Unfortunately, this approach would lead to a final control of order $\big(\mathbf{d}(\mathbf{x},\mathbf{x}')+(s-t)^{\frac12}\big)^\gamma\le C\big(\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')+(s-t)^{\frac{\gamma}{2}}\big)$ which is not enough on the considered integration set. Recall indeed that, in the diagonal regime $c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')\le(s-t)$ and the term $(s-t)^{\frac{\gamma}{2}}$ in the previous r.h.s. is *too big*. This means that for $\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{diag}}(t,T,\mathbf{x},\mathbf{x}')$, it would be more appropriate to consider the same spatial freezing point. In that case, taking $\boldsymbol{\xi}=\boldsymbol{\xi}'=\mathbf{x}$ and expanding the difference of the derivatives of the frozen Gaussian densities yields:

$$\begin{aligned}&\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\textbf{diag}}(t,T,\mathbf{x},\mathbf{x}')\\ =\ & -\sum_{j=1}^n\int_{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu D_{\mathbf{x}_j}D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}'-\mathbf{x})_j\\ &(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi})}_s)u_m(s,\mathbf{y})\\ =\ & -\sum_{j=1}^n\int_{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu D_{\mathbf{x}_j}D^2_{\mathbf{x}_1}\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}'-\mathbf{x})_j\\ &\left(\Delta_{1,\mathbf{F}_m,\sigma_m}(t,s,\boldsymbol{\theta}^m_{s,t}(\mathbf{x}),\mathbf{y},u_m)+\sum_{i=2}^n\big\langle\Delta_{i,\mathbf{F}_m,\sigma_m}(t,s,\boldsymbol{\theta}^m_{s,t}(\mathbf{x}),\mathbf{y})D_{\mathbf{y}_i}u_m(s,\mathbf{y})\big\rangle\right),\end{aligned} \tag{2.34}$$

using the notations introduced in (2.13) and (2.17) for the last equality. In the previous identities (2.34) and from now on, the symbol "·" between two tensors means the usual tensor contraction. In particular $D_{\mathbf{x}_j} D^2_{\mathbf{x}_1} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}'-\mathbf{x})_j$ is a $d\times d$ matrix.

In the current diagonal regime, it can be shown from (2.8) and the homogeneity of the distance $\mathbf{d}$ that there is $C>1$ such that for $(\tau,\boldsymbol{\xi})=(t,\mathbf{x})$:

$$
\begin{aligned}
& |D_{\mathbf{x}_j} D^2_{\mathbf{x}_1} \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})|\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\
\leq\ & \frac{C}{(s-t)^{j-\frac{1}{2}+1+\frac{n^2d}{2}}} \exp(-C^{-1}(s-t)|\mathbb{T}^{-1}_{s-t}(\mathbf{m}^{m,(\tau,\mathbf{x})}_{s,t}(\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}))-\mathbf{y}|^2)\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\
\leq\ & \frac{C}{(s-t)^{j-\frac{1}{2}+1+\frac{n^2d}{2}}} \exp(C^{-1}(s-t)|\mathbb{T}^{-1}_{s-t}(\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t}(\mathbf{x}-\mathbf{x}'))|^2) \\
& \times \exp(-\frac{C^{-1}}{2}(s-t)|\mathbb{T}^{-1}_{s-t}(\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t}(\mathbf{x}))-\mathbf{y}|^2)\Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\
\leq\ & \frac{C}{(s-t)^{j-\frac{1}{2}+1+\frac{n^2d}{2}}} \exp(-\frac{C^{-1}}{2}(s-t)|\mathbb{T}^{-1}_{s-t}(\boldsymbol{\theta}^m_{s,t}(\mathbf{x}))-\mathbf{y}|^2), \qquad (2.35)
\end{aligned}
$$

using for the last inequality that $\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t}(\mathbf{x})|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} = \boldsymbol{\theta}^m_{s,t}(\mathbf{x})$ and the fact that, from the linear structure of ODE satisfied by $\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t}(\mathbf{x})$ (which can be read from system (2.6) taking the expectation), $(s-t)^{\frac{1}{2}}|\mathbb{T}^{-1}_{s-t}(\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t})(\mathbf{x}-\mathbf{x}')|^2 \leq C(s-t)^{\frac{1}{2}}|\mathbb{T}^{-1}_{s-t}(\mathbf{x}-\mathbf{x}')|$. Since $c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')\leq s-t \Leftrightarrow c_0\mathbf{d}^2\big((s-t)^{\frac{1}{2}}\mathbb{T}^{-1}_{s-t}\mathbf{x},(s-t)^{\frac{1}{2}}\mathbb{T}^{-1}_{s-t}\mathbf{x}'\big)\leq 1$, we readily derive from the definition of $\mathbf{d}$ in (1.9) that $(s-t)|\mathbb{T}^{-1}_{s-t}(\mathbf{m}^{m,(\tau,\boldsymbol{\xi})}_{s,t})(\mathbf{x}-\mathbf{x}')|^2\leq C$. These points are thoroughly discussed in Sections 3.

From (2.34), (2.35) reproducing the previously described analysis, we finally derive:

$$
\begin{aligned}
& D^2_{\mathbf{x}_1}\Delta^{\tau,\boldsymbol{\xi},\boldsymbol{\xi}'}_{m,\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}') \\
\leq\ & \Lambda\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{j=1}^n\int_{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T \frac{ds}{(s-t)^{1+(j-\frac{1}{2})-\frac{\gamma}{2}}}|(\mathbf{x}-\mathbf{x}')_j| \\
\leq\ & \Lambda\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{j=1}^n \frac{|(\mathbf{x}-\mathbf{x}')_j|}{(c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}'))^{j-\frac{1}{2}-\frac{\gamma}{2}}} \\
\leq\ & \frac{\Lambda}{c_0^{n-\frac{1}{2}-\frac{\gamma}{2}}}\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'), \qquad (2.36)
\end{aligned}
$$

using again the definition of $\mathbf{d}$ in (1.9) for the last inequality and where Λ is an in Remark 3. We have again globally gained, thanks to the smoothness of the coefficients, a power $\frac{\gamma}{2}$ in the time singularities of equation (2.35).

From the previous discussion we now have to specify how to modify the freezing parameter depending on the position of the current time variable w.r.t. to the homogeneous spatial distance between the considered points. This can actually been done from the Duhamel formulation up to an additional discontinuity term. Restarting from (2.11) we can indeed rewrite for given $(t,\mathbf{x}')\in[0,T]\times\mathbb{R}^{nd}$ and all $r\in(t,T]$, $\boldsymbol{\xi}'\in\mathbb{R}^{nd}$:

$$
\begin{aligned}
u_m(t,\mathbf{x}') &= \tilde{P}^{m,(\tau,\boldsymbol{\xi}')}_{r,t}u_m(r,\mathbf{x}') + \tilde{G}^{m,(\tau,\boldsymbol{\xi}')}_{r,t}f_m(t,\mathbf{x}') \\
&\quad + \int_t^r ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s-\tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_s)u_m(s,\mathbf{y}), \qquad (2.37) \\
\forall 0\leq v<r\leq T,\ \tilde{G}^{m,(\tau,\boldsymbol{\xi}')}_{r,v}f_m(t,\mathbf{x}) &= \int_v^r ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})f_m(s,\mathbf{y}).
\end{aligned}
$$

Differentiating the above expression in $r\in(t,T]$ yields for any $\boldsymbol{\xi}'\in\mathbb{R}^{nd}$:

$$
\begin{aligned}
0 &= \partial_r\tilde{P}^{m,(\tau,\boldsymbol{\xi}')}_{r,t}u_m(r,\mathbf{x}') + \int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,r,\mathbf{x}',\mathbf{y})f_m(r,\mathbf{y}) \\
&\quad + \int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,r,\mathbf{x}',\mathbf{y})(L^m_r-\tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_r)u_m(r,\mathbf{y})d\mathbf{y}. \qquad (2.38)
\end{aligned}
$$

Integrating (2.38) between t and $t_0 \in (t,T]$ for a first given $\boldsymbol{\xi}'$ and between t_0 and T with a possibly different $\tilde{\boldsymbol{\xi}}'$ yields:

$$\begin{aligned}
0 &= \tilde{P}^{m,(\tau,\boldsymbol{\xi}')}_{t_0,t} u_m(t_0,\mathbf{x}') - u_m(t,\mathbf{x}') + \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y}) f_m(s,\mathbf{y}) \\
&+ \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s - \tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_s) u_m(s,\mathbf{y}) \\
&+ \tilde{P}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_{T,t} u_m(T,\mathbf{x}') - \tilde{P}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_{t_0,t} u_m(t_0,\mathbf{x}') + \int_{t_0}^{T} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}(t,s,\mathbf{x}',\mathbf{y}) f_m(s,\mathbf{y}) \\
&+ \int_{t_0}^{T} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s - \tilde{L}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_s) u_m(s,\mathbf{y}).
\end{aligned}$$

Recalling that $u_m(T,\mathbf{x}') = g_m(\mathbf{x}')$ (terminal condition), and with the notations of (2.37) the above equation becomes:

$$\begin{aligned}
u_m(t,\mathbf{x}') &= \tilde{P}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_{T,t} g_m(\mathbf{x}') + \tilde{G}^{m,(\tau,\boldsymbol{\xi}')}_{t_0,t} f_m(t,\mathbf{x}') + \tilde{G}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_{T,t_0} f_m(t,\mathbf{x}') \\
&+ \tilde{P}^{m,(\tau,\boldsymbol{\xi}')}_{t_0,t} u_m(t_0,\mathbf{x}') - \tilde{P}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_{t_0,t} u_m(t_0,\mathbf{x}') \\
&+ \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \Big(\mathbb{I}_{s\le t_0} \tilde{p}^{m,(\tau,\boldsymbol{\xi}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s - \tilde{L}^{m,(\tau,\boldsymbol{\xi}')}_s) \\
&\quad + \mathbb{I}_{s>t_0} \tilde{p}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}(t,s,\mathbf{x}',\mathbf{y})(L^m_s - \tilde{L}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_s) \Big) u_m(s,\mathbf{y}).
\end{aligned} \tag{2.39}$$

We see that for $\boldsymbol{\xi}' \neq \tilde{\boldsymbol{\xi}}'$ we have an additional discontinuity term deriving from the change of freezing point along the time variable. Of course expression (2.39) can be differentiated in space and taking **then**

$$t_0 = t + c_0 \mathbf{d}^2(\mathbf{x},\mathbf{x}'), \tag{2.40}$$

i.e. t_0 precisely corresponds to the critical time at which a change of regime occurs, and $\boldsymbol{\xi}' = \mathbf{x}'$, $\tilde{\boldsymbol{\xi}}' = \mathbf{x}$ precisely allows, when expanding $D^2_{\mathbf{x}_1} u_m(t,\mathbf{x}) - D^2_{\mathbf{x}_1} u_m(t,\mathbf{x}')$ using (2.12) for the first term and (2.39) for the second one, to exploit the previous analysis that led to (2.32) and (2.36) and which relied on the suitable choice of freezing point. We again insist on the fact that, in the analysis, t_0 is an additional freezing parameter, which is *a posteriori* chosen according to (2.40) as a function of $(t,\mathbf{x},\mathbf{x}')$. In particular the parameter t_0 does not intervene in the various possible differentiations of the considered perturbative expansions.

This approach eventually leads to:

$$\begin{aligned}
&|D^2_{\mathbf{x}_1} u_m(t,\mathbf{x}) - D^2_{\mathbf{x}_1} u_m(t,\mathbf{x}')| \\
&\le \mathbf{d}^\gamma(\mathbf{x},\mathbf{x}') \Big[C\big(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \big) + \Lambda\big(c_0^{-n+\frac12+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}} \big) \|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \Big] \\
&+ \Big| \big(D^2_{\mathbf{x}_1} \tilde{P}^{m,(\tau,\boldsymbol{\xi}')}_{t_0,t} u_m(t_0,\mathbf{x}') - D^2_{\mathbf{x}_1} \tilde{P}^{m,(\tau,\tilde{\boldsymbol{\xi}}')}_{t_0,t} u_m(t_0,\mathbf{x}') \big) \big|_{t_0 = t + c_0 \mathbf{d}^2(\mathbf{x},\mathbf{x}')} \Big|.
\end{aligned} \tag{2.41}$$

The last contribution can be controlled through cancellation techniques and the key estimate (2.33) on the difference of the flows. The specific choice of $t_0 = t + c_0 \mathbf{d}^2(\mathbf{x},\mathbf{x}')$ then precisely provides the required order leading to:

$$\begin{aligned}
&|D^2_{\mathbf{x}_1} u_m(t,\mathbf{x}) - D^2_{\mathbf{x}_1} u_m(t,\mathbf{x}')| \\
&\le \mathbf{d}^\gamma(\mathbf{x},\mathbf{x}') \Big[C\big(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \big) + \Big(\Lambda\big(c_0^{-n+\frac12+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}} \big) + C c_0^{\frac{\gamma}{2n-1}} \Big) \|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \Big].
\end{aligned}$$

We refer to Lemma 17 for results associated with the discontinuity term in (2.41).

2.5. **Conclusion and outline of the derivation of estimate in Theorem 1.** We have detailed up to now what happens with the second order derivatives w.r.t. the non-degenerate variables. The previous procedure can be applied as well to control the Hölder moduli w.r.t. the degenerate ones. We therefore end up with the following kind of estimate:

$$\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \le C\big(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \big) + \|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \Big[\Lambda\big(c_0^{-n+\frac12+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}} + T^{\frac{\gamma}{2}} \big) + C c_0^{\frac{\gamma}{2n-1}} \Big]. \tag{2.42}$$

Equation (2.42) would provide exactly the expected control if Λ and c_0 are small enough. On the one hand, the final parameter c_0 can always be chosen small enough (cutting threshold). On the other hand, it will appear from the proofs that the constant Λ in (2.42) actually depends on the Hölder norms of the considered coefficients

(see again Remark 3). If these quantities are small, i.e. the coefficients do not vary much and the components that transmit the noise are almost linear, then Λ will be small. For $\Big[\Lambda\big(c_0^{-n+\frac{1}{2}+\frac{\gamma}{2}}+c_0^{\frac{\gamma}{2}}+T^{\frac{\gamma}{2}}\big)+Cc_0^{\frac{\gamma}{2n-1}}\Big]\leq k_0<1$, we eventually derive:

$$\|u_m\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\leq\frac{C}{1-k_0}\big(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}+\|f\|_{L^\infty(C^{\gamma}_{b,\mathbf{d}})}\big), \tag{2.43}$$

which is precisely the expected control. The general case, is proved through a scaling argument which also allows to balance the opposite effects of c_0 (meant to be small, in particular $c_0\leq 1$) in the above bounds. This last point will be discussed in Section 6.

2.6. **Organization of this paper.** The remaining part of this article is organized as follows. We prove in Section 3 various properties for the density of the linearized Gaussian proxy: precise pointwise estimates for the density itself and its derivatives (see equation (2.8)) and some useful controls allowing cancellation arguments in our perturbative analysis. Section 4 is then devoted to the control of the supremum norms of the non-degenerate derivatives, corresponding to the previous equation (2.28). Section 5 addresses the issues of Hölder controls. Section 6 is concerned with the above mentioned scaling issues and we also conclude there the final proof of Theorem 1. Eventually, some auxiliary, but crucial, technical results are proved in Appendix A for the regularity of the flow and the mean, in Appendix B for the regularity of the resolvent and the covariance, and in Appendix C for the technical points related to the scaling analysis.

3. Gaussian proxy and associated controls

We first aim here at proving the control of equation (2.8). We recall that our point is to control the density of $(\tilde{\mathbf{X}}_s^{m,(\tau,\boldsymbol{\xi})})_{s\in(t,T]}$ satisfying (2.6).

WARNING: for notational simplicity, for the rest of the document we drop the sub and superscripts in m associated with the regularizations. We rewrite, with some notational abuse, for fixed $(\tau,\boldsymbol{\xi})\in[0,T]\times\mathbb{R}^{nd}$, the dynamics in (2.6) as:

$$\begin{aligned}d\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})}&=[\mathbf{F}(v,\boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))+D\mathbf{F}(v,\boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))(\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})}-\boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))]dv+B\sigma(v,\boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))dW_v,\\ &\forall v\in[t,s],\ \tilde{\mathbf{X}}_t^{(\tau,\boldsymbol{\xi})}=\mathbf{x},\end{aligned} \tag{3.1}$$

keeping in mind that $\mathbf{F},\boldsymbol{\theta},\sigma$ in (3.1) are *smooth* coefficients. We will give in the next subsection some key-controls to investigate the terms appearing the perturbative expansions (2.11) and (2.12).

3.1. **Controls for the frozen density.** We explicitly integrate (3.1) to obtain for any $v\in[t,s]$:

$$\begin{aligned}\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})}&=\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,t)\mathbf{x}+\int_t^v\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,u)\Big(\mathbf{F}(u,\boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi}))-D\mathbf{F}(u,\boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi}))\boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi})\Big)du\\ &\quad+\int_t^v\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,u)B\sigma(u,\boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi}))dW_u\\ &=:\mathbf{m}_{v,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})+\int_t^v\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,u)B\sigma(u,\boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi}))dW_u,\end{aligned}$$

where $(\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,u))_{t\leq u,v\leq s}$ stands for the resolvent associated with the collection of partial gradients in $(D\mathbf{F}(v,\boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi})))_{v\in[t,s]}$, introduced in (2.7), which satisfies for $v\in[t,s]$:

$$\partial_v\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,t)=D\mathbf{F}(v,\boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,t),\ \tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(t,t)=\mathbf{I}_{nd,nd}. \tag{3.2}$$

Note in particular that since the partial gradients are subdiagonal $\det(\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,t))=1$.

Also, for $v\in[t,s]$, we recall that $\mathbf{m}_{v,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})$ stands for the mean of $\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})}$ and corresponds as well to the solution of (3.1) when $\sigma=0$ and the starting point is $\mathbf{x}$. We write:

$$\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})}=\mathbf{m}_{v,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})+\int_t^v\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,u)B\sigma(u,\boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi}))dW_u,\ v\in[t,s]. \tag{3.3}$$

Importantly, we point out that $\mathbf{x}\in\mathbb{R}^{nd}\mapsto\mathbf{m}_{v,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})$ is **affine** w.r.t. the starting point $\mathbf{x}$. Precisely, for $\mathbf{x},\mathbf{x}'\in\mathbb{R}^{nd}$:

$$\mathbf{m}_{v,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}+\mathbf{x}')=\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,t)\mathbf{x}'+\mathbf{m}_{v,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}). \tag{3.4}$$

We first give in the next proposition a key estimate on the covariance matrix associated with (3.3) and its properties w.r.t. a suitable scaling of the system.

Proposition 2 (Good Scaling Properties of the Covariance Matrix)**.** *The covariance matrix of* $\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})}$ *in* (3.3) *writes:*

$$\tilde{\mathbf{K}}_{v,t}^{(\tau,\boldsymbol{\xi})} := \int_t^v \tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,u) Ba(u,\boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi})) B^* \tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(v,u)^* du.$$

Uniformly in $(\tau,\boldsymbol{\xi}) \in [0,T]\times\mathbb{R}^{nd}$ *and* $s\in[0,T]$*, it satisfies a good scaling property in the sense of Definition 3.2 in* [7] *(see also Proposition 3.4 of that reference). That is, for any fixed* $T>0$*, there exists* $C_{3.5} := C_{3.5}((\mathbf{A}),T) \geq 1$ *s.t. for all* $0\leq t<v\leq s\leq T$*,* $(\tau,\boldsymbol{\xi})\in[0,T]\times\mathbb{R}^{nd}$*:*

$$\forall \boldsymbol{\zeta}\in\mathbb{R}^{nd},\ C_{3.5}^{-1}(v-t)^{-1}|\mathbb{T}_{v-t}\boldsymbol{\zeta}|^2 \leq \langle \tilde{\mathbf{K}}_{v,t}^{(\tau,\boldsymbol{\xi})}\boldsymbol{\zeta},\boldsymbol{\zeta}\rangle \leq C_{3.5}(v-t)^{-1}|\mathbb{T}_{v-t}\boldsymbol{\zeta}|^2, \tag{3.5}$$

where for any $u>0$*, we denote by* $\mathbb{T}_u$ *the* intrinsic *scale matrix introduced in* (2.9)*. Namely:*

$$\mathbb{T}_u = \begin{pmatrix} u\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & u^2\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & u^n\mathbf{I}_{d,d} \end{pmatrix}.$$

The proof of the above proposition readily follows from Proposition 3.3 and Lemma 3.6 in [7]. We now state some important density bounds for the linearized model.

Proposition 3 (Density of the linearized dynamics)**.** *Under* **(A)***, we have that, for any* $s\in(t,T]$ *the random variable* $\tilde{\mathbf{X}}_s^{(\tau,\boldsymbol{\xi})}$ *in* (3.3) *admits a Gaussian density* $\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\cdot)$ *which writes for any* $\mathbf{y}\in\mathbb{R}^{nd}$*:*

$$\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) := \frac{1}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})})^{\frac12}} \exp\left(-\frac12\left\langle (\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})})^{-1}(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y}), \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y}\right\rangle\right), \tag{3.6}$$

with $\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})}$ *as in Proposition 2. Also, there exists* $C := C((\mathbf{A}),T) > 0$ *s.t. for each multi-index* $\vartheta = (\vartheta_1,\cdots,\vartheta_n)\in\mathbb{N}^n$*,* $|\vartheta|\leq 3$ *and denoting by* $D_{\mathbf{x}}^{\vartheta} := D_{\mathbf{x}_1}^{\vartheta_1}\cdots D_{\mathbf{x}_n}^{\vartheta_n}$*, we have:*

$$\begin{aligned} |D_{\mathbf{x}}^{\vartheta}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})| &\leq \frac{C}{(s-t)^{\sum_{i=1}^n \vartheta_i(i-\frac12)+\frac{n^2d}{2}}} \exp\left(-C^{-1}(s-t)\big|\mathbb{T}_{s-t}^{-1}\big(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y}\big)\big|^2\right) \\ &=: \frac{C}{(s-t)^{\sum_{i=1}^n \vartheta_i(i-\frac12)}}\bar{p}_{C^{-1}}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}), \end{aligned} \tag{3.7}$$

with $\int_{\mathbb{R}^{nd}} d\mathbf{y}\bar{p}_{C^{-1}}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})$*, up to a modification of the constants in* (3.7)*.*

Remark 4 (A slight abuse of notation)**.** *To ease the reading we denote, when there is no possible ambiguity,* $\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y}) := \bar{p}_{C^{-1}}^{(t,\mathbf{x})}(t,s,\mathbf{x},\mathbf{y})$*.*

Remark 5 (Regularizing effect of the quasi-distance)**.** *From equation* (3.7)*, we derive from the definition of* $\mathbf{d}$ *in* (1.9) *that for any given* $\beta>0$*, there exists* C_β *s.t.*

$$\mathbf{d}^{\beta}\big((s-t)^{\frac12}\mathbb{T}_{s-t}^{-1}\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}),(s-t)^{\frac12}\mathbb{T}_{s-t}^{-1}\mathbf{y}\big)|D_{\mathbf{x}}^{\vartheta}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})| \leq \frac{C_\beta}{(s-t)^{\sum_{i=1}^n\vartheta_i(i-\frac12)-\frac{\beta}{2}}}\bar{p}_{C_\beta^{-1}}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}), \tag{3.8}$$

i.e. equation (3.8) *quantifies the regularizing effect of the scaled arguments in the quasi-distance.*

Proof. Expression (3.6) readily follows from (3.2) and (3.3). Differentiating w.r.t. $\mathbf{x}$ recalling from (3.4) that $\mathbf{x}\mapsto\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})$ is affine yields:

$$D_{\mathbf{x}_j}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) = -\Big[\big[\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)\big]^*(\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})})^{-1}(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y})\Big]_j\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}). \tag{3.9}$$

The point is now to use scaling arguments. We can first rewrite

$$\big[\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)\big]^*(\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})})^{-1} = (s-t)\big[\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)\big]^*\mathbb{T}_{s-t}^{-1}(\widehat{\tilde{\mathbf{K}}_1^{s,t}})^{-1}\mathbb{T}_{s-t}^{-1}, \tag{3.10}$$

where $\widehat{\tilde{\mathbf{K}}_1^{s,t}}$ is the covariance matrix of the rescaled process $\big((s-t)^{\frac12}\mathbb{T}_{s-t}^{-1}\tilde{\mathbf{X}}_{t+v(s-t)}^{t,\mathbf{x}}\big)_{v\in[0,1]}$ at time 1. From the good-scaling property of Proposition 2, it is plain to derive that $\widehat{\tilde{\mathbf{K}}_1^{s,t}}$ a non-degenerate bounded matrix, i.e. there exists $\hat{C}\geq 1$ s.t. for any $\boldsymbol{\zeta}\in\mathbb{R}^{nd}$, $(\hat{C})^{-1}|\boldsymbol{\zeta}|^2 \leq \langle\widehat{\tilde{\mathbf{K}}_1^{s,t}}\boldsymbol{\zeta},\boldsymbol{\zeta}\rangle \leq \hat{C}|\boldsymbol{\zeta}|^2$. A similar rescaling argument yields on the deterministic system (3.2) of the resolvent yields that $\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)$ can also be written as:

$$[\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)]^* = \mathbb{T}_{s-t}^{-1}\Big[\widehat{\tilde{\mathbf{R}}}^{(\tau,\boldsymbol{\xi}),s,t}(1,0)\Big]^*\mathbb{T}_{s-t}, \tag{3.11}$$

where again $\widehat{\widehat{\mathbf{R}}}^{(\tau,\boldsymbol{\xi}),s,t}(1,0)$ is the resolvent at time 1 of the rescaled system

$$\Big(\mathbb{T}_{s-t}[\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(t+v(s-t),t)]^*\mathbb{T}_{s-t}^{-1}\Big)_{v\in[0,1]} = \Big(\big[\widehat{\widehat{\mathbf{R}}}^{(\tau,\boldsymbol{\xi}),s,t}(v,0)\big]^*\Big)_{v\in[0,1]}$$

associated with (3.2). From the analysis performed in Lemma 5.1 in [12] (see also the proof of Proposition 3.7 in [7]) one derives that there exists $\hat{C}_1$ s.t. for any $\boldsymbol{\zeta}\in\mathbb{R}^{nd}$,

$$|\big[\widehat{\widehat{\mathbf{R}}}^{(\tau,\boldsymbol{\xi}),s,t}(1,0)\big]^*\boldsymbol{\zeta}| \le \hat{C}_1|\boldsymbol{\zeta}|. \tag{3.12}$$

Equations (3.9), (3.10) and (3.11) therefore yield:

$$\begin{aligned}
&|D_{\mathbf{x}_j}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})|\\
\le\ & (s-t)^{-j+\frac{1}{2}}\left|\left(\Big[\widehat{\widehat{\mathbf{R}}}^{(\tau,\boldsymbol{\xi}),s,t}(1,0)\Big]^*\big(\widehat{\tilde{\mathbf{K}}_1^{s,t}}\big)^{-1}\big((s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y})\big)\right)_j\right|\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\\
\le\ & C(s-t)^{-j+\frac{1}{2}}(s-t)^{\frac{1}{2}}|\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y})|\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}).
\end{aligned}$$

From the explicit expression (3.6), Proposition 2 and the above equation, we eventually derive:

$$\begin{aligned}
&|D_{\mathbf{x}_j}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})|\\
\le\ & \frac{C}{(s-t)^{j-\frac{1}{2}}}\Big((s-t)^{\frac{1}{2}}|\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y})|\Big)\frac{1}{(s-t)^{\frac{n^2d}{2}}}\exp\Big(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})-\mathbf{y})|^2\Big)\\
\le\ & \frac{C}{(s-t)^{j-\frac{1}{2}}}\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y}),
\end{aligned}$$

up to a modification of C, which gives the statement for one partial derivative. The controls on the higher order derivatives are obtained similarly (see e.g. the proof of Lemma 5.5 of [7] for the bounds on $D^2_{\mathbf{x}_1}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})$). □

As a direct consequence of Proposition 3 we have the following result for the semi-group $\tilde{P}^{(\tau,\boldsymbol{\xi})}$ associated with (3.3):

Lemma 4. *For $\gamma\in(0,1)$, under* **(A)**, *there exists $C:=C((\mathbf{A}),T)$, s.t. for any function $\psi\in C^\gamma_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})$, and any given multi-index ϑ, $|\vartheta|\in[\![1,3]\!]$, for all $0\le t<s\le T$, $\mathbf{x}\in\mathbb{R}^{nd}$:*

$$|D^\vartheta_{\mathbf{x}}\tilde{P}^{(\tau,\boldsymbol{\xi})}_{s,t}\psi(\mathbf{x})|\big|_{\boldsymbol{\xi}=\mathbf{x}} \le C\|\psi\|_{C^\gamma_{\mathbf{d}}}(s-t)^{-\sum_{i=1}^n|\vartheta_i|(i-\frac{1}{2})+\frac{\gamma}{2}}. \tag{3.13}$$

Proof. It suffices to write:

$$D^\vartheta_{\mathbf{x}}\tilde{P}^{(\tau,\boldsymbol{\xi})}_{s,t}\psi(\mathbf{x}) = \int_{\mathbb{R}^{nd}}d\mathbf{y}D^\vartheta_{\mathbf{x}}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})\psi(\mathbf{y}) = \int_{\mathbb{R}^{nd}}d\mathbf{y}D^\vartheta_{\mathbf{x}}\tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})[\psi(\mathbf{y})-\psi(\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))],$$

so that from Proposition 3 and the smoothness of ψ:

$$|D^\vartheta_{\mathbf{x}}\tilde{P}^{(\tau,\boldsymbol{\xi})}_{s,t}\psi(\mathbf{x})| \le \frac{C}{(s-t)^{\sum_{i=1}^n\vartheta_i(i-\frac{1}{2})}}\|\psi\|_{C^\gamma_{\mathbf{d}}}\int_{\mathbb{R}^{nd}}\bar{p}^{(\tau,\boldsymbol{\xi})}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})\mathbf{d}^\gamma(\mathbf{y},\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})),$$

which yields the result thanks to Remark 5 recalling from the homogeneity of $\mathbf{d}$ that $\mathbf{d}^\gamma(\mathbf{y},\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})) = \big[(s-t)^{\frac{1}{2}}\mathbf{d}((s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}\mathbf{y},(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))\big]^\gamma$. □

Note carefully that in the above lemma $|\vartheta|\ge 1$. Indeed, if $|\vartheta|=0$ we cannot benefit from any regularizing effects which are precisely due to cancellation techniques.

We now give some useful controls involving the previous Gaussian kernel which will be used in our perturbative analysis. The main interest of the estimates below is that they precisely allow to exploit cancellation techniques.

Proposition 5. *For all $0 \le t \le s \le T$, $(\mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$, the following identities hold:*

$$\int_{\mathbb{R}^{nd}} \tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2} d\mathbf{y} = [\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})}]_{1,1}, \tag{3.14}$$

$$\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1} \tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))_1 d\mathbf{y} = \mathbf{0}_d, \tag{3.15}$$

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D^2_{\mathbf{x}_1} \tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) \cdot (\mathbf{y} - \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))_1 d\mathbf{y} = \mathbf{0}_{d,d},\ k \in [\![1,n]\!], \tag{3.16}$$

$$\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1} \tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) \mathrm{Tr}\Big(\mathbf{M}(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2}\Big) d\mathbf{y} = 2\mathbf{M},\ \mathbf{M} \in \mathbb{R}^d \otimes \mathbb{R}^d, \tag{3.17}$$

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D^2_{\mathbf{x}_1} \tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) \mathrm{Tr}\Big(\mathbf{M}(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2}\Big) d\mathbf{y} = \mathbf{0}_{d,d,d},\ k \in [\![1,n]\!], \mathbf{M} \in \mathbb{R}^d \otimes \mathbb{R}^d. \tag{3.18}$$

Where, in (3.14), we define for all $(i,j) \in [\![1,n]\!]^2$ and $M \in (\mathbb{R}^{nd})^{\otimes 2}$, $[M]_{i,j}$ is the $d \times d$ block matrix corresponding to the entry of M on the i^{th} line and the j^{th} column.

Proof. First of all, remark that equation (3.14) simply follows from a direct covariance computation.

Observe now that from Proposition 3, we have $\int_{\mathbb{R}^{nd}} \tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y})(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))_1 d\mathbf{y} = \mathbf{0}_d$. Differentiating twice this expression w.r.t. $\mathbf{x}_1$ and using the Leibniz formula (recalling as well the identity (3.4) which yields $D_{\mathbf{x}_1}[\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})]_1 = [\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)]_{1,1} = \mathbf{I}_{d,d}$) gives (3.15). Iterating the differentiation w.r.t. $D_{\mathbf{x}_k}$ then yields (3.16) (observing again that $D_{\mathbf{x}_k}[\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})]_1 = [\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)]_{1,k}$, i.e. $D_{\mathbf{x}_k}[\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})]_1) = \mathbf{I}_{d,d}$ if $k = 1$ and $\mathbf{0}_{d,d}$ for $k > 1$). Observe that $D_{\mathbf{x}_1} \int_{\mathbb{R}^{nd}} \mathrm{Tr}\Big(\mathbf{M}(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2}\Big) \tilde{p}^{(\tau,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) d\mathbf{y} = D_{\mathbf{x}_1} \mathrm{Tr}\Big(\mathbf{M}[\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})}]_{1,1}\Big) = \mathbf{0}_d$. Differentiating again w.r.t. $D_{\mathbf{x}_1}$, the Leibniz formula and identity $D_{\mathbf{x}_1}[\mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}(\mathbf{x})]_1 = [\tilde{\mathbf{R}}^{(\tau,\boldsymbol{\xi})}(s,t)]_{1,1} = \mathbf{I}_{d,d}$ yield (3.17). Eventually, (3.18) can be derived again from derivation or observing that the sum of the length of the multi-derivation index, here 3, and the power integrated, here 2, is an odd number. □

3.2. **Additional sensitivity controls : covariance, (mollified) flow, mean.** We now state three important estimates associated with our proxy. The first one concerns the sensitivity of the covariance w.r.t. the frozen point, the second and third one concern the linearization or sensitivity w.r.t. the initial point for the frozen (mollified) differential system (2.5). For the sake of simplicity, their proofs are postponed to appendixes A and B. We have:

Lemma 6 (Sensitivities of the covariance)**.** *There exists $\Lambda := \Lambda((\mathbf{A}), T)$ as in Remark 3 s.t. for given $(\boldsymbol{\xi}, \boldsymbol{\xi}') \in (\mathbb{R}^{nd})^2$ and $0 \le t < s \le T$, $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$:*

$$|[\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi})}]_{1,1} - [\tilde{\mathbf{K}}_{s,t}^{(\tau,\boldsymbol{\xi}')}]_{1,1}| \le \Lambda(s-t)\big(\mathbf{d}^\gamma(\boldsymbol{\xi}, \boldsymbol{\xi}') + (s-t)^{\frac{\gamma}{2}}\big). \tag{3.19}$$

The proof is given in Appendix B.2.1. We importantly point out that, in Lemma 6, the constant Λ mainly depends on the Hölder norms of the coefficients and is *small* provided the coefficients do not *vary much*. Precisely, it can be shown that Λ writes:

$$\Lambda := \tilde{C}\big(\|a\|_{L^\infty(C^\gamma_{\mathbf{d}})} + \sum_{i=2}^n \|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})}\big) \tag{3.20}$$

for some *universal* constant $\tilde{C}$, where, in the above equation, we write with the notation of Section 1.4:

$$\begin{aligned} \|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})} &:= \sup_{(t,\mathbf{z}) \in [0,T] \times \mathbb{R}^{(n-i+2)d}} \|(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1}(t,\mathbf{z},\cdot)\|_{C^{\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)} \\ &\quad + \sum_{j=i}^n \sup_{(t,\mathbf{z}) \in [0,T] \times \mathbb{R}^{(n-i+2)d}} \|(\mathbf{F}_i)_j(t,\mathbf{z},\cdot)\|_{C^{\frac{2i-3+\gamma}{2j-1}}(\mathbb{R}^d, \mathbb{R}^d)}. \end{aligned} \tag{3.21}$$

Namely, the quantity $\|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})}$ gathers the Hölder moduli of $\mathbf{F}_i$ at the intrinsic associated scales according to the distance $\mathbf{d}$ in the variables $j \in [\![i,n]\!]$ as well as the Hölder norm of the gradient w.r.t. the component which transmits the noise (but importantly not its supremum norm). Said differently, the $L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})$ semi-norm of $\mathbf{F}_i$ gathers the Hölder norms of the fractional parts of $\big(D^{\lfloor \frac{2i-3+\gamma}{2j-1} \rfloor}_{\mathbf{x}_j} \mathbf{F}_i\big)_{j \in [\![i-1,n]\!]}$ in the j^{th} variable with corresponding Hölder index $\frac{2i-3+\gamma}{2j-1} - \lfloor \frac{2i-3+\gamma}{2j-1} \rfloor$, where $\lfloor \cdot \rfloor$ stands for the integer part. Eventually, observe as well that, with respect to the notation (1.14) of Section 1.4, $\|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})} = \|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d}})} - \|D_{\mathbf{x}_{i-1}} \mathbf{F}_i\|_{L^\infty(L^\infty)}$.

We again refer to Appendix A for a precise statement and proof of this assertion (3.20) (See also Lemma 19 below and its proof for similar properties).

The second result is the following.

Lemma 7. *There exists $C := C((\mathbf{A}))$ s.t. for all $0 \le t \le s \le T$, $(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$:*

$$\mathbf{d}(\boldsymbol{\theta}_{s,t}(\mathbf{x}), \boldsymbol{\theta}_{s,t}(\mathbf{x}')) \le C\Big(\mathbf{d}(\mathbf{x}, \mathbf{x}') + (s-t)^{\frac{1}{2}}\Big).$$

The proof of Lemma 7 is postponed to Appendix A.1.1.

Eventually, this last lemma concerns the impact of the freezing point in the linearization procedure. Namely,

Lemma 8 (Sensitivity of the linearized flow w.r.t. the freezing parameter)**.** *There exists $C := C((\mathbf{A}))$ s.t. for all $\tau = t$, $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ at the change of regime time t_0 defined in (2.40) (i.e. $t_0 = (t + c_0\mathbf{d}^2(\mathbf{x}, \mathbf{x}')) \wedge T$):*

$$\mathbf{d}\big(\mathbf{m}_{t_0,t}^{(\tau,\mathbf{x})}(\mathbf{x}'), \mathbf{m}_{t_0,t}^{(\tau,\mathbf{x}')}(\mathbf{x}')\big) = \mathbf{d}\big(\mathbf{m}_{t_0,t}^{(\tau,\mathbf{x})}(\mathbf{x}'), \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')\big) \le C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}').$$

Again, the proof of Lemma 8 is postponed to Appendix A.1.2.

4. Control of the supremum of the derivatives w.r.t. the non-degenerate variables

WARNING: for notational simplicity, we drop from now on the sub and superscripts τ associated with the linearization since, as soon as the function u in (2.11) is evaluated at time t in $[0,T]$, we choose this parameter to be equal to t. For example, $\tilde{p}^{(\tau,\boldsymbol{\xi})}, \mathbf{m}_{s,t}^{(\tau,\boldsymbol{\xi})}, \boldsymbol{\theta}_{u,\tau}$ become respectively $\tilde{p}^{\boldsymbol{\xi}}, \mathbf{m}_{s,t}^{\boldsymbol{\xi}}, \boldsymbol{\theta}_{u,t}$. Recall as well that we decided to omit the dependence of such a function u in the regularization parameter m. This dependence is implicitly assumed and we will derive the desired control uniformly in m.

The result we aim at proving in this section concerns the supremum norm of the derivatives w.r.t. the non degenerate variables. Namely, we here prove the following result.

Proposition 9. *Let $\gamma \in (0,1)$ be given. Suppose that* **(A)** *is in force and that the terminal condition g and source term f of the Cauchy problem (1.1) satisfy: $g \in C_{b,\mathbf{d}}^{2+\gamma}(\mathbb{R}^{nd}, \mathbb{R})$ and $f \in L^\infty\big([0,T], C_{b,\mathbf{d}}^{\gamma}(\mathbb{R}^{nd}, \mathbb{R})\big)$. Then, there exist $C := C((\mathbf{A}), T)$ and Λ as in Remark 3 such that for any $(t, \mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$,*

$$|u(t,\mathbf{x})| + |D_{\mathbf{x}_1} u(t,\mathbf{x})| + |D_{\mathbf{x}_1}^2 u(t,\mathbf{x})| \le C\Big\{\|g\|_{C_{b,\mathbf{d}}^{2+\gamma}} + (T-t)^{\frac{\gamma}{2}} \|f\|_{L^\infty(C_{b,\mathbf{d}}^{\gamma})}\Big\} + \Lambda(T-t)^{\frac{\gamma}{2}} \|u\|_{L^\infty(C_{b,\mathbf{d}}^{2+\gamma})}. \tag{4.1}$$

Note first that the control for the function itself readily follows from **(A)** and the Feynman-Kac representation of the solution of (2.4) under **(A)**, i.e. recall from (2.3) that $u(t,\mathbf{x}) = \mathbb{E}[g(\mathbf{X}_T^{t,\mathbf{x}})] + \int_t^T ds \mathbb{E}[f(s, \mathbf{X}_s^{t,\mathbf{x}})]$ and that $T \le 1$.

For the derivatives, let us now start from (2.11) to control pointwise the second order derivative of u in the non-degenerate variable, i.e. $\|D_{\mathbf{x}_1}^2 u\|_{L^\infty}$. The first one can be controlled similarly and more directly. Write for any $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$:

$$\begin{aligned}
|D_{\mathbf{x}_1}^2 u(t,\mathbf{x})| &\le \|D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\boldsymbol{\xi}} g\|_{L^\infty} + \|D_{\mathbf{x}_1}^2 \tilde{G}^{\boldsymbol{\xi}} f\|_{L^\infty} \\
&+ \Big| \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \Delta_{1,\mathbf{F},\sigma}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y},u) d\mathbf{y} \Big| \\
&+ \Big| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{y}_i} \cdot \Big(\big(D_{\mathbf{x}_1}^2 \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \otimes \Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}) \big) \Big) u(s,\mathbf{y}) d\mathbf{y} \Big|,
\end{aligned} \tag{4.2}$$

where we recall from (2.13) and (2.17) that:

$$\begin{aligned}
\Delta_{1,\mathbf{F},\sigma}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y},u) &= \langle \big(\mathbf{F}_1(s,\mathbf{y}) - \mathbf{F}_1(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))\big), D_{\mathbf{y}_1} u(s,\mathbf{y}) \rangle \\
&\quad + \frac{1}{2} \mathrm{Tr}\Big(\big(a(s,\mathbf{y}) - a(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))\big) D_{\mathbf{y}_1}^2 u(s,\mathbf{y}) \Big), \\
\Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}) &= \mathbf{F}_i(s,\mathbf{y}) - \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}.
\end{aligned} \tag{4.3}$$

This section is then organized as follows: we first estimate the non degenerate part of (4.2) (first term in the r.h.s. of the equation) thanks to Lemma 10 in Section 4.1, we then estimate the degenerate part of (4.2) (second term in the r.h.s. of the equation) thanks to Lemma 11 in Section 4.2 and eventually estimate the remainder of (4.2) (third and fourth terms in the r.h.s. of the equation) thanks to Lemma 12 in Section 4.3. Proposition 13 then follows from the previous Lemmas. □

4.1. **Control of the non-degenerate part of the perturbative term.** The aim of this section is to prove identity (2.16) appearing in the detailed guide to the proof. To this end, we provide a general differentiation result, which will be useful as well in Section 5 to deal with the Hölder norms. Under the current assumptions on $a, \mathbf{F}$, the following lemma holds.

Lemma 10. *There exists $\Lambda := \Lambda((\mathbf{A}), T)$ as in Remark 3 s.t. for each multi-index $\vartheta = (\vartheta_1, \dots, \vartheta_n) \in \mathbb{N}^n$, $|\vartheta| \leq 3$:*

$$\Big| \int_{\mathbb{R}^{nd}} D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \Delta_{1,\mathbf{F},\sigma}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y},u) d\mathbf{y} \Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \leq \Lambda \|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} (s-t)^{-\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})+\frac{\gamma}{2}}. \tag{4.4}$$

Proof of Lemma 10. We first recall the control (2.14)

$$|\Delta_{1,\mathbf{F},\sigma}(t,s,\mathbf{y},\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),u)| \leq \Big([\mathbf{F}_1]_{\mathbf{d},\gamma} \|D_{\mathbf{x}_1} u(s,\cdot)\|_{L^\infty} + \frac{1}{2} [a(s,\cdot)]_{\mathbf{d},\gamma} \|D^2_{\mathbf{x}_1} u(s,\cdot)\|_{L^\infty} \Big) \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})).$$

From this control and Proposition 3, we directly obtain:

$$\begin{aligned}
&\Big| \int_{\mathbb{R}^{nd}} D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \Delta_{1,\mathbf{F},\sigma}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y},u) d\mathbf{y} \Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & \Big(\int_{\mathbb{R}^{nd}} |D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})|\ \mathbf{d}^\gamma\big(\mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})\big) \big(\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})} \|D_{\mathbf{x}_1} u\|_{L^\infty} + \|a\|_{L^\infty(C^\gamma_{\mathbf{d}})} \|D^2_{\mathbf{x}_1} u\|_{L^\infty} \big) d\mathbf{y} \Big)\Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & \Lambda \|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} (s-t)^{-\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})} \int_{\mathbb{R}^{nd}} \bar{p}_{C^{-1}}(s,t,\mathbf{x},\mathbf{y}) \mathbf{d}^\gamma(\boldsymbol{\theta}_{s,t}(\mathbf{x}),\mathbf{y}) d\mathbf{y} \\
\leq\ & \Lambda \|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} (s-t)^{-\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})+\frac{\gamma}{2}},
\end{aligned}$$

with the notations of Remark 4 for the last but one inequality. □

Equation (2.16) readily follows from Lemma 10 taking $\vartheta = (2, 0, \dots, 0)$. Namely:

$$\begin{aligned}
&\Big| \int_t^T ds \int_{\mathbb{R}^{nd}} D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \Delta_{1,\mathbf{F},\sigma}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y},u) d\mathbf{y} \Big| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & \Lambda \|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \int_t^T \frac{ds}{(s-t)^{1-\frac{\gamma}{2}}} \\
\leq\ & \Lambda \|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} (T-t)^{\frac{\gamma}{2}}.
\end{aligned} \tag{4.5}$$

4.2. **Control of the degenerate part of the perturbative term.** The point is here to control the terms $\sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \big\langle \Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}), D_{\mathbf{y}_i} u(s,\mathbf{y}) \big\rangle d\mathbf{y}$ appearing in equation (2.18) of the detailed guide to the proof. We precisely want to derive equation (2.25).

The bound will actually follow from the more general following result, which will again be useful for the Hölder norm in Section 5.

Lemma 11 (First Besov Control Lemma)**.** *There exists $\Lambda := \Lambda((\mathbf{A}), T)$ as in Remark 3 s.t. for each multi-index $\vartheta = (\vartheta_1, \dots, \vartheta_n) \in \mathbb{N}^n$, $|\vartheta| \leq 3$:*

$$\sum_{i=2}^n \Big| \int_{\mathbb{R}^{nd}} D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \big\langle \Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}), D_{\mathbf{y}_i} u(s,\mathbf{y}) \big\rangle d\mathbf{y} \Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \leq \Lambda \|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} (s-t)^{-\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})+\frac{\gamma}{2}}. \tag{4.6}$$

Proof of Lemma 11. Let us first emphasize that from the Hölder continuity assumption **(S)**-(iii) (w.r.t. the underlying homogeneous metric $\mathbf{d}$) on $\mathbf{F}_i$:

$$|\Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y})| \leq [\mathbf{F}_i(s,\cdot)]_{\mathbf{d},2i-3+\gamma} \mathbf{d}^{2i-3+\gamma}(\mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})). \tag{4.7}$$

Similarly to (2.19),we take for each $i \in [\![2, n]\!]$,

$$D_{\mathbf{y}_i} \cdot \big(\Theta^{\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y}) \big) = D_{\mathbf{y}_i} \cdot \big(D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \otimes \Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}) \big). \tag{4.8}$$

The contribution of the l.h.s. in (4.6) then rewrites:

$$\sum_{i=2}^n \Big| \int_{\mathbb{R}^{nd}} D^{\vartheta}_{\mathbf{x}} \tilde{p}^{(t,\boldsymbol{\xi})}(t,s,\mathbf{x},\mathbf{y}) \big\langle \Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}), D_{\mathbf{y}_i} u(s,\mathbf{y}) \big\rangle d\mathbf{y} \Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} = \sum_{i=2}^n \Big| \int_{\mathbb{R}^{nd}} \big[D_{\mathbf{y}_i} \cdot \big(\Theta^{\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y}) \big) \big] u(s,\mathbf{y}) d\mathbf{y} \Big|. \tag{4.9}$$

The point now is to observe that for any fixed $i \in [\![2, n]\!]$ and $\mathbf{z} = (\mathbf{z}_1, \cdots, \mathbf{z}_{i-1}, \mathbf{z}_{i+1}, \cdots, \mathbf{z}_n) \in \mathbb{R}^{(n-1)d}$ the mapping $\mathbf{y}_i \mapsto u(s, \mathbf{z}_{1:i-1}, \mathbf{y}_i, \mathbf{z}_{i+1:n})$ is in $C_b^{\frac{2+\gamma}{2i-1}}(\mathbb{R}^d) = B^{\frac{2+\gamma}{2i-1}}_{\infty,\infty}(\mathbb{R}^d)$ using the Besov space terminology, see e.g.

Triebel [31], uniformly in $s \in [0,T]$. We can hence put in duality the mappings $\mathbf{y}_i \mapsto u(s,\mathbf{y}_{1:i-1},\mathbf{y}_i,\mathbf{y}_{i+1:n})$ and

$$\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})} : \mathbf{y}_i \mapsto D_{\mathbf{y}_i}\cdot\big(\Theta^{\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y})\big), \tag{4.10}$$

see e.g. Proposition 3.6 in [20]. To do so, we thus have to prove that $\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}$ lies in the suitable Besov space, namely $\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})} \in B_{1,1}^{-\frac{2+\gamma}{2i-1}}(\mathbb{R}^d)$ and to control the associated norm. We will actually prove that those norms provide an integrable quantity w.r.t. $\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}$ as well as an integrable time singularity. This will be done through the thermic characterization of Besov spaces, see e.g. Section 2.6.4 in [31] as well as (2.21) above. Precisely, we recall that for a function $\psi:\mathbb{R}^d\to\mathbb{R}$ in $B_{1,1}^{-\tilde{\alpha}_i}(\mathbb{R}^d)$, $\tilde{\alpha}_i := \frac{2+\gamma}{2i-1}$ a quasi-norm is given by:

$$\|\psi\|_{B_{1,1}^{-\tilde{\alpha}_i}(\mathbb{R}^d)} := \|\varphi(D)\psi\|_{L^1(\mathbb{R}^d,\mathbb{R})} + \int_0^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v\star\psi\|_{L^1(\mathbb{R}^d,\mathbb{R})},\ \forall z\in\mathbb{R}^d,\ h_v(z) := \frac{1}{(2\pi v)^{\frac{d}{2}}}\exp\big(-\frac{|z|^2}{v}\big), \tag{4.11}$$

h being the usual *heat kernel* of $\mathbb{R}^d$, "$\star$" standing for the usual convolution on $\mathbb{R}^d$ for $\varphi\in C_0^\infty(\mathbb{R}^d,\mathbb{R})$ s.t. $\varphi(0)\neq 0$.

Taking $\psi = \Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}$ in the above characterization and from definition (4.10), the main advantage of using (4.11) consists in rebalancing the derivative appearing in the definition (4.10) to the heat kernel or to the smooth compactly supported function φ. Namely, focusing on the L^1 norm of the convolution product in (4.11), we write:

$$\begin{aligned}
\|h_v\star\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} &= \int_{\mathbb{R}^d}\Big|\int_{\mathbb{R}^d} h_v(z-\mathbf{y}_i)D_{\mathbf{y}_i}\cdot\big(\Theta^{\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y})\big)d\mathbf{y}_i\Big|dz \\
&= \int_{\mathbb{R}^d}\Big|\int_{\mathbb{R}^d} D^{\vartheta}_{\mathbf{x}}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})\big\langle\Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}),D_z h_v(z-\mathbf{y}_i)\big\rangle d\mathbf{y}_i\Big|dz.
\end{aligned} \tag{4.12}$$

To estimate $\|\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{B_{1,1}^{-\tilde{\alpha}_i}}$, we split the time integral in (4.11) into two parts writing:

$$\begin{aligned}
&\int_0^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v\star\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
=\ &\int_0^{(s-t)^{\beta_i}} \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v\star\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} + \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v\star\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})}
\end{aligned} \tag{4.13}$$

for a parameter $\beta_i>0$ to be specified. Precisely, in order to have a similar smoothing effect in time than for the terms appearing in (4.5), we now want to calibrate β_i to obtain:

$$\int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v\star\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \le \frac{\Lambda}{(s-t)^{\sum_{j=1}\vartheta_j(j-\frac12)-\frac{\gamma}{2}}}\hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})), \tag{4.14}$$

where introducing:

$$\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y}) =: \prod_{j=1}^n \mathcal{N}_{c(s-t)^{2j-1}}\big((\boldsymbol{\theta}_{s,t}(\mathbf{x})-\mathbf{y})_j\big),$$

where for $\varsigma>0$, $z\in\mathbb{R}^d$, $\mathcal{N}_\varsigma(z) = \frac{1}{(2\pi\varsigma)^{\frac{d}{2}}}\exp\big(-\frac{|z|^2}{2\varsigma}\big)$ stands for the standard Gaussian density of $\mathbb{R}^d$ with covariance matrix $\varsigma\mathbf{I}_{d,d}$, we introduce:

$$\hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) = \prod_{j\in[\![1,n]\!],j\neq i} \mathcal{N}_{c(s-t)^{2j-1}}\big((\boldsymbol{\theta}_{s,t}(\mathbf{x})-\mathbf{y})_j\big). \tag{4.15}$$

To choose properly the parameter β_i leading to (4.14), we now write from (4.12):

$$\begin{aligned}
&\int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \big\| h_v \star \Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})} \big\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
\leq\ & \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \Big| \int_{\mathbb{R}^d} D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \big\langle \Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}), D_z h_v(z-\mathbf{y}_i) \big\rangle d\mathbf{y}_i \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & \Lambda \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} \mathbf{d}^{2i-3+\gamma}(\boldsymbol{\theta}_{s,t}(\mathbf{x}),\mathbf{y}) \\
\leq\ & \Lambda \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} (s-t)^{\frac{2i-3+\gamma}{2}} \\
\leq\ & \Lambda \hat{q}_{c\setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \int_{(s-t)^{\beta_i}}^1 dv v^{-\frac{3}{2}+\frac{\tilde{\alpha}_i}{2}} (s-t)^{-\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})+\frac{2i-3+\gamma}{2}} \\
\leq\ & \Lambda \hat{q}_{c\setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) (s-t)^{[-\frac{1}{2}+\frac{\tilde{\alpha}_i}{2}]\beta_i - \sum_{j=1}^n \vartheta_j (j-\frac{1}{2})+\frac{2i-3+\gamma}{2}}
\end{aligned}$$

using (4.7), Proposition 3 for the third inequality (see also Remark 5, equation (3.8)) recalling as well that $0 \leq t < s \leq T$ is small.

To obtain (4.14), we then take:

$$[-\frac{1}{2}+\frac{\tilde{\alpha}_i}{2}]\beta_i - \sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) + \frac{2i-3+\gamma}{2} = -\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) + \frac{\gamma}{2} \iff \beta_i = \frac{(2i-3)(2i-1)}{2i-3-\gamma}. \tag{4.16}$$

The key point is now to check that the previous choice of β_i also yields a bound similar to (4.14) for the contribution in (4.13) associated with $v \in [0,(s-t)^{\beta_i}]$. To this end, we restart from identity (4.12), which allows to exploit partial cancellations w.r.t. the integration variable $\mathbf{y}_i$. Namely, write:

$$\begin{aligned}
&\int_{\mathbb{R}^d} h_v(z-\mathbf{y}_i) D_{\mathbf{y}_i} \cdot \Theta^{\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y}) d\mathbf{y}_i \\
=\ & \int_{\mathbb{R}^d} h_v(z-\mathbf{y}_i) D_{\mathbf{y}_i} \cdot \Big(\Theta^{\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y}) - \Theta^{\vartheta}_{i,(t,\mathbf{x})}(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}) \Big) d\mathbf{y}_i \\
=\ & \int_{\mathbb{R}^d} D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \big\langle \mathbf{F}_i(s,\mathbf{y}) - \mathbf{F}_i(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}), D_z h_v(z-\mathbf{y}_i) \big\rangle d\mathbf{y}_i \\
&+ \int_{\mathbb{R}^d} \big(D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) - D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}) \big) \\
&\big\langle \mathbf{F}_i(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}) - \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}, D_z h_v(z-\mathbf{y}_i) \big\rangle d\mathbf{y}_i \\
=:\ & \Big(\mathscr{T}_1 + \mathscr{T}_2 \Big)\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big),
\end{aligned} \tag{4.17}$$

using the definition in (4.8) and (4.12) for the last decomposition. Write now from Proposition 3 and the Hölder regularity assumed on $\mathbf{F}_i$ from **(S)**-(iii):

$$\begin{aligned}
|\mathscr{T}_1\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)| &\leq \Lambda \int_{\mathbb{R}^d} \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} |z-\mathbf{y}_i|^{\frac{2i-3+\gamma}{2i-1}} d\mathbf{y}_i \\
&\leq \Lambda \int_{\mathbb{R}^d} \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{2-\gamma}{4i-2}}} \frac{\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} d\mathbf{y}_i.
\end{aligned} \tag{4.18}$$

We thus derive from (4.18):

$$\begin{aligned}
&|\mathscr{T}_1\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)| \\
\leq\ & \frac{\Lambda}{v^{\frac{2-\gamma}{4i-2}} (s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} \hat{q}_{c\setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \mathcal{N}_{cv+(s-t)^{2i-1}}\big(z-\boldsymbol{\theta}_{s,t}(\mathbf{x})_i\big).
\end{aligned} \tag{4.19}$$

We now deal with the term $|\mathscr{T}_2\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)|$ in (4.17). From the Taylor formula applied to the i^{th} variable for the difference of the derivatives of the densities we obtain:

$$
\begin{aligned}
&|\mathscr{T}_2\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)|\\
\leq\;& C\int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}}\int_0^1 d\mu \frac{\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-1}{2}}}\\
&\times|\mathbf{y}_i-z|\bigg(\Big|\mathbf{F}_i(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})-\mathbf{F}_i(s,\mathbf{y}_{1:i-1},\boldsymbol{\theta}_{s,t}(\mathbf{x})_{i:n})\Big|\\
&+\Big|\mathbf{F}_i(s,\mathbf{y}_{1:i-1},\boldsymbol{\theta}_{s,t}(\mathbf{x})_{i:n})-\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))-D_{\mathbf{x}_{i-1}}\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{x}))(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1}\Big|\bigg)\\
\leq\;& \Lambda\int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\int_0^1 d\mu \frac{\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-1}{2}}}\\
&\times\bigg(|z-\boldsymbol{\theta}_{s,t}(\mathbf{x})_i|^{\frac{2i-3+\gamma}{2i-1}}+|(\boldsymbol{\theta}_{s,t}(\mathbf{x})-\mathbf{y})_{i-1}|^{1+\frac{\gamma}{2(i-1)-1}}+\sum_{k=i+1}^n|(\boldsymbol{\theta}_{s,t}(\mathbf{x})-\mathbf{y})_k|^{\frac{2i-3+\gamma}{2k-1}}\bigg).
\end{aligned}
$$

Writing, for any $\mu\in[0,1]$,

$$
|z-\boldsymbol{\theta}_{s,t}(\mathbf{x})_i|\leq\mu|z-\mathbf{y}_i|+|z+\mu(\mathbf{y}_i-z)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|,
$$

we thus derive

$$
\begin{aligned}
&|\mathscr{T}_2\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)|\\
\leq\;& \Lambda\int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\int_0^1 d\mu \frac{\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-1}{2}}}\\
&\times\Big(|\mathbf{y}_i-z|^{\frac{2i-3+\gamma}{2i-1}}+\mathbf{d}^{2i-3+\gamma}\Big(\boldsymbol{\theta}_{s,t}(\mathbf{x}),(\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})\Big)\Big)\\
\leq\;& \Lambda\int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\int_0^1 d\mu\bar{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})\\
&\times\bigg(\frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}}\bigg)\\
\leq\;& \Lambda\hat{q}_{c\backslash i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})\int_0^1 d\mu\int_{\mathbb{R}^d} h_{cv}(z-\mathbf{y}_i)\mathcal{N}_{c(s-t)^{2i-1}}(z+\mu(\mathbf{y}_i-z)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)d\mathbf{y}_i\\
&\times\bigg(\frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}}\bigg),
\end{aligned}
\tag{4.20}
$$

using again (3.8) for the second inequality. From (4.17), (4.19) and (4.20) we derive, with the notation introduced in (4.15):

$$
\begin{aligned}
&\|h_v\star\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1})}\|_{L^1(\mathbb{R}^d,\mathbb{R})}\\
\leq\;&\bigg(\frac{1}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})}}+\frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}}\bigg)\\
&\times\Lambda\hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))\int_0^1 d\mu\int_{\mathbb{R}^d} dz\int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\mathcal{N}_{c(s-t)^{2i-1}}(z+\mu(\mathbf{y}_i-z)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)\\
\leq\;&\Lambda\hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))\\
&\times\bigg(\frac{1}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})}}+\frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}}\bigg),
\end{aligned}
$$

using the change of variable $(w_1, w_2) = (z - \mathbf{y}_i, z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)$ for the last inequality. From the above computations and with the notations of (4.13), we derive:

$$\begin{aligned}
&\int_0^{(s-t)^{\beta_i}} dv v^{\frac{\tilde{\alpha}_i}{2}-1} \|h_v \star \Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
\leq\ & \Lambda \hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \int_0^{(s-t)^{\beta_i}} \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \\
&\times \left(\frac{1}{v^{\frac{2-\gamma}{4i-2}} (s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} + \frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \right) \\
=:\ & \Lambda \hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \mathcal{B}_{\vartheta,\beta_i}(t,s).
\end{aligned}$$

Let us now prove that for $\beta_i = \frac{(2i-3)(2i-1)}{2i-3-\gamma}$ defined in (4.16), we have:

$$\mathcal{B}_{\vartheta,\beta_i}(t,s) \leq \frac{C}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) - \frac{\gamma}{2}}}. \tag{4.21}$$

To prove (4.21), we now write:

$$\begin{aligned}
&\mathcal{B}_{\vartheta,\beta_i}(t,s) \\
\leq\ & C \Big[\frac{v^{\frac{\tilde{\alpha}_i}{2} - \frac{2-\gamma}{4i-2}}}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} + \frac{v^{\frac{\tilde{\alpha}_i}{2} + \frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{v^{\frac{\tilde{\alpha}_i}{2}}}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \Big]_{v=0}^{v=(s-t)^{\beta_i}} \\
\leq\ & C \Big[(s-t)^{\beta_i (\frac{\tilde{\alpha}_i}{2} - \frac{2-\gamma}{4i-2}) - \sum_{j=1}^n \vartheta_j (j-\frac{1}{2})} + (s-t)^{\beta_i (\frac{\tilde{\alpha}_i}{2} + \frac{2i-3+\gamma}{2(2i-1)}) - (\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) + \frac{2i-1}{2})} \\
&+ (s-t)^{\beta_i \frac{\tilde{\alpha}_i}{2} - \sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) - 1 + \frac{\gamma}{2}} \Big].
\end{aligned}$$

From the above equation, (4.21) holds as soon as β_i can be chosen so that the three following conditions hold:

$$\beta_i \left(\frac{\tilde{\alpha}_i}{2} - \frac{2-\gamma}{4i-2} \right) - \frac{\gamma}{2} \geq 0, \quad \beta_i \left(\frac{\tilde{\alpha}_i}{2} + \frac{2i-3+\gamma}{2(2i-1)} \right) - \frac{2i-1}{2} - \frac{\gamma}{2} \geq 0, \quad \beta_i \frac{\tilde{\alpha}_i}{2} - 1 \geq 0.$$

Recalling that $\frac{\tilde{\alpha}_i}{2} = \frac{1+\frac{\gamma}{2}}{2i-1}$ and for the previous choice of β_i, the above conditions rewrite:

$$\begin{aligned}
\left(\frac{(2i-3)(2i-1)}{2i-3-\gamma} \right) \left(\frac{1+\frac{\gamma}{2}}{2i-1} - \frac{1-\frac{\gamma}{2}}{2i-1} \right) - \frac{\gamma}{2} \geq 0 &\iff \frac{(2i-3)}{2i-3-\gamma}\gamma - \frac{\gamma}{2} \geq 0, \\
\left(\frac{(2i-3)(2i-1)}{2i-3-\gamma} \right) \left(\frac{2+\gamma}{2(2i-1)} + \frac{2i-3+\gamma}{2(2i-1)} \right) - \frac{2i-1}{2} - \frac{\gamma}{2} \geq 0 & \\
\iff \left(\frac{2i-3}{2i-3-\gamma} \right) (2i-1+2\gamma) - (2i-1+\gamma) \geq 0, & \\
\left(\frac{(2i-3)(2i-1)}{2i-3-\gamma} \right) \frac{1+\frac{\gamma}{2}}{2i-1} - 1 \geq 0 \iff \left(\frac{2i-3}{2i-3-\gamma} \right) (1+\frac{\gamma}{2}) - 1 \geq 0. &
\end{aligned}$$

All the above conditions are true for $i \in [\![2,n]\!]$, $\gamma \in (0,1]$. Note that the chosen β_i seems to be rather sharp in the sense that letting γ go to 0 the above constraints become equalities. This proves (4.21). We finally get:

$$\int_0^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \leq \frac{\Lambda}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) - \frac{\gamma}{2}}} \hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})). \tag{4.22}$$

Reproducing the previous computations we also write for a C^∞ compactly supported function φ:

$$\begin{aligned}
&\|\varphi(D) \Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
\leq\ & \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} D_{\mathbf{y}_i} \varphi^\vee(z-\mathbf{y}_i) \cdot \big(D^{\vartheta}_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y}) \otimes \Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}) \rangle d\mathbf{y}_i \Big| dz \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & \frac{\Lambda}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})}} \int_{\mathbb{R}^d} \overline{p}_{c^{-1}}(t,s,\mathbf{x},\mathbf{y}) \mathbf{d}^{2i-3+\gamma}(\boldsymbol{\theta}_{s,t}(\mathbf{x}),\mathbf{y}) d\mathbf{y}_i \\
\leq\ & \frac{\Lambda}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) - (i-\frac{3}{2}+\frac{\gamma}{2})}} \hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})).
\end{aligned}$$

From (4.11) and (4.22), we finally obtain:

$$\|\Psi^{\vartheta}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{B^{-\tilde{\alpha}_i}_{1,1}} \leq \frac{\Lambda}{(s-t)^{\sum_{j=1}^n \vartheta_j (j-\frac{1}{2}) - \frac{\gamma}{2}}} \hat{q}_{c\backslash i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})), \tag{4.23}$$

which together with (4.9) and (4.10) gives the result. □

Equation (2.26) now follows from Lemma 11 taking $\vartheta = (2, 0, \dots, 0)$. Namely,

$$\Big|\sum_{i=2}^{n}\int_t^T ds\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})\big\langle\Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}), D_{\mathbf{y}_i}u(s,\mathbf{y})\big\rangle\Big)\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \le \Lambda(T-t)^{\frac{\gamma}{2}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}. \tag{4.24}$$

4.3. **Non-degenerate derivatives for the frozen semi-group : terminal condition and source.** The main result of this section is the following lemma.

Lemma 12 (Derivatives of frozen semi-group and Green kernel)**.** *There exists a constant $C := C((\mathbf{A}))$ s.t. for any $(t,\mathbf{x}) \in [0,T]\times\mathbb{R}^{nd}$,*

$$\begin{aligned}|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x})|\Big|_{\boldsymbol{\xi}=\mathbf{x}} &\le C\|D^2_{\mathbf{x}_1}g\|_{L^\infty} \le C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}},\\ |D^2_{\mathbf{x}_1}\tilde{G}^{\boldsymbol{\xi}}f(t,\mathbf{x})|\Big|_{\boldsymbol{\xi}=\mathbf{x}} &\le C(T-t)^{\frac{\gamma}{2}}\|f\|_{L^\infty(C^{\gamma}_{b,\mathbf{d}})}.\end{aligned}$$

Proof of Lemma 12. Note first that,

$$\begin{aligned}\big|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x})\big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} &= \Big|\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[g(\mathbf{y}) - g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))]d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ &\le \Big|\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[g(\mathbf{y}) - g(\mathbf{y}_1,\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})_{2:n})]d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ &\quad + \Big|\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[g(\mathbf{y}_1,(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_{2:n}) - g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))]d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}.\end{aligned} \tag{4.25}$$

The first term in the r.h.s. of the previous identity is readily controlled thanks to Proposition 3

$$\begin{aligned}&\Big|\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[g(\mathbf{y}) - g(\mathbf{y}_1,(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_{2:n})]d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ &\le C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\int_{\mathbb{R}^{nd}}(T-t)^{-1}\bar{p}_{C^{-1}}(t,T,\mathbf{x},\mathbf{y})\mathbf{d}^{2+\gamma}(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}),\mathbf{y})d\mathbf{y}\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ &\le C(T-t)^{\frac{\gamma}{2}}\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}.\end{aligned} \tag{4.26}$$

The second term of (4.25) is more subtle. We need to expand $g(\mathbf{y}_1,(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_{2:n})$ in its non-degenerate variable to take advantage of the corresponding regularity of g. Namely, recalling from Proposition 5 that

$$\int_{\mathbb{R}^d} d\mathbf{y} D^2_{\mathbf{x}_1}\tilde{p}^{m,\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})\big\langle D_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})),(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_1\big\rangle\Big|_{\boldsymbol{\xi}=\mathbf{x}} = \mathbf{0}_{d,d},$$

we obtain

$$\begin{aligned}&\Big|\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[g(\mathbf{y}_1,(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_{2:n}) - g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))]d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ =\ &\Big|\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})\Big(\big\langle D_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})),(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_1\big\rangle\\ &+\int_0^1 d\mu(1-\mu)\mathrm{Tr}\Big(\mathrm{D}^2_{\mathbf{x}_1}\mathrm{g}\big(\mathbf{m}^{\boldsymbol{\xi}}_{\mathrm{T,t}}(\mathbf{x})_1+\mu(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{\mathrm{T,t}}(\mathbf{x}))_1,(\mathbf{m}^{\boldsymbol{\xi}}_{\mathrm{T,t}}(\mathbf{x}))_{2:\mathrm{n}}\big)\big(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{\mathrm{T,t}}(\mathbf{x})\big)^{\otimes 2}_1\Big)\Big)\mathrm{d}\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ \le\ &C\|D^2_{\mathbf{x}_1}g\|_{L^\infty}\int_{\mathbb{R}^{nd}}\frac{\bar{p}_{C^{-1}}(t,T,\mathbf{x},\mathbf{y})}{T-t}|(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_1|^2 d\mathbf{y}\Big|_{\boldsymbol{\xi}=x}\\ \le\ &C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}.\end{aligned} \tag{4.27}$$

Gathering identities (4.26), (4.27) into (4.25), we obtain the stated control for $|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x})|\Big|_{\boldsymbol{\xi}=\mathbf{x}}$.

Let us now turn to the Green kernel. We directly get from Proposition 3:

$$\begin{aligned}|D^2_{\mathbf{x}_1}\tilde{G}^{\boldsymbol{\xi}}f(t,\mathbf{x})|\Big|_{\boldsymbol{\xi}=\mathbf{x}} &\le \Big|\int_t^T ds\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})[f(s,\mathbf{y}) - f(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))]d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ &\le C\|f\|_{L^\infty(C^{\gamma}_{\mathbf{d}})}\int_t^T ds\int_{\mathbb{R}^{nd}}\frac{1}{s-t}\bar{p}_{C^{-1}}(s,t,\mathbf{x},\mathbf{y})\mathbf{d}(\boldsymbol{\theta}_{s,t}(\mathbf{x}),\mathbf{y})^{\gamma}d\mathbf{y}\\ &\le C\|f\|_{L^\infty(C^{\gamma}_{\mathbf{d}})}(T-t)^{\frac{\gamma}{2}},\end{aligned}$$

which gives the result.

□

5. Hölder controls

In this section, we aim at giving suitable controls on the Hölder moduli $[D^2_{\mathbf{x}_1}u(t,\cdot)]_{\gamma,\mathbf{d}}$ and $\sup_{z\in\mathbb{R}^d}[u(t,z,\cdot)]_{2+\gamma,\mathbf{d}}$ in order to derive our main Schauder estimate of Theorem 1. Namely, we want to establish the following result.

Proposition 13. *Let $\gamma\in(0,1)$ be given. Suppose that* **(A)** *is in force and that the terminal condition g and source term f of the Cauchy problem* (1.1) *satisfy: $g\in C^{2+\gamma}_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})$ and $f\in L^\infty\big([0,T],C^\gamma_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})\big)$. Then, there exists $C:=C((\mathbf{A}),T)$ and $\Lambda:=\Lambda((\mathbf{A}))$ as in Remark 3 such that for any $c_0\in(0,1]$:*

(5.1)

$$[D^2_{\mathbf{x}_1}u(t,\cdot)]_{\gamma,\mathbf{d}}+\sup_{z\in\mathbb{R}^d}[u(t,z,\cdot)]_{2+\gamma,\mathbf{d}}\le C\Big\{\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}+\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\Big\}+\Big(\Lambda(c_0^{-(n-\frac12)+\frac\gamma2}+c_0^{\frac\gamma2})+Cc_0^{\frac{\gamma}{2n-1}}\Big)\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}.$$

To prove this result the point is to consider for a fixed time $t\in[0,T]$ and for fixed $\mathbf{x}\in\mathbb{R}^{nd}$ the perturbative expansion (2.11) and for another spatial point $\mathbf{x}'\in\mathbb{R}^{nd}$, the possibly more refined version provided by equation (2.39) which precisely allows to take into account the various regimes depending on $\mathbf{d}(\mathbf{x},\mathbf{x}')$ and $(s-t)^{1/2}$, $s\in[t,T]$ detailed in the previous detailed guide to the proof (see Section 2 paragraph 2.4).

Hence, we will address separately two cases. For a constant c_0 to be specified later on (but formally meant to be *small*), we consider:

- The globally *off-diagonal* regime $T-t<c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$. In that case, the spatial points $\mathbf{x},\mathbf{x}'$ are globally *far* for the corresponding homogeneous distance $\mathbf{d}$ over the time horizon $s\in[t,T]$. Hence, there is no specific need to exploit (2.39). Expanding the quantities $D^2_{\mathbf{x}_1}u(t,\mathbf{x})-D^2_{\mathbf{x}_1}u(t,\mathbf{x}')$ and $u(t,z,\mathbf{x}_{2:n})-u(t,z,\mathbf{x}'_{2:n})$ with (2.11) is enough to get the result.

Indeed, writing from (2.11) the for any $(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2$ s.t. $(T-t)<c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$,

$$\begin{aligned}
&D^2_{\mathbf{x}_1}u(t,\mathbf{x})-D^2_{\mathbf{x}_1}u(t,\mathbf{x}')\\
=\;&\Big\{D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi}_{T,t}g(\mathbf{x})-D^2_{\mathbf{x}'_1}\tilde P^{\boldsymbol\xi'}_{T,t}g(\mathbf{x}')+D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi}f(t,\mathbf{x})-D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi'}f(t,\mathbf{x}')\\
&+\int_t^T ds\int_{\mathbb{R}^{nd}}\Big(D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol\xi}(t,s,\mathbf{x},\mathbf{y})(L_s-\tilde L^{\boldsymbol\xi}_s)u(s,\mathbf{y})-D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol\xi'}(t,s,\mathbf{x}',\mathbf{y})(L_s-\tilde L^{\boldsymbol\xi'}_s)u(s,\mathbf{y})\Big)d\mathbf{y}\Big\}\Big|_{(\boldsymbol\xi,\boldsymbol\xi')=(\mathbf{x},\mathbf{x}')},
\end{aligned}\tag{5.2}$$

and similarly, for any $z\in\mathbb{R}^d$,

$$\begin{aligned}
&u\big(t,(z,\mathbf{x}_{2:n})\big)-u\big(t,(z,\mathbf{x}'_{2:n})\big)\\
=\;&\Big\{\tilde P^{\boldsymbol\xi}_{T,t}g(z,\mathbf{x}_{2:n})-\tilde P^{\boldsymbol\xi'}_{T,t}g(z,\mathbf{x}'_{2:n})+\tilde G^{\boldsymbol\xi}f\big(t,(z,\mathbf{x}_{2:n})\big)-\tilde G^{\boldsymbol\xi'}f\big(t,(z,\mathbf{x}'_{2:n})\big)\\
&+\int_t^T ds\int_{\mathbb{R}^{nd}}\Big(\tilde p^{\boldsymbol\xi}(t,s,(z,\mathbf{x}_{2:n}),\mathbf{y})(L_s-\tilde L^{\boldsymbol\xi}_s)u(s,\mathbf{y})\\
&\qquad-\tilde p^{\boldsymbol\xi'}(t,s,(z,\mathbf{x}'_{2:n}),\mathbf{y})(L_s-\tilde L^{\boldsymbol\xi'}_s)u(s,\mathbf{y})\Big)d\mathbf{y}\Big\}\Big|_{(\boldsymbol\xi,\boldsymbol\xi')=\big((z,\mathbf{x}_{2:n}),(z,\mathbf{x}'_{2:n})\big)}.
\end{aligned}\tag{5.3}$$

In this case, we derive from Lemmas 10, 11, 12 and equation (4.2) that:

$$\begin{aligned}
&|D^2_{\mathbf{x}_1}u(t,\mathbf{x})-D^2_{\mathbf{x}_1}u(t,\mathbf{x}')|\\
\le\;&C\Big[|D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi}_{T,t}g(\mathbf{x})-D^2_{\mathbf{x}'_1}\tilde P^{\boldsymbol\xi'}_{T,t}g(\mathbf{x}')|\big|_{(\boldsymbol\xi,\boldsymbol\xi')=(\mathbf{x},\mathbf{x}')}+\big(\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}+\Lambda\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\big)(T-t)^{\frac\gamma2}\Big]\\
\le\;&C\Big[|D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi}_{T,t}g(\mathbf{x})-D^2_{\mathbf{x}'_1}\tilde P^{\boldsymbol\xi'}_{T,t}g(\mathbf{x}')|\big|_{(\boldsymbol\xi,\boldsymbol\xi')=(\mathbf{x},\mathbf{x}')}+\big(\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}+\Lambda\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\big)c_0^{\frac\gamma2}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\Big],
\end{aligned}$$

and

$$\begin{aligned}
&|u\big(t,(z,\mathbf{x}_{2:n})\big)-u\big(t,(z,\mathbf{x}'_{2:n})\big)|\\
\le\;&C\Big(|\tilde P^{\boldsymbol\xi}_{T,t}g(z,\mathbf{x}_{2:n})-\tilde P^{\boldsymbol\xi'}_{T,t}g(z,\mathbf{x}'_{2:n})|\big|_{(\boldsymbol\xi,\boldsymbol\xi')=\big((z,\mathbf{x}_{2:n}),(z,\mathbf{x}'_{2:n})\big)}\\
&+|\tilde G^{\boldsymbol\xi}f\big(t,(z,\mathbf{x}_{2:n})\big)-\tilde G^{\boldsymbol\xi'}f\big(t,(z,\mathbf{x}'_{2:n})\big)|\big|_{(\boldsymbol\xi,\boldsymbol\xi')=\big((z,\mathbf{x}_{2:n}),(z,\mathbf{x}'_{2:n})\big)}+\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}c_0^{\frac\gamma2}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\Big).
\end{aligned}$$

Therefore, these equations give the expected controls up to appropriate estimates for the Hölder moduli of the frozen semigroup and Green kernel (which are obtained below, for the so-called *mixed* regime).

- The *mixed* regime $T-t \geq c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$. In that case, up to the *transition time* t_0 defined in (2.40), $s-t < c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$, i.e. the *off-diagonal* regime holds for the times s *close* to t. Things are hence more involved. In particular, it is then crucial to exploit the more refined perturbative expansion (2.39) to derive suitable bounds for $D^2_{\mathbf{x}_1}u(t,\mathbf{x}) - D^2_{\mathbf{x}_1}u(t,\mathbf{x}')$ and $u\big(t,(z,\mathbf{x}_{2:n})\big) - u\big(t,(z,\mathbf{x}'_{2:n})\big)$. In particular, this leads to handle carefully the additional terms appearing from the change of freezing parameter.

To handle our controls in such a case, we will start from identity (2.39) for the expansion of $u(t,\mathbf{x}')$, where we have chosen $\boldsymbol{\xi}' = \mathbf{x}'$ and $\tilde{\boldsymbol{\xi}}' = \mathbf{x}$. Namely,

$$\begin{aligned} u(t,\mathbf{x}') &= \tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{T,t}g(\mathbf{x}') + \tilde{G}^{\boldsymbol{\xi}'}_{t_0,t}f(t,\mathbf{x}') + \tilde{G}^{\tilde{\boldsymbol{\xi}}'}_{T,t_0}f(t,\mathbf{x}') + \tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u(t_0,\mathbf{x}') - \tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u(t_0,\mathbf{x}') \\ &\quad + \int_t^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\Big(\mathbb{I}_{s\leq t_0}\tilde{p}^{\boldsymbol{\xi}'}(t,s,\mathbf{x}',\mathbf{y})(L_s - \tilde{L}^{\boldsymbol{\xi}'}_s) + \mathbb{I}_{s>t_0}p^{\tilde{\boldsymbol{\xi}}'}(t,s,\mathbf{x}',\mathbf{y})(L_s - \tilde{L}^{\tilde{\boldsymbol{\xi}}'}_s)\Big)u(s,\mathbf{y}). \end{aligned}$$

Again, t_0 must be here seen as a frozen parameter, which is *a posteriori*, i.e. after possible differentiation, chosen as in (2.40).

According to the notations of the detailed guide to the proof (see eq. (2.31) and (2.30) in Section 2), the terms to control then write for the Hölder norm of the derivatives w.r.t. the non-degenerate variables:

$$\begin{aligned} &D^2_{\mathbf{x}_1}u(t,\mathbf{x}) - D^2_{\mathbf{x}_1}u(t,\mathbf{x}') \\ &= \Big\{\Big(D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x}) - D^2_{\mathbf{x}'_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x}')\Big) + \Big(D^2_{\mathbf{x}_1}\tilde{G}^{\boldsymbol{\xi}}_{t_0,t}f(t,\mathbf{x}) - D^2_{\mathbf{x}_1}\tilde{G}^{\boldsymbol{\xi}'}_{t_0,t}f(t,\mathbf{x}')\Big) \\ &\quad + \Big(D^2_{\mathbf{x}_1}\tilde{G}^{\boldsymbol{\xi}}_{T,t_0}f(t,\mathbf{x}) - D^2_{\mathbf{x}_1}\tilde{G}^{\boldsymbol{\xi}}_{t_0,t}f(t,\mathbf{x}')\Big) + \Big(D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u(t_0,\mathbf{x}') - D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{t_0,t}u(t_0,\mathbf{x}')\Big) \\ &\quad + D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}') + D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}')\Big\}\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x}',\mathbf{x})}, \end{aligned} \tag{5.4}$$

with

$$\begin{aligned} \Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}') &= \int_t^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})\mathbb{I}_{s\leq t_0}(L_s - \tilde{L}^{\boldsymbol{\xi}}_s)u(s,\mathbf{y}) \\ &\quad - \int_t^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{\boldsymbol{\xi}'}(t,s,\mathbf{x}',\mathbf{y})\mathbb{I}_{s\leq t_0}(L_s - \tilde{L}^{\boldsymbol{\xi}'}_s)u(s,\mathbf{y}), \\ \Delta^{\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}') &= \int_t^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})\mathbb{I}_{s>t_0}(L_s - \tilde{L}^{\boldsymbol{\xi}}_s)u(s,\mathbf{y}) \\ &\quad - \int_t^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{\tilde{\boldsymbol{\xi}}'}(t,s,\mathbf{x}',\mathbf{y})\mathbb{I}_{s>t_0}(L_s - \tilde{L}^{\tilde{\boldsymbol{\xi}}'}_s)u(s,\mathbf{y}), \end{aligned} \tag{5.5}$$

and

$$\begin{aligned} &u\big(t,(z,\mathbf{x}_{2:n})\big) - u\big(t,(z,\mathbf{x}'_{2:n})\big) \\ &= \Big\{\Big(\tilde{P}^{\boldsymbol{\xi}}_{T,t}g\big(z,\mathbf{x}_{2:n}\big) - \tilde{P}^{\boldsymbol{\xi}}_{T,t}g\big(z,\mathbf{x}'_{2:n}\big)\Big) + \Big(\tilde{G}^{\boldsymbol{\xi}}_{t_0,t}f\big(t,(z,\mathbf{x}_{2:n})\big) - \tilde{G}^{\boldsymbol{\xi}'}_{t_0,t}f\big(t,(z,\mathbf{x}'_{2:n})\big)\Big) \\ &\quad + \Big(\tilde{G}^{\boldsymbol{\xi}}_{T,t_0}f\big(t,(z,\mathbf{x}_{2:n})\big) - \tilde{G}^{\boldsymbol{\xi}}_{t_0,t}f\big(t,(z,\mathbf{x}_{2:n})\big)\Big) + \Big(\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u\big(t_0,(z,\mathbf{x}'_{2:n})\big) - \tilde{P}^{\boldsymbol{\xi}}_{t_0,t}u\big(t_0,(z,\mathbf{x}'_{2:n})\big)\Big) \\ &\quad + \Delta^{\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}'}_{\mathbf{diag}}(t,T,(z,\mathbf{x}_{2:n}),(z,\mathbf{x}'_{2:n})) + \Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,(z,\mathbf{x}_{2:n}),(z,\mathbf{x}'_{2:n}))\Big\}\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=((z,\mathbf{x}_{2:n}),(z,\mathbf{x}'_{2:n}),(z,\mathbf{x}_{2:n}))}. \end{aligned} \tag{5.6}$$

for the Hölder moduli w.r.t. the degenerate variables according to the previously introduced notations.

To derive the expected bounds we will then devote a subsection to the Hölder controls for the frozen semi-group (see Lemma 14 in Section 5.1), for the frozen Green kernel (see Lemma 15 in Section 5.2), for the discontinuity term coming from the change of freezing point (see Lemma 17 in Section 5.4) and for the perturbative contribution (see Lemma 5.3 in Section 5.3). Aggregating the previously mentioned lemmas directly yields Proposition 13. □

5.1. **Hölder norms for the frozen semi-group.** We precisely want to establish the following result.

Lemma 14. *There exists $C := C((\mathbf{A}))$ s.t. for any $(t,\mathbf{x},\mathbf{x}') \in [0,T]\times\mathbb{R}^{nd}\times\mathbb{R}^{nd}$, taking*

$$(\boldsymbol{\xi},\boldsymbol{\xi}') = \begin{cases} (\mathbf{x},\mathbf{x}'), \text{ if } (T-t)^{1/2} < c_0\mathbf{d}(\mathbf{x},\mathbf{x}'), \\ (\mathbf{x},\mathbf{x}), \text{ if } (T-t)^{1/2} \geq c_0\mathbf{d}(\mathbf{x},\mathbf{x}'), \end{cases}$$

one has:

$$\begin{aligned} \left|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x}) - D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}'}_{T,t}g(\mathbf{x}')\right| &\leq C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{\gamma}(\mathbf{x},\mathbf{x}'), \\ \left|\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x}) - \tilde{P}^{\boldsymbol{\xi}'}_{T,t}g(\mathbf{x}')\right| &\leq C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'), \text{ for } \mathbf{x}_1 = \mathbf{x}'_1. \end{aligned}$$

Before entering into the proof of such a result, let us emphasize that, as suggested by the above off-diagonal and diagonal splitting, the constants C appearing in the r.h.s. of the above equations should depend on c_0. This is true, but since these terms are not planned to be passed in the r.h.s. of the final estimate, see (2.42) and the associated comments, we do not keep track of this dependence.

5.1.1. *Hölder norms of the derivatives w.r.t. the non-degenerate variables.* Let us deal with the first inequality of Lemma 14, i.e. the Hölder norms of the derivatives w.r.t. the non-degenerate variables $\mathbf{x}_1$. For the frozen semi-group, we say that the off-diagonal regime (resp. diagonal regime) holds when $T-t \leq c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$ (resp. $T-t \geq c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$).

- Off-diagonal regime. If $T-t \leq c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$, like in (4.25), we write:

$$\begin{aligned} & D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x}) - D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}'}_{T,t}g(\mathbf{x}') \\ = & \Big[\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[g(\mathbf{y}) - g(\mathbf{y}_1,(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_{2:n})]d\mathbf{y} \\ & - \int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,T,\mathbf{x}',\mathbf{y})[g(\mathbf{y}) - g(\mathbf{y}_1(,\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))_{2:n})]d\mathbf{y}\Big] \\ & + \Big[\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[g(\mathbf{y}_1,(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_{2:n}) - g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))]d\mathbf{y} \\ & - \int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,T,\mathbf{x}',\mathbf{y})[g(\mathbf{y}_1,(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))_{2:n}) - g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))]d\mathbf{y}\Big] \\ =: & \ \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}D^2_{\mathbf{x}_1}\tilde{P}_1 g(\mathbf{x},\mathbf{x}') + \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}D^2_{\mathbf{x}_1}\tilde{P}_2 g(\mathbf{x},\mathbf{x}'). \end{aligned} \tag{5.7}$$

The first term, which is associated with the degenerate variables, is controlled directly thanks to (4.26), which again readily follows from Proposition 3 (see as well Remark 3.8), for $(\boldsymbol{\xi},\boldsymbol{\xi}') = (\mathbf{x},\mathbf{x}')$. One hence gets:

$$\left|\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}D^2_{\mathbf{x}_1}\tilde{P}_1 g(\mathbf{x},\mathbf{x}')\right|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \leq 2C(T-t)^{\frac{\gamma}{2}}\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} \leq 2Cc_0^{\frac{\gamma}{2}}\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{\gamma}(\mathbf{x},\mathbf{x}'). \tag{5.8}$$

The second term is more delicate, we proceed like in (4.27):

$$\begin{aligned} & \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}D^2_{\mathbf{x}_1}\tilde{P}_2 g(\mathbf{x},\mathbf{x}') \\ = & \Big[\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})\int_0^1 d\mu(1-\mu) \\ & \mathrm{Tr}\Big(\big[D^2_{\mathbf{x}_1}g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})_1 + \mu(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_1, (\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_{2:n}\big) - D^2_{\mathbf{x}_1}g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})\big)\big]\big(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})\big)_1^{\otimes 2}\Big)d\mathbf{y} \\ - & \int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,T,\mathbf{x}',\mathbf{y})\int_0^1 d\mu(1-\mu) \\ & \mathrm{Tr}\Big(\big[D^2_{\mathbf{x}_1}g\big(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}')_1 + \mu(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))_1, (\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))_{2:n}\big) - D^2_{\mathbf{x}_1}g\big(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}')\big)\big]\big(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}')\big)_1^{\otimes 2}\Big)d\mathbf{y}\Big] \\ + & \Big[\frac{1}{2}\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_1^{\otimes 2}\Big)d\mathbf{y} \\ - & \frac{1}{2}\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,T,\mathbf{x}',\mathbf{y})\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))_1^{\otimes 2}d\mathbf{y}\Big] \\ =: & \ \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}D^2_{\mathbf{x}_1}\tilde{P}_{21}g(\mathbf{x},\mathbf{x}') + \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}D^2_{\mathbf{x}_1}\tilde{P}_{22}g(\mathbf{x},\mathbf{x}'). \end{aligned} \tag{5.9}$$

The first contribution of the previous identity is handled exploiting the smoothness of $D^2_{\mathbf{x}_1} g$ and Proposition 3. Namely,

$$
\begin{aligned}
|\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'} D^2_{\mathbf{x}_1} \tilde{P}_{21} g(\mathbf{x},\mathbf{x}')| \quad &\leq \quad C[D^2_{\mathbf{x}_1} g]^{\gamma}_{\mathbf{d}} \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(T-t)} \Big(\bar{p}^{\boldsymbol{\xi}}_{C^{-1}}(t,T,\mathbf{x},\mathbf{y}) |(\mathbf{y} - \mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_1|^{2+\gamma} \\
&\qquad + \bar{p}^{\boldsymbol{\xi}'}_{C^{-1}}(t,T,\mathbf{x}',\mathbf{y}) |(\mathbf{y} - \mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))_1|^{2+\gamma} \Big) \Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}.
\end{aligned}
$$

Hence,

$$
\Big| \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'} D^2_{\mathbf{x}_1} \tilde{P}_{21} g(\mathbf{x},\mathbf{x}') \Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \leq 2C(T-t)^{\frac{\gamma}{2}} \|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} \leq 2Cc_0^{\frac{\gamma}{2}} \|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} \mathbf{d}^{\gamma}(\mathbf{x},\mathbf{x}'). \tag{5.10}
$$

Let us now decompose the last contribution of (5.9):

$$
\begin{aligned}
&\Big| \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'} D^2_{\mathbf{x}_1} \tilde{P}_{22} g(\mathbf{x},\mathbf{x}') \Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \\
\leq \quad &\Big\{ \frac{1}{2} \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(s-t)} \bar{p}^{\boldsymbol{\xi}}_{C^{-1}}(t,T,\mathbf{x},\mathbf{y}) |D^2_{\mathbf{x}_1} g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})) - D^2_{\mathbf{x}_1} g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))| |(\mathbf{y} - \mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))_1|^2 \\
&+ \frac{1}{2} \Big| \int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y}) \mathrm{Tr}\Big(D^2_{\mathbf{x}_1} g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))(\mathbf{y} - \mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))^{\otimes 2}_1 \Big) \\
&- D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}'}(t,T,\mathbf{x}',\mathbf{y}) \mathrm{Tr}\Big(\langle D^2_{\mathbf{x}_1} g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))(\mathbf{y} - \mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))^{\otimes 2}_1 \Big) d\mathbf{y} \Big| \Big\} \Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \\
\leq \quad &C |D^2_{\mathbf{x}_1} g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})) - D^2_{\mathbf{x}_1} g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))| \big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \\
= \quad &C \big| \big(D^2_{\mathbf{x}_1} g\big)(\boldsymbol{\theta}_{T,t}(\mathbf{x})) - \big(D^2_{\mathbf{x}_1} g\big)(\boldsymbol{\theta}_{T,t}(\mathbf{x}')) \big|,
\end{aligned} \tag{5.11}
$$

exploiting Proposition 3 and equation (3.17) in Proposition 5 to observe that the second contribution of the first inequality above vanishes and recalling as well (2.15) for the last equality to identify the linearized flows, respectively frozen in $\boldsymbol{\xi} = \mathbf{x}$, $\boldsymbol{\xi}' = \mathbf{x}'$, with the initial non-linear ones.

From Lemma 7, we derive that for $T - t \leq c_0 \mathbf{d}^2(\mathbf{x},\mathbf{x}')$:

$$
\Big| \Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'} D^2_{\mathbf{x}_1} \tilde{P}_{22} g(\mathbf{x},\mathbf{x}') \Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \leq C \|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} \big(\mathbf{d}^{\gamma}(\mathbf{x},\mathbf{x}') + (T-t)^{\frac{\gamma}{2}} \big). \tag{5.12}
$$

Plugging (5.8), (5.10), (5.12) into (5.7) yields the result.

- Diagonal regime. If $T - t > c_0 \mathbf{d}^2(\mathbf{x},\mathbf{x}')$, we directly write:

$$
\begin{aligned}
&|D^2_{\mathbf{x}_1} \tilde{P}^{\boldsymbol{\xi}}_{T,t} g(\mathbf{x}) - D^2_{\mathbf{x}_1} \tilde{P}^{\boldsymbol{\xi}}_{T,t} g(\mathbf{x}')| \\
\leq \quad &\Big| \int_{\mathbb{R}^{nd}} [D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y}) - D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}',\mathbf{y})] g(\mathbf{y}) d\mathbf{y} \Big| \\
\leq \quad &\sum_{k=1}^{n} \Big| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}' + \mu(\mathbf{x}-\mathbf{x}'),\mathbf{y}) \cdot (\mathbf{x}-\mathbf{x}')_k g(\mathbf{y}) d\mathbf{y} \Big|
\end{aligned}
$$

This contribution is dealt through the cancellation tools of Proposition 5 (see equations (3.16), (3.18)). We get from the above estimate that:

$$
\begin{aligned}
&|D^2_{\mathbf{x}_1} \tilde{P}^{\boldsymbol{\xi}}_{T,t} g(\mathbf{x}) - D^2_{\mathbf{x}_1} \tilde{P}^{\boldsymbol{\xi}}_{T,t} g(\mathbf{x}')| \\
\leq \quad &\sum_{k=1}^{n} \Big| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}' + \mu(\mathbf{x}-\mathbf{x}'),\mathbf{y}) \cdot (\mathbf{x}-\mathbf{x}')_k \Big[g\big(\mathbf{y}\big) - g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}' + \mu(\mathbf{x}-\mathbf{x}')\big)\big) \\
&- \big\langle D_{\mathbf{x}_1} g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}' + \mu(\mathbf{x}-\mathbf{x}'))\big), \big(\mathbf{y} - \mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}' + \mu(\mathbf{x}-\mathbf{x}'))\big)_1 \big\rangle \\
&- \frac{1}{2} \mathrm{Tr}\Big(D^2_{\mathbf{x}_1} g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}' + \mu(\mathbf{x}-\mathbf{x}')\big)\big(\mathbf{y} - \mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}' + \mu(\mathbf{x}-\mathbf{x}'))\big)^{\otimes 2}_1 \Big) \Big] d\mathbf{y} \Big|.
\end{aligned} \tag{5.13}
$$

Because $g \in C^{2+\gamma}_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})$, we readily deduce, reproducing the Taylor expansion on g employed for equations (5.7)-(5.9) above, that:

$$\begin{aligned}&\Big|g\big(\mathbf{y}\big)-g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')\big)\big)-\big\langle D_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'))),(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')))_1\big\rangle\\&-\frac{1}{2}\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')\big)\big)\Big(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')\big)\Big)_1^{\otimes 2}\Big)\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\(5.14)\quad\leq\quad&\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}\big(\mathbf{y},\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')\big)\big)\Big|_{\boldsymbol{\xi}=\mathbf{x}}.\end{aligned}$$

Plugging this inequality into (5.13) yields:

$$\begin{aligned}&|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x})-D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x}')|\big|_{\boldsymbol{\xi}=\mathbf{x}}\\\leq\quad&C\|g\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n\int_{\mathbb{R}^{nd}}(T-t)^{-1-(k-\frac{1}{2})}\Big\{\bar{p}^{\boldsymbol{\xi}}_{C^{-1}}(t,T,\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'),\mathbf{y})|(\mathbf{x}-\mathbf{x}')_k|\\&\times\mathbf{d}^{2+\gamma}\big(\mathbf{y},\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'))\big)\Big\}\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\\leq\quad&C\|g\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n(T-t)^{-(k-\frac{1}{2})+\frac{\gamma}{2}}|(\mathbf{x}-\mathbf{x}')_k|\\(5.15)\qquad\leq\quad&C\|g\|_{C^{2+\gamma}},\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'),\end{aligned}$$

using (3.8) for the second inequality and recalling that, since $c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')<(T-t)$, we indeed have $(T-t)^{-(k-\frac{1}{2})+\frac{\gamma}{2}}|(\mathbf{x}-\mathbf{x}')_k|\leq(c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}'))^{-(k-\frac{1}{2})+\frac{\gamma}{2}}\mathbf{d}^{2k-1}(\mathbf{x},\mathbf{x}')\leq C\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')$.This concludes the proof of the first inequality of Lemma 14.

5.1.2. *Hölder control for the degenerate variables.* We are here interested in proving the second estimate in Lemma 14 relying on the Hölder regularity of the frozen semi-group w.r.t. the degenerate variables. This proof is also based on the previous techniques. In particular, we still take advantage of cancellation tools. For the whole paragraph we consider two arbitrary given spatial points $(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2$ s.t. $\mathbf{x}_1=\mathbf{x}'_1$, i.e. their first entry, corresponding to the non-degenerate variable, coincide.

• Off-diagonal regime. If $T-t\leq c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$, we proceed to an expansion similar to (5.7) for $D^2_{\mathbf{x}_1}\tilde{P}^{\mathbf{x}}_{T,t}g(\mathbf{x})$. In particular, with the notations introduced in (5.7), we write:

$$\tilde{P}^{\boldsymbol{\xi}}_{T,t}g(\mathbf{x})-\tilde{P}^{\boldsymbol{\xi}'}_{T,t}g(\mathbf{x}')=:\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_1g(\mathbf{x},\mathbf{x}')+\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}}\tilde{P}_2g(\mathbf{x},\mathbf{x}').\tag{5.16}$$

We directly obtain from the Proposition 3, similarly to (5.8), that:

$$\Big|\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_1g(\mathbf{x},\mathbf{x}')\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\leq 2C(T-t)^{\frac{2+\gamma}{2}}\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\leq 2C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').\tag{5.17}$$

We indeed recall that the difference w.r.t. (5.8) is that we do not have anymore the time-singularities coming therein from the spatial derivatives.

With the notations of (5.9), the second contribution of (5.16) writes:

$$\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_2g(\mathbf{x},\mathbf{x}')=\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_{21}g(\mathbf{x},\mathbf{x}')+\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_{22}g(\mathbf{x},\mathbf{x}').$$

Proposition 3 again yields, similarly to (5.10), that:

$$\Big|\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_{21}g(\mathbf{x},\mathbf{x}')\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\xi')}\leq 2C(T-t)^{\frac{2+\gamma}{2}}\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\leq 2C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').\tag{5.18}$$

On the other hand, we readily get from Proposition 5 that:

$$\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_{22}g(\mathbf{x},\mathbf{x}')=\frac{1}{2}\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))[\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{T,t}]_{1,1}-D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))[\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{T,t}]_{1,1}\Big).$$

Write now:

$$\begin{aligned}\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_{22}g(\mathbf{x},\mathbf{x}')\quad=\quad&\frac{1}{2}\mathrm{Tr}\Big([D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))-D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))][\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{T,t}]_{1,1}\Big)\\&+\frac{1}{2}\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))\big[[\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{T,t}]_{1,1}-[\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{T,t}]_{1,1}\big]\Big).\end{aligned}\tag{5.19}$$

Since $T-t \le c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$, recalling as well that for $\boldsymbol{\xi}=\mathbf{x}$, $\boldsymbol{\xi}'=\mathbf{x}'$, $\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x})=\boldsymbol{\theta}_{T,t}(\mathbf{x})$, $\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}')=\boldsymbol{\theta}_{T,t}(\mathbf{x}')$, we readily deduce from Proposition 2 and Lemma 7 that:

$$\begin{aligned}\frac{1}{2}\Big|[D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}))-D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}'}_{T,t}(\mathbf{x}'))][\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{T,t}]_{1,1}\Big|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} &\le C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{\gamma}\big(\boldsymbol{\theta}_{T,t}(\mathbf{x}),\boldsymbol{\theta}_{T,t}(\mathbf{x}')\big)(T-t)\\ &\le C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').\end{aligned}\tag{5.20}$$

For the last contribution, we directly obtain from Lemma 6 (equation (3.19) for $j=1$):

$$\begin{aligned}\big|\frac{1}{2}D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'))\big([\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{T,t}]_{1,1}-[\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{T,t}]_{1,1}\big)\big|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} &\le C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\big((T-t)^{\frac{2+\gamma}{2}}+(T-t)\mathbf{d}^{\gamma}(\mathbf{x},\mathbf{x}')\big)\\ &\le C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'),\end{aligned}\tag{5.21}$$

using again that $T-t\le c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$ for the last inequality. Plugging (5.20) and (5.21) into (5.19) gives

$$\Big|\Delta_{t,T,\boldsymbol{\xi},\boldsymbol{\xi}'}\tilde{P}_{22}g(\mathbf{x},\mathbf{x}')\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}\le C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').\tag{5.22}$$

Bringing together (5.18), (5.22) and (5.17) in (5.16) yields the result.

- Diagonal regime. If $T-t\ge c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$, we write:

$$\begin{aligned}&|\tilde{P}^{\mathbf{x}}_{T,t}g(\mathbf{x})-\tilde{P}^{\mathbf{x}}_{T,t}g(\mathbf{x}')|\\ \le\ &\Big|\int_{\mathbb{R}^{nd}}[\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})-\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}',\mathbf{y})]g(\mathbf{y})d\mathbf{y}\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ \le\ &\Big|\int_0^1 d\mu\int_{\mathbb{R}^{nd}}\langle\mathbf{D}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'),\mathbf{y}),(\mathbf{x}-\mathbf{x}')\rangle\Big[g\big(\mathbf{y}\big)-g\big(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')\big)\big)\\ &-\big\langle D_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'))),(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')))_1\big\rangle\\ &-\frac{1}{2}\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}g(\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')))(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')))_1^{\otimes 2}\Big)\Big]d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}},\end{aligned}$$

with the same cancellation argument as in (5.13). Observe anyhow that the cancellation involving the gradient in the above equation is possible precisely because $\mathbf{x}_1=\mathbf{x}'_1$ and therefore $\mathbf{D}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'),\mathbf{y})\cdot(\mathbf{x}-\mathbf{x}')=D_{\mathbf{x}_{2:n}}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'),\mathbf{y})\cdot(\mathbf{x}-\mathbf{x}')_{2:n}$. We then obtain thanks to the previous identity and (5.14):

$$\begin{aligned}&|\tilde{P}^{\mathbf{x}}_{T,t}g(\mathbf{x})-\tilde{P}^{\mathbf{x}}_{T,t}g(\mathbf{x}')|\\ \le\ &\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\sum_{k=2}^n\int_{\mathbb{R}^{nd}}\big|D_{\mathbf{x}_k}\tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'),\mathbf{y})\big||\mathbf{x}_k-\mathbf{x}'_k|\mathbf{d}^{2+\gamma}\big(\mathbf{y},\mathbf{m}^{\boldsymbol{\xi}}_{T,t}\big(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')\big)\big)d\mathbf{y}\\ \le\ &\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\sum_{k=2}^n\int_{\mathbb{R}^{nd}}\frac{C}{(T-t)^{1+(k-\frac{1}{2})-\frac{2+\gamma}{2}}}\bar{p}^{\mathbf{x}}_{C^{-1}}(t,T,\mathbf{m}^{\mathbf{x}}_{T,t}(\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')),\mathbf{y})|\mathbf{x}_k-\mathbf{x}'_k|d\mathbf{y}\\ \le\ &C\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'),\end{aligned}\tag{5.23}$$

reproducing the arguments used to establish (5.15) for the last inequality. This concludes the proof of the second assertion in Lemma 14.

5.2. **Hölder norms associated with the Green kernel.** Let us recall that in (2.39), for a source $f\in L^\infty\big([0,T],C^\gamma_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})\big)$, we have to control the Hölder norms of the Green kernel which we split into two parts according to the position of the time integration variable w.r.t. the change of regime time t_0 (see (2.40)) *a posteriori* chosen to be $t_0:=\big(t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')\big)\wedge T$. This is again the splitting according to the off-diagonal and diagonal regime. The point is that for the Green kernel, if $t_0<T$ both regimes appear.

Lemma 15. *Under* **(A)**, *for fixed spatial points* $(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2$, *we have that there exists a constant* $C:=C((\mathbf{A}),T)$, *s.t. for any* $f\in L^\infty\big([0,T],C^\gamma_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})\big)$:

$$\begin{aligned}&\sup_{t\in[0,T]}\big(|D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi}_{t_0,t}f(t,\mathbf{x})-D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi'}_{t_0,t}f(t,\mathbf{x}')|+|D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi}_{T,t_0}f(t,\mathbf{x})-D^2_{\mathbf{x}_1}\tilde G^{\tilde{\boldsymbol\xi}'}_{T,t_0}f(t,\mathbf{x}')|\big)\\ \le\;& C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'),\end{aligned}$$

and

$$\begin{aligned}&\sup_{t\in[0,T]}\big(|\tilde G^{\boldsymbol\xi}_{t_0,t}f(t,\mathbf{x})-\tilde G^{\boldsymbol\xi'}_{t_0,t}f(t,\mathbf{x}')|+|\tilde G^{\boldsymbol\xi}_{T,t_0}f(t,\mathbf{x})-\tilde G^{\tilde{\boldsymbol\xi}'}_{T,t_0}f(t,\mathbf{x}')|\big)\\ \le\;& C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'),\ \textit{if}\ \mathbf{x}_1=\mathbf{x}'_1,\end{aligned}$$

where $\boldsymbol\xi=\mathbf{x}$, $\boldsymbol\xi'=\mathbf{x}'$, $\tilde{\boldsymbol\xi}'=\mathbf{x}$.

Proof of Lemma 15. Let us begin with the statement concerning the second order derivatives of the frozen Green kernel w.r.t. the non-degenerate variable $\mathbf{x}_1$.

For the *off-diagonal regime*, involving the term $D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi'}_{t_0,t}f(t,\mathbf{x}')$, we readily get from Lemma 4 that

$$\begin{aligned}&\big|D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi}_{t_0,t}f(t,\mathbf{x})-D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi'}_{t_0,t}f(t,\mathbf{x}')\big|\Big|_{(\boldsymbol\xi,\boldsymbol\xi')=(\mathbf{x},\mathbf{x}')}\\ \le\;&\Big|\int_t^{t_0}ds D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi}_{s,t}f(s,\mathbf{x})\Big|\Big|_{\boldsymbol\xi=\mathbf{x}}+\Big|\int_t^{t_0}ds D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi'}_{s,t}f(s,\mathbf{x}')\Big|\Big|_{\boldsymbol\xi'=\mathbf{x}'}\\ \le\;& C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\int_t^{t_0}ds(s-t)^{-1+\frac\gamma2}\\ \le\;& C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}').\end{aligned}\tag{5.24}$$

For the *diagonal regime*, involving the term $D^2_{\mathbf{x}_1}\tilde G^{\tilde{\boldsymbol\xi}'}_{T,t_0}f(t,\mathbf{x}')$, we have to be more subtle and perform again a Taylor expansion of $D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi}_{s,t}f(s,\cdot)$. Namely:

$$\begin{aligned}&\big|D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi}_{T,t_0}f(t,\mathbf{x})-D^2_{\mathbf{x}_1}\tilde G^{\tilde{\boldsymbol\xi}'}_{T,t_0}f(t,\mathbf{x}')\big|\Big|_{(\boldsymbol\xi,\tilde{\boldsymbol\xi}')=(\mathbf{x},\mathbf{x})}\\ \le\;&\int_{t_0}^T ds\Big|\int_0^1 d\mu\mathbf{D}D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi}_{s,t}f(s,\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}'))\cdot(\mathbf{x}-\mathbf{x}')\Big|\Big|_{\boldsymbol\xi=\mathbf{x}}\\ \le\;&\sum_{i=1}^n|(\mathbf{x}-\mathbf{x}')_i|\int_{t_0}^T ds\int_0^1 d\mu\Big|D_{\mathbf{x}_i}D^2_{\mathbf{x}_1}\tilde P^{\boldsymbol\xi}_{s,t}f\big(s,\mathbf{x}'+\lambda(\mathbf{x}-\mathbf{x}')\big)\Big|\Big|_{\boldsymbol\xi=\mathbf{x}}\\ \le\;& C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\sum_{i=1}^n|(\mathbf{x}-\mathbf{x}')_i|\int_{t_0}^T ds(s-t)^{-1-(i-\frac12)+\frac\gamma2},\end{aligned}$$

using again Lemma 4 and the arguments of (5.15) for the last inequality. This finally yields, recalling the definition of $\mathbf{d}$ in (1.9) (especially that $|(\mathbf{x}-\mathbf{x}')_i|\le\mathbf{d}(\mathbf{x},\mathbf{x}')^{2i-1}$) and the fact that we chose now $t_0=\big(t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')\big)\wedge T$:

$$\begin{aligned}\big|D^2_{\mathbf{x}_1}\tilde G^{\boldsymbol\xi}_{T,t_0}f(t,\mathbf{x})-D^2_{\mathbf{x}'_1}\tilde G^{\boldsymbol\xi'}_{T,t_0}f(t,\mathbf{x}')\big|\Big|_{(\boldsymbol\xi,\boldsymbol\xi')=(\mathbf{x},\mathbf{x})}\;&\le\;C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\sum_{i=1}^n|(\mathbf{x}-\mathbf{x}')_i|(\mathbf{d}^2(\mathbf{x},\mathbf{x}'))^{-(i-\frac12)+\frac\gamma2}\\ &\le\;C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}').\end{aligned}\tag{5.25}$$

Gathering (5.24) and (5.25) gives the first estimate of the lemma.

Let us now turn to the Hölder controls on the degenerate variables. The idea is here again to perform a Taylor expansion at order one for $\mathbf{x}_1 = \mathbf{x}'_1$. Namely, from Lemma 4, the diagonal control is direct. We get

$$\begin{aligned}
&\big|\tilde{G}^{\boldsymbol{\xi}}_{T,t_0} f(t,\mathbf{x}) - \tilde{G}^{\tilde{\boldsymbol{\xi}}'}_{T,t_0} f(t,\mathbf{x}')\big|\Big|_{(\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x})} \\
\leq\ & \int_{t_0}^{T} ds \int_0^1 d\mu \big|\big\langle \mathbf{D}\tilde{P}^{\boldsymbol{\xi}}_{s,t} f\big(s,\mathbf{x}'+\mu(\mathbf{x}-\mathbf{x}')\big),(\mathbf{x}-\mathbf{x}')\big\rangle\big| \\
\leq\ & C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \int_{(t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}'))\wedge T}^{T} ds \sum_{i=2}^{n} |(\mathbf{x}-\mathbf{x}')_i| (s-t)^{-(i-\frac{1}{2})+\frac{\gamma}{2}} \\
\leq\ & C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \sum_{i=2}^{n} |(\mathbf{x}-\mathbf{x}')_i| \mathbf{d}(\mathbf{x},\mathbf{x}')^{2-(2i-1)+\gamma} \\
\leq\ & C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'). \qquad (5.26)
\end{aligned}$$

Now, for the off-diagonal bound, associated with the term $\tilde{G}^{\boldsymbol{\xi}}_{t_0,t} f(t,\cdot)$, we precisely need to exploit the smoothness of f associated to the fact that the semi-group $\tilde{P}^{\boldsymbol{\xi}}_{s,t}$ has a density. Indeed, we cannot take advantage of the cancellation tools of Lemma 4, but we have for all $\mathbf{x},\mathbf{x}' \in \mathbb{R}^{nd}$ s.t. $\mathbf{x}_1 = \mathbf{x}'_1$:

$$\begin{aligned}
&\Big|\int_t^{t_0} ds\Big(\tilde{P}^{\boldsymbol{\xi}}_{s,t} f(s,\mathbf{x}) - \tilde{P}^{\boldsymbol{\xi}'}_{s,t} f(s,\mathbf{x}')\Big)\Big|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \qquad (5.27) \\
\leq\ & \Big|\int_t^{t_0} ds \tilde{P}^{\boldsymbol{\xi}}_{s,t} f(s,\mathbf{x}) - f(s,\mathbf{m}^{\boldsymbol{\xi}}_{t_0,t}(\mathbf{x}))\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} + \Big|\int_t^{t_0} ds \tilde{P}^{\boldsymbol{\xi}'}_{s,t} f(s,\mathbf{x}') - f(s,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))\Big|\Big|_{\boldsymbol{\xi}'=\mathbf{x}'} \\
&+\Big|\int_t^{t_0} ds f(s,\mathbf{m}^{\boldsymbol{\xi}}_{t_0,t}(\mathbf{x})) - f(s,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))\Big|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}.
\end{aligned}$$

Note that the first two terms in the r.h.s. of inequality (5.27) are handled like in the previous section. Precisely, we write for the first contribution:

$$\begin{aligned}
&\Big|\int_t^{t_0} ds \tilde{P}^{\boldsymbol{\xi}}_{s,t} f(s,\mathbf{x}) - f(s,\mathbf{m}^{\boldsymbol{\xi}}_{t_0,t}(\mathbf{x})) d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
=\ & \Big|\int_t^{t_0} ds \int_{\mathbb{R}^{nd}} \tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y})[f(s,\mathbf{y}) - f(s,\boldsymbol{\theta}_{t_0,t}(\mathbf{x}))]\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & \|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} \tilde{p}^{\boldsymbol{\xi}}(t,T,\mathbf{x},\mathbf{y}) \mathbf{d}^\gamma(\mathbf{y},\boldsymbol{\theta}_{t_0,t}(\mathbf{x})) d\mathbf{y}\Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \int_t^{t_0} ds (s-t)^{\frac{\gamma}{2}} d\mathbf{y} \\
\leq\ & C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'),
\end{aligned}$$

by definition of t_0 in (2.40). The second term of (5.27) is handled similarly. We thus obtain:

$$\begin{aligned}
&\Big|\int_t^{t_0} ds \tilde{P}^{\boldsymbol{\xi}}_{s,t} f(s,\mathbf{x}) - f(s,\mathbf{m}^{\boldsymbol{\xi}}_{t_0,t}(\mathbf{x}) d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} + \Big|\int_t^{t_0} ds \tilde{P}^{\boldsymbol{\xi}'}_{s,t} f(s,\mathbf{x}') - f(s,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}') d\mathbf{y}\Big|\Big|_{\boldsymbol{\xi}'=\mathbf{x}'} \\
\leq\ & C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'). \qquad (5.28)
\end{aligned}$$

For the last contribution in (5.27), we have directly that:

$$\Big|\int_t^{t_0} ds f(s,\mathbf{m}^{\boldsymbol{\xi}}_{t_0,t}(\mathbf{x})) - f(s,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))\Big|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \leq C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \int_t^{t_0} ds \mathbf{d}^\gamma\big(\boldsymbol{\theta}_{t_0,t}(\mathbf{x}),\boldsymbol{\theta}_{t_0,t}(\mathbf{x}')\big). \qquad (5.29)$$

Lemma 7 and (5.29) eventually yield:

$$\begin{aligned}
\Big|\int_t^{t_0} ds f(s,\mathbf{m}^{\boldsymbol{\xi}}_{t_0,t}(\mathbf{x})) - f(s,\mathbf{m}^{\boldsymbol{\xi}}_{t_0,t}(\mathbf{x}'))\Big|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} &\leq C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\big((t_0-t)^{1+\frac{\gamma}{2}} + (t_0-t)\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\big) \\
&= 2C\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})} \mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').
\end{aligned}$$

This, together with (5.26) gives the second estimate of the lemma. □

5.3. **Hölder norms of the perturbative contribution.** This section is dedicated to the investigation of the spatial Hölder continuity of the perturbative term in (5.4) and (5.6). Recalling the notations (5.5) introduced at the beginning of this section, we prove the following lemma.

Lemma 16. *Under* **(A)**, *for fixed spatial points* $(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2$, *we have that there exists a constant* $\Lambda:=\Lambda((\mathbf{A}),T)$ *as in Remark 3, s.t.*

$$\begin{aligned}&|D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')|\big|_{(\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x})}+|D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off-diag}}(t,\mathbf{x},\mathbf{x}')|\big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}\\ \leq\;&\Lambda(c_0^{-(n-\frac12)+\frac\gamma2}+c_0^{\frac\gamma2})\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'),\end{aligned}\tag{5.30}$$

and, if we assume in addition that $\mathbf{x}_1=\mathbf{x}'_1$,

$$|\Delta^{\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')|\big|_{(\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x})}+|\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off-diag}}(t,\mathbf{x},\mathbf{x}')|\big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}\leq\Lambda(c_0^{-(n-\frac12)+\frac\gamma2}+c_0)\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}')\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}.\tag{5.31}$$

As already successfully used to establish in the previous sections to derive the regularity of the semi-group and the Green kernel we split the investigations into two parts: the first one is done when the system is in the *off-diagonal* regime (i.e. for time $s\leq t_0$) and the other one when the system is in the *diagonal* regime (i.e. for time $s>t_0$). We aslo recall that the critical time giving the change of regime is (chosen after potential differentiation) $t_0=t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')\wedge T$. We can assume here w.l.o.g. that $t_0<T$ (otherwise there is a globally off-diagonal regime and the analysis becomes easier).

• Control of (5.30). We decompose from definitions (2.31) and (2.39):

$$\begin{aligned}&|D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')|\Big|_{(\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x})}\\ \leq\;&\Big|\int_t^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\Big[D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})-D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}(t,s,\mathbf{x}',\mathbf{y})\Big]\Big[\langle\mathbf{F}_1(s,\mathbf{y})-\mathbf{F}_1(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})),D_{\mathbf{y}_1}\rangle\\ &+\frac12\mathrm{Tr}\Big(\big(a(s,\mathbf{y})-a(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))\big)D^2_{\mathbf{y}_1}\Big)\Big]u(s,\mathbf{y})\mathbb{I}_{s>t_0}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ &+\Big|\sum_{i=2}^n\int_t^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\Big[D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})-D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}(t,s,\mathbf{x}',\mathbf{y})\mathbb{I}_{s>t_0}\Big]\\ &\times\Big\langle\mathbf{F}_i(s,\mathbf{y})-[\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))+D_{\mathbf{x}_{i-1}}\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}],D_{\mathbf{y}_i}u(s,\mathbf{y})\Big\rangle\mathbb{I}_{s>t_0}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}},\end{aligned}$$

which readily yields with the notations of (4.3) that:

$$\begin{aligned}&|D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')|\Big|_{(\boldsymbol{\xi},\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x})}\\ \leq\;&\Big|\int_t^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu\mathbf{D}D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}-\mathbf{x}')\Delta_{1,\mathbf{F},\sigma}(t,s,\mathbf{y},\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),u)\mathbb{I}_{s>t_0}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ &+\Big|\int_t^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu\mathbf{D}D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}-\mathbf{x}')\big\langle\Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y}),D_{\mathbf{y}_i}u(s,\mathbf{y})\big\rangle\mathbb{I}_{s>t_0}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ =:\;&\big|D^2_{\mathbf{x}_1}\Delta^1_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')\big|+\big|D^2_{\mathbf{x}_1}\Delta^{2:n}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')\big|.\end{aligned}\tag{5.32}$$

We will now control the first term of the above right hand side. In other words, we specify the control of (2.16). We obtain directly thanks to the smoothness assumption **(S)** on the coefficients and Proposition 3 (see also equation (3.8)) that for each $k\in[\![1,n]\!]$:

$$\begin{aligned}&|\mathbf{F}_1(s,\mathbf{y})-\mathbf{F}_1(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))||D_{\mathbf{x}_k}D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ \leq\;&C\Big(\|\mathbf{F}_1(s,\cdot)\|_{C^\gamma_{\mathbf{d}}}\mathbf{d}^\gamma\big(\mathbf{y},\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})\big)\Big)\times(s-t)^{-1-(k-\frac12)}\bar p_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\ \leq\;&C\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})}(s-t)^{-1-(k-\frac12)+\frac\gamma2}\bar p_{C^{-1}}(t,s,\mathbf{x},\mathbf{y}).\end{aligned}\tag{5.33}$$

Similarly,

$$\begin{aligned}
&|a(s,\mathbf{y})-a(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))||D_{\mathbf{x}_k}D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & C\Big(\|a(s,\cdot)\|_{C^\gamma_{\mathbf{d}}}\mathbf{d}^\gamma\big(\mathbf{y},\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})\big)\Big)\times(s-t)^{-1-(k-\frac{1}{2})}\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & C\|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}(s-t)^{-1-(k-\frac{1}{2})+\frac{\gamma}{2}}\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y}). \qquad (5.34)
\end{aligned}$$

We carefully point-out that the indicated bound only depend on the supremum in time of the Hölder modulus of the coefficients (denoted $\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})}, \|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}$ respectively) and not on their supremum norm. In particular, we get from (5.33), (5.34):

$$\begin{aligned}
|D^2_{\mathbf{x}_1}\Delta^1_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')| \leq\ & \sum_{k=1}^n\int_{t_0}^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}(s-t)^{-1-(k-\frac{1}{2})+\frac{\gamma}{2}}\\
&\times\Big(\|D_{\mathbf{x}_1}u\|_{L^\infty}\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})}+\|D^2_{\mathbf{x}_1}u\|_{L^\infty}\|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}\Big)\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})|\mathbf{x}_k-\mathbf{x}'_k|\\
\leq\ & \Lambda(\|D_{\mathbf{x}_1}u\|_{L^\infty}+\|D^2_{\mathbf{x}_1}u\|_{L^\infty})\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'), \qquad (5.35)
\end{aligned}$$

where again the constant Λ is *small* provided the coefficients *do not vary much*. Thanks to Lemma 11 and the previous analysis of Section 4.2, we directly deduce:

$$\begin{aligned}
&|D^2_{\mathbf{x}_1}\Delta^{2:n}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')|\\
\leq\ & \Big|\sum_{i=2}^n\int_t^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu D_{\mathbf{y}_i}\cdot\Big\{\Big(\mathbf{D}D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}-\mathbf{x}')\Big)\\
&\otimes\Big(\mathbf{F}_i(s,\mathbf{y})-[\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))+D_{\mathbf{x}_{i-1}}\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}]\Big)\Big\}u(s,\mathbf{y})\mathbb{I}_{s>t_0}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & \Lambda\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n\int_{t_0}^T\frac{|\mathbf{x}_k-\mathbf{x}'_k|ds}{(s-t)^{1+(k-\frac{1}{2})-\frac{\gamma}{2}}}\\
\leq\ & \Lambda c_0^{-(n-\frac{1}{2})+\frac{\gamma}{2}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n\mathbf{d}^{\gamma-(2k-1)}(\mathbf{x},\mathbf{x}')|\mathbf{x}_k-\mathbf{x}'_k|\\
\leq\ & \Lambda c_0^{-(n-\frac{1}{2})+\frac{\gamma}{2}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'). \qquad (5.36)
\end{aligned}$$

Plugging (5.35) and (5.36) into (5.32) yields the stated control for the diagonal contribution.

Let us now turn to the control of $|D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}')|$ in (5.5) (or (2.31) in the detailed guide). In the off-diagonal case, we choose $\boldsymbol{\xi}=\mathbf{x}$ and $\boldsymbol{\xi}'=\mathbf{x}'$ and

$$\begin{aligned}
|D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}')|\big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')} \leq\ & \Big|\int_t^{t_0}ds\int_{\mathbb{R}^{nd}}d\mathbf{y}D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})(L_s-\tilde{L}^{\boldsymbol{\xi}}_s)u(s,\mathbf{y})\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
&+\Big|\int_t^{t_0}ds\int_{\mathbb{R}^{nd}}d\mathbf{y}D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,s,\mathbf{x}',\mathbf{y})(L_s-\tilde{L}^{\boldsymbol{\xi}'}_s)u(s,\mathbf{y})\Big|\Big|_{\boldsymbol{\xi}'=\mathbf{x}'}. \qquad (5.37)
\end{aligned}$$

We readily get thanks to Lemmas 10 and 11:

$$|D^2_{\mathbf{x}_1}\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}')|\big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}\leq\Lambda\sum_{i=2}^n\int_t^{t_0}\frac{ds}{(s-t)^{1-\frac{\gamma}{2}}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\leq\Lambda c_0^{\frac{\gamma}{2}}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}. \qquad (5.38)$$

We point out from (5.36) and (5.38) that there are opposite impacts of the threshold c_0 on the constants, depending on the diagonal and off-diagonal regimes at hand.
We eventually get the estimate (5.30) plugging (5.35), (5.36) into (5.32) and (5.38) recalling that $c_0\leq 1$.

• Control of (5.31). We proceed as above from definitions (5.5) considering spatial points $(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2$ s.t. $\mathbf{x}_1=\mathbf{x}'_1$. In the diagonal case, we also choose $\boldsymbol{\xi}=\tilde{\boldsymbol{\xi}}'=\mathbf{x}$ and we write similarly to (5.32):

$$
\begin{aligned}
&|\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x})}\\
&\le \Big|\int_t^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu\big\langle D_{\mathbf{x}}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y}),\mathbf{x}-\mathbf{x}'\big\rangle\Delta_{1,\mathbf{F},\sigma}(t,s,\mathbf{y},\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))\mathbb{I}_{s>t_0}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
&+\Big|\sum_{i=2}^n\int_t^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu\big\langle D_{\mathbf{x}}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y}),\mathbf{x}-\mathbf{x}'\big\rangle\Delta_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}),\mathbf{y})D_{\mathbf{y}_i}u(s,\mathbf{y})\mathbb{I}_{s>t_0}\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}.
\end{aligned}
\tag{5.39}
$$

We have an expression rather similar to the one that appeared for the control of $[D^2_{\mathbf{x}_1}u(t,\cdot)]_\gamma$ but with a weaker time singularity.

In other words, thanks to identities (5.33), (5.34) and Lemma 11 we obtain:

$$
\begin{aligned}
\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{diag}}(t,T,\mathbf{x},\mathbf{x}')| &\le \Lambda\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n\int_{t_0}^T ds(s-t)^{-(k-\frac12)+\frac\gamma2}|\mathbf{x}_k-\mathbf{x}'_k|\\
&\le \Lambda c_0^{-(n-\frac12)+\frac\gamma2}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n\mathbf{d}^{2(1-(k-\frac12)+\frac\gamma2)}(\mathbf{x},\mathbf{x}')|\mathbf{x}_k-\mathbf{x}'_k|\\
&\le \Lambda c_0^{-(n-\frac12)+\frac\gamma2}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'),
\end{aligned}
\tag{5.40}
$$

where again the constant C is small if the Hölder moduli of the coefficients are small.

For the off-diagonal contribution, we get for $\mathbf{x}_1=\mathbf{x}'_1$, $\boldsymbol{\xi}=\mathbf{x}$ and $\boldsymbol{\xi}'=\mathbf{x}'$:

$$
\begin{aligned}
&|\Delta^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,T,\mathbf{x},\mathbf{x}')|\big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}\\
&\le \Big|\int_t^{t_0}ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\tilde{p}^{\boldsymbol{\xi}}(t,s,\mathbf{x},\mathbf{y})(L_s-\tilde{L}^{\boldsymbol{\xi}}_s)u(s,\mathbf{y})\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}+\Big|\int_t^{t_0}ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\tilde{p}^{\boldsymbol{\xi}'}(t,s,\mathbf{x}',\mathbf{y})(L_s-\tilde{L}^{\boldsymbol{\xi}'}_s)u(s,\mathbf{y})\Big|\Big|_{\boldsymbol{\xi}'=\mathbf{x}'}\\
&\le \Lambda\sum_{k=2}^n\int_t^{t_0}ds(s-t)^{\frac\gamma2}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\\
&= \Lambda c_0\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}')\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}.
\end{aligned}
\tag{5.41}
$$

The last but one inequality is a consequence of Lemmas 10 and 11. Equations (5.40) and (5.41) yield (5.31). This concludes the proof of Lemma 16.

5.4. **Controls of the discontinuity terms arising from the change of freezing point.** It now remains to control the contribution arising from the change of freezing point in equation (5.4). The main result of this section is the following lemma.

Lemma 17 (Control of the discontinuity terms)**.** *There exists $C:=C((\mathbf{A}))$ s.t. for any $(t,\mathbf{x},\mathbf{x}')\in[0,T]\times\mathbb{R}^{nd}\times\mathbb{R}^{nd}$ taking $\boldsymbol{\xi}'=\mathbf{x}'$, $\tilde{\boldsymbol{\xi}}'=\mathbf{x}$,*

$$
\begin{aligned}
\big|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u(t_0,\mathbf{x}')-D^2_{\mathbf{x}_1}\tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u(t_0,\mathbf{x}')\big| &\le Cc_0^{\frac{\gamma}{(2n-1)}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'),\\
\big|\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u(t_0,\mathbf{x}')-\tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u(t_0,\mathbf{x}')\big| &\le Cc_0\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'),\ \textit{for } \mathbf{x}_1=\mathbf{x}'_1.
\end{aligned}
$$

We prove the above statement in the next paragraphs respectively dedicated to the control of the derivatives w.r.t. the non-degenerate variables (first estimate) and the control of the Hölder moduli associated with the degenerate ones (second estimate).

5.4.1. *Control of the derivatives w.r.t. the non-degenerate variables.* As done in (5.7) and (5.9), we can write:

$$
\begin{aligned}
&D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u(t_0,\mathbf{x}') - D^2_{\mathbf{x}_1}\tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u(t_0,\mathbf{x}')\\
=\quad&\Bigg[\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,t_0,\mathbf{x}',\mathbf{y})[u(t_0,\mathbf{y}) - u(t_0,\mathbf{y}_1,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')_{2:n})]d\mathbf{y}\\
&\quad-\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\tilde{\boldsymbol{\xi}}'}(t,t_0,\mathbf{x}',\mathbf{y})[u(t_0,\mathbf{y}) - u(t_0,\mathbf{y}_1,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')_{2:n})]d\mathbf{y}\Bigg]\\
&+\Bigg[\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,t_0,\mathbf{x}',\mathbf{y})\int_0^1 d\mu(1-\mu)\\
&\quad\mathrm{Tr}\Big(\big[D^2_{\mathbf{x}_1}u\big(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')_1+\mu(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))_1,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')_{2:n}\big) - D^2_{\mathbf{x}_1}u\big(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')\big)\big]\big(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')\big)_1^{\otimes 2}\Big)d\mathbf{y}\\
&\quad-\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\tilde{\boldsymbol{\xi}}'}(t,t_0,\mathbf{x}',\mathbf{y})\int_0^1 d\mu(1-\mu)\\
&\quad\mathrm{Tr}\Big(\big[D^2_{\mathbf{x}_1}u\big(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')_1+\mu(\mathbf{y}-\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))_1,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')_{2:n}\big) - D^2_{\mathbf{x}'_1}u\big(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')\big)\big]\big(\mathbf{y}-\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')\big)_1^{\otimes 2}\Big)d\mathbf{y}\Bigg]\\
&+\frac{1}{2}\Bigg[\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,t_0,\mathbf{x}',\mathbf{y})\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))_1^{\otimes 2}\Big)d\mathbf{y}\\
&\quad-\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\tilde{\boldsymbol{\xi}}'}(t,t_0,\mathbf{x}',\mathbf{y})\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))(\mathbf{y}-\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))_1^{\otimes 2}\Big)d\mathbf{y}\Bigg]\\
=:\quad&\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}D^2_{\mathbf{x}_1}\tilde{P}_1u(t_0,\mathbf{x},\mathbf{x}') + \Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}D^2_{\mathbf{x}_1}\tilde{P}_{21}u(t_0,\mathbf{x},\mathbf{x}') + \Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}D^2_{\mathbf{x}_1}\tilde{P}_{22}u(t_0,\mathbf{x},\mathbf{x}').
\end{aligned}
\tag{5.42}
$$

We first directly write from (5.8) and (5.10):

$$
\begin{aligned}
\Big|\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}D^2_{\mathbf{x}_1}\tilde{P}_1u(t_0,\mathbf{x},\mathbf{x}') + \Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}D^2_{\mathbf{x}_1}\tilde{P}_{21}u(t_0,\mathbf{x},\mathbf{x}')\Big|\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})} &\le 2C(t_0-t)^{\frac{\gamma}{2}}\|u\|_{C^{2+\gamma}_{b,\mathbf{d}}}\\
&\le Cc_0^{\frac{\gamma}{2}}\|u\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}').
\end{aligned}
\tag{5.43}
$$

Let us now deal with the last term in (5.42). We proceed similarly to equation (5.11) (control of the frozen semigroup). Write:

$$
\begin{aligned}
&\Big|\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}D^2_{\mathbf{x}_1}\tilde{P}_{22}u(t_0,\mathbf{x},\mathbf{x}')\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})}\\
\le\quad&\Bigg\{\frac{1}{2}\int_{\mathbb{R}^{nd}}\frac{d\mathbf{y}}{(s-t)}\bar{p}^{\boldsymbol{\xi}'}_{C^{-1}}(t,t_0,\mathbf{x}',\mathbf{y})|D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')) - D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))||(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))_1|^2\\
&\quad+\frac{1}{2}\Big|\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}(t,t_0,\mathbf{x}',\mathbf{y})\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))_1^{\otimes 2}\Big)\\
&\quad-D^2_{\mathbf{x}_1}\tilde{p}^{\tilde{\boldsymbol{\xi}}'}(t,t_0,\mathbf{x}',\mathbf{y})\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))(\mathbf{y}-\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))_1^{\otimes 2}\Big)d\mathbf{y}\Big|\Bigg\}\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})}\\
\le\quad&C|D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')) - D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))|\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})},
\end{aligned}
$$

exploiting Proposition 3 for the first contribution and identity (3.17) in Proposition 5 for the second contribution in the last inequality.

From Lemma 8, we derive that for $t_0 - t = c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')$:

$$
\Big|\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}D^2_{\mathbf{x}_1}\tilde{P}_{22}u(t_0,\mathbf{x},\mathbf{x}')\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})} \le C\|u(t_0,\cdot)\|_{C^{2+\gamma}_{b,\mathbf{d}}}c_0^{\frac{\gamma}{2n-1}}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}').
\tag{5.44}
$$

Eventually, from (5.43) and (5.44), we get the following control:

$$
\big|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u(t_0,\mathbf{x}') - D^2_{\mathbf{x}_1}\tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u(t_0,\mathbf{x}')\big|\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x}')} \le Cc_0^{\frac{\gamma}{2n-1}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'),
$$

which gives the first statement of Lemma 17 for the second order derivatives w.r.t. the non-degenerate variables.

5.4.2. *Hölder controls for the degenerate variables.* Again, for the Hölder norm w.r.t. the degenerate variables, the difficulty is that we cannot take any advantage of cancellation tools. We adapt here the arguments employed in Section 5.1.2 for the frozen semigroup. Precisely, for any $(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2$, $\mathbf{x}_1=\mathbf{x}'_1$, $\boldsymbol{\xi}'=\mathbf{x}'$, $\tilde{\boldsymbol{\xi}}'=\mathbf{x}$, we have similarly to (5.42) (but without the spatial derivatives $D^2_{\mathbf{x}_1}$):

$$\begin{aligned}\tilde P^{\boldsymbol{\xi}'}_{t_0,t}u(t_0,\mathbf{x}')-\tilde P^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u(t_0,\mathbf{x}') &= \Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}\tilde P_1u(t_0,\mathbf{x}',\mathbf{x}')+\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}\tilde P_{21}u(t_0,\mathbf{x}',\mathbf{x}')\\ &\quad+\frac12\mathrm{Tr}\Big([D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))-D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))][\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{t_0,t}]_{1,1}\Big)\\ &\quad+\frac12\mathrm{Tr}\Big(D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))\big([\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{t_0,t}]_{1,1}-[\tilde{\mathbf{K}}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}]_{1,1}\big)\Big),\end{aligned}\tag{5.45}$$

where accordingly with (5.42):

$$\begin{aligned}\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}\tilde P_1u(t_0,\mathbf{x}',\mathbf{x}') &:= \int_{\mathbb{R}^{nd}}\tilde p^{\boldsymbol{\xi}'}(t,t_0,\mathbf{x}',\mathbf{y})[u(t_0,\mathbf{y})-u(t_0,\mathbf{y}_1,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')_{2:n})]d\mathbf{y}\\ &\quad-\int_{\mathbb{R}^{nd}}\tilde p^{\tilde{\boldsymbol{\xi}}'}(t,t_0,\mathbf{x}',\mathbf{y})[u(t_0,\mathbf{y})-u(t_0,\mathbf{y}_1,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')_{2:n})]d\mathbf{y},\end{aligned}$$

and

$$\begin{aligned}&\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}\tilde P_{21}u(t_0,\mathbf{x}',\mathbf{x}')\\ :=\ &\int_{\mathbb{R}^{nd}}\tilde p^{\boldsymbol{\xi}'}(t,t_0,\mathbf{x}',\mathbf{y})\int_0^1 d\mu(1-\mu)\\ &\mathrm{Tr}\Big(\big[D^2_{\mathbf{x}_1}u\big(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')_1+\mu(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))_1,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')_{2:n}\big)-D^2_{\mathbf{x}_1}u\big(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')\big)\big]\big(\mathbf{y}-\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}')\big)^{\otimes2}_1\Big)d\mathbf{y}\\ &-\int_{\mathbb{R}^{nd}}\tilde p^{\tilde{\boldsymbol{\xi}}'}(t,t_0,\mathbf{x}',\mathbf{y})\int_0^1 d\mu(1-\mu)\\ &\mathrm{Tr}\Big(\big[D^2_{\mathbf{x}_1}u\big(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')_1+\mu(\mathbf{y}-\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))_1,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')_{2:n}\big)-D^2_{\mathbf{x}_1}u\big(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')\big)\big]\big(\mathbf{y}-\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}')\big)^{\otimes2}_1\Big)d\mathbf{y}.\end{aligned}$$

Reproducing the arguments that led to equations (5.17) and (5.18), we derive:

$$\begin{aligned}&\big|\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}\tilde P_1u(t_0,\mathbf{x}',\mathbf{x}')+\Delta_{t,t_0,\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}\tilde P_{21}u(t_0,\mathbf{x}',\mathbf{x}')\big|\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})}\\ \le\ &C(t_0-t)^{\frac{2+\gamma}{2}}\|u\|_{C^{2+\gamma}_{b,\mathbf{d}}}\\ \le\ &Cc_0^{\frac{2+\gamma}{2}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').\end{aligned}\tag{5.46}$$

Let us now turn to the last two contributions in (5.45). As done in (5.20), from Proposition 2 and Lemma 8 we obtain:

$$\begin{aligned}&\frac12\big|[D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\boldsymbol{\xi}'}_{t_0,t}(\mathbf{x}'))-D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))][\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{t_0,t}]_{1,1}\big|\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})}\\ \le\ &C\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}(t_0-t)\mathbf{d}^\gamma(\mathbf{m}^{\mathbf{x}}_{t_0,t}(\mathbf{x}'),\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))\\ \le\ &Cc_0^{1+\frac{\gamma}{2n-1}}\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').\end{aligned}\tag{5.47}$$

The last term of (5.45) is handled like in (5.21). Namely, by Lemma 6 (equation (3.19)), we obtain:

$$\begin{aligned}&\big|\frac12D^2_{\mathbf{x}_1}u(t_0,\mathbf{m}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}(\mathbf{x}'))\big([\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{t_0,t}]_{1,1}-[\tilde{\mathbf{K}}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}]_{1,1}\big)\big|\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x}',\mathbf{x})}\\ \le\ &C\|u(t_0,\cdot)\|_{C^{2+\gamma}_{b,\mathbf{d}}}\big((t_0-t)^{\frac{2+\gamma}{2}}+(t_0-t)\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\big)\\ \le\ &Cc_0\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}').\end{aligned}\tag{5.48}$$

Plugging (5.46), (5.47) and (5.48) into (5.45), recalling as well that $c_0<1$, we derive the second statement of Lemma 17.

Remark 6 (Concluding remark concerning the *a priori* estimates)**.** *From the results of Sections 4 and 5, i.e. Proposition 9 for the supremum norms and Proposition 13 for the Hölder norms, we actually derive the following bound. There exist constants $C:=C((\mathbf{A}))$ and $\Lambda:=\Lambda((\mathbf{A}),T)$ as in Remark 3 s.t.*

$$\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\le C(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}+\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})})+\Big(\Lambda\big((c_0^{-(n-\frac12)+\frac\gamma2}+c_0^{\frac\gamma2})+T^{\frac\gamma2}\big)+Cc_0^{\frac{\gamma}{2n-1}}\Big)\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}.\tag{5.49}$$

The above estimate then readily yields (2.43) *provided* c_0 *and* Λ *are small enough (also with* $\Lambda \ll c_0$*). Recalling from Remark 3 that* Λ *is small provided the coefficients do not vary much, the remaining delicate part consists in getting rid of the small Hölder moduli constraint. This can be done through suitable scaling arguments that are exposed in Section 6.*

6. Scaling issues and final proof of Theorem 1

The purpose of this section is to first introduce a suitable scaling procedure for the system with mollified coefficients for which we will be able to show equation (5.49). This intuitively means that the scaling has to make the Hölder moduli of the considered coefficients small. The expected control is then obtained going back to the initial variables through the *inverse* scaling procedure. Also, once the estimate is established for *small* final time horizon T, it can be deduced through iteration up to an arbitrary given time precisely because it provides a kind of stability for the solution in the space $L^\infty([0,T], C^{2+\gamma}_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R}))$. We then conclude the proof of our main result, Theorem 1 through compactness arguments.

6.1. **Scaling settings and controls.** We start here from the smooth solution u to equation (1.1) with mollified coefficients (that we again denote by a slight abuse of notation without the mollifying parameter m). For an additional parameter $\lambda > 0$ to be specified later (but meant to be *small*), introducing the scaled function $u^\lambda(t,\mathbf{x}) = u(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x})$, it is then clear that this latter satisfies

$$\begin{cases} \partial_t u^\lambda(t,\mathbf{x}) + \langle \mathbf{F}(t,\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}), \lambda^{1/2}\mathbb{T}_\lambda^{-1}\mathbf{D}u^\lambda(t,\mathbf{x})\rangle \\ +\frac{\lambda^{-1}}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1} u^\lambda(t,\mathbf{x}) a(t,\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x})\big) = -f(t,\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}),\ (t,\mathbf{x}) \in [0,T)\times\mathbb{R}^{nd}, \\ u^\lambda(T,\mathbf{x}) = g(\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}),\ \mathbf{x}\in\mathbb{R}^{nd}. \end{cases} \tag{6.1}$$

This choice of rescaling is natural in view of the invariance by dilatation property (1.7), i.e. each single variable $\mathbf{x}_i$ is scaled by the parameter λ according to its corresponding intrinsic scale.

We rewrite in short form the above equation as for any $\mathbf{x}\in\mathbb{R}^{nd}$

$$\begin{cases} \partial_t u^\lambda(t,\mathbf{x}) + \langle \mathbf{F}_\lambda(t,\mathbf{x}), \mathbf{D}u^\lambda(t,\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1} u^\lambda(t,\mathbf{x}) a_\lambda(t,\mathbf{x})\big) = -f_\lambda(t,\mathbf{x}),\ t\in[0,T), \\ u^\lambda(T,\mathbf{x}) = g_\lambda(\mathbf{x}), \end{cases} \tag{6.2}$$

where

$$\begin{aligned} f_\lambda(t,\mathbf{x}) &:= f(t,\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}),\ g_\lambda(\mathbf{x}) := g(\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}), \\ a_\lambda(t,\mathbf{x}) &:= \lambda^{-1} a(t,\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}), \\ \mathbf{F}_\lambda(t,\mathbf{x}) &:= \lambda^{1/2}\mathbb{T}_\lambda^{-1}\mathbf{F}(t,\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}). \end{aligned} \tag{6.3}$$

Accordingly, we introduce the spatial operator $(L^\lambda_s)_{s\geq 0}$ appearing in (6.2) which writes explicitly for any $\varphi\in C^2_0(\mathbb{R}^{nd},\mathbb{R})$ as:

$$L^\lambda_t\varphi = \langle \mathbf{F}_\lambda(t,\cdot), \mathbf{D}\varphi\rangle + \frac{1}{2}\mathrm{Tr}\Big(a_\lambda(t,\cdot) D^2_{\mathbf{x}_1}\varphi\Big),\ \lambda>0.$$

The dynamics of the SDE associated with the second order differential operator $(L^\lambda_t)_{t\in[0,T]}$ appearing in (6.1)-(6.2) writes for a given starting point $(t,\mathbf{x})\in[0,T]\times\mathbb{R}^{nd}$:

$$\mathbf{X}^{\lambda,t,\mathbf{x}}_s = \mathbf{x} + \int_t^s \mathbf{F}_\lambda(u,\mathbf{X}^{\lambda,t,\mathbf{x}}_u)du + \int_t^s B\sigma_\lambda(u,\mathbf{X}^{\lambda,t,\mathbf{x}}_u)dW_u,\ s\geq t, \tag{6.4}$$

where $(W_u)_{u\geq 0}$ is a d-dimensional Brownian motion on some filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$ and σ_λ is a square root of the diffusion matrix a_λ introduced in (6.3).

Equation (6.4) then naturally leads to consider, for fixed $(s,\mathbf{y})\in[t,T]\times\mathbb{R}^{nd}$, and with the notations of Section 3, the corresponding linearized model

$$\begin{aligned} d\tilde{\mathbf{X}}^{\boldsymbol{\xi},\lambda}_v &= [\mathbf{F}_\lambda(v,\boldsymbol{\theta}^\lambda_{v,t}(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_\lambda(v,\boldsymbol{\theta}^\lambda_{v,t}(\boldsymbol{\xi}^\lambda))(\tilde{\mathbf{X}}^{\boldsymbol{\xi},\lambda}_v - \boldsymbol{\theta}^\lambda_{v,t}(\boldsymbol{\xi}^\lambda))]dv + B\sigma_\lambda(v,\boldsymbol{\theta}^\lambda_{v,t}(\boldsymbol{\xi}^\lambda))dW_v,\ \forall v\in[t,s] \\ \tilde{\mathbf{X}}^{\boldsymbol{\xi},\lambda}_t &= \mathbf{x}, \end{aligned} \tag{6.5}$$

where $\boldsymbol{\theta}^\lambda_{v,t}(\boldsymbol{\xi}^\lambda) = \lambda^{\frac{1}{2}}\mathbb{T}^{-1}_\lambda\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}^\lambda)$, $\boldsymbol{\xi}^\lambda = \lambda^{-\frac{1}{2}}\mathbb{T}_\lambda\boldsymbol{\xi}$. The associated generator writes for $\varphi\in C^2_0(\mathbb{R}^{nd},\mathbb{R})$ and $(s,\mathbf{y})\in[t,T]\times\mathbb{R}^{nd}$ as:

$$\tilde{L}^{\lambda,\boldsymbol{\xi}}_s\varphi(\mathbf{y}) = \langle \mathbf{F}_\lambda(s,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_\lambda(s,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))(\mathbf{y}-\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)), \mathbf{D}\varphi(\mathbf{y})\rangle + \frac{1}{2}\mathrm{Tr}\big(a_\lambda(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))D^2_{\mathbf{y}_1}\varphi(\mathbf{y})\big).$$

Observe from (2.6) and (6.5) that the following very important correspondence holds:

$$\forall v\in[t,T],\ \tilde{\mathbf{X}}^{\boldsymbol{\xi},\lambda}_v := \lambda^{1/2}\mathbb{T}^{-1}_\lambda\tilde{\mathbf{X}}^{\boldsymbol{\xi}^\lambda}_v. \tag{6.6}$$

We thus derive from Proposition 3 the following important correspondence for the densities. Denoting by $\tilde{p}^{\boldsymbol{\xi}}_{\lambda}(t,s,\mathbf{x},\mathbf{y})$ the density of $\tilde{\mathbf{X}}^{\boldsymbol{\xi},\lambda}_{v}$ starting from $\mathbf{x}$ at time t at point $\mathbf{y}$ in s, and $\mathbf{x}^{\lambda}=\lambda^{-\frac{1}{2}}\mathbb{T}_{\lambda}\mathbf{x}$, we have

$$\begin{aligned}\tilde{p}^{\boldsymbol{\xi}}_{\lambda}(t,s,\mathbf{x},\mathbf{y}) &= \lambda^{\frac{n^2d}{2}}\tilde{p}^{\boldsymbol{\xi}^{\lambda}}(t,s,\lambda^{-\frac{1}{2}}\mathbb{T}_{\lambda}\mathbf{x},\lambda^{-\frac{1}{2}}\mathbb{T}_{\lambda}\mathbf{y})\\ &= \frac{\lambda^{\frac{n^2d}{2}}}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}^{\boldsymbol{\xi}^{\lambda}}_{s,t})^{\frac{1}{2}}}\exp\left(-\frac{\lambda^{-1}}{2}\left\langle\left[\mathbb{T}_{\lambda}(\tilde{\mathbf{K}}^{\boldsymbol{\xi}^{\lambda}}_{s,t})^{-1}\mathbb{T}_{\lambda}\right](\lambda^{\frac{1}{2}}\mathbb{T}_{\lambda}^{-1}\mathbf{m}^{\boldsymbol{\xi}^{\lambda}}_{s,t}(\mathbf{x}^{\lambda})-\mathbf{y}),\lambda^{\frac{1}{2}}\mathbb{T}_{\lambda}^{-1}\mathbf{m}^{\boldsymbol{\xi}^{\lambda}}_{s,t}(\mathbf{x}^{\lambda})-\mathbf{y}\right\rangle\right).\end{aligned}$$

(6.7)

In particular, for $\boldsymbol{\xi}=\mathbf{x}$ one derives:

$$\begin{aligned}\tilde{p}^{\boldsymbol{\xi}}_{\lambda}(t,s,\mathbf{x},\mathbf{y}) &= \lambda^{\frac{n^2d}{2}}\tilde{p}^{\boldsymbol{\xi}^{\lambda}}(t,s,\lambda^{-\frac{1}{2}}\mathbb{T}_{\lambda}\mathbf{x},\lambda^{-\frac{1}{2}}\mathbb{T}_{\lambda}\mathbf{y})\\ &= \frac{\lambda^{\frac{n^2d}{2}}}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}^{\boldsymbol{\xi}^{\lambda}}_{s,t})^{\frac{1}{2}}}\exp\left(-\frac{\lambda^{-1}}{2}\left\langle\left[\mathbb{T}_{\lambda}(\tilde{\mathbf{K}}^{\boldsymbol{\xi}^{\lambda}}_{s,t})^{-1}\mathbb{T}_{\lambda}\right](\boldsymbol{\theta}^{\lambda}_{s,t}(\mathbf{x}^{\lambda})-\mathbf{y}),\boldsymbol{\theta}^{\lambda}_{s,t}(\mathbf{x}^{\lambda})-\mathbf{y}\right\rangle\right).\end{aligned}$$

(6.8)

Equation (6.7) in turn yields the following important control:

$$\begin{aligned}|D^{\vartheta}_{\mathbf{x}}\tilde{p}^{\boldsymbol{\xi}}_{\lambda}(t,s,\mathbf{x},\mathbf{y})| &\leq C\left(\frac{\lambda}{(s-t)}\right)^{\sum_{i=1}^{n}\vartheta_i(i-\frac{1}{2})+\frac{n^2d}{2}}\exp\left(-C^{-1}\frac{(s-t)}{\lambda}\big|\mathbb{T}^{-1}_{\frac{s-t}{\lambda}}\left(\lambda^{\frac{1}{2}}\mathbb{T}_{\lambda}^{-1}\mathbf{m}^{\boldsymbol{\xi}^{\lambda}}_{s,t}(\mathbf{x}^{\lambda})-\mathbf{y}\right)\big|^2\right)\\ &=: C\left(\frac{\lambda}{(s-t)}\right)^{\sum_{i=1}^{n}\vartheta_i(i-\frac{1}{2})}\bar{p}^{\boldsymbol{\xi}}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y}).\end{aligned}\tag{6.9}$$

In the following, we will also denote $\bar{p}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y}):=\bar{p}^{\boldsymbol{\xi}}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y})\Big|_{\boldsymbol{\xi}=\mathbf{x}}$ in (6.9).

Remark 7. *We emphasize that* (6.9) *gives that, each derivation of the Gaussian kernel $\tilde{p}^{\boldsymbol{\xi}}_{\lambda}$ makes the small parameter λ appear. Hence, up to the additional time singularities, the iterated derivatives become smaller and smaller.*

6.2. **Scaling properties.** The point is now to reproduce the previous perturbative approach for the solution of (6.2) in order to obtain ad hoc versions of Propositions 9 (Section 4) and 13 (Section 5). Such versions then allow us to derive the analogous control of (5.49) in this rescaled setting involving precisely a positive exponent of λ in front of the term $\Lambda c_0^{-(n-\frac{1}{2})+\frac{\gamma}{2}}$. As underlined in Remark 6, this type of control will give the expected final bound provided the scaling parameter λ is small enough.

We will mainly focus on the Hölder norm of the remainder term (i.e. rescaled version of estimate (5.30) in Lemma 16) associated with the second order derivatives w.r.t. the non-degenerate variables. It can indeed be seen from the previous computations that the other contributions can be dealt similarly.

Hölder estimate on the rescaled remainder. According to the notations introduced in the beginning of Section 5, we denote for $\lambda>0$, $(t,\mathbf{x},\mathbf{x}',\mathbf{x}')\in[0,T]\times(\mathbb{R}^{nd})^2$ the quantity:

$$\begin{aligned}D^2_{\mathbf{x}_1}(\Delta^{\lambda})^{\boldsymbol{\xi},\boldsymbol{\xi}'}(t,\mathbf{x},\mathbf{x}') &:= \left(\int_t^{(t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}'))\wedge T}ds\int_{\mathbb{R}^{nd}}\Big(D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}_{\lambda}(t,s,\mathbf{x},\mathbf{y})(L^{\lambda}_s-\tilde{L}^{\lambda,\boldsymbol{\xi}}_s)u^{\lambda}(s,\mathbf{y})\right.\\ &\quad\left.-D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}_{\lambda}(t,s,\mathbf{x}',\mathbf{y})(L^{\lambda}_s-\tilde{L}^{\lambda,\boldsymbol{\xi}'}_s)u^{\lambda}(s,\mathbf{y})\Big)d\mathbf{y}\right)\Bigg|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}\\ &\quad+\left(\int_{(t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}'))\wedge T}^{T}ds\int_{\mathbb{R}^{nd}}\Big(D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}_{\lambda}(t,s,\mathbf{x},\mathbf{y})(L^{\lambda}_s-\tilde{L}^{\lambda,\boldsymbol{\xi}}_s)u^{\lambda}(s,\mathbf{y})\right.\\ &\quad\left.-D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}'}_{\lambda}(t,s,\mathbf{x}',\mathbf{y})(L^{\lambda}_s-\tilde{L}^{\lambda,\boldsymbol{\xi}'}_s)u^{\lambda}(s,\mathbf{y})\Big)d\mathbf{y}\right)\Bigg|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x})}\\ &=: D^2_{\mathbf{x}_1}(\Delta^{\lambda})^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}')+D^2_{\mathbf{x}_1}(\Delta^{\lambda})^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{diag}}(t,\mathbf{x},\mathbf{x}'),\end{aligned}\tag{6.10}$$

using again as in (5.4) different freezing points according to the spatial regime w.r.t. integration time s. Note however carefully that the cutting threshold here depends on the scaling parameter λ. This is very important in order to balance the various scales that will appear. Pay attention as well that the parameter c_0 also remains. Actually, a subtle balance between those two parameters will be needed to derive the expected control.

Proceeding as in Section 5.3, we aim at showing that there exist constants $C := C((\mathbf{A}), T)$ and $\Lambda := \Lambda((\mathbf{A}), T)$ as in Remark 3 s.t. for any $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$:

$$\sup_{t\in[0,T]} \frac{|D^2_{\mathbf{x}_1}(\Delta^\lambda)^{\boldsymbol{\xi},\boldsymbol{\xi}'}(t,\mathbf{x},\mathbf{x}')|}{\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')} \leq \Big(\Lambda(c_0^{-(n-\frac{1}{2})+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}})\lambda^{\frac{\gamma}{2}}\Big)\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}. \tag{6.11}$$

Proof. Let us first consider the diagonal term in (6.10) assuming w.l.o.g. that $t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}') \leq T$ (otherwise we only have the off-diagonal contribution) and recalling that in the *diagonal regime* we chose $\boldsymbol{\xi}' = \boldsymbol{\xi} = \mathbf{x}$. Write:

$$\begin{aligned}
&|D^2_{\mathbf{x}_1}(\Delta^\lambda)^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{diag}}(t,\mathbf{x},\mathbf{x}')|\Big|_{\boldsymbol{\xi}'=\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & \Big|\int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\Big[D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x},\mathbf{y}) - D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}',\mathbf{y})\Big]\Big[\langle \mathbf{F}_{\lambda,1}(t,\mathbf{y}) - \mathbf{F}_{\lambda,1}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)), D_{\mathbf{y}_1}\rangle\\
&\quad +\frac{1}{2}\mathrm{Tr}\Big(\big(a_\lambda(t,\mathbf{y}) - a_\lambda(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))\big)D^2_{\mathbf{y}_1}\Big)\Big]u^\lambda(s,\mathbf{y})\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
&+\Big|\sum_{i=2}^n\int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\Big[D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x},\mathbf{y}) - D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}',\mathbf{y})\Big]\\
&\quad \Big\langle \mathbf{F}_{\lambda,i}(t,\mathbf{y}) - [\mathbf{F}_{\lambda,i}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_{\lambda,i}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))(\mathbf{y}-\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))_{i-1}], D_{\mathbf{y}_i}u^\lambda(s,\mathbf{y})\Big\rangle\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
=\ & \Big|\int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\int_0^1 d\mu \mathbf{D}D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}-\mathbf{x}')\\
&\quad \Big(\Big\langle \mathbf{F}_{\lambda,1}(t,\mathbf{y}) - \mathbf{F}_{\lambda,1}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)), D_{\mathbf{y}_1}\Big\rangle + \frac{1}{2}\mathrm{Tr}\Big(\big(a_\lambda(t,\mathbf{y}) - a_\lambda(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))\big)D^2_{\mathbf{y}_1}\Big)\Big)u^\lambda(s,\mathbf{y})\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
&+\Big|\sum_{i=2}^n\int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\int_0^1 d\mu \mathbf{D}D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\cdot(\mathbf{x}-\mathbf{x}')\\
&\quad \Big\langle \mathbf{F}_{\lambda,i}(t,\mathbf{y}) - [\mathbf{F}_{\lambda,i}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_{\lambda,i}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))(\mathbf{y}-\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)_{i-1})], D_{\mathbf{y}_i}u^\lambda(s,\mathbf{y})\Big\rangle\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
=:\ & \Big\{D^2_{\mathbf{x}_1}(\Delta^\lambda)^{\boldsymbol{\xi},1}_{\mathbf{diag}}(t,\mathbf{x},\mathbf{x}') + D^2_{\mathbf{x}_1}(\Delta^\lambda)^{\boldsymbol{\xi},2:n}_{\mathbf{diag}}(t,\mathbf{x},\mathbf{x}')\Big\}\Big|_{\boldsymbol{\xi}=\mathbf{x}}.
\end{aligned} \tag{6.12}$$

We will now control the first term of the above right hand side. A key point for the analysis is to observe that, on the considered diagonal regime, we actually have from equations (6.8)-(6.9), recalling that $\mathbf{z} \mapsto \mathbf{m}^{\boldsymbol{\xi}^\lambda}_{s,t}(\mathbf{z})$ is affine and using the good scaling property in (3.5) and (3.11), that:

$$\begin{aligned}
|D_{\mathbf{x}_k}D^2_{\mathbf{x}_1}\tilde p^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})| &\leq \frac{C\lambda^{1+(k-\frac{1}{2})}}{(s-t)^{1+(k-\frac{1}{2})}}\bar p_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y})\exp\Big(\sum_{j=1}^n\frac{|(\mathbf{x}-\mathbf{x}')_j|^2\lambda^{2j-1}}{(s-t)^{2j-1}}\Big)\\
&\leq \frac{C\lambda^{1+(k-\frac{1}{2})}}{(s-t)^{1+(k-\frac{1}{2})}}\bar p_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y}).
\end{aligned} \tag{6.13}$$

To obtain the last inequality, we indeed observe from the homogeneity of the metric (see (1.9)) that $\lambda^{1/2}\mathbf{d}(\mathbf{x},\mathbf{x}') = \mathbf{d}\big(\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}, \lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x}'\big) = \sum_{j=1}^n \big|(\mathbf{x}-\mathbf{x}')_j\lambda^{(2j-1)/2}\big|^{1/(2j-1)}$. For the diagonal regime $\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}') \leq (s-t)$ in turn implies that for each $j \in [\![1,n]\!]$, $(s-t)^{-2j+1}|(\mathbf{x}-\mathbf{x}')_j|^2\lambda^{2j-1} \leq 1$.

Another key point is to observe that the contribution $(s-t)^{-1/2}\mathbf{d}\big(\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{y}, \lambda^{-1/2}\mathbb{T}_\lambda\boldsymbol{\theta}^\lambda_{s,t}(\mathbf{x}^\lambda)\big)$ is homogeneous to the argument of the exponential term in $\bar p_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y})$. Namely, for any given $\beta_0 > 0$ and $\beta \in (0,\beta_0]$, there exists C_{β_0} s.t.

$$\left(\frac{\mathbf{d}\big(\lambda^{-\frac{1}{2}}\mathbb{T}_\lambda\mathbf{y}, \lambda^{-\frac{1}{2}}\mathbb{T}_\lambda\boldsymbol{\theta}^\lambda_{s,t}(\mathbf{x}^\lambda)\big)}{(s-t)^{\frac{1}{2}}}\right)^\beta \bar p_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y}) \leq C_{\beta_0}\bar p_{C^{-1}_{\beta_0},\lambda}(t,s,\mathbf{x},\mathbf{y}), \tag{6.14}$$

Equation (6.14) is a direct consequence of the expression of $\bar p^{t,\mathbf{x}}_{C^{-1},\lambda}$ in Proposition 3 (see also (6.9)) and the definition of $\mathbf{d}$ in (1.9).

Hence, from the definition of $a_\lambda, \mathbf{F}_\lambda$ in (6.3), equations (6.13) and (6.14), we derive that, under **(A)**, for each $k \in [\![1,n]\!]$:

$$\begin{aligned}
&\Big\{|\mathbf{F}_{\lambda,1}(t,\mathbf{y}) - \mathbf{F}_{\lambda,1}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))||D_{\mathbf{x}_k} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})|\Big\}\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & C\Big\{\Big(\lambda^{-\frac12}\|\mathbf{F}_1(t,\cdot)\|_{C^\gamma_{\mathbf{d}}}\mathbf{d}^\gamma\big(\lambda^{-\frac12}\mathbb{T}_\lambda\mathbf{y},\lambda^{-\frac12}\mathbb{T}_\lambda\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)\big)\Big)\lambda^{1+k-\frac12}(s-t)^{-1-(k-\frac12)}\bar{p}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y})\Big\}\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
(6.15)\leq\ & C\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})}\lambda^k(s-t)^{-1-(k-\frac12)+\frac\gamma2}\bar{p}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y}).
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\Big\{|a_\lambda(t,\mathbf{y}) - a_\lambda(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))|D_{\mathbf{x}_k} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})|\Big\}\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & C\Big\{\Big(\lambda^{-1}\|a(t,\cdot)\|_{C^\gamma_{\mathbf{d}}}\mathbf{d}^\gamma\big(\lambda^{-\frac12}\mathbb{T}_\lambda\mathbf{y},\lambda^{-\frac12}\mathbb{T}_\lambda\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)\big)\Big)\lambda^{1+(k-\frac12)}(s-t)^{-1-(k-\frac12)}\bar{p}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y})\Big\}\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
(6.16)\leq\ & C\|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}\lambda^{k-\frac12}(s-t)^{-1-(k-\frac12)+\frac\gamma2}\bar{p}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y}).
\end{aligned}$$

Observe that both r.h.s. of (6.15) and (6.16) exhibit a positive power of λ. Hence, those quantities are *small* provided λ is. The key point in the above computations is that the potentially explosive Hölder norms of $\mathbf{F}_{\lambda,1}, a_\lambda$(when λ goes to zero) are compensated by the derivatives of $\tilde{p}^{\boldsymbol{\xi}}_\lambda(s,t,\mathbf{x},\mathbf{y})$ (see equation (6.9)). We again carefully point-out that the previous bounds only depend on the supremum in time of the Hölder moduli of the coefficients (denoted $\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})}, \|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}$ respectively) and not on their supremum norm. In particular, we get from (6.15), (6.16) with the notation of (6.12):

$$\begin{aligned}
&|D^2_{\mathbf{x}_1}(\Delta^\lambda)^{\boldsymbol{\xi},1}_{\mathbf{diag}}(t,\mathbf{x},\mathbf{x}')|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & \sum_{k=1}^n\int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x})}^T ds\int_{\mathbb{R}^{nd}} d\mathbf{y}\lambda^{k-\frac12}(s-t)^{-1-(k-\frac12)+\gamma/2}\\
&\times[\lambda^{\frac12}\|D_{\mathbf{x}_1}u^\lambda\|_{L^\infty}\|\mathbf{F}_1\|_{L^\infty(C^\gamma_{\mathbf{d}})} + \|D^2_{\mathbf{x}_1}u^\lambda\|_{L^\infty}\|a\|_{L^\infty(C^\gamma_{\mathbf{d}})}]\bar{p}_{C^{-1},\lambda}(t,s,\mathbf{x},\mathbf{y})|\mathbf{x}_k-\mathbf{x}'_k|\\
(6.17)\qquad \leq\ & c_0^{-(n-\frac12)+\frac\gamma2}\Lambda(\|D_{\mathbf{x}_1}u^\lambda\|_{L^\infty}+\|D^2_{\mathbf{x}_1}u^\lambda\|_{L^\infty})\lambda^{\frac\gamma2}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}').
\end{aligned}$$

Thanks to the inequality (4.24) and the previous analysis (to be performed according to the current scaling procedure replacing $(s-t)$ in Sections 4 and 5 by $(s-t)/\lambda$), we deduce:

$$\begin{aligned}
&|D^2_{\mathbf{x}_1}(\Delta^\lambda)^{\boldsymbol{\xi},2:n}_{\mathbf{diag}}(t,\mathbf{x},\mathbf{x}')|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & \Big|\sum_{i=2}^n\int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\int_{\mathbb{R}^{nd}}d\mathbf{y}\int_0^1 d\mu\Big\langle D_{\mathbf{y}_i}\Big\{\Big(\mathbf{F}_{\lambda,i}(t,\mathbf{y}) - [\mathbf{F}_{\lambda,i}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_{\lambda,i}(t,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))(\mathbf{y}-\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}^\lambda))_{i-1}]\Big)\\
&\times\Big(\big\langle\mathbf{x}-\mathbf{x}', D_{\mathbf{x}}D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x}+\mu(\mathbf{x}'-\mathbf{x}),\mathbf{y})\big\rangle\Big)\Big\}, u^\lambda(s,\mathbf{y})\Big\rangle\Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}}\\
\leq\ & \Lambda\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n\int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds\frac{\lambda^{k-\frac12}|\mathbf{x}_k-\mathbf{x}'_k|}{(s-t)^{1+(k-\frac12)-\frac\gamma2}}\\
\leq\ & \Lambda\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n c_0^{-(k-\frac12)+\frac\gamma2}\lambda^{(k-\frac12)-(k-\frac12)+\frac\gamma2}\mathbf{d}^{\gamma-(2k-1)}(\mathbf{x},\mathbf{x}')|\mathbf{x}_k-\mathbf{x}'_k|\\
\leq\ & \Lambda\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\sum_{k=1}^n c_0^{-(k-\frac12)+\frac\gamma2}\lambda^{\frac\gamma2}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}') \leq \Big(\Lambda c_0^{-(n-\frac12)+\frac\gamma2}\lambda^{\frac\gamma2}\Big)\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}').
\end{aligned}$$

(6.18)

Plugging (6.17) and (6.18) into (6.12) gives a diagonal control which precisely matches the r.h.s. of the expected final bound (6.11).

It therefore remains to handle the off-diagonal contributions. With the notations of (6.10), we deduce from the analysis of Section 5.3 (recall that in that case we chose $(\boldsymbol{\xi},\boldsymbol{\xi}') = (\mathbf{x},\mathbf{x}')$ as freezing points) and the previous arguments that:

$$\begin{aligned}
&|D^2_{\mathbf{x}_1}(\Delta^\lambda)^{\boldsymbol{\xi},\boldsymbol{\xi}'}_{\mathbf{off\text{-}diag}}(t,\mathbf{x},\mathbf{x}')|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}')=(\mathbf{x},\mathbf{x}')}\\
(6.19)\qquad \leq\ & \Lambda\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\int_t^{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}\frac{ds}{(s-t)^{1-\frac\gamma2}} \leq \Big(\Lambda c_0^{\frac\gamma2}\lambda^{\frac\gamma2}\Big)\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}').
\end{aligned}$$

In some sense the controls in (6.18) and (6.19) can be seen as a mere consequence of the intrinsic scaling of the system. This is indeed the case, but, in order to avoid any ambiguity, we provide in Appendix C a proof of the rescaled key Besov control of Lemma 11. The bound of equation (6.19) completes the proof of (6.11). □

Now, using the notations of (6.10), it can be deduced from the same procedure (exploiting the control (B.13) for the difference of the scaled covariances) that the following rescaled version of estimate (5.31) in Lemma 5 holds:

$$\sup_{t\in[0,T],(\mathbf{x},\mathbf{x}')\in(\mathbb{R}^{nd})^2,\ \mathbf{x}_1=\mathbf{x}_1'} \frac{|(\Delta^\lambda)^{\boldsymbol{\xi},\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}'}(t,\mathbf{x},\mathbf{x}')|\Big|_{(\boldsymbol{\xi},\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x}'\mathbf{x})}}{\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}')} \leq \Lambda\lambda^{\frac{\gamma}{2}}(c_0^{-(n-\frac{1}{2})+\frac{\gamma}{2}}+c_0)\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}. \tag{6.20}$$

Hölder estimate for the semi group at the discontinuity point. We would also derive from (5.45) and the same previous arguments by denoting $t_0 = t + c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')$ and using the bound of (A.15) for the difference of the scaled flows, that the analogous of estimates in Lemma 14 :

$$\big|D^2_{\mathbf{x}_1}\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u^\lambda(t_0,\mathbf{x}') - D^2_{\mathbf{x}_1}\tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u^\lambda(t_0,\mathbf{x}')\big|\Big|_{(\boldsymbol{\xi}',\tilde{\boldsymbol{\xi}}')=(\mathbf{x},\mathbf{x}')} \leq Cc_0^{\frac{\gamma}{2n-1}}\|u^\lambda\|_{C^{2+\gamma}_{b,\mathbf{d}}}\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}'). \tag{6.21}$$

and

$$\big|\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u^\lambda(t_0,\mathbf{x}') - \tilde{P}^{\tilde{\boldsymbol{\xi}}'}_{t_0,t}u^\lambda(t_0,\mathbf{x}')\big| \quad\leq\quad Cc_0\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'),\ \text{for } \mathbf{x}_1=\mathbf{x}_1',$$

hold. Note that above we denoted with a slight abuse of notation

$$\tilde{P}^{\boldsymbol{\xi}'}_{t_0,t}u^\lambda(t_0,\mathbf{x}') = \int_{\mathbb{R}^{nd}} d\mathbf{y}\tilde{p}^{\boldsymbol{\xi}'}_\lambda(t,t_0,\mathbf{x}',\mathbf{y})u^\lambda(t_0,\mathbf{y}).$$

Hölder estimates for the frozen semi-group and associated Green kernel. It easily follows from (1.8) that, in this rescaled setting, the ad hoc estimates in Lemmas 14 and 15 remain valid (i.e. with f_λ, g_λ intead of f, g and for the Green and frozen kernel associated with the rescaled system (6.1) therein).

Estimates of the supremum of the derivates w.r.t. the non degenerate variable.
We importantly point out that for the Hölder moduli, we have benefitted from a regularizing effect in the scaling parameter λ. Note carefully that this is not the case as far as supremum derivatives are concerned. We indeed get, using the same previous arguments, that

$$\begin{aligned}|D^2_{\mathbf{x}_1}u^\lambda(t,\mathbf{x})| &= |\int_t^T ds\int_{\mathbb{R}^{nd}} D^2_{\mathbf{x}_1}\tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x},\mathbf{y})(L^\lambda_s-\tilde{L}^{\lambda,\boldsymbol{\xi}}_s)u^\lambda(s,\mathbf{y})d\mathbf{y}|\\ &\leq C\big(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}+(T-t)^{\frac{\gamma}{2}}\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}\big)+\Lambda(T-t)^{\frac{\gamma}{2}}\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}.\end{aligned}$$

The difference w.r.t. e.g. equations (6.17), (6.18) is that we have no additional spatial differentiations and no cutting threshold in time which precisely allowed to make the scaling parameter appear. Equivalently, we have to integrate in time the control (6.19) of the previous off-diagonal regime but on the whole time interval $[t,T]$.

Conclusion: Schauder estimate for the rescaled system. Gathering the above estimates, we eventually derive similarly to (5.49), recalling that $c_0 < 1$, that:

$$\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \leq C(\|g_\lambda\|_{C^{2+\gamma}_{b,\mathbf{d}}}+\|f_\lambda\|_{L^\infty(C^\gamma_{b,\mathbf{d}})})+\Big[\Lambda\big(\lambda^{\frac{\gamma}{2}}(c_0^{-(n-\frac{1}{2})}+c_0^{\frac{\gamma}{2}})+T^{\frac{\gamma}{2}}\big)+Cc_0^{\frac{1}{2n-1}}\Big]\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})}. \tag{6.22}$$

Therefore, for T, $c_0 := c_0((\mathbf{A}),T)$ and $\lambda := \lambda((\mathbf{A}),T)$ small enough, with $\lambda \ll c_0$, i.e. such that $\bar{c}_0 := \Big[\Lambda\big(\lambda^{\frac{\gamma}{2}}(c_0^{-(n-\frac{1}{2})}+c_0^{\frac{\gamma}{2}})+T^{\frac{\gamma}{2}}\big)+Cc_0^{\frac{1}{2n-1}}\Big] < 1$, the expect final control holds:

$$\|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \leq \frac{C}{1-\bar{c}_0}(\|g_\lambda\|_{C^{2+\gamma}_{b,\mathbf{d}}}+\|f_\lambda\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}). \tag{6.23}$$

7. Conclusion: final proof of Theorem 1.

Equation (6.23) provides the expected Schauder estimate for the rescaled system with mollified coefficients for a *small* time horizon $T > 0$. The scaling parameter λ must precisely be tuned w.r.t. Λ (associated with the Hölder moduli of the coefficients, see Remark 3) in order that the above constant $\bar{c}_0$ be strictly less than one. Recalling that $u^\lambda(t,\mathbf{x}) = u(t,\lambda^{-1/2}\mathbb{T}_\lambda\mathbf{x})$, we then derive from (6.23) and (6.3) that for $T > 0$ and λ small enough there exists $\bar{C}_0 := \bar{C}_0((\mathbf{A}),T,\lambda) > 1$ s.t.

$$\|u\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} \leq \bar{C}_0(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}}+\|f\|_{L^\infty(C^\gamma_{b,\mathbf{d}})}), \tag{7.24}$$

which precisely provides the required estimate for the initial system with mollified coefficients.

The point is now to extend the previous bound to an arbitrary fixed time $T > 0$ not necessarily small. The stability resulting from estimate (7.24) allows to proceed by simple iterative application of the bound changing the final condition.

7.1. **Schauder estimates for the mollified system for a general time.** Equation (7.24) is valid for any $T < T_0$ with $T_0 \in (0, +\infty)$ sufficiently small. Now, for a given $T > 0$ (not necessary small), we can solve iteratively $N = \lceil \frac{T}{T_0} \rceil$ (where $\lceil \cdot \rceil$ is the ceiling function) Cauchy problems in the following way. Consider first:

$$\begin{cases} \partial_t u_{(1)}(t,\mathbf{x}) + \langle \mathbf{F}(t,\mathbf{x}), \mathbf{D}u_{(1)}(t,\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1} u_{(1)}(t,\mathbf{x}) a(t,\mathbf{x})\big) = -f(t,\mathbf{x}),\ (t,\mathbf{x}) \in [T(1-\frac{1}{N}), T) \times \mathbb{R}^{nd}, \\ u_{(1)}(T,\mathbf{x}) = g(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^{nd}. \end{cases}$$

In other words, our previous analysis, and the previous inequalities, are still available for $T - (1 - \frac{1}{N})T = \frac{1}{N}T \le T_0$ small enough. Precisely, from (7.24):

$$\|u_{(1)}\big(T(1-\frac{1}{N}),\cdot\big)\|_{C^{2+\gamma}_{b,\mathbf{d}}} \le \bar{C}_0(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty([T(1-\frac{1}{N}),T],C^\gamma_{b,\mathbf{d}})}). \tag{7.25}$$

Also, for mollified coefficients, it is plain from the Feynman-Kac formula to identify $u_{(1)}$ and u on $[T(1-\frac{1}{N}), T]$ where u solves (1.1) with mollified coefficients on $[0, T]$. Hence, (7.25) gives in particular that $u_{(1)}\big(T(1-\frac{1}{N}),\cdot\big) = u\big(T(1-\frac{1}{N}),\cdot\big) \in C^{2+\gamma}_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R})$ so that it is in particular natural to consider the following second Cauchy problem for any $\mathbf{x} \in \mathbb{R}^{nd}$:

$$\begin{cases} \partial_t u_{(2)}(t,\mathbf{x}) + \langle \mathbf{F}(t,\mathbf{x}), \mathbf{D}u_{(2)}(t,\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1} u_{(2)}(t,\mathbf{x}) a(t,\mathbf{x})\big) = -f(t,\mathbf{x}),\ t \in [(1-\frac{2}{N})T, (1-\frac{1}{N})T), \\ u_{(2)}((1-\frac{1}{N})T,\mathbf{x}) = u((1-\frac{1}{N})T,\mathbf{x}). \end{cases}$$

Hence $u_{(2)}$ satisfies identity (7.24) for the corresponding time interval and the associated source and final condition. It also coincides with u on $[T(1-\frac{2}{N}), T(1-\frac{1}{N})]$. We get:

$$\begin{aligned} & \|u_{(2)}\|_{L^\infty([T(1-\frac{2}{N}),T(1-\frac{1}{N})],C^{2+\gamma}_{b,\mathbf{d}})} \\ \le\ & \bar{C}_0(\|u(T(1-\frac{1}{N}),\cdot)\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty([T(1-\frac{2}{N}),T(1-\frac{1}{N})],C^\gamma_{b,\mathbf{d}})}) \\ \le\ & \bar{C}_0\Big(\bar{C}_0(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty([T(1-\frac{1}{N}),T],C^\gamma_{b,\mathbf{d}})}) + \|f\|_{L^\infty([T(1-\frac{2}{N}),T(1-\frac{1}{N})],C^\gamma_{b,\mathbf{d}})}\Big) \\ \le\ & \bar{C}_0^2(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty([T(1-\frac{2}{N}),T],C^\gamma_{b,\mathbf{d}})}). \end{aligned}$$

Repeating the analysis N-times, introducing for $k \in [\![3, n]\!]$ the auxiliary Cauchy problems for any $\mathbf{x} \in \mathbb{R}^{nd}$:

$$\begin{cases} \partial_t u_{(k)}(t,\mathbf{x}) + \langle \mathbf{F}(t,\mathbf{x}), \mathbf{D}u_{(k)}(t,\mathbf{x})\rangle + \frac{1}{2}\mathrm{Tr}\big(D^2_{\mathbf{x}_1} u_{(k)}(t,\mathbf{x}) a(t,\mathbf{x})\big) = -f(t,\mathbf{x}),\ t, \in [T(1-\frac{k}{N}), T(1-\frac{k-1}{N})), \\ u_{(k)}((1-\frac{k-1}{N})T,\mathbf{x}) = u((1-\frac{k-1}{N})T,\mathbf{x}), \end{cases}$$

we derive that

$$\|u\|_{L^\infty([0,T],C^{2+\gamma}_{b,\mathbf{d}})} \le \bar{C}_0^N(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty([0,T],C^\gamma_{b,\mathbf{d}})}). \tag{7.26}$$

This precisely gives our main estimate for the system (1.1) with mollified coefficients. Again, even though the coefficients are smooth, all the constants appearing in (7.26) only depend on the Hölder setting of assumption **(A)**.

Remark 8 (About the constants in the final estimate)**.** *We could actually have slightly better bounds than those in* (7.26)*. Namely a direct induction shows that the following control holds.*

$$\|u\|_{L^\infty([0,T],C^{2+\gamma}_{b,\mathbf{d}})} \le \bar{C}_0^N\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \sum_{i=1}^N C_0^i\|f\|_{L^\infty([T(1-\frac{N-i+1}{N}),T(1-\frac{N-i}{N})],C^\gamma_{b,\mathbf{d}})}). \tag{7.27}$$

We chose to write the bound in the form of equation (7.26) *for simplicity. Note however that in any case, equation* (7.26) *or* (7.27)*, we still have geometric constants coming from the iterative procedure. This is the main drawback of our approach, which anyhow seems, to the best of our knowledge, the only available one to consider degenerate Kolmogorov systems with non-linear drifts.*

7.2. **Compactness arguments.** We now make the mollifying parameter appear again using the notations introduced in the detailed guide (see equation (2.4)). Equation (7.24) rewrites in the following way. There exists a constant C_0 s.t. for each $m \in \mathbb{N}$:

$$\|u_m\|_{L^\infty([0,T],C^{2+\gamma}_{b,\mathbf{d}})} \leq \bar{C}_0(\|g\|_{C^{2+\gamma}_{b,\mathbf{d}}} + \|f\|_{L^\infty([0,T],C^{\gamma}_{b,\mathbf{d}})}). \tag{7.28}$$

From Ascoli's theorem we deduce that there exists a sequence $(u_{m_k})_{k\geq 1}$ of smooth solutions of (2.4), with $m_k \to_k +\infty$, satisfying (7.28) and s.t. $u_{m_k} \to_k u$ in $L^\infty([0,T], C^{2+\gamma-\varepsilon}_{b,\mathbf{d}})$, for all $0 < \varepsilon < \gamma$. We then deduce from the previous analysis that such a u also satisfies (7.28), and therefore lies in $L^\infty([0,T], C^{2+\gamma}_{b,\mathbf{d}})$. Since we also have from [4] that $u_{m_k}(t,\mathbf{x}) \to_k \mathbb{E}[g(\mathbf{X}^{t,\mathbf{x}}_T)] + \int_t^T \mathbb{E}[f(s,\mathbf{X}^{t,\mathbf{x}}_s)]ds$ where $(\mathbf{X}^{t,\mathbf{x}}_s)$ denotes the unique in law solution to (1.4), we deduce that $u(t,\mathbf{x}) = \mathbb{E}[g(\mathbf{X}^{t,\mathbf{x}}_T)] + \int_t^T \mathbb{E}[f(s,\mathbf{X}^{t,\mathbf{x}}_s)]ds$ corresponds to the *unique* martingale, or *mild*, solution of (1.1) which satisfies the stated Schauder estimate (7.24).

7.3. **From mild to weak solutions.** For this paragraph, we will consider the previous converging subsequence obtained from the compactness argument and will still denote it by $(u_m)_{m\in\mathbb{N}}$ for notational simplicity. Namely, $\|u - u_m\|_{L^\infty(C^{2+\gamma-\varepsilon}_{b,\mathbf{d}})} \underset{m}{\longrightarrow} 0,\ \varepsilon \in (0,\gamma)$.

Let φ be a smooth given function with compact support, i.e. $\varphi \in C^\infty_0(\mathbb{R}^{nd},\mathbb{R})$. It is clear that for the solution of (1.1) with mollified coefficients one indeed has:

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f_m(t,\mathbf{x})\varphi(t,\mathbf{x}) = \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}\varphi(t,\mathbf{x})\Big(\partial_t + L^m_t\Big)u_m(t,\mathbf{x}). \tag{7.29}$$

Indeed, both the solution and the coefficients are smooth. Integrating by parts yields:

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f_m(t,\mathbf{x})\varphi(t,\mathbf{x}) = \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}\Big(-\partial_t + (L^m_t)^*\Big)\varphi(t,\mathbf{x})u_m(t,\mathbf{x}),$$

where $(L^m_t)^*$ denotes the adjoint of L^m_t.

Write now:

$$\begin{aligned}\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f_m(t,\mathbf{x})\varphi(t,\mathbf{x}) &= \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f(t,\mathbf{x})\varphi(t,\mathbf{x}) + \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}(f_m - f)(t,\mathbf{x})\varphi(t,\mathbf{x}) \\ &=: \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f(t,\mathbf{x})\varphi(t,\mathbf{x}) + R_m(T,f).\end{aligned} \tag{7.30}$$

It is clear that under **(A)**, recall that $f \in L^\infty([0,T], C^\gamma_{b,\mathbf{d}}(\mathbb{R}^{nd},\mathbb{R}))$, $R_m(T,f) \underset{m}{\longrightarrow} 0$. On the other hand, we now decompose:

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}\Big(-\partial_t + (L^m_t)^*\Big)\varphi(t,\mathbf{x})u_m(t,\mathbf{x}) = \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}\Big(-\partial_t + (L_t)^*\Big)\varphi(t,\mathbf{x})u(t,\mathbf{x}) + R_m(T,u), \tag{7.31}$$

with

$$\begin{aligned}&R_m(T,u) \\ :=\ & \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}\Big((L^m_t)^* - L^*_t\Big)\varphi(t,\mathbf{x})u_m(t,\mathbf{x}) + \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}\Big(-\partial_t + (L_t)^*\Big)\varphi(t,\mathbf{x})(u_m(t,\mathbf{x}) - u(t,\mathbf{x})) \\ =:\ & (R^1_m + R^2_m)(T,u),\end{aligned} \tag{7.32}$$

where L^*_t is the formal adjoint of L_t. Observe first that:

$$R^2_{m,0}(T,u) := \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x}\partial_t\varphi(t,\mathbf{x})\big(u_m(t,\mathbf{x}) - u(t,\mathbf{x})\big) \underset{m}{\longrightarrow} 0,$$

since $\|u - u_m\|_{L^\infty(C^{2+\gamma-\varepsilon}_{b,\mathbf{d}})} \underset{m}{\longrightarrow} 0$ for any $0 < \varepsilon < \gamma$. For the terms of $R^2_m(T,u)$ which involve the adjoint, the point is again to use the Besov duality to control the remainders. Let $0 < \varepsilon < \gamma$, from the previous analysis we get that:

$$\begin{aligned}|R^2_{m,2:n}(T,u)| &:= |\sum_{i=2}^n \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} D_{\mathbf{x}_i}\big(\varphi(t,\mathbf{x})\mathbf{F}_i(t,\mathbf{x})\big)\big(u_m(t,\mathbf{x}) - u(t,\mathbf{x})\big)| \\ &\leq \sum_{i=2}^n \int_0^T dt \int_{\mathbb{R}^{(n-1)d}} \prod_{j\neq i} d\mathbf{x}_j \|D_{\mathbf{x}_i}\big(\varphi\mathbf{F}_i(t,\mathbf{x}_{\setminus i},\cdot)\big)\|_{B^{-\frac{2+\gamma-\varepsilon}{2i-1}}_{1,1}(\mathbb{R}^d,\mathbb{R})} \|(u_m - u)(t,\mathbf{x}_{\setminus i},\cdot)\|_{C_b^{\frac{2+\gamma-\varepsilon}{2i-1}}(\mathbb{R}^d,\mathbb{R})},\end{aligned}$$

denoting $\mathbf{x}_{\setminus i} = (\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n)$ and $\varphi \mathbf{F}_i(t, \mathbf{x}_{\setminus i}, \cdot) : \mathbf{x}_i \in \mathbb{R}^d \mapsto \varphi \mathbf{F}_i(t, \mathbf{x})$, $(u - u_m)(t, \mathbf{x}_{\setminus i}, \cdot) : \mathbf{x}_i \in \mathbb{R}^d \mapsto (u - u_m)(t, \mathbf{x})$. On the one hand we know that $\|u - u_m\|_{L^\infty(C^{2+\gamma-\varepsilon}_{b,\mathbf{d}})} \underset{m}{\longrightarrow} 0$. On the other hand, φ is a smooth test function and from **(S)** we have that $D_{\mathbf{x}_i} \mathbf{F}_i(t, \mathbf{x}_{\setminus i}, \cdot)$ belongs to $B^{\frac{-2+\gamma}{2i-1}}_{\infty,\infty}(\mathbb{R}^d, \mathbb{R})$. Since $-\frac{2-\gamma}{2i-1} \geq -\frac{2+\gamma-\varepsilon}{2i-1}$ for any $0 < \varepsilon < 2\gamma$ we deduce from the arguments of the proof of Lemma 11 that there exists C s.t. for each $i \in [\![2, n]\!]$, $\|D_{\mathbf{x}_i}\big(\varphi \mathbf{F}_i(t, \mathbf{x}_{\setminus i}, \cdot)\big)\|_{B_{1,1}^{-\frac{2+\gamma-\varepsilon}{2i-1}}(\mathbb{R}^d, \mathbb{R})} \leq C\psi_i(t, \mathbf{x}_{\setminus i})$ where ψ_i has compact support on $\mathbb{R}^{(n-1)d}$.

We thus readily derive $|R^2_{m,2:n}(T, u)| \underset{m}{\longrightarrow} 0$.

Eventually,

$$\begin{aligned} |R^2_{m,1}(T,u)| &:= \Big| \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \Big[D_{\mathbf{x}_1}\big(\varphi(t,\mathbf{x}) \mathbf{F}_1(t,\mathbf{x})\big) + D^2_{\mathbf{x}_1}\big(\varphi(t,\mathbf{x}) a(t,\mathbf{x})\big) \Big] \big(u_m(t,\mathbf{x}) - u(t,\mathbf{x})\big) \Big| \\ &= \Big| \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \Big[\big(\varphi(t,\mathbf{x}) \mathbf{F}_1(t,\mathbf{x})\big) D_{\mathbf{x}_1} + \big(\varphi(t,\mathbf{x}) a(t,\mathbf{x})\big) D^2_{\mathbf{x}_1} \Big] \big(u_m(t,\mathbf{x}) - u(t,\mathbf{x})\big) \Big|, \end{aligned}$$

which again tends to 0 with m since $\|u - u_m\|_{L^\infty(C^{2+\gamma-\varepsilon}_{b,\mathbf{d}})} \underset{m}{\longrightarrow} 0$ for any $0 < \varepsilon < \gamma$.

The contributions involving $\big((L^m_t)^* - L^*_t\big)\varphi$ in $R^1_m(T, u)$ defined in (7.32) can be handled as in the proof of Lemma 11 exploiting that $\|a - a_m\|_{C^\gamma_{b,\mathbf{d}}} + \|\mathbf{F}_1 - \mathbf{F}_{m,1}\|_{C^\gamma_{b,\mathbf{d}}} + \sum_{i=2}^n \|\mathbf{F}_i - \mathbf{F}_{m,i}\|_{C^{2i-3+\gamma}_{b,\mathbf{d}}} \to_m 0$. We now deduce from (7.31), (7.32) and the previous controls that $R_m(T, u) \to_m 0$. The same computations also give that the term $\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \Big(-\partial_t + (L_t)^* \Big)\varphi(t, \mathbf{x}) u(t, \mathbf{x})$ is well defined under **(A)**. From (7.30), (7.31), we thus finally derive:

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \varphi(t,\mathbf{x}) f(t,\mathbf{x}) = \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \big(-\partial_t + (L_t)^* \big)\varphi(t,\mathbf{x}) u(t,\mathbf{x}),$$

which gives the statement.

Appendix A. Proof of technical results

A.1. **Technical results associated with the flow.** We begin this paragraph stating and proving a key result for the sensitivity of Hölder flows, i.e. when the coefficients satisfy **(A)**. Those results are of course uniform w.r.t. a mollification procedure of the coefficients as the one previously considered from Section 3 to 5. Also, Lemma 7 is a direct consequence of Lemma 18 below and Young/convexity inequalities.

A.1.1. *A first sensitivity result for the flow.*

Lemma 18. *Under* **(A)**, *there exists* $C := C((\mathbf{A}), T)$ *s.t. for all* $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$, $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$, $0 \leq t < s \leq T$ *and* $i \in [\![1, n]\!]$*:*

$$|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \leq C\Big((s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x} - \mathbf{x}') \Big).$$

The previous bound can be interpreted as follows. We somehow have the expected bound involving the spatial points, up to an additional contribution in time, which is precisely due to the quasi-distance $\mathbf{d}$. Indeed, this type of result already appeared (for Lipschitz drifts) in Proposition 4.1 of [28]. Through an appropriate mollifying procedure, this result remains unchanged.

Proof. The main idea to prove this control relies on Grönwall's Lemma. However, under **(A)**, the function $\mathbf{F}$ is not Lipschitz (only Hölder continuous). We have then to mollify suitably $\mathbf{F}$. Let us denote by $\delta \in \mathbb{R}^n$, a vector with positive entries $\delta_i > 0$ for $i \in [\![2, n]\!]$. Define as well for all $v \in [0, T]$, $\mathbf{z} \in \mathbb{R}^{nd}$, $i \in [\![2, n]\!]$,

$$\mathbf{F}^\delta_i(v, \mathbf{z}^{i-1:n}) := \mathbf{F}_i(v, \cdot) \star \rho_{\delta_i}(\mathbf{z}) = \int_{\mathbb{R}^d} \mathbf{F}_i(v, \mathbf{z}_{i-1}, \mathbf{z}_i - w, \mathbf{z}_{i+1} \cdots, \mathbf{z}_n) \rho_{\delta_i}(w) dw, \tag{A.1}$$

with $\rho_{\delta_i}(w) := \frac{1}{\delta_i^d} \rho\Big(\frac{w}{\delta_i}\Big)$ where $\rho : \mathbb{R}^d \to \mathbb{R}_+$ is a usual mollifier, namely ρ has compact support and $\int_{\mathbb{R}^d} \rho(\mathbf{z}) d\mathbf{z} = 1$. Finally, we define $\mathbf{F}^\delta(v, \mathbf{z}) := (\mathbf{F}_1(v, \mathbf{z}), \mathbf{F}^\delta_2(v, \mathbf{z}), \cdots, \mathbf{F}^\delta_n(v, \mathbf{z}))$. In this definition, we make a slight abuse of notation since the first component $\mathbf{F}_1$ is not mollified. Due to the final control we want to prove and the intrinsic scale of the first component, the sublinearity of $\mathbf{F}_1$ (implied by its Hölder property) is enough and it is not needed to mollify this component.

To be at the *good* current time scale for the contributions associated with the mollification, we pick δ_i in order to have $C := C((\mathbf{A}), T) > 0$ s.t. for all $\mathbf{z} \in \mathbb{R}^{nd}$, $v \in [t, s]$:

$$\Big| (s-t)^{\frac{1}{2}} \mathbb{T}^{-1}_{s-t} \Big(\mathbf{F}(v, \mathbf{z}) - \mathbf{F}^\delta(v, \mathbf{z}) \Big) \Big| \leq C(s-t)^{-1}. \tag{A.2}$$

By the previous definition of $\mathbf{F}^\delta$ and assumptions **(A)**, identity (A.2) is equivalent to:

$$\sum_{i=2}^{n}(s-t)^{\frac{1}{2}-i}\delta_i^{\frac{2i-3+\gamma}{2i-1}} \leq C(s-t)^{-1}. \tag{A.3}$$

Hence, we choose from now on, for all $i \in [\![2,n]\!]$:

$$\delta_i = (s-t)^{(i-\frac{3}{2})\frac{2i-1}{2i-3+\gamma}}. \tag{A.4}$$

Next, let us control the last components of the flow. By the definition of $\boldsymbol{\theta}_{s,t}$ in (2.5), we get:

$$\begin{aligned}
&|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\
\leq\ & |(\mathbf{x}-\mathbf{x}')_n| + \int_t^s \Big(|\mathbf{F}_n^\delta(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n^\delta(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}'))| \\
&+|\mathbf{F}_n^\delta(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))| + |\mathbf{F}_n^\delta(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}')) - \mathbf{F}_n(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}'))|\Big)dv \\
\leq\ & |(\mathbf{x}-\mathbf{x}')_n| + C\int_t^s \Big(\big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_{n-1}\big| + \delta_n^{-1+\frac{2n-3+\gamma}{2n-1}}\big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_n\big|\Big)dv + (s-t)\delta_n^{\frac{2n-3+\gamma}{2n-1}}.
\end{aligned}$$

Hence by Grönwall' s Lemma, we get:

$$\begin{aligned}
&|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\
\leq\ & C\exp\Big(C(s-t)\delta_n^{-1+\frac{2n-3+\gamma}{2n-1}}\Big)\Big(|(\mathbf{x}-\mathbf{x}')_n| + (s-t)\delta_n^{\frac{2n-3+\gamma}{2n-1}} + \int_t^s \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_{n-1}\big|dv\Big) \\
\leq\ & C\exp\Big(C(s-t)^{\frac{\gamma}{2}}\Big)\Big(|(\mathbf{x}-\mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + \int_t^s \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_{n-1}\big|dv\Big),
\end{aligned} \tag{A.5}$$

using (A.3) for the last inequality. We proceed similarly for the $(n-1)^{\text{th}}$ component, but in this case we have to handle the non-Lipschitz continuity of $\mathbf{F}_{n-1}^\delta$ in its n^{th} variable.

For the rescaled flows see e.g. Lemma 2 in [4] this difficulty could also be circumvented through mollification, the situation is here slightly different and it seems that Young type controls are more appropriate. Write:

$$\begin{aligned}
&|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_{n-1}| \\
\leq\ & C\exp\Big(C(s-t)\delta_{n-1}^{-1+\frac{2(n-1)-3+\gamma}{2(n-1)-1}}\Big)\Big(|(\mathbf{x}-\mathbf{x}')_{n-1}| + (s-t)\delta_{n-1}^{\frac{2(n-1)-3+\gamma}{2(n-1)-1}} \\
&+ \int_t^s \Big\{\big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_{n-2}\big| + \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_n\big|^{\frac{2(n-1)-3+\gamma}{2n-1}}\Big\}dv\Big) \\
\leq\ & C\exp(C(s-t)^{\frac{\gamma}{2}})\Big(|(\mathbf{x}-\mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \Big\{\big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_{n-2}\big| \\
&+|(\mathbf{x}-\mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}} + (v-t)^{\frac{2(n-1)-3+\gamma}{2}} + \Big(\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)_{n-1}\big|dw\Big)^{\frac{2(n-1)-3+\gamma}{2n-1}}\Big\}dv\Big).
\end{aligned} \tag{A.6}$$

The last inequality is a consequence of our choice of δ_{n-1} in (A.4), identity (A.5) and convexity inequality.

We aim, now, to proceed with Grönwall's Lemma. To do so, we first of all need to use a Young inequality. Namely, we write for any $\tilde{\delta}_{n-1,n} > 0$ (where the two indexes in the subscript respectively denote the level of the chain, i.e. $n-1$, and the considered variable, i.e. n):

$$\begin{aligned}
&\Big(\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-1}\big|dw\Big)^{\frac{2(n-1)-3+\gamma}{2n-1}} \\
\leq\ & C\Big(\Big(\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-1}\big|dw\Big)\tilde{\delta}_{n-1,n}^{-\frac{2n-1}{2(n-1)-3+\gamma}} + \tilde{\delta}_{n-1,n}^{\frac{2n-1}{4-\gamma}}\Big).
\end{aligned}$$

In order to obtain the suitable time scale, we choose $\tilde{\delta}_n$ s.t.

$$\tilde{\delta}_{n-1,n}^{\frac{2n-1}{4-\gamma}} = (v-t)^{\frac{2(n-1)-3+\gamma}{2}} \iff \tilde{\delta}_{n-1,n} = (v-t)^{\frac{(2(n-1)-3+\gamma)(4-\gamma)}{2(2n-1)}},$$

which also yields that

$$\Big(\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-1}\big|dw\Big)\tilde{\delta}_{n-1,n}^{-\frac{2n-1}{2(n-1)-3+\gamma}} \leq \Big(\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-1}\big|dw\Big)(v-t)^{-\frac{4-\gamma}{2}}.$$

Hence we get from (A.6) and the previous controls that for any $\tilde{s} \in [t,s]$:

$$\begin{aligned}
&|(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_{n-1}| \\
\leq\ & C\exp(C(\tilde{s}-t)^{\frac{\gamma}{2}})\Big(|(\mathbf{x}-\mathbf{x}')_{n-1}| + (\tilde{s}-t)^{n-\frac{3}{2}} + \int_t^{\tilde{s}} \Big\{ \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)\big)_{n-2}\big| \\
&+|(\mathbf{x}-\mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}} + (v-t)^{\frac{2(n-1)-3+\gamma}{2}} + \Big(\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-1}\big| dw\Big)(v-t)^{-\frac{4-\gamma}{2}} \Big\} dv\Big)
\end{aligned}$$

The point is now to take the supremum in $\tilde{s} \in [s,t]$ in the above equation. This yields:

$$\begin{aligned}
&\sup_{\tilde{s}\in[t,s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_{n-1}| \\
\leq\ & C\exp(C(s-t)^{\frac{\gamma}{2}})\Big(|(\mathbf{x}-\mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)\big)_{n-2}\big| dv \\
&+|(\mathbf{x}-\mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}}(s-t) + (s-t)^{1+\frac{2(n-1)-3+\gamma}{2}} \\
&+\Big(\int_t^s \sup_{w\in[t,v]} \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-1}\big|(v-t)^{1-\frac{4-\gamma}{2}} dv\Big)\Big).
\end{aligned}$$

We get then by Grönwall's Lemma:

$$\begin{aligned}
&|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_{n-1}| \\
\leq\ & C\exp(C(s-t)^{\frac{\gamma}{2}}) \\
&\times\Big(|(\mathbf{x}-\mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)\big)_{n-2}\big| dv + |(\mathbf{x}-\mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}}(s-t)\Big) \\
\leq\ & C\exp(C(s-t)^{\frac{\gamma}{2}})\Big(|(\mathbf{x}-\mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)\big)_{n-2}\big| dv + |(\mathbf{x}-\mathbf{x}')_n|^{\frac{2n-3}{2n-1}}\Big),
\end{aligned} \tag{A.7}$$

using again the Young inequality $|(\mathbf{x}-\mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}}(s-t) \leq C((s-t)^{n-\frac{3}{2}} + |(\mathbf{x}-\mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}\frac{2n-3}{2n-5}})$ for the last identity, recalling as well that $\mathbf{d}(\mathbf{x},\mathbf{x}') \leq 1$, and therefore $|(\mathbf{x}-\mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-5}} \leq |(\mathbf{x}-\mathbf{x}')_n|$, for the last identity. The purpose of (A.7) is that each entry of the difference of the starting points appears at its intrinsic scale for the homogeneous distance $\mathbf{d}$.

Plugging the above inequality into (A.5) we derive:

$$\begin{aligned}
&|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\
\leq\ & C\exp\Big(C(s-t)^{\frac{\gamma}{2}}\Big)\Big(|(\mathbf{x}-\mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + |(\mathbf{x}-\mathbf{x}')_{n-1}|(s-t) \\
&+|(\mathbf{x}-\mathbf{x}')_n|^{\frac{2n-3}{2n-1}}(s-t) + \int_t^s\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-2}\big| dw dv\Big) \\
\leq\ & C\exp\Big(C(s-t)^{\frac{\gamma}{2}}\Big)\Big(|(\mathbf{x}-\mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + |(\mathbf{x}-\mathbf{x}')_{n-1}|^{\frac{2n-1}{2n-3}} + \int_t^s\int_t^v \big|\big(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}')\big)\big)_{n-2}\big| dw dv\Big),
\end{aligned}$$

using again the Young inequalities $|(\mathbf{x}-\mathbf{x}')_n|^{\frac{2n-3}{2n-1}}(s-t) \leq C(|(\mathbf{x}-\mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}})$ and $|(\mathbf{x}-\mathbf{x}')_{n-1}|(s-t) \leq C\big(|(\mathbf{x}-\mathbf{x}')_{n-1}|^{\frac{2n-1}{2n-3}} + (s-t)^{n-\frac{1}{2}}\big)$ for the last bound. Iterating these computations, we obtain:

$$|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \leq C\Big((s-t)^{n-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x}-\mathbf{x}')_j|^{\frac{2n-1}{2j-1}} + \int_t^{v_n=s} dv_{n-1}\ldots\int_t^{v_2} dv_1 \big|\big(\boldsymbol{\theta}_{v_1,t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1,t}(\mathbf{x})\big)_1\big|\Big). \tag{A.8}$$

Similarly, for $i \in [\![2,n]\!]$ we derive:

$$|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \leq C\Big((s-t)^{i-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x}-\mathbf{x}')_j|^{\frac{2i-1}{2j-1}} + \int_t^{v_i=s} dv_{i-1}\ldots\int_t^{v_2} dv_1 \big|\big(\boldsymbol{\theta}_{v_1,t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1,t}(\mathbf{x})\big)_1\big|\Big). \tag{A.9}$$

Remark 9. *We importantly point out that equations* (A.8) *and* (A.9) *actually hold not only for the fixed time s but also for any $v \in [t,T]$.*

The term for $i=1$ is treated slightly differently. Namely, for any $\tilde{s} \in [t,s]$, write:

$$|(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \leq |(\mathbf{x}-\mathbf{x}')_1| + C\sum_{j=1}^n \int_t^{\tilde{s}} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{\gamma}{2j-1}} dv,$$

which in turn implies, using (A.9) and Remark 9,

$$\begin{aligned}
&\sup_{\tilde{s}\in[t,s]}|(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x})-\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1|\\
\leq\;& |(\mathbf{x}-\mathbf{x}')_1|+C\Big((s-t)\big(\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|\big)^{\gamma}+\sum_{j=2}^{n}\int_t^s|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{\gamma}{2j-1}}dv\Big)\\
\leq\;& |(\mathbf{x}-\mathbf{x}')_1|+C\Bigg((s-t)\big(\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|\big)^{\gamma}\\
&+\sum_{j=2}^{n}C_j(s-t)\Big((s-t)^{j-\frac12}+\sum_{k=2}^{n}|(\mathbf{x}-\mathbf{x}')_k|^{\frac{2j-1}{2k-1}}+(s-t)^{j-1}\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|\Big)^{\frac{\gamma}{2j-1}}\Bigg)\\
\leq\;& |(\mathbf{x}-\mathbf{x}')_1|+C\Bigg((s-t)\big(\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|\big)^{\gamma}\\
&+\sum_{j=2}^{n}C_j(s-t)\Big((s-t)^{\frac{\gamma}{2}}+\sum_{k=2}^{n}|(\mathbf{x}-\mathbf{x}')_k|^{\frac{\gamma}{2k-1}}+(s-t)^{(j-1)\frac{\gamma}{2j-1}}\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|^{\frac{\gamma}{2j-1}}\Big)\Bigg),
\end{aligned}$$

using as well convexity inequalities for the last bound. We now write,

$$\begin{aligned}
&\sup_{\tilde{s}\in[t,s]}|(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x})-\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1|\\
\leq\;& C\Bigg(|(\mathbf{x}-\mathbf{x}')_1|+(s-t)^{1+\frac{\gamma}{2}}+(s-t)\sum_{k=2}^{n}|(\mathbf{x}-\mathbf{x}')_k|^{\frac{\gamma}{2k-1}}\\
&+\sum_{j=1}^{n}(s-t)^{1+(j-1)\frac{\gamma}{2j-1}}\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|^{\frac{\gamma}{2j-1}}\Bigg)\\
\leq\;& C(|(\mathbf{x}-\mathbf{x}')_1|+(s-t)+\sum_{k=2}^{n}|(\mathbf{x}-\mathbf{x}')_k|^{\frac{1}{2k-1}}) \qquad \text{(A.10)}
\end{aligned}$$

recalling that $(s-t)\leq T$ is small, and using again Young inequalities for the last bound. Namely,

$$(s-t)^{1+(j-1)\frac{\gamma}{2j-1}}\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|^{\frac{\gamma}{2j-1}}\leq C(s-t)\Big(1+\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|\Big),$$

and

$$(s-t)|(\mathbf{x}-\mathbf{x}')_k|^{\frac{\gamma}{2k-1}}\leq C((s-t)^{\frac{1}{1-\gamma}}+|(\mathbf{x}-\mathbf{x}')_k|^{\frac{1}{2k-1}}).$$

We eventually derive from (A.10) that:

$$\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|\leq C\big((s-t)^{\frac12}+\mathbf{d}(\mathbf{x},\mathbf{x}')\big),$$

which gives the stated bound for $i=1$. It now remains to plug this control into (A.9). We obtain for each $i\in[\![2,n]\!]$:

$$\begin{aligned}
|(\boldsymbol{\theta}_{s,t}(\mathbf{x})-\boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| &\leq C\big((s-t)^{i-\frac12}+\mathbf{d}^{2i-1}(\mathbf{x},\mathbf{x}')+(s-t)^{i-1}\sup_{v\in[t,s]}|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|\big)\\
&\leq C\Big((s-t)^{i-\frac12}+\mathbf{d}^{2i-1}(\mathbf{x},\mathbf{x}')+(s-t)^{i-1}\big((s-t)^{\frac12}+\mathbf{d}(\mathbf{x},\mathbf{x}')\big)\Big)\\
&\leq C\big((s-t)^{i-\frac12}+\mathbf{d}^{2i-1}(\mathbf{x},\mathbf{x}')\big),
\end{aligned}$$

using again the Young inequality to derive that $(s-t)^{i-1}\mathbf{d}(\mathbf{x},\mathbf{x}')\leq C\big((s-t)^{i-\frac12}+\mathbf{d}^{2i-1}(\mathbf{x},\mathbf{x}')\big)$. The proof is complete. □

Again, Lemma 7 is a direct consequence of the previous Lemma 18 and Young/convexity inequalities.

We are now in position to prove the sensitivity results for the linearized system w.r.t. the freezing parameter.

A.1.2. *Sensitivity results for the mean.*

Proof of the Technical Lemma 8. We assume w.l.o.g. that $\mathbf{d}(\mathbf{x},\mathbf{x}')\leq 1$. The idea of the proof is to separate the term to control into two contributions. Namely, we write:

$$\mathbf{m}^{\mathbf{x}}_{s,t}(\mathbf{x}')-\boldsymbol{\theta}_{s,t}(\mathbf{x}')=[\mathbf{m}^{\mathbf{x}}_{s,t}(\mathbf{x}')-\boldsymbol{\theta}_{s,t}(\mathbf{x})]+[\boldsymbol{\theta}_{s,t}(\mathbf{x})-\boldsymbol{\theta}_{s,t}(\mathbf{x}')]. \qquad \text{(A.11)}$$

The definition of the proxy (3.1) yields that the mean value of $\tilde{\mathbf{X}}_v^{m,\boldsymbol{\xi}}$, $\mathbf{m}_{v,t}^{\boldsymbol{\xi}}$ is s.t.

$$\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}) = \mathbf{x}' - \mathbf{x} + \int_t^s dv D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{v,t}(\mathbf{x})]. \tag{A.12}$$

The sub-triangular structure of $D\mathbf{F}$ yields that for each $i \in [\![2, n]\!]$:

$$\big(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x})\big)_i = \mathbf{x}'_i - \mathbf{x}_i + \int_t^s dv D_{i-1}\mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_{i-1} - \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1}].$$

Also, since $\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_1 = \mathbf{x}'_1 + \int_t^s \mathbf{F}_1(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))dv$, so that $[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_1 - \boldsymbol{\theta}_{v,t}(\mathbf{x})_1] = \mathbf{x}'_1 - \mathbf{x}_1$, we then obtain by iteration that:

$$\big(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x})\big)_i = \mathbf{x}'_i - \mathbf{x}_i + \sum_{k=2}^{i} \Big[\int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^{i} D_{j-1}\mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})) \Big][\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}],$$

with the convention that for $i = 1$, $\sum_{k=2}^{i} = 0$. From the above control, equation (A.11) and the dynamics of the flow, recalling that the starting points are the same, so that the contributions involving differences of the spatial points or flows only appear in iterated time integrals, we derive:

$$\begin{aligned}
&|\big(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}')\big)_i| \\
\leq\ & \Big| \sum_{k=2}^{i} \Big[\int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^{i} D_{j-1}\mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})) \Big][\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}] \Big| \\
&+ \int_t^s |\mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| dv \\
\leq\ & C\Big(\sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| + \int_t^s \Big(\sum_{j=i}^{n} \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_j\big|^{\frac{2i-3+\gamma}{2j-1}} + \big|\big(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}')\big)_{i-1}\big| \Big) dv \Big).
\end{aligned}$$

From the previous Lemma 18, we thus obtain:

$$\begin{aligned}
|\big(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}')\big)_i| \leq\ & C\Big(\sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| + (s-t)^{\frac{2i-3+\gamma}{2}+1} \\
&+ \mathbf{d}^{2i-3+\gamma}(\mathbf{x}, \mathbf{x}')(s-t) + \big((s-t)^{(i-1)-\frac{1}{2}} + \mathbf{d}^{2(i-1)-1}(\mathbf{x}, \mathbf{x}')\big)(s-t) \Big).
\end{aligned} \tag{A.13}$$

In particular, for $s = t_0 = t + c_0\mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ with $c_0 < 1$, the previous equation yields:

$$|\big(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')\big)_i| \leq C\Big(c_0\mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (c_0^{i-\frac{1}{2}+\frac{\gamma}{2}} + c_0)\mathbf{d}^{2i-1+\gamma}(\mathbf{x}, \mathbf{x}') + (c_0^{i-\frac{1}{2}} + c_0)\mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \Big).$$

So, after summing and by convexity inequalities, for $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$:

$$\mathbf{d}\big(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}'), \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')\big) \leq C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}').$$

□

A.2. **Sensitivities for the scaled flows.** For the scaling analysis of Section 6 we also need the scaled versions of the previous lemmas. Recalling the notations introduced therein, i.e. for $\lambda > 0$, $0 \leq t \leq v \leq T$, $\boldsymbol{\theta}_{v,t}^{\lambda}(\mathbf{x}^{\lambda}) = \lambda^{\frac{1}{2}}\mathbb{T}_{\lambda}^{-1}\boldsymbol{\theta}_{v,t}(\mathbf{x}^{\lambda})$, $\mathbf{x}^{\lambda} := \lambda^{-\frac{1}{2}}\mathbb{T}_{\lambda}\mathbf{x}$, we readily get:

$$\mathbf{d}\big(\lambda^{\frac{1}{2}}\mathbb{T}_{\lambda}^{-1}\mathbf{m}_{v,t}^{\mathbf{x}^{\lambda}}(\mathbf{x}'^{\lambda}), \boldsymbol{\theta}_{v,t}^{\lambda}(\mathbf{x}'^{\lambda})\big) = \lambda^{-\frac{1}{2}}\mathbf{d}\big(\mathbf{m}_{v,t}^{\mathbf{x}^{\lambda}}(\mathbf{x}'^{\lambda}), \boldsymbol{\theta}_{v,t}(\mathbf{x}'^{\lambda})\big). \tag{A.14}$$

Now, the discontinuity term leads to consider $v = t + c_0\lambda\mathbf{d}^2(\mathbf{x}, \mathbf{x}') = t + c_0\mathbf{d}^2(\mathbf{x}^{\lambda}, \mathbf{x}'^{\lambda})$. So, from (A.14), we can readily apply Lemma 8 to the quantity $\mathbf{d}\big(\mathbf{m}_{v,t}^{\mathbf{x}^{\lambda}}(\mathbf{x}'^{\lambda}), \boldsymbol{\theta}_{v,t}(\mathbf{x}'^{\lambda})\big)$ for the spatial points $\mathbf{x}^{\lambda}, \mathbf{x}'^{\lambda}$ and the corresponding critical time. This precisely yields $\mathbf{d}\big(\mathbf{m}_{v,t}^{\mathbf{x}^{\lambda}}(\mathbf{x}'^{\lambda}), \boldsymbol{\theta}_{v,t}(\mathbf{x}'^{\lambda})\big) \leq C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}^{\lambda}, \mathbf{x}'^{\lambda}) = C c_0^{\frac{1}{2n-1}} \lambda^{\frac{1}{2}} \mathbf{d}(\mathbf{x}, \mathbf{x}')$ which plugged into (A.14) finally leads to:

$$\mathbf{d}\big(\lambda^{\frac{1}{2}}\mathbb{T}_{\lambda}^{-1}\mathbf{m}_{v,t}^{\mathbf{x}^{\lambda}}(\mathbf{x}'^{\lambda}), \boldsymbol{\theta}_{v,t}^{\lambda}(\mathbf{x}'^{\lambda})\big) \leq C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}'). \tag{A.15}$$

In other words, Lemma 8 is invariant for the scaled flows.

Appendix B. Sensitivity results for the resolvent and covariance

B.1. Sensitivity Lemma for the Resolvent.

Lemma 19 (Controls of the Sensitivities for the Resolvent)**.** *There exists $\tilde{C}$ s.t. for all $0 \le t \le s \le T$, $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$, the following control holds. For all $1 \le j < i \le n$, with the notation of* (3.21)*:*

$$\begin{aligned}\big|\big[\tilde{\mathbf{R}}^{(\tau,\mathbf{x})}(t,s) - \tilde{\mathbf{R}}^{(\tau,\mathbf{x}')}(t,s)\big]_{i,j}\big| &\le \tilde{C}(s-t)^{i-j}\Big(\sum_{k=2}^{n}\|\mathbf{F}_k\|_{L^\infty(C^{2k-3+\gamma}_{\mathbf{d},\mathbf{H}})}\Big)\big((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\big)\\ &\le \Lambda(s-t)^{i-j}\big((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\big).\end{aligned}$$

Proof of Lemma 19. From the scaling properties of the resolvent, see e.g. the proof of Proposition 3 or Lemma 6.2 in [28], we have that:

$$\tilde{\mathbf{R}}^{\mathbf{x}}(s,t) = \mathbb{T}_{s-t}\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}}\mathbb{T}^{-1}_{s-t},\ \tilde{\mathbf{R}}^{\mathbf{x}'}(s,t) = \mathbb{T}_{s-t}\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}'}\mathbb{T}^{-1}_{s-t}, \tag{B.1}$$

where $\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}}$, $\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}'}$ are non-degenerate bounded matrices. We define then:

$$\Delta\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x},\mathbf{x}'} := \widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}} - \widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}'}. \tag{B.2}$$

Hence, from (B.1) and the definitions in (B.2):

$$\begin{aligned}|\Delta\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x},\mathbf{x}'}| &= |\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}} - \widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x}'}| = |\mathbb{T}^{-1}_{s-t}(\tilde{\mathbf{R}}^{\mathbf{x}} - \tilde{\mathbf{R}}^{\mathbf{x}'})(s,t)\mathbb{T}_{s-t}|\\ &= \Big|\mathbb{T}^{-1}_{s-t}\int_t^s\Big(D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))\tilde{\mathbf{R}}^{\mathbf{x}}(v,t) - D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}'))\tilde{\mathbf{R}}^{\mathbf{x}'}(v,t)\Big)dv\mathbb{T}_{s-t}\Big|\\ &\le \int_t^s|\mathbb{T}^{-1}_{s-t}D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))\mathbb{T}_{s-t}||\mathbb{T}^{-1}_{s-t}(\tilde{\mathbf{R}}^{\mathbf{x}} - \tilde{\mathbf{R}}^{\mathbf{x}'})(v,t)\mathbb{T}_{s-t}|dv\\ &\quad + \int_t^s\Big|\mathbb{T}^{-1}_{s-t}\Big(D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})) - D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}'))\Big)\mathbb{T}_{s-t}\Big||\mathbb{T}^{-1}_{s-t}\tilde{\mathbf{R}}^{\mathbf{x}'}(v,t)\mathbb{T}_{s-t}|dv\\ &= \int_t^s|\mathbb{T}^{-1}_{s-t}D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))\mathbb{T}_{s-t}||\mathbb{T}^{-1}_{s-t}\mathbb{T}_{v-t}\Delta\widehat{\tilde{\mathbf{R}}}_1^{v,t,\mathbf{x},\mathbf{x}'}\mathbb{T}^{-1}_{v-t}\mathbb{T}_{s-t}|dv\\ &\quad + \int_t^s\Big|\mathbb{T}^{-1}_{s-t}\Big(D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})) - D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}'))\Big)\mathbb{T}_{s-t}\Big||\mathbb{T}^{-1}_{s-t}\tilde{\mathbf{R}}^{\mathbf{x}'}(v,t)\mathbb{T}_{s-t}|dv.\end{aligned}$$

Using the Grönwall's Lemma and the structure of the resolvent, we get:

$$|\Delta\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x},\mathbf{x}'}| \le C\int_t^s(s-t)^{-1}|D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})) - D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}')|dv.$$

Pay attention that we only know from our smoothness assumption **(S)** that for all $i \in [\![2,n]\!]$, $\mathbf{z}^{i:n} = (\mathbf{z}_i,\cdots,\mathbf{z}_n) \in \mathbb{R}^{(n-i+1)d}$, $\mathbf{z}_{i-1} \mapsto D_{\mathbf{x}_{i-1}}\mathbf{F}_i(\mathbf{z}_{i-1},\mathbf{z}^{i:n})$ is $C^{\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d,\mathbb{R}^d\otimes\mathbb{R}^d)$-Hölder continuous for $\eta > 0$. Hence, we proceed carefully like in [4] and we obtain, from the above bound, that

$$\begin{aligned}|\Delta\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x},\mathbf{x}'}| &\le C\int_t^s(s-t)^{-1}\sum_{i=2}^n\Big(|D_{i-1}\mathbf{F}_i(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})) - D_{i-1}\mathbf{F}_i(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1},(\boldsymbol{\theta}_{v,t}(\mathbf{x}'))^{i:n})|\\ &\quad + \|(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i-1}\|_{L^\infty(C^\gamma_{\mathbf{d}})}|(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{i-1}|^{\eta_i}\Big)dv\\ &=: (R_1+R_2)(s,t,\mathbf{x},\mathbf{x}'),\end{aligned} \tag{B.3}$$

where $\eta_i := \frac{\gamma}{2(i-1)-1}$ and the notation $(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i-1}$ indicates that $D_{\mathbf{x}_{i-1}}\mathbf{F}_i$ is viewed as a function of its variable $(i-1)$ and the supremum is taken over the other ones. From Lemma 7 and the definition of $\mathbf{d}$ (see also Lemma 18), we readily get

$$\begin{aligned}|R_2(s,t,\mathbf{x},\mathbf{y})| &\le C\|D\mathbf{F}\|_{L^\infty(C^\gamma_{\mathbf{d}})}\sum_{i=2}^n\int_t^s(s-t)^{-1}\Big((v-t)^{(i-1-\frac{1}{2})} + \mathbf{d}^{2(i-1)-1}(\mathbf{x},\mathbf{x}')\Big)^{\eta_i}dv\\ &\le C\|D\mathbf{F}\|_{L^\infty(C^\gamma_{\mathbf{d}})}\big((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\big),\end{aligned} \tag{B.4}$$

denoting with a slight abuse of notation $\|D\mathbf{F}\|_{L^\infty(C^\gamma_{\mathbf{d}})} := \sum_{i=2}^n\|(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i-1}\|_{L^\infty(C^\gamma_{\mathbf{d}})}$. To control the difference of the gradients terms in $R_1(s,t,\mathbf{x},\mathbf{x}')$ in (B.3), we need the following result whose proof is postponed to Appendix B.3.

Lemma 20 (Reverse Taylor expansion)**.** *There is a constant $C>0$ s.t, for all $(\mathbf{z},\mathbf{z}')\in\mathbb{R}^{nd}\times\mathbb{R}^{nd}$, $v\in[0,T]$:*

$$\big|D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z})-D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}')\big|\le C\|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})}\mathbf{d}(\mathbf{z},\mathbf{z}')^\gamma,$$

with the notations of equation (3.21)*.*

From the reverse Taylor expansion of Lemma 20 and the definition in (B.3), we obtain:

$$\begin{aligned}|R_1(s,t,\mathbf{x},\mathbf{y})| &\le C(s-t)^{-1}\int_t^s dv\sum_{i=2}^n\|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})}\mathbf{d}^\gamma(\boldsymbol{\theta}_{v,t}(\mathbf{x}),\boldsymbol{\theta}_{v,t}(\mathbf{x}'))\\ &\le C\sum_{i=2}^n\|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})}\big((s-t)^{\frac{\gamma}{2}}+\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\big),\end{aligned}\tag{B.5}$$

using again Lemma 7 for the last inequality. Gathering (B.4), (B.5) into (B.3) and recalling the definition in (3.21), we obtain:

$$|\Delta\widehat{\tilde{\mathbf{R}}}_1^{s,t,\mathbf{x},\mathbf{x}'}|\le C\big(\sum_{i=2}^n\|\mathbf{F}_i\|_{L^\infty(C^{2i-3+\gamma}_{\mathbf{d},\mathbf{H}})}\big)\big((s-t)^{\frac{\gamma}{2}}+\mathbf{d}^\gamma(\mathbf{x},\mathbf{x}')\big).\tag{B.6}$$

The result follows from the previous bound, the definition in (B.2) and the scalings of equation (B.1). □

B.2. **Sensitivity Lemma for the covariances.**

B.2.1. *Proof of Lemma 6.* Let us first explicitly write the covariance matrices

$$\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{s,t}:=\int_t^s\tilde{\mathbf{R}}^{\boldsymbol{\xi}}(s,v)Ba(v,\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}))B^*\tilde{\mathbf{R}}^{\boldsymbol{\xi}}(s,v)^*dv.\tag{B.7}$$

So, we have to control the term

$$\begin{aligned}\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{s,t}-\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{s,t} &=: \Delta_1^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t)+\Delta_2^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t),\\ \Delta_1^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t) &:= \int_t^s dv\tilde{\mathbf{R}}^{\boldsymbol{\xi}}(s,v)B\Delta_v a(\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}),\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}'))B^*\tilde{\mathbf{R}}^{\boldsymbol{\xi}}(s,v)^*,\\ \Delta_v a(\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}),\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}')) &:= a(v,\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}))-a(v,\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}')),\\ \Delta_2^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t) &:= \int_t^s dv\Delta\tilde{\mathbf{R}}^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,v)Ba(v,\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}'))B^*\tilde{\mathbf{R}}^{\boldsymbol{\xi}}(s,v)^*\\ &\quad+\int_t^s dv\tilde{\mathbf{R}}^{\boldsymbol{\xi}'}(s,v)Ba(v,\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}'))B^*\Delta\tilde{\mathbf{R}}^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,v)^*,\\ \Delta\tilde{\mathbf{R}}^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,v) &= \tilde{\mathbf{R}}^{\boldsymbol{\xi}}(s,v)-\tilde{\mathbf{R}}^{\boldsymbol{\xi}'}(s,v).\end{aligned}\tag{B.8}$$

Hence, from the scalings of (B.1) and the definitions in (B.8), for all $1\le j\le i\le n$:

$$\begin{aligned}|[\Delta_1^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t)]_{i,j}| &\le C(s-t)^{-2}\int_t^s|[\mathbb{T}_{s-t}B\Delta_v a(\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}),\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}'))B^*\mathbb{T}_{s-t}]_{i,j}|dv\\ &\le \Lambda(s-t)^{i+j-2}\int_t^s\mathbf{d}(\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}),\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}'))^\gamma dv.\end{aligned}$$

We deduce by Lemma 7, that

$$|[\Delta_1^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t)]_{i,j}|\le\Lambda(s-t)^{i+j-1}\big((s-t)^{\frac{\gamma}{2}}+\mathbf{d}^\gamma(\boldsymbol{\xi},\boldsymbol{\xi}')\big).\tag{B.9}$$

Still from (B.1) and the definitions in (B.8), write now that:

$$\begin{aligned}|[\Delta_2^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t)]_{i,j}| &\le C(s-t)^{-2}\bigg(\int_t^s\Big(|[\mathbb{T}_{s-t}(\widehat{\tilde{\mathbf{R}}}_1^{s,v,\boldsymbol{\xi}}-\widehat{\tilde{\mathbf{R}}}_1^{s,v,\boldsymbol{\xi}'})Ba(v,\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}'))B^*(\widehat{\tilde{\mathbf{R}}}_1^{s,v,\boldsymbol{\xi}'})^*\mathbb{T}_{s-t}]_{i,j}|\\ &\quad+|[\mathbb{T}_{s-t}\widehat{\tilde{\mathbf{R}}}_1^{s,v,\boldsymbol{\xi}'}Ba(v,\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}'))B^*(\widehat{\tilde{\mathbf{R}}}_1^{s,v,\boldsymbol{\xi}'}-\widehat{\tilde{\mathbf{R}}}_1^{s,v,\boldsymbol{\xi}})^*\mathbb{T}_{s-t}]_{i,j}|\Big)dv\bigg).\end{aligned}\tag{B.10}$$

Thanks to equation (B.6) in the proof of Lemma 19, we thus obtain:

$$|[\Delta_2^{\boldsymbol{\xi},\boldsymbol{\xi}'}(s,t)]_{i,j}|\le\Lambda(s-t)^{i+j-1}\big((s-t)^{\frac{\gamma}{2}}+\mathbf{d}^\gamma(\boldsymbol{\xi},\boldsymbol{\xi}')\big).\tag{B.11}$$

Gathering (B.9) and (B.11) in (B.8) yields:

$$|[\tilde{\mathbf{K}}^{\boldsymbol{\xi}}_{s,t}]_{i,j}-[\tilde{\mathbf{K}}^{\boldsymbol{\xi}'}_{s,t}]_{i,j}|\le\Lambda(s-t)^{i+j-1}\big((s-t)^{\frac{\gamma}{2}}+\mathbf{d}^\gamma(\boldsymbol{\xi},\boldsymbol{\xi}')\big),\tag{B.12}$$

which precisely gives (3.19) for $i=j=1$ and then concludes the proof of Lemma 6. □

B.2.2. *Sensitivities for the scaled covariance.* In connection with Section 6, we recall the identity in law (6.6), i.e. $\tilde{\mathbf{X}}_v^{\boldsymbol{\xi},\lambda} := \lambda^{1/2}\mathbb{T}_\lambda^{-1}\tilde{\mathbf{X}}_v^{\boldsymbol{\xi}^\lambda}$, $v \in [t,T]$, which readily gives:

$$\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi},\lambda} := \mathrm{Cov}(\tilde{\mathbf{X}}_v^{\boldsymbol{\xi},\lambda}) = \lambda\mathbb{T}_\lambda^{-1}\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi}^\lambda}\mathbb{T}_\lambda^{-1}.$$

In particular, we thus derive from the analysis of the previous paragraph:

$$\begin{aligned}
[\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi},\lambda}]_{1,1} &\le C\lambda^{-1}(v-t),\\
\big|[\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi},\lambda}]_{1,1} - [\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi}',\lambda}]_{1,1}\big| &\le C\lambda^{-1}(v-t)\Big(\mathbf{d}^\gamma(\boldsymbol{\xi}^\lambda,\boldsymbol{\xi}'^\lambda) + (v-t)^{\frac{\gamma}{2}}\Big) \le C\lambda^{-1}(v-t)\Big(\lambda^{\frac{\gamma}{2}}\mathbf{d}^\gamma(\boldsymbol{\xi},\boldsymbol{\xi}') + (v-t)^{\frac{\gamma}{2}}\Big),
\end{aligned}$$

recalling that $\boldsymbol{\xi}^\lambda = \lambda^{-\frac{1}{2}}\mathbb{T}_\lambda\boldsymbol{\xi}$ and using the homogeneity properties of $\mathbf{d}$ for the last inequality. Also, for $v = t + c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')$ and taking $\boldsymbol{\xi} = \mathbf{x}$, $\boldsymbol{\xi}' = \mathbf{x}'$, the above controls rewrite:

$$[\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi},\lambda}]_{1,1} \le Cc_0\mathbf{d}^2(\mathbf{x},\mathbf{x}'),\ \big|[\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi},\lambda}]_{1,1} - [\tilde{\mathbf{K}}_{v,t}^{\boldsymbol{\xi}',\lambda}]_{1,1}\big| \le Cc_0\lambda^{\frac{\gamma}{2}}\mathbf{d}^{2+\gamma}(\mathbf{x},\mathbf{x}'). \tag{B.13}$$

Note that the sensitivity of the scaled covariance yields a contribution of the scaling coefficient in $\lambda^{\frac{\gamma}{2}}$. Unlike for the control of the scaled mean in (A.15), where, as previously noticed, we could not exploit the full regularity of $\mathbf{F}_i$ w.r.t. the $(i-1)^{\text{th}}$ variable, we can here precisely take advantage of such a regularity. Indeed, this follows from the expression of the covariance (B.7) which only involves $D_{i-1}\mathbf{F}_i$ so that one can exploit the associated $\frac{\gamma}{2i-3}$-Hölder regulariry w.r.t. $\mathbf{x}_{i-1}$.

B.3. **Reverse Taylor formula.**

Proof of Lemma 20. We assume here, for the sake of simplicity and without loss of generality, that $d = 1$ (scalar case). When $d > 1$, the proof below can be reproduced componentwise. Let us decompose the expression around the variables which do/do not transmit the noise. Namely, we write for any $\delta_i > 0$:

$$\begin{aligned}
&D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}) - D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}')\\
&= \int_0^1 d\mu\{D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}) - D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}_{i-1} + \mu\mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i},\mathbf{z}^{i:n})\}\\
&\quad +\{D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}_{i-1} + \mu\mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i},(\mathbf{z}')^{i:n}) - D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}')\}\\
&\quad +\{D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}_{i-1} + \mu\mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i},\mathbf{z}^{i:n}) - D_{\mathbf{x}_{i-1}}\mathbf{F}_i(v,\mathbf{z}_{i-1} + \mu\mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i},(\mathbf{z}')^{i:n})\}\\
&=: \sum_{\ell=1}^3 \Delta\mathbf{F}_i^\ell(v,\mathbf{z},\mathbf{z}').
\end{aligned} \tag{B.14}$$

The first two terms can be dealt directly. From **(A)** we get:

$$|\Delta\mathbf{F}_i^1(v,\mathbf{z},\mathbf{z}')| \le \|(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i-1}\|_{L^\infty(C_{\mathbf{d}}^\gamma)}\mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i\frac{\gamma}{2(i-1)-1}}. \tag{B.15}$$

Similarly,

$$\begin{aligned}
|\Delta\mathbf{F}_i^2(v,\mathbf{z},\mathbf{z}')| &\le \|(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i-1}\|_{L^\infty(C_{\mathbf{d}}^\gamma)}\big(\mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i\frac{\gamma}{2(i-1)-1}} + |(\mathbf{z}-\mathbf{z}')_{i-1}|^{\frac{\gamma}{2(i-1)-1}}\big)\\
&\le C\|(D_{\mathbf{x}_{i-1}}\mathbf{F}_i)_{i-1}\|_{L^\infty(C_{\mathbf{d}}^\gamma)}\mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i\frac{\gamma}{2(i-1)-1}}.
\end{aligned} \tag{B.16}$$

For $\Delta\mathbf{F}_i^3(t,\mathbf{z},\mathbf{z}')$, we use an explicit reverse Taylor expansion which yields together with the smoothness assumption of $\mathbf{F}_i$ in **(A)**:

$$\begin{aligned}
|\Delta\mathbf{F}_i^3(t,\mathbf{z},\mathbf{z}')| &= \mathbf{d}(\mathbf{z},\mathbf{z}')^{-\delta_i}\Big|\Big[\mathbf{F}_i(t,\mathbf{z}_{i-1} + \mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i},\mathbf{z}^{i:n}) - \mathbf{F}_i(t,\mathbf{z}_{i-1} + \mathbf{d}(\mathbf{z},\mathbf{z}')^{\delta_i},(\mathbf{z}')^{i:n})\\
&\quad +\mathbf{F}_i(t,\mathbf{z}_{i-1},\mathbf{z}^{i:n}) - \mathbf{F}_i(t,\mathbf{z}_{i-1},(\mathbf{z}')^{i:n})\Big]\Big|\\
&\le 2\|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d},\mathbf{H}}^{2i-3+\gamma})}\mathbf{d}(\mathbf{z},\mathbf{z}')^{2i-3+\gamma-\delta_i}.
\end{aligned} \tag{B.17}$$

Taking δ_i s.t. $\delta_i\frac{\gamma}{2(i-1)-1} = 2i-3+\gamma-\delta_i$, which implies that $\delta_i = 2i-3$, gives in (B.15), (B.16) and (B.17) a global bound of order $C\|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d},\mathbf{H}}^{2i-3+\gamma})}\mathbf{d}^\gamma(\mathbf{z},\mathbf{z}')$. The result then follows from (B.14).

□

Appendix C. Scaling Control of the degenerate part of the perturbative term

This section is dedicated to the proof of the scaled version of the key Besov control of Lemma 11. We recall that, with the definitions of Section 6, for all multi-index $\vartheta = (\vartheta_1, \dots, \vartheta_n) \in \mathbb{R}^{nd}$, $i \in [\![2, n]\!]$, we aim to control the terms

$$\int_t^T ds \int_{\mathbb{R}^{nd}} D^\vartheta D^2_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}_\lambda(t, s, \mathbf{x}, \mathbf{y}) \big\langle \Delta^\lambda_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \big\rangle d\mathbf{y},$$

with

$$\begin{aligned} &\sum_{i=2}^n \big\langle \mathbf{F}_{\lambda,i}(s, \mathbf{y}) - \mathbf{F}_{\lambda,i}(s, \boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_{\lambda,i}(s, \boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}))_{i-1}, D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \big\rangle \\ =: &\sum_{i=2}^n \big\langle \Delta^\lambda_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \big\rangle, \end{aligned}$$

which appear in equation (2.18) of the detailed guide for the scaled system. We precisely want to specify how the scaling procedure impacts the constants in equation (2.36).

This is exactly what equations (6.18) and (6.19) reflect. Those controls actually follow from the more general following result, which will again be useful for the Hölder norm in Section 5.

Lemma 21 (Scaled Besov Control Lemma)**.** *There exists $\Lambda := \Lambda((\mathbf{A}), T)$ as in Remark 3 s.t. for each multi-index $\vartheta = (\vartheta_1, \dots, \vartheta_n) \in \mathbb{N}^n$:*

$$\begin{aligned} &\sum_{i=2}^n \Big| \int_{\mathbb{R}^{nd}} D^\vartheta \tilde{p}^{\boldsymbol{\xi}}_\lambda(t, s, \mathbf{x}, \mathbf{y}) \big\langle \Delta^\lambda_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \big\rangle d\mathbf{y} \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\ \leq\ & \Lambda \lambda^{-1+\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})} \|u^\lambda\|_{L^\infty(C^{2+\gamma}_{b,\mathbf{d}})} (s-t)^{-\sum_{j=1}^n \vartheta_j (j-\frac{1}{2})+\frac{\gamma}{2}}. \end{aligned} \tag{C.1}$$

With Lemma 21 at hand, we indeed readily derive (6.18) and (6.19) taking $\vartheta = (2, 0, \cdots, 0) + e_k$ (where e_k stands for the k^{th} vector of the orthonormal basis) for each $k \in [\![2, n]\!]$ and $\vartheta = (2, 0, \cdots, 0)$ respectively. Let us now turn to the proof of the Lemma 21.

Proof of Lemma 21. The analysis of singularities is identical to the ones in the proof of Lemma 11. However, here, we have to track the scalling coefficient λ through the identites. Note carefully, we write the upper-script/sub-script λ to mean that we manage the scaled variables. In particular, we write:

$$\vartheta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s, \mathbf{y}) := D_{\mathbf{y}_i} \cdot \Big(D^\vartheta \tilde{p}^{\boldsymbol{\xi}}_\lambda(t, s, \mathbf{x}, \mathbf{y}) \otimes \Delta^\lambda_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}) \Big) =: D_{\mathbf{y}_i} \cdot \Theta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s, \mathbf{y}), \tag{C.2}$$

and

$$\Psi^{\vartheta,\lambda}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})} : \mathbf{y}_i \mapsto D_{\mathbf{y}_i} \cdot \big(\Theta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s, \mathbf{y}) \big). \tag{C.3}$$

With these notations, we have:

$$\sum_{i=2}^n \Big| \int_{\mathbb{R}^{nd}} D^\vartheta \tilde{p}^{(t,\boldsymbol{\xi})}_\lambda(t, s, \mathbf{x}, \mathbf{y}) \big\langle \Delta^\lambda_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \big\rangle d\mathbf{y} \Big|_{\boldsymbol{\xi}=\mathbf{x}} = \sum_{i=2}^n \Big| \int_{\mathbb{R}^{nd}} D_{\mathbf{y}_i} \cdot \big(\Theta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s, \mathbf{y}) \big) u^\lambda(s, \mathbf{y}) d\mathbf{y} \Big|. \tag{C.4}$$

The point is here again to control, for each $i \in [\![2, n]\!]$, the quantity $\| D_{\mathbf{y}_i} \cdot \big(\Theta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n}) \big) \|_{B^{\tilde{\alpha}_i}_{1,1}}$, $\tilde{\alpha}_i = \frac{2+\gamma}{2i-1}$ with the indicated bounds in the scaling parameter λ. Accordingly with what can be seen e.g. in (6.13), the previous analysis of the proof of the (non-scaled) Lemma 11 can be adapted replacing $(s-t)$ therein by $(s-t)/\lambda$ in the computations involving the thermic characterization of Besov spaces.

We also point out that, w.l.o.g., we assume that $T/\lambda \leq 1$ so that in particular for $0 \leq t < s \leq T$, $\lambda^{-1}(s-t) \leq 1$. Indeed, the parameter λ is meant to be *small* (at least $\lambda \leq 1$) but *macro* as well. From the previous analysis and the statement of Lemma 21 it can be seen that the *optimal* λ, i.e. the largest one, actually depends on the Hölder moduli of the coefficients. Hence, the condition $T/\lambda \leq 1$ is, up to a possible modification of T, not restrictive.

Let us first introduce some notation:

$$\hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y}) := \prod_{j=1}^n \mathcal{N}_{c\lambda^{\frac{2j-1}{2}}(s-t)^{2j-1}} \big((\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j \big) = \bar{p}_{c^{-1}}(t, s, \mathbf{x}, \mathbf{y}),$$

where for $\varsigma > 0,\ z \in \mathbb{R}^d$, like before $\mathcal{N}_\varsigma(z) = \frac{1}{(2\pi\varsigma)^{\frac{d}{2}}} \exp\big(-\frac{|z|^2}{2\varsigma}\big)$ is the standard Gaussian density of $\mathbb{R}^d$ with covariance matrix ςI_d, and:

$$\hat{q}_{c\backslash i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) = \prod_{j\in[\![1,n]\!],j\neq i} \mathcal{N}_{c\lambda^{\frac{2j-1}{2}}(s-t)^{2j-1}}\big((\boldsymbol{\theta}_{s,t}(\mathbf{x})-\mathbf{y})_j\big). \tag{C.5}$$

We recall from (4.16), that the parameter $\beta_i = \frac{(2i-3)(2i-1)}{2i-3-\gamma}$. The first contribution of the scaled Besov control is:

$$\begin{aligned}
&\int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \Psi^{\vartheta,\lambda}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
\leq\ & \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \Big| \int_{\mathbb{R}^d} D^\vartheta \tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x},\mathbf{y}) \big\langle \Delta^\lambda_{i,\mathbf{F}}(t,s,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}),\mathbf{y}), D_z h_v(z-\mathbf{y}_i)\big\rangle d\mathbf{y}_i \Big|\Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
\leq\ & \Lambda \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\lambda^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})} \hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})}} \\
&\times \lambda^{-i+\frac{1}{2}} \mathbf{d}^{2i-3+\gamma}(\lambda^{-1/2}\mathbb{T}_\lambda \boldsymbol{\theta}^\lambda_{s,t}(\mathbf{x}), \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{y}) \\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})-(i-\frac{1}{2})} \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})}} (s-t)^{\frac{2i-3+\gamma}{2}},
\end{aligned}$$

exploiting (6.13) for the last inequality. Then

$$\begin{aligned}
&\int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \Psi^{\vartheta,\lambda}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})-(i-\frac{1}{2})} \hat{q}_{c\backslash i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 dv v^{-\frac{3}{2}+\frac{\tilde{\alpha}_i}{2}} (s-t)^{-\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-3+\gamma}{2}} \\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})-(i-\frac{1}{2})} \hat{q}_{c\backslash i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \lambda^{[\frac{1}{2}-\frac{\tilde{\alpha}_i}{2}]\beta_i} (s-t)^{[-\frac{1}{2}+\frac{\tilde{\alpha}_i}{2}]\beta_i - \sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{2i-3+\gamma}{2}} \\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})-1} \hat{q}_{c\backslash i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))(s-t)^{-\sum_{j=1}^n \vartheta_j(j-\frac{1}{2})+\frac{\gamma}{2}},
\end{aligned} \tag{C.6}$$

the third inequality is a consequence of Proposition 3, and the last identity comes from the pick of β_i which in particular gives $-(i-\frac{1}{2}) + [1-\tilde{\alpha}_i]\frac{\beta_i}{2} = -1$.

Let us now consider the second contribution of the scaled Besov control, i.e. we take $v \in \big[0, [\lambda^{-1}(s-t)]^{\beta_i}\big]$. Write:

$$\begin{aligned}
&\int_{\mathbb{R}^d} h_v(z-\mathbf{y}_i) D_{\mathbf{y}_i} \cdot \big(\Theta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s,\mathbf{y})\big) d\mathbf{y}_i \\
=\ & \int_{\mathbb{R}^d} h_v(z-\mathbf{y}_i) D_{\mathbf{y}_i} \cdot \Big(\Theta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s,\mathbf{y}) - \Theta^{\vartheta,\lambda}_{i,(t,\mathbf{x})}(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\Big) d\mathbf{y}_i \\
=\ & \int_{\mathbb{R}^d} D^\vartheta \tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x},\mathbf{y}) \big\langle \mathbf{F}_{\lambda,i}(s,\mathbf{y}) - \mathbf{F}_{\lambda,i}(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}), D_z h_v(z-\mathbf{y}_i)\big\rangle d\mathbf{y}_i \\
&+ \int_{\mathbb{R}^d} \Big(D^\vartheta \tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x},\mathbf{y}) - D^\vartheta \tilde{p}^{\boldsymbol{\xi}}_\lambda(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\Big) \\
&\times \Big\langle \big(\mathbf{F}_{\lambda,i}(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}) - \mathbf{F}_{\lambda,i}(s,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}}\mathbf{F}_{\lambda,i}(s,\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}))(\mathbf{y}-\boldsymbol{\theta}^\lambda_{s,t}(\boldsymbol{\xi}))_{i-1}\big), D_z h_v(z-\mathbf{y}_i)\Big\rangle d\mathbf{y}_i \\
=:\ & \Big(\mathscr{T}_{\lambda,1} + \mathscr{T}_{\lambda,2}\Big)\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big),
\end{aligned} \tag{C.7}$$

thanks to the definition in (C.2) for the last identity.

$$
\begin{aligned}
&|\mathscr{T}_{\lambda,1}\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)|\\
\leq\ & \Lambda\int_{\mathbb{R}^d}\frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac12}}\frac{\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)}}\hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y})\lambda^{-i+\frac12}\big(\lambda^{\frac{2i-1}{2}}|z-\mathbf{y}_i|\big)^{\frac{2i-3+\gamma}{2i-1}}d\mathbf{y}_i\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)-(i-\frac12)+\frac{2i-3+\gamma}{2}}\int_{\mathbb{R}^d}\frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{2-\gamma}{4i-2}}}\frac{\hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)}}d\mathbf{y}_i.\\
\leq\ & \frac{\Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)-1+\frac{\gamma}{2}}}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)}}\hat{q}_{c\backslash i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))\mathcal{N}_{cv+\lambda^{\frac{2i-1}{2}}(s-t)^{2i-1}}\big(z-\boldsymbol{\theta}_{s,t}(\mathbf{x})_i\big).
\end{aligned}
\tag{C.8}
$$

Write now from (C.7):

$$
\begin{aligned}
&|\mathscr{T}_{\lambda,2}\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)|\\
\leq\ & \Lambda\int_{\mathbb{R}^d}d\mathbf{y}_i\frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac12}}\int_0^1 d\mu\frac{\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}\hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}\\
&\times|\mathbf{y}_i-z|\lambda^{-i+\frac12}\Big(\Big|\mathbf{F}_i\big(s,\lambda^{-\frac12}\mathbb{T}_\lambda(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)-\mathbf{F}_i\big(s,\lambda^{-\frac12}\mathbb{T}_\lambda(\mathbf{y}_{1:i-1},\boldsymbol{\theta}_{s,t}(\mathbf{x})_{i:n})\big)\Big|\\
&+\Big|\mathbf{F}_i\big(s,\lambda^{-\frac12}\mathbb{T}_\lambda(\mathbf{y}_{1:i-1},\boldsymbol{\theta}_{s,t}(\mathbf{x})_{i:n})\big)-\mathbf{F}_i\big(s,\lambda^{-\frac12}\mathbb{T}_\lambda\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})\big)\\
&-D_{\mathbf{x}_{i-1}}\mathbf{F}_i\big(s,\lambda^{-\frac12}\mathbb{T}_\lambda\boldsymbol{\theta}_{s,t}(\mathbf{x})\big)\big(\lambda^{-\frac12}\mathbb{T}_\lambda(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))\big)_{i-1}\Big|\Big)\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)}\int_{\mathbb{R}^d}d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\int_0^1 d\mu\frac{\hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}\\
&\times\Big((\lambda^{\frac{2i-1}{2}}|z-\boldsymbol{\theta}_{s,t}(\mathbf{x})_i|)^{\frac{2i-3+\gamma}{2i-1}}+(\lambda^{\frac{2(i-1)-1}{2}}|(\boldsymbol{\theta}_{s,t}(\mathbf{x})-\mathbf{y})_{i-1}|)^{1+\frac{\gamma}{2(i-1)-1}}\\
&+\sum_{k=i+1}^n(\lambda^{\frac{2j-1}{2}}|(\boldsymbol{\theta}_{s,t}(\mathbf{x})-\mathbf{y})_k|)^{\frac{2i-3+\gamma}{2k-1}}\Big).
\end{aligned}
$$

We have for any $\mu\in[0,1]$,

$$
|z-\boldsymbol{\theta}_{s,t}(\mathbf{x})_i|\leq\mu|z-\mathbf{y}_i|+|z+\mu(\mathbf{y}_i-z)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|,
$$

we thus derive

$$
\begin{aligned}
&|\mathscr{T}_{\lambda,2}\big(v,t,s,\mathbf{x},(\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})\big)|\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)}\int_{\mathbb{R}^d}d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\int_0^1 d\mu\frac{\hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}\\
&\times\Big(\lambda^{\frac{2i-3+\gamma}{2}}|\mathbf{y}_i-z|^{\frac{2i-3+\gamma}{2i-1}}+\mathbf{d}^{2i-3+\gamma}\Big(\lambda^{-1/2}\mathbb{T}_\lambda\boldsymbol{\theta}_{s,t}(\mathbf{x}),\lambda^{-1/2}\mathbb{T}_\lambda\Big(\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})\Big)\Big)\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)}\int_{\mathbb{R}^d}d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\int_0^1 d\mu\hat{q}_{c,\lambda}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\mu(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})\\
&\times\Big(\frac{\lambda^{\frac{2i-3+\gamma}{2}}v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+1-\frac{\gamma}{2}}}\Big)\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)}\hat{q}_{c\backslash i,\lambda}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})\\
&\times\int_0^1 d\mu\int_{\mathbb{R}^d}h_{cv}(z-\mathbf{y}_i)\mathcal{N}_{c\lambda^{\frac{2i-1}{2}}(s-t)^{2i-1}}(z+\mu(\mathbf{y}_i-z)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)d\mathbf{y}_i\\
&\times\Big(\frac{\lambda^{\frac{2i-3+\gamma}{2}}v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+1-\frac{\gamma}{2}}}\Big).
\end{aligned}
\tag{C.9}
$$

From (C.7), (C.8) and (C.9) we deduce, with the notation of (C.5):

$$\begin{aligned}
&\|h_v \star \Psi^{\vartheta,\lambda}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1})}\|_{L^1(\mathbb{R}^d,\mathbb{R})}\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n \vartheta_j(j-\frac12)}\hat q_{c\setminus i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))\\
&\times\Big(\frac{\lambda^{-1+\frac\gamma2}}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)}}+\frac{\lambda^{\frac{2i-3+\gamma}{2}}v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+1-\frac\gamma2}}\Big)\\
&\times\int_0^1 d\mu\int_{\mathbb{R}^d}dz\int_{\mathbb{R}^d}d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i)\mathcal{N}_{c\lambda^{\frac{2i-1}{2}}(s-t)^{2i-1}}(z+\mu(\mathbf{y}_i-z)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n \vartheta_j(j-\frac12)}\hat q_{c\setminus i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))\\
&\times\Big(\frac{\lambda^{-1+\frac\gamma2}}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)}}+\frac{\lambda^{\frac{2i-3+\gamma}{2}}v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+1-\frac\gamma2}}\Big).
\end{aligned}$$

The last identity is a again consequence of the change of variable $(w_1,w_2)=(z-\mathbf{y}_i,z+\mu(\mathbf{y}_i-z)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)$.

We now write:

$$\begin{aligned}
&\int_0^{[\lambda^{-1}(s-t)]^{\beta_i}} dv v^{\frac{\tilde\alpha_i}{2}-1}\|h_v\star\Psi^{\vartheta,\lambda}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1})}\|_{L^1(\mathbb{R}^d,\mathbb{R})}\\
\leq\ & \Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)}\hat q_{c\setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))\int_0^{[\lambda^{-1}(s-t)]^{\beta_i}}\frac{dv}{v}v^{\frac{\tilde\alpha_i}{2}}\\
&\times\Big(\frac{\lambda^{-1+\frac\gamma2}}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)}}+\frac{\lambda^{\frac{2i-3+\gamma}{2}}v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+\frac{2i-1}{2}}}+\frac{1}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)+1-\frac\gamma2}}\Big)\\
=:\ & \Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)}\hat q_{c\setminus i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}))\mathcal{B}^\lambda_{\vartheta,\beta_i}(t,s).
\end{aligned}\tag{C.10}$$

Let us now prove that

$$\mathcal{B}^\lambda_{\vartheta,\beta_i}(t,s)\leq\frac{C\lambda^{-1}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)-\frac\gamma2}}.\tag{C.11}$$

Integrating in v in (C.10) we derive:

$$\begin{aligned}
\mathcal{B}^\lambda_{\vartheta,\beta_i}(t,s)\ \leq\ & C(s-t)^{-\sum_{j=1}^n\vartheta_j(j-\frac12)}\Big[\lambda^{-1+\frac\gamma2}[\lambda^{-1}(s-t)]^{\beta_i(\frac{\tilde\alpha_i}{2}-\frac{2-\gamma}{4i-2})}\\
&+\lambda^{\frac{2i-3+\gamma}{2}}[\lambda^{-1}(s-t)]^{\beta_i(\frac{\tilde\alpha_i}{2}+\frac{2i-3+\gamma}{2(2i-1)})}(s-t)^{i-\frac12}+[\lambda^{-1}(s-t)]^{\beta_i\frac{\tilde\alpha_i}{2}}(s-t)^{-1+\frac\gamma2}\Big].
\end{aligned}$$

Recall now from the proof of Lemma 11 that:

$$\beta_i\Big(\frac{\tilde\alpha_i}{2}-\frac{2-\gamma}{4i-2}\Big)-\frac\gamma2\geq0,\ \beta_i\Big(\frac{\tilde\alpha_i}{2}+\frac{2i-3+\gamma}{2(2i-1)}\Big)-\frac{2i-1}{2}-\frac\gamma2\geq0,\ \beta_i\frac{\tilde\alpha_i}{2}-1\geq0,$$

with $\beta_i=\frac{(2i-3)(2i-1)}{2i-3-\gamma},\tilde\alpha_i=\frac{2+\gamma}{2i-1}$. Therefore, since $(s-t)/\lambda\leq1$:

$$\mathcal{B}^\lambda_{\vartheta,\beta_i}(t,s)\leq C(s-t)^{-\sum_{j=1}^n\vartheta_j(j-\frac12)}\Big[\lambda^{-1+\frac\gamma2}[\lambda^{-1}(s-t)]^{\frac\gamma2}+\lambda^{-1+\frac\gamma2}[\lambda^{-1}(s-t)]^{\frac\gamma2}+[\lambda^{-1}(s-t)](s-t)^{-1+\frac\gamma2}\Big],$$

which precisely gives (C.11).

Plugging (C.11) into (C.10) and from (C.6) we eventually get:

$$\int_0^1\frac{dv}{v}v^{\frac{\tilde\alpha_i}{2}}\|h_v\star\Psi^{\vartheta,\lambda}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})}\leq\frac{\Lambda\lambda^{\sum_{j=1}^n\vartheta_j(j-\frac12)-1}}{(s-t)^{\sum_{j=1}^n\vartheta_j(j-\frac12)-\frac\gamma2}}\hat q_{c\setminus i,\lambda}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})),$$

which is precisely the stated control. The term $\|\varphi(D)\Psi^{\vartheta,\lambda}_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}\|_{L^1(\mathbb{R}^d,\mathbb{R})}$ appearing in the Besov norm could be handled similarly. The result is complete. □

Paul-Éric Chaudru de Raynal. *Université de Savoie Mont Blanc, CNRS, LAMA, F-73000 Chambéry, France.* `E-mail:` *pe.deraynal@univ-smb.fr.*

Igor Honoré. *Laboratoire de Modélisation Mathématique d'Evry (LaMME), Université d'Evry Val d'Essonne, 23 Boulevard de France 91037 Evry.* `E-mail:` *igor.honore@univ-evry.fr.*

Stéphane Menozzi. *Laboratoire de Modélisation Mathématique d'Evry (LaMME, UMR CNRS 8071), Université d'Evry Val d'Essonne, Paris Saclay, 23 Boulevard de France 91037 Evry, France and Laboratory of Stochastic Analysis, National Research University Higher School of Economics, Pokrovsky Boulevard, 11, Moscow, Russian Federation.* `E-mail:` *stephane.menozzi@univ-evry.fr*

References

[1] D. R. Adams and L. I. Hedberg. *Function spaces and Potential Theory*. Springer, 1996.
[2] P.-É. Chaudru de Raynal. Strong existence and uniqueness for degenerate SDE with Hölder drift. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53(1):259–286, February 2017.
[3] P.-É. Chaudru de Raynal. Weak regularization by stochastic drift: result and counter example. *Discrete and Continuous Dynamical Systems (Series A)*, 38–3:1269–1291, 2018.
[4] P-É. Chaudru de Raynal and S. Menozzi. Regularization effects of a noise propagating through a chain of differential equations: an almost sharp result. *Transactions of the American Mathematical Society*, 2020. DOI:10.1090/tran/7947.
[5] P.-É. Chaudru de Raynal, S. Menozzi, and E. Priola. Schauder estimates for drifted fractional operators in the supercritical case. *J. Funct. Anal.*, 278(8):108425, 2020.
[6] P.É. Chaudru de Raynal, I. Honoré, and S. Menozzi. Strong Uniniqueness for a chain of degenerate Kolmogorov equations. working paper or preprint, 2018.
[7] 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.
[8] 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.
[9] E. Fedrizzi, F. Flandoli, E. Priola, and J. Vovelle. Regularity of stochastic kinetic equations. *Electronic Journal of Probability*, 22:♯ 48, 2017.
[10] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, 1964.
[11] Z. Hao, M. Wu, and X. Zhang. Schauder's estimate for nonlocal kinetic equations and its applications, 2019.
[12] L. Huang and S. Menozzi. A Parametrix Approach for some Degenerate Stable Driven SDEs. *Annales Instit. H. Poincaré*, 52(4):1925–1975, 2016.
[13] A. M. Il'in, A. S Kalashnikov, and O. A. Oleinik. Second-order linear equations of parabolic type. *Uspehi Mat. Nauk*, 17–3(105):3–146, 1962.
[14] V. Kolokoltsov. *Markov Processes, Semigroups and Generators*. De Gruyter, 2011.
[15] S. N. Kružkov, A. Castro, and M. Lopes. Schauder type estimates, and theorems on the existence of the solution of fundamental problems for linear and nonlinear parabolic equations. *Dokl. Akad. Nauk SSSR*, 220:277–280, 1975.
[16] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*. Graduate Studies in Mathematics 12. AMS, 1996.
[17] N. V. Krylov and E. Priola. Elliptic and parabolic second-order PDEs with growing coefficients. *Comm. Partial Differential Equations*, 35(1):1–22, 2010.
[18] N. V. Krylov and E. Priola. Poisson stochastic process and basic Schauder and Sobolev estimates in the theory of parabolic equations. *Arch. Ration. Mech. Anal.*, 225(3):1089–1126, 2017.
[19] H. Kunita. *Stochastic flows and Stochastic Differential Equations*. Cambridge Studies in Advanced Mathematics, 1997.
[20] P.-G. Lemarié-Rieusset. *Recent developments in the Navier-Stokes problem*. CRC Press, 2002.
[21] L. Lorenzi. Estimates of the derivatives for a class of parabolic degenerate operators with unbounded coefficients in $\mathbb{R}^N$. *Annali della Scuola Normale Superiore di Pisa. Classe di Scienze. Serie V*, 2, 01 2005.
[22] L. Lorenzi. Schauder estimates for degenerate elliptic and parabolic problems with unbounded coefficients in $\mathbb{R}^N$. *Differential Integral Equations*, 18–5:531–566, 2005.
[23] L. Lorenzi. Optimal Hölder regularity for nonautonomous Kolmogorov equations. *Discrete Contin. Dyn. Syst. Ser. S*, 4(1):169–191, 2011.
[24] A. Lunardi. Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbb{R}^n$. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 24(1):133–164, 1997.
[25] L. Marino. Schauder estimates for degenerate stable kolmogorov equations, 2019.
[26] H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1:43–69, 1967.
[27] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electronic Communications in Probability*, 17:234–250, 2011.
[28] S. Menozzi. Martingale problems for some degenerate Kolmogorov equations. *Stoc. Proc. Appl.*, 128-3:756–802, 2018.
[29] E. Priola. Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.*, 194(2):117–153, 2009.
[30] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.
[31] H. Triebel. *Theory of function spaces, II*. Birkhauser, 1983.
[32] X. Zhang. Stochastic Hamiltonian flows with singular coefficients. *Sci. China Math.*, 61(8):1353–1384, 2018.