Regularization effects of a noise propagating through a chain of differential equations: an almost sharp result

Paul-Eric Chaudru de Raynal, Stephane Menozzi

REGULARIZATION EFFECTS OF A NOISE PROPAGATING THROUGH A CHAIN OF DIFFERENTIAL EQUATIONS: AN ALMOST SHARP RESULT

PAUL-ERIC CHAUDRU DE RAYNAL AND STÉPHANE MENOZZI

ABSTRACT. We investigate the effects of the propagation of a non-degenerate Brownian noise through a chain of deterministic differential equations whose coefficients are *rough* and satisfy a weak like Hörmander structure (i.e. a non-degeneracy condition w.r.t. the components which transmit the noise). In particular we characterize, through suitable counter-examples, almost sharp regularity exponents that ensure that weak well posedness holds for the associated SDE. As a by-product of our approach, we also derive some density estimates of Krylov type for the weak solutions of the considered SDEs.

1. INTRODUCTION AND MAIN RESULTS

In this work we are interested in studying the weak regularization effects of a Brownian noise propagating through a chain of n d-dimensional oscillators. Namely, we establish weak uniqueness for Stochastic Differential Equations (SDEs in short) of the following type:

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

(1.1)

In the above equation, $(W_t)_{t \geq 0}$ stands for a d-dimensional Brownian motion and the components $(X_t^i)_{i \in [\![1,n]\!]}$ are $\mathbb{R}^d$-valued as well. We suppose that the $(F_i)_{i \in [\![2,n]\!]}$ satisfy a kind of weak Hörmander condition, i.e. the matrices $\left(D_{x_{i-1}} F_i(t, \cdot)\right)_{i \in [\![2,n]\!]}$ have full rank. However, the coefficients $(F_i)_{i \in [\![2,n]\!]}$ can be rather *rough* in their other entries. Namely, Hölder continuous or even in a suitable $L^q - L^p$ space for F_1, where the parameter q relates to the time integrability and p to the spatial one. We assume as well that the diffusion coefficient σ is bounded from above and below and spatially Hölder continuous.

We emphasize that, under these conditions, the Stroock and Varadhan Theory for weak uniqueness does not apply. This especially comes from the specific degenerate framework considered here: the noise in the ith component only comes from the

---

$(i-1)^{\text{th}}$ component, $2 \leq i \leq n$, through the non-degeneracy of the gradients $\left(D_{x_{i-1}} F_i(t, \cdot)\right)_{i \in [\![2,n]\!]}$ (components which transmit the noise). We nevertheless show that the system is well posed, in a weak sense, when the drift of the first component is Hölder continuous or bounded in supremum norm or in suitable $L^q - L^p$ norm and the drift functions of the other components are only Hölder continuous with respect to the variables that do not transmit the noise. Denoting by $(\beta_i^j)_{2 \leq i \leq j \leq n}$ the Hölder index of the drift of the ith component w.r.t. the jth variable we assume $\beta_i^j \in \left([(2i-3)/(2j-1)], 1\right]$. We also show that these thresholds are (almost) sharp thanks to appropriate counter examples.

Also, as a by-product of our analysis, we prove that the density of the unique weak solution of the system satisfies Krylov-like estimates.

Weak and strong regularization by noise. Strong and weak well posedness of stochastic systems outside the classical Cauchy-Lipschitz framework have motivated a lot of works since the last past four decades[1].

Concerning the strong well posedness, the first result in that direction is due to Zvonkin [Zvo74] who showed that one-dimensional non degenerate Brownian driven stochastic differential equations with bounded and measurable drift and Hölder continuous diffusion matrix are well posed for Hölder index strictly greater than 1/2. Then, Veretennikov [Ver80] generalized the result to the multidimensional case for a Lipschitz diffusion matrix. These results have been recently revisited in the work of Krylov and Röckner [KR05], where SDEs with additive Brownian noise and locally integrable drift are shown to be strongly well posed and Zhang [Zha10] who extended the Krylov and Röckner result to SDEs with multiplicative noise and weakly Lipschitz diffusion matrix (*i.e.* in Sobolev Sense). Similar issues are handled as well in [FF11]. Also, we can mention the recent work by Davie [Dav07] in which *path-by-path* uniqueness is proved for non degenerate Brownian SDEs with bounded drift and the approach of Catellier and Gubinelli [CG16] (which also relies on *path-by-path* uniqueness) where SDEs with additive fractional Brownian noise are investigated. Finally, let us mention the work [GO13] where the strong well-posedness of a particular one dimensional system with singular inhomogeneous drift is proved. We refer the reader to the Saint Flour Lecture notes of Flandoli [Fla11] where a very interesting and general account on the topic is given.

On the other hand, and still in the Brownian framework, it has been shown that non degenerate stochastic systems are well posed in a weak sense as soon as the drift function is measurable and bounded and the diffusion matrix only a continuous (in space) function. This is the celebrated theory of the *martingale problemma* put on complete mathematical framework by Stroock and Varadhan, see [SV79]. Weak well posedness results for non degenerate SDE with additive noise have also been explored recently: Flandoli, Issoglio and Russo showed in [FIR17] that multidimensional non degenerate SDEs with non-homogeneous distributional drift are well posed as soon as the regularity index is strictly greater than $-1/2$. At the same time, Delarue and Diel proved in [DD15] that the result still holds when the regularity index is strictly greater than $-2/3$ in the one-dimensional case. This last work has then been generalized by Cannizzaro and Chouk [CC18] to the

---

[1]In the presentation below, we will mainly focus on Brownian driven SDEs. We can refer to the recent work of Priola [Pri18] for the more general Lévy driven case in the non-degenerate framework.

multidimensional setting. Note however that, in the two last mentioned works, the Authors assumed that the drift can be enhanced into a rough path structure.

All the above strong and weak results deeply rely on the non-degeneracy assumption imposed to the noise and illustrate what is usually called, following the terminology of Flandoli, a *regularization by noise phenomenon*. Here, the regularization has to be understood as follows: while an ordinary differential system could be ill-posed when the drift is less than Lipschitz (or weakly Lipschitz [PL89]), the analogous non degenerate stochastic system is well posed (in a strong or a weak sense). To obtain this kind of result, the noise plays a central role. A striking example to understand the phenomenon is the Peano Example : while the deterministic scalar ODE

$$(1.2) \qquad \dot{Y}_t = \text{sign}(Y_t)|Y_t|^\alpha dt, \quad Y_0 = 0, \quad \alpha \in (0,1),$$

has an infinite number of solutions that could still be trapped in the singularity for any given time, the corresponding *Brownian* SDE is strongly well posed. In [DF14], Delarue and Flandoli put the phenomenom in light: in short time, the fluctuations of the noise dominate the system so that the solution leaves the singularity and in long time, the drift dominates and constrains the solution to fluctuate around one of the extremal solutions of the Peano Example. Hence, there is a strong competition, in short time, between the irregularity of the drift and the fluctuations of the noise.

Here, our result mostly emphasizes a *regularization phenomenon coming from a degenerate noise* (*i.e.* when $n \geq 2$ in (1.1)). In view of the above discussion, it is clear that the degeneracy may dramatically damage the *regularization by noise* properties and, in order to preserve some regularization effect, the noise still needs to have a way to propagate through the system. Such kind of behavior will typically hold when the system satisfies a so-called Hörmander condition for hypoellipticity, see Hörmander's seminal work on the topic [Hör67].

In our case, we suppose the drift of each component to be differentiable w.r.t. its first variable and the resulting gradient to be non-degenerate, but only Hölder continuous in the other variable. This last non-degeneracy assumption is the reason why this kind of condition is called *weak Hörmander condition*. Namely, up to some regularization of the diffusion coefficient, the drift is needed to span the space through Lie Bracketing. Also, in comparison with the *general* Hörmander setting, the specific drift structure we consider here is such that at each level of the chain we only require one additional Lie bracket to generate the corresponding directions, up to some regularization of the diffusion coefficient again. This setting allows us to recover some regularization effect of the noise at each level of the chain. We also refer for similar issues to Section 2 in [LP94].

Concerning the strong regularization effects of a degenerate noise in a weak Hörmander setting, one of the first result has been given by Chaudru de Raynal in [CdR17] and concerns strong well posedness of the above system (1.1) when $n = 2$. It is shown in that case that the system is well posed as soon as the drift coefficients are Hölder continuous with Hölder exponent strictly greater than 2/3 w.r.t. the degenerate variable and when the diffusion matrix is Lipschitz continuous in space. This result was then extended by Wang and Zhang [WZ16] under Hölder-Dini conditions with the same critical Hölder threshold 2/3. We also mention, again for two oscillators and when the degenerate component only depends linearly of the

non-degenerate variable and not on the degenerate component, the recent work by Fedrizzi, Flandoli, Priola and Vovelle [FFPV17] who address the case of a weakly differentiable non-degenerate drift with order of weak differentiation strictly greater than 2/3. The critical case corresponding to the exponent 2/3 has been discussed by Zhang [Zha18].

From the weak regularization by noise viewpoint, in our current weak Hörmander setting, one of the first results is the work by Menozzi [Men11]. The key-point there is to exploit some smoothing effects of a suitable parametrix kernel, associated with a Gaussian linearization of (1.1), which had already been used by Delarue and Menozzi in [DM10] to derive heat-kernel bounds for the solution of (1.1). In [Men11], it is shown that the system (1.1) is (weakly) well posed for a spatially Lipschitz continuous drift satisfying the previously mentioned non-degeneracy condition, and a spatially Hölder continuous diffusion coefficient. The result was then extended in [Men18] for a spatially continuous diffusion coefficient, following the martingale problem approach establishing some suitable Calderón-Zygmund estimates for a degenerate Gaussian kernel and appropriate non-standard localization arguments. Also, in the case of two oscillators, Zhang showed in [Zha18] that when the degenerate component only depends linearly on the non-degenerate variable and not on the degenerate component, the system is weakly well posed as soon as the drift of the first component satisfies some local integrability conditions and when the diffusion coefficient is continuous. At the same time, Chaudru de Raynal showed in [CdR18] that when $n = 2$ the system is well posed in a weak sense as soon as the drift of each component are at least $1/3$ Hölder continuous in the degenerate variable and showed that this result is (almost) sharp for the drift of the second oscillator thanks to an appropriate counter example.

Hence, the minimal threshold obtained for the Hölder regularity of the drift is not an artefact: this can be seen as the price to pay to balance the degeneracy of the noise. Especially, in view of the previous discussion on the Delarue and Flandoli work, it is related to the fact that the fluctuations of the degenerate noise are not strong enough to push the solution away from the singularity if the drift is too irregular. As said above, this is illustrated in [CdR18] where a counter example is built thanks to this observation. Namely, it is shown that any stochastic perturbation of the Peano Example (1.2) has to have (at least) fluctuations of order strictly lower than $1/(1-\alpha)$ in order to restore (weak) uniqueness. Hence, for two oscillators, assuming that the dynamics of the degenerate component is driven by (1.2) perturbed by the integral of the Brownian source plugged in the non-degenerate component in (1.1), we have that the typical variance of the noise is of order $t^{3/2}$ at time t. From the above condition, we indeed find $1/(1-\alpha) > 3/2 \iff \alpha > 1/3$.

Organization of our paper. The paper is organized as follows. Our assumptions and main results are stated at the end of the current Section. We present in Section 2 the main tools that allow to derive our results. Namely, a suitable Gaussian linearization of the initial model (1.1) around a deterministic Cauchy-Peano flow of the initial system of ODEs (corresponding to (1.1) taking $\sigma = 0$). In particular, since we consider *rough* coefficients, we establish therein measurability properties and bi-Lipschitz like regularity for such flows. The well posedness of the martingale problem for the operator associated with (1.1) is then obtained in Section 3. Section

4 is eventually dedicated to a class of counter examples which emphasize the *almost* sharpness of our results.

Assumptions and main result. Our assumptions are the following:

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

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

where $|\cdot|$ denotes the Euclidean norm, $\langle \cdot, \cdot \rangle$ is the inner product and * stands for the transpose.

(S) The coefficients $\sigma(t, \cdot)$, $\big(F_i(t, \mathbf{0})\big)_{i \in [\![2,n]\!]}$ are assumed to be bounded measurable in time. Also, the diffusion coefficient $\sigma(t, \cdot)$ is uniformly η-Hölder continuous in space, for some $\eta > 0$ uniformly in time. The drift entries $\big(F_i(t, \cdot)\big)_{i \in [\![2,n]\!]}$ are s.t. for all $(z_i, \cdots, z_n) \in \mathbb{R}^{d(n-(i-1))}$, the mapping $z \in \mathbb{R}^d \mapsto F_i(t, z, z_i, \cdots, z_n)$ is in $C^{1+\eta}(\mathbb{R}^d, \mathbb{R}^d)$ uniformly in time[2] and w.r.t. $(z_i, \ldots, z_n)$. Moreover, $D_{x_{i-1}}F_i(t, \cdot)$ is bounded. Eventually, the mappings $(z_i, \cdots, z_n) \in \mathbb{R}^{d(n-(i-1))} \mapsto F_i(t, z, z_i, \cdots, z_n)$ are, for $j \in [\![i, n]\!]$, β_i^j-Hölder continuous in the variable z_j, with $\beta_i^j > 0$, uniformly in time and in $z \in \mathbb{R}^d$.

(D) The first entry of the drift F_1 is supposed to satisfy one of the following assumptions:

(a) The measurable mapping $t \in \mathbb{R}_+ \mapsto F_1(t, \mathbf{0})$ is bounded and $F_1(t, \cdot)$ is Hölder continuous in space[3] uniformly in time.

(b) The measurable mapping $(t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^{nd} \mapsto F_1(t, \mathbf{x})$ is bounded.

(c) $F_1 \in L^q(\mathbb{R}_+, L^p(\mathbb{R}^{nd}))$, $\frac{n^2 d}{p} + \frac{2}{q} < 1, p \geq 2, q > 2$.

Observe that case (b) can be viewed as a particular case of (c), corresponding to $p = q = \infty$. Since the techniques used to address those two cases are rather different (see Section 3.2), we prefer to consider them separately.

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

We say that assumption **(A)** is in force whenever **(UE)**, **(S)**, **(H)** and at least one of the three items in **(D)** hold.

Theorem 1 (Weak Uniqueness and Hölder continuity indexes). *Assume **(A)** and that the following conditions hold:*

$$(1.3) \qquad \forall i \in [\![2, n]\!], \; j \in [\![i, n]\!], \beta_i^j \in \left(\frac{2i-3}{2j-1}, 1 \right].$$

Then, the martingale problem associated with $(L_t)_{t \geq 0}$ where for all $\phi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$, $\mathbf{x} \in \mathbb{R}^{nd}$,

$$(1.4) \qquad L_t\varphi(\mathbf{x}) = \langle F(t, \mathbf{x}), D_{\mathbf{x}}\phi(\mathbf{x}) \rangle + \frac{1}{2}\mathrm{tr}\big(a(t, \mathbf{x})D_{x_1}^2\phi(\mathbf{x})\big), \quad a := \sigma\sigma^*,$$

---

[2] For the sake of clarity we chose the same regularity index for σ and $(D_{x_{i-1}}F_i)_{i \in [\![2,n]\!]}$, but the result remains true for any η_σ Hölder continuous σ and $\eta_{\mathbf{F},i}$ Hölder continuous $D_{x_{i-1}}F_i$, provided η_σ and $\eta_{\mathbf{F},i}$ belong to $(0, 1]$.

[3] Actually one can assume that F_1 is β_1^j Hölder continuous in the jth variable for any β_1^j in $(0, 1]$.

is well posed, i.e. there exists a unique probability measure $\mathbb{P}$ *on* $C(\mathbb{R}_+, \mathbb{R}^{nd})$ *s.t. denoting by* $(\mathbf{X}_t)_{t\geq 0}$ *the associated canonical process, for every* $\varphi \in C_0^{1,2}(\mathbb{R}_+ \times \mathbb{R}^{nd}, \mathbb{R})$ *and conditionally to* $\mathbf{X}_t = \mathbf{x}$ *for* $t \geq 0$, $\left(\varphi(s, \mathbf{X}_s) - \varphi(t, \mathbf{x}) - \int_t^s (\partial_u + L_u)\varphi(u, \mathbf{X}_u) du \right)_{s\geq t}$ *is a* $\mathbb{P}$-*martingale. In particular weak uniqueness holds for the SDE* (1.1).

The transition probability $P(t, s, \mathbf{x}, \cdot)$, determined by $(L_s)_{s\geq 0}$, is s.t. for a given $T > 0$, almost all $s \in (t, T]$ and all $\Gamma \in \mathcal{B}(\mathbb{R}^{nd})$: $P(t, s, \mathbf{x}, \Gamma) = \int_\Gamma p(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y}$.

Furthermore, we have the following Krylov-like estimate: for all fixed $T > 0$ and every $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))$ with $\frac{n^2 d}{p'} + \frac{2}{q'} < 2, p' > 1, q' > 1, (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$:

$$(1.5) \qquad |\mathbb{E}^{\mathbb{P}_{t,\mathbf{x}}}[\int_t^T f(s, \mathbf{X}_s) ds]| \leq C\|f\|_{L^{q'}(([0,T], L^{p'}(\mathbb{R}^{nd})))},$$

where $\mathbb{E}^{\mathbb{P}_{t,\mathbf{x}}}$ *denotes the expectation w.r.t.*

$$\mathbb{P}_{t,\mathbf{x}}[\cdot] := \mathbb{P}[\cdot | \mathbf{X}_t = \mathbf{x}] \quad and \quad C := C((\mathbf{A}), p', q', T).$$

Hence, our Theorem allows to recover almost all the previously mentioned works on weak well posedness and provides an extension for the full chain. It also permits us to avoid any regularity assumption on the drift of the diffusion component so that we recover the Stroock and Varadhan result in the case $n = 1$ up to an arbitrary small Hölder exponent on the continuity of the diffusion matrix. Concerning this last point, we feel that using the localization strategy proposed by Menozzi in [Men18] we may be able to get rid of this assumption and only assume the diffusion coefficient to be continuous in space. Indeed, using our results (say Lemma 5 below together with condition (1.3)) should allow one to adapt the approach of [Men18] and extend Theorem 1 to continuous diffusion matrix.

We also underline that our result allows to deal with a large class of different drifts for the non degenerate component: the system can be globally with sub-linear growth (Assumption **(A)**-(a)), rough and bounded (Assumption **(A)**-(b)) or only suitably integrable and rough (Assumption **(A)**-(c)).

Moreover, the following result shows that Theorem 1 is *almost* sharp. By *almost*, we mean that the critical lower thresholds in (1.3) and in **(D)**-(c) are not yet handled. Namely we have:

Theorem 2 (Almost sharpness)**.** *There exists* $\mathbf{F}$ *satisfying* **(UE)**, **(S)**, **(H)** *and such that:*

$$(1.6) \qquad \exists i \in [\![2, n]\!], \ j \in [\![i, n]\!], \beta_i^j < \frac{2i - 3}{2j - 1},$$

or

$$(1.7) \qquad F_1 \in L^q([0, T], L^p(\mathbb{R}^{nd})), \quad \frac{n^2 d}{p} + \frac{2}{q} > 1, \quad p \geq 2, q > 2$$

for which weak uniqueness fails for the SDE (1.1).

We first emphasize that there are already some results in that direction: in [BFGM19] the Authors show that when $n = 1$ and when the integrability condition (1.7) is not satisfied (*i.e.* in the supercritical case) equation (1.1) does not have a weak solution. Another counter example to that case can be found in [GO13]. Note that in comparison with the results in [FIR17], [DD15], [CC18], the almost

sharpness of the integrability condition (1.7) has to be understood for drifts assumed to be functions and not distributions.

Secondly, it has been proved in [CdR18], that for all i in $[\![2, n]\!]$ the Hölder exponents β_i^i, are also almost sharp, thanks to a class of counter examples based on stochastic perturbations of the Peano example (1.2).

Thirdly, we feel that the Hölder continuity assumption assumed on $D_{x_{i-1}} F_i(t, \cdot)$ is a technical artefact. Nevertheless, relaxing this assumption to consider the $\left(D_{x_{i-1}} F_i(t, \cdot) \right)_{i \in [\![2, n]\!]}$ are just continuous functions of x_{i-1} is definitely more tricky. Indeed, in that case, our approach based on parametrix fails and the natural approach, relying on harmonic analysis techniques, seems very involved.

And last, but not least, let us notice that the two thresholds for the drift component (say (1.3) and condition (**A**)-(c)) will appear many times throughout this work as a minimal value for making our proof work (see the proofs of Lemmas 5, 6, 9 and 11). This underlines the sharpness of the exponent for the strategy we used and explains why the critical case of these conditions is not investigated here. It seems indeed clear for us that the critical case requires different tools as those presented here.

Remark 1. Before entering into the proof, we indicate that, from a more analytic viewpoint, the techniques we develop in the current work could also be used in order to derive well-posedness results in the mild sense (see [SV79], [Kol11]) for the corresponding degenerate parabolic PDE, which involves *rough* coefficients in a weak Hörmander setting, when the source term belongs to appropriate Lebesgue spaces.

2. STRATEGY AND KEY TOOLS

Our strategy relies on the martingale problem approach. Hence, we face two difficulties: firstly, we have to show the existence of a solution to the martingale problem in our current setting, which becomes quite tricky under (**D**)-(c) while it is quite obvious under (**D**) - (a) and (b); secondly we have to show that the solution is unique which is the real core of this paper.

About uniqueness. Usual approaches to uniqueness for the martingale problem associated with a given operator are based on a perturbative method. Let us detail two of the main strategies developed in the literature. The historical one due to Stroock and Varadhan [SV79] consists in exploiting some L^p controls on the derivatives of a suitable Gaussian heat kernel (*parametrix*). It allows, in the non-degenerate diffusive case, to establish well posedness provided the diffusion coefficient is solely continuous. As a by-product of this approach, Krylov like estimates of type (1.5) are obtained, emphasizing that the canonical process associated with the solution actually possesses a density which enjoys integrability properties up to a certain threshold. Extensions of these types of results to the chain (1.1) are available in [Men18].

On the other hand, a more recent approach is due to Bass and Perkins [BP09]. In the non-degenerate setting, under the stronger assumption of Hölder continuity of the diffusion coefficient, it only requires pointwise controls of an underlying *parametrix* kernel. This approach has then been successfully extended under the considered weak type Hörmander setting to a chain of type (1.1) in [Men11] in the diffusive case and in [HM16] for more general stable driven degenerate SDEs of type

(1.1) with dimension restrictions. It is actually more direct than the approach of Stroock and Varadhan. However, its drawback is that it does not provide *a priori* information on the density of the underlying canonical process.

Let us underline that in both cases, the *parametrix* plays a central role. This approach consists in expanding the generator of a given stochastic process around a suitable *proxy* generator which can be well handled. The point is then to control in a suitable way the associated approximation error. In our current degenerate diffusive setting, since the SDE is Brownian driven, the difficulty is to exhibit an appropriate Gaussian process that fulfills the previously indicated constraints.

When the drift **F** is *smooth* in addition to **(A)**, say globally Lipschitz continuous, it has been shown in [DM10], [Men11], [Men18] that a good proxy is provided by the linearization around the deterministic flow associated with (1.1) (*i.e.* when $\sigma = 0$ therein) leading to consider a multi-scale Gaussian process as parametrix. It is therefore a natural candidate for the current work. Anyhow, under **(A)**, we do not have anymore a deterministic flow in the usual Cauchy-Lipschitz sense. A first difficulty is therefore to deal with non-smooth and non-unique Cauchy-Peano flows. It actually turns out that any measurable flow solving (1.1) with $\sigma = 0$ is a good candidate to make our machinery work. The specific controls associated with those objects are presented in Section 2.2.

Also, in order to handle very rough drifts for the non degenerate component, from **(D)**-(c) $F_1 \in L^q([0,T], L^p(\mathbb{R}^{nd}))$, we are led to apply the Girsanov transform to the equation with $F_1 = 0$. To do so requires to have some *a priori* knowledge of the corresponding density. This is why, to achieve our goal, the Stroock and Varadhan approach leading to estimate (1.5), seems to be the natural framework.

In comparison with the approach based on the Zvonkin Transform initiated in [CdR18], our approach allows to obtain a clever analysis of the chain in the sense that we are here able to enlight the almost sharp regularity needed for each component of the drift and w.r.t. each variable. This last point is not possible *via* the Zvonkin Transfom which is more global and does not permit this distinction. Accordingly to the works [CdR17], [FFPV17], the Zvonkin approach seems more suited to derive strong uniqueness. In that case a global threshold appears for each variable at each level of the chain.

About existence. Concerning the existence part, our proof consists in adapting to our degenerate setting the idea introduced by Portenko [Por90] and used by Krylov and Röckner [KR05] as well to build local weak solutions in the non-degenerate case.

Usual notations. In what follows, we denote a quantity in $\mathbb{R}^{nd}$ by a bold letter: i.e. **0**, stands for zero in $\mathbb{R}^{nd}$ and we denote by $(X_t^1, \ldots, X_t^n)_{t \geq 0}$ the components of $(\mathbf{X}_t)_{t \geq 0}$. Introducing the embedding matrix B from $\mathbb{R}^d$ into $\mathbb{R}^{nd}$, i.e. $B = (I_{d \times d}, 0, \ldots, 0)^*$, where "$*$" stands for the transpose, we rewrite accordingly (1.1) in the shortened form

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

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

The deterministic backward flow. In the following, we will first assume for the sake of simplicity that assumption **(D)**-(a) is in force. The extension to cases

(b) and (c) will be discussed later on. Introduce now, for fixed $T > 0$, $\mathbf{y} \in \mathbb{R}^{nd}$ and $t \in [0, T]$ the backward flow:

$$(2.1) \qquad \dot{\boldsymbol{\theta}}_{t,T}(\mathbf{y}) = \mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y})), \ \boldsymbol{\theta}_{T,T}(\mathbf{y}) = \mathbf{y}.$$

Remark 2. We mention carefully that from the Cauchy-Peano theorem, a solution to (2.1) exists. Indeed, the coefficients are continuous and have at most linear growth.

2.1. Linearized Multi-scale Gaussian Process and Associated Controls.

We now want to introduce the forward linearized flow around a solution of (2.1). Namely, we consider for $s \geq 0$ the deterministic ODE

$$(2.2) \qquad \dot{\tilde{\boldsymbol{\phi}}}_s = \mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})) + D\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y}))[\tilde{\boldsymbol{\phi}}_s - \boldsymbol{\theta}_{s,T}(\mathbf{y})],$$

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

$$D\mathbf{F}(s, \mathbf{z}) = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ D_{z_1} F_2(s, \mathbf{z}) & 0 & \cdots & \cdots & 0 \\ 0 & D_{z_2} F_3(s, \mathbf{z}) & 0 & 0 & \vdots \\ \vdots & 0 & \ddots & \vdots & \\ 0 & \cdots & 0 & D_{z_{n-1}} F_n(s, \mathbf{z}) & 0 \end{pmatrix}$$

denotes the subdiagonal of the Jacobian matrix $D_{\mathbf{z}}\mathbf{F}(s, \cdot)$ at point $\mathbf{z}$. Introduce now for a given $(T, \mathbf{y}) \in \mathbb{R}_{+*} \times \mathbb{R}^{nd}$, the resolvent $(\tilde{\mathbf{R}}^{T,\mathbf{y}}(t, s))_{s,t \geq 0}$ associated with the partial gradients $(D\mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y})))_{t \geq 0}$ which satisfies for $(s, t) \in (\mathbb{R}_+)^2$:

$$(2.3) \qquad \partial_s \tilde{\mathbf{R}}^{T,\mathbf{y}}(s, t) = D\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})) \tilde{\mathbf{R}}^{T,\mathbf{y}}(s, t), \ \tilde{\mathbf{R}}^{T,\mathbf{y}}(t, t) = I_{nd \times nd}.$$

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

We consider now the stochastic linearized dynamics $(\tilde{\mathbf{X}}_s^{T,\mathbf{y}})_{s \in [t,T]}$:

$$d\tilde{\mathbf{X}}_s^{T,\mathbf{y}} = [\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})) + D\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y}))(\tilde{\mathbf{X}}_s^{T,\mathbf{y}} - \boldsymbol{\theta}_{s,T}(\mathbf{y}))]ds + B\sigma(s, \boldsymbol{\theta}_{s,T}(\mathbf{y}))dW_s,$$
$$(2.4) \quad \forall s \in [t, T], \ \tilde{\mathbf{X}}_t^{T,\mathbf{y}} = \mathbf{x}.$$

From equations (2.2) and (2.3) we explicitly integrate (2.4) to obtain for all $v \in [t, T]$:

$$(2.5)$$
$$\tilde{\mathbf{X}}_v^{T,\mathbf{y}} = \tilde{\mathbf{R}}^{T,\mathbf{y}}(v, t)\mathbf{x} + \int_t^v \tilde{\mathbf{R}}^{T,\mathbf{y}}(v, s)\Big(\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})) - D\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y}))\boldsymbol{\theta}_{s,T}(\mathbf{y})\Big)ds$$
$$+ \int_t^v \tilde{\mathbf{R}}^{T,\mathbf{y}}(v, s)B\sigma(s, \boldsymbol{\theta}_{s,T}(\mathbf{y}))dW_s.$$

Denoting by $\tilde{\boldsymbol{\theta}}_{s,t}^{T,\mathbf{y}}(\mathbf{x})$ the solution of (2.2) with starting point $\tilde{\boldsymbol{\theta}}_{t,t}^{T,\mathbf{y}}(\mathbf{x}) = \mathbf{x}$ we rewrite:

$$(2.6) \qquad \tilde{\mathbf{X}}_v^{T,\mathbf{y}} = \tilde{\boldsymbol{\theta}}_{v,t}^{T,\mathbf{y}}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{T,\mathbf{y}}(v, s)B\sigma(s, \boldsymbol{\theta}_{s,T}(\mathbf{y}))dW_s, \ v \in [t, T].$$

An important correspondence is now given by the following Proposition.

Proposition 3 (Density of the linearized dynamics)**.** Under **(A)**, we have that, for all $v \in (t, T]$ the random variable $\tilde{\mathbf{X}}_v^{T,\mathbf{y}}$ in (2.6) admits a Gaussian density $\tilde{p}^{T,\mathbf{y}}(t, v, \mathbf{x}, \cdot)$ which writes:

$$(2.7) \qquad\qquad\qquad\qquad \forall \mathbf{z} \in \mathbb{R}^{nd}, \ \tilde{p}^{T,\mathbf{y}}(t, v, \mathbf{x}, \mathbf{z})$$

$$:= \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{v,t}^{T,\mathbf{y}})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{v,t}^{T,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{v,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{z}), \tilde{\boldsymbol{\theta}}_{v,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{z} \right\rangle \right),$$

where

$$\tilde{\mathbf{K}}_{v,t}^{T,\mathbf{y}} := \int_t^v \tilde{\mathbf{R}}^{T,\mathbf{y}}(v, s) B a(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})) B^* \tilde{\mathbf{R}}^{T,\mathbf{y}}(v, s)^* ds,$$

$$a(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})) := \sigma\sigma^*(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})).$$

Also, there exists $C := C((\mathbf{A}), T) > 0$ s.t. for all $k \in [\![0, 2]\!]$, $i \in [\![1, n]\!]$,

$$|D_{x_i}^k \tilde{p}^{T,\mathbf{y}}(t, T, \mathbf{x}, \mathbf{y})|$$

$$\leq \frac{C}{(T-t)^{k\left((i-1)+\frac{1}{2}\right)+\frac{n^2 d}{2}}} \exp\left(-C^{-1}(T-t) \left| \mathbb{T}_{T-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y})) \right|^2 \right)$$

$$(2.8) \quad =: \ \frac{C}{(T-t)^{k\left((i-1)+\frac{1}{2}\right)}} \bar{p}_{C^{-1}}(t, T, \mathbf{x}, \mathbf{y}),$$

where for all $u > 0$, we denote by $\mathbb{T}_u$ the important scale matrix:

$$(2.9) \qquad\qquad \mathbb{T}_u = \begin{pmatrix} u I_{d \times d} & 0_{d \times d} & \cdots & 0_{d \times d} \\ 0_{d \times d} & u^2 I_{d \times d} & 0 & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0_{d \times d} & \cdots & 0 & u^n I_{d \times d} \end{pmatrix}.$$

Proof. Expression (2.7) readily follows from (2.6). We recall as well that, under **(A)**, the covariance matrix $\tilde{\mathbf{K}}_{v,t}^{T,\mathbf{y}}$ enjoys, uniformly in $\mathbf{y} \in \mathbb{R}^{nd}$ a *good scaling property* in the sense of Definition 3.2 in [DM10] (see also Proposition 3.4 of that reference). That is: for all fixed $T > 0$, there exists $C_{2.10} := C_{2.10}((\mathbf{A}), T) \geq 1$ s.t. for all $0 \leq v < t \leq T$, for all $\mathbf{y} \in \mathbb{R}^{nd}$:

$$(2.10) \quad \forall \boldsymbol{\xi} \in \mathbb{R}^{nd}, \ C_{2.10}^{-1}(v-t)^{-1} |\mathbb{T}_{v-t} \boldsymbol{\xi}|^2 \leq \langle \tilde{\mathbf{K}}_{v,t}^{T,\mathbf{y}} \boldsymbol{\xi}, \boldsymbol{\xi} \rangle \leq C_{2.10}(v-t)^{-1} |\mathbb{T}_{v-t} \boldsymbol{\xi}|^2.$$

Remark 3 (On the Hörmander like non-degeneracy assumption and the good scaling property)*.* We carefully point out that the boundedness and non-degeneracy conditions expressed on the derivatives $(D_{x_{i-1}} F_i(t, \cdot))_{i \in [\![2,n]\!]}$, as well as on the diffusion coefficient $\sigma(t, \cdot)$, in assumptions **(S)**, **(H)** and **(UE)** are precisely explicitly used to derive this bound. The natural analogue bounds of (2.10) also hold for the inverse matrices, see again [DM10], Definition 3.2 and Lemma 3.6 therein as well as the proof of (2.12) below.

Rewrite now from (2.5) and (2.6):

$$\left\langle (\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y}), \tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y} \right\rangle$$

$$(2.11) \qquad = \left\langle (\tilde{\mathbf{R}}^{T,\mathbf{y}}(T, t)^* (\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{-1} \tilde{\mathbf{R}}^{T,\mathbf{y}}(T, t)(\mathbf{x} - \tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y})), \mathbf{x} - \tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) \right\rangle,$$

where, accordingly with the previous notations for the forward linearized flow $\tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x})$, we denote:

$$\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) = \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,T)\mathbf{y} - \int_t^T \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,s)\Big(\mathbf{F}(s,\boldsymbol{\theta}_{s,T}(\mathbf{y})) - D\mathbf{F}(s,\boldsymbol{\theta}_{s,T}(\mathbf{y}))\boldsymbol{\theta}_{s,T}(\mathbf{y})\Big)ds.$$

for the corresponding linearized-backward flow starting from $\mathbf{y}$ at time T.

Observe now that $\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) = \boldsymbol{\theta}_{t,T}(\mathbf{y})$. Indeed, from (2.3):

$$\partial_t\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) = D\mathbf{F}(t,\boldsymbol{\theta}_{t,T}(\mathbf{y}))\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) + \Big(\mathbf{F}(t,\boldsymbol{\theta}_{t,T}(\mathbf{y})) - D\mathbf{F}(t,\boldsymbol{\theta}_{t,T}(\mathbf{y}))\boldsymbol{\theta}_{t,T}(\mathbf{y})\Big),$$

so that:

$$\partial_t\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) - \partial_t\boldsymbol{\theta}_{t,T}(\mathbf{y}) = D\mathbf{F}(t,\boldsymbol{\theta}_{t,T}(\mathbf{y}))\Big(\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) - \boldsymbol{\theta}_{t,T}(\mathbf{y})\Big).$$

Since $\tilde{\boldsymbol{\theta}}_{T,T}^{T,\mathbf{y}}(\mathbf{y}) = \boldsymbol{\theta}_{T,T}(\mathbf{y}) = \mathbf{y}$, we deduce from Gronwall's Lemma that $\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) = \boldsymbol{\theta}_{t,T}(\mathbf{y})$ for all $t \in [0,T]$.

We carefully point out that, even though the solution to the ODE (2.1) is not unique, once we have chosen a solution and consider the associated flow to construct our linearized Gaussian model, we precisely get the identification $\tilde{\boldsymbol{\theta}}_{t,T}^{T,\mathbf{y}}(\mathbf{y}) = \boldsymbol{\theta}_{t,T}(\mathbf{y})$ for all $t \in [0,T]$ with the **chosen** flow.

We thus get from the previous identification, equations (2.11), (2.10) and Remark 3 that there exists $C := C((\mathbf{A}),T) > 0$, s.t. for all $t \in [0,T)$,

$$(2.12) \quad C^{-1}(T-t)|\mathbb{T}_{T-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y}))|^2 \leq \Big\langle (\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y}), \tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y} \Big\rangle$$

$$\leq C(T-t)|\mathbb{T}_{T-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y}))|^2.$$

Indeed, from (2.10) it is easily derived that the spectrum of $\widehat{\tilde{\mathbf{K}}}_1^{T,t,T,\mathbf{y}} := (T-t)\mathbb{T}_{T-t}^{-1}\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}}\mathbb{T}_{T-t}^{-1}$ lies in $[C_{2.10}^{-1}, C_{2.10}]$. So does the spectrum of $(\widehat{\tilde{\mathbf{K}}}_1^{T,t,T,\mathbf{y}})^{-1} := (T-t)^{-1}\mathbb{T}_{T-t}(\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{-1}\mathbb{T}_{T-t}$. From Lemma 3.6 in [DM10] (see also Lemma 6.2 in [Men18] for notations closer to the current framework) we can also write $\tilde{\mathbf{R}}^{T,\mathbf{y}}(T,t) = \mathbb{T}_{T-t}\widehat{\tilde{\mathbf{R}}}^{T,t,T,\mathbf{y}}(1,0)\mathbb{T}_{T-t}^{-1}$ where $\widehat{\tilde{\mathbf{R}}}^{T,t,T,\mathbf{y}}(1,0)$ is a non-degenerate bounded matrix whose bounds do not depend on t, T, i.e. $\widehat{\tilde{\mathbf{R}}}^{T,t,T,\mathbf{y}}(1,0)$ is a non-degenerate *macro* matrix. Therefore, introducing $\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}} = \tilde{\mathbf{R}}^{T,\mathbf{y}}(t,T)\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}}\tilde{\mathbf{R}}^{T,\mathbf{y}}(t,T)^*$, we can rewrite from (2.11) that:

$$\Big\langle (\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y}), \tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y} \Big\rangle$$

$$= \Big\langle (\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1}(\mathbf{x} - \boldsymbol{\theta}_{T,t}(\mathbf{y})), \mathbf{x} - \boldsymbol{\theta}_{T,t}(\mathbf{y}) \Big\rangle,$$

$$(\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1} = \tilde{\mathbf{R}}^{T,\mathbf{y}}(T,t)^*(\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{-1}\tilde{\mathbf{R}}^{T,\mathbf{y}}(T,t)$$

$$(2.13) \qquad = (T-t)\mathbb{T}_{T-t}^{-1}\big(\widehat{\tilde{\mathbf{R}}}^{T,t,T,\mathbf{y}}(1,0)\big)^*(\widehat{\tilde{\mathbf{K}}}_1^{T,t,T,\mathbf{y}})^{-1}\widehat{\tilde{\mathbf{R}}}^{T,t,T,\mathbf{y}}(1,0)\mathbb{T}_{T-t}^{-1}.$$

The previous non-degeneracy properties of $(\widehat{\tilde{\mathbf{K}}}_1^{T,t,T,\mathbf{y}})^{-1}, \widehat{\tilde{\mathbf{R}}}^{T,t,T,\mathbf{y}}(1,0)$ then readily give (2.12). Put it differently, the matrix $\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}}$ also satisfies a *good scaling property*.

We now deduce from (2.13) and (2.10) that (2.8) holds for $k = 0$. We now write from (2.7):

$$D_{x_i}\tilde{p}^{T,\mathbf{y}}(t, T, \mathbf{x}, \mathbf{y})$$

$$=\frac{1}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{\frac{1}{2}}}D_{x_i}\exp\left(-\frac{1}{2}\left\langle(\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1}(\mathbf{x}-\boldsymbol{\theta}_{t,T}(\mathbf{y})), \mathbf{x}-\boldsymbol{\theta}_{t,T}(\mathbf{y})\right\rangle\right)$$

$$=\frac{\left(-(\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1}(\mathbf{x}-\boldsymbol{\theta}_{t,T}(\mathbf{y}))\right)_i}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left\langle(\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1}(\mathbf{x}-\boldsymbol{\theta}_{t,T}(\mathbf{y})), \mathbf{x}-\boldsymbol{\theta}_{t,T}(\mathbf{y})\right\rangle\right).$$

We thus derive from (2.12) and (2.13):

$$|D_{x_i}\tilde{p}^{T,\mathbf{y}}(t, T, \mathbf{x}, \mathbf{y})| \le \frac{C}{(T-t)^{(i-1)+\frac{1}{2}+\frac{n^2d}{2}}}\exp\left(-C^{-1}(T-t)\left|\mathbb{T}_{T-t}^{-1}\left(\mathbf{x}-\boldsymbol{\theta}_{t,T}(\mathbf{y})\right)\right|^2\right),$$

which proves (2.8) for $k = 1$. The case $k = 2$ is derived similarly. $\qquad\square$

2.2. Regularity and measurability of the Cauchy-Peano flow. Let us recall, as indicated before equation (2.1), that we work first under assumption **(D)**-(a). In this setting, we mention that the delicate part here consists in dealing with the nonlinear flow $\boldsymbol{\theta}_{t,s}(\mathbf{y})$. Because of our low Hölder regularity, we face two problems: one has to choose a measurable flow of (2.1) (which is very important to make licit any integration of this flow along the terminal condition) and this flow must have the appropriate regularity to deal with our parametrix kernel, say e.g. bi-Lipschitz as in [DM10], [Men11], [Men18].

The first issue is addressed by Lemma 4 below. The second problem is quite involved and requires also a careful analysis. Indeed our approach, based on parametrix kernel, makes an intensive use of the gradient estimate of the frozen transition density $\tilde{p}^{T,\mathbf{y}}$ given in (2.7). This leads us to study the space integral of the Gaussian like function $\bar{p}_{C^{-1}}$ defined by (2.8) w.r.t. the backward variable $\mathbf{y}$. The crucial point is that such an integral then involves the backward flow with argument the integration variable. In a smooth setting, such a problem is easily handled through a change of variable. When working with non-continuously differentiable coefficients, one may also use the bi-Lipschitz property of the flow to change its argument from the integration variable to the fixed initial one (see e.g. [DM10] where $\mathbf{F}$ is Lipschitz in space). In the current setting the flow is not smooth enough either to perform a change of variable nor to use the bi-Lipschitz estimate. Nevertheless, using a careful regularization procedure which precisely works under the condition (1.3) on the Hölder continuity exponents, we show in Lemma 5 below that the chosen flow satisfies an *approximate* bi-Lipschitz estimate. This *approximate* bi-Lipschitz estimate is sufficient to deal with our parametrix kernel.

Lemma 4. *For a given $T > 0$, there exists a measurable mapping $(s, t, \mathbf{x}) \in [0, T]^2 \times \mathbb{R}^{nd} \mapsto \boldsymbol{\theta}_{t,s}(\mathbf{x})$ s.t. $\boldsymbol{\theta}_{t,s}(\mathbf{x}) = \mathbf{x} + \int_s^t \mathbf{F}(v, \boldsymbol{\theta}_{v,s}(\mathbf{x}))dv$.*

Proof. The proof follows from the result of [Zub12] and usual covering arguments. $\qquad\square$

From now on, we choose by simplicity to work with a given measurable flow $\boldsymbol{\theta}_{t,s}(\mathbf{x})$ which exists by the previous lemma.

Lemma 5. *There exist constants* $(C_5, C_5') := (C_5, C_5')((\mathbf{A}), T) \geq 1$ *s.t. for all* $0 \leq t < s \leq T$ *small enough:*

$$C_5^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 - C_5' \leq (s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2$$

$$(2.14) \hspace{4cm} \leq C_5(s-t)|\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 + C_5'.$$

Also, for any measurable flow $\check{\boldsymbol{\theta}}_{s,t}$ *satisfying the integral equation in Lemma 4 and possibly different from the chosen one* $\boldsymbol{\theta}_{s,t}$*, it also holds that:*

$$C_5^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\check{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 - C_5' \leq (s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2$$

$$(2.15) \hspace{4cm} \leq C_5(s-t)|\mathbb{T}_{s-t}^{-1}(\check{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 + C_5'.$$

Lemma 5 is a key tool for our analysis. It roughly says that, even though the drift coefficient is not smooth, we can still expect a kind of equivalence of the rescaled forward and backward flows (which has been thoroughly used in the papers [DM10], [Men11], [Men18] for Lipschitz drifts) up to an additional constant contribution. This is precisely the result of equation (2.14). Also, since uniqueness here possibly fails for the flows, equation (2.15) gives that the bound still holds for two arbitrary measurable flows. This specific property will be used later on in the proof of Lemma 6 below in Appendix A.

It turns out that, the new contribution in (2.14), (2.15) does not perturb the analysis of the parametrix kernels associated with the density of $\tilde{\mathbf{X}}_T^{T,\mathbf{y}}$ starting from $\mathbf{x}$ at time $t \in [0, T)$ given in Proposition 3. We refer to Section 3.1 for details.

Proof. We focus on the proof of (2.15). Indeed, (2.14) is derived as a special case taking the same flows, i.e. $\check{\boldsymbol{\theta}} = \boldsymbol{\theta}$. Considering now two measurable flows $\boldsymbol{\theta}, \check{\boldsymbol{\theta}}$ provided by the integral equation in Lemma 4, we write from the integral dynamics:

$$(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))$$

$$= (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}\Big[(\check{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) - \mathbf{y}) - \int_t^s \Big(\mathbf{F}(u, \check{\boldsymbol{\theta}}_{u,t}(\mathbf{x})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y}))\Big) du\Big].$$

$$(2.16) \hspace{1cm} = (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\check{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) - \mathbf{y}) - \mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y}),$$

$$\mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y}) := (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}\int_t^s \Big(\mathbf{F}(u, \check{\boldsymbol{\theta}}_{u,t}(\mathbf{x})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y}))\Big) du.$$

We aim at establishing that

$$(2.17) \quad |\mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y})| \leq C\Big\{1 + (s-t)^{-1}\int_t^s (s-t)^{\frac{1}{2}}|\mathbb{T}_{s-t}^{-1}(\check{\boldsymbol{\theta}}_{u,t}(\mathbf{x}) - \boldsymbol{\theta}_{u,s}(\mathbf{y}))|du\Big\},$$

which together with (2.16) and the Gronwall lemma gives the r.h.s. of (2.15). The l.h.s. could be derived similarly to the analysis we now perform.

Since the function $\mathbf{F}$ is not Lispchitz, we will thoroughly use, as crucial auxiliary tool, some appropriate mollified flows. We first denote by $\delta \in \mathbb{R}^n \otimes \mathbb{R}^n$, a matrix whose entry δ_{ij} is strictly positive for indexes $i \in [\![2, n]\!]$ and $j \geq i$. We then define for all $v \in [0, T]$, $\mathbf{z} \in \mathbb{R}^{nd}$, $i \in [\![2, n]\!]$,

$$(2.18) \quad F_i^\delta(v, \mathbf{z}^{i-1,n}) := F_i(v, \cdot) \star \rho_{\delta_{i,\cdot}}(\mathbf{z})$$

$$= \int_{\mathbb{R}^{d(n-(i-1))}} F_i(v, z_{i-1}, z_i - w_i, \cdots, z_n - w_n)\rho_{\delta_{i,\cdot}}(w)dw.$$

Here, for all $w = (w_i, \cdots, w_n) \in \mathbb{R}^{d(n-(i-1))}$,

$$\rho_{\delta_{i,.}}(w) := \frac{1}{\prod_{j=i}^{n} \delta_{ij}^{d}} \rho_i \left(\frac{w_i}{\delta_{ii}}, \frac{w_{i+1}}{\delta_{i(i+1)}}, \cdots, \frac{w_n}{\delta_{in}} \right),$$

where $\rho_i : \mathbb{R}^{d(n-(i-1))} \to \mathbb{R}_+$ is a standard mollifier, i.e. ρ_i is smooth, has compact support and $\int_{\mathbb{R}^{d(n-(i-1))}} \rho_i(\mathbf{z}) d\mathbf{z} = 1$. We denote

$$\mathbf{F}^{\delta}(v, \mathbf{z}) = (F_1(v, \mathbf{z}), F_2^{\delta}(v, \mathbf{z}), \cdots, F_n^{\delta}(v, \mathbf{z})).$$

Note carefully that, under the considered assumptions **(A)**, the $(F_i^{\delta})_{i \in [\![2,n]\!]}$ are thus Lipschitz continuous functions, with explosive Lipschitz constant w.r.t. the mollifying procedure in the variables $(x_i, \cdots, x_n)$. More precisely, standard arguments from approximation theory give that, under the current assumptions, the Lipschitz constant of F_i^{δ} w.r.t. its jth variable blows up at rate $\delta_{ij}^{-1+\beta_i^j}$.

Controls associated with the mollification procedure. The first key point is that the regularized drift $\mathbf{F}^{\delta}$ only appears in a time integral for our analysis (see equations (2.16) and (2.23)). So, the parameters δ_{ij} only have to satisfy that there exists $C := C((\mathbf{A}), T) > 0$ such that for all $\mathbf{z}$ in $\mathbb{R}^{nd}$, for all u in $[t, s]$:

$$(2.19) \qquad \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \Big(\mathbf{F}(u, \mathbf{z}) - \mathbf{F}^{\delta}(u, \mathbf{z}) \Big) \right| \leq C(s-t)^{-1}.$$

From the definition of our regularization procedure in (2.18) this means that δ_{ij} must be such that

$$(2.20) \qquad \sum_{i=2}^{n} (s-t)^{\frac{1}{2}-i} \sum_{j=i}^{n} \delta_{ij}^{\beta_i^j} \leq C(s-t)^{-1}.$$

Hence, one can choose $\delta_{ij} = (s-t)^{(i-\frac{3}{2})\frac{1}{\beta_i^j}}$ which yields $(s-t)^{\frac{1}{2}-i} \delta_{ij}^{\beta_i^j} = (s-t)^{-1}$. This choice of δ_{ij} will be kept for the rest of the proof.

The second key point relies on the fact that, for this choice of the regularization parameter, the rescaled drift $\mathbf{F}^{\delta}$ satisfies an *approximate* Lipschitz property whose Lipschitz constant, once the drift is integrated, does not yield any additional singularity. Namely, there exists a $C := C((\mathbf{A}), T)$ such that for all u in $[s, t]$, for all $\mathbf{z}, \mathbf{z}'$ in $\mathbb{R}^{nd}$

$$(2.21) \qquad \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \Big(\mathbf{F}^{\delta}(u, \mathbf{z}) - \mathbf{F}^{\delta}(u, \mathbf{z}') \Big) \right|$$

$$\leq C \left((s-t)^{-\frac{1}{2}} + (s-t)^{-1} | (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} (\mathbf{z} - \mathbf{z}') | \right).$$

Indeed, as already underlined at the end of the previous paragraph, the $(F_i^{\delta})_{i \in [\![2,n]\!]}$ are Lipschitz continuous functions (with potentially explosive Lipschitz constant in the variables $(x_i, \cdots, x_n)$ for F_i^{δ} because of the regularization procedure) and F_1 is $\beta_1^j > 0$ Hölder continuous in the jth variable for arbitrary $(\beta_1^j)_{j \in [\![1,n]\!]}$ in $(0, 1]$. The Young inequality then yields that there exists $C_{\beta_1^j} > 0$ s.t. for all

$\mathbf{x} \in \mathbb{R}^{nd}$, $|\mathbf{x}|^{\beta_1^j} \leq C_{\beta_1^j}(1 + |\mathbf{x}|)$. Hence,

$$\left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \Big(\mathbf{F}^\delta(v, \mathbf{z}) - \mathbf{F}^\delta(v, \mathbf{z}') \Big) \right|$$

$$\leq C \Bigg((s-t)^{-\frac{1}{2}}(1 + |(\mathbf{z} - \mathbf{z}')|)$$

$$+ \sum_{i=2}^{n} (s-t)^{\frac{1}{2}-i} \Big(|(\mathbf{z} - \mathbf{z}')_{i-1}| + \sum_{j=i}^{n} \frac{|(\mathbf{z} - \mathbf{z}')_j|}{\delta_{ij}^{1-\beta_i^j}} \Big) \Bigg)$$

$$\leq C \Bigg((s-t)^{-\frac{1}{2}} + |(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\mathbf{z} - \mathbf{z}')|$$

$$\times \Big(1 + (s-t)^{-1} + \sum_{i=2}^{n} \sum_{j=i}^{n} \frac{(s-t)^{j-i}}{\delta_{ij}^{1-\beta_i^j}} \Big) \Bigg).$$

Hence (2.21) follows from the fact that, from our previous choice of δ_{ij}, one gets

$$(2.22) \qquad \frac{(s-t)^{j-i}}{\delta_{ij}^{1-\beta_i^j}} = (s-t)^{(j-i)-(i-\frac{3}{2})\frac{1}{\beta_i^j}(1-\beta_i^j)} \leq C(s-t)^{-1},$$

since from the assumption (1.3) on the indexes of Hölder continuity:

$$\beta_i^j > \frac{2i-3}{2j-1} \iff (j-i) - (i-\frac{3}{2})(1-\beta_i^j)/\beta_i^j > -1.$$

Derivation of the final bound. We are now in position to bound the term $\mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y})$ defined in (2.16). Under **(A)** we have:

$$|\mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y})| \leq \int_t^s du |(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \big(\mathbf{F}(u, \check{\boldsymbol{\theta}}_{u,t}(\mathbf{x})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y})) \big)|$$

$$\leq \int_t^s du \Big| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \Big(\mathbf{F}(u, \check{\boldsymbol{\theta}}_{u,t}(\mathbf{x})) - \mathbf{F}^\delta(u, \check{\boldsymbol{\theta}}_{u,t}(\mathbf{x})) \Big) \Big|$$

$$+ \int_t^s du \Big| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \Big(\mathbf{F}^\delta(u, \check{\boldsymbol{\theta}}_{u,t}(\mathbf{x})) - \mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,s}(\mathbf{y})) \Big) \Big|$$

$$+ \int_t^s du \Big| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \Big(\mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,s}(\mathbf{y})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y})) \Big) \Big|$$

$$(2.23) \qquad =: \mathcal{I}_{s,t}^1(\mathbf{x}, \mathbf{y}) + \mathcal{I}_{s,t}^2(\mathbf{x}, \mathbf{y}) + \mathcal{I}_{s,t}^3(\mathbf{x}, \mathbf{y}).$$

Using (2.19), we obtain that there exists $C := C((\mathbf{A}), T)$ s.t. for all $0 \leq t < s \leq T$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{nd}$:

$$(2.24) \qquad |\mathcal{I}_{s,t}^1(\mathbf{x}, \mathbf{y})| + |\mathcal{I}_{s,t}^3(\mathbf{x}, \mathbf{y})| \leq C.$$

Finally, one can use (2.21) to derive that for all $0 \leq t < s \leq T$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{nd}$:

$$|\mathcal{I}_{s,t}^2(\mathbf{x}, \mathbf{y})| \leq C \Bigg[(s-t)^{\frac{1}{2}} + \int_t^s (s-t)^{-1} \big((s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\check{\boldsymbol{\theta}}_{u,t}(\mathbf{x}) - \boldsymbol{\theta}_{u,s}(\mathbf{y}))| \big) du \Bigg]$$

$$\leq C \Big\{ 1 + (s-t)^{-1} \int_t^s (s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\check{\boldsymbol{\theta}}_{u,t}(\mathbf{x}) - \boldsymbol{\theta}_{u,s}(\mathbf{y}))| du \Big\},$$

up to a modification of $C := C((\mathbf{A}), T)$ for the last inequality. From this last equation together with (2.24), we therefore derive (2.17). The proof is complete. $\square$

2.3. **Frozen Green kernels and associated PDEs.** In this paragraph we introduce useful tools for the analysis of the martingale problem. Namely, we consider suitable green kernels associated with the previously defined frozen process and establish the *Cauchy problemma* which it solves.

For all $f \in C_0^{1,2}([0,T] \times \mathbb{R}^{nd}, \mathbb{R})$, and $\varepsilon \geq 0$ meant to be small, we define the Green function:

$$(2.25) \quad \forall (t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}, \tilde{G}^\varepsilon f(t,\mathbf{x}) = \int_{(t+\varepsilon)\wedge T}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y}) f(s,\mathbf{y}).$$

We point out that the measurability of the flow in $(s,\mathbf{y})$ established in Lemma 4 precisely gives that $\tilde{p}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y})$ is a measurable function of these parameters and ensures that the above Green function is properly defined. Denote by $(\tilde{L}_t^{s,\mathbf{y}})_{t\in[0,s]}$ the generator of $(\tilde{\mathbf{X}}_t^{s,\mathbf{y}})_{t\in[0,s]}$, i.e. for all $\varphi \in C_0^\infty(\mathbb{R}^{nd}, \mathbb{R})$, $\mathbf{x} \in \mathbb{R}^{nd}$,

$$
\begin{aligned}
\tilde{L}_t^{s,\mathbf{y}}\varphi(\mathbf{x}) \quad := \quad & \langle \mathbf{F}(t,\boldsymbol{\theta}_{t,s}(\mathbf{y})) + D\mathbf{F}(t,\boldsymbol{\theta}_{t,s}(\mathbf{y}))(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y})), D_\mathbf{x}\varphi(\mathbf{x})\rangle \\
(2.26) \quad & + \frac{1}{2}\mathrm{tr}(\sigma\sigma^*(t,\boldsymbol{\theta}_{t,s}(\mathbf{y}))D_{x_1}^2\varphi(\mathbf{x})).
\end{aligned}
$$

One now easily checks that:

$$(2.27) \qquad \forall (t,\mathbf{x},\mathbf{z}) \in [0,s) \times (\mathbb{R}^{nd})^2, \left(\partial_t + \tilde{L}_t^{s,\mathbf{y}}\right)\tilde{p}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{z}) = 0.$$

However, we carefully mention that some care is needed to establish the following lemma, whose proof is postponed to Appendix A, which is crucial to derive that $\tilde{G}f := \tilde{G}^0 f$ actually solves an appropriate Cauchy *like* problem.

Lemma 6 (Dirac convergence of the frozen density). *For all bounded continuous function $f : \mathbb{R}^{nd} \to \mathbb{R}, \mathbf{x} \in \mathbb{R}^{nd}$, setting for all (ε,t) in $(\mathbb{R}_+\backslash\{0\}) \times [0,T]$, $f_{\varepsilon,t}(\mathbf{x}) := \int_{\mathbb{R}^{nd}} f(\mathbf{y})\tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}$, we have:*

$$(2.28) \qquad\qquad \sup_{t\in[0,T]} |f_{\varepsilon,t}(\mathbf{x}) - f(\mathbf{x})| \xrightarrow[\varepsilon\downarrow 0]{} 0.$$

We emphasize that the above lemma is not a direct consequence of the convergence of the law of the frozen process towards the Dirac mass (see e.g. (2.27)). Indeed, the integration parameter is also the freezing parameter which makes things more subtle.

We also have the following related result again proved in Appendix A.

Lemma 7 ($L^{q'} - L^{p'}$ convergence for the mollification with the frozen density). *For all $f \in L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))$, $q',p' < \infty$, setting, for any $\varepsilon > 0$, $f_\varepsilon(t,\mathbf{x}) := \int_{\mathbb{R}^{nd}} f(t+\varepsilon,\mathbf{y})\mathbb{I}_{t\in[0,T-\varepsilon]}\tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}$, we have for all $p',q' > 1$:*

$$(2.29) \qquad\qquad \|f_\varepsilon - f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \xrightarrow[\varepsilon\downarrow 0]{} 0.$$

Introducing for all $f \in C_0^{1,2}([0,T] \times \mathbb{R}^{nd}, \mathbb{R})$, $\varepsilon \geq 0$ and $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$ the quantity:

$$(2.30) \qquad \tilde{M}_t^\varepsilon f(t,\mathbf{x}) = \int_{(t+\varepsilon)\wedge T}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{L}_t^{s,\mathbf{y}} \tilde{p}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y}) f(s,\mathbf{y}),$$

we derive from (2.27) and Proposition 3 that the following equality holds for all $\varepsilon > 0$ small enough:

$$(2.31) \qquad \partial_t \tilde{G}^\varepsilon f(t,\mathbf{x}) + \tilde{M}_t^\varepsilon f(t,\mathbf{x}) = -f_\varepsilon(t,\mathbf{x}), \ \forall (t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd},$$

denoting by $f_\varepsilon(t, \mathbf{x}) = \int_{\mathbb{R}^{nd}} f(t + \varepsilon, \mathbf{y}) \mathbb{I}_{t \in [0, T-\varepsilon]} \tilde{p}^{t+\varepsilon, \mathbf{y}}(t, t + \varepsilon, \mathbf{x}, \mathbf{y}) d\mathbf{y}$, i.e. the time argument of f is also shifted and truncated above $T - \varepsilon$. Observe here that the localization w.r.t. ε is precisely needed to exploit directly (2.27) and derive (2.31), for a given fixed $\varepsilon > 0$, by usual domination arguments. We mention that when $\varepsilon = 0$, it follows from the definition (2.25) of our Green kernel that the smoothness on f is not a sufficient condition to derive the smoothness of $\tilde{G}f = \tilde{G}^0 f$. This comes from the dependence of the covariance matrix in $\tilde{p}^{s, \mathbf{y}}$ w.r.t. the integration variable (see (2.7)).

Proposition 8. Pointwise control of the Green Kernel. There exists $C(T) := C((\mathbf{A}), T) \underset{T \to 0}{\longrightarrow} 0$ s.t. for all $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ and all $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))$ s.t. $\frac{n^2 d}{p'} + \frac{2}{q'} < 2, p' > 1, q' > 1, \varepsilon \in [0, T - t]$:

$$(2.32) \qquad |\tilde{G}^\varepsilon f(t, \mathbf{x})| \le C(T) \|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}.$$

Proof. From (2.8) with $k = 0$ and Lemma 5 we have that for all $\varepsilon \in [0, T - t]$,

$$(2.33) \qquad |\tilde{G}^\varepsilon f(t, \mathbf{x})|$$

$$\le C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y})$$

$$\le C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \frac{\exp\left(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2\right)}{(s-t)^{\frac{n^2 d}{2}}}$$

$$\le C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \frac{\exp\left(-C^{-1}\left[(s-t)|\mathbb{T}_{s-t}^{-1} \boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y}|^2 + 1\right]\right)}{(s-t)^{\frac{n^2 d}{2}}},$$

up to a modification of C. So, the result follows from the Hölder inequality and the condition on the exponents p' and q'. Denoting by $\tilde{p}'$, $\tilde{q}'$ the conjugate of p' and q' respectively we indeed have

$$\int_t^T ds \left| \int_{\mathbb{R}^{nd}} d\mathbf{y} \left| \frac{\exp\left(-C^{-1}(s-t)(|\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})|^2)\right)}{(s-t)^{\frac{n^2 d}{2}}} \right|^{\tilde{p}'} \right|^{\frac{\tilde{q}'}{\tilde{p}'}}$$

$$\le C \int_t^T ds \frac{1}{(s-t)^{\frac{n^2 d}{2}(\tilde{p}'-1)(\frac{\tilde{q}'}{\tilde{p}'})}} < +\infty \Leftrightarrow \frac{n^2 d(\tilde{p}' - 1)}{2} \frac{\tilde{q}'}{\tilde{p}'} < 1 \Leftrightarrow \frac{n^2 d}{2p'} + \frac{1}{q'} < 1.$$

$\square$

3. Well posedness of the corresponding martingale problem

We have now given the main tools needed to prove our main results: the well posedness of the martingale problem associated with $(L_t)_{t \ge 0}$ defined in (1.4) and the corresponding Krylov-type estimates. This section is organized as follows: we first investigate the well posedness under Assumption **(D)**-(a). In that case, the existence part is not a challenge, since it readily follows from previously known results based on compactness arguments that exploit the sublinear structure of the drift **F**, while the uniqueness part is quite more delicate. As a by-product of our approach to uniqueness we derive the Krylov like estimate.

The scheme used for proving uniqueness under **(D)**-(a) will be a major tool to extend our result for uniqueness under **(D)**-(b) (the existence part under that

assumption being a trivial application of the Girsanov Theorem) and then for existence and uniqueness under **(D)**-(c). Indeed, under this last assumption, even the existence part requires to derive first some Krylov type estimates. We will precisely exploit those established under **(A)**-(a) considering first $F_1 = 0$ and then cope with the true $L^q - L^p$ drift through a Girsanov argument. The approach is in some sense similar to the one of Krylov and Röckner [KR05] or Fedrizzi *et al.* [FFPV17] for the Girsanov part. The main difference is that in the quoted work the required Krylov like estimate readily followed from the explicit density of the unperturbed process at hand. The Brownian motion in [KR05], the joint density of the Brownian motion and its integral in [FFPV17]. We here precisely show that, first under **(A)**-(a), the solution to the martingale problem has a density which satisfies a similar Krylov type estimate. We actually prove that *any* solution to the martingale problem satisfies such an estimate (see equation (3.5) below).

It is precisely to deal with $L^q - L^p$ drifts (under Assumption **(D)**-(c)) that we have chosen an approach inspired by the Stroock and Varadhan original arguments which explicitly provides the required Krylov like estimates. Before going into the proof, let us briefly explain the main differences between our analysis and the strategy of [SV79]. In particular, our approach differs from the original one because of the specific structure of our problem.

In the original non degenerate setting with bounded drifts considered by Stroock and Varadhan, the Girsanov Theorem allows them to deal with the diffusive part of the equation only. Their main idea to obtain the desired control on their perturbed kernel goes through regularization arguments. The key point allowing them to get the estimation at the limit are: the strong convergence of the driftless Euler scheme (to keep track on pointwise estimate) and a localization argument.

In our current setting things are a bit different: we are not allowed anymore to get rid of the drift, because of our degenerate structure.

Thus, our strategy is the following. In the Hölder framework of case **(D)**-(a), we manage to prove directly the existence of the density of the canonical process for any solution to the martingale problem through the associated Krylov-type estimate, for $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))$ and p', q' *large enough*. This is a consequence of the pointwise controls established in Lemma 9 below. This is enough to derive the well posedness of the martingale problem. In a second time, we complete the proof of the Krylov estimate (1.5) on the indicated range for p', q' through a regularization argument. Namely, we regularize the drift coefficient **F** through convolution. For the regularized drift, it follows from [DM10], [Men11] that the corresponding process has a density. It then follows from the previous analysis that the process with mollified coefficients satisfies uniformly w.r.t. the mollification parameter the Krylov estimate. The final statement then follows letting the mollification parameter tend to zero from the well posedness of the martingale problem. Cases (b) and (c) are handled from case (a) through an additional Girsanov type argument.

3.1. **Well posedness with full Hölder drift, Assumption (D)-(a).**

3.1.1. *Existence under Assumption* **(D)**-(a). The first step is to establish that there exists a solution to the martingale problem defined in Theorem 1. From the definition of $(L_t)_{t \geq 0}$ in (1.4) it is easily seen that, under **(A)**, existence is obtained adapting to our current framework Theorem 6.1.7 in [SV79]. The strategy is clear.

An Euler like scheme can be considered. Fix first $T > 0$ and consider the grid $\Lambda_m([0, T]) := \{(t_i := ih)_{i \in [\![0, m]\!]}\}, h = T/m, m \in \mathbb{N}$. Introduce the corresponding "discretization" scheme:

$$(3.1) \qquad \mathbf{X}_s^m = \mathbf{X}_{t_i}^m + \int_{t_i}^s \Big(\mathbf{F}(t_i, \mathbf{X}_{t_i}^m) + D\mathbf{F}(t_i, \mathbf{X}_{t_i}^m)(\mathbf{X}_u^m - \mathbf{X}_{t_i}^m) \Big) du$$
$$+ B\sigma(t_i, \mathbf{X}_{t_i}^m)(W_s - W_{t_i}), \ s \in [t_i, t_{i+1}].$$

This scheme defines a sequence of measures $(\mathbb{P}^m)_{m \geq 1}$ on $C([0, T], \mathbb{R}^{nd})$ which is tight and for which the continuity assumption on the coefficients and the sub-linearity of the drift allow to identify that any limit $\mathbb{P}$ solves the martingale associated with $(L_t)_{t \geq 0}$ on $[0, T]$. We refer to Section 6.1 of [SV79] for details. To derive the existence of a solution to the martingale problem on the whole positive line, we can rely on a usual chaining in time argument, see e.g. Chapter 6 in [SV79].

3.1.2. *Uniqueness under Assumption* **(D)**-(a). To establish uniqueness and the Krylov type control of Theorem 1, the key ingredient is to prove that an operator involving L and a suitable associated perturbation (based on the frozen process/generators of Section 2) satisfy appropriate estimates. Namely, for $\varepsilon \geq 0$ introduce:

$$(3.2) \quad R^\varepsilon f(t, \mathbf{x}) \quad := \quad (L_t \tilde{G}^\varepsilon f - \tilde{M}_t^\varepsilon f)(t, \mathbf{x})$$
$$= \quad \int_{(t+\varepsilon) \wedge T}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_t - \tilde{L}_t^{s, \mathbf{y}}) \tilde{p}^{iss, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}),$$

with $\tilde{G}^\varepsilon f, \tilde{M}^\varepsilon f$ and $\tilde{p}^{s, \mathbf{y}}$ defined in (2.25), (2.30) and (2.7) respectively. From now on, we also set $Rf(t, \mathbf{x}) = R^0 f(t, \mathbf{x})$. The aforementioned estimates are summarized in the following general Lemma whose proof is postponed to the end of the current section.

Lemma 9 (Pointwise and $L^{q'} - L^{p'}$ Control for R^ε).
There exists $q_0' := q_0'((\mathbf{A})), p_0' := p_0'((\mathbf{A}))$ s.t. for $q' \geq q_0', p \geq p_0'$, it holds that for all $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ and R^ε as in (3.2):

$$(3.3) \qquad\qquad |R^\varepsilon f(t, \mathbf{x})| \leq C \|f\|_{L^{q'}([0, T] \times \mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))},$$

with $C := C((\mathbf{A}), T, p_0', q_0')$.

Also, for $q', p' > 1$ s.t. $\frac{n^2 d}{p'} + \frac{2}{q'} < 2$, we have that for all $f \in L^{q'}([0, T] \times \mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))$, uniformly in $\varepsilon \in [0, \varepsilon_0], \varepsilon_0$ small enough:

$$(3.4) \qquad \|R^\varepsilon f\|_{L^{q'}([0, T] \times \mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))} \leq C \|f\|_{L^{q'}([0, T] \times \mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))},$$

with $C := C((\mathbf{A}), T, p', q') \xrightarrow[T \to 0]{} 0$. In particular, equation (3.4) implies that the operator $I - R^\varepsilon$ is invertible, with bounded inverse in $L^{q'} - L^{p'}$, provided T is small enough.

Having this result at hand, we are now in position to derive uniqueness of the martingale problem.

The first step of our approach consists in showing that any solution to the martingale problem satisfies the Krylov like density estimate of Theorem 1 for p' and

q' therein *large enough* but finite. Indeed, the first condition on p', q' precisely allows us to take benefit from the pointwise control (3.3) in Lemma 9. Note that in this setting, thanks to an approximation argument, the Krylov type estimate (1.5) reduces to
(3.5)

$$\forall f \in C_0^\infty([0,T] \times \mathbb{R}^{nd}, \mathbb{R}), \quad \left| \mathbb{E}^{\mathbb{P}}\left[\int_t^T f(s, \mathbf{X}_s^{t,\mathbf{x}})ds \right] \right| \le C(T)\|f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))}.$$

To establish the above control, we apply the Itô formula on the Green kernel $\tilde{G}^\varepsilon f$, for f in $C_0^\infty([0,T] \times \mathbb{R}^{nd}, \mathbb{R})$, for ε small enough (depending on the support of f) and the process $\mathbf{X}_s^{t,\mathbf{x}}$:

$$\tilde{G}^\varepsilon f(t,\mathbf{x}) + \mathbb{E}\left[\int_t^T (\partial_s + L_s)\tilde{G}^\varepsilon f(s, \mathbf{X}_s^{t,\mathbf{x}})ds \right]$$

$$= \tilde{G}^\varepsilon f(t,\mathbf{x}) + \int_t^T \int_{\mathbb{R}^{nd}} (\partial_s + L_s)\tilde{G}^\varepsilon f(s,\mathbf{y})P_{\mathbf{X}_s^{t,\mathbf{x}}}(d\mathbf{y})ds = 0,$$

where $P_{\mathbf{X}_s^{t,\mathbf{x}}}$ denotes the law of $\mathbf{X}_s^{t,\mathbf{x}}$. We exploit (2.31) to write:

$$(3.6) \qquad \tilde{G}^\varepsilon f(t,\mathbf{x}) - \int_t^T \int_{\mathbb{R}^{nd}} f_\varepsilon(s,\mathbf{y})P_{\mathbf{X}_s^{t,\mathbf{x}}}(d\mathbf{y})ds$$

$$+ \int_t^T \int_{\mathbb{R}^{nd}} (L_s\tilde{G}^\varepsilon f - \tilde{M}_s^\varepsilon f)(s,\mathbf{y})P_{\mathbf{X}_s^{t,\mathbf{x}}}(d\mathbf{y})ds = 0,$$

where f_ε is defined in Lemma 7. Recall that thanks to Proposition 8 we have that there exists $C(T) := C((\mathbf{A}), T) \xrightarrow[T \to 0]{} 0$ s.t. for all $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$ and all $f \in L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))$:

$$(3.7) \qquad\qquad |\tilde{G}^\varepsilon f(t,\mathbf{x})| \le C(T)\|f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))}.$$

As we assume that p', q' are *large enough*, the pointwise control (3.3) of Lemma 9 holds. From (3.6), (3.7) and (3.3) we hence readily get:

$$(3.8) \qquad\qquad |\int_t^T \mathbb{E}[f_\varepsilon(\mathbf{X}_s^{t,\mathbf{x}})]ds| \le C\|f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))}.$$

Letting ε go to zero, we thus derive from Lemma 6 that, for any solution $\mathbb{P}$ of the martingale problem, provided p', q' are large enough, inequality (3.5) holds.

We are now in position to prove uniqueness to the martingale problem. Let $\mathbb{P}$ be a solution and denote, thanks to the Krylov type estimate established above, its

density by $p(t, s, \mathbf{x}, \mathbf{y})$. We have from (3.6) that: for f in $C_0^\infty([0, T] \times \mathbb{R}^{nd}, \mathbb{R})$

$$-\tilde{G}^\varepsilon f(t, \mathbf{x})$$
$$= -\mathbb{E}^{\mathbb{P}}\left[\int_t^T f_\varepsilon(s, \mathbf{X}_s^{t, \mathbf{x}}) ds\right] + \mathbb{E}^{\mathbb{P}}\left[\int_t^T R^\varepsilon f(s, \mathbf{X}_s^{t, \mathbf{x}}) ds\right]$$
$$= -\int_t^T \int_{\mathbb{R}^{nd}} f_\varepsilon(s, \mathbf{y}) p(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds + \int_t^T \int_{\mathbb{R}^{nd}} R^\varepsilon f(s, \mathbf{y}) p(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds$$
$$(3.9) = -\int_t^T \int_{\mathbb{R}^{nd}} (I^\varepsilon - R^\varepsilon) f(s, \mathbf{y}) p(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds,$$

denoting by $I^\varepsilon f(s, \mathbf{y}) := f_\varepsilon(s, \mathbf{y})$. It thus easily follows from Proposition 8, Lemmas 7 and 9 eq. (3.4) that both sides are continuous with respect to $L^{q'} - L^{p'}$ norm, uniformly in ε. In particular, (3.9) holds for all $f \in L^{q'} - L^{p'}$. From Lemmas 7 and 9 we can apply (3.9) to $(I^\varepsilon - R^\varepsilon)^{-1} f$ and deduce, still from these Lemmas, that

$$(3.10) \qquad \mathbb{E}^{\mathbb{P}}\left[\int_t^T f(s, \mathbf{X}_s^{t, \mathbf{x}}) ds\right] = \tilde{G} \circ (I - R)^{-1} f(t, \mathbf{x}),$$

by letting ε go to zero. This gives uniqueness for T small enough. Global well-posedness is derived from a chaining in time argument.

To conclude the proof of Theorem 1, it now only remains to derive the Krylov like density estimate (1.5) in full generality *i.e.* for any p', q' satisfying $\frac{n^2 d}{p'} + \frac{2}{q'} < 2$ (whence not assuming that they are *large enough* and finite).

Assume first that both p' and q' are not *large enough*. We now consider, for a parameter $\delta > 0$, the SDE (1.1) with drift $\mathbf{F}^\delta(t, \mathbf{x}) := (\mathbf{F}(t, \cdot) \star \Phi_\delta)(\mathbf{x})$, where $\Phi_\delta(\cdot) := \delta^{-nd} \Phi(\cdot/\delta)$ with $\Phi \in C_0^\infty(\mathbb{R}^{nd}, \mathbb{R}^+)$, $\int_{\mathbb{R}^{nd}} \Phi(\mathbf{z}) d\mathbf{z} = 1$. It is known that, denoting by $\mathbb{P}^\delta$ the associated solution to the martingale problem, the canonical process enjoys two-sided multi-scale Gaussian bounds similar to those of (2.8) (with $k = 0$), see indeed again [DM10], [Men11], with constants possibly depending on δ.

Denoting by $p^\delta(t, s, \mathbf{x}, \mathbf{y})$ its density, similarly to (3.6) we get for any f in $C_0^\infty([0, T] \times \mathbb{R}^{nd}, \mathbb{R})$,

$$(3.11) \qquad \tilde{G}^{\delta, \varepsilon} f(t, \mathbf{x}) - \int_t^T \int_{\mathbb{R}^{nd}} f_\varepsilon(s, \mathbf{y}) p^\delta(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds$$
$$+ \int_t^T \int_{\mathbb{R}^{nd}} (L_s^\delta \tilde{G}^\varepsilon f - \tilde{M}_s^{\delta, \varepsilon} f)(s, \mathbf{y}) p^\delta(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds = 0,$$

where in the above equation, $\tilde{G}^{\delta, \varepsilon} f, L_s^\delta, \tilde{M}_s^{\delta, \varepsilon}$ denote the frozen Green kernel, generator and frozen generator with mollified drift. Importantly, the pointwise bound (3.7) on the Green kernel and the controls of Lemma 9 are uniform with respect to this additional mollifying parameter.

Then, from (3.11) and Lemma 9, we deduce that

$$C(T) \|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \left(1 + \|p^\delta\|_{L^{\tilde{q}'}([0,T], L^{\tilde{p}'}(\mathbb{R}^{nd}))}\right)$$
$$\geq \left|\int_t^T \int_{\mathbb{R}^{nd}} f_\varepsilon(s, \mathbf{y}) p^\delta(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds\right|,$$

where $\tilde{q}', \tilde{p}'$ are the conjugate exponents of q', p' respectively.

Using now Lemma 7 we get, for any $f \in C_0^\infty([0,T] \times \mathbb{R}^{nd}, \mathbb{R})$

$$
\left| \int_t^T \int_{\mathbb{R}^{nd}} f(s, \mathbf{y}) p^\delta(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds \right|
$$

$$
\leq C(T) \|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \left(1 + \|p^\delta\|_{L^{\tilde{q}'}([0,T], L^{\tilde{p}'}(\mathbb{R}^{nd}))} \right)
$$

$$
+ \|f - f_\varepsilon\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \|p^\delta\|_{L^{\tilde{q}'}([0,T], L^{\tilde{p}'}(\mathbb{R}^{nd}))}
$$

$$
\leq C(T) \|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \left(1 + \|p^\delta\|_{L^{\tilde{q}'}([0,T], L^{\tilde{p}'}(\mathbb{R}^{nd}))} \right)
$$

$$
+ \frac{1}{2} \|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \|p^\delta\|_{L^{\tilde{q}'}([0,T], L^{\tilde{p}'}(\mathbb{R}^{nd}))}
$$

as soon as ε, possibly depending on f, is small enough. Up to an additional density argument, this yields in particular from the Riesz representation theorem that, for T small enough, $\|p^\delta\|_{L^{\tilde{q}'}([0,T], L^{\tilde{p}'}(\mathbb{R}^{nd}))} \leq C(T)$. Thus, for any f in $L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))$:

$$
\left| \int_t^T \int_{\mathbb{R}^{nd}} f(s, \mathbf{y}) p^\delta(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} ds \right| = \left| \int_t^T \mathbb{E}^{\mathbb{P}^\delta}[f(s, \mathbf{X}_s^{t, \mathbf{x}})] ds \right|
$$

$$
\leq C(T) \|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))}.
$$

Estimate (3.5) is then obtained letting δ go to zero from the well posedness of the martingale problem and Theorem 11.1.4 in [SV79].

Next, the case when $q' = p' = \infty$ is direct and the remaining cases are thus $q' < +\infty$, $p' = \infty$ and $q' = \infty$, $p' < +\infty$. The first one is again direct. The second follows from the fact that for any $q' \in (2 + \infty]$, $\|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \leq \|f\|_{L^\infty([0,T], L^{p'}(\mathbb{R}^{nd}))} T^{1/q'}$. $\qquad \square$

Proof of Lemma 9. We focus on the proof for $Rf = R^0 f$. The parameter ε does not play here any role for the estimates. We have, by definition

$$
Rf(t, \mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_t - \tilde{L}_t^{s, \mathbf{y}}) \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y})
$$

(3.12)

$$
=: \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} H(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}).
$$

Here, the operator H is the so-called backward *parametrix kernel*, see [MS67]. It already appeared, in a similar form but under stronger smoothness assumptions in [DM10, Men11].

The bound (2.8) of Proposition 3 now yields that there exists $C := C((\mathbf{A}))$ such that:

$$
|Rf(t, \mathbf{x})|
$$

$$
\leq \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \Bigg\{ \Big| F_1(t, \mathbf{x}) - F_1(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) \Big| |D_{x_1} \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y})|
$$

$$
+ \sum_{i=2}^n \Big\{ \Big| F_i(t, \mathbf{x}) - \Big(F_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) - D_{x_{i-1}} F_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y}))(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_{i-1} \Big) \Big|
$$

$$
\times |D_{x_i} \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y})| \Big\} + |a(t, \mathbf{x}) - a(t, \boldsymbol{\theta}_{t,s}(\mathbf{y}))| |D_{x_1}^2 \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y})| \Bigg\}
$$

$$
\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \Bigg\{ \frac{\left(1 + |\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y})| \right)}{(s-t)^{\frac{1}{2}}} + \sum_{i=2}^n \Big\{ \Big[\Big| F_i(t, \mathbf{x}) - F_i(t, x_{i-1}, \boldsymbol{\theta}_{t,s}(\mathbf{y})^{i:n}) \Big|
$$

$$
+ \Big| F_i(t, x_{i-1}, \boldsymbol{\theta}_{t,s}(\mathbf{y})^{i:n}) - \Big(F_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) - D_{x_{i-1}} F_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y}))(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_{i-1} \Big) \Big| \Big]
$$

$$
\times \frac{1}{(s-t)^{(i-1)+\frac{1}{2}}} \Big\} + \frac{|\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y})|^\eta}{s-t} \Bigg\} \frac{\exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2 \right)}{(s-t)^{\frac{n^2 d}{2}}},
$$

where we have denoted for $\mathbf{z} \in \mathbb{R}^{nd}, \mathbf{z}^{i:n} = (z_i, \cdots, z_n) \in \mathbb{R}^{(n-(i-1))d}$. From $(\mathbf{A})$ and (1.3) we thus derive, up to a modification of C:

$$
|Rf(t, \mathbf{x})| \leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \Bigg\{ \left((s-t)^{-\frac{1}{2}} + (s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))| \right)
$$

$$
+ \sum_{i=2}^n \Bigg\{ \Big[\sum_{j=i}^n \Big(\frac{|(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_j|}{(s-t)^{j-\frac{1}{2}}} \Big)^{\beta_i^j} (s-t)^{\beta_i^j (j-\frac{1}{2})}
$$

$$
+ \Big(\frac{|(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_{i-1}|}{(s-t)^{(i-1)-\frac{1}{2}}} \Big)^{1+\eta} (s-t)^{(i-\frac{3}{2})(1+\eta)} \Big] \frac{1}{(s-t)^{i-\frac{1}{2}}} \Bigg\}
$$

$$
+ \frac{\left((s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))| \right)^\eta}{(s-t)^{1-\frac{\eta}{2}}} \Bigg\} \frac{\exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2 \right)}{(s-t)^{\frac{n^2 d}{2}}}.
$$

In the above equation we put each contribution coming from the difference of the generators at its *intrinsic scale* w.r.t. exponential bounds. In other words, the terms:

$$
\Big(\frac{|(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_j|}{(s-t)^{j-\frac{1}{2}}} \Big)^{\beta_i^j} \leq \left((s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))| \right)^{\beta_i^j},
$$

$$
\Big(\frac{|(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_{i-1}|}{(s-t)^{(i-1)-\frac{1}{2}}} \Big)^{1+\eta} \leq \left((s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))| \right)^{(1+\eta)},
$$

can be absorbed by the exponential. Therefore,

$$
\begin{aligned}
|Rf(t,\mathbf{x})| &\leq C\Bigg\{ \sum_{i=2}^{n}\sum_{j=i}^{n} \int_{t}^{T} \frac{ds}{(s-t)^{i-\frac{1}{2}-\beta_i^j(j-\frac{1}{2})}} \int_{\mathbb{R}^{nd}} d\mathbf{y}\, \bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})|f(s,\mathbf{y})| \\
&\quad + \int_{t}^{T} \frac{ds}{(s-t)^{1-\frac{\eta}{2}}} \int_{\mathbb{R}^{nd}} d\mathbf{y}\, \bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})|f(s,\mathbf{y})| \\
&\quad + \sum_{i=2}^{n} \int_{t}^{T} \frac{ds}{(s-t)^{1-\eta(i-\frac{3}{2})}} \int_{\mathbb{R}^{nd}} d\mathbf{y}\, \bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})|f(s,\mathbf{y})| \Bigg\}.
\end{aligned}
\tag{3.13}
$$

Note carefully that the condition (1.3) precisely gives that for all $i \in [\![2,n]\!], j \in [\![i,n]\!]$, $1 - \{(i-1) + \frac{1}{2} - \beta_i^j(j-1+\frac{1}{2}) > 0$ so that all the above time singularity are integrable. Note also that, thanks to Lemma 5 (almost equivalence on the flows):

$$
\int_{\mathbb{R}^{nd}} d\mathbf{y}\, \bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^{nd}} d\mathbf{y}\, \frac{\exp\left(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x}-\boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2\right)}{(s-t)^{\frac{n^2 d}{2}}} \leq C_{3.14}.
\tag{3.14}
$$

The results in (3.3) and (3.4) now follow from (3.13) and the following key Lemma.

Lemma 10 ($L^{q'} - L^{p'}$ Controls for the singularized Green kernel)**.** *Introduce for* $f \in L^{q'}\left([0,T], L^{p'}(\mathbb{R}^{nd})\right)$, *with* $p', q' > 1$ *and some* $\gamma \in [0,1)$ *the quantify:*

$$
N_\gamma f(t,\mathbf{x}) := \int_{t}^{T} \frac{ds}{(s-t)^{\gamma}} \int_{\mathbb{R}^{nd}} d\mathbf{y}\, \bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})|f(s,\mathbf{y})|.
\tag{3.15}
$$

There exists $q_0' := q_0'(\gamma) \geq 1$, $p_0' := p_0'(\gamma) \geq 1$ *and* $C := C((\mathbf{A}),T,p_0',q_0')$ *s.t. for* $q' \geq q_0', p' \geq p_0'$, *it holds that for all* $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$:

$$
N_\gamma f(t,\mathbf{x}) \leq C\|f\|_{L^{q'}([0,T]\times\mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))}.
$$

Also, there exists $C(T) := C(T,(\mathbf{A}),\gamma) \underset{T\to 0}{\longrightarrow} 0$ *s.t. for all* $f \in L^{q'}\left([0,T], L^{p'}(\mathbb{R}^{nd})\right)$,

$$
\|N_\gamma f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \leq C(T)\|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))}.
$$

$\square$

Proof of Lemma 10. Let us first start with the pointwise estimate. For all $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$, denoting by $\tilde{p}', \tilde{q}'$ the conjugate exponents of p', q', write:

$$
\begin{aligned}
N_\gamma f(t,\mathbf{x}) &\leq \left(\int_{t}^{T} \frac{ds}{(s-t)^{\tilde{q}'\gamma}} \left(\int_{\mathbb{R}^{nd}} d\mathbf{y}\, \bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})^{\tilde{p}'} \right)^{\frac{\tilde{q}'}{\tilde{p}'}} \right)^{\frac{1}{\tilde{q}'}} \|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \\
&\leq C \int_{t}^{T} \frac{ds}{(s-t)^{\tilde{q}'(\gamma+\frac{n^2 d}{2}(1-\frac{1}{\tilde{p}'}))}} \|f\|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))}.
\end{aligned}
$$

It is thus clear that the *pointwise* statement of the lemma is fulfilled if p', q' are *large* enough, in order to guarantee that, for $(\tilde{p}')^{-1} + (p')^{-1} = 1$, $(\tilde{q}')^{-1} + (q')^{-1} = 1$, $\tilde{p}', \tilde{q}'$ are sufficiently close to 1 to have $\tilde{q}'(\gamma + \frac{n^2 d}{2}(1-\frac{1}{\tilde{p}'})) < 1$.

Let us now turn to the estimate in $L^{q'}\left([0,T], L^{p'}(\mathbb{R}^{nd})\right)$ norm. Setting

$$
K_{t,s,\gamma} \star f(s,\mathbf{x}) := \frac{1}{(s-t)^{\gamma}} \int_{\mathbb{R}^{nd}} d\mathbf{y}\, \bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})f(s,\mathbf{y})d\mathbf{y},
$$

we derive from (3.15) and the triangle inequality:
(3.16)
$$\left\|N_\gamma f(t,\cdot)\right\|_{L^{p'}(\mathbb{R}^{nd})} = \left\|\int_t^T ds K_{t,s,\gamma}\star f(s,\cdot)\right\|_{L^{p'}(\mathbb{R}^{nd})} \le \int_t^T ds \|K_{t,s,\gamma}\star f(s,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}.$$

From usual $L^1 - L^{p'}$ convolution arguments we also get:

$$\|K_{t,s,\gamma}\star f(s,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})} \le \frac{C}{(s-t)^\gamma}\|f(s,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}.$$

Plugging this estimate into (3.16) and using the Hölder inequality with exponents $q',\tilde{q}' > 1$ s.t. $(q')^{-1} + (\tilde{q}')^{-1} = 1$ we obtain:

$$
\begin{aligned}
\int_0^T dt\|N_\gamma f(t,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}^{q'} &\le C\int_0^T dt\Big(\int_t^T \frac{ds}{(s-t)^\gamma}\|f(s,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}\Big)^{q'} \\
&\le C_T\int_0^T dt\int_t^T ds\frac{C}{(s-t)^\gamma}\|f(s,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}^{q'},
\end{aligned}
$$

where $C_T := C((\mathbf{A}), p', q', T)$ denotes a constant which is small as T is. From the Fubini Theorem we eventually derive

$$
\begin{aligned}
\int_0^T dt\|N_\gamma f(t,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}^{q'} &\le C_T\int_0^T ds\|f(s,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}^{q'}\int_0^s dt\frac{C}{(s-t)^\gamma} \\
&\le C_T\|f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))}^{q'},
\end{aligned}
$$

up to a modification of C_T in the last inequality. $\qquad\square$

Hence, using the above control in (3.13) yields:

$$\|Rf\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))} \le C_T\|f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))}.$$

This concludes the proof of Lemma 9. $\qquad\square$

3.2. **Existence and uniqueness under (D)-(b) and (c).** Since in that setting no continuity is assumed on F_1 we will derive the well posedness through Girsanov arguments. It therefore clearly appears that the dynamics with 0 drift on the first non-degenerate component, and its associated linearization, plays a key role. We first introduce some notations used throughout this section.

We first define for all $\mathbf{x} \in \mathbb{R}^{nd}$, $\bar{\mathbf{F}}(t,\mathbf{x}) := (\mathbf{0}, F_2(t,\mathbf{x}), \cdots, F_n(t,\mathbf{x}))$, where $\bar{\mathbf{F}}$ satisfies **(S)** and **(H)**. Recall from Theorem 1 that, under **(A)**-(a), weak uniqueness holds for the SDE:

(3.17)
$$d\bar{\mathbf{X}}_t = \bar{\mathbf{F}}(t,\bar{\mathbf{X}}_t)dt + B\sigma(t,\bar{\mathbf{X}}_t)dW_t.$$

For fixed $(T,\mathbf{y}) \in \mathbb{R}^+ \times \mathbb{R}^{nd}$, we consider the following deterministic system to define our Gaussian proxy:

(3.18)
$$\dot{\bar{\boldsymbol{\theta}}}_{t,T}(\mathbf{y}) = \bar{\mathbf{F}}(t,\bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})), \ \bar{\boldsymbol{\theta}}_{T,T}(\mathbf{y}) = \mathbf{y},$$

and

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

Again, in (3.18), we consider a Cauchy-Peano flow furnished by Lemma 4 and which also satisfies the equivalence of rescaled norms of Lemma 5. The dynamics of the linearized Gaussian process associated with (3.17) writes:

(3.19)
$$d\tilde{\bar{\mathbf{X}}}_t^{T,\mathbf{y}} = \Big(\bar{\mathbf{F}}(t,\bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})) + D\bar{\mathbf{F}}(t,\bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y}))(\tilde{\bar{\mathbf{X}}}_t^{T,\mathbf{y}} - \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})) \Big) dt + B\sigma(t,\bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y}))dW_t,$$

and we denote the associated generator by $(\tilde{\bar{L}}_t^{T,\mathbf{y}})_{t\in[0,T]}$ and by $\tilde{\bar{p}}^{T,\mathbf{y}}(t,s,\mathbf{x},\cdot)$ the corresponding density at times $s>t$ when the process starts in $\mathbf{x}$ at time t.

We point out that, with respect to the previously used notations, we choose to keep track of the driftless dynamics for F_1 adding bars on the associated objects: dynamics, generators, density.

For our strategy, recall that we aim at proving uniqueness for the initial SDE (1.1) through the well posedness of the martingale problem associated with $(L_t)_{t\geq0}$. Once existence is known, the point is that we use a different Gaussian proxy than previously, namely the one considered in (3.19) associated with the driftless dynamics on the first component.

3.2.1. *Existence and Uniqueness under Assumption* **(D)**-(b). Under (b) (bounded measurable drift F_1 on the non-degenerate component), existence is a direct consequence of the Girsanov theorem. We thus now focus on uniqueness.

Repeating the previous approach (see subsection 3.1.2), using the family of random variables $(\tilde{\bar{\mathbf{X}}}_s^{s,\mathbf{y}})_{s\in[t,T]}$ defined in (3.19) with $\tilde{\bar{\mathbf{X}}}_t^{s,\mathbf{y}} = \mathbf{x}$ as Gaussian proxys, we have to bound analogously to the estimate of Lemma 9:

$$
\begin{aligned}
(3.20)\quad Rf(t,\mathbf{x}) &:= \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y}(L_t - \tilde{\bar{L}}_t^{s,\mathbf{y}})\tilde{\bar{p}}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y})f(s,\mathbf{y}) \\
&= \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y}\Big\{(L_t - \bar{L}_t) + (\bar{L}_t - \tilde{\bar{L}}_t^{s,\mathbf{y}})\Big\}\tilde{\bar{p}}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y})f(s,\mathbf{y}) \\
&=: \bar{\mathcal{R}}f(t,\mathbf{x}) + \bar{R}f(t,\mathbf{x}),
\end{aligned}
$$

where

$$\bar{R}f(t,\mathbf{x}) := \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y}(\bar{L}_t - \tilde{\bar{L}}_t^{s,\mathbf{y}})\tilde{\bar{p}}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y})f(s,\mathbf{y}),$$

and

$$
\begin{aligned}
(3.21)\quad \bar{\mathcal{R}}f(t,\mathbf{x}) &:= (R - \bar{R})f(t,\mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y}(L_t - \bar{L}_t)\tilde{\bar{p}}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y})f(s,\mathbf{y}) \\
&= \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y}\langle F_1(t,\mathbf{x}), D_{x_1}\tilde{\bar{p}}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y})\rangle f(s,\mathbf{y}).
\end{aligned}
$$

From Lemma 9, we have already shown that $\bar{R}f(t,\mathbf{x})$ is controlled in $L^{q'} - L^{p'}$ norm. It thus suffice to investigate the behavior of the $L^{q'} - L^{p'}$ norm of $\bar{\mathcal{R}}f(t,\mathbf{x})$ defined by (3.21). Namely, our goal is to prove that

$$(3.22)\qquad \|\bar{\mathcal{R}}f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))} \leq C(T)\|f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))}, \text{ with } C(T) \underset{T\to0}{\longrightarrow} 0.$$

Since from (2.8) and (3.21) we have for all $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$:

$$|\bar{\mathcal{R}}f(t,\mathbf{x})| \leq C\int_t^T \frac{ds}{(s-t)^{\frac{1}{2}}} \int_{\mathbb{R}^{nd}} d\mathbf{y} \frac{\exp\left(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \bar{\boldsymbol{\theta}}_{t,s}(\mathbf{y}))|^2\right)}{(s-t)^{\frac{n^2 d}{2}}}|f(s,\mathbf{y})|.$$

The estimate (3.22) then readily follows from Lemma 10.

3.2.2. *Existence and Uniqueness under Assumption* **(D)**-(c). We choose in this paragraph to address first the uniqueness, which is a rather direct extension of our previous approach, whereas the existence is a bit involved and requires to exploit the Krylov like inequality (1.5) that has been established for the process $(\bar{\mathbf{X}}_t)_{t \geq 0}$ with 0 drift in the non-degenerate component introduced in (3.17).

Uniqueness under Assumption (D)-(c). With the notations of the previous paragraph it remains to control, in $L^{q'} - L^{p'}$ norm, the contribution $\bar{\mathcal{R}}f$ introduced in (3.21). The term $\bar{R}f$ in (3.20) is again controlled as under assumption **(D)**-(a). Similarly to the previous paragraph (see also (2.8) and (3.21)), we have:

$$|D_{x_1}\tilde{\bar{p}}^{s,\mathbf{y}}(t,s,\mathbf{x},\mathbf{y})| \quad \leq \quad \frac{C}{(s-t)^{\frac{n^2 d}{2}+\frac{1}{2}}} \times \exp\left(-\frac{C^{-1}}{2}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x}-\bar{\boldsymbol{\theta}}_{t,s}(\mathbf{y}))|^2\right)$$

$$(3.23) \qquad\qquad\qquad \leq \quad \frac{C}{(s-t)^{\frac{1}{2}}}\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y}).$$

Uniqueness then follows from the following lemma which can be viewed as a refinement of Lemma 10 and explicitly exploits the condition on p and q stated in **(D)**-(c).

Lemma 11 (Refined $L^q - L^p$ control of singularized Green kernels)**.**
 Introduce for all $f, F_1 \in L^q\big([0,T], L^p(\mathbb{R}^{nd})\big)$ and $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$,

$$(3.24) \qquad u_{F_1}(t,\mathbf{x}) := F_1(t,\mathbf{x})\int_t^T \frac{ds}{(s-t)^{\frac{1}{2}}}\int_{\mathbb{R}^{nd}}\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})f(s,\mathbf{y})d\mathbf{y}.$$

Then, for all $p \geq 2, q > 2$ s.t. $\frac{n^2 d}{p} + \frac{2}{q} < 1$, there exists $C(T) := C(T,(\mathbf{A}),p,q) \xrightarrow[T \to 0]{} 0$ s.t. for all $f, F_1 \in L^q\big([0,T], L^p(\mathbb{R}^{nd})\big)$

$$\|u_{F_1}\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))} \leq C(T)\|F_1\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}\|f\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}.$$

$\square$

Proof of Lemma 11. With the notations of Lemma 10 rewrite:

$$u_{F_1}(t,\mathbf{x}) = F_1(t,\mathbf{x})\int_t^T \frac{ds}{(s-t)^{\frac{1}{2}}}\int_{\mathbb{R}^{nd}}\bar{p}_{C^{-1}}(t,s,\mathbf{x},\mathbf{y})f(s,\mathbf{y})d\mathbf{y}$$

$$(3.25) \qquad\qquad =: F_1(t,\mathbf{x})\int_t^T ds K_{t,s,\frac{1}{2}} \star f(s,\mathbf{x}).$$

The triangle inequality yields:
(3.26)
$$\left\|\int_t^T ds F_1(t,\cdot)K_{t,s,\frac{1}{2}} \star f(s,\cdot)\right\|_{L^p(\mathbb{R}^{nd})} \leq \int_t^T ds\|F_1(t,\cdot)K_{t,s,\frac{1}{2}} \star f(s,\cdot)\|_{L^p(\mathbb{R}^{nd})}.$$

The idea is here to reproduce the computations of Lemma 10 integrating directly the singularized heat-kernel, i.e. $K_{t,s,\frac{1}{2}}$, in the $\mathbf{y}$ variable when performing the Hölder inequality in order to make the product of the norms $\|f(s,\cdot)\|_{L^p(\mathbb{R}^{nd})}\|F_1(t,\cdot)\|_{L^p(\mathbb{R}^{nd})}$

appear. Precisely:

$$\|F_1(t,\cdot)K_{t,s,\frac{1}{2}} \star f(s,\cdot)\|_{L^p(\mathbb{R}^{nd})}$$

$$=\Big(\int_{\mathbb{R}^{nd}} d\mathbf{x}|F_1(t,\mathbf{x})|^p\Big|\int_{\mathbb{R}^{nd}} d\mathbf{y}K_{t,s,\frac{1}{2}}(\mathbf{x},\mathbf{y})f(s,\mathbf{y})\Big|^p\Big)^{\frac{1}{p}}$$

$$\leq\Big(\int_{\mathbb{R}^{nd}} d\mathbf{x}|F_1(t,\mathbf{x})|^p\Big\{\int_{\mathbb{R}^{nd}} d\mathbf{y}|f(s,\mathbf{y})|^p\Big\}\Big\{\int_{\mathbb{R}^{nd}} d\mathbf{y}|K_{t,s,\frac{1}{2}}(\mathbf{x},\mathbf{y})|^{\tilde{p}}\Big\}^{\frac{p}{\tilde{p}}}\Big)^{\frac{1}{p}},$$

where again $p^{-1} + (\tilde{p})^{-1} = 1$. Observe now that usual Gaussian calculations give that there exists $C_p > 0$ such that, for all $\mathbf{x}$ in $\mathbb{R}^{nd}$:

$$\Big\{\int_{\mathbb{R}^{nd}} d\mathbf{y}|K_{t,s,\frac{1}{2}}(\mathbf{x},\mathbf{y})|^{\tilde{p}}\Big\}^{\frac{1}{\tilde{p}}} \leq \frac{C_p}{(s-t)^{\frac{n^2d}{2}\frac{1}{p}+\frac{1}{2}}}.$$

Thus,

(3.27)
$$\|F_1(t,\cdot)K_{t,s,\frac{1}{2}} \star f(s,\cdot)\|_{L^p(\mathbb{R}^{nd})} \leq \|F_1(t,\cdot)\|_{L^p(\mathbb{R}^{nd})}\|f(s,\cdot)\|_{L^p(\mathbb{R}^{nd})}\frac{C_p}{(s-t)^{\frac{n^2d}{2}\frac{1}{p}+\frac{1}{2}}},$$

which yields from (3.26) that:

$$\Big\|\int_t^T dsF_1(t,\cdot)K_{t,s,\frac{1}{2}} \star f(s,\cdot)\Big\|_{L^p(\mathbb{R}^{nd})}$$

$$\leq C_p\int_t^T \frac{ds}{(s-t)^{\frac{n^2d}{2}\frac{1}{p}+\frac{1}{2}}}\|F_1(t,\cdot)\|_{L^p(\mathbb{R}^{nd})}\|f(s,\cdot)\|_{L^p(\mathbb{R}^{nd})}.$$

From the definition in (3.25) we eventually derive that:

$$\int_0^T dt\|u_{F_1}(t,\cdot)\|_{L^p(\mathbb{R}^{nd})}^q$$

$$\leq C_p^q\int_0^T \|F_1(t,\cdot)\|_{L^p(\mathbb{R}^{nd})}^q\Big(\int_t^T \frac{ds}{(s-t)^{(\frac{n^2d}{2}\frac{1}{p}+\frac{1}{2})\tilde{q}}}\Big)^{\frac{q}{\tilde{q}}}\int_t^T ds\|f(s,\cdot)\|_{L^p(\mathbb{R}^{nd})}^q$$

(3.28)
$$\leq C_p^q\|F_1\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}^q\|f\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}^q\Big(\int_t^T \frac{ds}{(s-t)^{(\frac{n^2d}{2}\frac{1}{p}+\frac{1}{2})\tilde{q}}}\Big)^{\frac{q}{\tilde{q}}},$$

with $q^{-1} + (\tilde{q})^{-1} = 1$. Let us now show that the remaining time integral in the above equation gives a small contribution in times. To do so, it suffices to show that $(\frac{n^2d}{2}\frac{1}{p} + \frac{1}{2})\tilde{q} < 1$. Since $(\tilde{q})^{-1} = 1 - \frac{1}{q}$, we have that:

$$\Big(\frac{n^2d}{2}\frac{1}{p} + \frac{1}{2}\Big)\tilde{q} < 1 \iff \frac{n^2d}{2}\frac{1}{p} + \frac{1}{2} < 1 - \frac{1}{q} \iff \frac{n^2d}{p} + \frac{2}{q} < 1,$$

which is precisely the condition appearing in **(D)**-(c) and assumed in the current Lemma. $\qquad\square$

Existence under Assumption (D)-(c). We here consider a function $F_1 \in L^q([0, T],$ $L^p(\mathbb{R}^{nd}))$, where p, q are as in **(D)**-(c). To prove the existence, the strategy is to exploit the idea introduced by Portenko [Por90] and used by Krylov and Röckner [KR05] as well to build local weak solutions (before they also establish that they are actually strong solutions) in the non-degenerate case. We also refer for perturbed degenerate Ornstein-Uhlenbeck dynamics to [FFPV17] and [Zha18]. We adapt a bit this approach.

Recall that for the process $(\bar{\mathbf{X}}_t)_{t \geq 0}$ introduced in (3.17) we have from Theorem 1, equation (1.5), the following density estimate.

Denoting by $\bar{P}(t, s, \mathbf{x}, \cdot)$ the transition probability determined by $(\bar{L}_s)_{s \geq 0}$, it is s.t. for a given $T > 0$, almost all $s \in (t, T]$ and all $\Gamma \in \mathcal{B}(\mathbb{R}^{nd})$: $\bar{P}(t, s, \mathbf{x}, \Gamma) = \int_\Gamma \bar{p}(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y}$. More specifically, for any $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd})), \frac{n^2 d}{p'} + \frac{2}{q'} < 2, p' > 1, q' > 1$, and $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$:
(3.29)
$$\left| \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{t, \mathbf{x}}} \left[\int_t^T f(s, \bar{\mathbf{X}}_s) ds \right] \right| \leq C_{3.29} \|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}, \quad C_{3.29} := C_{3.29}((\mathbf{A}), p', q', T).$$

We now state an exponential integrability result for the unique weak solution of (3.17). Such types of estimates were first proved by Khas'minskii in [Kha59] for the Brownian motion. We can also refer to Lemma 2.1 in Chapter 1 of the monograph by Sznitman [Szn98]. Since the proof only relies on the Markov property, it readily extends to the current inhomogeneous and non-Brownian framework.

Lemma 12 (Khas'minskii's type exponential integrability). *Let $(\bar{\mathbf{X}}_t)_{t \geq 0}$ be the (unique-weak) solution to (3.17). Then, for any fixed $T > 0$ and a positive Borel function $f : [0, T] \times \mathbb{R}^{nd} \to \mathbb{R}_+$ s.t.*

$$\alpha_T := \sup_{\mathbf{x} \in \mathbb{R}^{nd}} \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{0, \mathbf{x}}} \left[\int_0^T f(s, \bar{\mathbf{X}}_s) ds \right] < 1,$$

one also has:

$$\sup_{\mathbf{x} \in \mathbb{R}^{nd}} \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{0, \mathbf{x}}} \left[\exp \left(\int_0^T f(s, \bar{\mathbf{X}}_s) ds \right) \right] < \frac{1}{1 - \alpha_T}.$$

As a corollary to the previous Lemma we obtain the following proposition which will allow to apply the Girsanov Theorem to derive the existence of a solution to the martingale problem under the assumption **(D)**-(c) through a change of probability.

Proposition 13 (Exponential integrability). *Let $(\bar{\mathbf{X}}_t)_{t \geq 0}$ be the (unique-weak) solution to (3.17). Let $F_1 \in L^q([0, T], L^p(\mathbb{R}^{nd}))$ with $\frac{n^2 d}{p} + \frac{2}{q} < 1, p \geq 2, q > 2$. Then, for any $\lambda > 0$, there exists $K_{F_1, \lambda} := K_{F_1, \lambda}((\mathbf{A}), T)$ s.t.*

$$(3.30) \qquad \sup_{\mathbf{x} \in \mathbb{R}^{nd}} \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{0, \mathbf{x}}} \left[\exp \left(\lambda \int_0^T |(\sigma^{-1} F_1)(s, \bar{\mathbf{X}}_s)|^2 ds \right) \right] \leq K_{F_1, \lambda}.$$

The constant $K_{F_1, \lambda}$ depends continuously on $\|F_1\|_{L^q([0, T], L^p(\mathbb{R}^{nd}))}$ and λ.

We point out that in the above Proposition, the case $p = 2$ can only be considered in the scalar non-degenerate case $d = n = 1$.

Proof. Observe that for $a > 1$ s.t. $a(\frac{n^2 d}{p} + \frac{2}{q}) < 1$, setting $p' = \frac{p}{2a}$, $q' = \frac{q}{2a}$, so that indeed $\frac{n^2 d}{p'} + \frac{2}{q'} < 2$, one has:

$$\| \, |\sigma^{-1} F_1|^{2a} \, \|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} = \left(\int_0^T dt \Big\{ \int_{\mathbb{R}^{nd}} d\mathbf{x} (|(\sigma^{-1} F_1)(t,\mathbf{x})|^{2a})^{p'} \Big\}^{\frac{q'}{p'}} \right)^{\frac{1}{q'}}$$

$$\leq \kappa^a \left(\int_0^T dt \Big\{ \int_{\mathbb{R}^{nd}} d\mathbf{x} |F_1(t,\mathbf{x})|^p \Big\}^{\frac{q}{p}} \right)^{\frac{2a}{q}} \leq \kappa^a \|F_1\|_{L^q([0,T], L^p(\mathbb{R}^{nd}))}^{2a}.$$

From equation (3.29) we thus derive that for all $\mathbf{x} \in \mathbb{R}^{nd}$:

$$(3.31) \quad \left| \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{0,\mathbf{x}}} \Big[\int_0^T |\sigma^{-1} F_1(s, \bar{\mathbf{X}}_s)|^{2a} ds \Big] \right|$$

$$\leq C_{3.29} \| \, |\sigma^{-1} F_1|^{2a} \, \|_{L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))} \leq C_{3.29} \kappa^a \|F_1\|_{L^q([0,T], L^p(\mathbb{R}^{nd}))}^{2a}.$$

For a as above, write now for $\varepsilon \in (0,1)$ and from the Young inequality:

$$\bar{\mathbb{E}}^{\bar{\mathbb{P}}_{0,\mathbf{x}}} \left[\exp \left(\lambda \int_0^T |(\sigma^{-1} F_1)(s, \bar{\mathbf{X}}_s)|^2 ds \right) \right]$$

$$\leq \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{0,\mathbf{x}}} \left[\exp \left(\lambda \int_0^T |(\sigma^{-1} F_1)(s, \bar{\mathbf{X}}_s)|^{2a} \frac{\varepsilon}{a} ds \right) \right] C(T, a, \varepsilon, \lambda).$$

The statement now directly follows from the above equation, (3.31) and Lemma 12 taking $\varepsilon := (2\lambda C_{3.29} \kappa^a \| \, F_1 \|_{L^q([0,T], L^p(\mathbb{R}^{nd}))}^{2a})^{-1}$.   □

Proposition 13 ensures that the Novikov condition is fulfilled in order to prove existence for the martingale problem associated with $(L_t)_{t \geq 0}$ for F_1 satisfying **(D)**-(c) starting from $\bar{\mathbb{P}}_{0,\mathbf{x}}$ and the associated dynamics (3.17) of the canonical process. Set,

$$\tilde{W}_t := W_t - \int_0^t (\sigma^{-1} F_1)(s, \bar{\mathbf{X}}_s) ds, \ t \in [0, T].$$

From Proposition 13, we derive that:

$$M_t := \exp \left(\int_0^t (\sigma^{-1} F_1)(s, \bar{\mathbf{X}}_s) dW_s - \frac{1}{2} \int_0^t |(\sigma^{-1} F_1)(s, \bar{\mathbf{X}}_s)|^2 ds \right), \ t \in [0, T]$$

is a $\bar{\mathbb{P}}$-$\mathcal{F}_t$ martingale (here $(\mathcal{F}_s)_{s \in [0,T]}$ stands for the natural filtration associated with the canonical process $(\bar{\mathbf{X}}_s)_{s \in [0,T]}$ under $\bar{\mathbb{P}}$). It follows from the Girsanov theorem that $(\tilde{W}_t)_{t \in [0,T]}$ is a Wiener process on $(\Omega, \mathcal{F}_T, (\mathcal{F}_s)_{s \in [0,T]}, \mathbb{P})$ where $[d\mathbb{P}/d\bar{\mathbb{P}}]\big|_{\mathcal{F}_T} := M_T$. The dynamics of $(\bar{\mathbf{X}}_s)_{s \in [0,T]}$, writes under $\mathbb{P}$:

$$d\bar{\mathbf{X}}_t = \left(BF_1(t, \bar{\mathbf{X}}_t) + \bar{\mathbf{F}}(t, \bar{\mathbf{X}}_t) \right) dt + B\sigma(t, \bar{\mathbf{X}}_t) d\tilde{W}_t = \mathbf{F}(t, \bar{\mathbf{X}}_t) dt + B\sigma(t, \bar{\mathbf{X}}_t) d\tilde{W}_t,$$

that is $(\bar{\mathbf{X}}_t)_{t \in [0,T]}$ solves (1.1) under $\mathbb{P}$.

3.3. **Krylov bounds under (D)-(b) and (c).**
We aim here at proving estimate (1.5) under assumptions **(D)**-(b) and (c). We mainly focus on case (c) which is the more involved. Case (b) can indeed be derived from the methodology developed under **(D)**-(a) and the controls obtained in Section 3.2.1 or alternatively following the approach presented below for **(D)**-(c) which remains valid for a bounded drift on the non-degenerate component.

Thanks to Proposition 13, the proof of the Krylov bound (1.5) is similar to the proof of Lemma 3.3 in [KR05]. We provide a complete proof in our context for the sake of completeness.

Write now for $f \in L^{q'}([0,T], L^{p'}(\mathbb{R}^{nd}))$, $\frac{n^2 d}{p'} + \frac{2}{q'} < 2, p' > 1, q' > 1$, $(t, \mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$ with the notations of the previous paragraph:

$$
\begin{aligned}
|\mathbb{E}^{\mathbb{P}_{t,\mathbf{x}}}[\int_t^T f(s, \mathbf{X}_s) ds]| &= \left| \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{t,\mathbf{x}}} \Big[\frac{M_T}{M_t} \int_t^T f(s, \mathbf{X}_s) ds \Big] \right| \\
&\leq \bar{\mathbb{E}}^{\bar{\mathbb{P}}_{t,\mathbf{x}}} \Big[\Big(\frac{M_T}{M_t} \Big)^\alpha \Big]^{\frac{1}{\alpha}} T^{\frac{1}{\alpha}} \Big(\bar{\mathbb{E}}^{\bar{\mathbb{P}}_{t,\mathbf{x}}} \Big[\int_t^T |f(s, \bar{\mathbf{X}}_s)|^\beta ds \Big] \Big)^{\frac{1}{\beta}},
\end{aligned}
$$

where $\alpha^{-1} + \beta^{-1} = 1, \alpha > 1, \beta > 1$. We have from Proposition 13 that the exponential martingale has moments of all orders so that there exists $C_T > 0$, with $C_T \to 0$ when $T \to 0$ such that:

$$
\begin{aligned}
|\mathbb{E}^{\mathbb{P}_{t,\mathbf{x}}}[\int_t^T f(s, \mathbf{X}_s) ds]| &\leq C_T \Big(\bar{\mathbb{E}}^{\bar{\mathbb{P}}_{t,\mathbf{x}}} \Big[\int_t^T |f(s, \bar{\mathbf{X}}_s)|^\beta ds \Big] \Big)^{\frac{1}{\beta}} \\
&\leq C_T \||f|^\beta\|_{L^{q''}([0,T], L^{p''}(\mathbb{R}^{nd}))}^{\frac{1}{\beta}} \\
&\leq C_T \|f\|_{L^{q''\beta}([0,T], L^{p''\beta}(\mathbb{R}^{nd}))},
\end{aligned}
$$

if $p'', q'' \in [1, +\infty]$ are s.t. $n^2 d/p'' + 2/q'' < 2$. Taking $\beta > 1$ sufficiently close to 1 s.t. the previous condition holds for $p'' = p'/\beta, q'' = q'/\beta$ eventually yields the result.

4. Counter example

This section is devoted to the almost sharpness of the thresholds appearing in Theorem 1. This is the purpose of Theorem 2 which we now prove. We only focus here on the statement concerning the almost sharpness of the Hölder exponents β_i^j, $i \leq j$ in $[\![2, n]\!]^2$ in (1.3). Indeed, as emphasized in the introduction, the sharpness of $L^q - L^p$ integrability conditions on the first component F_1 of the drift follows from Example 69 of [BFGM19] (see also Proposition 3.3 of [GO13]). We eventually recall that, the almost sharpness of the coefficients β_i^i, i in $[\![2, n]\!]$, has been already proved in [CdR18].

Let us first introduce the main idea of our counter example. As we already discussed, the Peano system (1.2) is ill posed as soon as α is in $(0, 1)$ and Y starts from 0 and well posed (in a strong sense) as soon as it is suitably perturbed. In [DF14], the authors show that, in order to regularize, there must exist a transition time strictly less than one such that, before this time, the noise dominates in the dynamics of the system and therefore allows the solution to leave the singularity. This competition can be written explicitly and gives the following (heuristic) rule: the fluctuations of order γ of the noise added in the system has to be strictly lower than $1/(1 - \alpha)$. We formalize these facts with our Proposition 14 below.

This proposition will be the key tool to handle each Hölder threshold which depends on the component and the variable. Hence, in order to exhibit the (almost) optimal threshold for the drift of the ith component with respect to the jth variable we need to build an *ad hoc* Peano-like example. Focusing on the ith component

and the jth variable, for $i \in [\![2, n]\!]$ and $j \in [\![i, n]\!]$, we consider:

$$
\begin{cases}
\dot{x}_t^1 = \dot{x}_t^2 = \ldots = \dot{x}_t^{i-1} = 0 \\[2mm]
\dot{x}_t^i = \text{sign}(x_t^j)|x_t^j|^{\beta_i^j} \\[2mm]
\dot{x}_t^{i+1} = x_t^i, \ldots, \dot{x}_t^j = x_t^{j-1} \\[2mm]
\dot{x}_t^{j+1} = \dot{x}_t^{j+2} = \ldots = \dot{x}_t^n = 0
\end{cases}
\qquad t \geq 0,
$$

and $x_0^l = 0$ for all l in $[\![1, j]\!]$. Each entry $(x_t^k)_{t \geq 0, k \in [\![1,n]\!]}$ of the above dynamics is scalar. It is well seen that the global well posedness of this system relies on the well posedness of the ith equation whose extremal solutions write $\pm c_{\beta_i^j}^{ij} t^{((j-i)\beta_i^j+1)/(1-\beta_i^j)}$, for some positive $c_{\beta_i^j}^{ij}$. For the considered initial point $\mathbf{x}_0 = (x_0^1, \cdots, x_0^n) = (0, \cdots, 0)$, the above dynamics can be rewritten in short as:

$$
d\mathbf{x}_t = [\mathbf{A}\mathbf{x}_t + F_i^j(\mathbf{x}_t)]dt, \quad \mathbf{A} = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & 0 & \vdots \\ 0 & 1 & 0 & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix},
$$

where $F_i^j(\mathbf{x}) = e_i \text{sign}(x_j)|x_j|^{\beta_i^j}$ and e_i is the ith vector of the canonical basis of $\mathbb{R}^n$. In that case, the corresponding stochastic perturbation has the form

(4.1) $$dX_t = [\mathbf{A}X_t + F_i^j(X_t)]dt + BdW_t,$$

This in particular means that the perturbation of the ith component is done by the $(i-1)^{\text{th}}$ iterated integrals (in time) of Brownian motion. Thus, focusing on the ith level of the chain and the jth component this means that we are interested in the following type of SDE

(4.2)

$$
Z_t^{i,j} = x + \mathcal{W}_t^i
$$

$$
+ \int_0^t \text{sign}\left(\int_0^s \ldots \int_0^{s_{i+2}} Z_{s_{i+1}}^{i,j} ds_{i+1} \ldots ds_j \right) \left| \int_0^s \ldots \int_0^{s_{i+2}} Z_{s_{i+1}}^{i,j} ds_{i+1} \ldots ds_j \right|^{\beta_i^j} ds
$$

where $\mathcal{W}^i$ will be chosen as the $(i-1)^{\text{th}}$ iterated integral in time of Brownian motion. The non-uniqueness in law for equation (4.2) will then follow from the next proposition.

Proposition 14 (Failure of the well posedness for the regularized Peano system). Let $\mathcal{W}$ be a random process with continuous paths satisfying, in law, an invariance by symmetry and a self-similarity property of order $\gamma > 0$. Namely:

$$
(\mathcal{W}_t, \ t \geq 0) \overset{(\text{law})}{=} (-\mathcal{W}_t, \ t \geq 0), \ \forall \rho > 0, (\mathcal{W}_{\rho t})_{t \geq 0} \overset{(\text{law})}{=} (\rho^\gamma \mathcal{W}_t)_{t \geq 0}.
$$

Suppose moreover that $\mathbb{E}[\sup_{s\in[0,1]}|\mathcal{W}_s|] < +\infty$ and that $\mathcal{W}$ and $\alpha < 1$ are such that there exists a weak solution to the following SDE:

(4.3)
$$Z_t = x + \int_0^t \text{sign}\left(\int_0^s \ldots \int_0^{s_2} Z_{s_1} ds_1 \ldots ds_k\right)\left|\int_0^s \ldots \int_0^{s_2} Z_{s_1} ds_1 \ldots ds_k\right|^\alpha ds + \mathcal{W}_t,$$

for any initial condition $x \geq 0$ where $k \in \mathbb{N}$ is given and that it satisfies the Kolmogorov criterion.

Then, if $\alpha < (\gamma - 1)/(k + \gamma)$, uniqueness in law fails for (4.3).

Turning now to our claim, it is clear that for $\beta_i^j > 0$ (4.2), admits for all initial condition $x \geq 0$ at least one solution which satisfies the Kolmogorov Criterion. Our statement concerning the non uniqueness in law for the solution of (4.2) then readily follows from Proposition 14. Taking $\alpha = \beta_i^j$, $\mathcal{W} = \mathcal{W}^i$, which corresponds to the $(i-1)^{\text{th}}$ iterated integrals (in time) of Brownian motion and therefore induces to take $\gamma = i - \frac{1}{2}$, and $k = j - i$, we deduce that weak uniqueness fails as soon as

$$\beta_i^j < \frac{2i - 3}{2j - 1}.$$

It now remains to prove Proposition 14.

Proof of Proposition 14. Consider the extremal solutions of the deterministic solutions associated with (4.3), that correspond to the case $\mathcal{W}_t = 0$ and write $\pm c_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)}$. The crucial point consists in comparing the fluctuations of the noise in the dynamics of (4.3) with the extremal solution of the associated Peano Example. The proof follows the lines of [CdR18] but we decide to reproduce it here for the sake of completeness.

For a given parameter $\beta \in (0,1)$, we define for any continuous path Y from $\mathbb{R}^+$ to $\mathbb{R}$ the variable $\tau(Y)$ as

$$\tau(Y) = \inf\{t \geq 0 \ : \ Y_t \leq (1-\beta)c_{\alpha,k}t^{(k\alpha+1)/(1-\alpha)}\}.$$

The stopping time $\tau(Y)$ then corresponds to the first passage of Y below a threshold related to the (positive) extremal solution of the deterministic Peano system. Then, the key point to the proof of Proposition 14 is the following Lemma.

Lemma 15. *Let Z be a weak solution of (4.3) starting from some $x > 0$ and suppose that $\alpha < (\gamma - 1)/(k + \gamma)$. Then, there exists a positive ρ, depending on α, β, γ and $\mathbb{E}|\mathcal{W}_1|$ only, such that*

(4.4)
$$\mathbb{P}_x(\tau(Z) \geq \rho) \geq 3/4.$$

Roughly speaking, the result tells us that, when the noise in the system is not strong enough, the solution from above the (positive) extremal solution will remain above with great probability. If weak uniqueness holds, the symmetry property implies that any solution from below the (negative) extremal solution will remain below with great probability. Letting the starting point tend to 0 (i.e. to the singularity) this leads to a contradiction. Together with the above Lemma, this last fact will allow us to conclude our counter-example.

Let $(Z, \mathcal{W})$ be a weak solution of (4.3) with the initial condition $x = 0$. Then, $(-Z, -\mathcal{W})$ is also a weak solution of (4.3) so that, if uniqueness in law holds, Z and $-Z$ have the same law.

Let Z^n be a sequence of weak solutions of (4.3) starting from $1/n$, n being a positive integer and $(\mathbb{P}_{1/n})_{n \geq 0}$ its law. Thanks to Kolmogorov's criterion, we can extract a converging subsequence $(\mathbb{P}_{1/n_k})_{k \geq 0}$ that converges to $\mathbb{P}_0$, the law of the weak solution Z of (1.2) starting from 0. Since the bound in (4.4) does not depend on the initial condition we get that

$$\mathbb{P}_0(\tau(Z) \geq \rho) \geq 3/4,$$

and, thanks to uniqueness in law

$$\mathbb{P}_0(\tau(-Z) \geq \rho) \geq 3/4,$$

which is obviously impossible.

$\square$

Proof of Lemma 15. Let Z be a weak solution of (4.3) starting from $x > 0$. Since it has continuous path, we have almost surely that $\tau(Z) > 0$. Then, note that for $t \in [0, \tau(Z)]$ we have:

$$
\begin{aligned}
Z_t &= x + \int_0^t \mathrm{sign}\left(\int_0^s \ldots \int_0^{s_2} Z_{s_1} ds_1 \ldots ds_k\right) \left|\int_0^s \ldots \int_0^{s_2} Z_{s_1} ds_1 \ldots ds_k\right|^\alpha ds + \mathcal{W}_t \\
&\geq (1-\beta)^\alpha \bar{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)} + \mathcal{W}_t,
\end{aligned}
$$

for some positive constant $\bar{c}_{\alpha,k}$. Indeed, from the very definition of $\tau(Z)$, one gets $Z_{s_1} \geq (1-\beta)c_{\alpha,k} s_1^{(k\alpha+1)/(1-\alpha)}$, which once integrated in time and taking the αth power yields a lower bound of the form $(1-\beta)^\alpha c_{\alpha,k}^\alpha \mathfrak{c}_{\alpha,k}^\alpha t^{\alpha(k+(k\alpha+1)/(1-\alpha))} = (1-\beta)^\alpha c_{\alpha,k}^\alpha \mathfrak{c}_{\alpha,k}^\alpha t^{\alpha(k+1)/(1-\alpha)}$ for some additional constant $\mathfrak{c}_{\alpha,k} > 0$ related to the iterated time integration, recalling that $\alpha < 1$ and that t is small for the last inequality. One sets eventually $\bar{c}_{\alpha,k} := c_{\alpha,k}^\alpha \mathfrak{c}_{\alpha,k}^\alpha$.

Hence, choosing η such that $(1-\eta) = [(1-\beta)^\alpha + (1-\beta)]/2$, we observe that $\beta - \eta + 1 - \eta = (1-\beta)^\alpha$ and we get that:

$$Z_t \geq (1-\eta)\bar{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)} + (\beta - \eta)\bar{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)} + \mathcal{W}_t,$$

for all t in $[0, \tau(Z)]$.

Let now ρ be a positive number to be specified later on. Set $\tilde{c}_{\alpha,k} = (\beta - \eta)\bar{c}_{\alpha,k}$ and define

$$A = \left\{\tilde{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)} + \mathcal{W}_t > 0 \text{ for all } t \text{ in } (0, \rho]\right\}.$$

The event A allows us to compare the fluctuations of the noise and those of the (positive) extremal solution. More precisely, it is the set of realizations for which the amount of noise in the system is lacking. Note that on A we have

$$Z_t \geq (1-\eta)\bar{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)} \geq (1-\beta)\bar{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)}$$

for all t in $[0, \rho]$, recalling that $\alpha, \beta \in (0,1)$ for the last inequality. But this is compatible only with the event $\{\tau(Z) \geq \rho\}$ so that $A \subset \{\tau(Z) \geq \rho\}$. Hence

(4.5) $$\mathbb{P}(\tau(Z) \geq \rho) \geq \mathbb{P}(A).$$

It remains to choose $\rho > 0$ such that $\mathbb{P}(A) \geq 3/4$. Write:

$$
\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}[\forall t \in (0, \rho], \ \tilde{c}_{\alpha,k} t^{\frac{k\alpha+1}{1-\alpha}} + \mathcal{W}_t > 0] = \mathbb{P}[\forall t \in (0, 1], \ \tilde{c}_{\alpha,k}(\rho t)^{\frac{k\alpha+1}{1-\alpha}} + \mathcal{W}_{\rho t} > 0] \\
&= \mathbb{P}[\forall t \in (0, 1], \ \tilde{c}_{\alpha,k}(\rho t)^{\frac{k\alpha+1}{1-\alpha}} + \rho^\gamma \mathcal{W}_t > 0] \\
&= \mathbb{P}[\forall t \in (0, 1], \ \tilde{c}_{\alpha,k} \rho^{\frac{k\alpha+1}{1-\alpha}-\gamma} + t^{-\frac{k\alpha+1}{1-\alpha}} \mathcal{W}_t > 0],
\end{aligned}
$$

from the self-similarity assumption on $\mathcal{W}$. Since by assumption $\alpha < \frac{\gamma-1}{\gamma+k} \iff \frac{k\alpha+1}{1-\alpha} - \gamma < 0$, the statement will follow taking ρ small enough as soon as we prove the process $\mathcal{R}_t := t^{-\frac{k\alpha+1}{1-\alpha}} \mathcal{W}_t$, $t \in (0, 1]$, which is continuous on the open set $(0, 1]$, can be extended by continuity in 0 with $\mathcal{R}_0 = 0$. Observe that $\mathbb{E}[|\mathcal{R}_t|] = t^{\gamma - \frac{k\alpha+1}{1-\alpha}} \mathbb{E}[|\mathcal{W}_1|] \xrightarrow[t \to 0]{} 0$. Setting $\delta := \gamma - \frac{k\alpha+1}{1-\alpha} > 0$ and introducing $t_n := n^{-1/\delta(1+\eta)}$, $\eta > 0$, we get that for all $\varepsilon > 0$,

$$
\mathbb{P}[|\mathcal{R}_{t_n}| \geq \varepsilon] \leq \varepsilon^{-1} \mathbb{E}[|\mathcal{R}_{t_n}|] = \varepsilon^{-1} t_n^\delta \mathbb{E}[|\mathcal{W}_1|] = \varepsilon^{-1} n^{-(1+\eta)} \mathbb{E}[|\mathcal{W}_1|].
$$

We thus get from the Borel-Cantelli lemma that $\mathcal{R}_{t_n} \xrightarrow[n, \ a.s.]{} 0$. Namely, we have almost sure convergence along the subsequence t_n going to zero with n. It now remains to prove that the process $\mathcal{R}_t$ does not fluctuate much between two successive times t_n and t_{n+1}. Write for $t \in [t_{n+1}, t_n]$:

$$
\begin{aligned}
|\mathcal{R}_t| := |t^{-\frac{k\alpha+1}{1-\alpha}} \mathcal{W}_t| &\leq t_{n+1}^{-\frac{k\alpha+1}{1-\alpha}} \left(|\mathcal{W}_{t_{n+1}}| + \sup_{s \in [t_{n+1}, t_n]} |\mathcal{W}_s - \mathcal{W}_{t_{n+1}}| \right) \\
(4.6) \qquad\qquad &\leq t_{n+1}^{-\frac{k\alpha+1}{1-\alpha}} \left(2|\mathcal{W}_{t_{n+1}}| + \sup_{s \in [0, t_n]} |\mathcal{W}_s| \right).
\end{aligned}
$$

The first term of the above left hand side tends almost surely to zero with n. Observe as well that, from the scaling properties of $\mathcal{W}$, for any $\varepsilon > 0$:

$$
\mathbb{P}[t_{n+1}^{-\frac{k\alpha+1}{1-\alpha}} \sup_{s \in [0, t_n]} |\mathcal{W}_s| \geq \varepsilon] = \mathbb{P}[t_{n+1}^{-\frac{k\alpha+1}{1-\alpha}} t_n^\gamma \sup_{s \in [0, 1]} |\mathcal{W}_s| \geq \varepsilon]
$$

$$
\leq \varepsilon^{-1} t_n^\delta \left(\frac{t_n}{t_{n+1}}\right)^{\frac{k\alpha+1}{1-\alpha}} \mathbb{E}[\sup_{s \in [0,1]} |\mathcal{W}_s|] \leq C\varepsilon^{-1} n^{-(1+\eta)} \mathbb{E}[\sup_{s \in [0,1]} |\mathcal{W}_s|],
$$

which again gives from the Borel-Cantelli lemma the a.s. convergence with n of the second term in the r.h.s of (4.6). We eventually derive that $\mathcal{R}_t \xrightarrow[t \to 0, \ a.s.]{} 0$. Again, the key point is that we normalize the process $\mathcal{W}$ at a rate, $t^{\frac{k\alpha+1}{1-\alpha}}$, which is lower than its own characteristic time scale, t^γ. This is precisely what leaves some margin to establish continuity.

$\square$

Appendix A. Proof of the Technical Lemmas 6 and 7

A.1. **Proof of Lemma 6 (Dirac convergence of the frozen density).** For the proof of this Lemma we are somehow faced with the same type of difficulties as for Lemma 5. Namely, recalling the expression of $\tilde{p}^{t+\varepsilon, \mathbf{y}}(t, t+\varepsilon, \mathbf{x}, \mathbf{y})$ derived from (2.7), we have a dependence of the covariance matrix $\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon, \mathbf{y}}$ and of the linearized flow $\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon, \mathbf{y}}(\mathbf{x})$ in the integration variable $\mathbf{y}$.

Let $(\boldsymbol{\theta}_{u,t}(\mathbf{x}))_{u\in[t,t+\varepsilon]}$ be the forward flow provided by Lemma 4. To study the sensitivity of the covariance matrix w.r.t. the flows we now introduce, for a given point $\mathbf{x}\in\mathbb{R}^{nd}$, the linear Gaussian diffusion $(\bar{\mathbf{X}}_u)_{u\in[t,t+\varepsilon]}$ with dynamics:

$$(A.1) \qquad d\bar{\mathbf{X}}_u = D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))\bar{\mathbf{X}}_u du + B\sigma(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))dW_u.$$

The associated covariance matrix between t and $t+\varepsilon$ writes:

$$(A.2) \qquad \bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}} = \int_t^{t+\varepsilon} \bar{\mathbf{R}}^{t,\mathbf{x}}(t+\varepsilon,u)B\sigma\sigma^*(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))B^*\bar{\mathbf{R}}^{t,\mathbf{x}}(t+\varepsilon,u)^* du,$$

where $(\bar{\mathbf{R}}^{t,\mathbf{x}}(v,u))_{t\le u,v\le t+\varepsilon}$ stands for the resolvent associated with

$$(D\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{x})))_{u\in[t,t+\varepsilon]}.$$

We point out that $(\bar{\mathbf{R}}^{t,\mathbf{x}}(v,u))_{t\le u,v\le t+\varepsilon}$ and $(\tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(v,u))_{t\le u,v\le t+\varepsilon}$ are *similar resolvents*, in the sense that they actually only differ in the flow considered in the linear dynamics. The flow is forward for $\bar{\mathbf{R}}^{t,\mathbf{x}}$ and backward for $\tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}$.

Observe now that, from **(H)**, $\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}}$ satisfies the *good scaling property* (2.10). Write now:

$$\int_{\mathbb{R}^{nd}} \tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})f(\mathbf{y})d\mathbf{y}$$

$$= \left[\int_{\mathbb{R}^{nd}} \left\{ \frac{\exp\left(-\frac{1}{2}\left\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})-\mathbf{y}),\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})-\mathbf{y}\right\rangle\right)}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}} \right.\right.$$

$$\left.\left. -\frac{\exp\left(-\frac{1}{2}\left\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}),\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}\right\rangle\right)}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}} \right\}f(\mathbf{y})d\mathbf{y}\right]$$

$$+\left[\int_{\mathbb{R}^{nd}} \left\{ \frac{\exp\left(-\frac{1}{2}\left\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}),\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}\right\rangle\right)}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}} \right.\right.$$

$$\left.\left. -\frac{\exp\left(-\frac{1}{2}\left\langle(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}),\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}\right\rangle\right)}{(2\pi)^{\frac{nd}{2}}\det(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}} \right\}f(\mathbf{y})d\mathbf{y}\right]$$

$$+\left[\int_{\mathbb{R}^{nd}} \frac{\exp\left(-\frac{1}{2}\left\langle(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}),\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}\right\rangle\right)}{(2\pi)^{\frac{nd}{2}}\det(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}}f(\mathbf{y})d\mathbf{y}\right]$$

$$(A.3) =: \sum_{i=1}^{3} \boldsymbol{\Xi}_i^\varepsilon(t,\mathbf{x}).$$

It is directly seen from the dominated convergence theorem that $\boldsymbol{\Xi}_3^\varepsilon(t,\mathbf{x}) \underset{\varepsilon\downarrow 0}{\to} f(\mathbf{x})$.

It remains to prove that $\boldsymbol{\Xi}_1^\varepsilon(t,\mathbf{x}),\boldsymbol{\Xi}_2^\varepsilon(t,\mathbf{x})$ can be viewed as remainders as $\varepsilon\downarrow 0$.

Let us write

$$(A.4) \quad |\boldsymbol{\Xi}_1^\varepsilon(t,\mathbf{x})| \quad \le \quad \|f\|_\infty \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(2\pi)^{\frac{nd}{2}}\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}} \int_0^1 d\lambda |(\varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda)|,$$

where $\lambda \in [0,1]$:

$$
\begin{aligned}
\varphi_{t,\mathbf{x},\mathbf{y}}^{\varepsilon}(\lambda) \;=\; & \exp\Bigg(-\frac{1}{2}\Big\{\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}),\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}\rangle + \\
& \lambda\Big[\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})-\mathbf{y}),\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})-\mathbf{y}\rangle \\
& -\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}),\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}\rangle\Big]\Big\}\Bigg), \\
|(\varphi_{t,\mathbf{x},\mathbf{y}}^{\varepsilon})'(\lambda)| \;\leq\; & \left|(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-\frac{1}{2}}\Big\{(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})+(\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})-\mathbf{y})\Big\}\right| \\
\text{(A.5)} \qquad & \times\left|(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-\frac{1}{2}}\Big\{(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})-(\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})-\mathbf{y})\Big\}\right|\varphi_{t,\mathbf{x},\mathbf{y}}^{\varepsilon}(\lambda),
\end{aligned}
$$

using the Cauchy-Schwarz inequality for the last assertion. A key quantity to control for the analysis is now the linearization error $|(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{x}})^{-\frac{1}{2}}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}))|$. From (2.10) we readily have:

$$
\text{(A.6)} \qquad \left|(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{x}})^{-\frac{1}{2}}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}))\right| \leq C\varepsilon^{\frac{1}{2}}|\mathbb{T}_{\varepsilon}^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}))|.
$$

To bound the above r.h.s. we first introduce for $\mathbf{z}\in\mathbb{R}^{nd}$, $u\in[t,t+\varepsilon]$,

$$
\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u,\mathbf{z}) := \big(F_1(u,\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})),F_2(u,z_1,(\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))^{2,n}),
$$
$$
\text{(A.7)} \qquad F_3(u,z_2,(\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))^{3,n}),\cdots,F_n(s,z_{n-1},(\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))_n)\big).
$$

We then write:

$$
\begin{aligned}
& \varepsilon^{\frac{1}{2}}\mathbb{T}_{\varepsilon}^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})) \\
& := \varepsilon^{\frac{1}{2}}\mathbb{T}_{\varepsilon}^{-1}\Bigg\{\int_t^{t+\varepsilon}du\Bigg[\bigg(\mathbf{F}(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))-\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))\bigg) \\
& \quad +\bigg(D\mathbf{F}(u,\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\tilde{\boldsymbol{\theta}}_{u,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}))\bigg) \\
& \quad +\bigg(\int_0^1 d\lambda\,\big(D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u,\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})+\lambda(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))) \\
& \quad -D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u,\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))\big)\,(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))\bigg)\Bigg]\Bigg\} \\
\text{(A.8)} \qquad & := (\mathcal{I}_{t+\varepsilon,t}^1+\mathcal{I}_{t+\varepsilon,t}^2+\mathcal{I}_{t+\varepsilon,t}^3)(\mathbf{x},\mathbf{y}),
\end{aligned}
$$

where, according to the notations of (A.7), for $(u,\mathbf{z})\in[t,t+\varepsilon]\times\mathbb{R}^{nd}$, $D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u,\mathbf{z})$ is the $(nd)\times(nd)$ matrix with only non zero $d\times d$ matrix entries $(D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u,\mathbf{z}))_{j,j-1}$ $:= D_{x_{j-1}}F_j(u,z_{j-1},\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})^{j,n})$, $j\in[\![2,n]\!]$. In particular $D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u,\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})) = D\mathbf{F}(u,\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))$.

Observe now that, from **(S)**:

$$
\begin{aligned}
|\mathcal{I}_{t+\varepsilon,t}^3(\mathbf{x},\mathbf{y})| \;\leq\; & C\sum_{i=2}^n\int_t^{t+\varepsilon}du\varepsilon^{-(i-1/2)}|(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))_{i-1}|^{1+\eta} \\
\;\leq\; & C\int_t^{t+\varepsilon}du\varepsilon^{-1+\eta/2}\Big(\varepsilon^{1/2}|\mathbb{T}_{\varepsilon}^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))|\Big)^{1+\eta}.
\end{aligned}
$$

From Lemma 5 (almost equivalence of the flows) we now derive:

$$
\begin{aligned}
|\mathcal{I}^3_{t+\varepsilon,t}(\mathbf{x},\mathbf{y})| &\leq C\int_t^{t+\varepsilon} du\, \varepsilon^{-1+\eta/2}\Big(\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|+1\Big)^{1+\eta}\\
&\leq C\varepsilon^{\eta/2}\Big(\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|+1\Big)^{1+\eta}.
\end{aligned}
$$

(A.9)

Let us now deal with $\mathcal{I}^1_{t+\varepsilon,t}(\mathbf{x},\mathbf{y})$. From the previous definition of $\mathbf{F}^{t+\varepsilon,\mathbf{y}}$ in (A.7), the key idea is to use the sub-linearity of $\mathbf{F}$ and the appropriate Hölder exponents. Namely, using the Young inequality we derive:

$$
\begin{aligned}
|\mathcal{I}^1_{t+\varepsilon,t}(\mathbf{x},\mathbf{y})| &\leq C\sum_{i=1}^n \varepsilon^{-i+1/2}\sum_{j=i}^n \int_t^{t+\varepsilon} du\,|(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))_j|^{\beta_i^j}\\
&\leq C\Bigg(\varepsilon^{-1/2}\int_t^{t+\varepsilon} du\Big(|(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))|+1\Big)\\
&\quad+\sum_{i=2}^n \varepsilon^{-i+1/2}\sum_{j=i}^n \int_t^{t+\varepsilon} du\Big\{\Big(\frac{|((\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))_j)|}{\varepsilon^{\gamma_i^j}}\Big)+\varepsilon^{\gamma_i^j\frac{\beta_i^j}{1-\beta_i^j}}\Big\}\Bigg),
\end{aligned}
$$

for some parameters $\gamma_i^j>0$ to be specified below. Hence,

$$
\begin{aligned}
|\mathcal{I}^1_{t+\varepsilon,t}(\mathbf{x},\mathbf{y})| &\leq C\Bigg(\int_t^{t+\varepsilon} du\, \varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))|+\varepsilon^{1/2}\\
&\quad+\sum_{i=2}^n\sum_{j=i}^n \int_t^{t+\varepsilon} du\Big\{\varepsilon^{-i+j-\gamma_i^j}\Big(\frac{|((\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))_j)|}{\varepsilon^{j-1/2}}\Big)+\varepsilon^{-i+1/2+\gamma_i^j\frac{\beta_i^j}{1-\beta_i^j}}\Big\}\Bigg)\\
&\leq C\Bigg(\int_t^{t+\varepsilon} du\, \varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))|+\varepsilon^{1/2}\\
&\quad+\sum_{i=2}^n\sum_{j=i}^n \int_t^{t+\varepsilon} du\Big\{\varepsilon^{-i+j-\gamma_i^j}\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))|+\varepsilon^{-i+1/2+\gamma_i^j\frac{\beta_i^j}{1-\beta_i^j}}\Big\}\Bigg).
\end{aligned}
$$

We now use Lemma 5 to derive

$$
\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))|\leq C(\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|+1).
$$

We emphasize here that in our current framework we should *a priori* write $\boldsymbol{\theta}_{t+\varepsilon,u}(\boldsymbol{\theta}_{u,t}(\mathbf{x}))$ in the above equation since we do not have a priori the flow property. Anyhow, since Lemma 5 is valid for any flow starting from $\boldsymbol{\theta}_{u,t}(\mathbf{x})$ at time u associated with the ODE (see equation (2.15)) we can proceed along the previous one, i.e. $(\boldsymbol{\theta}_{v,t}(\mathbf{x}))_{v\in[u,t+\varepsilon]}$. This yields:

(A.10)

$$
|\mathcal{I}^1_{t+\varepsilon,t}(\mathbf{x},\mathbf{y})|
$$
$$
\leq C\Big[\varepsilon^{1/2}+(\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|+1)\varepsilon\Big(1+\sum_{i=2}^n\sum_{j=i}^n\Big\{\varepsilon^{-i+j-\gamma_i^j}+\varepsilon^{-i+1/2+\gamma_i^j\frac{\beta_i^j}{1-\beta_i^j}}\Big\}\Big)\Big].
$$

Choose now for $i \in [\![2, n]\!]$ and $j \in [\![i, n]\!]$,

$$-i + j - \gamma_i^j = -i + 1/2 + \gamma_i^j \frac{\beta_i^j}{1 - \beta_i^j} \iff \gamma_i^j = (j - \frac{1}{2})(1 - \beta_i^j),$$

to balance the two previous contributions associated with the indexes i, j. To obtain a global smoothing effect w.r.t ε in (A.10) we need to impose:

$$-i + j - \gamma_i^j > -1 \iff \beta_i^j > \frac{2i - 3}{2j - 1}.$$

Hence, under (1.3), we have that there exists $\zeta := \zeta\big((\mathbf{A}), (\beta_i^j)_{i \in [\![1,n]\!], j \in [\![i,n]\!]}\big) \in (0, 1)$ s.t.:

$$(\text{A.11}) \qquad |\mathcal{I}_{t+\varepsilon,t}^1(\mathbf{x}, \mathbf{y})| \leq C\varepsilon^\zeta\big(1 + \varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|\big).$$

We now get from (A.6), (A.8), (A.9), (A.11) and the Gronwall lemma that:

$$\left|(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{x}})^{-\frac{1}{2}}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}))\right|$$

$$\leq C\varepsilon^{\frac{1}{2}}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}))|$$

$$(\text{A.12}) \quad \leq C\varepsilon^{\eta/2 \wedge \zeta}\big(1 + \varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + (\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|)^{1+\eta}\big).$$

Hence, recalling from (2.12) and Lemma 5 that

$$\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}) - \mathbf{y}), \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}) - \mathbf{y}\rangle \geq C^{-1}\varepsilon|\mathbb{T}_\varepsilon^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,t+\varepsilon}(\mathbf{y}))|^2$$

$$\geq C^{-1}(\varepsilon|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|^2 - 1),$$

we get from the definition in (A.5) that for all $\lambda \in [0, 1]$:

$$\varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon(\lambda) \leq C\exp(-C^{-1}\varepsilon|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|^2).$$

We finally obtain from (A.5) and (A.12) that there exists $C_2 := C_2(T, (\mathbf{A})) \geq 1$ s.t.:

$$(\text{A.13}) \quad |(\varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda)| \leq C_2\varepsilon^{\eta/2 \wedge \zeta}\exp\left(-C_2^{-1}\varepsilon|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|^2\right).$$

Plugging Equation (A.13) into (A.4), we derive that, since $\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}$ satisfies (2.10), $|\Xi_1^\varepsilon(t, \mathbf{x})| \underset{\varepsilon \downarrow 0}{\to} 0$.

Let us now consider the term $\Xi_2^\varepsilon(t, \mathbf{x})$ in (A.3). Write first $\Xi_2^\varepsilon(t, \mathbf{x}) := (\Xi_{21}^\varepsilon + \Xi_{22}^\varepsilon)(t, \mathbf{x})$ where:

$$\Xi_{21}^\varepsilon(t, \mathbf{x}) := \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(2\pi)^{\frac{nd}{2}}} \frac{f(\mathbf{y})}{\det(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}} \int_0^1 d\lambda (\psi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda),$$

$$\forall \lambda \in [0, 1], \ \psi_{t,\mathbf{x},\mathbf{y}}^\varepsilon(\lambda) := \exp\left(-\frac{1}{2}\bigg\{\langle(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}\rangle\right.$$

$$+ \lambda\bigg[\langle(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}\rangle$$

$$\left.-\langle(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}\rangle\bigg]\bigg\}\right),$$

$$(\text{A.14}) \qquad \Xi_{22}^\varepsilon(t, \mathbf{x}) := \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(2\pi)^{\frac{nd}{2}}} f(\mathbf{y})\left[\frac{1}{\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}} - \frac{1}{\det(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}}\right](\psi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)(1).$$

Observe that for all $\lambda \in [0,1]$,

$$|(\psi^{\varepsilon}_{t,\mathbf{x},\mathbf{y}})'(\lambda)| \leq \left|\langle((\tilde{\mathbf{K}}^{t+\varepsilon,\mathbf{y}}_{t+\varepsilon,t})^{-1} - (\bar{\mathbf{K}}^{t,\mathbf{x}}_{t+\varepsilon,t})^{-1})(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}\rangle\right| \psi^{\varepsilon}_{t,\mathbf{x},\mathbf{y}}(\lambda).$$

Equation (2.10), which holds for $\bar{\mathbf{K}}^{t,\mathbf{x}}_{t+\varepsilon,t}$ as well, yields:

$$\begin{aligned}
|(\psi^{\varepsilon}_{t,\mathbf{x},\mathbf{y}})'(\lambda)| &\leq C\left|\langle((\tilde{\mathbf{K}}^{t+\varepsilon,\mathbf{y}}_{t+\varepsilon,t})^{-1} - (\bar{\mathbf{K}}^{t,\mathbf{x}}_{t+\varepsilon,t})^{-1})(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}\rangle\right| \\
&\quad \times \exp(-C\varepsilon|\mathbb{T}^{-1}_{\varepsilon}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|^2) \\
\text{(A.15)} &=: C|Q_{\varepsilon}| \exp(-C\varepsilon|\mathbb{T}^{-1}_{\varepsilon}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|^2),
\end{aligned}$$

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

Now, the covariance matrices explicitly write

$$\begin{aligned}
\tilde{\mathbf{K}}^{t+\varepsilon,\mathbf{y}}_{t+\varepsilon,t} &= \int_t^{t+\varepsilon} du\, \tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(t+\varepsilon,u) Ba(u, \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})) B^* \tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(t+\varepsilon,u)^*, \\
\bar{\mathbf{K}}^{t,\mathbf{x}}_{t+\varepsilon,t} &= \int_t^{t+\varepsilon} du\, \bar{\mathbf{R}}^{t,\mathbf{x}}(t+\varepsilon,u) Ba(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) B^* \bar{\mathbf{R}}^{t,\mathbf{x}}(t+\varepsilon,u)^*,
\end{aligned}$$

where $\tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}$, $\bar{\mathbf{R}}^{t,\mathbf{x}}$ respectively denote the resolvents associated with the linear parts of equations (2.2) and (A.1). Thus, setting:

$$\text{(A.16)} \qquad \tilde{\mathbf{K}}^{t+\varepsilon,\mathbf{y}}_{t+\varepsilon,t} = \varepsilon^{-1}\mathbb{T}_{\varepsilon}\widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t+\varepsilon,\mathbf{y}}_1 \mathbb{T}_{\varepsilon}, \; \bar{\mathbf{K}}^{t,\mathbf{x}}_{t+\varepsilon,t} = \varepsilon^{-1}\mathbb{T}_{\varepsilon}\widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t,\mathbf{x}}_1 \mathbb{T}_{\varepsilon},$$

we write:

$$\begin{aligned}
&\left|\langle(\tilde{\mathbf{K}}^{t+\varepsilon,\mathbf{y}}_{t+\varepsilon,t} - \bar{\mathbf{K}}^{t,\mathbf{x}}_{t+\varepsilon})(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}\rangle\right| \\
&= \left|\langle(\widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t+\varepsilon,\mathbf{y}}_1 - \widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t,\mathbf{x}}_1)(\varepsilon^{-1/2}\mathbb{T}_{\varepsilon}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})), \varepsilon^{-1/2}\mathbb{T}_{\varepsilon}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})\rangle\right| \\
&\leq C|\widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t+\varepsilon,\mathbf{y}}_1 - \widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t,\mathbf{x}}_1| \times \varepsilon^{-1}|\mathbb{T}_{\varepsilon}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|^2.
\end{aligned}$$

It remains to control the term

$$\widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t+\varepsilon,\mathbf{y}}_1 - \widehat{\tilde{\mathbf{K}}}^{t+\varepsilon,t,t,\mathbf{x}}_1 = (\Delta^{t+\varepsilon,t}_1 - \Delta^{t+\varepsilon,t}_2)(\mathbf{x},\mathbf{y}),$$

$$\Delta^{t+\varepsilon,t}_1(\mathbf{x},\mathbf{y}) := \varepsilon \int_t^{t+\varepsilon} du\, \mathbb{T}^{-1}_{\varepsilon}\tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(t+\varepsilon,u) B\Delta a(u,t+\varepsilon)(\mathbf{x},\mathbf{y}) B^* \tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(t+\varepsilon,u)^* \mathbb{T}^{-1}_{\varepsilon},$$

$$\Delta a(u,t+\varepsilon)(\mathbf{x},\mathbf{y}) = \Big(a(u,\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})) - a(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))\Big),$$

$$\Delta^{t+\varepsilon,t}_2(\mathbf{x},\mathbf{y}) := \varepsilon \int_t^{t+\varepsilon} du\, \mathbb{T}^{-1}_{\varepsilon}\Delta\tilde{\tilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x},\mathbf{y}}(t+\varepsilon,u) Ba(u,\boldsymbol{\theta}_{u,t}(\mathbf{x})) B^* \bar{\mathbf{R}}^{t,\mathbf{x}}(t+\varepsilon,u)^* \mathbb{T}^{-1}_{\varepsilon}$$

$$-\varepsilon \int_t^{t+\varepsilon} du\Big\{\mathbb{T}^{-1}_{\varepsilon}\tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(t+\varepsilon,u) Ba(u,\boldsymbol{\theta}_{u,t}(\mathbf{x})) B^* \Big(\Delta\tilde{\tilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x},\mathbf{y}}(t+\varepsilon,u)\Big)^* \mathbb{T}^{-1}_{\varepsilon}\Big\},$$

(A.17)

$$\Delta\tilde{\tilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x},\mathbf{y}}(t+\varepsilon,u) = \Big(\bar{\mathbf{R}}^{t,\mathbf{x}}(t+\varepsilon,u) - \tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(t+\varepsilon,u)\Big).$$

From the scaling properties of the resolvent, see e.g. Lemma 6.2 in [Men18] for details, we have that:

$$\text{(A.18)} \quad \bar{\mathbf{R}}^{t,\mathbf{x}}(t+\varepsilon,u) = \mathbb{T}_{\varepsilon}\widehat{\tilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x}}_1 \mathbb{T}^{-1}_{\varepsilon}, \; \tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(t+\varepsilon,u) = \mathbb{T}_{\varepsilon}\widehat{\tilde{\mathbf{R}}}^{t+\varepsilon,u,t+\varepsilon,\mathbf{y}}_1 \mathbb{T}^{-1}_{\varepsilon},$$

where $\widehat{\mathbf{R}}_1^{t+\varepsilon,u,t,\mathbf{x}}$, $\widehat{\mathbf{R}}_1^{t+\varepsilon,u,t+\varepsilon,\mathbf{y}}$ are non-degenerate bounded matrices, uniformly in $u \in [t, t+\varepsilon]$. Hence, from (A.18) and the definitions in (A.17):

$$|\Delta_1^{t+\varepsilon,t}(\mathbf{x},\mathbf{y})|$$

$$(\text{A.19}) \quad \leq C\varepsilon^{-1}\int_t^{t+\varepsilon}|\Delta a(u,t+\varepsilon)(\mathbf{x},\mathbf{y})|du \leq C\varepsilon^{-1}\int_t^{t+\varepsilon}|\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})|^\eta du$$

$$\leq C\varepsilon^{\eta/2-1}\int_t^{t+\varepsilon}du|\varepsilon^{1/2}\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x})-\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))|^\eta$$

$$\leq C\varepsilon^{\eta/2}(|\varepsilon^{1/2}\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|^\eta+1),$$

using again Lemma 5 for the last inequality. Still from (A.18), the definitions in (A.17) and recalling as well that $\mathbb{T}_\varepsilon^{-1}Ba(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))B^*\mathbb{T}_\varepsilon^{-1} = \varepsilon^{-2}Ba(u,\boldsymbol{\theta}_{u,t}(\mathbf{x}))B^*$, we now write that:

$$(\text{A.20}) \quad |\Delta_2^{t+\varepsilon,t}(\mathbf{x},\mathbf{y})| \leq C\varepsilon^{-1}\int_t^{t+\varepsilon}|\mathbb{T}_\varepsilon^{-1}\Delta\widetilde{\widetilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x},\mathbf{y}}(t+\varepsilon,u)\mathbb{T}_\varepsilon|du.$$

Note then:

$$|\mathbb{T}_\varepsilon^{-1}\Delta\widetilde{\widetilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x},\mathbf{y}}(t+\varepsilon,u)\mathbb{T}_\varepsilon|$$

$$= \left|\mathbb{T}_\varepsilon^{-1}\int_u^{t+\varepsilon}\left(D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))\bar{\mathbf{R}}^{t,\mathbf{x}}(v,u)-D\mathbf{F}(v,\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))\tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(v,u)\right)dv\mathbb{T}_\varepsilon\right|$$

$$\leq \int_u^{t+\varepsilon}|\mathbb{T}_\varepsilon^{-1}D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))\mathbb{T}_\varepsilon||\mathbb{T}_\varepsilon^{-1}\Delta\widetilde{\widetilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x},\mathbf{y}}(v,u)\mathbb{T}_\varepsilon|dv$$

$$+ \int_u^{t+\varepsilon}\left|\mathbb{T}_\varepsilon^{-1}\left(D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))-D\mathbf{F}(v,\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))\right)\mathbb{T}_\varepsilon\right||\mathbb{T}_\varepsilon^{-1}\tilde{\mathbf{R}}^{t+\varepsilon,\mathbf{y}}(v,t)\mathbb{T}_\varepsilon|dv$$

$$\leq C\int_u^{t+\varepsilon}\varepsilon^{-1}|D\mathbf{F}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))-D\mathbf{F}(v,\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))|dv,$$

using the Gronwall lemma and the structure of the resolvent for the last inequality.

Pay attention that we only know from **(S)** that for all $i \in [\![2,n]\!], \forall\mathbf{z}^{i:n} = (z_i,\cdots,z_n) \in \mathbb{R}^{(n-i+1)d}$, $z_{i-1} \mapsto D_{x_{i-1}}F_i(z_{i-1},\mathbf{z}^{i:n})$ is $C^\eta(\mathbb{R}^d,\mathbb{R}^d\otimes\mathbb{R}^d)$-Hölder continuous for $\eta > 0$. We thus have to handle the above term with some care. Write with the notations of (A.7):

$$|\mathbb{T}_\varepsilon^{-1}\Delta\widetilde{\widetilde{\mathbf{R}}}^{t+\varepsilon,u,t,\mathbf{x},\mathbf{y}}(t+\varepsilon,u)\mathbb{T}_\varepsilon|$$

$$\leq C\int_u^{t+\varepsilon}\varepsilon^{-1}\sum_{i=2}^n\left(|D_{x_{i-1}}F_i(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))-D_{x_{i-1}}F_i^{t+\varepsilon,\mathbf{y}}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))|\right.$$

$$\left.+|D_{x_{i-1}}F_i^{t+\varepsilon,\mathbf{y}}(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))-D_{x_{i-1}}F_i(v,\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))|\right)dv$$

$$\leq C\int_u^{t+\varepsilon}\varepsilon^{-1}\sum_{i=2}^n\left(|D_{x_{i-1}}F_i(v,\boldsymbol{\theta}_{v,t}(\mathbf{x}))-D_{x_{i-1}}F_i(v,\boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1},(\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))^{i:n})|\right.$$

$$\left.+|(\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))_{i-1}|^\eta\right)dv =: (R_1+R_2)(t+\varepsilon,u,t,\mathbf{x},\mathbf{y}).$$

We get $|R_2(t+\varepsilon,u,t,\mathbf{x},\mathbf{y})| \leq C\int_u^{t+\varepsilon}\varepsilon^{-1}|\boldsymbol{\theta}_{v,t}(\mathbf{x})-\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y})|^\eta dv$ which can be handled similarly to (A.19). This yields:

$$|R_2(t+\varepsilon,u,t,\mathbf{x},\mathbf{y})| \leq C\varepsilon^{\eta/2}\left(\varepsilon^{1/2}|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|+1\right)^\eta.$$

On the other hand, using a reverse Taylor expansion, for positive parameters $(\delta_i)_{i \in [\![2,n]\!]}$ to be specified:

$$|R_1(t+\varepsilon, u, t, \mathbf{x}, \mathbf{y})|$$

$$\leq C \int_u^{t+\varepsilon} \varepsilon^{-1} \sum_{i=2}^n \Big(\Big| \big\{ F_i\big(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1} + \delta_i, (\boldsymbol{\theta}_{v,t}(\mathbf{x}))^{i:n}\big) - F_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) \big\}$$

$$- \big\{ F_i\big(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1} + \delta_i, (\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))^{i:n}\big) - F_i\big(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1}, (\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))^{i:n}\big) \big\} \Big| \delta_i^{-1} + \delta_i^\eta \Big) dv$$

$$\leq C \int_u^{t+\varepsilon} \varepsilon^{-1} \sum_{i=2}^n \Big(\sum_{j=i}^n |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))_j|^{\beta_i^j} \delta_i^{-1} + \delta_i^\eta \Big)$$

$$\leq C \Big\{ \sum_{i=2}^n \sum_{j=i}^n \Big(|\varepsilon^{1/2}(\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}))|^{\beta_i^j} + 1 \Big) \varepsilon^{(j-1/2)\beta_i^j} \delta_i^{-1} + \max_{i \in [\![2,n]\!]} \delta_i^\eta \Big\},$$

using again Lemma 5 for the last inequality. For this contribution to be a remainder it therefore suffices to choose $\delta_i = \max_{j \in [\![i,n]\!]} \varepsilon^{(j-1/2)\beta_i^j - \gamma}$, for $\gamma > 0$ small enough. From the above computations we eventually derive that there exists $\zeta' := \zeta'((\mathbf{A}), (\beta_i^j)_{i \in [\![1,n]\!], j \in [\![i,n]\!]}) \in (0,1)$ s.t.

$$|\mathbb{T}_\varepsilon^{-1} \Delta \widetilde{\widehat{\mathbf{R}}}^{t+\varepsilon, u, t, \mathbf{x}, \mathbf{y}}(t+\varepsilon, u) \mathbb{T}_\varepsilon| \leq C \varepsilon^{\frac{\eta}{2} \wedge \zeta'} \big(\varepsilon^{\frac{1}{2}} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + 1 \big).$$

Plugging this bound into (A.20), we then derive from (A.19) and (A.17) that:

$$(\text{A.21}) \qquad |\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} - \widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t, \mathbf{x}}| \leq C \varepsilon^{\frac{\eta}{2} \wedge \zeta'} \big(\varepsilon^{\frac{1}{2}} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + 1 \big).$$

Recalling now that $\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}}$, $\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t, \mathbf{x}}$ are because of (2.10) and (A.16) non-degenerate uniformly w.r.t. the parameter ε, we deduce that the inverse matrices $\big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} \big)^{-1}, \big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t, \mathbf{x}} \big)^{-1}$ have the same Hölder regularity. Indeed,

$$\big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} \big)^{-1} - \big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t, \mathbf{x}} \big)^{-1}$$

$$= \big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t, \mathbf{x}} \big)^{-1} \big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t, \mathbf{x}} - \widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} \big) \big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} \big)^{-1},$$

and we eventually conclude from (A.21). Hence, from the definition in (A.15)

$$|Q_\varepsilon| := \Big| \langle ((\widetilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1} - (\bar{\mathbf{K}}_{t+\varepsilon,}^{t,\mathbf{x}})^{-1})(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y} \rangle \Big|$$

$$= \Big| \langle \big(\big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} \big)^{-1} - \big(\widehat{\widetilde{\mathbf{K}}}_1^{t+\varepsilon, t, t, \mathbf{x}} \big)^{-1} \big) (\varepsilon^{1/2} \mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}))$$

$$, \varepsilon^{1/2} \mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}) \rangle \Big|$$

$$\leq C \varepsilon^{\eta/2 \wedge \zeta'} \big(\varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + 1 \big) \varepsilon |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|^2,$$

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

We eventually get:

$$|\Xi_{21}^{\varepsilon}(t,\mathbf{x})|$$

$$\leq C\varepsilon^{\eta/2\wedge\zeta'}\int_{\mathbb{R}^{nd}}\frac{d\mathbf{y}}{\varepsilon^{\frac{n^2d}{2}}}(\varepsilon^{1/2}|\mathbb{T}_{\varepsilon}^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|+1)\exp(-C\varepsilon|\mathbb{T}_{\varepsilon}^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|^2)$$

(A.22)

$$\leq C\varepsilon^{\eta/2\wedge\zeta'},\ C:=C((\mathbf{A}),T).$$

This yields $|\Xi_{21}^{\varepsilon}(t,\mathbf{x})|\underset{\varepsilon\to0}{\to}0$. Arguments similar to those employed for $\Xi_{21}^{\varepsilon}(t,\mathbf{x})$ can be used to prove that for the term $\Xi_{22}^{\varepsilon}(t,\mathbf{x})$ defined in (A.14). Namely,

$$|\Xi_{22}^{\varepsilon}(t,\mathbf{x})| \leq \|f\|_{\infty}\int_{\mathbb{R}^{nd}}\frac{d\mathbf{y}}{(2\pi)^{\frac{nd}{2}}}\left|\frac{1}{\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}}-\frac{1}{\det(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}}\right|(\psi_{t,\mathbf{x},\mathbf{y}}^{\varepsilon})(1)$$

$$\leq C\|f\|_{\infty}\int_{\mathbb{R}^{nd}}d\mathbf{y}\left|1-\frac{\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}}{\det(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}}\right|\bar{p}_{C^{-1}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})$$

(A.23) $$\leq C\|f\|_{\infty}\int_{\mathbb{R}^{nd}}d\mathbf{y}\left|1-\det\left(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}\right)^{\frac{1}{2}}\right|\bar{p}_{C^{-1}}(t,t+\varepsilon,\mathbf{x},\mathbf{y}).$$

Using again (A.16), we now write:

$$\det\left(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}\right) = \det\left(\widehat{\widehat{\mathbf{K}}}_1^{t+\varepsilon,t,t+\varepsilon,\mathbf{y}}(\widehat{\widehat{\mathbf{K}}}_1^{t+\varepsilon,t,t,\mathbf{x}})^{-1}\right)$$

$$= \det\left(I+(\widehat{\widehat{\mathbf{K}}}_1^{t+\varepsilon,t,t+\varepsilon,\mathbf{y}}-\widehat{\widehat{\mathbf{K}}}_1^{t+\varepsilon,t,t,\mathbf{x}})(\widehat{\widehat{\mathbf{K}}}_1^{t+\varepsilon,t,t,\mathbf{x}})^{-1}\right).$$

Plugging this identity into (A.23), we thus derive from (A.21) that $\Xi_{22}^{\varepsilon}(t,\mathbf{x})\underset{\varepsilon\to0}{\to}0$. The proof is complete. $\square$

A.2. **Proof of Lemma 7.** Up to approximation argument, we can suppose without loss of generality that $f\in C_0^{1,2}([0,T)\times\mathbb{R}^{nd},\mathbb{R})$. Write for $p',q'>1$,

$$\|f_{\varepsilon}-f\|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))}^{q'} = \int_0^T\|f_{\varepsilon}(t,\cdot)-f(t,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}^{q'}dt.$$

Note that, up to a middle point type argument, the indicator part in the very definition of f_{ε} can be easily dealt. With a slight abuse of notation, we thus start from the following expression for f_{ε}:

$$\forall(t,\mathbf{x})\in[0,T]\times\mathbb{R}^{nd},\quad\int_{\mathbb{R}^{nd}}f(t+\varepsilon,\mathbf{y})\tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}.$$

Now,

$$\|f_{\varepsilon}(t,\cdot)-f(t,\cdot)\|_{L^{p'}(\mathbb{R}^{nd})}^{p'}$$

$$=\int_{\mathbb{R}^{nd}}\left|\int_{\mathbb{R}^{nd}}f(t+\varepsilon,\mathbf{y})\tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}-f(t,\mathbf{x})\right|^{p'}d\mathbf{x}$$

$$\leq 2^{p'-1}\Big(\int_{\mathbb{R}^{nd}}\left|\int_{\mathbb{R}^{nd}}f(t+\varepsilon,\mathbf{y})\tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}-f(t+\varepsilon,\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}))\right|^{p'}d\mathbf{x}$$

$$+\int_{\mathbb{R}^{nd}}|f(t+\varepsilon,\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}))-f(t,\mathbf{x})|^{p'}d\mathbf{x}\Big)=:2^{p'-1}(I_1^{\varepsilon}(t)+I_2^{\varepsilon}(t)).$$

Recalling that f is smooth and has compact support in time and space, we readily get $I_2^\varepsilon(t) \underset{\varepsilon \to 0}{\longrightarrow} 0$. Let us now turn to $I_1^\varepsilon(t)$. Write:

$$I_1^\varepsilon(t)$$

$$\leq 2^{p'-1}\Big(\int_{\mathbb{R}^{nd}} \Big|\int_{\mathbb{R}^{nd}} \Big(f(t+\varepsilon,\mathbf{y}) - f(t+\varepsilon,\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}))\Big)\tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}\Big|^{p'} d\mathbf{x}$$

$$+ \int_{\mathbb{R}^{nd}} \Big|f(t+\varepsilon,\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}))\int_{\mathbb{R}^{nd}} \big(\tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{x},\mathbf{y}) - \bar{p}(t,t+\varepsilon,\mathbf{x},\mathbf{y})\big)d\mathbf{y}\Big|^{p'} d\mathbf{x}\Big)$$

$$=: I_{11}^\varepsilon(t) + I_{12}^\varepsilon(t),$$

where $\bar{p}(t,t+\varepsilon,\mathbf{x},\mathbf{y}) := \frac{\exp\big(-\frac{1}{2}\big\langle (\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}),\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}\big\rangle\big)}{(2\pi)^{\frac{nd}{2}}\det(\bar{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}}$ is a *true* density in $\mathbf{y}$ and already appears in the term $\boldsymbol{\Xi}_3^\varepsilon$ defined in (A.3) in the proof of Lemma 6. The previous analysis of this term readily gives $I_{12}^\varepsilon(t) \underset{\varepsilon \to 0}{\longrightarrow} 0$. On the other hand, from the good scaling property of equation (2.10), Lemma 5 and the Hölder inequality, there exists $C := C(p',(\mathbf{A}))$ s.t.:

$$I_{11}^\varepsilon(t) \leq C \int_{\mathbb{R}^{nd}} \int_{\mathbb{R}^{nd}} \Big|f(t+\varepsilon,\mathbf{y}) - f(t+\varepsilon,\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}))\Big|^{p'} \bar{p}_{C^{-1}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}$$

$$\leq C \int_{\mathbb{R}^{nd}} \big(\mathbb{I}_{d(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}),\text{supp}(f(t+\varepsilon,\cdot))) \leq \delta} + \mathbb{I}_{d(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}),\text{supp}(f(t+\varepsilon,\cdot))) > \delta}\big)$$

$$\times \int_{\mathbb{R}^{nd}} \Big|f(t+\varepsilon,\mathbf{y}) - f(t+\varepsilon,\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}))\Big|^{p'} \bar{p}_{C^{-1}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}$$

$$\leq C\Big(\varepsilon^{p'/2}\|Df\|_\infty + \int_{\mathbb{R}^{nd}} \mathbb{I}_{d(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}),\text{supp}(f(t+\varepsilon,\cdot))) > \delta}$$

$$\times \int_{\mathbb{R}^{nd}} |f(t+\varepsilon,\mathbf{y})|^{p'} \bar{p}_{C^{-1}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}\Big)$$

$$\leq C\Big(\varepsilon^{p'/2}\|Df\|_\infty + \|f\|_\infty \exp\Big(-C^{-1}\frac{\delta^2}{\varepsilon}\Big)\int_{\mathbb{R}^{nd}} \mathbb{I}_{\mathbf{y}\in\text{supp}(f(t+\varepsilon,\cdot))}$$

$$\times \int_{\mathbb{R}^{nd}} \bar{p}_{C^{-1}}(t,t+\varepsilon,\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}\Big)$$

$$\leq C\Big(\varepsilon^{p'/2}\|Df\|_\infty + \|f\|_\infty \exp\Big(-C^{-1}\frac{\delta^2}{\varepsilon}\Big)\Big),$$

where in the above computations we recall that $\bar{p}_{C^{-1}}$ introduced in (2.8) is actually a density w.r.t. $\mathbf{x}$. We conclude the proof from the above convergence of $I_1^\varepsilon(t)$, $I_2^\varepsilon(t)$ thanks to the dominated convergence theorem.

Acknowledgments.

References

[BFGM19] L. Beck, F. Flandoli., M. Gubinelli, and M. Maurelli. Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electronic Journal of Probability*, 24:♯136, 2019.

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

[CC18]   G. Cannizzaro and K. Chouk. Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential. *Annals of Probability*, 46–3:1710–1763, 2018.

[CdR17]   P.-E. Chaudru de Raynal. Strong existence and uniqueness for degenerate SDE with Höder drift. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53(1):259–286, February 2017.

[CdR18]   P.-E. Chaudru de Raynal. Weak regularization by stochastic drift: result and counter example. *Discrete and Continuous Dynamical Systems (Series A)*, 38–3:1269–1291, 2018.

[CG16]   R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of ODEs. *Stoch. Proc and Appl.*, 126–8:2323–2366, 2016.

[Dav07]   A. M. Davie. Uniqueness of solutions of stochastic differential equations. *International Mathematics Research Notices. IMRN*, (24):Art. ID rnm124, 26, 2007.

[DD15]   F. Delarue and R. Diel. Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probability Theory and Related Fields*, pages 1–63, April 2015.

[DF14]   F. Delarue and F. Flandoli. The transition point in the zero noise limit for a 1d Peano example. *Discrete and Continuous Dynamical Systems*, 34(10):4071–4083, April 2014.

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

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

[FFPV17]   E. Fedrizzi, F. Flandoli, E. Priola, and J. Vovelle. Regularity of stochastic kinetic equations. *Electronic Journal of Probability*, 22:♯ 48, 2017.

[FIR17]   F. Flandoli, E. Issoglio, and F. Russo. Multidimensional stochastic differential equations with distributional drift. *Transactions of the American Mathematical Society*, 369:1665–1688, 2017.

[Fla11]   F. Flandoli. *Random perturbation of PDEs and fluid dynamic models*, volume 2015 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010.

[GO13]   Mihai Gradinaru and Yoann Offret. Existence and asymptotic behaviour of some time-inhomogeneous diffusions. *Annales Instit. H. Poincaré, Probabilités et Statistiques*, 49(1):182D207, 2013.

[HM16]   L. Huang and S. Menozzi. A Parametrix Approach for some Degenerate Stable Driven SDEs. *Annales Instit. H. Poincaré*, 52(4):1925–1975, 2016.

[Hör67]   L. Hörmander. Hypoelliptic second order differential operators. *Acta. Math.*, 119:147–171, 1967.

[Kha59]   R.Z. Khasminskii. On positive solutions of the equation $au + vu = 0$. *Theoret. Probab. Appl.*, 4:309D318, 1959.

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

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

[LP94]   E. Lanconelli and S. Polidoro. On a class of hypoelliptic evolution operators. volume 52, pages 29–63. 1994. Partial differential equations, II (Turin, 1993).

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

[Men18]   S. Menozzi. Martingale problems for some degenerate Kolmogorov equations. *Stoc. Proc. Appl.*, 128-3:756–802, 2018.

[MS67]   H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1:43–69, 1967.

[PL89]   R. Di Perna and P. L. Lions. Ordinary Differential Equations, transport theory and Sobolev Spaces. *Inv. Math.*, 98:511–547, 1989.

[Por90]   N.I. Portenko. *Generalized Diffusion Processes*. English translation: Amer. Math. Soc., Providence, Rhode Island, 1990.

[Pri18]   E. Priola. Davie's type uniqueness for a class of SDEs with jumps. *Ann. Inst. Henri Poincaré*, 54(2):694–725, 2018.

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

[Szn98]   A. S. Sznitman. *Obstacles and random media*. Springer Monographs in Mathematics, 1998.

[Ver80]   A. Y. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Matematicheski Sbornik. Novaya Seriya*, 111(153)(3):434–452, 480, 1980.

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

[Zha10]   X. Zhang. Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Revista Matematica Iberoamericana*, 29(1), 2010.

[Zha18]   X. Zhang. Stochastic Hamiltonian flows with singular coefficients. *Sci. China Math.*, 61(8):1353–1384, 2018.

[Zub12]   O. Zubelevich. Several Notes on Existence Theorem of Peano. *Funkcialaj Ekvacioj*, 55(1):89–97, 2012.

[Zvo74]   A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

UMR CNRS 6629, Laboratoire de Mathématiques Jean Leray, 2, rue de la Houssinière 44322 Nantes cedex 3, France.
  Email address: `Paul-Eric.Deraynal@univ-nantes.fr`

UMR CNRS 8071, Laboratoire de Modélisation Mathématique d'Evry (LaMME), Université d'Evry Val d'Essonne, Paris-Saclay, 23 Boulevard de France 91037 Evry, France and Laboratory of Stochastic Analysis, HSE University, Pokrovsky Blvd, 11, Moscow, Russian Federation
  Email address: `stephane.menozzi@univ-evry.fr`