



**HAL**  
open science

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

Paul-Éric Chaudru de Raynal, Stéphane Menozzi

## ► To cite this version:

Paul-Éric Chaudru de Raynal, Stéphane Menozzi. Regularization effects of a noise propagating through a chain of differential equations: an almost sharp result. Transactions of the American Mathematical Society, 2021, 375 (01), pp.1-45. 10.1090/tran/7947 . hal-01613679v1

**HAL Id: hal-01613679**

**<https://hal.science/hal-01613679v1>**

Submitted on 9 Oct 2017 (v1), last revised 31 Oct 2021 (v3)

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

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

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

Paul-Éric Chaudru de Raynal\* and Stéphane Menozzi†

October 9, 2017

## 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.

**Keywords:** regularization by noise, martingale problem, Kolmogorov hypoelliptic PDEs, density estimates, parametrix.

**MSC:** Primary: 60H10, 34F05; Secondary: 60H30.

## 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:

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

In the above equation,  $(W_t)_{t \geq 0}$  stands for a  $d$ -dimensional Brownian motion and the components  $(X_t^i)_{i \in \llbracket 1, n \rrbracket}$  are  $\mathbb{R}^d$ -valued as well. We suppose that the  $(F_i)_{i \in \llbracket 2, n \rrbracket}$  satisfy a kind of weak Hörmander condition, i.e. the matrices  $(D_{x_{i-1}} F_i(t, \cdot))_{i \in \llbracket 2, n \rrbracket}$  have full rank. However, the coefficients  $(F_i)_{i \in \llbracket 2, n \rrbracket}$

---

\*Univ. Savoie Mont Blanc, CNRS, LAMA, F-73000 Chambéry, France. pe.deraynal@univ-smb.fr

†Laboratoire de Modélisation Mathématique d'Evry (LaMME), Université d'Evry Val d'Essonne, 23 Boulevard de France 91037 Evry, France and Laboratory of Stochastic Analysis, HSE, Shabolovka 31, Moscow, Russian Federation. stephane.menozzi@univ-evry.fr

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  $\sigma$  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  $i^{\text{th}}$  component only comes from the  $(i - 1)^{\text{th}}$  component,  $2 \leq i \leq n$ , through the non-degeneracy of the gradients  $(D_{x_{i-1}} F_i(t, \cdot))_{i \in \llbracket 2, n \rrbracket}$  (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  $i^{\text{th}}$  component w.r.t. the  $j^{\text{th}}$  variable we assume  $\beta_i^j \in ([ (2i - 3) / (2j - 1) ], 1]$ . 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 estimate.

## 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<sup>1</sup>.

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 problem* 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

---

<sup>1</sup>In the presentation below, we will mainly focus on Brownian driven SDEs. We can refer to the recent work of Priola [Pri17] for the more general Lévy driven case in the non-degenerate framework.

degenerate SDE 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 [CC15] to the multidimensional setting.

All the above strong and weak results strongly 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 ODE

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

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 phenomenon 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 extreme 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, 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 stage of the chain we only require one additional Lie bracket to generate the corresponding directions. This setting allows us to recover some regularization effect of the noise at each stage of the chain.

Concerning the strong regularization effects of a degenerate noise, 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 [Zha16].

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 [Men17] 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 [Zha16] that when the degenerate component only depends linearly of 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 [CdR16] 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 [CdR16] 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  $\sigma(t, \cdot), (F_i(t, \cdot))_{i \in \llbracket 2, n \rrbracket}$  are assumed to be bounded measurable in time. Also, the diffusion coefficient  $\sigma(t, \cdot)$  is uniformly  $\eta$ -Hölder continuous in space, for some  $\eta > 0$  uniformly in time. The drift entries  $(F_i(t, \cdot))_{i \in \llbracket 2, n \rrbracket}$  are s.t. for all  $(z_i, \dots, z_n) \in \mathbb{R}^{d(n-(i-1))}$ , the mapping  $z \in \mathbb{R}^d \mapsto F_i(t, z, z_i, \dots, z_n)$  is in  $C^{1+\eta}(\mathbb{R}^d, \mathbb{R}^d)$  uniformly in time<sup>2</sup>. Eventually, the mappings  $(z_i, \dots, z_n) \in \mathbb{R}^{d(n-(i-1))} \mapsto F_i(t, z, z_i, \dots, z_n)$  are  $\beta_i^j$ -Hölder continuous, 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 be bounded measurable in time and to satisfy spatially one of the following assumptions:

- (a) to be such that  $F_1(t, \cdot)$  is Hölder continuous in space<sup>3</sup> uniformly in time.
- (b)  $F_1(t, \cdot)$  is measurable bounded uniformly in time.
- (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  $(\mathbf{x}_{i-1}, \dots, \mathbf{x}_n) \in \mathbb{R}^{(n-i+2)d}$ ,  $D_{\mathbf{x}_{i-1}} F_i(t, \mathbf{x}_{i-1}, \dots, \mathbf{x}_n) \in \mathcal{E}_{i-1}$ . For example,  $\mathcal{E}_i, i \in \llbracket 1, n-1 \rrbracket$ , 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:*

$$\forall i \in \llbracket 2, n \rrbracket, j \in \llbracket i, n \rrbracket, \beta_i^j \in \left( \frac{2i-3}{2j-1}, 1 \right]. \quad (1.3)$$

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

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

*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).*

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

<sup>3</sup>Actually one can assume that  $F_1$  is  $\beta_1^j$  Hölder continuous in the  $j^{\text{th}}$  variable for any  $\beta_1^j$  in  $(0, 1]$ .

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  $t \in (s, T]$  and all  $\Gamma \in \mathcal{B}(\mathbb{R}^{nd})$ :  $P(t, s, \mathbf{x}, \Gamma) = \int_{\Gamma} p(s, t, \mathbf{x}, \mathbf{y}) d\mathbf{y}$ .

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  $\frac{n^2 d}{p'} + \frac{2}{q'} < 2, p' > 1, q' > 1, (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

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

where  $\mathbb{E}^{\mathbb{P}_{t, \mathbf{x}}}$  denotes the expectation w.r.t.  $\mathbb{P}_{t, \mathbf{x}}[\cdot] := \mathbb{P}[\cdot | \mathbf{X}_t = \mathbf{x}]$  and  $C := C((\mathbf{A}), p', q', T)$ .

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 [Men17] 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 2 below together with condition (1.3)) should allow one to adapt the approach of [Men17] 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 threshold 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:*

$$\exists i \in \llbracket 2, n \rrbracket, j \in \llbracket i, n \rrbracket, \beta_i^j < \frac{2i - 3}{2j - 1}, \quad (1.6)$$

or

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

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

We first emphasize that there are already some results in that direction: in [BFGM14] 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], [CC15], 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 [CdR16], that for all  $i$  in  $\llbracket 2, n \rrbracket$  the Hölder exponents  $\beta_i^i$ , 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  $(D_{x_{i-1}} F_i(t, \cdot))_{i \in \llbracket 2, n \rrbracket}$  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 2, 3, 4 and 6). 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.

## 2 Strategy and key tools

Our strategy relies on the martingale problem approach. Hence, we face two problems: 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 [Men17].

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 to a hypoelliptic 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 restriction. It is actually more direct than the approach of Stroock and Varadhan. However, its drawback is that it does not provide *a priori* informations 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* process 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  $\mathbf{F}$  is *smooth* in addition to **(A)**, say globally Lipschitz continuous, it has been shown in [DM10], [Men11], [Men17] 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 [CdR16], 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 Transform 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 stage 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.  $\mathbf{0}$ , stands for zero in  $\mathbb{R}^{nd}$  and the solution  $(X_t^1, \dots, X_t^n)_{t \geq 0}$  to (1.1) is denoted by  $(\mathbf{X}_t)_{t \geq 0}$ . Introducing the embedding matrix  $B$  from  $\mathbb{R}^d$  into  $\mathbb{R}^{nd}$ , i.e.  $B = (I_{d \times d}, 0, \dots, 0)^*$ , where “\*” stands for the transpose, we rewrite (1.1) in the shortened form

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

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

**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:

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

**Remark 1.** *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

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

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

$$D\mathbf{F}(s, \mathbf{z}) = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ D_{\mathbf{z}_1} F_2(s, \mathbf{z}) & 0 & \dots & \dots & 0 \\ 0 & D_{\mathbf{z}_2} F_3(s, \mathbf{z}) & 0 & 0 & \vdots \\ \vdots & 0 & \ddots & \vdots & \\ 0 & \dots & 0 & D_{\mathbf{z}_{n-1}} F_n(s, \mathbf{z}) & 0 \end{pmatrix}$$

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  $(s, t) \in (\mathbb{R}_+)^2$ :

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

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  $(\tilde{\mathbf{X}}_s^{T, \mathbf{y}})_{s \in [t, T]}$ :

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

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

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

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:

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

An important correspondence is now given by the following Proposition.

**Proposition 1** (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:*

$$\forall \mathbf{z} \in \mathbb{R}^{nd}, \quad \tilde{p}^{T, \mathbf{y}}(t, v, \mathbf{x}, \mathbf{z}) := \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{v, t}^{T, \mathbf{y}})^{\frac{1}{2}}} \exp \left( -\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{v, t}^{T, \mathbf{y}})^{-1} (\tilde{\boldsymbol{\theta}}_{v, t}^{T, \mathbf{y}}(\mathbf{x}) - \mathbf{z}), \tilde{\boldsymbol{\theta}}_{v, t}^{T, \mathbf{y}}(\mathbf{x}) - \mathbf{z} \right\rangle \right), \quad (2.7)$$

where

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

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

$$\begin{aligned} |D_{x_i}^k \tilde{p}^{T, \mathbf{y}}(t, T, \mathbf{x}, \mathbf{y})| &\leq \frac{C}{(T-t)^{k((i-1)+\frac{1}{2})+\frac{n^2 d}{2}}} \exp\left(-C^{-1}(T-t)|\mathbb{T}_{T-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y}))|^2\right) \\ &=: \frac{C}{(T-t)^{k((i-1)+\frac{1}{2})}} \bar{p}_{C^{-1}}(t, T, \mathbf{x}, \mathbf{y}), \end{aligned} \quad (2.8)$$

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

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

*Proof.* Expression (2.7) readily follows from (2.6). We recall as well that, under  $(\mathbf{A})$ , 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  $\mathbf{y} \in \mathbb{R}^{nd}$ :

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

Rewrite now from (2.5) and (2.6):

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

where we denote:

$$\tilde{\boldsymbol{\theta}}_{t,T}^{T, \mathbf{y}}(\mathbf{y}) = \tilde{\mathbf{R}}^{T, \mathbf{y}}(t, T)\mathbf{y} - \int_t^T \tilde{\mathbf{R}}^{T, \mathbf{y}}(t, s) \left( \mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y})) - D\mathbf{F}(s, \boldsymbol{\theta}_{s,T}(\mathbf{y}))\boldsymbol{\theta}_{s,T}(\mathbf{y}) \right) ds. \quad (2.12)$$

Observe that (2.12) also solves (2.1). Indeed, from (2.3):

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

so that:

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

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 and equations (2.11), (2.10) that there exists  $C := C((\mathbf{A}), T) > 0$ , s.t. for all  $t \in [0, T)$ ,

$$\begin{aligned} C^{-1}(T-t)|\mathbb{T}_{T-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y}))|^2 &\leq \left\langle (\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})^{-1}(\tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y}), \tilde{\boldsymbol{\theta}}_{T,t}^{T,\mathbf{y}}(\mathbf{x}) - \mathbf{y} \right\rangle \\ &\leq C(T-t)|\mathbb{T}_{T-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y}))|^2. \end{aligned} \quad (2.13)$$

We then deduce from the above equation and (2.10) that (2.8) holds for  $k = 0$ .

Introduce now  $\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}} = \tilde{\mathbf{R}}^{T,\mathbf{y}}(t, T)\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}}\tilde{\mathbf{R}}^{T,\mathbf{y}}(t, T)^*$ . We point out that, by construction, since  $(D\mathbf{F}(v, \boldsymbol{\theta}_{v,T}(\mathbf{y})))_{v \in [t, T]}$  is strictly subdiagonal,  $\det(\tilde{\mathbf{R}}^{T,\mathbf{y}}(t, T)) = 1$ . Hence,  $\det(\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}}) = \det(\tilde{\mathbf{K}}_{T,t}^{T,\mathbf{y}})$ . We then deduce from (2.7):

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

We derive from (2.10):

$$|D_{x_i}\tilde{p}^{T,\mathbf{y}}(t, T, \mathbf{x}, \mathbf{y})| \leq \frac{C|(\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1/2})_{i,\cdot}|}{(T-t)^{\frac{n^2d}{2}}} \exp\left(-\frac{C^{-1}}{2} \left\langle (\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y})), \mathbf{x} - \boldsymbol{\theta}_{t,T}(\mathbf{y}) \right\rangle\right),$$

where  $((\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1/2})_{i,\cdot}$  stands for the  $i^{\text{th}}$  line of a square root of  $(\tilde{\mathbf{H}}_{T,t}^{T,\mathbf{y}})^{-1}$ . Still by (2.10), we deduce that:

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

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

## 2.2 Regularity and measurability of the Cauchy-Peano flow

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: it is non-trivial that one can choose a measurable flow of (2.1) (which is very important to make licit any integration of this flow along the terminal condition) and that this flow has the appropriate regularity to deal with our parametrix kernel, say e.g. bi-Lipschitz as in [DM10], [Men11], [Men17].

To solve the first issue we propose two approaches : the first one consists in assuming the Axiom of choice. This permits to choose a measurable flow of (2.1) (see Lemma 1 below). The second one does not involve the Axiom of choice and relies on compactness arguments to prove that there exists a measurable Cauchy-Peano flow satisfying (2.1) almost everywhere. Using a “weak” solution of (2.1) does not damage or even modify our study, but we have to be more careful when making pointwise estimates of quantities involving such a flow. Especially, we will need to switch from this measurable flow defined almost everywhere to another Cauchy Peano flow defined everywhere to complete our task. Hence, in order to keep things as clear as possible, we first assume the Axiom of choice and

give in Appendix B the procedure to remove that assumption and use the measurable flow of (2.1) defined almost everywhere.

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  $\tilde{p}$  given in (2.7). This leads us to study space integral of the Gaussian like function  $\tilde{p}_{C^{-1}}$  defined by (2.8) w.r.t. the backward variable. The crucial point is that such an integral is then done along the backward flow and that it is not clear that it is still bounded. To avoid this problem the main idea consists in using the bi-Lipschitz property to transfer the flow from the integration variable to the initial (and then fixed) variable. Here, we were not able to prove such a 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 2 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 1.** *For a given  $T > 0$  and all  $(t, s) \in [0, T]^2$ , there exists a measurable mapping  $\mathbf{x} \in \mathbb{R}^{nd} \mapsto \boldsymbol{\theta}_{t,s}(\mathbf{x})$  s.t.  $\boldsymbol{\theta}_{t,s}(\mathbf{x}) = \mathbf{x} + \int_s^t \mathbf{F}(v, \boldsymbol{\theta}_{v,s}(\mathbf{x})) dv$ .*

*Proof.* Assuming the Axiom of choice, the proof follows from the result of [Zub12] and usual covering arguments.  $\square$

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 2.** *There exists a constant  $C_2 := C_2((\mathbf{A}), T)$  s.t. for all  $0 \leq t < s \leq T$  small enough:*

$$C_2\{(s-t)|\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 - 1\} \leq (s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2 \leq C_2\{(s-t)|\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 + 1\}. \quad (2.14)$$

Lemma 2 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], [Men17] for Lipschitz drifts) up to an additional constant contribution. It turns out that, this new contribution does not perturb the analysis of the parametrix kernels associated with the density of  $\mathbf{X}_T^{T,\mathbf{y}}$  starting from  $\mathbf{x}$  at time  $t \in [0, T]$  given in Proposition 1. We refer to Section 3.1 for details.

*Proof.* Considering the flow provided by Lemma 1, we write from the differential dynamics:

$$\begin{aligned} (s-t)^{1/2}\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y})) &= (s-t)^{1/2}\mathbb{T}_{s-t}^{-1}\left[(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y}) - \int_t^s \left(\mathbf{F}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y}))\right) du\right]. \\ &= (s-t)^{1/2}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y}) - \mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y}), \\ \mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y}) &= (s-t)^{1/2}\mathbb{T}_{s-t}^{-1} \int_t^s \left(\mathbf{F}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y}))\right) du. \end{aligned} \quad (2.15)$$

We aim at establishing that

$$|\mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y})| \leq C \left\{ |(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})| + 1 \right\}, \quad (2.16)$$

which together with (2.15) gives the r.h.s. of (2.14). The l.h.s. could be derived similarly to the analysis we now perform.

Since the function  $\mathbf{F}$  is not Lipschitz, 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  $\delta_{ij}$  is strictly positive for indexes  $i \in \llbracket 2, n \rrbracket$  and  $j \geq i$ . We then define for all  $v \in [0, T]$ ,  $\mathbf{z} \in \mathbb{R}^{nd}$ ,  $i \in \llbracket 2, n \rrbracket$ ,

$$\mathbf{F}_i^\delta(v, \mathbf{z}^{i-1, n}) := \mathbf{F}_i(v, \cdot) \star \rho_{\delta_{i,\cdot}}(\mathbf{z}) = \int_{\mathbb{R}^{d(n-(i-1))}} \mathbf{F}_i(v, \mathbf{z}_{i-1}, \mathbf{z}_i - w_i, \dots, \mathbf{z}_n - w_n) \rho_{\delta_{i,\cdot}}(w) dw. \quad (2.17)$$

Here, for all  $w = (w_i, \dots, w_n) \in \mathbb{R}^{d(n-(i-1))}$ ,  $\rho_{\delta_{i,\cdot}}(w) := \frac{1}{\prod_{j=i}^n \delta_{ij}^d} \rho_i\left(\frac{w_i}{\delta_{ii}}, \frac{w_{i+1}}{\delta_{i(i+1)}}, \dots, \frac{w_n}{\delta_{in}}\right)$  where  $\rho_i : \mathbb{R}^{d(n-(i-1))} \rightarrow \mathbb{R}_+$  is a standard mollifier, i.e.  $\rho_i$  has compact support and  $\int_{\mathbb{R}^{d(n-(i-1))}} \rho_i(\mathbf{z}) d\mathbf{z} = 1$ . We denote  $\mathbf{F}^\delta(v, \mathbf{z}) = (\mathbf{F}_1(v, \mathbf{z}), \mathbf{F}_2^\delta(v, \mathbf{z}), \dots, \mathbf{F}_n^\delta(v, \mathbf{z}))$  and consider the two auxiliary mollified dynamics:

$$\begin{aligned} \dot{\boldsymbol{\theta}}_{v,t}^\delta(\mathbf{x}) &= \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})), \quad v \in [t, s], \quad \boldsymbol{\theta}_{t,t}^\delta(\mathbf{x}) = \mathbf{x}, \\ \dot{\boldsymbol{\theta}}_{v,s}^\delta(\mathbf{y}) &= \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,s}^\delta(\mathbf{y})), \quad v \in [t, s], \quad \boldsymbol{\theta}_{s,s}^\delta(\mathbf{y}) = \mathbf{y}. \end{aligned} \quad (2.18)$$

**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 (2.15) and (2.25)). So, the parameters  $\delta_{ij}$  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  $[t, s]$ :

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

From the definition of our regularization procedure in (2.17) this means that  $\delta_{ij}$  must be such that

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

Hence, one can choose  $\delta_{ij} = (s-t)^{\frac{(i-3)}{2} \frac{1}{\beta_i^j}}$  which yields  $(s-t)^{\frac{1}{2}-i} \delta_{ij}^{\beta_j^i} = (s-t)^{-1}$ . This choice of  $\delta_{ij}$  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  $\mathbb{R}^{nd}$

$$\begin{aligned} & \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}^\delta(u, \mathbf{z}) - \mathbf{F}^\delta(u, \mathbf{z}') \right) \right| \\ & \leq C \left( (s-t)^{-\frac{1}{2}} + (s-t)^{-1} |(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\mathbf{z} - \mathbf{z}')| \right). \end{aligned} \quad (2.21)$$

Indeed, the  $(\mathbf{F}_i^\delta)_{i \in \llbracket 2, n \rrbracket}$  are Lipschitz continuous functions (with potentially explosive Lipschitz constant in the variables  $(x_i, \dots, x_n)$  for  $\mathbf{F}_i^\delta$  because of the regularization procedure) and  $\mathbf{F}_1$  is  $\beta_1^j > 0$

Hölder continuous in the  $j^{\text{th}}$  variable for arbitrary  $(\beta_1^j)_{j \in \llbracket 1, n \rrbracket}$  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,

$$\begin{aligned} \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}^\delta(v, \mathbf{z}) - \mathbf{F}^\delta(v, \mathbf{z}') \right) \right| &\leq C \left( (s-t)^{-\frac{1}{2}} (1 + |(\mathbf{z} - \mathbf{z}')|) \right. \\ &\quad \left. + \sum_{i=2}^n (s-t)^{\frac{1}{2}-i} \left( |(\mathbf{z} - \mathbf{z}')_{i-1}| + \sum_{j=i}^n \frac{|(\mathbf{z} - \mathbf{z}')_j|}{\delta_{ij}^{1-\beta_i^j}} \right) \right) \\ &\leq C \left( (s-t)^{-\frac{1}{2}} + |(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\mathbf{z} - \mathbf{z}')| \right. \\ &\quad \left. \times \left( 1 + (s-t)^{-1} + \sum_{i=2}^n \sum_{j=i}^n \frac{(s-t)^{j-i}}{\delta_{ij}^{1-\beta_i^j}} \right) \right). \end{aligned}$$

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

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

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

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

**Difference between the initial and mollified flow.** Write:

$$\begin{aligned} |(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}^\delta(\mathbf{x}))| &\leq \int_t^u \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})) \right) \right| dv \\ &\leq \int_t^u \left\{ \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) \right) \right| \right. \\ &\quad \left. + \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})) \right) \right| \right\} dv \end{aligned}$$

Using (2.19) for the first term in the above r.h.s. and (2.21) for the second term in the r.h.s. we obtain from the Gronwall lemma that:

$$\exists C := C((\mathbf{A}), T), \forall u \in [s, t], |(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{u,t}(\mathbf{x}) - \boldsymbol{\theta}_{u,t}^\delta(\mathbf{x}))| \leq C, \quad (2.23)$$

thanks to the time integration and recalling that  $u - t \leq s - t$ . We would derive similarly that  $|(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{u,s}(\mathbf{y}) - \boldsymbol{\theta}_{u,s}^\delta(\mathbf{y}))| \leq C$ .

Now for the regularized flow, we have a Lipschitz property of the rescaled flow, which is here explosive with  $\delta$  but once integrated does not yield any additional singularity. Namely,

$$\begin{aligned} &|(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{u,t}^\delta(\mathbf{x}) - \boldsymbol{\theta}_{u,s}^\delta(\mathbf{y}))| \\ &\leq |(s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}^\delta(\mathbf{x}) - \mathbf{y})| + \int_u^s dv \left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})) - \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,s}^\delta(\mathbf{y})) \right) \right|. \end{aligned}$$

From (2.21) and the Gronwall lemma we get:

$$\begin{aligned}
|(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{u,t}^\delta(\mathbf{x}) - \boldsymbol{\theta}_{u,s}^\delta(\mathbf{y}))| &\leq C|(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}^\delta(\mathbf{x}) - \mathbf{y})| \\
&\leq C\left\{|(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})| \right. \\
&\quad \left. + |(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}^\delta(\mathbf{x}))|\right\} \\
&\leq C\left\{|(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})| + 1\right\}, \tag{2.24}
\end{aligned}$$

using (2.23) for the last inequality.

**Derivation of the final bound.** We are now in position to bound the term  $\mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y})$  defined in (2.15). Under **(A)** we have:

$$\begin{aligned}
|\mathcal{I}_{s,t}(\mathbf{x}, \mathbf{y})| &\leq \int_t^s du |(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{F}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y})))| \\
&\leq \int_t^s du \left| (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{F}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,t}^\delta(\mathbf{x}))) \right| \\
&\quad + \int_t^s du \left| (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,t}^\delta(\mathbf{x})) - \mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,s}^\delta(\mathbf{y}))) \right| \\
&\quad + \int_t^s du \left| (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,s}^\delta(\mathbf{y})) - \mathbf{F}(u, \boldsymbol{\theta}_{u,s}(\mathbf{y}))) \right| \\
&=: \mathcal{I}_{s,t}^1(\mathbf{x}, \mathbf{y}) + \mathcal{I}_{s,t}^2(\mathbf{x}, \mathbf{y}) + \mathcal{I}_{s,t}^3(\mathbf{x}, \mathbf{y}). \tag{2.25}
\end{aligned}$$

Observe that the contribution  $\mathcal{I}_{s,t}^1(\mathbf{x}, \mathbf{y})$  can be split as

$$\begin{aligned}
\mathcal{I}_{s,t}^1(\mathbf{x}, \mathbf{y}) &= \int_t^s du \left| (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{F}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,t}^\delta(\mathbf{x}))) \right| \\
&\leq \int_t^s du \left| (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{F}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,t}(\mathbf{x}))) \right| \\
&\quad + \int_t^s du \left| (s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}^\delta(u, \boldsymbol{\theta}_{u,t}^\delta(\mathbf{x}))) \right|
\end{aligned}$$

Using (2.19) for the first term in the r.h.s. above and (2.21) together with (2.23) for the second term in the r.h.s., we obtain that there exists  $C := C((\mathbf{A}), T)$  s.t. for all  $0 \leq t < s \leq T$ ,  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{nd}$ :

$$|\mathcal{I}_{s,t}^1(\mathbf{x}, \mathbf{y})| \leq C. \tag{2.26}$$

Similarly, we derive that there exists  $C' := C'((\mathbf{A}), T)$  s.t. for all  $0 \leq t < s \leq T$ ,  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{nd}$ :

$$|\mathcal{I}_{s,t}^3(\mathbf{x}, \mathbf{y})| \leq C'. \tag{2.27}$$

Finally, one can use (2.21) and then (2.24) to derive that there exists  $C'' := C''((\mathbf{A}), T)$  s.t. for all  $0 \leq t < s \leq T$ ,  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{nd}$ :

$$\begin{aligned}
|\mathcal{I}_{s,t}^2(\mathbf{x}, \mathbf{y})| &\leq C \left[ (s-t)^{\frac{1}{2}} + \left\{ |(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})| + 1 \right\} \right] \\
&\leq C \left\{ |(s-t)^{\frac{1}{2}}\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})| + 1 \right\}.
\end{aligned}$$



From this last equation together with (2.26) and (2.27) we therefore derive (2.16). The proof is complete.  $\square$

### 2.3 Frozen Green kernels and associated PDEs

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

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

$$\forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \tilde{G}f(t, \mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}). \quad (2.28)$$

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  $\varphi \in C_0^\infty(\mathbb{R}^{nd}, \mathbb{R})$ ,  $\mathbf{x} \in \mathbb{R}^{nd}$ ,

$$\begin{aligned} \tilde{L}_t^{s, \mathbf{y}} \varphi(\mathbf{x}) &:= \langle \mathbf{F}(t, \boldsymbol{\theta}_{t, s}(\mathbf{y})) + D\mathbf{F}(t, \boldsymbol{\theta}_{t, s}(\mathbf{y}))(\mathbf{x} - \boldsymbol{\theta}_{t, s}(\mathbf{y})), D_{\mathbf{x}} \varphi(\mathbf{x}) \rangle \\ &\quad + \frac{1}{2} \text{tr}(\sigma \sigma^*(t, \boldsymbol{\theta}_{t, s}(\mathbf{y})) D_{\mathbf{x}_1}^2 \varphi(\mathbf{x})). \end{aligned} \quad (2.29)$$

One now easily checks that:

$$\forall (t, \mathbf{x}, \mathbf{z}) \in [0, s] \times (\mathbb{R}^{nd})^2, \left( \partial_t + \tilde{L}_t^{s, \mathbf{y}} \right) \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{z}) = 0. \quad (2.30)$$

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

**Lemma 3** (Dirac convergence of the frozen density). *For all bounded continuous function  $f : \mathbb{R}^{nd} \rightarrow \mathbb{R}$ ,  $\mathbf{x} \in \mathbb{R}^{nd}$ :*

$$\left| \int_{\mathbb{R}^{nd}} f(\mathbf{y}) \tilde{p}^{t+\varepsilon, \mathbf{y}}(t, t+\varepsilon, \mathbf{x}, \mathbf{y}) d\mathbf{y} - f(\mathbf{x}) \right| \xrightarrow{\varepsilon \downarrow 0} 0. \quad (2.31)$$

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. (2.30)). Indeed, the integration parameter is also the freezing parameter which makes things more subtle.

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

$$\tilde{M}_t f(t, \mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{L}_t^{s, \mathbf{y}} \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}), \quad (2.32)$$

we derive from (2.30) and Lemma 3 that the following equality holds:

$$\partial_t \tilde{G}f(t, \mathbf{x}) + \tilde{M}_t f(t, \mathbf{x}) = -f(t, \mathbf{x}), \quad \forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}. \quad (2.33)$$

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

$$|\tilde{G}f(t, \mathbf{x})| \leq C(T) \|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}. \quad (2.34)$$

*Proof.* From (2.8) with  $k = 0$  and Lemma 2 we have that

$$\begin{aligned}
|\tilde{G}f(t, \mathbf{x})| &\leq \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) \\
&\leq \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \frac{\exp(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2)}{(s-t)^{\frac{n^2d}{2}}} \\
&\leq \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \frac{\exp(-C^{-1}[(s-t)|\mathbb{T}_{s-t}^{-1}\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y}]^2 + 1))}{(s-t)^{\frac{n^2d}{2}}},
\end{aligned} \tag{2.35}$$

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

$$\begin{aligned}
&\int_t^T ds \left| \int_{\mathbb{R}^{nd}} d\mathbf{y} \left| \frac{\exp(-C^{-1}(s-t)(|\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})|^2))}{(s-t)^{\frac{n^2d}{2}}} \right| \right|^{\tilde{p}'} \\
&\leq C \int_t^T ds \frac{1}{(s-t)^{\frac{n^2d}{2}(\tilde{p}'-1)(\frac{\tilde{q}'}{p'})}} < +\infty \Leftrightarrow \frac{n^2d(\tilde{p}'-1)\tilde{q}'}{2\tilde{p}'} < 1 \Leftrightarrow \frac{n^2d}{2p'} + \frac{1}{q'} < 1.
\end{aligned}$$

□

### 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  $(L_t)_{t \geq 0}$  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  $\mathbf{F}$ , 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 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  $\mathbf{F}_1 = 0$  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 (3.7) 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 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. We then have to keep it in our Euler scheme: the drift is needed for the Brownian noise to propagate into the whole system. The dramatic point is that, contrary to [Men17], it is not clear that the strong convergence of the Euler scheme still holds, because of the lack of Lipschitz regularity of the drift  $\mathbf{F}$  and we are not able to localize it, due to our hypoelliptic structure. We then have to find another kind of regularization.

This is the reason why we decided to smooth the marginal law of our canonical process by means of Gaussian convolution (see the part of Section 3.1.2 following equation (3.5)). Doing this regularization allows us to recover the desired estimate established on our parametrix kernel and to propagate it to our original process (canonical process associated with the solution of the martingale problem). It seems to us that this, rather new, method is quite robust, in some sense more than the original one, and may be applied in other settings where the strong convergence of the Euler scheme does not hold anymore.

### 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  $\Lambda_m([0, T]) := \{(t_i := ih)_{i \in [0, m]}\}$ ,  $h = T/m$ . Introduce the corresponding “discretization” scheme:

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

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 Assumption (D)-(a)

We first want to establish that any solution to the martingale problem satisfies the suitable Krylov like density estimate of Theorem 1. To do so, 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) has small  $L^{q'} - L^{p'}$  norm when the fixed final horizon  $T$  is small. Namely, for

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

with  $\tilde{G}f, \tilde{M}f$  and  $\tilde{p}$  defined in (2.28), (2.32) and (2.7) respectively we have the following Lemma whose proof is postponed to the end of the current section:

**Lemma 4** ( $L^{q'} - L^{p'}$  Control for  $R$ ). *For  $R$  as in (3.2) and  $q', p' > 1$  s.t.  $\frac{n^2 d}{p'} + \frac{2}{q'} < 2$  we have that for all  $f \in L^{q'}([0, T] \times \mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))$ :*

$$\|Rf\|_{L^{q'}([0, T] \times \mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))} \leq C \|f\|_{L^{q'}([0, T] \times \mathbb{R}^{nd}, L^{p'}(\mathbb{R}^{nd}))}, \quad (3.3)$$

with  $C := C((\mathbf{A}), T, p', q') \xrightarrow{T \rightarrow 0} 0$ . In particular, equation (3.3) implies that the operator  $I - R$  is invertible, with bounded inverse in  $L^{q'} - L^{p'}$ , provided  $T$  is small enough.

In order to derive the previously mentioned Krylov-like estimate, we now need to perturb the canonical process  $(\mathbf{X}_t)_{t \in [0, T]}$  associated with  $\mathbb{P}$ . When doing so we have to be careful and to regularize the marginals of the law.

The first step of our procedure consists in applying Itô's formula on the Green kernel  $\tilde{G}f$ , for  $f$  in  $C_0^\infty([0, T] \times \mathbb{R}^{nd}, \mathbb{R})$  and the process  $\mathbf{X}_s^{t, \mathbf{x}}$ . However, it follows from the definition (2.28) of our Green kernel that the smoothness on  $f$  is not a sufficient condition to derive the smoothness of  $\tilde{G}f$ . This comes from the dependence of the covariance matrix of  $\tilde{p}$  w.r.t. the integration variable (see (2.7)). The strategy hence consists in mollifying the Green kernel with a suitable mollifier  $\rho_\delta$ , in applying Itô's formula on this regularized version and then show that the estimate can be obtained uniformly on the regularization procedure. This fact can be easily checked from the computations done below so that, in order to preserve the exposition of our proof as clear as possible, we do *as if* we could apply Itô's formula to  $\tilde{G}f$  whereas we should apply the previously exposed strategy.

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

where  $P_{\mathbf{X}_s^{t, \mathbf{x}}}$  denotes the law of  $\mathbf{X}_s^{t, \mathbf{x}}$ . We exploit (2.33) to write:

$$\tilde{G}f(t, \mathbf{x}) - \int_t^T \int_{\mathbb{R}^{nd}} f(s, \mathbf{y}) P_{\mathbf{X}_s^{t, \mathbf{x}}} (d\mathbf{y}) ds + \int_t^T \int_{\mathbb{R}^{nd}} (L_s \tilde{G}f - \tilde{M}f)(s, \mathbf{y}) P_{\mathbf{X}_s^{t, \mathbf{x}}} (d\mathbf{y}) ds = 0. \quad (3.4)$$

Recall that thanks to Proposition 2 we have that there exists  $C(T) := C((\mathbf{A}), T) \xrightarrow{T \rightarrow 0} 0$  s.t. for all  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$  and all  $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))$ :

$$|\tilde{G}f(t, \mathbf{x})| \leq C(T) \|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}. \quad (3.5)$$

We then regularize  $P_{\mathbf{X}_s^{t,\mathbf{x}}}$  through convolution considering  $P_{\mathbf{X}_s^{t,\mathbf{x}}}^\delta := P_{\mathbf{X}_s^{t,\mathbf{x}}} \star G_\delta$ , where  $G_\delta$  denotes the law of the standard Gaussian random variable on  $\mathbb{R}^{nd}$  with covariance matrix  $\delta I_{nd}$ . In particular  $P_{\mathbf{X}_s^{t,\mathbf{x}}}^\delta(d\mathbf{y}) = p^\delta(t, s, \mathbf{x}, \mathbf{y})d\mathbf{y}$ . From (3.4) we get

$$\tilde{G}f(t, \mathbf{x}) - \int_t^T \int_{\mathbb{R}^{nd}} f(s, \mathbf{y})p^\delta(t, s, \mathbf{x}, \mathbf{y})d\mathbf{y}ds + \int_t^T \int_{\mathbb{R}^{nd}} (L_s \tilde{G}f - \tilde{M}f)(s, \mathbf{y})p^\delta(t, s, \mathbf{x}, \mathbf{y})d\mathbf{y}ds = 0. \quad (3.6)$$

From (3.5), (3.4) and Lemma 4 we deduce that for all  $f \in C_0^\infty([0, T] \times \mathbb{R}^{nd}, \mathbb{R})$ :

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

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

$$\left| \int_t^T \int_{\mathbb{R}^{nd}} f(s, \mathbf{y})p^\delta(t, s, \mathbf{x}, \mathbf{y})d\mathbf{y}ds \right| = \left| \int_t^T \int_{\mathbb{R}^{nd}} f(s, \mathbf{y})P_{\mathbf{X}_s^{t,\mathbf{x}}}^\delta(d\mathbf{y})ds \right| \leq C(T)\|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}.$$

The Krylov like estimate (1.5) is then obtained by the bounded convergence theorem.

We have established that any solution of the martingale problem is s.t. for all  $f \in C_0^\infty([0, T] \times \mathbb{R}^{nd}, \mathbb{R})$ :

$$\left| \mathbb{E}^\mathbb{P} \left[ \int_t^T f(s, \mathbf{X}_s^{t,\mathbf{x}})ds \right] \right| \leq C(T)\|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}, \quad (3.7)$$

which is precisely the estimate (1.5) of Theorem 1.

Let now  $\mathbb{P}$  be any solution of the martingale problem and denote, thanks to (1.5), its density by  $p(t, s, \mathbf{x}, \mathbf{y})$ . We have that:

$$\begin{aligned} -\tilde{G}f(t, \mathbf{x}) &= -\mathbb{E}^\mathbb{P} \left[ \int_t^T f(s, \mathbf{X}_s^{t,\mathbf{x}})ds \right] + \mathbb{E} \left[ \int_t^T Rf(s, \mathbf{X}_s^{t,\mathbf{x}})ds \right] \\ &= -\int_t^T \int_{\mathbb{R}^{nd}} f(s, \mathbf{y})p(t, s, \mathbf{x}, \mathbf{y})d\mathbf{y}ds + \int_t^T \int_{\mathbb{R}^{nd}} Rf(s, \mathbf{y})p(t, s, \mathbf{x}, \mathbf{y})d\mathbf{y}ds \\ &= -\int_t^T \int_{\mathbb{R}^{nd}} (I - R)f(s, \mathbf{y})p(t, s, \mathbf{x}, \mathbf{y})d\mathbf{y}ds. \end{aligned} \quad (3.8)$$

Since both sides are continuous with respect to  $L^{q'} - L^{p'}$  norm, we conclude that (3.8) holds for all  $f \in L^{q'} - L^{p'}$ . We then conclude from Lemma 4 that:

$$\mathbb{E}^\mathbb{P} \left[ \int_t^T f(s, \mathbf{X}_s^{t,\mathbf{x}})ds \right] = \tilde{G} \circ (I - R)^{-1}f(t, \mathbf{x}), \quad (3.9)$$

which gives uniqueness for  $T$  small enough. Global well-posedness is again derived from a chaining in time argument. The proof of Theorem 1 is now complete under **(A)**-(a).

**Proof of Lemma 4** We have, by definition

$$\begin{aligned} Rf(t, \mathbf{x}) &= \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_t - \tilde{L}_t^{s, \mathbf{y}}) \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) \\ &=: \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} H(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}). \end{aligned} \quad (3.10)$$

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

The bound (2.8) of Proposition 1 now yields that there exists  $C := C(\mathbf{A})$  such that:

$$\begin{aligned} |Rf(t, \mathbf{x})| &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \left\{ \left| \mathbf{F}_1(t, \mathbf{x}) - \mathbf{F}_1(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) \right| |D_{x_1} \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y})| \right. \\ &\quad + \sum_{i=2}^n \left\{ \left| \mathbf{F}_i(t, \mathbf{x}) - \left( \mathbf{F}_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) - D_{x_{i-1}} \mathbf{F}_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) (\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_{i-1} \right) \right| \right. \\ &\quad \left. \left. \times |D_{x_i} \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y})| \right\} + |a(t, \mathbf{x}) - a(t, \boldsymbol{\theta}_{t,s}(\mathbf{y}))| |D_{x_1}^2 \tilde{p}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y})| \right\} \\ &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})| \left\{ \frac{(1 + |\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y})|)}{(s-t)^{\frac{1}{2}}} + \sum_{i=2}^n \left\{ \left| \mathbf{F}_i(t, \mathbf{x}) - \mathbf{F}_i(t, \mathbf{x}_{i-1}, \boldsymbol{\theta}_{t,s}(\mathbf{y}))^{i:n} \right| \right. \right. \\ &\quad \left. \left. + \left| \mathbf{F}_i(t, \mathbf{x}_{i-1}, \boldsymbol{\theta}_{t,s}(\mathbf{y}))^{i:n} - \left( \mathbf{F}_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) - D_{x_{i-1}} \mathbf{F}_i(t, \boldsymbol{\theta}_{t,s}(\mathbf{y})) (\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))_{i-1} \right) \right| \right\} \right. \\ &\quad \left. \times \frac{1}{(s-t)^{(i-1) + \frac{1}{2}}} \right\} + \frac{|\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y})|^\eta}{s-t} \left\} \frac{\exp(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2)}{(s-t)^{\frac{n^2 d}{2}}}. \end{aligned}$$

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

$$\begin{aligned} |Rf(t, \mathbf{x})| &\leq C \left\{ \sum_{i=2}^n \sum_{j=i}^n \int_t^T \frac{ds}{(s-t)^{(2i-1)/2 - \beta_i^j (2j-1)/2}} \int_{\mathbb{R}^{nd}} d\mathbf{y} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) |f(s, \mathbf{y})| \right. \\ &\quad + \int_t^T \frac{ds}{(s-t)^{1-\eta/2}} \int_{\mathbb{R}^{nd}} d\mathbf{y} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) |f(s, \mathbf{y})| \\ &\quad \left. + \sum_{i=2}^n \int_t^T \frac{ds}{(s-t)^{1-\eta(i-3/2)}} \int_{\mathbb{R}^{nd}} d\mathbf{y} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) |f(s, \mathbf{y})| \right\}. \end{aligned} \quad (3.11)$$

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

$$\int_{\mathbb{R}^{nd}} d\mathbf{y} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) = \int_{\mathbb{R}^{nd}} d\mathbf{y} \frac{\exp(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2)}{(s-t)^{\frac{n^2 d}{2}}} \leq C_{3.12}. \quad (3.12)$$

The result now follows from (3.11) and the following key Lemma.

**Lemma 5** ( $L^{q'} - L^{p'}$  Controls for the singularized Green kernel). *Introduce for  $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))$  and some  $\gamma \in (0, 1]$  the quantify:*

$$N_\gamma f(t, \mathbf{x}) := \int_t^T \frac{ds}{(s-t)^\gamma} \int_{\mathbb{R}^{nd}} d\mathbf{y} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) |f(s, \mathbf{y})|. \quad (3.13)$$

There exists  $C(T) := C(T, (\mathbf{A}), \gamma) \xrightarrow{T \rightarrow 0} 0$  s.t. for all  $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))$ ,

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

□

*Proof of Lemma 5.* 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 (3.13) and the triangle inequality:

$$\left\| N_\gamma f(t, \cdot) \right\|_{L^{p'}(\mathbb{R}^{nd})} = \left\| \int_t^T ds K_{t,s,\gamma} \star f(s, \cdot) \right\|_{L^{p'}(\mathbb{R}^{nd})} \leq \int_t^T ds \|K_{t,s,\gamma} \star f(s, \cdot)\|_{L^{p'}(\mathbb{R}^{nd})}. \quad (3.14)$$

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

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

Plugging this estimate into (3.14) and using the Hölder inequality with exponents  $q', \tilde{q}' > 1$  s.t.  $(q')^{-1} + (\tilde{q}')^{-1} = 1$  we obtain:

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

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

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

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

Hence, using the above control in (3.3) yields:

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

This conclude the proof. □

### 3.2 Existence and uniqueness under (D)-(b) and (c)

Since under the considered assumption 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  $\mathbf{x} \in \mathbb{R}^{nd}$ ,  $\bar{\mathbf{F}}(t, \mathbf{x}) := (\mathbf{0}, \mathbf{F}_2(t, \mathbf{x}), \dots, \mathbf{F}_n(t, \mathbf{x}))$ , where  $\bar{\mathbf{F}}$  satisfies (S) and (H). Recall from Theorem 1 under (A)-(a) that weak uniqueness holds for the SDE:

$$d\bar{\mathbf{X}}_t = \bar{\mathbf{F}}(t, \bar{\mathbf{X}}_t)dt + B\sigma(t, \bar{\mathbf{X}}_t)dW_t. \quad (3.15)$$

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

$$\dot{\bar{\boldsymbol{\theta}}}_{t,T}(\mathbf{y}) = \bar{\mathbf{F}}(t, \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})), \quad \bar{\boldsymbol{\theta}}_{T,T}(\mathbf{y}) = \mathbf{y}, \quad (3.16)$$

and

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

Again, in (3.16), we consider the Cauchy-Peano flow furnished by Lemma 1 and which also satisfies the equivalence of rescaled norms of Lemma 2. The dynamics of the linearized Gaussian process associated with (3.15) writes:

$$d\tilde{\mathbf{X}}_t^{T,\mathbf{y}} = \left( \bar{\mathbf{F}}(t, \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})) + D\bar{\mathbf{F}}(t, \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y}))(\tilde{\mathbf{X}}_t^{T,\mathbf{y}} - \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y})) \right) dt + B\sigma(t, \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y}))dW_t, \quad (3.17)$$

and we denote the associated generator by  $(\tilde{L}_t^{T,\mathbf{y}})_{t \in [0,T]}$  and by  $\tilde{p}^{T,\mathbf{y}}(t, s, \mathbf{x}, \cdot)$  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  $L$ . Once existence is known, the point is that we use a different Gaussian proxy than previously, namely the one considered in (3.17) 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  $(\tilde{\mathbf{X}}_s^{s,\mathbf{y}})_{s \in [t,T]}$  defined in (3.17) with  $\tilde{\mathbf{X}}_t^{s,\mathbf{y}} = \mathbf{x}$  as Gaussian proxys, we have to bound analogously to the estimate of Lemma 4:

$$\begin{aligned} Rf(t, \mathbf{x}) &:= \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_t - \tilde{L}_t^{s,\mathbf{y}}) \tilde{p}^{s,\mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) \\ &= \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\{ (L_t - \bar{L}_t^{s,\mathbf{y}}) + (\bar{L}_t - \tilde{L}_t^{s,\mathbf{y}}) \right\} \tilde{p}^{s,\mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) \\ &=: \bar{\mathcal{R}}f(t, \mathbf{x}) + \bar{R}f(t, \mathbf{x}), \end{aligned} \quad (3.18)$$



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

$$\begin{aligned} \bar{\mathcal{R}}f(t, \mathbf{x}) &:= (R - \bar{R})f(t, \mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_t - \bar{L}_t) \tilde{\bar{p}}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) \\ &= \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \langle \mathbf{F}_1(t, \mathbf{x}), D_{x_1} \tilde{\bar{p}}^{s, \mathbf{y}}(t, s, \mathbf{x}, \mathbf{y}) \rangle f(s, \mathbf{y}). \end{aligned} \quad (3.19)$$

From Lemma 4, 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 (3.19). Namely, our goal is to prove that

$$\|\bar{\mathcal{R}}f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))} \leq C(T) \|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}, \quad \text{with } C(T) \xrightarrow{T \rightarrow 0} 0. \quad (3.20)$$

Since from (2.8) and (3.19) we have for all  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

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

The estimate (3.20) then readily follows from Lemma 5.

### 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 (3.15).

**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 (3.19). The term  $\bar{R}f$  in (3.18) is again controlled as under assumption (A)-(a). Similarly to the previous paragraph (see also (2.8) and (3.19)), we have:

$$\begin{aligned} |D_{x_1} \tilde{\bar{p}}^{T, \mathbf{y}}(t, T, \mathbf{x}, \mathbf{y})| &\leq \frac{C}{(T-t)^{\frac{n^2d}{2} + \frac{1}{2}}} \times \exp\left(-\frac{C^{-1}}{2}(T-t)|\mathbb{T}_{T-t}^{-1}(\mathbf{x} - \bar{\boldsymbol{\theta}}_{t,T}(\mathbf{y}))|^2\right) \\ &\leq \frac{C}{(s-t)^{\frac{1}{2}}} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned} \quad (3.21)$$

Uniqueness then follows from the following lemma which can be viewed as a refinement of Lemma 5 and explicitly exploits the condition on  $p$  and  $q$  stated in (A)-(c).

**Lemma 6** (Refined  $L^q - L^p$  control of singularized Green kernels).

Introduce for all  $f, F_1 \in L^q([0, T], L^p(\mathbb{R}^{nd}))$  and  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ ,

$$u_{F_1}(t, \mathbf{x}) := F_1(t, \mathbf{x}) \int_t^T \frac{ds}{(s-t)^{\frac{1}{2}}} \int_{\mathbb{R}^{nd}} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) d\mathbf{y}. \quad (3.22)$$

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

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

□

*Proof of Lemma 6.* With the notations of Lemma 5 rewrite:

$$u_{F_1}(t, \mathbf{x}) = F_1(t, \mathbf{x}) \int_t^T \frac{ds}{(s-t)^{\frac{1}{2}}} \int_{\mathbb{R}^{nd}} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) d\mathbf{y} =: F_1(t, \mathbf{x}) \int_t^T ds K_{t, s, \frac{1}{2}} \star f(s, \mathbf{x}). \quad (3.23)$$

The triangle inequality yields:

$$\left\| \int_t^T ds F_1(t, \cdot) K_{t, s, \frac{1}{2}} \star f(s, \cdot) \right\|_{L^p(\mathbb{R}^{nd})} \leq \int_t^T ds \|F_1(t, \cdot) K_{t, s, \frac{1}{2}} \star f(s, \cdot)\|_{L^p(\mathbb{R}^{nd})}. \quad (3.24)$$

The idea is here to reproduce the computations of Lemma 5 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:

$$\begin{aligned} \|F_1(t, \cdot) K_{t, s, \frac{1}{2}} \star f(s, \cdot)\|_{L^p(\mathbb{R}^{nd})} &= \left( \int_{\mathbb{R}^{nd}} d\mathbf{x} |F_1(t, \mathbf{x})|^p \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{y} K_{t, s, \frac{1}{2}}(\mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) \right\}^p \right)^{\frac{1}{p}} \\ &\leq \left( \int_{\mathbb{R}^{nd}} d\mathbf{x} |F_1(t, \mathbf{x})|^p \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{y} |f(s, \mathbf{y})|^p \right\} \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{y} |K_{t, s, \frac{1}{2}}(\mathbf{x}, \mathbf{y})|^{\tilde{p}} \right\}^{\frac{p}{\tilde{p}}} \right)^{\frac{1}{p}} \\ &\leq \|F_1(t, \cdot)\|_{L^p(\mathbb{R}^{nd})} \|f(s, \cdot)\|_{L^p(\mathbb{R}^{nd})} \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{y} |K_{t, s, \frac{1}{2}}(\mathbf{x}, \mathbf{y})|^{\tilde{p}} \right\}^{\frac{1}{\tilde{p}}}, \end{aligned} \quad (3.25)$$

where again  $p^{-1} + (\tilde{p})^{-1} = 1$ . Observe now that usual Gaussian calculations give that:

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

which plugged into (3.25) yields from (3.24) that:

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

From the definition in (3.23) we eventually derive that:

$$\begin{aligned} \int_0^T dt \|u_{F_1}(t, \cdot)\|_{L^p(\mathbb{R}^{nd})}^q &\leq \int_0^T \|F_1(t, \cdot)\|_{L^p(\mathbb{R}^{nd})}^p \left( \int_t^T \frac{ds}{(s-t)^{\frac{n^2d}{2} \frac{1}{p} + \frac{1}{2}} \tilde{q}} \right)^{\frac{q}{\tilde{q}}} \int_t^T ds \|f(s, \cdot)\|_{L^p(\mathbb{R}^{nd})}^q \\ &\leq \|F_1\|_{L^q([0, T], L^p(\mathbb{R}^{nd}))}^q \|f\|_{L^q([0, T], L^p(\mathbb{R}^{nd}))}^q \left( \int_t^T \frac{ds}{(s-t)^{\frac{n^2d}{2} \frac{1}{p} + \frac{1}{2}} \tilde{q}} \right)^{\frac{q}{\tilde{q}}}, \end{aligned} \quad (3.26)$$

with  $q^{-1} + (\tilde{q})^{-1} = 1$ . Hence, in order to derive our final bound on  $\|\mathcal{R}f\|_{L^q([0,T], L^p(\mathbb{R}^{nd}))}$ , it suffices to prove that the remaining time integral in the above equation gives a small contribution in times. It therefore remains 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:

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

which is precisely the condition appearing in **(A)**-(c) and assumed for the Lemma. The conclusion then readily follows from (3.26).  $\square$

**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 here to use 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 [Zha16]. We adapt here a bit the approach.

Recall that for the process  $(\bar{\mathbf{X}}_t)_{t \geq 0}$  introduced in (3.15) 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  $t \in (s, T]$  and all  $\Gamma \in \mathcal{B}(\mathbb{R}^{nd})$ :  $\bar{P}(t, s, x, \Gamma) = \int_{\Gamma} \bar{p}(s, t, \mathbf{x}, \mathbf{y}) d\mathbf{y}$ . More specifically, for any  $f \in L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))$ ,  $\frac{n^2d}{p'} + \frac{2}{q'} < 2$ ,  $p' > 1$ ,  $q' > 1$ , and  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

$$\left| \bar{\mathbb{E}}^{\bar{P}_{t, \mathbf{x}}} \left[ \int_t^T f(s, \bar{\mathbf{X}}_s) ds \right] \right| \leq C_{3.27} \|f\|_{L^{q'}([0, T], L^{p'}(\mathbb{R}^{nd}))}, \quad C_{3.27} := C_{3.27}(\mathbf{A}, T, p', q'). \quad (3.27)$$

We now state an exponential integrability result for the unique weak solution of (3.15). Such types of estimates were first proved by Khas'minskii in [Kha59] for the Brownian motion. Since the proof only relies on the Markov property, it does not require any specific modification.

**Lemma 7** (Khas'minskii's type exponential integrability). *Let  $(\bar{\mathbf{X}}_t)_{t \geq 0}$  be the (unique-weak) solution to (3.15). Then, for any  $T > 0$  s.t.*

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

one also has:

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

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 3** (Exponential integrability). *Let  $(\bar{\mathbf{X}}_t)_{t \geq 0}$  be the (unique-weak) solution to (3.15). Let  $\mathbf{F}_1 \in L^p([0, T], L^q(\mathbb{R}^{nd}))$  with  $\frac{n^2d}{p} + \frac{2}{q} < 1$ ,  $p \geq 2$ ,  $q > 2$ . Then, there exists  $K_{\mathbf{F}_1} := K_{\mathbf{F}_1}(\mathbf{A}, T)$  s.t.*

$$\sup_{\mathbf{x} \in \mathbb{R}^{nd}} \bar{\mathbb{E}}^{\bar{P}_{0, \mathbf{x}}} \left[ \exp \left( \int_0^T |(\sigma^{-1} \mathbf{F}_1)(s, \bar{\mathbf{X}}_s)|^2 ds \right) \right] \leq K_{\mathbf{F}_1}. \quad (3.28)$$

The constant  $K_{\mathbf{F}_1}$  depends continuously on  $\|\mathbf{F}_1\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}$ .

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

$$\begin{aligned} \|\sigma^{-1}\mathbf{F}_1\|^{2a} \|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))} &= \left( \int_0^T dt \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{x} |(\sigma^{-1}\mathbf{F}_1)(t, \mathbf{x})|^{2a p'} \right\}^{\frac{q'}{p'}} \right)^{\frac{1}{q'}} \\ &\leq C \left( \int_0^T dt \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{x} |\mathbf{F}_1(t, \mathbf{x})|^p \right\}^{\frac{q}{p}} \right)^{\frac{2a}{q}} \leq C \|\mathbf{F}_1\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}^{2a}. \end{aligned}$$

From equation (3.27) we thus derive that for all  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$\left| \mathbb{E}^{\bar{\mathbb{P}}_{0,\mathbf{x}}} \left[ \int_0^T |\sigma^{-1}\mathbf{F}_1(s, \bar{\mathbf{X}}_s)|^{2a} ds \right] \right| \leq C_{3.27} \|\sigma^{-1}\mathbf{F}_1\|^{2a} \|_{L^{q'}([0,T],L^{p'}(\mathbb{R}^{nd}))} \leq C_{3.27} C \|\mathbf{F}_1\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}^{2a}. \quad (3.29)$$

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

$$\mathbb{E}^{\bar{\mathbb{P}}_{0,\mathbf{x}}} \left[ \exp \left( \int_0^T |(\sigma^{-1}\mathbf{F}_1)(s, \bar{\mathbf{X}}_s)|^2 ds \right) \right] \leq \mathbb{E}^{\bar{\mathbb{P}}_{0,\mathbf{x}}} \left[ \exp \left( \int_0^T |(\sigma^{-1}\mathbf{F}_1)(s, \bar{\mathbf{X}}_s)|^{2a} \frac{\varepsilon}{a} ds \right) \right] C(T, a, \varepsilon).$$

The statement now directly follows from the above equation, (3.29) and Lemma 7 taking  $\varepsilon := (2C_{3.27}C \|\mathbf{F}_1\|_{L^q([0,T],L^p(\mathbb{R}^{nd}))}^{2a})^{-1}$ .  $\square$

Proposition 3 ensures that the Novikov condition is fulfilled in order to prove existence for the martingale problem associated with  $(L_t)_{t \geq 0}$  for  $\mathbf{F}_1$  satisfying **(D)**-(c) starting from  $\bar{\mathbb{P}}_{0,\mathbf{x}}$  and the associated dynamics (3.15) of the canonical process. Set,

$$\tilde{W}_t := W_t - \int_0^t (\sigma^{-1}\mathbf{F}_1)(s, \bar{\mathbf{X}}_s) ds, \quad t \in [0, T].$$

From Proposition 3, we derive that:

$$M_t := \exp \left( \int_0^t (\sigma^{-1}\mathbf{F}_1)(s, \bar{\mathbf{X}}_s) dW_s - \frac{1}{2} \int_0^t |(\sigma^{-1}\mathbf{F}_1)(s, \bar{\mathbf{X}}_s)|^2 ds \right), \quad t \in [0, T]$$

is an  $\bar{\mathbb{P}}$ - $\mathcal{F}_t$  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  $\frac{d\mathbb{P}}{d\bar{\mathbb{P}}}\Big|_{\mathcal{F}_T} := M_T$ . The dynamics of  $(\bar{\mathbf{X}}_s)_{s \in [0,T]}$ , writes under  $\mathbb{P}$ :

$$d\bar{\mathbf{X}}_t = \left( B\mathbf{F}_1(t, \bar{\mathbf{X}}_t) + \bar{\mathbf{F}}(t, \bar{\mathbf{X}}_t) \right) dt + B\sigma(t, \bar{\mathbf{X}}_t) d\tilde{W}_t = \mathbf{F}(t, \bar{\mathbf{X}}_t) dt + B\sigma(t, \bar{\mathbf{X}}_t) d\tilde{W}_t,$$

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

## 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  $\beta_i^j$ ,  $i \leq j$  in  $\llbracket 2, n \rrbracket^2$  in (1.3) since, as emphasized in the introduction, we recall that when the condition on the  $L^q - L^p$  integrability of  $\mathbf{F}_1$  we have from Example 69 of [BFGM14] that even weak existence can fail or weak uniqueness could not hold (see Proposition 3.3 of [GO13]). We eventually recall that, the almost sharpness of the coefficients  $\beta_i^i$ ,  $i$  in  $\llbracket 2, n \rrbracket$ , has been already proved in [CdR16].

Let us first introduce the main idea of our counter example. As we already discussed, the system (1.2) is ill posed as soon as  $\alpha$  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  $\gamma$  of the noise added in the system has to be strictly lower than  $1/(1 - \alpha)$ . We formalize these facts with our Proposition 4 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 stress the (almost) optimal threshold for the drift of the  $i^{\text{th}}$  component with respect to the  $j^{\text{th}}$  variable we need to build an *ad hoc* Peano-like example. Focusing on the  $i^{\text{th}}$  component and the  $j^{\text{th}}$  variable, for  $i \in \llbracket 2, n \rrbracket$  and  $j \in \llbracket i, n \rrbracket$ , we consider:

$$\begin{cases} \dot{x}_t^1 = \dot{x}_t^2 = \dots = \dot{x}_t^{i-1} = 0 \\ \dot{x}_t^i = \text{sign}(x_t^j) |x_t^j|^{\beta_i^j} dt \\ \dot{x}_t^{i+1} = x_t^i dt, \dots, \dot{x}_t^j = x_t^{j-1} dt \\ \dot{x}_t^{j+1} = x_t^{j+2} dt = \dots = \dot{x}_t^n = 0 dt \end{cases} \quad t \geq 0,$$

and  $x_0^l = 0$  for all  $l$  in  $\llbracket 1, j \rrbracket$ . Each entry  $(x_t^k)_{t \geq 0, k \in \llbracket 1, n \rrbracket}$  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  $i^{\text{th}}$  equation whose extreme solutions write  $\pm c \frac{x^j}{\beta_i^j} t^{((j-i)\beta_i^j + 1)/(1 - \beta_i^j)}$ . In that case, our stochastic perturbation is on the form

$$d\mathbf{X}_t = \mathbf{A}\mathbf{X}_t dt + B dW_t + \mathbf{F}_i^j(\mathbf{X}_t) dt, \quad \mathbf{A} = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ 1 & 0 & \dots & 0 & \vdots \\ 0 & 1 & 0 & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}, \quad (4.1)$$

where  $\mathbf{F}_i^j(\mathbf{x}) = e_i \text{sign}(\mathbf{x}_j) |\mathbf{x}_j|^{\beta_i^j}$ ,  $e_i = (0 \dots 1 \dots 0)^*$ . This in particular means that the perturbation of the  $i^{\text{th}}$  component is done by the  $(i - 1)^{\text{th}}$  iterated integrals (in time) of Brownian motion. Thus,

focusing on the  $i^{\text{th}}$  stage of the chain and the  $j^{\text{th}}$  component this means that we are interested in the following type of SDE

$$Z_t^{i,j} = x + \int_0^t \text{sign} \left( \int_0^s \dots \int_0^{s_{i+2}} Z_{s_{i+1}}^{i,j} ds_{i+1} \dots ds_j \right) \left| \int_0^s \dots \int_0^{s_{i+2}} Z_{s_{i+1}}^{i,j} ds_{i+1} \dots ds_j \right|^{\beta_i^j} ds + \mathcal{W}_t^i, \quad (4.2)$$

The non-uniqueness in law for equation (4.2) will then follow from the next proposition.

**Proposition 4** (Failure of the well posedness for the regularized Peano system). *Let  $\mathcal{W}$  be a random process with continuous path satisfying, in law, an invariance by symmetry and a self-similarity property of order  $\gamma > 0$ . Namely:*

$$(\mathcal{W}_t, t \geq 0) = (-\mathcal{W}_t, t \geq 0), \quad \forall t \geq 0, t^\gamma \mathcal{W}_1 = \mathcal{W}_t.$$

Suppose moreover that  $\mathbb{E}|\mathcal{W}_1| < +\infty$  and that  $\mathcal{W}$  and  $\alpha < 1$  are such that there exists a weak solution to the following SDE:

$$Z_t = x + \int_0^t \text{sign} \left( \int_0^s \dots \int_0^{s_{i+2}} Z_{s_1} ds_1 \dots ds_k \right) \left| \int_0^s \dots \int_0^{s_{i+2}} Z_{s_1} ds_1 \dots ds_k \right|^\alpha ds + \mathcal{W}_t, \quad (4.3)$$

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 4. 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 4.

*Proof of Proposition 4.* Consider the maximal 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 extreme solution of the associated Peano Example. The proof follows the lines of [CdR16] 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

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

Again  $\tau(Y)$  corresponds to the first passage of  $Y$  below a threshold related to the extreme solution of the deterministic Peano system.

The key point to the proof of Proposition 4 is the following Lemma.

**Lemma 8.** *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  $\rho$ , depending on  $\alpha, \beta, \gamma$  and  $\mathbb{E}|\mathcal{W}_1|$  only, such that*

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

Note now that if  $(Z, \mathcal{W})$  is 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 a contradiction. □

*Proof of Lemma 8.* 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:

$$\begin{aligned} Z_t &= x + \int_0^t \text{sign} \left( \int \dots \int Z_{s_1} ds_1 \dots ds_k \right) \left| \int \dots \int Z_{s_1} ds_1 \dots ds_k \right|^\alpha dt + \mathcal{W}_t \\ &\geq (1 - \beta)^\alpha c_{\alpha, k} t^{(k\alpha+1)/(1-\alpha)} + \mathcal{W}_t. \end{aligned}$$

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

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

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

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

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

Note that on  $A$  we have

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

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

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

We are now going to bound from below the probability of the event  $A$ . We have

$$\begin{aligned}
\mathbb{P}(A^c) &= \mathbb{P}\left(\exists t \in (0, \rho] : \tilde{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)} + \mathcal{W}_t \leq 0\right) \\
&\leq \mathbb{P}\left(\exists t \in (0, \rho] : |\mathcal{W}_t| \geq \tilde{c}_{\alpha,k} t^{(k\alpha+1)/(1-\alpha)}\right) \\
&= \mathbb{P}\left(\exists t \in (0, 1] : (\rho t)^\gamma |\mathcal{W}_1| \geq \tilde{c}_{\alpha,k} (\rho t)^{(k\alpha+1)/(1-\alpha)}\right) \\
&= \mathbb{P}\left(\exists t \in (0, 1] : |\mathcal{W}_1| \geq \tilde{c}_{\alpha,k} (\rho t)^{-\delta}\right),
\end{aligned}$$

where  $\delta = \gamma - (k\alpha + 1)/(1 - \alpha)$ . Since  $\alpha < (\gamma - 1)/(k + \gamma)$ , we get that  $\delta > 0$  and we obtain from the previous computations that

$$\mathbb{P}(A^c) \leq \mathbb{P}\left(|\mathcal{W}_1| \geq \tilde{c}_{\alpha,k} \rho^{-\delta}\right) \leq \mathbb{E}|\mathcal{W}_1| \tilde{c}_{\alpha,k}^{-1} \rho^\delta,$$

from the Markov inequality. Thus

$$\mathbb{P}(\tau(Z) \geq \rho) \geq \mathbb{P}(A) \geq 1 - \mathbb{E}|\mathcal{W}_1| \tilde{c}_{\alpha,k}^{-1} \rho^\delta,$$

so that there exists a positive  $\rho$  such that  $\mathbb{P}(\tau(Z) \geq \rho) \geq 3/4$ . □

## A Proof of the Technical Lemma 3

For the proof of this Lemma we are somehow faced with the same type of difficulties as for Lemma 2. 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 1. 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:

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

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

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

where  $(\bar{\mathbf{R}}^{t, \mathbf{x}}(v, u))_{t \leq u, v \leq t+\varepsilon}$  stands for the resolvent associated with  $(D\mathbf{F}(u, \mathbf{x}))_{u \in [t, t+\varepsilon]}$ . Observe that,



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

$$\begin{aligned}
& \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] =: \sum_{i=1}^3 \Xi_i^\varepsilon(t, \mathbf{x}).
\end{aligned}$$

It is directly seen that  $\Xi_3^\varepsilon(t, \mathbf{x}) \xrightarrow{\varepsilon \downarrow 0} f(\mathbf{x})$ . It remains to prove that  $\Xi_1^\varepsilon(t, \mathbf{x}), \Xi_2^\varepsilon(t, \mathbf{x})$  can be viewed as remainders as  $\varepsilon \downarrow 0$ .

Let us write

$$|\Xi_1^\varepsilon(t, \mathbf{x})| \leq \|f\|_\infty \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}} \int_0^1 d\lambda |(\varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda)|, \quad (\text{A.3})$$

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

$$\begin{aligned}
\varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon(\lambda) &= \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1} (\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y} \right\rangle + \right. \\
& \quad \lambda \left[ \left\langle (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1} (\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}) - \mathbf{y}), \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}) - \mathbf{y} \right\rangle \right. \\
& \quad \left. \left. - \left\langle (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1} (\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y} \right\rangle \right] \right), \\
|(\varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda)| &\leq \left| (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-\frac{1}{2}} \left\{ (\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}) + (\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}) - \mathbf{y}) \right\} \right| \\
& \quad \times \left| (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-\frac{1}{2}} \left\{ (\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y}) - (\tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}) - \mathbf{y}) \right\} \right| \varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon(\lambda), \quad (\text{A.4})
\end{aligned}$$

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:

$$\left| (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{x}})^{-\frac{1}{2}} (\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})) \right| \leq C\varepsilon^{\frac{1}{2}} |\mathbb{T}_\varepsilon^{-1} (\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x}))|. \quad (\text{A.5})$$

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

$$\begin{aligned}
\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u, \mathbf{z}) &:= (\mathbf{F}_1(u, \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})), \mathbf{F}_2(u, \mathbf{z}_1, (\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))^{2,n}), \mathbf{F}_3(u, \mathbf{z}_2, (\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))^{3,n}), \dots, \\
& \quad \mathbf{F}_n(s, \mathbf{z}_{n-1}, (\boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))_n). \quad (\text{A.6})
\end{aligned}$$

We then write:

$$\begin{aligned}
\varepsilon^{\frac{1}{2}} \mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})) &:= \varepsilon^{\frac{1}{2}} \mathbb{T}_\varepsilon^{-1} \left\{ \int_t^{t+\varepsilon} du \left[ \left( \mathbf{F}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) - \mathbf{F}^{t+\varepsilon,\mathbf{y}}(u, \boldsymbol{\theta}_{u,t}(\mathbf{x})) \right) \right. \right. \\
&\quad + \left( D\mathbf{F}(u, \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})) (\boldsymbol{\theta}_{u,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{u,t}^{t+\varepsilon,\mathbf{y}}(\mathbf{x})) \right) \\
&\quad + \left( \int_0^1 d\lambda \left( D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u, \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}) + \lambda(\boldsymbol{\theta}_{u,t}(\mathbf{x}) - \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y}))) \right. \right. \\
&\quad \left. \left. - D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u, \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})) \right) (\boldsymbol{\theta}_{u,t}(\mathbf{x}) - \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})) \right) \left. \right] \Big\} \\
&:= (\mathcal{I}_{t+\varepsilon,t}^1 + \mathcal{I}_{t+\varepsilon,t}^2 + \mathcal{I}_{t+\varepsilon,t}^3)(\mathbf{x}, \mathbf{y}), \tag{A.7}
\end{aligned}$$

where, according to the notations of (A.6), 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  $(D\mathbf{F}^{t+\varepsilon,\mathbf{y}}(u, \mathbf{z}))_{j,j-1} := D_{\mathbf{x}_{j-1}} \mathbf{F}_j(u, \mathbf{z}_{j-1}, \boldsymbol{\theta}_{u,t+\varepsilon}(\mathbf{y})^{j,n})$ ,  $j \in \llbracket 2, n \rrbracket$ . 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):

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

From Lemma 2 we now derive:

$$\begin{aligned}
|\mathcal{I}_{t+\varepsilon,t}^3(\mathbf{x}, \mathbf{y})| &\leq C \int_t^{t+\varepsilon} du \varepsilon^{-1+\eta/2} \left( \varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + 1 \right)^{1+\eta} \\
&\leq C \varepsilon^{\eta/2} \left( \varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + 1 \right)^{1+\eta}. \tag{A.8}
\end{aligned}$$

Let us now deal with  $\mathcal{I}_{t+\varepsilon,t}^1(\mathbf{x}, \mathbf{y})$ . From the previous definition of  $\mathbf{F}^{t+\varepsilon,\mathbf{y}}$  in (A.6), the key idea is to use the sub-linearity of  $\mathbf{F}$  and the appropriate Hölder exponents. Namely, using the Young inequality we derive:

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

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

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

We now use Lemma 2 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 2 is valid for any flow starting from  $\boldsymbol{\theta}_{u,t}(\mathbf{x})$  at time  $u$  associated with the ODE (see Appendix B below) we can proceed along the previous one, i.e.  $(\boldsymbol{\theta}_{v,t}(\mathbf{x}))_{v \in [u, t+\varepsilon]}$ . This yields:

$$\begin{aligned}
&|\mathcal{I}_{t+\varepsilon,t}^1(\mathbf{x}, \mathbf{y})| \tag{A.9} \\
&\leq C \left[ \varepsilon^{1/2} + (\varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + 1) \varepsilon \left( 1 + \sum_{i=2}^n \sum_{j=i}^n \left\{ \varepsilon^{-i+j-\gamma_i^j} + \varepsilon^{-i+1/2+\gamma_i^j \frac{\beta_i^j}{1-\beta_i^j}} \right\} \right) \right].
\end{aligned}$$

Choose now for  $i \in \llbracket 2, n \rrbracket$  and  $j \in \llbracket i, n \rrbracket$ ,

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

to balance the two previous contributions associated with the indexes  $i, j$ . To obtain a global smoothing effect w.r.t  $\varepsilon$  in (A.9) 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  $\zeta := \zeta((\mathbf{A}), (\beta_i^j)_{i \in \llbracket 1, n \rrbracket, j \in \llbracket i, n \rrbracket}) \in (0, 1)$  s.t.:

$$|\mathcal{I}_{t+\varepsilon,t}^1(\mathbf{x}, \mathbf{y})| \leq C \varepsilon^\zeta (1 + \varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|). \tag{A.10}$$

We now get from (A.5), (A.7), (A.8), (A.10) and the Gronwall lemma that:

$$\begin{aligned}
&\left| (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon, \mathbf{x}})^{-\frac{1}{2}} (\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon, \mathbf{y}}(\mathbf{x})) \right| \\
&\leq C \varepsilon^{\frac{1}{2}} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}_{t+\varepsilon,t}^{t+\varepsilon, \mathbf{y}}(\mathbf{x}))| \\
&\leq C \varepsilon^{\eta/2 \wedge \zeta} (1 + \varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})| + (\varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x}) - \mathbf{y})|)^{1+\eta}). \tag{A.11}
\end{aligned}$$

Hence, recalling from (2.13) and Lemma 2 that

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

we get from the definition in (A.4) 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.4) and (A.11) that there exists  $C_2 := C_2(T, (\mathbf{A})) \geq 1$  s.t.:

$$|(\varphi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda)| \leq C_2\varepsilon^{\eta/2\wedge\zeta} \exp(-C_2^{-1}\varepsilon|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|^2). \quad (\text{A.12})$$

Plugging Equation (A.12) into (A.3), we derive that, since  $\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}}$  satisfies (2.10),  $|\Xi_1^\varepsilon(t, \mathbf{x})| \xrightarrow{\varepsilon \downarrow} 0$ .

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

$$\begin{aligned} \Xi_{21}^\varepsilon(t, \mathbf{x}) &:= \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(2\pi)^{\frac{nd}{2}}} \frac{f(\mathbf{y})}{\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}} \int_0^1 d\lambda (\psi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda), \\ \forall \lambda \in [0, 1], \psi_{t,\mathbf{x},\mathbf{y}}^\varepsilon(\lambda) &:= \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y} \right\rangle \right. \\ &\quad \left. + \lambda \left[ \left\langle (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y} \right\rangle \right. \right. \\ &\quad \left. \left. - \left\langle (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y} \right\rangle \right] \right\}, \\ \Xi_{22}^\varepsilon(t, \mathbf{x}) &:= \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(2\pi)^{\frac{nd}{2}}} f(\mathbf{y}) \left[ \frac{1}{\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{\frac{1}{2}}} - \frac{1}{\det(\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{\frac{1}{2}}} \right] (\psi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)(1). \end{aligned} \quad (\text{A.13})$$

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

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

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

$$\begin{aligned} |(\psi_{t,\mathbf{x},\mathbf{y}}^\varepsilon)'(\lambda)| &\leq C \left| \left\langle ((\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t+\varepsilon,\mathbf{y}})^{-1} - (\tilde{\mathbf{K}}_{t+\varepsilon,t}^{t,\mathbf{x}})^{-1})(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y} \right\rangle \right| \\ &\quad \times \exp(-C\varepsilon|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|^2) \\ &=: C|Q_\varepsilon| \exp(-C\varepsilon|\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon,t}(\mathbf{x})-\mathbf{y})|^2), \end{aligned} \quad (\text{A.14})$$

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

Now, the covariance matrices explicitly write

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

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:

$$\tilde{\mathbf{K}}_{t+\varepsilon, t}^{t+\varepsilon, \mathbf{y}} = \varepsilon^{-1} \mathbb{T}_\varepsilon \widehat{\mathbf{K}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} \mathbb{T}_\varepsilon, \quad \bar{\mathbf{K}}_{t+\varepsilon, t}^{t, \mathbf{x}} = \varepsilon^{-1} \mathbb{T}_\varepsilon \widehat{\mathbf{K}}_1^{t+\varepsilon, t, t, \mathbf{x}} \mathbb{T}_\varepsilon,$$

we write:

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

It remains to control the term

$$\begin{aligned} \widehat{\mathbf{K}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} - \widehat{\mathbf{K}}_1^{t+\varepsilon, t, t, \mathbf{x}} &= (\Delta_1^{t+\varepsilon, t} - \Delta_2^{t+\varepsilon, t})(\mathbf{x}, \mathbf{y}), \\ \Delta_1^{t+\varepsilon, t}(\mathbf{x}, \mathbf{y}) &:= \varepsilon \int_t^{t+\varepsilon} du \mathbb{T}_\varepsilon^{-1} \tilde{\mathbf{R}}^{t+\varepsilon, \mathbf{y}}(t+\varepsilon, u) B \Delta a(u, t+\varepsilon)(\mathbf{x}, \mathbf{y}) B^* \tilde{\mathbf{R}}^{t+\varepsilon, \mathbf{y}}(t+\varepsilon, u)^* \mathbb{T}_\varepsilon^{-1}, \\ \Delta a(u, t+\varepsilon)(\mathbf{x}, \mathbf{y}) &= \left( a(u, \boldsymbol{\theta}_{u, t+\varepsilon}(\mathbf{y})) - a(u, \boldsymbol{\theta}_{u, t}(\mathbf{x})) \right), \\ \Delta_2^{t+\varepsilon, t}(\mathbf{x}, \mathbf{y}) &:= \varepsilon \int_t^{t+\varepsilon} du \mathbb{T}_\varepsilon^{-1} \widetilde{\Delta \mathbf{R}}^{t+\varepsilon, u, t, \mathbf{x}, \mathbf{y}}(t+\varepsilon, u) B a(u, \boldsymbol{\theta}_{u, t}(\mathbf{x})) B^* \bar{\mathbf{R}}^{t, \mathbf{x}}(t+\varepsilon, u)^* \mathbb{T}_\varepsilon^{-1} \\ &\quad - \varepsilon \int_t^{t+\varepsilon} du \left\{ \mathbb{T}_\varepsilon^{-1} \tilde{\mathbf{R}}^{t+\varepsilon, \mathbf{y}}(t+\varepsilon, u) B a(u, \boldsymbol{\theta}_{u, t}(\mathbf{x})) B^* \right. \\ &\quad \left. \times \left( \widetilde{\Delta \mathbf{R}}^{t+\varepsilon, u, t, \mathbf{x}, \mathbf{y}}(t+\varepsilon, u) \right)^* \mathbb{T}_\varepsilon^{-1} \right\} \\ \widetilde{\Delta \mathbf{R}}^{t+\varepsilon, u, t, \mathbf{x}, \mathbf{y}}(t+\varepsilon, u) &= \left( \bar{\mathbf{R}}^{t, \mathbf{x}}(t+\varepsilon, u) - \tilde{\mathbf{R}}^{t+\varepsilon, \mathbf{y}}(t+\varepsilon, u) \right). \end{aligned} \tag{A.15}$$

From the scaling properties of the resolvent, see e.g. Lemma 6.2 in [Men17] for details, we have that:

$$\bar{\mathbf{R}}^{t, \mathbf{x}}(t+\varepsilon, u) = \mathbb{T}_\varepsilon \widehat{\mathbf{R}}_1^{t+\varepsilon, u, t, \mathbf{x}} \mathbb{T}_\varepsilon^{-1}, \quad \tilde{\mathbf{R}}^{t+\varepsilon, \mathbf{y}}(t+\varepsilon, u) = \mathbb{T}_\varepsilon \widehat{\mathbf{R}}_1^{t+\varepsilon, u, t, \mathbf{x}} \mathbb{T}_\varepsilon^{-1}, \tag{A.16}$$

where  $\widehat{\mathbf{R}}_1^{t+\varepsilon, u, t, \mathbf{x}}$ ,  $\widehat{\mathbf{R}}_1^{t+\varepsilon, u, t, \mathbf{x}}$  are non-degenerate bounded matrices, uniformly in  $u \in [t, t+\varepsilon]$ . Hence, from (A.16) and the definitions in (A.15):

$$\begin{aligned} & |\Delta_1^{t+\varepsilon, t}(\mathbf{x}, \mathbf{y})| \\ &\leq C \varepsilon^{-1} \int_t^{t+\varepsilon} |\Delta a(u, t+\varepsilon)(\mathbf{x}, \mathbf{y})| du \leq C \varepsilon^{-1} \int_t^{t+\varepsilon} |\boldsymbol{\theta}_{u, t}(\mathbf{x}) - \boldsymbol{\theta}_{u, t+\varepsilon}(\mathbf{y})|^\eta du \\ &\leq C \varepsilon^{\eta/2-1} \int_t^{t+\varepsilon} du |\varepsilon^{1/2} \mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{u, t}(\mathbf{x}) - \boldsymbol{\theta}_{u, t+\varepsilon}(\mathbf{y}))|^\eta \leq C \varepsilon^{\eta/2} (|\varepsilon^{1/2} \mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})|^\eta + 1), \end{aligned} \tag{A.17}$$

using again Lemma 2 for the last inequality. Still from (A.16) and the definitions in (A.15), write now that:

$$|\Delta_2^{t+\varepsilon, t}(\mathbf{x}, \mathbf{y})| \leq C \varepsilon^{-1} \int_t^{t+\varepsilon} |\mathbb{T}_\varepsilon^{-1} \widetilde{\Delta \mathbf{R}}^{t+\varepsilon, u, t, \mathbf{x}, \mathbf{y}}(t+\varepsilon, u) \mathbb{T}_\varepsilon| du. \tag{A.18}$$

Note then:

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

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  $i \in \llbracket 2, n \rrbracket, \forall \mathbf{z}^{i:n} = (\mathbf{z}_i, \dots, \mathbf{z}_n) \in \mathbb{R}^{(n-i+1)d}, \mathbf{z}_{i-1} \mapsto D_{\mathbf{x}_{i-1}} \mathbf{F}_i(\mathbf{z}_{i-1}, \mathbf{z}^{i:n})$  is  $C^\eta(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ -Hölder continuous for  $\eta > 0$ . We thus have to handle the above term with some care. Write with the notations of (A.6):

$$\begin{aligned}
& |\mathbb{T}_\varepsilon^{-1} \Delta \tilde{\mathbf{R}}^{\tilde{t}+\varepsilon, u, t, \mathbf{x}, \mathbf{y}}(t + \varepsilon, u) \mathbb{T}_\varepsilon| \\
&\leq C \int_u^{t+\varepsilon} \varepsilon^{-1} \sum_{i=2}^n \left( |D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i^{t+\varepsilon, \mathbf{y}}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))| \right. \\
&\quad \left. + |D_{\mathbf{x}_{i-1}} \mathbf{F}_i^{t+\varepsilon, \mathbf{y}}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))| \right) dv \\
&\leq C \int_u^{t+\varepsilon} \varepsilon^{-1} \sum_{i=2}^n \left( |D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1}, (\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))^{i:n})| \right. \\
&\quad \left. + |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))_{i-1}|^\eta \right) dv =: (R_1 + R_2)(t + \varepsilon, u, t, \mathbf{x}, \mathbf{y}).
\end{aligned}$$

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.17). This yields:  $|R_2(t + \varepsilon, u, t, \mathbf{x}, \mathbf{y})| \leq C \varepsilon^{\eta/2} (\varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})| + 1)^\eta$ . On the other hand, using a reverse Taylor expansion, for positive parameters  $(\delta_i)_{i \in \llbracket 2, n \rrbracket}$  to be specified:

$$\begin{aligned}
& |R_1(t + \varepsilon, u, t, \mathbf{x}, \mathbf{y})| \\
&\leq C \int_u^{t+\varepsilon} \varepsilon^{-1} \sum_{i=2}^n \left( \left| \{ \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1} + \delta_i, (\boldsymbol{\theta}_{v,t}(\mathbf{x}))^{i:n}) - \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) \} \right. \right. \\
&\quad \left. \left. - \{ \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1} + \delta_i, (\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))^{i:n}) - \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1}, (\boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))^{i:n}) \} \right| \delta_i^{-1} + \delta_i^\eta \right) dv \\
&\leq C \int_u^{t+\varepsilon} \varepsilon^{-1} \sum_{i=2}^n \left( \sum_{j=i}^n |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t+\varepsilon}(\mathbf{y}))_j|^{\beta_i^j} \delta_i^{-1} + \delta_i^\eta \right). \\
&\leq C \left\{ \sum_{i=2}^n \sum_{j=i}^n \left( |\varepsilon^{1/2} (\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y}))|^{\beta_i^j} + 1 \right) \varepsilon^{(j-1/2)\beta_i^j} \delta_i^{-1} + \max_{i \in \llbracket 2, n \rrbracket} \delta_i^\eta \right\},
\end{aligned}$$

using again Lemma 2 for the last inequality. For this contribution to be a remainder it therefore suffices to choose  $\delta_i = \max_{j \in \llbracket i, n \rrbracket} \varepsilon^{(j-1/2)\beta_i^j - \gamma}$ , 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.

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

Plugging this bound into (A.18), we then derive from (A.17) and (A.15) that:

$$|\widehat{\mathbf{K}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}} - \widehat{\mathbf{K}}_1^{t+\varepsilon, t, t, \mathbf{x}}| \leq C \varepsilon^{\frac{\eta}{2} \wedge \zeta'} (\varepsilon^{\frac{1}{2}} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})| + 1). \quad (\text{A.19})$$

Because of the non degeneracy of  $a$ , the inverse matrices  $(\widehat{\mathbf{K}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}})^{-1}$ ,  $(\widehat{\mathbf{K}}_1^{t+\varepsilon, t, t, \mathbf{x}})^{-1}$  have the same Hölder regularity. Indeed, up to a change of coordinates one can assume that one of the two matrices is diagonal at the point considered and that the other has dominant diagonal if  $|\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y}|$  is small enough (depending on the ellipticity bound and the dimension). This reduces to the scalar case. Hence, from the definition in (A.14)

$$\begin{aligned} |Q_\varepsilon| &:= \left| \langle ((\tilde{\mathbf{K}}_{t+\varepsilon, t}^{t+\varepsilon, \mathbf{y}})^{-1} - (\bar{\mathbf{K}}_{t+\varepsilon, t}^{t, \mathbf{x}})^{-1})(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y}), \boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y} \rangle \right| \\ &= \left| \langle \left( (\widehat{\mathbf{K}}_1^{t+\varepsilon, t, t+\varepsilon, \mathbf{y}})^{-1} - (\widehat{\mathbf{K}}_1^{t+\varepsilon, t, t, \mathbf{x}})^{-1} \right) (\varepsilon^{1/2} \mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})), \varepsilon^{1/2} \mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y}) \rangle \right| \\ &\leq C \varepsilon^{\eta/2 \wedge \zeta'} (\varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})| + 1) \varepsilon |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})|^2, \quad C := C((\mathbf{A}), T). \end{aligned}$$

We eventually get:

$$\begin{aligned} |\Xi_{21}^\varepsilon(t, \mathbf{x})| &\leq C \varepsilon^{\eta/2 \wedge \zeta'} \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{\varepsilon^{\frac{n^2 d}{2}}} (\varepsilon^{1/2} |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})| + 1) \exp(-C \varepsilon |\mathbb{T}_\varepsilon^{-1}(\boldsymbol{\theta}_{t+\varepsilon, t}(\mathbf{x}) - \mathbf{y})|^2) \\ &\leq C \varepsilon^{\eta/2 \wedge \zeta'}, \quad C := C((\mathbf{A}), T). \end{aligned} \quad (\text{A.20})$$

This yields  $|\Xi_{21}^\varepsilon(t, \mathbf{x})| \xrightarrow{\varepsilon \rightarrow 0} 0$ . 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.13) we have  $\Xi_{22}^\varepsilon(t, \mathbf{x}) \xrightarrow{\varepsilon \rightarrow 0} 0$ . The proof is complete.  $\square$

## B On the proof of Theorem 1 without the Axiom of choice

In this part we explain how our result still holds without assuming the Axiom of choice. As already emphasized, the price to pay when working in that setting is the existence of a measurable Cauchy-Peano flow for the deterministic ODE associated with (1.1) under  $(\mathbf{A})$ -(a) (obtained taking  $\sigma = 0$  in (1.1)) which is only defined almost everywhere in space. Namely, instead of Lemma 1 we have the following Lemma whose proof is postponed at the end of the current section.

**Lemma 9.** *For a given  $T > 0$  sufficiently small and all  $(t, s) \in [0, T]^2$ , there exists a measurable mapping  $\mathbf{x} \in \mathbb{R}^{nd} \mapsto \boldsymbol{\theta}_{t, s}(\mathbf{x})$  s.t. for almost every  $\mathbf{x}$  in  $\mathbb{R}^{nd}$ ,  $\boldsymbol{\theta}_{t, s}(\mathbf{x}) = \mathbf{x} + \int_s^t \mathbf{F}(v, \boldsymbol{\theta}_{v, s}(\mathbf{x})) dv$ .*

Working with this “weak” Cauchy-Peano flow however does not affect the main lines of our proof but we have to pay attention at some stage. Actually, from Lemma 9, we would derive that the bounds of Lemma 2 now hold for almost every  $\mathbf{x}$  and  $\mathbf{y}$  uniformly in  $(s, t) \in [0, T]^2$ . Such a restriction does not modify the controls involving  $L^{q'} - L^{p'}$  norms of the Green kernel and its derivatives, see

e.g. Lemma 4. On the other hand, it can be seen from the proof of Lemma 2 that one can consider two different flows associated with the ODE deriving from (1.1). Namely, it holds that for the flow of Lemma 9, for almost every  $\mathbf{y}$  and any  $\mathbf{x}$ , uniformly in  $(s, t) \in [0, T]^2$ , for any Peano solution  $\bar{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) = \mathbf{x} + \int_t^s \mathbf{F}(v, \bar{\boldsymbol{\theta}}_{v,t}(\mathbf{x}))dv$ , one has:

$$C_2\{(s-t)|\mathbb{T}_{s-t}^{-1}(\bar{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 - 1\} \leq (s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \boldsymbol{\theta}_{t,s}(\mathbf{y}))|^2 \leq C_2\{(s-t)|\mathbb{T}_{s-t}^{-1}(\bar{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) - \mathbf{y})|^2 + 1\}, \quad (\text{B.1})$$

which allows us to switch from one flow (defined everywhere but not necessarily measurable) from another (measurable but defined only almost everywhere) and to preserve all the results.

This property is crucial to guarantee the Green kernel (2.28) defined in Section 2.3 to be well defined for any  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ , which was a key point to derive the well posedness of the martingale problem. It indeed suffices, instead of using Lemma 2 in the second line of (2.35), to invoke (B.1) switching from the measurable Cauchy-Peano flow defined almost everywhere in space provided by Lemma 9, to any other Cauchy-Peano flow  $(\bar{\boldsymbol{\theta}}_{s,t}(\mathbf{x}))_{s \in [t, T]}$  solving for a given  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ , and for all  $s \in [t, T]$ ,  $\bar{\boldsymbol{\theta}}_{s,t}(\mathbf{x}) = \mathbf{x} + \int_t^s \mathbf{F}(v, \bar{\boldsymbol{\theta}}_{v,t}(\mathbf{x}))dv$ .

*Proof of Lemma 9.* Our proof is as follows. We first show the existence of such a solution on every compact subset  $K$  of  $\mathbb{R}^{nd}$  so that the Lemma follows from covering argument.

Let  $(\boldsymbol{\theta}_{s,t}^\delta)_{t \leq s \leq T}$  be the mollified forward flow defined by (2.18). Then, for any  $(t, s) \in [0, T]^2$ , on any compact subset  $K$  of  $\mathbb{R}^{nd}$  we have from the sub linear growth assumed on  $\mathbf{F}$  that

$$\begin{aligned} \|\boldsymbol{\theta}_{s,t}^\delta(\cdot)\|_{L^2(K)}^2 &\leq \int_K d\mathbf{x} \left( |\mathbf{x}|^2 + C|s-t|\text{sgn}(s-t) \int_t^s dv \{1 + |\boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})|\} \right) \\ &\leq \hat{C} + CT^2 \left( |K| + \sup_{(s',t') \in [0,T]^2} \|\boldsymbol{\theta}_{s',t'}^\delta(\cdot)\|_{L^2(K)}^2 \right), \end{aligned}$$

where  $\hat{C} := \hat{C}(|K|)$ . Hence, taking the supremum in  $s, t$  in the above l.h.s. and for  $T$  small enough, we obtain:

$$\sup_{(s,t) \in [0,T]^2} \|\boldsymbol{\theta}_{s,t}^\delta(\cdot)\|_{L^2(K)} \leq \tilde{C}, \quad \tilde{C} := \tilde{C}(T, (\mathbf{A}), K), \quad (\text{B.2})$$

that is, the sequence  $(\boldsymbol{\theta}_{\cdot,\cdot}^\delta)_{\delta>0}$  is uniformly bounded on  $\mathcal{C}([0, T]^2, L^2(K))$ . Note as well that similar arguments yield that, for all  $(s, t), (s', t') \in [0, T]^2$ :

$$\begin{aligned} \|\boldsymbol{\theta}_{s,t}^\delta(\cdot) - \boldsymbol{\theta}_{s',t'}^\delta(\cdot)\|_{L^2(K)} &\leq \|\boldsymbol{\theta}_{s,t}^\delta(\cdot) - \boldsymbol{\theta}_{s',t}^\delta(\cdot)\|_{L^2(K)} + \|\boldsymbol{\theta}_{s',t}^\delta(\cdot) - \boldsymbol{\theta}_{s',t'}^\delta(\cdot)\|_{L^2(K)} \\ &\leq \left( \int_K d\mathbf{x} \left| \int_s^{s'} \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))dv \right|^2 \right)^{\frac{1}{2}} + \left( \int_K d\mathbf{x} \left| \int_t^{t'} \mathbf{F}^\delta(v, \boldsymbol{\theta}_{s',v}(\mathbf{x}))dv \right|^2 \right)^{\frac{1}{2}} \\ &\leq \bar{C}(|s-s'| + |t-t'|), \end{aligned}$$

where  $\bar{C}$  also depends on the uniform bound  $\tilde{C}$  on  $\|\boldsymbol{\theta}_{s,t}^\delta(\cdot)\|_{L^2(K)}$  derived from (B.2). The mapping  $(t, s) \in [0, T]^2 \mapsto \boldsymbol{\theta}_{s,t}^\delta(\cdot) \in L^2(K)$  is so  $\bar{C}$ -Lipschitz. Hence, from the Arzelà-Ascoli theorem, one can extract a subsequence, still denoted by  $(\boldsymbol{\theta}_{\cdot,\cdot}^\delta)_{\delta>0}$ , that converges uniformly in  $(t, s)$  for almost every  $\mathbf{x}$  in  $K$  to some flow  $\boldsymbol{\theta}$  and such that  $\|\boldsymbol{\theta}_{s,t}^\delta(\cdot) - \boldsymbol{\theta}_{s,t}(\cdot)\|_{L^2(K)}$  goes to 0 with  $\delta$  uniformly in  $(t, s)$ . Let us now show that such a  $\boldsymbol{\theta}$  solves (2.3) in the  $L^2$  sense on any compact subset  $K$  of  $\mathbb{R}^{nd}$ . Let  $(t, s)$  in  $[0, T]^2$ , we have:



$$\begin{aligned}
\|\boldsymbol{\theta}_{s,t}^\delta(\cdot) - \boldsymbol{\theta}_{s,t}(\cdot)\|_{L^2(K)}^2 &= \int_K d\mathbf{x} \left| \mathbf{x} + \int_t^s \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})) dv - \boldsymbol{\theta}_{s,t}(\mathbf{x}) \right|^2 \\
&= \int_K d\mathbf{x} \left| \left( \int_t^s \mathbf{F}^\delta(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})) - \mathbf{F}(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})) dv \right) \right. \\
&\quad \left. + \left( \int_t^s \mathbf{F}(v, \boldsymbol{\theta}_{v,t}^\delta(\mathbf{x})) - \mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) dv \right) \right. \\
&\quad \left. + \left( \mathbf{x} + \int_t^s \mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) dv - \boldsymbol{\theta}_{s,t}(\mathbf{x}) \right) \right|^2 \\
&:= \left\| \mathcal{M}_1^\delta(t, s, \cdot) + \mathcal{M}_2^\delta(t, s, \cdot) + \mathcal{M}_3(t, s, \cdot) \right\|_{L^2(K)}^2.
\end{aligned} \tag{B.3}$$

It is clear from the definition of our regularization procedure that for all  $(t, s, \mathbf{x})$  in  $[t, T]^2 \times \mathbb{R}^{nd}$ ,  $\mathcal{M}_1^\delta \rightarrow 0$  as  $\delta \rightarrow 0$ . Let us now consider the term  $\mathcal{M}_2^\delta$ . From **(A)**, there exists  $C := C(\mathbf{A})$  such that, for all  $v$  in  $[t, T]$

$$\begin{aligned}
|\mathbf{F}(v, \boldsymbol{\theta}_{v,t}^\delta(\cdot)) - \mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\cdot))| &\leq C \left\{ \sum_{j=1}^n |(\boldsymbol{\theta}_{v,t}^\delta(\cdot) - \boldsymbol{\theta}_{v,t}(\cdot))_j|^{\beta_1^j} + \sum_{i=2}^n \left\{ |(\boldsymbol{\theta}_{v,t}^\delta(\cdot) - \boldsymbol{\theta}_{v,t}(\cdot))_{i-1}| \right. \right. \\
&\quad \left. \left. + \sum_{j=i}^n |(\boldsymbol{\theta}_{v,t}^\delta(\cdot) - \boldsymbol{\theta}_{v,t}(\cdot))_j|^{\beta_i^j} \right\} \right\}.
\end{aligned} \tag{B.4}$$

Now, for every  $\gamma \in (0, 1]$  and  $j$  in  $\llbracket 1, n \rrbracket$  we obtain, thanks to Hölder inequality and Fubini Theorem, that there exists a positive  $C' := C'(\mathbf{A}, T, |K|)$  such that for all  $(s, t) \in [0, T]^2$ :

$$\begin{aligned}
\int_K d\mathbf{x} \left( \int_t^s dv |(\boldsymbol{\theta}_{v,t}^\delta(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}))_j|^\gamma \right)^2 &\leq |t - s| \int_K d\mathbf{x} \operatorname{sgn}(s - t) \int_t^s dv |(\boldsymbol{\theta}_{v,t}^\delta(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}))_j|^{2\gamma} \\
&\leq |t - s|^2 |K|^{1-\gamma} \sup_{(s', t') \in [0, T]^2} \left( \int_K d\mathbf{x} |(\boldsymbol{\theta}_{s', t'}^\delta(\mathbf{x}) - \boldsymbol{\theta}_{s', t'}(\mathbf{x}))_j|^2 \right)^\gamma \\
&\leq T |K|^{1-\gamma} \sup_{(s', t') \in [0, T]^2} \|\boldsymbol{\theta}_{s', t'}^\delta(\cdot) - \boldsymbol{\theta}_{s', t'}(\cdot)\|_{L^2(K)}^{2\gamma}.
\end{aligned} \tag{B.5}$$

We deduce from (B.4) and (B.5) that  $\mathcal{M}_2^\delta \xrightarrow{C([0, T]^2, L^2(K))} 0$  as  $\delta \rightarrow 0$ .

Hence, taking the supremum over all the  $(t, s)$  in  $[0, T]^2$  and passing to the limit in (B.3), we eventually derive that, on any compact subset  $K$  of  $\mathbb{R}^{nd}$ :

$$\sup_{(t, s) \in [0, T]^2} \|\mathcal{M}_3(t, s, \cdot)\|_{L^2(K)}^2 = \sup_{(t, s) \in [0, T]^2} \int_K d\mathbf{x} \left| \mathbf{x} + \int_t^s \mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) dv - \boldsymbol{\theta}_{v,t}(\mathbf{x}) \right|^2 = 0.$$

Hence, we obtain that for all compact subset  $K$  in  $\mathbb{R}^{nd}$ ,  $[0, T]^2 \ni (s, t) \mapsto \boldsymbol{\theta}_{s,t}(\cdot) \in L^2(K)$  is a Lipschitz function that solves the ODE for almost every  $\mathbf{x}$  in  $K$ . The result can then be extended to  $\mathbb{R}^{nd}$  writing  $\mathbb{R}^{nd} = \bigcup_{(k_1, \dots, k_{nd}) \in \mathbb{Z}^{nd}} [k_i, k_i + 1]$ . □

## Acknowledgments.

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

## References

- [BFGM14] L. Beck, F. Flandoli., M. Gubinelli, and M. Maurelli. Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *arXiv:1401.1530 [math]*, January 2014.
- [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.
- [CC15] G. Cannizzaro and K. Chouk. Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential. *arXiv:1501.04751 [math]*, to appear in *Annals of Probability*, January 2015.
- [CdR16] P.-E. Chaudru de Raynal. Weak regularization by stochastic drift: result and counter example. *arXiv:1606.05458*, to appear in *Discrete and Continuous Dynamical Systems (Series A)*, 2016.
- [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.
- [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):182–207, 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.*, 4:309–318, 1959.
- [KR05] N. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Prob. Theory Rel. Fields*, 131:154–196, 2005.
- [Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electronic Communications in Probability*, 17:234–250, 2011.
- [Men17] S. Menozzi. Martingale problems for some degenerate Kolmogorov equations. *arXiv:1404.0461v3*, to appear in *Stoc. Proc. Appl.*, <https://doi.org/10.1016/j.spa.2017.06.001>, 2017.
- [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.
- [Pri17] E. Priola. Davie’s type uniqueness for a class of SDEs with jumps. *arXiv:1509.07448*, To appear in *Ann. Inst. Henri Poincaré*, 2017.
- [SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.
- [Ver80] A. Y. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Matematicheskii 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 Matemática Iberoamericana*, 29(1), 2010.
- [Zha16] X. Zhang. Stochastic Hamiltonian flows with singular coefficients. *arXiv:1606.04360*, 2016.
- [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.