Rough Path and Regularity Structure Approaches to Stochastic Analysis: Rough Volatility and Singular SPDEs

RYOJI TAKANO

SEPTEMBER, 2025

Rough Path and Regularity Structure Approaches to Stochastic Analysis: Rough Volatility and Singular SPDEs

A dissertation submitted to
THE GRADUATE SCHOOL OF ENGINEERING SCIENCE,
The University of Osaka

in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN SCIENCE

RYOJI TAKANO

SEPTEMBER, 2025

Abstract

In this thesis, we analyze rough volatility models and parabolic Anderson model, which arise in mathematical finance and physics respectively, using rough path theory and the theory of regularity structures. Chapter 1 and Chapter 2 are devoted to rough volatility models, where we establish a large deviation principle (LDP) under weaker assumptions on the coefficients than in previous works.

Chapter 1 introduces partial rough paths, functionals of the noise $(\hat{X}, X)$, where X is a Brownian motion and $\hat{X}$ is typically Riemann-Liouville fractional Brownian motion. The structure of such partial rough paths is determined from approximations of integrals of the form $\int f(\hat{X})\,\mathrm{d}X$, obtained via Taylor expansion of a smooth function f. Using this framework, we prove an LDP for rough volatility models and derive the asymptotic behavior of the implied volatility, consistent with the power-law behavior observed in equity option markets.

Chapter 2 analyzes one-dimensional rough volatility models within the standard rough path framework. To this end, we focus on stochastic integrals whose integrand is given by the volatility process, constructing rough paths via Young pairing. Establishing an LDP for these stochastic integrals yields an LDP for rough volatility models, again characterizing the asymptotic behavior of the implied volatility.

Although both chapter share the common goal of analyzing rough volatility models through rough path techniques, their methodology differs. Chapter 1 emphasizes advancing rough path theory itself through the concept of partial rough paths, with largely deterministic tools. Chapter 2 takes a more probabilistic route, focusing on LDPs for stochastic integrals. Both approaches offer advantages over previous studies by requiring weaker assumptions on volatility coefficients though the precise conditions differ slightly between them.

Chapter 3 turns to singular stochastic partial differential equations, studied via the theory of regularity structures. A central object of interest is the construction of solution maps, achieved by combining the reconstruction theorem with the multi-level Schauder estimate. As an application, we construct local-in-time solutions to the two-dimensional parabolic Anderson model with a non-translation-invariant differential operator, again under weaker assumptions on the coefficients than those in previous works.

A main theme throughout the thesis is the analysis of random time evolutions through modern analytic frameworks, rough path theory and the theory of regularity structures. These methods are inherently pathwise, yet in finite-dimension settings they remain consistent with Itô calculus. Rough analysis thus presents a formulation that looks quite different from the classical Itô calculus, offering new insights. The results presented in this thesis contribute to broadening the mathematical foundations of rough path analysis and regularity structures in the study of stochastic models in finance and physics. It is hoped that further development of such analytic techniques will continue to deepen our understanding of stochastic analysis and its applications.

Contents

Note that the meanings of symbols vary from chapter to chapter. The content of Chapter 1 is based on joint work with Professor Masaaki Fukasawa (The University of Osaka). The content of Chapter 3 is based on joint work with Professor Masato Hoshino (Institute of Science Tokyo).

Chapter 1

A partial rough path space for rough volatility*

1.1 Introduction

A rough volatility model is a stochastic volatility model for an asset price process with volatility being rough, meaning that the Hölder regularity of the volatility path is less than half. Recently, such models have been attracting attention in mathematical finance because of their unique consistency to market data. Indeed, rough volatility models are the only class of continuous price models that are consistent to a power law of implied volatility term structure typically observed in equity option markets, as shown by [31]. One way to derive the power law under rough volatility models is to prove a large deviation principle (LDP) as done by many authors [23, 9, 8, 24, 25, 48, 50, 51, 62, 42, 63, 49] using various methods. An introduction to LDP and some of its applications to finance and insurance problems can be found in [71, 26]. In the context of the implied volatility, a short-time LDP under local volatility models provides a validity proof for a precise approximation known as the BBF formula [12, 1]. The SABR formula, which is of daily use in financial practice, is also proved as a valid approximation under the SABR model by means of LDP [69]. From these successes in classical (non-rough) volatility models, we expect LDP for rough volatility models to provide in particular a useful implied volatility approximation formula for financial practice such as model calibration.

For the classical models that are described by standard stochastic differential equations (SDEs), an elegant way to prove an LDP is to apply the contraction principle in the framework of rough path analysis [29, 30]. Under rough volatility models, the volatility of an asset price has a lower Hölder regularity than the asset price process. The stochastic integrands are therefore not controlled by the stochastic integrators in the sense of [43]. Hence, a rough volatility model is beyond the scope of rough path

*Reproduction of a joint work with Professor Masaaki Fukasawa (The University of Osaka), first published in Electronic Journal of Probability, Vol.29, No.18, pp.1–28, (2024), https://doi.org/10.1214/24-EJP1080.

theory, which motivated [8] to develop a regularity structure for rough volatility. For classical SDEs, the Freidlin–Wentzell LDP can be obtained as a consequence of the continuity of the solution map (the Lyons–Itô map) that is the core of rough path theory. In [8], the LDP for rough volatility models is obtained using the continuity of Hairer's reconstruction map. Herein, we take an approach that is similar to that of [8] in spirit but differs somewhat. Instead of embedding a rough volatility model into the abstract framework of regularity structure, we develop a minimal extension of rough path theory to incorporate rough volatility models. Besides the relatively elementary construction, an advantage of our theory is that it ensures the continuity of the integration map between rough path spaces, which enables us to treat a more general model than [8].

We focus on a model of the following form:

$$dS_t = \sigma(S_t) f(\hat{X}_t, t) dX_t, \quad S_0 \in \mathbb{R}, \tag{1.1}$$

where X is a d-dimensional Brownian motion, $\hat{X}$ is an e-dimensional stochastic process of which components include $\int_0^t \kappa(t-s) dX_s$ with a deterministic L^2 kernel κ. The stochastic integration is in the Itô sense. An example is the rough Bergomi model ($\kappa = \kappa_H$ is the Riemann–Liouville kernel (1.4), f is exponential, and $\sigma(s) = s$ in (1.1)) introduced by [7]. When $\kappa = \kappa_H$ or more generally κ has a similar singularity to κ_H with $H < 1/4$, beyond the case of $\sigma(s) = 1$ or $\sigma(s) = s$, no LDP is available in the literature so far, including [8]. As mentioned above, the difference between classical SDEs and (1.1) is that the volatility process $\hat{X}$ is not controlled by X because of its lower regularity. From empirical evidence, we are particularly interested in the case where $\hat{X}$ is correlated with X and $H < 1/4$ [40, 11, 36, 14]. Unfortunately, the application of existing rough path theory involves iterated integrals of $\hat{X}$ while, as is well-known, the standard rough path lift of $(X, \hat{X})$ that is amenable to LDP does not work when $H < 1/4$; see e.g., [30].

Our idea, inspired by [8], is to consider a partial rough path space in which we lack the iterated integrals of $\hat{X}$ but are still able to treat (1.1). More precisely, we define the space of a triplet of iterated integrals driven by X (we do not consider iterated integrals driven by $\hat{X}$) and rederive analytical results obtained in existing rough path theory. The notion of a partial rough path was introduced in [46] to prove the existence of global solutions for differential equations driven by a rough path with vector fields of linear growth. Our motivation is different and requires a space of higher-level paths. In contrast to [8], our method does not rely on the theory of regularity structure and enables us to treat not only the rough Bergomi model but also the following rough volatility models:

- the rough SABR model [35, 67, 34, 32];
- the mixed rough Bergomi model [15];
- rough local stochastic volatility [61];
- the two-factor fractional volatility model [38].

To the best of our knowledge, no LDP for these models is established so far in the literature.

To explain the idea of the partial rough path, here, we argue for how such a partial rough path space should be. Suppose that $d, e \geqq 1$, $x : [0,T] \to \mathbb{R}^d$, $\hat{x} : [0,T] \to \mathbb{R}^e$, and $f : \mathbb{R}^e \to \mathbb{R}$ are good enough. By the Taylor expansion, for $s < t$ (which are close enough), we have

$$\int_s^t f(\hat{x}_r)\mathrm{d}x_r \approx f(\hat{x}_s)(x_t - x_s) + \sum_{|i|\leqq n} \frac{1}{i!}\partial^i f(\hat{x}_s)\left[\int_s^t (\hat{x}_r - \hat{x}_s)^i \mathrm{d}x_r\right]$$

and

$$\begin{aligned}&\int_s^t \left(\int_s^r \mathrm{d}y_u\right) \otimes \mathrm{d}y_r \\ &\approx \sum_{|j+k|\leqq n} \frac{1}{j!k!}\partial^j f(\hat{x}_s)\partial^k f(\hat{x}_s)\left[\int_s^t (\hat{x}_r - \hat{x}_s)^k \left(\int_s^r (\hat{x}_u - \hat{x}_s)^j \mathrm{d}x_u\right) \otimes \mathrm{d}x_r\right],\end{aligned}$$

where $y_t := \int_0^t f(\hat{x}_r)\mathrm{d}x_r$, i, j, k are multi-indices, and we use the following notation:

$$|i| := \sum_{l=1}^e i_l, \quad i! := \prod_{l=1}^e i_l!, \quad x^i := \prod_{l=1}^e (x_l)^{i_l}, \quad \partial^i := \prod_{l=1}^e \left(\frac{\partial}{\partial x_l}\right)^{i_l}$$

for $i = (i_1, ..., i_e)$, $x = (x_1, ..., x_e)$. Therefore, following the idea of rough path theory, we would be able to define a rough path integral $\int f(\hat{x}_r)\mathrm{d}x_r$ if we could define

$$X_{st}^{(i)} := \frac{1}{i!}\int_s^t (\hat{X}_{sr})^i \mathrm{d}x_r, \quad \mathbf{X}_{st}^{(jk)} := \frac{1}{k!}\int_s^t (\hat{X}_{sr})^k X_{sr}^{(j)} \otimes \mathrm{d}x_r$$

for $\hat{X}_{sr} := \hat{x}_r - \hat{x}_s$. By the linearity of the integration and the binomial theorem (see Section 8.1 in [22]), $X^{(i)}$ and $\mathbf{X}^{(jk)}$ should satisfy the following formulas respectively: for any $i, j, k \in \mathbb{Z}_+^e$ and $s \leqq u \leqq t$,

$$X_{st}^{(i)} = X_{su}^{(i)} + \sum_{p\leqq i} \frac{1}{(i-p)!}(\hat{X}_{su})^{i-p} X_{ut}^{(p)} \tag{1.2}$$

and

$$\begin{aligned}\mathbf{X}_{st}^{(jk)} = \mathbf{X}_{su}^{(jk)} &+ \sum_{q\leqq k} \frac{1}{(k-q)!}(\hat{X}_{su})^{k-q} X_{su}^{(j)} \otimes X_{ut}^{(q)} \\ &+ \sum_{p\leqq j}\sum_{q\leqq k} \frac{1}{(j-p)!(k-q)!}(\hat{X}_{su})^{j+k-p-q}\mathbf{X}_{ut}^{(pq)},\end{aligned} \tag{1.3}$$

where, for $i, j \in \mathbb{Z}_+^e$, $i \leqq j$ means for all $l \in \{1, ..., e\}$, $i_l \leqq j_l$, and $\mathbb{Z}_+$ is the set of the nonnegative integers. Our partial rough space is a space for $\hat{X}$, $X^{(i)}$ and $\mathbf{X}^{(jk)}$, where the formulas (1.2) and (1.3) should play the role of Chen's identity.

In Section 1.2, we formulate such a partial rough path space and state some fundamental properties including the continuity of the integration map. In Section 1.3, we construct a rough path lift of our rough volatility model and state an LDP. Proofs are relegated to Section 1.4.

1.2 A partial rough path space

1.2.1 Definition

Throughout this article, we fix $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, $\beta \in (0, \frac{1}{2})$, $T > 0$ and denote

$$\Delta_T := \{(s,t) | 0 \leqq s \leqq t \leqq T\}, \quad I := \{i \in \mathbb{Z}_+^e | \ |i|\beta + \alpha \leqq 1\},$$

and

$$J := \{(j,k) \in \mathbb{Z}_+^e \times \mathbb{Z}_+^e | \ |j+k|\beta + 2\alpha \leqq 1\}.$$

Extending the notion of an α-Hölder rough path in rough path theory, here we define an (α, β) rough path.

Definition 1.2.1. An (α,β) rough path $\mathbb{X} = \left(\hat{X}, X^{(i)}, \mathbf{X}^{(jk)}\right)_{i \in I, (j,k) \in J}$ is a triplet of functions on Δ_T satisfying the following conditions for any $i \in I, (j,k) \in J$, and $s \leqq u \leqq t$.

(i) $\hat{X}$ is $\mathbb{R}^e$-valued, $X^{(i)}$ is $\mathbb{R}^d$-valued, and $\mathbf{X}^{(jk)}$ is $\mathbb{R}^d \otimes \mathbb{R}^d$-valued.

(ii) *Modified Chen's relation*: $\hat{X}_{st} = \hat{X}_{su} + \hat{X}_{ut}$, and $X^{(i)}$ and $\mathbf{X}^{(jk)}$ satisfy (1.2) and (1.3), respectively.

(iii) *Hölder regularity*:

$$|\hat{X}_{st}| \lesssim |t-s|^{\beta}, \quad |X_{st}^{(i)}| \lesssim |t-s|^{|i|\beta+\alpha}, \quad |\mathbf{X}_{st}^{(jk)}| \lesssim |t-s|^{|j+k|\beta+2\alpha}.$$

Let $\Omega_{(\alpha,\beta)\text{-Hld}}$ denote the set of (α, β) rough paths. We define a metric function $d_{(\alpha,\beta)}$ on $\Omega_{(\alpha,\beta)\text{-Hld}}$ and a homogeneous norm $|||\mathbb{X}|||_{(\alpha,\beta)}$ respectively by

$$\begin{aligned} & d_{(\alpha,\beta)}(\mathbb{X}, \mathbb{Y}) \\ & := ||\hat{X} - \hat{Y}||_{\beta\text{-Hld}} + \sum_{i \in I, (j,k) \in J} ||X^{(i)} - Y^{(i)}||_{|i|\beta+\alpha\text{-Hld}} + ||\mathbf{X}^{(jk)} - \mathbf{Y}^{(jk)}||_{|j+k|\beta+2\alpha\text{-Hld}} \end{aligned}$$

and

$$\begin{aligned} & |||\mathbb{X}|||_{(\alpha,\beta)} \\ & := ||\hat{X}||_{\beta\text{-Hld}} + \sum_{i \in I, (j,k) \in J} \left(||X^{(i)}||_{|i|\beta+\alpha\text{-Hld}}\right)^{1/(|i|+1)} + \left(||\mathbf{X}^{(jk)}||_{|j+k|\beta+2\alpha\text{-Hld}}\right)^{1/(|j+k|+2)}, \end{aligned}$$

where $\| \cdot \|_{\gamma\text{-Hld}}$ is the γ-Hölder norm for two-parameter functions for $\gamma \in (0,1]$:

$$||X||_{\gamma\text{-Hld}} := \sup_{0 \leqq s < t \leqq T} \frac{|X_{st}|}{|t-s|^{\gamma}}.$$

Remark 1.2.2. The modified Chen's relation and the Hölder regularity of $X^{(i)}$ and $\mathbf{X}^{(jk)}$ are from the following correspondence:

$$X_{st}^{(i)} \leftrightarrow \frac{1}{i!} \int_s^t \left(\hat{X}_{sr}\right)^i \mathrm{d}X_r^{(0)}, \quad \mathbf{X}_{st}^{(jk)} \leftrightarrow \frac{1}{k!} \int_s^t \left(\hat{X}_{sr}\right)^k X_{sr}^{(j)} \otimes \mathrm{d}X_r^{(0)}$$

when $X^{(0)}$ and $\hat{X}$ have Hölder regularity α and β, respectively. Note also that $\left(X^{(0)}, \mathbf{X}^{(00)}\right)$ is an α-Hölder rough path with the first level $X^{(0)}$ and the second level $\mathbf{X}^{(00)}$ in the usual rough path terminology. An (α, β) rough path has two first-level paths: $X^{(0)}$ and $\hat{X}$.

Remark 1.2.3. Our modified Chen's relation is a particular form of the algebraic structure of branched rough paths studied in [44]. However, because $\hat{X}$ is not a controlled path of X, the novel framework of (α, β) rough paths is essential for establishing the rough path integral stated in the Introduction.

Remark 1.2.4 (A comparison with [8])**.** The iterated integral $X^{(i)}_{st} = \frac{1}{i!}\int_s^t \left(\hat{X}_{sr}\right)^i \mathrm{d}X^{(0)}_r$ plays a key role also in [8] (see Section 3.1 in [8], where $X^{(i)}_{st} = \mathbb{W}^i_{st}$ in their notation). In [8], its derivative $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{W}^i_{st}$ appears in the structure space of regularity structure. Our (α, β) rough path consists of not only $X^{(i)}_{st}$ but also $\mathbf{X}^{(jk)}_{st}$. The latter is required to construct a rough path integral as an element of a rough path space, while in [8] the corresponding integral is constructed as merely a distribution and such terms as $\mathbf{X}^{(jk)}_{st}$ are not necessary for that purpose. As mentioned in Introduction, the key to treat (1.1) with a general function σ is to construct $\int f(\hat{X}_t, t)\mathrm{d}X_t$ as an element of a rough path space.

1.2.2 (α, β) rough path integration

Extending the rough path integration, here we introduce an integration with respect to an (α, β) rough path.

Definition 1.2.5. Fix $\mathbb{X} \in \Omega_{(\alpha,\beta)\text{-Hld}}$. We define $Y^{(1)}$ and $Y^{(2)}$ as follows if they exist:

$$Y^{(1)}_{st} := \lim_{|\mathcal{P}|\searrow 0} \sum_{p=1}^{N} \sum_{i\in I} \partial^i f(\hat{x}_{t_{p-1}}) X^{(i)}_{t_{p-1}t_p},$$

$$Y^{(2)}_{st} := \lim_{|\mathcal{P}|\searrow 0} \sum_{p=1}^{N} \left(Y^{(1)}_{t_0 t_{p-1}} \otimes Y^{(1)}_{t_{p-1}t_p} + \sum_{(j,k)\in J} \partial^j f(\hat{x}_{t_{p-1}}) \partial^k f(\hat{x}_{t_{p-1}}) \mathbf{X}^{(jk)}_{t_{p-1}t_p} \right),$$

where $\hat{x}_s := \hat{X}_{0s}$, and $\mathcal{P} = \{s = t_0 < t_1 < \ldots < t_N = t\}$ is a partition of the interval $[s, t]$. The mesh size $|\mathcal{P}|$ is defined by $|\mathcal{P}| = \max_p |t_p - t_{p-1}|$. If they exist on Δ_T, we denote $(Y^{(1)}, Y^{(2)})$ by $\int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X}$, and we call this the (α, β) rough path integral of f.

Denote by $\Omega_{\alpha\text{-Hld}}$ the α-Hölder rough path space, and denote by d_α the metric function on $\Omega_{\alpha\text{-Hld}}$; see [27], for example. Here, we state our first main result, the proof of which is given in Section 1.4.1.

Theorem 1.2.6. Let $n := \max\{|i| : i \in I\}$ and assume that $f : \mathbb{R} \to \mathbb{R}$ is C^{n+2}.

(i) For any $\mathbb{X} \in \Omega_{(\alpha,\beta)\text{-Hld}}$, the (α, β) rough path integral $\int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X}$ is well-defined, and $\int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X} \in \Omega_{\alpha\text{-Hld}}$.

(ii) The integration map $\int : \Omega_{(\alpha,\beta)\text{-Hld}} \to \Omega_{\alpha\text{-Hld}}$ is locally Lipschitz continuous. More precisely, for any $M > 0$, the map $\int |_{\mathcal{E}_M}$, restricted on the set

$$\mathcal{E}_M := \left\{ \mathbb{X} \in \Omega_{(\alpha,\beta)\text{-Hld}} \middle| \ |||\mathbb{X}|||_{(\alpha,\beta)} \leqq M \right\},$$

is Lipschitz continuous; that is, there exists a positive constant $C > 0$ such that

$$d_\alpha \left(\int f(\hat{\mathbb{V}})\mathrm{d}\mathbb{V}, \int f(\hat{\mathbb{W}})\mathrm{d}\mathbb{W} \right) \leqq C d_{(\alpha,\beta)} (\mathbb{V}, \mathbb{W}), \quad \mathbb{V}, \mathbb{W} \in \mathcal{E}_M.$$

1.3 Large deviation

1.3.1 A lift to the partial rough path space

We now construct an (α, β) rough path, which plays an important role in this paper. For notational simplicity we focus on a low dimensional case (both κ and W below are one-dimensional) but extensions to higher dimensional cases are straightforward. The proof is deferred to Section 1.4.2. Let $\kappa : (0, T] \to [0, \infty)$ as

$$\kappa(t) := g(t)t^{\zeta-\gamma}, \quad t \in (0, T],$$

where $\gamma, \zeta \in (0, 1)$ and g is a Lipschitz function. For example, the Riemann–Liouville kernel

$$\kappa_H(t) := \frac{t^{H-1/2}}{\Gamma(H + 1/2)}, \quad t \in (0, T], \ H \in (0, 1/2) \tag{1.4}$$

has the above form ($\zeta = H - \delta, \gamma = 1/2 - \delta, g(t) = 1/\Gamma(H + 1/2)$, where $\delta \in (0, 1/2)$). For $\alpha \in (0, 1]$, let $C^{\alpha\text{-Hld}}$ denote the space of α-Hölder continuous functions on $[0, T]$. Let $\mathcal{K} : C^{\gamma\text{-Hld}} \to C^{\zeta\text{-Hld}}$ as

$$\begin{aligned}\mathcal{K}f(t) &:= \lim_{\epsilon\to 0} \left\{ [\kappa(t - \cdot)(f(\cdot) - f(t))]_0^{t-\epsilon} + \int_0^{t-\epsilon} (f(s) - f(t))\kappa'(t - s)\mathrm{d}s \right\} \\ &= \kappa(t)(f(t) - f(0)) + \int_0^t (f(s) - f(t))\kappa'(t - s)\mathrm{d}s.\end{aligned}$$

Proposition 1.3.1. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geqq 0})$ be a filtered probability space, and fix $\alpha \in (1/3, 1/2]$, $\beta \in (0, 1/2)$, and $\gamma, \zeta \in (0, 1)$ with $\gamma < 1/2$, $\beta < \zeta$. Suppose that $X = (X^1, ..., X^d)$ is a d-dimensional (possibly correlated) Brownian motion, and W is a one-dimensional Brownian motion possibly correlated to X. Using the Itô integration, define $\hat{X}$, $X^{(i)}$, and $\mathbf{X}^{(jk)}$ as follows: for $(s, t) \in \Delta_T$, $i \in I$ and $(j, k) \in J$,

$$\hat{X}^{(1)}_{st} := \mathcal{K}W(t) - \mathcal{K}W(s),$$

$$\hat{X}^{(2)}_{st} := t^\zeta - s^\zeta,$$

$$X^{(i)}_{st} := \frac{1}{i!} \int_s^t \left(\hat{X}_{sr}\right)^i \mathrm{d}X_r, \quad \mathbf{X}^{(jk)}_{st} := \frac{1}{k!} \int_s^t \left(\hat{X}_{sr}\right)^k X^{(j)}_{sr} \otimes \mathrm{d}X_r.$$

Let $\kappa_{st}(r) := \left(\kappa(t-r) - \kappa(s-r)1_{(0,s)}(r)\right) 1_{(0,t)}(r)$ and assume that

$$||\kappa_{st}||^2_{L^2(\mathbb{R}_+)} \lesssim C|t-s|^{2(\zeta-\gamma)+1}.$$

Then we have the following.

(i) For a.s. $\omega \in \Omega$, $\mathbb{X}(\omega) := \left(\hat{X}(\omega), X^{(i)}(\omega), \mathbf{X}^{(jk)}(\omega)\right)_{i\in I,(j,k)\in J}$ is an (α, β) rough path.

(ii) It holds that

$$\left(\int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X}\right)^{(1)}_{0t} = \int_0^t f(\hat{X}_{0r})\mathrm{d}X_r \quad a.s.,$$

where the left-hand side is the first level of the (α, β) rough path integral and the right-hand side is the Itô integral.

1.3.2 The large deviation principle on $\Omega_{(\alpha,\beta)\text{-Hld}}$

We now discuss the LDP on $\Omega_{(\alpha,\beta)\text{-Hld}}$. Following [62, 42], we use Garcia's theorem [39]. Let $(W, W^{\perp})$ be a two-dimensional standard Brownian motion and $X := \rho W + \sqrt{1-\rho^2}W^{\perp}$, $\rho \in [-1, 1]$. Define $\hat{X}, X^{(i)}, \mathbf{X}^{(jk)}$ as in Proposition 1.3.1 with $d = 1$, $e = 2$. We state our second main result, the proof of which is given in Section 1.4.3.

Theorem 1.3.2. Let $\mathbb{X} = (\hat{X}, X^{(i)}, \mathbf{X}^{(jk)})$ be the random variable taking values on $(\Omega_{(\alpha,\beta)\text{-Hld}}, d_{(\alpha,\beta)})$ defined as above. Then, the sequence of triplets

$$\mathbb{X}^{\epsilon} := \left(\epsilon^{1/2}\hat{X}, \epsilon^{(|i|+1)/2}X^{(i)}, \epsilon^{(|j+k|+2)/2}\mathbf{X}^{(jk)}\right)$$

satisfies the LDP on $(\Omega_{(\alpha,\beta)\text{-Hld}}, d_{(\alpha,\beta)})$ with speed ϵ^{-1} with good rate function

$$\tilde{I}^{\#\#}(\hat{x}, x^{(i)}, \mathbf{x}^{(jk)}) := \inf\left\{\tilde{I}^{\#}(\tilde{v}) \middle| \tilde{v} \in \mathcal{H},\ (\hat{x}, x^{(i)}, \mathbf{x}^{(jk)}) = \mathbb{L} \circ \mathbb{K}(\tilde{v})\right\},$$

where $\mathcal{H}$ is the Cameron–Martin space from $[0, T]$ to $\mathbb{R}^2$,

$$\mathbb{K}(\tilde{v}) := \left(\left(\int_0^{\cdot} \kappa(\cdot - r)\mathrm{d}\tilde{v}^{(1)}_r, 0\right), \rho\tilde{v}^{(1)} + \sqrt{1-\rho^2}\tilde{v}^{(2)}\right)$$

and

$$\mathbb{L}(u, v) := (\delta u, u \cdot v, u * v), \quad u, v \in C_{[0,T]},\ v \in \mathcal{H},$$

$\delta u_{st} := u_t - u_s$, $u \cdot v = (u \cdot_i v)$, $u * v = (u *_{jk} v)$, and

$$(u \cdot_i v)_{st} := \int_s^t (u_r - u_s)^i \mathrm{d}v_r, \quad (u *_{jk} v)_{st} := \int_s^t (u \cdot_j v)_{sr}(u_r - u_s)^k \mathrm{d}v_r.$$

Here, $\tilde{I}^{\#} : C \to [0, \infty)$ is the rate function of two-dimensional Brownian motion:

$$\tilde{I}^{\#}(\tilde{v}) := \begin{cases} \frac{1}{2}||\tilde{v}||^2_{\mathcal{H}}, & \tilde{v} \in \mathcal{H}, \\ \infty, & \text{otherwise.} \end{cases}$$

Theorem 1.3.3. The sequence of the processes $\left\{Y^{\epsilon} := \int f(\hat{\mathbb{X}}^{\epsilon})\mathrm{d}\mathbb{X}^{\epsilon}\right\}_{\epsilon \geqq 0}$ satisfies the LDP on $(\Omega_{\alpha\text{-Hld}}, d_{\alpha})$ with speed ϵ^{-1} with good rate function

$$\tilde{I}^{\#\#\#}(y) := \inf\left\{\tilde{I}^{\#\#}(\mathbb{X}) \middle| \mathbb{X} \in \Omega_{(\alpha,\beta)\text{-Hld}},\ y = \int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X}\right\}$$
$$= \inf\left\{\tilde{I}^{\#}(\tilde{v}) \middle| \tilde{v} \in \mathcal{H},\ (u,v) = \mathbb{K}(\tilde{v}),\ y = \int f(\hat{\mathbb{L}}(u,v))\mathrm{d}\mathbb{L}(u,v)\right\},$$

where $\tilde{I}^{\#\#}$ is defined in Theorem 1.3.2.

Proof. By Theorems 1.2.6 and 1.3.2 together with the contraction principle, we have the claim. □

1.3.3 Rough differential equations and their LDP

We now discuss the following type of rough differential equation (RDE) (in Lyons' sense; see Section 8.8 of [27], for example):

$$\bar{S}_t = \int_0^t \bar{\sigma}(\bar{S}_u)\mathrm{d}Y_u, \tag{1.5}$$

where $\bar{S}_t = S_t - S_0$, $\bar{\sigma}(s) = \sigma(S_0 + s)$ and

$$Y = \int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X} \in \Omega_{\alpha\text{-Hld}}([0,T],\mathbb{R}^d),\ \ \mathbb{X} \in \Omega_{(\alpha,\beta)\text{-Hld}}. \tag{1.6}$$

Theorem 1.3.4. Let $\sigma \in C_b^3$.

(i) RDE (1.5) driven by (1.6) has a unique solution $\Phi(Y) = (Y, \bar{S})$, where

$$\Phi : \Omega_{\alpha\text{-Hld}}([0,T],\mathbb{R}^d) \times \mathbb{R} \to \Omega_{\alpha\text{-Hld}}([0,T],\mathbb{R}^{d+1})$$

is the solution map of (1.5) that is locally Lipschitz continuous with respect to d_{α}.

(ii) The first level of the last component $\bar{S}$ of the solution to RDE (1.5) for (1.6) with $\mathbb{X} = \mathbb{X}(\omega)$ defined in Proposition 1.3.1 gives the solution $S(\omega) = S_0 + \bar{S}$ to the Itô SDE (1.1).

Proof. (i) is a standard result from rough path theory; see e.g., Theorem 1 in [65] or Chapter 8 in [27]. (ii) follows from Proposition 1.3.1; see Chapter 9 in [27]. □

Theorem 1.3.5. Let $\sigma \in C_b^3$ and $\bar{S}^{\epsilon} := \Phi(Y^{\epsilon})$, where Φ is the solution map of Theorem 1.3.4. Then the sequence of the processes $\{\bar{S}^{\epsilon}\}_{\epsilon \geqq 0}$ satisfies the LDP on $\Omega_{\alpha\text{-Hld}}$ with speed ϵ^{-1} with good rate function

$$\tilde{I}(\bar{s}) := \inf\left\{\tilde{I}^{\#\#\#}(Y) \middle| Y \in \Omega_{\alpha\text{-Hld}},\ \bar{s} = \Phi(Y)\right\}$$
$$= \inf\left\{\tilde{I}^{\#}(\tilde{v}) \middle| \tilde{v} \in \mathcal{H},\ (u,v) = \mathbb{K}(\tilde{v}),\ \bar{s} = \int \bar{\sigma}(\bar{s}) f(\hat{\mathbb{L}}(u,v))\mathrm{d}\mathbb{L}(u,v)\right\}.$$

Proof. Because the solution map Φ is continuous, Theorem 1.3.4 and the contraction theorem imply the claim. □

1.3.4 Short-time asymptotics

We consider the case $\kappa = \kappa_H$ (see (1.4)). By the scaling property of the Riemann–Liouville fractional Brownian motion $\hat{X}$ and the standard Brownian motion X, we have

$$\hat{X}_{\epsilon t} \sim \epsilon^H \hat{X}_t, \quad X_{\epsilon t} \sim \epsilon^{1/2} X_t.$$

This implies

$$\tilde{Y}^\epsilon_t := \epsilon^{H-1/2} \int_0^{\epsilon t} f(\hat{X}_u) \mathrm{d}X_u \sim \int_0^t f(\hat{X}^\epsilon_u) \mathrm{d}X^\epsilon_u,$$

where $(\hat{X}^\epsilon, X^\epsilon) = \epsilon^H (\hat{X}, X)$, of which the rough path lift is $\mathbb{X}^\epsilon$ of Theorem 1.3.2. Letting

$$\tilde{S}^\epsilon_t = \frac{S_{\epsilon t} - S_0}{\epsilon^{1/2-H}}, \quad \tilde{\sigma}^\epsilon(s) = \sigma(S_0 + \epsilon^{1/2-H} s),$$

we have

$$\tilde{S}^\epsilon_t = \int_0^t \tilde{\sigma}^\epsilon(\tilde{S}^\epsilon_u) \mathrm{d}\tilde{Y}^\epsilon_u,$$

and we can derive an LDP for $\tilde{S}^\epsilon$ by an extended contraction principle [72].

Theorem 1.3.6. Let $\sigma \in C^3_b$. Then $\{\tilde{S}^\epsilon\}_{0<\epsilon\leqq 1}$ satisfies the LDP on $\Omega_{\alpha\text{-Hld}}$ as $\epsilon \to 0$ with speed ϵ^{-2H} with good rate function

$$\tilde{J}(\tilde{s}) := \inf \left\{ \tilde{I}^\#(\tilde{v}) \middle| \tilde{v} \in \mathcal{H}, \ (u, v) = \mathbb{K}(\tilde{v}), \ \tilde{s} = \sigma(S_0) \int f(\hat{\mathbb{L}}(u, v)) \mathrm{d}\mathbb{L}(u, v) \right\}.$$

Proof. Denote by Φ_ϵ the solution map of the RDE (1.5) with $\bar{\sigma} = \tilde{\sigma}^\epsilon$. We are going to show that Φ_ϵ is locally equicontinuous. Because for all $i \in \mathbb{Z}_+$,

$$||\nabla^i \tilde{\sigma}^\epsilon||_\infty \leqq (1+\epsilon)^i ||\nabla^i \sigma||_\infty \leqq 2^i ||\nabla^i \sigma||_\infty,$$

the local Lipschitz constants of Φ_ϵ can be taken uniformly in ϵ by Theorem 4 in [65]. Therefore Φ_ϵ is equicontinuous on bounded sets, and we conclude $\Phi_\epsilon(Y_\epsilon) \to \Phi_0(Y)$ for any converging sequence $Y_\epsilon \to Y$ for any Y with $\tilde{I}^{\#\#\#}(Y) < \infty$. Then by Theorem 1.3.3 and an extended contraction principle [72][Theorem 2.1], we have the desired results. □

Remark 1.3.7. By the usual argument, adding a drift term to the above RDE is straightforward. The result then generalizes the existing LDP for the rough Bergomi model:

$$\mathrm{d} \log S_t = -\frac{1}{2} f^2(\hat{X}_t) \mathrm{d}t + f(\hat{X}_t) \mathrm{d}X_t$$

in [23, 8, 48, 62, 42]. To deal with the mixed rough Bergomi model [15] or the two-factor fractional volatility model [38], we need an extension with higher dimensional κ and W that is also straightforward.

An LDP for the marginal distribution $\tilde{S}^\epsilon_1$ follows from the contraction principle, and the corresponding one-dimensional rate function extends the one obtained by [23] as follows.

Theorem 1.3.8. Assume $\sigma \in C_b^3$ and $|\rho| < 1$. Then $t^{H-1/2}\bar{S}_t$ satisfies the LDP as $t \to 0$ with speed t^{-2H} with good rate function

$$\tilde{J}^{\#}(z) := \inf_{g \in L^2([0,1])} \left[\frac{1}{2} \int_0^1 |g_r|^2 \mathrm{d}r + \frac{\left\{z - \rho\sigma(S_0) \int_0^1 f\left(K_H g(r), 0\right) g_r \mathrm{d}r\right\}^2}{2(1-\rho^2)\sigma(S_0)^2 \int_0^1 f\left(K_H g(r), 0\right)^2 \mathrm{d}r} \right],$$

where $K_H g(t) = \int_0^t \kappa_H(t-r) g_r \mathrm{d}r$.

Proof. See Section 1.6. □

A short-time asymptotic formula of the implied volatility (regarding S as a price or a log-price process) then follows from Theorem 1.3.8 as in [23]. From the rate function of Theorem 1.3.8, we observe that the effect of the function σ to the short-time asymptotics is only through the constant $\sigma(S_0)$. In particular, the local volatility function σ does not add any flexibility to the asymptotic shape of the implied volatility surface.

1.4 Proofs of main theorems

1.4.1 Proof of Theorem 1.2.6

Proof. By a localizing argument, we can assume without loss of generality that the derivatives of f are bounded. For brevity, let $K := \|f\|_{C_b^{n+2}}$ and $M := |||\mathbb{X}|||_{(\alpha,\beta)}$. Let

$$J^{(1)}_{st} = J^{(1)}(\mathbb{X})_{st} := \sum_{i \in I} \partial^i f(\hat{x}_s) X^{(i)}_{st}, \quad J^{(2)}_{st} = J^{(2)}(\mathbb{X})_{st} := \sum_{(j,k) \in J} \partial^j f(\hat{x}_s) \partial^k f(\hat{x}_s) \mathbf{X}^{(jk)}_{st},$$

where $\hat{x}_s := \hat{X}_{0s}$. Below, we follow the standard argument of rough path theory with Chen's identity replaced by our modified version (1.2), (1.3).

(Claim 1) The first level of the (α, β) rough path integral $Y^{(1)}_{st}$ is well-defined and has the following inequality:

$$|Y^{(1)}_{st}| \leqq K C_1 |t-s|^{\alpha}, \tag{1.7}$$

where

$$C_1 := (n+1)^{2e}(1+M)^{n+2}(1+T)^{(n+1)\beta}\left\{1 + 2^{(n+1)\beta+\alpha}\zeta((n+1)\beta+\alpha)\right\},$$

and $\zeta(r) := \sum_{p=1}^{\infty} \frac{1}{p^r}$.

Proof. By Taylor expansion, we have

$$\begin{aligned}\sum_{i\in I}\partial^i f(\hat{x}_u)X^{(i)}_{ut} &= \sum_{i\in I}\left\{\sum_{|p|\leqq n-|i|}\frac{1}{p!}\partial^{i+p}f(\hat{x}_s)\left(\hat{X}_{su}\right)^p X^{(i)}_{ut} + R_i X^{(i)}_{ut}\right\}\\ &= \sum_{i\in I}\partial^i f(\hat{x}_s)\left\{\sum_{p\leqq i}\frac{1}{(i-p)!}\left(\hat{X}_{su}\right)^{i-p}X^{(p)}_{ut}\right\} + \sum_{i\in I}R_i X^{(i)}_{ut},\end{aligned} \tag{1.8}$$

where

$$\begin{aligned}R_i &= R(\mathbb{X})_i\\ &= \sum_{|p|=n+1-|i|}\left(\int_0^1\frac{(1-\theta)^{n+1-|i|}(n+1-|i|)}{p!}\partial^p f(\hat{x}_s+\theta\hat{X}_{su})\mathrm{d}\theta\right)(\hat{X}_{su})^p\end{aligned} \tag{1.9}$$

By the modified Chen's relation (1.2) and (1.8), for any $s \leqq u \leqq t$,

$$\begin{aligned}&J^{(1)}_{su} + J^{(1)}_{ut} - J^{(1)}_{st}\\ &= \sum_{i\in I}\partial^i f(\hat{x}_s)\left(X^{(i)}_{su} - X^{(i)}_{st}\right) + \sum_{i\in I}\partial^i f(\hat{x}_u)X^{(i)}_{ut}\\ &= -\sum_{i\in I}\partial^i f(\hat{x}_s)\left\{\sum_{p\leqq i}\frac{1}{(i-p)!}\left(\hat{X}_{su}\right)^{i-p}X^{(p)}_{ut}\right\} + \sum_{i\in I}\partial^i f(\hat{x}_u)X^{(i)}_{ut}\\ &= \sum_{i\in I}R_i X^{(i)}_{ut}.\end{aligned} \tag{1.10}$$

Because for all $i \in I$,

$$\left|R_i X^{(i)}_{ut}\right| \leqq K\sum_{|p|=n+1-|i|}\left|(\hat{X}_{su})^p X^{(i)}_{ut}\right| \leqq K(n+1)^e(1+M)^{n+2}|t-s|^{(n+1)\beta+\alpha},$$

we have

$$\left|J^{(1)}_{su} + J^{(1)}_{ut} - J^{(1)}_{st}\right| \leqq K(n+1)^{2e}(1+M)^{n+2}|t-s|^{(n+1)\beta+\alpha}.$$

For any partition $\mathcal{P} = \{s = t_0 < t_1 < \ldots < t_N = t\}$, let $J^{(1)}_{st}(\mathcal{P}) := \sum_{p=1}^N J^{(1)}_{t_{p-1}t_p}$. By Lemma 1.5.1, there exists $p \in \{1, \ldots, N\}$ such that

$$|t_{p+1} - t_{p+1}| \leqq \frac{2}{N-1}|t-s|. \tag{1.11}$$

Then we have

$$\begin{aligned}&\left|J^{(1)}_{st}(\mathcal{P}) - J^{(1)}_{st}(\mathcal{P}\backslash\{t_p\})\right|\\ &= \left|J^{(1)}_{t_{p-1}t_p} + J^{(1)}_{t_p t_{p+1}} - J^{(1)}_{t_{p-1}t_{p+1}}\right|\\ &\leqq K(n+1)^{2e}(1+M)^{n+2}|t_{p+1} - t_{p-1}|^{(n+1)\beta+\alpha}\\ &\leqq K(n+1)^{2e}(1+M)^{n+2}\left(\frac{2}{N-1}\right)^{(n+1)\beta+\alpha}|t-s|^{(n+1)\beta+\alpha},\end{aligned}$$

and this implies (note that $(n+1)\beta+\alpha > 1$)

$$\left|J_{st}^{(1)}(\mathcal{P}) - J_{st}^{(1)}\right| \\ \leqq K(n+1)^{2e}(1+M)^{n+2}2^{(n+1)\beta+\alpha}\zeta\left((n+1)\beta+\alpha\right)|t-s|^{(n+1)\beta+\alpha}. \quad (1.12)$$

(Claim 1a) $\{J_{st}^{(1)}(\mathcal{P})\}_{\mathcal{P}}$ is a Cauchy sequence with $|\mathcal{P}| \searrow 0$.

Let $\mathcal{Q}$ be any subdivision of $\mathcal{P}$: $\mathcal{Q} = \{s = \tau_0 < \tau_1 < \ldots < \tau_L = t,\}, L > N$. Consider the subsequence $\{\tau_{l_0} < \tau_{l_1} < \ldots < \tau_{l_N}\}$ with $\tau_{l_p} = t_p$, and let $\mathcal{Q}_p := \mathcal{Q} \cap [t_{p-1}, t_p]$. Then $\mathcal{Q}_p$ is a partition of $[t_{p-1}, t_p]$. By using (1.12), we have that

$$\begin{aligned}
&\left|J_{st}^{(1)}(\mathcal{Q}) - J_{st}^{(1)}(\mathcal{P})\right| \\
&\leqq \sum_{p=1}^{N}\left|J_{st}^{(1)}(\mathcal{Q}_p) - J_{t_{p-1}t_p}^{(1)}\right| \\
&\leqq K(n+1)^{2e}(1+M)^{n+2}2^{(n+1)\beta+\alpha}\zeta\left((n+1)\beta+\alpha\right)\sum_{p=1}^{N}|t_p - t_{p-1}|^{(n+1)\beta+\alpha} \\
&\leqq K(n+1)^{2e}(1+M)^{n+2}2^{(n+1)\beta+\alpha}\zeta\left((n+1)\beta+\alpha\right)T\left(\sup_{t-s\leqq|\mathcal{P}|}|t-s|^{(n+1)\beta+\alpha-1}\right).
\end{aligned}$$

Hence for any partition $\mathcal{P}, \mathcal{P}'$ with $|\mathcal{P}| \vee |\mathcal{P}'| \leqq \delta$, we have that

$$\begin{aligned}
&\left|J_{st}^{(1)}(\mathcal{P}) - J_{st}^{(1)}(\mathcal{P}')\right| \\
&\leqq \left|J_{st}^{(1)}(\mathcal{P}) - J_{st}^{(1)}(\mathcal{P}\cup\mathcal{P}')\right| + \left|J_{st}^{(1)}(\mathcal{P}\cup\mathcal{P}') - J_{st}^{(1)}(\mathcal{P}')\right| \\
&\leqq K(n+1)^{2e}(1+M)^{n+2}2^{(n+1)\beta+\alpha+1}\zeta\left((n+1)\beta+\alpha\right)T\left(\sup_{t-s\leqq\delta}|t-s|^{(n+1)\beta+\alpha-1}\right),
\end{aligned}$$

and because $(n+1)\beta+\alpha > 1$, we conclude that $\{J_{st}^{(1)}(\mathcal{P})\}_{\mathcal{P}}$ is a Cauchy sequence. Therefore, $Y_{st}^{(1)}$ is well-defined. Furthermore, by (1.12), we have

$$|Y_{st}^{(1)}| \leqq |J_{st}^{(1)}| + |Y_{st}^{(1)} - J_{st}^{(1)}| \leqq KC_1|t-s|^{\alpha}.$$

Thus we have proved the statement of Claim 1. □

(Claim 2) Let $m := \max_{(j,k)\in J}|j+k|$. Then the second level of the (α,β) rough path integral $Y_{st}^{(2)}$ is well-defined and has the following inequality:

$$|Y_{st}^{(2)}| \leqq K^2C_2|t-s|^{2\alpha},$$

where

$$C_2 := (1+m)^{2e}M(1+T)^{m\beta} + \left(\tilde{C}_2 + 2C_1^2T^{(n-m)\beta}\right)2^{(m+1)\beta+2\alpha}\zeta\left((m+1)\beta+2\alpha\right),$$

and

$$\tilde{C}_2 := 2(1+n+m)^{4e}(1+M)^{m+3}(1+T)^{(2n-m-1)\beta}.$$

In particular, we have $\int f(\hat{\mathbb{X}})d\mathbb{X} \in \Omega_{\alpha\text{-Hld}}$.

Proof. By the modified Chen's relation (1.3), for all $s \leqq u \leqq t$,

$$\begin{aligned}
&J^{(2)}_{su} + J^{(2)}_{ut} + J^{(1)}_{su} \otimes J^{(1)}_{ut} - J^{(2)}_{st} \\
&= J^{(1)}_{su} \otimes J^{(1)}_{ut} + \sum_{(j,k)\in J} \left[\partial^j f(\hat{x}_s)\partial^k f(\hat{x}_s)\left(\mathbf{X}^{(jk)}_{su} - \mathbf{X}^{(jk)}_{st}\right) + \partial^j f(\hat{x}_u)\partial^k f(\hat{x}_u)\mathbf{X}^{(jk)}_{ut}\right] \\
&=: S_1 + S_2,
\end{aligned}$$

where

$$\begin{aligned}
&S_1 = S_1(\mathbb{X}) \\
&:= J^{(1)}_{su} \otimes J^{(1)}_{ut} - \sum_{(j,k)\in J} \partial^j f(\hat{x}_s)\partial^k f(\hat{x}_s)\left(\sum_{q\leqq k} \frac{1}{(k-q)!}\left(\hat{X}_{su}\right)^{k-q} X^{(j)}_{su} \otimes X^{(q)}_{ut}\right)
\end{aligned}$$

and

$$\begin{aligned}
&S_2 = S_2(\mathbb{X}) \\
&:= \sum_{(j,k)\in J} \partial^j f(\hat{x}_u)\partial^k f(\hat{x}_u)\mathbf{X}^{(jk)}_{ut} \\
&\quad - \sum_{(j,k)\in J} \partial^j f(\hat{x}_s)\partial^k f(\hat{x}_s)\left(\sum_{p\leqq j}\sum_{q\leqq k} \frac{1}{(j-p)!(k-q)!}\left(\hat{X}_{su}\right)^{j+k-p-q}\mathbf{X}^{(pq)}_{ut}\right).
\end{aligned}$$

Note that

$$\begin{aligned}
J^{(1)}_{su} \otimes J^{(1)}_{ut} &= \left(\sum_{j\in I} \partial^j f(\hat{x}_s) X^{(j)}_{su}\right) \otimes \left(\sum_{k\in I} \partial^k f(\hat{x}_u) X^{(k)}_{ut}\right) \\
&= \sum_{j\in I}\sum_{k\in I} \partial^j f(\hat{x}_s)\partial^k f(\hat{x}_u) X^{(j)}_{su} \otimes X^{(k)}_{ut} \\
&= \sum_{(j,k)\in J} \partial^j f(\hat{x}_s)\partial^k f(\hat{x}_u) X^{(j)}_{su} \otimes X^{(k)}_{ut} + \sum_{(j,k)\in I\times I\setminus J} \partial^j f(\hat{x}_s)\partial^k f(\hat{x}_u) X^{(j)}_{su} \otimes X^{(k)}_{ut}.
\end{aligned}$$

By Taylor expansion, we have

$$\begin{aligned}
&\sum_{(j,k)\in J} \partial^j f(\hat{x}_s)\partial^k f(\hat{x}_u) X^{(j)}_{su} \otimes X^{(k)}_{ut} \\
&= \sum_{(j,k)\in J}\left\{\sum_{|q|\leqq m-|j+k|} \frac{1}{q!}\partial^j f(\hat{x}_s)\partial^{k+q} f(\hat{x}_s)(\hat{X}_{su})^q X^{(j)}_{su} \otimes X^{(k)}_{ut} + \partial^j f(\hat{x}_s) R^{(1)}_{jk} X^{(j)}_{su} \otimes X^{(k)}_{ut}\right\} \\
&= \sum_{(j,k)\in J}\left\{\partial^j f(\hat{x}_s)\partial^k f(\hat{x}_s)\left(\sum_{q\leqq k} \frac{1}{(k-q)!}\left(\hat{X}_{su}\right)^{k-q} X^{(j)}_{su} \otimes X^{(q)}_{ut}\right) + \partial^j f(\hat{x}_s) R^{(1)}_{jk} X^{(j)}_{su} \otimes X^{(k)}_{ut}\right\},
\end{aligned}$$

where

$$R_{jk}^{(1)} = R_{jk}^{(1)}(\mathbb{X})$$
$$:= \sum_{|q|=m+1-|j+k|} \left(\int_0^1 \frac{(1-\theta)^{m+1-|j+k|}(m+1-|j+k|)}{q!} \partial^{k+q} f(\hat{x}_s + \theta \hat{X}_{su}) \mathrm{d}\theta \right) (\hat{X}_{su})^q,$$

and so we obtain that

$$\begin{aligned} S_1 &= J_{su}^{(1)} \otimes J_{ut}^{(1)} - \sum_{(j,k)\in J} \partial^j f(\hat{x}_s) \partial^k f(\hat{x}_s) \left(\sum_{q \leqq k} \frac{1}{(k-q)!} \left(\hat{X}_{su}\right)^{k-q} X_{su}^{(j)} \otimes X_{ut}^{(q)} \right) \\ &= \sum_{(j,k)\in J} \partial^j f(\hat{x}_s) R_{jk}^{(1)} X_{su}^{(j)} \otimes X_{ut}^{(k)} + \sum_{(j,k)\in I\times I\setminus J} \partial^j f(\hat{x}_s) \partial^k f(\hat{x}_u) X_{su}^{(j)} \otimes X_{ut}^{(k)} \end{aligned} \tag{1.13}$$

and

$$\begin{aligned} &|S_1| \\ &\leqq \sum_{(j,k)\in J} \left| \partial^j f(\hat{x}_s) R_{jk}^{(1)} X_{su}^{(j)} \otimes X_{ut}^{(k)} \right| + \sum_{(j,k)\in I\times I\setminus J} \left| \partial^j f(\hat{x}_s) \partial^k f(\hat{x}_u) X_{su}^{(j)} \otimes X_{ut}^{(k)} \right| \\ &\leqq \sum_{(j,k)\in J} \sum_{|q|=m+1-|j+k|} K^2 |(\hat{X}_{su})^q X_{su}^{(j)} \otimes X_{ut}^{(k)}| + \sum_{(j,k)\in I\times I\setminus J} K^2 |X_{su}^{(j)} \otimes X_{ut}^{(k)}| \\ &\leqq K^2(1+m)^{3e}(1+M)^{m+3}|t-s|^{(m+1)\beta+2\alpha} \\ &\quad + K^2(1+n)^{2e}(1+M)^2(1+T)^{(2n-m-1)\beta}|t-s|^{(m+1)\beta+2\alpha} \\ &\leqq 2K^2(1+n+m)^{3e}(1+M)^{m+3}(1+T)^{(2n-m-1)\beta}|t-s|^{(m+1)\beta+2\alpha}. \end{aligned} \tag{1.14}$$

Here we use $m \leqq n$ (because $(n+1)\beta + \alpha > 1$, we have $(n+1)\beta + 2\alpha > 1$, and the definition of m implies $m \leqq n$).

On the other hand, one can show that

$$S_2 = \sum_{(j,k)\in J} \left\{ \sum_{|p|\leqq m-|j+k|} \frac{1}{p!} \partial^{j+p} f(\hat{x}_s) R_{jkp}^{(3)} (\hat{X}_{su})^p \mathbf{X}_{ut}^{(jk)} + \partial^k f(\hat{x}_u) R_{jk}^{(2)} \mathbf{X}_{ut}^{(jk)} \right\} \tag{1.15}$$

by using the Taylor expansion

$$
\begin{aligned}
&\sum_{(j,k)\in J} \partial^j f(\hat{x}_u)\partial^k f(\hat{x}_u)\mathbf{X}_{ut}^{(jk)} \\
&= \sum_{(j,k)\in J} \left\{ \sum_{|p|\leqq m-|j+k|} \frac{1}{p!}\partial^{j+p} f(\hat{x}_s)(\hat{X}_{su})^p + R_{jk}^{(2)} \right\} \partial^k f(\hat{x}_u)\mathbf{X}_{ut}^{(jk)} \\
&= \sum_{(j,k)\in J} \sum_{|p|\leqq m-|j+k|} \sum_{|q|\leqq m-|j+k+p|} \frac{1}{p!q!}\partial^{j+p} f(\hat{x}_s)\partial^{k+q} f(\hat{x}_s)(\hat{X}_{su})^{p+q}\mathbf{X}_{ut}^{(jk)} \\
&\quad + \sum_{(j,k)\in J} \sum_{|p|\leqq m-|j+k|} \frac{1}{p!}\partial^{j+p} f(\hat{x}_s) R_{jkp}^{(3)}(\hat{X}_{su})^p \mathbf{X}_{ut}^{(jk)} \\
&\quad + \sum_{(j,k)\in J} R_{jk}^{(2)}\partial^k f(\hat{x}_u)\mathbf{X}_{ut}^{(jk)},
\end{aligned}
$$

where

$$
\begin{aligned}
&R_{jk}^{(2)} = R_{jk}^{(2)}(\mathbb{X}) \\
&:= \sum_{|p|=m+1-|j+k|} \left(\int_0^1 \frac{(1-\theta)^{m+1-|j+k|}(m+1-|j+k|)}{p!} \partial^{j+p} f(\hat{x}_s+\theta\hat{X}_{su})\mathrm{d}\theta \right) (\hat{X}_{su})^p,
\end{aligned}
$$

$$
\begin{aligned}
&R_{jkp}^{(3)} = R_{jkp}^{(3)}(\mathbb{X}) \\
&:= \sum_{|q|=m+1-|j+k+p|} \left(\int_0^1 \frac{(1-\theta)^{m+1-|j+k+p|}(m+1-|j+k+p|)}{q!} \partial^{k+q} f(\hat{x}_s+\theta\hat{X}_{su})\mathrm{d}\theta \right) (\hat{X}_{su})^q.
\end{aligned}
$$

Because for all $(j,k)\in J$ and $0 \leqq |p| \leqq m-|j+k|$,

$$
\begin{aligned}
\left|\partial^{j+p} f(\hat{x}_s) R_{jkp}^{(3)}(\hat{X}_{su})^p \mathbf{X}_{ut}^{(jk)}\right| &\leqq K^2 \sum_{|q|=m+1-|j+k+p|} |(\hat{X}_{su})^{p+q}\mathbf{X}_{ut}^{(jk)}| \\
&\leqq K^2(1+m)^e(1+M)^{m+2}|t-s|^{(m+1)\beta+2\alpha},
\end{aligned}
$$

and

$$
\begin{aligned}
\left|R_{jk}^{(2)}\partial^k f(\hat{x}_u)\mathbf{X}_{ut}^{(jk)}\right| &\leqq K^2 \sum_{|p|=m+1-|j+k|} |(\hat{X}_{su})^p\mathbf{X}_{ut}^{(jk)}| \\
&\leqq K^2(1+m)^e(1+M)^{m+2}|t-s|^{(m+1)\beta+2\alpha},
\end{aligned}
$$

we have

$$
\begin{aligned}
&|S_2| \\
&\leqq \sum_{(j,k)\in J} \sum_{|p|\leqq m-|j+k|} |\partial^{j+p} f(\hat{x}_s) R_{jkp}^{(3)}(\hat{X}_{su})^p\mathbf{X}_{ut}^{(jk)}| + \sum_{(j,k)\in J} |\partial^k f(\hat{x}_u) R_{jk}^{(2)}\mathbf{X}_{ut}^{(jk)}| \\
&\leqq K^2(1+m)^{4e}(1+M)^{m+2}|t-s|^{(m+1)\beta+2\alpha} + K^2(1+m)^{3e}(1+M)^{m+1}|t-s|^{(m+1)\beta+2\alpha} \\
&\leqq 2K^2(1+m)^{4e}(1+M)^{m+2}|t-s|^{(m+1)\beta+2\alpha}. \qquad (1.16)
\end{aligned}
$$

By (1.14) and (1.16), we have

$$\left|J_{su}^{(2)} + J_{ut}^{(2)} + J_{su}^{(1)} \otimes J_{ut}^{(1)} - J_{st}^{(2)}\right| \leqq |S_1| + |S_2| \leqq K^2 \tilde{C}_2 |t-s|^{(m+1)\beta + 2\alpha},$$

where $\tilde{C}_2 = 2(1+n+m)^{4e}(1+M)^{m+3}(1+T)^{(2n-m-1)\beta}$. Moreover, by (1.7) and (1.12), we have

$$\begin{aligned}
&\left|Y_{su}^{(1)} \otimes Y_{ut}^{(1)} - J_{su}^{(1)} \otimes J_{ut}^{(1)}\right| \\
&\leqq \left|Y_{su}^{(1)}\right|\left|Y_{ut}^{(1)} - J_{ut}^{(1)}\right| + \left|Y_{su}^{(1)} - J_{su}^{(1)}\right|\left|J_{ut}^{(1)}\right| \leqq 2K^2 C_1^2 |t-s|^{(n+1)\beta + 2\alpha}.
\end{aligned}$$

Let $J_{st}^{(2)}(\mathcal{P}) := \sum_{p=1}^{n} Y_{t_0 t_{p-1}}^{(1)} \otimes Y_{t_{p-1}t_p}^{(1)} + J_{t_{p-1}t_p}^{(2)}$. By Lemma 1.5.1, there exists $p \in \{1, \ldots, N\}$ such that (1.11) holds. Note that $m \leqq n$. Then, the above inequalities imply that

$$\begin{aligned}
&\left|J_{st}^{(2)}(\mathcal{P}) - J_{st}^{(2)}(\mathcal{P}\backslash\{t_p\})\right| \\
&\leqq \left|J_{t_{p-1}t_p}^{(2)} + J_{t_p t_{p+1}}^{(2)} + Y_{t_{p-1}t_p}^{(1)} \otimes Y_{t_p t_{p+1}}^{(1)} - J_{t_{p-1}t_{p+1}}^{(2)}\right| \\
&\leqq \left|J_{t_{p-1}t_p}^{(2)} + J_{t_p t_{p+1}}^{(2)} + J_{t_{p-1}t_p}^{(1)} \otimes J_{t_p t_{p+1}}^{(1)} - J_{t_{p-1}t_{p+1}}^{(2)}\right| + \left|Y_{t_{p-1}t_p}^{(1)} \otimes Y_{t_p t_{p+1}}^{(1)} - J_{t_{p-1}t_p}^{(1)} \otimes J_{t_p t_{p+1}}^{(1)}\right| \\
&\leqq K^2 \tilde{C}_2 |t_{p+1} - t_{p-1}|^{(m+1)\beta + 2\alpha} + 2K^2 C_1^2 |t_{p+1} - t_{p-1}|^{(n+1)\beta + 2\alpha} \\
&\leqq K^2 \left(\tilde{C}_2 + 2C_1^2 T^{(n-m)\beta}\right)\left(\frac{2}{N-1}\right)^{(m+1)\beta + 2\alpha} |t-s|^{(m+1)\beta + 2\alpha}.
\end{aligned}$$

This implies that (note that $(m+1)\beta + 2\alpha > 1$)

$$|J_{st}^{(2)}(\mathcal{P}) - J_{st}^{(2)}| \leqq K^2 C_2 |t-s|^{(m+1)\beta + 2\alpha}. \tag{1.17}$$

This shows that $\{J_{st}^{(2)}(\mathcal{P})\}_{\mathcal{P}}$ is a Cauchy sequence when $|\mathcal{P}| \searrow 0$ (one can adapt the argument of Claim 1a in the proof of Claim 1 by using (1.17) instead of (1.12)). Hence, $Y_{st}^{(2)}$ is well-defined. We also obtain that

$$|Y_{st}^{(2)}| \leqq |J_{st}^{(2)}| + |Y_{st}^{(2)} - J_{st}^{(2)}| \leqq K^2 C_2 |t-s|^{2\alpha}.$$

Next, we prove that $\int f(\hat{\mathbb{X}}) \mathrm{d}\mathbb{X}$ satisfies Chen's relation. Fix $\epsilon > 0$ and $s < u < t$. By taking a partition $\mathcal{P} = \{s = t_0 < t_1 < \ldots < t_N = t\}$ of $[s,t]$ small enough (which has the point $t_{\tilde{N}} = u$), we have

$$\begin{aligned}
&\left|Y_{st}^{(1)} - Y_{su}^{(1)} - Y_{ut}^{(1)}\right| \\
&\leqq \left|Y_{st}^{(1)} - \sum_{p=1}^{N} J_{t_{p-1}t_p}^{(1)}\right| + \left|Y_{su}^{(1)} - \sum_{p=1}^{\tilde{N}} J_{t_{p-1}t_p}^{(1)}\right| + \left|Y_{ut}^{(1)} - \sum_{p=\tilde{N}+1}^{N} J_{t_{p-1}t_p}^{(1)}\right| \\
&\leqq 3\epsilon
\end{aligned}$$

and so the first level of $\int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X}$ satisfies Chen's relation. Note that this result implies that

$$\sum_{q=1}^{N} Y^{(1)}_{t_0 t_{q-1}} \otimes Y^{(1)}_{t_{q-1}t_q} = \sum_{0<p<q\leqq N} Y^{(1)}_{t_{p-1}t_p} \otimes Y^{(1)}_{t_{q-1}t_q}.$$

Note also that

$$\begin{aligned}
&Y^{(1)}_{su} \otimes Y^{(1)}_{ut} \\
&= \left(\sum_{p=1}^{\tilde{N}} Y^{(1)}_{t_{p-1}t_p}\right) \otimes \left(\sum_{q=\tilde{N}+1}^{N} Y^{(1)}_{t_{q-1}t_q}\right) \\
&= \sum_{0<p<q\leqq N} Y^{(1)}_{t_{p-1}t_p} \otimes Y^{(1)}_{t_{q-1}t_q} - \sum_{0<p<q\leqq \tilde{N}} Y^{(1)}_{t_{p-1}t_p} \otimes Y^{(1)}_{t_{q-1}t_q} - \sum_{\tilde{N}<p<q\leqq N} Y^{(1)}_{t_{p-1}t_p} \otimes Y^{(1)}_{t_{q-1}t_q} \\
&= \sum_{p=1}^{N} Y^{(1)}_{t_0 t_{p-1}} \otimes Y^{(1)}_{t_{p-1}t_p} - \sum_{p=1}^{\tilde{N}} Y^{(1)}_{t_0 t_{p-1}} \otimes Y^{(1)}_{t_{p-1}t_p} - \sum_{p=\tilde{N}+1}^{N} Y^{(1)}_{t_{\tilde{N}} t_{p-1}} \otimes Y^{(1)}_{t_{p-1}t_p},
\end{aligned}$$

and so we have

$$\begin{aligned}
&\left| Y^{(2)}_{st} - Y^{(2)}_{su} - Y^{(2)}_{ut} - Y^{(1)}_{su} \otimes Y^{(1)}_{ut} \right| \\
&\leqq \left| Y^{(2)}_{st} - \mathcal{S}_{st} \right| + \left| Y^{(2)}_{su} - \mathcal{S}_{su} \right| + \left| Y^{(2)}_{ut} - \mathcal{S}_{ut} \right| \leqq 3\epsilon,
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{S}_{st} &:= \sum_{p=1}^{N} \left(Y^{(1)}_{t_0 t_{p-1}} \otimes Y^{(1)}_{t_{p-1}t_p} + \sum_{(j,k)\in J} \partial^j f(\hat{x}_{t_{p-1}}) \partial^k f(\hat{x}_{t_{p-1}}) \mathbf{X}^{(jk)}_{t_{p-1}t_p} \right), \\
\mathcal{S}_{su} &:= \sum_{p=1}^{\tilde{N}} \left(Y^{(1)}_{t_0 t_{p-1}} \otimes Y^{(1)}_{t_{p-1}t_p} + \sum_{(j,k)\in J} \partial^j f(\hat{x}_{t_{p-1}}) \partial^k f(\hat{x}_{t_{p-1}}) \mathbf{X}^{(jk)}_{t_{p-1}t_p} \right), \\
\mathcal{S}_{ut} &:= \sum_{p=\tilde{N}+1}^{N} \left(Y^{(1)}_{t_M t_{p-1}} \otimes Y^{(1)}_{t_{p-1}t_p} + \sum_{(j,k)\in J} \partial^j f(\hat{x}_{t_{p-1}}) \partial^k f(\hat{x}_{t_{p-1}}) \mathbf{X}^{(jk)}_{t_{p-1}t_p} \right).
\end{aligned}$$

Therefore, the second level of $\int f(\hat{\mathbb{X}})\mathrm{d}\mathbb{X}$ also satisfies Chen's relation. The above argument proves statement (i) of Theorem 1.2.6. □

(Claim 3) Suppose that there exist $M > 0$ and $\epsilon > 0$ such that

$$|\hat{V}_{st}| \vee |\hat{W}_{st}| \leqq M|t-s|^{\beta}, \quad |V^{(i)}_{st}| \vee |W^{(i)}_{st}| \leqq M|t-s|^{|i|\beta+\alpha},$$

$$|\mathbf{V}^{(jk)}_{st}| \vee |\mathbf{W}^{(jk)}_{st}| \leqq M|t-s|^{|j+k|\beta+2\alpha}, \quad |\hat{V}_{st} - \hat{W}_{st}| \leqq \epsilon|t-s|^{\beta},$$

and

$$|V^{(i)}_{st} - W^{(i)}_{st}| \leqq \epsilon|t-s|^{|i|\beta+\alpha}, \quad |\mathbf{V}^{(jk)}_{st} - \mathbf{W}^{(jk)}_{st}| \leqq \epsilon|t-s|^{|j+k|\beta+2\alpha}.$$

Then, there exists $C_3 > 0$ such that

$$\left|\left(\int f(\hat{\mathbb{V}})\mathrm{d}\mathbb{V}\right)^{(1)}_{st} - \left(\int f(\hat{\mathbb{W}})\mathrm{d}\mathbb{W}\right)^{(1)}_{st}\right| \leqq K\epsilon C_3|t-s|^{\alpha}, \tag{1.18}$$

where

$$C_3 := (1+n)^{2e+1}(1+T)^{(n+1)\beta}\{1+(3e+2)(1+M)^{n+2}2^{(n+1)\beta+\alpha}\zeta((n+1)\beta+\alpha)\}.$$

Proof. By the assumption and the mean value theorem, we have

$$\begin{aligned}
&\left|J^{(1)}(\mathbb{V})_{st} - J^{(1)}(\mathbb{W})_{st}\right| \\
&= \left|\sum_{i\in I}\partial^i f(\hat{v}_s)V^{(i)}_{st} - \sum_{i\in I}\partial^i f(\hat{w}_s)W^{(i)}_{st}\right| \\
&\leqq \sum_{i\in I}\left\{|\partial^i f(\hat{v}_s) - \partial^i f(\hat{w}_s)||V^{(i)}_{st}| + |\partial^i f(\hat{w}_s)||V^{(i)}_{st} - W^{(i)}_{st}|\right\} \\
&\leqq K\epsilon(1+eM)(1+n)^e(1+T)^{(n+1)\beta}|t-s|^{\alpha}.
\end{aligned} \tag{1.19}$$

By (1.9), (1.10), and the mean value theorem, for all $s \leqq u \leqq t$,

$$\begin{aligned}
&\left|J^{(1)}(\mathbb{V})_{su} + J^{(1)}(\mathbb{V})_{ut} - J^{(1)}(\mathbb{V})_{st} - \left\{J^{(1)}(\mathbb{W})_{su} + J^{(1)}(\mathbb{W})_{ut} - J^{(1)}(\mathbb{W})_{st}\right\}\right| \\
&\leqq \sum_{i\in I}\left|R_i(\mathbb{V})V^{(i)}_{ut} - R_i(\mathbb{W})W^{(i)}_{ut}\right| \\
&\leqq \sum_{i\in I}\left|R_i(\mathbb{V}) - R_i(\mathbb{W})||V^{(i)}_{ut}| + |R_i(\mathbb{W})||V^{(i)}_{ut} - W^{(i)}_{ut}\right| \\
&\leqq (2e+1)K\epsilon(1+n)^{2e+1}(1+T)^{\beta}(1+M)^{n+2}|t-s|^{(n+1)\beta+\alpha} \\
&\quad + K\epsilon(1+n)^{2e}(1+M)^{n+1}|t-s|^{(n+1)\beta+\alpha} \\
&\leqq (2e+2)K\epsilon(1+n)^{2e+1}(1+T)^{\beta}(1+M)^{n+2}|t-s|^{(n+1)\beta+\alpha}.
\end{aligned}$$

By Lemma 1.5.1, there exists $p \in \{1, ..., N\}$ such that (1.11) holds. By the above inequality, we have that

$$\begin{aligned}
&\left|J^{(1)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(1)}(\mathbb{V})_{st}(\mathcal{P}\backslash\{t_p\}) - \left\{J^{(1)}(\mathbb{W})_{st}(\mathcal{P}) - J^{(1)}(\mathbb{W})_{st}\left(\mathcal{P}\backslash\{t_p\}\right)\right\}\right| \\
&= \Big|J^{(1)}(\mathbb{V})_{t_{p-1}t_p} + J^{(1)}(\mathbb{V})_{t_pt_{p+1}} - J^{(1)}(\mathbb{V})_{t_{p-1}t_{p+1}} \\
&\qquad - \left\{J^{(1)}(\mathbb{W})_{t_{p-1}t_p} + J^{(1)}(\mathbb{W})_{t_pt_{p+1}} - J^{(1)}(\mathbb{W})_{t_{p-1}t_{p+1}}\right\}\Big| \\
&\leqq (2e+2)K\epsilon(1+n)^{2e+1}(1+T)^{\beta}(1+M)^{n+2}|t_{p+1} - t_{p-1}|^{(n+1)\beta+\alpha} \\
&\leqq (2e+2)K\epsilon(1+n)^{2e+1}(1+T)^{\beta}(1+M)^{n+2}\left(\frac{2}{N-1}\right)^{(n+1)\beta+\alpha}|t-s|^{(n+1)\beta+\alpha}.
\end{aligned}$$

This implies that (note that $(n+1)\beta + \alpha > 1$)

$$\begin{aligned}&\left|J^{(1)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(1)}(\mathbb{V})_{st} - \{J^{(1)}(\mathbb{W})_{st}(\mathcal{P}) - J^{(1)}(\mathbb{W})_{st}\}\right| \\ &\leqq (2e+2)K\epsilon(1+n)^{2e+1}(1+T)^{\beta}(1+M)^{n+2}2^{(n+1)\beta+\alpha}\zeta((n+1)\beta+\alpha)|t-s|^{(n+1)\beta+\alpha}.\end{aligned} \tag{1.20}$$

Therefore, by (1.19) and (1.20), we conclude that

$$\begin{aligned}&\left|J^{(1)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(1)}(\mathbb{W})_{st}(\mathcal{P})\right| \\ &\leqq \left|J^{(1)}(\mathbb{V})_{st} - J^{(1)}(\mathbb{W})_{st}\right| + \left|J^{(1)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(1)}(\mathbb{V})_{st} - \left\{J^{(1)}(\mathbb{W})_{st}(\mathcal{P}) - J^{(1)}(\mathbb{W})_{st}\right\}\right| \\ &\leqq K\epsilon(1+eM)(1+n)^{e}(1+T)^{(n+1)\beta}|t-s|^{\alpha} \\ &\quad + (2e+2)K\epsilon(1+n)^{2e+1}(1+T)^{\beta}(1+M)^{n+2}2^{(n+1)\beta+\alpha}\zeta((n+1)\beta+\alpha)|t-s|^{(n+1)\beta+\alpha} \\ &\leqq K\epsilon C_3|t-s|^{\alpha}.\end{aligned}$$

Taking $|\mathcal{P}| \searrow 0$, we prove (1.18). □

(Claim 4) Suppose that there exist $M > 0$ and $\epsilon > 0$ such that

$$|\hat{V}_{st}| \vee |\hat{W}_{st}| \leqq M|t-s|^{\beta}, \quad |V^{(i)}_{st}| \vee |W^{(i)}_{st}| \leqq M|t-s|^{|i|\beta+\alpha},$$

$$|\mathbf{V}^{(jk)}_{st}| \vee |\mathbf{W}^{(jk)}_{st}| \leqq M|t-s|^{|j+k|\beta+2\alpha}, \quad |\hat{V}_{st} - \hat{W}_{st}| \leqq \epsilon|t-s|^{\beta},$$

and

$$|V^{(i)}_{st} - W^{(i)}_{st}| \leqq \epsilon|t-s|^{|i|\beta+\alpha}, \ \ |\mathbf{V}^{(jk)}_{st} - \mathbf{W}^{(jk)}_{st}| \leqq \epsilon|t-s|^{|j+k|\beta+2\alpha}.$$

Then

$$\left|\left(\int f(\hat{\mathbb{V}})\mathrm{d}\mathbb{V}\right)^{(2)}_{st} - \left(\int f(\hat{\mathbb{W}})\mathrm{d}\mathbb{W}\right)^{(2)}_{st}\right| \leqq K^2\epsilon C_4|t-s|^{2\alpha}, \tag{1.21}$$

where

$$C_4 := (1+m)^{2e}(1+2eM)(1+T)^{(m+1)\beta} + (1+T^{(n-m)\beta})(\tilde{C}_4 + 4C_1C_3)2^{(m+1)\beta+2\alpha}\zeta((m+1)\beta+2\alpha),$$

$$\tilde{C}_4 := (15e+7)(1+n+m)^{3e}(1+M)^{m+3}(1+T)^{(2n-m)\beta}.$$

In particular, the integration map is Lipschitz continuous.

Proof. The assumption and the mean value theorem imply that

$$\begin{aligned}&\left|J^{(2)}(\mathbb{V})_{st} - J^{(2)}(\mathbb{W})_{st}\right| \\ &\leqq \sum_{(j,k)\in J}\left|\partial^{j} f(\hat{v}_s)\partial^{k} f(\hat{v}_s)\mathbf{V}^{(jk)}_{st} - \partial^{j} f(\hat{w}_s)\partial^{k} f(\hat{w}_s)\mathbf{W}^{(jk)}_{st}\right| \\ &\leqq K^2\epsilon(1+m)^{2e}(2eM+1)(1+T)^{(m+1)\beta}|t-s|^{2\alpha}.\end{aligned} \tag{1.22}$$

On the other hand, by (1.13) and (1.15), we can calculate

$$
\begin{aligned}
&|S_1(\mathbb{V}) - S_1(\mathbb{W})| \\
&\leqq \sum_{(j,k)\in J} |\partial^j f(\hat{v}_s) R^{(1)}_{jk}(\mathbb{V}) V^{(j)}_{su} \otimes V^{(k)}_{ut} - \partial^j f(\hat{w}_s) R^{(1)}_{jk}(\mathbb{W}) W^{(j)}_{su} \otimes W^{(k)}_{ut}| \\
&\quad + \sum_{(j,k)\in I\times I\setminus J} |\partial^j f(\hat{v}_s)\partial^k f(\hat{v}_u) V^{(j)}_{su} \otimes V^{(k)}_{ut} - \partial^j f(\hat{w}_s)\partial^k f(\hat{w}_u) W^{(j)}_{su} \otimes W^{(k)}_{ut}| \\
&\leqq K^2\epsilon(1+m)^{3e}(1+M)^{m+3}(1+T)^{\beta}(5e+2)|t-s|^{(m+1)\beta+2\alpha} \\
&\quad + K^2\epsilon(1+n)^{2e}(1+M)^2(1+T)^{(2n-m)\beta}(2e+2)|t-s|^{(m+1)\beta+2\alpha} \\
&\leqq K^2\epsilon(1+n+m)^{3e}(1+M)^{m+3}(1+T)^{(2n-m)\beta}(7e+4)|t-s|^{(m+1)\beta+2\alpha},
\end{aligned}
$$

and

$$
\begin{aligned}
&|S_2(\mathbb{V}) - S_2(\mathbb{W})| \\
&\leqq \sum_{(j,k)\in J} \sum_{|p|\leqq m-|j+k|} \left|\partial^{j+p} f(\hat{v}_s) R^{(3)}_{jkp}(\mathbb{V})(\hat{V}_{su})^p \mathbf{V}^{(jk)}_{ut} - \partial^{j+p} f(\hat{w}_s) R^{(3)}_{jkp}(\mathbb{W})(\hat{W}_{su})^p \mathbf{W}^{(jk)}_{ut}\right| \\
&\quad + \sum_{(j,k)\in J} \sum_{|p|\leqq m-|j+k|} \left|\partial^k f(\hat{v}_u) R^{(2)}_{jk}(\mathbb{V}) \mathbf{V}^{(jk)}_{ut} - \partial^k f(\hat{w}_u) R^{(2)}_{jk}(\mathbb{W}) \mathbf{W}^{(jk)}_{ut}\right| \\
&\leqq K^2\epsilon(5e+2)(1+m)^{3e+1}(1+T)^{\beta}(1+M)^{m+2}|t-s|^{(m+1)\beta+2\alpha} \\
&\quad + K^2\epsilon(3e+1)(1+m)^{3e+1}(1+M)^{m+2}|t-s|^{(m+1)\beta+2\alpha} \\
&\leqq K^2\epsilon(8e+3)(1+m)^{3e+1}(1+T)^{\beta}(1+M)^{m+2}|t-s|^{(m+1)\beta+2\alpha}.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&|\Sigma(\mathbb{V})_{sut} - \Sigma(\mathbb{W})_{sut}| \\
&\leqq |S_1(\mathbb{V}) - S_1(\mathbb{W})| + |S_2(\mathbb{V}) - S_2(\mathbb{W})| \leqq K^2\epsilon\tilde{C}_4|t-s|^{(m+1)\beta+2\alpha},
\end{aligned}
$$

where

$$
\Sigma_{sut}(\mathbb{V}) := J^{(2)}(\mathbb{V})_{su} + J^{(2)}(\mathbb{V})_{ut} + J^{(1)}(\mathbb{V})_{su} \otimes J^{(1)}(\mathbb{V})_{ut} - J^{(2)}(\mathbb{V})_{st}, \quad s \leqq u \leqq t
$$

and

$$
\tilde{C}_4 = (15e+7)(1+n+m)^{3e}(1+M)^{m+3}(1+T)^{(2n-m)\beta}.
$$

Let

$$
\Gamma(\mathbb{V})_{sut} := Y^{(1)}(\mathbb{V})_{su} \otimes Y^{(1)}(\mathbb{V})_{ut} - J^{(1)}(\mathbb{V})_{su} \otimes J^{(1)}(\mathbb{V})_{ut}, \quad s \leqq u \leqq t.
$$

Then by (1.7), (1.12), (1.18), and (1.20), we have

$$
\begin{aligned}
&|\Gamma(\mathbb{V})_{sut} - \Gamma(\mathbb{W})_{sut}| \\
&\leqq \left|Y^{(1)}(\mathbb{V})_{su} \otimes (Y^{(1)}(\mathbb{V})_{ut} - J^{(1)}(\mathbb{V})_{ut}) - Y^{(1)}(\mathbb{W})_{su} \otimes (Y^{(1)}(\mathbb{W})_{ut} - J^{(1)}(\mathbb{W})_{ut})\right| \\
&\quad + \left|(Y^{(1)}(\mathbb{V})_{su} - J^{(1)}(\mathbb{V})_{su}) \otimes J^{(1)}(\mathbb{V})_{ut} - (Y^{(1)}(\mathbb{W})_{su} - J^{(1)}(\mathbb{W})_{su}) \otimes J^{(1)}(\mathbb{W})_{ut}\right| \\
&\leqq \left|Y^{(1)}(\mathbb{V})_{su}\right| \left|Y^{(1)}(\mathbb{V})_{ut} - J^{(1)}(\mathbb{V})_{ut} - Y^{(1)}(\mathbb{W})_{ut} + J^{(1)}(\mathbb{W})_{ut}\right| \\
&\quad + \left|Y^{(1)}(\mathbb{V})_{su} - Y^{(1)}(\mathbb{W})_{su}\right| \left|Y^{(1)}(\mathbb{W})_{ut} - J^{(1)}(\mathbb{W})_{ut}\right| \\
&\quad + \left|Y^{(1)}(\mathbb{V})_{ut} - J^{(1)}(\mathbb{V})_{ut} - Y^{(1)}(\mathbb{W})_{ut} + J^{(1)}(\mathbb{W})_{ut}\right| \left|J^{(1)}(\mathbb{V})_{ut}\right| \\
&\quad + \left|Y^{(1)}(\mathbb{W})_{ut} - J^{(1)}(\mathbb{W})_{ut}\right| \left|J^{(1)}(\mathbb{V})_{ut} - J^{(1)}(\mathbb{W})_{ut}\right| \\
&\leqq K^2 \epsilon 4 C_1 C_3 |t-s|^{(n+1)\beta + 2\alpha}.
\end{aligned}
$$

By Lemma 1.5.1, there exists $p \in \{1, ..., N\}$ such that (1.11) holds. Then we have

$$
\begin{aligned}
&\left|J^{(2)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(2)}(\mathbb{V})_{st}(\mathcal{P}\backslash\{t_p\}) - \left\{J^{(2)}(\mathbb{W})_{st}(\mathcal{P}) - J^{(2)}(\mathbb{W})_{st}(\mathcal{P}\backslash\{t_p\})\right\}\right| \\
&\leqq \left|\Sigma(\mathbb{V})_{t_{p-1}t_p t_{p+1}} - \Sigma(\mathbb{W})_{t_{p-1}t_p t_{p+1}}\right| + \left|\Gamma(\mathbb{V})_{t_{p-1}t_p t_{p+1}} - \Gamma(\mathbb{V})_{t_{p-1}t_p t_{p+1}}\right| \\
&\leqq K^2 \epsilon \tilde{C}_4 |t_{i+1} - t_{i-1}|^{(m+1)\beta + 2\alpha} + K^2 \epsilon 4 C_1 C_3 |t_{i+1} - t_{i-1}|^{(n+1)\beta + 2\alpha} \\
&\leqq K^2 \epsilon (1 + T^{(n-m)\beta})(\tilde{C}_4 + 4 C_1 C_3) \left(\frac{2}{N-1}\right)^{(m+1)\beta + 2\alpha} |t-s|^{(m+1)\beta + 2\alpha}.
\end{aligned}
$$

This implies that (note that $(m+1)\beta + 2\alpha > 1$)

$$
\begin{aligned}
&\left|J^{(2)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(2)}(\mathbb{V})_{st} - \left\{J^{(2)}(\mathbb{W})_{st}(\mathcal{P}) - J^{(2)}(\mathbb{W})_{st}\right\}\right| \\
&\leqq K^2 \epsilon (1 + T^{(n-m)\beta})(\tilde{C}_4 + 4 C_1 C_3) 2^{(m+1)\beta + 2\alpha} \zeta((m+1)\beta + 2\alpha) |t-s|^{(m+1)\beta + 2\alpha}.
\end{aligned}
\tag{1.23}
$$

Therefore, by (1.22) and (1.23) we conclude that

$$
\begin{aligned}
&\left|J^{(2)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(2)}(\mathbb{W})_{st}(\mathcal{P})\right| \\
&\leqq \left|J^{(2)}(\mathbb{V})_{st} - J^{(2)}(\mathbb{W})_{st}\right| \\
&\quad + \left|J^{(2)}(\mathbb{V})_{st}(\mathcal{P}) - J^{(2)}(\mathbb{V})_{st} - \{J^{(2)}(\mathbb{W})_{st}(\mathcal{P}) - J^{(2)}(\mathbb{W})_{st}\}\right| \\
&\leqq K^2 \epsilon C_4 |t-s|^{2\alpha}.
\end{aligned}
$$

Taking $|\mathcal{P}| \searrow 0$, we have (1.21).

For any $\mathbb{V}, \mathbb{W} \in \mathcal{E}_M$, take $\epsilon := d_{(\alpha,\beta)}(\mathbb{V}, \mathbb{W})$. Then we have

$$
|\hat{V}_{st}| \vee |\hat{W}_{st}| \leqq M|t-s|^{\beta}, \quad |V_{st}^{(i)}| \vee |W_{st}^{(i)}| \leqq M|t-s|^{|i|\beta + \alpha},
$$

$$|\mathbf{V}^{(jk)}_{st}| \vee |\mathbf{W}^{(jk)}_{st}| \leqq M|t-s|^{|j+k|\beta+2\alpha}, \quad |\hat{V}_{st} - \hat{W}_{st}| \leqq \epsilon |t-s|^{\beta},$$

and

$$|V^{(i)}_{st} - W^{(i)}_{st}| \leqq \epsilon |t-s|^{|i|\beta+\alpha}, \;\; |\mathbf{V}^{(jk)}_{st} - \mathbf{W}^{(jk)}_{st}| \leqq \epsilon |t-s|^{|j+k|\beta+2\alpha}.$$

Therefore, by (1.18) and (1.21) we conclude that for all $\mathbb{V}, \mathbb{W} \in \mathcal{E}_M$,

$$\begin{aligned} d_\alpha\left(\int f(\hat{\mathbb{V}})\mathrm{d}\mathbb{V}, \int f(\hat{\mathbb{W}})\mathrm{d}\mathbb{W}\right) &\leqq KC_3\epsilon + K^2C_4\epsilon \\ &\leqq K(C_3 + KC_4)d_{(\alpha,\beta)}(\mathbb{V}, \mathbb{W}), \end{aligned}$$

and this is the claim. □

Claims 1–4 complete the proof of Theorem 1.2.6. □

1.4.2 Proof of Proposition 1.3.1

We use the following lemmas.

Lemma 1.4.1 ([68] Proposition 1.1.2)**.**

$$I_1(g)I_p(g^{\otimes p}) = I_{p+1}(g^{\otimes(p+1)}) + p||g||^2_{L^2} I_{p-1}(g^{\otimes(p-1)}), \quad g \in L^2(\mathbb{R}_+),\ p \geqq 1.$$

Lemma 1.4.2 ([60] Corollary 9.7)**.** Let Y belong to the m-th Wiener chaos and $p \geqq 2$. Then we have

$$||Y||_p \leqq \sqrt{m+1}(p-1)^{m/2}||Y||_2.$$

Proof of Proposition 1.3.1. (i) Because $\gamma < 1/2$, $\hat{X}$ is well-defined and one can prove that $\mathcal{K}W(t) = \int_0^t \kappa(t-r)\mathrm{d}W_r$. The modified Chen's relation follows from the binomial theorem as illustrated in the Introduction. For the Hölder property, by Kolmogorov's continuity theorem (see Theorem 3.1 in [27]), it is sufficient to prove the following inequalities: for $p \geqq 2$, $i \in I$, $(j,k) \in J$, and $(s,t) \in \Delta_T$

$$||X^{(i)}_{st}||_p \leqq C|t-s|^{|i|\zeta+1/2}, \quad ||\mathbf{X}^{(jk)}_{st}||_p \leqq C|t-s|^{|j+k|\zeta+1}.$$

Fix $s < r < t$. Note that $\hat{X}^{(1)}_{sr} = \int_0^r \kappa_{sr}(u)\mathrm{d}W_u$. Then by using Lemma 1.4.1 repeatedly, we have that for all $m \in \mathbb{Z}_+$,

$$\left(\hat{X}^{(1)}_{sr}\right)^{2m} = \sum_{l=0}^{m} \tilde{c}_{l,m} I_{2l}(\kappa_{sr}^{\otimes 2l})||\kappa_{sr}||^{2m-2l}_{L^2},$$

$$\left(\hat{X}^{(1)}_{sr}\right)^{2m+1} = \sum_{l=0}^{m} c_{l,m} I_{2l+1}(\kappa_{sr}^{\otimes(2l+1)})||\kappa_{sr}||^{2m-2l}_{L^2},$$

where $c_{0,0} = 1$,

$$\tilde{c}_{l,m} = \begin{cases} c_{0,m-1} & l = 0, \\ c_{l-1,m-1} + (2l+1)c_{l,m-1} & l = 1, \ldots, m-1, \\ 1 & \text{otherwise}, \end{cases}$$

and

$$c_{l,m} = \begin{cases} \tilde{c}_{l,m} + 2(l+1)\tilde{c}_{l+1,m} & l = 0, ..., m-1, \\ 1 & \text{otherwise.} \end{cases}$$

Then the assumption $\gamma < 1/2$ and Lemma 1.4.2 imply that for all $m \in \mathbb{Z}_+$,

$$\begin{aligned} &\left\| \int_s^t \left(\hat{X}_{sr}^{(1)}\right)^{2m} \left(\hat{X}_{sr}^{(2)}\right)^{i_2} \mathrm{d}X_r \right\|_p \\ &\leqq \sum_{l=0}^m \tilde{c}_{l,m} \left\| \int_s^t I_{2l}(\kappa_{sr}^{\otimes 2l}) ||\kappa_{sr}||_{L^2}^{2m-2l} |r-s|^{\zeta i_2} \mathrm{d}X_r \right\|_p \\ &\leqq \sum_{l=0}^m \tilde{c}_{l,m} p^{(2l+1)/2} \left\| \int_s^t I_{2l}(\kappa_{sr}^{\otimes 2l}) ||\kappa_{sr}||_{L^2}^{2m-2l} |r-s|^{\zeta i_2} \mathrm{d}X_r \right\|_2 \\ &\leqq p^{(i_1+1)/2} \left(\sum_{l=0}^m \tilde{c}_{l,m} p^{l-m} \right) |t-s|^{|i|\zeta + 1/2 - i_1/2(2\gamma-1)} \\ &\leqq Cp^{(i_1+1)/2} \left(\sum_{l=0}^m \tilde{c}_{l,m} p^{l-m} \right) |t-s|^{|i|\zeta+1/2}, \end{aligned}$$

and

$$\begin{aligned} &\left\| \int_s^t \left(\hat{X}_{sr}^{(1)}\right)^{2m+1} \left(\hat{X}_{sr}^{(2)}\right)^{i_2} \mathrm{d}X_r \right\|_p \\ &\leqq \sum_{l=0}^m c_{l,m} \left\| \int_s^t I_{2l+1}(\kappa_{sr}^{\otimes(2l+1)}) ||\kappa_{sr}||_{L^2}^{2m-2l} |r-s|^{\zeta i_2} \mathrm{d}X_r \right\|_p \\ &\leqq \sum_{l=0}^m c_{l,m} p^{(2l+2)/2} \left\| \int_s^t I_{2l+1}(\kappa_{sr}^{\otimes(2l+1)}) ||\kappa_{sr}||_{L^2}^{2m-2l} |r-s|^{\zeta i_2} \mathrm{d}X_r \right\|_2 \\ &\leqq p^{(i_1+1)/2} \left(\sum_{l=0}^m c_{l,m} p^{l-m} \right) |t-s|^{|i|\zeta + 1/2 - i_1/2(2\gamma-1)} \\ &\leqq Cp^{(i_1+1)/2} \left(\sum_{l=0}^m c_{l,m} p^{l-m} \right) |t-s|^{|i|\zeta+1/2}. \end{aligned}$$

Therefore, we conclude that for all $i = (i_1, i_2) \in \mathbb{Z}_+^2$,

$$||X_{st}^{(i)}||_p = \left\| \int_s^t \left(\hat{X}_{sr}^{(1)}\right)^{i_1} \left(\hat{X}_{sr}^{(2)}\right)^{i_2} \mathrm{d}X_r \right\|_p \leqq Cp^{(i_1+1)/2} |t-s|^{|i|\zeta+1/2}, \tag{1.24}$$

and this implies the claim. By the same argument, we have

$$||\mathbf{X}_{st}^{(jk)}||_p \leqq Cp^{(j_1+k_1+2)/2} |t-s|^{|j+k|\zeta+1}, \quad (j,k) = ((j_1,j_2),(k_1,k_2)) \in \mathbb{Z}_+^2 \times \mathbb{Z}_+^2. \tag{1.25}$$

(ii) By (i) and Theorem 1.2.6, for a.s. ω, the limit

$$\left(\int f(\hat{\mathbb{X}}) \mathrm{d}\mathbb{X} \right)_{st}^{(1)} = \lim_{N\to\infty} \sum_{q=1}^N \sum_{i \in I} \partial^i f(\hat{X}_{t_{q-1}}) X_{t_{q-1}t_q}^{(i)}$$

exists. Because

$$\int_s^t f(\hat{X}_r)\mathrm{d}X_r = \lim_{N\to\infty}\sum_{q=1}^N f(\hat{X}_{t_{q-1}})X^{(0)}_{t_{q-1}t_q}$$

in the sense of the convergence in probability, it is sufficient to prove that for all $i \in I\backslash\{0\}$,

$$\lim_{N\to\infty}\sum_{q=1}^N \partial^i f(\hat{X}_{t_{q-1}})X^{(i)}_{t_{q-1}t_q} = 0$$

in probability. Fix $i \in I\backslash\{0\}$. We can assume $f \in C_b^{n+2}$ without loss of generality. By the result (i), we have

$$\mathsf{E}\left[\left(X^{(i)}_{st}\right)^2\right] = C|t-s|^{2|i|\zeta+1} < \infty,$$

and so taking $K := ||f||_{C_b^{n+2}}$, we conclude that

$$\begin{aligned}
\mathsf{E}\left[\left(\sum_{q=1}^N \partial^i f(\hat{x}_{t_{q-1}})X^{(i)}_{t_{q-1},t_q}\right)^2\right] &= \sum_{q=1}^N \mathsf{E}\left[\left(\partial^i f(\hat{x}_{t_{q-1}})X^{(i)}_{t_{q-1}t_q}\right)^2\right] \\
&\leqq K^2\sum_{q=1}^N |t_q - t_{q-1}|^{2|i|\zeta+1} \\
&= K^2\left(\sup_{|t-s|\leqq|\mathcal{P}|}|t-s|\right)^{2|i|\zeta} T \\
&\to 0 \quad (as\ |\mathcal{P}| \searrow 0),
\end{aligned}$$

and this indicates the L^2 convergence. □

1.4.3 Proof of Theorem 1.3.2

Denote by $C_{[0,T]}$ the set of the $\mathbb{R}$-valued continuous functions on $[0,T]$ equipped with the uniform topology. Let C_{Δ_T} be the set of continuous functions on Δ_T, taking values in $\mathbb{R}^D$, equipped with the uniform topology for the metric

$$d(X,Y) := \sup_{(s,t)\in\Delta_T}|X_{st} - Y_{st}|, \quad X,Y \in C_{\Delta_T}.$$

We use the same notation C_{Δ_T} for different dimensions D, more specifically any one of $D = 1$, $D = \max\{|i||\ i \in I\}$, or $D = \max\{|j+k||\ (j,k) \in J\}$. Let $\mathcal{S}_0$ be the set of the $\mathbb{R}$-valued $\{\mathcal{F}_t\}$-adapted simple processes on $[0,T]\times\Omega$ and

$$\mathcal{S} := \left\{Z \in \mathcal{S}_0 \middle| \sup_{t\in[0,T]}|Z_t| \leqq 1\right\}.$$

Definition 1.4.3 ([39]). Let $\{V^n\}$ be a sequence of $\mathbb{R}$-valued semimartingales on $[0,T]\times\Omega$. We say that the sequence is uniformly exponentially tight (UET) if for every $T > 0$ and every $a > 0$ there is $K_{T,a}$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \sup_{Z\in\mathcal{S}} \mathsf{P}\left[\sup_{t\in[0,T]} |(Z_{-}\cdot V^n)_t| \geqq K_{T,a}\right] \leqq -a, \tag{1.26}$$

where $Z_{-}\cdot V$ is the Itô integral of Z with respect to V:

$$(Z_{-}\cdot V)_t := \int_0^t Z_{r-}\mathrm{d}V_r,$$

For a one-dimensional Brownian motion W, $V^n = n^{-1/2}W$ is an example of a UET sequence; see Lemma 2.4 of [39].

Theorem 1.4.4. Let $\{U^n\}$ be a UET sequence of $\mathbb{R}$-valued semimartingales and $\{V^n\}$ a sequence of $\mathbb{R}$-valued continuous adapted processes. Assume that the sequence $\{(U^n,V^n)\}$ satisfies the LDP on $C_{[0,T]}\times C_{[0,T]}$ with speed n^{-1} and good rate function $\tilde{J}^*$. Then the sequence $\{(U^n,V^n,(U^n\cdot_i V^n)_{i\in I})\}$ satisfies the LDP on $C_{[0,T]}\times C_{[0,T]}\times C_{\Delta_T}$ with speed n^{-1} and good rate function

$$\begin{aligned}\tilde{J}^{**}(u,v,x) &:= \begin{cases}\tilde{J}^*(u,v), & v\in \mathrm{BV},\ \forall i\in I, x^{(i)} = u\cdot_i v,\\ \infty, & \text{otherwise,}\end{cases}\\ &= \inf\left\{\tilde{J}^*(u,v)\middle| u,v\in C_{[0,T]}, v\in\mathrm{BV},\ \forall i\in I, x^{(i)} = u\cdot_i v\right\},\end{aligned} \tag{1.27}$$

where BV is the set of the functions of bounded variation on $[0,T]$, $x = (x^{(i)})_{i\in I}\in C_{\Delta_T}$ and

$$(u\cdot_i v)_{st} := \int_s^t (u_r - u_s)^i \mathrm{d}v_r.$$

Proof. By the assumption and the contraction principle, $\{(U^n,V^n,((U^n)^i)_{i\in I})\}$ satisfies the LDP with good rate function

$$\Lambda_1(u,v,\varphi) = \inf\left\{\tilde{J}^*(u,v)\middle| \forall i\in I, \varphi^{(i)} = u^i\right\}.$$

Therefore, by [39][Theorem 1.2], we have that $\{(U^n,V^n,((U^n)^i,U^n\odot_i V^n)_{i\in I})\}$ satisfies the LDP with good rate function

$$\Lambda_2(u,v,\varphi,x) = \inf\left\{\tilde{J}^*(u,v)\middle| u,v\in C_{[0,T]},\ v\in\mathrm{BV},\ (\varphi^{(i)},x^{(i)}) = (u^i, u\odot_i v)\right\},$$

where $(u\odot_i v)_t := (u\cdot_i v)_{0t}$. Note that by the modified Chen's relation (1.2), we have

$$(u\cdot_i v)_{st} = (u\odot_i v)_t - (u\odot_i v)_s - \sum_{p<i}\frac{1}{(i-p)!}(u_s-u_0)^{i-p}(u\cdot_p v)_{st}.$$

Hence, by the contraction principle again with the aid of induction, we see that $\{(U^n,V^n,(U^n\cdot_i V^n)_{i\in I})\}$ satisfies the LDP with good rate function (1.27). □

Theorem 1.4.5. Under the same conditions as in Theorem 1.4.4, the sequence

$$\{(\delta U^n, (U^n \cdot_i V^n)_{i\in I}, (U^n *_{jk} V^n)_{(j,k)\in J})\}$$

satisfies the LDP on $C_{\Delta_T} \times C_{\Delta_T} \times C_{\Delta_T}$ with speed n^{-1} with good rate function

$$\tilde{J}^{***}(\hat{x}, x, \mathbf{x}) = \inf \left\{ \tilde{J}^*(u,v) \middle| \begin{array}{c} u, v \in C_{[0,T]},\ v \in \mathrm{BV}, \\ \forall i \in I, \forall (j,k) \in J, (\hat{x}, x^{(i)}, \mathbf{x}^{(jk)}) = (\delta u, u \cdot_i v, u *_{jk} v) \end{array} \right\}, \tag{1.28}$$

where $(\delta u)_{st} := u_t - u_s$ and

$$(u *_{jk} v)_{st} := \int_s^t (u \cdot_j v)_{sr} (u_r - u_s)^k \mathrm{d}v_r.$$

Proof. By Theorem 1.4.4 and the contraction principle, the sequence

$$\left\{ \left(U^n, V^n, (U^n \cdot_i V^n)_{i\in I}, ((U^n \odot_j V^n)(U^n)^k)_{(j,k)\in J} \right) \right\}$$

satisfies the LDP with good rate function

$$\Lambda_3(u,v,x,\varphi) = \inf \left\{ \tilde{J}^*(u,v) \middle| u, v \in C_{[0,T]},\ v \in \mathrm{BV}, (x^{(i)}, \varphi^{(jk)}) = (u \cdot_i v, (u \odot_j v) u^k) \right\}.$$

Therefore, by [39][Theorem 1.2], we have that

$$\left\{ \left(U^n, (U^n \cdot_i V^n)_{i\in I}, (U^n \circledast_{jk} V^n)_{(j,k)\in J} \right) \right\}$$

satisfies the LDP with good rate function

$$\Lambda_4(u,x,\varphi) = \inf \left\{ \tilde{J}^*(u,v) \middle| u, v \in C_{[0,T]},\ v \in \mathrm{BV},\ (x^{(i)}, \varphi^{(jk)}) = (u \cdot_i v, u \circledast_{jk} v) \right\},$$

where $(U \circledast_{jk} V)_t = (U *_{jk} V)_{0t}$. Note that by the modified Chen's relation (1.3), we have

$$\begin{aligned} (U^n *_{jk} V^n)_{st} &= (U^n \circledast_{jk} V^n)_t - (U^n \circledast_{jk} V^n)_{0s} \\ &\quad - \sum_{q \leq k} \frac{1}{(k-q)!} (U^n_{0s})^{k-q} (U^n \cdot_j V^n)_{0s} \otimes (U^n \cdot_q V^n)_{st} \\ &\quad - \sum_{p+q<j+k} \frac{1}{(j-p)!(k-q)!} (U^n_{0s})^{j+k-p-q} (U^n *_{pq} V^n)_{st}. \end{aligned}$$

Hence, by the contraction principle again with the aid of induction, we see that $\{(\delta U^n, (U^n \cdot_i V^n)_{i\in I}, (U^n *_{jk} V^n)_{(j,k)\in J})\}$ satisfies the LDP on $C_{\Delta_T} \times C_{\Delta_T} \times C_{\Delta_T}$ with good rate function (1.28). □

Lemma 1.4.6. (i) The (α,β) rough path $\mathbb{X}$ of Theorem 1.3.2 has exponential integrability, i.e., there exists $\eta > 0$ such that

$$\mathsf{E}\left[\exp\left\{ \eta |||\mathbb{X}|||^2_{(\alpha,\beta)} \right\} \right] < \infty.$$

(ii) Assume that the family of random variables

$$\mathbb{X}^\epsilon = (\epsilon^{1/2}\hat{X}, \epsilon^{(|i|+1)/2}X^{(i)}, \epsilon^{(|j+k|+2)/2}\mathbf{X}^{(jk)})$$

taking values in $\Omega_{(\alpha,\beta)\text{-Hld}}$ satisfies the LDP on $C_{\Delta_T}\times C_{\Delta_T}\times C_{\Delta_T}$ (with the uniform topology). Then, $\mathbb{X}^\epsilon$ satisfies the LDP on $\Omega_{(\alpha,\beta)\text{-Hld}}$ (in the $d_{(\alpha,\beta)}$ topology) with the same good rate function.

Proof. (i) Let $Z := |||\mathbb{X}|||_{(\alpha,\beta)}$. By the inequality (1.24), (1.25), we have that for all $p \in [2,\infty)$,

$$||X^{(i)}_{st}||_p \leqq Cp^{(i_1+1)/2}|t-s|^{|i|\zeta+1/2} \quad ||\mathbf{X}^{(jk)}_{st}||_p \leqq Cp^{(j_1+k_1+2)/2}|t-s|^{|j+k|\zeta+1},$$

and this inequality and Kolmogorov's continuity theorem (see Theorem 3.1 in [27]) imply that for $p \geqq \xi$,

$$\left\|||\hat{X}||_{\beta\text{-Hld}}\right\|_p \leqq \tilde{c}\sqrt{p}, \quad \left\|||X^{(i)}||_{|i|\beta+\alpha\text{-Hld}}\right\|_p \leqq \tilde{c}p^{(i_1+1)/2},$$

and

$$\left\|||\mathbf{X}^{(jk)}||_{|j+k|\beta+2\alpha\text{-Hld}}\right\|_p \leqq \tilde{c}p^{(j_1+k_1+2)/2},$$

where $\xi := \lceil\hat{\xi}^{-1}\rceil + \max_{i\in I}\lceil\xi_i^{-1}\rceil + \max_{(j,k)\in J}\lceil\xi_{jk}^{-1}\rceil$, $\hat{\xi} := \zeta-\beta$, $\xi_i := |i|(\zeta-\beta)+(1/2-\alpha)$, $\xi_{jk} := |j+k|(\zeta-\beta)+(1-2\alpha)$, $\tilde{c} := \hat{c} + \max_{i\in I} c_i + \max_{(j,k)\in J} c_{jk}$, and

$$\hat{c} := \frac{2C}{1-(1/2)^{(\hat{\xi}-\xi^{-1})}}, \quad c_i := \frac{2C}{1-(1/2)^{(\xi_i-\xi^{-1})}}, \quad c_{jk} := \frac{2C}{1-(1/2)^{(\xi_{jk}-\xi^{-1})}}.$$

Then Jensen's inequality implies that

$$\left\|\left(||X^{(i)}||_{|i|\beta+\alpha\text{-Hld}}\right)^{1/(|i|+1)}\right\|_p \leqq \left\|||X^{(i)}||_{|i|\beta+\alpha\text{-Hld}}\right\|_p^{1/(|i|+1)} \leqq \tilde{c}^{1/(|i|+1)}\sqrt{p},$$

and similarly

$$\left\|\left(||\mathbf{X}^{(jk)}||_{|j+k|\beta+2\alpha\text{-Hld}}\right)^{1/(|j+k|+2)}\right\|_p \leqq \tilde{c}^{1/(|j+k|+2)}\sqrt{p}.$$

Therefore, we have that

$$||Z||_p \leqq c\sqrt{p}, \quad p \geqq \xi,$$

where $c := \tilde{c} + \sum_{i\in I}\tilde{c}^{1/(|i|+1)} + \sum_{(j,k)\in J}\tilde{c}^{1/(|j+k|+2)}$. Then we have that

$$\mathsf{E}\left[\exp\{\eta Z^2\}\right] = \sum_{n=0}^{\infty}\frac{\eta^n}{n!}||Z||_{2n}^{2n} \leqq \sum_{2n\leqq\xi}\frac{\eta^n}{n!}||Z||_{2n}^{2n} + \sum_{2n>\xi}\frac{(2c^2\eta)^n}{n!}n^n,$$

and so taking $\eta > 0$ small enough ($2c^2\eta e < 1$), Stirling's formula implies the claim.

(ii) We adapt the argument of [30][Proposition 13.43]. By the inverse contraction principle (see Theorem 4.2.4 of [21]), it is sufficient to prove that $\{\mathbb{X}^\epsilon\}$ is exponentially tight on $\Omega_{(\alpha,\beta)\text{-Hld}}$. By (i), there exists $c > 0$ such that

$$\mathsf{P}\left[|||\mathbb{X}|||_{(\alpha',\beta')} > l\right] \leqq \exp(-cl^2)$$

for any $\alpha' \in (\alpha, 1/2)$ and $\beta' \in (\beta, 1/2)$, and this implies that for all $M > 0$, there exists a precompact set

$$K_M = \left\{\mathbb{X} \in \Omega_{(\alpha,\beta)\text{-Hld}} \middle| \, |||\mathbb{X}|||_{(\alpha',\beta')} \leqq \sqrt{M/c}\right\}$$

on $\Omega_{(\alpha,\beta)\text{-Hld}}$ such that

$$\begin{aligned}\epsilon \log \mathsf{P}\left[\mathbb{X}^\epsilon \in K_M^c\right] &= \epsilon \log \mathsf{P}\left[|||\mathbb{X}^\epsilon|||_{(\alpha',\beta')} > \sqrt{\frac{M}{c}}\right] \\ &= \epsilon \log \mathsf{P}\left[|||\mathbb{X}|||_{(\alpha',\beta')} > \sqrt{\frac{M}{c\epsilon}}\right] \leqq -M,\end{aligned}$$

from which we conclude the claim. □

The inverse contraction principle (see Theorem 4.2.4 of [21]) implies that $\{\epsilon^{1/2}(W, W^\perp)\}$ satisfies the LDP on $C^{\gamma\text{-Hld}}$ with speed ϵ^{-1} with good rate function $\tilde{I}^\#$ (note that $\gamma \in (0, 1/2)$). By Theorem 1 in [33], the map $f \mapsto \mathcal{K}f$ is continuous from $C^{\gamma\text{-Hld}}$ to $C^{\zeta\text{-Hld}}$. Then the contraction principle implies that $\{\epsilon^{1/2}(\hat{X}^{(1)}, X) = \epsilon^{1/2}(\mathcal{K}W, \rho W + \sqrt{1-\rho^2}W^\perp)\}$ satisfies the LDP on $C_{[0,T]} \times C_{[0,T]}$ with speed ϵ^{-1} with good rate function

$$\tilde{I}^{(1)}(w,v) = \inf\left\{\tilde{I}^\#(\tilde{v}) \middle| \tilde{v} \in \mathcal{H}, (w,v) = \left(\int_0^\cdot \kappa(\cdot - r)\mathrm{d}\tilde{v}_r^{(1)}, \rho\tilde{v}^{(1)} + \sqrt{1-\rho^2}\tilde{v}^{(2)}\right)\right\}.$$

Let $F_\epsilon : C_{[0,T]} \times C_{[0,T]} \to C_{[0,T]} \times C_{[0,T]} \times C_{[0,T]}$ and $F : C_{[0,T]} \times C_{[0,T]} \to C_{[0,T]} \times C_{[0,T]} \times C_{[0,T]}$ as $F_\epsilon(w,v)_t := ((w_t, \epsilon^{1/2}t), v_t)$ and $F(w,v)_t := ((w_t, 0), v_t)$. Then F is continuous and $F_\epsilon(w^\epsilon, v^\epsilon) \to F(w,v)$ for all converging sequences $(w^\epsilon, v^\epsilon) \to (w,v)$ with $\tilde{I}^{(1)}(w,v) < \infty$. Hence the extended contraction principle [72][Theorem 2.1] implies that $\{\epsilon^{1/2}(\hat{X}, X)\}$ satisfies the LDP on $C_{[0,T]} \times C_{[0,T]} \times C_{[0,T]}$ with speed ϵ^{-1} with good rate function

$$\tilde{J}^*(u,v) := \inf\left\{\tilde{I}^\#(\tilde{v}) \middle| \tilde{v} \in \mathcal{H}, \ (u,v) = \mathbb{K}(\tilde{v})\right\}.$$

As mentioned earlier, $\{X^\epsilon = \epsilon^{1/2}X\}$ is UET by Lemma 2.4 of [39] with $n = \epsilon^{-1}$. Therefore, by Lemma 1.4.6 (ii) and Theorem 1.4.5 (regarding $n = \epsilon^{-1}$, $U = \hat{X}$ and $V = X$), we have proved Theorem 1.3.2.

1.5 A lemma from rough path theory

Lemma 1.5.1 ([60] Proposition 1.6)**.** Let ω be a control function, i.e.,

$$\omega(s,u) + \omega(u,t) \leqq \omega(s,t), \quad 0 \leqq s \leqq u \leqq t \leqq T,$$

and $\mathcal{P} = \{s = t_0 < t_1 < \dots < t_N = t\}$ be a partition on $[s,t]$ $(N \geqq 2)$. Then there exists an integer i $(1 \leqq i \leqq N)$ such that

$$\omega(t_{i-1}, t_{i+1}) \leqq \frac{2}{N-1}\omega(s,t).$$

Proof. By the definition of ω, we have

$$\sum_{p=1}^{N-1} \omega(t_{i-1}, t_{i+1}) = \sum_{i:\text{odd}} \omega(t_{i-1}, t_{i+1}) + \sum_{i:\text{even}} \omega(t_{i-1}, t_{i+1}) \leqq 2\omega(s,t).$$

Therefore, there exists such i that satisfies the desired inequality. □

1.6 Proof of Theorem 1.3.8

Proof. For brevity, let $\sigma := \sigma(S_0)$. By Theorem 1.3.6 and the contraction principle, $t^{H-1/2}\bar{S}_t$ satisfies the LDP with speed t^{-2H} with good rate function

$$\tilde{J}(\tilde{s}) := \inf\left\{ \tilde{I}^{\#}(\tilde{v}) \middle| \tilde{v} \in \mathcal{H},\ \tilde{s} = \left(\sigma \int f(\widehat{\mathbb{L} \circ \mathbb{K}}(\tilde{v}) \mathrm{d}\mathbb{L} \circ \mathbb{K}(\tilde{v})\right)^{(1)}_{01} \right\}.$$

Let $\tilde{v} = (h^1, h^2) \in \mathcal{H}(\mathbb{R}) \times \mathcal{H}(\mathbb{R})$. Then

$$\begin{aligned}
\tilde{s} &= \left(\sigma \int f(\widehat{\mathbb{L} \circ \mathbb{K}}(\tilde{v}))\mathbb{L} \circ \mathbb{K}(\tilde{v})\right)^{(1)}_{01} \\
&= \sigma \int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}^1_r \mathrm{d}r, 0\right) \mathrm{d}\left(\rho h^1_t + \sqrt{1-\rho^2} h^2_t\right) \\
&= \rho\sigma \int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}^1_r \mathrm{d}r, 0\right) \mathrm{d}h^1_t + \sqrt{1-\rho^2}\sigma \int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}^1_r \mathrm{d}r, 0\right) \mathrm{d}h^2_t,
\end{aligned}$$

and so

$$\frac{\tilde{s} - \rho\sigma \int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}^1_r \mathrm{d}r, 0\right) \mathrm{d}h^1_t}{\sqrt{1-\rho^2}} = \sigma \int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}^1_r \mathrm{d}r, 0\right) \mathrm{d}h^2_t. \tag{1.29}$$

Fix h_1, and minimize $\frac{1}{2}||\tilde{v}||^2_{\mathcal{H}(\mathbb{R}^2)}$ with respect to $h_2 \in \mathcal{H}(\mathbb{R})$ under the condition (1.29). Let $\tilde{h}$ be the minimizer. Take $\epsilon > 0$ and $\hat{h} \in \mathcal{H}(\mathbb{R})$, and consider $\tilde{h} + \epsilon\hat{h}$. Because $\tilde{h}$ satisfies the condition (1.29),

$$\int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}^1_r \mathrm{d}r, 0\right) \mathrm{d}\hat{h}_t = 0. \tag{1.30}$$

Because $\tilde{h}$ is the minimizer, we have

$$\left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \frac{1}{2}\int_0^1 (\dot{\tilde{h}}_r + \epsilon\dot{\hat{h}}_r)^2 \mathrm{d}r = 0, \text{ i.e., } \int_0^1 \dot{\tilde{h}}_r \dot{\hat{h}}_r \mathrm{d}r = 0,$$

for any $\hat{h}$ with (1.30). Therefore, there exists $c \in \mathbb{R}$ such that

$$\dot{\hat{h}} = cf\left(\int_0^{\cdot} \kappa_H(\cdot - r)\dot{h}_r^1 \mathrm{d}r, 0\right).$$

Hence

$$\frac{\tilde{s} - \rho\sigma \int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}_r^1 \mathrm{d}r, 0\right) \mathrm{d}h_t^1}{\sqrt{1-\rho^2}} = c\sigma \int_0^1 f^2\left(\int_0^t \kappa_H(t-r)\dot{h}_r^1 \mathrm{d}r, 0\right) \mathrm{d}t,$$

and we conclude that

$$\tilde{J}(\tilde{s}) = \tilde{I}^{\#}(\tilde{v}) = \frac{1}{2}\int_0^1 |\dot{h}_r^1|^2 \mathrm{d}s + \frac{\left\{\tilde{s} - \rho\sigma \int_0^1 f\left(\int_0^t \kappa_H(t-r)\dot{h}_r^1 \mathrm{d}r, 0\right) \mathrm{d}h_t^1\right\}^2}{2(1-\rho^2)\sigma^2 \int_0^1 f^2\left(\int_0^t \kappa_H(t-r)\dot{h}_r^1 \mathrm{d}r, 0\right) \mathrm{d}t},$$

which is the claim. □

Chapter 2

LDP for stochastic differential equations driven by stochastic integrals†

2.1 Introduction

A rough volatility model is a stochastic volatility model for an asset price process in which the Hölder regularity of volatility processes is less than half. In recent years, such a model has attracted attention, because as shown by [31], rough volatility models are the only class of continuous price models that are consistent to a power law of implied volatility term structure typically observed in equity option markets. Proving a large deviation principle (LDP) is one way to derive the power law under rough volatility models as done by many authors using various methods [23, 9, 8, 24, 25, 48, 50, 51, 62, 42, 63, 49, 37]. An introduction to LDP and some of its applications to finance and insurance problems are discussed in [71, 26]. One precise approximation formula for implied volatility is the BBF formula [12, 1], which follows from short-time LDP under local volatility models. On the other hand, the SABR formula, which is of daily use in financial practice, is also proved for a valid approximation under the SABR model by means of LDP [69]. From these relations between LDP and precise approximation under classical (non-rough) volatility models, we expect LDP for rough volatility models to provide in particular a useful implied volatility approximation formula for financial practice such as model calibration.

For the proof for pathwise LDP of standard stochastic differential equations (SDEs), an elegant method using rough path theory was proposed [29, 30]. The continuity of the solution map on rough path spaces is key to derive the pathwise LDP for such SDEs. However, when considering rough volatility models, the usual rough path theory does not work because the regularity of the volatility process is lower than that of asset prices,

†Reproduction of a work, first published in SIAM Journal on Financial Mathematics, Vol.16, No.2, pp480-515, (2025), https://doi.org/10.1137/24M1653306.

and so stochastic integrands are not controlled by the stochastic integrators in the sense of [43]. Nevertheless, methods which are analogue of rough path theory have been proposed to prove the pathwise LDP for rough volatility model, one uses the theory of regularity structure [8], and another uses a variant of rough path theory [37]. In [62, 8], the following Itô SDE is discussed (here Y represents the dynamics of the logarithm of a stock price process):

$$\mathrm{d}Y_t = f(\hat{X}_t, t)\mathrm{d}X_t - \frac{1}{2}f^2(\hat{X}_t, t)\mathrm{d}t,$$

where X is a Brownian motion, $\hat{X}$ is the Riemann-Liouville type fractional Brownian motion with Hurst index $H \in (0, 1/2)$, and f is a smooth function. This SDE is called rough Bergomi model [7]. In [8], the authors proved the short-time LDP for rough Bergomi models and by using the continuity of Hairer's reconstruction map. The point is that its proof comes down to the small-noise LDP for "models" which construct the solution of the rough Bergomi model. On the other hand, this result was extended to situations where rough volatility models have local volatility in [37];

$$\mathrm{d}Y = \sigma(Y)f(\hat{X}, t)\mathrm{d}X - \frac{1}{2}\sigma^2(Y)f^2(\hat{X}, t)\mathrm{d}t, \tag{2.1}$$

where X is a Brownian motion, $\hat{X} := \int_0^\cdot \kappa(\cdot - s)\mathrm{d}W_s$ (where κ is a deterministic singular kernel and W is a Brownian motion), σ and f are smooth functions respectively. If $\sigma = 1$ and $\kappa = \kappa_H$ (κ_H is the Riemann-Liouville kernel, see (2.10)), (2.2) is the rough Bergomi model. In [37], partial rough path spaces lacking the iterated integral of $\hat{X}$ were considered, and a partial rough path integration map was constructed. By using the continuity property of this integration map, the small-noise and short-time LDP for (2.1) were proved based on the pathwise LDP for the canonical noises constructed by $(X, \hat{X})$ on partial rough path spaces. Compared with [8], the framework of [37] is more elementary and one can prove that not only the LDP for rough Bergomi model but also that for many rough volatility models, see the lists of Introduction in [37]. However, the continuity property of the integration map in [37] relies on the smoothness of the coefficient f, because the higher order Taylor expansion of f is needed to cover the low regularity of $\hat{X}$. For these reason, although the previous work [37] is widely applicable, it goes beyond the framework of it when f is not smooth. For example generalized rough volatility models discussed in [58] or when $f(y, t) := \sqrt{y}$ [23].

Inspired by above previous research, we will discuss the following SDE in one dimension in this paper:

$$\mathrm{d}Y^\epsilon = \sigma(Y^\epsilon)A^\epsilon_t \mathrm{d}X^\epsilon - \frac{1}{2}\sigma^2(Y^\epsilon)(A^\epsilon)^2_t \mathrm{d}t \tag{2.2}$$

Here $X^\epsilon := \epsilon^{1/2}X$, X is a one dimensional Brownian motion, A^ϵ is an adapted continuous process and σ is a smooth function. If $A^\epsilon = f(\hat{X}^\epsilon, \cdot)$, (2.2) coincides with (2.1). In this paper, we will discuss the pathwise LDP for (2.2).

Now we consider how to prove the pathwise LDP for (2.2). Let $A \cdot X$ be the Itô stochastic integral for A with respect to X and $\Lambda(t) := t$. Let also $Z^\epsilon := (A^\epsilon \cdot X^\epsilon, (A^\epsilon)^2 \cdot \Lambda)$ and we will regard Z^ϵ as the driver for (2.2). Then we define the Young pairing (see Section 9.4 in [30]) $\mathbb{Z}^\epsilon$ for Z^ϵ and we regard $\mathbb{Z}^\epsilon$ as the canonical lift for rough path spaces. Since X^ϵ are one dimensional paths, the mapping $Z^\epsilon \mapsto \mathbb{Z}^\epsilon$ is continuous. Combining to the usual rough path theory, we finally can construct the solution Y^ϵ of (2.2) from Z^ϵ:

$$\begin{array}{cccc} G\Omega_{\alpha\text{-Hld}} & \mathbb{Z}^\epsilon & \xrightarrow{\text{sol. map}} & \mathbb{Y}^\epsilon \\ & \text{Young pair}\uparrow & & \downarrow\text{projection} \\ C^{\alpha\text{-Hld}} & Z^\epsilon & & Y^\epsilon \end{array}$$

Here for $\alpha \in (1/3, 1/2]$, $C^{\alpha\text{-Hld}}$ is Hölder spaces, and $G\Omega_{\alpha\text{-Hld}}$ is rough path spaces, and "sol. map" in the above diagram means the solution map in the sense of rough differential equations. Therefore, the pathwise LDP for (2.2) can be proved from the small noise LDP for $\{Z^\epsilon\}_{\epsilon>0}$ on Hölder spaces. This idea enables us to avoid adherence to use the smoothness of coefficient f which is the essential condition to cover the low regularity of $\hat{X}$ in [8, 37]. We also note that our approach does not use a variant of rough path theory or regularity structure theory, which means we are able to obtain simpler proof. Although the small noise LDP for stochastic integrals with respect to the uniform topology was proved in [39], this results cannot be applied for our methods, because our idea requires the small noise LDP for stochastic integrals with respect to "Hölder topology".

Our method also allows for a unified treatment of pathwise LDP for rough volatility models, compared with [8, 37]. For example, the pathwise LDP for rough volatility models were discussed under the different assumptions which are not mutually inclusive [23, 58, 47, 37, 8], but these results indeed are included in our setting. To the best of the author's knowledge, no such pathwise LDP for these models is known in the literature.

In the perspective of applications for mathematical finance, it is important to derive the asymptotic formula of the implied volatility because of the pricing of put/call options. Moreover, the formula is applicable to check whether models are consistent to the power law of implied volatility or not. For example, generalized rough volatility models discussed in [58] are widely applicable, in the sense that the authors of [58] provide us how to make a numerical approximation of such models. Although one reason for using and studying such models is that it is expected to be consistent with the power laws of the implied volatility observed in the market, there is no justifications of this expectation in the literature. As an application of our analysis, we will prove the short-time LDP of them (actually one can treat more general models) and derive an asymptotic formula of the implied volatility which tells us the models are consistent to the power law of the implied volatility (Corollary 2.3.12). This formula is described as a generalization of Forde and Zhang's work [23].

In section 2.2, we first discuss the pathwise LDP for stochastic integrals on Hölder spaces, this is Theorem 2.2.6. We next discuss the pathwise LDP for (2.2) (actually we will discuss the Stratonovich SDEs (2.5) corresponding to (2.2)), see Theorem 2.2.11. In section 2.3, we will first show how to apply main results to the pathwise LDP for rough volatility models, see Theorem 2.3.2 for small-noise LDP and see Theorem 2.3.7 for short-time LDP. Then we will derive an asymptotic formula for implied volatility, see Corollary 2.3.12. In section 2.4, we will prove the main theorem in order.

2.2 Main results

2.2.1 Large deviation principle for stochastic integrals

We first review the LDP for stochastic integrals with respect to the uniform topology discussed in [39]. In [39], the index set of a sequence of stochastic processes is the natural number set $n \in \mathbb{N}$, but regarding as $n = \epsilon^{-1}$, we consider the family of stochastic processes $\{X^\epsilon\}_{\epsilon>0}$ which means that the index set of it is $(0,1]$. We say that a real function $x : [0,\infty) \to \mathbb{R}$ is cadlag if x is continuous on the right and has limits on the left. Throughout this paper, we fix a filtered probability space satisfying the usual conditions $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$.

Definition 2.2.1 (Definition 1.1 [39])**.** Let $\{X^\epsilon\}_{\epsilon>0}$ be a family of real valued cadlag semi-martingales. We say that the family $\{X^\epsilon\}_{\epsilon>0}$ is uniformly exponentially tight with speed ϵ^{-1} if for every $t > 0$ and every $M > 0$, there is $K_{M,t} > 0$ such that

$$\limsup_{\epsilon\searrow 0} \epsilon \log \sup_{U\in\mathcal{S}} \mathsf{P}\left[\sup_{s\leq t} |(U_- \cdot X^\epsilon)_s| \geq K_{M,t}\right] \leq -M, \tag{2.3}$$

where $\mathcal{S}$ be the set of all simple adapted processes U with $\sup_{t\geq 0}|U_t| \leq 1$ and $(U_-)_t := \lim_{s\to t-} U_s$. In this paper, denote $U \cdot X$ by the stochastic integral for U with respect to a semi-martingale X in Itô sense:

$$(U \cdot X)_t := \int_0^t U_r \mathrm{d}X_r.$$

Definition 2.2.2 (Section 1.2 [21])**.** Let $(E, \mathcal{B}(E))$ be a metric space with a Borel σ-algebra $\mathcal{B}(E)$.

(*i*) We say that a function $I : E \to [0,\infty]$ be a good rate function if, for all $\lambda \in [0,\infty)$, the set

$$\{x \in E : I(x) \leq \lambda\},$$

is compact on E.

(*ii*) We say that the family of measures $\{\mu_\epsilon\}_{\epsilon>0}$ on E satisfies the LDP with speed ϵ^{-1} with good rate function I if, for all $\Gamma \in \mathcal{B}(E)$,

$$-\inf_{x\in\Gamma^o} I(x) \leq \liminf_{\epsilon\searrow 0} \epsilon \log \mu_\epsilon(\Gamma) \leq \limsup_{\epsilon\searrow 0} \epsilon \log \mu_\epsilon(\Gamma) \leq -\inf_{x\in\overline{\Gamma}} I(x),$$

where $\overline{\Gamma}$ is the closure of Γ, and Γ^o is the interior of Γ.

In this paper, " with speed ϵ^{-1}" is omitted. Denote by $D([0,\infty),\mathbb{R})$ the space of all cadlag functions and denote by d_D the Skorohod topology (see Chapter 3 in [13]).

Lemma 2.2.3 (Theorem 1.2 [39]). Let $\{X^\epsilon\}_{\epsilon>0}$ be a uniformly exponentially tight family of cadlag adapted semi-martingales on $\mathbb{R}$ and $\{A^\epsilon\}_{\epsilon>0}$ be a family of real valued cadlag adapted processes. Assume that $\{(A^\epsilon, X^\epsilon)\}_{\epsilon>0}$ satisfies the LDP on $(D([0,\infty),\mathbb{R}), d_D) \times (D([0,\infty),\mathbb{R}), d_D)$ with good rate function $I^\#$. Then the family of triples $\{(A^\epsilon, X^\epsilon, A^\epsilon \cdot X^\epsilon)\}_{\epsilon>0}$ satisfies the LDP on $(D([0,\infty),\mathbb{R}), d_D) \times (D([0,\infty),\mathbb{R}), d_D) \times (D([0,\infty),\mathbb{R}), d_D)$ with good rate function

$$I(a,x,z) = \begin{cases} I^\#(a,x), & z = a \cdot x,\ x \in \mathbf{BV}, \\ +\infty, & \text{otherwise}, \end{cases}$$

where **BV** is the set of bounded variation and $a \cdot x$ means the Riemann-Stieltjes integral for a with respect to x.

We wii improve on this result in terms of Hölder topology. For $\alpha \in (0,1]$, denote by $C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ the Hölder space with the Hölder norm

$$\|x\|_{\alpha\text{-Hld}} := |x_0| + \sup_{0\le s<t\le 1} \frac{|x_t - x_s|}{|t-s|^\alpha},$$

and let

$$C_0^{\alpha\text{-Hld}}([0,1],\mathbb{R}) := \{x \in C^{\alpha\text{-Hld}}([0,1],\mathbb{R}) : \lim_{\delta\searrow 0} w_\alpha(\delta, x) = 0\},$$

where

$$w_\alpha(\delta, x) := \sup_{|t-s|\le\delta} \frac{|x_t - x_s|}{|t-s|^\alpha}.$$

Note that $C_0^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ is a separable Banach space, see [56]. We next introduce a concept of α-Uniformly Exponentially Tight.

Definition 2.2.4. We fix $\alpha \in (0,1]$. Let $\{X^\epsilon\}_{\epsilon>0}$ be a family of real valued continuous semi-martingales on $[0,1]$. We say that the family $\{X^\epsilon\}_{\epsilon>0}$ is *α-Uniformly Exponentially Tight* if, for all $M>0$, there exists $K_M>0$ such that

$$\limsup_{\epsilon\searrow 0} \epsilon \log \sup_{U\in\mathcal{B}([0,1],\mathbb{R})} \mathsf{P}\left[\|U \cdot X^\epsilon\|_{\alpha\text{-Hld}} \ge K_M\right] \le -M, \tag{2.4}$$

where $\mathcal{B}([0,1],\mathbb{R})$ is the set of all adapted, left continuous with right limits processes U on $[0,1]$ such that $\sup_{t\in[0,1]} |U_t| \le 1$.

Remark 2.2.5. α-Uniformly Exponentially Tight is stronger than uniformly exponentially tight in the following sense. Assume that $\{X^\epsilon\}_{\epsilon>0}$ is α-Uniformly Exponentially Tight. Note that for all $U \in \mathcal{S}$, $U_- \in \mathcal{B}([0,1],\mathbb{R})$. For all $M>0$, take $K_M>0$ such that (2.4) holds. Then we have that for all $t\in(0,1)$,

$$\begin{aligned} &\limsup_{\epsilon\searrow 0} \epsilon \log \sup_{U\in\mathcal{S}} \mathsf{P}\left[\sup_{s\le t} |(U_- \cdot X^\epsilon)_s| \ge K_M\right] \\ &\le \limsup_{\epsilon\searrow 0} \epsilon \log \sup_{\tilde{U}\in\mathcal{B}([0,1],\mathbb{R})} \mathsf{P}\left[\|\tilde{U} \cdot X^\epsilon\|_{\alpha\text{-Hld}} \ge K_M\right] \\ &\le -M. \end{aligned}$$

Hence we conclude that $\{X^\epsilon\}_{\epsilon>0}$ satisfies (2.3) when $t\in(0,1)$.

Let $C([0,1],\mathbb{R})$ be the set of all real valued continuous functions on $[0,1]$ equipped with the uniform topology. Here, we state our first main result, the proof is given in Section 2.4.1.

Theorem 2.2.6. We fix $\alpha \in (0,1]$ and $\beta < \alpha$. Let $\{X^\epsilon\}_{\epsilon>0}$ be a family of real valued α-Hölder continuous semi-martingales on $[0,1]$ and $\{A^\epsilon\}_{\epsilon>0}$ be a family of real valued adapted continuous processes on $[0,1]$ such that $A^\epsilon \cdot X^\epsilon \in C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$. Assume that $\{(A^\epsilon, X^\epsilon)\}_{\epsilon>0}$ satisfies the LDP on $C([0,1],\mathbb{R}) \times C_0^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ with good rate function $I^\#$.

Then if $\{X^\epsilon\}_{\epsilon>0}$ is α-Uniformly Exponentially Tight, $\{(A^\epsilon, X^\epsilon, A^\epsilon \cdot X^\epsilon)\}_{\epsilon>0}$ satisfies the LDP on $C([0,1],\mathbb{R}) \times C_0^{\alpha\text{-Hld}}([0,1],\mathbb{R}) \times C_0^{\beta\text{-Hld}}([0,1],\mathbb{R})$ with good rate function I;

$$I(a,x,z) = \begin{cases} I^\#(a,x), & z = a \cdot x,\ x \in \mathbf{BV}, \\ +\infty, & \text{otherwise}. \end{cases}$$

Remark 2.2.7. Let V is an adapted continuous process. Note that if $V \in C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ and $\beta < \alpha$, then $V \in C_0^{\beta\text{-Hld}}([0,1],\mathbb{R})$. Note also that since V is an adapted continuous process, we have that

$$\|V\|_{\beta\text{-Hld}} = \sup_{0 \le s < t \le 1, s,t \in \mathbb{Q}} \frac{|V_t - V_s|}{|t-s|^\beta}$$

is $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable. Since $C_0^{\beta\text{-Hld}}([0,1],\mathbb{R})$ is a separable Banach space, we conclude that V is $\mathcal{F}/\mathcal{B}(C_0^{\beta\text{-Hld}}([0,1],\mathbb{R}))$-measurable.

One of the most important family of α-Uniformly Exponentially Tight semi-martingales is constructed from scaled Brownian motions. The proof is deferred to Section 2.4.1.

Proposition 2.2.8. We fix $\alpha \in [1/3, 1/2)$. Let B be an $\mathbb{R}$-valued standard Brownian motion on $[0,\infty)$ and assume that B is $(\mathcal{F}_t)$-adapted. Let $B^\epsilon := \epsilon^{1/2} B$ and $\bar{B}^\epsilon := B_{\epsilon\cdot}$. Then we have that:

(i) for all $(\mathcal{F}_t)$-adapted continuous processes A on $[0,1]$, we have $A \cdot B^\epsilon \in C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$, and $\{B^\epsilon\}_{\epsilon>0}$ is α-Uniformly Exponentially Tight: for all $M > 0$, there exists $K_M > 0$ such that

$$\limsup_{\epsilon \searrow 0} \epsilon \log \sup_{U \in \mathcal{B}([0,1],\mathbb{R})} \mathsf{P}\left[\|U \cdot B^\epsilon\|_{\alpha\text{-Hld}} \ge K_M\right] \le -M.$$

(ii) Let $\mathcal{F}_t^\epsilon := \mathcal{F}_{\epsilon t}$. Then for all $(\mathcal{F}_t^\epsilon)$-adapted continuous processes $\bar{A}$ on $[0,1]$, we have $\bar{A} \cdot \bar{B}^\epsilon \in C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ and $\{\bar{B}^\epsilon\}_{\epsilon>0}$ is α-Uniformly Exponentially Tight.

2.2.2 LDP for SDE driven by stochastic integrals

In this section, we will discuss how to derive the LDP for SDEs driven by stochastic integrals in one dimension from Theorem 2.2.6. Consider the Stratonovich SDEs in one dimension:

$$\mathrm{d}Y_t = \sigma_1(Y_t) \circ A_t \mathrm{d}X_t + \sigma_2(Y_t)\tilde{A}_t \mathrm{d}t, \quad Y_0 \in \mathbb{R}, \tag{2.5}$$

where $\sigma_1, \sigma_2 \in C_b^3$, X is a one-dimensional standard Brownian motion, and A and $\tilde{A}$ are real valued adapted continuous processes respectively. Note that we regard (2.5) as the equation driven by a stochastic integral $A \cdot X$ and $\tilde{A} \cdot \Lambda$ where $\Lambda(t) := t$.

Let

$$Z = (Z^{(1)}, Z^{(2)}) := (A \cdot X, \tilde{A} \cdot \Lambda),$$

where $\cdot$ means the Itô integral. Let also that

$$\mathbb{Z}_{st} := (1, Z_{st}, \mathbf{Z}_{st}), \quad 0 \le s < t \le 1, \tag{2.6}$$

where for $i, j \in \{1, 2\}$,

$$Z_{st} := Z_t - Z_s, \quad \mathbf{Z}_{st}^{(ij)} := \begin{cases} 2^{-1}(Z_{st}^{(1)})^2, & i = j = 1, \\ \int_s^t (Z_r^{(i)} - Z_s^{(i)})\mathrm{d}Z_r^{(j)}, & \text{otherwise}, \end{cases}$$

and $\mathbf{Z}$ is defined by the Young integral (see also Section 9.4 in [30], this is the Young pairing). Note that by Proposition 2.2.8 (i), for $\alpha \in (1/3, 1/2)$, $Z \in C^{\alpha\text{-Hld}} \times C^{1\text{-Hld}}$ and so the Young integral is well-defined. For $\alpha \in (1/3, 1/2]$, denote by $G\Omega^{\alpha\text{-Hld}}([0,1], \mathbb{R}^2)$ the geometric rough path space and d_α the metric function on $G\Omega^{\alpha\text{-Hld}}([0,1], \mathbb{R}^2)$ (see Section 2.2 in [27]). One can prove that for $\alpha \in (1/3, 1/2)$, $\mathbb{Z} \in G\Omega^{\alpha\text{-Hld}}([0,1], \mathbb{R}^2)$, see the proof of Theorem 2.2.11.

We now discuss the following type of rough differential equation (RDE) (in Lyons' sense; see Section 8.8 of [27], for example):

$$\bar{Y}_t = \int_0^t \bar{\sigma}(\bar{Y}_u)\mathrm{d}\mathbb{Z}_u, \tag{2.7}$$

where $\bar{Y}_t = Y_t - Y_0$, $\bar{\sigma}(y) = (\sigma_1(Y_0 + y), \sigma_2(Y_0 + y))$.

Theorem 2.2.9. Let $\sigma_1, \sigma_2 \in C_b^3$.

(i) RDE (2.7) driven by (2.6) has a unique solution $\bar{Y} = \Phi(\mathbb{Z}, y_0)$, where

$$\Phi : \Omega_{\alpha\text{-Hld}}([0,T], \mathbb{R}^2) \times \mathbb{R} \to C^{\alpha\text{-Hld}}([0,T], \mathbb{R})$$

is the solution map of (2.7) that is locally Lipschitz continuous with respect to d_α.

(ii) The first level of the last component $\bar{Y}$ of the solution to RDE (2.7) for (2.6) gives the solution $Y(\omega) = y_0 + \bar{Y}$ to the stratonovich SDE (2.5).

Proof. These are standard results from rough path theory; see e.g., Theorem 1 in [65] or Chapter 8 in [27] for (i) and Chapter 9 in [27] or Theorem 17.3 in [30] for (ii). □

Remark 2.2.10. Although the solution $\bar{Y}$ is one-dimension, the noise Z is a two dimensional path and so it is not trivial whether $\bar{Y}$ can be constructed from Z or not, and this is why we need to consider rough paths $\mathbb{Z}$ of Z.

Let $X^\epsilon := \epsilon^{1/2}X$, and $A^\epsilon, \tilde{A}^\epsilon$ are $(\mathcal{F}_t)$ adapted continuous processes respectively (these correspond scaled processes of $A, \tilde{A}$ respectively). Let

$$Z^\epsilon = ((Z^{(1)})^\epsilon, (Z^{(2)})^\epsilon) := (A^\epsilon \cdot X^\epsilon, \tilde{A}^\epsilon \cdot \Lambda),$$

and we define $\mathbb{Z}^\epsilon$ like (2.6). We now consider the following scaled SDEs:

$$\mathrm{d}Y_t^\epsilon = \sigma_1(Y_t^\epsilon) \circ A_t^\epsilon \mathrm{d}X_t^\epsilon + \sigma_2(Y_t^\epsilon)\tilde{A}_t^\epsilon \mathrm{d}t, \tag{2.8}$$

We state the second main result, the proof is given in Section 2.4.2.

Theorem 2.2.11. We fix $\alpha \in [1/3, 1/2)$. Assume that there exists $\alpha' \in [1/3, 1/2)$ with $\alpha' > \alpha$ such that $\{(A^\epsilon, \tilde{A}^\epsilon, X^\epsilon)\}_{\epsilon>0}$ satisfies the LDP on $C([0,1],\mathbb{R}) \times C([0,1],\mathbb{R}) \times C_0^{\alpha'\text{-Hld}}([0,1],\mathbb{R})$ with good rate function $J^\#$.

Then $\{Y^\epsilon\}_{\epsilon>0}$ satisfies the LDP on $C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ with good rate function

$$J(y) := \inf\left\{J^\#(a,\tilde{a},x) : y = \Phi \circ F(a\cdot x, \tilde{a}\cdot\Lambda),\ x \in \mathbf{BV}\right\},$$

where

$$F(z)_{st} := (1, z_{st}, \mathbf{z}_{st}), \tag{2.9}$$

and for $i, j \in \{1,2\}$,

$$z_{st}^{(i)} := z_t^{(i)} - z_s^{(i)}, \quad \mathbf{z}_{st}^{(ij)} := \begin{cases} 2^{-1}(z_{st}^{(i)})^2, & i=j=1, \\ \int_s^t (z_r^{(i)} - z_s^{(i)})\mathrm{d}z_r^{(j)}, & \text{otherwise,} \end{cases}$$

and $\mathbf{z}$ is defined by the Young integral.

2.3 An application for mathematical finance

2.3.1 Small noise asymptotics for rough volatility models (2.12)

We now discuss an application of Theorem 2.2.11. Let $\kappa : (0,1] \to [0,\infty)$ as

$$\kappa(t) := g(t)t^{\gamma-\alpha}, \quad t \in (0,1],$$

where $\alpha, \gamma \in (0,1)$ and g is a Lipschitz function. Let $\mathcal{K} : C^{\alpha\text{-Hld}}([0,1],\mathbb{R}) \to C^{\gamma\text{-Hld}}([0,1],\mathbb{R})$ as

$$\begin{aligned}\mathcal{K}f(t) &:= \lim_{\epsilon\searrow 0}\left\{[\kappa(t-\cdot)(f(\cdot)-f(t))]_0^{t-\epsilon} + \int_0^{t-\epsilon}(f(s)-f(t))\kappa'(t-s)\mathrm{d}s\right\} \\ &= \kappa(t)(f(t)-f(0)) + \int_0^t (f(s)-f(t))\kappa'(t-s)\mathrm{d}s.\end{aligned}$$

This map is called the fractional integral for $\gamma > \alpha$ and the fractional derivative for $\gamma < \alpha$, see [33] for details. For simplicity, let $\mu := \gamma - \alpha$.

Remark 2.3.1. Because of the existence of Lipschitz part g, κ has sufficient generality for applications. For example, we can take the following singular kernels.

1. the Riemann–Liouville kernel
$$\kappa_H(t) := t^{H-1/2}, \quad t \in (0,1],\ H \in (0,1/2) \tag{2.10}$$
has the above form ($\mu = H - 1/2$).

2. the Gamma fractional
$$\kappa(t) := t^{\mu} \exp(ct), \quad t \in (0,1],\ \mu \in (-1,1),\ c < 0,$$

3. Power-law
$$\kappa(t) := t^{\mu}(1+t)^{\beta-\mu}, \quad t \in (0,1],\ \mu \in (-1,1),\ \beta < -1.$$

For convenience, we denote $\mathcal{K}_0$ by $\mathcal{K}$ associated with the Riemann–Liouville kernel κ_H, which means $\mathcal{K}_0$ is the usual fractional operator.

We fix $\alpha \in (0,1/2)$ and $\gamma \in (0,1)$ (α and γ are the parameters of $\mathcal{K}$ respectively). Denote by $(W, W^{\perp})$ a two-dimensional standard Brownian motion. Set
$$X := \rho W + \sqrt{1-\rho^2}W^{\perp}, \quad V := \Psi(\mathcal{K}A), \quad \rho \in [-1,1], \tag{2.11}$$
where A is the solution to the SDE
$$\mathrm{d}A_t = b(A_t)\mathrm{d}t + a(A_t)\mathrm{d}W_t, \quad A_0 \in \mathbb{R},$$
$a, b \in C_b^4$, and $\Psi : \mathbb{R} \to \mathbb{R}$ is a nice function (see in Remark 2.3.3). Consider the following Itô SDEs (here Y represents the dynamics of the logarithm of a stock price process):
$$\mathrm{d}Y_t = \sigma(Y_t)f(V_t,t)\mathrm{d}X_t - \frac{1}{2}\sigma^2(Y_t)f^2(V_t,t)\mathrm{d}t, \quad Y_0 \in \mathbb{R} \tag{2.12}$$
where $f : \mathbb{R}\times[0,1] \to [0,\infty)$ be a nice function (see in Remark 2.3.3), and $\sigma : \mathbb{R} \to \mathbb{R}$ is in C_b^4. In this paper, we call (2.12) rough volatility models. The equation can be rewrite in the Stratonovich sense:
$$\mathrm{d}Y_t = \sigma(Y_t) \circ f(V_t,t)\mathrm{d}X_t - \frac{1}{2}\left\{\sigma^2(Y_t) + \sigma(Y_t)\sigma'(Y_t)\right\} f^2(V_t,t)\mathrm{d}t$$
Note that we regard this SDE as the equation driven by a stochastic integral $f(V,\cdot)\cdot X$ and $f^2(V,\cdot)\cdot\Lambda$. For $\epsilon > 0$, let $(X^{\epsilon}, V^{\epsilon}) := (\epsilon^{1/2}X, \epsilon^{1/2}V)$ and consider the following SDEs:
$$\mathrm{d}Y_t^{\epsilon} = \sigma(Y_t^{\epsilon}) \circ f(V_t^{\epsilon},t)\mathrm{d}X_t^{\epsilon} - \frac{1}{2}\left\{\sigma^2(Y_t^{\epsilon}) + \sigma(Y_t^{\epsilon})\sigma'(Y_t^{\epsilon})\right\} f^2(V_t^{\epsilon},t)\mathrm{d}t, \quad \epsilon > 0, \tag{2.13}$$

Let

$$Z^{\epsilon} := (f(V^{\epsilon},\cdot)\cdot X^{\epsilon}, f^2(V^{\epsilon},\cdot)\cdot \Lambda), \quad \epsilon > 0.$$

We state an application of Theorem 2.2.11 for rough volatility models, the proof is given in Section 2.4.2.

Theorem 2.3.2. We fix $\alpha \in [1/3, 1/2)$ and $\gamma \in (0,1)$. Assume that $x \mapsto f(x,\cdot)$ is continuous map on $C([0,1],\mathbb{R})$, $x \mapsto \Psi(x)$ continuous map from $C_0^{\gamma\text{-Hld}}([0,1],\mathbb{R})$ into $C([0,1],\mathbb{R})$ and let Y^{ϵ} is the solution of (2.13). Then $\{Y^{\epsilon}\}_{\epsilon>0}$ satisfies the LDP on $C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ with good rate function

$$\tilde{J}(y) := \inf\left\{\frac{1}{2}\|(w,w^{\perp})\|_{\mathcal{H}}^2 : y = \Phi\circ F\circ F_f \circ \mathbb{K}(w,w^{\perp}),\ (w,w^{\perp})\in\mathcal{H}\right\},$$

where $\mathcal{H}$ is the Cameron-Martin space on $\mathbb{R}^2$,

$$\mathbb{K}w := (\Psi\mathcal{K}(a(A_0)w^{(1)}), \rho w^{(1)} + \sqrt{1-\rho^2}w^{(2)}), \quad F_f(v,x) := \Big(f(v,\cdot)\cdot x, f(v,\cdot)^2\cdot\Lambda\Big), \tag{2.14}$$

F is defined as (2.11).

Remark 2.3.3 (assumptions for f and Ψ)**.** If $f : \mathbb{R}\times[0,\infty) \to \mathbb{R}$ is in C^1-class or $f(x,t)$ does not depend on $t \in [0,\infty)$ and locally β-Hölder continuous with respect to $x \in \mathbb{R}$ ($\beta \in (0,1]$), then f satisfies the assumption of Theorem. Similarly, if $\Psi : \mathbb{R} \to \mathbb{R}$ is β-Hölder continuous, then Ψ satisfies the assumption of Theorem. As will be discussed later (Remark 2.3.8 and Table 2.1), these sufficient assumptions are weaker than these in previous works [8, 37, 23, 62].

Remark 2.3.4 (comparison with previous studies : small-noise LDP)**.** Although there are few previous works about small-noise LDP for rough volatility models when compered with short-time LDP, this Theorem is a natural extension of the previous work [37]. In [37], the authors discussed in the case of generalized rough SABR models (the case when $f \in C^{\infty}$, $\sigma \in C_b^4$, $\Psi :=$ id, and A is a Brownian motion). The main difference is that the proof of Theorem 2.3.2 is much simpler than that of [37], because our method only uses the standard rough path theory while the method of [37] need to use a partial rough path theory which is further developments of it. Also, our theorem is more flexible than the previous result in terms of f, $\mathcal{K}$, A, and Ψ. For example, one can derives the small-noise LDP for generalized rough Heston models discussed in [58] (see also Remark 2.3.8).

2.3.2 Short time asymptotics for rough volatility models (2.12)

In this subsection, we prove the LDP for $\{t^{\mu}(Y_t - Y_0)\}_{t>0}$ on $\mathbb{R}$ when $t \searrow 0$, where Y is the solution for (2.12), $\mu := \gamma - \alpha$, and γ, α are the parameter for a kernel κ. To do so, let

$$\tilde{Y}_t^{\epsilon} := \epsilon^{\mu}(Y_{\epsilon t} - Y_0).$$

Note that

$$V_{\epsilon t} = \Psi(\mathcal{K}A_{\epsilon t}) = \Psi(\mathcal{K}^\epsilon(\epsilon^\mu(A_{\epsilon\cdot} - A_0)_t) =: V_t^\epsilon,$$

where

$$\mathcal{K}^\epsilon f(t) := \epsilon^{-\mu}\left[\kappa_\epsilon(t)(f(t) - f(0)) + \int_0^t (f(s) - f(t))\frac{\mathrm{d}}{\mathrm{d}t}\kappa_\epsilon(t-s)\mathrm{d}s\right], \quad \epsilon > 0,$$

and $\kappa_\epsilon(t) := \kappa(\epsilon t)$. Here we use the relation $\mathcal{K}A(\epsilon t) = \mathcal{K}^\epsilon(\epsilon^\mu A_{\epsilon\cdot})$ in the second equality. By using the change of variables for stochastic integrals and Riemann integrals, one has that

$$\begin{aligned}
\tilde{Y}_t^\epsilon &= \epsilon^\mu\left\{\int_0^{\epsilon t}\sigma(Y_u)f(V_u,u)\mathrm{d}X_u - \frac{1}{2}\int_0^{\epsilon t}\sigma^2(Y_u)f^2(V_u,u)\mathrm{d}u\right\} \\
&= \int_0^t \tilde{\sigma}^\epsilon(\tilde{Y}_u^\epsilon)\mathrm{d}(Z^{(1)})_u^\epsilon - \frac{1}{2}\int_0^t(\tilde{\sigma}^\epsilon)^2(\tilde{Y}_u^\epsilon)\mathrm{d}(Z^{(2)})_u^\epsilon,
\end{aligned}$$

where

$$\begin{aligned}
(Z^{(1)})_t^\epsilon &:= \int_0^t f(V_u^\epsilon, \epsilon u)\mathrm{d}(\epsilon^\mu X_{\epsilon u}), \\
(Z^{(2)})_t^\epsilon &:= \epsilon^{\mu+1}\int_0^t f^2(V_u^\epsilon, \epsilon u)\mathrm{d}u,
\end{aligned}$$

and

$$\tilde{\sigma}^\epsilon(s) := \sigma(Y_0 + \epsilon^{-\mu}s).$$

Note that $Z^{(1)}$ is well-defined since V^ϵ and $X_{\epsilon\cdot}$ are $(\mathcal{F}_t^\epsilon) := (\mathcal{F}_{\epsilon t})$-adapted respectively. Indeed, we can derive an LDP for $\{\tilde{Y}^\epsilon\}$ under the following Hypothesis 2.3.5.

Hypothesis 2.3.5. We assume that κ satisfies the following conditions:

1. $\alpha \in (0, 1/2)$, and $\mu < 0$, and $\sigma \in C_b^4$,
2. the Lipschitz part g of κ is in C_b^2, and $\sup_{t\in[0,1]}|g(\epsilon t) - 1| \to 0$ as $\epsilon \searrow 0$.
3. $f : \mathbb{R}\times[0,\infty) \to \mathbb{R}$ is a continuous function and satisfies the following conditions: for all $v^n, v \in C([0,1],\mathbb{R})$ with $v^n \to v$ in $C([0,1],\mathbb{R})$,

$$\sup_{\epsilon\in(0,1)}\sup_{t\in[0,1]}|f(v_t,\epsilon t) - f(v_t^n,\epsilon t)| \to 0, \quad \text{as } n \to \infty.$$

$$\sup_{t\in[0,1]}|f(v_t,\epsilon t) - f(v_t,0)| \to 0, \quad \text{as } \epsilon \searrow 0.$$

Note that the assumption 2 is harmless in the sense that all examples which appear in Remark 2.3.1 satisfy them. Note also that the assumption 3 is harmless in the sense that the all functions discussed in Remark 2.3.3 satisfy this condition.

Theorem 2.3.6. Assume that Hypothesis 2.3.5.

Then $\{\tilde{Y}^\epsilon\}_{0<\epsilon\leqq 1}$ satisfies the LDP on $C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ as $\epsilon \searrow 0$ with speed $\epsilon^{-(2\mu+1)}$ with good rate function

$$\tilde{J}(\tilde{y}) := \inf \left\{ \frac{1}{2}\|(w,w^\perp)\|_{\mathcal{H}}^2 : \begin{array}{l} x = \rho w + \sqrt{1-\rho^2}w^\perp, \\ \tilde{y} = \sigma(y_0)\displaystyle\int_0^\cdot f(\Psi\mathcal{K}_0(a(A_0)w)_r, 0)\mathrm{d}x_r,\ (w,w^\perp) \in \mathcal{H} \end{array} \right\},$$

The proof of Theorem 2.3.6 is given in Section 2.4.3.

An LDP for the marginal distribution $\tilde{Y}_1^\epsilon$ follows from the contraction principle, and the corresponding one-dimensional rate function as follows.

Theorem 2.3.7. Assume Hypothesis 2.3.5. Then $\{t^\mu(Y_t - Y_0)\}_{1\geq t>0}$ satisfies the LDP on $\mathbb{R}$ with $t \searrow 0$ with speed $t^{-(2\mu+1)}$ with good rate function

$$\tilde{J}^\#(z) := \inf_{g\in L^2([0,1])} \left[\frac{1}{2}\int_0^1 |g_r|^2 \mathrm{d}r + \frac{\left\{z - \rho\sigma(Y_0)\int_0^1 f(v(g)_r,0)\, g_r \mathrm{d}r\right\}^2}{2(1-\rho^2)\sigma(Y_0)^2 \int_0^1 f(v(g)_r,0)^2\,\mathrm{d}r} \right], \quad (2.15)$$

where $v(g) = a(A_0)\Psi(\mathcal{K}_0 g)$, and $\mathcal{K}_0 g := \int_0^\cdot \kappa_H(t-r)g_r \mathrm{d}r$.

Proof. By the contraction principle and the previous theorem, $\{t^\mu(Y_t - Y_0)\}_{1\geq t>0}$ satisfies the LDP on $\mathbb{R}$ with $t \searrow 0$ with speed $t^{-(2\mu+1)}$ with good rate function

$$\tilde{J}^\#(z) := \inf \left\{ \frac{1}{2}\|(w,w^\perp)\|_{\mathcal{H}}^2 : \begin{array}{l} x = \rho w + \sqrt{1-\rho^2}w^\perp, \\ z = \sigma(y_0)\displaystyle\int_0^1 f(\Psi\mathcal{K}_0(a(A_0)w)_t, 0)\mathrm{d}x_r,\ (w,w^\perp) \in \mathcal{H} \end{array} \right\}.$$

By using the argument in Theorem 3.8 in [37], one can prove that $\tilde{J}^\#$ has the above representation (2.15). ◻

Remark 2.3.8 (comparison with previous studies : short-time LDP)**.** The theorem is a natural extension for the results in [23, 62, 8, 37] because if Ψ = id, $a = 1$, and $b = 0$, then the statement (in particular, the rate function (2.15)) corresponds to that appeared in the previous works. First we will compare the assumption for the parameter of models (see Table 2.1). In view of the local volatility function σ, our method outperforms [23, 62, 8], although we have to assume a slightly stronger smoothness than [37] because we transform (2.12) into stratonovich SDEs. On the other hand, our assumption of the stochastic volatility function f is the most general in the sense of Remark 2.3.3. Also, our method is more flexible than the others in the sense of the fractional operator $\mathcal{K}$, A, and Ψ. Our method allows us to add a Lipschitz part of κ, a diffusion process A, and some transformation Ψ of volatility processes.

As mentioned Remark 2.3.4, our first contribution is that the proof is much simpler than that of previous works, because our method only uses the standard rough path theory, while the method of [8, 37] need to use regularity structures or a partial rough path theory which are further developments of it ([23, 62] is somewhat less applicable).

Table 2.1: Short time asymptotics for rough volatility models (2.12)

method	σ	f	$\mathcal{K}$	A	Ψ
Forde & Zhang [23]	1	Hölder continuous	$\mathcal{K}_0$	$a = 1, b = 0$	id
Jacquier et al. [62]	1	$\sqrt{\exp(x - t^{2H}/2)}$	$\mathcal{K}_0$	$a = 1, b = 0$	Id
Bayer et al. [8]	1	C^∞	$\mathcal{K}_0$	$a = 1, b = 0$	id
Fukasawa & T [37]	C_b^3	C^∞	$\mathcal{K}_0$	$a = 1, b = 0$	id
Our method	C_b^4	general (see Remark 2.3.3)	general	general	general

Moreover, our method allows for a unified treatment of short-time LDP for rough volatility models. For example, the result obtained by [37] does not contain the Forde & Zhang' result [23] since Hölder continuous function is not smooth in general.

Furthermore, one can derives the short-time LDP for generalized rough Heston models discussed in [58]:

$$\begin{aligned} \mathrm{d}Y_t &= -\frac{1}{2}V_t\mathrm{d}t + \sqrt{V_t}\mathrm{d}X_t, \quad Y_0 = 0, \\ V_t &= \tilde{\Psi}(\mathcal{K}A)_t, \end{aligned} \tag{2.16}$$

where $x \mapsto \tilde{\Psi}(x)$ is a continuous map from $C^{\gamma\text{-Hld}}([0,1])$ into $C_+^{\gamma\text{-Hld}}([0,1])$. The equation (2.16) coincides with the equation (2.12) with $\sigma = 1$ and $f(v,t) := \sqrt{v}$.

(2.16) are widely applicable, in the sense that the authors of [58] provide us how to make a numerical approximation of the solution of (2.16). Although a reason for using and studying such models is that it is expected to be consistent with the power laws of implied volatility observed in the market, there is no justifications of this expectation in the literature because f is not smooth and $\mathcal{K}$, A, or Ψ are general. One can remedy this problem in the sense that the approximation formula (2.18) which is given later and consistent to the power law of the implied volatility in the market is obtained.

Remark 2.3.9 (flexibility of the parameters of (2.12))**.** Although we adopt the generalized fractional operator $\mathcal{K}$, it is $\mathcal{K}_0$ that appears in the rate function (2.15) for the short time asymptotics of (2.12) and so the effect of $\mathcal{K}$ is partial in this sense. In other words, the generality of Lipschitz part g of κ does not affect it. On the other hand, Ψ and $a(A_0)$ truly affects the rate function for the short time asymptotics of (2.12). Also the scaling order for Y does depends on the parameter μ, and these suggests that μ does affect the implied volatility skew.

2.3.3 Short time asymptotics for put/call options and implied volatility

Let

$$\Lambda^*(x) := \begin{cases} \inf_{y>x} \tilde{J}^\#(y), & x \geq 0, \\ \inf_{y<x} \tilde{J}^\#(y), & x \leq 0. \end{cases}$$

where $\tilde{J}^{\#}$ is defined in Theorem 2.3.7. Let $(x)_+ := x \vee 0$.

Theorem 2.3.10. Under the Hypothesis 2.3.5, we have the following:

(i) we have the following small-time behavior for out of the money put option on $S_t = \exp(Y_t)$ with $S_0 = 1$:

$$-\lim_{t \searrow 0} t^{2(\mu+1/2)} \log \mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+\right] = \Lambda^*(x), \quad x \leq 0,$$

where $x := \log K$ is the log moneyness.

(ii) Moreover, if we assume that

$$\limsup_{t \searrow 0} t^{2\mu+1} \log \mathsf{E}\left[S_t^q\right] = 0, \quad q > 1, \tag{2.17}$$

then we also have the following small-time behavior for out of the money call option on $S_t = \exp(Y_t)$ with $S_0 = 1$:

$$-\lim_{t \searrow 0} t^{2(\mu+1/2)} \log \mathsf{E}\left[(S_t - \exp(xt^{-\mu}))_+\right] = \Lambda^*(x), \quad x \geq 0.$$

Proof. See Section 2.6 □

Remark 2.3.11. Because it is difficult to check the exponentially integrability of S in general, the assumption (2.17) is natural, see Assumption 2.4 in [9], for example.

Corollary 2.3.12. Denote $\Sigma(x, t)$ by the implied volatility at the log moneyness x and the maturity t. Then for the rough volatility models (2.12), we have

$$\lim_{t \searrow 0} \Sigma(xt^{-\mu}, t) = \frac{|x|}{\sqrt{2\Lambda^*(x)}}, \quad x < 0. \tag{2.18}$$

Moreover, if we assume that (2.17), then

$$\lim_{t \searrow 0} \Sigma(xt^{-\mu}, t) = \frac{|x|}{\sqrt{2\Lambda^*(x)}}, \quad x > 0.$$

Proof. One can adapt the same argument as Corollary 4.13 in [23]. □

Remark 2.3.13 (observation and future works for (2.18))**.** The corollary is an extension for Corollary 4.15 in [23], and our result outperforms the previous results in the sense that σ, f, $\mathcal{K}$ and A is more general. The dependence of these parameters is determined by how the rate function $\tilde{J}^{\#}$ depends on them (see Remark 2.3.9). Note that if $-\mu$ is negative, then the steepness of the implied volatility smile is infinite V-shape as $t \searrow 0$, while it is flat when $-\mu$ is positive.

Although justifications are left in the future, I think there are several chances to apply the asymptotic formula (2.18) for practical applications. For one thing, the approximate formula (2.18) has a slightly different structure compared to previous studies [23] because of the generality of (2.12). As a result, the right-hand side of

(2.18) may be explicitly solved. Even if it cannot be explicitly solved, further precise approximation formula may be found here. Furthermore, (2.18) is more flexible than that of [23] in terms of the parameters of (2.12). This flexibility may have an impact on approximation accuracy when compared with the previous work [23]. Moreover, the formula may suggest a reasonable and concrete model which is consistent to the power law of the implied volatility. Finally, by using a numerical implementation method suggested by [23], one can apply (2.18) to the pricing of put/call options with being consistent to the power law of the implied volatility.

2.4 Proof of main theorems

2.4.1 Proof of Theorem 2.2.6 and 2.2.8

We fix $\alpha \in (0,1]$ and $\beta < \alpha$. For simplicity, we will write $C_0^{\alpha\text{-Hld}}([0,1],\mathbb{R})$ as $C_0^{\alpha\text{-Hld}}([0,1])$, and $\{X^\epsilon\}_{\epsilon>0}$ as $\{X^\epsilon\}_\epsilon$.

Definition 2.4.1. For $\delta > 0$, we define the map $G_\delta : C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1]) \to C_0^{\beta\text{-Hld}}([0,1])$ as follows:

$$G_\delta(a,x)_t := \sum_{k=1}^{\infty} a_{\tau_{k-1}^\delta} \left(x_{t\wedge\tau_k^\delta} - x_{t\wedge\tau_{k-1}^\delta} \right), \quad t \in [0,1],$$

where $\tau_0^\delta = 0$ and

$$\tau_k^\delta = \tau_k^\delta(a) := \inf\{t > \tau_{k-1}^\delta : |a_t - a_{\tau_{k-1}^\delta}| > \delta\} \wedge 1, \quad k \in \mathbb{N}.$$

Remark 2.4.2. We fix $a \in C([0,1])$. By the definition of $\{\tau_k^\delta(a)\}_k$, for all $k \in \mathbb{N}$, $\tau_k^\delta(a) < \tau_{k+1}^\delta(a)$ or $\tau_k^\delta(a) = 1$. Moreover, there exists $k \in \mathbb{N}$ such that $\tau_k^\delta(a) = 1$.

Remark 2.4.3. We fix $a \in C([0,1])$ and $\delta > 0$. Let k_0' is the smallest integer such that $\tau_{k_0'}^\delta(a) = 1$. For all $0 \le s < t \le 1$, take the smallest number $l, l' \in \mathbb{N}$ such that $s \in [\tau_l^\delta, \tau_{l+1}^\delta]$ and $t \in [\tau_{l'}^\delta, \tau_{l'+1}^\delta]$ $(l \le l' \le k_0' - 1)$. Because

$$\begin{aligned} |G_\delta(a,x)_t - G_\delta(a,x)_s| &\le \|a\|_\infty \left\{ \sum_{k=l+2}^{l'} |x_{\tau_k^\delta} - x_{\tau_{k-1}^\delta}| + |x_t - x_{\tau_{l'}^\delta}| + |x_{\tau_{l+1}^\delta} - x_s| \right\} \\ &\le \|a\|_\infty \|x\|_{\alpha\text{-Hld}} k_0' |t-s|^\alpha, \end{aligned}$$

we conclude that $G_\delta(a,x)$ belongs to $C^{\alpha\text{-Hld}}([0,1])$. Since $\beta < \alpha$, we have that $G_\delta(a,x) \in C_0^{\beta\text{-Hld}}([0,1])$.

Lemma 2.4.4. We fix $\delta > 0$. For $a(n), a \in C([0,1])$, assume that $a(n) \to a$ in $C([0,1])$. Then there exists a subsequence $\{a(n')\}_{n'}$ of $\{a(n)\}_n$ and a sequence $\{r_k^\delta\}_k$ on $[0,1]$ such that the following properties hold;

(i) for all $k \ge 1$, $\tau_k^\delta(a(n')) \to r_k^\delta$ as $n' \to \infty$, and $\{r_k^\delta\}_k$ is non-decreasing,

(ii) for all $k \geq 1$, either $r_k^\delta < \tau_k^\delta(a(n'))$ for all $n' \geq k$, or $r_k^\delta \geq \tau_k^\delta(a(n'))$ for all $n' \geq k$,

(iii) for all $k \geq 1$, $a(n')_{\tau_k^\delta(a(n'))} \to a_{r_k^\delta}$ as $n' \to \infty$,

(iv) for all $k \geq 1$, $r_k^\delta = 1$ or $r_{k-1}^\delta < r_k^\delta$,

(v) there exists $k' \in \mathbb{N}$ such that $r_{k'}^\delta = 1$.

Proof. We adapt the argument of Theorem 6.5 in [39]. We fix $\delta > 0$. For brevity, we write τ_k^δ as τ_k (we apply the same notation to r_k^δ). Since $\{\tau_1(a(n))\}_n$ is a sequence on $[0,1]$, there exists a subsequence $\{a(n^{(1)})\}_{n^{(1)}}$ of $\{a(n)\}_n$ and $r_1 \in [0,1]$ such that $\tau_1(a(n^{(1)})) \to r_1$ as $n^{(1)} \to \infty$. Since $\{\tau_2(a(n^{(1)}))\}_{n^{(1)}}$ is a sequence on $[0,1]$, there exists a subsequence $\{a(n^{(2)})\}_{n^{(2)}}$ of $\{a(n^{(1)})\}_{n^{(1)}}$ and $r_2 \in [0,1]$ such that $\tau_2(a(n^{(2)})) \to r_2$ as $n^{(2)} \to \infty$. By using the same argument, for all $k \in \mathbb{N}$, there exists a subsequence

$$\{a(n^{(k)})\}_{n^{(k)}} \subset \{a(n^{(k-1)})\}_{n^{(k-1)}} \subset \ldots \subset \{a(n^{(1)})\}_{n^{(1)}} \subset \{a(n)\}_n$$

and $r_k \in [0,1]$ such that $\tau_k(a(n^{(k)})) \to r_k$ as $n^{(k)} \to \infty$. Let $a(n'_k) := a(n_k^{(k)})$ (here n'_k means the k-th number of n'). Then we have that $\{a(n')\}_{n'}$ is a subsequence of $\{a(n)\}_n$, and for all $k \geq 1$, $\{a(n'_j)\}_{j \geq k} \subset \{a(n^{(k)})\}_{n^{(k)}}$. So we have that for $k \geq 1$, $\tau_k(a(n')) \to r_k$ as $n' \to \infty$. Since $\tau_{k-1}(a(n')) \leq \tau_k(a(n'))$, we have $r_{k-1} \leq r_k$. So this is a subsequence $\{a(n')\}_{n'}$ of $\{a(n)\}_n$ such that $\{a(n')\}_{n'}$ satisfies (i). We can also select a subsequence $\{a(n'')\}_{n''}$ of $\{a(n')\}_{n'}$ satisfies (ii). We fix this subsequence $\{a(n'')\}_{n''}$ and rewrite $\{a(n'')\}_{n''}$ as $\{a(n')\}_{n'}$ for brevity.

It is straightforward to show (iii) because of the fact that $\tau_k(a(n')) \to r_k$ and the uniform convergence of $\{a(n')\}_{n'}$.

Let us verify the property (iv). If this were not true, there exists $k \in \mathbb{N}$ such that $r = r_{k-1} = r_k < 1$. From (iii), we have

$$a(n')_{\tau_k(a(n'))} \to a_r, \quad a(n')_{\tau_{k-1}(a(n'))} \to a_r, \quad \text{as } n' \to \infty.$$

In particular, there exists $N'(\delta) \in \mathbb{N}$ such that if $n' \geq N'(\delta)$,

$$|a(n')_{\tau_k(a(n'))} - a(n')_{\tau_{k-1}(a(n'))}| < \delta.$$

On the other hand, by using (ii), we can prove that there exists $N''(k) \geq 1$ such that if $n' \geq N''(k)$, $\tau_{k-1}(a(n')) < \tau_k(a(n'))$. This is because in the case of $\tau_{k-1}(a(n')) < r_{k-1}$, $r_{k-1} < 1$ implies that $\tau_{k-1}(a(n')) < 1$ and so by Remark 2.4.2, we have that $\tau_{k-1}(a(n')) < \tau_k(a(n'))$. In the case of $\tau_{k-1}(a(n')) \geq r_{k-1}$, since $r_{k-1} < 1$, there exists $N''(k-1)$ such that if $n' \geq N''(k-1)$, $\tau_{k-1}(a(n')) < 1$ and so $\tau_{k-1}(a(n')) < \tau_k(a(n'))$. Then the definition of τ_k implies that if $n' \geq N''(k)$,

$$\delta \leq |a(n')_{\tau_k(a(n'))} - a(n')_{\tau_{k-1}(a(n'))}|,$$

and this is a contradiction.

It remains to verify (v). If this were not true, for all $k \geq 1$, $r_k < 1$. By using the same argument in the proof of (iv), we can prove that for all $k \geq 1$, there exists

$N'(k) \geq 1$ such that if $n' \geq N''(k)$, $\tau_{k-1}(a(n')) < \tau_k(a(n'))$. Then by the definition of τ_k, if $n \geq N''(k)$,

$$\delta \leq |a(n')_{\tau_k(a(n'))} - a(n')_{\tau_{k-1}(a(n'))}|.$$

Since $r_k < 1$, (iv) implies that $r_{r-1} < r_k$. Then (iii) implies that

$$\delta \leq |a_{r_k} - a_{r_{k-1}}|.$$

On the other hand, since $\{r_k\}_k$ is a non-decreasing and bounded sequence, there exists $R \in [0, 1]$ such that $r_k \to R$ as $k \to \infty$ and so the continuity of a implies that $a_{r_k} \to a_R$ as $k \to \infty$. In particular, $\{a_{r_k}\}$ is a Cauchy sequence. However, this is the contradiction. ◻

Definition 2.4.5 (Definition 6.1 [39]). Let E_1, E_2 be a metric space respectively. We say that a function $G : E_1 \to E_2$ is almost compact if for all $x \in E_1$ and $\{x(n)\}_n \subset E_1$ with $x(n) \to x$ in E_1, there exists a subsequence $\{x(n_k)\}_k$ and $y \in E_2$ such that $G(x(n_k)) \to y$ in E_2.

Lemma 2.4.6. For all $\delta > 0$, G_δ is almost compact.

Proof. We fix $\delta > 0$. Assume that $(a(n), x(n)) \to (a, x)$ in $C([0, 1]) \times C_0^{\alpha\text{-Hld}}([0, 1])$. Take a subsequence $\{a(n')\}$ of $\{a(n)\}$, and $\{r_k\}_k$ such that the properties of Lemma 2.4.4 hold. Let k_0 be the smallest number such that (v) holds in Lemma 2.4.4 and let

$$z_t := \sum_{k=1}^{\infty} a_{r_{k-1}} \left(x_{t \wedge r_k} - x_{t \wedge r_{k-1}}\right), \quad t \in [0, 1].$$

By using the same argument as $G_\delta(a, x)$ in Remark 2.4.3, one can show that

$$|z_t - z_s| \leq k_0 \|a\|_\infty \|x\|_{\alpha\text{-Hld}} |t - s|^\alpha,$$

and this implies that $z \in C_0^{\beta\text{-Hld}}([0, 1])$.

We will show that $G_\delta(a(n'), x(n')) \to z$ in $C_0^{\beta\text{-Hld}}([0, 1])$. We fix $\eta > 0$ and we will prove that there exists $N(\eta) \in \mathbb{N}$ such that if $n \geq N(\eta)$,

$$\|G_\delta(a(n'), x(n')) - z\|_{\beta\text{-Hld}} \lesssim \eta.$$

To show the assertion, we fix $0 \leq s < t \leq 1$ and take the smallest number $l', l \in \mathbb{N}$ such that $s \in [r_{l'}, r_{l'+1}]$ and $t \in [r_l, r_{l+1}]$ $(l' \leq l \leq k_0 - 1)$.

Let $\tilde{\delta} = \tilde{\delta}(k_0, \eta) := (\min_{0 \leq k \leq k_0} |r_k - r_{k-1}|) \wedge 1$ and $N(\eta) := \max_{0 \leq k \leq k_0} N(k, \eta)$, where $N(k, \eta)$ is a number such that if $n' \geq N(\eta, k)$: for $0 \leq k \leq k_0$,

(I) $|\tau_k(a(n')) - r_k| < \tilde{\delta}/2$,

(II) $|a(n')_{\tau_k(a(n'))} - a_{r_k}| < \eta$,

(III) $\|a(n') - a\|_\infty < \eta$ and $\|x(n') - x\|_\infty < \eta$.

Now we fix $n' \geq N(\eta)$. For brevity, we write $\tau_k(a(n'))$ as τ_k. By (I), we can prove that $r_{k+1} < \tau_{k+2}$ and $\tau_{k-1} < r_k$ for all $0 \leq k \leq k_0$. Hence we consider the following nine cases.

(Case1) $s \in [\tau_{l'}, \tau_{l'+1}]$ and $t \in [\tau_l, \tau_{l+1}]$.

$$\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\leq |a(n')_{\tau_l}(x(n')_t - x(n')_{\tau_l}) - a_{r_l}(x_t - x_{r_l})| \\
&\quad + \sum_{k=l'+2}^{l} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'}}(x(n')_{\tau_{l'+1}} - x(n')_s) - a_{r_{l'}}(x_{r_{l'+1}} - x_s)| =: T_{11} + T_{12} + T_{13}.
\end{aligned}$$

Since $|\tau_l - r_l| < |t-s|$, $|\tau_l - t| < |t-s|$, and $|t - r_l| < |t-s|$, (I) to (III) imply that

$$\begin{aligned}
T_{11} &\leq |a(n')_{\tau_l} - a_{r_l}||x(n')_t - x(n')_{\tau_l}| + |a_{r_l}||x(n')_t - x(n')_{\tau_l} - (x_t - x_{r_l})| \\
&\leq |a(n')_{\tau_l} - a_{r_l}|\|x(n')\|_{\alpha\text{-Hld}}|t - \tau_l|^\alpha \\
&\quad + \|a\|_\infty \left\{|x(n')_t - x(n')_{r_l} - (x_t - x_{r_l})| + |x(n')_{r_l} - x(n')_{\tau_l}|\right\} \\
&\leq \Big[|a(n')_{\tau_l} - a_{r_l}|\|x(n')\|_{\alpha\text{-Hld}} + \|a\|_\infty\|x(n') - x\|_{\alpha\text{-Hld}} \\
&\quad + \|a\|_\infty\|x(n')\|_{\alpha\text{-Hld}}|\tau_l - r_l|^{\alpha-\beta}\Big]|t-s|^\beta \\
&\lesssim (\eta + \eta^{\alpha-\beta})|t-s|^\beta.
\end{aligned}$$

Since $|r_k - r_{k-1}| < |t-s|$, $|\tau_k - \tau_{k-1}| < |t-s|$, and $|\tau_k - r_k| < |t-s|$ for all $l'+2 \leq k \leq l$, (I) to (III) imply that

$$\begin{aligned}
T_{12} &\leq |a(n')_{\tau_{k-1}} - a_{r_{k-1}}||x(n')_{\tau_k} - x(n')_{\tau_{k-1}}| \\
&\qquad + |a_{r_{k-1}}||x(n')_{\tau_k} - x(n')_{\tau_{k-1}} - (x_{r_k} - x_{r_{k-1}})| \\
&\leq |a(n')_{\tau_{k-1}} - a_{r_{k-1}}|\|x(n')\|_{\alpha\text{-Hld}}|\tau_k - \tau_{k-1}|^\alpha \\
&\qquad + \|a\|_\infty\Big\{|x(n')_{r_k} - x(n')_{r_{k-1}} - (x_{r_k} - x_{r_{k-1}})| \\
&\qquad + |x(n')_{\tau_k} - x(n')_{\tau_{k-1}} - (x(n')_{r_k} - x(n')_{r_{k-1}})|\Big\} \\
&\leq \Big\{|a(n')_{\tau_{k-1}} - a_{r_{k-1}}|\|x(n')\|_{\alpha\text{-Hld}} + \|a\|_\infty\|x(n') - x\|_{\alpha\text{-Hld}} \\
&\qquad + \|a\|_\infty\|x(n')\|_{\alpha\text{-Hld}}\{|\tau_k - r_k|^{\alpha-\beta} + |\tau_{k-1} - r_{k-1}|^{\alpha-\beta}\}\Big\}|t-s|^\beta \\
&\lesssim (\eta + \eta^{\alpha-\beta})|t-s|^\beta,
\end{aligned}$$

and we can estimate T_{12}. Since $|\tau_{l'+1} - s| < |t-s|$, $|r_{l'+1} - s| < |t-s|$, and

$|\tau_{l'+1} - r_{l'+1}| < |t-s|$,

$$\begin{aligned}
T_{13} &\le |a(n')_{\tau_{l'}} - a_{r_{l'}}||x(n')_{\tau_{l'+1}} - x(n')_s| \\
&\quad + |a_{r_{l'}}||x(n')_{\tau_{l'+1}} - x(n')_s - (x_{r_{l'+1}} - x_s)| \\
&\le |a(n')_{\tau_{l'}} - a_{r_{l'}}|\|x(n')\|_{\alpha\text{-Hld}}|\tau_{l'+1} - s|^{\alpha} \\
&\quad + \|a\|_\infty \Big\{ |x(n')_{r_{l'+1}} - x(n')_s - (x_{r_{l'+1}} - x_s)| + |x(n')_{r_{l'+1}} - x(n')_{\tau_{l'+1}}| \Big\} \\
&\le \Big\{ |a(n')_{\tau_{l'}} - a_{r_{l'}}|\|x(n')\|_{\alpha\text{-Hld}} \\
&\quad + \|a\|_\infty \{ \|x(n') - x\|_{\alpha\text{-Hld}} + \|x(n')\|_{\alpha\text{-Hld}}|r_{l'+1} - \tau_{l'+1}|^{\alpha-\beta} \} \Big\} |t-s|^{\beta} \\
&\lesssim (\eta + \eta^{\alpha-\beta})|t-s|^{\beta},
\end{aligned}$$

and so by using these inequalities, we have that

$$|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \lesssim (\eta + \eta^{\alpha-\beta})|t-s|^{\beta}.$$

(Case2) $s \in [\tau_{l'}, \tau_{l'+1}]$ and $t \in [\tau_{l+1}, \tau_{l+2}]$.

$$\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\le |a(n')_{\tau_{l+1}}(x(n')_t - x(n')_{\tau_{l+1}})| + |a(n')_{\tau_l}(x(n')_{\tau_{l+1}} - x(n')_{\tau_l}) - a_{r_l}(x_t - x_{r_l})| \\
&\quad + \sum_{k=l'+2}^{l} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'}}(x(n')_{\tau_{l'+1}} - x(n')_s) - a_{r_{l'}}(x_{r_{l'+1}} - x_s)| =: T_{21} + T_{22} + T_{23} + T_{24}.
\end{aligned}$$

We can estimate T_{23} and T_{24} as the same argument of T_{12} and T_{13} in (Case1). To estimate T_{21}, let us note that (I) implies

$$\tau_{l+1} \le t < r_{l+1} < \tau_{l+1} + \tilde{\delta}/2,$$

and so $|t - \tau_{l+1}| < \tilde{\delta}/2$. Then $|\tau_{l+1} - t| < |t-s|$ implies that

$$T_{21} \le \|a(n')\|_\infty \|x(n')\|_{\alpha\text{-Hld}}|t - \tau_{l+1}|^{\alpha-\beta}|t-s|^{\beta} \lesssim \eta^{\alpha-\beta}|t-s|^{\beta}.$$

Since $|\tau_{l+1} - \tau_l| < |t-s|$, $|r_l - \tau_l| < |t-s|$, and $|\tau_{l+1} - t| < |t-s|$, (I) to (III) imply that

$$\begin{aligned}
T_{22} &\le |a(n')_{\tau_l} - a_{r_l}||x(n')_{\tau_{l+1}} - x(n')_{\tau_l}| + |a_{r_l}||x(n')_{\tau_{l+1}} - x(n')_{\tau_l} - (x_t - x_{r_l})| \\
&\le |a(n')_{\tau_l} - a_{r_l}|\|x(n')\|_{\alpha\text{-Hld}}|\tau_{l+1} - \tau_l|^{\alpha} \\
&\quad + \|a\|_\infty \Big\{ |x(n')_{\tau_{l+1}} - x(n')_{\tau_l} - (x_{\tau_{l+1}} - x_{\tau_l})| + |x_t - x_{r_l} - (x_{\tau_{l+1}} - x_{\tau_l})| \Big\} \\
&\le \Big\{ |a(n')_{\tau_l} - a_{r_l}|\|x(n')\|_{\alpha\text{-Hld}} \\
&\quad + \|a\|_\infty \{ \|x(n') - x\|_{\alpha\text{-Hld}} + \|x\|_{\alpha\text{-Hld}}\{|t - \tau_{l+1}|^{\alpha-\beta} + |\tau_l - r_l|^{\alpha-\beta}\}\} \Big\} |t-s|^{\beta} \\
&\lesssim (\eta + \eta^{\alpha-\beta})|t-s|^{\beta}.
\end{aligned}$$

(Case3) $s \in [\tau_{l'}, \tau_{l'+1}]$ and $t \in [\tau_{l-1}, \tau_l]$.

$$
\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\leq |a_{r_l}(x_t - x_{r_l})| + |a(n')_{\tau_{l-1}}(x(n')_t - x(n')_{\tau_{l-1}}) - a_{r_{l-1}}(x_{r_l} - x_{r_{l-1}})| \\
&\quad + \sum_{k=l'+2}^{l-1} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'}}(x(n')_{\tau_{l'+1}} - x(n')_s) - a_{r_{l'}}(x_{r_{l'+1}} - x_s)| =: T_{31} + T_{32} + T_{33} + T_{34}.
\end{aligned}
$$

We can estimate T_{33} and T_{34} as the same argument T_{12} and T_{13} in (Case1). To estimate T_{31}, let us note that (I) implies

$$
r_l - \tilde{\delta}/2 < t \leq \tau_l < r_l + \tilde{\delta}/2
$$

and so $|t - r_l| < \tilde{\delta}/2$. Since $|t - r_l| < |t - s|$, (II) implies that

$$
T_{31} \leq \|a\|_\infty |x_t - x_{r_l}| \leq \|a\|_\infty \|x\|_{\alpha\text{-Hld}} |t - r_l|^\alpha \lesssim \eta^{\alpha-\beta} |t - s|^\beta.
$$

On the other hand, since $|t - r_l| < |t - s|$, $|t - \tau_{l-1}| < |t - s|$, and $|\tau_{l-1} - r_{l-1}| < |t - s|$, (I) to (III) imply that

$$
\begin{aligned}
T_{32} &\leq |a_{r_{l-1}} - a(n')_{\tau_{l-1}}||x(n')_t - x(n')_{\tau_{l-1}}| + |a_{r_{l-1}}||x(n')_t - x(n')_{\tau_{l-1}} - (x_{r_l} - x_{r_{l-1}})| \\
&\leq |a_{r_{l-1}} - a(n')_{\tau_{l-1}}| \|x(n')\|_{\alpha\text{-Hld}} |t - \tau_{l-1}|^\alpha \\
&\quad + \|a\|_\infty \Big\{ |x_t - x_{r_{l-1}} - (x(n')_t - x(n')_{r_{l-1}})| + |(x(n')_{r_{l-1}} - x(n')_{\tau_{l-1}})| + |x_{r_l} - x_t| \Big\} \\
&\leq \Big\{ |a_{r_{l-1}} - a(n')_{\tau_{l-1}}| \|x(n')\|_{\alpha\text{-Hld}} + \|a\|_\infty \Big\{ \|x(n') - x\|_{\alpha\text{-Hld}} \\
&\quad + \|x(n')\|_{\alpha\text{-Hld}} |r_{l-1} - \tau_{l-1}|^{\alpha-\beta} + \|x\|_{\alpha\text{-Hld}} |t - r_l|^{\alpha-\beta} \Big\} |t - s|^\beta \\
&\lesssim (\eta + \eta^{\alpha-\beta}) |t - s|^\beta.
\end{aligned}
$$

(Case4) $s \in [\tau_{l'+1}, \tau_{l'+2}]$ and $t \in [\tau_l, \tau_{l+1}]$.

$$
\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\leq |a(n')_{\tau_l}(x(n')_t - x(n')_{\tau_l}) - a_{r_l}(x_t - x_{r_l})| \\
&\quad + \sum_{k=l'+3}^{l} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'+1}}(x(n')_{\tau_{l'+2}} - x(n')_s) - a_{r_{l'+1}}(x_{r_{l'+2}} - x_{r_{l'+1}})| + |a_{\tau_{l'}}(x_{r_{l'+1}} - x_s)| \\
&=: T_{41} + T_{42} + T_{43} + T_{44}.
\end{aligned}
$$

We can estimate T_{41} and T_{42} as the same argument T_{11} and T_{12} in (Case1). To estimate T_{43} and T_{44}, let us note that by using the same argument in (Case2), $|s - r_{l'+1}| < \tilde{\delta}/2$. Since $|r_{l'+2} - s| < |t - s|$, $|r_{l'+1} - s| < |t - s|$, and $|r_{l'+2} - \tau_{l'+2}| <$

$|t-s|$, (I) to (III) imply that

$$
\begin{aligned}
T_{43} &\le |a(n')_{\tau_{l'+1}} - a_{r_{l'+1}}||x(n')_{\tau_{l'+2}} - x(n')_s| + |a_{r_{l'+1}}||x(n')_{\tau_{l'+2}} - x(n')_s - (x_{r_{l'+2}} - x_{r_{l'+1}})| \\
&\le |a(n')_{\tau_{l'+1}} - a_{r_{l'+1}}|\|x(n')\|_{\alpha\text{-Hld}}|\tau_{l'+2} - s|^{\alpha} \\
&\quad + \|a\|_\infty \Big\{|x(n')_{r_{l'+2}} - x(n')_s - (x_{r_{l'+2}} - x_s)| + |x_{r_{l'+1}} - x_s| + |x(n')_{r_{l'+2}} - x(n')_{\tau_{l'+2}}|\Big\} \\
&\le \Big\{|a(n')_{\tau_{l'+1}} - a_{r_{l'+1}}|\|x(n')\|_{\alpha\text{-Hld}} + \|a\|_\infty \Big\{\|x(n') - x\|_{\alpha\text{-Hld}} \\
&\quad + \|x\|_{\alpha\text{-Hld}}|r_{l'+1} - s|^{\alpha-\beta} + \|x(n')\|_{\alpha\text{-Hld}}|r_{l'+2} - \tau_{l'+2}|^{\alpha-\beta}\Big\}\Big\}|t-s|^{\beta} \\
&\lesssim (\eta + \eta^{\alpha-\beta})|t-s|^{\beta}.
\end{aligned}
$$

Since $|r_{l'+1} - s| < |t-s|$,

$$
T_{44} \le \|a\|_\infty \|x\|_{\alpha\text{-Hld}}|r_{l'+1} - s|^{\alpha} \lesssim \eta^{\alpha-\beta}|t-s|^{\beta}.
$$

(Case5) $s \in [\tau_{l'+1}, \tau_{l'+2}]$ and $t \in [\tau_{l+1}, \tau_{l+2}]$.

$$
\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\le |a(n')_{\tau_{l+1}}(x(n')_t - x(n')_{\tau_{l+1}})| \\
&\quad + |a(n')_{\tau_l}(x(n')_{\tau_{l+1}} - x(n')_{\tau_l}) - a_{r_l}(x_t - x_{r_l})| \\
&\quad + \sum_{k=l'+3}^{l} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'+1}}(x(n')_{\tau_{l'+2}} - x(n')_s) - a_{r_{l'+1}}(x_{r_{l'+2}} - x_{r_{l'+1}})| \\
&\quad + |a_{\tau_{l'}}(x_{r_{l'+1}} - x_s)| =: T_{51} + T_{52} + T_{53} + T_{54} + T_{55}.
\end{aligned}
$$

We can estimate T_{51} and T_{52} as the same argument T_{21} and T_{22} in (Case2) and T_{53} as the same argument T_{12} in (Case1). We can also estimate T_{54} and T_{55} as the same argument T_{43} and T_{44} in (Case4).

(Case6) $s \in [\tau_{l'+1}, \tau_{l'+2}]$ and $t \in [\tau_{l-1}, \tau_l]$.

$$
\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\le |a_{r_l}(x_t - x_{r_l})| + |a(n')_{\tau_{l-1}}(x(n')_t - x(n')_{\tau_{l-1}}) - a_{r_{l-1}}(x_{r_l} - x_{r_{l-1}})| \\
&\quad + \sum_{k=l'+3}^{l-1} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'+1}}(x(n')_{\tau_{l'+2}} - x(n')_s) - a_{r_{l'+1}}(x_{r_{l'+2}} - x_{r_{l'+1}})| + |a_{\tau_{l'}}(x_{r_{l'+1}} - x_s)| \\
&= T_{61} + T_{62} + T_{63} + T_{64} + T_{65}.
\end{aligned}
$$

We can estimate T_{61} and T_{62} as the same argument T_{31} and T_{32} in (Case3) and T_{63} as the same argument T_{12} in (Case1). We can also estimate T_{64} and T_{65} as the same argument T_{43} and T_{44} in (Case4).

(Case7) $s \in [\tau_{l'-1}, \tau_{l'}]$ and $t \in [\tau_l, \tau_{l+1}]$.

$$
\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\le |a(n')_{\tau_l}(x(n')_t - x(n')_{\tau_l}) - a_{r_l}(x_t - x_{r_l})| \\
&\quad + \sum_{k=l'+2}^{l} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'}}(x(n')_{\tau_{l'+1}} - x(n')_{\tau_{l'}}) - a_{r_{l'}}(x_{r_{l'+1}} - x_s)| \\
&\quad + |a(n')_{\tau_{l'-1}}(x(n')_{\tau_{l'}} - x(n')_s)| =: T_{71} + T_{72} + T_{73} + T_{74}.
\end{aligned}
$$

We can estimate T_{71} and T_{72} as the same argument T_{11} and T_{12} in (Case1). To estimate T_{73} and T_{74}, let us note that by using the same argument in (Case2), $|s - \tau_{l'}| < \tilde{\delta}/2$. Since $|\tau_{l'+1} - \tau_{l'}| < |t - s|$, $|\tau_{l'+1} - r_{l'+1}| < |t - s|$, and $|\tau_{l'} - s| < |t - s|$, (I) to (III) imply that

$$
\begin{aligned}
T_{73} &\le |a(n')_{\tau_{l'}} - a_{r_{l'}}||x(n')_{\tau_{l'+1}} - x(n')_{\tau_{l'}}| \\
&\quad + |a_{r_{l'}}||x(n')_{\tau_{l'+1}} - x(n')_{\tau_{l'}} - (x_{r_{l'+1}} - x_s)| \\
&\le |a(n')_{\tau_{l'}} - a_{r_{l'}}|\|x(n')\|_{\alpha\text{-Hld}}|\tau_{l'+1} - \tau_{l'}|^\alpha \\
&\quad + \|a\|_\infty \Big\{ |x(n')_{\tau_{l'+1}} - x(n')_{\tau_{l'}} - (x_{\tau_{l'+1}} - x_{\tau_{l'}}) + |x_{\tau_{l'+1}} - x_{\tau_{l'}} - (x_{r_{l'+1}} - x_s)| \Big\} \\
&\le \Big\{ |a(n')_{\tau_{l'}} - a_{r_{l'}}|\|x(n')\|_{\alpha\text{-Hld}} + \|a\|_\infty \Big\{ \|x(n') - x\|_{\alpha\text{-Hld}} \\
&\quad + \|x\|_{\alpha\text{-Hld}}(|\tau_{l'+1} - r_{l'+1}|^{\alpha-\beta} + |s - \tau_{l'}|^{\alpha-\beta}) \Big\} \Big\} |t - s|^\beta \lesssim (\eta + \eta^{\alpha-\beta})|t - s|^\beta.
\end{aligned}
$$

Since $|\tau_{l'} - s| < |t - s|$,

$$
T_{74} \le \|a(n')\|_\infty \|x(n')\|_{\alpha\text{-Hld}}|\tau_{l'} - s|^\alpha \lesssim \eta^{\alpha-\beta}|t - s|^\beta.
$$

(Case8) $s \in [\tau_{l'-1}, \tau_{l'}]$ and $t \in [\tau_{l+1}, \tau_{l+2}]$.

$$
\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\le |a(n')_{\tau_{l+1}}(x(n')_t - x(n')_{\tau_{l+1}})| + |a(n')_{\tau_l}(x(n')_{\tau_{l+1}} - x(n')_{\tau_l}) - a_{r_l}(x_t - x_{r_l})| \\
&\quad + \sum_{k=l'+2}^{l} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'}}(x(n')_{\tau_{l'+1}} - x(n')_{\tau_{l'}}) - a_{r_{l'}}(x_{r_{l'+1}} - x_s)| + |a(n')_{\tau_{l'-1}}(x(n')_{\tau_{l'}} - x(n')_s)| \\
&= T_{81} + T_{82} + T_{83} + T_{84} + T_{85}.
\end{aligned}
$$

We can estimate T_{81} and T_{82} as the same argument T_{21} and T_{22} in (Case2) and T_{83} as the same argument T_{12} in (Case1). We can also estimate T_{84} and T_{85} as the same argument T_{73} and T_{74} in (Case7).

(Case9) $s \in [\tau_{l'-1}, \tau_{l'}]$ and $t \in [\tau_{l-1}, \tau_l]$.

$$
\begin{aligned}
&|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \\
&\leq |a_{r_l}(x_t - x_{r_l})| + |a(n')_{\tau_{l-1}}(x(n')_t - x(n')_{\tau_{l-1}}) - a_{r_{l-1}}(x_{r_l} - x_{r_{l-1}})| \\
&\quad + \sum_{k=l'+2}^{l-1} |a(n')_{\tau_{k-1}}(x(n')_{\tau_k} - x(n')_{\tau_{k-1}}) - a_{r_{k-1}}(x_{r_k} - x_{r_{k-1}})| \\
&\quad + |a(n')_{\tau_{l'}}(x(n')_{\tau_{l'+1}} - x(n')_{\tau_{l'}}) - a_{r_{l'}}(x_{r_{l'+1}} - x_s)| + |a(n')_{\tau_{l'-1}}(x(n')_{\tau_{l'}} - x(n')_s)| \\
&=: T_{91} + T_{92} + T_{93} + T_{94} + T_{95}.
\end{aligned}
$$

We can estimate T_{91} and T_{92} as the same argument T_{31} and T_{32} in (Case3) and T_{93} as the same argument T_{12} in (Case1). We can also estimate T_{94} and T_{95} as the same argument T_{73} and T_{74} in (Case7).

By using the all estimation of (Case1)-(Case9), we conclude that

$$
|G_\delta(a(n'), x(n'))_t - G_\delta(a(n'), x(n'))_s - (z_t - z_s)| \lesssim (\eta + \eta^{\alpha-\beta})|t-s|^\beta,
$$

and this is the claim. □

Definition 2.4.7 (Definition 4.2.14 in [21])**.** Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space and $\delta, \epsilon > 0$. Let $Z^{\delta,\epsilon}$ and Z^ϵ be random functions on a metric space (E, d_E) respectively. We say that $\{Z^{\delta,\epsilon}\}_{\delta,\epsilon>0}$ are exponentially good approximation of $\{Z^\epsilon\}_{\epsilon>0}$ if for every $\eta > 0$,

$$
\{\omega \in \Omega : d_E(Z^{\delta,\epsilon}(\omega), Z^\epsilon(\omega)) > \eta\} \in \mathcal{F},
$$

and

$$
\lim_{\delta \searrow 0} \limsup_{\epsilon \searrow 0} \epsilon \log \mathsf{P}\left[d_E(Z^{\delta,\epsilon}, Z^\epsilon) > \eta\right] = -\infty.
$$

Now let us note that one can represent G_δ as a stochastic integral. For $\{\tau_k^\delta\} = \{\tau_k^\delta(A^\epsilon)\}$, let

$$
\Psi_\delta(A^\epsilon)_t := \sum_{k=1}^\infty A^\epsilon_{\tau^\delta_{k-1}} 1_{(\tau^\delta_{k-1}, \tau^\delta_k]}(t), \quad t \in [0,1].
$$

Then the definition of $\{\tau_k^\delta\}$, the process $\{\Psi_\delta(A^\epsilon)\}_{\delta,\epsilon}$ is a family of adapted, left continuous with right limits processes on $[0,1]$. Therefore, we have that

$$
G_\delta(A^\epsilon, X^\epsilon)_t = \int_0^t \Psi_\delta(A^\epsilon)_r \mathrm{d}X^\epsilon_r, \quad t \in [0,1], \tag{2.19}
$$

where the integration in the right hand side is Itô integral. By Remark 2.4.3, we have $\Psi_\delta(A^\epsilon) \cdot X^\epsilon \in C_0^{\beta\text{-Hld}}([0,1])$.

Lemma 2.4.8. If $\{X^\epsilon\}_\epsilon$ is α-Uniformly Exponentially Tight, $\{\Psi_\delta(A^\epsilon) \cdot X^\epsilon\}_{\epsilon,\delta}$ is exponentially good approximation of $\{A^\epsilon \cdot X^\epsilon\}_\epsilon$ on $C_0^{\beta\text{-Hld}}([0,1])$: for all $\eta > 0$ and $M > 0$, there exists $\delta(\eta, M) > 0$ such that if $0 < \delta < \delta(\eta, M)$,

$$
\limsup_{\epsilon \searrow 0} \epsilon \log \mathsf{P}\left[\| (\Psi_\delta(A^\epsilon) - A^\epsilon) \cdot X^\epsilon \|_{\beta\text{-Hld}} > \eta\right] \leq -M.
$$

Proof. The measurability requirement in Definition 2.4.7 is satisfied by the fact that $A^\epsilon \cdot X^\epsilon$ is an adapted process and

$$\sup_{0\le s<t\le 1} \frac{|x_t - x_s|}{|t-s|^\alpha} = \sup_{0\le s<t\le 1, s,t\in\mathbb{Q}} \frac{|x_t - x_s|}{|t-s|^\alpha}, \quad x \in C^\alpha([0,1]).$$

To verify the remaining assertion, fix $M > 0$ and $\eta > 0$. Take $K_M > 0$ such that (2.4) holds. For this η and K_M, taking δ small enough ($\eta\delta^{-1} > K_M$). Note that the definition of $\{\tau_k^\delta(A^\epsilon)\}$ implies that $|A_t^\epsilon - \Psi_\delta(A^\epsilon)_t| \le \delta$, and so $\delta^{-1}(A^\epsilon - \Psi_\delta(A^\epsilon)) \in \mathcal{B}([0,1],\mathbb{R})$. Then one has that

$$\begin{aligned} \mathsf{P}\left[\|(A^\epsilon - \Psi_\delta(A^\epsilon)) \cdot X^\epsilon\|_{\beta\text{-Hld}} > \eta\right] &= \mathsf{P}\left[\left\|\delta^{-1}(A^\epsilon - \Psi_\delta(A^\epsilon)) \cdot X^\epsilon\right\|_{\beta\text{-Hld}} > \eta\delta^{-1}\right] \\ &\le \mathsf{P}\left[\left\|\delta^{-1}(A^\epsilon - \Psi_\delta(A^\epsilon)) \cdot X^\epsilon\right\|_{\beta\text{-Hld}} > K_M\right] \\ &\le \sup_{U\in\mathcal{B}([0,1],\mathbb{R})} \mathsf{P}\left[\|U \cdot X^\epsilon\|_{\beta\text{-Hld}} > K_M\right], \end{aligned}$$

and so (2.4) implies the claim. □

Definition 2.4.9. Let X^ϵ is a random function taking value a Banach space E. We say that $\{X^\epsilon\}_{\epsilon>0}$ is exponentially tight if for all $M > 0$, there exists a compact set K_M on E such that

$$\limsup_{\epsilon\searrow 0} \epsilon \log \mathsf{P}\left[X^\epsilon \in K_M^c\right] \le -M.$$

Proof of Theorem 2.2.6. We will first prove that $\{(A^\epsilon, X^\epsilon, A^\epsilon \cdot X^\epsilon)\}_\epsilon$ is exponentially tight on $C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1]) \times C_0^{\beta\text{-Hld}}([0,1])$. Since $C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ is Polish space, the assumption implies that $\{(A^\epsilon, X^\epsilon)\}_\epsilon$ is exponentially tight on $C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ (see Exercise 4.1.10 in [21]). Lemma 2.4.6 implies that $(a,x) \mapsto, G_\delta(a,x)$ is almost compact from $C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ into $C_0^{\beta\text{-Hld}}([0,1])$. Since $\{X^\epsilon\}_\epsilon$ is α-uniformly exponentially tight, (2.19) and Lemma 2.4.8 imply that $\{G_\delta(A^\epsilon, X^\epsilon)\}_{\delta,\epsilon}$ is exponentially good approximation of $\{A^\epsilon \cdot X^\epsilon\}_\epsilon$ on $C_0^{\beta\text{-Hld}}([0,1])$. Therefore Theorem 7.1 in [39] implies that $\{(A^\epsilon, X^\epsilon, A^\epsilon \cdot X^\epsilon)\}_\epsilon$ is exponentially tight on $C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1]) \times C_0^{\beta\text{-Hld}}([0,1])$.

Let $C([0,\infty))$ is the set of all continuous function with the metric

$$d_\infty(x,y) := \sum_{n=1}^\infty \frac{1}{2^n}(1 \wedge \sup_{t\in[0,n]} |x_t - y_t|), \quad x, y \in C([0,\infty)),$$

and let $(D([0,\infty)), d_\infty)$ is the set of all cadlag function equipped with d_∞. Let $F_1 : C([0,1]) \to C([0,\infty))$ as $F_1(x)_t := x_{t\wedge 1}$ and let $F_2 : C([0,1]) \to (D([0,\infty)), d_\infty)$ as $F_2(x)_t := x_t 1_{[0,1)}(t)$. Since F_1 and F_2 are continuous and injective respectively, the contraction principle implies that $\{(F_2(A^\epsilon), F_1(X^\epsilon))\}$ satisfies the LDP on $(D([0,\infty)), d_\infty) \times C([0,\infty))$ with good rate function

$$I^{(1)}(\tilde{a},\tilde{x}) := \begin{cases} I^\#(a,x), & \exists (a,x) \in C([0,1]) \times C([0,1]) \text{ s.t. } (\tilde{a},\tilde{x}) = (F_2(a), F_1(x)), \\ \infty, & \text{otherwise.} \end{cases}$$

Note that for a real valued adapted left continuous with right limits process H and a real valued semi-martingale V, we have that

$$H \cdot F_1(V) = F_2(H) \cdot F_1(V) = F_1(H \cdot V), \tag{2.20}$$

see Theorem 5.6 in [64], for example. Then we have that for all $t \in [0, \infty)$, $U \in \mathcal{S}$, and $K > 0$,

$$\begin{aligned}
\mathsf{P}\left[\sup_{s \le t} |(U_- \cdot F_1(X^\epsilon))_s| > K\right] &= \mathsf{P}\left[\sup_{s \le t} |F_1(U_- \cdot X^\epsilon)_s| > K\right] \\
&\le \mathsf{P}\left[\sup_{s \le 1} |F_1(U_- \cdot X^\epsilon)_s| > K\right] \\
&\le \mathsf{P}\left[\|(U_-|_{[0,1]}) \cdot X^\epsilon\|_{\alpha\text{-Hld}} > K\right] \\
&\le \sup_{\tilde{U} \in \mathcal{B}([0,1],\mathbb{R})} \mathsf{P}\left[\|\tilde{U} \cdot X^\epsilon\|_{\alpha\text{-Hld}} > K\right],
\end{aligned}$$

where $U_-|_{[0,1]}$ is the restriction of U_- to $[0, 1]$. Because $\{X^\epsilon\}$ is the α-Uniformly Exponentially tight, $\{F_1(X^\epsilon)\}$ is a martingale satisfying (2.3). Then Lemma 2.2.3 implies that

$$\{(F_2(A^\epsilon), F_1(X^\epsilon), F_2(A^\epsilon) \cdot F_1(X^\epsilon))\}_\epsilon = \{(F_2(A^\epsilon), F_1(X^\epsilon), F_1(A^\epsilon \cdot X^\epsilon))\}_\epsilon$$

satisfies the LDP on $(D([0, \infty)), d_\infty) \times C([0, \infty)) \times C([0, \infty))$ with good rate function

$$\begin{aligned}
I^{(2)}(a, x, \tilde{z}) &= \begin{cases} I^\#(a, x), & \tilde{z} = F_2(a) \cdot F_1(x),\ x \in \mathbf{BV} \\ \infty, & \text{otherwise} \end{cases} \\
&= \begin{cases} I^\#(a, x), & \tilde{z} = F_1(a \cdot x),\ x \in \mathbf{BV} \\ \infty, & \text{otherwise.} \end{cases}
\end{aligned}$$

By using Lemma 4.1.5 (b) in [21], it is straightforward to prove that $\{(F_2(A^\epsilon), F_1(X^\epsilon), F_1(A^\epsilon \cdot X^\epsilon))\}_\epsilon$ satisfies the LDP on $(F_2(C([0, 1])), d_\infty) \times (F_1(C([0, 1])), d_\infty) \times (F_1(C([0, 1])), d_\infty)$ with good rate function $I^{(2)}$.

Let $\mathcal{E} := (F_2(C([0, 1])), d_\infty) \times (F_1(C([0, 1])), d_\infty) \times (F_1(C([0, 1])), d_\infty)$ and let $F_3 : \mathcal{E} \to C([0, 1]) \times C([0, 1]) \times C([0, 1])$ as

$$F_3(\tilde{a}, \tilde{x}, \tilde{z})_t := \begin{cases} (\tilde{a}_t, \tilde{x}_t, \tilde{z}_t) & t \in [0, 1) \\ (\lim_{t \nearrow 1} \tilde{a}_t, \tilde{x}_1, \tilde{z}_1) & t = 1. \end{cases}$$

Because F_3 is continuous and injective, the contraction principle implies that $\{(A^\epsilon, X^\epsilon, A^\epsilon \cdot X^\epsilon)\}_\epsilon$ satisfies the LDP on $C([0, 1]) \times C([0, 1]) \times C([0, 1])$ with good rate function

$$I(a, x, z) = \begin{cases} I^\#(a, x) & z = a \cdot x,\ x \in \mathbf{BV}, \\ \infty, & \text{otherwise.} \end{cases}$$

Therefore the inverse contraction principle (Theorem 4.2.4 in [21]) implies that

$$\{(A^\epsilon, X^\epsilon, A^\epsilon \cdot X^\epsilon)\}_\epsilon$$

satisfies the LDP on $C([0, 1]) \times C_0^{\alpha\text{-Hld}}([0, 1]) \times C_0^{\beta\text{-Hld}}([0, 1])$ with good rate function I, and this is the claim. ◻

Proof of Proposition 2.2.8. To verify $A \cdot B^\epsilon \in C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$, we fix any $(\mathcal{F}_t)$-adapted continuous processes A on $[0,1]$ and $\alpha \in [1/3, 1/2)$. For brevity, we assume that $\epsilon = 1$. Let $F_2 : C([0,\infty)) \to D([0,\infty))$ as $F_2(x)_t := x_t 1_{[0,1)}(t) + 1_{[1,\infty)}(t)$ and $\tau_n := \inf\{t \geq 0 : |F_2(A_t)| > n\}$. Then we have that τ_n is a $(\mathcal{F}_t)$-stopping time, $\tau_n \leq \tau_{n+1}$ a.s. for all $n \in \mathbb{N}$, and $\tau_n \to \infty$ as $n \to \infty$ a.s. One can also prove that $\sup_{t\in[0,1\wedge\tau_n]} |A_t| \leq n$. Let $\|x\|_{\alpha\text{-Hld},[0,c]} := |x_0| + \sup_{0\leq s<t\leq c} \frac{|x_t - x_s|}{|t-s|^\alpha}$. Then we have that

$$\begin{aligned}\mathsf{P}\left[\|A \cdot B\|_{\alpha\text{-Hld},[0,1]} < \infty\right] &\geq \mathsf{P}\left[\cap_{n=1}^\infty \{\|A \cdot B\|_{\alpha\text{-Hld},[0,1\wedge\tau_n]} < \infty\}\right] \\ &= \lim_{n\to\infty} \mathsf{P}\left[\|A \cdot B\|_{\alpha\text{-Hld},[0,1\wedge\tau_n]} < \infty\right],\end{aligned}$$

and so it is sufficient to prove that for all $n \in \mathbb{N}$,

$$\mathsf{P}\left[\|A \cdot B\|_{\alpha\text{-Hld},[0,1\wedge\tau_n]} < \infty\right] = 1. \tag{2.21}$$

Let $\tilde{A}_t^{(n)} := A_t 1_{[0,1\wedge\tau_n]}(t) + 1_{(1\wedge\tau_n,\infty)}(t)$, then $\|A \cdot B\|_{\alpha,[0,1\wedge\tau_n]} = \|\tilde{A}^{(n)} \cdot B\|_{\alpha,[0,1\wedge\tau_n]}$. Since

$$\langle \tilde{A}^{(n)} \cdot B\rangle_t = \int_0^t (\tilde{A}^{(n)})_r^2 \mathrm{d}r \to \infty, \quad t \to \infty,$$

and $\tilde{A}^{(n)} \cdot B$ is a local continuous martingale, the Dambis-Dubins-Schwarz's Theorem implies that there exists a Brownian motion $\tilde{B}$ such that

$$(\tilde{A}^{(n)} \cdot B)_t = \tilde{B}_{\langle \tilde{A}^{(n)} \cdot B\rangle_t}, \quad t \in [0,\infty). \tag{2.22}$$

We now restrict $\tilde{A}^{(n)} \cdot B$ on $[0,1]$. Since $\sup_{t\in[0,1\wedge\tau_n]} |\tilde{A}_t^{(n)}| = \sup_{t\in[0,1\wedge\tau_n]} |A_t| \leq n$, one has that $\langle \tilde{A}^{(n)} \cdot B\rangle_t \leq n^2$ for $t \in [0,1]$, and this implies that for $0 \leq s < t \leq 1$,

$$\begin{aligned}\left|\tilde{B}_{\langle \tilde{A}^{(n)} \cdot B\rangle_t} - \tilde{B}_{\langle \tilde{A}^{(n)} \cdot B\rangle_s}\right| &\leq \|\tilde{B}\|_{\alpha\text{-Hld},[0,n^2]} \left|\langle \tilde{A}^{(n)} \cdot B\rangle_t - \langle \tilde{A}^{(n)} \cdot B\rangle_s\right|^\alpha \\ &\leq n^{2\alpha} \|\tilde{B}\|_{\alpha\text{-Hld},[0,n^2]} |t-s|^\alpha.\end{aligned}$$

Combined with (2.22), we have

$$\|A \cdot B\|_{\alpha\text{-Hld},[0,1\wedge\tau_n]} = \|\tilde{A}^{(n)} \cdot B\|_{\alpha\text{-Hld},[0,1]} \leq n^{2\alpha} \|\tilde{B}\|_{\alpha\text{-Hld},[0,n^2]}, \quad \text{a.s.}$$

and this implies that for all n, we have (2.21). Hence we conclude that $A \cdot B^\epsilon \in C^{\alpha\text{-Hld}}([0,1],\mathbb{R})$,

To verify the α-Uniformly Exponentially Tightness of $\{B^\epsilon\}$, we fix $M > 0$ and $U \in \mathcal{B}([0,1],\mathbb{R})$. To regard U as a process on $[0,\infty)$, define $U_t = 1$, $t \in (1,\infty)$. Then we have

$$\langle U \cdot B\rangle_t = \int_0^t U_r^2 \mathrm{d}r \to \infty, \quad t \to \infty.$$

Since $U \cdot B$ is a continuous martingale, the Dambis-Dubins-Schwarz's Theorem implies that there exists a Brownian motion $\bar{B}'$ such that

$$(U \cdot B)_t = \bar{B}'_{\langle U \cdot B\rangle_t}, \quad t \in [0,\infty). \tag{2.23}$$

We now restrict $U \cdot B$ on $[0,1]$. Since $\sup_{t\in[0,1]} |U_t| \le 1$, one has that $\langle U \cdot B\rangle_t \le t$, and this implies that for $0 \le s < t \le 1$,

$$\begin{aligned} \left|\bar{B}'_{\langle U\cdot B\rangle_t} - \bar{B}'_{\langle U\cdot B\rangle_s}\right| &\le \|\bar{B}'\|_{\alpha\text{-Hld}} \left|\langle U \cdot B\rangle_t - \langle U \cdot B\rangle_s\right|^{\alpha} \\ &\le \|\bar{B}'\|_{\alpha\text{-Hld}} |t-s|^{\alpha} . \end{aligned}$$

Combined with (2.23), we have

$$\|U \cdot B\|_{\alpha\text{-Hld}} \le \|\bar{B}'\|_{\alpha\text{-Hld}}.$$

Since $\tilde{B}$ is an one dimensional Brownian motion, $\|\tilde{B}\|_{\alpha\text{-Hld}}$ has a Gaussian tail (Corollary 13.14 in [30]). Therefore, there exists $c > 0$ such that for all $K > 0$,

$$\begin{aligned} \mathsf{P}\left[||U \cdot B^{\epsilon}||_{\alpha\text{-Hld}} > K\right] &= \mathsf{P}\left[||U \cdot B||_{\alpha\text{-Hld}} > \epsilon^{-1/2}K\right] \\ &\le \mathsf{P}\left[\|\bar{B}'\|_{\alpha\text{-Hld}} > \epsilon^{-1/2}K\right] \le c^{-1}\exp\left(-c\epsilon^{-1}K^2\right). \end{aligned}$$

This implies that

$$\limsup_{\epsilon\searrow 0} \epsilon \log \sup_{U\in\mathcal{B}([0,1],\mathbb{R})} \mathsf{P}\left[||U \cdot B^{\epsilon}||_{\alpha\text{-Hld}} > K\right] \le -cK^2,$$

and so take K_M large enough ($cK_M^2 > M$), then we conclude that

$$\limsup_{\epsilon\searrow 0} \epsilon \log \sup_{U\in\mathcal{B}([0,1],\mathbb{R})} \mathsf{P}\left[||U \cdot B^{\epsilon}||_{\alpha\text{-Hld}} > K_M\right] \le -M,$$

and this is the claim.

It remains to verify (ii). The proof of $\bar{A} \cdot \bar{B}^{\epsilon} \in C^{\alpha\text{-Hld}}$ follows from a simple modification of (i) and so we will focus on α-Uniformly Exponentially Tightness of $\{\bar{B}^{\epsilon}\}$. We fix $U \in \mathcal{B}([0,1],\mathbb{R})$ (note that U is an $(\mathcal{F}^{\epsilon})$-adapted process). Let $\tilde{U}_t := U_t 1_{[0,1]}(t) + 1_{(1,\infty)}(t)$. Then we have that for all $\epsilon > 0$,

$$\langle \tilde{U} \cdot \bar{B}^{\epsilon}\rangle_t = \int_0^t (\tilde{U})_r^2 \mathrm{d}\langle \bar{B}^{\epsilon}\rangle_r = \int_0^t (\tilde{U})_r^2 \mathrm{d}\epsilon r \to \infty, \quad t \to \infty,$$

and so for each $\epsilon > 0$, there exists a Brownian motion $\tilde{B}^{(\epsilon)}$ such that

$$(\tilde{U} \cdot \bar{B}^{\epsilon})_t = \tilde{B}^{(\epsilon)}_{\langle \tilde{U}\cdot\bar{B}^{\epsilon}\rangle_t} = \tilde{B}^{(\epsilon)}_{\epsilon\left(\int_0^t (\tilde{U})_r^2 \mathrm{d}r\right)}, \quad t \in [0,\infty).$$

Since $\sup_{t\in[0,1]} |\int_0^t (\tilde{U})_r^2 \mathrm{d}r| \le 1$, we have that

$$\begin{aligned} \left|\tilde{B}^{(\epsilon)}_{\epsilon\left(\int_0^t (\tilde{U})_r^2 \mathrm{d}r\right)} - \tilde{B}^{(\epsilon)}_{\epsilon\left(\int_0^s (\tilde{U})_r^2 \mathrm{d}r\right)}\right| &\le \|\tilde{B}^{(\epsilon)}_{\epsilon\cdot}\|_{\alpha\text{-Hld},[0,1]} \left|\int_0^t (\tilde{U})_r^2 \mathrm{d}r - \int_0^s (\tilde{U})_r^2 \mathrm{d}r\right|^{\alpha} \\ &\le \|\tilde{B}^{(\epsilon)}_{\epsilon\cdot}\|_{\alpha\text{-Hld},[0,1]} |t-s|^{\alpha} , \end{aligned}$$

and so one has

$$\|U \cdot \bar{B}^{\epsilon}\|_{\alpha,[0,1]} \le \|\tilde{B}^{(\epsilon)}_{\epsilon\cdot}\|_{\alpha\text{-Hld},[0,1]}, \quad \text{a.s.}$$

Since $\tilde{B}^{(\epsilon)}$ is a Brownian motion, one can prove that

$$\mathsf{E}\left[|\tilde{B}^{(\epsilon)}_{\epsilon t} - \tilde{B}^{(\epsilon)}_{\epsilon s}|^p\right]^{1/p} \leq \sqrt{\epsilon}\sqrt{p}|t-s|^{1/2}.$$

Then the argument of Gaussian tails in Lemma A.17 in [30], one can prove that there exists $c > 0$ (ϵ-uniform) such that

$$\mathsf{P}\left[\|\tilde{B}^{(\epsilon)}_{\epsilon\cdot}\|_{\alpha\text{-Hld},[0,1]} \geq K\right] \leq c\exp\left(-\frac{K^2}{8ec_{\alpha}\epsilon}\right),$$

and so

$$\mathsf{P}\left[\left\|U\cdot \bar{B}^{\epsilon}\right\|_{\alpha\text{-Hld},[0,1]} > K\right] \leq \mathsf{P}\left[\|\tilde{B}^{(\epsilon)}_{\epsilon\cdot}\|_{\alpha\text{-Hld},[0,1]} > K\right] \leq c\exp\left(-\frac{K^2}{8ec_{\alpha}\epsilon}\right),$$

and we have the claim. □

2.4.2 Proof of Theorem 2.2.11 and 2.3.2

Proof of Theorem 2.2.11. We fix $\alpha \in [1/3, 1/2)$ and $1/2 > \alpha' > \alpha$ such that $\{(A^\epsilon, \tilde{A}^\epsilon, X^\epsilon)\}_\epsilon$ satisfies the LDP on $C([0,1]) \times C([0,1]) \times C_0^{\alpha'\text{-Hld}}([0,1])$ with good rate function $J^\#$. Take α'' with $\alpha' > \alpha'' > \alpha$. Note that $A^\epsilon \cdot X^\epsilon \in C^{\alpha'\text{-Hld}}([0,1])$ and $\{X^\epsilon\}_\epsilon$ is α'-Uniformly Exponentially Tight by Proposition 2.2.8 (i). Then Theorem 2.2.6 implies that $\{(A^\epsilon, \tilde{A}^\epsilon, X^\epsilon, A^\epsilon \cdot X^\epsilon)\}_\epsilon$ satisfies the LDP on $C([0,1]) \times C([0,1]) \times C_0^{\alpha'\text{-Hld}}([0,1]) \times C_0^{\alpha''\text{-Hld}}([0,1])$ with good rate function

$$J^{(1)}(a,\tilde{a},x,z) := \begin{cases} J^\#(a,\tilde{a},x) & z = a\cdot x,\ x \in \mathbf{BV} \\ \infty, & \text{otherwise.} \end{cases}$$

Since $x \mapsto \int_0^\cdot x_r \mathrm{d}r$ is continuous from $C([0,1])$ to $C^{1\text{-Hld}}([0,1])$, the contraction principle implies that $\{Z^\epsilon := (A^\epsilon\cdot X^\epsilon, \tilde{A}^\epsilon\cdot\Lambda)\}_\epsilon$ satisfies the LDP on $C_0^{\alpha''\text{-Hld}}([0,1]) \times C^{1\text{-Hld}}([0,1])$ with good rate function

$$J^{(2)}(z^{(1)}, z^{(2)}) := \inf\left\{J^\#(a,\tilde{a},x) : (z^{(1)}, z^{(2)}) = (a\cdot x, \tilde{a}\cdot\Lambda),\ x \in \mathbf{BV}\right\}.$$

We define $F : C_0^{\alpha''\text{-Hld}}([0,1],\mathbb{R}) \times C^{1\text{-Hld}}([0,1],\mathbb{R}) \to G\Omega^{\alpha\text{-Hld}}([0,1],\mathbb{R}^2)$ as (2.9):

$$F(z)_{st} := (1, z_{st}, \mathbf{z}_{st}), \quad z \in C_0^{\alpha''\text{-Hld}}(\mathbb{R}) \times C^{1\text{-Hld}}(\mathbb{R}).$$

We first prove $F(z) \in G\Omega^{\alpha\text{-Hld}}([0,1],\mathbb{R}^2)$. It is straightforward to show that $F(z) = (1, z, \mathbf{z})$ has the Chen's relation: for $s \leq u \leq t$,

$$z_{st} = z_{su} + z_{ut}, \quad \mathbf{z}_{st} = \mathbf{z}_{su} + \mathbf{z}_{ut} + z_{su}\otimes z_{ut}.$$

We also have that

$$\sup_{0\leq s<t\leq 1}\frac{|z_{st}|}{|t-s|^\alpha} < \infty, \quad \sup_{0\leq s<t\leq 1}\frac{|\mathbf{z}_{st}|}{|t-s|^{2\alpha}} < \infty,$$

by the estimate of Young integral (Theorem 6.8 in [30])

$$\left|\int_s^t (z_r^{(i)} - z_s^{(i)}) \mathrm{d}z_r^{(j)}\right| \lesssim \|z^{(i)}\|_{\alpha^{(i)}\text{-Hld}} \|z^{(j)}\|_{\alpha^{(j)}\text{-Hld}} |t-s|^{2\alpha}, \tag{2.24}$$

where $\alpha^{(i)} = \alpha''$ if $i = 1$, otherwise $\alpha^{(i)} = 1$. By using Theorem 5.25 in [30] and the estimate of Young integral, one can prove that $F(z) \in G\Omega^{\alpha\text{-Hld}}([0,1], \mathbb{R}^2)$. Now we will show that F is continuous. Assume that $z(n) \to z$ in $C_0^{\alpha''\text{-Hld}}([0,1], \mathbb{R}) \times C^{1\text{-Hld}}([0,1], \mathbb{R})$. It is sufficient to consider the continuity of $\mathbf{z}^{(ij)}$. It is obvious when $i = j = 1$. In the other case, by using (2.24), we have that

$$\begin{aligned}
|\mathbf{z}(n)_{st}^{(ij)} - \mathbf{z}_{st}^{(ij)}| &\leq \left|\int_s^t \{z(n)_r^{(i)} - z(n)_s^{(i)} - z_r^{(i)} + z_s^{(i)}\} \mathrm{d}z(n)_r^{(j)}\right| \\
&\quad + \left|\int_s^t \{z_r^{(i)} - z_s^{(i)}\} \mathrm{d}\{z(n)_r^{(j)} - z_r^{(j)}\}\right| \\
&\lesssim \|z(n)^{(i)} - z^{(i)}\|_{\alpha^{(i)}\text{-Hld}} \|z(n)^{(j)}\|_{\alpha^{(j)}\text{-Hld}} |t-s|^{2\alpha} \\
&\quad + \|z^{(i)}\|_{\alpha^{(i)}\text{-Hld}} \|z(n)^{(j)} - z^{(j)}\|_{\alpha^{(j)}\text{-Hld}} |t-s|^{2\alpha},
\end{aligned}$$

and so we have that F is continuous. Hence the contraction principle implies that $\{\mathbb{Z}^\epsilon = F(Z^\epsilon)\}_\epsilon$ satisfies the LDP on $G\Omega^{\alpha\text{-Hld}}([0,1], \mathbb{R}^2)$ with good rate function

$$J^{(3)}(\tilde{\mathbf{z}}) := \inf\left\{J^\#(a, \tilde{a}, x) : \tilde{\mathbf{z}} = F(a \cdot x, \tilde{a} \cdot \Lambda),\ x \in \mathbf{BV}\right\},$$

Because the solution map $\Phi : G\Omega^{\alpha\text{-Hld}}([0,1], \mathbb{R}^2) \to C^{\alpha\text{-Hld}}([0,1])$ is continuous, the contraction principle implies that $\{Y^\epsilon = \Phi \circ F(Z^\epsilon)\}_\epsilon$ satisfies the LDP on $C^{\alpha\text{-Hld}}([0,1])$ with good rate function

$$J(y) := \inf\left\{J^\#(a, \tilde{a}, x) : y = \Phi \circ F(a \cdot x, \tilde{a} \cdot \Lambda),\ x \in \mathbf{BV}\right\},$$

and so Theorem 2.2.9 implies the claim. □

Lemma 2.4.10. We fix $\alpha \in (1/3, 1/2)$, $\gamma \in (0,1)$, and let X and $\hat{X}$ be a stochastic process defined as (2.11) respectively. Then $\left\{\left(f(\hat{X}^\epsilon, \cdot), f^2(\hat{X}^\epsilon, \cdot), X^\epsilon\right)\right\}_\epsilon$ satisfies the LDP on $C([0,1]) \times C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ with good rate function

$$\tilde{J}^\#(a, \tilde{a}, x) = \inf\left\{\frac{1}{2}\|(w, w^\perp)\|_{\mathcal{H}}^2 : (a, \tilde{a}) = F_f \circ \mathbb{K}(w, w^\perp),\ (w, w^\perp) \in \mathcal{H}\right\}.$$

Proof. It is well-known that $\{\epsilon^{1/2}(W, W^\perp)\}_\epsilon$ satisfies the LDP on $C([0,1]) \times C([0,1])$ with good rate function

$$I^{(0)}(w, w^\perp) := \begin{cases} \frac{1}{2}\|(w, w^\perp)\|_{\mathcal{H}}^2, & (w, w^\perp) \in \mathcal{H}, \\ \infty, & \text{otherwise.} \end{cases}$$

Since $\alpha \in (0, 1/2)$ and $\|(W, W^\perp)\|_{\alpha\text{-Hld}}$ has a Gaussian tails, the inverse contraction principle (see Theorem 4.2.4 in [21]) implies that $\{\epsilon^{1/2}(W, W^\perp)\}_\epsilon$ satisfies the LDP on $C_0^{\alpha\text{-Hld}}([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ with good rate function $I^{(0)}$ (here we use the argument of Proposition 13.43 in [30]). By Theorem 1 in [33], the map $f \mapsto \mathcal{K}f$ is continuous from $C^{\alpha\text{-Hld}}([0,1])$ to $C^{\gamma\text{-Hld}}([0,1])$. Then the contraction principle implies that $\left\{\epsilon^{1/2}(\hat{X}, X) = \epsilon^{1/2}(\mathcal{K}W, \rho W + \sqrt{1-\rho^2}W^{\perp})\right\}_\epsilon$ satisfies the LDP on $C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ with good rate function

$$\tilde{J}^{(1)}(\hat{x}, x) = \inf\left\{\frac{1}{2}\|(w, w^{\perp})\|_{\mathcal{H}}^2 \middle| w \in \mathcal{H}, (\hat{x}, x) = \mathbb{K}(w, w^{\perp}),\ (w, w^{\perp}) \in \mathcal{H}\right\}.$$

Hence the contraction principle again,

$$\left\{\left(f(\hat{X}^\epsilon, \cdot), f^2(\hat{X}^\epsilon, \cdot), X^\epsilon\right)\right\}_\epsilon$$

satisfies the LDP on $C([0,1]) \times C([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ with good rate function

$$\tilde{J}^{\#}(a, \tilde{a}, x) = \inf\left\{\frac{1}{2}\|(w, w^{\perp})\|_{\mathcal{H}}^2 : (a, \tilde{a}) = F_f \circ \mathbb{K}(w, w^{\perp}),\ (w, w^{\perp}) \in \mathcal{H}\right\}.$$

and this is the claim. □

Proof of Theorem 2.3.2. Since $\sigma \in C_b^4$, the coefficient of drift term $\frac{1}{2}(\sigma^2 + \sigma\sigma')$ in (2.13) is in C_b^3. Then by Lemma 2.4.10 and Proposition 2.2.8 (i), one can apply Theorem 2.2.11 by taking

$$(A^\epsilon, \tilde{A}^\epsilon, X^\epsilon) = (f(\hat{X}^\epsilon, \cdot), f^2(\hat{X}^\epsilon, \cdot), X^\epsilon)$$

and the rate function is given by

$$\begin{aligned}\tilde{J}(y) &:= \inf\left\{\tilde{J}^{\#}(a, \tilde{a}, x) : y = \Phi \circ F\left(a \cdot x, \tilde{a} \cdot \Lambda\right),\ x \in \mathbf{BV},\right\} \\ &= \inf\left\{\frac{1}{2}\|(w, w^{\perp})\|_{\mathcal{H}}^2 : y = \Phi \circ F \circ F_f \circ \mathbb{K}(w, w^{\perp}),\ (w, w^{\perp}) \in \mathcal{H}\right\},\end{aligned}$$

and this is the claim. □

2.4.3 Proof of Theorem 2.3.6

Proof. Since $\{\epsilon^\mu(W_{\epsilon\cdot}, W^\perp_{\epsilon\cdot})\}$ and $\{\epsilon^{\mu+1/2}(W, W^\perp)\}$ are the same law, one can show that $\{\epsilon^\mu(W_{\epsilon\cdot}, W^\perp_{\epsilon\cdot})\}$ satisfies the LDP on $C([0,1]) \times C([0,1])$ with speed $\epsilon^{2\mu+1}$ with good rate function

$$I^{(0)}(w, w^\perp) := \begin{cases} \frac{1}{2}\|(w, w^\perp)\|_{\mathcal{H}}^2, & (w, w^\perp) \in \mathcal{H}, \\ \infty, & \text{otherwise.}\end{cases}$$

Since $\|(W_{\epsilon\cdot}, W^\perp_{\epsilon\cdot})\|_{\alpha\text{-Hld}}$ has a Gaussian tails (see the proof of Proposition 2.2.8 (ii)), one can also prove that $\{\epsilon^\mu(W_{\epsilon\cdot}, W^\perp_{\epsilon\cdot})\}$ satisfies the LDP on $C_0^{\alpha\text{-Hld}}([0,1]) \times C_0^{\alpha\text{-Hld}}([0,1])$ with speed $\epsilon^{2\mu+1}$ with good rate function $I^{(0)}$. Let $(W^\epsilon, (W^\perp)^\epsilon) :=$

$(\epsilon^{\mu} W_{\epsilon\cdot}, \epsilon^{\mu} W^{\perp}_{\epsilon\cdot})$ and $X^{\epsilon} := \rho W^{\epsilon} + \sqrt{1-\rho^2}(W^{\perp})^{\epsilon}$. Then the contraction principle implies that $\{(X^{\epsilon}, F(W^{\epsilon}, \Lambda))\}$ satisfies the LDP on $C_0^{\alpha\text{-Hld}}(\mathbb{R}) \times G\Omega^{\alpha\text{-Hld}}(\mathbb{R}^2)$ with speed $\epsilon^{2\mu+1}$ with good rate function

$$I^{(1)}(x, X) := \inf\{\frac{1}{2}\|(w, w^{\perp})\|^2_{\mathcal{H}} : x = \rho w + \sqrt{1-\rho^2} w^{\perp},\ X = F(w, \Lambda),\ (w, w^{\perp}) \in \mathcal{H}\},$$

where F is the Young pair, see (2.9). Let $\tilde{A}^{\epsilon}_t := \epsilon^{\mu}(A_{\epsilon t} - A_0)$. By using the change of variable for stochastic integrals and Riemann-integrals, one can prove that $\tilde{A}^{\epsilon}$ is the solution of the following Itô SDE:

$$\tilde{A}^{\epsilon}_t = \int_0^t \tilde{a}^{\epsilon}(\tilde{A}^{\epsilon}_u) \mathrm{d}(\epsilon^{\mu} W_{\epsilon u}) + \int_0^t \tilde{b}^{\epsilon}(\tilde{A}^{\epsilon}_u) \mathrm{d}u,$$

where

$$\tilde{a}^{\epsilon}(y) := a(A_0 + \epsilon^{-\mu} y), \quad \tilde{b}^{\epsilon}(y) := \epsilon^{1+\mu} b(A_0 + \epsilon^{-\mu} y).$$

Then one can show that $\tilde{A}^{\epsilon}$ is the solution of the following Stratonovich SDE,

$$\tilde{A}^{\epsilon}_t = \int_0^t \tilde{a}^{\epsilon}(\tilde{A}^{\epsilon}_u) \circ \mathrm{d}(\epsilon^{\mu} W_{\epsilon u}) - \frac{1}{2} \int_0^t \tilde{a}^{\epsilon}(\tilde{a}^{\epsilon})'(\tilde{A}^{\epsilon}_u) \mathrm{d}(\epsilon^{2\mu+1} u) + \int_0^t \tilde{b}^{\epsilon}(\tilde{A}^{\epsilon}_u) \mathrm{d}u,$$

and by Theorem 2.2.9, $\tilde{A}^{\epsilon}$ is the solution of RDE with a coefficient $(\tilde{a}^{\epsilon}, -\frac{\epsilon^{2\mu+1}}{2} \tilde{a}^{\epsilon}(\tilde{a}^{\epsilon})' + \tilde{b}^{\epsilon})$.

Let $\tilde{\Phi}_{\epsilon}$ be the solution map of RDE with the coefficient $(\tilde{a}^{\epsilon}, -\frac{\epsilon^{2\mu+1}}{2} \tilde{a}^{\epsilon}(\tilde{a}^{\epsilon})' + \tilde{b}^{\epsilon})$ i.e. $\tilde{A}^{\epsilon} = \tilde{\Phi}_{\epsilon} \circ F(W, \Lambda)$. Note that

$$\|\tilde{a}^{\epsilon}\|_{C^3_b} \lesssim \|a\|_{C^3_b}, \quad \| - (\epsilon^{2\mu+1}/2) \tilde{a}^{\epsilon}(\tilde{a}^{\epsilon})' + \tilde{b}^{\epsilon}/2\|_{C^3_b} \lesssim (\|a\|_{C^4_b} + \|b\|_{C^3_b},$$

where the proportional constant does not depend on ϵ. Let $\tilde{\Phi}_0$ be the solution map of RDE with the coefficient $(a(A_0), 0)$. Since the upper bound of $\|\cdot\|_{C^3_b}$ norm for the coefficient $(\tilde{a}^{\epsilon}, -\frac{\epsilon^{2\mu+1}}{2} \tilde{a}^{\epsilon}(\tilde{a}^{\epsilon})' + \tilde{b}^{\epsilon})$ is ϵ-uniform, we can show that $\{\tilde{\Phi}_{\epsilon}\}_{\epsilon}$ is equicontinuous, and for any $(x, X) \in C_0^{\alpha\text{-Hld}}([0,1]) \times G\Omega_{\alpha\text{-Hld}}(\mathbb{R}^2)$ with $I^{(1)}(x, X) < \infty$, $\tilde{\Phi}_{\epsilon}(X) \to \tilde{\Phi}(X)$. Then we have that for any converging sequence $(x_{\epsilon}, X_{\epsilon}) \to (x, X)$ with $I^{(1)}(x, X) < \infty$, $(x_{\epsilon}, \tilde{\Phi}_{\epsilon}(X_{\epsilon}))$ converges to $(x, \tilde{\Phi}_0(X))$, and so the extended contraction principle (Theorem 2.1 in [72]) implies that $\{(X^{\epsilon}, \tilde{A}^{\epsilon})\}$ satisfies the LDP on $C^{\alpha\text{-Hld}}(\mathbb{R}^2)$ with speed $\epsilon^{2\mu+1}$ with good rate function

$$I^{(2)}(x, \tilde{a}) := \inf\left\{\frac{1}{2}\|(w, w^{\perp})\|^2_{\mathcal{H}} : x = \rho w + \sqrt{1-\rho^2} w^{\perp},\ \tilde{a} = a(A_0) w\right\}.$$

Since $\{\mathcal{K}_{\epsilon}\}_{\epsilon>0}$ is equicontinuous and converge to the usual fractional kernel $\mathcal{K}_0$ (see Section 2.5), the extended contraction principle implies that $\{(X^{\epsilon}, V^{\epsilon})\}$ satisfies the LDP on $C^{\alpha\text{-Hld}}(\mathbb{R}) \times C(\mathbb{R})$ with speed $\epsilon^{2\mu+1}$ with good rate function

$$I^{(3)}(x, v) := \inf\left\{\frac{1}{2}\|(w, w^{\perp})\|^2_{\mathcal{H}} : x = \rho w + \sqrt{1-\rho^2} w^{\perp},\ v = \Psi \mathcal{K}_0(a(A_0) w)\right\},$$

By using the assumption of f, and Proposition 2.2.8 (ii), we can apply the contraction principle and Theorem 2.2.6, and so $\{F(Z^\epsilon) := F((Z^{(1)})^\epsilon, (Z^{(2)})^\epsilon)\}$ satisfies the LDP on $G\Omega_{\alpha\text{-Hld}}(\mathbb{R}^2)$ with speed $\epsilon^{2\mu+1}$ with good rate function

$$I^{(4)}(X) := \inf \left\{ \frac{1}{2}\|(w, w^\perp)\|^2_{\mathcal{H}} : \begin{array}{l} x = \rho w + \sqrt{1-\rho^2} w^\perp, \\ X = F(f(\Psi\mathcal{K}_0(a(A_0)w), 0) \cdot x, 0) \end{array} \right\},$$

Let Φ_ϵ be the solution map of RDE with the coefficient $(\tilde{\sigma}^\epsilon, -\frac{1}{2}\{\tilde{\sigma}^\epsilon(\tilde{\sigma}^\epsilon)' + (\tilde{\sigma}^\epsilon)^2\})$ i.e. $\tilde{Y}^\epsilon = \Phi_\epsilon \circ F(Z^\epsilon)$. Since the same reason as $\tilde{\Phi}_\epsilon$, $\{\Phi_\epsilon\}_{\epsilon\in(0,1]}$ is equcontinuous, and one can prove that for any sequence with $X_\epsilon \to X$ with $I^{(4)}(X) < \infty$, $\tilde{\Phi}_\epsilon(X_\epsilon) \to \tilde{\Phi}_0(X)$. Therefore, the extended contraction principle [72] implies that $\{\tilde{Y}^\epsilon\}$ satisfies the LDP on $C^{\alpha\text{-Hld}}$ with speed $\epsilon^{2\mu+1}$ with good rate function

$$\tilde{J}(\tilde{y}) := \inf \left\{ \frac{1}{2}\{(w, w^\perp)\|^2_{\mathcal{H}} : \begin{array}{l} x = \rho w + \sqrt{1-\rho^2} w^\perp, \\ \tilde{y} = \sigma(y_0) \displaystyle\int_0^\cdot f(\Psi\mathcal{K}_0(a(A_0)w)_r, 0) \mathrm{d}x_r \end{array} \right\},$$

and this is the claim. □

2.5 Some properties for $\mathcal{K}^\epsilon$

Proposition 2.5.1. We fix $\alpha, \gamma \in (0, 1)$. Under the Hypothesis 2.3.5, we have the following:

1. $\{\mathcal{K}^\epsilon\}_{\epsilon\in(0,1]}$ is equicontinuous,
2. for all $f \in C^\alpha$, $\mathcal{K}^\epsilon f$ converges to $\mathcal{K}_0 f$.

Proof. (i) Since $\mathcal{K}^\epsilon$ is linear, it is enough to show that for any $f \in C^{\alpha\text{-Hld}}$, there exists a constant $C > 0$ (uniformly ϵ and f) such that $\|\mathcal{K}^\epsilon f\|_\gamma \le C\|f\|_\alpha$. First note that

$$|\kappa_\epsilon(t)| \lesssim \epsilon^\mu t^\mu, \quad \left|\frac{\mathrm{d}}{\mathrm{d}t}\kappa_\epsilon(t)\right| \lesssim \epsilon^\mu t^{\mu-1}, \quad \left|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\kappa_\epsilon(t)\right| \lesssim \epsilon^\mu t^{\mu-2}.$$

We will estimate the first term. Let $\phi_\epsilon(t) := \epsilon^{-\mu}\kappa_\epsilon(t)(f(t) - f(0))$. For $\forall t \in [0, 1]$ and $\forall h \in (0, 1-t]$, we have that

$$\begin{aligned} |\phi_\epsilon(t+h) - \phi_\epsilon(t)| &\le \epsilon^{-\mu}\{|\kappa_\epsilon(t+h)||f(t+h) - f(t)| + |f(t) - f(0)||\kappa_\epsilon(t+h) - \kappa_\epsilon(t)|\} \\ &\le \|f\|_\alpha h^\alpha (t+h)^\mu + \epsilon^{-\mu}\|f\|_\alpha t^\alpha \left(\int_t^{t+h} \left|\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r)\right| \mathrm{d}r\right) \\ &\lesssim \|f\|_\alpha \{|h|^\gamma + t^\alpha(t^\mu - (t+h)^\mu)\} \lesssim \|f\|_\alpha |h|^\gamma. \end{aligned}$$

Here we use that

$$t^\gamma(t+h)^\mu((t+h)^{-\mu} - t^{-\mu}) \lesssim h^\gamma, \tag{2.25}$$

in the final inequality (see [67][Chapter1, Page 15]). The case $h < 0$ is analogous.

We now consider the second term. Let

$$\varphi_\epsilon(t) := \epsilon^{-\mu}\int_0^t (f(t)-f(s))\frac{\mathrm{d}}{\mathrm{d}t}\kappa_\epsilon(t-s)\mathrm{d}s = \epsilon^{-\mu}\int_0^t (f(t)-f(t-r))\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r)\mathrm{d}r.$$

Then the change of variables implies that

$$\begin{aligned}\varphi_\epsilon(t+h)-\varphi_\epsilon(t) &= \epsilon^{-\mu}\int_0^t (f(t)-f(t-r))\left(\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r+h)-\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r)\right)\mathrm{d}r\\ &\quad+\epsilon^{-\mu}\int_0^t (f(t+h)-f(t))\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r+h)\mathrm{d}r\\ &\quad+\epsilon^{-\mu}\int_{-h}^0 (f(t+h)-f(t-r))\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r+h)\mathrm{d}r =: I_1+I_2+I_3.\end{aligned}$$

Then we have that

$$\begin{aligned}|I_1| &\le \epsilon^{-\mu}\int_0^t\int_r^{r+h}|f(t)-f(t-r)|\left|\frac{\mathrm{d}^2}{\mathrm{d}u^2}\kappa_\epsilon(u)\right|\mathrm{d}u\mathrm{d}r\\ &\le \|f\|_\alpha\int_0^t\int_r^{r+h} r^\alpha u^{\mu-2}\mathrm{d}u\mathrm{d}r\\ &\lesssim \|f\|_\alpha|h|^\gamma\left(\int_0^{t/h} r^\alpha[r^{\mu-1}-(1+r)^{\mu-1}]\mathrm{d}r\right)\lesssim \|f\|_\alpha|h|^\gamma.\end{aligned}$$

Here we use that the function $y\mapsto 1-(1+y)^{\mu-1}+(\mu-1)y$ is concave with a maximum value of 0 at $y=0$ in the last inequality. One can also obtain that

$$|I_2| \le \epsilon^{-\mu}\|f\|_\alpha\int_0^t h^\alpha\left|\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r+h)\right|\mathrm{d}r \le \|f\|_\alpha h^\alpha\int_0^t (r+h)^{\mu-1}\mathrm{d}r \lesssim \|f\|_\alpha h^\gamma$$

$$|I_3| \le \epsilon^{-\mu}\|f\|_\alpha\int_{-h}^0 (r+h)^\alpha\left|\frac{\mathrm{d}}{\mathrm{d}r}\kappa_\epsilon(r+h)\right|\mathrm{d}r \le \|f\|_\alpha\int_{-h}^0 (u+h)^{\alpha+\mu-1}\mathrm{d}r \lesssim \|f\|_\alpha h^\gamma$$

and these inequalities imply the first assertion.

(ii) The simple calculation implies that

$$\begin{aligned}&|\mathcal{K}^\epsilon f(t)-\mathcal{K}_0 f(t)-\mathcal{K}^\epsilon f(s)+\mathcal{K}_0 f(s)|\\ &\le |\phi_\epsilon(t)-(f(t)-f(0))t^\mu-\phi_\epsilon(s)+(f(s)-f(0))s^\mu|\\ &\quad+\left|\varphi_\epsilon(t)-\alpha\int_0^t (f(t)-f(u))(t-u)^{\mu-1}\mathrm{d}u-\varphi_\epsilon(s)+\alpha\int_0^s (f(s)-f(u))(s-u)^{\mu-1}\mathrm{d}u\right|\\ &\le |\epsilon^{-\mu}(f(t)-f(s))\kappa_\epsilon(t)-(f(t)-f(s))t^\mu|\\ &\quad+|\epsilon^{-\mu}(f(s)-f(0))(\kappa_\epsilon(t)-\kappa_\epsilon(s))-(f(s)-f(0))(t^\mu-s^\mu)|\\ &\quad+\left|\int_0^s (f(s)-f(t))\left(\epsilon^{-\mu}\frac{\mathrm{d}}{\mathrm{d}t}\kappa_\epsilon(t-r)-\mu(t-r)^{\mu-1}\right)\mathrm{d}r\right|\\ &\quad+\left|\int_0^s (f(r)-f(s))\left(\epsilon^{-\mu}\frac{\mathrm{d}}{\mathrm{d}t}\kappa_\epsilon(t-r)-\mu(t-r)^{\mu-1}-\epsilon^{-\mu}\frac{\mathrm{d}}{\mathrm{d}s}\kappa_\epsilon(s-r)+\mu(s-r)^{\mu-1}\right)\mathrm{d}r\right|\\ &\quad+\left|\int_s^t (f(r)-f(t))\left(\epsilon^{-\mu}\frac{\mathrm{d}}{\mathrm{d}t}\kappa_\epsilon(t-r)-\mu(t-r)^{\mu-1}\right)\mathrm{d}r\right| =: T_1+T_2+T_3+T_4+T_5,\end{aligned}$$

and so we will estimate them. T_1 is simply estimated by

$$\begin{aligned}
|T_1| &\leq \|f\|_\alpha |t-s|^\alpha |\epsilon^{-\mu}\kappa_\epsilon(t) - t^\mu| \\
&\leq \|f\|_\alpha |t-s|^\gamma \left(\frac{|t-s|}{t}\right)^{-\mu} \left(\sup_{t\in[0,1]} |g(\epsilon t)-1|\right) \lesssim \|f\|_\alpha |t-s|^\gamma \left(\sup_{t\in[0,1]} |g(\epsilon t)-1|\right).
\end{aligned}$$

We also have that

$$\begin{aligned}
|T_2| &= |(f(s)-f(0))\{(g(\epsilon t)-g(\epsilon s))t^\mu + (g(\epsilon s)-1)(t^\mu - s^\mu)\}| \\
&\leq |f(s)-f(0)| \left|\int_{\epsilon s}^{\epsilon t} \frac{\mathrm{d}}{\mathrm{d}r} g(r)\mathrm{d}r\right| t^\mu + |f(s)-f(0)||g(\epsilon s)-1||t^\mu - s^\mu| \\
&\leq \|f\|_\alpha \left\{\epsilon|t-s|s^\alpha t^\mu + s^\alpha \left(\sup_{t\in[0,1]} |g(\epsilon t)-1|\right)|t^\mu - s^\mu|\right\} \\
&\leq \|f\|_\alpha \left\{\epsilon|t-s|^\gamma + \left(\sup_{t\in[0,1]} |g(\epsilon t)-1|\right) s^\alpha (s^\mu - t^\mu)\right\} \\
&\lesssim \|f\|_\alpha \left(\epsilon + \sup_{t\in[0,1]} |g(\epsilon t)-1|\right)|t-s|^\gamma.
\end{aligned}$$

Here we use (2.25) in the last inequality. T_3 is estimated by

$$\begin{aligned}
|T_3| &\leq \|f\|_\alpha |t-s|^\alpha \int_0^s \left|\epsilon^{-\mu}\frac{\mathrm{d}}{\mathrm{d}t}\kappa_\epsilon(t-r) - \mu(t-r)^{\mu-1}\right| \mathrm{d}r \\
&\leq \|f\|_\alpha |t-s|^\alpha \int_0^s \left|\epsilon g'(\epsilon(t-r))(t-r)^\mu + \mu g(\epsilon(t-r))(t-r)^{\mu-1} - \mu(t-r)^{\mu-1}\right| \mathrm{d}r \\
&\leq \|f\|_\alpha |t-s|^\alpha \left\{\epsilon \int_0^s (t-r)^\mu \mathrm{d}r + |\mu| \int_0^s |g(\epsilon(t-r))-1|(t-r)^{\mu-1}\mathrm{d}r\right\} \\
&\lesssim \|f\|_\alpha |t-s|^\gamma \left(\epsilon + \sup_{t\in[0,1]} |g(\epsilon t)-1|\right).
\end{aligned}$$

To estimate T_4, we decompose it as follows:

$$
\begin{aligned}
|T_4| &\leq \|f\|_\alpha \int_0^s (s-r)^\alpha \Big\{ |\epsilon g'(\epsilon(t-r))(t-r)^\mu - \epsilon g'(\epsilon(s-r))(s-r)^\mu| \\
&\qquad + |\mu| \left| g(\epsilon(t-r))(t-r)^{\mu-1} - (t-r)^{\mu-1} - g(\epsilon(s-r))(s-r)^{\mu-1} + (s-r)^{\mu-1} \right| \Big\} \mathrm{d}r \\
&\lesssim \int_0^s (s-r)^\alpha \, |\epsilon g'(\epsilon(t-r))(t-r)^\mu - \epsilon g'(\epsilon(s-r))(s-r)^\mu| \, \mathrm{d}r \\
&\quad + \int_0^s (s-r)^\alpha |g(\epsilon(t-r)) - g(\epsilon(s-r))||(t-r)^{\mu-1} - (s-r)^{\mu-1}| \mathrm{d}r \\
&\quad + \int_0^s (s-r)^{\gamma-1} |g(\epsilon(t-r)) - g(\epsilon(s-r))| \mathrm{d}r \\
&\quad + \int_0^s (s-r)^\alpha |g(\epsilon(s-r)) - 1||(t-r)^{\mu-1} - (s-r)^{\mu-1}| \mathrm{d}r =: T_{41} + T_{42} + T_{43} + T_{44},
\end{aligned}
$$

The estimations of them are obtained as follows:

$$
\begin{aligned}
|T_{41}| &= \epsilon \int_0^s r^\alpha \, |\{g'(\epsilon(t-s+r)) - g'(\epsilon r)\}(t-s+r)^\mu + g'(\epsilon r)((t-s+r)^\mu - r^\mu)| \, \mathrm{d}r \\
&\leq \epsilon \int_0^s r^\alpha \left| \left(\int_{\epsilon r}^{\epsilon(t-s+r)} g''(u) \mathrm{d}u \right) (t-s+r)^\mu \right| \mathrm{d}r + \epsilon \int_0^s r^\alpha |g'(\epsilon r)||(t-s+r)^\mu - r^\mu| \mathrm{d}r \\
&\leq \epsilon^2 |t-s| \int_0^s r^\alpha (t-s+r)^\mu \mathrm{d}r + \epsilon \int_0^s r^\alpha (r^\mu - (t-s+r)^\mu) \mathrm{d}r \\
&\leq \epsilon^2 |t-s|^{1+\mu} \left(\int_0^s r^\alpha \mathrm{d}r \right) + \epsilon |t-s|^\gamma \int_0^s \mathrm{d}r \\
&\lesssim (\epsilon + \epsilon^2) |t-s|^\gamma
\end{aligned}
$$

$$
\begin{aligned}
|T_{42}| &\leq \int_0^s (s-r)^\alpha \left| \int_{\epsilon(s-r)}^{\epsilon(t-r)} g'(u) \mathrm{d}u \right| |(t-r)^{\mu-1} - (s-r)^{\mu-1}| \mathrm{d}r \\
&\leq \epsilon |t-s| \int_0^s r^\alpha |r^{\mu-1} - (t-s+r)^{\mu-1}| \mathrm{d}r \lesssim \epsilon |t-s|^\gamma,
\end{aligned}
$$

Here we use the same argument as the estimation of I_1 in (i) in the final equality.

$$
|T_{43}| \leq \int_0^s (s-r)^{\gamma-1} \left| \int_{\epsilon(s-r)}^{\epsilon(t-r)} g'(u) \mathrm{d}u \right| \mathrm{d}r \leq \epsilon |t-s| \int_0^s (s-r)^{\gamma-1} \mathrm{d}r \lesssim \epsilon |t-s|^\gamma,
$$

$$
\begin{aligned}
|T_{44}| &\leq \left(\sup_{t \in [0,1]} |g(\epsilon t) - 1| \right) \int_0^s (s-r)^\alpha |(t-r)^{\mu-1} - (s-r)^{\mu-1}| \mathrm{d}r \\
&\leq \left(\sup_{t \in [0,1]} |g(\epsilon t) - 1| \right) \int_0^s r^\alpha |r^{\mu-1} - (t-s+r)^{\mu-1}| \mathrm{d}r \lesssim \epsilon |t-s|^\gamma,
\end{aligned}
$$

Here we use again the same argument as the estimation of I_1 in (i) in the final equality. These inequalities imply the desired estimate of T_4. Finally, one has that

$$
\begin{aligned}
|T_5| &\leq \int_s^t |f(t)-f(r)| \left|\epsilon g'(\epsilon(t-r))(t-r)^{\mu} + \mu g(\epsilon(t-r))(t-r)^{\mu-1} - \mu(t-r)^{\mu-1}\right| \mathrm{d}r \\
&\leq \|f\|_{\alpha} \left\{ \epsilon \int_s^t (t-r)^{\gamma} \mathrm{d}r + |\mu| \int_s^t (t-r)^{\gamma-1} |g(\epsilon(t-r)) - 1| \mathrm{d}r \right\} \\
&\lesssim \epsilon |t-s|^{\gamma+1} + \left(\sup_{t\in[0,1]} |g(\epsilon t) - 1| \right) \int_s^t (t-r)^{\gamma-1} \mathrm{d}r \\
&\leq \left(\epsilon + \sup_{t\in[0,1]} |g(\epsilon t) - 1| \right) |t-s|^{\gamma}
\end{aligned}
$$

and so we have the claim. □

2.6 Proof of Theorem 2.3.10

Proof. (i) Lower bound

For $x \leq 0$ and $\delta > 0$, one has that

$$
\begin{aligned}
\mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+\right] &\geq \mathsf{E}\left[1_{\{\exp(x(1+\delta)t^{-\mu}) > S_t\}} (\exp(xt^{-\mu}) - S_t)\right] \\
&\geq (\exp(xt^{-\mu}) - \exp(x(1+\delta)t^{-\mu}))\, \mathsf{P}\left[\exp(x(1+\delta)t^{-\mu}) > S_t\right] \\
&\geq \exp(x(1+\delta)t^{-\mu})(-x\delta t^{-\mu}) \mathsf{P}\left[\exp(x(1+\delta)t^{-\mu}) > S_t\right].
\end{aligned}
$$

Since $\lim_{t\searrow 0} t^{2\mu+1} \log t = 0$, one has that

$$
\begin{aligned}
&\liminf_{t\searrow 0} t^{2\mu+1} \log \mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+\right] \\
&\geq \liminf_{t\searrow 0} t^{2\mu+1} \left(x(1+\delta)t^{-\mu} + \log(-x) + \log\delta + -\mu \log t + \log \mathsf{P}\left[\exp(x(1+\delta)t^{-\mu}) > S_t\right]\right) \\
&= \liminf_{t\searrow 0} t^{2\mu+1} \log \mathsf{P}\left[\exp(x(1+\delta)t^{-\mu}) > S_t\right]
\end{aligned}
$$

and so Theorem 2.3.7 implies that

$$
\liminf_{t\searrow 0} t^{2\mu+1} \log \mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+\right] \geq -\Lambda^*(x(1+\delta)),
$$

and so the continuity of Λ^* implies the lower bound.

(ii) Upper bound.

For all $q > 1$, the Hölder inequality implies that

$$
\begin{aligned}
\mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+\right] &= \mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+ 1_{\{\exp(xt^{-\mu}) > S_t\}}\right] \\
&\leq \mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+^q\right]^{1/q} \mathsf{E}\left[1_{\{\exp(xt^{-\mu}) > S_t\}}\right]^{1-1/q}
\end{aligned}
$$

and so

$$t^{2\mu+1} \log \mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+\right]$$
$$\leq \frac{t^{2\mu+1}}{q} \log \mathsf{E}\left[(\exp(xt^{-\mu}) - S_t)_+^q\right] + t^{2\mu+1}(1-1/q)\mathsf{P}\left[1_{\{\exp(xt^{-\mu})>S_t\}}\right] =: T_t^{(1)} + T_t^{(2)}.$$

Since S is positive,

$$\limsup_{t\searrow 0} T_t^{(1)} \leq \limsup_{t\searrow 0} \frac{t^{2\mu+1}}{q} \log \mathsf{E}\left[\exp(qxt^{-\mu})\right] = \limsup_{t\searrow 0} xt^{\mu+1} = 0. \tag{2.26}$$

By Theorem 2.3.7, it is also true that

$$\limsup_{t\searrow 0} T_t^{(2)} \leq -\Lambda^*(x).$$

Combined with above two inequalities one can obtain the upper bound. By (i), (ii), one can obtain the first assertion.

In the second assertion, one has to improve the estimate of (2.26) because $(\cdot - \exp(xt^{-\mu}))_+$ is not bounded. By using the assumption (2.17), one can estimate $\mathsf{E}\left[S_t^q\right]$ instead of $\mathsf{E}\left[\exp(qxt^{-\mu})\right]$ and the same argument as above implies the required bound.

□

Chapter 3

A semigroup approach to the reconstruction theorem and its applications‡

3.1 Introduction

The theory of regularity structures established by Hairer [52] provides a robust framework adapted to a wide class of (subcritical) singular stochastic PDEs. One of the most important concepts in this theory is the notion of *modelled distributions*, which are considered as "generalized Taylor expansions" of the solutions to the underlying equations. The analytic core of the theory is to prove two key theorems for modelled distributions: the *reconstruction theorem* [52, Theorem 3.10] and the *multilevel Schauder estimate* [52, Theorem 5.12]. The former theorem constructs a global distribution by gluing local distributions derived from a given modelled distribution together. The latter translates an integral operator such as the convolution operator with Green function into the operator on the space of modelled distributions. Since Hairer first proved the reconstruction theorem, some alternative proofs have been proposed using various approaches, such as Littlewood–Paley theory [45], the heat semigroup approach [70, 4], the mollification approach [75], and the convolution approach [27]. Inspired by [70], the first author of this paper proved both theorems by using the operator semigroup in [59]. On the other hand, Caravenna and Zambotti [19] introduced the notion of *germs* to describe the analytic core of the proof of the reconstruction theorem, and later, they and Broux [16] proved the multilevel Schauder estimate at the level of germs. See also [53, 20, 57, 66, 73, 17, 76, 55] for extensions of the theorems into different settings, such as Besov or Triebel–Lizorkin norms, or Riemannian manifolds. See also [28] for a Besov extension of the sewing lemma, which plays a role similar to the reconstruction

‡Reproduction of a joint work with Professor Masato Hoshino (Institute of Science Tokyo), first published in Stochastics and Partial Differential Equations: Analysis and Computations, online first, (2025), https://doi.org/10.1007/s40072-025-00352-5. Reproduced with permission from Springer Nature.

theorem in rough path theory.

In the aforementioned literatures, modelled distributions are often defined on the entire space $\mathbb{R}^d$ to avoid technical difficulties related to boundary conditions. However, it is not sufficient for applications. To apply the theory of regularity structures to parabolic equations, it is necessary to define modelled distributions on the time-space region $(0,\infty)\times\mathbb{R}^d$ allowing a singularity at the hyperplane $\{0\}\times\mathbb{R}^d$. This modified version of modelled distributions is called *singular modelled distributions*. In [52, Section 6], the reconstruction theorem and the multilevel Schauder estimate were extended to the class of singular modelled distributions. An extension to Besov norms is demonstrated in [54], and boundary conditions on both time and space variables are considered in [41]. However, compared to the case of modelled distributions without boundary conditions, there seems to be a less number of studies on alternative proofs and extensions. It should be mentioned that, in the context of rough path theory, the sewing lemma is extended into the singular path spaces allowing a singularity at time $t = 0$ by [10].

The aim of this paper is to extend the semigroup approach used in [59] and provide alternative proofs of the reconstruction theorem (see Corollary 3.3.9) and the multilevel Schauder estimate (see Corollary 3.4.6) for singular modelled distributions. The proofs use arguments similar to [59], but require the following technical modifications.

(i) Following [59], we define Besov norms using the operator semigroup $\{Q_t\}_{t>0}$. The associated integral kernel $Q_t(x,y)$ is inhomogeneous and has restricted regularities with respect to x and y in general. Hence the equivalence between the norm associated with $\{Q_t\}_{t>0}$ and the standard norm defined from Littlewood–Paley theory is uncertain. For this reason, we need some nontrivial arguments to prove the uniqueness of the reconstruction.

(ii) Since Q_t is an integral operator defined over the entire spacetime, we always require global bounds on models and modelled distributions, unlike the original definitions in [52] that assume only local bounds. Consequently, in addition to the definition of singular modelled distributions (see Definition 3.3.4) which is closer to the original one, we use a different definition that assumes global bounds (see Proposition 3.3.5-(iii)). For this reason, as for the existence of the reconstruction, we assume a stronger condition "$\eta-\gamma>-\mathfrak{s}_1$" for the parameters appearing in the definition of singular modelled distributions than the condition "$\eta>-\mathfrak{s}_1$" as in [52]. It is not actually a serious problem in applications because we can switch to a small γ to apply the reconstruction theorem.

Moreover, as an application, we discuss the parabolic Anderson model (PAM)

$$\big(\partial_1 - a(x)\Delta\big)u(t,x) = b\big(u(t,x)\big)\xi(x) \qquad ((t,x)\in(0,\infty)\times\mathbb{T}^2)$$

with a spatial white noise ξ. Here $b:\mathbb{R}\to\mathbb{R}$ is in the class C_b^3 and $a:\mathbb{T}^2\to\mathbb{R}$ is an α-Hölder continuous function for some $\alpha\in(0,1)$ and satisfies

$$C_1 \le a(x) \le C_2 \qquad (x\in\mathbb{T}^2)$$

for some constants $0<C_1<C_2$. When a is a constant, the above equation is one of the simplest examples of subcritical singular stochastic PDEs, as studied in [52, 20].

We show that the equation with general coefficients as above can be renormalized, with the spacetime dependent renormalization function (see Theorem 3.5.14). Such "non-translation invariant" equations are more generally studied by [3, 74]. The aim of this paper is to deepen the analytic core of [3], which uses the semigroup approach. On the other hand, [74] is a direct extension of [52]. One of the differences between this paper and [74] is in the requirements of the smoothness of coefficients. In [74], a bit of smoothness is required, but in this paper the coefficients only need to have positive Hölder continuities.

This paper is organized as follows. In Section 3.2, we recall from [59] Besov norms associated with the operator semigroup, and prove important inequalities used throughout this paper. In Section 3.3, we recall the basics of regularity structures and prove the reconstruction theorem for singular modelled distributions. Section 3.4 is devoted to the proof of the multilevel Schauder estimate for singular modelled distributions. In Section 3.5, we discuss an application to the two-dimensional PAM.

Notations

The symbol $\mathbb{N}$ denotes the set of all nonnegative integers. Until Section 3.4, we fix an integer $d \geq 1$, the *scaling* $\mathfrak{s} = (\mathfrak{s}_1, \dots, \mathfrak{s}_d) \in [1,\infty)^d$, and a number $\ell > 0$. We define $|\mathfrak{s}| = \sum_{i=1}^d \mathfrak{s}_i$. For any multiindex $\mathbf{k} = (k_i)_{i=1}^d \in \mathbb{N}^d$, any $x = (x_i)_{i=1}^d \in \mathbb{R}^d$, and any $t > 0$, we use the following notations.

$$\mathbf{k}! := \prod_{i=1}^d k_i!, \quad |\mathbf{k}|_{\mathfrak{s}} := \sum_{i=1}^d \mathfrak{s}_i k_i, \quad \|x\|_{\mathfrak{s}} := \sum_{i=1}^d |x_i|^{1/\mathfrak{s}_i},$$
$$x^{\mathbf{k}} := \prod_{i=1}^d x_i^{k_i}, \quad t^{\mathfrak{s}/\ell} x := (t^{\mathfrak{s}_i/\ell} x_i)_{i=1}^d, \quad t^{-\mathfrak{s}/\ell} x := (t^{-\mathfrak{s}_i/\ell} x_i)_{i=1}^d.$$

We define the set $\mathbb{N}[\mathfrak{s}] := \{|\mathbf{k}|_{\mathfrak{s}}\,; \mathbf{k} \in \mathbb{N}^d\}$, which will be used in Section 3.4. The parameter t is not a physical time variable, but an auxiliary variable used to define regularities of distributions. For multiindices $\mathbf{k} = (k_i)_{i=1}^d$ and $\mathbf{l} = (l_i)_{i=1}^d$, we write $\mathbf{l} \leq \mathbf{k}$ if $l_i \leq k_i$ for any $1 \leq i \leq d$, and then define $\binom{\mathbf{k}}{\mathbf{l}} := \prod_{i=1}^d \binom{k_i}{l_i}$.

We use the notation $A \lesssim B$ for two functions $A(x)$ and $B(x)$ of a variable x, if there exists a constant $c > 0$ independent of x such that $A(x) \leq cB(x)$ for any x.

3.2 Preliminaries

In this section, we introduce some function spaces and prove important inequalities used throughout this paper. Until Section 3.4, we fix a nonnegative measurable function $G : \mathbb{R}^d \to \mathbb{R}$ and define for any $t > 0$,

$$G_t(x) = t^{-|\mathfrak{s}|/\ell} G\big(t^{-\mathfrak{s}/\ell} x\big).$$

3.2.1 Weighted Besov space

In this subsection, we recall from [59] some basics of Besov norms associated with the operator semigroup. For simplicity, we consider only L^∞ type norms.

Definition 3.2.1. A continuous function $w : \mathbb{R}^d \to [0,1]$ which is strictly positive outside a set of Lebesgue measure 0 is called a *weight*. For any weight w, we define the weighted L^∞ norm of a measurable function $f : \mathbb{R}^d \to \mathbb{R}$ by

$$\|f\|_{L^\infty(w)} := \|fw\|_{L^\infty(\mathbb{R}^d)}.$$

We denote by $L^\infty(w)$ the space of all measurable functions with finite $L^\infty(w)$ norms, and define $C(w) = C(\mathbb{R}^d) \cap L^\infty(w)$.

While we assumed that $w(x) > 0$ for every $x \in \mathbb{R}^d$ in [59], we impose a weaker condition to consider a weight vanishing on the hyperplane $\{0\}\times\mathbb{R}^{d-1}$ in next subsection. Note that $\|\cdot\|_{L^\infty(w)}$ is nondegenerate because $w(x) > 0$ for almost every $x \in \mathbb{R}^d$. If $w(x) > 0$ for any $x \in \mathbb{R}^d$, then $C(w)$ is a closed subspace of $L^\infty(w)$.

Definition 3.2.2. A weight w is said to be *G-controlled* if $w(x) > 0$ for any $x \in \mathbb{R}^d$ and there exists a continuous function $w^* : \mathbb{R}^d \to [1,\infty)$ such that

$$w(x+y) \le w^*(x)w(y) \tag{3.1}$$

for any $x, y \in \mathbb{R}^d$ and

$$\sup_{0<t\le T} \sup_{x\in\mathbb{R}^d} \left\{ \|x\|_{\mathfrak{s}}^n \, w^*\big(t^{\mathfrak{s}/\ell}x\big) G(x) \right\} < \infty \tag{3.2}$$

for any $n \ge 0$ and $T > 0$.

From the properties (3.1) and (3.2), we have that

$$\|G_t * f\|_{L^\infty(w)} \lesssim \|f\|_{L^\infty(w)} \tag{3.3}$$

uniformly over $f \in L^\infty(w)$ and $t \in (0,T]$ for any $T > 0$. This is a particular case of [59, Lemma 2.4]. Next we introduce a semigroup of integral operators.

Definition 3.2.3. We call a family of continuous functions $\{Q_t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}\}_{t>0}$ a *G-type semigroup* if it satisfies the following properties.

(i) (Semigroup property) For any $0 < s < t$ and $x, y \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} Q_{t-s}(x,z)Q_s(z,y)dz = Q_t(x,y).$$

(ii) (Conservativity) For any $x \in \mathbb{R}^d$,

$$\lim_{t\downarrow 0} \int_{\mathbb{R}^d} Q_t(x,y)dy = 1.$$

(iii) (Upper G-type estimate) There exists a constant $C_1 > 0$ such that, for any $t > 0$ and $x, y \in \mathbb{R}^d$,

$$|Q_t(x,y)| \le C_1 G_t(x-y).$$

(iv) (Time derivative) For any $x, y \in \mathbb{R}^d$, $Q_t(x,y)$ is differentiable with respect to t. Moreover, there exists a constant $C_2 > 0$ such that, for any $t > 0$ and $x, y \in \mathbb{R}^d$,

$$|\partial_t Q_t(x,y)| \le C_2 t^{-1} G_t(x-y).$$

We fix a G-type semigroup $\{Q_t\}_{t>0}$ until Section 3.4. If w is a G-controlled weight, the linear operator on $L^\infty(w)$ defined by

$$(Q_t f)(x) =: Q_t(x,f) := \int_{\mathbb{R}^d} Q_t(x,y) f(y) dy \qquad (f \in L^\infty(w),\ x \in \mathbb{R}^d)$$

is bounded in $L^\infty(w)$ uniformly over $t \in (0,1]$, by Definition 3.2.3-(iii) and the inequality (3.3). As an important fact, $Q_t f$ is a continuous function for any $f \in L^\infty(w)$ and $t > 0$. Moreover, if $f \in C(w)$, we have

$$\lim_{t \downarrow 0} (Q_t f)(x) = f(x) \tag{3.4}$$

for any $x \in \mathbb{R}^d$. See [59, Proposition 2.8] for the proofs.

Definition 3.2.4. Let w be a G-controlled weight and let $\{Q_t\}_{t>0}$ be a G-type semigroup. For every $\alpha \le 0$, we define the Besov space $C^{\alpha,Q}(w)$ as the completion of $C(w)$ under the norm

$$\|f\|_{C^{\alpha,Q}(w)} := \sup_{0<t\le 1} t^{-\alpha/\ell} \|Q_t f\|_{L^\infty(w)}.$$

By the property (3.4), the norm $\|\cdot\|_{C^{\alpha,Q}(w)}$ is nondegenerate on $C(w)$. When $\mathfrak{s} = (1,1,\dots,1)$, $\ell = 2$, and Q_t is the heat semigroup $e^{t\Delta}$, the above norm (with $\alpha < 0$ and $w = 1$) is equivalent to the classical Besov norm in Euclidean setting, see e.g., [2, Theorem 2.34]. For more general semigroups, a similar equivalence is obtained by [18, Theorem 5.1] when the adjoint operator of Q_t also satisfies some conditions similar to those in Definition 3.2.3. As far as the authors know, without such an additional assumption for the semigroup, it is unclear whether the equivalence holds even for the case of isotropic scaling and no weight.

Remark 3.2.5. As stated in [59, Proposition 2.14], for any $\alpha_1 < \alpha_2 \le 0$, the identity $\iota_{\alpha_1} : C(w) \hookrightarrow C^{\alpha_1,Q}(w)$ is uniquely extended to the continuous injection

$$\iota_{\alpha_1}^{\alpha_2} : C^{\alpha_2,Q}(w) \hookrightarrow C^{\alpha_1,Q}(w).$$

Moreover, for any $\alpha \le 0$, the operator $Q_t : C(w) \to C(w)$ is continuously extended to the operator $Q_t^\alpha : C^{\alpha,Q}(w) \to C(w)$ and they satisfy the relation

$$Q_t^{\alpha_1} \circ \iota_{\alpha_1}^{\alpha_2} = Q_t^{\alpha_2}$$

for any $\alpha_1 < \alpha_2 \le 0$. For this compatibility, we can omit the letter α and use the notation Q_t to mean its extension Q_t^α regardless of its domain.

3.2.2 Temporal weights

In what follows, the first variable x_1 in $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ is regarded as the temporal variable, and the others $(x_2, \dots, x_d)$ are spatial variables, denoted by $x' = (x_2, \dots, x_d)$. Accordingly, we denote $\mathfrak{s}' = (\mathfrak{s}_2, \dots, \mathfrak{s}_d)$. The aim of this paper is to extend the results in [59] to norms allowing a singularity at the hyperplane $\{0\}\times\mathbb{R}^{d-1}$. We define the weight $\varphi : \mathbb{R}^d \to [0,1]$ by

$$\varphi(x) := |x_1|^{1/\mathfrak{s}_1} \wedge 1$$

and set $\varphi(x,y) := \varphi(x) \wedge \varphi(y)$. The following inequalities are used frequently throughout this paper.

Lemma 3.2.6. Let w be a G-controlled weight. For any $\alpha \geq 0$ and $\beta \in [0, \mathfrak{s}_1)$, there exists a constant C such that, for any $t \in (0,1]$ and $x \in \mathbb{R}^d$ we have

$$\int_{\mathbb{R}^d} \varphi(y)^{-\beta} \|x-y\|_{\mathfrak{s}}^{\alpha} w^*(x-y) G_t(x-y)dy \leq C t^{\alpha/\ell} \left\{ \varphi(x)^{-\beta} \wedge t^{-\beta/\ell} \right\}$$

and

$$\int_{\mathbb{R}^d} \varphi(x,y)^{-\beta} \|x-y\|_{\mathfrak{s}}^{\alpha} w^*(x-y) G_t(x-y)dy \leq C t^{\alpha/\ell} \varphi(x)^{-\beta}.$$

Proof. The second inequality immediately follows from the first one because of the trivial inequality $\varphi(x,y)^{-\beta} \leq \varphi(x)^{-\beta} + \varphi(y)^{-\beta}$. Hence we focus on the first inequality. To obtain the bound $Ct^{(\alpha-\beta)/\ell}$, we divide the integral into two parts. In the region $\{|y_1|^{1/\mathfrak{s}_1} > t^{1/\ell}\}$, since $\varphi(y)^{-\beta} \leq t^{-\beta/\ell}$ we have

$$\begin{aligned}
&\int_{|y_1|^{1/\mathfrak{s}_1} > t^{1/\ell}} \varphi(y)^{-\beta} \|x-y\|_{\mathfrak{s}}^{\alpha} w^*(x-y) G_t(x-y)dy \\
&\leq t^{-\beta/\ell} \int_{\mathbb{R}^d} \|z\|_{\mathfrak{s}}^{\alpha} w^*(z) G_t(z) dz \\
&\leq t^{(\alpha-\beta)/\ell} \int_{\mathbb{R}^d} \|z\|_{\mathfrak{s}}^{\alpha} w^*\big(t^{\mathfrak{s}/\ell} z\big) G(z) dz \lesssim t^{(\alpha-\beta)/\ell}.
\end{aligned}$$

In the region $\{|y_1|^{1/\mathfrak{s}_1} \leq t^{1/\ell}\}$, by treating the temporal variable and spatial variables separately, we have

$$\begin{aligned}
&\int_{|y_1|^{1/\mathfrak{s}_1} \leq t^{1/\ell}} \varphi(y)^{-\beta} \|x-y\|_{\mathfrak{s}}^{\alpha} w^*(x-y) G_t(x-y)dy \\
&\leq \left(\int_{|y_1|^{1/\mathfrak{s}_1} \leq t^{1/\ell}} |y_1|^{-\beta/\mathfrak{s}_1} dy_1 \right) \left(\int_{\mathbb{R}^{d-1}} \sup_{z_1 \in \mathbb{R}} \|(z_1, z')\|_{\mathfrak{s}}^{\alpha} w^*(z_1, z') G_t(z_1, z') dz' \right) \\
&\lesssim (t^{\mathfrak{s}_1/\ell})^{1-\beta/\mathfrak{s}_1} \\
&\qquad \times \left(t^{-\mathfrak{s}_1/\ell} \int_{\mathbb{R}^{d-1}} \sup_{z_1 \in \mathbb{R}} \|(t^{\mathfrak{s}_1/\ell} z_1, t^{\mathfrak{s}'/\ell} z')\|_{\mathfrak{s}}^{\alpha} w^*(t^{\mathfrak{s}_1/\ell} z_1, t^{\mathfrak{s}'/\ell} z') G(z_1, z') dz' \right) \\
&= (t^{\mathfrak{s}_1/\ell})^{1-\beta/\mathfrak{s}_1} \left(t^{-\mathfrak{s}_1/\ell + \alpha/\ell} \int_{\mathbb{R}^{d-1}} \sup_{z_1 \in \mathbb{R}} \|(z_1, z')\|_{\mathfrak{s}}^{\alpha} w^*(t^{\mathfrak{s}_1/\ell} z_1, t^{\mathfrak{s}'/\ell} z') G(z_1, z') dz' \right) \\
&\lesssim t^{(\alpha-\beta)/\ell}.
\end{aligned}$$

Therefore, we obtain the upper bound $Ct^{(\alpha-\beta)/\ell}$. Moreover, by decomposing

$$\varphi(x)^\beta \lesssim |x_1 - y_1|^{\beta/\mathfrak{s}_1} + \varphi(y)^\beta \lesssim \|x-y\|_{\mathfrak{s}}^\beta + \varphi(y)^\beta, \tag{3.5}$$

we have

$$\begin{aligned}
&\varphi(x)^\beta \int_{\mathbb{R}^d} \varphi(y)^{-\beta} \|x-y\|_{\mathfrak{s}}^\alpha \, w^*(x-y) G_t(x-y) dy \\
&\lesssim \int_{\mathbb{R}^d} \{\varphi(y)^{-\beta} \|x-y\|_{\mathfrak{s}}^{\alpha+\beta} + \|x-y\|_{\mathfrak{s}}^\alpha\} \, w^*(x-y) G_t(x-y) dy \\
&\lesssim t^{\alpha/\ell}.
\end{aligned}$$

This yields another bound $Ct^{\alpha/\ell}\varphi(x)^{-\beta}$. □

From the above lemma, we obtain an inequality similar to (3.3).

Corollary 3.2.7. Let w be a G-controlled weight. For any $\beta \in [0, \mathfrak{s}_1)$, there exists a constant C such that, for any $f \in L^\infty(\varphi^\beta w)$ we have

$$\sup_{0<t\le 1} \|G_t * f\|_{L^\infty(\varphi^\beta w)} + \sup_{0<t\le 1} t^{\beta/\ell} \|G_t * f\|_{L^\infty(w)} \le C\|f\|_{L^\infty(\varphi^\beta w)}.$$

Proof. By Lemma 3.2.6, we have

$$\begin{aligned}
w(x)|(G_t * f)(x)| &\le \int_{\mathbb{R}^d} \varphi(y)^{-\beta} w^*(x-y) G_t(x-y) \varphi(y)^\beta w(y) |f(y)| dy \\
&\le C\big\{\varphi(x)^{-\beta} \wedge t^{-\beta/\ell}\big\} \|f\|_{L^\infty(\varphi^\beta w)}.
\end{aligned}$$

□

We obtain the following assertions by arguments similar to [59].

Proposition 3.2.8. Let w be a G-controlled weight and let $\{Q_t\}_{t>0}$ be a G-type semigroup. We consider the weight $\tilde{w} := \varphi^\beta w$ for any fixed $\beta \in [0, \mathfrak{s}_1)$.

(i) For any $f \in L^\infty(\tilde{w})$ and $t > 0$, the function $Q_t f$ belongs to $C(w)$.

(ii) For any $\alpha \le 0$, the Besov norm

$$\|f\|_{C^{\alpha,Q}(\tilde{w})} := \sup_{0<t\le 1} t^{-\alpha/\ell} \|Q_t f\|_{L^\infty(\tilde{w})}$$

is nondegenerate on $C(\tilde{w})$, so we can define $C^{\alpha,Q}(\tilde{w})$ as the completion of $C(\tilde{w})$ under this norm.

(iii) For any $\alpha_1 < \alpha_2 \le 0$, the identity $\tilde{\iota}_{\alpha_1} : C(\tilde{w}) \hookrightarrow C^{\alpha_1,Q}(\tilde{w})$ is uniquely extended to the continuous injection $\tilde{\iota}_{\alpha_1}^{\alpha_2} : C^{\alpha_2,Q}(\tilde{w}) \hookrightarrow C^{\alpha_1,Q}(\tilde{w})$. For any $\alpha \le 0$, the operator $Q_t : C(\tilde{w}) \to C(\tilde{w})$ is continuously extended to the operator $\tilde{Q}_t^\alpha : C^{\alpha,Q}(\tilde{w}) \to \overline{C(\tilde{w})}$, where $\overline{C(\tilde{w})}$ is the closure of $C(\tilde{w})$ under the norm $\|\cdot\|_{L^\infty(\tilde{w})}$. Moreover, they satisfy $\tilde{Q}_t^{\alpha_1} \circ \tilde{\iota}_{\alpha_1}^{\alpha_2} = \tilde{Q}_t^{\alpha_2}$ for any $\alpha_1 < \alpha_2 \le 0$.

(iv) For any $\alpha \le 0$, the identity $i : C(w) \hookrightarrow C(\tilde{w})$ is uniquely extended to the continuous injection $i_\alpha : C^{\alpha,Q}(w) \hookrightarrow C^{\alpha,Q}(\tilde{w})$. Moreover, the extensions $\tilde{Q}_t^\alpha : C^{\alpha,Q}(\tilde{w}) \to \overline{C(\tilde{w})}$ and $Q_t^\alpha : C^{\alpha,Q}(w) \to C(w)$ defined in (iii) and Remark 3.2.5 satisfy the relation

$$i \circ Q_t^\alpha = \tilde{Q}_t^\alpha \circ i_\alpha.$$

Consequently, we can use the same notation Q_t to denote both Q_t^α and $\tilde{Q}_t^\alpha$.

(v) For any $\alpha \le 0$, there exists a constant $C > 0$ such that, for any $f \in C^{\alpha,Q}(\tilde{w})$, $t \in (0,1]$, and $\varepsilon \in [0,\ell]$, we have

$$\|(Q_t - \mathrm{id})f\|_{C^{\alpha-\varepsilon,Q}(\tilde{w})} \le C\, t^{\varepsilon/\ell} \|f\|_{C^{\alpha,Q}(\tilde{w})}.$$

The norm $C^{\alpha,Q}(\varphi^\beta w)$ is used in the proof of Theorem 3.3.7.

Proof. (i) We have $Q_t f \in L^\infty(w)$ by Corollary 3.2.7. To show the continuity of $(Q_t f)(x)$ with respect to x, it is sufficient to consider the case $t = 1$. By the property (3.2), for any fixed $R > 0$ and $n \ge 0$, the inequalities

$$w(x)|Q_1(x,y)f(y)| \lesssim w^*(x-y)w(y)G(x-y)|f(y)| \lesssim \frac{\varphi(y)^{-\beta}}{1+\|y\|_{\mathfrak{s}}^n}\|f\|_{L^\infty(\tilde{w})}$$

hold uniformly over $\|x\|_{\mathfrak{s}} \le R$ and $y \in \mathbb{R}^d$. Since $\int_{\mathbb{R}^d} \varphi(y)^{-\beta}/(1+\|y\|_{\mathfrak{s}}^n)dy < \infty$ for $n > |\mathfrak{s}|$, we have

$$\lim_{z\to x}(Q_1 f)(z)w(z) = \int_{\mathbb{R}^d} \lim_{z\to x} Q_1(z,y)f(y)w(z)dy = (Q_1 f)(x)w(x)$$

by Lebesgue's convergence theorem. Since w is strictly positive and continuous, we have $\lim_{z\to x}(Q_1 f)(z) = (Q_1 f)(x)$.

(ii) It is sufficient to show that

$$\lim_{t\downarrow 0}(Q_t f)(x) = f(x)$$

for any $f \in C(\tilde{w})$ and $x \in \mathbb{R}^d$. For any $\varepsilon > 0$, we can choose $\delta > 0$ such that $|f(y) - f(x)| < \varepsilon$ if $\|y-x\|_{\mathfrak{s}} < \delta$, and have

$$\begin{aligned}
&|w(x)(Q_t f - f)(x)| \\
&= w(x)\left| \int_{\mathbb{R}^d} Q_t(x,y)\big(f(y)-f(x)\big)dy + \left(\int_{\mathbb{R}^d} Q_t(x,y)dy - 1\right) f(x)\right| \\
&\le w(x)\varepsilon \int_{\|y-x\|_{\mathfrak{s}}<\delta} G_t(x-y)dy + w(x)\int_{\|y-x\|_{\mathfrak{s}}\ge\delta} G_t(x-y)|f(y)|dy \\
&\quad + w(x)|f(x)|\int_{\|y-x\|_{\mathfrak{s}}\ge\delta} G_t(x-y)dy + w(x)|f(x)|\left|\int_{\mathbb{R}^d} Q_t(x,y)dy - 1\right|.
\end{aligned}$$

In the far right-hand side, the only nontrivial part is the second term. We bound it from above by

$$\begin{aligned}
&\int_{\|y-x\|_{\mathfrak{s}}\geq\delta} G_t(x-y)w^*(x-y)|f(y)|w(y)dy \\
&\leq \|f\|_{L^\infty(\tilde{w})}\int_{\|y-x\|_{\mathfrak{s}}\geq\delta}\varphi(y)^{-\beta}w^*(x-y)G_t(x-y)dy \\
&\leq \|f\|_{L^\infty(\tilde{w})}\delta^{-\mathfrak{s}_1}\int_{\mathbb{R}^d}\|y-x\|_{\mathfrak{s}}^{\mathfrak{s}_1}\varphi(y)^{-\beta}w^*(x-y)G_t(x-y)dy \\
&\lesssim \|f\|_{L^\infty(\tilde{w})}\delta^{-\mathfrak{s}_1}t^{(\mathfrak{s}_1-\beta)/\mathfrak{s}_1}.
\end{aligned}$$

Since $\beta < \mathfrak{s}_1$, we obtain the convergence as $t \downarrow 0$.

The proofs of (iii) and (iv) are similar to [59, Proposition 2.14], and the proof of (v) is similar to [59, Lemma 2.15]. □

3.3 Reconstruction of singular modelled distributions

In this section, we recall from [52] the definitions of regularity structures, models, and singular modelled distributions, and prove the reconstruction theorem for singular modelled distributions using the operator semigroup. For simplicity, we consider only regularity structures, rather than general regularity-integrability structures as in [59]. Throughout this and next sections, we fix a G-type semigroup $\{Q_t\}_{t>0}$.

3.3.1 Regularity structures and models

Definition 3.3.1. A *regularity structure* $\mathscr{T} = (\mathbf{A}, \mathbf{T}, \mathbf{G})$ consists of the following objects.

(1) (Index set) $\mathbf{A}$ is a locally finite subset of $\mathbb{R}$ bounded below.

(2) (Model space) $\mathbf{T} = \bigoplus_{\alpha\in\mathbf{A}} \mathbf{T}_\alpha$ is an algebraic sum of Banach spaces $(\mathbf{T}_\alpha, \|\cdot\|_\alpha)$.

(3) (Structure group) $\mathbf{G}$ is a group of continuous linear operators on $\mathbf{T}$ such that, for any $\Gamma \in \mathbf{G}$ and $\alpha \in \mathbf{A}$,

$$(\Gamma - \mathrm{id})\mathbf{T}_\alpha \subset \mathbf{T}_{<\alpha} := \bigoplus_{\beta\in\mathbf{A},\,\beta<\alpha} \mathbf{T}_\beta.$$

The smallest element α_0 of $\mathbf{A}$ is called the *regularity* of $\mathscr{T}$. For any $\alpha \in \mathbf{A}$, we denote by $P_\alpha : \mathbf{T} \to \mathbf{T}_\alpha$ the canonical projection and write

$$\|\tau\|_\alpha := \|P_\alpha\tau\|_\alpha$$

for any $\tau \in \mathbf{T}$, by abuse of notation.

Following [59], we define the topology on the space of models by using $\{Q_t\}_{t>0}$. For two Banach spaces X and Y, we denote by $\mathcal{L}(X,Y)$ the Banach space of all continuous linear operators $X \to Y$. When $Y = \mathbb{R}$, we write $X^* := \mathcal{L}(X,\mathbb{R})$.

Definition 3.3.2. Let w be a G-controlled weight. A *smooth model* $M = (\Pi, \Gamma)$ is a pair of two families of continuous linear operators $\Pi = \{\Pi_x : \mathbf{T} \to C(w)\}_{x\in\mathbb{R}^d}$ and $\Gamma = \{\Gamma_{xy}\}_{x,y\in\mathbb{R}^d} \subset \mathbf{G}$ with the following properties.

(1) (Algebraic conditions) $\Pi_x\Gamma_{xy} = \Pi_y$, $\Gamma_{xx} = \mathrm{id}$, and $\Gamma_{xy}\Gamma_{yz} = \Gamma_{xz}$ for any $x, y, z \in \mathbb{R}^d$.

(2) (Analytic conditions) For any $\gamma \in \mathbb{R}$,

$$\begin{aligned}\|\Pi\|_{\gamma,w} &:= \max_{\alpha\in\mathbf{A},\,\alpha<\gamma}\sup_{0<t\le 1}\sup_{x\in\mathbb{R}^d}\Big(t^{-\alpha/\ell}w(x)\big\|Q_t\big(x,\Pi_x(\cdot)\big)\big\|_{\mathbf{T}_\alpha^*}\Big)\\ &= \max_{\alpha\in\mathbf{A},\,\alpha<\gamma}\sup_{0<t\le 1}\sup_{x\in\mathbb{R}^d}\sup_{\tau\in\mathbf{T}_\alpha\setminus\{0\}}\left(t^{-\alpha/\ell}w(x)\frac{|Q_t(x,\Pi_x\tau)|}{\|\tau\|_\alpha}\right) < \infty\end{aligned}$$

and

$$\begin{aligned}\|\Gamma\|_{\gamma,w} &:= \max_{\substack{\alpha,\beta\in\mathbf{A}\\ \beta<\alpha<\gamma}}\sup_{x,y\in\mathbb{R}^d,\,x\neq y}\frac{w(x)\|\Gamma_{yx}\|_{\mathcal{L}(\mathbf{T}_\alpha,\mathbf{T}_\beta)}}{w^*(y-x)\|y-x\|_{\mathfrak{s}}^{\alpha-\beta}}\\ &= \max_{\substack{\alpha,\beta\in\mathbf{A}\\ \beta<\alpha<\gamma}}\sup_{x,y\in\mathbb{R}^d,\,x\neq y}\sup_{\tau\in\mathbf{T}_\alpha\setminus\{0\}}\frac{w(x)\|\Gamma_{yx}\tau\|_\beta}{w^*(y-x)\|y-x\|_{\mathfrak{s}}^{\alpha-\beta}\|\tau\|_\alpha} < \infty.\end{aligned}$$

We write $\|\!|M|\!\|_{\gamma,w} := \|\Pi\|_{\gamma,w} + \|\Gamma\|_{\gamma,w}$. In addition, for any two smooth models $M^{(i)} = (\Pi^{(i)}, \Gamma^{(i)})$ with $i \in \{1,2\}$, we define the pseudo-metrics

$$\|\!|M^{(1)}; M^{(2)}|\!\|_{\gamma,w} := \|\Pi^{(1)} - \Pi^{(2)}\|_{\gamma,w} + \|\Gamma^{(1)} - \Gamma^{(2)}\|_{\gamma,w}$$

by replacing Π and Γ above with $\Pi^{(1)} - \Pi^{(2)}$ and $\Gamma^{(1)} - \Gamma^{(2)}$ respectively. Finally, we define the space $\mathscr{M}_w(\mathscr{T})$ as the completion of the set of all smooth models, under the pseudo-metrics $\|\!|\cdot;\cdot|\!\|_{\gamma,w}$ for all $\gamma \in \mathbb{R}$. We call each element of $\mathscr{M}_w(\mathscr{T})$ a *model* for $\mathscr{T}$. We still use the notation $M = (\Pi, \Gamma)$ to denote a generic model.

When $\ell = 2$ and Q_t is the heat semigroup $e^{t\Delta}$, the above definition essentially coincides with the original definition of models [52, Definition 2.17] if we ignore the difference between local and global bounds. For more general semigroups, such an equivalence is unclear by the same reason as the case of Besov norms.

Remark 3.3.3. As stated in [59, Proposition 3.3], if there exist two G-controlled weights w_1 and w_2 that satisfy

$$\sup_{x\in\mathbb{R}^d}\left\{\|x\|_{\mathfrak{s}}^n\, w^*(x)w_1(x)\right\} + \sup_{x\in\mathbb{R}^d}\left\{\|x\|_{\mathfrak{s}}^n\, w_1^*(x)w_2(x)\right\} < \infty$$

for any $n \geq 0$, and such that ww_1 and ww_2 are also G-controlled, then we can regard Π_x as a continuous linear operator from $\mathbf{T}$ to $C^{\alpha_0\wedge 0,Q}(ww_1)$, where α_0 is the regularity of $\mathscr{T}$. More precisely, for any $\alpha < \gamma$ and $\tau \in \mathbf{T}_\alpha$ we have

$$\sup_{x\in\mathbb{R}^d} (ww_2)(x)\|\Pi_x\tau\|_{C^{\alpha_0\wedge 0,Q}(ww_1)} \lesssim \|\Pi\|_{\gamma,w}(1+\|\Gamma\|_{\gamma,w})\|\tau\|_\alpha.$$

In what follows, we assume the existence of w_1 and w_2 as above, and regard $\Pi_x\tau$ as an element of $C^{\alpha_0\wedge 0,Q}(ww_1)$ for any $\tau \in \mathbf{T}$.

3.3.2 Singular modelled distributions

Throughout the rest of this section, we fix a regularity structure $\mathscr{T}$ of regularity α_0, and also fix G-controlled weights w and v such that wv is also G-controlled. Recall the definitions of functions $\varphi(x)$ and $\varphi(x,y)$ from Section 3.2.2.

Definition 3.3.4. Let $M = (\Pi,\Gamma) \in \mathscr{M}_w(\mathscr{T})$. For any $\gamma \in \mathbb{R}$ and $\eta \leq \gamma$, we define $\mathcal{D}_v^{\gamma,\eta}(\Gamma)$ as the space of all functions $f : (\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1} \to \mathbf{T}_{<\gamma}$ such that

$$\llparenthesis f \rrparenthesis_{\gamma,\eta,v} := \max_{\alpha<\gamma} \sup_{x\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1}} \frac{v(x)\|f(x)\|_\alpha}{\varphi(x)^{(\eta-\alpha)\wedge 0}} < \infty,$$

$$\|f\|_{\gamma,\eta,v} := \max_{\alpha<\gamma} \sup_{\substack{x,y\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1},\, x\neq y\\ \|y-x\|_\mathfrak{s}\leq\varphi(x,y)}} \frac{v(x)\|\Delta^\Gamma_{yx}f\|_\alpha}{\varphi(x,y)^{\eta-\gamma}\, v^*(x-y)\|y-x\|_\mathfrak{s}^{\gamma-\alpha}} < \infty,$$

where $\Delta^\Gamma_{yx}f := f(y) - \Gamma_{yx}f(x)$. We write $|||f|||_{\gamma,\eta,v} := \llparenthesis f \rrparenthesis_{\gamma,\eta,v} + \|f\|_{\gamma,\eta,v}$. We call each element of $\mathcal{D}_v^{\gamma,\eta}(\Gamma)$ a *singular modelled distribution*.

In addition, for any two models $M^{(i)} = (\Pi^{(i)},\Gamma^{(i)}) \in \mathscr{M}_w(\mathscr{T})$ and singular modelled distributions $f^{(i)} \in \mathcal{D}_v^{\gamma,\eta}(\Gamma^{(i)})$ with $i \in \{1,2\}$, we define $|||f^{(1)};f^{(2)}|||_{\gamma,\eta,v} := \llparenthesis f^{(1)}-f^{(2)} \rrparenthesis_{\gamma,\eta,v} + \|f^{(1)};f^{(2)}\|_{\gamma,\eta,v}$ by

$$\llparenthesis f^{(1)}-f^{(2)} \rrparenthesis_{\gamma,\eta,v} := \max_{\alpha<\gamma} \sup_{x\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1}} \frac{v(x)\|f^{(1)}(x)-f^{(2)}(x)\|_\alpha}{\varphi(x)^{(\eta-\alpha)\wedge 0}},$$

$$\|f^{(1)};f^{(2)}\|_{\gamma,\eta,v} := \max_{\alpha<\gamma} \sup_{\substack{x,y\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1},\, x\neq y\\ \|y-x\|_\mathfrak{s}\leq\varphi(x,y)}} \frac{v(x)\|\Delta^\Gamma_{yx}f^{(1)}-\Delta^\Gamma_{yx}f^{(2)}\|_\alpha}{\varphi(x,y)^{\eta-\gamma}\, v^*(x-y)\|y-x\|_\mathfrak{s}^{\gamma-\alpha}}.$$

In [52], the topologies of the space of models and the space of modelled distributions are defined by the family of pseudo-metrics parametrized by compact subsets K of $\mathbb{R}^d$, where x and y in the above definitions are restricted within K. In this paper, we employ weight functions w and v instead of such local bounds.

We consider the relations between $\mathcal{D}_v^{\gamma,\eta}$ under varying parameters γ,η, as well as the relation between $\mathcal{D}_v^{\gamma,\eta}$ and a variant. We say that the function $u : \mathbb{R}^d \to \mathbb{R}$ is *symmetric* if $u(-x) = u(x)$ for any $x \in \mathbb{R}^d$.

Proposition 3.3.5. Let $M = (\Pi,\Gamma) \in \mathscr{M}_w(\mathscr{T})$ and $\eta \leq \gamma$.

(i) For any $\theta \le \eta$, we have the continuous embedding $\mathcal{D}_v^{\gamma,\eta}(\Gamma) \hookrightarrow \mathcal{D}_v^{\gamma,\theta}(\Gamma)$.

(ii) Assume that w^* is symmetric. For each $\alpha \in \mathbb{R}$, we denote by $P_{<\alpha} : \mathbf{T} \to \mathbf{T}_{<\alpha}$ the canonical projection. For any $\eta \le \delta \le \gamma$, the map $P_{<\delta}$ extends to a continuous linear map $\mathcal{D}_v^{\gamma,\eta}(\Gamma) \to \mathcal{D}_{wv}^{\delta,\eta}(\Gamma)$. To be precise, we have the inequality

$$\|P_{<\delta} f\|_{\delta,\eta,wv} \lesssim \|\Gamma\|_{\gamma,w} (\!| f |\!)_{\gamma,\eta,v} + \|f\|_{\gamma,\eta,v}.$$

(iii) Instead of the norm $\|f\|_{\gamma,\eta,v}$, we define

$$\|f\|_{\gamma,\eta,v}^{\#} := \max_{\alpha<\gamma} \sup_{x,y\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1},\, x\neq y} \frac{v(x)\|\Delta_{yx}^{\Gamma} f\|_\alpha}{\varphi(x,y)^{\eta-\gamma}\, v^*(x-y)\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}}.$$

Then the inequality $\|f\|_{\gamma,\eta,v} \le \|f\|_{\gamma,\eta,v}^{\#}$ obviously holds. Conversely, if w^* is symmetric, then we also have

$$\|f\|_{\gamma,\eta\wedge\alpha_0,wv}^{\#} \lesssim (1 + \|\Gamma\|_{\gamma,w})(\!| f |\!)_{\gamma,\eta,v} + \|f\|_{\gamma,\eta,v}.$$

Proof. (i) The assertion immediately follows from the inequalities $\varphi(x)^{(\eta-\alpha)\wedge 0} \le \varphi(x)^{(\theta-\alpha)\wedge 0}$ and $\varphi(x,y)^{\eta-\gamma} \le \varphi(x,y)^{\theta-\gamma}$.

(ii) For any $x, y \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}^{d-1}$ such that $\|y-x\|_{\mathfrak{s}} \le \varphi(x,y)$ and any $\alpha < \delta$, we decompose

$$\begin{aligned}(wv)(x)\|\Delta_{yx}^{\Gamma} P_{<\delta} f\|_\alpha &\le v(x)\|\Delta_{yx}^{\Gamma} f\|_\alpha + (wv)(x) \sum_{\beta\in[\delta,\gamma)} \|\Gamma_{yx} P_\beta f(x)\|_\alpha \\ &=: A_1 + A_2.\end{aligned}$$

For A_1, by definition of the norm $\|f\|_{\gamma,\eta,v}$ we have

$$\begin{aligned}A_1 &\le \|f\|_{\gamma,\eta,v}\, v^*(x-y)\varphi(x,y)^{\eta-\gamma}\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha} \\ &\le \|f\|_{\gamma,\eta,v}\, v^*(x-y)\varphi(x,y)^{\eta-\delta}\|y-x\|_{\mathfrak{s}}^{\delta-\alpha}.\end{aligned}$$

For A_2, by definitions of the model and the norm $(\!| f |\!)_{\gamma,\eta,v}$ we have

$$\begin{aligned}A_2 &\le \sum_{\beta\in[\delta,\gamma)} w(x)\|\Gamma_{yx}\|_{\mathcal{L}(\mathbf{T}_\beta,\mathbf{T}_\alpha)} v(x)\|f(x)\|_\beta \\ &\le \|\Gamma\|_{\gamma,w}(\!| f |\!)_{\gamma,\eta,v}\, w^*(y-x) \sum_{\beta\in[\delta,\gamma)} \|y-x\|_{\mathfrak{s}}^{\beta-\alpha}\varphi(x)^{\eta-\beta} \\ &\le \|\Gamma\|_{\gamma,w}(\!| f |\!)_{\gamma,\eta,v}\, w^*(x-y)\|y-x\|_{\mathfrak{s}}^{\delta-\alpha}\varphi(x,y)^{\eta-\delta}.\end{aligned}$$

Thus we obtain the desired inequality for $\|P_{<\delta} f\|_{\delta,\eta,wv}$.

(iii) It is sufficient to show the estimate of $\Delta_{yx}^{\Gamma} f$ on the region $\|y-x\|_{\mathfrak{s}} > \varphi(x,y)$. For any $\alpha < \gamma$ we decompose

$$(wv)(x)\|\Delta_{yx}^{\Gamma} f\|_\alpha \le v(x)\|f(y)\|_\alpha + (wv)(x) \sum_{\beta\in[\alpha,\gamma)} \|\Gamma_{yx} P_\beta f(x)\|_\alpha =: B_1 + B_2.$$

For B_1, by definition of the norm $(\!| f |\!)_{\gamma,\eta,v}$ we have

$$\begin{aligned}
B_1 &\le v^*(x-y)v(y)\|f(y)\|_\alpha \\
&\le (\!| f |\!)_{\gamma,\eta,v}\, v^*(x-y)\varphi(y)^{(\eta-\alpha)\wedge 0} \\
&\le (\!| f |\!)_{\gamma,\eta,v}\, v^*(x-y)\varphi(x,y)^{(\eta-\alpha)\wedge 0} \\
&\le (\!| f |\!)_{\gamma,\eta,v}\, v^*(x-y)\varphi(x,y)^{\eta\wedge\alpha-\gamma}\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}.
\end{aligned}$$

For B_2, by an argument similar to A_2 in the proof of (ii), we have

$$\begin{aligned}
B_2 &\le \|\Gamma\|_{\gamma,w}(\!| f |\!)_{\gamma,\eta,v}\, w^*(y-x)\sum_{\beta\in[\alpha,\gamma)}\|y-x\|_{\mathfrak{s}}^{\beta-\alpha}\varphi(x)^{(\eta-\beta)\wedge 0} \\
&\lesssim \|\Gamma\|_{\gamma,w}(\!| f |\!)_{\gamma,\eta,v}\, w^*(x-y)\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}\varphi(x,y)^{\eta\wedge\alpha-\gamma}.
\end{aligned}$$

Thus we obtain the desired inequality. ◻

We also recall the definition of reconstruction.

Definition 3.3.6. Let $M=(\Pi,\Gamma)\in\mathscr{M}_w(\mathscr{T})$. For any $\eta\le\gamma$ and $f\in\mathcal{D}_v^{\gamma,\eta}(\Gamma)$, we say that $\Lambda\in C^{\zeta,Q}(wv)$ with some $\zeta\le 0$ is a *reconstruction* of f for M, if it satisfies

$$[\![\Lambda]\!]_{\gamma,\eta,wv} := \sup_{0<t\le 1}\sup_{x\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1}}\Big(t^{-\gamma/\ell}\varphi(x)^{\gamma-\eta}(wv)(x)|Q_t(x,\Lambda_x)|\Big)<\infty,$$

where $\Lambda_x := \Lambda - \Pi_x f(x)$. Furthermore, for any $M^{(i)}=(\Pi^{(i)},\Gamma^{(i)})\in\mathscr{M}_w(\mathscr{T})$, $f^{(i)}\in\mathcal{D}_v^{\gamma,\eta}(\Gamma^{(i)})$, and any reconstructions $\Lambda^{(i)}\in C^{\zeta,Q}(wv)$ of $f^{(i)}$ for $M^{(i)}$ with $i\in\{1,2\}$, we define

$$\begin{aligned}
&[\![\Lambda^{(1)};\Lambda^{(2)}]\!]_{\gamma,\eta,wv} \\
&:= \sup_{0<t\le 1}\sup_{x\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1}}\Big(t^{-\gamma/\ell}\varphi(x)^{\gamma-\eta}(wv)(x)\big|Q_t\big(x,\Lambda_x^{(1)}-\Lambda_x^{(2)}\big)\big|\Big),
\end{aligned}$$

where $\Lambda_x^{(i)} := \Lambda^{(i)} - \Pi_x^{(i)} f^{(i)}(x)$ for each $i\in\{1,2\}$.

3.3.3 Reconstruction Theorem

In this subsection, we provide a short proof of the reconstruction theorem. First, we prove the theorem for the subclass $\mathcal{D}_v^{\gamma,\eta}(\Gamma)^\#$ of $\mathcal{D}_v^{\gamma,\eta}(\Gamma)$ consisting of all functions $f:(\mathbb{R}\setminus\{0\})\times\mathbb{R}^{d-1}\to\mathbf{T}_{<\gamma}$ such that

$$|\!|\!| f |\!|\!|^\#_{\gamma,\eta,v} := (\!| f |\!)_{\gamma,\eta,v} + \|f\|^\#_{\gamma,\eta,v} < \infty.$$

In addition, for any $M^{(i)}=(\Pi^{(i)},\Gamma^{(i)})\in\mathscr{M}_w(\mathscr{T})$ and $f^{(i)}\in\mathcal{D}_v^{\gamma,\eta}(\Gamma^{(i)})^\#$ with $i\in\{1,2\}$, we define $|\!|\!| f^{(1)};f^{(2)} |\!|\!|^\#_{\gamma,\eta,v} := (\!| f^{(1)}-f^{(2)} |\!)_{\gamma,\eta,v} + \|f^{(1)};f^{(2)}\|^\#_{\gamma,\eta,v}$ similarly to definition 3.3.4.

Theorem 3.3.7. Let $\gamma > 0$ and $\eta \in (\gamma - \mathfrak{s}_1, \gamma]$. Then for any $M = (\Pi, \Gamma) \in \mathscr{M}_w(\mathscr{T})$ and $f \in \mathcal{D}_v^{\gamma,\eta}(\Gamma)^\#$, there exists a unique reconstruction $\mathcal{R}f \in C^{\zeta,Q}(wv)$ of f for M with $\zeta := \eta \wedge \alpha_0 \wedge 0$ and it holds that

$$\|\mathcal{R}f\|_{C^{\zeta,Q}(wv)} \lesssim \|\Pi\|_{\gamma,w} \|\!|f|\!\|^\#_{\gamma,\eta,v}, \tag{3.6}$$

$$[\![\mathcal{R}f]\!]_{\gamma,\eta,wv} \lesssim \|\Pi\|_{\gamma,w} \|f\|^\#_{\gamma,\eta,v}. \tag{3.7}$$

Moreover, there is an affine function $C_R > 0$ of $R > 0$ such that

$$\|\mathcal{R}f^{(1)} - \mathcal{R}f^{(2)}\|_{C^{\zeta,Q}(wv)} \le C_R\big(\|\Pi^{(1)} - \Pi^{(2)}\|_{\gamma,w} + \|\!|f^{(1)}; f^{(2)}|\!\|^\#_{\gamma,\eta,v}\big),$$

$$[\![\mathcal{R}f^{(1)}; \mathcal{R}f^{(2)}]\!]_{\gamma,\eta,wv} \le C_R\big(\|\Pi^{(1)} - \Pi^{(2)}\|_{\gamma,w} + \|f^{(1)}; f^{(2)}\|^\#_{\gamma,\eta,v}\big)$$

for any $M^{(i)} = (\Pi^{(i)}, \Gamma^{(i)}) \in \mathscr{M}_w(\mathscr{T})$ and $f^{(i)} \in \mathcal{D}_v^{\gamma,\eta}(\Gamma^{(i)})$ with $i \in \{1, 2\}$ such that $\|\!|M^{(i)}|\!\|_{\gamma,w} \le R$ and $\|\!|f^{(i)}|\!\|^\#_{\gamma,\eta,v} \le R$.

Proof. The proof is carried out by a method similar to that of [59, Theorem 4.1], but we have to treat the temporal weight more carefully. For $t > 0$ and $0 < s \le t \wedge 1$, we define the functions

$$\mathcal{R}_s^t f(x) := \begin{cases} \displaystyle\int_{\mathbb{R}^d} Q_{t-s}(x,y) Q_s\big(y, \Pi_y f(y)\big) dy, & s < t, \\ Q_t\big(x, \Pi_x f(x)\big), & s = t. \end{cases}$$

Note that

$$\begin{aligned} (wv)(x)\big|Q_t\big(x, \Pi_x f(x)\big)\big| &\le \sum_{\alpha<\gamma} w(x) \big\|Q_t\big(x, \Pi_x(\cdot)\big)\big\|_{\mathbf{T}_\alpha^*} v(x) \|f(x)\|_\alpha \\ &\le \|\Pi\|_{\gamma,w} (\!| f |\!)_{\gamma,\eta,v} \sum_{\alpha<\gamma} t^{\alpha/\ell} \varphi(x)^{(\eta-\alpha)\wedge 0}. \end{aligned}$$

Thus, by Proposition 3.2.8-(i), for any $s \in (0,t)$ we have $\mathcal{R}_s^t f \in C(wv)$ and

$$\|\mathcal{R}_s^t f\|_{L^\infty(wv)} \lesssim \|\Pi\|_{\gamma,w} (\!| f |\!)_{\gamma,\eta,v} \sum_{\alpha<\gamma} s^{\alpha/\ell} (t-s)^{(\eta-\alpha)\wedge 0}. \tag{3.8}$$

We separate the proof into four steps.

(1) Cauchy property. Set $F_x := \Pi_x f(x)$. By the definition of norms, we have

$$\begin{aligned} &(wv)(y)|Q_t(x, F_y - F_x)| \\ &= (wv)(y)\big|Q_t\big(x, \Pi_x\big\{\Gamma_{xy} f(y) - f(x)\big\}\big)\big| \\ &\le w^*(y-x) \sum_{\alpha<\gamma} w(x) \|Q_t(x, \Pi_x(\cdot))\|_{\mathbf{T}_\alpha^*} v(y) \|\Gamma_{xy} f(y) - f(x)\|_\alpha \\ &\le \|\Pi\|_{\gamma,w} \|f\|^\#_{\gamma,\eta,v} (w^* v^*)(y-x) \sum_{\alpha<\gamma} t^{\alpha/\ell} \varphi(x,y)^{\eta-\gamma} \|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}. \end{aligned} \tag{3.9}$$

By the semigroup property, for any $0 < u < s < t \wedge 1$ we have

$$
\begin{aligned}
&(wv)(x)|\mathcal{R}^t_s f(x) - \mathcal{R}^t_u f(x)| \\
&\le \int_{(\mathbb{R}^d)^2} (w^* v^*)(x-y)(wv)(y)|Q_{t-s}(x,y)Q_{s-u}(y,z)Q_u(z, F_y - F_z)|dydz \\
&\lesssim \|\Pi\|_{\gamma,w}\|f\|^{\#}_{\gamma,\eta,v} \sum_{\alpha<\gamma} u^{\alpha/\ell} \int_{(\mathbb{R}^d)^2} (w^* v^*)(x-y)(w^* v^*)(y-z) \\
&\qquad\qquad\qquad\qquad \times G_{t-s}(x-y)G_{s-u}(y-z)\varphi(y,z)^{\eta-\gamma}\|y-z\|_{\mathfrak{s}}^{\gamma-\alpha} dydz.
\end{aligned}
$$

By applying the second inequality of Lemma 3.2.6 to the integral with respect to z and then applying the first inequality of Lemma 3.2.6 to the integral with respect to y, we obtain

$$
\begin{aligned}
&(wv)(x)|\mathcal{R}^t_s f(x) - \mathcal{R}^t_u f(x)| \\
&\lesssim \|\Pi\|_{\gamma,w}\|f\|^{\#}_{\gamma,\eta,v} \\
&\qquad \times \sum_{\alpha<\gamma} u^{\alpha/\ell}(s-u)^{(\gamma-\alpha)/\ell} \int_{(\mathbb{R}^d)^2} (w^* v^*)(x-y)G_{t-s}(x-y)\varphi(y)^{\eta-\gamma} dy \\
&\lesssim \|\Pi\|_{\gamma,w}\|f\|^{\#}_{\gamma,\eta,v} \sum_{\alpha<\gamma} u^{\alpha/\ell}(s-u)^{(\gamma-\alpha)/\ell}\varphi(x)^{\eta-\gamma}.
\end{aligned}
$$

Consequently, when $u \in [s/2, s)$ we have the inequality

$$
(wv)(x)|\mathcal{R}^t_s f(x) - \mathcal{R}^t_u f(x)| \lesssim \|\Pi\|_{\gamma,w}\|f\|^{\#}_{\gamma,\eta,v}\, \varphi(x)^{\eta-\gamma} s^{\gamma/\ell}. \tag{3.10}
$$

Similarly to the proof of [59, Theorem 4.1], we can also extend it into $u \in (0, s/2)$ by decomposing

$$
|\mathcal{R}^t_s f(x) - \mathcal{R}^t_u f(x)| \le \sum_{n=0}^{\infty} |\mathcal{R}^t_{(s/2^n)\wedge u} f(x) - \mathcal{R}^t_{(s/2^{n+1})\wedge u} f(x)|.
$$

The same inequality for the case $s = t \le 1$ can be obtained by a similar argument. In the end, the inequality (3.10) holds for any $0 < u < s \le t \wedge 1$.

(2) Convergence as $s \downarrow 0$. Note that $Q_s \mathcal{R}^t_u f = \mathcal{R}^{t+s}_u f$ follows from the semigroup property. By the inequality (3.10), for any $0 < u < s \le t/2$ we have

$$
\begin{aligned}
&(wv)(x)|\mathcal{R}^t_s f(x) - \mathcal{R}^t_u f(x)| \\
&\le \int_{\mathbb{R}^d} (w^* v^*)(x-y)(wv)(y)|Q_{t/2}(x,y)(\mathcal{R}^{t/2}_s f - \mathcal{R}^{t/2}_u f)(y)|dy \\
&\lesssim \|\Pi\|_{\gamma,w}\|f\|^{\#}_{\gamma,\eta,v}\, s^{\gamma/\ell} \int_{\mathbb{R}^d} (w^* v^*)(x-y)G_{t/2}(x-y)\varphi(y)^{\eta-\gamma} dy \\
&\lesssim \|\Pi\|_{\gamma,w}\|f\|^{\#}_{\gamma,\eta,v}\, s^{\gamma/\ell} t^{(\eta-\gamma)/\ell}.
\end{aligned}
\tag{3.11}
$$

Since $\gamma > 0$, this implies that $\{\mathcal{R}^t_s f\}_{0<s\le t/2}$ is Cauchy in $C(wv)$ as $s \downarrow 0$. We denote its limit by

$$
\mathcal{R}^t_0 f := \lim_{s\downarrow 0} \mathcal{R}^t_s f.
$$

We also have $Q_s \mathcal{R}_0^t f = \mathcal{R}_0^{t+s} f$ by taking the limit $u \downarrow 0$ in $Q_s \mathcal{R}_u^t f = \mathcal{R}_u^{t+s} f$.

(3) Convergence as $t \downarrow 0$. Combining the Cauchy property (3.11) and the bound (3.8) with $s = t/2$, we have

$$\begin{aligned}\|\mathcal{R}_0^t f\|_{L^\infty(wv)} &\le \|\mathcal{R}_{t/2}^t f\|_{L^\infty(wv)} + \|\mathcal{R}_{t/2}^t f - \mathcal{R}_0^t f\|_{L^\infty(wv)} \\ &\lesssim \|\Pi\|_{\gamma,w} \|\!|f|\!\|_{\gamma,\eta,v}^{\#}\, t^{(\eta\wedge\alpha_0)/\ell}.\end{aligned}$$

Since $Q_s \mathcal{R}_0^t f = \mathcal{R}_0^{t+s} f$, this implies

$$\sup_{0<t\le 1} \|\mathcal{R}_0^t f\|_{C^{\eta\wedge\alpha_0\wedge 0,Q}(wv)} \lesssim \|\Pi\|_{\gamma,w} \|\!|f|\!\|_{\gamma,\eta,v}^{\#}.$$

From here onward, in exactly the same way as the part (4) of the proof of [59, Theorem 4.1], we can show the existence of $\mathcal{R}f \in C^{\zeta,Q}(wv)$ with $\zeta = \eta \wedge \alpha_0 \wedge 0$ which satisfies the bound (3.6) and

$$\lim_{t\downarrow 0} \|\mathcal{R}f - \mathcal{R}_0^t f\|_{C^{\zeta-\varepsilon,Q}(wv)} = 0$$

for any $\varepsilon \in (0, \ell]$. Moreover, we have $Q_t \mathcal{R}f = \mathcal{R}_0^t f$ by taking the limit $s \downarrow 0$ in $Q_t \mathcal{R}_0^s f = \mathcal{R}_0^{t+s} f$. We have another bound (3.7) by letting $u \downarrow 0$ and $s = t$ in the inequality (3.10).

(4) Uniqueness. Let $\Lambda, \Lambda' \in C^{\zeta,Q}(wv)$ be reconstructions of f for M. By the property of reconstruction, $g := \Lambda - \Lambda'$ satisfies

$$\sup_{x\in\mathbb{R}^d} \varphi(x)^{\gamma-\eta}(wv)(x)|Q_t g(x)| \lesssim t^{\gamma/\ell}.$$

Set $\tilde{w} := \varphi^{\gamma-\eta} wv$. By Proposition 3.2.8-(iv) and (v), for any $\varepsilon \in (0, \ell]$ we have

$$\begin{aligned}\|g\|_{C^{\zeta-\varepsilon,Q}(\tilde{w})} &\le \|(Q_t - \mathrm{id})g\|_{C^{\zeta-\varepsilon,Q}(\tilde{w})} + \|Q_t g\|_{C^{\zeta-\varepsilon,Q}(\tilde{w})} \\ &\lesssim t^{\varepsilon/\ell} \|g\|_{C^{\zeta,Q}(\tilde{w})} + \|Q_t g\|_{L^\infty(\tilde{w})} \\ &\lesssim t^{\varepsilon/\ell} \|g\|_{C^{\zeta,Q}(wv)} + t^{\gamma/\ell}.\end{aligned}$$

By taking the limit $t \downarrow 0$, we have $g = 0$ in $C^{\zeta-\varepsilon,Q}(\tilde{w})$. By Proposition 3.2.8-(iii) and (iv), we also have $g = 0$ in $C^{\zeta,Q}(wv)$. □

The following result is used in Section 3.5.

Proposition 3.3.8. In addition to the setting of Theorem 3.3.7, we assume that the model M is smooth in the sense of Definition 3.3.2 and

$$\sup_{x\in\mathbb{R}^d} \sup_{\tau\in\mathbf{T}_\alpha\setminus\{0\}} w(x) \frac{|(\Pi_x \tau)(x)|}{\|\tau\|_\alpha} < \infty$$

for any $\alpha \in \mathbf{A}$. Then the reconstruction $\mathcal{R}f$ of $f \in \mathcal{D}_v^{\gamma,\eta}(\Gamma)^{\#}$ is realized as a continuous function on $(\mathbb{R} \setminus \{0\}) \times \mathbb{R}^{d-1}$ such that

$$(\mathcal{R}f)(x) = \big(\Pi_x f(x)\big)(x)$$

for any $x \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}^{d-1}$.

Proof. Set $\Lambda(x) = \big(\Pi_x f(x)\big)(x)$. Since $(\Pi_x\tau)(x) = \lim_{t\downarrow 0} Q_t(x,\Pi_x\tau) = 0$ if $\tau \in \mathbf{T}_\alpha$ with $\alpha > 0$, we have

$$\begin{aligned}(wv)(x)|\Lambda(x)| &\le \sum_{\alpha\le 0} w(x)\big\|\big(\Pi_x(\cdot)\big)(x)\big\|_{\mathbf{T}_\alpha^*} v(x)\|f(x)\|_\alpha \\ &\lesssim \sum_{\alpha\le 0} \varphi(x)^{(\eta-\alpha)\wedge 0} \lesssim \varphi(x)^{\eta\wedge 0}.\end{aligned}$$

Since $\eta > -\mathfrak{s}_1$, we have $\Lambda \in C^{\eta\wedge 0,Q}(wv) \subset C^{\zeta,Q}(wv)$ by Corollary 3.2.7. Moreover, since

$$\begin{aligned}&(wv)(x)|Q_t(x,\Lambda_x)| = (wv)(x)\left|\int_{\mathbb{R}^d} Q_t(x,y)\Pi_y\big(f(y)-\Gamma_{yx}f(x)\big)(y)dy\right| \\ &\lesssim \sum_{\alpha\le 0}\int_{\mathbb{R}^d} w^*(x-y)G_t(x-y)w(y)\big\|\big(\Pi_y(\cdot)\big)(y)\big\|_{\mathbf{T}_\alpha^*} v(x)\|f(y)-\Gamma_{yx}f(x)\|_\alpha dy \\ &\lesssim \sum_{\alpha\le 0}\int_{\mathbb{R}^d} (w^*v^*)(x-y)G_t(x-y)\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}\varphi(x,y)^{\eta-\gamma}dy \\ &\lesssim \sum_{\alpha\le 0} t^{(\gamma-\alpha)/\ell}\varphi(x)^{\eta-\gamma} \lesssim t^{\gamma/\ell}\varphi(x)^{\eta-\gamma},\end{aligned}$$

we have $[\![\Lambda]\!]_{\gamma,\eta,wv} < \infty$. Hence $\mathcal{R}f = \Lambda$ by the uniqueness of the reconstruction. □

Combining Theorem 3.3.7 with Proposition 3.3.5-(iii), we have the following result.

Corollary 3.3.9. Assume that w^2v is also G-controlled. If $\gamma > 0$ and $\eta\wedge\alpha_0 \in (\gamma-\mathfrak{s}_1,\gamma]$, then for any $M = (\Pi,\Gamma) \in \mathscr{M}_w(\mathscr{T})$ and $f \in \mathcal{D}_v^{\gamma,\eta}(\Gamma)$, there exists a unique reconstruction $\mathcal{R}f \in C^{\eta\wedge\alpha_0\wedge 0,Q}(w^2v)$ of f for M and it holds that

$$\begin{aligned}\|\mathcal{R}f\|_{C^{\eta\wedge\alpha_0\wedge 0,Q}(w^2v)} &\lesssim \|\Pi\|_{\gamma,w}(1+\|\Gamma\|_{\gamma,w})|\!|\!|f|\!|\!|_{\gamma,\eta,v}, \\ [\![\mathcal{R}f]\!]_{\gamma,\eta\wedge\alpha_0,w^2v} &\lesssim \|\Pi\|_{\gamma,w}(1+\|\Gamma\|_{\gamma,w})\|f\|_{\gamma,\eta,v}.\end{aligned}$$

The local Lipschitz estimates similar to the latter part of Theorem 3.3.7 also hold.

3.4 Multilevel Schauder estimate

This section is devoted to the proof of the multilevel Schauder estimate for singular modelled distributions. After recalling from [59] the basics of regularizing kernels in the first subsection, we prove the multilevel Schauder estimate in the second subsection.

3.4.1 Regularizing kernels

We recall from [59, Section 5.1] the definition of regularizing kernels.

Definition 3.4.1. Let $\bar\beta > 0$. A *$\bar\beta$-regularizing (integral) kernel admissible for* $\{Q_t\}_{t>0}$ is a family of continuous functions $\{K_t : \mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}\}_{t>0}$ which satisfies the following properties for some constants $\delta > 0$ and $C_K > 0$.

(i) (Convolution with Q) For any $0 < s < t$ and $x, y \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} K_{t-s}(x,z)Q_s(z,y)dz = K_t(x,y).$$

(ii) (Upper estimate) For any $\mathbf{k} \in \mathbb{N}^d$ with $|\mathbf{k}|_{\mathfrak{s}} < \delta$, the $\mathbf{k}$-th partial derivative of $K_t(x,y)$ with respect to x exists, and we have for any $t > 0$ and $x, y \in \mathbb{R}^d$,

$$|\partial_x^{\mathbf{k}} K_t(x,y)| \le C_K t^{(\bar{\beta}-|\mathbf{k}|_{\mathfrak{s}})/\ell-1} G_t(x-y).$$

(iii) (Hölder continuity) For any $\mathbf{k} \in \mathbb{N}^d$ with $|\mathbf{k}|_{\mathfrak{s}} < \delta$, any $t > 0$ and $x, y, h \in \mathbb{R}^d$ with $\|h\|_{\mathfrak{s}} \le t^{1/\ell}$,

$$\begin{aligned} &\left|\partial_x^{\mathbf{k}} K_t(x+h,y) - \sum_{|\mathbf{l}|_{\mathfrak{s}} < \delta - |\mathbf{k}|_{\mathfrak{s}}} \frac{h^{\mathbf{l}}}{\mathbf{l}!} \partial_x^{\mathbf{k}+\mathbf{l}} K_t(x,y)\right| \\ &\le C_K \|h\|_{\mathfrak{s}}^{\delta-|\mathbf{k}|_{\mathfrak{s}}} t^{(\bar{\beta}-\delta)/\ell-1} G_t(x-y). \end{aligned}$$

We fix a $\bar{\beta}$-regularizing kernel $\{K_t\}_{t>0}$ throughout this section. For any $f \in L^\infty(w)$ with a G-controlled weight w and any $|\mathbf{k}|_{\mathfrak{s}} < \delta$, we define

$$(\partial^{\mathbf{k}} K_t f)(x) =: \partial^{\mathbf{k}} K_t(x,f) := \int_{\mathbb{R}^d} \partial_x^{\mathbf{k}} K_t(x,y) f(y) dy.$$

Moreover, we write $\partial^{\mathbf{k}} K f := \int_0^1 \partial^{\mathbf{k}} K_t f dt$ if the integral makes sense.

Lemma 3.4.2. Let w and v be G-controlled weights such that w^2 and wv are also G-controlled. Let $\mathscr{T} = (\mathbf{A}, \mathbf{T}, \mathbf{G})$ be a regularity structure and let $M = (\Pi, \Gamma) \in \mathscr{M}_w(\mathscr{T})$.

(i) [59, Lemma 5.4] For any $\alpha \le 0$, $|\mathbf{k}|_{\mathfrak{s}} < \delta$, and $f \in L^\infty(w)$, we have

$$\|\partial^{\mathbf{k}} K_t f\|_{L^\infty(w)} \lesssim C_K\, t^{(\alpha+\bar{\beta}-|\mathbf{k}|_{\mathfrak{s}})/\ell-1} \|f\|_{C^{\alpha,Q}(w)},$$

where the implicit proportional constant depends only on G and w. Consequently, if $|\mathbf{k}|_{\mathfrak{s}} < (\alpha + \bar{\beta}) \wedge \delta$, the integral $\partial^{\mathbf{k}} K f := \int_0^1 \partial^{\mathbf{k}} K_t f dt$ converges in $C(w)$.

(ii) [59, Lemma 5.6] For any $\alpha < \gamma$, $\tau \in \mathbf{T}_\alpha$, $|\mathbf{k}|_{\mathfrak{s}} < \delta$, and $t \in (0,1]$, we have

$$\|\partial^{\mathbf{k}} K_t(x, \Pi_x \tau)\|_{L_x^\infty(w^2)} \lesssim C_K\, t^{(\alpha+\bar{\beta}-|\mathbf{k}|_{\mathfrak{s}})/\ell-1} \|\Pi\|_{\gamma,w}(1 + \|\Gamma\|_{\gamma,w}) \|\tau\|_\alpha,$$

where the implicit proportional constant depends only on G, w, and $\mathbf{A}$. Consequently, if $|\mathbf{k}|_{\mathfrak{s}} < (\alpha + \bar{\beta}) \wedge \delta$, the integral $\partial^{\mathbf{k}} K(x, \Pi_x \tau) := \int_0^1 \partial^{\mathbf{k}} K_t(x, \Pi_x \tau) dt$ converges for any $x \in \mathbb{R}^d$.

(iii) Let $\gamma \in \mathbb{R}$, $\eta \in (\gamma - \mathfrak{s}_1, \gamma]$, and $\zeta \le 0$. For any $f \in \mathcal{D}_v^{\gamma,\eta}(\Gamma)^\#$ and its reconstruction $\Lambda \in C^{\zeta,Q}(wv)$, $|\mathbf{k}|_{\mathfrak{s}} < \delta$, and $t \in (0,1]$, we have

$$\begin{aligned} &(wv)(x) |\partial^{\mathbf{k}} K_t(x, \Lambda_x)| \\ &\lesssim C_K\, t^{(\gamma+\bar{\beta}-|\mathbf{k}|_{\mathfrak{s}})/\ell-1} \varphi(x)^{\eta-\gamma} \big([\![\Lambda]\!]_{\gamma,\eta,wv} + \|\Pi\|_{\gamma,w} \|f\|_{\gamma,\eta,v}^\# \big), \end{aligned}$$

where the implicit proportional constant depends only on G, w, v, and $\mathbf{A}$. Consequently, if $|\mathbf{k}|_{\mathfrak{s}} < (\gamma + \bar{\beta}) \wedge \delta$, the integral $\partial^{\mathbf{k}} K(x, \Lambda_x) := \int_0^1 \partial^{\mathbf{k}} K_t(x, \Lambda_x) dt$ converges for any $x \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}^{d-1}$.

Proof. We prove only (iii). By Definition 3.4.1-(i), we can decompose

$$|\partial^{\mathbf{k}} K_t(x, \Lambda_x)| \le \left| \int_{\mathbb{R}^d} \partial^{\mathbf{k}} K_{t/2}(x, y) Q_{t/2}(y, \Lambda_y) dy \right| + \left| \int_{\mathbb{R}^d} \partial^{\mathbf{k}} K_{t/2}(x, y) Q_{t/2}\big(y, \Pi_y f(y) - \Pi_x f(x)\big) dy \right|.$$

For the first term, by Definition 3.4.1-(ii) and by the property of reconstruction, we have

$$\begin{aligned}
&(wv)(x) \left| \int_{\mathbb{R}^d} \partial^{\mathbf{k}} K_{t/2}(x, y) Q_{t/2}(y, \Lambda_y) dy \right| \\
&\lesssim C_K\, t^{(\bar{\beta} - |\mathbf{k}|_{\mathfrak{s}})/\ell - 1} \int_{\mathbb{R}^d} (w^* v^*)(x - y) G_{t/2}(x - y) (wv)(y) |Q_{t/2}(y, \Lambda_y)| dy \\
&\lesssim C_K\, t^{(\gamma + \bar{\beta} - |\mathbf{k}|_{\mathfrak{s}})/\ell - 1} [\![\Lambda]\!]_{\gamma, \eta, wv} \int_{\mathbb{R}^d} \varphi(y)^{\eta - \gamma} (w^* v^*)(x - y) G_{t/2}(x - y) dy \\
&\lesssim C_K\, t^{(\gamma + \bar{\beta} - |\mathbf{k}|_{\mathfrak{s}})/\ell - 1}\, \varphi(x)^{\eta - \gamma}\, [\![\Lambda]\!]_{\gamma, \eta, wv}.
\end{aligned}$$

For the second term, by using the inequality (3.9) obtained in the proof of Theorem 3.3.7 with x and y swapped, we have

$$\begin{aligned}
&(wv)(x) \left| \int_{\mathbb{R}^d} \partial^{\mathbf{k}} K_{t/2}(x, y) Q_{t/2}\big(y, \Pi_y f(y) - \Pi_x f(x)\big) dy \right| \\
&\lesssim C_K\, t^{(\bar{\beta} - |\mathbf{k}|_{\mathfrak{s}})/\ell - 1} \int_{\mathbb{R}^d} G_{t/2}(x - y) (wv)(x) \big| Q_{t/2}\big(y, \Pi_y f(y) - \Pi_x f(x)\big) \big| dy \\
&\lesssim C_K \|\Pi\|_{\gamma, w} \|f\|^{\#}_{\gamma, \eta, v} \sum_{\alpha < \gamma} t^{(\alpha + \bar{\beta} - |\mathbf{k}|_{\mathfrak{s}})/\ell - 1} \int_{\mathbb{R}^d} \varphi(x, y)^{\eta - \gamma} \|y - x\|_{\mathfrak{s}}^{\gamma - \alpha} \\
&\qquad\qquad \times (w^* v^*)(x - y) G_{t/2}(x - y) dy \\
&\lesssim C_K\, t^{(\gamma + \bar{\beta} - |\mathbf{k}|_{\mathfrak{s}})/\ell - 1}\, \varphi(x)^{\eta - \gamma}\, \|\Pi\|_{\gamma, w} \|f\|^{\#}_{\gamma, \eta, v}.
\end{aligned}$$

□

3.4.2 Compatible models and multilevel Schauder estimate

We recall from [59, Section 5.2] the notions of abstract integrations and compatible models. Hereafter, we use the *polynomial structure* generated by dummy variables $X_1, \dots, X_d$ as in [52, Section 2].

Definition 3.4.3. Let $\bar{\mathscr{T}} = (\bar{\mathbf{A}}, \bar{\mathbf{T}}, \bar{\mathbf{G}})$ be a regularity structure satisfying the following properties.

(1) $\mathbb{N}[\mathfrak{s}] \subset \bar{\mathbf{A}}$.

(2) For each $\alpha \in \mathbb{N}[\mathfrak{s}]$, the space $\bar{\mathbf{T}}_\alpha$ contains all $X^{\mathbf{k}} := \prod_{i=1}^d X_i^{k_i}$ with $|\mathbf{k}|_\mathfrak{s} = \alpha$.

(3) The subspace $\mathrm{span}\{X^{\mathbf{k}}\}_{\mathbf{k}\in\mathbb{N}^d}$ of $\bar{\mathbf{T}}$ is closed under $\bar{\mathbf{G}}$-actions.

Let $\mathscr{T} = (\mathbf{A}, \mathbf{T}, \mathbf{G})$ be another regularity structure. A continuous linear operator $\mathcal{I} : \mathbf{T} \to \bar{\mathbf{T}}$ is called an *abstract integration* of order $\beta \in (0, \bar{\beta}]$ if

$$\mathcal{I} : \mathbf{T}_\alpha \to \bar{\mathbf{T}}_{\alpha+\beta}$$

for any $\alpha \in \mathbf{A}$. For a fixed G-controlled weight w, we say that the pair $(M, \bar{M})$ of two models $M = (\Pi, \Gamma) \in \mathscr{M}_w(\mathscr{T})$ and $\bar{M} = (\bar{\Pi}, \bar{\Gamma}) \in \mathscr{M}_w(\bar{\mathscr{T}})$ is *compatible for* $\mathcal{I}$ if it satisfies the following properties.

(i) For any $\mathbf{k} \in \mathbb{N}^d$,

$$(\bar{\Pi}_x X^{\mathbf{k}})(\cdot) = (\cdot - x)^{\mathbf{k}}, \qquad \bar{\Gamma}_{yx} X^{\mathbf{k}} = \sum_{\mathbf{l}\le\mathbf{k}} \binom{\mathbf{k}}{\mathbf{l}} (y-x)^{\mathbf{l}} X^{\mathbf{k}-\mathbf{l}}.$$

(ii) For each $x \in \mathbb{R}^d$, we define the linear map $\mathcal{J}(x) : \mathbf{T}_{<\delta-\beta} \to \mathrm{span}\{X^{\mathbf{k}}\}_{|\mathbf{k}|_\mathfrak{s}<\delta} \subset \bar{\mathbf{T}}$ by setting

$$\mathcal{J}(x)\tau = \sum_{|\mathbf{k}|_\mathfrak{s}<\alpha+\beta} \frac{X^{\mathbf{k}}}{\mathbf{k}!} \partial^{\mathbf{k}} K(x, \Pi_x \tau) \tag{3.12}$$

for any $\alpha \in \mathbf{A}$ such that $\alpha + \beta < \delta$ and $\tau \in \mathbf{T}_\alpha$. Then for any $\tau \in \mathbf{T}_{<\delta-\beta}$,

$$\bar{\Gamma}_{yx}\big(\mathcal{I} + \mathcal{J}(x)\big)\tau = \big(\mathcal{I} + \mathcal{J}(y)\big)\Gamma_{yx}\tau.$$

In addition, if the regularity α_0 of $\mathscr{T}$ is greater than $-\bar{\beta}$ and

$$(\bar{\Pi}_x \mathcal{I}\tau)(\cdot) = K(\cdot, \Pi_x\tau) - \sum_{|\mathbf{k}|_\mathfrak{s}<\alpha+\beta} \frac{(\cdot - x)^{\mathbf{k}}}{\mathbf{k}!} \partial^{\mathbf{k}} K(x, \Pi_x\tau) \tag{3.13}$$

for any $\tau \in \mathbf{T}_\alpha$ with $\alpha + \beta < \delta$, then we say that the pair $(M, \bar{M})$ is K*-admissible for* $\mathcal{I}$.

In (3.12) and (3.13), the function $K(\cdot, \Pi_x\tau)$ and the coefficients $\partial^{\mathbf{k}} K(x, \Pi_x\tau)$ are well-defined by Lemma 3.4.2. The following theorem is the second main result of this paper.

Theorem 3.4.4. Let $\mathscr{T}$ and $\bar{\mathscr{T}}$ be regularity structures satisfying the setting of Definition 3.4.3 and let $\mathcal{I} : \mathbf{T} \to \bar{\mathbf{T}}$ be an abstract integration of order $\beta \in (0, \bar{\beta}]$. Let w and v be G-controlled weights such that $w^2 v$ is also G-controlled. Given $(\Pi, \Gamma) \in \mathscr{M}_w(\mathscr{T})$, $f \in \mathcal{D}_v^{\gamma,\eta}(\Gamma)^\#$ with $\gamma + \bar{\beta} < \delta$ and $\eta \in (\gamma - \mathfrak{s}_1, \gamma]$, and its reconstruction $\Lambda \in C^{\zeta,Q}(wv)$, we define the functions

$$\mathcal{N}(x; f, \Lambda) = \sum_{|\mathbf{k}|_\mathfrak{s}<\gamma+\beta} \frac{X^{\mathbf{k}}}{\mathbf{k}!} \partial^{\mathbf{k}} K(x, \Lambda_x)$$

and

$$\mathcal{K} f(x) := \mathcal{I} f(x) + \mathcal{J}(x) f(x) + \mathcal{N}(x; f, \Lambda)$$

for $x \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}^{d-1}$. We assume $\zeta \le \eta \wedge \alpha_0$ and either of the following conditions.

(1) $\beta < \bar{\beta}$.

(2) $\beta = \bar{\beta}$ and $\{\alpha + \bar{\beta}\,;\, \alpha \in \mathbf{A} \cup \{\gamma, \zeta\}\} \cap \mathbb{N}[\mathfrak{s}] = \emptyset$.

Then for any compatible pair of models $\big(M = (\Pi, \Gamma), \bar{M} = (\bar{\Pi}, \bar{\Gamma})\big) \in \mathscr{M}_w(\mathscr{T}) \times \mathscr{M}_w(\bar{\mathscr{T}})$ and any singular modelled distribution $f \in \mathcal{D}_v^{\gamma,\eta}(\Gamma)^\#$, the function $\mathcal{K}f$ belongs to $\mathcal{D}_{w^2 v}^{\gamma+\beta,\zeta+\beta}(\bar{\Gamma})^\#$, and we have

$$(\!|\mathcal{K}f|\!)_{\gamma+\beta,\zeta+\beta,w^2v} \lesssim \|\mathcal{I}\| (\!|f|\!)_{\gamma,\eta,v} + C_K \big\{ \|\Pi\|_{\gamma,w}(1 + \|\Gamma\|_{\gamma,w}) |\!|\!| f |\!|\!|^\#_{\gamma,\eta,v} + \|\Lambda\|_{C^{\zeta,Q}(wv)} + [\![\Lambda]\!]_{\gamma,\eta,wv} \big\}, \tag{3.14}$$

$$\|\mathcal{K}f\|^\#_{\gamma+\beta,\zeta+\beta,w^2v} \lesssim \|\mathcal{I}\| \|f\|^\#_{\gamma,\eta,v} + C_K \big\{ \|\Pi\|_{\gamma,w}(1 + \|\Gamma\|_{\gamma,w}) \|f\|^\#_{\gamma,\eta,v} + [\![\Lambda]\!]_{\gamma,\eta,wv} \big\}, \tag{3.15}$$

where $\|\mathcal{I}\|$ is the operator norm from $\mathbf{T}_{<\gamma}$ to $\bar{\mathbf{T}}_{<\gamma+\beta}$, and the implicit proportional constant depends only on G, w, v, γ, η, and $\mathbf{A}$. Moreover, there is a quadratic function $C_R > 0$ of $R > 0$ such that

$$|\!|\!| \mathcal{K}f^{(1)}; \mathcal{K}f^{(2)} |\!|\!|^\#_{\gamma+\beta,\zeta+\beta,w^2v} \le C_R \Big(|\!|\!| M^{(1)}; M^{(2)} |\!|\!|_{\gamma,w} + |\!|\!| f^{(1)}; f^{(2)} |\!|\!|^\#_{\gamma,\eta,v} \Big),$$

for any $M^{(i)} = (\Pi^{(i)}, \Gamma^{(i)}) \in \mathscr{M}_w(\mathscr{T})$ and $\bar{M}^{(i)} = (\bar{\Pi}^{(i)}, \bar{\Gamma}^{(i)}) \in \mathscr{M}_w(\bar{\mathscr{T}})$ such that $(M^{(i)}, \bar{M}^{(i)})$ is compatible and any $f^{(i)} \in \mathcal{D}_v^{\gamma,\eta}(\Gamma^{(i)})$ with $i \in \{1,2\}$ such that $|\!|\!| M^{(i)} |\!|\!|_{\gamma,w} \le R$ and $|\!|\!| f^{(i)} |\!|\!|^\#_{\gamma,\eta,v} \le R$.

Proof. The proof is carried out by a method similar to that of [59, Theorem 5.12], but we have to prove (3.14) more carefully than [59]. For the $\mathcal{I}$ term, by the continuity of $\mathcal{I}$ we immediately have

$$v(x)\|\mathcal{I}f(x)\|_\alpha \le v(x)\|\mathcal{I}\|\|f(x)\|_{\alpha-\beta} \le \|\mathcal{I}\| (\!|f|\!)_{\gamma,\eta,v}\, \varphi(x)^{(\eta+\beta-\alpha)\wedge 0}$$

for any $\alpha < \gamma + \beta$. For the $\mathcal{J}$ and $\mathcal{N}$ terms, we decompose

$$\mathcal{J}(x)f(x) + \mathcal{N}(x, f; \Lambda) = \sum_{|\mathbf{k}|_\mathfrak{s} < \gamma+\beta} \frac{X^\mathbf{k}}{\mathbf{k}!} \mathcal{A}^\mathbf{k}(x),$$

where

$$\mathcal{A}^\mathbf{k}(x) = \sum_{\alpha \in [\alpha_0, \gamma),\, |\mathbf{k}|_\mathfrak{s} < \alpha+\beta} \partial^\mathbf{k} K\big(x, \Pi_x P_\alpha f(x)\big) + \partial^\mathbf{k} K(x, \Lambda_x).$$

We further define the decomposition $\mathcal{A}^\mathbf{k}(x) = \int_0^1 \mathcal{A}_t^\mathbf{k}(x)dt$ according to the integral form $K = \int_0^1 K_t dt$, where $\mathcal{A}_t^\mathbf{k}$ is defined in the same way as $\mathcal{A}^\mathbf{k}$ with K replaced by K_t. By using Lemma 3.4.2-(ii) for $\partial^\mathbf{k} K_t\big(x, \Pi_x P_\alpha f(x)\big)$ and (iii) for $\partial^\mathbf{k} K_t(x, \Lambda_x)$, we have

$$(w^2 v)(x)|\mathcal{A}_t^\mathbf{k}(x)| \lesssim L_1 \sum_{\alpha \in [\alpha_0, \gamma],\, |\mathbf{k}|_\mathfrak{s} < \alpha+\beta} \varphi(x)^{(\eta-\alpha)\wedge 0}\, t^{(\alpha+\bar{\beta}-|\mathbf{k}|_\mathfrak{s})/\ell - 1}$$

where $L_1 := C_K\big\{\|\Pi\|_{\gamma,w}(1+\|\Gamma\|_{\gamma,w})\|\!| f\|\!|^{\#}_{\gamma,\eta,v} + [\![\Lambda]\!]_{\gamma,\eta,wv}\big\}$. Since all powers of t above are greater than -1, we have

$$(w^2v)(x)\int_0^{\varphi(x)^\ell} |\mathcal{A}_t^{\mathbf{k}}(x)|dt \lesssim L_1 \sum_{\alpha\in[\alpha_0,\gamma],\,|\mathbf{k}|_{\mathfrak{s}}<\alpha+\beta} \varphi(x)^{\eta\wedge\alpha+\bar\beta-|\mathbf{k}|_{\mathfrak{s}}}$$
$$\lesssim L_1\,\varphi(x)^{(\eta\wedge\alpha_0+\beta-|\mathbf{k}|_{\mathfrak{s}})\wedge 0}.$$

For the integral over $\varphi(x)^\ell < t \le 1$, we use another decomposition

$$\mathcal{A}_t^{\mathbf{k}}(x) = \partial^{\mathbf{k}}K_t\big(x,\Lambda\big) - \sum_{\alpha\in[\alpha_0,\gamma),\,|\mathbf{k}|_{\mathfrak{s}}\ge\alpha+\beta} \partial^{\mathbf{k}}K_t\big(x,\Pi_x P_\alpha f(x)\big)$$

and consider the two terms in the right hand side separately. For the first term, by the assumption that $\Lambda\in C^{\zeta,Q}(wv)$ and by Lemma 3.4.2-(i), we have

$$(wv)(x)|\partial^{\mathbf{k}}K_t\big(x,\Lambda\big)| \lesssim C_K\|\Lambda\|_{C^{\zeta,Q}(wv)}\, t^{(\zeta+\bar\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1}.$$

If $\zeta+\bar\beta-|\mathbf{k}|_{\mathfrak{s}}\neq 0$, we have

$$\int_{\varphi(x)^\ell}^1 t^{(\zeta+\bar\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1}dt \lesssim \varphi(x)^{(\zeta+\bar\beta-|\mathbf{k}|_{\mathfrak{s}})\wedge 0} \lesssim \varphi(x)^{(\zeta+\beta-|\mathbf{k}|_{\mathfrak{s}})\wedge 0}.$$

Otherwise, since $\zeta+\beta-|\mathbf{k}|_{\mathfrak{s}} < \zeta+\bar\beta-|\mathbf{k}|_{\mathfrak{s}} = 0$ by assumption we have

$$\int_{\varphi(x)^\ell}^1 t^{(\zeta+\bar\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1}dt \lesssim \int_{\varphi(x)^\ell}^1 t^{(\zeta+\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1}dt \lesssim \varphi(x)^{\zeta+\beta-|\mathbf{k}|_{\mathfrak{s}}}.$$

In either case, we obtain the desired estimate. For the remaining term, by Lemma 3.4.2-(ii) we have

$$(w^2v)(x)\sum_{\alpha\in[\alpha_0,\gamma),\,|\mathbf{k}|_{\mathfrak{s}}\ge\alpha+\beta} \Big|\partial^{\mathbf{k}}K_t\big(x,\Pi_x P_\alpha f(x)\big)\Big|$$
$$\lesssim L_2 \sum_{\alpha\in[\alpha_0,\gamma),\,|\mathbf{k}|_{\mathfrak{s}}\ge\alpha+\beta} \varphi(x)^{(\eta-\alpha)\wedge 0}\, t^{(\alpha+\bar\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1},$$

where $L_2 := C_K\|\Pi\|_{\gamma,w}(1+\|\Gamma\|_{\gamma,w})(\!| f |\!)_{\gamma,\eta,v}$. For α such that $|\mathbf{k}|_{\mathfrak{s}} > \alpha+\beta$, we easily have

$$\varphi(x)^{(\eta-\alpha)\wedge 0}\int_{\varphi(x)^\ell}^1 t^{(\alpha+\bar\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1}dt \lesssim \varphi(x)^{(\eta-\alpha)\wedge 0}\int_{\varphi(x)^\ell}^1 t^{(\alpha+\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1}dt$$
$$\lesssim \varphi(x)^{\eta\wedge\alpha+\beta-|\mathbf{k}|_{\mathfrak{s}}}.$$

If there exists α such that $|\mathbf{k}|_{\mathfrak{s}} = \alpha+\beta$, then since $0 = \alpha+\beta-|\mathbf{k}|_{\mathfrak{s}} < \alpha+\bar\beta-|\mathbf{k}|_{\mathfrak{s}}$ by assumption, we have

$$\varphi(x)^{(\eta-\alpha)\wedge 0}\int_{\varphi(x)^\ell}^1 t^{(\alpha+\bar\beta-|\mathbf{k}|_{\mathfrak{s}})/\ell-1}dt \lesssim \varphi(x)^{(\eta-\alpha)\wedge 0} = \varphi(x)^{\eta\wedge\alpha+\beta-|\mathbf{k}|_{\mathfrak{s}}}.$$

Consequently, we obtain

$$(w^2v)(x)\int_{\varphi(x)^\ell}^1 |\mathcal{A}_t^{\mathbf{k}}(x)|dt \lesssim \{C_K\|\Lambda\|_{C^{\zeta,Q}(wv)} + L_2\}\varphi(x)^{(\zeta+\beta-|\mathbf{k}|_{\mathfrak{s}})\wedge 0}.$$

The proof of (3.15) is completely the same as that of [59, Theorem 5.12] except the existence of the factor $\varphi(x,y)^{\eta-\gamma}$. □

The following theorem is obtained similarly to [59, Theorem 5.13], so we omit the proof.

Theorem 3.4.5. In addition to the setting of Theorem 3.4.4, we assume that $\zeta+\bar{\beta}>0$ and that $(M,\bar{M})$ is K-admissible for $\mathcal{I}$. Then $K\Lambda\in C(wv)$ is a reconstruction of $\mathcal{K}f\in\mathcal{D}_{w^2v}^{\gamma+\beta,\zeta+\beta}(\bar{\Gamma})^\#$ and

$$[\![K\Lambda]\!]_{\gamma+\beta,\zeta+\beta,w^2v} \lesssim C_K\big([\![\Lambda]\!]_{\gamma,\eta,wv} + \|\Pi\|_{\gamma,w}\|f\|_{\gamma,\eta,v}^\#\big).$$

A similar local Lipschitz estimate to the latter part of Theorem 3.4.4 also holds.

Combining Theorem 3.4.4 with Proposition 3.3.5-(iii), we have the following result.

Corollary 3.4.6. In addition to the setting of Theorem 3.4.4, assume that w^3v is G-controlled and that $\alpha_0>\gamma-\mathfrak{s}_1$. Then for any compatible pair of models $\big(M=(\Pi,\Gamma),\bar{M}=(\bar{\Pi},\bar{\Gamma})\big)\in\mathscr{M}_w(\mathscr{T})\times\mathscr{M}_w(\bar{\mathscr{T}})$ and any singular modelled distribution $f\in\mathcal{D}_v^{\gamma,\eta}(\Gamma)$, the function $\mathcal{K}f$ belongs to $\mathcal{D}_{w^3v}^{\gamma+\beta,\zeta+\beta}(\bar{\Gamma})$, and we have

$$\begin{aligned}
(\!|\mathcal{K}f|\!)_{\gamma+\beta,\zeta+\beta,w^3v} &\lesssim \|\mathcal{I}\|(\!|f|\!)_{\gamma,\eta,v} + C_K\big\{\|\Pi\|_{\gamma,w}(1+\|\Gamma\|_{\gamma,w})^2|\!|\!|f|\!|\!|_{\gamma,\eta,v} \\
&\qquad + \|\Lambda\|_{C^{\zeta,Q}(wv)} + [\![\Lambda]\!]_{\gamma,\eta,wv}\big\},\\
\|\mathcal{K}f\|_{\gamma+\beta,\zeta+\beta,w^3v} &\lesssim \|\mathcal{I}\|\big\{\|\Gamma\|_{\gamma,w}(\!|f|\!)_{\gamma,\eta,v} + \|f\|_{\gamma,\eta,v}\big\}\\
&\qquad + C_K\big\{\|\Pi\|_{\gamma,w}(1+\|\Gamma\|_{\gamma,w})^2\|f\|_{\gamma,\eta,v} + [\![\Lambda]\!]_{\gamma,\eta,wv}\big\}.
\end{aligned}$$

A similar local Lipschitz estimate to the latter part of Theorem 3.4.4 also holds.

3.5 Parabolic Anderson model

In this section, we study the parabolic Anderson model (PAM)

$$\big(\partial_1 - a(x')\Delta + c\big)u(x) = b\big(u(x)\big)\xi(x') \qquad (x\in(0,\infty)\times\mathbb{T}^2) \tag{3.16}$$

with a spatial white noise ξ defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Recall that x_1 in $x=(x_1,x_2,x_3)$ denotes the temporal variable and $x'=(x_2,x_3)$ denotes the spatial variables. Throughout this section, we fix the function $b:\mathbb{R}\to\mathbb{R}$ in the class C_b^3, and the function $a:\mathbb{T}^2\to\mathbb{R}$ which is α-Hölder continuous for some $\alpha\in(0,1)$ and satisfies

$$C_1\le a(x')\le C_2 \qquad (x'\in\mathbb{T}^2)$$

for some constants $0<C_1<C_2$. The constant $c>0$ in the left hand side of (3.16) is fixed later (see Proposition 3.5.1 and 3.5.2). We prove the renormalizability of (3.16) in Section 3.5.6. We fix $\alpha\in(0,1)$, $d=3$, $\mathfrak{s}=(2,1,1)$, and $\ell=4$ throughout this section.

3.5.1 Preliminaries

We denote by $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, and $e_3 = (0,0,1)$ the canonical basis vectors of $\mathbb{R}^3$. We define $C_b(\mathbb{R}\times\mathbb{T}^2)$ as the set of all bounded continuous functions $f : \mathbb{R}^3 \to \mathbb{R}$ such that

$$f(x+e_i) = f(x)$$

for any $x \in \mathbb{R}^3$ and $i \in \{2,3\}$. For any $\beta > 0$, we define $C_{\mathfrak{s}}^{\beta}(\mathbb{R}\times\mathbb{T}^2)$ as the set of all elements $f \in C_b(\mathbb{R}\times\mathbb{T}^2)$ such that $\partial_x^k f \in C_b(\mathbb{R}\times\mathbb{T}^2)$ for any $|k|_{\mathfrak{s}} < \beta$, and if $|k|_{\mathfrak{s}} < \beta \le |k|_{\mathfrak{s}} + \mathfrak{s}_i$, we have

$$|\partial^k f(x+he_i) - \partial^k f(x)| \lesssim |h|^{(\beta-|k|_{\mathfrak{s}})/\mathfrak{s}_i}$$

for any $x \in \mathbb{R}^3$ and $h \in \mathbb{R}$.

We denote by $P_{x_1}(x', y')$ the fundamental solution of the parabolic operator $\partial_1 - a\Delta + c$. Moreover, we introduce the anisotropic elliptic operator

$$\mathcal{L} := \big(\partial_1 - a(x')\Delta\big)(\partial_1 + \Delta)$$

on $\mathbb{R}^3$ and denote by $Q_t(x,y)$ the fundamental solution of $\partial_t - \mathcal{L} + c$ with an additional variable $t > 0$. We recall from [6, Appendix A] some properties of $P_{x_1}(x', y')$ and $Q_t(x,y)$.

Proposition 3.5.1 ([6, Theorem 57])**.** For any $C > 0$, we define the function $G^{(C)}$ on $\mathbb{R}^3$ by

$$G^{(C)}(x) = \exp\big\{-C\big(|x_1|^2 + |x_2|^{4/3} + |x_3|^{4/3}\big)\big\}.$$

For sufficiently large $c > 0$, $\{Q_t\}_{t>0}$ is a $G^{(C)}$-type semigroup for some constant $C > 0$, in the sense of Definition 3.2.3.

In what follows, we fix $C > 0$ and write $G = G^{(C)}$. For any G-controlled weight w and any $\zeta \le 0$, we can define the Besov space $C^{\zeta,Q}(w)$ in the sense of Definition 3.2.4. We denote by $C^{\zeta,Q}(\mathbb{R}\times\mathbb{T}^2)$ the closure of $C_b(\mathbb{R}\times\mathbb{T}^2)$ in the space $C^{\zeta,Q}(1)$ with the flat weight $w = 1$.

Proposition 3.5.2. For sufficiently large $c > 0$, we have the following.

(i) [6, Theorems 61 and Proposition 62] Let $\beta \in (0,\alpha)$. For any $g \in C_{\mathfrak{s}}^{\beta}(\mathbb{R}\times\mathbb{T}^2)$, we can define the function on $\mathbb{R}\times\mathbb{T}^2$ by

$$\big((\partial_1 - a\Delta + c)^{-1} g\big)(x) := \int_{(-\infty,x_1]\times\mathbb{R}^2} P_{x_1-y_1}(x', y') g(y) dy.$$

Then $h = (\partial_1 - a\Delta + c)^{-1} g$ is the unique solution of $(\partial_1 - a\Delta + c)h = g$ such that $h \in C_{\mathfrak{s}}^{\beta+2}(\mathbb{R}\times\mathbb{T}^2)$ and $\lim_{x_1\to-\infty} h(x) = 0$.

(ii) [6, Theorem 63] The operator $c - \mathcal{L}$ has an inverse of the form

$$(c-\mathcal{L})^{-1} f = \int_0^\infty Q_t f\, dt = \int_0^1 Q_t f\, dt + Q_1 (c-\mathcal{L})^{-1} f.$$

For any $\zeta \in (-4,0)\setminus\mathbb{Z}$, the map $(c-\mathcal{L})^{-1}$ uniquely extends to a continuous operator from $C^{\zeta,Q}(\mathbb{R}\times\mathbb{T}^2)$ to $C_{\mathfrak{s}}^{\zeta+4}(\mathbb{R}\times\mathbb{T}^2)$.

(iii) [6, Theorem 6] We can decompose $(\partial_1 - a\Delta + c)^{-1} = K + S$, where

$$K :=: \int_0^1 K_t\, dt := -\int_0^1 (\partial_1 + \Delta) Q_t\, dt$$

and

$$S := K_1(c - \mathcal{L})^{-1} + c(\partial_1 - a\Delta + c)^{-1}(1 + \partial_1 + \Delta)(c - \mathcal{L})^{-1}.$$

Then $\{K_t\}_{t>0}$ is a 2-regularizing kernel admissible for $\{Q_t\}_{t>0}$ in the sense of Definition 3.4.1, where $\delta \in (2, 2 + \alpha)$ in the condition (iii). Moreover, for any $\zeta \in (-2, 0) \setminus \{-1\}$ and $\varepsilon > 0$, S is continuous from $C^{\zeta,Q}(\mathbb{R} \times \mathbb{T}^2)$ to $C_{\mathfrak{s}}^{\alpha\wedge(\zeta+2)+2-\varepsilon}(\mathbb{R} \times \mathbb{T}^2)$.

Remark 3.5.3. One needs to pick a constant $c > 0$ large enough to construct the inverse operator $(c - \mathcal{L})^{-1}$, see the proof of Theorem 63 in [6]. However, in the equation (3.16), c can be an arbitrary constant. This is because we can replace the c on the left-hand side with a larger constant c' by adding a linear correction term $(c' - c)u(x)$ to the right-hand side. This correction term has no serious influences on the discussion in this section.

3.5.2 Regularity structure associated with PAM

Following [52], we prepare the regularity structure associated with PAM (3.16).

Definition 3.5.4. For any fixed $\varepsilon \in (0, 1/2)$, we define the regularity structure $\mathscr{T} = (\mathbf{A}, \mathbf{T}, \mathbf{G})$ of regularity $\alpha_0 := -1 - \varepsilon$ as follows.

(1) (Index set) $\mathbf{A} = \{-1 - \varepsilon,\ -2\varepsilon,\ -\varepsilon,\ 0,\ 1 - \varepsilon,\ 1,\ 2 - 2\varepsilon,\ 2 - \varepsilon\}$.

(2) (Model space) $\mathbf{T}$ is an eleven dimensional linear space spanned by the symbols

$$\Xi,\ \mathcal{I}(\Xi)\Xi,\ X_2\Xi,\ X_3\Xi,\ \mathbf{1},\ \mathcal{I}(\Xi),\ X_2,\ X_3,\ \mathcal{I}(\mathcal{I}(\Xi)\Xi),\ \mathcal{I}(X_2\Xi),\ \mathcal{I}(X_3\Xi).$$

The direct sum decomposition $\mathbf{T} = \bigoplus_{\alpha\in\mathbf{A}} \mathbf{T}_\alpha$ is given by

$$\begin{aligned} \mathbf{T}_{-1-\varepsilon} &= \mathrm{span}\{\Xi\}, & \mathbf{T}_{-2\varepsilon} &= \mathrm{span}\{\mathcal{I}(\Xi)\Xi\}, \\ \mathbf{T}_{-\varepsilon} &= \mathrm{span}\{X_i\Xi\}_{i\in\{2,3\}}, & \mathbf{T}_0 &= \mathrm{span}\{\mathbf{1}\}, \\ \mathbf{T}_{1-\varepsilon} &= \mathrm{span}\{\mathcal{I}(\Xi)\}, & \mathbf{T}_1 &= \mathrm{span}\{X_i\}_{i\in\{2,3\}}, \\ \mathbf{T}_{2-2\varepsilon} &= \mathrm{span}\{\mathcal{I}(\mathcal{I}(\Xi)\Xi)\}, & \mathbf{T}_{2-\varepsilon} &= \mathrm{span}\{\mathcal{I}(X_i\Xi)\}_{i\in\{2,3\}}. \end{aligned}$$

(3) (Structure group) $\mathbf{G}$ is a group of continuous linear operators on $\mathbf{T}$ such that, for any $\Gamma \in \mathbf{G}$ and $\alpha \in \mathbf{A}$,

$$(\Gamma - \mathrm{id})\mathbf{T}_\alpha \subset \mathbf{T}_{<\alpha}.$$

Although the above structure group is a more generic one copied from Definition 3.3.1 than the more particular one defined in [52, Section 8], we use the above definition to avoid preparing algebraic matters such as Hopf algebras and comodules. The above one is sufficient for the discussion in this section. The admissible model defined later is also realized in the particular structure group defined in [52, Section 8]. In what follows, let $\mathscr{T}$ be the regularity structure given in Definition 3.5.4 with fixed ε.

We consider the models and modelled distributions as in Section 3.3 with slight modifications. For any $r \geq 0$, we define the weight function

$$v_r(x) = e^{-r|x_1|}.$$

It is easy to see that v_r satisfies the inequality (3.1) with $v_r^*(x) := e^{r|x_1|}$ and v_r is G-controlled. Moreover, v_r satisfies the assumption of Remark 3.3.3 with $w_1(x) = e^{-2r\|x\|}$ and $w_2(x) = e^{-3r\|x\|}$, where $\|x\| := \sum_{i=1}^3 |x_i|$.

Definition 3.5.5. We say that a smooth model $M \in \mathscr{M}_{v_r}(\mathscr{T})$ (defined on $\mathbb{R}^3$) is *admissible* if it satisfies the following properties.

(i) For any $x, y \in \mathbb{R}^3$ and $i \in \{2,3\}$, we have

$$\big(\Pi_{x+e_i}(\cdot)\big)(y+e_i) = \big(\Pi_x(\cdot)\big)(y), \qquad \Gamma_{(y+e_i)(x+e_i)} = \Gamma_{yx}.$$

(ii) We write $\Pi\Xi = \Pi_x\Xi$ since it is independent of x. For any $x \in \mathbb{R}^3$, we have

$$\Pi_x\mathbf{1} = 1, \qquad \Pi_x X_i = (\cdot)_i - x_i, \qquad \Pi_x\mathcal{I}(\Xi) = K(\cdot,\Pi\Xi) - K(x,\Pi\Xi),$$

and

$$\Pi_x\mathcal{I}(\tau\Xi) = K(\cdot,\Pi_x\tau\Xi) - K(x,\Pi_x\tau\Xi) - \sum_{i\in\{2,3\}} ((\cdot)_i - x_i)\partial_i K(x,\Pi_x\tau\Xi),$$

where $\tau \in \{\mathcal{I}(\Xi), X_2, X_3\}$.

(iii) For any $x, y \in \mathbb{R}^3$, we have

$$\Gamma_{yx}\mathbf{1} = \mathbf{1}, \qquad \Gamma_{yx}X_i = X_i + (y_i - x_i)\mathbf{1},$$
$$\Gamma_{yx}\Xi = \Xi, \qquad \Gamma_{yx}\mathcal{I}(\Xi) = \mathcal{I}(\Xi) + \big(K(y,\Pi\Xi) - K(x,\Pi\Xi)\big)\mathbf{1},$$

and

$$\begin{aligned}
\Gamma_{yx}(\tau\Xi) &= \tau\Xi + (\Pi_x\tau)(y)\Xi,\\
\Gamma_{yx}\mathcal{I}(\tau\Xi) &= \mathcal{I}(\tau\Xi) + (\Pi_x\tau)(y)\mathcal{I}(\Xi)\\
&\quad + \Big(K(y,\Pi_x\tau\Xi) - K(x,\Pi_x\tau\Xi) - \sum_{i\in\{2,3\}} (y_i - x_i)\partial_i K(x,\Pi_x\tau\Xi)\Big)\mathbf{1}\\
&\quad + \sum_{i\in\{2,3\}} \big(\partial_i K(y,\Pi_y\tau\Xi) - \partial_i K(x,\Pi_x\tau\Xi)\big)X_i,
\end{aligned}$$

where $\tau \in \{\mathcal{I}(\Xi), X_2, X_3\}$.

(iv) For any $\tau \in \{\Xi, \mathcal{I}(\Xi)\Xi, X_2\Xi, X_3\Xi, \mathbf{1}\}$, we have

$$\sup_{x\in\mathbb{R}^d} v_r(x)|(\Pi_x\tau)(x)| < \infty.$$

We define the closed subspace $\mathscr{M}_r^{\mathrm{ad}}(\mathscr{T})$ of $\mathscr{M}_{v_r}(\mathscr{T})$ as the completion of the set of smooth admissible models.

By definition, the subspace

$$\mathbf{S} := \mathrm{span}\{\mathbf{1}, \mathcal{I}(\Xi), X_2, X_3, \mathcal{I}(\mathcal{I}(\Xi)\Xi), \mathcal{I}(X_2\Xi), \mathcal{I}(X_3\Xi)\}$$

is invariant under the action of admissible models. In the sense of Definition 3.4.3, the linear operator $\mathcal{I} : \mathbf{T} \to \mathbf{S}$ defined by

$$\mathcal{I}\tau = \begin{cases} \mathcal{I}\tau & (\tau \in \{\Xi, \mathcal{I}(\Xi), X_2\Xi, X_3\Xi\}) \\ 0 & (\tau \in \{\mathbf{1}, \mathcal{I}(\Xi), X_2, X_3, \mathcal{I}(\mathcal{I}(\Xi)\Xi), \mathcal{I}(X_2\Xi), \mathcal{I}(X_3\Xi)\}) \end{cases}$$

is an abstract integration of order 2, and for any $M \in \mathscr{M}_r^{\mathrm{ad}}(\mathscr{T})$, the pair (M, M) is K-admissible for $\mathcal{I}$. Therefore, we can define the operator $\mathcal{K}$ by Corollary 3.4.6.

The weight function v_r is used only to ensure the global bound of the model M defined from the white noise. For the definition of singular modelled distributions, the flat weight $v_0 = 1$ is sufficient since we study the local-in-time solution theory of (3.16).

Definition 3.5.6. For any interval $I \subset \mathbb{R}$ and any $\eta \le \gamma$, we define $\mathcal{D}^{\gamma,\eta}(I;\Gamma)$ as the space of all functions $f : (I \setminus \{0\}) \times \mathbb{T}^2 \to \mathbf{T}_{<\gamma}$ such that

$$⦅ f ⦆_{\gamma,\eta;I} := \max_{\alpha<\gamma} \sup_{x\in(I\setminus\{0\})\times\mathbb{T}^2} \frac{\|f(x)\|_\alpha}{\varphi(x)^{(\eta-\alpha)\wedge 0}} < \infty,$$

$$\|f\|_{\gamma,\eta;I} := \max_{\alpha<\gamma} \sup_{\substack{x,y\in(I\setminus\{0\})\times\mathbb{T}^2,\, x\neq y \\ \|y-x\|_{\mathfrak{s}}\le\varphi(x,y)}} \frac{\|\Delta_{yx}^\Gamma f\|_\alpha}{\varphi(x,y)^{\eta-\gamma}\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}} < \infty.$$

We denote by $\mathcal{D}^{\gamma,\eta}(I, \mathbf{S}; \Gamma)$ the subspace of $\mathbf{S}$-valued functions in the class $\mathcal{D}^{\gamma,\eta}(I;\Gamma)$.

3.5.3 Convolution operators

We can rewrite the equation (3.16) in the form

$$u(x) = \int_{\mathbb{R}^2} P_{x_1}(x', y')u_0(y')dy' + (\partial_1 - a\Delta + c)^{-1}\big\{\mathbf{1}_{(0,\infty)\times\mathbb{R}^2} b(u)\xi\big\}(x), \tag{3.17}$$

where u_0 is the initial value of u at $x_1 = 0$. In this subsection, we prepare some operators to reformulate the equation (3.17) at the level of singular modelled distributions.

First, the function $Pu_0(x) := \int_{\mathbb{R}^2} P_{x_1}(x', y')u_0(y')dy'$ can be lifted to the singular modelled distribution taking values in the polynomial structure. For any sufficiently regular function f on $(\mathbb{R} \setminus \{0\}) \times \mathbb{R}^2$, we define the $\mathbf{T}$-valued function

$$Lf(x) := f(x)\mathbf{1} + (\partial_2 f)(x)X_2 + (\partial_3 f)(x)X_3 \qquad (x \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}^2).$$

Lemma 3.5.7 ([6, Lemma 29])**.** Let $\theta \in (0, 1)$ and $u_0 \in C^\theta(\mathbb{T}^2)$. Then the lift $L(Pu_0)$ of the function $\mathbf{1}_{x_1>0}Pu_0(x)$ is in the class $\mathcal{D}^{\gamma,\theta}$ for any $\gamma \in (0, 2)$ and we have

$$\|L(Pu_0)\|_{\gamma,\theta;(0,t)} \lesssim \|u_0\|_{C^\theta(\mathbb{T}^2)}$$

for any $t > 0$.

Next, to lift the second term on the right hand side of (3.17), we prepare two lemmas. The first one is used to "extend" the domain of singular modelled distributions from $(0,t)\times\mathbb{T}^2$ to $\mathbb{R}\times\mathbb{T}^2$.

Lemma 3.5.8. We fix a smooth non-increasing function $\chi:(0,\infty)\to[0,1]$ such that

$$\chi(t)=\begin{cases}1 & (0<t\le 1),\\ 0 & (t\ge 2).\end{cases}$$

For each $t>0$, we define the function $\chi_t:\mathbb{R}^3\to\mathbb{R}$ by setting $\chi_t(x)=\mathbf{1}_{x_1>0}\chi(x_1/t)$. Let $M=(\Pi,\Gamma)\in\mathscr{M}_r^{\mathrm{ad}}(\mathscr{T})$ with some $r>0$ and let $\gamma\in(0,1-2\varepsilon)$ and $\eta\le\gamma$. For any $t\in(0,1]$ and any $f\in\mathcal{D}^{\gamma,\eta}((0,2t);\Gamma)$, we define the function

$$(E_tf)(x)=P_{<\gamma}\big((L\chi_t)(x)\cdot f(x)\big),$$

where the (partial) product $(\cdot)$ on $\mathbf{T}$ is defined by

$$\mathbf{1}\cdot\tau=\tau\quad(\tau\in\{\Xi,\mathcal{I}(\Xi)\Xi,X_2\Xi,X_3\Xi,\mathbf{1}\}),\qquad X_i\cdot\Xi=X_i\Xi\quad(i\in\{2,3\}).$$

(Other products do not appear due to the assumption on γ.) Then the function E_tf belongs to $\mathcal{D}^{\gamma,\eta\wedge\alpha_0}(\mathbb{R};\Gamma)$ and satisfies

$$|\!|\!| E_tf|\!|\!|_{\gamma,\eta\wedge\alpha_0;\mathbb{R}}\le C(1+\|\Gamma\|_{\gamma,v_r})|\!|\!| f|\!|\!|_{\gamma,\eta;(0,2t)}$$

for some constant $C>0$ independent of t. Moreover, $(E_tf)|_{(0,t]\times\mathbb{T}^2}=f|_{(0,t]\times\mathbb{T}^2}$.

Proof. We can check that $|\!|\!| L\chi_t|\!|\!|_{\gamma',0;\mathbb{R}}\lesssim 1$ for any $\gamma'\in(1,2)$ by definition, so by applying the continuity of the multiplication of modelled distributions [52, Proposition 6.12], we have

$$|\!|\!| E_tf|\!|\!|_{\gamma,\eta\wedge\alpha_0;(0,2t)}\lesssim|\!|\!| f|\!|\!|_{\gamma,\eta;(0,2t)}.$$

We can extend it into $|\!|\!| E_tf|\!|\!|_{\gamma,\eta\wedge\alpha_0;(0,2t]}\lesssim|\!|\!| f|\!|\!|_{\gamma,\eta;(0,2t)}$ by the uniform continuity. To show that $E_tf\in\mathcal{D}^{\gamma,\eta\wedge\alpha_0}((0,\infty);\mathbb{R})$, we pick $x\in[2t,\infty)\times\mathbb{T}^2$ and $y\in(0,2t)\times\mathbb{T}^2$. By setting $z=(2t,y')$ we have

$$\begin{aligned}&\|(E_tf)(y)-\Gamma_{yx}(E_tf)(x)\|_\alpha\\ &\le\|(E_tf)(y)-\Gamma_{yz}(E_tf)(z)\|_\alpha+\|\Gamma_{yz}(E_tf)(z)-\Gamma_{yx}(E_tf)(x)\|_\alpha\\ &\le\|E_tf\|_{\gamma,\eta\wedge\alpha_0;(0,2t]}\,\varphi(y)^{\eta\wedge\alpha_0-\gamma}\|y-z\|_{\mathfrak{s}}^{\gamma-\alpha}\\ &\lesssim|\!|\!| f|\!|\!|_{\gamma,\eta;(0,2t)}\,\varphi(x,y)^{\eta\wedge\alpha_0-\gamma}\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}.\end{aligned}$$

In the second inequality, we use the fact that $(E_tf)(z)=(E_tf)(x)=0$ because of the definition of E_t. For the case that $x\in(0,2t)\times\mathbb{T}^2$ and $y\in[2t,\infty)\times\mathbb{T}^2$, by the properties of models we have

$$\begin{aligned}&v_r(x)\|(E_tf)(y)-\Gamma_{yx}(E_tf)(x)\|_\alpha=v_r(x)\|\Gamma_{yx}\{\Gamma_{xy}(E_tf)(y)-(E_tf)(x)\}\|_\alpha\\ &\le\|\Gamma\|_{\gamma,v_r}v_r^*(y-x)\sum_{\alpha\le\beta<\gamma}\|y-x\|_{\mathfrak{s}}^{\beta-\alpha}\|\Gamma_{xy}(E_tf)(y)-(E_tf)(x)\|_\beta\\ &\lesssim\|\Gamma\|_{\gamma,v_r}|\!|\!| f|\!|\!|_{\gamma,\eta;(0,2t)}\,v_r^*(y-x)\varphi(x,y)^{\eta\wedge\alpha_0-\gamma}\|y-x\|_{\mathfrak{s}}^{\gamma-\alpha}.\end{aligned}$$

Note that the supremum in the definition of the norm $\|\cdot\|_{\gamma,\eta;I}$ is taken over $\|y-x\|_{\mathfrak{s}} \le \varphi(x,y)$. Since $|y_1| \le 1+|x_1| \le 3$ in this region, the factors $v_r(x)$ and $v_r^*(y-x)$ are bounded both above and below. Thus we can ignore these weights and have $E_t f \in \mathcal{D}^{\gamma,\eta\wedge\alpha_0}((0,\infty);\Gamma)$. On the other hand, $E_t f \in \mathcal{D}^{\gamma,\eta\wedge\alpha_0}((-\infty,0);\Gamma)$ is obvious from the definition. Since $\|y-x\|_{\mathfrak{s}} \le \varphi(x,y)$ implies that x_1 and y_1 have the same sign, we obtain the assertion. □

Remark 3.5.9. Although the norm of Π-parts of models is perhaps different from the original one in [52], the norms of Γ-part and modelled distributions are not different since the semigroup $\{Q_t\}$ is not used for them. Because of this, here and in some places below (Lemma 3.5.10 and Theorem 3.20), we can use the continuity results of modelled distribution obtained in [52].

Next, we recall from [52] a different norm of singular modelled distributions. The following result holds for any singular modelled distributions on $\mathbb{R}^d$ taking values in arbitrary regularity structures and any models.

Lemma 3.5.10 ([52, lemma 6.5])**.** Let $\eta \le \gamma$ and $r \ge 0$, and let $I \subset \mathbb{R}$ be an interval. For any function $f : (I\setminus\{0\})\times\mathbb{T}^2 \to \mathbf{T}_{<\gamma}$, we define

$$(\!| f |\!)^{\circ}_{\gamma,\eta;I} := \max_{\alpha<\gamma} \sup_{x\in(I\setminus\{0\})\times\mathbb{T}^2} \frac{\|f(x)\|_\alpha}{\varphi(x)^{\eta-\alpha}}.$$

Then the inequality $(\!| f |\!)_{\gamma,\eta;I} \le (\!| f |\!)^{\circ}_{\gamma,\eta;I}$ obviously holds. Conversely, if

$$\lim_{x_1\to 0} P_\alpha f(x) = 0$$

holds for any $\alpha < \eta$, then there exists a polynomial $p(\cdot)$ such that, for any $M \in \mathscr{M}_r^{\mathrm{ad}}(\mathscr{T})$ and $f \in \mathcal{D}^{\gamma,\eta}(I;\Gamma)$, we have

$$(\!| f |\!)^{\circ}_{\gamma,\eta;I} \lesssim p(\|\Gamma\|_{\gamma,v_r}) |\!|\!| f |\!|\!|_{\gamma,\eta;I}.$$

In the end, we can lift the operator $(\partial_1 - a\Delta + c)^{-1}$ to the level of singular modelled distributions. Recall the decomposition $(\partial_1 - a\Delta + c)^{-1} = K + S$ from Proposition 3.5.2-(iii).

Theorem 3.5.11. Let $\gamma \in (0, \alpha\wedge(1-2\varepsilon))$, $\eta \in (\gamma-2,\gamma]$, $r \ge 0$, and $t \in (0,1]$. For any $M = (\Pi,\Gamma) \in \mathscr{M}_r^{\mathrm{ad}}(\mathscr{T})$, $f \in \mathcal{D}^{\gamma,\eta}((0,2t);\Gamma)$, and $\delta \in (0,\gamma+2]$, we define the function

$$\mathcal{P}_t^\delta f := P_{<\delta}\{\mathcal{K}(E_t f) + L(S(\mathcal{R}E_t f))\}.$$

Then $\mathcal{P}_t^\delta f \in \mathcal{D}^{\delta,\eta\wedge\alpha_0+2}_{v_{3r}}(\mathbb{R};\Gamma)$. If M is smooth and admissible in the sense of Definition 3.5.5, then we have

$$\mathcal{R}(\mathcal{P}_t^\delta f)(x) = (\partial_1 - a\Delta + c)^{-1}(\mathcal{R}E_t f)(x). \tag{3.18}$$

Moreover, there exists a polynomial $p(\cdot)$ such that, for any $\kappa \ge 0$ we have

$$|\!|\!| \mathcal{P}_t^\delta f |\!|\!|_{\delta,\eta\wedge\alpha_0+2-\kappa;(0,2t)} \le p(|\!|\!| M |\!|\!|_{\gamma,v_r})\, t^{\kappa/2} |\!|\!| f |\!|\!|_{\gamma,\eta;(0,2t)}. \tag{3.19}$$

Finally, there exists a polynomial $q(\cdot)$ such that

$$\begin{aligned}&|\!|\!|\mathcal{P}_t^\delta f^{(1)};\mathcal{P}_t^\delta f^{(2)}|\!|\!|_{\delta,\eta\wedge\alpha_0+2-\kappa;(0,2t)}\\&\le q(R)\,t^{\kappa/2}\big(|\!|\!|M^{(1)};M^{(2)}|\!|\!|_{\gamma,v_r}+|\!|\!|f^{(1)};f^{(2)}|\!|\!|_{\gamma,\eta;(0,2t)}\big)\end{aligned}$$

for any $M^{(i)}\in\mathscr{M}_r^{\rm ad}(\mathscr{T})$ and $f^{(i)}\in\mathcal{D}^{\gamma,\eta}((0,2t);\Gamma^{(i)})$ with $i\in\{1,2\}$ such that $|\!|\!|M^{(i)}|\!|\!|_{\gamma,v_r}\le R$ and $|\!|\!|f^{(i)}|\!|\!|_{\gamma,\eta;(0,2t)}\le R$.

Proof. In the proof of inequalities, due to the density argument, we can assume that the model M is smooth.

We know $\mathcal{K}E_tf\in\mathcal{D}_{v_{3r}}^{\gamma+2,\eta\wedge\alpha_0+2}(\mathbb{R};\Gamma)$ from Corollary 3.4.6, and $\mathcal{R}E_tf\in C^{\eta\wedge\alpha_0,\mathcal{Q}}(v_{2r})$ from Corollary 3.3.9. Moreover, since $E_tf(x)$ vanishes outside $[0,2]\times\mathbb{T}^2$, we also obtain $\mathcal{R}E_tf\in C^{\eta\wedge\alpha_0,\mathcal{Q}}(\mathbb{R}\times\mathbb{T}^2)$ by modifying the proof of Theorem 3.3.7. Then by Proposition 3.5.2-(iii), we have $S(\mathcal{R}E_tf)\in C_{\mathfrak{s}}^{\gamma+2}(\mathbb{R}\times\mathbb{T}^2)$ and thus $L(S(\mathcal{R}E_tf))\in\mathcal{D}^{\gamma+2,\gamma+2}(\mathbb{R};\Gamma)$. Therefore, $\mathcal{P}_t^\delta f\in\mathcal{D}_{v_{3r}}^{\delta,\eta\wedge\alpha_0+2}(\Gamma)$ by Proposition 3.3.5-(ii). The identity (3.18) follows from Theorem 3.4.5 and the definition of $L(S(\mathcal{R}E_tf))$.

Note that $|\!|\!|\mathcal{P}_t^\delta f|\!|\!|_{\delta,\eta\wedge\alpha_0+2;(0,2t)}\le C_r|\!|\!|\mathcal{P}_t^\delta f|\!|\!|_{\delta,\eta\wedge\alpha_0+2,v_{3r}}$ for some r-dependent constant C_r. We show (3.19) for $\kappa>0$ by applying Lemma 3.5.10. By definition, the only index $\alpha\in\mathbf{A}$ of elements in $\mathbf{S}$ smaller than $\eta\wedge\alpha_0+2$ ($\le 1-\varepsilon$) is $\alpha=0$. Since M is smooth, by Proposition 3.3.8, the $\mathbf{T}_0$-component of $\mathcal{P}_t^\delta f(x)$ is equal to

$$\big(\Pi_x(\mathcal{P}_t^\delta f)(x)\big)(x)=(\mathcal{R}\mathcal{P}_t^\delta f)(x)=(\partial_1-a\Delta+c)^{-1}(\mathcal{R}E_tf)(x).$$

Since $(\mathcal{R}E_tf)(y)=\big(\Pi_y(E_tf)(y)\big)(y)=0$ vanishes on $y\in(-\infty,0)\times\mathbb{T}^2$, we also have

$$(\partial_1-a\Delta+c)^{-1}(\mathcal{R}E_tf)(x)=\int_{[0,x_1]\times\mathbb{R}^2}P_{x_1-y_1}(x',y')(\mathcal{R}E_tf)(y)dy.$$

Note that, in the proof of Proposition 3.3.8, we obtained

$$|\mathcal{R}E_tf(y)|\lesssim\varphi(y)^{\eta\wedge\alpha_0}.$$

Since $\eta\wedge\alpha_0>-2$, we can show that

$$|(\partial_1-a\Delta+c)^{-1}(\mathcal{R}E_tf)(x)|\lesssim\int_0^{x_1}|y_1|^{(\eta\wedge\alpha_0)/2}dy_1\to 0$$

as $x_1\downarrow 0$. Therefore, by Lemma 3.5.10 we have

$$\begin{aligned}|\!|\!|\mathcal{P}_t^\delta f|\!|\!|_{\gamma,\eta\wedge\alpha_0+2-\kappa;(0,2t)}&\lesssim|\!|\!|\mathcal{P}_t^\delta f|\!|\!|^\circ_{\gamma,\eta\wedge\alpha_0+2-\kappa;(0,2t)}\\&\lesssim t^{\kappa/2}|\!|\!|\mathcal{P}_t^\delta f|\!|\!|^\circ_{\gamma,\eta\wedge\alpha_0+2;(0,2t)}\lesssim t^{\kappa/2}|\!|\!|\mathcal{P}_t^\delta f|\!|\!|_{\gamma,\eta\wedge\alpha_0+2;(0,2t)},\end{aligned}$$

where $|\!|\!|\cdot|\!|\!|^\circ_{\gamma,\eta;I}:=(\!|\cdot|\!)^\circ_{\gamma,\eta;I}+\|\cdot\|_{\gamma,\eta;I}$. The proof of the local Lipschitz estimate is a slight modification. □

3.5.4 Solution theory for PAM

We show the local-in-time well-posedness of the equation

$$U = L(Pu_0) + \mathcal{P}_t^{\gamma}\big(b(U)\Xi\big) \tag{3.20}$$

in the class $\mathcal{D}^{\gamma,\eta}((0,2t),\mathbf{S};\Gamma)$ with some appropriate choices of γ and η. The term $L(Pu_0)$ and the operator $\mathcal{P}_t^{\gamma}$ was defined in the previous subsection. The only undefined object $b(U)$ is the lift of the composition map $u \mapsto b(u)$ defined in [52, Proposition 6.13]. In the present case, for sufficiently small ε and any $U \in \mathcal{D}^{\gamma,\eta}((0,2t),\mathbf{S};\Gamma)$ with $\gamma \in (1, 2-2\varepsilon)$ and $\eta \in [0,\gamma]$ of the form

$$U(x) = u(x)\mathbf{1} + v(x)\mathcal{I}(\Xi) + u_2(x)X_2 + u_3(x)X_3,$$

we can define $b(U) \in \mathcal{D}^{\gamma,\eta}((0,2t),\mathbf{S};\Gamma)$ by the concrete form

$$b(U)(x) = b(u(x))\mathbf{1} + b'(u(x))\{v(x)\mathcal{I}(\Xi) + u_2(x)X_2 + u_3(x)X_3\}.$$

Then the map $U \mapsto b(U)$ is locally Lipschitz continuous.

Theorem 3.5.12. Assume $\varepsilon \in (0, \alpha \wedge (1/4))$ and let $\theta \in (0, 1-\varepsilon)$. Then there exists a function $t_0 : (0,\infty)^2 \to (0,1]$ such that, the following assertion holds for any $R_1, R_2 > 0$: For any $u_0 \in C^{\theta}(\mathbb{T}^2)$ such that $\|u_0\|_{C^{\theta}(\mathbb{T}^2)} \le R_1$, and any $M \in \mathscr{M}_r^{\mathrm{ad}}(\mathscr{T})$ such that $\|M\|_{\gamma,\nu_r} \le R_2$, the equation (3.20) with $t = t_0(R_1,R_2)$ and $\gamma = 1+2\varepsilon$ has a unique solution U in the class $\mathcal{D}^{1+2\varepsilon,\theta}((0,2t),\mathbf{S};\Gamma)$. Moreover, the mapping

$$S_t : (u_0, M) \mapsto U$$

is Lipschitz continuous on the space $\{u_0\,;\, \|u_0\|_{C^{\theta}(\mathbb{T}^2)} \le R_1\} \times \{M\,;\, \|M\|_{\gamma,\nu_r} \le R_2\}$.

Proof. The proof is a standard fixed point argument. Note that, the following operators are well-defined and locally Lipschitz continuous.

- ([52, Proposition 6.13]) $U \in \mathcal{D}^{1+2\varepsilon,\theta}((0,2t),\mathbf{S};\Gamma) \mapsto b(U) \in \mathcal{D}^{1+2\varepsilon,\theta}((0,2t),\mathbf{S};\Gamma)$.
- ([52, Proposition 6.12]) $V \in \mathcal{D}^{1+2\varepsilon,\theta}((0,2t),\mathbf{S};\Gamma) \mapsto V\Xi \in \mathcal{D}^{\varepsilon,\theta-1-\varepsilon}((0,2t);\Gamma)$.
- (Theorem 3.5.11) $W \in \mathcal{D}^{\varepsilon,\theta-1-\varepsilon}((0,2t);\Gamma) \mapsto \mathcal{P}_t^{1+2\varepsilon}W \in \mathcal{D}^{1+2\varepsilon,1-\varepsilon} \in ((0,2t),\mathbf{S};\Gamma)$.

Therefore, by setting $F(U) = L(Pu_0) + \mathcal{P}_t^{1+2\varepsilon}\big(b(U)\Xi\big)$, we have

$$\begin{aligned}
|\!|\!| F(U) |\!|\!|_{1+2\varepsilon,\theta;(0,2t)} &\lesssim \|u_0\|_{C^{\theta}} + t^{(1-\varepsilon-\theta)/2} |\!|\!| b(U)\Xi |\!|\!|_{\varepsilon,\theta-1-\varepsilon} \\
&\lesssim \|u_0\|_{C^{\theta}} + t^{(1-\varepsilon-\theta)/2} |\!|\!| b(U) |\!|\!|_{1+2\varepsilon,\theta} \\
&\lesssim \|u_0\|_{C^{\theta}} + t^{(1-\varepsilon-\theta)/2} p(|\!|\!| U |\!|\!|_{1+2\varepsilon,\theta})
\end{aligned}$$

for some polynomial $p(\cdot)$. From this inequality, we can find a large $R > 0$ depending on u_0 and M and show that F maps a ball of radius R in $\mathcal{D}^{1+2\varepsilon,\theta}((0,2t),\mathbf{S};\Gamma)$ into itself. From here onward, we can show the assertion by an argument similar to [52, Theorem 7.8]. $\square$

3.5.5 Convergence of models

In this subsection, we define the sequence of smooth admissible models associated with regularized noises and show its probabilistic convergence. We fix an even function $\varrho : \mathbb{R}^2 \to [0,1]$ in the Schwartz class and such that $\int_{\mathbb{R}^2} \varrho(x)dx = 1$, and set $\varrho_n(x) = 2^{2n}\varrho(2^n x)$ for each $n \in \mathbb{N}$. We define the smooth approximation of the spatial white noise ξ by

$$\xi_n(x) = \int_{\mathbb{T}^2} \widetilde{\rho_n}(x-y)\xi(y)dy, \qquad (x \in \mathbb{T}^2)$$

where $\widetilde{\rho_n}$ denotes the spatial periodization of ρ_n defined by $\widetilde{\rho_n}(x) := \sum_{k\in\mathbb{Z}^2} \rho_n(x+k)$. For such ξ_n, we can define the unique smooth admissible model $M^n = (\Pi^n, \Gamma^n) \in \mathscr{M}_r^{\mathrm{ad}}(\mathscr{T})$ by the properties

$$\begin{aligned} &(\Pi_x^n \Xi)(y) = \xi_n(y'), \qquad (\Pi_x^n X_i \Xi)(y) = (y_i - x_i)\xi_n(y'), \\ &\big(\Pi_x^n \mathcal{I}(\Xi)\Xi\big)(y) = \big(K\xi_n(y) - K\xi_n(x)\big)\xi_n(y') - C_n(y), \end{aligned}$$

where the function C_n is defined by

$$C_n(x) = \mathbb{E}\big[(K\xi_n)(x)\xi_n(x')\big] = \int_{\mathbb{R}^3} K(x,y)c_n(x'-y')dy$$

with $c_n(x'-y') := \mathbb{E}[\xi_n(x')\xi_n(y')] = \widetilde{\varrho_n^{*2}}(x'-y')$.

Theorem 3.5.13. For any $r > 0$ and $p \in [1,\infty)$, the sequence $\{M^n\}_{n\in\mathbb{N}}$ of models defined above converges in $L^p(\Omega, \mathscr{M}_r^{\mathrm{ad}}(\mathscr{T}))$.

Proof. In view of the inductive proof as in [5], it is sufficient to show the uniform bounds

$$\big|\mathbb{E}\big[Q_t(x, \Pi_x^n \tau)\big]\big| \lesssim t^{\beta/4} \tag{3.21}$$

for any $\beta \in \{-1-\varepsilon, -2\varepsilon, -\varepsilon\}$ and $\tau \in \mathbf{T}_\beta$. Note that the assumptions in [5] are more restrictive: the kernel $Q_t(x,y)$ is homogeneous in the sense that it depends only on $x-y$, and the renormalization model is defined from an x-independent preparation map. However, the first restriction is used only to prove the above estimate in [5], so if we can establish this estimate in some alternative way, we can still follow the discussion in [5]. Moreover, the second restriction is also not problematic, as the algebraic relations derived from preparation maps can be easily adapted to include x-dependent preparation maps. Such a modification is carried out in [6].

Since ξ is a centered Gaussian, we have only to show (3.21) for $\tau = \mathcal{I}(\Xi)\Xi$. By definition,

$$\begin{aligned} \mathbb{E}\big[Q_t(x, \Pi_x^n \tau)\big] &= -\int_{\mathbb{R}^3} Q_t(x,y)\mathbb{E}[(K\xi_n)(x)\xi_n(y')]dy \\ &= -\int_{(\mathbb{R}^3)^2} Q_t(x,y)K(x,z)c_n(z'-y')dydz. \end{aligned}$$

To estimate this integral, we decompose $K = \int_0^1 K_s ds$ and set

$$I_{t,s}^n(x) = -\int_{(\mathbb{R}^3)^2} Q_t(x,y)K_s(x,z)c_n(z'-y')dydz.$$

By the Gaussian estimates of Q_t and K_s, their time integral is estimated as

$$\int_{\mathbb{R}} |Q_t(x,y)|dy_1 \lesssim h_t^{(C)}(x'-y'), \qquad \int_{\mathbb{R}} |K_s(x,z)|dz_1 \lesssim s^{-1/2} h_s^{(C)}(x'-z'),$$

for some constant $C > 0$, where $h_t^{(C)}(x') := t^{-1/2} e^{-C\{(|x_2|^4/t)^{1/3} + (|x_3|^4/t)^{1/3}\}}$. Thus we have

$$|I_{t,s}^n(x)| \lesssim s^{-1/2}(h_t^{(C)} * h_s^{(C)} * |c_n|)(0).$$

Since $|h_t^{(C)} * h_s^{(C)}(x)| \lesssim h_{t+s}^{(c)}(x)$ for some constant $c \in (0,C)$ (see [6, Lemma 55] for instance), we have

$$|I_{t,s}^n(x)| \lesssim s^{-1/2}(t+s)^{-1/2}.$$

Since we have

$$\int_0^1 |I_{t,s}^n(x)|ds \lesssim \int_0^t s^{-1/2} t^{-1/2} ds + \int_t^1 s^{-1} ds \lesssim -\log t \lesssim t^{-\varepsilon/2}$$

for any $\varepsilon > 0$, we obtain the estimate (3.21) for $\tau = \mathcal{I}(\Xi)\Xi$. □

3.5.6 Renormalization of PAM

For a fixed initial condition $u_0 \in C^\theta(\mathbb{T}^2)$ and the sequence of random models $\{M^n\}$ constructed in the previous subsection, we denote by

$$U_n = S_t(u_0, M^n)$$

the solution of the equation (3.20) with $\gamma = 1 + 2\varepsilon$ and with the random time

$$t = t_0\bigg(\|u_0\|_{C^\theta(\mathbb{T}^2)}, \sup_{n\in\mathbb{N}} \|M^n\|_{\gamma, v_r}\bigg).$$

Combining Theorem 3.5.13 with Theorem 3.5.12, we have the following theorem.

Theorem 3.5.14. For each $n \in \mathbb{N}$, we denote by $\mathcal{R}^n$ the reconstruction operator associated with M^n. Then the function $u_n = \mathcal{R}^n(E_t U_n)$ converges in $L^\infty((0,t)\times\mathbb{T}^2)$ in probability as $n \to \infty$ and coincides with the unique solution of the equation

$$\big(\partial_1 - a(x')\Delta + c\big)u_n(x) = b\big(u_n(x)\big)\xi_n(x') - C_n(x)(bb')\big(u_n(x)\big) \tag{3.22}$$

with the initial value $u_0 \in C^\theta(\mathbb{T}^2)$ on $x \in (0,t)\times\mathbb{T}^2$.

As noted in Remark 3.5.3, the constant c in the equation (3.22) can be arbitrary.

Proof. On the region $x \in (0,t)\times\mathbb{T}^2$, since $u_n(x) = \big(\Pi_x^n U_n(x)\big)(x)$, we can assume that U_n is of the form

$$U_n(x) = u_n(x)\mathbf{1} + v_n(x)\mathcal{I}(\Xi) + u_{2,n}(x)X_2 + u_{3,n}(x)X_3. \tag{3.23}$$

The convergence of $\{u_n\}$ in $L^\infty((0,t)\times\mathbb{T}^2)$ follows from the convergence of $\{U_n\}$ and the definition of the norm $(\!| \cdot |\!)_{\gamma,\eta;(0,t)}$.

Finally, we show that u_n satisfies the equation (3.22) on the region $(0,t)\times\mathbb{T}^2$. For any $x\in(0,t)\times\mathbb{T}^2$, the function $b(U_n)(x)$ is of the form

$$b(U_n)(x) = b(u_n(x))\mathbf{1} + b'(u_n(x))\{v_n(x)\mathcal{I}(\Xi) + u_{2,n}(x)X_2 + u_{3,n}(x)X_3\},$$

and then $\mathcal{P}_t^{1+2\varepsilon}\big(b(U)\Xi\big)$ is of the form

$$\mathcal{P}_t^{1+2\varepsilon}\big(b(U)\Xi\big)(x) = w_n(x)\mathbf{1} + b(u_n(x))\mathcal{I}(\Xi) + w_{2,n}(x)X_2 + w_{3,n}(x)X_3$$

for some functions $w_n, w_{2,n}$, and $w_{3,n}$. For U_n to solve the equation (3.20), the coefficient $v_n(x)$ in (3.23) must be equal to $b(u_n(x))$ for any $x\in(0,t)\times\mathbb{T}^2$. By Theorem 3.5.11, the function u_n satisfies

$$u_n(x) = Pu_0(x) + \int_{[0,x_1]\times\mathbb{R}^2} P_{x_1-y_1}(x',y')(\mathcal{R}^n E_t f(U_n)\Xi)(y)dy.$$

Since $y\in(0,t)\times\mathbb{T}^2$, from the definition of $\Pi_x^n\mathcal{I}(\Xi)\Xi$, we obtain

$$\begin{aligned}(\mathcal{R}E_t b(U_n)\Xi)(y) &= \big(\Pi_y E_t b(U_n)(y)\Xi\big)(y) = \big(\Pi_y b(U_n)(y)\Xi\big)(y)\\ &= b(u_n(y))\xi_n(y') - C_n(y)(bb')(u_n(y)).\end{aligned}$$

This implies that u_n satisfies the equation (3.22) (in mild sense) on $(0,t)\times\mathbb{T}^2$. □

We also have a stronger convergence result.

Corollary 3.5.15. In the setting of Theorem 3.5.14, the convergence of $\{u_n\}$ also holds in the space $C_{\mathfrak{s}}^\theta((0,t)\times\mathbb{T}^2)$.

Proof. We only show the uniform bounds of $\{u_n\}$ in the θ-Hölder norm, since the proof of the convergence is a simple modification. First, we set $\bar{U}_n = \mathcal{P}_t^{1+2\varepsilon}\big(b(U_n)\Xi\big) \in \mathcal{D}^{1+2\varepsilon,\theta}((0,2t),\mathbf{S};\Gamma^n)$ and decompose

$$u_n = Pu_0 + \bar{u}_n, \qquad \bar{u}_n := \mathcal{R}^n(E_t\bar{U}_n).$$

Since the uniform bounds of $\{Pu_0\}$ in the θ-Hölder norm is more elementary (see e.g. [6, Proposition 62]), we focus on the remaining term. By definition, for any $x\in(0,t)\times\mathbb{T}^2$, $\bar{u}_n(x)$ coincides with the $\mathbf{1}$-component of $\bar{U}_n(x)$, and also with that of $P_{<\theta}\bar{U}_n(x)$. Since $\{P_{<\theta}\bar{U}_n\}$ is uniformly bounded in the norm $\|\cdot\|_{\theta,\theta;(0,t)}$ by Proposition 3.3.5-(ii), we have

$$|\bar{u}_n(y) - \bar{u}_n(x)| \lesssim \|y-x\|_{\mathfrak{s}}^\theta$$

for any $x,y\in(0,t)\times\mathbb{T}^2$ such that $\|y-x\|_{\mathfrak{s}} \le \varphi(x,y)$. Here and in what follows, we omit proportional constants polynomially depending on the norms of $\{\bar{U}_n\}$ and $\{\Gamma^n\}$, which are uniform over n. It remains to show the same Hölder-type inequality in the region $\varphi(x,y) < \|y-x\|_{\mathfrak{s}}$. In this region, by using the inequality (3.5), we have $\varphi(x)\vee\varphi(y) \lesssim \|y-x\|_{\mathfrak{s}}$. On the other hand, we also have that $\{\bar{U}_n\}$ is uniformly bounded in the norm $(\!|\cdot|\!)^\circ_{1+2\varepsilon,\theta;(0,t)}$ by Lemma 3.5.10. Hence

$$|\bar{u}_n(y) - \bar{u}_n(x)| \le |\bar{u}_n(y)| + |\bar{u}_n(x)| \lesssim \varphi(y)^\theta + \varphi(x)^\theta \lesssim \|y-x\|_{\mathfrak{s}}^\theta$$

in the region $\varphi(x,y) < \|y-x\|_{\mathfrak{s}}$. This completes the proof. □

Bibliography

[1] M. Avellaneda, D. Boyer-Olson, J. Busca, and P. Friz, Application of large deviation methods to the pricing of index options in finance, *Comptes Rendus Mathematique* 336 (3), 263-266 (2003).

[2] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*, Springer, Heidelberg (2011).

[3] I. Bailleul and Y. Bruned, *Locality for singular stochastic PDEs*, arXiv:2109.00399.

[4] I. Bailleul and M. Hoshino, *A tourist's guide to regularity structures*, arXiv:2006.03524.

[5] I. Bailleul and M. Hoshino, *Random models on regularity-integrability structures*. arXiv:2310.10202.

[6] I. Bailleul, M. Hoshino, and S. Kusuoka, *Regularity structures for quasilinear singular SPDEs*, Arch. Ration. Mech. Anal. **248** (2024), no. 6, Paper No. 127.

[7] C. Bayer, P. Friz, and J. Gatheral, Pricing under rough volatility, *Quantitative Finance*, 16 (2016), pp. 887-904.

[8] C. Bayer, P. Friz, P. Gassiat, J. Martin and B. Stemper, A regularity structure for rough volatility, *Mathematical Finance*, 30, 782–832, (2020).

[9] C. Bayer, P.K. Friz, A. Gulisashvili, B. Horvath and B. Stemper, Short-time near-the-money skew in rough fractional volatility models, *Quantitative Finance*, 19, 779–798, (2019).

[10] C. Bellingeri, P. K. Friz, and M. Gerencsér, *Singular paths spaces and applications*, Stoch. Anal. Appl. **40** (2022), 1126–1149.

[11] M. Bennedsen, A. Lunde, and M.S. Pakkanen, Decoupling the Short- and Long-Term Behavior of Stochastic Volatility, *Journal of Financial Econometrics*, nbaa049, https://doi.org/10.1093/jjfinec/nbaa049

[12] H. Berestycki, J. Busca, and I. Florent, Asymptotics and calibration of local volatility models, *Quantitative Finance* 2 (1), 61-69, (2002).

[13] P.Billingsley, *Convergence of Probability Measures*, 2.nd, Wiley,(1999).

[14] A.E. Bolko, K. Christensen, M.S. Pakkanen, and B. Veliyev, A GMM approach to estimate the roughness of stochastic volatility, to appear in Journal of Econometrics (2022).

[15] F. Bourgey, S. De Marco, and E. Gobet, Weak approximations and VIX option price expansions in forward variance curve models, Quantitative Finance , Vol.23, Issue 9, (2023).

[16] L. Broux, F. Caravenna, and L. Zambotti, *Hairer's multilevel Schauder estimates without Regularity Structures*. arXiv:2301.07517.

[17] L. Broux and D. Lee, *Besov reconstruction*. Potential Anal. **59** (2023), 1875–1912.

[18] H-Q. Bui, X. T. Duong and L. Yan, *Calderón reproducing formulas and new Besov spaces associated with operators*, Adv. in Math. **229** (2012), 2449–2502.

[19] F. Caravenna and L. Zambotti, *Hairer's reconstruction theorem without regularity structures*. EMS Surv. Math. Sci. **7** (2020), 207–251.

[20] A. Dahlqvist, J. Diehl, and B. K. Driver, *The parabolic Anderson model on Riemann surfaces*. Probab. Theory Related Fields **174** (2019), 369–444.

[21] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed., Springer, 1998.

[22] G. Folland, *Real Analysis Modern Techniques and Their Applications*, 2nd., Wiley, (1999).

[23] M. Forde and H. Zhang, Asymptotics for rough stochastic volatility models, *SIAM Journal on Financial Mathematics* 8, 114–145 (2017).

[24] P. Friz, P. Gassiat and P. Pigato, Precise asymptotics: robust stochastic volatility models, *The Annals of Applied Probability*, 31, 896–940, (2021).

[25] P. Friz, P. Gassiat and P. Pigato, Short-dated smile under rough volatility: asymptotics and numerics, *Quantitative Finance*, 21, 1–18, (2021).

[26] P. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, and J. Teichmann, *Large Deviations and Asymptotic Methods in Finance*, Springer, (2015).

[27] P. Friz and M. Hairer, *A Course on Rough Paths*, Springer (2014).

[28] P. K. Friz and B. Seeger, *Besov rough path analysis. With an appendix by Z.-K. Pavel*. J. Differential Equations **339** (2022), 152–231.

[29] P. Friz and N. Victoir, Approximations of the Brownian rough path with applications to stochastic analysis, *Annales de l'Institut Henri Poincare (B) Probability and Statistics* 41 (4), 703-724, (2005).

[30] P. Friz and N. Victoir, *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*, Cambridge Univarsity Press (2010).

[31] M. Fukasawa, Volatility has to be rough, *Quantitative Finance* 21, 1-8 (2021).

[32] M. Fukasawa, On asymptotically arbitrage-free approximations of the implied volatility, Frontiers of Mathematical Finance Vol. 1, No. 4, pp. 525-537, (2022).

[33] M. Fukasawa, Wiener Spiral for Volatility Modeling, *Theory of probability and its application*, Vol.68, Iss.3, (2023).

[34] M. Fukasawa and J. Gatheral, A rough SABR formula, *Frontiers of Mathematical Finance*, 1 (2022), 81-97.

[35] M. Fukasawa, B. Horvath and P. Tankov, Hedging under rough volatility, arXiv:2105.04073.

[36] M. Fukasawa, T. Takabatake and R. Westphal, Consistent estimation for fractional stochastic volatility model under high-frequency asymptotics, *Mathematical Finance*, 32 (2022), 1086-1132.

[37] M. Fukasawa and R. Takano, A partial rough path for rough volatility, *Electronic Journal of Probability*, Vol. 29, no. 18, 1-28, (2024).

[38] H. Funahashi and M. Kijima, A Solution to the Time-Scale Fractional Puzzle in the Implied Volatility, Fractal and Fractional,1, no. 1: 14, https://doi.org/10.3390/fractalfract1010014, (2017).

[39] J. Garcia, A Large Deviation Principle for Stochastic Integrals, *Journal of Theoretical Probability* 21, 476-501, (2008).

[40] J. Gatheral, T. Jaisson, and M. Rosenbaum, Volatility is rough, *Quantitative Finance*, 18 (2018), pp. 933-949.

[41] M. Gerencsér and M. Hairer, *Singular SPDEs in domains with boundaries*. Probab. Theory Relat. Fields **173** (2019), 697–758.

[42] S. Gerhold, A. Jacquier, M. Pakkanen, H. Stone and T. Wagenhofer, Pathwise large deviations for the rough Bergomi model: Corrigendum, *Journal of Applied Probability*, 58, 849–850, (2021).

[43] M. Gubinelli, Controlling rough paths, *Journal of Functional Analysis* 216, no.1, (2004) .

[44] M. Gubinelli, Ramification of rough paths, *Journal of Differential Equations*, 248, no.4, (2010) .

[45] M. Gubinelli, P. Imkeller, and N. Perkowski, *Paracontrolled distributions and singular PDEs*. Forum Math. Pi **3** (2015), e6, 75 pp.

[46] M. Gubinelli and A. Lejay, Global existence for rough differential equations under linear growth conditions, hal:00384327, (2009).

[47] H. Guennoun, A. Jacquier, P. Roome and F. Shi, Asymptotic behaviour of the fractional Heston model, *SIAM Journal on Financial Mathematics*, 9(3), 1017-1045, (2018).

[48] A, Gulisashvili, Large deviation principle for Volterra type fractional stochastic volatility models, *SIAM Journal on Financial Mathematics*, 9, 1102–1136, (2018).

[49] A. Gulisashvili, Multivariate Stochastic Volatility Models and Large Deviation Principles, arXiv:2203.09015, (2022).

[50] A. Gulisashvili, F. Viens, and X. Zhang, Extreme-strike asymptotics for general Gaussian stochastic volatility models, *Annals of Finance*, 15, 59–101, (2019).

[51] A. Gulisashvili, F. Viens, and X. Zhang, Small-time asymptotics for Gaussian self-similar stochastic volatility models, *Applied Mathematics & Optimization*, 82, 183–223, (2020).

[52] M. Hairer, *A theory of regularity structures*, Invent. Math. **198** (2014), 269–504.

[53] M. Hairer and C. Labbé, *The reconstruction theorem in Besov spaces*. J. Funct. Anal. **273** (2017), 2578–2618.

[54] M. Hairer and C. Labbé, *Multiplicative stochastic heat equations on the whole space*. J. Eur. Math. Soc. **20** (2018), 1005–1054.

[55] M. Hairer and H. Singh, *Regularity Structures on Manifolds and Vector Bundles*. arXiv:2308.05049.

[56] D. Hamadouche. Invariance principles in Hölder spaces. *Portugaliae Mathematica*, 57(2):127-151, (2000).

[57] S. Hensel and T. Rosati, *Modelled distributions of Triebel–Lizorkin type*. Studia Math. **252** (2020), 251–297.

[58] B. Horvath, A. Jacquier,and A. Muguruza, Functional Central Limit Theorems for Rough Volatility, SSRN3078743, (2017).

[59] M. Hoshino, *A semigroup approach to the reconstruction theorem and the multilevel Schauder estimate*, arXiv:2310.07396.

[60] Y. Inahana, *Rough path theory* (in Japanese), Iwanami (2022).

[61] A. Jacquier and M. Oumgari, Deep PPDEs for rough local stochastic volatility, SSRN 3400035, (2019).

[62] A. Jacquier, M. S. Pakkanen and H. Stone, Pathwise Large Deviations for the Rough Bergomi Model, *Journal of Applied Probability*, 55,1078-1092, (2018).

[63] A. Jacquier and A. Panner, Large and moderate deviations for stochastic Volterra systems, *Stochastic Processes and their Applications*, 149, 142-187, (2022).

[64] J. Le Gall, *Brownian Motion, Martingales, and Stochastic Calculus*, Springer, 2013.

[65] A. Lejay, On rough differential equations, Electronic Journal of Probability 14, 341-364 (2009).

[66] C. Liu, D. J. Prömel, and J. Teichmann, *Stochastic analysis with modelled distributions*. Stoch PDE: Anal Comp **9** (2021), 343–379.

[67] N. Muskhelishvili, *Singular Integral Equations. Boundary Problems of Functions Theory and Their Applications to Mathematical Physics*, edited by J.R.M. Radok. Reprinted. Wolters-Noordhoff Publishing, Groningen (1972).

[68] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer, (2005).

[69] Y. Osajima, General asymptptics of Wiener functionals and applications to implied volatilities, In: *Large Deviations and Asymptotic Methods in Finance*, Springer, (2015).

[70] F. Otto and H. Weber, *Quasilinear SPDEs via Rough Paths*. Arch Rational Mech Anal **232** (2019), 873–950.

[71] H. Pham, H. Some Applications and Methods of Large Deviations in Finance and Insurance. In: *Paris-Princeton Lectures on Mathematical Finance 2004*, Lecture Notes in Mathematics, vol 1919. Springer, Berlin, Heidelberg, (2007).

[72] A. Puhalksii, Large deviation analysis of the single server queue, *Queueing Systems* 21, 5-66, (1995).

[73] P. Rinaldi and F. Sclavi, *Reconstruction theorem for germs of distributions on smooth manifolds*. J. Math. Anal. Appl. **501** (2021), Paper No. 125215, 14 pp.

[74] H. Singh, *Canonical solutions to non-translation invariant singular SPDEs*. arXiv:2310.01085.

[75] H. Singh and J. Teichmann, *An elementary proof of the reconstruction theorem*. arXiv:1812.03082.

[76] P. Zorin-Kranich, *The reconstruction theorem in quasinormed spaces*. Rev. Mat. Iberoam. **39** (2023), 1233–1246.

Acknowledgment

I truly appreciate Professor Masaaki Fukasawa for his fruitful guidance and comments for my research during my doctoral course. I would also like to thank Professor Masato Hoshino, a collaborator in my research, for his valuable comments and the opportunity for discussion.

The author was supported by JST (Grant Number JPMJSP2138) from April 2023 to March 2025. The author is currently supported by JSPS KAKENHI (Grant Number 25KJ1763) from April 2025.

List of Works

This thesis is based on the following works:

Publications

1. M.Hoshino and R. Takano, A semigroup approach to the reconstruction theorem and the multilevel Schauder estimate for singular modelled distributions, Stoch PDE: Anal Comp (2025), https://doi.org/10.1007/s40072-025-00352-5, arXiv:2408.04322.

2. R. Takano, Large Deviation Principle for Stochastic Differential Equations Driven by Stochastic Integrals, SIAM Journal on Financial Mathematics Vol.16, 2, p480-515, (2025), arXiv:2403.14321.

3. M. Fukasawa, R. Takano, A partial rough path space for rough volatility, Electron. J. Probab. 29 (2024), article no. 18, 1–28, arXiv:2205.09958.

Talks

1. A semigroup approach to the reconstruction theorem and its applications,Far East Probability Workshop 2025, Hokkaido university, June 2025. [invited]

2. A semigroup approach to the reconstruction theorem and its applications, Third China-Japan-Korea Joint Probability Workshop, Kyoto university, May 2025.

3. A semigroup approach to the reconstruction theorem and its applications, 東北大学数学教室確率論セミナー, 東北大学, 2025 年 4 月. [invited]

4. A semigroup approach to the reconstruction theorem and its applications, 2024 年度日本数学会年会, 早稲田大学, 2025 年 3 月.

5. Large deviation principle for rough volatility models, Winter Workshop on Operations Research, Finance and Mathematics, 2025, Hokkaido, Feb. 2025.

6. A semigroup approach to the reconstruction theorem and its applications, New Trends in Rough Path Analysis, Osaka university, Feb, 2025.

7. Large deviation principle for rough volatility models, Ritsumeikan University Probability and Mathematical Finance Seminar, Ritsumeikan University, Jan. 2025. [invited]

8. A semigroup approach to the reconstruction theorem and the multilevel Schauder estimate for singular modelled distributions, 2024 年度確率論シンポジウム, 京都大学, 2024 年 12 月.

9. Large deviation principle for rough volatility models, UQ-Osaka Seminar on Financial Mathematics and Economics, online, Dec. 2024. [invited]

10. Large deviation principle for stochastic differential equations driven by stochastic integrals, 確率解析とその周辺, ソニックシティビル, 2024 年 12 月.

11. Large deviation principle for rough volatility models, TMU Workshop on Finance 2024, Tokyo Metropolitan University, Sep. 2024. [invited]

12. A semigroup approach to the reconstruction theorem and the multilevel Schauder estimate for singular modelled distributions, 2024 Open German-Japanese Conference on Stochastic Analysis and Applications, Hokkaido University, Sep. 2024, poster session.

13. 確率積分が駆動する確率微分方程式に対する大偏差原理, 日本数学会 2024 年度秋季総合分科会, 大阪大学, 2024 年 9 月.

14. A semigroup approach to the reconstruction theorem and the multilevel Schauder estimate for singular modelled distributions, 2024 年度確率論ヤングサマーセミナー, 西谷津温泉宮本の湯, 2024 年 8 月.

15. Large deviation principle for stochastic differential equations driven by stochastic integrals, The Eighth Asian Quantitative Finance Conference, National Taipei University of Technology, Aug. 2024.

16. Large deviation principle for stochastic differential equations driven by stochastic integrals, 九州大学確率論セミナー, 九州大学, 2024 年 6 月. [invited]

17. ラフボラティリティモデルに対する大偏差原理, 丸の内 QF セミナー, 東京都立大学, 2024 年 5 月. [invited]

18. Large deviation principle for stochastic differential equations driven by stochastic integrals, 関西確率論セミナー, 京都大学, 2024 年 5 月. [invited]

19. A partial rough path space for rough volatility, Winter Workshop on Operations Research, Finance and Mathematics 2024, Hokkaido, Mar. 2024.

20. A partial rough path space for rough volatility, 関西大学確率論セミナー, 関西大学, 2024 年 2 月. [invited]

21. A partial rough path space for rough volatility, 16 Bachelier Colloquium, France, Jan. 2024

22. A partial rough path space for rough volatility, 大阪大学確率論セミナー, 大阪大学, 2023 年 4 月,

23. The large deviation for the Lyons-Victoir extension, 確率論早春セミナー 2023, 奈良女子大学, 2023 年 3 月.

24. The large deviation for the Lyons-Victoir extension, 日本数学会 2023 年度年会, 中央大学, 2023 年 3 月.

25. The large deviation for the Lyons-Victoir extension, Osaka-UCL Mini-Workshop on Stochastics, Numerics and Risk, Osaka University, Feb. 2023.

26. The large deviation for the Lyons-Victoir extension, 2022 年度確率論シンポジウム, 京都大学, 2022 年 12 月.
27. A partial rough path space for rough volatility, 2022 年度確率論ヤングサマーセミナー, 京都大学, 2022 年 8 月.
28. A partial rough path space for rough volatility, 日本数学会 2022 年度年会, 2022 年 3 月,（コロナウイルス蔓延措置有）
29. A partial rough path space for rough volatility, 日本応用数理学会第 18 回研究部会連合発表会, オンライン, 2022 年 3 月
30. A partial rough path space for rough volatility, 2021 年度確率論シンポジウム, オンライン, 2021 年 12 月.
31. Rough Path について, 2021 年度確率論若手セミナー, オンライン, 2021 年 8 月.