
Article:
Lê, K. orcid.org/0000-0002-7654-7139 and Ling, C. (Accepted: 2024) Taming singular stochastic differential equations: A numerical method. Annals of Probability. ISSN 0091-1798 (In Press)

Submitted to the Annals of Probability

TAMING SINGULAR STOCHASTIC DIFFERENTIAL EQUATIONS: A NUMERICAL METHOD

BY KHOA LÊ[1,a] AND CHENGCHENG LING[2,b]

[1]*School of Mathematics, University of Leeds, Leeds, United Kingdom;* [a]*k.le@leeds.ac.uk*

[2]*Institut für Mathematik, Universität Augsburg, Augsburg, Germany;* [b]*chengcheng.ling@uni-a.de*

We consider a generic and explicit tamed Euler–Maruyama scheme for multidimensional time-inhomogeneous stochastic differential equations with multiplicative Brownian noise. The diffusive coefficient is uniformly elliptic, Hölder continuous and weakly differentiable in the spatial variables while the drift satisfies the strict Ladyzhenskaya–Prodi–Serrin condition, as considered by Krylov and Röckner (2005). In the discrete scheme, the drift is tamed by replacing it by an approximation. A strong rate of convergence of the scheme is provided in terms of the approximation error of the drift in a suitable and possibly very weak topology. A few examples of approximating drifts are discussed in detail. The parameters of the approximating drifts can vary and—under suitable conditions— be fine-tuned to achieve a strong convergence rate which is arbitrarily close to the benchmark 0.5 rate. The result is then applied to provide numerical solutions for stochastic transport equations with singular vector fields satisfying the aforementioned condition.

CONTENTS

1. Introduction. The aim of this article is to devise a numerical scheme and obtain its strong convergence rate for stochastic differential equations (SDEs) with integrable drift coefficients and elliptic regular diffusive coefficients. We consider the SDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad X_0 = x_0, \quad t \in [0,1], \tag{1.1}$$

where $d \geqslant 1$, $b : [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ is a Borel measurable function satisfying

$$\int_0^1 \left[\int_{\mathbb{R}^d} |b(t,x)|^p dx \right]^{\frac{q}{p}} dt < \infty \quad \text{with} \quad q, p \in [2, \infty) \quad \text{and} \quad \frac{d}{p} + \frac{2}{q} < 1, \tag{1.2}$$

MSC2020 subject classifications: Primary 60H35, 60H10; secondary 60H50, 60L90, 35B65.

Keywords and phrases: Singular SDEs, strong approximation, tamed Euler scheme, regularization by noise, stochastic sewing, Zvonkin's transformation, quantitative Khasminskii's lemma.

and $\sigma : [0,1] \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ is a bounded Borel measurable function, continuous in the spatial variables and uniformly elliptic, $(B_t)_{t\geqslant 0}$ is a d-dimensional standard Brownian motion defined on some complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ and x_0 is a $\mathcal{F}_0$-random variable. With $\leqslant 1$ in place of < 1, (1.2) is known in the fluid dynamics' literature as the Ladyzhenskaya–Prodi–Serrin condition.

In the seminal paper [41], Krylov and Röckner, building upon [74, 68], show that (1.1) has a unique strong solution assuming that σ is the identity matrix and b satisfies (1.2). This result is later extended by Zhang [71] (complemented by [69])[1] for variable diffusive coefficients which are weakly differentiable, uniformly elliptic, uniformly bounded and uniformly continuous in x locally uniformly in time.

While theoretical solutions of (1.1) are well understood since [41], numerical analysis of (1.1) under condition (1.2) has been an open problem. At the moment of writing, we are aware of two publications on the topic. Jourdain and Menozzi consider in [35] the case σ is the identity matrix and show that the marginal density of a tamed Euler–Maruyama scheme with truncated drifts converges at the rate $\frac{1}{2} - \frac{d}{2p} - \frac{1}{q}$. Gyöngy and Krylov in [26] recently show that the tamed Euler–Maruyama scheme with truncated drifts converges in probability to the exact solution, albeit without any rate. Needless to say, a strong convergence rate is desirable and is of independent interest. For this purpose, we consider the discrete scheme defined by

$$dX_t^n = b^n(t, X^n_{k_n(t)})dt + \sigma(t, X^n_{k_n(t)})dB_t, \quad X_0^n = x_0^n, \quad t \in [0,1], \tag{1.3}$$

where x_0^n is a $\mathcal{F}_0$-random variable and b^n is an approximation of the vector field b and

$$k_n(t) = \frac{j}{n} \text{ whenever } \frac{j}{n} \leqslant t < \frac{j+1}{n} \text{ for some integer } j \geqslant 0.$$

We note that (1.3) with the choice $b^n = b$ is the usual Euler–Maruyama scheme, which, however, is not well-behaved for a merely integrable function b even when b is replaced by $b\mathbf{1}_{(|b|<\infty)}$. This is because the simulation for the usual Euler–Maruyama scheme may enter a neighborhood of a singularity of b, making the scheme unstable and uncontrollable. We thus have to tame the vector field b, replacing it by a suitable approximation b^n. Henceforth, we call (1.3) a *tamed Euler–Maruyama scheme*. The terminology is borrowed from [32], who consider a specific case of (1.3) to approximate SDEs with regular but super-linear drifts. The name "tamed Euler–Maruyama" thus should be understood in a broad sense, and in particular, (1.3) also includes the "truncated Euler–Maruyama" scheme considered in [54].

Natural choices for b^n are the truncated vector fields

$$b_r^n(x) = b_r(x)\mathbf{1}_{(|b_r(x)|\leqslant Cn^\chi \|b_r\|_{L_p(\mathbb{R}^d)})}, \tag{1.4}$$

$$b_r^n(x) = b_r(x)\mathbf{1}_{(|b_r(x)|\leqslant Cn^\chi)}, \tag{1.5}$$

for some constants $C, \chi > 0$. Another practical choice is the regularized vector field

$$b_r^n(x) = p_{1/n^\chi} * b_r(x), \tag{1.6}$$

where $\chi > 0$, $p_t(x)$ is the Gaussian density of variance t and $*$ is the spatial convolution. Alternatively, multiresolution approximations by wavelet ([56]) or the truncated discrete φ-transform ([19]) could be used whenever desirable.

The main results of the article, Theorems 2.2 and 2.3 below, assert the strong convergence of (1.3) to (1.1) with an explicit rate under some mild regularity conditions on σ and on the approximating drifts b^n. When the approximating drifts take one of the forms (1.4)-(1.6), Corollary 2.4 expresses the convergence rates which are proportional to χ. For each form

[1] [71, Theorem 5.1] is non-trivial whose proof is provided in [69].

of b^n, a suitable validity range of χ is identified in terms of the parameters p, q, d. For the approximating drifts (1.5) and Lipschitz σ, the parameter χ can be tuned within the interval $(0, 1/2)$ to obtain a strong convergence rate χ, which is arbitrarily close to the benchmark 0.5 rate. For the approximating drifts (1.4) and (1.6), such sharp rate can be achieved under some restricted conditions on p, q.

We expect that Theorems 2.2 and 2.3 are useful in algorithm designs when the vector field b is not explicitly available but rather arises from another analytic system which itself needs to be numerically evaluated. Such situations appear in hydrodynamic-type equations due to their fundamental connection with singular SDEs, see for instance, [12, 70, 73] where the SDE (1.1) is coupled with another analytic constraint on b. In such scenarios, b^n does not have an explicit form, but nevertheless, Theorems 2.2 and 2.3 could be implemented. While we leave this problem for future investigations, herein we focus on a simpler application to stochastic transport equations with vector fields satisfying (1.2) (see Eq. (8.1)). While theoretical solutions for such equations have been considered in [18, 17, 60, 7], singularity of the coefficients have prevented the study of numerical solutions by standard tools ([11, 61]). We propose in Theorem 8.1 an explicit numerical scheme with rate for such equations, based upon the method of characteristics.

Literatures on convergence of Euler–Maruyama schemes for SDEs is vast and expanding, for which we provide a brief and personalized overview. When the coefficients are continuous, convergence rates of the Euler–Maruyama scheme are well-studied. For Lipschitz continuous coefficients and non-trivial diffusive coefficients, the optimal strong rate of convergence is $1/2$, as shown in [43, 34]. Results on the strong rate of convergence for Hölder / Dini continuous drifts are discussed in [27, 5, 55] and only settled recently by Dareiotis and Gerencsér in [13], who obtain the $L_2(\Omega)$-rate $1/2 - \varepsilon$, for any $\varepsilon \in (0, 1/2)$, when b is Dini continuous and σ is the identity matrix. This result is extended in [10] for the case when b is Hölder continuous and σ is uniformly elliptic and twice continuously differentiable. For an in-depth overview and more complete lists of other contributions, see [9, 38, 39, 57] and the references therein. Results for discontinuous drifts are more sparse but are attracting attention. The case of piecewise Lipschitz drifts are considered in [49, 50, 58]. [59] considers one-dimensional SDEs with additive noise and bounded measurable drifts with a positive Sobolev–Slobodecki-regularity. [6] considers bounded measurable drifts with a certain Gaussian–Besov-regularity. For merely bounded measurable drifts without any regularity, the recent article [14] obtains the $L_p(\Omega)$-rate $1/2 - \varepsilon$, for any $p \geqslant 2$ and $\varepsilon \in (0, 1/2)$, extending the results of [13, 10]. At last, we mention the work [2] who consider similar tamed Euler–Maruyama schemes for one dimensional SDEs with distributional drifts. For comparison, our approach is different, our results are in a multidimensional setting and allow completely generic approximating drifts b^n. Furthermore, we emphasize that one dimensional SDEs are more specific, often well-posed even for distributional drifts, and usually require tailored techniques, [44, 29]. This list is surely not exhaustive.

The article is organized as follows. In Section 2, we state our standing assumptions and the main results. Section 3 contains auxiliary results which are collected and adapted from previous works. Section 4 is pivotal and contains a case study of moment estimation for some relevant functionals of Brownian motion. While some results in this section will not be used directly to prove the main results, the section showcases our main estimates in a simpler setting. Sections 5 and 6 extend the moment estimates in Section 4 respectively to functionals of the solutions of (1.3) and (1.1). The two sections contain most of the technical estimates of the paper which build up a foundation for the proofs of the main results. In Section 7, we give the proofs of Theorems 2.2 and 2.3, using the moment estimates from the prior sections. The application to numerical solutions for stochastic transport equations is discussed in Section 8. The appendix contains maximal regularity estimates for parabolic equations with variable coefficients and distributional forcing, which are needed but independent from the main text.

2. Main results. We first fix a few notation. Let $p, q \in [1, \infty]$ be some fixed parameters. $L_p(\mathbb{R}^d)$ and $L_p(\Omega)$ denote the Lebesgue spaces respectively on $\mathbb{R}^d$ and Ω. The expectation with respect to $\mathbb{P}$ is denoted by $\mathbb{E}$. For each $\nu \in \mathbb{R}$, $L_{\nu,p}(\mathbb{R}^d) := (\mathbb{I} - \Delta)^{-\nu/2}\big(L^p(\mathbb{R}^d)\big)$ is the usual Bessel potential space on $\mathbb{R}^d$ equipped with the norm $\|f\|_{L_{\nu,p}(\mathbb{R}^d)} := \|(\mathbb{I} - \Delta)^{\nu/2} f\|_{L_p(\mathbb{R}^d)}$, where $(\mathbb{I} - \Delta)^{\nu/2} f$ is defined through Fourier's transform. $\mathbb{L}^q_{\nu,p}([0,1])$ denotes the space of measurable function $f : [0,1] \to L_{\nu,p}(\mathbb{R}^d)$ such that $\|f\|_{\mathbb{L}^q_{\nu,p}([0,1])}$ is finite. Here, for each $s, t \in [0,1]$ satisfying $s \leqslant t$, we denote

$$\|f\|_{\mathbb{L}^q_{\nu,p}([s,t])} := \left(\int_s^t \|f(r,\cdot)\|^q_{L_{\nu,p}(\mathbb{R}^d)} dr \right)^{\frac{1}{q}}$$

with obvious modification when $q = \infty$. When $\nu = 0$, we simply write $\mathbb{L}^q_p([0,1])$ instead of $\mathbb{L}^q_{0,p}([0,1])$. In particular, $\mathbb{L}^q_p([0,1])$ contains Borel measurable functions $f : [0,1] \times \mathbb{R}^d \to \mathbb{R}$ such that $\int_0^1 \left[\int_{\mathbb{R}^d} |f(t,x)|^p dx \right]^{q/p} dt$ is finite. For each $\rho \in (0,1)$, $L_\rho(\mathbb{R}^d)$ denotes the space of all measurable functions f on $\mathbb{R}^d$ such that $\|f\|_{L_\rho(\mathbb{R}^d)} := (\int_{\mathbb{R}^d} |f(x)|^\rho dx)^{1/\rho}$ is finite. Note that in this case, $\|\cdot\|_{L_\rho(\mathbb{R}^d)}$ is not a norm.

For each $\mathcal{X} \in \{\mathbb{L}^q_{\nu,p}([0,1]), L_p(\mathbb{R}^d), L_p(\Omega)\}$, an $\mathbb{R}^m$-valued function $f = (f^1, \ldots, f^m)$ belongs to $\mathcal{X}$, if all components $f^1, \ldots, f^m$ belong to $\mathcal{X}$, and we put $\|f\|_{\mathcal{X}} = \max_{i=1,\ldots,m}\{\|f^i\|_{\mathcal{X}}\}$. Since we only deal with either scalars or $\mathbb{R}^d$-valued functions and random variables, we conventionally drop the dimension of the range in the notation of the spaces $\mathbb{L}^q_{\nu,p}([0,1]), L_p(\mathbb{R}^d), L_p(\Omega)$.

Put $D_n = \{i/n : i = 0, \ldots, n\}$. For each $S \leqslant T$, we put $\Delta([S,T]) = \{(s,t) \in [S,T]^2 : s \leqslant t\}$ and $\Delta_2([S,T]) = \{(s,u,t) \in [S,T]^3 : s \leqslant u \leqslant t\}$. We abbreviate $\Delta = \Delta([0,1])$ and $\Delta_2 = \Delta_2([0,1])$. We say that a function $w : \Delta([S,T]) \to [0,\infty)$ is a control if $w(s,u) + w(u,t) \leqslant w(s,t)$ for every $(s,u,t) \in \Delta_2([S,T])$. For a $d \times d$-matrix P, P^* denotes its transpose and $\|P\|$ denotes its Hilbert–Schmidt norm. The following conditions are enforced throughout unless noted otherwise.

CONDITION $\mathfrak{A}$. The diffusion coefficient σ is a $d \times d$-matrix-valued measurable function on $[0,1] \times \mathbb{R}^d$. There exists a constant $K_1 \in [1, \infty)$ such that for every $s \in [0,1]$ and $x \in \mathbb{R}^d$

$$K_1^{-1} I \leqslant (\sigma\sigma^*)(s,x) \leqslant K_1 I. \tag{2.1}$$

Furthermore, the following conditions hold.

1. There are constants $\alpha \in (0,1]$ and $K_2 \in (0,\infty)$ such that for every $s \in [0,1]$ and $x, y \in \mathbb{R}^d$

$$|(\sigma\sigma^*)(s,x) - (\sigma\sigma^*)(s,y)| \leqslant K_2 |x-y|^\alpha.$$

2. $\sigma(s,\cdot)$ is weakly differentiable for a.e. $s \in [0,1]$ and there are constants $p_0 \in [2,\infty)$, $q_0 \in (2,\infty]$ and $K_3 \in (0,\infty)$ such that

$$\frac{d}{p_0} + \frac{2}{q_0} < 1 \quad \text{and} \quad \|\nabla\sigma\|_{\mathbb{L}^{q_0}_{p_0}([0,1])} \leqslant K_3.$$

CONDITION $\mathfrak{B}$. x_0 belongs to $L_p(\Omega, \mathcal{F}_0)$ and b belongs to $\mathbb{L}^q_p([0,1])$ for some $p, q \in [2,\infty)$ satisfying $\frac{d}{p} + \frac{2}{q} < 1$. For each n, x_0^n belongs to $L_p(\Omega, \mathcal{F}_0)$ and b^n belongs to $\mathbb{L}^q_p([0,1]) \cap \mathbb{L}^q_\infty([0,1])$ with p, q as above. Furthermore, there exist finite positive constants K_4, θ and continuous controls $\{\mu^n\}_n$ such that $\sup_{n \geqslant 1}(\|b^n\|_{\mathbb{L}^q_p([0,1])} + \mu^n(0,1)) \leqslant K_4$ and

$$(1/n)^{\frac{1}{2} - \frac{1}{q}} \|b^n\|_{\mathbb{L}^q_\infty([s,t])} \leqslant \mu^n(s,t)^\theta \quad \forall\, t - s \leqslant 1/n. \tag{2.2}$$

In the above, I denotes the identity matrix. If one replaces Hölder continuity by uniform continuity, Conditions $\mathfrak{A}$-$\mathfrak{B}$ are comparable to those from [71, 69], who show strong uniqueness for (1.1). Hence, hereafter, we assume that the solution to (1.1) exists and is strongly unique.[2] Next, we define an important quantity which controls the strong convergence rate.

DEFINITION 2.1. Let $\lambda > 0$ be a fixed number which is sufficiently large. Let $U = (U^1, \dots, U^d)$ where for each $h = 1, \dots, d$, U^h is the solution to the following equation

$$\partial_t U^h + \sum_{i,j=1}^{d} \frac{1}{2}(\sigma\sigma^*)^{ij}\partial^2_{ij} U^h + b^{n,h} \cdot \nabla U^h = \lambda U^h - b^{n,h}, \quad U^h(1,\cdot) = 0. \tag{2.3}$$

Let X be the solution to (1.1). For each $\bar{p} \in [1, \infty)$, we put

$$\varpi_n(\bar{p}) = \Big\| \sup_{t\in[0,1]} \Big| \int_0^t (1 + \nabla U)[b - b^n](r, X_r)dr \Big| \Big\|_{L_{\bar{p}}(\Omega)}.$$

In the above and hereafter, we omit the dependence of U on n. Equation (2.3) arises from a Zvonkin transformation, which we postpone to Section 7 for the details. It is known that when λ is sufficiently large, equation (2.3) has a unique solution, see Lemma 7.1 below.

THEOREM 2.2. *Assume that Conditions $\mathfrak{A}$-$\mathfrak{B}$ hold. Let $(X^n_t)_{t\in[0,1]}$ be the solution to (1.3) and $(X_t)_{t\in[0,1]}$ be the solution to SDE (1.1). Then for any $\bar{p} \in (1, p) \cap (1, \frac{2}{d}(p \wedge p_0))$ and any $\gamma \in (0, 1)$, there exists a finite constant $N(K_1, K_2, K_3, K_4, \alpha, p_0, q_0, p, q, d, \bar{p}, \gamma)$ such that*

$$\| \sup_{t\in[0,1]} |X^n_t - X_t| \|_{L_{\gamma\bar{p}}(\Omega)} \leqslant N \left[\|x^n_0 - x_0\|_{L_{\bar{p}}(\Omega)} + (1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}} \log(n) + \varpi_n(\bar{p}) \right]. \tag{2.4}$$

Actually, by adding an exponential weight, moments up to order p-th can be estimated, see Proposition 7.3 below. The condition $\bar{p} < \frac{2}{d}(p \wedge p_0)$ ensures finiteness of the moments of the exponential weight and therefore deduces (2.4) by an application of Hölder inequality.

Under uniform ellipticity and Hölder regularity of σ, pathwise uniqueness for (1.1), $p > d/\alpha$, $\lim_n b^n = b$ in $\mathbb{L}^q_p([0,1])$ and the following condition

$$\sup_{n\geqslant 1} (1/n)^{\frac{1}{2}-\frac{1}{q}} \|b^n\|_{\mathbb{L}^q_\infty([0,1])} < \infty, \tag{2.5}$$

[26, Theorem 2.11] recently shows that the tamed Euler–Maruyama scheme (1.3) converges in probability to the solution of (1.1). It is evident that (2.5) implies (2.2) (with the choice $\mu^n(s,t)^{\frac{1}{q}} = (1/n)^{\frac{1}{2}-\frac{1}{q}} \|b^n\|_{\mathbb{L}^q_\infty([s,t])}$). However, because of the interchangeability between $t - s$ and $1/n$ in (2.2), truncated vector fields with higher truncation levels, which yield better convergence rates, satisfy (2.2) but not (2.5), see Corollary 2.4 below. Under Conditions $\mathfrak{A}$-$\mathfrak{B}$, the above result provides an upper bound for the moments of $\sup_{t\in[0,1]} |X^n_t - X_t|$ which depends on n and ϖ_n. When $\lim_n b^n = b$ in $\mathbb{L}^q_p([0,1])$ as in the setting of [26], one can show that $\lim_n \varpi_n = 0$ (cf. Corollary 2.4). However, the topology of $\mathbb{L}^q_p$ does not provide any explicit rate. There are, of course, many other topologies for $\lim_n b^n = b$ so that one can actually obtain an explicit rate. The choice of a suitable topology depends on the approximating vector fields b^n. Our next main result relates ϖ_n with the convergence of b^n to b with respective to the topologies of $\mathbb{L}^{q_1}_{p_1}$ (for some $p_1, q_1 \in [1, \infty]$) and $\mathbb{L}^q_{-\nu,p}$ (for some $\nu \in [0, 1)$).

[2] Actually the results from [71, 69] are for deterministic $x_0 \in \mathbb{R}^d$, however, can be easily extended to our case by conditioning and utilizing Markov property of Brownian motion. See also Remark 2.6 below.

THEOREM 2.3. *Assume that Conditions $\mathfrak{A}$-$\mathfrak{B}$ hold.*

(i) Let $p_1, q_1 \in [1,\infty]$ be such that $\frac{d}{p_1} + \frac{2}{q_1} < 2$. Then for every $m \geqslant 1$, there exists a constant N depending on K_1, K_2, K_3, K_4, α, p_0, q_0, p_1, q_1, p, q, d, m such that

$$\varpi_n(m) \leqslant N \|b - b^n\|_{\mathbb{L}^{q_1}_{p_1}([0,1])}. \tag{2.6}$$

(ii) Assuming furthermore that $q_0 = \infty$ and $\frac{1}{p} + \frac{1}{p_0} < 1$. Let $\nu \in [0,1)$ be such that

$$\nu < \frac{3}{2} - \frac{d}{2p} - \frac{2}{q}. \tag{2.7}$$

Then for every $\bar{p} \in [1,p)$, there exists a constant N depending on K_1, K_2, K_3, K_4, α, p_0, p, q, d, $\bar{p}$, ν such that

$$\varpi_n(\bar{p}) \leqslant N \|b - b^n\|_{\mathbb{L}^q_{-\nu,p}([0,1])}. \tag{2.8}$$

(iii) Assuming furthermore that $q_0 = \infty$, $\frac{1}{p} + \frac{1}{p_0} < 1$ and

$$\frac{d}{p} + \frac{4}{q} < 1. \tag{2.9}$$

Suppose that there exists a continuous control w_0 on Δ and a constant $\Gamma \geqslant 0$ such that

$$\|b - b^n\|_{\mathbb{L}^q_{-1,p}([s,t])} \leqslant \Gamma w_0(s,t)^{\frac{1}{q}} \quad \text{and} \quad \|b - b^n\|_{\mathbb{L}^q_p([s,t])} \leqslant w_0(s,t)^{\frac{1}{q}} \tag{2.10}$$

for every $(s,t) \in \Delta$. Then for every $\bar{p} \in [1,p)$, there exists a constant N depending on K_1, K_2, K_3, K_4, α, p_0, p, q, d, $\bar{p}$ such that

$$\varpi_n(\bar{p}) \leqslant N\Gamma\left(1 + |\log\Gamma|\right) w_0(0,1)^{\frac{1}{q}}. \tag{2.11}$$

Using Theorems 2.2 and 2.3, we can derive explicit strong convergence rates for the scheme (1.3) when the approximating vector field b^n take one of the forms (1.4)-(1.6).

COROLLARY 2.4. *Assume that Condition $\mathfrak{A}$ holds. Let b, x_0, x_0^n be as in Condition $\mathfrak{B}$; and $\bar{p}, \gamma$ be as in Theorem 2.2.*

(a) Let $C > 0$ and $\chi \in (0, 1/2 - 1/q]$ be constants and define b^n by (1.4). Let $\rho \in (1,p]$ be a number such that $\rho\frac{d}{p} + \frac{2}{q} < 2$. Then there exists a constant N depending on K_1, K_2, K_3, K_4, α, p_0, q_0, p, q, d, $\bar{p}$, γ, ρ, χ, C such that

$$\begin{aligned} &\| \sup_{t\in[0,1]} |X_t^n - X_t| \|_{L_{\gamma\bar{p}}(\Omega)} \\ &\qquad \leqslant N\left[\|x_0 - x_0^n\|_{L_{\bar{p}}(\Omega)} + (1/n)^{\chi(\rho-1)} + (1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}}\log(n)\right]. \end{aligned} \tag{2.12}$$

(b) Let $C > 0$ and $\chi \in (0, 1/2)$ be constants and define b^n by (1.5). Then there exists a constant N depending on K_1, K_2, K_3, K_4, α, p_0, q_0, p, q, d, $\bar{p}$, γ, ρ, χ, C such that (2.12) holds for any $\rho \in (1, p \wedge q]$ satisfying $\rho\left(\frac{d}{p} + \frac{2}{q}\right) < 2$.

(c) Let $\chi \in \left(0, \frac{p}{d}\left(1 - \frac{2}{q}\right)\right]$ and define b^n by (1.6). Let $\nu \in (0,1)$ be any number satisfying (2.7). Assume furthermore that $q_0 = \infty$ and $\frac{1}{p} + \frac{1}{p_0} < 1$. Then there exists a constant N depending on K_1, K_2, K_3, K_4, α, p_0, p, q, d, $\bar{p}$, ν, γ, χ such that

$$\| \sup_{t\in[0,1]} |X_t^n - X_t| \|_{L_{\gamma\bar{p}}(\Omega)} \leqslant N\left[\|x_0 - x_0^n\|_{L_{\bar{p}}(\Omega)} + (1/n)^{\chi\frac{\nu}{2}} + (1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}}\log(n)\right]. \tag{2.13}$$

PROOF. In view of Theorem 2.2, it suffices to estimate $\varpi_n(\bar{p})$. (a) It is obvious that $\|b^n\|_{\mathbb{L}^q_p([0,1])} \leqslant \|b\|_{\mathbb{L}^q_p([0,1])}$. From the inequality $\|b^n_r\|_{L_\infty(\mathbb{R}^d)} \leqslant Cn^\chi \|b_r\|_{L_p(\mathbb{R}^d)}$, we see that $\|b^n\|_{\mathbb{L}^q_\infty([s,t])} \lesssim n^\chi \|b\|_{\mathbb{L}^q_p([s,t])}$. It follows that for every $0 \leqslant t-s \leqslant 1/n$,

$$(1/n)^{\frac{1}{2}-\frac{1}{q}} \|b^n\|_{\mathbb{L}^q_\infty([s,t])} \lesssim (1/n)^{\frac{1}{2}-\frac{1}{q}-\chi} \|b\|_{\mathbb{L}^q_p([s,t])} \lesssim \|b\|_{\mathbb{L}^q_p([s,t])},$$

verifying Condition $\mathfrak{B}$ with $\mu^n(s,t) = \|b\|^q_{\mathbb{L}^q_p([s,t])}$ and $\theta = 1/q$. Furthermore,

$$|b_r(x) - b^n_r(x)| \leqslant |b_r(x)| \mathbf{1}_{(|b_r(x)| > Cn^\chi \|b_r\|_{L_p(\mathbb{R}^d)})} \leqslant C^{1-\rho} n^{-\chi(\rho-1)} \|b_r\|^{1-\rho}_{L_p(\mathbb{R}^d)} |b_r(x)|^\rho.$$

The function $(r,x) \mapsto \|b_r\|^{1-\rho}_{L_p(\mathbb{R}^d)} |b_r(x)|^\rho$ belongs to $\mathbb{L}^q_{p/\rho}([0,1])$ and hence,

$$\|b - b^n\|_{\mathbb{L}^q_{p/\rho}([0,1])} \lesssim n^{-\chi(\rho-1)} \|b\|_{\mathbb{L}^q_p([0,1])}.$$

It follows from (2.6) that $\varpi_n \lesssim (1/n)^{\chi(\rho-1)} \|b\|_{\mathbb{L}^q_p([0,1])}$. The stated estimate is then a consequence of (2.4).

(b) For the vector field b^n defined by (1.5), we have $\|b^n_r\|_{L_\infty(\mathbb{R}^d)} \lesssim n^\chi$ so that for every $0 \leqslant t-s \leqslant 1/n$,

$$(1/n)^{\frac{1}{2}-\frac{1}{q}} \|b^n\|_{\mathbb{L}^q_\infty([s,t])} \lesssim (1/n)^{\frac{1}{2}-\frac{1}{q}-\chi} (t-s)^{\frac{1}{q}} \lesssim (t-s)^\theta$$

for any $\theta > 0$ such that $\theta \leqslant \min(1/q, 1/2-\chi)$. This verifies Condition $\mathfrak{B}$ with $\mu^n(s,t) = t-s$. On the other hand, $|b-b^n| \lesssim n^{-\chi(\rho-1)} |b|^\rho$ so that $\|b-b^n\|_{\mathbb{L}^{q/\rho}_{p/\rho}([0,1])} \lesssim n^{-\chi(\rho-1)} \|b\|^\rho_{\mathbb{L}^q_p([0,1])}$. It follows from (2.6) that $\varpi_n \lesssim n^{-\chi(\rho-1)} \|b\|^\rho_{\mathbb{L}^q_p([0,1])}$.

(c) We have $\|b^n_r\|_{L_\infty(\mathbb{R}^d)} \lesssim n^{\chi\frac{d}{2p}} \|b_r\|_{L_p(\mathbb{R}^d)}$ and hence, for every $0 \leqslant t-s \leqslant 1/n$,

$$(1/n)^{\frac{1}{2}-\frac{1}{q}} \|b^n\|_{\mathbb{L}^q_\infty([s,t])} \lesssim (1/n)^{\frac{1}{2}-\frac{1}{q}-\chi\frac{d}{2p}} \|b\|_{\mathbb{L}^q_p([s,t])} \lesssim \|b\|_{\mathbb{L}^q_p([s,t])},$$

verifying condition (2.2) with $\mu^n(s,t) = \|b\|^q_{\mathbb{L}^q_p([s,t])}$ and $\theta = 1/q$. We also have

$$\|b - b^n\|_{\mathbb{L}^q_{-\nu,p}} \lesssim (1/n)^{\chi\nu/2} \|b\|_{\mathbb{L}^q_p}.$$

Applying (2.8), we have $\varpi_n(\bar{p}) \lesssim (1/n)^{\chi\frac{\nu}{2}}$. □

REMARK 2.5. Similar truncated vector fields to (1.5) with the values $\chi = 1/2$ and $\chi = d/(2p) + 1/q$ were considered in [35], in which a weak rate of convergence of order $\frac{1}{2} - \frac{d}{2p} - \frac{1}{q}$ was obtained. While Corollary 2.4(b) excludes the value $\chi = 1/2$, by choosing $\rho = 2$, it yields the strong rate

$$\|x_0 - x^n_0\|_{L_{\bar{p}}(\Omega)} + (1/n)^\chi + (1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}} \log(n),$$

in which χ can be as close as one desires to $1/2$.

Similar (but different) regularized vector fields to (1.6) was considered in [2] in a different setting.

REMARK 2.6. The proof of Theorem 2.2 actually works for any adapted solution to (1.1), see Remark 6.5 and Section 7. Consequently, Theorem 2.2 yields an alternative proof ([71, 69]) of pathwise uniqueness for (1.1) under Conditions $\mathfrak{A}$-$\mathfrak{B}$.

The restriction on the unit time interval in Theorems 2.2 and 2.3 is of course artificial and it is straightforward to extend the above results on arbitrary finite time intervals. In such case, the constants in our estimates also depend on the length the time interval. The logarithmic factor in (2.4) arises from the stochastic Davie–Grönwall lemma with critical exponents (see

[20] or Lemma 3.2 herein). The explicit estimation for square moments from [13] suggests that the logarithmic factor in (2.4) could be improved. Because of the role of the stochastic Davie–Grönwall lemma in the study of rough/stochastic ordinary/partial differential equations ([20, 4, 15, 16]), it is an important problem to identify the sharpness of the logarithmic factor. However, we do not pursue this direction herein.

Let us briefly explain our general method and strategy. Starting from (1.1) and (1.3), we decompose the difference $X_t - X_t^n$ into three types of differences:

- differences between functionals of $b(t, X_t)$ and $b^n(t, X_t)$,
- differences between functionals of X_t and functionals of X_t^n,
- differences between functionals of X_t^n and functionals of $X_{k_n(t)}^n$.

At this stage, our strategy aligns with the classical works [43, 34] for SDE's with Lipschitz coefficients. However, in order to utilize the regularizing effect of the noise in compensation for the lack of regularity of the drift, our treatments for these functionals are different and follow the recent approach of [14]. The differences of the first type can be easily estimated from above by ϖ_n. For the differences of the second type, we use a Zvonkin-type transformation to show that they depend on $\sup_{t\in[0,1]} |X_t - X_t^n|$ in a Lipschitz sense. The differences of the last type contain, for instance, the functional

$$\sup_{t\in[0,1]} \left| \int_0^t [b^n(s, X_{k_n(s)}^n) - b^n(s, X_s^n)] ds \right|.$$

Because b and b^n are not continuous (uniformly in n), estimation for the above functional is a challenging problem and one has to utilize the regularizing effect from the noise, an important observation made by Dareiotis and Gerencsér in [13]. For these differences, we use stochastic sewing techniques—originated from [45] and further extended in [20, 47]—to estimate them by a constant multiple of $(1/n)^{\alpha/2} + (1/n)^{1/2} \log(n)$. From here, we obtain an integral inequality for the moment of $\sup_{t\in[0,1]} |X_t - X_t^n|$. An application of the stochastic Grönwall inequality yields the desired estimate in Theorem 2.2. From this analysis, one observes that the strong rate of convergence for (1.3) is deduced from the rates of the estimations for the differences of the first and the last types. The estimates for ϖ_n in Theorem 2.3 are obtained by mean of Krylov estimates, Khasminskii estimates and stochastic sewing techniques, utilizing statistical properties of the solution to (1.1).

We make a few observations comparing with previous works. Setting technicalities aside, our proof of Theorem 2.2 follows the approach of [14]; and similar to [10], we also apply stochastic sewing techniques to obtain moment estimates for the differences of the last type. However, the works [10, 14] crucially rely on the fact that the drifts are either continuous or bounded, which is not available under Conditions $\mathfrak{A}$-$\mathfrak{B}$. In particular, the stochastic lemmas in [10, 14] cannot be applied under Conditions $\mathfrak{A}$-$\mathfrak{B}$ because the resulting Hölder exponents are strictly below $1/2$; and even if control functions were employed, one would end up with a regularity exponent of exactly $1/2$ (cf. Propositions 4.2 and 5.12).[3] In other words, the situations considered herein are at the border line and are critical to a certain extent. To successfully adapt the method above to the current setting, to overcome criticality and to remove the ε-loss in the obtained rate, we have benefited from the recent stochastic Davie–Grönwall lemma with critical exponents from [20], the analysis for singular paths from [8] and novel usage of control functions inspired by Lyon's theory of rough paths [52]. To the authors' knowledge, these tools, which are developed within rough path theory, have not been utilized previously in stochastic numerics. Lastly, in order to verify the hypotheses for

[3] We recall that an exponent of $1/2 + \varepsilon$ is required in these stochastic sewing lemmas.

stochastic sewing and of independent interests, we have obtained some new and improved analytic estimates ([37, 69, 35, 26, 48, 6]) for the probability laws of the solutions to the discrete scheme (1.3) and to equation (1.1) (see Sections 5 and 6).

While this article was under review, progress had been made in improving our main results. Namely, [46] removes the moment restrictions in Theorems 2.2 and 2.3 and consequently in Corollary 2.4. This is accomplished by taking advantage of two other recent developments (after the first appearance of the current article). One is the stability results from Galeati and the second named author in [22]. The other is the John–Nirenberg inequality for stochastic processes of bounded mean oscillations as discussed in [46].

Convention. Whenever convenience, we place temporal variables into subscript right after the function, e.g. $f_t(x) = f(t,x)$. The relation $A \lesssim B$ means that $A \leqslant CB$ for some finite constant $C \geqslant 0$. The implicit constants C may change from one inequality to another and their values may depend on other parameters which are clear from the context. We will also make use of Einstein's convention of summation over repeated indices.

3. Preliminaries.

In the current section, we collect and enhance some relevant results which appear separately in previous works from various authors. These results form a useful toolbox which is used in later sections to prove our main results.

For any one-parameter process $t \mapsto Y_t$ and any two-parameter process $(s,t) \mapsto A_{s,t}$, we denote $\delta Y_{s,t} = Y_t - Y_s$ and $\delta A_{s,u,t} = A_{s,t} - A_{s,u} - A_{u,t}$ for every $s \leqslant u \leqslant t$. We say that Y (resp. A) is L_m-integrable if $\|Y_t\|_{L_m(\Omega)}$ (resp. $\|A_{s,t}\|_{L_m(\Omega)}$) is finite for each t (resp. (s,t)); we say A is adapted if $A_{s,t}$ is $\mathcal{F}_t$-measurable whenever $s \leqslant t$. Let $v \in [0,1]$ and let $\mathbb{P}|\mathcal{F}_v$ be the probability measure conditioned on $\mathcal{F}_v$. We denote by $L_p(\Omega|\mathcal{F}_v)$ the space of random variables Z such that

$$\|Z\|_{L_p(\Omega|\mathcal{F}_v)} := \operatorname*{ess\,sup}_{\omega} [\mathbb{E}(|Z|^p|\mathcal{F}_v)]^{1/p} < \infty.$$

The advantages of considering the conditional moment norms over the usual moment norms are summarized in the following result, which is implicit in [14, 20].

LEMMA 3.1. *Let $\mathcal{A} = (\mathcal{A}_t)_{t\in[0,1]}$ be a continuous adapted stochastic process and let $p, N \in (0,\infty)$ be some fixed constants. Assume that $\mathcal{A}_0 = 0$ and*

$$\sup_{0\leqslant s\leqslant t\leqslant 1} \|\delta\mathcal{A}_{s,t}\|_{L_p(\Omega|\mathcal{F}_s)} \leqslant N.$$

Then the following statements hold.

(i) *There exists a constant $c(p)$ such that $\|\mathcal{A}_\tau\|_{L_p(\Omega)} \leqslant c(p)N$ for any stopping time $\tau \leqslant 1$.*
(ii) *For every $\bar{p} \in (0,p)$, there exists a constant $c(\bar{p},p)$ such that*

$$\| \sup_{t\in[0,1]} |\mathcal{A}_t| \|_{L_{\bar{p}}(\Omega)} \leqslant c(\bar{p},p)N.$$

PROOF. Let τ be a stopping time taking finitely many values $\{t_j\} \subset [0,1]$. By assumption, we have

$$\mathbb{E}(|\delta\mathcal{A}_{\tau,1}|^p) = \mathbb{E}\sum_j \mathbf{1}_{(\tau=t_j)}\mathbb{E}(|\delta\mathcal{A}_{t_j,1}|^p|\mathcal{F}_{t_j}) \leqslant \mathbb{E}\sum_j \mathbf{1}_{(\tau=t_j)}N^p \leqslant N^p.$$

Using the elementary inequality $(a+b)^p \lesssim a^p + b^p$, we have

$$\|\mathcal{A}_\tau\|_{L_p(\Omega)} \lesssim \|\delta\mathcal{A}_{\tau,1}\|_{L_p(\Omega)} + \|\mathcal{A}_1\|_{L_p(\Omega)} \lesssim N.$$

By approximations and continuity of $\mathcal{A}$, the above inequality also holds for all stopping times $\tau \leqslant 1$. This shows (i). Part (ii) is a consequence of part (i) and Lenglart inequality. □

The next result is a variant of the stochastic Davie–Grönwall lemma from [20] and is closely related to the stochastic sewing lemmas from [45, 47].

LEMMA 3.2 (Stochastic sewing). *Let $\varepsilon > 0$; $v, S, T, C_1, C_2, C_3, \Gamma_1, \Gamma_2 \geqslant 0$ be fixed numbers such that $0 \leqslant v < S < T$. Let w be a deterministic control on $\Delta([S,T])$ which is continuous. Let J be a L_m-integrable adapted process indexed by $\Delta([S,T])$ such that*

$$\|J_{s,t}\|_{L_m(\Omega|\mathcal{F}_v)} \leqslant C_2 w(s,t)^{\frac{1}{2}+\varepsilon}, \quad \|\mathbb{E}_s J_{s,t}\|_{L_m(\Omega|\mathcal{F}_v)} \leqslant C_1 w(s,t)^{1+\varepsilon}, \tag{3.1}$$

$$\|\delta J_{s,u,t}\|_{L_m(\Omega|\mathcal{F}_v)} \leqslant \Gamma_2 w(s,t)^{\frac{1}{2}} + C_3\Gamma_2 w(s,t)^{\frac{1}{2}+\varepsilon} \tag{3.2}$$

and

$$\|\mathbb{E}_s \delta J_{s,u,t}\|_{L_m(\Omega|\mathcal{F}_v)} \leqslant \Gamma_1 w(s,t)^{1+\varepsilon} \tag{3.3}$$

for every (s,u,t) in $\Delta_2([S,T])$. Then there exists a constant $N = N(\varepsilon, m)$, in particular independent from $\Gamma_1, \Gamma_2, C_1, C_2, S, T, v$ and w, such that for every $(s,t) \in \Delta([S,T])$

$$\begin{aligned}\|J_{s,t}\|_{L_m(\Omega|\mathcal{F}_v)} \leqslant N\Gamma_2 \Big[(1+|\log\Gamma_2|)w(s,t)^{\frac{1}{2}} + C_1 w(s,t)^{1+\varepsilon} + (C_2+C_3)w(s,t)^{\frac{1}{2}+\varepsilon}\Big] \\ + N\Gamma_1 w(s,t)^{1+\varepsilon}.\end{aligned} \tag{3.4}$$

PROOF. For each $(s,t) \in \Delta([S,T])$, define

$$u = \inf\{r \in [s,t] : w(s,r) \geqslant \frac{1}{2} w(s,t)\}$$

and call u the w-midpoint of $[s,t]$. Since t trivially belongs to the set defining u above, such a point always exists and uniquely defined. If u is a w-midpoint of $[s,t]$, then it follows from continuity of w that

$$w(s,u) \leqslant \frac{1}{2} w(s,t) \quad \text{and} \quad w(u,t) \leqslant \frac{1}{2} w(s,t).$$

See [47] for more detail. For convenience, we denote $(s|t)$ for the w-midpoint of $[s,t]$.

Let (s,t) be in $\Delta([S,T])$. Define $d_0^0(s,t) = s$ and $d_1^0(s,t) = t$. For each integers $h \geqslant 0$ and $i = 0, \dots, 2^{h+1}$, we set $d_i^{h+1}(s,t) = d_{i/2}^h(s,t)$ if i is even and $d_i^{h+1}(s,t)$ equal to the w-midpoint of $[d_{(i-1)/2}^h(s,t), d_{(i+1)/2}^h(s,t)]$ if i is odd. Set $D_w^h(s,t) := \{d_i^h(s,t)\}_{i=0}^{2^h}$ for each $h \geqslant 0$. It is readily checked that for every integers $h \geqslant 0$ and $i = 0, \dots, 2^h - 1$, we have

$$D_w^h(s,t) \subset D_w^{h+1}(s,t), \tag{3.5}$$

$$[d_i^h(s,t), d_{i+1}^h(s,t)] = [d_{2i}^{h+1}(s,t), d_{2i+1}^{h+1}(s,t)] \cup [d_{2i+1}^{h+1}(s,t), d_{2i+2}^{h+1}(s,t)], \tag{3.6}$$

$$w(d_i^h(s,t), d_{i+1}^h(s,t)) \leqslant 2^{-h} w(s,t). \tag{3.7}$$

Herein, we abbreviate $\|\cdot\|$ for $\|\cdot\|_{L_m(\Omega|\mathcal{F}_v)}$. The implicit constants below only depend on ε and m. By triangle inequality, we have

$$\|J_{s,t}\| \leqslant \Big\| \sum_{[u,v] \in D_w^h(s,t)} J_{u,v} \Big\| + \Big\| J_{s,t} - \sum_{[u,v] \in D_w^h(s,t)} J_{u,v} \Big\|.$$

We estimate the first term using conditional BDG inequality ([47, Section 2]), condition (3.1) and (3.7),

$$\begin{aligned}\| \sum_{[u,v] \in D_w^h(s,t)} J_{u,v} \| &\lesssim \sum_{[u,v] \in D_w^h(s,t)} \|\mathbb{E}_u J_{u,v}\| + \left(\sum_{[u,v] \in D_w^h(s,t)} \|J_{u,v}\|^2 \right)^{1/2} \\ &\lesssim C_1 2^{-h\varepsilon} w(s,t)^{1+\varepsilon} + C_2 2^{-h\varepsilon} w(s,t)^{\frac{1}{2}+\varepsilon}.\end{aligned}$$

For the second term, we derive from (3.6) (cf. [47, Lemma 3.6]) and conditional BDG inequality that for $h \geqslant 1$

$$\|J_{s,t} - \sum_{[u,v]\in D_w^h(s,t)} J_{u,v}\| = \|\sum_{k=0}^{h-1} \sum_{[u,v]\in D_w^k(s,t)} \delta J_{u,(u|v),v}\|$$

$$\lesssim \sum_{k=0}^{h-1} \sum_{[u,v]\in D_w^k(s,t)} \|\mathbb{E}_u \delta J_{u,(u|v),v}\| + \sum_{k=0}^{h-1} \left(\sum_{[u,v]\in D_w^k(s,t)} \|\delta J_{u,(u|v),v}\|^2 \right)^{1/2}.$$

Applying (3.3), (3.2) and (3.7), we have

$$\sum_{[u,v]\in D_w^k(s,t)} \|\mathbb{E}_u \delta J_{u,(u|v),v}\| \lesssim 2^{-k\varepsilon} \Gamma_1 w(s,t)^{1+\varepsilon}$$

and

$$\left(\sum_{[u,v]\in D_w^k(s,t)} \|\delta J_{u,(u|v),v}\|^2 \right)^{\frac{1}{2}} \lesssim \Gamma_2 w(s,t)^{\frac{1}{2}} + 2^{-k\varepsilon} C_3 \Gamma_2 w(s,t)^{\frac{1}{2}+\varepsilon}.$$

Summing in k, we have

$$\Big\|J_{s,t} - \sum_{[u,v]\in D_w^h(s,t)} J_{u,v}\Big\| \lesssim \Gamma_1 w(s,t)^{1+\varepsilon} + h\Gamma_2 w(s,t)^{\frac{1}{2}} + C_3 \Gamma_2 w(s,t)^{\frac{1}{2}+\varepsilon}.$$

Combining the previous estimates, we have shown that for every integer $h \geqslant 1$ and each $(s,t) \in \Delta$

$$\|J_{s,t}\| \lesssim 2^{-h\varepsilon} \left[C_1 w(s,t)^{1+\varepsilon} + C_2 w(s,t)^{\frac{1}{2}+\varepsilon} \right] + \Gamma_1 w(s,t)^{1+\varepsilon} + h\Gamma_2 w(s,t)^{\frac{1}{2}} + C_3 \Gamma_2 w(s,t)^{\frac{1}{2}+\varepsilon}. \tag{3.8}$$

If $\Gamma_2 \geqslant 1$, we choose $h = 1$ while if $\Gamma_2 < 1$, we choose h such that $2^{-h\varepsilon} \approx \Gamma_2$. In both cases, we obtain (3.4) from (3.8). □

Some controls which are relevant for our purpose are given below.

EXAMPLE 3.3. (a) For any $\phi \in L_q([0,1])$, $q \in [1,\infty)$, $w(s,t) = \|\phi\|^q_{L_q([s,t])}$ is a continuous control on $\Delta([0,1])$. (b) For any $\nu \geqslant 0$, $w(s,t) = s^{-\nu}(t-s)$ is a continuous control on $\Delta([S,T])$ for any $0 < S \leqslant T$. (c) For any controls w_1, w_2 and any number $\theta \in [0,1]$, $w = w_1^\theta w_2^{1-\theta}$ is another control. For further examples and basic properties of controls, we refer to [21, Chapter 5].

The following result is an excerpt from [8, Lemma 2.3].

LEMMA 3.4. *Let $(\mathcal{E}, \|\cdot\|)$ be a normed vector space, $s_{-1}, \tau_i, \eta_i \in [0,1]$, $i = 1, \dots, h$ be fixed numbers and let $Y : (0,1] \to \mathcal{E}$ be a function such that*

$$\|Y_t - Y_s\| \leqslant \sum_{i=1}^h C_i s^{-\eta_i} (t-s)^{\tau_i} \quad \forall s_{-1} \leqslant s \leqslant t \leqslant 1, s \neq 0 \tag{3.9}$$

for some constant $C_1, \dots, C_h \geqslant 0$. Assume that $\tau_i - \eta_i > 0$ for each i. Then

$$\|Y_t - Y_s\| \leqslant \sum_{i=1}^h (1 - 2^{\eta_i - \tau_i})^{-1} C_i (t-s)^{\tau_i - \eta_i} \quad \forall s_{-1} \leqslant s \leqslant t \leqslant 1, s \neq 0.$$

PROOF. Observe that (3.9) implies that Y is continuous on $[s_{-1},1]\setminus\{0\}$. We fix $s_{-1}\leqslant s<t\leqslant 1$, $s\neq 0$, and put $s_n=s+(t-s)2^{-n}$ for each integer $n\geqslant 0$. By continuity and triangle inequality, we have

$$\|Y_t-Y_s\|\leqslant\sum_{n=0}^{\infty}\|Y_{s_n}-Y_{s_{n+1}}\|\leqslant\sum_{n=0}^{\infty}\sum_{i=1}^{h}C_i s_{n+1}^{-\eta_i}(s_n-s_{n+1})^{\tau_i}.$$

Note that $s_{n+1}\geqslant(t-s)2^{-n-1}$ and $s_n-s_{n+1}=(t-s)2^{-n-1}$. Hence, from the previous estimate, we have

$$\|Y_t-Y_s\|\leqslant\sum_{i=1}^{h}\sum_{n=0}^{\infty}C_i(t-s)^{\tau_i-\eta_i}2^{-(n+1)(\tau_i-\eta_i)}.$$

Because $\sum_{n=0}^{\infty}2^{-(n+1)(\tau_i-\eta_i)}\leqslant(1-2^{\eta_i-\tau_i})^{-1}$ for each i, this yields the stated estimate. □

The following result is the Khasminskii's lemma[4] enhanced with some quantitative estimates.

LEMMA 3.5 (Quantitative Khasminskii's lemma). *Let S,T be such that $0\leqslant S\leqslant T$ and let $\{\beta(t)\}_{t\in[S,T]}$ be a nonnegative measurable $(\mathcal{F}_t)$-adapted process. Assume that for all $S\leqslant s\leqslant t\leqslant T$,*

$$\left\|\int_s^t\beta(r)dr\right\|_{L_1(\Omega|\mathcal{F}_s)}\leqslant\rho(s,t), \tag{3.10}$$

where $(s,t)\mapsto\rho(s,t)$ is a nonrandom function on $\Delta([S,T])$ satisfying the following conditions:

(i) $\rho(t_1,t_2)\leqslant\rho(t_3,t_4)$ *if* $(t_1,t_2)\subset(t_3,t_4)$,
(ii) $\lim_{h\downarrow 0}\sup_{S\leqslant s<t\leqslant T,|t-s|\leqslant h}\rho(s,t)=\kappa$, $\kappa\geqslant 0$.

Then for any real $\lambda<\kappa^{-1}$, (if $\kappa=0$, then $\kappa^{-1}=\infty$), and any integer $m\geqslant 1$

$$\mathbb{E}\exp\left(\lambda\int_S^T\beta(r)dr\right)<\infty\quad\text{and}\quad\left\|\int_S^T\beta(r)dr\right\|_{L_m(\Omega)}\leqslant(m!)^{\frac{1}{m}}\rho(S,T).$$

Suppose additionally that there exist $\gamma>0$ and a continuous control w on $\Delta([S,T])$ such that

(iii) $\rho(s,t)\leqslant w(s,t)^{\gamma}$ *for each* $(s,t)\in\Delta([S,T])$.

Then for every $\lambda>0$,

$$\mathbb{E}\exp\left(\lambda\int_S^T\beta(r)dr\right)\leqslant 2^{1+(2\lambda)^{1/\gamma}w(S,T)}.$$

PROOF. The former statement is an excerpt from [63, pg. 1 Lemma 1.1.], which gives the following estimate

$$\mathbb{E}\exp\left(\lambda\int_S^T\beta(r)dr\right)\leqslant\prod_{k=1}^{n}(1-\lambda\rho(t_{k-1},t_k))^{-1}. \tag{3.11}$$

[4]This result goes back at least to the paper [36] of Khasminskii, although in a less general form and with a smallness condition, then rediscovered without the smallness condition by Portenko [62], who considered (iii) with $w(s,t)=t-s$. The general version here is based on [63]. For a bit of history, see [1, pg. 214].

In the above, $S = t_0 < t_1 < \ldots < t_n = T$ are chosen so that $\sup_{k=1,\ldots,n} \lambda\rho(t_{k-1}, t_k) < 1$.

To obtain the estimate in $L_m(\Omega)$-norm, we apply Tonelli theorem and the assumption to see that

$$\begin{aligned}\mathbb{E}\left(\int_S^T \beta(r)dr\right)^m &= m!\mathbb{E}\int_{S<r_1<\ldots<r_m<T} \beta(r_1)\ldots\beta(r_m)dr_1\ldots dr_m \\ &\leqslant m!\rho(0,T)\mathbb{E}\int_{S<r_1<\ldots<r_{m-1}<T} \beta(r_1)\ldots\beta(r_{m-1})dr_1\ldots dr_{m-1}.\end{aligned}$$

Iterating the above inequality, we obtain the stated estimate for $\|\int_S^T \beta(r)dr\|_{L_m(\Omega)}$.

Under the additional condition (iii), we can choose $t_0 = S$ and for each $k \geqslant 1$,

$$t_k = \sup\{t \in [t_{k-1}, T] : \lambda w(t_{k-1}, t)^\gamma \leqslant 1/2\}.$$

With this choice, we have $\lambda w(t_{k-1}, t_k)^\gamma = 1/2$ for $k = 1, \ldots, n-1$ and $\lambda w(t_{n-1}, t_n)^\gamma \leqslant 1/2$. By definition of controls, we have

$$\frac{n-1}{(2\lambda)^{\frac{1}{\gamma}}} \leqslant \sum_{k=1}^n w(t_{k-1}, t_k) \leqslant w(S,T),$$

which yields $n \leqslant 1 + (2\lambda)^{1/\gamma} w(S,T)$. Hence, from (3.11), we have

$$\mathbb{E}\exp\left(\lambda\int_S^T \beta(r)dr\right) \leqslant 2^n \leqslant 2^{1+(2\lambda)^{1/\gamma} w(S,T)},$$

completing the proof. □

REMARK 3.6. In the setting of Lemma 3.5, if for each $(s,t) \in \Delta([S,T])$, $\rho(s,t) \leqslant w_1(s,t)^{\gamma_1} + w_2(s,t)^{\gamma_2}$ for some continuous controls w_1, w_2 and some constants $0 < \gamma_1 \leqslant \gamma_2$. Then we have

$$\mathbb{E}\exp\left(\lambda\int_S^T \beta(r)dr\right) \leqslant 2^{1+(4\lambda)^{1/\gamma_1}\left(w_1(S,T)+w_2(S,T)^{\gamma_2/\gamma_1}\right)}. \tag{3.12}$$

Indeed, the function $w = w_1 + w_2^{\gamma_2/\gamma_1}$ is a control (see [21, Excersice 1.10]) and we have $\rho(s,t) \leqslant 2w(s,t)^{\gamma_1}$. Then Lemma 3.5(iii) implies (3.12).

REMARK 3.7. In Lemma 3.5, we can assume without loss of generality that $\gamma \leqslant 1$—for otherwise, condition (3.10) implies the trivial identification $\beta \equiv 0$. Furthermore, Lemma 3.5(iii) implies that for every $\kappa > 0$ and every $\rho \in (0, \frac{1}{1-\gamma})$, with $\frac{1}{1-\gamma} = \infty$ if $\gamma = 1$, we have

$$\mathbb{E}\exp\left(\kappa\left(\int_S^T \beta(r)dr\right)^\rho\right) < \infty.$$

This follows from the same argument used in Lemma 7.4 below.

The next result is a kind of stochastic Grönwall inequality, which is of independent interest.

LEMMA 3.8 (Stochastic Grönwall inequality). *Let ξ_t, V_t be nonnegative nondecreasing processes, let A_t be a continuous nondecreasing $\mathcal{F}_t$-adapted process with $A_0 = 0$, and let M_t be $\mathcal{F}_t$-local martingale with $M_0 = 0$. Suppose that there exists a constant $\theta \in (0,\infty)$ such that with probability one,*

$$\xi_t \leqslant \left(\int_0^t \xi_s^{1/\theta} dA_s\right)^\theta + M_t + V_t, \quad \forall t \geqslant 0. \tag{3.13}$$

Then for any bounded stopping time τ, we have

$$\mathbb{E}2^{-2^{1/\theta}A_\tau}\xi_\tau \leqslant 2\mathbb{E}V_\tau \quad \textit{when} \quad \theta \leqslant 1 \tag{3.14}$$

and

$$\mathbb{E}2^{-2A_\tau^\theta}\xi_\tau \leqslant 2\mathbb{E}V_\tau \quad \textit{when} \quad \theta > 1. \tag{3.15}$$

PROOF. We put $G = M + V$ and consider two cases.
Case 1: when $\theta \leqslant 1$. Define

$$\bar{\xi}_t = \left(\int_0^t \xi_s^{1/\theta} dA_s\right)^\theta + G_t \quad \text{so that} \quad 0 \leqslant \xi_t \leqslant \bar{\xi}_t.$$

We assume first that M is a uniformly integrable martingale. For any $t \geqslant s \geqslant 0$, we have

$$\delta\bar{\xi}_{s,t} = \left(\int_0^t \xi_r^{1/\theta} dA_r\right)^\theta - \left(\int_0^s \xi_r^{1/\theta} dA_r\right)^\theta + \delta G_{s,t}.$$

We use the inequality $a^\theta - b^\theta \leqslant (a-b)^\theta$ (valid for any $a \geqslant b \geqslant 0$) to obtain from the previous identity that

$$\delta\bar{\xi}_{s,t} \leqslant \left(\int_s^t \xi_r^{1/\theta} dA_r\right)^\theta + \delta G_{s,t}. \tag{3.16}$$

Define $t_0 = 0$ and for each integer $j \geqslant 1$, the stopping time

$$t_j = \inf\{t > t_{j-1} \,:\, A_t - A_{t_j - 1} \geqslant 2^{-1/\theta}\}.$$

Let $j \geqslant 1$ be fixed. For every $t \in [t_{j-1}, t_j]$, we derive from (3.16) that

$$\delta\bar{\xi}_{t_{j-1},t} \leqslant \frac{1}{2}\xi_t + \delta G_{t_{j-1},t} \leqslant \frac{1}{2}\bar{\xi}_t + \delta G_{t_{j-1},t}$$

which yields $\bar{\xi}_t \leqslant 2\bar{\xi}_{t_{j-1}} + 2\delta G_{t_{j-1},t}$. By iteration and the fact that $\bar{\xi}_0 = V_0$, we have

$$\bar{\xi}_t \leqslant 2^j V_0 + \sum_{i=1}^{j} 2^{j-i+1}\delta G_{t_{i-1},t_i\wedge t}, \quad \forall t \in [t_{j-1}, t_j]. \tag{3.17}$$

Next, let τ be a bounded stopping time and let N be an (random) integer such that $t_N > \tau$. We have

$$2^{-2^{1/\theta}A_\tau}\bar{\xi}_\tau = \sum_{j=1}^{N} \mathbf{1}_{[t_{j-1},t_j)}(\tau) 2^{-2^{1/\theta}A_\tau}\bar{\xi}_\tau \leqslant \sum_{j=1}^{N} \mathbf{1}_{[t_{j-1},t_j)}(\tau) 2^{-(j-1)}\bar{\xi}_\tau.$$

Using (3.17), we have

$$\begin{aligned}
2^{-2^{1/\theta}A_\tau}\bar{\xi}_\tau &\leqslant 2\sum_{j=1}^{N} \mathbf{1}_{[t_{j-1},t_j)}(\tau)\left(V_0 + \sum_{i=1}^{j} 2^{1-i}\delta G_{t_{i-1},\tau\wedge t_i}\right) \\
&= 2V_0 + 2\sum_{i=1}^{N}\sum_{j=i}^{N} \mathbf{1}_{[t_{j-1},t_j)}(\tau) 2^{1-i}\delta G_{t_{i-1},\tau\wedge t_i} \\
&= 2V_0 + 2\sum_{i=1}^{N} \mathbf{1}_{[t_{i-1},t_N)}(\tau) 2^{1-i}\delta G_{t_{i-1},\tau\wedge t_i},
\end{aligned}$$

which is rewritten as

$$2^{-2^{1/\theta}A_\tau}\bar{\xi}_\tau \leqslant 2V_0 + 2\sum_{i=1}^{\infty}\mathbf{1}_{[t_{i-1},\infty)}(\tau)2^{1-i}\delta V_{\tau\wedge t_{i-1},\tau\wedge t_i} + 2\sum_{i=1}^{\infty}\mathbf{1}_{[t_{i-1},\infty)}(\tau)2^{1-i}\delta M_{\tau\wedge t_{i-1},\tau\wedge t_i}. \tag{3.18}$$

By martingale property, boundedness of τ and uniform integrability, $\mathbb{E}|\delta M_{\tau\wedge t_{i-1},\tau\wedge t_i}| \leqslant \mathbb{E}|M_{\tau\wedge t_{i-1}}| + \mathbb{E}|M_{\tau\wedge t_i}| \leqslant 2\sup_{t\geqslant 0}\mathbb{E}|M_t| < \infty$. Hence, by Fubini theorem and martingale property,

$$\mathbb{E}\sum_{i=1}^{\infty}\mathbf{1}_{[t_{i-1},\infty)}(\tau)2^{1-i}\delta M_{\tau\wedge t_{i-1},\tau\wedge t_i} = \mathbb{E}\sum_{i=1}^{\infty}\mathbf{1}_{[t_{i-1},\infty)}(\tau)2^{1-i}\mathbb{E}(\delta M_{\tau\wedge t_{i-1},\tau\wedge t_i}|\mathcal{F}_{\tau\wedge t_{i-1}}) = 0.$$

Taking expectation in (3.18) gives

$$\begin{aligned}\mathbb{E}2^{-2^{1/\theta}A_\tau}\bar{\xi}_\tau &\leqslant \mathbb{E}\left(2V_0 + 2\sum_{i=1}^{\infty}\mathbf{1}_{[t_{i-1},\infty)}(\tau)2^{1-i}\delta V_{\tau\wedge t_{i-1},\tau\wedge t_i}\right)\\ &\leqslant \mathbb{E}\left(2V_0 + 2\sum_{i=1}^{N}\delta V_{\tau\wedge t_{i-1},\tau\wedge t_i}\right) = 2\mathbb{E}V_\tau.\end{aligned}$$

In the general case when M is a local martingale, let $\{\tau_n\}$ be a sequence of increasing stopping times such that $\lim_n \tau_n = \infty$ a.s. and for each n, $M_{\tau_n\wedge\cdot}$ is a uniformly integrable martingale. For a bounded stopping time τ, the previous case implies that

$$\mathbb{E}2^{-2^{1/\theta}A_{\tau\wedge\tau_n}}\bar{\xi}_{\tau\wedge\tau_n} \leqslant 2\mathbb{E}V_{\tau\wedge\tau_n}.$$

Sending $n\to\infty$ yields (3.14).

Case 2: when $\theta > 1$. Using Hölder inequality and integration by parts

$$\left(\int_0^t \xi^{1/\theta}dA\right)^\theta \leqslant \left(\int_0^t \xi dA\right)A_t^{\theta-1} = \int_0^t\int_0^s \xi dA(\theta-1)A_s^{\theta-2}dA_s + \int_0^t A^{\theta-1}\xi dA.$$

By monotonicity, $\int_0^s \xi dA \leqslant \xi_s A_s$ so that

$$\int_0^t\int_0^s \xi dA(\theta-1)A_s^{\theta-2}dA_s \leqslant \int_0^t \xi_s(\theta-1)A_s^{\theta-1}dA_s.$$

Hence, we have

$$\left(\int_0^t \xi^{1/\theta}dA\right)^\theta \leqslant \int_0^t \xi_s\theta A_s^{\theta-1}dA_s = \int_0^t \xi_s dA_s^\theta.$$

Together with (3.13), we have

$$\xi_t \leqslant \int_0^t \xi dA^\theta + M_t + V_t, \quad \forall t\geqslant 0.$$

Using the result from the previous case, we obtain (3.15). □

REMARK 3.9. Stochastic Grönwall inequality is useful in applications to obtain moment estimates for solutions to SDEs. Starting from [64], there have been several extensions, for instance [30, 31, 53]. The setting of Lemma 3.8 is similar to that of [53], however, our result comes with less stringent hypotheses and stronger conclusions. In particular,estimates (3.14) and (3.15) hold for any $\theta\in(0,\infty)$ and do not depend on the quadratic variation of the martingale part.

LEMMA 3.10. *Let $\varepsilon > 0$, $s \in D_n$ and $r > s$. Then*

$$\int_s^r (r - k_n(\theta))^{-1-\varepsilon} d\theta \leqslant N_\varepsilon [\min(r-s, 1/n)]^{-\varepsilon} + \mathbf{1}_{(r \notin D_n)} (r - k_n(r))^{-\varepsilon}, \tag{3.19}$$

$$\int_s^r (r - k_n(\theta))^{-1} d\theta \leqslant \log(n(k_n(r) - s)) + 2, \tag{3.20}$$

$$\int_s^r (r - k_n(\theta))^{-1+\varepsilon} d\theta \leqslant N_\varepsilon (r-s)^{\varepsilon}. \tag{3.21}$$

PROOF. If $r - s \leqslant 1/n$, we have

$$\int_s^r (r - k_n(\theta))^{-1-\varepsilon} d\theta = \int_s^r (r-s)^{-1-\varepsilon} d\theta = (r-s)^{-\varepsilon}.$$

We now assume that $r - s > 1/n$. If $r \in D_n$ then

$$\int_s^r (r - k_n(\theta))^{-1-\varepsilon} d\theta \leqslant (1/n)^{-\varepsilon} \sum_{j=1}^{\infty} j^{-1-\varepsilon} = N(1/n)^{-\varepsilon}.$$

If $r \notin D_n$, then we have $r > k_n(r)$ and

$$\begin{aligned} \int_s^r (r - k_n(\theta))^{-1-\varepsilon} d\theta &= \int_s^{k_n(r)} (r - k_n(\theta))^{-1-\varepsilon} d\theta + \int_{k_n(r)}^r (r - k_n(\theta))^{-1-\varepsilon} d\theta \\ &\leqslant N(1/n)^{-\varepsilon} + (r - k_n(r))^{-\varepsilon}. \end{aligned}$$

This shows (3.19).

When $\varepsilon = 0$, we argue analogously. The only notable difference is the following estimate

$$\int_s^{k_n(r)} (r - k_n(\theta))^{-1} d\theta \leqslant \sum_{j=1}^{n(k_n(r)-s)} j^{-1} \leqslant \log(n(k_n(r) - s)) + 1.$$

This shows (3.20).

Since $r - k_n(\theta) \geqslant r - \theta$, estimate (3.21) is obvious. □

LEMMA 3.11 ([14, Proposition 2.7]). *Let $K > 0$ be a constant and let $\Sigma, \bar{\Sigma}$ be symmetric invertible matrices such that $K^{-1} I \leqslant \Sigma \bar{\Sigma}^{-1} \leqslant K I$. Then for all $x \in \mathbb{R}^d$, one has the bound*

$$|p_\Sigma(x) - p_{\bar{\Sigma}}(x)| \leqslant N \|I - \Sigma \bar{\Sigma}^{-1}\| \left(p_{\Sigma/2}(x) + p_{\bar{\Sigma}/2}(x) \right) \tag{3.22}$$

where N is a constant depending only on d, K.

4. Regularizing properties of Brownian paths—a case study. We obtain various moment estimates for the following functionals of Brownian paths

$$\int_s^t [f(r, B_r) - f(r, B_{k_n(r)})] dr \quad \text{and} \quad \int_s^t g(r, B_r) dr$$

where f, g are measurable functions in $\mathbb{L}_p^q([0,1])$. In typical applications herein, we take $f = b^n$ and $g = b - b^n$. Hence, f usually has an additional property of being in $\mathbb{L}_\infty^q([0,1])$. To extract a rate from $b - b^n$, one has to measure g with respect to a norm which is weaker than that of $\mathbb{L}_p^q$. For this purpose, we measure g in $\mathbb{L}_{-\nu,p}^q([0,1])$ for some $\nu \in [0,1]$ or in $\mathbb{L}_{p_1}^{q_1}([0,1])$ for some $q_1 \leqslant q$ and some $p_1 \leqslant p$.

In later sections, analogous functionals of the solutions to (1.1) and (1.3) will play a central role in the proofs of Theorems 2.2 and 2.3. While not being applied directly in proving the main results, the analysis in the current section are relatively simpler, mostly due to the fact that statistical properties of Brownian motion are well-understood. In addition, some of the arguments generalizes directly when B is replaced by another stochastic process which has similar analytic estimates. Therefore, we present these results at an early stage in the hope of easing out the technicalities and outlining our method. Readers who are familiar with the stochastic sewing techniques, of course, may go directly to the following sections.

Let $p_t(x) := (2\pi t)^{-d/2} e^{-|x|^2/(2t)}$ and $P_{s,t} f(x) := p_{t-s} * f(x)$.

LEMMA 4.1. *Let $p \in [1,\infty]$, $\delta \in (0,1)$, for $0 < s < t$, there exists $N = N(d,p,\delta) > 0$ such that for any $f \in L_p(\mathbb{R}^d)$ and $0 < s < t$,*

$$\|f(B_t)\|_{L_p(\Omega)} \leqslant N t^{-\frac{d}{2p}} \|f\|_{L_p(\mathbb{R}^d)} \tag{4.1}$$

and

$$\|P_{0,t} f - P_{0,s} f\|_{L_p(\mathbb{R}^d)} \leqslant N|t-s|^{\delta} |s|^{-\delta} \|f\|_{L_p(\mathbb{R}^d)}. \tag{4.2}$$

PROOF. Inequality (4.1) is taken from [14, Lemma 2.5]. We only show (4.2). First

$$\|\nabla^2 P_t f\|_{L_p(\mathbb{R}^d)} \leqslant \|\nabla^2 p_t\|_{L_1(\mathbb{R}^d)} \|f\|_{L_p(\mathbb{R}^d)} \lesssim t^{-1} \|f\|_{L_p(\mathbb{R}^d)}.$$

Then for $\delta \in (0,1)$, we have

$$\begin{aligned} \|P_{0,t} f - P_{0,s} f\|_{L_p(\mathbb{R}^d)} &\leqslant \int_s^t \|\partial_t P_{0,r} f\|_{L_p(\mathbb{R}^d)} dr = \int_s^t \|\Delta P_{0,r} f\|_{L_p(\mathbb{R}^d)} dr \\ &\lesssim \int_s^t r^{-1+\delta} r^{-\delta} dr \|f\|_{L_p(\mathbb{R}^d)} \lesssim s^{-\delta} (t-s)^{\delta} \|f\|_{L_p(\mathbb{R}^d)}, \end{aligned}$$

completing the proof. □

While not being used directly, the following result is pivotal.

PROPOSITION 4.2. *Let $f \in \mathbb{L}_p^q([0,1])$, with $p,q \in [2,\infty)$ satisfying $\frac{d}{p} + \frac{2}{q} < 1$. Then for all $2/n \leqslant S \leqslant T \leqslant 1$ and $n \in \mathbb{N}$ one has the bounds*

$$\begin{aligned} &\| \int_S^T (f(r, B_r) - f(r, B_{k_n(r)})) dr \|_{L_p(\Omega)} \\ &\qquad \leqslant N (1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([S,T])} \left[S^{-\frac{d}{2p}} |T-S|^{\frac{1}{2} - \frac{1}{q}} + S^{-\frac{d}{p}} |T-S|^{1-\frac{2}{q}} \right] \end{aligned} \tag{4.3}$$

and

$$\| \int_S^T (f(r, B_r) - f(r, B_{k_n(r)})) dr \|_{L_p(\Omega)} \leqslant N (1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([S,T])} |T-S|^{\frac{1}{2} - \frac{1}{q} - \frac{d}{2p}}. \tag{4.4}$$

PROOF. (4.4) is a direct consequence of (4.3) and Lemma 3.4. We show (4.3) below. Let $2/n \leqslant S \leqslant T \leqslant 1$ be fixed. By linearity, we can assume that $\|f\|_{\mathbb{L}_p^q([S,T])} = 1$. For $S \leqslant s \leqslant t \leqslant T$, let

$$A_{s,t} := \mathbb{E}_s \int_s^t (f(r, B_r) - f(r, B_{k_n(r)})) dr.$$

We treat two cases $t \leqslant k_n(s) + \frac{2}{n}$ and $t > k_n(s) + \frac{2}{n}$ separately as following.

Case 1. For $t \in (s, k_n(s) + \frac{2}{n}]$, by triangle inequality and (4.1) we have

$$\begin{aligned}\|A_{s,t}\|_{L_p(\Omega)} &\leqslant \int_s^t \|f(r, B_r)\|_{L_p(\Omega)} + \|f(r, B_{k_n(r)})\|_{L_p(\Omega)} dr \\ &\lesssim \int_s^t k_n(r)^{-\frac{d}{2p}} \|f(r,\cdot)\|_{L_p(\mathbb{R}^d)} dr.\end{aligned}$$

Note that $k_n(r) \geqslant k_n(s) \geqslant s/2$, applying Hölder inequality and the fact that $t - s \leqslant 2/n$, we have

$$\begin{aligned}\int_s^t k_n(r)^{-\frac{d}{2p}} \|f(r,\cdot)\|_{L_p(\mathbb{R}^d)} dr &\lesssim s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{1-\frac{1}{q}} \\ &\lesssim (1/n)^{1/2} s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{\frac{1}{2}-\frac{1}{q}}.\end{aligned}$$

This gives

$$\|A_{s,t}\|_{L_p(\Omega)} \lesssim (1/n)^{\frac{1}{2}} s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{\frac{1}{2}-\frac{1}{q}}. \tag{4.5}$$

Case 2. When $t \in (k_n(s) + \frac{2}{n}, 1]$, by triangle inequality,

$$\begin{aligned}\|A_{s,t}\|_{L_p(\Omega)} \leqslant &\int_s^{k_n(s)+\frac{2}{n}} \|\mathbb{E}_s(f(r, B_r) - f(r, B_{k_n(r)}))\|_{L_p(\Omega)} dr \\ &+ \int_{k_n(s)+\frac{2}{n}}^t \|\mathbb{E}_s(f(r, B_r) - f(r, B_{k_n(r)}))\|_{L_p(\Omega)} dr =: I_1 + I_2.\end{aligned}$$

For I_1, from (4.5) we know that

$$I_1 \lesssim (1/n)^{\frac{1}{2}} s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} \left(k_n(s) - s + \frac{2}{n}\right)^{\frac{1}{2}-\frac{1}{q}}.$$

Because $k_n(s) - s + \frac{2}{n} \leqslant t - s$, we get

$$I_1 \lesssim (1/n)^{\frac{1}{2}} s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{\frac{1}{2}-\frac{1}{q}}.$$

By (4.1) and (4.2) we have for I_2

$$\begin{aligned}I_2 &\lesssim \int_{k_n(s)+\frac{2}{n}}^t \|P_{s,r} f(r, B_s) - P_{s,k_n(r)} f(r, B_s)\|_{L_p(\Omega)} dr \\ &\lesssim \int_{k_n(s)+\frac{2}{n}}^t s^{-\frac{d}{2p}} \|P_{s,r} f(r,\cdot) - P_{s,k_n(r)} f(r,\cdot)\|_{L_p(\mathbb{R}^d)} dr \\ &\lesssim \int_{k_n(s)+\frac{2}{n}}^t s^{-\frac{d}{2p}} n^{-\frac{1}{2}} (k_n(r) - s)^{-\frac{1}{2}} \|f(r,\cdot)\|_{L_p(\mathbb{R}^d)} dr \\ &\lesssim (1/n)^{\frac{1}{2}} s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{\frac{1}{2}-\frac{1}{q}}.\end{aligned}$$

Combining these two cases together we obtain that for $S \leqslant s \leqslant t \leqslant T$,

$$\|A_{s,t}\|_{L_p(\Omega)} \lesssim (1/n)^{\frac{1}{2}} s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{\frac{1}{2}-\frac{1}{q}}. \tag{4.6}$$

Furthermore, for $u \in (s,t)$, we have $\mathbb{E}_s \delta A_{s,u,t} = 0$. Let w be the continuous control defined by

$$w(s,t) = \left[s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{\frac{1}{2}-\frac{1}{q}}\right]^2 + s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{1-\frac{1}{q}}.$$

(See Example 3.3 for a justification that w is a control.) Denote

$$\mathcal{A}_t := \int_0^t (f(r, B_r) - f(r, B_{k_n(r)}))dr, \quad J_{s,t} := \delta\mathcal{A}_{s,t} - A_{s,t}.$$

Using similar estimates leading to (4.5), we have

$$\|J_{s,t}\|_{L_p(\Omega)} \lesssim s^{-\frac{d}{2p}} \|f\|_{\mathbb{L}_p^q([s,t])} (t-s)^{1-\frac{1}{q}} \lesssim w(s,t).$$

Furthermore, $\delta J_{s,u,t} = -\delta A_{s,u,t}$ and we derive from (4.6) that

$$\|\delta J_{s,u,t}\|_{L_p(\Omega)} \lesssim (1/n)^{\frac{1}{2}} w(s,t)^{\frac{1}{2}}.$$

It is obvious that $\mathbb{E}_s J_{s,t} = 0$ and hence $\mathbb{E}_s \delta J_{s,u,t} = 0$. Applying Lemma 3.2, we have

$$\|J_{s,t}\|_{L_p(\Omega)} \lesssim (1/n)^{\frac{1}{2}} \log(n) \left[w(s,t)^{\frac{1}{2}} + w(s,t) \right]$$

for every $S \leqslant s \leqslant t \leqslant T$. By triangle inequality and (4.6), this implies that

$$\|\delta\mathcal{A}_{s,t}\|_{L_p(\Omega)} \lesssim (1/n)^{\frac{1}{2}} \log(n) \left[w(s,t)^{\frac{1}{2}} + w(s,t) \right].$$

Because $\|f\|_{\mathbb{L}_p^q([s,t])} \leqslant \|f\|_{\mathbb{L}_p^q([S,T])} = 1$ and $t - s \leqslant 1$, we have $w(s,t) \leqslant 2s^{-\frac{d}{p}}(t-s)^{1-\frac{2}{q}}$. Hence, we deduce (4.3) from the above estimate. □

In the following result, the $L_p(\Omega)$-norm in (4.4) is improved to $L_p(\Omega|\mathcal{F}_v)$-norm.

PROPOSITION 4.3. *Let $f \in \mathbb{L}_p^q([0,1])$, with $p, q \in [2, \infty)$ satisfying $\frac{d}{p} + \frac{2}{q} < 1$. Let $v \in [0, 1 - 2/n]$ be a fixed number. Then for all $v + 2/n \leqslant S \leqslant T \leqslant 1$ and all n, one has the bound*

(4.7)
$$\| \int_S^T [f(r, B_r) - f(r, B_{k_n(r)})]dr \|_{L_p(\Omega|\mathcal{F}_v)} \leqslant N(1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([S,T])} |T-S|^{\frac{1}{2}-\frac{1}{q}-\frac{d}{2p}},$$

where $N = N(p,d)$ is a constant.

PROOF. We follow the argument used in Proposition 4.2, replacing the $L_p(\Omega)$-norm by the $L_p(\Omega|\mathcal{F}_v)$-norm. The estimate (4.1) used therein (whose purpose is to deduce the analytic $L_p(\mathbb{R}^d)$-norm from the probabilistic $L_p(\Omega)$-norm) is replaced by the following estimate

$$\|g(B_t)\|_{L_p(\Omega|\mathcal{F}_v)} \leqslant N(t-v)^{-\frac{d}{2p}} \|g\|_{L_p(\mathbb{R}^d)} \quad \forall g \in L_p(\mathbb{R}^d), t > v, \tag{4.8}$$

with the same constant N as in (4.1). This yields the following estimate, which corresponds to (4.3),

$$\begin{aligned} &\| \int_S^T [f(r, B_r) - f(r, B_{k_n(r)})]dr \|_{L_p(\Omega|\mathcal{F}_v)} \\ &\quad \leqslant N(1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([S,T])} \left[(S-v)^{-\frac{d}{2p}} |T-S|^{\frac{1}{2}-\frac{1}{q}} + (S-v)^{-\frac{d}{p}} |T-S|^{1-\frac{2}{q}} \right]. \end{aligned} \tag{4.9}$$

Applying Lemma 3.4, we obtain (4.7). □

An advantage of the conditional norms over the usual ones is realized the next result.

PROPOSITION 4.4. *Let f be a Borel function in $\mathbb{L}_p^q([0,1]) \cap \mathbb{L}_\infty^q([0,1])$ for some $p, q \in [2,\infty)$ satisfying $\frac{d}{p} + \frac{2}{q} < 1$.*

We put $\beta_n(f) = \sup_{r \in D_n} \|f\|_{\mathbb{L}_\infty^q([r,r+1/n])}$. Then for any $\bar{p} \in (0,p)$, there exists a constant $N = N(d,p,q,\bar{p})$ such that

$$\| \sup_{t\in[0,1]} | \int_0^t [f(r,B_r) - f(r,B_{k_n(r)})]dr| \|_{L_{\bar{p}}(\Omega)} \leqslant N \left[(1/n)^{1-\frac{1}{q}} \beta_n(f) + (1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([0,1])} \right].$$

PROOF. Put $\mathcal{A}_t = \int_0^t (f(r,B_r) - f(r,B_{k_n(r)}))dr$ which has continuous sample paths by (4.11). In view of Lemma 3.1, it suffices to show that there exists a constant $N = N(d,p,q)$ such that

$$\|\delta\mathcal{A}_{s,t}\|_{L_p(\Omega|\mathcal{F}_s)} \leqslant N \left[(1/n)^{1-\frac{1}{q}} \beta_n(f) + (1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([0,1])} \right] \tag{4.10}$$

for every $(s,t) \in \Delta$.

Indeed, by assumption and Hölder inequality, we have

$$|\delta\mathcal{A}_{s,t}| \leqslant 2 \int_s^t \|f_r\|_{L_\infty(\mathbb{R}^d)} dr \lesssim \|f\|_{\mathbb{L}_\infty^q([s,t])} (t-s)^{1-\frac{1}{q}} \tag{4.11}$$

for every $(s,t) \in \Delta$. For every $(s,t) \in \Delta$ satisfying $t - s \geqslant 2/n$, we obtain from (4.11) and (4.7) that

$$\begin{aligned} \|\delta\mathcal{A}_{s,t}\|_{L_p(\Omega|\mathcal{F}_s)} &\leqslant \|\delta\mathcal{A}_{s,s+2/n}\|_{L_p(\Omega|\mathcal{F}_s)} + \|\delta\mathcal{A}_{s+2/n,t}\|_{L_p(\Omega|\mathcal{F}_s)} \\ &\lesssim (1/n)^{1-\frac{1}{q}} \|f\|_{\mathbb{L}_\infty^q([s,s+2/n])} + (1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([0,1])}. \end{aligned}$$

Note that $\|f\|_{\mathbb{L}_\infty^q([s,s+2/n])} \lesssim \beta_n(f)$. For every $(s,t) \in \Delta$ satisfying $t - s \leqslant 2/n$, (4.11) trivially implies that

$$\|\delta\mathcal{A}_{s,t}\|_{L_p(\Omega|\mathcal{F}_s)} \lesssim (1/n)^{1-\frac{1}{q}} \|f\|_{\mathbb{L}_\infty^q([s,s+2/n])} \lesssim (1/n)^{1-\frac{1}{q}} \beta_n(f).$$

Hence, in both cases, we have obtained (4.10). □

Next, we turn to the functional $\int_s^t g(r,B_r)dr$.

LEMMA 4.5. *Let g be a function in $\mathbb{L}_p^q([0,1])$ for some $p, q \in [1,\infty]$ satisfying $\frac{d}{p} + \frac{2}{q} < 2$. Then for every $m \geqslant 1$, there exists a constant $N = N(m,d,p,q)$ such that for every $(s,t) \in \Delta$,*

$$\| \int_s^t g(r,B_r)dr \|_{L_m(\Omega)} \leqslant N \|g\|_{\mathbb{L}_p^q([s,t])} (t-s)^{1-\frac{d}{2p}-\frac{1}{q}}.$$

PROOF. We can assume that g is nonnegative. Using standard estimate for the heat kernel and Hölder inequality, we have for every $(s,t) \in \Delta$ that

$$\begin{aligned} \mathbb{E}_s \int_s^t g(r,B_r)dr = \int_s^t P_{s,r} g(r,B_s)dr &\lesssim \int_s^t (r-s)^{-\frac{d}{2p}} \|g_r\|_{L_p(\mathbb{R}^d)} dr \\ &\lesssim \|g\|_{\mathbb{L}_p^q([s,t])} (t-s)^{1-\frac{d}{2p}-\frac{1}{q}}. \end{aligned}$$

Applying Lemma 3.5, we obtain the stated estimate for every integer $m \geqslant 1$. This and the Hölder interpolation inequality implies the estimate for any real $m \geqslant 1$. □

PROPOSITION 4.6. *Let $p \in [2,\infty)$ and $q \in [2,\infty]$. Let Γ be a positive number, w_0 be a continuous control on Δ and g be a function in $\mathbb{L}_p^q([0,1])$ such that for any $(s,t) \in \Delta$,*

$$\|g\|_{\mathbb{L}^q_{-1,p}([s,t])} \leqslant \Gamma w_0(s,t)^{\frac{1}{q}} \quad \text{and} \quad \|g\|_{\mathbb{L}^q_p([s,t])} \leqslant w_0(s,t)^{\frac{1}{q}}.$$

(a) Then for every $0 \leqslant v < S \leqslant T \leqslant 1$

$$\|\int_S^T g(r,B_r)dr\|_{L_p(\Omega|\mathcal{F}_v)} \leqslant N w_0(S,T)^{\frac{1}{q}}\Gamma(1+|\log(\Gamma)|) \times \left[(S-v)^{-\frac{d}{2p}}(T-S)^{\frac{1}{2}-\frac{1}{q}} + (S-v)^{-\frac{d}{p}}(T-S)^{1-\frac{2}{q}}\right]. \tag{4.12}$$

(b) If furthermore p,q satisfy $\frac{d}{p}+\frac{2}{q}<1$, then for any $\bar{p} \in (0,p)$, there exists a constant $N = N(d,p,q,\bar{p})$ such that

$$\|\sup_{t\in[0,1]}|\int_0^t g(r,B_r)dr|\|_{L_{\bar{p}}(\Omega)} \leqslant N\Gamma(1+|\log(\Gamma)|)w_0(0,1).$$

PROOF. (a) Let v,S,T be fixed such that $0 \leqslant v < S < T \leqslant 1$. We can assume without loss of generality that $w_0(S,T)=1$. For each $(s,t) \in \Delta([S,T])$, let

$$A_{s,t} = \mathbb{E}_s \int_s^t g(r,B_r)dr = \int_s^t P_{s,r}g(r,B_s)dr \quad \text{and} \quad J_{s,t} = \int_s^t g(r,B_r)dr - A_{s,t}.$$

In the above, we can interchange the conditional expectation and the integration due to Fubini theorem and Lemma 4.5. Define the continuous control w on $\Delta([S,T])$ by

$$w(s,t) = \left[(s-v)^{-\frac{d}{2p}}(t-s)^{\frac{1}{2}-\frac{1}{q}}w_0(s,t)^{\frac{1}{q}}\right]^2 + (s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{1}{q}}w_0(s,t)^{\frac{1}{q}}.$$

Applying Minkowski inequality, (4.1) and Hölder inequality, we have

$$\begin{aligned}\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} &\leqslant 2\int_s^t \|g(r,B_r)\|_{L_p(\Omega|\mathcal{F}_v)}dr \lesssim \int_s^t (r-v)^{-\frac{d}{2p}}\|g(r,\cdot)\|_{L_p(\mathbb{R}^d)}dr \\ &\lesssim (s-v)^{-\frac{d}{2p}}\|g\|_{\mathbb{L}^q_p([s,t])}(t-s)^{1-\frac{1}{q}} \lesssim w(s,t).\end{aligned}$$

Furthermore, $\mathbb{E}_s J_{s,t}=0$, showing that (3.1) is satisfied.

On the other hand, we have

$$\begin{aligned}\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} &\leqslant \int_s^t \|P_{s,r}g(r,B_s)\|_{L_p(\Omega|\mathcal{F}_v)}dr \\ &\lesssim \int_s^t (s-v)^{-\frac{d}{2p}}\|P_{s,r}g(r,\cdot)\|_{L_p(\mathbb{R}^d)}dr \\ &\lesssim \int_s^t (s-v)^{-\frac{d}{2p}}(r-s)^{-\frac{1}{2}}\|g(r,\cdot)\|_{L_{-1,p}(\mathbb{R}^d)}dr \\ &\lesssim (s-v)^{-\frac{d}{2p}}(t-s)^{\frac{1}{2}-\frac{1}{q}}\|g\|_{\mathbb{L}^q_{-1,p}([s,t])}.\end{aligned}$$

Combining with our assumption on g leads to $\|A_{s,t}\|_{L_p(\Omega)} \lesssim \Gamma w(s,t)^{1/2}$. Since $\delta J_{s,u,t} = -\delta A_{s,u,t}$, this implies that J satisfies (3.2). The condition (3.3) is trivial because $\mathbb{E}_s J_{s,t}=0$. Applying Lemma 3.2, we have for every $(s,t) \in \Delta([S,T])$,

$$\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim \Gamma(1+|\log\Gamma|)w(s,t)^{\frac{1}{2}} + \Gamma w(s,t)$$

and by triangle inequality,

$$\|\int_s^t g(r,B_r)dr\|_{L_p(\Omega)} \lesssim \Gamma(1+|\log\Gamma|)w(s,t)^{\frac{1}{2}} + \Gamma w(s,t).$$

Because $w_0(s,t) \leqslant 1$ and $t-s \leqslant 1$, we have $w(s,t) \leqslant 2(s-v)^{-\frac{d}{p}}(t-s)^{1-\frac{2}{q}}$. Hence, we deduce (4.12) from the above estimate by taking $(s,t)=(S,T)$.

(b) Applying Lemma 3.4 and part (a), we have

$$\|\int_s^t g(r,B_r)dr\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim \Gamma(1+|\log(\Gamma)|)w_0(0,1)$$

for every $v<s\leqslant t\leqslant 1$. In view of Lemma 4.5 and Kolmogorov continuity theorem, it is easy to see that the process $t\mapsto\int_0^t g(r,B_r)dr$ has a continuous version. Hence, the above inequality holds for every $0\leqslant v=s\leqslant t\leqslant 1$. Applying Lemma 3.1, we obtain the result. □

PROPOSITION 4.7. *Let $\nu\in[0,1)$, $p\in[2,\infty)$ and $q\in[2,\infty]$, $\frac{d}{p}+\frac{2}{q}+\nu<2$. Let g be a function in $\mathbb{L}_{\bar{p}}^q([0,1])\cap\mathbb{L}_{-\nu,p}^q([0,1])$ Then for any $\bar{p}\in(0,p)$, there exists a constant $N=N(\nu,d,p,q,\bar{p})$ such that*

$$\|\sup_{t\in[0,1]}|\int_0^t g(r,B_r)dr|\|_{L_{\bar{p}}(\Omega)} \leqslant N\|g\|_{\mathbb{L}^q_{-\nu,p}([0,1])}.$$

PROOF. In view of Lemma 3.1, it suffices to show that

$$\sup_{(s,t)\in\Delta}\|\int_s^t g(r,B_r)dr\|_{L_p(\Omega|\mathcal{F}_s)} \lesssim \|g\|_{\mathbb{L}^q_{-\nu,p}([0,1])}. \tag{4.13}$$

The proof is similar to that of Proposition 4.6, however, the control can be chosen in a simpler way. Let $v\in[0,1]$ be fixed but arbitrary. For each $(s,t)\in\Delta$, $s>v$, define $A_{s,t}$ and $J_{s,t}$ as in the proof of Proposition 4.6. As in the aforementioned proof, we have for every $v<s\leqslant u\leqslant t$, $\mathbb{E}_s\delta A_{s,u,t}=0$ and

$$\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}+\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim (s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{\nu}{2}-\frac{1}{q}}\|g\|_{\mathbb{L}^q_{-\nu,p}([s,t])}.$$

Let w be the control on $\Delta((v,1])$ defined by

$$w(s,t)=\left[(s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{\nu}{2}-\frac{1}{q}}\|g\|_{\mathbb{L}^q_{-\nu,p}([s,t])}\right]^{1/(1-\nu/2)}.$$

The previous estimate yields $\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}+\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}\lesssim w(s,t)^{1-\nu/2}$. It is evident that $\mathbb{E}_sJ_{s,t}=0$. Noting that $1-\nu/2>1/2$ and applying Lemma 3.2

$$\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim w(s,t)^{1-\frac{\nu}{2}}.$$

By triangle inequality and the previous estimate for $\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}$, we have for every $(s,t)\in\Delta$, $s>v$,

$$\|\int_s^t g(r,B_r)dr\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim w(s,t)^{1-\frac{\nu}{2}} \lesssim (s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{\nu}{2}-\frac{1}{q}}\|g\|_{\mathbb{L}^q_p([0,1])}.$$

An application of Lemma 3.4 gives

$$\|\int_s^t g(r,B_r)dr\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim \|g\|_{\mathbb{L}^q_p([0,1])}$$

for every $(s,t)\in\Delta$, $s>v$. In view of Lemma 4.5 and Kolmogorov continuity theorem, the process $t\mapsto\int_0^t g(r,B_r)dr$ has a continuous version. For this version, we see that the previous estimate holds for every $(s,t)\in\Delta$ and $v=s$, which shows (4.13). □

5. Analysis of the discrete paths. We extend the results of Section 4 to functionals of the solution to the discrete scheme (1.3).

THEOREM 5.1. *Assume that Conditions* $\mathfrak{A}\mathfrak{1}$ *and* $\mathfrak{B}$ *hold. Let* X^n *be the solution to* (1.3) *and let* $f \in \mathbb{L}_p^q([0,1]) \cap \mathbb{L}_\infty^q([0,1])$ *and* $g \in \mathbb{L}_{1,p}^q([0,1]) \cap \mathbb{L}_\infty^\infty([0,1])$. *Define* $\beta_n(f) = \sup_{t\in D_n} \|f\|_{\mathbb{L}_\infty^q([t,t+1/n])}$. *Then for any* $\bar{p} \in (0,p)$, *there exists a constant* $N = N(d,p,q,\bar{p})$ *such that*

$$\begin{aligned}(5.1)\quad &\Big\| \sup_{t\in[0,1]} \Big| \int_0^t g(r,X_r^n)[f(r,X_r^n) - f(r,X_{k_n(r)}^n)]dr \Big| \Big\|_{L_{\bar{p}}(\Omega)} \\ &\leqslant N \left[\|g\|_{\mathbb{L}_\infty^\infty([0,1])} + \|g\|_{\mathbb{L}_{1,p}^q([0,1])} \right] \\ &\quad \times \left[(1/n)^{1-\frac{1}{q}} \beta_n(f) + (1/n)^{\frac{\alpha}{2}} \|f\|_{\mathbb{L}_p^q([0,1])} + (1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}_p^q([0,1])} \right].\end{aligned}$$

The rest of the current section is devoted for the proof of Theorem 5.1. We follow the idea described in Section 4. First we derive some analytic estimates on the transition operators associated the discrete Euler–Maruyama scheme without drift. These estimates are similar to the ones in Lemma 4.1. By means of the stochastic sewing techniques (Lemma 3.2) and Girsanov theorem, these analytic estimates are utilized to obtain the desired moment bound. In what follows, we carry out this program in more detail. Conditions $\mathfrak{A}\mathfrak{1}$ and $\mathfrak{B}$ are enforced throughout the current section unless indicated otherwise.

5.1. *Analytic estimates.* For each $s \in D_n$ and $x \in \mathbb{R}^d$, let $\bar{X}^n(s,x)$ be the solution to the following Euler–Maruyama scheme

$$(5.2)\qquad \bar{X}_t^n = x + \int_s^t \sigma(r, \bar{X}_{k_n(r)}^n) dB_r, \quad t \geqslant s.$$

For each $t \geqslant s$ and bounded measurable function f, we define the function $Q_{s,t}^n f$ by

$$Q_{s,t}^n f(x) = \mathbb{E} f(\bar{X}_t^n(s,x)).$$

Let

$$(5.3)\qquad T_{s,t}^* f(y) = \mathbb{E} f\left(y + \int_s^t \sigma(r,y) dB_r \right),$$

and let the operator $T_{s,t}$ be conjugate to $T_{s,t}^*$ in L_2-sense, which can be computed explicitly

$$(5.4)\qquad T_{s,t} f(x) = \int_{\mathbb{R}^d} f(y) p_{\Sigma_{s,t}(y)}(y-x) dy, \quad \text{when} \quad s < t$$

and $T_{s,s} f(x) = f(x)$, where $\Sigma_{s,t}^{ij} = \frac{1}{2}\int_s^t a_r^{ij} dr$, $a_r = \sigma_r \sigma_r^*$. Whenever $s < t$, the function $T_{s,t} f(x)$ is infinitely differentiable and satisfies

$$\partial_s T_{s,t} f(x) = -\partial_{x_i x_j}^2 T_{s,t}[a_s^{ij} f](x).$$

We also define for every $r < t$

$$\eta_r(x) = \int_{k_n(r)}^r \sigma(v,x) dB_v$$

and

$$H_{r,t}^n f(x) = \mathbb{E}\left[a_r^{ij}(x) (\partial_{x_i x_j}^2 T_{r,t} f)(x + \eta_r(x)) - (\partial_{x_i x_j}^2 T_{r,t}[a_r^{ij} f])(x + \eta_r(x)) \right].$$

The function η also depends on n, however, we omit this dependence in the notation. By direct computations (see also [25, p.153] or [26, p.11]), we have

$$H^n_{r,t}f(x) = \int_{\mathbb{R}^d} K^n_{r,t}(x,y)f(y)dy$$

where

$$K^n_{r,t}(x,y) = [a^{ij}_r(x) - a^{ij}_r(y)]\partial^2_{ij}p_{\Sigma_{k_n(r),r}(x)+\Sigma_{r,t}(y)}(y-x)$$
$$= [a^{ij}_r(x) - a^{ij}_r(y)][(A_{r,t}(x,y)z)^i(A_{r,t}(x,y)z)^j - A^{ij}_{r,t}(x,y)]p_{\Sigma_{k_n(r),r}(x)+\Sigma_{r,t}(y)}(z)\Big|_{z=y-x}$$

and

$$A_{r,t}(x,y) = (\Sigma_{k_n(r),r}(x) + \Sigma_{r,t}(y))^{-1}.$$

The relation between Q^n and T, H^n is described in Lemma 5.4, which is a kind of Duhamel formula. It is, however, convenient to obtain first analytic estimates for H^n and T. We make use of the following result, inspired by [25, Lemma 4.1].[5]

LEMMA 5.2. *Let $\lambda, \ell, \varepsilon > 0$, $\alpha \in [0,1]$ be fixed numbers, let a_1 be a symmetric $d \times d$ matrix and let $a(x), \bar{a}(x)$ be $d \times d$ matrix-valued functions. Assume that for each x, $a(x), \bar{a}(x)$ are symmetric, $\lambda^{-1}\ell I \leqslant a(x) + a_1, \bar{a}(x) + a_1 \leqslant \lambda\ell I$, where I is the $d \times d$ unit matrix, and $\sup_y \|\bar{a}(y) - a(y)\| \leqslant \varepsilon$. Let $g(x)$ be a real function such that $|g(x) - g(y)| \leqslant \lambda|x-y|^\alpha$ for all x, y. Let ξ and η be independent d-dimensional Gaussian vectors with zero means. Assume that ξ has distribution $\mathcal{N}(0, I)$ and η has distribution $\mathcal{N}(0, \sqrt{a_1})$. Define an operator T^* by the formula $T^*f(y) = \mathbb{E}f(y + \sqrt{a(y)}\xi)$ and let T be the conjugate for T^* in L_2-sense. Let $1 \leqslant i, j \leqslant d$ be fixed and define*

$$H(x) = \mathbb{E}\left[g(x)(\partial^2_{x_ix_j}Tf)(x+\eta) - (\partial^2_{x_ix_j}T[gf])(x+\eta)\right].$$

Define $\bar{T}^*, \bar{T}$ and $\bar{H}(x)$ analogously with $\bar{a}$ replacing a. Then for any $p \in [1, \infty]$ and bounded Borel f,

(5.5)
$$\sup_{x\in\mathbb{R}^d} |H(x)| \leqslant N\|f\|_{L_p(\mathbb{R}^d)}\ell^{\frac{\alpha}{2}-1-\frac{d}{2p}}, \quad \|H\|_{L_p(\mathbb{R}^d)} \leqslant N\|f\|_{L_p(\mathbb{R}^d)}\ell^{\frac{\alpha}{2}-1},$$

(5.6)
$$|\bar{H}(x) - H(x)| \leqslant N\|f\|_{L_p(\mathbb{R}^d)}\varepsilon\ell^{\frac{\alpha}{2}-2-\frac{d}{2p}} \quad \text{and} \quad \|\bar{H} - H\|_{L_p(\mathbb{R}^d)} \leqslant N\|f\|_{L_p(\mathbb{R}^d)}\varepsilon\ell^{\frac{\alpha}{2}-2},$$

where the constant N depends only on λ, d, p, α.

PROOF. By direct computations (see also [25, p.153] or [26, p.11]), we have

$$H(x) = \int_{\mathbb{R}^d} [g(x) - g(y)]f(y)[(A(y)z)^i(A(y)z)^j - A^{ij}(y)]p_{a(y)+a_1}(z)\Big|_{z=y-x}dy,$$

where $A(y) = (a(y) + a_1)^{-1}$. A similar formula for $\bar{H}(x)$ is valid with $\bar{A}(y) = (\bar{a}(y) + a_1)^{-1}$. By ellipticity of $a(y) + a_1$ and Hölder continuity of g, we have

$$|H(x)| \lesssim \int_{\mathbb{R}^d} |x-y|^\alpha |f(y)| \left(\ell^{-2}|y-x|^2 + \ell^{-1}\right) p_{2\lambda\ell}(y-x)dy.$$

[5] [26] recently corrected [25, Lemma 4.1]. Our assumptions and conclusions are different from both works.

Setting $q = \frac{p}{p-1}$ and applying Hölder inequality, we get

$$|H(x)| \lesssim \|f\|_{L_p(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} |y|^{q\alpha} (\ell^{-2}|y|^2 + \ell^{-1})^q p_{2\lambda\ell}(y)^q dy \right)^{\frac{1}{q}} \lesssim \|f\|_{L_p(\mathbb{R}^d)} \ell^{\frac{\alpha}{2}-1-\frac{d}{2p}}.$$

This shows the first estimate in (5.5). The second estimate in (5.5) is obtained by a similar argument, using instead the Minkowski inequality.

It is evident that $\|A(y)\|, \|\bar{A}(y)\| \lesssim \ell^{-1}$ uniformly in y. For two invertible matrices C, D, we have $C^{-1} - D^{-1} = C^{-1}(D - C)D^{-1}$ so that $\|C^{-1} - D^{-1}\| \leqslant \|C^{-1}\|\|D^{-1}\|\|D - C\|$. Thus $\|\bar{A}(y) - A(y)\| \leqslant \|A(y)\|\|\bar{A}(y)\|\|\bar{a}(y) - a(y)\| \lesssim \ell^{-2}\varepsilon$ and similarly

$$\|(\bar{A}(y)z)^i(\bar{A}(y)(y-x))^j - (A(y)z)^i(A(y)(y-x))^j\| \lesssim |z|^2\ell^{-3}\varepsilon.$$

Using our assumptions, it is straightforward to verify that $K^{-1}I \leqslant (a(y)+a_1)(\bar{a}(y)+a_1)^{-1} \leqslant KI$ and $\|I - (a(y) + a_1)(\bar{a}(y) + a_1)^{-1}\| \leqslant K\ell^{-1}\varepsilon$ for some constant $K > 0$. Hence, from (3.22), we have

$$\begin{aligned} |p_{\bar{a}(y)+a_1}(z) - p_{a(y)+a_1}(z)| &\lesssim \ell^{-1}\varepsilon \left(p_{(\bar{a}(y)+a_1)/2}(z) + p_{(a(y)+a_1)/2}(z) \right) \\ &\lesssim \ell^{-1}\varepsilon p_{\lambda\ell}(z), \end{aligned}$$

where we have used the fact that $\lambda^{-1}t \leqslant a(y) + a_1, \bar{a}(y) + a_1 \leqslant \lambda t$. It follows that

$$|\bar{H}(x) - H(x)| \lesssim \varepsilon \int_{\mathbb{R}^d} |f(y)||y-x|^\alpha \left(|y-x|^2\ell^{-3} + \ell^{-2} \right) p_{\lambda\ell}(y-x)dy.$$

From here, we apply Hölder inequality and Minkowski inequality as previously to obtain (5.6). □

LEMMA 5.3. *Let $p_1, p_2 \in [1, \infty]$ be such that $p_1 \leqslant p_2$ and let f be a function in $L_{p_1}(\mathbb{R}^d)$. For every $r < t \leqslant 1$, we have*

$$\|T_{r,t}f\|_{L_{p_2}(\mathbb{R}^d)} \leqslant N(t-r)^{\frac{d}{2p_2}-\frac{d}{2p_1}} \|f\|_{L_{p_1}(\mathbb{R}^d)} \tag{5.7}$$

and

$$\|H^n_{r,t}f\|_{L_{p_2}(\mathbb{R}^d)} \leqslant N(t-k_n(r))^{\frac{\alpha}{2}-1+\frac{d}{2p_2}-\frac{d}{2p_1}} \|f\|_{L_{p_1}(\mathbb{R}^d)}, \tag{5.8}$$

where the constant N depends only on d, p_1, p_2, K_1, K_2.

PROOF. By uniform ellipticity, there exists a constant $\lambda > 0$ such that for every $x, y \in \mathbb{R}^d$,

$$\lambda^{-1}(t - k_n(r)) \leqslant \int_r^t a_\theta(y)d\theta + \int_{k_n(r)}^r a_\theta(x)d\theta \leqslant \lambda(t - k_n(r)). \tag{5.9}$$

From here, we can derive (5.7) using standard Gaussian estimates.

Applying (5.5), we have

$$\|H^n_{r,t}f\|_{L_\infty(\mathbb{R}^d)} \lesssim (t-k_n(r))^{\frac{\alpha}{2}-1-\frac{d}{2p_1}} \|f\|_{L_{p_1}(\mathbb{R}^d)}$$

and

$$\|H^n_{r,t}f\|_{L_{p_1}(\mathbb{R}^d)} \lesssim (t-k_n(r))^{\frac{\alpha}{2}-1} \|f\|_{L_{p_1}(\mathbb{R}^d)}.$$

From the above estimates and the Hölder interpolation inequality

$$\|H\|_{L_{p_2}(\mathbb{R}^d)} \leqslant \|H\|^{\frac{p_1}{p_2}}_{L_{p_1}(\mathbb{R}^d)} \|H\|^{1-\frac{p_1}{p_2}}_{L_\infty(\mathbb{R}^d)},$$

we deduce (5.8). □

The estimate (5.8) shows in particular that whenever $r<t$, $H^n_{r,t}$ maps bounded measurable functions to bounded measurable functions. It is evident from their definitions that $Q^n_{r,t}, T_{r,t}$ also map bounded measurable functions to bounded measurable functions.

LEMMA 5.4. *Let $s\in D_n$ and f be a bounded uniformly continuous function. Then, for every $t>s$ and $x\in\mathbb{R}^d$*

$$Q^n_{s,t}f(x)=T_{s,t}f(x)+\int_s^t Q^n_{s,k_n(r)}[H^n_{r,t}f](x)dr. \tag{5.10}$$

PROOF. Let $\bar{X}^n=\bar{X}^n(s,x)$ and $\tau\in(s,t)$. Applying Itô formula for $r\mapsto T_{r,t}f(\bar{X}^n_r)$, for any $t>s$, we obtain that

$$\begin{aligned}&\mathbb{E}T_{\tau,t}f(\bar{X}^n_\tau)\\&\quad=T_{s,t}f(\bar{X}^n_s)+\mathbb{E}\int_s^\tau\left[a^{ij}_r(\bar{X}^n_{k_n(r)})(\partial^2_{x_ix_j}T_{r,t}f)(\bar{X}^n_r)-(\partial^2_{x_ix_j}T_{r,t}[a^{ij}_rf])(\bar{X}^n_r)\right]dr.\end{aligned}$$

Writing $\bar{X}^n_r=\bar{X}^n_{k_n(r)}+\eta_r(\bar{X}^n_{k_n(r)})$, we take conditional expectation given $\mathcal{F}_{k_n(r)}\supset\mathcal{F}_s$. This yields

$$\mathbb{E}T_{\tau,t}f(\bar{X}^n_\tau)=T_{s,t}f(\bar{X}^n_s)+\int_s^\tau\mathbb{E}H^n_{r,t}f(\bar{X}^n_{k_n(r)})dr. \tag{5.11}$$

We now take the limit $\tau\uparrow t$ in the above formula. By uniform continuity of f and a.s. continuity of $\bar{X}^n$, $\lim_{\tau\uparrow t}T_{\tau,t}f(\bar{X}^n_\tau)=f(\bar{X}^n_t)$. From (5.8), we have

$$\int_\tau^t|\mathbb{E}H^n_{r,t}f(\bar{X}^n_{k_n(r)})|dr\lesssim\int_\tau^t(t-k_n(r))^{\frac{\alpha}{2}-1}\|f\|_{L_\infty(\mathbb{R}^d)}dr\lesssim\|f\|_{L_\infty(\mathbb{R}^d)}(t-\tau)^{\frac{\alpha}{2}},$$

which allows one to apply the limit $\tau\uparrow t$ to the last term in (5.11). Hence, we have

$$\mathbb{E}f(\bar{X}^n_t)=T_{s,t}f(\bar{X}^n_s)+\int_s^t\mathbb{E}H^n_{r,t}f(\bar{X}^n_{k_n(r)})dr,$$

which deduces to (5.10). □

THEOREM 5.5. *Assume that Condition 2.1 holds. Let $p_1,p_2\in[1,\infty]$, $p_1\leqslant p_2$, $p_1<\infty$ and let f be a function in $L_{p_1}(\mathbb{R}^d)$. There exists a constant $N=N(d,p_1,p_2,\alpha,K_1,K_2)$ such that for every $s\in D_n$ and $t\in(s,1]$, we have*

$$\|Q^n_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)}\leqslant N(t-s)^{\frac{d}{2p_2}-\frac{d}{2p_1}}\|f\|_{L_{p_1}(\mathbb{R}^d)}. \tag{5.12}$$

PROOF. We put $\rho=\frac{d}{2p_1}-\frac{d}{2p_2}$.

Step 1. We show some rough estimates for $\|Q^n_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)}$ in terms of $\|f\|_{L_{p_1}(\mathbb{R}^d)}$. Assume first that f is a bounded uniformly continuous function. From (5.10) and Lemma 5.3, we have for every $t\in[s,s+1/n]$

$$\begin{aligned}\|Q^n_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)}&\leqslant\|T_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)}+\int_s^t\|H^n_{r,t}f\|_{L_{p_2}(\mathbb{R}^d)}dr\\&\lesssim(t-s)^{-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)}+\int_s^t(t-k_n(r))^{\frac{\alpha}{2}-1-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)}dr\\&\lesssim(t-s)^{-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)},\end{aligned}$$

where the last inequality follows from the fact that $k_n(r) = s$ for $r \in [s, s+1/n)$. Since smooth functions are dense in $L_{p_1}(\mathbb{R}^d)$, it follows that that $\|Q^n_{s,t}\|_{L_{p_2}(\mathbb{R}^d)} \lesssim (t-s)^{-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)}$ for any function f in $L_{p_1}(\mathbb{R}^d)$.

We proceed inductively. Let $j \geqslant 1$ be an integer. Suppose that for every $t \in [s, s+j/n]$ and every function $f \in L_{p_1}(\mathbb{R}^d)$,

$$\|Q^n_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)} \leqslant C_j(t-s)^{-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)} \tag{5.13}$$

for some constant C_j, independent from n, s, t, f.

Let f be a bounded uniformly continuous function. Then for each $t \in (s+j/n, s+(j+1)/n]$, we obtain from (5.10), Lemma 5.3 and the inductive hypothesis that

$$\begin{aligned}\|Q^n_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)} \lesssim \|T_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)} + \int_s^{s+1/n} \|H^n_{r,t}f\|_{L_{p_2}(\mathbb{R}^d)}dr \\ + C_j\int_{s+1/n}^t (k_n(r)-s)^{-\rho}\|H^n_{r,t}f\|_{L_{p_1}(\mathbb{R}^d)}dr.\end{aligned} \tag{5.14}$$

The first two terms are estimated as previously, $\|T_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)} \lesssim (t-s)^{-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)}$ and

$$\int_s^{s+1/n} \|H^n_{r,t}f\|_{L_{p_2}(\mathbb{R}^d)}dr \lesssim (1/n)^{-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)}.$$

Using Lemma 5.3, we have

$$\|H^n_{r,t}f\|_{L_{p_1}(\mathbb{R}^d)}) \lesssim (t-k_n(r))^{\frac{\alpha}{2}-1}\|f\|_{L_{p_1}(\mathbb{R}^d)}.$$

Using the above estimate and the fact that $k_n(r) - s \geqslant 1/n$ for any $r \geqslant s+1/n$ and Lemma 3.10, we have

$$\begin{aligned}\int_{s+1/n}^t (k_n(r)-s)^{-\rho}\|H^n_{r,t}f\|_{L_{p_1}(\mathbb{R}^d)}dr &\lesssim (1/n)^{-\rho}\int_{s+1/n}^t (t-k_n(r))^{\frac{\alpha}{2}-1}dr\|f\|_{L_{p_1}(\mathbb{R}^d)} \\ &\lesssim (1/n)^{-\rho}\|f\|_{L_{p_1}(\mathbb{R}^d)}.\end{aligned}$$

Observe furthermore that $(1/n)^{-\rho} \leqslant (j+1)^\rho(t-s)^{-\rho}$. Putting these estimates in (5.14), we see that (5.13) holds for $t \in (s+j/n, s+(j+1)/n]$ with a bounded uniformly continuous function f and some constant C_{j+1}. By approximation, we can extend the inequality to all functions $f \in L_{p_1}(\mathbb{R}^d)$.

To show that (5.13) actually holds uniformly in j, we proceed as follows.

Step 2. Assuming that $\rho = \frac{d}{2p_2} - \frac{d}{2p_1} < 1$, we show that there exists a constant $C > 0$, independent from n, such that

$$\|Q^n_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)} \leqslant C(t-s)^{\frac{d}{2p_2}-\frac{d}{2p_1}}\|f\|_{L_{p_1}(\mathbb{R}^d)} \text{ for every } t > s \text{ and function } f \in L_{p_1}(\mathbb{R}^d). \tag{5.15}$$

In view of (5.13), it suffices to consider $t \geqslant s+2/n$. Let $\lambda > 1$ be a constant to be chosen later. For each $t \in [s+2/n, 1]$, define

$$m_t = e^{-\lambda(t-s)}(t-s)^\rho \sup_{g \in L_{p_1}(\mathbb{R}^d)\,:\,\|g\|_{L_{p_1}(\mathbb{R}^d)}=1} \|Q^n_{s,t}g\|_{L_{p_2}(\mathbb{R}^d)}$$

and $m^*_t = \sup_{r\in[s+2/n,t]} m_r$, which are finite by the previous step. In particular, for every $t \geqslant s+2/n$ and $g \in L_{p_1}(\mathbb{R}^d)$, we have (the case $\|g\|_{L_{p_1}(\mathbb{R}^d)} = 0$ is treated by (5.13))

$$\|Q^n_{s,t}g\|_{L_{p_2}(\mathbb{R}^d)} \leqslant m^*_t e^{\lambda(t-s)}(t-s)^{-\rho}\|g\|_{L_{p_1}(\mathbb{R}^d)}. \tag{5.16}$$

Let $t \geqslant s + 2/n$ and f be a bounded uniformly continuous function, $\|f\|_{L_{p_1}(\mathbb{R}^d)} = 1$. From (5.10), (5.13) and (5.16), we have that

$$\|Q^n_{s,t} f\|_{L_{p_2}(\mathbb{R}^d)} \lesssim \|T_{s,t} f\|_{L_{p_2}(\mathbb{R}^d)} + \int_s^{s+2/n} \|H^n_{r,t} f\|_{L_{p_2}(\mathbb{R}^d)} dr + m^*_t \int_{s+2/n}^t e^{\lambda(k_n(r)-s)} (k_n(r) - s)^{-\rho} \|H^n_{r,t} f\|_{L_{p_1}(\mathbb{R}^d)} dr. \tag{5.17}$$

From Lemma 5.3, we have $\|T_{s,t} f\|_{L_{p_2}(\mathbb{R}^d)} \lesssim (t-s)^{-\rho}$ and

$$\int_s^{s+2/n} \|H^n_{r,t} f\|_{L_{p_2}(\mathbb{R}^d)} dr \lesssim \int_s^{s+2/n} (t - k_n(r))^{\frac{\alpha}{2}-1-\rho} dr \lesssim (1/n)(t - s - 1/n)^{\frac{\alpha}{2}-1-\rho} \lesssim (t-s)^{-\rho},$$

where we have used the fact that $1/n \leqslant t - s \leqslant 1$ and $t - s - 1/n \geqslant (t-s)/2$. Similarly, using Lemma 5.3, we have

$$\begin{aligned} &\int_{s+2/n}^t e^{\lambda(k_n(r)-s)} (k_n(r) - s)^{-\rho} \|H^n_{r,t} f\|_{L_{p_1}(\mathbb{R}^d)} dr \\ &\lesssim \int_{s+2/n}^t e^{\lambda(k_n(r)-s)} (k_n(r) - s)^{-\rho} (t - k_n(r))^{\frac{\alpha}{2}-1} dr \\ &\lesssim \int_{s+1/n}^t e^{\lambda(r-s)} (r - s - 1/n)^{-\rho} (t - r)^{\frac{\alpha}{2}-1} dr. \end{aligned}$$

To estimate the integrals on the right-hand sides above, we split them into two regions, putting $\bar{s} = s + 1/n$,

$$\int_{\bar{s}}^{(\bar{s}+t)/2} e^{-\lambda(t-r)} (r - \bar{s})^{-\rho} (t - r)^{\frac{\alpha}{2}-1} dr \lesssim e^{-\frac{\lambda}{2}(t-\bar{s})} (t - \bar{s})^{\frac{\alpha}{2}-1} \int_{\bar{s}}^{(\bar{s}+t)/2} (r - \bar{s})^{-\rho} dr \lesssim e^{-\frac{\lambda}{2}(t-\bar{s})} (t - \bar{s})^{\frac{\alpha}{2}-\rho} \lesssim \lambda^{-\frac{\alpha}{2}} (t - \bar{s})^{-\rho}$$

and

$$\int_{(\bar{s}+t)/2}^t e^{-\lambda(t-r)} (r - \bar{s})^{-\rho} (t - r)^{\frac{\alpha}{2}-1} dr \leqslant (t - \bar{s})^{-\rho} \int_{(\bar{s}+t)/2}^t e^{-\lambda(t-r)} (t - r)^{\frac{\alpha}{2}-1} dr \leqslant (t - \bar{s})^{-\rho} \int_0^\infty e^{-\lambda u} u^{\frac{\alpha}{2}-1} du \lesssim \lambda^{-\frac{\alpha}{2}} (t - \bar{s})^{-\rho}.$$

In the above, all integrals are finite because $\alpha \in (0, 1]$ and $\rho < 1$. Observe furthermore that $(t-s)/(t-\bar{s}) \leqslant 2$. Thus we have

$$\int_{s+2/n}^t e^{\lambda(k_n(r)-s)} (k_n(r) - s)^{-\rho} (t - k_n(r))^{\frac{\alpha}{2}-1} dr \lesssim \lambda^{-\frac{\alpha}{2}} e^{\lambda(t-s)} (t-s)^{-\rho}.$$

Putting the previous estimates altogether into (5.17), we have

$$\|Q^n_{s,t} f\|_{L_{p_2}(\mathbb{R}^d)} \lesssim (t-s)^{-\rho} (1 + m^*_t \lambda^{-\frac{\alpha}{2}} e^{\lambda(t-s)}).$$

By approximations, the above estimate also holds for any function $f \in L_{p_1}(\mathbb{R}^d)$ with $\|f\|_{L_{p_1}(\mathbb{R}^d)} = 1$. It follows that $m_t \lesssim 1 + \lambda^{-\frac{\alpha}{2}} m^*_t$ for every $t \geqslant s + 2/n$. By choosing λ

sufficiently large, we conclude that m_1^* is bounded by a constant independent from n and thus obtain (5.15).

Step 3. We remove the restriction $\rho < 1$ in the previous step.

Suppose that $\rho := \frac{d}{2p_1} - \frac{d}{2p_2} < 2$. Define $p_3 \in [p_1, p_2]$ by $\frac{d}{p_3} = \frac{d}{2p_1} + \frac{d}{2p_2}$ so that $\frac{d}{2p_1} - \frac{d}{2p_3} = \frac{d}{2p_3} - \frac{d}{2p_2} = \frac{\rho}{2} < 1$. Let $t \geqslant s + 4/n$ and $u = (s+t)/2$. Then by Markov property of the Euler–Maruyama scheme, we have $Q^n_{s,t} = Q^n_{k_n(u),t} Q^n_{s,k_n(u)}$. It is easy to see that $t > k_n(u) > s$ so that by (5.15), we have for every $f \in L_{p_1}(\mathbb{R}^d)$ that

$$\begin{aligned}\|Q^n_{s,t} f\|_{L_{p_2}(\mathbb{R}^d)} &\lesssim (t - k_n(u))^{-\frac{\rho}{2}} \|Q^n_{s,k_n(u)} f\|_{L_{p_3}(\mathbb{R}^d)} \\ &\lesssim (t - k_n(u))^{-\frac{\rho}{2}} (k_n(u) - s)^{-\frac{\rho}{2}} \|f\|_{L_{p_1}(\mathbb{R}^d)}.\end{aligned}$$

It is straightforward to see that $t - k_n(u) \geqslant (t-s)/2$ and $k_n(u) - s \geqslant (t-s)/4$. Hence, from the above estimate, we deduce that $\|Q^n_{s,t} f\|_{L_{p_2}(\mathbb{R}^d)} \lesssim (t-s)^{-\rho} \|f\|_{L_{p_1}(\mathbb{R}^d)}$ for any $t \geqslant s + 4/n$. Combining with (5.13) from step 1., we see that (5.15) holds for any $p_1, p_2 \in [1, \infty]$ satisfying $\frac{d}{2p_1} - \frac{d}{2p_2} < 2$. We iterate the argument. After $\lfloor \log_2(d/2) \rfloor + 1$ iterations, we see that (5.15) holds whenever $\rho \leqslant d/2$, which is trivially satisfied for any $p_1, p_2 \in [1, \infty]$. Hence, we have shown (5.12). □

REMARK 5.6. Theorem 5.5 complements previous works. It is shown in [48] that for each $s, t \in D_n$, $s < t$, the operator $Q^n_{s,t}$ has a kernel density which has Gaussian upper bounds. From here, one can deduce estimate (5.12) for discrete times $s, t \in D_n$. Gaussian upper bounds for the kernel density of $Q^n_{s,t}$, which are valid for all times $t > s$, $s \in D_n$, are also established in [6, 35] under Lipschitz regularity of a. Some related estimates are also obtained in [25, 26] under the additional condition $\alpha > d/p$. We are able to remove this condition herein mainly due to (5.8), which was known previously with the factor $(t - r)$ on its right-hand side.

LEMMA 5.7. *Let $p_1 \in [1, \infty]$ and let $r < k_n(t)$. Then for every $f \in L_{p_1}(\mathbb{R}^d)$,*

$$\|T_{r,t} f - T_{r,k_n(t)} f\|_{L_{p_1}(\mathbb{R}^d)} \leqslant N(1/n)(t-r)^{-1} \|f\|_{L_{p_1}(\mathbb{R}^d)} \tag{5.18}$$

and

$$\|H^n_{r,t} f - H^n_{r,k_n(t)} f\|_{L_{p_1}(\mathbb{R}^d)} \leqslant N(1/n)(t - k_n(r))^{\frac{\alpha}{2} - 2} \|f\|_{L_{p_1}(\mathbb{R}^d)}, \tag{5.19}$$

where the constant N depends only on d, p_1, K_1, K_2.

PROOF. We make use of (5.9). Observe furthermore that $1/2 \leqslant (k_n(t) - k_n(r))/(t - k_n(r)) \leqslant 1$. It is then straightforward to verify the hypothesis of Lemma 3.11 for $\Sigma = \Sigma_{r,t}(y)$ and $\bar{\Sigma} = \Sigma_{r,k_n(t)}(y)$. In addition

$$\|I - \Sigma \bar{\Sigma}^{-1}\| \lesssim (t - k_n(t))(t - k_n(r))^{-1} \lesssim (1/n)(t - k_n(r))^{-1}.$$

From the definition of T, we apply (3.22) to get that

$$|T_{r,t} f(x) - T_{r,k_n(t)} f(x)| \lesssim (1/n)(t-r)^{-1} \int_{\mathbb{R}^d} p_{c(t-r)}(y) |f(y - x)| dy$$

for some universal constant c. From here, we apply Minkowski inequality to obtain (5.18). (5.19) is obtained analogously, using (5.6). □

COROLLARY 5.8. *Assuming Condition 2.1. Let f be a function in $L_{p_1}(\mathbb{R}^d)$, $p_1 \in [1, \infty)$. Then for any $s \in D_n$ and $t \in (s + 2/n, 1]$,*

$$\|Q^n_{s,t} f - Q^n_{s,k_n(t)} f\|_{L_{p_1}(\mathbb{R}^d)} \leqslant N(1/n)^{\frac{\alpha}{2}} \|f\|_{L_{p_1}(\mathbb{R}^d)} (t-s)^{-\frac{\alpha}{2}},$$

where the constant N depends only on d, p_1, K_1, K_2.

PROOF. By approximation, we can assume that f is bounded and uniformly continuous. We put $u = k_n(t)$. From (5.10), $Q^n_{s,t}f - Q^n_{s,u}f = I_1 + I_2 + I_3$, where

$$I_1 = T_{s,t}f - T_{s,u}f,$$
$$I_2 = \int_s^u Q^n_{s,k_n(r)}[H^n_{r,t}f - H^n_{r,u}f]dr, \quad I_3 = \int_u^t Q^n_{s,k_n(r)}[H^n_{r,t}f]dr.$$

We estimate each I_1, I_2, I_3 below. Note that $s < t - 1/n$ implies $s < k_n(t)$. Applying (5.18), we have $\|I_1\|_{L_p(\mathbb{R}^d)} \lesssim (1/n)(t-s)^{-1}\|f\|_{L_p(\mathbb{R}^d)}$.

From Theorem 5.5, we have $\|Q^n_{s,k_n(r)}f\|_{L_p(\mathbb{R}^d)} \lesssim \|f\|_{L_p(\mathbb{R}^d)}$ for every $r > s$. It follows that

$$\|I_2\|_{L_p(\mathbb{R}^d)} \lesssim \int_s^u \|H^n_{r,t}f - H^n_{r,u}f\|_{L_p(\mathbb{R}^d)}dr.$$

Applying (5.19) and Lemma 3.10 (noting that $u - s \geqslant 1/n$), we have

$$\|I_2\|_{L_p(\mathbb{R}^d)} \lesssim \int_s^u (1/n)(u - k_n(r))^{\frac{\alpha}{2}-2}dr\|f\|_{L_p(\mathbb{R}^d)} \lesssim (1/n)^{\frac{\alpha}{2}}\|f\|_{L_p(\mathbb{R}^d)}.$$

Applying Theorem 5.5 and Lemma 5.3, we have

$$\|I_3\|_{L_p(\mathbb{R}^d)} \lesssim \int_u^t \|H^n_{r,t}f\|_{L_p(\mathbb{R}^d)}dr \lesssim \int_u^t (t - k_n(r))^{\frac{\alpha}{2}-1}\|f\|_{L_p(\mathbb{R}^d)}dr \lesssim (1/n)^{\frac{\alpha}{2}}\|f\|_{L_p(\mathbb{R}^d)}.$$

Combining the previous estimates, we obtain the result. □

COROLLARY 5.9. *Let g be a function in $L_{1,p}(\mathbb{R}^d)$ for some $p \in [1, \infty)$*

$$\|Q^n_{s,t}g - g\|_{L_p(\mathbb{R}^d)} \lesssim \|g\|_{L_{1,p}(\mathbb{R}^d)}(t-s)^{\frac{\alpha}{2}}. \tag{5.20}$$

PROOF. By approximation, we can assume g is continuously differentiable and has bounded derivatives. From (5.10), we have $Q^n_{s,t}g - g = I_1 + I_2$, where $I_1 = T_{s,t}g - g$ and $I_2 = \int_s^t Q^n_{s,k_n(r)}[H^n_{r,t}g]dr$. From (5.4), we have

$$\begin{aligned} I_1 &= \int_{\mathbb{R}^d} [p_{\Sigma_{s,t}(y)}(y-x)g(y) - p_{\Sigma_{s,t}(x)}(y-x)g(x)]dy \\ &= \int_{\mathbb{R}^d} [p_{\Sigma_{s,t}(y)}(y-x) - p_{\Sigma_{s,t}(x)}(y-x)]g(y)dy + \int_{\mathbb{R}^d} [g(y) - g(x)]p_{\Sigma_{s,t}(x)}(y-x)dy \\ &= I_{11} + I_{12}. \end{aligned}$$

Using Condition 2.1, it is straightforward to verify that $\lambda^{-1}I \leqslant \Sigma_{s,t}(y)\Sigma_{s,t}(x)^{-1} \leqslant \lambda I$ and $\|I - \Sigma_{s,t}(x)\Sigma_{s,t}(y)^{-1}\| \leqslant \lambda|x-y|^\alpha$ for some finite constant λ. Then, we apply Lemma 3.11 to get

$$|I_{11}| \lesssim \int_{\mathbb{R}^d} |x-y|^\alpha p_{N(t-s)}(x-y)|g(y)|dy.$$

Using Minkowski inequality, we obtain from the above that $\|I_{11}\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{\frac{\alpha}{2}}\|g\|_{L_p(\mathbb{R}^d)}$. From the Hardy–Littlewood maximal inequality, there is a non-negative function $h \in L_p(\mathbb{R}^d)$ such that $\|h\|_{L_p(\mathbb{R}^d)} \lesssim \|\nabla g\|_{L_p(\mathbb{R}^d)}$ and

$$|g(y) - g(x)| \leqslant |x-y|(h(x) + h(y)), \quad \text{a.e.} \quad x, y \in \mathbb{R}^d.$$

Using ellipticity and the above estimate, we have that

$$|I_{12}| \leqslant \int_{\mathbb{R}^d} |x-y|(h(x)+h(y))p_{\Sigma_{s,t}(x)}(x-y)dy$$
$$\lesssim h(x)(t-s)^{\frac{1}{2}} + \int_{\mathbb{R}^d} |x-y|h(y)p_{N(t-s)}(x-y)dy.$$

Applying Minkowski inequality, we obtain that $\|I_{12}\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{1/2}\|h\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{1/2}\|\nabla g\|_{L_p(\mathbb{R}^d)}$. Applying Theorem 5.5 and Lemma 5.3, we have

$$\|I_2\|_{L_p(\mathbb{R}^d)} \lesssim \int_s^t \|H^n_{r,t}g\|_{L_p(\mathbb{R}^d)}dr \lesssim \int_s^t (t-r)^{\frac{\alpha}{2}-1}\|g\|_{L_p(\mathbb{R}^d)}dr \lesssim (t-s)^{\frac{\alpha}{2}}\|g\|_{L_p(\mathbb{R}^d)}.$$

Combining the estimates for I_{11}, I_{12} and I_2, we obtain the result. □

5.2. *Moment estimates.* We consider the Euler–Maruyama scheme

$$\bar{X}^n_t = x + \int_0^t \sigma(s, \bar{X}^n_{k_n(s)})dB_s, \tag{5.21}$$

where x_0 is a $\mathcal{F}_0$-random variable. By Markov property, for every $s \in D_n$ and bounded measurable f, we have $\mathbb{E}[f(\bar{X}^n_t)|\mathcal{F}_s] = Q^n_{s,t}f(\bar{X}^n_s)$.

PROPOSITION 5.10. *Let $\bar{X}^n$ be the solution to (5.21).*

(i) Let h be a measurable function such that $\|h\|_{L_\rho(\mathbb{R}^d)}$ is finite for some $\rho \in (0,\infty]$. Then for every $r, v \in [0,1]$, $r - v \geqslant 2/n$,

$$\|h(\bar{X}^n_r)\|_{L_\rho(\Omega|\mathcal{F}_v)} \leqslant N\|h\|_{L_\rho(\mathbb{R}^d)}(r-v)^{-\frac{d}{2\rho}}. \tag{5.22}$$

(ii) Let f be a function in $L_p(\mathbb{R}^d)$, g be a function in $L_{1,p}(\mathbb{R}^d) \cap L_\infty(\mathbb{R}^d)$ for some $p \in [1,\infty)$. Then, for every $r, s, v \in [0,1]$ such that $r - v \geqslant 2/n$, $r > k_n(s) + 3/n$ and $s - v \geqslant 2/n$,

$$\|\mathbb{E}_s(f(\bar{X}^n_r) - f(\bar{X}^n_{k_n(r)}))\|_{L_p(\Omega|\mathcal{F}_v)} \leqslant N(1/n)^{\frac{\alpha}{2}}\|f\|_{L_p(\mathbb{R}^d)}(s-v)^{-\frac{d}{2p}}\left(r-s-\frac{2}{n}\right)^{-\frac{\alpha}{2}} \tag{5.23}$$

and

$$\|\mathbb{E}_s(g(\bar{X}^n_r)f(\bar{X}^n_r) - g(\bar{X}^n_r)f(\bar{X}^n_{k_n(r)}))\|_{L_p(\Omega|\mathcal{F}_v)} \leqslant N(1/n)^{\frac{\alpha}{2}}\|f\|_{L_p(\mathbb{R}^d)}(s-v)^{-\frac{d}{2p}}$$
$$\times\left[\|g\|_{L_\infty(\mathbb{R}^d)}\left(r-s-\frac{2}{n}\right)^{-\frac{\alpha}{2}} + \|g\|_{L_{1,p}(\mathbb{R}^d)}\left(r-s-\frac{2}{n}\right)^{-\frac{d}{2p}}\right]. \tag{5.24}$$

PROOF. (i) Put $\bar{v} = k_n(v) + 1/n$. In the case when $\rho < \infty$, applying Theorem 5.5 (with the choice $p_1 = 1$ and $p_2 = \infty$), we have

$$\mathbb{E}[|h(\bar{X}^n_r)|^\rho|\mathcal{F}_{\bar{v}}] = Q^n_{\bar{v},r}[|h|^\rho](\bar{X}^n_{\bar{v}}) \lesssim (r-\bar{v})^{-\frac{d}{2}}\||h|^\rho\|_{L_1(\mathbb{R}^d)}.$$

Noting that $r - \bar{v} \geqslant (r-v)/2$, we obtain (5.22) for any $\rho \in (0,\infty)$ from the above. When $\rho = \infty$, (5.22) is trivial.

(ii) Put $\bar{s} = k_n(s) + 1/n$. Applying (5.22) and Corollary 5.8 (noting that $\bar{s} - v \geqslant 2/n$ and $r - \bar{s} > 2/n$), we have

$$\|\mathbb{E}_{\bar{s}}(f(\bar{X}^n_r) - f(\bar{X}^n_{k_n(r)}))\|_{L_p(\Omega|\mathcal{F}_v)} = \|Q^n_{\bar{s},r}f(\bar{X}^n_{\bar{s}}) - Q^n_{\bar{s},k_n(r)}f(\bar{X}^n_{\bar{s}})\|_{L_p(\Omega|\mathcal{F}_v)}$$
$$\lesssim (\bar{s}-v)^{-\frac{d}{2p}}\|Q^n_{\bar{s},r}f - Q^n_{\bar{s},k_n(r)}f\|_{L_p(\mathbb{R}^d)}$$
$$\lesssim (1/n)^{\frac{\alpha}{2}}\|f\|_{L_p(\mathbb{R}^d)}(\bar{s}-v)^{-\frac{d}{2p}}(k_n(r)-\bar{s})^{-\frac{\alpha}{2}}.$$

We observe that $\bar{s}-v \geqslant s-v$, $k_n(r)-\bar{s} \geqslant r-s-2/n$ and

$$\|\mathbb{E}_s(f(\bar{X}^n_r)-f(\bar{X}^n_{k_n(r)}))\|_{L_p(\Omega|\mathcal{F}_v)} \leqslant \|\mathbb{E}_{\bar{s}}(f(\bar{X}^n_r)-f(\bar{X}^n_{k_n(r)}))\|_{L_p(\Omega|\mathcal{F}_v)}.$$

From here, we obtain (5.23) by combining the previous estimates.

Lastly, we show (5.24). We write

$$\begin{aligned} &g(\bar{X}^n_r)f(\bar{X}^n_r)-g(\bar{X}^n_r)f(\bar{X}^n_{k_n(r)}) \\ &=[(gf)(\bar{X}^n_r)-(gf)(\bar{X}^n_{k_n(r)})]+[(g(\bar{X}^n_{k_n(r)})-g(\bar{X}^n_r))f(\bar{X}^n_{k_n(r)})]. \end{aligned}$$

We observe that $\|fg\|_{L_p(\mathbb{R}^d)} \leqslant \|f\|_{L_p(\mathbb{R}^d)}\|g\|_{L_\infty(\mathbb{R}^d)}$ and apply (5.23) to see that

$$\left\|\mathbb{E}_s\left((fg)(\bar{X}^n_r)-(fg)(\bar{X}^n_{k_n(r)})\right)\right\|_{L_p(\Omega|\mathcal{F}_v)}$$

is smaller than the right-hand side of (5.24). It suffices to estimate the $L_p(\Omega|\mathcal{F}_v)$-norm of

$$A := \mathbb{E}_{\bar{s}}[(g(\bar{X}^n_{k_n(r)})-g(\bar{X}^n_r))f(\bar{X}^n_{k_n(r)})].$$

By conditioning on $\mathcal{F}_{k_n(r)}$, we have

$$A=\mathbb{E}_{\bar{s}}[h(\bar{X}^n_{k_n(r)})]=Q^n_{\bar{s},k_n(r))}h(\bar{X}^n_{\bar{s}}), \quad \text{where} \quad h=(g-Q^n_{k_n(r),r}g)f.$$

Applying (5.22),

$$\|A\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim (\bar{s}-v)^{-\frac{d}{2p}}\|Q^n_{\bar{s},k_n(r)}h\|_{L_p(\mathbb{R}^d)}.$$

We continue by applying Theorem 5.5 (with $p_1=p/2$ and $p_2=p$),

$$\|A\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim (\bar{s}-v)^{-\frac{d}{2p}}(k_n(r)-\bar{s})^{-\frac{d}{2p}}\|h\|_{L_{p/2}(\mathbb{R}^d)}.$$

By Hölder inequality

$$\|h\|_{L_{p/2}(\mathbb{R}^d)} \leqslant \|Q_{k_n(r),r}g-g\|_{L_p(\mathbb{R}^d)}\|f\|_{L_p(\mathbb{R}^d)}$$

and by (5.20),

$$\|Q_{k_n(r),r}g-g\|_{L_p(\mathbb{R}^d)} \lesssim (r-k_n(r))^{\frac{\alpha}{2}}\|g\|_{L_{1,p}(\mathbb{R}^d)}.$$

Combining the previous estimates, we see that $\|A\|_{L_p(\Omega|\mathcal{F}_v)}$ is also smaller than the right-hand side of (5.24), finishing the proof. □

REMARK 5.11. Concerning Proposition 5.10(i), if $v \in D_n$, then there exists a constant $N=N(d,\rho,\alpha,K_1,K_2)$ such that for every $r>v$, we have

$$\|h(\bar{X}^n_r)\|_{L_\rho(\Omega|\mathcal{F}_v)} \leqslant N\|h\|_{L_\rho(\mathbb{R}^d)}(r-v)^{-\frac{d}{2\rho}}. \tag{5.25}$$

Indeed, the inequality is trivial when $\rho=\infty$. When $\rho<\infty$ and when $v \in D_n$, Theorem 5.5 is applied directly, which yields

$$\mathbb{E}[|h(\bar{X}^n_r)|^\rho|\mathcal{F}_v]=Q^n_{v,r}[|h|^\rho](\bar{X}^n_v) \lesssim (r-v)^{-\frac{d}{2}}\,\||h|^\rho\|_{L_1(\mathbb{R}^d)}.$$

From here, we deduce (5.25).

PROPOSITION 5.12. *Let $\bar{X}^n$ be the solution to (5.21). Let $f \in \mathbb{L}^q_p([0,1])$ and $g \in \mathbb{L}^q_{1,p}([0,1]) \cap \mathbb{L}^\infty_\infty([0,1])$, with $p,q \in [2,\infty)$ satisfying $\frac{d}{p}+\frac{2}{q}<1$. Let $v \in [0,1-4/n]$ be a fixed number, $n \geqslant 4$ is an integer. Then for every $v+\frac{4}{n} \leqslant S \leqslant T \leqslant 1$, one has the bound*

$$\begin{aligned} &\|\int_S^T g(r,\bar{X}^n_r)[f(r,\bar{X}^n_r)-f(r,\bar{X}^n_{k_n(r)})]dr\|_{L_p(\Omega|\mathcal{F}_v)} \\ &\qquad \leqslant N[(1/n)^{\frac{\alpha}{2}}+(1/n)^{\frac{1}{2}}\log(n)](\|g\|_{\mathbb{L}^\infty_\infty([S,T])}+\|g\|_{\mathbb{L}^q_{1,p}([S,T])})\|f\|_{\mathbb{L}^q_p([S,T])}, \end{aligned} \tag{5.26}$$

where $N=N(d,p,q,K_1,K_2)$ is a constant.

PROOF. Let S,T be such that $v+4/n\leqslant S\leqslant T\leqslant 1$. By linearity, we can assume that $\|f\|_{\mathbb{L}_p^q([S,T])}=\|g\|_{\mathbb{L}_\infty^\infty([S,T])}+\|g\|_{\mathbb{L}_{1,p}^q([S,T])}=1$.

For each $(s,t)\in\Delta_2([S,T])$, define

$$A_{s,t}:=\mathbb{E}_s\int_s^t g(r,\bar{X}_r^n)(f(r,\bar{X}_r^n)-f(r,\bar{X}_{k_n(r)}^n))dr.$$

We treat two cases $t\leqslant k_n(s)+\frac{4}{n}$ and $t>k_n(s)+\frac{4}{n}$ separately as following.

Case 1. For $t\in(s,k_n(s)+\frac{4}{n}]$, by triangle inequality and (5.22) (note that $k_n(r)-v\geqslant k_n(s)-v\geqslant s-v-1/n\geqslant 2/n$) we have

$$\begin{aligned}\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}&\leqslant\|g\|_{\mathbb{L}_\infty^\infty([S,T])}\int_s^t\|f(r,\bar{X}_r^n)\|_{L_p(\Omega|\mathcal{F}_v)}+\|f(r,\bar{X}_{k_n(r)}^n)\|_{L_p(\Omega|\mathcal{F}_v)}dr\\&\lesssim\int_s^t(k_n(r)-v)^{-\frac{d}{2p}}\|f(r,\cdot)\|_{L_p(\mathbb{R}^d)}dr.\end{aligned}$$

Note that $k_n(r)-v\geqslant k_n(s)-v\geqslant(s-v)/2$, applying Hölder inequality and the fact that $t-s\leqslant 4/n$, we have

$$\begin{aligned}\int_s^t(k_n(r)-v)^{-\frac{d}{2p}}\|f(r,\cdot)\|_{L_p(\mathbb{R}^d)}dr&\lesssim(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{1}{q}}\\&\lesssim(1/n)^{\frac{1}{2}}(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{\frac{1}{2}-\frac{1}{q}}.\end{aligned}$$

This gives

$$\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}\lesssim(1/n)^{\frac{1}{2}}(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{\frac{1}{2}-\frac{1}{q}}.\tag{5.27}$$

Case 2. When $t\in(k_n(s)+\frac{4}{n},1]$, by triangle inequality,

$$\begin{aligned}\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}&\leqslant\|A_{s,k_n(s)+4/n}\|_{L_p(\Omega|\mathcal{F}_v)}\\&\quad+\int_{k_n(s)+\frac{4}{n}}^t\|\mathbb{E}_s[g(r,\bar{X}_r^n)(f(r,\bar{X}_r^n)-f(r,\bar{X}_{k_n(r)}^n))]\|_{L_p(\Omega|\mathcal{F}_v)}dr\\&=:I_1+I_2.\end{aligned}$$

For I_1, from (5.27) we know that

$$I_1\lesssim(1/n)^{\frac{1}{2}}(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}\left(k_n(s)+\frac{4}{n}-s\right)^{\frac{1}{2}-\frac{1}{q}}.$$

Because $k_n(s)+\frac{4}{n}-s\leqslant t-s$, we get

$$I_1\lesssim(1/n)^{\frac{1}{2}}(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{\frac{1}{2}-\frac{1}{q}}.$$

Applying (5.24) and Hölder inequality, we have for I_2

$$\begin{aligned}I_2&\lesssim(1/n)^{\frac{\alpha}{2}}(s-v)^{-\frac{d}{2p}}\|g\|_{\mathbb{L}_\infty^\infty([S,T])}\int_{k_n(s)+\frac{4}{n}}^t\left(r-s-\frac{2}{n}\right)^{-\frac{\alpha}{2}}\|f(r,\cdot)\|_{L_p(\mathbb{R}^d)}dr\\&\quad+(1/n)^{\frac{\alpha}{2}}(s-v)^{-\frac{d}{2p}}\int_{k_n(s)+\frac{4}{n}}^t\left(r-s-\frac{2}{n}\right)^{-\frac{d}{2p}}\|g(r,\cdot)\|_{L_{1,p}(\mathbb{R}^d)}\|f(r,\cdot)\|_{L_p(\mathbb{R}^d)}dr\\&\lesssim(1/n)^{\frac{\alpha}{2}}(s-v)^{-\frac{d}{2p}}\left[\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{\alpha}{2}-\frac{1}{q}}+\|g\|_{\mathbb{L}_{1,p}^q([s,t])}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{d}{2p}-\frac{2}{q}}\right].\end{aligned}$$

Combining these two cases together we obtain that for $v+4/n \leqslant s \leqslant t \leqslant 1$,

$$\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim (1/n)^{\frac{1}{2}}(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{\frac{1}{2}-\frac{1}{q}} \tag{5.28}$$
$$+(1/n)^{\frac{\alpha}{2}}(s-v)^{-\frac{d}{2p}}\left[\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{\alpha}{2}-\frac{1}{q}}+\|g\|_{\mathbb{L}_{1,p}^q([s,t])}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{d}{2p}-\frac{2}{q}}\right].$$

Furthermore, for $u \in (s,t)$, we have $\mathbb{E}_s\delta A_{s,u,t}=0$. Let w be the continuous control on $\Delta([v+4/n,1])$ defined by

$$\begin{aligned} w(s,t) = &\left[(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{\frac{1}{2}-\frac{1}{q}}\right]^2 \\ &+\left[(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{\alpha}{2}-\frac{1}{q}}\right]^{1/(1-\frac{\alpha}{2})} \\ &+\left[(s-v)^{-\frac{d}{2p}}\|g\|_{\mathbb{L}_{1,p}^q([s,t])}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{d}{2p}-\frac{2}{q}}\right]^{1/(1-\frac{d}{2p})} \\ &+(s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{1}{q}}. \end{aligned}$$

Denote

$$\mathcal{A}_t := \int_0^t (f(r,\bar{X}_r^n)-f(r,\bar{X}_{k_n(r)}^n))dr, \quad J_{s,t} := \delta\mathcal{A}_{s,t}-A_{s,t}.$$

Using similar estimates leading to (5.27), we have

$$\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim (s-v)^{-\frac{d}{2p}}\|f\|_{\mathbb{L}_p^q([s,t])}(t-s)^{1-\frac{1}{q}} \lesssim w(s,t).$$

Furthermore, $\delta J_{s,u,t}=-\delta A_{s,u,t}$ and we derive from (5.28) that

$$\|\delta J_{s,u,t}\|_{L_p(\Omega|\mathcal{F}_v)} \lesssim (1/n)^{\frac{1}{2}}w(s,t)^{\frac{1}{2}}+(1/n)^{\frac{\alpha}{2}}\left[w(s,t)^{1-\frac{\alpha}{2}}+w(s,t)^{1-\frac{d}{2p}}\right].$$

It is obvious that $\mathbb{E}_s J_{s,t}=0$ and hence $\mathbb{E}_s\delta J_{s,u,t}=0$. Applying Lemma 3.2, we have

$$\begin{aligned} &\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} \\ &\quad\lesssim [(1/n)^{\frac{\alpha}{2}}+(1/n)^{\frac{1}{2}}\log(n)]\left[w(s,t)^{\frac{1}{2}}+w(s,t)^{1-\frac{\alpha}{2}}+w(s,t)^{1-\frac{d}{2p}}+w(s,t)\right]. \end{aligned}$$

By triangle inequality and (5.28), this implies that

$$\begin{aligned} &\|\delta\mathcal{A}_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)} \\ &\quad\lesssim [(1/n)^{\frac{\alpha}{2}}+(1/n)^{\frac{1}{2}}\log(n)]\left[w(s,t)^{\frac{1}{2}}+w(s,t)^{1-\frac{\alpha}{2}}+w(s,t)^{1-\frac{d}{2p}}+w(s,t)\right]. \end{aligned}$$

Because $\|f\|_{\mathbb{L}_p^q([s,t])} \leqslant \|f\|_{\mathbb{L}_p^q([S,T])}=1$, we have

$$\begin{aligned} w(s,t) \leqslant &\left[(s-v)^{-\frac{d}{2p}}(t-s)^{\frac{1}{2}-\frac{1}{q}}\right]^2 \\ &+\left[(s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{\alpha}{2}-\frac{1}{q}}\right]^{1/(1-\frac{\alpha}{2})}+\left[(s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{d}{2p}-\frac{2}{q}}\right]^{1/(1-\frac{d}{2p})} \\ &+(s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{1}{q}} \end{aligned} \tag{5.29}$$

and hence

$$w(s,t)^{\frac{1}{2}}+w(s,t)^{1-\frac{\alpha}{2}}+w(s,t)^{1-\frac{d}{2p}}+w(s,t) \lesssim \sum_{i=1}^{16}(s-v)^{-\eta_i}(t-s)^{\tau_i},$$

where for each $i = 1, \dots, 16$; $\eta_i, \tau_i \in [0,1]$ are some constants such that $\tau_i - \eta_i > 0$. The constants η_i, τ_i's can be calculated explicitly by applying the powers $1/2$, $1-\alpha/2$, $1-d/(2p)$ and 1 to the singularity exponents and Hölder exponents in the right-hand side of (5.29), however, their exact values are non-essential. That $\tau_i - \eta_i$ is positive for each i because the sums of the Hölder exponents and the corresponding singular exponents of each factor on the right-hand side of (5.29) are positive. Hence, we deduce from the above estimate that

$$\| \int_S^T g(r, \bar{X}^n_r)(f(r, \bar{X}^n_r) - f(r, \bar{X}^n_{k_n(r)}))dr \|_{L_p(\Omega|\mathcal{F}_v)} \\ \leqslant N[(1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}} \log(n)] \sum_{i=1}^{16} (S-v)^{-\eta_i} (T-S)^{\tau_i},$$

which holds for every $v + 4/n \leqslant S \leqslant T \leqslant 1$. We then apply Lemma 3.4 to obtain (5.26). □

PROPOSITION 5.13. *Let $\bar{X}^n$ be the solution to (5.21). Let $f \in \mathbb{L}^q_p([0,1]) \cap \mathbb{L}^q_\infty([0,1])$ and $g \in \mathbb{L}^q_{1,p}([0,1]) \cap \mathbb{L}^\infty_\infty([0,1])$, with $p, q \in [2, \infty)$ satisfying $\frac{d}{p} + \frac{2}{q} < 1$.*

As in Proposition 4.4, we put $\beta_n(f) = \sup_{r \in D_n} \|f\|_{\mathbb{L}^q_\infty([r, r+1/n])}$. Then for any $\bar{p} \in (0, p)$, there exists a constant $N = N(d, p, q, \bar{p})$ such that

$$\| \sup_{t \in [0,1]} | \int_0^t g(r, \bar{X}^n_r)[f(r, \bar{X}^n_r) - f(r, \bar{X}^n_{k_n(r)})]dr | \|_{L_{\bar{p}}(\Omega)} \leqslant N \left[\|g\|_{\mathbb{L}^\infty_\infty([0,1])} + \|g\|_{\mathbb{L}^q_{1,p}([0,1])} \right] \\ \times \left[(1/n)^{1-\frac{1}{q}} \beta_n(f) + (1/n)^{\frac{\alpha}{2}} \|f\|_{\mathbb{L}^q_p([0,1])} + (1/n)^{\frac{1}{2}} \log(n) \|f\|_{\mathbb{L}^q_p([0,1])} \right].$$

PROOF. This result is a consequence of Proposition 5.12 and Lemma 3.1. The proof is analogous to that of Proposition 4.4, hence, omitted. □

LEMMA 5.14. *Let $\bar{X}^n$ be the solution to (5.21) and f be a function in $\mathbb{L}^{q_1}_{p_1}([0,1])$ for some $p_1, q_1 \in [1, \infty]$ satisfying $\frac{d}{p_1} + \frac{2}{q_1} < 2$. Then*

$$\mathbb{E} \exp \left(\int_0^1 f(r, \bar{X}^n_r) dr \right) \leqslant 2 \exp \left(N \|f\|^{1/(1-d/(2p_1))}_{\mathbb{L}^{q_1}_{p_1}([0,1])} \right), \tag{5.30}$$

where N depends only on $d, \alpha, p_1, q_1, K_1, K_2$.

Assume additionally that there are continuous control w_0 on Δ and positive constants M, γ_0 such that

$$(1/n)^{1-\frac{1}{q_1}} \|f\|_{\mathbb{L}^{q_1}_\infty([s,t])} \leqslant w_0(s,t)^{\gamma_0} \quad \forall\, 0 \leqslant t - s \leqslant 1/n \tag{5.31}$$

and

$$\|f\|_{\mathbb{L}^{q_1}_{p_1}([0,1])} + w_0(0,1) \leqslant M.$$

Then there exists a finite constant $\bar{N}$ which depends only on $M, \gamma_0, d, \alpha, p_1, q_1, K_1, K_2$ such that

$$\mathbb{E} \exp \left(\int_0^1 f(r, \bar{X}^n_{k_n(r)}) dr \right) \leqslant \bar{N}. \tag{5.32}$$

PROOF. We can assume without loss of generality that f is nonnegative. We rely on Lemma 3.5. For each $(s,t) \in \Delta$, define $\bar{s} = k_n(s) + 1/n$. For each $r \in (s, \bar{s})$, we write

$\bar{X}^n_r = \bar{X}^n_s + \int_s^r \sigma(\theta, \bar{X}^n_{k_n(s)})dB_\theta$ so that

$$\mathbb{E}_s \int_s^{\bar{s}\wedge t} f(r, \bar{X}^n_r)dr = \int_s^{\bar{s}\wedge t} P_{\Sigma_{s,r}(\bar{X}^n_{k_n(s)})} f_r(\bar{X}^n_s)dr.$$

Using ellipticity of $\Sigma_{s,r}$ and Gaussian estimates, we see that $\sup_y \|P_{\Sigma_{s,r}(y)} f_r\|_{L_\infty(\mathbb{R}^d)} \lesssim (r-s)^{-d/(2p_1)}\|f_r\|_{L_{p_1}(\mathbb{R}^d)}$. Hence, using this estimate and Hölder inequality, we have

$$\mathbb{E}_s \int_s^{\bar{s}\wedge t} f(r, \bar{X}^n_r)dr \lesssim \int_s^t (r-s)^{-\frac{d}{2p_1}} \|f_r\|_{L_{p_1}(\mathbb{R}^d)}dr \lesssim \|f\|_{\mathbb{L}^{q_1}_{p_1}([s,t])}(t-s)^{1-\frac{d}{2p_1}-\frac{1}{q_1}}.$$

On the interval $(\bar{s}\wedge t, t)$, we use (5.25) and Hölder inequality to see that

$$\mathbb{E}_s \int_{\bar{s}\wedge t}^t f(r, \bar{X}^n_r)dr \lesssim \int_{\bar{s}\wedge t}^t (r-\bar{s})^{-\frac{d}{2p_1}} \|f_r\|_{L_{p_1}(\mathbb{R}^d)}dr \lesssim \|f\|_{\mathbb{L}^{q_1}_{p_1}([s,t])}(t-s)^{1-\frac{d}{2p_1}-\frac{1}{q_1}}.$$

It follows that

$$\mathbb{E}_s \int_s^t f(r, \bar{X}^n_r)dr \lesssim \|f\|_{\mathbb{L}^{q_1}_{p_1}([s,t])}(t-s)^{1-\frac{d}{2p_1}-\frac{1}{q_1}}.$$

Observe that w defined by $w(s,t)^{1-\frac{d}{2p_1}} = \|f\|_{\mathbb{L}^{q_1}_{p_1}([s,t])}(t-s)^{1-\frac{d}{2p_1}-\frac{1}{q_1}}$ is a continuous control on Δ. Applying Lemma 3.5, we obtain (5.30).

The second part is obtained in an analogous way. For each $(s,t)\in\Delta$, define $\tilde{s} = k_n(s) + 2/n$, using Hölder inequality, (5.31) and (5.22), we have

$$\mathbb{E}_s \int_s^{\tilde{s}\wedge t} f(r, \bar{X}^n_{k_n(r)})dr \leqslant \int_s^{\tilde{s}\wedge t} \|f_r\|_{L_\infty(\mathbb{R}^d)}dr \lesssim \|f\|_{\mathbb{L}^{q_1}_\infty([s,\tilde{s}\wedge t])}(1/n)^{1-\frac{1}{q_1}} \lesssim w_0(s,t)^{\gamma_0}$$

and

$$\begin{aligned}\mathbb{E}_s \int_{\tilde{s}\wedge t}^t f(r, \bar{X}^n_{k_n(r)})dr &= \int_{\tilde{s}\wedge t}^t \mathbb{E}_s f(r, \bar{X}^n_{k_n(r)})dr \lesssim \int_{\tilde{s}\wedge t}^t (k_n(r)-s)^{-\frac{d}{2p_1}} \|f_r\|_{L_{p_1}(\mathbb{R}^d)}dr \\ &\lesssim \int_{\tilde{s}\wedge t}^t (r-s)^{-\frac{d}{2p_1}} \|f_r\|_{L_{p_1}(\mathbb{R}^d)}dr \lesssim \|f\|_{\mathbb{L}^{q_1}_{p_1}([s,t])}(t-s)^{1-\frac{d}{2p_1}-\frac{1}{q_1}}.\end{aligned}$$

Hence, $\mathbb{E}_s \int_s^t f(r, \bar{X}^n_{k_n(r)})dr \lesssim w_0(s,t)^{\gamma_0} + w(s,t)^{1-\frac{d}{2p_1}}$, where w is the control defined previously. Applying Lemma 3.5 and Remark 3.6, we obtain (5.32). □

REMARK 5.15. From Remark 3.6, one can compute $\bar{N}$ explicitly, however (5.32) is sufficient for our considerations.

Proof of Theorem 5.1. For any continuous process Z, we define

$$\mathcal{H}(Z) = \sup_{t\in[0,1]} \Big| \int_0^t g(r, Z_r)[f(r, Z_r) - f(r, Z_{k_n(r)})]dr \Big|.$$

Let $\bar{X}^n$ be the solution to (5.21) and

$$\rho := \exp\left(-\int_0^1 (\sigma^{-1}b^n)(r, \bar{X}^n_{k_n(r)})dB_r - \frac{1}{2}\int_0^1 \left|(\sigma^{-1}b^n)(r, \bar{X}^n_{k_n(r)})\right|^2 dr \right).$$

Using the fact that $\sigma^{-1}b^n \in \mathbb{L}^q_\infty([0,1])$, we see that ρ is a probability density. It follows from Girsanov theorem (see e.g. [33, IV Corollary of Theorem 4.2]) and Hölder inequality for $\frac{1}{\gamma'} + \frac{1}{\gamma} = 1$ with $\gamma > 1$ close enough to 1 such that $\gamma\bar{p} < p$

$$\mathbb{E}\mathcal{H}(X^n)^{\bar{p}} = \mathbb{E}(\rho\mathcal{H}(\bar{X}^n)^{\bar{p}}) \leqslant [\mathbb{E}\mathcal{H}(\bar{X}^n)^{\gamma\bar{p}}]^{1/\gamma}[\mathbb{E}\rho^{\gamma'}]^{1/\gamma'}. \tag{5.33}$$

From Proposition 5.13 we immediately get that

$$(5.34)\quad (\mathbb{E}\mathcal{H}(\bar{X}^n)^{\gamma\bar{p}})^{\frac{1}{\gamma\bar{p}}} = \|\mathcal{H}(\bar{X}^n)\|_{L_{\gamma\bar{p}}(\Omega)} \leqslant N\left[\|g\|_{\mathbb{L}^\infty_\infty([0,1])} + \|g\|_{\mathbb{L}^q_{1,p}([0,1])}\right] \times \left[(1/n)^{1-\frac{1}{q}}\beta_n(f) + (1/n)^{\frac{\alpha}{2}}\|f\|_{\mathbb{L}^q_p([0,1])} + (1/n)^{\frac{1}{2}}\log(n)\|f\|_{\mathbb{L}^q_p([0,1])}\right].$$

Using Cauchy–Schwarz inequality, we have

$$\begin{aligned}\mathbb{E}\rho^{\gamma'} &= \mathbb{E}\exp\left(-\gamma'\int_0^1 (\sigma^{-1}b^n)(r,\bar{X}^n_{k_n(r)})dB_r - \frac{\gamma'}{2}\int_0^1 \left|(\sigma^{-1}b^n)(r,\bar{X}^n_{k_n(r)})\right|^2 dr\right)\\ &\leqslant \left[\mathbb{E}\exp\left(-2\gamma'\int_0^1 (\sigma^{-1}b^n)(r,\bar{X}^n_{k_n(r)})dB_r - 2\gamma'^2\int_0^1 \left|(\sigma^{-1}b^n)(r,\bar{X}^n_{k_n(r)})\right|^2 dr\right)\right]^{\frac{1}{2}}\\ &\quad\times\left[\mathbb{E}\exp\left((2\gamma'^2-\gamma')\int_0^1 \left|(\sigma^{-1}b^n)(r,\bar{X}^n_{k_n(r)})\right|^2 dr\right)\right]^{\frac{1}{2}}.\end{aligned}$$

In the right-hand side above, the first factor is identical to 1 by martingale properties. For the second factor, we recall Condition $\mathfrak{B}$ and the uniform ellipticity of σ, which imply that the function $f := |\sigma^{-1}b^n|^2$ belongs to $\mathbb{L}^{q/2}_{p/2}([0,1]) \cap \mathbb{L}^{q/2}_\infty([0,1])$ and satisfies

$$(1/n)^{1-\frac{2}{q}}\|f\|_{\mathbb{L}^{q/2}_\infty([s,t])} \lesssim \left[(1/n)^{\frac{1}{2}-\frac{1}{q}}\|b^n\|_{\mathbb{L}^q_\infty([s,t])}\right]^2 \lesssim \mu(s,t)^{2\theta}, \quad \forall\, 0\leqslant t-s\leqslant 1/n.$$

Applying Lemma 5.14, we see that $\mathbb{E}\exp\left((2\gamma'^2-\gamma')\int_0^1 |(\sigma^{-1}b^n)(r,\bar{X}^n_{k_n(r)})|^2 dr\right)$ is bounded uniformly by a finite constant. Hence, we have shown that $\mathbb{E}\rho^{\gamma'}$ is bounded uniformly in n. Combining with (5.33) and (5.34), we obtain (5.1). □

6. Analysis of the continuum paths. To show Theorem 2.3, we need the following result, extending the results of Section 4 to functionals of solutions to (1.1).

THEOREM 6.1. *Let X be the solution to (1.1).*

(i) Assuming Conditions $\mathfrak{A}\mathbf{1}$ and $\mathfrak{B}$. Let h be a function in $\mathbb{L}^{q_1}_{p_1}([0,1])$ for some $p_1, q_1 \in [1,\infty]$ satisfying $\frac{d}{p_1}+\frac{2}{q_1}<2$. Then for every $m\geqslant 1$, there exists a constant $N = N(d,p,q,m)$ such that

$$\|\int_0^1 h(r,X_r)dr\|_{L_m(\Omega)} \leqslant N\|h\|_{\mathbb{L}^{q_1}_{p_1}([0,1])}.$$

(ii) Assuming Conditions $\mathfrak{A}$-$\mathfrak{B}$ with $q_0=\infty$ and $\frac{1}{p}+\frac{1}{p_0}<1$. Let g be a function in $\mathbb{L}^{q_2}_{p_2}([0,1])$ and let $\nu\in[0,1)$ such that $\frac{d}{p_2}+\frac{2}{q_2}+\nu<2$. Then for any $\bar{p}\in(0,p_2)$, there exists a constant $N = N(\nu,d,p,q,p_2,q_2,\bar{p})$ such that

$$\|\sup_{t\in[0,1]}\left|\int_0^t g(r,X_r)dr\right|\|_{L_{\bar{p}}(\Omega)} \leqslant N\|g\|_{\mathbb{L}^{q_2}_{-\nu,p_2}([0,1])}.$$

(iii) Assuming Conditions $\mathfrak{A}$-$\mathfrak{B}$ with $q_0=\infty$ and $\frac{1}{p}+\frac{1}{p_0}<1$. Let g be a function in $\mathbb{L}^{q_2}_{p_2}([0,1])$ with $\frac{d}{p_2}+\frac{2}{q_2}<1$, Γ be a nonnegative number and w_0 be a continuous control on Δ. We assume that for every $(s,t)\in\Delta$,

$$\|g\|_{\mathbb{L}^{q_2}_{-1,p_2}([s,t])} \leqslant \Gamma w_1(s,t)^{\frac{1}{q_2}} \quad \text{and} \quad \|g\|_{\mathbb{L}^{q_2}_{p_2}([s,t])} \leqslant w_1(s,t)^{\frac{1}{q_2}}.$$

Then for any $\bar{p} \in (0, p_2)$, there exists a constant $N = N(\nu, d, p_2, q_2, p, q, \bar{p})$ such that

$$\| \sup_{t\in[0,1]} | \int_0^t g(r, X_r)dr| \|_{L_{\bar{p}}(\Omega)} \leqslant N\Gamma(1 + |\log(\Gamma)|)w_1(0, 1)^{\frac{1}{q_2}}.$$

Similar to the methods in Sections 4 and 5, first we derive some analytic estimates on the transition operators associated to solutions of (1.1) without drift. Using these estimates, one can apply stochastic sewing techniques (Lemma 3.2) and Girsanov theorem to obtain the desired moment bounds.

We begin with moment bounds on the solutions to the driftless SDEs. Let $\bar{X}$ be a solution to SDE

$$d\bar{X}_t = \sigma(t, \bar{X}_t)dB_t, \quad \bar{X}_0 = x \in \mathbb{R}^d. \tag{6.1}$$

Under Condition $\mathfrak{A}$, it is well-known (see [66]) that the probability law of $\bar{X}$ is unique and Markov. In fact, solutions to equation (6.1) are strongly unique under Condition $\mathfrak{A}$. This follows from the arguments in the proof of Theorem 2.2 in the following section. However, only the law of $\bar{X}$ is relevant to our considerations herein. Let $Q_{s,t}$ be the transition operator associated to $\bar{X}$. In particular, we have $\mathbb{E}(f(\bar{X}_t)|\mathcal{F}_s) = Q_{s,t}f(\bar{X}_s)$ for any bounded measurable function f.

LEMMA 6.2. *Assuming Condition $\mathfrak{A}\mathbf{1}$. Let $p_1, p_2 \in [1, \infty]$, $p_1 \leqslant p_2$. There exists a constant $N = N(\alpha, d, p_1, p_2, K_1, K_2)$ such that for every $f \in L_{p_1}(\mathbb{R}^d)$ and $s \leqslant t$,*

$$\|Q_{s,t}f\|_{L_{p_2}(\mathbb{R}^d)} \leqslant N(t - s)^{\frac{d}{2p_2} - \frac{d}{2p_1}} \|f\|_{L_{p_1}(\mathbb{R}^d)}. \tag{6.2}$$

PROOF. Let $\bar{X}^n$ be the solution to the Euler–Maruyama scheme (5.21). It suffices to show that the laws of $\bar{X}^n$ converge to the law of $\bar{X}$ for (6.2) is then derived from Theorem 5.5. Let $\mathbf{P}^n$ be the probability law of $\bar{X}^n$ on $C([0, 1])$. Here $C([0, 1])$ is the space of continuous functions $\omega : [0, 1] \to \mathbb{R}^d$ equipped with the topology of uniform convergence, the Borel σ-algebra and the filtration $t \mapsto \mathcal{G}_t = \sigma\{\omega_s : s \in [0, t]\}$. Let ϕ be a smooth function with bounded derivatives. By Itô formula, we see that

$$M_t^n(\omega) = \phi(\omega_t) - \phi(x) - \frac{1}{2}\int_0^t a^{ij}(r, \omega_{k_n(r)})\partial_{ij}^2\phi(\omega_r)dr$$

is a martingale under $\mathbf{P}^n(d\omega)$. Define

$$M_t(\omega) = \phi(\omega_t) - \phi(x) - \frac{1}{2}\int_0^t a^{ij}(r, \omega_r)\partial_{ij}^2\phi(\omega_r)dr.$$

It is easy to see that $\|\bar{X}_t^n - \bar{X}_s^n\|_{L_p(\Omega)} \lesssim (t-s)^{1/2}$ for any $p \geqslant 2$ and $s \leqslant t$. This implies that the probability laws $\{\mathbf{P}^n\}_n$ are tight. Let $\mathbf{P}$ be a probability measure such that $\mathbf{P}^n$ converges to $\mathbf{P}$ through a subsequence, which we still denote by $\mathbf{P}^n$. Let $s \leqslant t$ be fixed and $G \in \mathcal{G}_s$. We have

$$\begin{aligned}\int \delta M_{s,t}\mathbf{1}_G d\mathbf{P} &= \int \delta M_{s,t}\mathbf{1}_G(d\mathbf{P} - d\mathbf{P}^n) + \int (\delta M_{s,t} - \delta M_{s,t}^n)\mathbf{1}_G d\mathbf{P}^n + \int \delta M_{s,t}^n \mathbf{1}_G d\mathbf{P}^n \\ &=: I_1 + I_2 + I_3.\end{aligned}$$

It is evident that $\lim_n I_1 = 0$. Using Hölder continuity of a we have

$$\begin{aligned}|I_2| \lesssim \int \left[\int_s^t |\omega_r - \omega_{k_n(r)}|^\alpha dr\right] d\mathbf{P}^n(\omega) &\lesssim \int_s^t \mathbb{E}|\bar{X}_r^n - \bar{X}_{k_n(r)}^n|^\alpha dr \\ &\lesssim \int_s^t |r - k_n(r)|^{\alpha/2} dr.\end{aligned}$$

This implies that $\lim_n I_2 = 0$. Because M^n is a martingale under $\mathbf{P}^n$, $I_3 = 0$. It follows that $\int \delta M_{s,t} \mathbf{1}_G d\mathbf{P} = 0$, and hence M is a martingale under $\mathbf{P}$. In other words, $\mathbf{P}$ is a solution to the martingale problem associated to equation (6.1), which is unique ([66]). We have shown that $\{\mathbf{P}^n\}_n$ has exactly one accumulating point, which is the law of (6.1). This also means that $\bar{X}^n$ converges weakly to $\bar{X}$. □

LEMMA 6.3. *Assuming Condition 𝔄1. Let h be a measurable function on $\mathbb{R}^d$ such that $\|h\|_{L_\rho(\mathbb{R}^d)}$ is finite for some $\rho \in (0,\infty]$. Then there exists a finite constant $N = N(\alpha, d, \rho, K_1, K_2)$ such that for every $r, v \in [0,1]$, $r > v$,*

$$\|h(\bar{X}_r)\|_{L_\rho(\Omega|\mathcal{F}_v)} \leqslant N(r-v)^{-\frac{d}{2\rho}} \|h\|_{L_\rho(\mathbb{R}^d)}. \tag{6.3}$$

PROOF. This is a direct consequence of (6.2). The argument is similar to that of Proposition 5.10(i), hence, is omitted. □

The next result is a special case of Theorem 6.1 when $b = 0$, which is an analogue of Propositions 4.6 and 4.7.

PROPOSITION 6.4. *Let $p \in (1,\infty)$, $q \in (2,\infty)$. Let $\bar{X}$ be a solution to (6.1).*

(i) Assuming Condition 𝔄1. Let h be a function in $\mathbb{L}^{q_1}_{p_1}([0,1])$ for some $p_1, q_1 \in [1,\infty]$ satisfying $\frac{d}{p_1} + \frac{2}{q_1} < 2$. Then for every $m \geqslant 1$, there exists a constant $N = N(d, p_1, q_1, m)$ such that

$$\| \int_0^1 h(r, \bar{X}_r) dr \|_{L_m(\Omega)} \leqslant N \|h\|_{\mathbb{L}^{q_1}_{p_1}([0,1])}.$$

(ii) Assuming Condition 𝔄 with $q_0 = \infty$ and $\frac{1}{p} + \frac{1}{p_0} < 1$. Let g be a function in $\mathbb{L}^q_p([0,1])$ and let $\nu \in [0,1)$ such that $\frac{d}{p} + \frac{2}{q} + \nu < 2$. Then for any $\bar{p} \in (0,p)$, there exists a constant $N = N(\nu, d, p, q, \bar{p})$ such that

$$\| \sup_{t\in[0,1]} | \int_0^t g(r, \bar{X}_r) dr | \|_{L_{\bar{p}}(\Omega)} \leqslant N \|g\|_{\mathbb{L}^q_{-\nu,p}([0,1])}.$$

(iii) Assuming Condition 𝔄 with $q_0 = \infty$, $\frac{1}{p} + \frac{1}{p_0} < 1$ and $\frac{d}{p} + \frac{2}{q} < 1$. Let g be a function in $\mathbb{L}^q_p([0,1])$, Γ be a nonnegative number and w_1 be a continuous control on Δ. We assume that for every $(s,t) \in \Delta$,

$$\|g\|_{\mathbb{L}^q_{-1,p}([s,t])} \leqslant \Gamma w_1(s,t)^{\frac{1}{q}} \quad \text{and} \quad \|g\|_{\mathbb{L}^q_p([s,t])} \leqslant w_1(s,t)^{\frac{1}{q}}.$$

Then for any $\bar{p} \in (0,p)$, there exists a constant $N = N(\nu, d, p, q, \bar{p})$ such that

$$\| \sup_{t\in[0,1]} | \int_0^t g(r, \bar{X}_r) dr | \|_{L_{\bar{p}}(\Omega)} \leqslant N\Gamma(1 + |\log(\Gamma)|) w_1(0,1)^{\frac{1}{q}}.$$

PROOF. (i) Using Minkowski inequality, Lemma 6.3 and Hölder inequality, we have

$$\begin{aligned} \| \int_s^t h(r, \bar{X}_r) dr \|_{L_{p_1}(\Omega|\mathcal{F}_s)} &\leqslant \int_s^t \|h(r, \bar{X}_r)\|_{L_{p_1}(\Omega|\mathcal{F}_s)} dr \\ &\lesssim \int_s^t (r-s)^{-\frac{d}{2p_1}} \|h_r\|_{L_{p_1}(\mathbb{R}^d)} dr \lesssim (t-s)^{1-\frac{d}{2p_1}-\frac{1}{q_1}} \|h\|_{\mathbb{L}^{q_1}_{p_1}([s,t])}. \end{aligned}$$

This implies that

$$\mathbb{E}_s \int_s^t |h(r,\bar{X}_r)|dr \lesssim (t-s)^{1-\frac{d}{2p_1}-\frac{1}{q_1}} \|h\|_{\mathbb{L}^{q_1}_{p_1}([s,t])}. \tag{6.4}$$

From here, we apply Lemma 3.5 to obtain part (i).

The proofs of parts (ii,iii) are similar to those of Propositions 4.6 and 4.7. The statistical estimates for Brownian motion are replaced by those obtained in Theorem A.11. The condition $\frac{d}{p}+\frac{2}{q}+\nu<2$ appears when applying Lemma 3.4. We only provide proof of (iii) while the proof of (ii) is left to the readers.

(iii) For each $(s,t)\in\Delta$, put

$$A_{s,t}=\mathbb{E}_s\int_s^t g(r,\bar{X}_r)dr \quad \text{and} \quad J_{s,t}=\int_s^t g(r,\bar{X}_r)dr - A_{s,t}.$$

Define the control w by

$$w(s,t)=\Big[(s-v)^{-\frac{d}{2p}}(t-s)^{\frac{1}{2}-\frac{1}{q}}w_1(s,t)^{\frac{1}{q}}\Big]^2+(s-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{1}{q}}w_1(s,t)^{\frac{1}{q}}.$$

Let $u^t\in\mathbb{L}^q_{2,p}([0,t])$ be the solution ([51, Theorem 2.1]) to

$$(\partial_s+\frac{1}{2}a^{ij}\partial^2_{ij})u+g=0, \quad u(t,\cdot)=0.$$

Applying Itô formula for non-degenerate diffusions (see [69, Lemma 4.1]), we see that $A_{s,t}=u^t_s(\bar{X}_s)$. Applying (6.3) and Theorem A.11, we have

$$\|A_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}\lesssim (s-v)^{-\frac{d}{2p}}\|u^t_s\|_{L_p(\mathbb{R}^d)}\lesssim (s-v)^{-\frac{d}{2p}}(t-s)^{\frac{1}{2}-\frac{1}{q}}\|g\|_{\mathbb{L}^q_{-1,p}([s,t])}.$$

By our assumption, the previous estimate implies that $\|\delta J_{s,u,t}\|_{L_p(\Omega|\mathcal{F}_v)}\lesssim \Gamma w(s,t)^{1/2}$ for every $v<s\leqslant u\leqslant t\leqslant 1$. It is evident that $\mathbb{E}_s J_{s,t}=0$ and hence $\mathbb{E}_s\delta J_{s,u,t}=0$. This verifies the conditions (3.2) and (3.3) of Lemma 3.2.

On the other hand, using Minkowski inequality, (6.3) and Hölder inequality, we have

$$\begin{aligned}\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}&\leqslant 2\int_s^t\|g(r,\bar{X}_r)\|_{L_p(\Omega|\mathcal{F}_v)}dr\lesssim\int_s^t(r-v)^{-\frac{d}{2p}}\|g_r\|_{L_p(\mathbb{R}^d)}dr\\&\lesssim (r-v)^{-\frac{d}{2p}}(t-s)^{1-\frac{1}{q}}\|g\|_{\mathbb{L}^q_p([s,t])}\lesssim w(s,t),\end{aligned}$$

verifying condition (3.1). An application of Lemma 3.2 yields that $\|J_{s,t}\|_{L_p(\Omega|\mathcal{F}_v)}\lesssim\Gamma(1+|\log(\Gamma)|)w(s,t)^{1/2}+\Gamma w(s,t)$ for every $v<s\leqslant u\leqslant t\leqslant 1$. From here, the argument follows analogously as in the proof of Proposition 4.6, using Lemma 3.4 to remove the singularity near v then using Lemma 3.1 to obtain the desired estimate for the supremum. □

Proof of Theorem 6.1. Let $\bar{X}$ be a solution to (6.1). Similar to the proof of Theorem 5.1, we use Girsanov transformation to deduce the moment estimates for X from those obtained in Proposition 6.4. Indeed, define the measure $\bar{\mathbb{P}}:=\rho\mathbb{P}$ where

$$\rho:=\exp\left(-\int_0^1(\sigma^{-1}b)(r,\bar{X}_r)dB_r-\frac{1}{2}\int_0^1\left|(\sigma^{-1}b)(r,\bar{X}_r)\right|^2dr\right).$$

From [69, Lemma 4.1] and Novikov criterion, we see that $\mathbb{E}\rho^r<\infty$ for every $r\in\mathbb{R}$. By Girsanov theorem, $\bar{\mathbb{P}}$ is a probability measure and the law of X under $\bar{\mathbb{P}}$ is the same as the law of $\bar{X}$ under $\mathbb{P}$. From here, we deduce Theorem 6.1 from Proposition 6.4, using similar computations as in the proof of Theorem 5.1. □

REMARK 6.5. Observe that pathwise uniqueness is not necessary and only weak uniqueness of (1.1) is used in the above proof. In addition, reasoning as in [41, Lemmas 3.2, 3.3 and Remark 3.5], one can derive weak uniqueness for (1.1) from Proposition 6.4(i). Consequently, Theorem 6.1 holds for any adapted solution to (1.1).

7. Proof of the main results. We present in the current section the proofs of Theorems 2.2 and 2.3. We state a maximal regularity result for parabolic equations, which is a direct consequence of [69, Theorem 3.2].

LEMMA 7.1. *Assume Conditions $\mathfrak{A}$-$\mathfrak{B}$. Let $f \in \mathbb{L}_p^q([0,1])$ and $M > 0$ be such that*

$$\|f\|_{\mathbb{L}_p^q([0,1])} + \|b\|_{\mathbb{L}_p^q([0,1])} \leqslant M.$$

Then there exists $\lambda_0 = \lambda_0(M, a) \geqslant 1$ such that that for all $\lambda \geqslant \lambda_0$, there is a unique solution u in $\mathbb{L}_{2,p}^q([0,1])$ to the equation

$$\partial_t u + \frac{1}{2} a^{ij} \partial_{ij}^2 u + b \cdot \nabla u + f = \lambda u, \quad u(1,\cdot) = 0. \tag{7.1}$$

Furthermore, for any $\gamma \in [0,2)$, $p_1 \in [p,\infty)$, $q_1 \in [q,\infty)$ with $\frac{d}{p} + \frac{2}{q} < 2 - \gamma + \frac{d}{p_1} + \frac{2}{q_1}$, there is a constant $C = C(M, \gamma, p_1, q_1) > 0$ such that for any $\lambda \geqslant \lambda_0$,

$$\lambda^{\frac{1}{2}(2-\gamma+\frac{d}{p_1}+\frac{2}{q_1}-\frac{d}{p}-\frac{2}{q})} \|u\|_{\mathbb{L}_{\gamma,p_1}^{q_1}([0,1])} + \|\partial_t u\|_{\mathbb{L}_p^q([0,1])} + \|u\|_{\mathbb{L}_{2,p}^q([0,1])} \leqslant C \|f\|_{\mathbb{L}_p^q([0,1])}.$$

LEMMA 7.2. *Let X^n be the solution to (1.3). Then for every $m \geqslant 2$,*

$$\sup_{t\in[0,1]} \|X_t^n - X_{k_n(t)}^n\|_{L_m(\Omega)} \lesssim (1/n)^{\frac{1}{2}}.$$

PROOF. We have

$$X_t^n - X_{k_n(t)}^n = \int_{k_n(t)}^t b^n(r, X_{k_n(t)}^n) dr + \int_{k_n(t)}^t \sigma(r, X_{k_n(t)}^n) dB_r.$$

Using BDG inequality and Hölder, we have

$$\|X_t^n - X_{k_n(t)}^n\|_{L_m(\Omega)} \lesssim (t - k_n(t))^{1-\frac{1}{q}} \|b^n\|_{\mathbb{L}_\infty^q([k_n(t),t])} + (t - k_n(t))^{1/2}.$$

Using $t - k_n(t) \leqslant 1/n$ and Condition $\mathfrak{B}$,

$$(t - k_n(t))^{1-\frac{1}{q}} \|b^n\|_{\mathbb{L}_\infty^q([k_n(t),t])} \lesssim (1/n)^{\frac{1}{2}}.$$

Combining these estimates, we obtain the result. □

Recall that U is the solution to the equation (2.3). Under Conditions $\mathfrak{A}$-$\mathfrak{B}$, there exists $\lambda_0 > 0$ such that for every $\lambda \geqslant \lambda_0$, ∇U is bounded Hölder continuous on $[0,1] \times \mathbb{R}^d$,

$$\sup_n \left[\|\partial_t U\|_{\mathbb{L}_p^q([0,1])} + \|U\|_{\mathbb{L}_{2,p}^q([0,1])} \right] < \infty \quad \text{and} \quad \sup_n \sup_{(t,x)\in[0,1]\times\mathbb{R}^d} |\nabla U(t,x)| = o_\lambda(1), \tag{7.2}$$

where $o_\lambda(1)$ denotes any constant such that $\lim_{\lambda\to\infty} o_\lambda(1) = 0$. Indeed, existence and uniqueness and the first estimate in (7.2) follow from Lemma 7.1. Hölder continuity of ∇U and the second estimate in (7.2) follow by applying [41, Lemma 10.2] and the estimate in Lemma 7.1. Let $\mathcal{M}$ be the Hardy–Littlewood maximal operator defined as

$$\mathcal{M}f(x) := \sup_{0<r<\infty} \frac{1}{|\mathcal{B}_r|} \int_{\mathcal{B}_r} f(x+y) dy, \quad \mathcal{B}_r := \{x \in \mathbb{R}^d : |x| < r\}, \quad r > 0.$$

It is well-known that $\mathcal{M}$ is bounded on $L_p(\mathbb{R}^d)$. Thus we have

$$\|\mathcal{M}f\|_{\mathbb{L}_p^q([0,1])} \lesssim \|f\|_{\mathbb{L}_p^q([0,1])}. \tag{7.3}$$

Define

$$A_t^n := t + \int_0^t \left[\mathcal{M}|\nabla^2 U|(s, X_s) + \mathcal{M}|\nabla^2 U|(s, X_s^n)\right]^2 ds + \int_0^t \left[\mathcal{M}|\nabla\sigma|(s, X_s) + \mathcal{M}|\nabla\sigma|(s, X_s^n)\right]^2 ds. \tag{7.4}$$

PROPOSITION 7.3. *For every $\bar{p} \in (1, p)$, there exists a finite positive constant $c_{\bar{p}}$ such that*

$$\left\| e^{-c_{\bar{p}}|A_1^n|^{\max(\bar{p}/2,1)}} \sup_{t\in[0,1]} |X_t - X_t^n| \right\|_{L_{\bar{p}}(\Omega)} \lesssim \|x_0 - x_0^n\|_{L_{\bar{p}}(\Omega)} + \varpi_n(\bar{p}) + (1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}} \log(n).$$

PROOF. Applying Itô's formula ([69, Lemma 4.1]) for $U(t, X_t)$, we obtain that

$$\int_0^t b^n(r, X_r)dr = U(0, X_0) - U(t, X_t) + \lambda \int_0^t U(r, X_r)dr + \int_0^t \nabla U(r, X_r)[b(r, X_r) - b^n(r, X_r)]dr + \int_0^t (\nabla U \cdot \sigma)(r, X_r)dB_r, \tag{7.5}$$

and similarly,

$$\begin{aligned}\int_0^t b^n(r, X_r^n)dr &= U(0, X_0^n) - U(t, X_t^n) + \lambda \int_0^t U(r, X_r^n)dr \\ &\quad + \int_0^t \nabla U(r, X_r^n)[b^n(r, X_{k_n(r)}^n) - b^n(r, X_r^n)]dr \\ &\quad + \int_0^t \nabla^2 U(r, X_r)[a(r, X_{k_n(r)}^n) - a(r, X_r^n)]dr \\ &\quad + \int_0^t \nabla U(r, X_r^n)\sigma(r, X_{k_n(r)}^n)dB_r. \end{aligned} \tag{7.6}$$

From equations (1.1) and (1.3), we have

$$\begin{aligned} X_t - X_t^n &= x_0 - x_0^n + \int_0^t [b^n(r, X_r) - b^n(r, X_r^n)]dr \\ &\quad + \int_0^t [b(r, X_r) - b^n(r, X_r)]dr + \int_0^t [b^n(r, X_r^n) - b^n(r, X_{k_n(r)}^n)]dr \\ &\quad + \int_0^t [\sigma(r, X_r) - \sigma(r, X_r^n)]dB_r + \int_0^t [\sigma(r, X_r^n) - \sigma(r, X_{k_n(r)}^n)]dB_r. \end{aligned}$$

We plug (7.5) and (7.6) into the previous identity, raise to $\bar{p}$-th power to find that

$$\xi_t := \sup_{s\in[0,t]} |X_s - X_s^n|^{\bar{p}} \lesssim |x_0 - x_0^n|^{\bar{p}} + V_t^0 + \sum_{i=1}^3 V_t^i + \sum_{i=1}^3 I_t^i,$$

where

$$V_t^0 = |U(0,x_0) - U(0,x_0^n)|^{\bar{p}} + \sup_{s\in[0,t]} |U(s,X_s) - U(s,X_s^n)|^{\bar{p}} + \lambda^{\bar{p}} \Big| \int_0^t |U(r,X_r) - U(r,X_r^n)| dr \Big|^{\bar{p}},$$

$$V_t^1 = \sup_{s\in[0,t]} \Big| \int_0^s [I + \nabla U(r,X_r)][b(r,X_r) - b^n(r,X_r)] dr \Big|^{\bar{p}},$$

$$V_t^2 = \sup_{s\in[0,t]} \Big| \int_0^s [I + \nabla U(r,X_r^n)][b^n(r,X_r^n) - b^n(r,X_{k_n(r)}^n)] dr \Big|^{\bar{p}},$$

$$V_t^3 = \sup_{s\in[0,t]} \Big| \int_0^s \nabla^2 U(r,X_r)[a(r,X_{k_n(r)}^n) - a(r,X_r^n)] dr \Big|^{\bar{p}},$$

$$I_t^1 = \sup_{s\in[0,t]} \Big| \int_0^s [I + \nabla U(r,X_r^n)][\sigma(r,X_r) - \sigma(r,X_r^n)] dB_r \Big|^{\bar{p}},$$

$$I_t^2 = \sup_{s\in[0,t]} \Big| \int_0^s [I + \nabla U(r,X_r^n)][\sigma(r,X_r^n) - \sigma(r,X_{k_n(r)}^n)] dB_r \Big|^{\bar{p}},$$

$$I_t^3 = \sup_{s\in[0,t]} \Big| \int_0^s [\nabla U(r,X_r) - \nabla U(r,X_r^n)] \cdot \sigma(r,X_r) dB_r \Big|^{\bar{p}}.$$

Using (7.2) and Cauchy–Schwarz inequality

$$V_t^0 \lesssim |x_0 - x_0^n|^{\bar{p}} + o_\lambda(1) \sup_{s\in[0,t]} |X_s - X_s^n|^{\bar{p}} + \left(\int_0^t \xi_r^{2/\bar{p}} dr \right)^{\frac{\bar{p}}{2}}.$$

To estimate I^1, I^2, I^3, we will utilize a special case of the pathwise Burkholder–Davis–Gundy (BDG) inequality of [65, Theorem 5]. Namely, there exists a constant $C = C(\bar{p}, d)$ such that for any cádlág martingale $\bar{M}$, there exists a local martingale $\widetilde{M}$ such that with probability one,

$$\sup_{s\in[0,t]} |\bar{M}_s|^{\bar{p}} \leqslant C[\bar{M}]_t^{\frac{\bar{p}}{2}} + \widetilde{M}_t, \quad \forall t.$$

In the above, $[\bar{M}]$ is the quadratic variation of $\bar{M}$. We now estimate I^1. By property of maximal function (see [28, Proposition C.1]) and continuity of σ, we have for every $r \in [0,1]$ and every $x, y \in \mathbb{R}^d$

$$|\sigma(r,x) - \sigma(r,y)| \lesssim |x-y|(\mathcal{M}|\nabla\sigma(r,x)| + \mathcal{M}|\nabla\sigma(r,y)|).$$

Using this, (7.2) and the pathwise BDG inequality, we can find a local martingale M^1 such that

$$I_t^1 \lesssim \left(\int_0^t |I + \nabla U(r,X_r^n)|^2 |X_r - X_r^n|^2 \left(\mathcal{M}|\nabla\sigma|(r,X_r) + \mathcal{M}|\nabla\sigma|(r,X_r^n)\right)^2 dr \right)^{\frac{\bar{p}}{2}} + M_t^1$$

$$\lesssim \left(\int_0^t \xi_r^{2/\bar{p}} dA_r^n \right)^{\frac{\bar{p}}{2}} + M_t^1.$$

Similarly,

$$I_t^3 \lesssim \left(\int_0^t \xi_r^{2/\bar{p}} dA_r^n \right)^{\frac{\bar{p}}{2}} + M_t^3$$

for some local martingale M^3. Using Lemma 7.2, (7.2), Hölder continuity of σ and the pathwise BDG inequality, we have

$$I_t^2 \lesssim \left(\int_0^t |X_r^n - X_{k_n(r)}^n|^{2\alpha} dr\right)^{\frac{\bar{p}}{2}} + M_t^2 \lesssim (1/n)^{\bar{p}\frac{\alpha}{2}} + M_t^2.$$

It follows that

$$\xi_t \lesssim o_\lambda(1)\xi_t + \left(\int_0^t \xi^{2/\bar{p}} dA^n\right)^{\frac{\bar{p}}{2}} + V_t + M_t$$

where $V = |x_0 - x_0^n|^{\bar{p}} + V^1 + V^2 + V^3$ and $M = M^1 + M^2 + M^3$. By choosing λ sufficiently large, this deduces to

$$\xi_t \lesssim \left(\int_0^t \xi^{2/\bar{p}} dA^n\right)^{\frac{\bar{p}}{2}} + V_t + M_t$$

Applying stochastic Grönwall lemma, Lemma 3.8, we have

$$\mathbb{E} e^{-c_{\bar{p}}|A_1^n|^{\max(\bar{p}/2,1)}} \xi_1 \leqslant \mathbb{E} V_1 \tag{7.7}$$

for some finite positive constant $c_{\bar{p}}$. In view of Definition 2.1, it is evident that

$$\mathbb{E} V_1^1 \lesssim \varpi_n(\bar{p})^{\bar{p}}.$$

Using Theorem 5.1, (7.2) and Condition $\mathfrak{B}$, we have

$$\mathbb{E} V_1^2 \lesssim \left[(1/n)^{1-\frac{1}{q}} \beta_n(b^n) + (1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}} \log(n)\right]^{\bar{p}} \lesssim \left[(1/n)^{\frac{\alpha}{2}} + (1/n)^{\frac{1}{2}} \log(n)\right]^{\bar{p}}.$$

Using Condition $\mathfrak{A}$ and Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\mathbb{E} V_1^3 &\lesssim \mathbb{E}\Big| \int_0^1 |\nabla^2 U(r, X_r)||X_r^n - X_{k_n(r)}^n|^\alpha dr\Big|^{\bar{p}} \\
&\leqslant \mathbb{E}\left(\int_0^1 |\nabla^2 U(r, X_r)|^2 dr\right)^{\frac{\bar{p}}{2}} \left(\int_0^1 |X_r^n - X_{k_n(r)}^n|^{2\alpha} dr\right)^{\frac{\bar{p}}{2}} \\
&\leqslant \left[\mathbb{E}\left(\int_0^1 |\nabla^2 U(r, X_r)|^2 dr\right)^{\bar{p}}\right]^{1/2} \left[\mathbb{E}\left(\int_0^1 |X_r^n - X_{k_n(r)}^n|^{2\alpha} dr\right)^{\bar{p}}\right]^{1/2}.
\end{aligned}$$

In view of (7.2), Theorem 6.1(i) and Lemma 7.2

$$\mathbb{E} V_1^3 \lesssim (1/n)^{\bar{p}\frac{\alpha}{2}}.$$

The previous estimates for $\mathbb{E} V_1^i$'s and (7.7) yield the result. □

LEMMA 7.4. *Let $\rho \in (0, \frac{p \wedge p_0}{d})$ and $\kappa > 0$ be some fixed constants. Then $\sup_n \mathbb{E} e^{\kappa |A_1^n|^\rho}$ is finite.*

PROOF. We observe that

$$\mathbb{E}_s \delta A_{s,t}^n \lesssim (t-s) + (t-s)^{1-\frac{d}{p}-\frac{2}{q}} \|\mathcal{M}|\nabla^2 u|\|_{\mathbb{L}_p^q([s,t])}^2 + (t-s)^{1-\frac{d}{p_0}-\frac{2}{q_0}} \|\mathcal{M}|\nabla\sigma|\|_{\mathbb{L}_{p_0}^{q_0}([s,t])}^2.$$

Indeed, the estimates for functionals of X^n follow from Proposition 5.10 and Girsanov theorem. The estimates for the functionals of X follows from (6.4), Lemma 3.5 and Girsanov theorem;

or alternatively can be derived from those of X^n and the weak convergence of X^n to X. In view of Remark 3.6, this implies that for any $\lambda \geqslant 0$

$$\mathbb{E}e^{\lambda A_1^n} \lesssim e^{c\lambda^a}, \quad \frac{1}{a} = 1 - \frac{d}{p \wedge p_0},$$

where c is some universal positive constant. For simplicity, we write A for A_1^n below. For every $x > 0$, using Chebyshev inequality, we have

$$\mathbb{P}(A > x) = \mathbb{P}(e^{\lambda A} > e^{\lambda x}) \leqslant e^{-\lambda x}\mathbb{E}e^{\lambda A} \lesssim e^{-\lambda x + c\lambda^a}.$$

One can optimize in λ to obtain that for every x bounded away from 0,

$$\mathbb{P}(A > x) \lesssim e^{-cx^{a'}}, \quad \frac{1}{a} + \frac{1}{a'} = 1,$$

where c is another positive constant. In view of layer cake representation

$$\mathbb{E}e^{\kappa A^\rho} = \kappa\rho \int_0^\infty e^{\kappa x^\rho} x^{\rho-1}\mathbb{P}(A > x)dx,$$

we see that $\mathbb{E}e^{\kappa A^\rho}$ is finite if $\rho < a'$, completing the proof because $a' = (p \wedge p_0)/d$. □

Proof of Theorem 2.2. For $\bar{p} \in (0, 2\frac{p\wedge p_0}{d})$, we obtain from Lemma 7.4 that the quantity $\sup_n \mathbb{E}e^{\kappa|A_1^n|^{\max(\gamma/2,1)}}$ is finite for any constant positive κ. From here, we obtain (2.4) from Proposition 7.3 and Hölder inequality. □

Proof of Theorem 2.3. We put $g = \nabla U$. From Lemma 7.1, we have

$$\sup_n(\|g\|_{\mathbb{L}^{q_2}_{\nu,p_2}([0,1])} + \|g\|_{\mathbb{L}^\infty_\infty([0,1])} + \|g\|_{\mathbb{L}^q_{1,p}([0,1])}) < \infty$$

for all $p_2 \in [p, \infty), q_2 \in [q, \infty)$ and $\nu \in [0, 1)$ with

$$\frac{d}{p} + \frac{2}{q} + \nu - 1 < \frac{d}{p_2} + \frac{2}{q_2}. \tag{7.8}$$

Part (i). Let p_1, q_1 be as in Theorem 2.3(i). From Theorem 6.1(i), we have

$$\varpi_n(m) \lesssim \|(1+g)(b-b^n)\|_{\mathbb{L}^{q_1}_{p_1}} \lesssim (1 + \|g\|_{\mathbb{L}^\infty_\infty})\|b - b^n\|_{\mathbb{L}^{q_1}_{p_1}}$$

which shows (2.6).

Part (ii). Define q_3 by $\frac{1}{q_3} = \frac{1}{q_2} + \frac{1}{q}$. For each $\nu \in [0, 1]$ and $p_3 \in (1, \infty)$ satisfying

$$\frac{1}{p} \leqslant \frac{1}{p_3} \leqslant \frac{1}{p} + \frac{1}{p_2} < \frac{1}{p_3} + \frac{\nu}{d}, \tag{7.9}$$

an application of Lemma A.3(ii) and Hölder inequality shows that the pointwise multiplication is a continuous bilinear map

$$\mathbb{L}^q_{-\nu,p}([0,1]) \times \mathbb{L}^{q_2}_{\nu,p_2}([0,1]) \to \mathbb{L}^{q_3}_{-\nu,p_3}([0,1]).$$

If p_3 can be chosen such that

$$\frac{d}{p_3} + \frac{2}{q_2} + \frac{2}{q} < 2 - \nu, \tag{7.10}$$

then Theorem 6.1(ii) can be applied, which gives for every $\bar{p} \in (0, p_3)$,

$$\|\sup_{t\in[0,t]}|\int_0^t g(b-b^n)(r, X_r)dr|\|_{L_{\bar{p}}(\Omega)} \lesssim \|g(b-b^n)\|_{\mathbb{L}^{q_3}_{-\nu,p_3}} \lesssim \|b-b^n\|_{\mathbb{L}^q_{-\nu,p}},$$

To obtain the last step, we apply the multiplication result above to see that

$$\|g(b-b^n)\|_{\mathbb{L}^{q_3}_{-\nu,p_3}} \lesssim \|g\|_{\mathbb{L}^{q_2}_{\nu,p_2}} \|b-b^n\|_{\mathbb{L}^{q}_{-\nu,p}} \lesssim \|b-b^n\|_{\mathbb{L}^{q}_{-\nu,p}}.$$

On the other hand, applying Theorem 6.1(ii), we have

$$\|\sup_{t\in[0,1]} |\int_0^t (b-b^n)(r,X_r)dr|\|_{L_{\bar{p}}(\Omega)} \lesssim \|b-b^n\|_{\mathbb{L}^{q}_{-\nu,p}}.$$

These estimates verify (2.8).

Next, we verify that it is possible to choose p_2, q_2, p_3 satisfying all the above conditions. Given p_2, q_2, p, q, ν, there exists $p_3 \in (1,\infty)$ satisfying (7.9) and (7.10) iff

$$\begin{cases} \dfrac{d}{p} < 2-\nu-\dfrac{2}{q_2}-\dfrac{2}{q} \\ \dfrac{d}{p}+\dfrac{d}{p_2}-\nu < 2-\nu-\dfrac{2}{q_2}-\dfrac{2}{q} \end{cases}$$

which is deduced to

$$\begin{cases} \dfrac{2}{q_2} < 2-\nu-\dfrac{d}{p}-\dfrac{2}{q} \\ \dfrac{d}{p_2}+\dfrac{2}{q_2} < 2-\dfrac{d}{p}-\dfrac{2}{q}. \end{cases} \tag{7.11}$$

Given p, q, ν, the existence $p_2 \in [p,\infty), q_2 \in [q,\infty)$ satisfying (7.8) and (7.11) is equivalent to the problem (P): given real numbers $a>0, b>0$ and $c_1<c_2$, find $x\in(0,a), y\in(0,b)$ such that $c_1 < x+y < c_2$. Here, we have put $x = d/p_2$, $y = 2/q_2$, $a = d/p$, $b = \min(2/q, 2-\nu-\zeta)$, $c_1 = \zeta+\nu-1$, $c_2 = 2-\zeta$ and $\zeta = d/p+2/q$. With some plotting aid, it is seen that this problem has a solution (x,y) iff $c_1 < a+b$ and $c_2>0$. This deduces to the condition (2.7).

Since (2.8) is valid for all $\bar{p}\in(0,p_3)$, it remains to identify the largest possible value for p_3, denoted by p_3^*. From (7.9) and (7.10), we see that

$$\frac{1}{p_3^*} = \max\left(\frac{1}{p}, \frac{1}{p}+\frac{1}{p_2}-\frac{\nu}{d}\right).$$

We observe that the problem (P) with the additional constraint $d/p_2 \leqslant \nu$ has a solution (x,y) iff (2.7) holds. In other words, we can choose p_2 so that $d/p_2 \leqslant \nu$ and hence $p_3^* = p$. This shows that (2.8) holds for every $\bar{p}\in(0,p)$.

Part (iii). When $\nu = 1$, we ought to take $p_2 = p$, $q_2 = 2$. We can also choose $p_3 = p$ and $q_3 = q/2$. Condition (7.9) is trivially satisfied while condition (7.10) is verified by (2.9). Define the control w_1 by

$$w_1(s,t) = \left(\|g\|^{q/2}_{\mathbb{L}^q_{-1,p}([s,t])} + \|g\|^{q/2}_{\mathbb{L}^q_{\infty}([s,t])}\right) w_0(s,t)^{1/2}$$

Then by the multiplication result above, Hölder inequality and (2.10), we have

$$\|g(b-b^n)\|_{\mathbb{L}^{q/2}_{-1,p}([s,t])} \lesssim \|g\|_{\mathbb{L}^q_{-1,p}([s,t])} \|b-b^n\|_{\mathbb{L}^q_{-1,p}([s,t])} \lesssim \Gamma w_1(s,t)^{\frac{2}{q}},$$

$$\|g(b-b^n)\|_{\mathbb{L}^{q/2}_{p}([s,t])} \lesssim \|g\|_{\mathbb{L}^q_{\infty}([s,t])} \|b-b^n\|_{\mathbb{L}^q_{p}([s,t])} \lesssim w_1(s,t)^{\frac{2}{q}}.$$

We then apply Theorem 6.1(iii) to get

$$\begin{aligned} \|\sup_{t\in[0,t]} |\int_0^t g(b-b^n)(r,X_r)dr|\|_{L_{\bar{p}}(\Omega)} &\lesssim \Gamma(1+|\log\Gamma|) w_1(0,1)^{\frac{2}{q}} \\ &\lesssim \Gamma(1+|\log\Gamma|) w_0(0,1)^{\frac{1}{q}}. \end{aligned}$$

On the other hand, applying Theorem 6.1(iii) under condition (2.10) yields

$$\| \sup_{t\in[0,1]} | \int_0^t (b-b^n)(r,X_r)dr| \|_{L_{\bar{p}}(\Omega)} \lesssim \Gamma(1+|\log\Gamma|) w_0(s,t)^{\frac{1}{q}}.$$

Combining the previous two estimates, we obtain (2.11). □

8. Application: stochastic transport equations. Let (W_t) be a standard d-dimensional Brownian motion on a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,1]},\mathbb{P})$ and let $b:[0,1]\times\mathbb{R}^d\to\mathbb{R}^d$ be a Borel measurable function satisfying (1.2). In this section, we propose a numerical scheme for the following (forward) stochastic linear transport equation

$$\partial_t u + b\cdot\nabla u + \nabla u \circ dW_t = 0, \quad u(0,x)=\rho(x), \tag{8.1}$$

where $\rho\in\cap_{r\geqslant 1}L_{1,r}(\mathbb{R}^d)$ and $\nabla u\circ dW_t$ is interpreted in Stratonovich sense. As in [17], we say u is a *weakly differentiable solution* to (8.1) if

- $u:\Omega\times[0,1]\times\mathbb{R}^d\to\mathbb{R}$ is measurable, $\int_{\mathbb{R}^d}u(t,x)\psi(x)dx$ is progressively measurable for each $\psi\in C_0^\infty(\mathbb{R}^d)$;
- $\mathbb{P}(u(t,\cdot)\in\cap_{r\geqslant1}L^{loc}_{1,r}(\mathbb{R}^d))=1$ for $t\in[0,1]$ and both u and ∇u are in $C([0,1];\cap_{r\geqslant1}L^r(\mathbb{R}^d\times\Omega))$;
- for any $\psi\in C_0^\infty(\mathbb{R}^d)$ and $t\in[0,1]$ with probability one the following holds

$$\begin{aligned}\int_{\mathbb{R}^d}u(t,x)\psi(x)dx+\int_0^t\int_{\mathbb{R}^d}b(s,x)\cdot\nabla u(s,x)\psi(x)dxds\\ =\int_{\mathbb{R}^d}\rho(x)\psi(x)dx+\sum_{i=1}^d\int_0^t\Big(\int_{\mathbb{R}^d}u(s,x)\partial_{x_i}\psi(x)dx\Big)dW^i_s\\ +\frac{1}{2}\int_0^t\int_{\mathbb{R}^d}u(s,x)\Delta\psi(x)dxds.\end{aligned}$$

It is known from [17, Theorems 10, 11] that a weakly differentiable solution u to (8.1) exists uniquely and has the representation $u(\tau,x)=\rho(\phi_0^\tau(x))$, where $\phi_0^\tau(x)$ is the inverse of the stochastic flow of homeomorphisms generated by the solution $(X_\tau(x))_{\tau\in[0,1]}$ to the SDE

$$dX_\tau(x)=b(\tau,X_\tau(x))d\tau+dW_\tau, \quad \tau\in[0,1], \quad X_0(x)=x. \tag{8.2}$$

For each fixed $\tau\in[0,1]$, consider the backward-in-time SDE

$$X_{\tau,s}(x)=x-\int_s^\tau b(r,X_{\tau,r}(x))dr+W_s-W_\tau, \quad 0\leqslant s\leqslant\tau, \quad X_{\tau,\tau}(x)=x. \tag{8.3}$$

The inverse flow $\phi_0^\tau(x)$ is directly related to the solution to the previous SDE through the relation $\phi_0^\tau(x)=X_{\tau,0}(x)$. Indeed, when b is a smooth bounded function with bounded derivatives, this relation is classical, see [42, Theorem 3.7.1]. When b belongs to $\mathbb{L}_p^q$, one can approximate it by smooth functions. Because both $\phi_0^\tau(x), X_{\tau,0}(x)$ are stable under this approximation (see [17, Lemma 3] and [22, Theorem 1.2] respectively), the relation holds true in this case. Alternatively, another argument for this fact is recently provided in [3] utilizing *path-by-path uniqueness* of (8.2).

To devise a numerical scheme for (8.1), it convenes to introduce $X_s^\tau(x):=X_{\tau,\tau-s}(x)$. By a change of variables, we find that

$$X_s^\tau(x)=x-\int_0^s b^\tau(r,X_r^\tau(x))dr+W_s^\tau, \quad 0\leqslant s\leqslant\tau, \tag{8.4}$$

where $b^\tau(r,x) := b(\tau - r, x)$ and $W_r^\tau := W_\tau - W_{\tau - r}$ for $r \in [0,\tau]$. Observe that $(W_t^\tau)_{t\in[0,\tau]}$ is a $(\mathcal{F}_t^\tau)$-Brownian motion with $\mathcal{F}_t^\tau := \sigma(W_{\tau-r_1} - W_{\tau-r_2}, 0 \leqslant r_1 \leqslant r_2 \leqslant t)$ for $t \in [0,\tau]$. Hence, we have the representation $u(\tau,x) = \rho(X_\tau^\tau(x))$. This naturally suggests the numerical scheme

$$u^n(\tau,x) = \rho(X_\tau^{\tau,n}(x)),$$

where for each $\tau \in (0,1]$, $(X_s^{\tau,n})_{s\in[0,\tau]}$ is the tamed Euler–Maruyama approximation for (8.4), namely

$$X_s^{\tau,n}(x) = x - \int_0^s b^{\tau,n}(r, X_{k_n(r)}^{\tau,n}(x))dr + W_s^\tau, \quad 0 \leqslant s \leqslant \tau. \tag{8.5}$$

Here, $b^{\tau,n}(r,x) := b^n(\tau - r, x)$ for $r \in [0,\tau]$ and b^n is an approximation for b satisfying Condition $\mathfrak{B}$.

THEOREM 8.1. *Suppose that Condition $\mathfrak{B}$ holds. Let $\nu \in [0,1)$ satisfy (2.7) and $p_1, q_1 \in [1,\infty]$ satisfy $\frac{d}{p_1} + \frac{2}{q_1} < 2$. Then for any $l \in (1, p \wedge \frac{2p}{d})$, any $\bar r \in (1,\infty)$ satisfying $\frac{1}{\bar r} < \frac{1}{l} - \frac{1}{p}(1 \vee \frac{d}{2})$, we have*

$$\sup_{(\tau,x)\in[0,1]\times\mathbb{R}^d} \tau^{\frac{d}{2\bar r}} \|u^n(\tau,x) - u(\tau,x)\|_{L_l(\Omega)}$$

$$\leqslant N\|\nabla\rho\|_{L_{\bar r}(\mathbb{R}^d)}((1/n)^{\frac12}\log n + \min(\|b - b^n\|_{\mathbb{L}_{p_1}^{q_1}([0,1])}, \|b - b^n\|_{\mathbb{L}_{-\nu,p}^q([0,1])})), \tag{8.6}$$

where N depends on K_4, p, d, l, p_1, q_1 and $\bar r$.

PROOF. We put $b(r,x) = 0$ and $b^n(r,x) = 0$ whenever $r \in \mathbb{R} \setminus [0,1]$ so that $b^\tau(r,x)$ and $b^{\tau,n}(r,x)$ are well-defined functions on $[0,1]\times\mathbb{R}^d$. Let $(\hat W_t)_{t\in[0,1]}$ be a standard $(\hat{\mathcal{F}}_t)$-Brownian motion which is independent from $(W_t)_{t\in[0,1]}$ and define

$$W_r^\tau := \begin{cases} W_\tau - W_{\tau-r} & \text{if } r \in [0,\tau], \\ W_\tau + \hat W_r - \hat W_\tau & \text{if } r \in (\tau,1], \end{cases} \tag{8.7}$$

which is a standard Brownian motion with respect to the filtration $(\mathcal{G}_t^\tau)_{t\in[0,1]} := (\mathcal{F}_{t\wedge\tau}^\tau \vee \hat{\mathcal{F}}_{t\vee\tau})_{t\in[0,1]}$. Equations (8.4) and (8.5) (for each fixed $\tau \in [0,1]$) are extended uniquely over the whole time interval $[0,1]$.

By property of maximal function and continuity of ρ (see [28, Proposition C.1]) we have for every $x, y \in \mathbb{R}^d$,

$$|\rho(x) - \rho(y)| \lesssim |x - y|(\mathcal{M}|\nabla\rho|(x) + \mathcal{M}|\nabla\rho|(y)).$$

It follows that

$$|u(\tau,x) - u_n(\tau,x)| \lesssim |X_\tau^\tau(x) - X_\tau^{\tau,n}(x)|(\mathcal{M}|\nabla\rho|(X_\tau^\tau(x)) + \mathcal{M}|\nabla\rho|(X_\tau^{\tau,n}(x))). \tag{8.8}$$

Next, we estimate the terms on the right-hand side of the previous inequality.

In order to apply Theorem 2.2 to obtain estimates for $X^{\tau(x)} - X^{\tau,n}(x)$, we verify that b^τ and $b^{\tau,n}$ fulfill Condition $\mathfrak{B}$ for each τ. Indeed, it is evident that $b^\tau \in \mathbb{L}_p^q([0,1])$, $b^{\tau,n} \in \mathbb{L}_p^q([0,1]) \cap \mathbb{L}_\infty^q([0,1])$ and that $\|b^{\tau,n}\|_{\mathbb{L}_p^q([0,1])} \leqslant \|b^n\|_{\mathbb{L}_p^q([0,1])}$ which is bounded uniformly in n. In addition, define $\mu^{\tau,n}(s,t) = \mu^n((\tau - t)\vee 0, (\tau - s)\vee 0)$, which is a continuous control on the simplex $\Delta([0,1])$. We have $\|b^{\tau,n}\|_{\mathbb{L}_\infty^q([s,t])} = \|b^n\|_{\mathbb{L}_\infty^q([(\tau-t)\vee0,(\tau-s)\vee0])}$ and hence by (2.2)

$$(1/n)^{\frac12 - \frac1q}\|b^{\tau,n}\|_{\mathbb{L}_\infty^q([s,t])} \leqslant \mu^{\tau,n}(s,t)^\theta.$$

Similarly, $\|b^{\tau,n}\|_{\mathbb{L}^q_p([0,1])} = \|b^n\|_{\mathbb{L}^q_p([0,\tau])} \leqslant \|b^n\|_{\mathbb{L}^q_p([0,1])}$ and $\mu^{\tau,n}(0,1) = \mu^n(0,\tau) \leqslant \mu^n(0,1)$ so that

$$\sup_{n\geqslant 1}\left(\|b^{\tau,n}\|_{\mathbb{L}^q_p([0,1])} + \mu^{\tau,n}(0,1)\right) \leqslant K_4,$$

where K_4 is the constant in Condition $\mathfrak{B}$.

Hence Theorem 2.3(i-ii) yields that for $\bar{p} \in [1,p)$

$$\varpi_n(\bar{p}) \leqslant N\min(\|b-b^n\|_{\mathbb{L}^{q_1}_{p_1}([0,1])}, \|b-b^n\|_{\mathbb{L}^q_{-\nu,p}([0,1])}).$$

Theorem 2.2 yields for any $\gamma \in (0,1)$ and $\bar{p} \in [1, p \wedge \frac{2p}{d})$,

$$\begin{aligned}&\sup_{(t,x)\in[0,1]\times\mathbb{R}^d} \|X^\tau_t(x) - X^{\tau,n}_t(x)\|_{L_{\gamma\bar{p}}(\Omega)}\\ &\leqslant N\left(n^{-1/2}\log n + \min(\|b-b^n\|_{\mathbb{L}^{q_1}_{p_1}([0,1])}, \|b-b^n\|_{\mathbb{L}^q_{-\nu,p}([0,1])})\right). \end{aligned} \tag{8.9}$$

The constant N depends on $K_4, \nu, p, q, p_1, q_1, \bar{p}$.

Similar to the arguments used in the proofs of Theorems 5.1 and 6.1, using Girsanov theorem, one can deduce the estimates for $X^{\tau,n}, X^\tau$ to estimates for Brownian motion, which are obtained in Lemma 4.1. Hence, for any $\bar{r} > r > 1$, we have that

$$\|\mathcal{M}|\nabla\rho|(X^\tau_\tau(x))\|_{L_r(\Omega)} \lesssim \|\mathcal{M}|\nabla\rho|(B_\tau)\|_{L_{\bar{r}}(\Omega)} \lesssim \tau^{-\frac{d}{2\bar{r}}}\|\mathcal{M}|\nabla\rho|\|_{L_{\bar{r}}(\mathbb{R}^d)} \lesssim \tau^{-\frac{d}{2\bar{r}}}\|\nabla\rho\|_{L_{\bar{r}}(\mathbb{R}^d)},$$

and similarly,

$$\|\mathcal{M}|\nabla\rho|(X^{\tau,n}_\tau(x))\|_{L_r(\Omega)} \lesssim \tau^{-\frac{d}{2\bar{r}}}\|\nabla\rho\|_{L_{\bar{r}}(\mathbb{R}^d)}.$$

Given $l, \bar{r}$ as in the statement, we can choose $r \in (1,\bar{r})$, $\bar{p} \in [1,p)$, $\gamma \in (0,1)$ such that $\frac{1}{r} + \frac{1}{\gamma\bar{p}} = \frac{1}{l}$. From (8.8), applying Hölder inequality, we have

$$\begin{aligned}&\|u_n(\tau,x) - u(\tau,x)\|_{L_l(\Omega)}\\ &\quad\lesssim \|X^\tau_\tau(x) - X^{\tau,n}_\tau(x)\|_{L_{\gamma\bar{p}}(\Omega)}(\|\mathcal{M}|\nabla\rho|(X^\tau_\tau(x))\|_{L_r(\Omega)} + \|\mathcal{M}|\nabla\rho|(X^{\tau,n}_\tau(x))\|_{L_r(\Omega)}).\end{aligned}$$

Combining with the estimates obtained previously, we obtain that

$$\begin{aligned}&\|u_n(\tau,x) - u(\tau,x)\|_{L_l(\Omega)}\\ &\quad\lesssim \tau^{-\frac{d}{2\bar{r}}}\|\nabla\rho\|_{L_{\bar{r}}(\mathbb{R}^d)}\left(n^{-1/2}\log n + \min(\|b-b^n\|_{\mathbb{L}^{q_1}_{p_1}([0,1])}, \|b-b^n\|_{\mathbb{L}^q_{-\nu,p}([0,1])})\right),\end{aligned}$$

which implies (8.6). □

APPENDIX: PARABOLIC EQUATIONS WITH DISTRIBUTIONAL FORCING

In this section we show the well-posedness and regularity estimates for the solutions for a class parabolic equations with distributional forcing, which are used in Section 6 and have their own interests. Although such equations have been considered extensively in literatures, for instance in [40], [37], [72] and [69], the available results therein are valid under different hypotheses and are not applicable to our situations.

For each $r \in [1,\infty]$, we denote its Hölder conjugate by r', i.e. $\frac{1}{r} + \frac{1}{r'} = 1$. For each Banach space $\mathcal{E}$, we denote its topological dual by $\mathcal{E}^*$, and the dual paring between $\mathcal{E}$ and $\mathcal{E}^*$ by $\langle\cdot,\cdot\rangle_{\mathcal{E}^*,\mathcal{E}}$. We consider the parabolic partial differential equations (PDEs)

$$(\partial_s + a^{ij}\partial^2_{ij})u = f, \quad u(1,\cdot) = 0 \tag{A.10}$$

and

$$\partial_t v - \partial^2_{ij}(a^{ij}v) + g = 0, \quad v(0,\cdot) = 0 \tag{A.11}$$

under the following assumptions:[6]

CONDITION $\mathfrak{A}'$.

1. a is a $d \times d$-symmetric matrix-valued measurable function on $[0,1] \times \mathbb{R}^d$. There exists a constant $k_1 \in [1,\infty)$ such that for every $s \in [0,1]$ and $x \in \mathbb{R}^d$

$$k_1^{-1} I \leqslant a(s,x) \leqslant k_1 I. \tag{A.12}$$

Furthermore, $a(s,\cdot)$ is weakly differentiable for a.e. $s \in [0,1]$ and $k_3 := \|\nabla a\|_{\mathbb{L}^\infty_{p_0}([0,1])}$ is finite for some $p_0 \in (d,\infty)$.

2. $f \in \mathbb{L}^q_{-1,p}([0,1])$ and $g \in \mathbb{L}^{q'}_{-1,p'}([0,1])$ for some $p, q \in (1,\infty)$ satisfying $\frac{1}{p} + \frac{1}{p_0} < 1$.

DEFINITION A.2. A measurable function $u : [0,1] \times \mathbb{R}^d \to \mathbb{R}$ is a solution to (A.10) if $u \in \mathbb{L}^q_{1,p}([0,1])$, $\partial_s u \in \mathbb{L}^q_{-1,p}([0,1])$, $u(1,\cdot) = 0$ and equation (A.10) holds in $\mathbb{L}^{q'}_{-1,p'}([0,1])$, i.e. for every $\phi \in \mathbb{L}^{q'}_{1,p'}([0,1])$

$$\int_0^1 \langle (\partial_s + a^{ij}\partial^2_{ij})u_t, \phi_t \rangle_{L_{-1,p}(\mathbb{R}^d) \times L_{1,p'}(\mathbb{R}^d)} dt = \int_0^1 \langle f_t, \phi_t \rangle_{L_{-1,p}(\mathbb{R}^d) \times L_{1,p'}(\mathbb{R}^d)} dt. \tag{A.13}$$

Likewise, a measurable function $v : [0,1] \times \mathbb{R}^d$ is a solution to (A.11) if $v \in \mathbb{L}^{q'}_{1,p'}([0,1])$, $\partial_t v \in \mathbb{L}^{q'}_{-1,p'}([0,1])$, $v(0,\cdot) = 0$ and equation (A.11) holds in $\mathbb{L}^q_{-1,p}([0,1])$, i.e. for every $\phi \in \mathbb{L}^q_{1,p}([0,1])$

$$\int_0^1 \langle \partial_t v_s - \partial^2_{ij}(a^{ij}v_s), \phi_s \rangle_{L_{-1,p'}(\mathbb{R}^d) \times L_{1,p}(\mathbb{R}^d)} ds + \int_0^1 \langle g_s, \phi_s \rangle_{L_{-1,p'}(\mathbb{R}^d) \times L_{1,p}(\mathbb{R}^d)} ds = 0. \tag{A.14}$$

In the above definitions, we have implicitly understood that $a^{ij}\partial^2_{ij}u$ and $\partial^2_{ij}(a^{ij}v)$ are well-defined distributions in $\mathbb{L}^{q'}_{-1,p'}([0,1])$ and $\mathbb{L}^q_{-1,p}([0,1])$ respectively. To see this, we need the following multiplication result:

LEMMA A.3. *Let p, p_1, p_2 be real numbers in $(1,\infty)$ and let $\nu \in (0,1]$.*
(i) Assume that $p_1, p_2 \geqslant p$ and that $\frac{1}{p} \leqslant \frac{1}{p_1} + \frac{1}{p_2} < \frac{1}{p} + \frac{\nu}{d}$. Then the pointwise multiplication is a continuous bilinear map

$$L_{\nu,p_1}(\mathbb{R}^d) \times L_{\nu,p_2}(\mathbb{R}^d) \to L_{\nu,p}(\mathbb{R}^d).$$

(ii) Assume that $p_1 \geqslant p$, that $p_2 \geqslant p_1'$ and that $\frac{1}{p} \leqslant \frac{1}{p_1} + \frac{1}{p_2} < \frac{1}{p} + \frac{\nu}{d}$. Then the pointwise multiplication is a continuous bilinear map

$$L_{-\nu,p_1}(\mathbb{R}^d) \times L_{\nu,p_2}(\mathbb{R}^d) \to L_{-\nu,p}(\mathbb{R}^d).$$

[6] Parabolic PDEs with distributional forcing have been considered by Kim [37]. However, his result is applicable to (A.10) and (A.11) only when a is continuously differentiable in the spatial variables with bounded derivatives and $p \leqslant q$ hence Condition $\mathfrak{A}'$ is excluded.

(iii) Let g be a bounded measurable function such that $\nabla g \in L_{p_0}(\mathbb{R}^d)$ for some $p_0 \in (d,\infty)$. Let f be in $L_{1,p'}(\mathbb{R}^d)$, h be in $L_{-1,p}(\mathbb{R}^d)$ and assume that $\frac{1}{p}+\frac{1}{p_0}<1$. Then fg belongs to $L_{1,p'}(\mathbb{R}^d)$, gh belongs to $L_{-1,p}(\mathbb{R}^d)$ and

$$\|fg\|_{L_{1,p'}(\mathbb{R}^d)} \lesssim (\|g\|_{L_\infty(\mathbb{R}^d)} + \|\nabla g\|_{L_{p_0}(\mathbb{R}^d)})\|f\|_{L_{1,p'}(\mathbb{R}^d)}, \tag{A.15}$$

$$\|gh\|_{L_{-1,p}(\mathbb{R}^d)} \lesssim (\|g\|_{L_\infty(\mathbb{R}^d)} + \|\nabla g\|_{L_{p_0}(\mathbb{R}^d)})\|h\|_{L_{-1,p}(\mathbb{R}^d)}. \tag{A.16}$$

PROOF. (i-ii) are consequences of [72, Lemma 2.2]. Concerning (iii), define p_3 by $\frac{1}{p'} = \frac{1}{p_0}+\frac{1}{p_3}$. Then by Hölder inequality

$$\|\nabla gf\|_{L_{p'}(\mathbb{R}^d)} \leqslant \|\nabla g\|_{L_{p_0}(\mathbb{R}^d)}\|f\|_{L_{p_3}(\mathbb{R}^d)}.$$

The embedding $L_{1,p'}(\mathbb{R}^d) \hookrightarrow L_{p_3}(\mathbb{R}^d)$ is valid if $\frac{1}{p'}-\frac{1}{d} \leqslant \frac{1}{p_3} \leqslant \frac{1}{p'}$, which is justified by our assumption. It follows that $\|\nabla gf\|_{L_{p'}(\mathbb{R}^d)} \lesssim \|\nabla g\|_{L_{p_0}(\mathbb{R}^d)}\|f\|_{L_{1,p'}(\mathbb{R}^d)}$. It is evident that

$$\|gf\|_{L_{p'}(\mathbb{R}^d)} + \|g\nabla f\|_{L_{p'}(\mathbb{R}^d)} \lesssim \|g\|_{L_\infty(\mathbb{R}^d)}\|f\|_{L_{1,p'}(\mathbb{R}^d)}.$$

From here, we obtain (A.15). To show (A.16), we note that by duality and (A.15),

$$\begin{aligned}\|fgh\|_{L_1(\mathbb{R}^d)} &\lesssim \|h\|_{L_{-1,p}(\mathbb{R}^d)}\|fg\|_{L_{1,p'}(\mathbb{R}^d)}\\ &\lesssim \|h\|_{L_{-1,p}(\mathbb{R}^d)}(\|g\|_{L_\infty(\mathbb{R}^d)} + \|\nabla g\|_{L_{p_0}(\mathbb{R}^d)})\|f\|_{L_{1,p'}(\mathbb{R}^d)}.\end{aligned}$$

This implies (A.16) by duality. □

PROPOSITION A.4. *For every $u \in \mathbb{L}^q_{1,p}([0,1])$ and $v \in \mathbb{L}^{q'}_{1,p'}([0,1])$, under Condition $\mathfrak{A}'$ we have*

$$a^{ij}\partial^2_{ij}u \in \mathbb{L}^q_{-1,p}([0,1]) \quad \text{and} \quad \partial^2_{ij}(a^{ij}v) \in \mathbb{L}^{q'}_{-1,p'}([0,1]).$$

PROOF. Using Lemma A.3(iii), we see that

$$\begin{aligned}\|a^{ij}\partial^2_{ij}u\|_{L_{-1,p}(\mathbb{R}^d)} &\lesssim (\|a^{ij}\|_{L_\infty(\mathbb{R}^d)} + \|\nabla a^{ij}\|_{L_{p_0}(\mathbb{R}^d)})\|\partial^2_{ij}u\|_{L_{-1,p}(\mathbb{R}^d)}\\ &\lesssim (\|a^{ij}\|_{L_\infty(\mathbb{R}^d)} + \|\nabla a^{ij}\|_{L_{p_0}(\mathbb{R}^d)})\|u\|_{L_{1,p}(\mathbb{R}^d)}\end{aligned}$$

and

$$\|\partial^2_{ij}(a^{ij}v)\|_{L_{-1,p'}(\mathbb{R}^d)} \lesssim \|a^{ij}v\|_{L_{1,p'}(\mathbb{R}^d)} \lesssim (\|a^{ij}\|_{L_\infty(\mathbb{R}^d)} + \|\nabla a^{ij}\|_{L_{p_0}(\mathbb{R}^d)})\|v\|_{L_{1,p'}(\mathbb{R}^d)}.$$

These estimates imply the result. □

THEOREM A.5. *Under Condition $\mathfrak{A}'$, there exist a unique solution u to (A.10) and a unique solution v to (A.11). Furthermore, we have*

$$\|u\|_{\mathbb{L}^q_{1,p}([0,1])} + \|\partial_s u\|_{\mathbb{L}^q_{-1,p}([0,1])} \leqslant N\|f\|_{\mathbb{L}^q_{-1,p}([0,1])}, \tag{A.17}$$

$$\|v\|_{\mathbb{L}^{q'}_{1,p'}([0,1])} + \|\partial_t v\|_{\mathbb{L}^{q'}_{-1,p'}([0,1])} \leqslant N\|g\|_{\mathbb{L}^{q'}_{-1,p'}([0,1])}, \tag{A.18}$$

where N is a finite positive constant depending on d,p,q,p_0,k_1,k_3.

Before giving the proof of the above theorem, we show several auxiliary results.

LEMMA A.6 ([72, Lemma 4.1]). *Let ζ be a nonzero smooth function with compact support. Define $\zeta^z(x)=\zeta(x-z)$. For any $\nu \in \mathbb{R}$ and $p \in (1,\infty)$, there exists a constant $C \geqslant 1$ depending only on ν,p,ζ such that for any $f \in L_{\nu,p}(\mathbb{R}^d)$,*

$$C^{-1}\|f\|_{L_{\nu,p}(\mathbb{R}^d)} \leqslant \left(\int_{\mathbb{R}^d}\|f\zeta^z\|^p_{L_{\nu,p}(\mathbb{R}^d)}dz\right)^{1/p} \leqslant C\|f\|_{L_{\nu,p}(\mathbb{R}^d)}.$$

LEMMA A.7 ([37, Lemma 2.5]). *For $k = 1, \dots, n$, let $a^k : \mathbb{R}_+ \to \mathbb{R}^d \times \mathbb{R}^d$ be a measurable function satisfying (A.12). For fixed $\nu \in \mathbb{R}$, $p \in (1, \infty)$, let $u^k \in \mathbb{L}^q_{\nu,p}([0,1])$ solve the following PDE*

$$(\partial_s + a^{ij}_k \partial^2_{ij})u^k = f^k, \quad u(1, \cdot) = 0.$$

Then

$$\int_0^1 \prod_{k=1}^n \|\nabla^2 u^k(t)\|^p_{L_{\nu,p}(\mathbb{R}^d)} dt \leqslant N \sum_{k=1}^n \int_0^1 \|f^k\|^p_{L_{\nu,p}(\mathbb{R}^d)} \prod_{j \neq k} \|\nabla^2 u^j(t)\|^p_{L_{\nu,p}(\mathbb{R}^d)} dt.$$

We will make use the following:

Convention. For a parameter $\rho > 0$, we write o_ρ and C_ρ for any constants whose exact values depend on ρ and may change from one instance to another, but it is always enforced that $\lim_{\rho \to 0} o_\rho = 0$. In particular, the inequality $A \lesssim D + o_\rho F + C_\rho E$ means that $A \leqslant cD + co_\rho F + cC_\rho E$ for some constant c independent from ρ.

LEMMA A.8. *Assuming Condition $\mathfrak{A}'$. Let ψ_1 be a smooth function supported in the ball $\mathcal{B}_1 := \{x \in \mathbb{R}^d : |x| \leqslant 1\}$. For each $\rho > 0$ and $z \in \mathbb{R}^d$, define $\mathcal{B}^z_\rho := \{x \in \mathbb{R}^d : |x - z| \leqslant \rho\}$, $\psi^z_\rho(x) := \psi_1(\frac{x-z}{\rho})$ and $a(z)(t,x) = \frac{1}{|\mathcal{B}^z_\rho|} \int_{\mathcal{B}^z_\rho} a(t,y)dy$. Then we have*

$$\sup_{(t,z) \in [0,1] \times \mathbb{R}^d} \|(a - a(z))\psi^z_\rho\|_{L_{1,p_0}(\mathbb{R}^d)} \lesssim o_\rho \tag{A.19}$$

Here $\psi^z_\rho(x) := \psi_1(\frac{x-z}{\rho})$, $x, z \in \mathbb{R}^d$, $\rho > 0$.

PROOF. Observe that

$$I_1 := \sup_{z \in \mathbb{R}^d} \|(a - a(z))\psi^z_\rho\|_{L_{p_0}(\mathbb{R}^d)} \lesssim \|\psi_1\|_{L_\infty(\mathbb{R}^d)} \|a\|_{L_\infty(\mathbb{R}^d)} \|\mathbf{1}_{\mathcal{B}_{2\rho}}\|_{L_{p_0}(\mathbb{R}^d)}$$

and

$$I_2 := \sup_{z \in \mathbb{R}^d} \|\nabla a \psi^z_\rho\|_{L_{p_0}(\mathbb{R}^d)} \lesssim \|\psi_1\|_{L_\infty(\mathbb{R}^d)} \|\nabla a \cdot \mathbf{1}_{\mathcal{B}_{2\rho}}\|_{L_{p_0}(\mathbb{R}^d)}.$$

Using Poincaré inequality ([24, Theorem 3.17])

$$I_3 := \sup_{z \in \mathbb{R}^d} \|(a - a(z))\nabla\psi^z_\rho\|_{L_{p_0}(\mathbb{R}^d)} \lesssim \|\nabla\psi_1\|_{L_\infty(\mathbb{R}^d)} \|\nabla a \cdot \mathbf{1}_{\mathcal{B}_{2\rho}}\|_{L_{p_0}(\mathbb{R}^d)}.$$

Condition $\mathfrak{A}'$ implies that $I_1 + I_2 + I_3 \leqslant o_\rho$. Furthermore, because

$$\sup_{z \in \mathbb{R}^d} \|(a - a(z))\psi^z_\rho\|_{L_{1,p_0}(\mathbb{R}^d)} \lesssim (I_1 + I_2 + I_3),$$

we obtain the desired result. □

LEMMA A.9. *Assuming Condition $\mathfrak{A}'$ and additionally that $q \geqslant p$. Then there exists a unique solution u to (A.10) which satisfies (A.17).*

PROOF. By Marcinkiewicz interpolation theorem, it suffices to show the result when $q = \bar{n}p$ for any integer $\bar{n} \geqslant 1$. Let $\bar{n} \geqslant 1$ be a fixed integer. By the method of continuity (e.g. see [23, Theorem 5.2]), it suffices to show that there exists positive $N = N(d, p, q, p_0, k_1, k_3)$ such that whenever u is a solution to (A.10),

$$\|u\|_{\mathbb{L}^{\bar{n}p}_{1,p}([0,1])} \leqslant N \|f\|_{\mathbb{L}^{\bar{n}p}_{-1,p}([0,1])}. \tag{A.20}$$

Note that if u is a solution to (A.10), then using Proposition A.4, the above estimate implies that

$$\|\partial_t u\|_{\mathbb{L}^{\bar{n}p}_{-1,p}([0,1])} \lesssim \|f\|_{\mathbb{L}^{\bar{n}p}_{-1,p}([0,1])}.$$

Let $\rho > 0$ be a fixed constant and ϕ be a nonnegative smooth function such that ϕ is supported in the ball $\mathcal{B}_\rho := \{x \in \mathbb{R}^d : |x| \leqslant \rho\}$ and $\|\phi\|_{L_p(\mathbb{R}^d)} = 1$. For each $z \in \mathbb{R}^d$, define $a(z)$ as in Lemma A.8 and

$$\phi^z(x) := \phi(x - z), \quad u^z(s,x) := u(s,x)\phi^z(x), \quad f^z(s,x) := f(s,x)\phi^z(x).$$

Then u^z satisfies the relation

$$\partial_t u^z + a^{ij}(z)\partial^2_{ij} u^z = F^z, \quad u^z(1,\cdot) = 0, \tag{A.21}$$
$$F^z := f\phi^z + 2a^{ij}\partial_i u \partial_j \phi^z + a^{ij} u \partial^2_{ij}\phi^z + (a^{ij}(z) - a^{ij})\partial^2_{ij} u^z.$$

The proof is now divided into several steps.

Step 1. We show that for each $t \in [0,1]$,

$$\left(\int_{\mathbb{R}^d} \|F^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)} dz\right)^{1/p} \lesssim \|f_t\|_{L_{-1,p}(\mathbb{R}^d)} + C_\rho \|u_t\|_{L_p(\mathbb{R}^d)} + o_\rho \|u_t\|_{L_{1,p}(\mathbb{R}^d)}. \tag{A.22}$$

Applying Lemma A.6 (with $\zeta = \phi, \partial_j\phi, \partial^2_{ij}\phi$ respectively) and Lemma A.3, noting that $\|\nabla\phi\|_\infty + \|\nabla^2\phi\|_\infty \lesssim C_\rho$, we obtain that

$$\begin{aligned}
&\left(\int_{\mathbb{R}^d} \|f\phi^z + 2a^{ij}\partial_i u \partial_j \phi^z + a^{ij} u \partial^2_{ij}\phi^z\|^p_{L_{-1,p}(\mathbb{R}^d)} dz\right)^{1/p} \\
&\quad \lesssim \|f\|_{L_{-1,p}(\mathbb{R}^d)} + \|\nabla\phi\|_\infty \|a^{ij}\partial_i u\|_{L_{-1,p}(\mathbb{R}^d)} + |\nabla^2\phi\|_\infty \|a^{ij} u\|_{L_{-1,p}(\mathbb{R}^d)} \\
&\quad \lesssim \|f\|_{L_{-1,p}(\mathbb{R}^d)} + C_\rho \|u\|_{L_p(\mathbb{R}^d)}.
\end{aligned}$$

Let ψ be a smooth function on $\mathbb{R}^d$ such that $\psi(x) = 1$ if $|x| \leqslant \rho$ and $\psi(x) = 0$ if $|x| \geqslant 2\rho$. Define $\psi^z(x) = \psi(x - z)$. Applying Lemma A.3 (whose hypothesis is justified because $p_0 > p'$ and $p_0 > d$), we have

$$\|(a^{ij}(z) - a^{ij})\partial^2_{ij} u^z\|_{L_{-1,p_0}(\mathbb{R}^d)} \lesssim \|(a^{ij}(z) - a^{ij})\psi^z\|_{L_{1,p_0}(\mathbb{R}^d)} \|\partial^2_{ij} u^z\|_{L_{-1,p}(\mathbb{R}^d)}.$$

By Lemma A.8, we see that

$$\sup_{z \in \mathbb{R}^d} \|(a^{ij}(z) - a^{ij})\psi^z\|_{L_{1,p_0}(\mathbb{R}^d)} \leqslant o_\rho.$$

Hence,

$$\|(a^{ij}(z) - a^{ij})\partial^2_{ij} u^z\|_{L_{-1,p_0}(\mathbb{R}^d)} \leqslant o_\rho \|u^z\|_{L_{1,p}(\mathbb{R}^d)}.$$

It is easy to see that $\int_{\mathbb{R}^d} \|h\phi^z\|^p_{L_p(\mathbb{R}^d)} dz = \|h\|_{L_p(\mathbb{R}^d)}$ for any $h \in L_p(\mathbb{R}^d)$. Hence, by Minkowski inequality and Lemma A.6, we have

$$\begin{aligned}
&\left(\int_{\mathbb{R}^d} \|u^z\|^p_{L_{1,p}} dz\right)^{1/p} \leqslant \left(\int_{\mathbb{R}^d} \|\nabla u^z\|^p_{L_p} dz\right)^{1/p} + \left(\int_{\mathbb{R}^d} \|u^z\|^p_{L_p} dz\right)^{1/p} \\
&\leqslant \left(\int_{\mathbb{R}^d} \|\nabla u \phi^z\|^p_{L_p} dz\right)^{1/p} + \left(\int_{\mathbb{R}^d} \|u \nabla \phi^z\|^p_{L_p} dz\right)^{1/p} + \left(\int_{\mathbb{R}^d} \|u^z\|^p_{L_p} dz\right)^{1/p} \\
&\lesssim \|\nabla u\|_{L_p(\mathbb{R}^d)} + C_\rho \|u\|_{L_p(\mathbb{R}^d)}.
\end{aligned} \tag{A.23}$$

This shows that

$$\left(\int_{\mathbb{R}^d}\|(a^{ij}(z)-a^{ij})\partial^2_{ij}u^z\|^p_{L_{-1,p}(\mathbb{R}^d)}dz\right)^{1/p}\lesssim o_\rho\|u\|_{L_{1,p}(\mathbb{R}^d)}+C_\rho\|u\|_{L_p(\mathbb{R}^d)}.$$

Hence, we have (A.22).

Step 2. We show that for every integer $1\leqslant n\leqslant \bar{n}$ and every $s\in[0,1]$,

$$\|u\|_{\mathbb{L}^{np}_{1,p}([s,1])}\lesssim\|f\|_{\mathbb{L}^{np}_{-1,p}([s,1])}+\|u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}. \tag{A.24}$$

Since

$$\|u\|_{\mathbb{L}^{np}_{1,p}([s,1])}\approx\|\nabla^2u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}+\|u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}, \tag{A.25}$$

it suffices to estimate $\|\nabla^2u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}$. From Lemma A.6, we have

$$\|\nabla^2u\|^{np}_{\mathbb{L}^{np}_{-1,p}([s,1])}\lesssim\int_s^1\left(\int_{\mathbb{R}^d}\|\nabla^2u^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)}dz\right)^n dt+\int_s^1\|u_t\|^{np}_{L_p(\mathbb{R}^d)}dt. \tag{A.26}$$

From Tonelli's theorem, Lemma A.7 and (A.22),

$$\begin{aligned}&\int_s^1\left(\int_{\mathbb{R}^d}\|\nabla^2u^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)}dz\right)^n dt\\&\lesssim\int_s^1\left(\int_{\mathbb{R}^d}\|\nabla^2u^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)}dz\right)^{n-1}\left(\|f_t\|_{L_{-1,p}(\mathbb{R}^d)}+\|u_t\|_{L_p(\mathbb{R}^d)}+o_\rho\|u_t\|_{L_{1,p}(\mathbb{R}^d)}\right)^p dt.\end{aligned}$$

Applying Hölder inequality, we have

$$\begin{aligned}\int_s^1\left(\int_{\mathbb{R}^d}\|\nabla^2u^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)}dz\right)^n dt&\lesssim\left[\int_s^1\left(\int_{\mathbb{R}^d}\|\nabla^2u^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)}dz\right)^n dt\right]^{1-\frac{1}{n}}\\&\quad\times\left[\int_s^1\left(\|f\|_{L_{-1,p}(\mathbb{R}^d)}+C_\rho\|u\|_{L_p(\mathbb{R}^d)}+o_\rho\|u\|_{L_{1,p}(\mathbb{R}^d)}\right)^{np}dt\right]^{\frac{1}{n}},\end{aligned}$$

which yields that

$$\begin{aligned}&\int_s^1\left(\int_{\mathbb{R}^d}\|\nabla^2u^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)}dz\right)^n dt\\&\lesssim\int_s^1\left(\|f\|^{np}_{L_{-1,p}(\mathbb{R}^d)}+C_\rho\|u\|^{np}_{L_p(\mathbb{R}^d)}+o_\rho\|u\|^{np}_{L_{1,p}(\mathbb{R}^d)}\right)dt.\end{aligned}$$

Putting this into (A.26), we obtain that

$$\|\nabla^2u\|^{np}_{\mathbb{L}^{np}_{-1,p}([s,1])}\lesssim\|f\|^{np}_{\mathbb{L}^{np}_{-1,p}([s,1])}+C_\rho\|u\|^{np}_{\mathbb{L}^{np}_{p}([s,1])}+o_\rho\|u\|^{np}_{\mathbb{L}^{np}_{1,p}([s,1])}.$$

Using interpolation inequality

$$\|u\|_{\mathbb{L}^{np}_{p}([s,1])}\lesssim C_\rho\|u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}+o_\rho\|u\|_{\mathbb{L}^{np}_{1,p}([s,1])}, \tag{A.27}$$

we get

$$\|\nabla^2u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}\lesssim\|f\|_{\mathbb{L}^{np}_{-1,p}([s,1])}+C_\rho\|u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}+o_\rho\|u\|_{\mathbb{L}^{np}_{1,p}([s,1])}.$$

In view of (A.25), we have

$$\|u\|_{\mathbb{L}^{np}_{1,p}([s,1])}\lesssim\|f\|_{\mathbb{L}^{np}_{-1,p}([s,1])}+C_\rho\|u\|_{\mathbb{L}^{np}_{-1,p}([s,1])}+o_\rho\|u\|_{\mathbb{L}^{np}_{1,p}([s,1])}.$$

By choosing ρ sufficiently small, we derive derive (A.24) from the above estimate.

Step 3. We show that

$$\|u\|_{\mathbb{L}^{\infty}_{-1,p}([0,1])} \lesssim \|f\|_{\mathbb{L}^{p}_{-1,p}([0,1])}. \tag{A.28}$$

From (A.21), we have

$$u^z_s = \int_s^1 P_{\Sigma_{s,t}(z)} F^z_t dt, \quad \text{where} \quad \Sigma_{s,t}(z) = 2\int_s^t a(r,z)dr.$$

Then Minkowski inequality and [67, Theorem 5.30] yield

$$\|u^z_s\|_{L_{-1,p}(\mathbb{R}^d)} \lesssim \int_s^1 \|F^z_t\|_{L_{-1,p}(\mathbb{R}^d)} dt.$$

Applying Hölder inequality, Lemma A.6, (A.22) and the interpolation inequality (A.27), we obtain from the above that

$$\begin{aligned}\|u_s\|^p_{L_{-1,p}(\mathbb{R}^d)} &\lesssim \int_s^1 \int_{\mathbb{R}^d} \|F^z_t\|^p_{L_{-1,p}(\mathbb{R}^d)} dz dt \\ &\lesssim \int_s^1 \left[\|f_t\|^p_{L_{-1,p}(\mathbb{R}^d)} + \|u_t\|^p_{L_{-1,p}(\mathbb{R}^d)} + \|u_t\|^p_{L_{1,p}(\mathbb{R}^d)}\right] dt.\end{aligned}$$

Applying (A.24) (with $n = 1$), we have

$$\|u_s\|^p_{L_{-1,p}(\mathbb{R}^d)} \lesssim \|f\|^p_{\mathbb{L}^p_{1,p}([0,1])} + \int_s^1 \|u_t\|^p_{L_{-1,p}(\mathbb{R}^d)} dt,$$

which implies (A.28) by Grönwall inequality.

Step 4. Using (A.28) in (A.24) yields

$$\|u\|_{\mathbb{L}^{np}_{1,p}([0,1])} \lesssim \|f\|_{\mathbb{L}^{np}_{1,p}([0,1])} + \|f\|_{\mathbb{L}^{p}_{1,p}([0,1])} \lesssim \|f\|_{\mathbb{L}^{np}_{1,p}([0,1])},$$

which implies (A.17). □

LEMMA A.10. *Assuming Condition 2.1′ and additionally that $q' \geqslant p'$. Then there exists a unique solution v to (A.11) which satisfies (A.18).*

PROOF. The proof is similar to that of Lemma A.9. The main differences are the computations in step 1 of the proof of Lemma A.9, which we will explain below. Let v be a solution to (A.11). Define $\phi^z, a(z)$ as in the proof of Lemma A.9. In addition, define $g^z(t,z) = g(t,x)\phi^z(x)$ and $v^z(t,x) = v(t,x)\phi^z(x)$. Then v^z satisfies the parabolic equation

$$\partial_t v^z - a^{ij}(z)\partial^2_{ij} v^z + G^z = 0, \quad v^z(0,\cdot) = 0,$$

where

$$\begin{aligned}G^z &= g^z + a^{ij}(z)\partial^2_{ij} v^z - \partial^2_{ij}(a^{ij}v)\phi^z \\ &= g^z - \partial^2_{ij}\left((a^{ij} - a^{ij}(z))v^z\right) + 2\partial_i(a^{ij}v)\partial_j\phi^z + a^{ij}v\partial^2_{ij}\phi^z.\end{aligned}$$

Let ψ be a smooth function on $\mathbb{R}^d$ such that $\psi(x) = 1$ if $|x| \leqslant \rho$ and $\psi(x) = 0$ if $|x| \geqslant 2\rho$. Define $\psi^z(x) = \psi(x - z)$. Applying Lemmas A.3 and A.8

$$\begin{aligned}\|\partial_{ij}((a^{ij} - a^{ij}(z))v^z)\|_{L_{-1,p'}(\mathbb{R}^d)} &\lesssim \|(a^{ij} - a^{ij}(z))v^z\|_{L_{1,p'}(\mathbb{R}^d)} \\ &\lesssim \|(a^{ij} - a^{ij}(z))\psi^z\|_{L_{1,p_0}(\mathbb{R}^d)} \|v^z\|_{L_{1,p'}(\mathbb{R}^d)} \leqslant o_\rho \|v^z\|_{L_{1,p'}(\mathbb{R}^d)}\end{aligned}$$

and

$$\|a^{ij}v\partial^2_{ij}\phi^z\|_{L_{-1,p'}(\mathbb{R}^d)} \lesssim \|a^{ij}\psi^z\|_{L_{1,p_0}(\mathbb{R}^d)}\|v\partial^2_{ij}\phi^z\|_{L_{-1,p'}(\mathbb{R}^d)} \lesssim \|v\partial^2_{ij}\phi^z\|_{L_{-1,p'}(\mathbb{R}^d)}.$$

Similar to (A.23), we have

$$\left(\int_{\mathbb{R}^d} \|v^z\|^{p'}_{L_{1,p}(\mathbb{R}^d)}\right)^{1/p'} \leqslant \|v\|_{L_{1,p'}(\mathbb{R}^d)} + C_\rho\|v\|_{L'_p(\mathbb{R}^d)}.$$

This yields

$$\left(\int_{\mathbb{R}^d} \|\partial_{ij}((a^{ij}-a^{ij}(z))v^z) + a^{ij}v\partial^2_{ij}\phi^z\|^{p'}_{L_{-1,p'}(\mathbb{R}^d)}dz\right)^{1/p'} \lesssim C_\rho\|v\|_{L_{p'}(\mathbb{R}^d)} + o_\rho\|v\|_{L_{1,p'}(\mathbb{R}^d)}.$$

Applying Lemma A.6, we have

$$\begin{aligned}\left(\int_{\mathbb{R}^d} \|g^z + 2\partial_i(a^{ij}v)\partial_j\phi^z\|^{p'}_{L_{-1,p'}(\mathbb{R}^d)}dz\right)^{1/p'} &\lesssim \|g\|_{L_{-1,p'}(\mathbb{R}^d)} + C_\rho\|\partial_i(a^{ij}v)\|_{L_{-1,p'}(\mathbb{R}^d)}\\ &\lesssim \|g\|_{L_{-1,p'}(\mathbb{R}^d)} + C_\rho\|v\|_{L_{p'}(\mathbb{R}^d)}.\end{aligned}$$

These estimates imply that

$$\left(\int_{\mathbb{R}^d} \|G^z\|^{p'}_{L_{-1,p'}(\mathbb{R}^d)}dz\right)^{1/p'} \lesssim \|g\|_{L_{-1,p'}(\mathbb{R}^d)} + C_\rho\|v\|_{L_{p'}(\mathbb{R}^d)} + o_\rho\|v\|_{L_{1,p'}(\mathbb{R}^d)}.$$

Using the interpolation inequality

$$\|v\|_{L_{p'}(\mathbb{R}^d)} \leqslant C_\rho\|v\|_{L_{-1,p'}(\mathbb{R}^d)} + o_\rho\|v\|_{L_{1,p'}(\mathbb{R}^d)},$$

we obtain from the previous estimate that

$$\left(\int_{\mathbb{R}^d} \|G^z\|^{p'}_{L_{-1,p'}(\mathbb{R}^d)}dz\right)^{1/p'} \lesssim \|g\|_{L_{-1,p'}(\mathbb{R}^d)} + C_\rho\|v\|_{L_{-1,p'}(\mathbb{R}^d)} + o_\rho\|v\|_{L_{1,p'}(\mathbb{R}^d)}.$$

One can now follow steps 2,3 of the proof of Lemma A.9 to obtain (A.18). □

Proof of Theorem A.5. Concerning equation (A.10), by the method of continuity it suffices to show (A.17) whenever u is a solution to (A.10). The case $q \geqslant p$ has been treated in Lemma A.9. Consider the case $q < p$, which is equivalent to $q' > p'$. For each $g \in \mathbb{L}^{q'}_{-1,p'}([0,1])$, let v be the solution to (A.10), which exists uniquely by Lemma A.10. We take $\phi = v$ in (A.13) and use the equation (A.11) for v to see that

$$\int_0^1 \langle u_s, g_s\rangle_{L_{1,p}(\mathbb{R}^d)\times L_{-1,p'}(\mathbb{R}^d)}ds = \int_0^1 \langle v_t, f_t\rangle_{L_{1,p'}(\mathbb{R}^d)\times L_{-1,p}(\mathbb{R}^d)}dt.$$

Applying Hölder inequality and (A.18), we have

$$\begin{aligned}\left|\int_0^1 \langle v_t, f_t\rangle_{L_{1,p'}(\mathbb{R}^d)\times L_{-1,p}(\mathbb{R}^d)}dt\right| &\leqslant \|v\|_{\mathbb{L}^{q'}_{1,p'}([0,1])}\|f\|_{\mathbb{L}^{q'}_{-1,p'}([0,1])}\\ &\lesssim \|g\|_{\mathbb{L}^{q'}_{-1,p'}([0,1])}\|f\|_{\mathbb{L}^{q'}_{-1,p'}([0,1])},\end{aligned}$$

and hence

$$\left|\int_0^1 \langle u_s, g_s\rangle_{L_{1,p}(\mathbb{R}^d)\times L_{-1,p'}(\mathbb{R}^d)}ds\right| \lesssim \|g\|_{\mathbb{L}^{q'}_{-1,p'}([0,1])}\|f\|_{\mathbb{L}^{q'}_{-1,p'}([0,1])}.$$

Since g is arbitrary, this implies (A.17). The result for equation (A.11) follows from similar arguments. □

In the remaining, we consider the parabolic differential equation

$$\partial_s u + \frac{1}{2} a^{ij} \partial^2_{ij} u = f, \quad u(t,\cdot) = 0 \tag{A.29}$$

where $f \in \mathbb{L}^q_{-1,p}([0,1])$, $t \in (0,1]$ is fixed. Whenever the dependence on t plays a role, we write $u^t_s(x)$ for the solution to (A.29) evaluated at (s,x), $s \leqslant t$, $x \in \mathbb{R}^d$. We quantify the dependence on the terminal time of various quantities related to u^t_s under Condition $\mathfrak{A}'$. These estimates are used in Section 6 where we particularly take $(a^{ij}) := \sigma\sigma^*$ for a σ satisfying Condition $\mathfrak{A}$ with $q_0 = \infty$.

THEOREM A.11. *Assuming Condition $\mathfrak{A}'$. Let $q \in (2,\infty)$ and $p \in (1,\infty)$ be such that $\frac{1}{p} + \frac{1}{p_0} < 1$. Then for every $\nu \in [0,1]$, every $0 \leqslant s \leqslant t \leqslant 1$ and $f \in \mathbb{L}^q_{-\nu,p}([0,1])$, we have*

$$\|u^t_s\|_{L_p(\mathbb{R}^d)} \leqslant N(t-s)^{1-\frac{\nu}{2}-\frac{1}{q}} \|f\|_{\mathbb{L}^q_{-\nu,p}([s,t])}. \tag{A.30}$$

LEMMA A.12. *Let $q \in (2,\infty)$ and $p \in (1,\infty)$. Let $t \in [0,1]$. If $u(r,x) = 0$ for $r \in [t,1]$, then for every $s \in [0,t]$*

$$\|u(s)\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{\frac{1}{2}-\frac{1}{q}} \|\partial_s u + \frac{1}{2}\Delta u\|_{\mathbb{L}^q_{-1,p}([s,t])} \text{ whenever } u \in \mathbb{L}^q_{1,p}([0,1]) \tag{A.31}$$

and

$$\|u(s)\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{1-\frac{1}{q}} \|\partial_s u + \frac{1}{2}\Delta u\|_{\mathbb{L}^q_p([s,t])} \text{ whenever } u \in \mathbb{L}^q_{2,p}([0,1]). \tag{A.32}$$

PROOF. By approximation, we can assume that u is a smooth function on $[0,1]\times\mathbb{R}^d$ with compact support. Put $g := \partial_s u + \frac{1}{2}\Delta u$. Then by Duhamel's formula $u(s,x) = \int_s^t P_{s,r} g(r,x) dr$. Applying Minkowski inequality and [67, Theorem 5.30], we have

$$\|u_s\|_{L_p(\mathbb{R}^d)} \leqslant \int_s^t \|P_{s,r} g_r\|_{L_p(\mathbb{R}^d)} dr \lesssim \int_s^t (r-s)^{-\frac{1}{2}} \|g_r\|_{L_{-,1,p}(\mathbb{R}^d)} dr.$$

Using Hölder inequality, we have $\|u_s\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{\frac{1}{2}-\frac{1}{q}} \|g\|_{\mathbb{L}^q_{-1,p}([s,t])}$. This shows (A.31). Inequality (A.32) is obtained in the same way. □

Proof of Theorem A.11. By interpolation, it suffices to show that

$$\|u^t_s\|_{L_p(\mathbb{R}^d)} \leqslant N(t-s)^{\frac{1}{2}-\frac{1}{q}} \|f\|_{\mathbb{L}^q_{-1,p}([s,t])} \quad \text{for } f \in \mathbb{L}^q_{-1,p}([0,1]), \tag{A.33}$$

$$\|u^t_s\|_{L_p(\mathbb{R}^d)} \leqslant N(t-s)^{1-\frac{1}{q}} \|f\|_{\mathbb{L}^q_p([s,t])} \quad \text{for } f \in \mathbb{L}^q_p([0,1]). \tag{A.34}$$

From Theorem A.5, we have $\|u^t\|_{\mathbb{L}^q_{1,p}([s,t])} \lesssim \|f\|_{\mathbb{L}^q_{-1,p}([s,t])}$ for every $0 \leqslant s \leqslant t \leqslant 1$. From (A.29), we have $\partial_s u + \frac{1}{2}\Delta u = f + \frac{1}{2}(\Delta - a^{ij}\partial^2_{ij})u$. It follows from Lemma A.12 that

$$\|u^t_s\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{\frac{1}{2}-\frac{1}{q}} (\|f\|_{\mathbb{L}^q_{-1,p}([s,t])} + \|(\delta^{ij} - a^{ij})\partial^2_{ij} u^t\|_{\mathbb{L}^q_{-1,p}([s,t])}),$$

where $\delta^{ij} = 1$ if $i = j$ and $\delta^{ij} = 0$ otherwise. From Lemma A.3(iii) and the hypotheses, we have

$$\begin{aligned}\|(\delta^{ij} - a^{ij})\partial^2_{ij} u^t\|_{\mathbb{L}^q_{-1,p}([s,t])} &\lesssim (\|a^{ij}\|_{\mathbb{L}^\infty_\infty([s,t])} + \|\nabla a^{ij}\|_{\mathbb{L}^\infty_{p_0}([s,t])}) \|\partial^2_{ij} u^t\|_{\mathbb{L}^q_{-1,p}([s,t])} \\ &\lesssim \|u^t\|_{\mathbb{L}^q_{1,p}([s,t])}.\end{aligned}$$

Combining the previous estimates, we obtain (A.33).

Inequality (A.34) is shown analogously. Indeed, using Lemma A.12, we have

$$\|u^t_s\|_{L_p(\mathbb{R}^d)} \lesssim (t-s)^{1-\frac{1}{q}}(\|f\|_{\mathbb{L}^q_p([s,t])} + \|(\delta^{ij}-a^{ij})\partial^2_{ij}u^t\|_{\mathbb{L}^q_p([s,t])})$$
$$\lesssim (t-s)^{1-\frac{1}{q}}(\|f\|_{\mathbb{L}^q_p([s,t])} + \|\partial^2_{ij}u^t\|_{\mathbb{L}^q_p([s,t])}).$$

It is known ([51, Theorem 2.1]) that $\|u^t\|_{\mathbb{L}^q_{2,p}([s,t])} \lesssim \|f\|_{\mathbb{L}^q_p([s,t])}$. These estimates imply (A.34). □

Acknowledgments. The authors would like to thank the editors, associated editors and the referees for their constructive suggestions.

Funding. The first author was supported by the Humboldt fellowship and the second author was supported by the DFG via Research Unit FOR 2402 while they were at TU Berlin where this work was carried out.

REFERENCES

[1] AIZENMAN, M. and SIMON, B. (1982). Brownian motion and Harnack inequality for Schrödinger operators. *Comm. Pure Appl. Math.* **35** 209–273. https://doi.org/10.1002/cpa.3160350206 MR644024

[2] DE ANGELIS, T., GERMAIN, M. and ISSOGLIO, E. (2022). A numerical scheme for stochastic differential equations with distributional drift. *Stochastic Processes and their Applications* **154** 55-90. https://doi.org/10.1016/j.spa.2022.09.003

[3] ANZELETTI, L., LÊ, K. and LING, C. (2023). Path-by-path uniqueness for stochastic differential equations under Krylov-Röckner condition. arXiv:2304.06802.

[4] ATHREYA, S., BUTKOVSKY, O., LÊ, K. and MYTNIK, L. (2024). Well-posedness of stochastic heat equation with distributional drift and skew stochastic heat equation. *Communications on Pure and Applied Mathematics* **77** 2708-2777. https://doi.org/10.1002/cpa.22157

[5] BAO, J., HUANG, X. and YUAN, C. (2019). Convergence rate of Euler-Maruyama scheme for SDEs with Hölder-Dini continuous drifts. *J. Theoret. Probab.* **32** 848–871. https://doi.org/10.1007/s10959-018-0854-9 MR3959630

[6] BAO, J., HUANG, X. and ZHANG, S.-Q. (2022). Convergence rate of the EM algorithm for SDEs with low regular drifts. *Journal of Applied Probability* **59** 447–470. https://doi.org/10.1017/jpr.2021.56

[7] BECK, L., FLANDOLI, F., GUBINELLI, M. and MAURELLI, M. (2019). Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electron. J. Probab.* **24** Paper No. 136, 72. https://doi.org/10.1214/19-ejp379 MR4040996

[8] BELLINGERI, C., FRIZ, P. K. and GERENCSÉR, M. (2021). Singular paths spaces and applications. *Stochastic Analysis and Applications* 1–24.

[9] BURRAGE, K., BURRAGE, P. M. and TIAN, T. (2004). Numerical methods for strong solutions of stochastic differential equations: an overview. **460** 373–402. Stochastic analysis with applications to mathematical finance. https://doi.org/10.1098/rspa.2003.1247 MR2052268

[10] BUTKOVSKY, O., DAREIOTIS, K. and GERENCSÉR, M. (2021). Approximation of SDEs: a stochastic sewing approach. *Probability Theory and Related Fields* 1–60.

[11] CHORIN, A. J. (1980). Vortex models and boundary layer instability. *SIAM Journal on Scientific and Statistical Computing* **1** 1–21.

[12] CONSTANTIN, P. and IYER, G. (2008). A stochastic Lagrangian representation of the three-dimensional incompressible Navier-Stokes equations. *Comm. Pure Appl. Math.* **61** 330–345. https://doi.org/10.1002/cpa.20192 MR2376844

[13] DAREIOTIS, K. and GERENCSÉR, M. (2020). On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift. *Electron. J. Probab.* **25** Paper No. 82, 18. https://doi.org/10.1214/20-ejp479 MR4125787

[14] DAREIOTIS, K., GERENCSÉR, M. and LÊ, K. (2023). Quantifying a convergence theorem of Gyöngy and Krylov. *The Annals of Applied Probability* **33** 2291 – 2323. https://doi.org/10.1214/22-AAP1867

[15] DAVIE, A. M. (2007). Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not. IMRN* **24** Art. ID rnm124, 26. https://doi.org/10.1093/imrn/rnm124 MR2377011

[16] DAVIE, A. M. (2008). Differential equations driven by rough paths: an approach via discrete approximation. *Appl. Math. Res. Express. AMRX* Art. ID abm009, 40. [Issue information previously given as no. 2 (2007)]. MR2387018

[17] FEDRIZZI, E. and FLANDOLI, F. (2013). Noise prevents singularities in linear transport equations. *J. Funct. Anal.* **264** 1329–1354. https://doi.org/10.1016/j.jfa.2013.01.003 MR3017266
[18] FLANDOLI, F., GUBINELLI, M. and PRIOLA, E. (2010). Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* **180** 1–53. https://doi.org/10.1007/s00222-009-0224-4 MR2593276
[19] FRAZIER, M., JAWERTH, B. and WEISS, G. (1991). *Littlewood-Paley theory and the study of function spaces*. *CBMS Regional Conference Series in Mathematics* **79**. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI. https://doi.org/10.1090/cbms/079 MR1107300
[20] FRIZ, P. K., HOCQUET, A. and LÊ, K. (2021). Rough stochastic differential equations. arXiv:2106.10340.
[21] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional stochastic processes as rough paths*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge University Press, Cambridge Theory and applications. https://doi.org/10.1017/CBO9780511845079 MR2604669
[22] GALEATI, L. and LING, C. (2023). Stability estimates for singular SDEs and applications. *Electron. J. Probab.* **28** Paper No. 24, 31. https://doi.org/10.1214/23-ejp913 MR4546633
[23] GILBARG, D. and TRUDINGER, N. S. (2001). *Elliptic partial differential equations of second order*. *Classics in Mathematics*. Springer-Verlag, Berlin Reprint of the 1998 edition. MR1814364
[24] GIUSTI, E. (2003). *Direct Methods in the Calculus of Variations*. World Scientific Pub.
[25] GYÖNGY, I. and KRYLOV, N. (1996). Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Theory Related Fields* **105** 143–158. https://doi.org/10.1007/BF01203833 MR1392450
[26] GYÖNGY, I. and KRYLOV, N. V. (2022). Existence of strong solutions for Itô's stochastic equations via approximations: revisited. *Stoch. Partial Differ. Equ. Anal. Comput.* **10** 693–719. https://doi.org/10.1007/s40072-022-00273-7 MR4491500
[27] GYÖNGY, I. and RÁSONYI, M. (2011). A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. *Stochastic Process. Appl.* **121** 2189–2200. https://doi.org/10.1016/j.spa.2011.06.008 MR2822773
[28] HINZ, M., TÖLLE, J. M. and VIITASAARI, L. (2023). Variability of paths and differential equations with -coefficients. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **59** 2036 – 2082. https://doi.org/10.1214/22-AIHP1308
[29] HU, Y., LÊ, K. and MYTNIK, L. (2017). Stochastic differential equation for Brox diffusion. *Stochastic Process. Appl.* **127** 2281–2315. https://doi.org/10.1016/j.spa.2016.10.010 MR3652414
[30] HUDDE, A., HUTZENTHALER, M. and MAZZONETTO, S. (2021). A stochastic Gronwall inequality and applications to moments, strong completeness, strong local Lipschitz continuity, and perturbations. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 603–626. https://doi.org/10.1214/20-aihp1064 MR4260476
[31] HUTZENTHALER, M. and JENTZEN, A. (2020). On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with nonglobally monotone coefficients. *Ann. Probab.* **48** 53–93. https://doi.org/10.1214/19-AOP1345 MR4079431
[32] HUTZENTHALER, M., JENTZEN, A. and KLOEDEN, P. E. (2012). Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22** 1611–1641. https://doi.org/10.1214/11-AAP803 MR2985171
[33] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes* **1**. North-Holland Publishing Company.
[34] JACOD, J. and PROTTER, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26** 267–307. https://doi.org/10.1214/aop/1022855419 MR1617049
[35] JOURDAIN, B. and MENOZZI, S. (2024). Convergence rate of the Euler-Maruyama scheme applied to diffusion processes with $L^q - L^\rho$ drift coefficient and additive noise. *Ann. Appl. Probab.* **34** 1663–1697. https://doi.org/10.1214/23-aap2006 MR4700268
[36] KHASMINSKII, R. Z. (1959). On positive solutions of the equation $\mathcal{U}u + Vu = 0$. *Theor. Probability Appl.* **4** 309–318. MR123373
[37] KIM, K.-H. (2008). $L_q(L_p)$-theory of parabolic PDEs with variable coefficients. *Bull. Korean Math. Soc.* **45** 169–190. https://doi.org/10.4134/BKMS.2008.45.1.169 MR2391465
[38] KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical solution of stochastic differential equations*. *Applications of Mathematics (New York)* **23**. Springer-Verlag, Berlin. https://doi.org/10.1007/978-3-662-12616-5 MR1214374
[39] KLOEDEN, P. E., PLATEN, E. and SCHURZ, H. (1994). *Numerical solution of SDE through computer experiments*. *Universitext*. Springer-Verlag, Berlin. https://doi.org/10.1007/978-3-642-57913-4 MR1260431
[40] KRYLOV, N. V. (2001). The heat equation in $L_q((0,T),L_p)$-spaces with weights. *SIAM J. Math. Anal.* **32** 1117–1141. https://doi.org/10.1137/S0036141000372039 MR1828321

[41] KRYLOV, N. V. and RÖCKNER, M. (2005). Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** 154–196. https://doi.org/10.1007/s00440-004-0361-z MR2117951

[42] KUNITA, H. (2019). *Stochastic flows and jump-diffusions*. *Probability Theory and Stochastic Modelling* **92**. Springer, Singapore. https://doi.org/10.1007/978-981-13-3801-4 MR3929750

[43] KURTZ, T. G. and PROTTER, P. (1991). Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic analysis* 331–346. Academic Press, Boston, MA. MR1119837

[44] LE GALL, J. F. (1984). One-dimensional stochastic differential equations involving the local times of the unknown process. In *Stochastic analysis and applications (Swansea, 1983)*. *Lecture Notes in Math.* **1095** 51–82. Springer, Berlin. https://doi.org/10.1007/BFb0099122 MR777514

[45] LÊ, K. (2020). A stochastic sewing lemma and applications. *Electron. J. Probab.* **25** Paper No. 38, 55. https://doi.org/10.1214/20-ejp442 MR4089788

[46] LÊ, K. (2022). Quantitative John–Nirenberg inequality for stochastic processes of bounded mean oscillation. *arXiv preprint arXiv:2210.15736*.

[47] LÊ, K. (2023). Stochastic sewing in Banach spaces. *Electronic Journal of Probability* **28** 1 – 22. https://doi.org/10.1214/23-EJP918

[48] LEMAIRE, V. and MENOZZI, S. (2010). On some non asymptotic bounds for the Euler scheme. *Electron. J. Probab.* **15** no. 53, 1645–1681. https://doi.org/10.1214/ejp.v15-814 MR2735377

[49] LEOBACHER, G. and SZÖLGYENYI, M. (2017). A strong order 1/2 method for multidimensional SDEs with discontinuous drift. *Ann. Appl. Probab.* **27** 2383–2418. https://doi.org/10.1214/16-AAP1262 MR3693529

[50] LEOBACHER, G. and SZÖLGYENYI, M. (2018). Convergence of the Euler-Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. *Numer. Math.* **138** 219–239. https://doi.org/10.1007/s00211-017-0903-9 MR3745015

[51] LING, C. and XIE, L. (2021). Strong Solutions of Stochastic Differential Equations with Coefficients in Mixed-Norm Spaces. *Potential Analysis*. https://doi.org/10.1007/s11118-021-09913-4

[52] LYONS, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14** 215–310. https://doi.org/10.4171/RMI/240 MR1654527

[53] MAKASU, C. (2020). Extension of a stochastic Gronwall lemma. *Bull. Pol. Acad. Sci. Math.* **68** 97–104. https://doi.org/10.4064/ba190524-7-5 MR4165506

[54] MAO, X. (2015). The truncated Euler-Maruyama method for stochastic differential equations. *J. Comput. Appl. Math.* **290** 370–384. https://doi.org/10.1016/j.cam.2015.06.002 MR3370415

[55] MENOUKEU PAMEN, O. and TAGUCHI, D. (2017). Strong rate of convergence for the Euler-Maruyama approximation of SDEs with Hölder continuous drift coefficient. *Stochastic Process. Appl.* **127** 2542–2559. https://doi.org/10.1016/j.spa.2016.11.008 MR3660882

[56] MEYER, Y. (1992). *Wavelets and operators*. *Cambridge Studies in Advanced Mathematics* **37**. Cambridge University Press, Cambridge Translated from the 1990 French original by D. H. Salinger. MR1228209

[57] MILSTEIN, G. N. and TRETYAKOV, M. V. (2004). *Stochastic numerics for mathematical physics*. *Scientific Computation*. Springer-Verlag, Berlin. https://doi.org/10.1007/978-3-662-10063-9 MR2069903

[58] MÜLLER-GRONBACH, T. and YAROSLAVTSEVA, L. (2020). On the performance of the Euler-Maruyama scheme for SDEs with discontinuous drift coefficient. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 1162–1178. https://doi.org/10.1214/19-AIHP997 MR4076779

[59] NEUENKIRCH, A. and SZÖLGYENYI, M. (2021). The Euler-Maruyama scheme for SDEs with irregular drift: convergence rates via reduction to a quadrature problem. *IMA J. Numer. Anal.* **41** 1164–1196. https://doi.org/10.1093/imanum/draa007 MR4246871

[60] NEVES, W. and OLIVERA, C. (2015). Wellposedness for stochastic continuity equations with Ladyzhenskaya-Prodi-Serrin condition. *NoDEA Nonlinear Differential Equations Appl.* **22** 1247–1258. https://doi.org/10.1007/s00030-015-0321-6 MR3399177

[61] POPE, S. B. (2002). Stochastic Lagrangian models of velocity in homogeneous turbulent shear flow. *Physics of fluids* **14** 1696–1702.

[62] PORTENKO, N. I. (1975). Diffusion processes with an unbounded drift coefficient. *Teor. Verojatnost. i Primenen.* **20** 29–39. MR0375483

[63] PORTENKO, N. I. (1990). *Generalized diffusion processes*. *Translations of Mathematical Monographs* **83**. American Mathematical Society, Providence, RI Translated from the Russian by H. H. McFaden. https://doi.org/10.1090/mmono/083 MR1104660

[64] SCHEUTZOW, M. (2013). A stochastic Gronwall lemma. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **16** 1350019, 4. https://doi.org/10.1142/S0219025713500197 MR3078830

[65] SIORPAES, P. (2018). Applications of pathwise Burkholder-Davis-Gundy inequalities. *Bernoulli* **24** 3222–3245. https://doi.org/10.3150/17-BEJ958 MR3788172

[66] STROOCK, D. W. and VARADHAN, S. R. S. (2006). *Multidimensional diffusion processes*. *Classics in Mathematics*. Springer-Verlag, Berlin Reprint of the 1997 edition. MR2190038

[67] TRIEBEL, H. (2013). *Local function spaces, heat and Navier-Stokes equations*. *EMS Tracts in Mathematics* **20**. European Mathematical Society (EMS), Zürich. https://doi.org/10.4171/123 MR3086433

[68] VERETENNIKOV, A. J. (1979). Strong solutions of stochastic differential equations. *Teor. Veroyatnost. i Primenen.* **24** 348–360. MR532447

[69] XIA, P., XIE, L., ZHANG, X. and ZHAO, G. (2020). $L^q(L^p)$-theory of stochastic differential equations. *Stochastic Process. Appl.* **130** 5188–5211. https://doi.org/10.1016/j.spa.2020.03.004 MR4108486

[70] ZHANG, X. (2010). A stochastic representation for backward incompressible Navier-Stokes equations. *Probab. Theory Related Fields* **148** 305–332. https://doi.org/10.1007/s00440-009-0234-6 MR2653231

[71] ZHANG, X. (2011). Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** no. 38, 1096–1116. https://doi.org/10.1214/EJP.v16-887 MR2820071

[72] ZHANG, X. and ZHAO, G. (2017). Heat kernel and ergodicity of SDEs with distributional drifts. arXiv:1710.10537.

[73] ZHANG, X. and ZHAO, G. (2021). Stochastic Lagrangian path for Leray's solutions of 3D Navier-Stokes equations. *Comm. Math. Phys.* **381** 491–525. https://doi.org/10.1007/s00220-020-03888-w MR4207449

[74] ZVONKIN, A. K. (1974). A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)* **93(135)** 129–149, 152. MR0336813