QUANTITATIVE JOHN–NIRENBERG INEQUALITY FOR STOCHASTIC PROCESSES OF BOUNDED MEAN OSCILLATION

KHOA Lê

Abstract. Stroock and Varadhan in 1997 and Geiss in 2005 independently introduced stochastic processes with bounded mean oscillation (BMO) and established their exponential integrability with some unspecified exponential constant. This result is an analogue of the John–Nirenberg inequality for functions of bounded mean oscillation. In this work, we quantify the size of the exponential constant by the modulus of mean oscillation. Some new applications of BMO processes in rough stochastic differential equations, numerical approximations and regularization by noise are discussed.

Mathematics Subject Classification (2020): Primary 60G07, Secondary 60H50, 60H35, 60H10.

Keywords: BMO processes; VMO processes; John–Nirenberg inequality; regularization by noise

1. Introduction

A real-valued locally integrable function \( f \) defined on \( \mathbb{R}^d \) is of bounded mean oscillation (BMO) if

\[
\|f\|_{\text{BMO}} := \sup_{Q} \frac{1}{|Q|} \int_Q |f(x) - f_Q| \, dx < \infty,
\]

where the supremum is taken over all cubes \( Q \) in \( \mathbb{R}^d \), \( |Q| \) denotes the Lebesgue measure of \( Q \) and \( f_Q = \frac{1}{|Q|} \int_Q f(x) \, dx \). For such function, John and Nirenberg in [JN61] show that

\[
\sup_{Q} \frac{1}{|Q|} \int_Q e^{\lambda|f - f_Q|} \, dx < \infty \tag{1.1}
\]

for some constant \( \lambda > 0 \). The largest constant \( \lambda \) such that (1.1) holds is denoted by \( \lambda(f) \) and can be quantified by the distance between \( f \) and the space of essentially bounded functions \( L^\infty(\mathbb{R}^d) \) via the Garnett–Jones theorem [GJ78] which asserts that

\[
\frac{1}{C(d)} \frac{1}{\lambda(f)} \leq \inf_{g \in L^\infty(\mathbb{R}^d)} \|f - g\|_{\text{BMO}} \leq C(d) \frac{1}{\lambda(f)} \tag{1.2}
\]

for some constant \( C(d) \).

Stochastic processes of bounded mean oscillation are considered by Stroock and Varadhan [SV06] and independently by S. Geiss in [Gei05]. To give the precise definition, let \( (\Omega, \mathcal{G}, \mathbb{P}) \) be a probability space equipped with a filtration \( \{\mathcal{F}_t\}_{t \geq 0} \) satisfying the usual conditions. Let
\( \tau > 0 \) be a fixed number. For each stopping time \( S \), \( \mathbb{E}_S \) denotes the conditional expectation with respect to \( \mathcal{F}_S \) and for each \( G \in \mathcal{G} \), \( \mathbb{P}_S(G) := \mathbb{E}_S(1_G) \). For each \( m \in [1, \infty] \), \( \| \cdot \|_m \) denotes the norm in \( L^m(\Omega, \mathcal{G}, \mathbb{P}) \).

**Definition 1.1.** Let \((V_t)_{t \in [0, \tau]}\) be a real valued adapted right continuous process with left limits (RCLL). \( V \) is of bounded mean oscillation (BMO) if

\[
[V]_{\text{BMO}} := \sup_{0 \leq S \leq T \leq \tau} \| \mathbb{E}_S[V_T - V_S] \|_\infty < \infty
\]  

(1.3)

where the supremum is taken over all stopping times \( S, T \); \( V_{S-} = \lim_{r \uparrow S} V_r \) and we set \( V_{0-} = V_0 \) by convention.

**Remark 1.2.** In [Gei05], BMO processes are defined so that (1.3) holds with \( T = \tau \). This definition is equivalent to ours due to triangle inequality and the fact that \( \mathbb{E}_S|V_T - V_S| = \lim_{\epsilon \downarrow 0} \mathbb{E}_S|V_T - V_{S+\epsilon-}| \).

**Remark 1.3.** Herein, we focus on estimations for moments of BMO processes and for this problem, there is no loss of generality when restricting to real valued processes. Indeed, if \( Z \) is an adapted RCLL process taking values in some metric space \((\mathcal{E}, d)\) such that

\[
\sup_{s \leq T \leq \tau} \| \mathbb{E}_s d(V_{S-}, V_T) \|_\infty < \infty
\]

then the processes \( Z = d(V_0, V) \) is a real valued BMO process as of Definition 1.1. This is an immediate consequence of the triangle inequality. In fact, the maximal process \( Z_t^\tau = \sup_{s \leq t} |Z_s| \) is also of BMO, see Proposition 2.8 below.

For BMO processes, Geiss shows in [Gei05, Theorem 1] that

\[
\sup_{s \in [0, \tau]} \| \mathbb{E}_s e^{\lambda \sup_{t \in [0, \tau]} |V_t - V_{S-}|} \|_\infty < \infty
\]

(1.4)

for some constant \( \lambda > 0 \), which is an analogue of the John–Nirenberg inequality (1.1) for BMO functions. For continuous BMO processes, (1.4) was shown earlier in the last pages of the book [SV06]. Stroock–Varadhan called out John–Nirenberg inequality however they did not name BMO processes. The largest constant \( \lambda \) such that (1.4) holds is denoted by \( \lambda(V) \). Quantitative estimates for \( \lambda(V) \) have not been considered, as far as the author’s knowledge. Nevertheless, Varopoulos in [Var80] shows the following estimates for BMO martingales

\[
\frac{C_1}{\lambda(V)} \leq \inf_{\psi \in L^\infty(\Omega)} \| V - \mathbb{E}_V \|_{\text{BMO}} \leq \frac{C_2}{\lambda(V)}.
\]

(1.5)

for some constants \( C_1, C_2 > 0 \), which is a probabilistic Garnett–Jones theorem.

BMO processes and the John–Nirenberg inequality (1.4) have been utilized effectively in theory of singular integrals by Stroock [Str73], in financial mathematics and backward SDEs by Geiss and coauthors in [GY20, Gei05, GN20]. It is remarkable that BMO processes appear inconspicuously in many other problems, in [Dav07, ABLM20, Lê22, GG22] on regularization by noise phenomenon, in [DG20, LL21, DGL22] on strong convergence rate of numerical methods for stochastic differential equations, in [FHL21] on rough stochastic differential equations. In
these occurrences, BMO property was not identified and the connection with John–Nirenberg inequality was not established. We give two examples of BMO processes which arise from these applications.

**Example 1.4.** (a) Let \((X^x_t)_{t \geq 0, x \in \mathbb{R}^d}\) be a Markov process and let \(f : [0, \tau] \times \mathbb{R}^d \to \mathbb{R}\) be a measurable function. Suppose that one has the uniform Krylov estimate

\[
\sup_{0 \leq s \leq t \leq \tau} \sup_{x \in \mathbb{R}^d} \mathbb{E} |\int_s^t f(r, X^x_{r-s})dr| \leq C
\]

for some finite constant \(C\). If furthermore the process \(K_t := \int_0^t f(r, X^0_r)dr\) is a.s. continuous then it is BMO. (Indeed, by Markov property, \(\mathbb{E}_t[K_t - K_s] = \mathbb{E}\left[\int_s^t f(r, X^x_{r-s})dr\right]_{x = X^0_t} \leq C\) so that \(K\) is BMO by Proposition 2.2 herein.) Krylov estimate is an important tool in the study of stochastic differential equations (SDEs), see for instance [RZ21, Zha16].

(b) Let \(f : [0, 1] \times \mathbb{R}^d \to \mathbb{R}\) be a Borel function which is uniformly bounded by 1 and \(B\) be a standard Brownian motion in \(\mathbb{R}^d\). Define for each integer \(n \geq 1\), the process \(V^n_t = \int_0^t [f(r, B_r) - f(r, B_{[nr]/n})]dr, \ t \in [0, 1]\), which corresponds to the error of the quadrature rule approximating the functional \(\int_0^1 f(r, B_r)dr\). Then, \((V^n_t)_{t \in [0, 1]}\) is BMO with \([V^n]_{BMO} \leq (N \log(n+1)/n)^{1/2}\) for some finite constant \(N\). Indeed, it is shown by Dareiotis and Gerencsér in [DG20] (see Lemma 2.1 therein) that

\[
\text{ess sup}_\omega (\mathbb{E}_\omega |V^n_s - V^n_t|^2)^{1/2} \leq (N \log(n+1)/n)^{1/2}\text{ for every } s \leq t \leq 1.
\]

Without the conditional expectation, (1.7) is also obtained by Altmeyer in [Alt21]. The authors of both works rely on explicit moment computations and therefore are restricted to the second moment. Quadrature error estimates such as the above are directly related to the strong convergence rate of the Euler–Maruyama scheme for SDEs. It is of important interests to have quadrature error estimates for all moments. We will revisit these processes in Example 2.6.

The current article provides two main contributions.

**I.** We provide a lower bound for \(\lambda(V)\) in terms of the *modulus of mean oscillation* of \(V\), which is new even for BMO martingales (see Remark 2.4). This kind of result is inspired by Portenko’s formulation of the Khasminskii’s lemma for increasing processes, see Chapter I.1 in [Por90]. While Garnett–John inequality (1.5) is a beautiful theoretical result, our estimate is a practical one because moduli of mean oscillation are readily available in most applications and therefore can be applied directly. Based upon this estimate, we provide quantitative exponential integrability for processes with *vanishing mean oscillation*, which is inspired in part by Lyons’ precise estimate for multiplicative functionals in [Lyo98]. The obtained results are new and lie outside the scope of (1.4) and (1.5).

**II.** We bring to light the role of BMO processes and John–Nirenberg inequality in various problems in regularization-by-noise phenomenon, numerical approximation for SDEs and rough stochastic differential equations. This is far from being mundane and we illustrate by three applications, in which knowledge of BMO processes effectively shorten existing proofs.
and improve results. The first one is about exponential moments of a class of stochastic controlled rough paths considered in [FHL21]. In the second application, we explain that Davie’s exponential estimate ([Dav07]) can be simplified and deduced from a single estimate for the second moment. The last application is about strong convergence rate of the tamed Euler–Maruyama scheme for SDEs with integrable drifts. Such rate has been obtained previously in [LL21] however only for small moments. Here, we are able to derive the same rate for all moments using John–Nirenberg inequality in conjunction with the recent stability estimate of Galeati and Ling in [GL22].

Further literature. Fefferman in [Fef71] identified the space of BMO functions as the dual of the Hardy space $H_1$. Getoor and Sharpe introduced among other things in [GS72] continuous time BMO martingales and established the duality of $H_1$ and BMO martingales, which is a probabilistic analogue of Fefferman’s earlier result. This duality for discrete time martingales is due to Fefferman and Stein [FS72], Garsia and Herz [Gar73a, Her74]. Results for continuous time BMO martingales are summarized in [Kaz94]. See [BN08, DMS+97, Gei05] for further applications in financial mathematics.

Open problems. Several challenging questions remain open. Geiss considered in [Gei05] weighted BMO and established a John–Nirenberg-type inequality for these processes. It is interesting to obtain analogous results presented herein to these processes. Another problem is to establish the Garnett–John theorem for BMO processes, extending [Var80]. Furthermore, Fefferman-type duality for BMO processes seems unexplored.

We conclude the introduction with the layout of the paper. Results for BMO and VMO processes are presented in Sections 2 and 3 respectively. We also provide another proof of Geiss’ estimate (1.4). The applications are discussed in Section 4. The appendix contains two auxiliary results which are well-known but are adjusted to our setting.

2. BMO processes

We recall Definition 1.1 of BMO processes. For each BMO process $V$, we define its modulus of mean oscillation $\rho(V) : \{(s,t) \in [0,\tau]^2 : s \leq t\} \rightarrow [0,\infty)$ by

$$
\rho_{s,t}(V) = \sup_{s \leq S \leq T \leq t} \|E_S[V_T - V_S^-]\|_\infty, \quad 0 \leq s \leq t \leq \tau,
$$

where the supremum is taken over all stopping times $S,T$ satisfying $s \leq S \leq T \leq t$. The function $\rho(V)$ is monotone in the sense that

$$
\rho_{u,v}(V) \leq \rho_{s,t}(V) \quad \text{whenever} \quad [u,v] \subset [s,t].
$$

(2.1)

We define the function $\kappa(V) : \{(s,t) \in [0,\tau]^2 : s \leq t\} \rightarrow [0,\infty)$ by the relation

$$
\kappa_{s,t}(V) = \lim_{h \downarrow 0} \sup_{s \leq u \leq t, v-u \leq h} \rho_{u,v}(V).
$$

(2.2)

It is easy to see that the above limit always exists, that $\kappa(V)$ is monotone and $\kappa_{s,t}(V) \in [0,\rho_{s,t}(V)]$ for each $s \leq t$.

An immediate consequence of (1.3) is that all jumps of a BMO process are uniformly bounded by $\kappa$. 

4
**Proposition 2.1.** If $V$ is BMO then $\| \sup_{t \in [0, \tau]} |V_t - V_{t-}| \|_\infty \leq \kappa_{0, \tau}(V)$.

**Proof.** By definitions, $\| \mathbb{E}_S |V_T - V_{S-}| \|_\infty \leq \rho_{0, \tau}(V)$ for all stopping times in $[0, \tau]$. Taking $T = S$ yields that $\sup_S |V_S - V_{S-}|(\omega) \leq \rho_{0, \tau}(V)$ for all $\omega$ in a set of full probability. For such $\omega$ and for each $r \in [0, \tau]$ such that $V_r(\omega) \neq V_{r-}(\omega)$, there is a stopping time $S$ such that $S(\omega) = r$ (Prop. 2.26 [KS91]). With this stopping time, one deduces that $|V_T - V_{T-}|(\omega) \leq \rho_{0, \tau}(V)$. The argument can be applied over arbitrary sub-intervals of $[0, \tau]$. Hence, for all $\omega$ in a set of full probability, for every rationals $s \leq t$ in $[0, \tau]$ and every $r \in [s, t]$, we have $|V_r - V_{r-}|(\omega) \leq \rho_{s,t}(V)$. This implies the result. □

If $V$ is RCLL and has uniformly bounded jumps, one can replace the stopping times in (1.3) by deterministic times.

**Proposition 2.2.** Let $V$ be an adapted RCLL process and assume that

$$\sup_{s \leq t \leq \tau} \| \mathbb{E}_S |V_t - V_{S-}| \|_\infty \leq B \quad \text{and} \quad \sup_{t \in [0, \tau]} |V_t - V_{t-}| \leq C$$

for some finite constants $B, C$. Then for every stopping times $S \leq T \leq \tau$, one has

$$\mathbb{E}_S |V_T - V_{S-}| \leq 2B + 3C.$$  \hfill (2.3)

Consequently, $V$ is BMO.

**Proof.** Let $S \leq \tau$ be a stopping time and assume that $S$ takes finitely many values $\{s_1 < \ldots < s_k\}$. We have

$$\mathbb{E}_S |V_T - V_{S-}| = \sum_j \mathbb{E}_S [|V_T - V_{S-}| 1_{(S= s_j)}] = \sum_j 1_{(S= s_j)} \mathbb{E}_{s_j} [|V_T - V_{S_j-}|] \leq B.$$  

Every general stopping time $S$ is the decreasing limit of a sequence of discrete stopping times $S^n$. Without loss of generality, we assume that $S^n \leq \tau + 1$ and define $V_t = V_T$ for all $t \geq \tau$. Then by triangle inequality

$$\mathbb{E}_S [|V_{T+1} - V_{S-} | \wedge N] \leq \mathbb{E}_S [|V_{T+1} - V_{S^n-} |] + \mathbb{E}_S [|V_{S^n-} - V_{S-} | \wedge N] \leq B + \mathbb{E}_S [|V_{S^n-} - V_{S-} | \wedge N].$$

Note that $\lim_n V_{S^n-} = V_S$ so that by Fatou lemma and Lebesgue dominated convergence theorem,

$$\mathbb{E}_S [|V_{T+1} - V_{S-} | \wedge N] \leq B + \mathbb{E}_S [|V_S - V_{S-} | \wedge N] \leq B + C.$$  

Sending $N \to \infty$ yields

$$\mathbb{E}_S |V_T - V_{S-}| \leq B + C.$$  \hfill (2.4)

Let $S \leq T \leq \tau$ be stopping times. We have by triangle inequality that

$$\mathbb{E}_S |V_T - V_{S-}| \leq \mathbb{E}_S |V_T - V_{S-}| + \mathbb{E}_S |V_T - V_{T-}| + \mathbb{E}_S |V_T - V_{T-}|.$$  

Hence, applying (2.4) and the assumptions, we obtain (2.3). □
Theorem 2.3 (John–Nirenberg inequality). Let $V$ be a BMO process and let $r$ be a fixed number in $[0, \tau]$. Then

$$\|\mathbb{E}_r \sup_{r \leq t \leq \tau} |V_t - V_r|^p\|_\infty \leq p!(11\rho_{r, \tau}(V))^p$$

for every integer $p \geq 1$, \hspace{1cm} (2.5)

and

$$\mathbb{E}_r e^{\lambda \sup_{r \leq t \leq \tau} |V_t - V_r|} < \infty \text{ for every } \lambda < (11\kappa_{r, \tau}(V))^{-1}. \hspace{1cm} (2.6)$$

Remark 2.4. Comparing with (1.4), the estimate (2.6) is more precise because it relates the range of the exponential constant $\lambda$ with the function $\kappa$, which leads to new exponential estimates in the following section. The role of $\kappa$ in (2.6) is intrinsic to stochastic processes over finite time domains. To be more precise, we recall that a continuous martingale $(X_t)_{t \geq 0}$ is BMO if

$$[X]^{2}_{\text{BMO}} := \sup_S \|\mathbb{E}_S [X_\infty - X_S]^2\|_\infty < \infty,$$

where the supremum is taken over all stopping times $S$. Osękowski shows in [Ose15] that the inequality

$$\mathbb{E} e^{\lambda \sup_{t \geq 0} |X_t - X_0|} \leq \int_0^\infty e^{\lambda [X]^{2}_{\text{BMO}} e^{-t}} dt \hspace{1cm} (\lambda > 0) \hspace{1cm} (2.7)$$

is true and sharp, i.e. there is a martingale $X$ with $0 < [X]^2_{\text{BMO}} < \infty$ for which both sides are equal. One sees that $\kappa(X)$ plays no role whatsoever for BMO processes over infinite time horizon. In the other direction, let $(V_t)_{t \in [0, \tau]}$ be a continuous BMO martingale and define $X_t = V_{t\wedge \tau}$. Then $(X_t)_{t \geq 0}$ is a continuous BMO martingale on the whole positive axis and is subjected to (2.7). In particular, one sees that $\mathbb{E} e^{\lambda \sup_{t \in [0, \tau]} |V_t - V_0|}$ is finite if $\lambda [V_{\tau}]^{2}_{\text{BMO}} < 1$. Since $[V_{\tau}]^{2}_{\text{BMO}}$ is comparable to $\rho_{[0, \tau]}(V)$ (by John–Nirenberg inequality), we deduce that $\mathbb{E} e^{\lambda \sup_{t \in [0, \tau]} |V_t - V_0|}$ is finite if $\lambda \rho_{[0, \tau]}(V) < c$ for some universal constant $c$. This is much more restrictive than the condition $\lambda \kappa_{[0, \tau]}(V) < 11$ provided by Theorem 2.3, especially for processes of vanishing mean oscillation (cf. Theorem 3.4).

Combining with (1.5) and recalling the notation $\lambda(V)$ from the introduction, one immediately obtains

Corollary 2.5. There is a constant $C_3 > 0$ such that for all BMO martingales $V$,

$$\lambda(V) \geq (11\kappa_{0, \tau}(V))^{-1} \hspace{1cm} \text{and} \hspace{1cm} \inf_{\psi \in L^\infty(\Omega)} \|V - \mathbb{E}_r \psi\|_{\text{BMO}} \leq C_3 \kappa_{0, \tau}(V).$$

We revisit the example from the introduction and discuss the implication of Theorem 2.3.

Example 2.6. We recall Example 1.4.

(a) From Theorem 2.3, we have

$$\text{ess sup}_{\omega} \mathbb{E}_s \left( \sup_{r \in [s, \tau]} \left| \int_s^t f(r, X_r^0) dr \right|^p \right) \leq p!(11C)^p$$

for every integer $p \geq 1$ and every $s \in [0, \tau]$. The John–Nirenberg inequality for BMO processes can thus be considered as a passage from uniform Krylov estimate to moment estimates of
all orders. Such passage has been known previously only when $f$ is non-negative through Khasminskii’s lemma, or when $f$ is a distribution through the stochastic sewing lemma from [Le20] under some additional constraints on the modulus of mean oscillation (see also Remark 5.3 in [ABLM20]).

(b) Theorem 2.3 implies the following new estimate without any book-keeping calculations of high moments

$$\text{ess sup}_{\omega} \left( \mathbb{E}_s \sup_{s \leq t \leq 1} \left| \int_s^t [f(r, B_r) - f(r, B_{[nr]/n})] dr \right|^p \right) \leq p!(121N \log(n + 1)/n)^{\frac{p}{2}} \tag{2.9}$$

for every integer $p \geq 1$ and every $s \leq 1$.

Later in Theorem 3.4, we will see that when the modulus of mean oscillation can be quantified, one can improve the growth constant $p$! in (2.8) and (2.9).

The proof of Theorem 2.3 relies on the following two intermediate results.

Proposition 2.7. Let $(A_r)_{r \in [0, \tau]}$ be a BMO process which is non-decreasing. Then for every $r \in [0, \tau]$ and every $\lambda < (\kappa_{r, r}(A))^{-1}$, there is a finite constant $c = c(\lambda, \rho(A)|_{[r, 1]^2})$ such that $\mathbb{E}_r e^{\lambda (A_r - A_s)} \leq c$.

Proof. It suffices to show the result for $r = 0$. By assumption, $\|\mathbb{E}_s (A_t - A_{S_t})\|_{\infty} \leq \rho_{s,t}(A)$ for every times $s \leq t \leq \tau$ and every stopping time $s \leq S \leq t$. We apply the energy inequality, Lemma A.2, to obtain that $\|\mathbb{E}_s[(A_t - A_s)^p] \leq p!(\rho_{s,t}(A))^p$ for every $s \leq t \leq \tau$ and every integer $p \geq 1$. For each $\lambda < (\kappa_{0, \tau}(A))^{-1}$, there is an $h_0 > 0$ such that $\lambda \rho_{s,t}(A) < 1$ whenever $t - s \leq h_0$. For such $s$, $t$, we have by Taylor’s expansion that

$$\|\mathbb{E}_s e^{\lambda (A_t - A_s)}\|_{\infty} \leq (1 - \lambda \rho_{s,t}(A))^{-1} < \infty.$$ 

Now partition $[0, \tau]$ by points $0 = t_0 < t_1 < \ldots < t_n = \tau$ so that $\max_{1 \leq k \leq n}(t_k - t_{k-1}) \leq h_0$. Then

$$\mathbb{E}_r e^{\lambda (A_t - A_r)} = \mathbb{E}_r e^{\lambda (A_{t_1} - A_r)} e^{\lambda (A_{t_1} - A_{t_0})} = \mathbb{E}_r e^{\lambda (A_{t_1} - A_r)} \mathbb{E}_{t_1} e^{\lambda (A_{t_1} - A_{t_0})} \leq \mathbb{E}_r e^{\lambda (A_{t_1} - A_r)} \|\mathbb{E}_{t_1} e^{\lambda (A_{t_1} - A_{t_0})}\|_{\infty}.$$ 

Iterating the previous inequality yields

$$\|\mathbb{E}_r e^{\lambda (A_t - A_r)}\|_{\infty} \leq \|\mathbb{E}_r e^{\lambda (A_{t_j} - A_r)}\|_{\infty} \prod_{k=j+1}^n \|\mathbb{E}_{t_{k-1}} e^{\lambda (A_{t_k} - A_{t_{k-1}})}\|_{\infty},$$

where $j$ is such that $t_{j-1} \leq r < t_j$. This implies the bound

$$\|\mathbb{E}_r e^{\lambda (A_t - A_r)}\|_{\infty} \leq \prod_{k=1}^n (1 - \lambda \rho_{t_{k-1}, t_k}(A))^{-1},$$

which yields the result. □

Proposition 2.8. Let $V$ be a BMO process. Define $V^*_t = \sup_{s \leq t} |V_s - V_0|$. Then $V^*$ is BMO with $\rho(V^*) \leq 11 \rho(V)$.
Proof. Fix \( s \leq t \). For stopping times \( s \leq S \leq t \), we have

\[
\mathbb{E}_s [V_T - V_{S-}] \leq c \text{ with } c = \rho_{s,t}(V).
\]

We define \( D_{s,t} = \sup_{s \leq r \leq t} |V_r - V_{s-}| \) and apply Lemma A.1 to obtain that

\[
\beta \mathbb{P}_s ((D_{s,t} - \beta)^+ \geq \alpha) \leq \beta \mathbb{P}_s (D_{s,t} \geq \alpha + \beta) \leq \rho_{s,t}(V) \mathbb{P}_s (D_{s,t} \geq \alpha)
\]

for every \( \alpha, \beta > 0 \). It follows that \( \beta \mathbb{E}_s [(D_{s,t} - \beta)^+] \leq c \mathbb{E}_s D_{s,t} \). Choosing \( \beta = 2c \), we have

\[
2c \mathbb{E}_s (D_{s,t} - 2c) \leq c \mathbb{E}_s D_{s,t}, \text{ that is } \mathbb{E}_s \sup_{s \leq r \leq t} |V_r - V_{s-}| \leq 4 \rho_{s,t}(V).
\]

Combining with the elementary estimate \( V_{t-} - V_{s-} \leq \sup_{s \leq r \leq t} |V_t - V_{r-}| \), we obtain that

\[
\mathbb{E}_s |V_t^* - V_{s-}^*| \leq 4 \rho_{s,t}(V).
\]

We also have \( V_{t-}^* - V_{s-}^* \leq V_t - V_s - \) for every \( t \leq \tau \) and hence an application of Proposition 2.2 yields the result. \( \square \)

Proof of Theorem 2.3. It suffices to show the result for \( r = 0 \). Define \( A_t = \sup_{\tau \in [0,t]} |V_{\tau} - V_0| \). By Proposition 2.8, \( A \) is BMO and \( \rho(A) \leq 11 \rho(V) \). We obtain (2.5) and (2.6) by applying Lemma A.2 and Proposition 2.7 respectively. \( \square \)

3. VMO processes

Definition 3.1. A BMO process \((V_t)_{t \in [0,\tau]}\) is VMO if \( \kappa_{0,\tau}(V) = 0 \).

An immediate consequence of Proposition 2.1 is that every VMO process has continuous sample paths. It is straightforward to see that the class of VMO processes starting from 0 forms a closed subspace of the space of BMO processes starting from 0. To quantify the regularity of VMO processes, we propose two additional subclasses.

Definition 3.2. Let \((V_t)_{t \in [0,\tau]}\) be a VMO process, \( p \in [1,\infty) \) and \( \alpha \in (0,1) \) be some fixed numbers. We say that \( \rho(V) \) contains only constant pro-

\[
[V]_{\text{VMO}^p_{\text{var}},[0,\tau]} := \left( \sup_{\pi \in \mathcal{P}([0,\tau])} \left[ \sum_{[s,t] \in \pi} |\rho_{s,t}(V)|^p \right] \right)^{1/p} < \infty
\]

where \( \mathcal{P}([0,\tau]) \) is the set of all partitions on \([0,\tau]\). We say that \( \rho(V) \) contains only constant pro-

\[
[V]_{\text{VMO}^\alpha,[0,\tau]} := \sup_{0 \leq s < t \leq \tau} \frac{\rho_{s,t}(V)}{(t-s)^\alpha} < \infty.
\]

It is evident that \( \text{VMO}^\alpha \subset \text{VMO}^{1/\alpha}_{\text{var}} \) and that

\[
[V]_{\text{VMO}^{1/\alpha}_{\text{var}},[0,\tau]} \leq \tau^\alpha [V]_{\text{VMO}^\alpha,[0,\tau]}.
\]

When \( \alpha > 1 \) and \( 0 < p < 1 \), the spaces \( \text{VMO}^\alpha \) and \( \text{VMO}^p_{\text{var}} \) contain only constant pro-

processes. This can be verified rather directly or alternatively using (3.2) below.
Proposition 3.3. Let \((V_t)_{t \in [0, T]}\) be a process in \(\text{VMO}^{p-\text{var}}\) and define the function

\[ (s, t) \mapsto w_{s,t}(V) := ([V]_{\text{VMO}^{p-\text{var}}; [s, t]})^p. \]

Then \(w(V) : \{(s, t) \in [0, T]^2 : s \leq t\} \rightarrow [0, \infty)\) is a control, i.e. \(w(V)\) is continuous and satisfies \(w_{s,u}(V) + w_{u,t}(V) \leq w_{s,t}(V)\) whenever \(s \leq u \leq t\) (super-additivity).

Proof. That \(w_{s,u}(V) + w_{u,t}(V) \leq w_{s,t}(V)\) whenever \(s \leq u \leq t\) is evident from definitions. Next, we verify that \(w(V)\) is continuous in several steps.

Step 1. We explain that \(\rho(V)\) is continuous. Whenever \(s \leq u \leq t\), we have by triangle inequality that

\[ \|E_t|V_t - V_s\|_\infty \leq \|E_s|V_u - V_s\|_\infty + \|E_u|V_t - V_u\|_\infty. \]

This implies that

\[ \rho_{s,t}(V) \leq \rho_{s,u}(V) + \rho_{u,t}(V). \quad (3.1) \]

Hence,

\[ |\rho_{s,t}(V) - \rho_{s+h,t-h}(V)| = \lim_{u \downarrow s} |\rho_{s,t}(V) - \rho_{u,t}(V)| \leq \lim_{u \downarrow s} \rho_{s,u}(V) = 0 \]

and

\[ |\rho_{s-h,t-h}(V) - \rho_{s,t}(V)| = \lim_{r \uparrow s} |\rho_{r,t}(V) - \rho_{s,t}(V)| \leq \lim_{r \uparrow s} \rho_{r,t}(V) = 0 \]

which show that \(\rho(V)\) is continuous in the former argument. Similarly, one can show continuity in the later argument. Thus \(\rho(V)\) is continuous.

Step 2. We show that \(w(V)\) is continuous from the inside, i.e. \(w_{s,t}(V) = w_{s+\varepsilon,t}(V)\). Fix \(s < u < t\) and a small number \(h > 0\). We have by super-additivity

\[ w_{s+\varepsilon,u-h}(V) + w_{u+\varepsilon,t-h}(V) \leq w_{s+\varepsilon,t}(V). \]

Sending \(h\) to 0, we see that \(\tilde{w}_{s,t} := w_{s+\varepsilon,t}(V)\) is super-additive. From the estimate \(\rho_{s+\varepsilon,t-h}(V) \leq w_{s+\varepsilon,t-h}(V)\), we also have \(\rho_{s,t}(V) \leq \tilde{w}_{s,t}\). From definition of \(w(V)\), this implies that \(w(V) \leq \tilde{w}\). It is obvious that \(\tilde{w} \leq w(V)\) and hence \(\tilde{w} = w(V)\), showing continuity from the inside.

Step 3. We show that \(w_{s,s+\varepsilon} = 0\). From \(w_{s,u}(V) + w_{u,t}(V) \leq w_{s,t}(V)\), we send \(u \downarrow s\) to get \(w_{s+\varepsilon}(V) + w_{s+\varepsilon,t}(V) \leq w_{s,t}(V)\). Using continuity from the inside, we obtain the claim.

Step 4. We show that \(w(V)\) is continuous from the outside, i.e. \(w_{s,t}(V) = w_{s-\varepsilon,t}(V)\). We fix \(s < t\) and \(h, \varepsilon > 0\) and consider \(\pi = \{t_i\}_{i=1}^n \in \mathcal{P}([s, t + h])\) such that

\[ \sum_{i=1}^n |\rho_{t_i,t_{i+1}}(V)|^p > w_{s,t+h}(V) - \varepsilon. \]

Let \(j\) be such that \(t_j < t \leq t_{j+1}\). From the above inequality, we have

\[ w_{s,t}(V) + |\rho_{t_j,t_{j+1}}(V)|^p - |\rho_{t_j,t}(V)|^p + w_{t,t+h}(V) > w_{s,t+h}(V) - \varepsilon. \]

We send \(h\) to 0, noting that (by (3.1)) \(|\rho_{t_j,t_{j+1}}(V)|^p - |\rho_{t_j,t}(V)|^p \leq |\rho_{t,t+h}(V)|^p \leq \rho_{t,t+h}(V)\) vanishing, to obtain that \(w_{s,t}(V) + w_{t,t+h}(V) \geq w_{s,t+h}(V)\). By the previous step, we have \(w_{s,t}(V) \geq w_{s,t+h}(V)\).
The reverse inequality is obvious by monotonicity so that \( w_{s,t}(V) = w_{s-t}(V) \). In an analogous way, one has \( w_{s,t}(V) = w_{s,t}(V) \), and hence continuity from the outside. \( \square \)

As an immediate consequence, we have for each \( V \) in \( \text{VMO}^{p-\text{var}} \) that

\[
\| \mathbb{E}_s |V_t - V_s| \|_\infty \leq |w_{s,t}(V)|^{1/p} \quad \forall \ 0 \leq s \leq t \leq \tau. \tag{3.2}
\]

**Theorem 3.4.** Let \( (V_t)_{t \in [0,\tau]} \) be \( \text{VMO} \). Then

\[
\mathbb{E}_r e^{\lambda \sup_{t \in [r,\tau]} |V_t - V_r|} < \infty \quad \text{for every } r \in [0,\tau] \text{ and every } \lambda > 0. \tag{3.3}
\]

If \( V \) belongs to \( \text{VMO}^{p-\text{var}} \) then

\[
\sup_{r \in [0,\tau]} \mathbb{E}_r e^{\lambda \sup_{t \in [r,\tau]} |V_t - V_r|} \leq 2^{1+(22\lambda)^p} w_0(V) \quad \text{for every } \lambda > 0. \tag{3.4}
\]

**Proof.** The estimate (3.3) is a direct consequence of Theorem 2.3.

Suppose now that \( V \) is \( \text{VMO}^{p-\text{var}} \) and put \( \rho = 11w(V)^{1/p} \). We define \( t_0 = 0 \) and for each integer \( k \geq 1 \),

\[ t_k = \sup \{ t \in [t_{k-1}, \tau] : \lambda \rho t_{k-1, t} \leq 1/2 \}. \]

By continuity of \( w \), we have \( \lambda \rho t_{k-1, t_k} = 1/2 \) for \( k = 1, \ldots, n-1 \) and \( \lambda \rho t_{n-1, t_n} \leq 1/2 \). By definition of controls, we have

\[
\frac{n-1}{(22\lambda)^p} \leq \sum_{k=1}^{n} w_{t_{k-1}, t_k}(V) \leq w_0(V),
\]

which yields \( n \leq 1 + (22\lambda)^p w_0(V) \). Fix \( r \in [0, \tau] \) and define \( A_r = \sup_{t \in [r, \tau]} |V_t - V_r| \). Proposition 2.8 shows that \( A \) is \( \text{BMO} \) with \( \rho(A) \leq \rho \). Let \( j \) be such that \( t_{j-1} \leq r < t_j \). Following the proof of Proposition 2.7, we have

\[
\mathbb{E}_r e^{\lambda A_r} \leq \prod_{k=j}^{n} (1 - \lambda \rho t_{k-1, t_k}^{1/2})^{-1} \leq 2^n.
\]

These estimates imply (3.4). \( \square \)

**Corollary 3.5.** Let \( (V_t)_{t \in [0,\tau]} \) be \( \text{VMO}^{p-\text{var}} \). If \( p \in (1,\infty) \), then there are constants \( c_p, C_p \) such that for all \( \lambda \) satisfying \( \lambda (w_0(V))^{1/p} < c_p \),

\[
\mathbb{E} \exp \left( \lambda \sup_{t \leq \tau} |V_t - V_0|^{\rho} \right) < \infty, \tag{3.5}
\]

and for all real number \( m \geq 1 \),

\[
\| \mathbb{E}_s [ \sup_{r \in [s,\tau]} |V_r - V_s|^m ] \|_\infty \leq C_p \Gamma(m(1-1/p) + 1)(w_{s,t}(V))^{m/p}, \tag{3.6}
\]

where \( \Gamma(z) = \int_0^\infty u^{z-1}e^{-u} du \) is the Gamma function.

If \( p = 1 \) then

\[
\mathbb{P} \left( |V_t - V_s| \leq 22w_{s,t}(V) \text{ for all } s \leq t \leq \tau \right) = 1. \tag{3.7}
\]
Proof. Consider first the case $p \in (1, \infty)$. Define $Z = \sup_{t \leq r} |V_t - V_0|$ and $a = 1/a$. By Chebyshev inequality and (3.4), we have

$$\mathbb{P}(Z > x) = \mathbb{P}(e^{\lambda Z} > e^{\lambda x}) \leq e^{-\lambda x} \mathbb{E}e^{\lambda Z} \leq Ce^{-\lambda x + c\beta^p}$$

where $\beta = w_{0,r}(V)$ and $c = c(p), C = C(p)$ are some universal positive constants. One can optimize in $\lambda$ to obtain that for every $x$ bounded away from 0,

$$\mathbb{P}(Z > x) \leq Ce^{-c\beta^p \frac{1}{p' - 1} x^{p'}} \cdot \frac{1}{p} + \frac{1}{p'} = 1,$$

where $c_p, C$ are some other positive constants. In view of the layer cake representation

$$\mathbb{E}e^{\lambda Z^{p'}} = \lambda p' \int_0^\infty e^{\lambda x^{p'}} x^{p'-1} \mathbb{P}(Z > x)\,dx,$$

we see that $\mathbb{E}e^{\lambda Z^{p'}}$ is finite if $\lambda \beta^{1/(p-1)} < c_p$. We obtain (3.5) by observing that $p' = p/(p - 1)$.

The estimate (3.6) is obtained in an analogous way. Define $Y = \sup_{t \in [s,t]} |V_t - V_s|$. Reasoning as previously,

$$\mathbb{P}_s(Y > x) \leq Ce^{-c\beta^p \frac{1}{p' - 1} x^{p'}}.$$

By the layer cake representation,

$$\mathbb{E}_s Y^m = m \int_0^\infty x^{m-1} \mathbb{P}_s(Y > x)\,dx \leq Cm \int_0^\infty x^{m-1} e^{-c\beta^{1/(p-1)} x^{p'}}\,dx.$$

After the change of variable $y = c_p \beta^{1/(p-1)} x^{p'}$, using the identity $\Gamma(z + 1) = z\Gamma(z)$, we arrive at (3.6).

In the case $p = 1$, a similar argument with Chebyshev inequality and (3.4) leads to

$$\mathbb{P}(|V_t - V_s| > x) \leq e^{-\lambda x + 2w_{s,t}(V)}$$

for every $x > 0$ and $\lambda > 0$. When $x > 22w_{s,t}(V)$, we can send $\lambda \to \infty$ to obtain that $\mathbb{P}(|V_t - V_s| > x) = 0$. This implies that $\mathbb{P}(|V_t - V_s| \leq 22w_{s,t}(V)) = 1$. Since $V$ has continuous sample paths, this implies (3.7).

Estimate (3.6) is inspired by the precise estimate from Lyons’ first extension theorem (inequality (2.21) in [Lyo98]), which is obtained through the so-called neo-classical inequality. Note that there is a smallness condition in the aforementioned article, which is not present in Corollary 3.5.

We note that (3.5) cannot be derived from the Garnett–John inequality (1.5). Let $\lambda_p(V)$ denote the largest constant such that (3.5) holds. An interesting open question is whether one could characterize the size of $\lambda_p(V)$ by the distance of $V$ to a certain subspace of $\text{VMO}^{p-\text{var}}$. 

---

[Note: The document appears to contain mathematical expressions and inequalities, which might require context-specific knowledge to fully understand. The text is formatted in a way that suggests it is part of a larger mathematical or scientific discussion, possibly related to probability theory or stochastic processes.]
4. Applications

4.1. Rough stochastic differential equations. In [FHL21], the authors consider a hybrid rough stochastic differential equation of the type

\[ dY_t = b_t(Y_t)dt + \sigma_t(Y_t)dB_t + (f_t, f'_t)(Y_t)dX_t \]  

(4.1)

where \( B \) is a standard Brownian motion and \( X = (X, X) \) is a Hölder rough path. The coefficients \( b, \sigma, f, f' \) are progressively measurable and regular in the \( Y \)-component. Under some natural regularity conditions, [FHL21] shows that (4.1) has a unique continuous solution in a certain class of stochastic controlled rough paths denoted by \( D^X_{\alpha,L_{m,\infty}} \) for some \( \alpha \in (1/3, 1/2] \) and \( m \geq 2 \). Such processes are adapted and satisfy

\[ \sup_{s < t \leq \tau} \frac{\|\mathbb{E}S|Y_t - Y_s|^m\|_\infty^{1/m}}{(t - s)^\alpha} < \infty, \]  

(4.2)

together with some controllness conditions. As is shown in [FHL21], this class of stochastic controlled rough paths are stable under composition with smooth vector fields and rough stochastic integration. Because most of these properties are irrelevant for our purpose, we refer to the cited reference for further details.

Based upon earlier sections, any continuous adapted process satisfying the property (4.2) is \( \text{VMO}^\alpha \). Hence, such stochastic controlled rough paths are subjected to the John–Nirenberg inequality. Although [FHL21] also discusses exponential estimates for the solution of (4.1) by means of Lyons’ multiplicative functionals, their result comes with some additional restrictions on \( m, \alpha \) and the connection with VMO processes was not present there. On the other hand, our results actually accommodate minimal conditions that \( \alpha \in (0, 1] \) and \( m = 1 \).

**Theorem 4.1.** Let \((Y_t)_{t \in [0, \tau]}\) be a continuous process in \( D^X_{\alpha,L_{1,\infty}} \) (see Section 3 of [FHL21] for the precise definition) for some \( \alpha \in (0, 1) \). Then \( Y \) is \( \text{VMO}^\alpha \) and

\[ \mathbb{E}e^{\lambda \sup_{t \in [0, \tau]} |Y_t - Y_0|} \leq 2^{1+2[2\alpha][\text{VMO}^\alpha, \lambda]^{1/\alpha}} \] 

for every \( \lambda > 0 \).

**Proof.** As discussed earlier, \( Y \) satisfies (4.2) with \( m = 1 \). By Proposition 2.2 and sample path continuity, \( Y \) is necessarily \( \text{VMO}^\alpha \). The exponential estimate is a direct consequence of (3.4). \( \square \)

Another class of processes introduced in [FHL21] is \( C^{\alpha,L_{m,\infty}} \) (with \( \alpha \in (0, 1] \) and \( m \geq 1 \)) consisting of adapted processes \( Y \) such that

\[ \sup_{t \in [0, \tau]} \|Y_t\|_m + \sup_{s < t \leq \tau} \frac{(\|\mathbb{E}S|Y_t - Y_s|^m\|_\infty)^{1/m}}{(t - s)^\alpha} < \infty. \]

We can classify these classes as VMO processes in the following way.

**Proposition 4.2.** Let \( \alpha \in (0, 1] \) and \( m \in [1, \infty) \) be some fixed numbers and \((Y_t)_{t \in [0, \tau]}\) be a continuous adapted process. \( Y \) belongs to \( C^{\alpha,L_{m,\infty}} \) if and only if \( Y_0 \) is \( L_m \)-integrable and \( Y \) is \( \text{VMO}^\alpha \).

**Proof.** Straightforward from Proposition 2.2. \( \square \)
4.2. **Davie’s estimates.** Let \((B_t)_{t \geq 0}\) be a standard Brownian motion and \(g : [0, \tau] \times \mathbb{R}^d \to \mathbb{R}\) be a bounded Borel measurable function such that \(|g(r, y)| \leq 1\) for all \((r, y) \in [0, \tau] \times \mathbb{R}^d\). Davie shows in [Dav07] that for any even integer positive integer \(m\) and \(x \in \mathbb{R}^d\),
\[
\mathbb{E} \left( \int_0^1 [g(t, B_t + x) - g(t, B_t)] \, dt \right)^m \leq C^m \Gamma (m/2 + 1)|x|^m, \tag{4.3}
\]
where \(C\) is an absolute constant. This inequality exhibits the regularization effect of Brownian motion through temporal integration. Such regularization effect is developed further into the framework of nonlinear Young integration by Catellier and Gubinelli in [CG16]. Davie’s estimate is an important one and has been reproduced in different forms under other conditions and setups [Sha16, RZ21, Rez14, Lê22, Lê20, ABLM20].

Davie shows (4.3) by first expanding the moment into an iterated multiple integral. The lack of regularity in \(g\) is compensated by the smooth density of the Brownian motion through integration by parts. This procedure produces a sum of \(2^{p-1}\) iterated multiple integrals involving derivatives of the Gaussian density. He then estimates each of these multiple integrals carefully to obtain (4.3). Davie’s proof is beautiful yet intricate because of its analysis of high moments. We now explain how John–Nirenberg inequality in Theorems 2.3 and 3.4 could be utilized in this context. Indeed, following Davie’s proof in [Dav07], one has
\[
\mathbb{E} \left( \int_s^t [g(r, B_r + x) - g(r, B_r)] \, dr \right)^2 \leq C^2|x|^2(t - s).
\]
Since Brownian motion has independent increments, we can upgrade the above inequality to the following estimate
\[
\mathbb{E} \left( \int_s^t [g(r, B_r + x) - g(r, B_r)] \, dr \right)^2 \leq C^2|x|^2(t - s).
\]
This shows that the process \(V_t = \int_0^t [g(r, B_r + x) - g(r, B_r)] \, dr\) is VMO\(^{1/2}\) with \([V]_{\text{VMO}^{1/2}} \leq C|x|\). The estimate (3.6) gives
\[
\mathbb{E} \left( \sup_{t \in [0,1]} \left| \int_0^t [g(r, B_r + x) - g(r, B_r)] \, dr \right| \right)^m \leq C^m \Gamma (m/2 + 1)|x|^m. \tag{4.4}
\]
This estimate is comparable to or perhaps stronger than (4.3) because of the supremum in its left-hand side.

4.3. **Quadrature error estimates and strong convergence rate of Euler method.** Consider the stochastic differential equation
\[
dx_t = b(t, x_t) dt + \sigma(t, x_t) dB_t, \quad x_0 = x_0, \quad t \in [0, 1], \tag{4.5}
\]
where \(d \geq 1, b : [0, 1] \times \mathbb{R}^d \to \mathbb{R}^d, \sigma : [0, 1] \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d\) are Borel measurable functions, \((B_t)_{t \geq 0}\) is a \(d\)-dimensional standard Brownian motion defined on some complete
filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) and \(x_0\) is an \(\mathcal{F}_0\)-random variable. The tamed Euler–Maruyama scheme associated to (4.5) is

\[
dX^n_t = b^n(t, X^n_{k_n(t)}) \, dt + \sigma(t, X^n_{k_n(t)}) \, dB_t, \quad X^n_0 = x^n_0, \quad t \in [0, 1],
\]

(4.6)

where \(x^n_0\) is a \(\mathcal{F}_0\)-random variable and \(b^n\) is an approximation of the vector field \(b\) and

\[k_n(t) = \frac{j}{n} \quad \text{whenever} \quad \frac{j}{n} \leq t < \frac{j+1}{n} \quad \text{for some integer } j \geq 0.
\]

We note that (4.6) with the choice \(b^n = b\) is the usual Euler–Maruyama scheme, which, however, is not well-defined for a merely integrable function \(b\) even when \(b\) is replaced by \(b1_{\{|b|<\infty\}}\). 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.

The recent article [LL21] establishes strong rate of convergence of the tamed Euler–Maruyama scheme (4.6) to (4.5) under some integrability condition of the drift \(b\). To state their result, we first recall some notation from [LL21]. 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 \(\Omega\). For each \(v \in \mathbb{R}, L_{v,p}(\mathbb{R}^d) := (1 - \Delta)^{-v/2}(L^p(\mathbb{R}^d))\) is the usual Bessel potential space on \(\mathbb{R}^d\) equipped with the norm \(\|f\|_{L_{v,p}(\mathbb{R}^d)} := \| (1 - \Delta)^{-v/2} f \|_{L^p(\mathbb{R}^d)}\), where \((1 - \Delta)^{v/2} f\) is defined through Fourier’s transform. \(L_{v,p}^q([0,1])\) denotes the space of measurable function \(f : [0,1] \to L_{v,p}(\mathbb{R}^d)\) such that \(\|f\|_{L_{v,p}^q([0,1])}\) is finite. Here, for each \(s, t \in [0, 1]\) satisfying \(s \leq t\), we denote

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

with obvious modification when \(q = \infty\). When \(v = 0\), we simply write \(L_p^q([0,1])\) instead of \(L_{0,p}^q([0,1])\). In particular, \(L_p^q([0,1])\) 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. As in [LL21], we assume the following conditions.

**Condition \(\mathcal{C}\).** The diffusion coefficient \(\sigma\) 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 \(x \in \mathbb{R}^d\)

\[
K_1^{-1} I \leq (\sigma \sigma^*) (s, x) \leq K_1 I,
\]

(4.7)

where \(I\) denotes the identity matrix. 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 \(x, y \in \mathbb{R}^d\)

\[
|(\sigma \sigma^*)(s, x) - (\sigma \sigma^*)(s, y)| \leq K_2 |x - y|^{\alpha}.
\]

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

\[
\frac{d}{p_0} + \frac{2}{q_0} < 1 \quad \text{and} \quad \|\nabla \sigma\|_{L_{p_0}^{q_0}([0,1])} \leq K_3.
\]
Condition \( \mathcal{B} \). \( x_0 \) belongs to \( L_p(\Omega, \mathcal{F}_0) \) and \( b \) belongs to \( L_p^q([0, 1]) \) for some \( p, q \in [2, \infty) \) satisfying \( \frac{2}{p} + \frac{2}{q} < 1 \). For each \( n \), \( x_0^n \) belongs to \( L_p(\Omega, \mathcal{F}_0) \) and \( b^n \) belongs to \( L_p^q([0, 1]) \cap L_{\infty}^q([0, 1]) \) with \( p, q \) as above. Furthermore, there exist finite positive constants \( K_4, \theta \) and continuous controls \( \{\mu^n\}_n \) such that \( \sup_{n \geq 1} (\|b^n\|_{L_p^q([0, 1])} + \mu^n(0, 1)) \leq K_4 \) and

\[
(1/n)^{1/2} \|b^n\|_{L_{\infty}^q([s,t])} \leq \mu^n(s,t)^\theta \quad \forall \ t - s \leq 1/n.
\] (4.8)

Definition 4.3. Let \( \lambda > 0 \) be a fixed number which is sufficiently large. Let \( U = (U^1, \ldots, U^d) \) where for each \( h = 1, \ldots, d \), \( U^h \) is the solution to the following equation

\[
\partial_t U^h + \sum_{i,j=1}^d \frac{1}{2}(\sigma^*_i v)^{ij} \partial_{ij} U^h + b^h \cdot \nabla U^h = \lambda U^h - b^h, \quad U^h(1, \cdot) = 0.
\] (4.9)

Let \( X \) be the solution to (4.5). For each \( \tilde{p} \in [1, \infty) \), we put

\[
\omega_n(\tilde{p}) = \left\| \sup_{t \in [0,1]} \left| \int_0^t (1 + \nabla U)[b - b^n](r, X_r) dr \right| \right\|_{L_p(\Omega)}.
\]

Note that we have changed the definition of \( \omega_n \) from [LL21] by replacing \( b^{n,h} \) with \( b^h \) in (4.9).

The main result of [LL21] (Theorems 2.2 and 2.3 therein) asserts that for any \( \tilde{p} \in (1, p) \cap (1, \frac{2}{\theta}(p \wedge p_0)) \) and any \( y \in (0, 1) \), there exists a finite constant \( N \) such that

\[
\| \sup_{t \in [0,1]} |X^n_t - X_t| \|_{L_{\tilde{p}}(\Omega)} \leq N \left[ \|x^n_0 - x_0\|_{L_p(\Omega)} + (1/n)^{\theta/2} + (1/n)^{1/2} \log(n) + \omega_n(\tilde{p}) \right].
\] (4.10)

To obtain estimate, [LL21] first utilizes stability results for (4.5) to show that the strong convergence rate is bounded by \( \omega_n(\tilde{p}) \) and the quadrature error of the type

\[
\| \sup_{t \in [0,1]} \int_0^t g(r, X^n_r)(f(r, X^n_r) - f(r, X^n_{k_0(r)})) dr \|_{L_p(\Omega)}
\]

where \( f \in L_p^q \cap L_{\infty}^q \) and \( g \in L_1^q \cap L_{\infty}^q \). [LL21] then applies stochastic sewing techniques (introduced in [Lé20]) to obtain the rate \( (1/n)^{1/2} \log(n) \) for the quadrature error and to bound \( \omega_n \) by a suitable distance between \( b \) and \( b^n \).

While (4.10) produces the best available rate in the literature, it comes with an unnatural constraint on \( \tilde{p} \). This restriction is purely technical and is necessary for both stability analysis and stochastic sewing arguments described previously. In the more recent article [GL22], a stability estimate which is valid for all moments has been obtained. In the present article, we utilize the John–Nirenberg inequality (Theorem 2.3) to overcome the moment restriction in the stochastic sewing arguments used to estimate \( \omega_n \) and the quadrature error. Our main contribution are the following two results, which remove the moment restrictions from Theorems 2.2 and 2.3 of [LL21].
Theorem 4.4. Assume that Conditions $\mathfrak{A}$-$\mathfrak{B}$ hold. Let $(X^n_t)_{t \in [0,1]}$ be the solution to (4.6) and $(X_t)_{t \in [0,1]}$ be the solution to (4.5). Then for any $\hat{p} \in (1, \infty)$ and any $\gamma > 1$, there exists a finite constant $N(K_1, K_2, K_3, K_4, \alpha, p_0, q_0, p, q, d, \hat{p}, \gamma)$ such that

$$
\| \sup_{t \in [0,1]} |X^n_t - X_t| \|_{L^p(\Omega)} \leq N \left[ \| x_0^n - x_0 \|_{L^p(\Omega)} + (1/n)^{\frac{d}{2}} + (1/n)^{\frac{2}{q}} \log(n) + \omega_n(\gamma \hat{p}) \right].
$$

(4.11)

Theorem 4.5. Assume that Conditions $\mathfrak{A}$-$\mathfrak{B}$ hold with $q_0 = \infty$ and $\frac{1}{p} + \frac{1}{p_0} < 1$. Let $\nu \in [0, 1)$ be such that $\nu < \frac{3}{2} - \frac{d}{2p} - \frac{2}{q}$. Then for every $\hat{p} \in (1, \infty)$, there exists a constant $N$ depending on $K_1, K_2, K_3, K_4, \alpha, p_0, p, q, d, \hat{p}, \nu$ such that

$$
\omega_n(\hat{p}) \leq N \| b - b^n \|_{L^{q,p}([0,1])}.
$$

(4.13)

[LL21] also considers the case $\nu = 1$ in Theorem 4.5. Our argument also works in this case without much effort. We therefore leave it for interested readers.

Proposition 4.6. Let $p \in (1, \infty)$, $q \in (2, \infty)$ and assume that Condition $\mathfrak{A}$ holds with $q_0 = \infty$ and $\frac{1}{p} + \frac{1}{p_0} < 1$. Let $X$ be a solution to (4.5). Let $g$ be a function in $L^q_p([0,1])$ and let $\nu \in [0, 1)$ such that $\frac{d}{p} + \frac{2}{q} + \nu < 2$. Then for any $\hat{p} \in [1, \infty)$, there exists a constant $N = N(\nu, d, p, q, \hat{p})$ such that

$$
\| \sup_{t \in [0,1]} \left| \int_0^t g(r, X_r)dr \right| \|_{L^p(\Omega)} \leq N \| g \|_{L^{q,p}([0,1])}.
$$

(4.14)

Proof. By Girsanov transformation, we can assume without loss of generality that $b = 0$ (see the argument in the proof of Theorem 5.1 in [LL21]). Put $V_t = \int_0^t g(r, X_r)dr$. We note that by Krylov estimate, $\| V_t - V_s \|_{L^m(\Omega)} \leq \| g \|_{L^q_p(t-s)^{1-\frac{d}{2p} - \frac{1}{q}}} for all $m \geq 1$ (see inequality (6.11) and Lemma 3.4 from [LL21]). Consequently, $V$ is a.s. continuous. The proof of Proposition 6.6 from [LL21] shows that

$$
\sup_{s \leq t \leq 1} \left( \| E_s \|_{L^p(\Omega)} \right)^{1/p} \leq \| g \|_{L^{q,p}([0,1])}.
$$

This shows that $V$ is BMO and hence estimate (4.14) follows from Theorem 2.3. \hfill \square

Proposition 4.7. Assume that Conditions $\mathfrak{A}$1 and $\mathfrak{B}$ hold. Let $X^n$ be the solution to (4.6) and let $f, g$ be measurable functions on $[0, 1] \times \mathbb{R}^d$. Assume that $\| f \|_{L^q_p([0,1])} = \| g \|_{L^q_p([0,1])} = 1$ and $\beta_n(f) = \sup_{0 \leq j \leq n-1} \| f \|_{L^q_p((j/n,(j+1)/n))} is finite. Then for any $\hat{p} \geq 1$, there exists a constant $N = N(d, p, q, \hat{p})$ such that

$$
\left\| \sup_{t \in [0,1]} \left| \int_0^t g(r, X^n_r) \left[ f(r, X^n_r) - f(r, X^n_{k_n(r)}) \right] dr \right| \right\|_{L^p(\Omega)}
$$

16
\[
\leq N \left[ (1/n)^{1-\frac{1}{4}} \beta_n(f) + (1/n)^{\frac{d}{2}} + (1/n)^{\frac{5}{4}} \log(n) \right]. 
\] (4.15)

**Proof.** By Girsanov transformation, we can assume without loss of generality that \( b = 0 \) (see the argument in the proof of Theorem 5.1 in [LL21]). Define
\[
V_t = \int_0^t g(r, X^n_r) \left[ f(r, X^n_r) - f(r, X^n_{k_n(r)}) \right] dr.
\]

Proposition 5.12 of [LL21] and its proof shows that for every \( v + 4/n \leq s \leq t \leq 1, \)
\[
\| \mathbf{E}_0 [V_t - V_s] \|^{1/p}_{\infty} \leq [(1/n)^{\frac{d}{2}} + (1/n)^{\frac{5}{4}} \log(n)].
\]

By assumption and Hölder inequality, we have for every \( s \leq t \)
\[
|V_t - V_s| \leq \|f\|_{L^2([s,t])} (t-s)^{1-\frac{1}{4}}.
\]

Combining the previous two estimates yields that for every \( s \leq t \)
\[
\| \mathbf{E}_0 [V_t - V_s] \|^{1/p}_{\infty} \leq [(1/n)^{\frac{d}{2}} + (1/n)^{\frac{5}{4}} \log(n) + (1/n)^{1-\frac{1}{4}} \beta_n(f)].
\]

Since \( V \) is continuous, from Proposition 2.2, it is BMO (in fact VMO) and hence by applying Theorem 2.3, we obtain (4.15). \( \Box \)

**Proof of Theorem 4.4.** The stability argument in [LL21] comes with a restriction on the moment and we must replace it by the recent stability estimate from [GL22]. Indeed, from Section 3.4.1 of the aforementioned reference, we have
\[
\| \sup_{t \in [0,1]} |X_t - X^n_t| \|_{L^p(\Omega)} \leq \|x_0 - x^n_0\|_{L^p(\Omega)} + \| \sup_{t \in [0,1]} |V_t| \|_{L^p(\Omega)},
\]

where
\[
V_t = \int_0^t \left( \frac{1}{2} \left[ (\mathbf{R}^2 + \sigma)(\mathbf{R}^2 + \sigma)^* - \sigma \sigma^* \right] : D^2U + R^1 \cdot (I + \nabla U) (r, X^n_r) dr \right.
\]
\[
+ \left. \int_0^t [R^2(I + \nabla u)(r, Y_r) dB_r, \right.
\]
\[
R^1_t = b^n(t, X^n_{k_n(t)}) - b(t, X^n_t), \quad R^2_t = \sigma(t, X^n_{k_n(t)}) - \sigma(t, X^n_t).
\]

Since \( \sigma \) is Hölder continuous, the moments of terms with \( R^2 \) are bounded by a constant multiple of \( (1/n)^{\alpha/2} \) (see Section 7 of [LL21] for some analogous estimates). To treat the term with \( R^1 \), we note that
\[
| \int_0^t R^1 \cdot (I + \nabla U)(r, X^n_r) dr | \leq \| \sup_{t \in [0,1]} |V_t| \|_{L^p(\Omega)}.
\]

To treat the first term, we apply Proposition 4.7, regularity of \( U \) (Lemma 7.1 of [LL21]) and Condition \( \mathbb{B} \) to have

17
\[ \left\| \sup_{t \in [0,1]} \left| \int_0^t (b^n(r, X^n_{k_n(r)}) - b^n(X^n_r)) \cdot (I + \nabla U)(r, X^n_r) \, dr \right|_{L_p(\Omega)} \right\| \leq (1/n)^{1 - \frac{1}{d}} \beta_n(b^n) + (1/n)^{\frac{d}{2}} + (1/n)^{\frac{d}{2}} \log(n) \leq (1/n)^{\frac{d}{2}} + (1/n)^{\frac{d}{2}} \log(n). \]

Moment of the second term is directly related to \( \omega_n(\hat{p}) \) through Definition 4.3. This leads us to the following estimate

\[ \left\| \sup_{t \in [0,1]} |V_t| \right\|_{L_p(\Omega)} \leq (1/n)^{\frac{d}{2}} + (1/n)^{\frac{d}{2}} \log(n) + \omega_n(\hat{p}). \]

Combining with (4.16), we obtain (4.11). \( \square \)

**Proof of Theorem 4.5.** The proof follows in exactly the same way as the proof of Theorem 2.3 in [LL21] (Section 7 therein). The restriction \( \hat{p} < p \) there is now lifted thanks to Proposition 4.6. \( \square \)

**Appendix A. Auxiliary results**

**Lemma A.1** (Garsia’s upcrossing lemma, [Gar73b, Str73]). Let \((X_t)_{t \in [0,T]}\) be a right continuous adapted process with left limits and let \( s \) be a fixed time in \([0,T]\). Suppose that there is a nonnegative integrable random variable \( U \) such that

\[ E_S|X_T - X_{S-}| \leq E_S U \]  

(A.1)

for any pair \( S, T \) of stopping times with \( s \leq S \leq T \leq \tau \). Let \( Y \) be an \( F_s \)-random variable and define \( X^* = \sup_{s \leq r \leq \tau} |X_r - Y| \). Then for every \( \alpha, \beta > 0 \), one has

\[ \beta P_S(X^* \geq \alpha + \beta) \leq E_S(U1_{(X^* \geq \alpha)}). \]  

(A.2)

**Proof.** We adopt the arguments from [Kaz94]. Let \( \alpha, \beta > 0 \) be given and \( G \in F_s \). We set \( X_t = X_{\tau} \) for \( t > \tau \) and define

\[ S = \inf\{t \geq s : |X_t - Y| \geq \alpha\}, \quad T = \inf\{t \geq s : |X_t - Y| \geq \alpha + \beta\}, \]

with the standard convention that \( \inf(\emptyset) = \infty \). Clearly \( S \) and \( T \) are stopping times and \( s \leq S \leq T \). We also have from the above definitions,

\[ \{X^* \geq \alpha + \beta\} \subset \{|X_T - X_{S-}| \geq \beta, |X_S - Y| \geq \alpha\}. \]  

(A.3)

It follows that

\[ E(1_{(X^* \geq \alpha + \beta)}1_G) \leq E(1_{|X_T - X_{S-}| \geq \beta}1_{(|X_S - Y| \geq \alpha)}1_G) \leq \frac{1}{\beta} E(|X_T - X_{S-}|1_{(|X_S - Y| \geq \alpha)}1_G) \leq \frac{1}{\beta} E(U1_{(|X_S - Y| \geq \alpha)}1_G) \]

which implies the result. \( \square \)
We note that right-continuity of the filtration is necessary so that $S, T$ defined in the previous proof are stopping times. In addition, the inclusion (A.3) does not hold if one replaces $X_{S-}$ by $X_S$ in (A.1). These technical conditions become irrelevant when dealing with continuous processes.

**Lemma A.2 (Energy inequality).** Let $c$ be a deterministic constant and $(A_t)_{t \geq 0}$ be an adapted, right-continuous, non-decreasing process. Let $\tau > 0$ be fixed and suppose that
\[
\|E_S(A_\tau - A_{S-})\|_\infty \leq c \text{ for every stopping time } S \leq \tau.
\] (A.4)

Then for every $s \in [0, \tau]$ and every integer $p \geq 1$,
\[
\|E_s(A_\tau - A_s)^p\|_\infty \leq p!c^p.
\] (A.5)

**Proof.** When $A_t$ takes the specific form $\int_0^t \beta(r)dr$ for some $\beta \geq 0$, this result deduces to the Khasminskii’s lemma ([Kha59]). In the general form, it is known as energy inequality and can be found in [Mey66, Kik92]. Our statement is for processes over finite time intervals which differs from previous ones and needs justifications.

Let $s \in [0, \tau]$ be fixed and $G$ be an event in $\mathcal{F}_s$. For each $r \geq 0$, define $\tilde{A}_r = 1_G(A_{(r+s)}^{A,r} - A_r)$. The process $\tilde{A}$ is adapted with respect to the filtration $\tilde{\mathcal{F}} := \{\mathcal{F}_{r+s}, r \geq 0\}$, right-continuous, satisfies $\tilde{A}_0 = 0$ and $\|E_S(\tilde{A}_\tau - \tilde{A}_{S-})\|_\infty \leq c$ for all $\tilde{\mathcal{F}}$-stopping times $S$. Applying Theorem 4 of [Kik92] to the process $\tilde{A}$, we get that $E(1_G(A_\tau - A_s)^p) \leq p!c^p$. Since $G$ is arbitrary, this implies (A.5).

\[\square\]

**Acknowledgment**

The author thanks Peter Friz for motivating discussions on exponential integrability and his interest in this work. He is grateful to Stefan Geiss for explaining the works [GY20, GN20, Gei05] and equivalent definitions of BMO processes.

**Funding**

The author was supported by Alexander von Humboldt Research Fellowship during the early stage of this work.

**References**

[AABLM20] Siva Athreya, Oleg Butkovsky, Khoa Lê, and Leonid Mytnik. Well-posedness of stochastic heat equation with distributional drift and skew stochastic heat equation. *arXiv preprint arXiv:2011.13498*, 2020.

[Alt21] Randolf Altmeyer. Approximation of occupation time functionals. *Bernoulli*, 27(4):2714–2739, 2021.

[BN08] Jocelyne Bion-Nadal. Dynamic risk measures: time consistency and risk measures from BMO martingales. *Finance Stoch.*, 12(2):219–244, 2008.

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

[Dav07] A. M. Davie. Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not. IMRN*, (24):Art. ID rnm124, 26, 2007.

[DG20] Konstantinos Dareiotis and Máté Gerencsér. On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift. *Electron. J. Probab.*, 25:Paper No. 82, 18, 2020.
[DGL22] Konstantinos Dareiotis, Máté Gerencsér, and Khoa Lê. Quantifying a convergence theorem of Gyöngy and Krylov. _Annals of Applied Probability_, 2022+.

[DMS’97] Freddy Delbaen, Pascale Monat, Walter Schachermayer, Martin Schweizer, and Christophe Stricker. Weighted norm inequalities and hedging in incomplete markets. _Finance and Stochastics_, 1(3):181–227, 1997.

[Fe71] Charles Fefferman. Characterizations of bounded mean oscillation. _Bull. Amer. Math. Soc._, 77:587–588, 1971.

[FHL21] Peter Friz, Antoine Hocquet, and Khoa Lê. Rough stochastic differential equations. _arXiv preprint arXiv:2106.10340_, 2021.

[FS72] C. Fefferman and E. M. Stein. $H^p$ spaces of several variables. _Acta Math._, 129(3-4):137–193, 1972.

[Gar73a] Adriano M. Garsia. The Burgess Davis inequalities via Fefferman’s inequality. _Ark. Mat._, 11:229–237, 1973.

[Gar73b] Adriano M. Garsia. _Martingale inequalities: Seminar notes on recent progress_. Mathematics Lecture Note Series. W. A. Benjamin, Inc., Reading, Mass.-London-Amsterdam, 1973.

[Gei05] Stefan Geiss. Weighted BMO and discrete time hedging within the Black-Scholes model. _Probab. Theory Related Fields_, 132(1):13–38, 2005.

[GG22] Lucio Galeati and Máté Gerencsér. Solution theory of fractional SDEs in complete subcritical regimes. _arXiv preprint arXiv:2207.03475_, 2022.

[GJ78] John B. Garnett and Peter W. Jones. The distance in BMO to $L^\infty$. _Ann. of Math. (2)_., 108(2):373–393, 1978.

[GL22] Lucio Galeati and Chengcheng Ling. Stability estimates for singular SDEs and applications. _arXiv preprint arXiv:2208.03670_, 2022.

[GN20] STEFAN Geiss and T Nguyen. On riemann–liouville operators, bmo, gradient estimates in the lévy–itô space, and approximation. _arXiv preprint arXiv:2009.00899_, 2020.

[GS72] R. K. Getoor and M. J. Sharpe. Conformal martingales. _Invent. Math._, 16:271–308, 1972.

[GY20] Stefan Geiss and Juha Ylinen. Weighted bounded mean oscillation applied to backward stochastic differential equations. _Stochastic Process. Appl._, 130(6):3711–3752, 2020.

[Her74] Carl Herz. Bounded mean oscillation and regulated martingales. _Trans. Amer. Math. Soc._, 193:199–215, 1974.

[JN61] F. John and L. Nirenberg. On functions of bounded mean oscillation. _Comm. Pure Appl. Math._, 14:415–426, 1961.

[Kaz94] Norihiko Kazamaki. _Continuous exponential martingales and BMO_, volume 1579 of _Lecture Notes in Mathematics_. Springer-Verlag, Berlin, 1994.

[Kha59] R. Z. Khasminskii. On positive solutions of the equation $Uu + Vu = 0$. _Theor. Probability Appl._, 4:309–318, 1959.

[Kik92] Masato Kikuchi. A note on the energy inequalities for increasing processes. In _Séminaire de Probabilités, XXVI_, volume 1526 of _Lecture Notes in Math._, pages 533–539. Springer, Berlin, 1992.

[KS91] Ioannis Karatzas and Steven E. Shreve. _Brownian motion and stochastic calculus_, volume 113 of _Graduate Texts in Mathematics_. Springer-Verlag, New York, second edition, 1991.

[Lê20] Khoa Lê. A stochastic sewing lemma and applications. _Electron. J. Probab._, 25:Paper No. 38, 55, 2020.

[Lê22] Khoa Lê. Stochastic sewing in Banach space. _arXiv preprint arXiv:2105.09364_, 2022.

[LL21] Khoa Lê and Chengcheng Ling. Taming singular stochastic differential equations: A numerical method. _arXiv preprint arXiv:2110.01343_, 2021.

[Lyo98] Terry J. Lyons. Differential equations driven by rough signals. _Rev. Mat. Iberoamericana_, 14(2):215–310, 1998.

[Mey66] Paul-A. Meyer. _Probability and potentials_. Blaisdell Publishing Co. [Ginn and Co.], Waltham, Mass.-Toronto, Ont.-London, 1966.
[Ose15] Adam Osękowski. Sharp maximal estimates for BMO martingales. Osaka J. Math., 52(4):1125–1142, 2015.

[Por90] N. I. Portenko. Generalized diffusion processes, volume 83 of Translations of Mathematical Monographs. American Mathematical Society, Providence, RI, 1990. Translated from the Russian by H. H. McFaden.

[Rez14] Fraydoun Rezakhanlou. Regular flows for diffusions with rough drifts. arXiv preprint arXiv:1405.5856, 2014.

[RZ21] Michael Röckner and Guohuan Zhao. Sdes with critical time dependent drifts: strong solutions. arXiv preprint arXiv:2103.05803, 2021.

[Sha16] A. V. Shaposhnikov. Some remarks on Davie’s uniqueness theorem. Proc. Edinb. Math. Soc. (2), 59(4):1019–1035, 2016.

[Str73] Daniel W. Stroock. Applications of Fefferman-Stein type interpolation to probability theory and analysis. Comm. Pure Appl. Math., 26:477–495, 1973.

[SV06] Daniel W. Stroock and S. R. Srinivasa Varadhan. Multidimensional diffusion processes. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[Var80] N. Th. Varopoulos. A probabilistic proof of the Garnett-Jones theorem on BMO. Pacific J. Math., 90(1):201–221, 1980.

[Zha16] Xicheng Zhang. Stochastic differential equations with Sobolev diffusion and singular drift and applications. Ann. Appl. Probab., 26(5):2697–2732, 2016.

School of Mathematics, University of Leeds, U.K.

Email address: k. le@leeds. ac. uk