Bernoulli **21**(1), 2015, 437–464
DOI: 10.3150/13-BEJ574

Integrability and concentration of the truncated variation for the sample paths of fractional Brownian motions, diffusions and Lévy processes

WITOLD MAREK BEDNORZ[1] and RAFAŁ MARCIN ŁOCHOWSKI[2,3]

[1]*Institute of Mathematics, Warsaw University, Banacha 2, 02-097 Warszawa, Poland.*
E-mail: wbednorz@mimuw.edu.pl
[2]*Department of Mathematics and Mathematical Economics, Warsaw School of Economics, Madalińskiego 6/8, 02-513 Warszawa, Poland.*
[3]*Department of Core Mathematics and Social Sciences, Prince Mohammad Bin Fahd University, P.O. Box 1664, Al Khobar 31952, Saudi Arabia.*
E-mail: rlocho314@gmail.com; url: http://akson.sgh.waw.pl/~rlocho

For a real càdlàg function f defined on a compact interval, its truncated variation at the level $c > 0$ is the infimum of total variations of functions uniformly approximating f with accuracy $c/2$ and (in opposite to the total variation) is always finite. In this paper, we discuss exponential integrability and concentration properties of the truncated variation of fractional Brownian motions, diffusions and Lévy processes. We develop a special technique based on chaining approach and using it we prove Gaussian concentration of the truncated variation for certain class of diffusions. Further, we give sufficient and necessary condition for the existence of exponential moment of order $\alpha > 0$ of truncated variation of Lévy process in terms of its Lévy triplet.

Keywords: diffusions; Gaussian processes; Lévy processes; sample boundedness; truncated variation

1. Introduction

Let $X = (X(t))_{t\geq 0}$ be a real valued stochastic process with càdlàg trajectories. In general, the total path variation of X on the compact interval $[a, b] \subset [0, +\infty)$, defined as

$$\mathrm{TV}\big(X, [a, b]\big) = \sup_n \sup_{a\leq t_0<t_1<\cdots<t_n\leq b} \sum_{i=1}^{n} \big|X(t_i) - X(t_{i-1})\big|,$$

may be (and in many most important cases is) almost surely infinite. However, in the neighborhood of every càdlàg path we may easily find a function with finite total variation.

Let f be a càdlàg function $f : [a, b] \to \mathbb{R}$ and let $c > 0$. The natural question arises, what is the smallest possible (or the greatest lower bound for the) total variation of functions from the ball $B(f, c/2) = \{g : \|f - g\|_\infty \leq c/2\}$, where $\|f - g\|_\infty := \sup_{s\in[a,b]} |f(s) - g(s)|$. Some bound from below reads as

$$\mathrm{TV}\big(g,[a,b]\big) \geq \mathrm{TV}^c\big(f,[a,b]\big),$$

where

$$\mathrm{TV}^c\big(f,[a,b]\big) := \sup_n \sup_{a\leq t_0<t_1<\cdots<t_n\leq b} \sum_{i=1}^n \max\big\{\big|f(t_i)-f(t_{i-1})\big|-c,0\big\} \tag{1.1}$$

and follows immediately from the inequality

$$\big|g(t_i)-g(t_{i-1})\big| \geq \max\big\{\big|f(t_i)-f(t_{i-1})\big|-c,0\big\}.$$

It is possible to show (cf. Łochowski [11]) that in fact we have equality

$$\inf\big\{\mathrm{TV}\big(g,[a,b]\big) : \|f-g\|_\infty \leq c/2\big\} = \mathrm{TV}^c\big(f,[a,b]\big) \tag{1.2}$$

attained for some function f^c from the ball $B(f,c/2)$.

Remark 1. Since we deal with càdlàg functions, a more natural setting of our problem would be the investigation of

$$\inf\big\{\mathrm{TV}\big(g,[a,b]\big) : g-\text{càdlàg}, d_D(f,g)\leq c/2\big\},$$

where d_D denotes the Skorohod metric. Since the total variation does not depend on the (continuous and strictly increasing) change of argument and the function f^c minimizing $\mathrm{TV}(g,[a,b])$ appears to be a càdlàg one, solutions of both problems coincide.

The quantity (1.1) is called truncated variation and it is finite for any càdlàg function, since every such a function may be uniformly approximated by step functions. Moreover, the truncated variation is a continuous and convex function of the parameter $c>0$ (cf. Łochowski [11]) and it obviously tends to the total variation as $c\downarrow 0$. For a process with paths with almost surely infinite total variation may be of interest to assess the rate at which TV^c diverges to infinity.

This was done so far for continuous semimartingales and it appears (cf. Łochowski and Miłoś [12]) that for any continuous semimartingale X we have that

$$c\cdot\mathrm{TV}^c\big(X,[a,b]\big) \to_{c\downarrow 0} \langle X\rangle_b - \langle X\rangle_a \qquad \text{almost surely}, \tag{1.3}$$

where $\langle\cdot\rangle$ denotes the quadratic variation of X. The truncated variation appears also implicitly in the paper Picard [14] where it corresponds to the double Lebesgue measure L^c of a trimmed tree at the level c, associated with a càdlàg path. In Picard [14] there were established deep connections of this measure, the variation index and the upper box (or Minkowski) dimension, as well as the counterparts of (1.3) in terms of L^c for fractional Brownian motions and stable Lévy processes.

For $t\geq 0$ denote $\mathrm{TV}^c(X,t)=\mathrm{TV}^c(X,[0,t])$. For X being the unique strong solution of the equation $X_0=0$, $\mathrm{d}X_t=\mu(X_t)\,\mathrm{d}t+\sigma(X_t)\,\mathrm{d}W_t$, $t\in[0,S]$, driven by a standard Brownian motion

W, with μ and σ satisfying some linear growth conditions, we have second order convergence result (cf. Łochowski and Miłoś [12], Theorem 10)

$$\mathrm{TV}^c(X,t) - \frac{\langle X\rangle_t}{c} \Rightarrow_{c\downarrow 0} \tilde{W}_{\langle X\rangle_t/3}, \tag{1.4}$$

where $\tilde{W}$ is a standard Brownian motion independent from W and the convergence "$\Rightarrow$" is understood as the weak functional convergence in $C([0,S],\mathbb{R})$ topology.

The truncated variation is more informative than p-variation, since the latter may be described in terms of the asymptotic properties of TV^c as $c \downarrow 0$ but for any fixed $c > 0$, $\mathrm{TV}^c(X,S)$ is a proper random variable and it is possible to consider its distribution. For $X = W$ and fixed $S, c > 0$ convergence result (1.4) seems to indicate very strong concentration of $\mathrm{TV}^c(W,S)$ around S/c, but it still does not tell anything about the tail probabilities of the functional considered.

These observations motivated us to study the integrability and concentration properties of the truncated variation in greater detail. Some investigation into this direction was already undertaken in Łochowski [10], where the existence of the moment generating function of the truncated variation of Brownian motion with drift on the whole real line was proven. In this paper, we obtain much stronger – Gaussian concentration result, by which we mean the integrability of $\exp(\alpha\,\mathrm{TV}^c(X,S)^2)$ for some positive α.

Another incentive for the study of the magnitude of truncated variation for possibly broad class of processes is the pathwise approach to stochastic integration. In Łochowski [13], it was shown that when both – integrand and integrator are semimartingales then it is possible to define the stochastic integral, with some correction term, as an almost sure limit of pathwise Lebesgue–Stieltjes integrals. The construction utilizes uniform approximation of the integrator with finite variation processes. The truncated variation gives the magnitude of such integrals, more precisely

$$\inf_{\|X-X^c\|_\infty \le c/2} \sup_{\|Y\|_\infty \le 1} \int_0^S Y_-\,\mathrm{d}X^c = \inf_{\|X-X^c\|_\infty \le c/2} \mathrm{TV}\big(X^c,S\big) = \mathrm{TV}^c(X,S),$$

where the supremum is over all càdlàg processes Y with absolute value uniformly bounded by 1 and the infimums are over all pathwise càdlàg approximations X^c of X such that $\|X - X^c\|_\infty := \sup_{t\ge 0}|X(t) - X^c(t)| \le c/2$.

In this paper, we study the magnitude of the truncated variation for a broad class of stochastic processes, including Gaussian processes, among them fractional Brownian motions, and diffusions. Further we also consider Lévy processes. Our main goal is to describe the tail behavior of $\mathrm{TV}^c(X,S)$ assuming that X satisfies some increment condition. We use various techniques depending on the assumption we make.

At the beginning, we use the chaining concept, we assume that X satisfies some exponential integrability condition on increments and deduce the exponential integrability of the truncated variation (e.g., diffiusions with bounded covariance and drift coefficients). The chaining approach was first used to study problems of sample boundedness of processes on the general index space Fernique [4,5]. The method was developed to give the full description of classes of processes that are sample bounded, under certain integrability condition Bednorz [1,2], Bednorz [3], Ledoux and Talagrand [8], Talagrand [17], and the small ball probability Li and Shao [9].

For a comprehensive study where many analytical examples are given, see Talagrand [18]. In our study, we need some modification of this idea, since we are interested in bounding the supremum of special sums of increments, not the supremum over increments itself. Therefore, we have to invent a special random variable of exponential integrability that bounds the truncated variation.

Our main guiding example is the class of fractional Brownian motions, that is, centered Gaussian processes W_H, $H \in (0,1)$, starting from 0 and such that $\mathbf{E}|W_H(t) - W_H(s)|^2 = |t-s|^{2H}$. One of the corollaries we get is the following concentration inequality

$$\mathbf{P}\bigl(\mathrm{TV}^c(W_H, S) \geq c^{(H-1)/H} S(A_H + B_H u)\bigr) \leq C_H \exp\bigl(-u^{2H}\bigr), \qquad \text{for } u \geq 0,$$

where A_H, B_H, C_H are constants; moreover, for $H \geq \frac{1}{2}$ one can set $C_H = 1$. By the homogeneity of increments, we deduce that for $Sc^{-1/H} \geq 2$, $\mathbf{E}\,\mathrm{TV}^c(W_H, S)$ is comparable with $c^{(H-1)/H} S$ and in this way we prove that for $u \geq 0$,

$$\mathbf{P}\bigl(\mathrm{TV}^c(W_H, S) \geq \mathbf{E}\,\mathrm{TV}^c(W_H, S)(\bar{A}_H + \bar{B}_H u)\bigr) \leq \bar{C}_H \exp\bigl(-u^{2H}\bigr), \tag{1.5}$$

for some constants $\bar{A}_H, \bar{B}_H, \bar{C}_H$ (again $\bar{C}_H = 1$ for $H \geq \frac{1}{2}$). In fact, any process with similar properties as the fractional Brownian motion, that is, satisfying some boundedness condition of the increments (inequality (2.2)) may be treated by our method.

Next, we turn to investigate the standard Brownian motion, that is, $W = W_{1/2}$, and diffusions driven by it. Here we can improve our result using the Markov property. It turns out that for Markov processes with moderate growth some local exponential integrability can be extended to the global one. Note that (1.5) implies the existence of the Laplace transform $\mathbf{E}\exp(\alpha\,\mathrm{TV}^c(W, S))$ for sufficiently small $\alpha > 0$; assuming the Markov property for diffusions with moderate growth we get the estimate for the Laplace transform of their truncated variations on the whole real line. The main result we get this way is Theorem 2, which for a standard Brownian motion and $Sc^{-2} \geq 2$ implies the following concentration inequality

$$\mathbf{P}\bigl(\mathrm{TV}^c(W, S) \geq \bar{A}\mathbf{E}\,\mathrm{TV}^c(W, S) + \bar{B}\sqrt{S}u\bigr) \leq \exp\bigl(-u^2\bigr), \qquad \text{for } u \geq 0,$$

here $\bar{A}$, $\bar{B}$ are universal constants. Therefore, the Gaussian concentration holds for the truncated variation of the standard Brownian motion. Our result gives better understanding of the already mentioned result (1.4) from which follows that $S^{-1/2}(\mathrm{TV}^c(W, S) - S/c)$ converges in distribution to $\mathcal{N}(0, 1/3)$ as $c \downarrow 0$.

We conclude the paper by proving sufficient and necessary condition for the finiteness of $\mathbf{E}\exp(\alpha\,\mathrm{TV}^c(X, S))$ for a Lévy process X, in terms of its generating triplet. Here we apply the method of level crossing stopping times.

The structure of the paper is as follows. In Section 2, we introduce the chaining approach which will lead us to the main result on the concentration for processes with increments of exponential decay. Then in Section 2.3, we discuss the application of the developed methodology to the fractional Brownian motions and then, in Section 3 its improvement for a standard Wiener process and diffusions with moderate growth. In Section 4 we deal with truncated variation of Lévy processes.

Remark 2. In the whole paper, any dependence of a nonnegative constant on some parameters is always indicated by listing them in brackets or in subscripts, for example, $C(n, \varepsilon)$ or $C_{n,\varepsilon}$.

2. The chaining approach

In this section, we prove the fundamental Theorem 1, which will allow us to establish integrability and concentration properties of the truncated variation for a broad class of processes satisfying some increment condition.

For simplicity, we consider processes indexed by a parameter from the metric space (T, d), where T is the compact interval $[0, S]$, $S > 0$, equipped with the distance $d(s,t) = |s - t|^q$, $s, t \in T$, where $0 < q < 1$. Further, we introduce an Orlicz function $\varphi : [0, +\infty) \to \mathbb{R}$ – convex, even, satisfying $\varphi(0) = 0$, $\varphi(1) = 1$, strictly increasing and such that there exists $L < +\infty$ such that for any $x, y \geq 0$,

$$\varphi^{-1}(xy) \leq L\big(\varphi^{-1}(x) + \varphi^{-1}(y)\big). \tag{2.1}$$

Moreover, we will require that $x \mapsto \varphi(x^q)$, $x \geq 0$, is also convex.

Remark 3. The convexity assumptions of φ may be weakened in such a way that φ is convex on some interval $[C_\varphi, \infty)$, where $C_\varphi \geq 0$, and $\varphi(x^q)$ is convex on some interval $[C_{\varphi,q}, \infty)$, where $C_{\varphi,q} \geq 0$.

The standard example of functions with properties mentioned are $\varphi_p(x) = 2^{x^p} - 1$, $p > 0$, for which condition (2.1) holds with $L_p = \max\{1, 2^{(1-p)/p}\}$. Note that when $p \geq 1$, φ_p is convex on whole interval $[0, +\infty)$ but when $0 < p < 1$, φ_p is convex only on the interval $[C_p; +\infty)$ where $C_p = (\frac{1-p}{p \ln 2})^{1/p}$. Clearly $\varphi_p(x^q) = \varphi_{pq}(x)$ and therefore this function is convex on the whole interval $[0; +\infty)$ if $pq \geq 1$ and convex on the interval $[C_{p,q}; +\infty)$, where $C_{p,q} = C_{pq}$, if $pq < 1$. We use the notation $C_p, C_{p,q}$ for all $p > 0$, $0 < q < 1$, setting $C_p = 0$ for $p \geq 1$ (thus $C_{p,q} = 0$ for $pq \geq 1$). Further, we denote $D_p = \varphi_p(C_p)$, $D_{p,q} = \varphi_p(C_{p,q}^q) = \varphi_{pq}(C_{pq})$. Note that $D_{p,q} = 0$ for $pq \geq 1$. In more general case, we will denote $D_\varphi = \varphi(C_\varphi)$ and $D_{\varphi,q} = \varphi(C_{\varphi,q}^q)$.

Let now $X(t)$, $t \in T$, be a stochastic process with increments controlled by φ. Namely

$$\mathbf{E}\varphi\left(\frac{|X(s) - X(t)|}{Cd(s,t)}\right) \leq 1 \tag{2.2}$$

for $s, t \in T$, $s \neq t$, where $0 < C < \infty$ is a universal constant.

Remark 4. In fact, in (2.2) one may consider any distance d of the form $d(s,t) = \eta(|s - t|)$, where η is positive, concave, increasing to ∞ and such that $\eta(0) = 0$. We choose $\eta(x) = x^q$, $0 < q < 1$, for the sake of simplicity, however we stress that our results can be easily extended to a more general η.

Condition (2.2) enables us to control the magnitude of the increments of the process X, while the truncated variation takes into account only increments greater than c (cf. formula (1.1)). Note that as the consequence of (2.2) and the compactness of T we obtain the existence of a separable modification of $X(t)$, $t \in T$. Then by the linear order of T we can define the càdlàg modification of X which we refer to from now on.

The fundamental result of this paper, from which exponential integrability and concentration properties will follow, is the following theorem.

Theorem 1. *Let $X(t)$, $t \in T$, satisfies (2.2). Then there exist random variables $Z_1, Z_2 \geq 0$ such that $\mathbf{E}Z_1, \mathbf{E}Z_2 \leq 1$ and for some universal constants $K_1(q), K_2(\varphi, q) < \infty$ the following estimate holds*

$$\mathrm{TV}^c(X, S) \leq c^{(q-1)/q} S\big[K_1(\varphi, q)\varphi^{-1}(Z_1 + D_\varphi) + K_2(\varphi, q)\big[\varphi^{-1}(Z_2 + D_{\varphi,q})\big]^{1/q}\big].$$

Remark 5. The main reason why the result holds is that (2.2) gives an exponential decay of increments with large jumps. Therefore, we can show a global upper bound on increments in the defined set approximation of the truncated variation. Such an idea is used to bound suprema of processes, for example, Bednorz [1], Fernique [4], Kwapień and Rosiński [6] and Talagrand [17]. In this paper, the main technical contribution is to invent a common upper bound for an arbitrary sum of truncated increments.

The meaning of the result the that for suitable φ and $0 < q < 1$ there holds some concentration inequality. To formulate results in an elegant way, observe that there exists $E_q \in [0; 1]$ such that $E_q + x^{1/q} \geq x$ for $x \geq 0$ and hence we get

$$E_q + \big[\varphi^{-1}\big(x + \max\{D_\varphi, D_{\varphi,q}\}\big)\big]^{1/q} \geq \varphi^{-1}(x + D_\varphi) \qquad \text{for } x \geq 0. \tag{2.3}$$

As a consequence of Theorem 1, (2.3) and Jensen's inequality we get the following corollary.

Corollary 1. *Under the assumptions of Theorem 1 there exist r.v. Z such that $Z \geq 0$, $\mathbf{E}Z \leq 1$ and for some constants $A_{\varphi,q}$, $B_{\varphi,q}$ the following estimate holds*

$$\mathrm{TV}^c(X, S) \leq c^{(q-1)/q} S\big[A_{\varphi,q} + B_{\varphi,q}\big[\varphi^{-1}\big(Z + \max\{D_\varphi, D_{\varphi,q}\}\big)\big]^{1/q}\big].$$

For $\varphi = \varphi_p$ let us denote $A_{p,q} = A_{\varphi,q}$ and $B_{p,q} = B_{\varphi,q}$. Applying Corollary 1, the Markov inequality and the fact that $D_{p,q} \geq D_p$ we obtain:

Corollary 2. *Let $X(t)$, $t \in T$, satisfies (2.2) with $\varphi = \varphi_p$. The following inequality holds*

$$\mathbf{P}\big(\mathrm{TV}^c(X, S) \geq c^{(q-1)/q} S[\bar{A}_{p,q} + \bar{B}_{p,q} u]\big) \leq \bar{D}_{p,q} \exp\big(-u^{pq}\big), \qquad \textit{for } u > 0,$$

where $\bar{A}_{p,q}, \bar{B}_{p,q}$ are universal constants, $\bar{A}_{p,q} = A_{p,q} + (2/\ln 2)^{1/(pq)} B_{p,q}$, $\bar{B}_{p,q} = (2/\ln 2)^{1/(pq)} B_{p,q}$ and $\bar{D}_{p,q} = D_{p,q} + 1$. In particular, $\bar{D}_{p,q} = 1$ for $pq \geq 1$.

To prove Theorem 1, we start with the construction of finite sets approximating T.

2.1. Approximating sequence

The first tool we need is a proper geometric approximation of the set T. The approximation consists of a sequence of finite sets $(T_n)_{n=0}^{\infty}$, $T_n \subset T$ constructed in such a way that for each point $t \in T$ and $n = 0, 1, 2, \ldots$, there exists a point $s \in T_n$, such that $s \le t$ and $d(s,t) \le r^{-n} S^q$. Here, we fix $r \ge 4$. One of possible constructions is the following

$$T_n = \left\{ k r^{-n/q} S : k = 0, 1, 2, \ldots \right\} \cap T. \tag{2.4}$$

For T_n defined by (2.4) and $t \in T$, by $\pi_n(t)$ we denote the unique point $s \in T_n$ such that $s \le t$ and $d(s,t) < r^{-n} S^q$. This way we define the function $\pi_n : T \to T_n$. We have $d(t, \pi_n(t)) < r^{-n} S^q$ for all $t \in T$ and $\pi_n(s) \le \pi_n(t)$ if $s \le t$. Note also that for $s, t \in T_n$, $s \ne t$, $d(s,t) \ge r^{-n} S^q$. Clearly

$$r^{n/q} < |T_n| = \left\lfloor r^{n/q} \right\rfloor + 1 \le r^{n/q} + 1. \tag{2.5}$$

Moreover for any $m = 1, 2, \ldots$

$$\sum_{n=0}^{m} r^{-n} |T_{n+1}| \le \sum_{n=0}^{m} r^{-n} \left(r^{(n+1)/q} + 1 \right) \le A(r,q) r^{m(1-q)/q}, \tag{2.6}$$

where $A(r,q) := r^{(2-q)/q} (r^{(1-q)/q} - 1)^{-1}$ (note that $r \ge 2$). For each $t \in T_{n+1}$ let $I_{n+1}(t)$ denote the set of the nearest neighbors of t in T_{n+1}, namely

$$I_{n+1}(t) = \left\{ s \in T_{n+1} : d(s,t) \le 2 r^{-n} S^q \right\}. \tag{2.7}$$

Observe that since $|s - t| \ge r^{-(n+1)/q} S$ for $s, t \in T_{n+1}$, $s \ne t$,

$$\left| I_{n+1}(t) \right| \le \frac{2^{1/q} r^{-n/q} S}{r^{-(n+1)/q} S} + 1 = 2^{1/q} r^{1/q} + 1 =: B(r,q). \tag{2.8}$$

2.2. Proof of the main theorem

The plan of the proof is the following. After having constructed the set approximation of T, we use this approximation to build a type of discretization of any given partition and derive a chaining bound on the truncated variation (Lemma 1). Then we turn to estimate each increment in the partition bound (Lemma 5) and finally apply the bounds as well as some technical observations (Lemmas 3, 4 and 2) to derive the required bounds (Lemmas 6, 7).

Our first step is to analyze a given partition $\Pi_n = \{t_0, t_1, \ldots, t_n\}$, where $0 \le t_0 < t_1 < \cdots < t_n \le S$. We decompose the set $\{1, \ldots, n\}$ into subsets J_m, $m = 0, 1, 2, \ldots$, defined in the following way

$$J_m = \left\{ i \in \{1, \ldots, n\} : r^{-m-1} S^q < d(t_{i-1}, t_i) \le r^{-m} S^q \right\}.$$

Let $M_0 := 12CL$, where L and C are constants appearing in (2.1) and (2.2). The level $m_0 \in \{0, 1, 2, \ldots\}$ such that

$$r^{-m_0-1} S^q < c/M_0 \le r^{-m_0} S^q$$

will be of particular meaning in the proof. Since Π_n is finite, $J_m = \varnothing$ for m large enough, say $m \ge N_0 \ge m_0$. We will use different bounds for $i \in J_m$ with $m > m_0$ and for $i \in J_m$ with $m \le m_0$.

Therefore, let us make the trivial separation

$$\sum_{i=1}^{n}\left(\left|X(t_i)-X(t_{i-1})\right|-c\right)_+ \le \sum_{m=0}^{m_0}\sum_{i\in J_m}\left(\left|X(t_i)-X(t_{i-1})\right|-c\right)_+ + \sum_{m=m_0+1}^{\infty}\sum_{i\in J_m}\left(\left|X(t_i)-X(t_{i-1})\right|-c\right)_+. \tag{2.9}$$

Now we turn to describe the chaining method which is the main tool in the proof. First, we fix $N \ge N_0$ and define $t_i^{N+1} = \pi_{N+1}(t_i)$, then for $l \in \{0, 1, \dots, N\}$ we put by the reverse induction $t_i^l = \pi_l(t_i^{l+1})$. Note that by the construction of π_l we preserve the order of the projections, namely $t_0^l \le t_1^l \le \dots \le t_n^l$ for any $0 \le l \le N+1$. Moreover since $N \ge N_0$ points $\{t_0^{N+1}, t_1^{N+1}, \dots, t_n^{N+1}\}$ are separated, that is, $t_i^{N+1} \ne t_{i-1}^{N+1}$, $i \in \{1, \dots, n\}$. Let us denote $\bar{m} = \max\{m, m_0\}$. For $i \in J_m$ with $m > m_0$, we estimate

$$\begin{aligned}\left(\left|X(t_i)-X(t_{i-1})\right|-c\right)_+ &\le \left(\left|X\left(t_i^{m+1}\right)-X\left(t_{i-1}^{m+1}\right)\right|-\frac{c}{3}\right)_+ + \sum_{s\in\{i-1,i\}}\left|X\left(t_s^{N+1}\right)-X(t_s)\right| \\ &\quad + \sum_{l=m+1}^{N}\sum_{s\in\{i-1,i\}}\left(\left|X\left(t_s^{l}\right)-X\left(t_s^{l+1}\right)\right|-2^{-l+\bar{m}}\frac{c}{3}\right)_+\end{aligned} \tag{2.10}$$

and for $i \in J_m$ with $m \le m_0$ we have

$$\begin{aligned}\left(\left|X(t_i)-X(t_{i-1})\right|-c\right)_+ &\le \left|X\left(t_i^{m+1}\right)-X\left(t_{i-1}^{m+1}\right)\right| \\ &\quad + \sum_{s\in\{i-1,i\}}\left|X\left(t_s^{N+1}\right)-X(t_s)\right| + \sum_{l=m+1}^{m_0}\sum_{s\in\{i-1,i\}}\left|X\left(t_s^{l}\right)-X\left(t_s^{l+1}\right)\right| \\ &\quad + \sum_{l=m_0+1}^{N}\sum_{s\in\{i-1,i\}}\left(\left|X\left(t_s^{l}\right)-X\left(t_s^{l+1}\right)\right|-2^{-l+\bar{m}}\frac{c}{3}\right)_+.\end{aligned} \tag{2.11}$$

Putting together estimates (2.9), (2.10) and (2.11), we obtain the following decomposition lemma.

Lemma 1. *For any partition $\Pi_n = \{t_0, \dots, t_n\}$, where $n \ge 0$, $0 \le t_0 < t_1 < \dots < t_n \le S$ and $N > m_0$ the following estimate holds*

$$\sum_{i=1}^{n}\left(\left|X(t_i)-X(t_{i-1})\right|-c\right)_+ \le V_1+V_2+W_1+W_2+\sum_{i=1}^{n}\sum_{s\in\{i-1,i\}}\left|X(t_s)-X\left(t_s^{N+1}\right)\right|,$$

where

$$V_1 := \sum_{m=0}^{m_0} \sum_{i \in J_m} \sum_{l=m+1}^{m_0} \sum_{s \in \{i-1,i\}} \left|X\left(t_s^l\right) - X\left(t_s^{l+1}\right)\right|;$$

$$W_1 := \sum_{m=0}^{m_0} \sum_{i \in J_m} \left|X\left(t_i^{m+1}\right) - X\left(t_{i-1}^{m+1}\right)\right|;$$

$$V_2 := \sum_{m=0}^{\infty} \sum_{i \in J_m} \sum_{l=\bar{m}+1}^{N} \sum_{s \in \{i-1,i\}} \left(\left|X\left(t_s^l\right) - X\left(t_s^{l+1}\right)\right| - 2^{-l+\bar{m}} \frac{c}{3}\right)_+;$$

$$W_2 := \sum_{m=m_0+1}^{\infty} \sum_{i \in J_m} \left(\left|X\left(t_i^{m+1}\right) - X\left(t_{i-1}^{m+1}\right)\right| - \frac{c}{3}\right)_+.$$

For each $i \in J_m$, $m \geq 0$ we say that $t_s^{m+1}, t_s^{m+2}, \ldots, t_s^{N+1}$, $s = i-1, i$ are path approximations of t_{i-1} and t_i, respectively (see Figure 1). Note that for $i \in \{1, \ldots, n-1\}$ there are two path approximations of t_i, one from the pair t_{i-1}, t_i and the second from the pair t_i, t_{i+1}, that coincide, starting from some point, yet may differ on the length since $i \in J_m$, $i+1 \in J_{m'}$ and numbers m and m' may be different. The fundamental property of the path approximation is that for

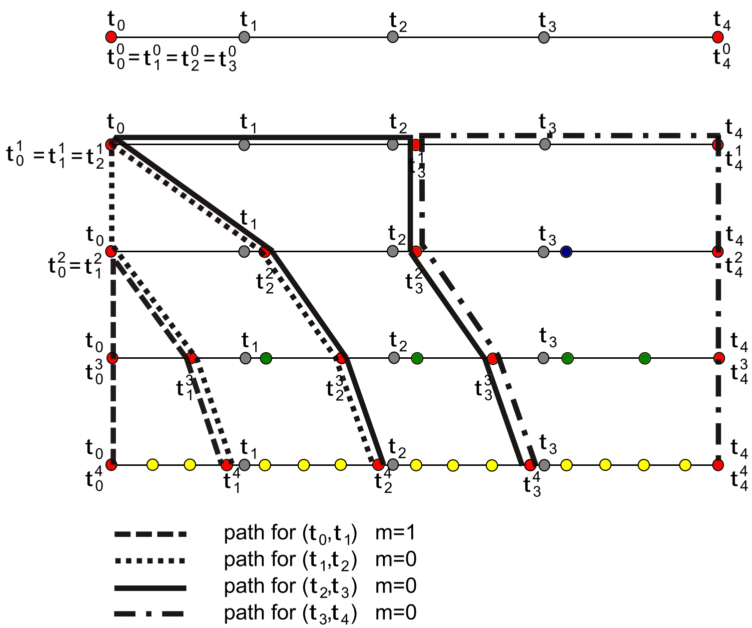

Figure 1. Path approximations.

a given $u \in T_{l+1}$ the step $\pi_l(u), u$ may occur in at most two path approximations of some t_i, $i \in \{0, 1, \ldots, n\}$.

Lemma 2. *Consider $u \in T_{l+1}$, $l \in \{0, 1, \ldots, n\}$. The step $\pi_l(u), u$ may occur in at most two path approximations of some t_i, $i \in \{0, 1, \ldots, n\}$, that is, there exits no more than one $i \in \{0, 1, \ldots, n\}$ such that $i \in J_m$, $m + 1 \le l$ and $t_i^l = \pi_l(u)$, $t_i^{l+1} = u$ or $i + 1 \in J_{m'}$, $m' + 1 \le l$ and $t_i^l = \pi_l(u)$, $t_i^{l+1} = u$ for some $m, m' = 0, 1, 2, \ldots, N$.*

Proof. Recall that $r \ge 4$. It suffices to prove that for a given $i \in J_m$, $l \ge m+1$ points t_i^{l+1} and t_{i-1}^{l+1} are different. Indeed since $t_0^{l+1} \le t_1^{l+1} \le \cdots \le t_n^{l+1}$ the property implies that there can be at most one $i \in \{0, 1, \ldots, n\}$ such that $t_i^{l+1} = u$. To prove the assertion, we use $d(t_i, t_{i-1}) > r^{-m-1} S^q$ which implies that for $l \ge m + 1$

$$\begin{aligned} d\big(t_i^{l+1}, t_{i-1}^{l+1}\big) &\ge r^{-m-1} S^q - d\big(t_{i-1}^{l+1}, t_{i-1}\big) - d\big(t_i^{l+1}, t_i\big) \\ &\ge r^{-m-1} S^q - 2 \sum_{j=l+1}^{\infty} r^{-j} S^q \ge r^{-m-1} S^q - 2\frac{r^{-m-2} S^q}{1 - r^{-1}} > 0. \end{aligned}$$

□

In the sequel, we will use two simple observations concerning increasing function ψ that is convex starting from some $C_0 \ge 0$, that is, convex for $x \ge C_0$.

Lemma 3. *Let $\psi : [0; +\infty) \to [0; +\infty)$ be a strictly increasing function. Assume that ψ is convex on the interval $[C_0; +\infty)$ where $C_0 \ge 0$, then for any nonnegative $x_1, \ldots, x_k$ and positive $\alpha_1, \ldots, \alpha_k$ such that $\sum_{i=1}^k \alpha_i \le M$ we have*

$$\sum_{i=1}^k \alpha_i x_i \le M \psi^{-1}\left(M^{-1} \sum_{i=1}^k \alpha_i \psi(x_i) + \psi(C_0)\right). \tag{2.12}$$

Proof. Observe that the function $\bar{\psi}(x) = \psi(x + C_0) - \psi(C_0)$ for $x \ge 0$ is convex, strictly increasing and such that $\bar{\psi}(0) = 0$. Consequently, $\bar{\psi}^{-1}(y) = \psi^{-1}(y + \psi(C_0)) - C_0$ is concave with $\bar{\psi}^{-1}(0) = 0$ and we have

$$\begin{aligned} \sum_{i=1}^k \alpha_i x_i &\le \sum_{i=1}^k \alpha_i \psi^{-1}\big(\psi(x_i) + \psi(C_0)\big) \\ &= \sum_{i=1}^k \alpha_i \big(\bar{\psi}^{-1}\big(\psi(x_i)\big) + C_0\big) \le M C_0 + M \sum_{i=1}^k \frac{\alpha_i}{M} \bar{\psi}^{-1}\big(\psi(x_i)\big) \\ &\le M C_0 + M \bar{\psi}^{-1}\left(\sum_{i=1}^k \frac{\alpha_i}{M} \psi(x_i)\right) \\ &= M \psi^{-1}\left(M^{-1} \sum_{i=1}^k \alpha_i \psi(x_i) + \psi(C_0)\right), \end{aligned}$$

where the last inequality follows from Jensen's inequality

$$\left(1-\sum_{i=1}^{k}\frac{\alpha_i}{M}\right)\bar{\psi}^{-1}(0)+\sum_{i=1}^{k}\frac{\alpha_i}{M}\bar{\psi}^{-1}\big(\psi(x_i)\big)\le\bar{\psi}^{-1}\left(\sum_{i=1}^{k}\frac{\alpha_i}{M}\psi(x_i)\right).$$

□

Further, we also have the following lemma.

Lemma 4. *For any strictly increasing function* $\psi:[0;+\infty)\to[0;+\infty)$ *such that* ψ *is convex on the interval* $[C_0;+\infty)$ *where* $C_0\ge 0$ *and for any* $M>0$ *and* $y\ge 0$, *we have*

$$\psi^{-1}\big(y+\psi(C_0)\big)\le\max\{M,1\}\psi^{-1}\big(y/M+\psi(C_0)\big). \tag{2.13}$$

Proof. Again, we consider the function $\bar{\psi}^{-1}$. If $M<1$, then (2.13) follows from the monotonicity of $\bar{\psi}^{-1}$. Now assume that $M\ge 1$. By concavity and $\bar{\psi}^{-1}(0)=0$, for $y\ge 0$ and $M\ge 1$, we get

$$M\bar{\psi}^{-1}(y/M)\ge\bar{\psi}^{-1}(y),$$

which reads as

$$M\big(\psi^{-1}\big(y/M+\psi(C_0)\big)-C_0\big)\ge\psi^{-1}\big(y+\psi(C_0)\big)-C_0,$$
$$M\psi^{-1}\big(y/M+\psi(C_0)\big)\ge\psi^{-1}\big(y+\psi(C_0)\big)+(M-1)C_0$$

and which gives

$$\psi^{-1}\big(y+\psi(C_0)\big)\le M\psi^{-1}\big(y/M+\psi(C_0)\big).$$

□

Now we formulate some basic bounds on increments in the chaining argument. For simplicity, we use the following notation

$$\Delta(u,v)=\varphi\left(\frac{|X(u)-X(v)|}{Cd(u,v)}\right),\qquad\text{for all } u,v\in T.$$

Recall that $\bar{m}=\max\{m,m_0\}$.

Lemma 5. *Suppose that* $i\in J_m$, $m\ge 0$ *then*

1. *for any* $m\le m_0$, $l\in\{m+1,\dots,m_0\}$

$$\big|X\big(t_i^l\big)-X\big(t_i^{l+1}\big)\big|\le Cr^{-l}S^q\varphi^{-1}\big(\Delta\big(t_i^l,t_i^{l+1}\big)\big);$$

2. *for any* $m\ge 0$, $l\in\{\bar{m}+1,\dots,N\}$

$$\left(\big|X\big(t_i^l\big)-X\big(t_i^{l+1}\big)\big|-2^{-l+\bar{m}}\frac{c}{3}\right)_+\le b_{l,\bar{m}}\big[\varphi^{-1}\big(a_{l,\bar{m}}^{-1}\Delta\big(t_i^l,t_i^{l+1}\big)\big)\big]^{1/q},$$

where

$$a_{l,\bar{m}} = \varphi\left(\frac{2^{-l+\bar{m}} r^l c}{6CLS^q}\right), \qquad b_{l,\bar{m}} = \left(\frac{c}{6} 2^{-l+\bar{m}}\right)^{(q-1)/q} \left(CLr^{-l} S^q\right)^{1/q};$$

3. *for* $m \le m_0$

$$\left|X\left(t_i^{m+1}\right) - X\left(t_{i-1}^{m+1}\right)\right| \le 2Cr^{-m} S^q \varphi^{-1}\left(\Delta\left(t_i^{m+1}, t_{i-1}^{m+1}\right)\right);$$

4. *and for* $m > m_0$

$$\left(\left|X\left(t_i^{m+1}\right) - X\left(t_{i-1}^{m+1}\right)\right| - \frac{c}{3}\right)_+ \le b_m\left[\varphi^{-1}\left(a_m^{-1} \Delta\left(t_i^{m+1}, t_{i-1}^{m+1}\right)\right)\right]^{1/q},$$

where

$$a_m = \varphi\left(\frac{r^m c}{12CLS^q}\right), \qquad b_m = \left(\frac{c}{6}\right)^{(q-1)/q} \left(2CLr^{-m} S^q\right)^{1/q}.$$

Remark 6. Note that the choice of M_0 in the definition of m_0 guarantees that $a_{l,\bar{m}} \ge 1$ and $a_m \ge 1$ for $m > m_0$. Moreover, for φ convex on the whole real line, that is, $C_\varphi = 0$ one can deduce $a_{l,\bar{m}}^{-1} \le (r/2)^{-l+\bar{m}}$ and $a_m^{-1} \le r^{-m+m_0+1}$.

Proof of Lemma 5. Let us denote $u = t_i^l$, then the construction of approximation paths implies that $t_i^{l+1} = \pi_l(u)$. Clearly $d(u, \pi_l(u)) \le r^{-l} S^q$ and hence

$$\left|X\left(\pi_l(u)\right) - X(u)\right| \le Cr^{-l} S^q \varphi^{-1}\left(\Delta\left(\pi_l(u), u\right)\right). \tag{2.14}$$

To prove the second assertion, we use (2.14) to get

$$\left(\left|X\left(\pi_l(u)\right) - X(u)\right| - 2^{-l+\bar{m}} \frac{c}{3}\right)_+ \le \left(Cr^{-l} S^q \varphi^{-1}\left(\Delta\left(\pi_l(u), u\right)\right) - 2^{-l+m} \frac{c}{3}\right)_+. \tag{2.15}$$

Now we rewrite (2.15) using $a_{l,\bar{m}}$

$$\begin{aligned} &\left(\left|X\left(\pi_l(u)\right) - X(u)\right| - 2^{-l+\bar{m}} \frac{c}{3}\right)_+ \\ &\quad \le \left[Cr^{-l} S^q \left(\varphi^{-1}\left(\Delta\left(\pi_l(u), u\right)\right) - L\varphi^{-1}(a_{l,\bar{m}})\right)_+ - 2^{-l+\bar{m}} \frac{c}{6}\right]_+. \end{aligned}$$

Therefore, we can apply (2.1) and see

$$\begin{aligned} &\left(\left|X\left(\pi_l(u)\right) - X(u)\right| - 2^{-l+\bar{m}} \frac{c}{3}\right)_+ \\ &\quad \le \left(CLr^{-l} S^q \varphi^{-1}\left(a_{l,\bar{m}}^{-1} \Delta\left(\pi_l(u), u\right)\right) - 2^{-l+\bar{m}} \frac{c}{6}\right)_+. \end{aligned} \tag{2.16}$$

Using the inequality $(x-1)_+ \le x^{1/q}$ valid for $x \ge 0$, we get

$$\begin{aligned}
&\left(\left|X\left(\pi_l(u)\right) - X(u)\right| - 2^{-l+\bar{m}}\frac{c}{3}\right)_+ \\
&\quad \le \frac{c}{6}2^{-l+\bar{m}}\left(6c^{-1}2^{l-\bar{m}}CLr^{-l}S^q\varphi^{-1}\left(a_{l,\bar{m}}^{-1}\Delta\left(\pi_l(u),u\right)\right)-1\right)_+ \\
&\quad \le \frac{c}{6}2^{-l+\bar{m}}\left(6c^{-1}2^{l-\bar{m}}CLr^{-l}S^q\right)^{1/q}\left[\varphi^{-1}\left(a_{l,\bar{m}}^{-1}\Delta\left(\pi_l(u),u\right)\right)\right]^{1/q} \\
&\quad = b_{l,\bar{m}}\left[\varphi^{-1}\left(a_{l,\bar{m}}^{-1}\Delta\left(\pi_l(u),u\right)\right)\right]^{1/q}.
\end{aligned}$$

To prove the third assertion, we first observe that $d(t_i,t_{i-1}) \le r^{-m}S^q$ for $i \in J_m$ and hence

$$\begin{aligned}
d\left(t_i^{m+1},t_{i-1}^{m+1}\right) &\le d(t_i,t_{i-1}) + d\left(t_i^{N+1},t_i\right) + d\left(t_{i-1}^{N+1},t_{i-1}\right) + \sum_{l=m+1}^{N}\left[d\left(t_i^{l+1},t_i^l\right)+d\left(t_{i-1}^{l+1},t_{i-1}^l\right)\right] \\
&\le r^{-m}S^q + 2\sum_{l=m+1}^{\infty} r^{-l}S^q,
\end{aligned}$$

so

$$d\left(t_i^{m+1},t_{i-1}^{m+1}\right) \le 2r^{-m}S^q. \tag{2.17}$$

Denoting $u = t_i^{m+1}$ and $v = t_{i-1}^{m+1}$ we get in the same way as (2.14) that

$$\left|X(u)-X(v)\right| \le 2Cr^{-m}S^q\varphi^{-1}\left(\Delta(u,v)\right).$$

Then using the same idea as for the second assertion we deduce the remaining inequality. □

We turn to apply the above lemmas to bound increments in the chaining bound formulated in Lemma 1. First, we consider a bound on $V_1 + W_1$.

Lemma 6. *There exists a universal constant $K_1(\varphi,r,q) < \infty$ and a random variable $Z_1 \ge 0$ independent from the partition Π_n, such that $\mathbf{E}Z_1 \le 1$ and for V_1 and W_1 defined in Lemma 1 one has*

$$V_1 + W_1 \le K_1(\varphi,r,q)c^{(q-1)/q}S\varphi^{-1}(Z_1 + D_\varphi).$$

Proof. By Lemma 2 and the first bound in Lemma 5, we get

$$\begin{aligned}
V_1 &\le 2\sum_{l=0}^{m_0}\sum_{u\in T_{l+1}}\left|X(u) - X\left(\pi_l(u)\right)\right| \\
&\le \mathcal{V}_1 := 2C\sum_{l=0}^{m_0} r^{-l}S^q \sum_{u\in T_{l+1}}\varphi^{-1}\left(\Delta\left(\pi_l(u),u\right)\right).
\end{aligned} \tag{2.18}$$

To bound W_1 we use (2.17), that is, that $d(t_i^{m+1}, t_{i-1}^{m+1}) \le 2r^{-m}S^q$ for $i \in J_m$. Using the already defined sets $I_m(u) = \{v \in T_{m+1} : d(u,v) \le 2r^{-m}S^q\}$, and the third bound in Lemma 5

$$\begin{aligned} W_1 &\le \sum_{l=0}^{m_0} \sum_{u \in T_{l+1}} \sum_{v \in I_{l+1}(u)} \big|X(u) - X(v)\big| \\ &\le \mathcal{W}_1 := C \sum_{l=0}^{m_0} 2r^{-l} S^q \sum_{u \in T_{l+1}} \sum_{v \in I_{l+1}} \varphi^{-1}\big(\Delta(u,v)\big). \end{aligned} \tag{2.19}$$

We calculate the sum of all weights appearing in (2.18) and (2.19). By (2.8) for each $u \in T_{m+1}$ we have $|I_{l+1}(u)| \le B(r,q)$ and hence, using also (2.6)

$$\begin{aligned} M_1 &:= \sum_{l=0}^{m_0} r^{-l} S^q \Big[|T_{l+1}| + \sum_{u \in T_{l+1}} \big|I_{l+1}(u)\big| \Big] \\ &\le \big[1 + B(r,q)\big] S^q \sum_{l=0}^{m_0} r^{-l} |T_{l+1}| \\ &\le A(r,q)\big[1 + B(r,q)\big] r^{m_0(1-q)/q} S^q. \end{aligned}$$

Therefore by $c \le M_0 r^{-m_0} S^q$ we get $M_0 r^{m_0(1-q)/q} S^q \le M_0^{1/q} c^{(q-1)/q} S$ and hence

$$M_1 \le M_0^{1/q} A(r,q)\big[1 + B(r,q)\big] c^{(q-1)/q} S.$$

Using Lemma 3 for φ which is convex above C_φ we get

$$\mathcal{V}_1 + \mathcal{W}_1 \le 2CM_1 \varphi^{-1}\big(Z_1 + \varphi(C_\varphi)\big) \le K_1(r,q) c^{(q-1)/q} S \varphi^{-1}\big(Z_1 + \varphi(C_\varphi)\big), \tag{2.20}$$

where $K_1(\varphi, r, q) := 2CM_0^{1/q} A(r,q)[1 + B(r,q)]$ (the dependence on φ is through C and L) and

$$Z_1 = M_1^{-1} \sum_{l=0}^{m_0} r^{-l} S^q \sum_{u \in T_{l+1}} \Big(\Delta\big(\pi_l(u), u\big) + \sum_{v \in I_{l+1}(u)} \Delta(u,v) \Big).$$

Obviously $Z_1 \ge 0$ and $\mathbf{E} Z_1 \le 1$ by (2.2) and the definition of M_1. Combining (2.18), (2.19) and (2.20) we get the result. □

Our second goal is to prove a bound for $V_2 + W_2$ in Lemma 1 above the level m_0.

Lemma 7. *There exists a universal constant $K_2(\varphi, r, q) < \infty$ and a random variable $Z_2 \ge 0$ independent from the partition Π_n such that $\mathbf{E} Z_2 \le 1$ and for V_2 and W_2 defined in Lemma 1 the following inequality holds*

$$V_2 + W_2 \le K_2(\varphi, r, q)\big[\varphi^{-1}(Z_2 + D_{\varphi,q})\big]^{1/q}.$$

Proof. First, we prove a bound for V_2. We analyze the increment

$$\left(\left|X\left(t_i^{l+1}\right)-X\left(t_i^l\right)\right|-2^{\bar{m}-l}\frac{c}{3}\right)_+,\qquad l>\bar{m}, i\in J_m, m\geq 0.$$

Using the second inequality in Lemma 5, we obtain that

$$\left(\left|X\left(t_i^{l+1}\right)-X\left(t_i^l\right)\right|-2^{-l+\bar{m}}\frac{c}{3}\right)_+\leq b_{l,\bar{m}}\left[\varphi^{-1}\left(a_{l,\bar{m}}^{-1}\Delta\left(t_i^{l+1},t_i^l\right)\right)\right]^{1/q}.$$

Now observe that $|t_i-t_{i-1}|\geq r^{-(\bar{m}+1)/q}S$ for $i\in J_m$, $m\geq 0$. Therefore,

$$\begin{aligned}\sum_{l=\bar{m}+1}^{N} b_{l,\bar{m}} &= \sum_{l=\bar{m}+1}^{N}\left(\frac{c}{6}2^{l-\bar{m}}\right)^{(q-1)/q}\left(CLr^{-l}S^q\right)^{1/q}\\ &\leq \left(6^{1-q}CL\right)^{1/q}\sum_{l=\bar{m}+1}^{\infty}\left(2^{(q-1)/q}r^{-1/q}\right)^{l-\bar{m}}c^{(q-1)/q}r^{-\bar{m}/q}S\\ &\leq M_2c^{(q-1)/q}|t_i-t_{i-1}|,\end{aligned}\tag{2.21}$$

where $M_2=M_2(\varphi,r,q)$ is defined by

$$\begin{aligned}\left(6^{1-q}CL\right)^{1/q}r^{1/q}\sum_{l=\bar{m}+1}^{\infty}\left(2^{(q-1)/q}r^{-1/q}\right)^{l-\bar{m}} &\leq \left((12)^{1-q}CL\right)^{1/q}\sum_{l'=0}^{\infty}\left(2^{(1-q)/q}r^{-1/q}\right)^{l'}\\ &= \left((12)^{1-q}CL\right)^{1/q}\left(1-2^{(1-q)/q}r^{-1/q}\right)^{-1}=:M_2.\end{aligned}$$

Consequently

$$\sum_{m=0}^{\infty}\sum_{i\in J_m}\sum_{l=\bar{m}+1}^{N}\sum_{s\in\{i-1,i\}}b_{l,\bar{m}}\leq 2M_2c^{(q-1)/q}\sum_{i=1}^{n}|t_i-t_{i-1}|=2M_2c^{(q-1)/q}S.$$

Thus we can apply Lemma 3 for $\varphi(x^q)$ which is convex above $C_{\varphi,q}$ and get

$$V_2\leq 2M_2c^{(q-1)/q}S\left[\varphi^{-1}(\bar{V}_2+D_{\varphi,q})\right]^{-1},\tag{2.22}$$

where

$$\bar{V}_2:=(2M_2)^{-1}\sum_{m=0}^{\infty}\sum_{i\in J_m}\sum_{l=\bar{m}+1}^{N}\sum_{s\in\{i-1,i\}}\frac{\bar{b}_{l,\bar{m}}}{a_{l,\bar{m}}}\Delta\left(t_i^l,t_i^{l+1}\right)$$

and

$$\bar{b}_{l,\bar{m}}=\left(c^{(q-1)/q}S\right)^{-1}b_{l,\bar{m}}=\left(6^{-1}2^{-l+\bar{m}}\right)^{(q-1)/q}\left(CLr^{-l}\right)^{1/q}.$$

Observe that

$$\bar{b}_{l,\bar{m}} \le \bar{b}_{l,m_0}, a_{l,\bar{m}} \ge a_{l,m_0},$$

which implies $\bar{b}_{l,\bar{m}}/a_{l,\bar{m}} \le \bar{b}_{l,m_0}/a_{l,m_0}$. Hence by Lemma 2, we have

$$\bar{V}_2 \le \mathcal{V}_2 := M_2^{-1} \sum_{l=m_0+1}^{\infty} \frac{\bar{b}_{l,m_0}}{a_{l,m_0}} \sum_{u \in T_{l+1}} \Delta\bigl(\pi_l(u), u\bigr). \tag{2.23}$$

By the construction, $a_{l,m_0} \ge 1$, furthermore by (2.5)

$$\begin{aligned} M_2^{-1} \sum_{l=m_0+1}^{\infty} \frac{\bar{b}_{l,m_0}}{a_{l,m_0}} |T_{l+1}| &\le M_2^{-1} \sum_{l=m_0+1}^{\infty} \frac{(6^{-1}2^{-l+m_0})^{(q-1)/q}(CLr^{-l})^{1/q}}{\varphi((2^{-l+m_0}r^l c)/(6CLS^q))} \bigl(r^{(l+1)/q}+1\bigr) \\ &\le 2M_2^{-1}\bigl((12)^{1-q}CL\bigr)^{1/q} \sum_{l'=0}^{\infty} \frac{2^{l'(1-q)/q}}{\varphi(2^{-l'}r^{l'})} =: M_3, \end{aligned}$$

where we have used the fact that $(r^{(l+1)/q}+1) \le 2r^{(l+1)/q}$ and the definition of m_0, that is, $r^{-m_0-1}S^q < c/M_0$, $M_0 = 12CL$, together with the monotonicity of φ. Note that by Remark 6, for convex φ, that is, $C_\varphi = 0$,

$$\sum_{l'=0}^{\infty} \frac{2^{l'(1-q)/q}}{\varphi(2^{-l'}r^{l'})} \le \sum_{l'=0}^{\infty} 4^{l'} r^{-l'} = \bigl(1-4r^{-1}\bigr)^{-1}. \tag{2.24}$$

For φ which is convex for $x \ge C_\varphi$ basically the same argument works but l' must be large enough to apply the convexity. Indeed, using that $\psi(x) = \varphi(x+C_\varphi) - \varphi(C_\varphi)$ is convex and $\psi(0) = 0$ we deduce $\psi(2^{-l'}r^{l'}x) \ge 2^{-l'}r^{l'}\psi(x)$ for $x \ge 0$ and thus for all $x \ge 0$,

$$\varphi\bigl(2^{-l'}r^{l'}x + C_\varphi\bigr) \ge 2^{-l'}r^{l'}\bigl(\varphi(x+C_\varphi) - \varphi(C_\varphi)\bigr) + \varphi(C_\varphi). \tag{2.25}$$

Now choosing a suitable x one can get a bound similar to (2.24) yet for general φ. Note that in this case the bounding constant may depend on φ. It proves that $M_3 < \infty$. Finally, by (2.22), (2.23) and Lemma 4 we get

$$V_2 \le 2M_2 \max\{M_3, 1\} c^{(q-1)/q} S\bigl[\varphi^{-1}(\mathcal{V}_2/M_3 + D_{\varphi,q})\bigr]^{1/q}. \tag{2.26}$$

Clearly, by (2.2) and the definition of M_3 we have $\mathbf{E}\mathcal{V}_2/M_3 \le 1$.

A similar argument can be used to bound increments in W_2. Namely using the forth inequality in Lemma 5 we get that for $m > m_0$ and $i \in J_m$

$$\Bigl(\bigl|X\bigl(t_i^{m+1}\bigr) - X\bigl(t_{i-1}^{m+1}\bigr)\bigr| - \frac{c}{3}\Bigr)_+ \le b_m\bigl[\varphi^{-1}\bigl(a_m^{-1}\Delta\bigl(t_i^{m+1}, t_{i-1}^{m+1}\bigr)\bigr)\bigr]^{1/q}.$$

Using that $r^{-(m+1)/q}S \le |t_i - t_{i-1}| \le r^{-m/q}S$ we get

$$b_m = \left(\frac{c}{6}\right)^{(q-1)/q} \left(2CLr^{-m}S^q\right)^{1/q} \le M_4 c^{(q-1)/q}|t_i - t_{i-1}|,$$

where $M_4 = (2 \cdot 6^{1-q}CLr^{-1})^{1/q}$. Therefore,

$$\sum_{m=m_0+1}^{\infty} \sum_{i \in J_m} b_m \le M_4 c^{(q-1)/q} \sum_{i=1}^{n} |t_i - t_{i-1}| = M_4 c^{(q-1)/q} S,$$

and thus using Lemma 3 for $\varphi(x^q)$ we get

$$W_2 \le M_4 c^{(q-1)/q} S\left[\varphi^{-1}(\bar{W}_2 + D_{\varphi,q})\right]^{1/q}, \tag{2.27}$$

where

$$\bar{W}_2 := M_4^{-1} \sum_{m=m_0+1}^{\infty} \sum_{i \in J_m} \frac{\bar{b}_m}{a_m} \Delta\left(t_i^{m+1}, t_{i-1}^{m+1}\right)$$

and $\bar{b}_m = (c^{(q-1)/q}S)^{-1}b_m = (2 \cdot 6^{1-q}CLr^{-m})^{1/q}$. By (2.17) we have $d(t_i^{m+1}, t_{i-1}^{m+1}) \le 2r{-}mS^q$ and thus using the definition of the set $I_{m+1}(u)$ for each $m > m_0$ and $u \in T_{m+1}$

$$\bar{W}_2 \le \mathcal{W}_2 = M_4^{-1} \sum_{m=m_0+1}^{\infty} \frac{\bar{b}_m}{a_m} \sum_{u \in T_{m+1}} \sum_{v \in I_{m+1}(u)} \Delta(u, v).$$

Note that by (2.5), (2.8)

$$\sum_{u \in T_{m+1}} \left|I_{m+1}(u)\right| \le 2^{-1}B(r, q)\left(r^{(m+1)/q} + 1\right) \le B(r, q)r^{(m+1)/q}.$$

Hence

$$M_4^{-1}B(r, q) \sum_{m=m_0+1}^{\infty} \frac{\bar{b}_m}{a_m} r^{(m+1)/q} M_4^{-1}B(r, q) \sum_{m=m_0+1}^{\infty} \frac{(2 \cdot 6^{1-q}CLr)^{1/q}}{\varphi((r^m c)/(12CLS^q))}$$

$$\le M_4^{-1}B(r, q)\left(2 \cdot 6^{1-q}CLr\right)^{1/q} \sum_{m'=0}^{\infty} \left(\varphi\left(r^{m'}\right)\right)^{-1} =: M_5,$$

where in the last line we used that $r^{-m_0-1}S^q < c/M_0$, $M_0 = 12CL$. The same argument as for M_3 proves that $M_5 < \infty$. Note that in the case of convex φ we can easily bound $\sum_{m'=0}^{\infty}(\varphi(r^{m'}))^{-1}$ by $(1 - r^{-1})^{-1}$. By (2.27) and Lemma 4, we get

$$W_2 \le M_4 \max\{M_5, 1\} c^{(q-1)/q} S\left[\varphi^{-1}(\mathcal{W}_2/M_5 + D_{\varphi,q})\right]^{1/q}. \tag{2.28}$$

Obviously $\mathbf{E}\mathcal{W}_2/M_5 \leq 1$, consequently by (2.26), (2.28) and Jensen's inequality we obtain the desired result. □

Now we are ready to finish the proof of Theorem 1.

Proof of Theorem 1. Note that for fixed q and φ we may minimize constants $K_1(\varphi, r, q)$ and $K_2(\varphi, r, q)$ appearing Lemmas 6, 7 with respect to $r \geq 4$. It is clear from our discussion about the finitness of M_3, M_5 that one can set $r = 4$ in the case of convex φ. If $C_\varphi > 0$ the choice of $r \geq 4$ may be of meaning as we have explained in (2.25). Such minimal constants depend only on φ and q, and we will denote them by $K_1(\varphi, q)$ and $K_2(\varphi, q)$ respectively. Now it suffices to use Lemma 1, then universal bounds given in Lemmas 6, 7 and finally let $N \to \infty$. Recall that by the construction variables Z_1 and Z_2 of Lemmas 6, 7 do not depend on N and $\lim_{N\to\infty} d(t, \pi_{N+1}(t)) = 0$ for any $t \in T$. From condition (2.2), for a given partition $\Pi_n = \{t_0, t_1, \dots, t_n\}$ we get $R_N := \sum_{i=1}^n |X(t_i) - X(t_i^{N+1})| \to 0$ in probability as $N \uparrow +\infty$. Taking subsequence N_k such that $R_{N_k} \to 0$ almost surely, we get the universal bound for the sum $\sum_{i=1}^n (|X(t_i) - X(t_{i-1})| - c)_+$. Since Π_n was arbitrary we get the result for $\mathrm{TV}^c(X, S)$. □

2.3. Application to the fractional Brownian motion

Let $W_H(t)$, $t \geq 0$, be a fractional Brownian motion of the Hurst parameter $H \in (0, 1)$, that is, a centered Gaussian process which has the following covariance function

$$\mathbf{E}\big(W_H(s)W_H(t)\big) = \tfrac{1}{2}\big(s^{2H} + t^{2H} - |s-t|^{2H}\big). \tag{2.29}$$

Let us consider $T = [0, S]$ with distance $d(s, t) = |t - s|^H$. From (2.29), it follows that $W_H(t) - W_H(s) \sim \mathcal{N}(0, |t-s|^{2H})$ and thus, for some constant $C(H)$,

$$\mathbf{E}\varphi_2\left(\frac{|W_H(t) - W_H(s)|}{C(H)|t-s|^H}\right) \leq 1, \qquad \text{for } s, t \in T, s \neq t.$$

Consequently, all assumptions of Corollary 2 are satisfied with $p = 2$, $q = H$ and we get the following corollary.

Corollary 3. *For any fractional Brownian motion $W_H(t)$, $t \in T$, the following inequality holds*

$$\mathbf{P}\big(\mathrm{TV}^c(W_H, S) \geq c^{(H-1)/H} S(A_H + B_H u)\big) \leq C_H \exp\big(-u^{2H}\big), \qquad \textit{for } u > 0,$$

where A_H, B_H, C_H are universal constants and $C_H = 1$ for $H \geq 1/2$.

Note that Corollary 3 implies that $\mathbf{E}\,\mathrm{TV}^c(W_H, S) \leq K_H c^{(H-1)/H} S$, where $K_H < \infty$. On the other hand $c^{(H-1)/H} S$ is also the proper lower bound for $\mathbf{E}\,\mathrm{TV}^c(W_H, S)$ when $Sc^{-1/H}$ is not too small. Indeed, let us consider the partition $\Pi = \{0 \leq t_0 < t_1 < \cdots < t_N \leq S\}$ given by $t_i = ic^{1/H}$, $i = 0, 1, 2, \dots, N = \lfloor Sc^{-1/H} \rfloor$. We have

$$\mathrm{TV}^c(W_H, S) \geq \sum_{i=1}^N \big(\big|W_H(t_i) - W_H(t_{i-1})\big| - c\big)_+.$$

Clearly, for $Sc^{-1/H} \geq 2$, $N > Sc^{-1/H} - 1 \geq Sc^{-1/H}/2$ and $\mathbf{E}(|W_H(t_i) - W_H(t_{i-1})| - c)_+ \geq k_H c$ for some positive constant k_H. It proves that when $Sc^{-1/H} \geq 2$, $c^{(H-1)/H} S$ is comparable with $\mathbf{E}\,\mathrm{TV}^c(W_H, S)$ up to a constant depending only on H. Therefore, we have another formulation of Corollary 3.

Corollary 4. *Assume that $Sc^{-1/H} \geq 2$. For any fractional Brownian motion $W_H(t)$, $t \in T$, the following inequality holds*

$$\mathbf{P}\big(\mathrm{TV}^c(W_H, S) \geq \mathbf{E}\,\mathrm{TV}^c(W_H, S)(\bar{A}_H + \bar{B}_H u)\big) \leq \bar{C}_H \exp\big(-u^{2H}\big), \qquad \text{for } u > 0,$$

where $\bar{A}_H, \bar{B}_H, \bar{C}_H < \infty$ are universal constants. Moreover $\bar{C}_H = 1$ for $H \geq 1/2$.

3. Application to the standard Brownian motion and diffusions

For a standard Brownian motion $W = W_{1/2}$, which is the only fractional Brownian motion with independent increments one may, using this property, strengthen the results obtained for general fBm and obtain Gaussian concentration of $\mathrm{TV}^c(W, S)$. The generalization of this result for diffusions with moderate growth, driven by W, is also possible.

Let us assume that X_t, $t \geq 0$, is a one-dimensional diffusion satisfying

$$X(t) = x_0 + \int_0^t \mu\big(s, X(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(s, X(s)\big)\,\mathrm{d}W(s). \tag{3.1}$$

We assume that $\sigma : [0; +\infty) \times \mathbb{R} \to [-R; R]$ is measurable and bounded (i.e., $0 < R < +\infty$) and $\mu : [0; +\infty) \times \mathbb{R} \to \mathbb{R}$ is measurable and satisfying the following linear growth condition: there exists $C, D \geq 0$ such that for all $t \geq 0$

$$\big|\mu(t, x)\big| \leq C + D|x|. \tag{3.2}$$

We will also need the natural assumption that X is a Markov process. With this assumption, we have the following theorem.

Theorem 2. *For X being a Markov process satisfying (3.1) with μ and σ as above and $\lambda \geq 0$ one has*

$$\begin{aligned}\mathbf{E}\exp\big(\lambda\,\mathrm{TV}^c(X, S)\big) &\leq 2\exp\big(\lambda^2 S\alpha_R + \lambda S c^{-1}\beta_R + \lambda\gamma_{x_0,C,D,S}\big)\\ &\quad \times \big(1 + 8\lambda\eta_{D,R,S}\exp\big(\lambda^2\eta_{D,R,S}^2\big)\big),\end{aligned}$$

where $\gamma_{x_0,C,D,S} = (C + D|x_0|)S\mathrm{e}^{DS}$, $\delta_{D,S} = DS\mathrm{e}^{DS}$ and $\eta_{D,R,S} = \delta_{D,S} R\sqrt{S/2}$. In particular, when $D = 0$ we get

$$\mathbf{E}\exp\big(\lambda\,\mathrm{TV}^c(X, S)\big) \leq 2\exp\big(\lambda^2 S\alpha_R + \lambda S\big(c^{-1}\beta_R + C\big)\big)$$

and for the standard Brownian motion $X = W$ we get

$$\mathbf{E}\exp\big(\lambda\,\mathrm{TV}^c(W, S)\big) \leq 2\exp\big(\lambda^2 S\alpha + \lambda S c^{-1}\beta\big), \tag{3.3}$$

where α, β are universal constants.

Proof. Let us define

$$M(t) := \int_0^t \mu\big(s, X(s)\big)\, \mathrm{d}s, \qquad Y(t) := \int_0^t \sigma\big(s, X(s)\big)\, \mathrm{d}W(s)$$

and $Y^* = \sup_{0\le s\le S} |Y(s)|$. We have $X(t) = x_0 + M(t) + Y(t)$, and due to (3.2) we estimate

$$\begin{aligned}\big|M(t)\big| &\le \int_0^t \big|\mu\big(s, X(s)\big)\big|\, \mathrm{d}s \le \int_0^t C + D\big|X(s)\big|\, \mathrm{d}s \\ &\le \int_0^t C + D|x_0| + D\big|M(s)\big| + DY^*\, \mathrm{d}s \\ &\le \big(C + D|x_0| + DY^*\big)S + D\int_0^t \big|M(s)\big|\, \mathrm{d}s.\end{aligned} \tag{3.4}$$

Hence, from Gronwall's lemma (cf. Revuz and Yor [15], Appendix §1), we get

$$\big|M(t)\big| \le \big(C + D|x_0| + DY^*\big)S\mathrm{e}^{Dt}. \tag{3.5}$$

Notice that due to (3.5) M is adapted, absolute continuous process with locally bounded total variation. Indeed, repeating estimates (3.4) and using (3.5) we get

$$\begin{aligned}\mathrm{TV}(M, S) &\le \int_0^S \big|\mu\big(s, X(s)\big)\big|\, \mathrm{d}s \\ &\le \big(C + D|x_0| + DY^*\big)S + D\int_0^S \big|M(t)\big|\, \mathrm{d}s \\ &\le \big(C + D|x_0| + DY^*\big)S + D\big(C + D|x_0| + DY^*\big)S\int_0^S \mathrm{e}^{Dt}\, \mathrm{d}s \\ &= \big(C + D|x_0|\big)S\mathrm{e}^{DS} + DS\mathrm{e}^{DS}Y^*.\end{aligned} \tag{3.6}$$

($\mathrm{TV} = \mathrm{TV}^0$ denotes here the total variation.)

By Łochowski and Miłoś [12], Fact 17, we have

$$\mathrm{TV}^c(X, S) \le \mathrm{TV}(M, S) + \mathrm{TV}^c(Y, S). \tag{3.7}$$

Now we will investigate $\mathrm{TV}^c(Y, S)$.

First, let us prove that Y satisfies condition (2.1) with $\varphi = \varphi_2$ and $d(s,t) = |s-t|^{1/2}$. Indeed, let us fix $0 \le s < t \le S$ and consider the following martingale $Z(u) := Y(s+u) - Y(s)$, $u \in [0; t-s]$. We have

$$Z(u) = \int_s^{s+u} \sigma\big(\tau, X(\tau)\big)\, \mathrm{d}W(\tau)$$

and

$$\langle Z\rangle(u) = \int_s^{s+u} \sigma\big(\tau, X(\tau)\big)^2 \, \mathrm{d}\tau \le R^2(t-s).$$

Hence, by Bernstein's inequality (cf. Revuz and Yor [15], Chapter IV, Exercise 3.16), we have

$$\begin{aligned}
\mathbf{P}\big(\big|Y(t)-Y(s)\big| \ge x\big) &\le 2\mathbf{P}\Big(\sup_{u\in[0;t-s]} Z(u) \ge x\Big) \\
&= 2\mathbf{P}\Big(\sup_{u\in[0;t-s]} Z(u) \ge x, \langle Z\rangle(t-s) \le R^2(t-s)\Big) \\
&\le 2\exp\big(-x^2/\big(2R^2(t-s)\big)\big).
\end{aligned} \tag{3.8}$$

From (3.8), we immediately get that Y satisfies condition (2.1) for $\varphi = \varphi_2$ and $d(s,t) = |s-t|^{1/2}$. Hence, from Corollary 2 we obtain the following bound on the tails of $\mathrm{TV}^c(Y, S)$:

$$\mathbf{P}\big(\mathrm{TV}^c(Y, S) \ge c^{-1} S(A + Bu)\big) \le \mathrm{e}^{-u}, \tag{3.9}$$

where $A = A(R)$ and $B = B(R)$ depend on R only. Notice that for $\delta > 0$ applying Bernstein's inequality to Y^* we get $\mathbf{P}(Y^* \ge x) \le 2\exp(-x^2/(2R^2 S))$ and using integration by parts we have

$$\mathbf{E}\exp\big(\delta Y^*\big) \le 1 + 2\delta \int_0^\infty \mathrm{e}^{\delta y} \mathrm{e}^{-y^2/(2R^2 S)} \, \mathrm{d}y \le 1 + 8\delta R\sqrt{S/2}\mathrm{e}^{\delta^2 R^2 S/2}. \tag{3.10}$$

Now, we will strengthen estimate (3.9) using the Markov property of X. First, using (3.9) and integration by parts we have

$$\mathbf{E}\exp\big(\lambda\big[\mathrm{TV}^c(Y, S) - c^{-1} S A\big]\big) \le \frac{1}{1 - \lambda S B/c} \tag{3.11}$$

for $\lambda < c(SB)^{-1}$. Let now $S = S_1 + S_2$, where $S_1, S_2 > 0$. Using the inequality $\mathrm{TV}^c(Y, S) \le \mathrm{TV}^c(Y, S_1) + c + \mathrm{TV}^c(Y, [S_1, S])$, which follows easily from the estimate:

$$\big(\big|Y(t) - Y(u)\big| - c\big)_+ \le \big(\big|Y(t) - Y(S_1)\big| - c\big)_+ + \big(\big|Y(S_1) - Y(u)\big| - c\big)_+ + c$$

for $0 \le t < S_1 < u \le S$, and then the Markov property of X we get

$$\begin{aligned}
&\mathbf{E}\exp\big(\lambda\big[\mathrm{TV}^c(Y, S) - c^{-1} S A\big]\big) \\
&\quad \le \mathbf{E}\exp\big(\lambda \,\mathrm{TV}^c(Y, S_1) + \lambda c + \lambda\, \mathrm{TV}^c\big(Y, [S_1, S]\big) - \lambda c^{-1} S A\big) \\
&\quad = \mathrm{e}^{\lambda c}\mathbf{E}\big(\mathrm{e}^{\lambda \,\mathrm{TV}^c(Y, S_1) - \lambda c^{-1} S_1 A}\mathbf{E}\big[\mathrm{e}^{\lambda\,\mathrm{TV}^c(Y,[S_1,S]) - \lambda c^{-1} S_2 A} | X(S_1)\big]\big) \\
&\quad \le \mathrm{e}^{\lambda c}\frac{1}{1 - \lambda S_1 B/c}\frac{1}{1 - \lambda S_2 B/c}.
\end{aligned} \tag{3.12}$$

The last inequality follows by (3.11), since the right-hand side of (3.11) does not depend on x_0, and using the Markov property in similar way we have the universal estimate for the conditional expectation

$$\mathbf{E}\big(\exp\big\{\lambda\,\mathrm{TV}^c\big(Y,[S_1,S]\big)-\lambda c^{-1}S_2A\big\}\big|X(S_1)=x_1\big)\le\frac{1}{1-\lambda S_2B/c}$$

(note that the length of interval $[S_1,S]$ is S_2). Notice now that from (3.12) it follows that $\mathbf{E}\exp(\lambda[\mathrm{TV}^c(Y,S)-c^{-1}S\bar{A}])<+\infty$ for $\lambda<\min\{c(S_1B)^{-1},c(S_2B)^{-1}\}$. Let us fix integer $n\ge 1$. Iterating (3.12) we obtain

$$\mathbf{E}\exp\big(\lambda\big[\mathrm{TV}^c(Y,S)-c^{-1}SA\big]\big)\le \mathrm{e}^{\lambda c(n-1)}\bigg(\frac{1}{1-\lambda SB(cn)^{-1}}\bigg)^n \tag{3.13}$$

for $\lambda<cn(SB)^{-1}$, which gives that $\mathbf{E}\exp(\lambda[\mathrm{TV}^c(Y,S)-c^{-1}SA])<+\infty$ for any $\lambda\in\mathbb{R}$. Now, let us fix $\lambda>0$ and set $n=\lceil 2\lambda SBc^{-1}\rceil$. Using (3.13), we get

$$\begin{aligned}\mathbf{E}\exp\big(\lambda\big[\mathrm{TV}^c(Y,S)-c^{-1}SA\big]\big)&\le \mathrm{e}^{\lambda c(n-1)}2^n\\ &\le 2\exp\big(2\lambda^2SB+2(\ln 2)\lambda SBc^{-1}\big)\end{aligned}$$

and thus

$$\begin{aligned}\mathbf{E}\exp\big(\lambda\,\mathrm{TV}^c(Y,S)\big)&\le 2\exp\big(2\lambda^2SB+\lambda Sc^{-1}\big(A+2(\ln 2)B\big)\big)\\ &=2\exp\big(\lambda^2S\alpha_R+\lambda Sc^{-1}\beta_R\big),\end{aligned} \tag{3.14}$$

where $\alpha_R=2B=2B(R)$ and $\beta_R=A+2(\ln 2)B=A(R)+2(\ln 2)B(R)$. Now, from (3.7), (3.6) and (3.14) we get

$$\begin{aligned}\mathbf{E}\exp\big(\lambda\,\mathrm{TV}^c(X,S)\big)&\le E\exp\big(\lambda\,\mathrm{TV}(M,S)+\lambda\,\mathrm{TV}^c(Y,S)\big)\\ &\le 2\exp\big(\lambda^2S\alpha_R+\lambda Sc^{-1}\beta_R+\lambda\gamma_{x_0,C,D,S}\big)\mathbf{E}\exp\big(\lambda\delta_{D,S}Y^*\big),\end{aligned}$$

where $\gamma_{x_0,C,D,S}=(C+D|x_0|)S\mathrm{e}^{DS}$, $\delta_{D,S}=DS\mathrm{e}^{DS}$. Finally, using (3.10) with $\delta=\lambda\delta_{D,S}$ we get

$$\begin{aligned}\mathbf{E}\exp\big(\lambda\,\mathrm{TV}^c(X,S)\big)&\le 2\exp\big(\lambda^2S\alpha_R+\lambda Sc^{-1}\beta_R+\lambda\gamma_{x_0,C,D,S}\big)\\ &\quad\times\big(1+8\lambda\eta_{D,R,S}\exp\big(\lambda^2\eta_{D,R,S}^2\big)\big),\end{aligned}$$

where $\eta_{D,R,S}=\delta_{D,S}R\sqrt{S/2}$. □

Remark 7. Let us notice that the condition that σ is bounded is essential for obtaining the Gaussian concentration of $\mathrm{TV}^c(X,S)$. To see this it is enough to consider the equation $dX(t)=2^{-1}X(t)\,\mathrm{d}t+X(t)\,\mathrm{d}W(t)$ with the starting condition $X(0)=1$. Notice that $\mathrm{TV}^c(X,S)\ge(X(S)-X(0)-c)_+$ and that $(X(S)-X(0)-c)_+=(\exp W(S)-1-c)_+$ does not reveal the Gaussian concentration.

Remark 8. Notice that for the standard Brownian motion $X = W$ and $Sc^{-2} \geq 2$, Sc^{-1} is comparable up to a universal constant with $\mathbf{E}\,\mathrm{TV}^c(W, S)$. Hence, from (3.3) we obtain that for $c > 0$ such that $Sc^{-2} \geq 2$, there exist universal constants $\bar{A}, \bar{B} < +\infty$ such that the Gaussian concentration holds

$$\mathbf{P}\big(\mathrm{TV}^c(W, S) \geq \bar{A}\mathbf{E}\,\mathrm{TV}^c(W, S) + \bar{B}\sqrt{S}u\big) \leq \exp\big(-u^2\big), \qquad \text{for } u \geq 0.$$

4. Existence of moment-generating functions of the truncated variation of Lévy processes

In this section, we will deal with the existence of finite exponential moments of the truncated variation of a Lévy process X. We will state the necessary and sufficient condition for the finiteness of $\mathbf{E}\exp(\alpha\,\mathrm{TV}^c(X, S))$ in terms of the generating triplet of the process X (cf. Sato [16], Chapter 2, Section 11). The methodology used here is very similar to the methodology used in Łochowski [10] for a Wiener process W, where the existence of $\mathbf{E}\exp(\alpha\,\mathrm{TV}^c(W, S))$ for any complex α was proved.

We start with the following lemma.

Lemma 8. *Let X be a Lévy process. For any $c > 0$ and $\alpha > 0$ one has $\mathbf{E}\exp(\alpha\,\mathrm{TV}^c(X, S)) < +\infty$ if and only if*

$$\mathbf{E}\exp\Big(\alpha \sup_{0\leq s\leq S}\big|X(s)\big|\Big) < +\infty.$$

Proof. The 'only if' part follows from the inequality

$$\begin{aligned}\mathrm{TV}^c(X, S) &\geq \sup_{0\leq s\leq S} \max\big\{\big|X(s) - X(0)\big| - c, 0\big\} \\ &= \max\Big\{\sup_{0\leq s\leq S}\big|X(s)\big| - c, 0\Big\} \geq \sup_{0\leq s\leq S}\big|X(s)\big| - c.\end{aligned}$$

To prove the opposite implication let us define $T_0^c = 0$ and for $i = 1, 2, \ldots$

$$T_i^c = \inf\big\{t > T_{i-1}^c : \big|X(t) - X\big(T_{i-1}^c\big)\big| > c/2\big\} \wedge \big(S + T_{i-1}^c\big).$$

Observe that $T_1^c = \inf\{t > 0 : |X(t)| > c/2\} \wedge S \leq S$ and that $(X(t))_{t\geq 0} \stackrel{d}{=} (X(t) - X(T_1^c))_{t\geq T_1^c}$, where "$\stackrel{d}{=}$" denotes the equality of distributions. Now let us define

$$X_t^c = \sum_{i=0}^{\infty} X\big(T_i^c\big) I_{[T_i^c, T_{i+1}^c)}(t).$$

Since $\|X^c - X\|_\infty \leq c/2$, we have

$$\mathrm{TV}^c(X, S) \leq \mathrm{TV}\big(X^c, S\big) \tag{4.1}$$

and since X^c is piecewise constant with the first jump at $T_1^c \le S$, denoting $\Delta X^c(T_1^c) = X^c(T_1^c) - X^c(T_1^c-)$ we have

$$\begin{aligned}\mathrm{TV}\big(X^c, S\big) &= \big|\Delta X^c\big(T_1^c\big)\big| + \mathrm{TV}\big(X^c, \big[T_1^c, S\big]\big) \\ &\le \sup_{0\le s\le T_1^c} \big|X(s)\big| + \mathrm{TV}\big(X^c, \big[T_1^c, S\big]\big).\end{aligned} \tag{4.2}$$

Let now $\delta \in (0; S)$ be such a small number that

$$\begin{aligned}&\mathbf{E}\Big[\exp\Big(\alpha \sup_{0\le s\le S} \big|X(s)\big|\Big); T_1^c \le \delta\Big] \\ &:= \mathbf{E}\Big[\exp\Big(\alpha \sup_{0\le s\le S} \big|X(s)\big|\Big) I_{\{T_1^c\le\delta\}}\Big] < 1.\end{aligned} \tag{4.3}$$

Note that such a number exists, since we assume that $\mathbf{E}\exp(\alpha \sup_{0\le s\le S} |X(s)|) < +\infty$ and from the càdlàg property and stochastic continuity of X it follows that $\mathbf{P}(T_1^c \le \delta) = \mathbf{P}(\sup_{0\le s\le\delta} |X(s)| > c/2) \downarrow 0$ as $\delta \downarrow 0$.

Let us fix $M > 0$. Note that on the set $\{T_1^c > \delta\}$ we have $\mathrm{TV}(X^c, \delta) = 0$, hence

$$\begin{aligned}\mathbf{E}\exp\big(\alpha\,\mathrm{TV}\big(X^c, \delta\big) \wedge M\big) &= \mathbf{E}\big[\exp\big(\alpha\,\mathrm{TV}\big(X^c, \delta\big) \wedge M\big); T_1^c \le \delta\big] \\ &\quad + \mathbf{E}\big[\exp(0 \wedge M); T_1^c > \delta\big] \\ &= \mathbf{E}\big[\exp\big(\alpha\,\mathrm{TV}\big(X^c, \delta\big) \wedge M\big); T_1^c \le \delta\big] \\ &\quad + \mathbf{P}\big(T_1^c > \delta\big).\end{aligned} \tag{4.4}$$

Now, applying (4.2), the independence of the process $X(t) - X(T_1^c), t \ge T_1^c$, and the two-dimensional r.v. $(\sup_{0\le s\le T_1^c} |X(s)|, T_1^c)$ (to see this notice that T_1^c is a stopping time and use the strong Markov property of Lévy processes) and the equality of distributions of $\mathrm{TV}(X^c, s)$ and $\mathrm{TV}(X^c, [T_1^c; T_1^c + s])$ for any $s \ge 0$, we have

$$\begin{aligned}&\mathbf{E}\big[\exp\big(\alpha\,\mathrm{TV}\big(X^c, \delta\big) \wedge M\big); T_1^c \le \delta\big] \\ &\quad \le \mathbf{E}\Big[\exp\Big(\alpha\Big(\sup_{0\le s\le T_1^c} \big|X(s)\big| + \mathrm{TV}\big(X^c; \big[T_1^c; \delta\big]\big)\Big) \wedge M\Big); T_1^c \le \delta\Big] \\ &\quad \le \mathbf{E}\Big[\exp\Big(\alpha \sup_{0\le s\le T_1^c} \big|X(s)\big| + \alpha\,\mathrm{TV}\big(X^c; \big[T_1^c; \delta + T_1^c\big]\big) \wedge M\Big); T_1^c \le \delta\Big] \\ &\quad = \mathbf{E}\Big[\exp\Big(\alpha \sup_{0\le s\le T_1^c} \big|X(s)\big|\Big); T_1^c \le \delta\Big]\mathbf{E}\exp\big(\alpha\,\mathrm{TV}\big(X^c, \delta\big) \wedge M\big) + \mathbf{P}\big(T_1^c > \delta\big) \\ &\quad \le \mathbf{E}\Big[\exp\Big(\alpha \sup_{0\le s\le S} \big|X(s)\big|\Big); T_1^c \le \delta\Big]\mathbf{E}\exp\big(\alpha\,\mathrm{TV}\big(X^c, \delta\big) \wedge M\big).\end{aligned}$$

By this and by (4.4), (4.3) we have

$$\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,\delta\bigr)\wedge M\bigr)\le\frac{\mathbf{P}(T_1^c>\delta)}{1-\mathbf{E}[\exp(\alpha\sup_{0\le s\le S}|X(s)|);\,T_1^c\le\delta]}.\tag{4.5}$$

Using similar arguments as before (i.e., (4.2), independence of $X(t)-X(T_1^c), t\ge T_1^c$, and $(\sup_{0\le s\le T_1^c}|X(s)|, T_1^c)$ and the equality of distributions of $\mathrm{TV}(X^c,s)$ and $\mathrm{TV}(X^c,[T_1^c;T_1^c+s])$ for $s\ge 0$) we obtain

$$\begin{aligned}
&\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S\bigr)\wedge M\bigr)\\
&\quad\le\mathbf{E}\Bigl[\exp\Bigl(\alpha\sup_{0\le s\le T_1^c}\bigl|X(s)\bigr|+\alpha\,\mathrm{TV}\bigl(X^c;\bigl[T_1^c;S+T_1^c\bigr]\bigr)\wedge M\Bigr);\,T_1^c\le\delta\Bigr]\\
&\qquad+\mathbf{E}\Bigl[\exp\Bigl(\alpha\sup_{0\le s\le T_1^c}\bigl|X(s)\bigr|+\alpha\,\mathrm{TV}\bigl(X^c;\bigl[T_1^c;S+T_1^c-\delta\bigr]\bigr)\wedge M\Bigr);\,T_1^c>\delta\Bigr]\\
&\quad=\mathbf{E}\Bigl[\exp\Bigl(\alpha\sup_{0\le s\le T_1^c}\bigl|X(s)\bigr|\Bigr);\,T_1^c\le\delta\Bigr]\mathbf{E}\bigl[\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S\bigr)\wedge M\bigr)\bigr]\\
&\qquad+\mathbf{E}\Bigl[\exp\Bigl(\alpha\sup_{0\le s\le T_1^c}\bigl|X(s)\bigr|\Bigr);\,T_1^c>\delta\Bigr]\mathbf{E}\bigl[\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S-\delta\bigr)\wedge M\bigr)\bigr]\\
&\quad\le\mathbf{E}\Bigl[\exp\Bigl(\alpha\sup_{0\le s\le S}\bigl|X(s)\bigr|\Bigr);\,T_1^c\le\delta\Bigr]\mathbf{E}\bigl[\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S\bigr)\wedge M\bigr)\bigr]\\
&\qquad+\mathbf{E}\exp\Bigl(\alpha\sup_{0\le s\le S}\bigl|X(s)\bigr|\Bigr)\mathbf{E}\bigl[\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S-\delta\bigr)\wedge M\bigr)\bigr].
\end{aligned}$$

From this, we have

$$\begin{aligned}
&\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S\bigr)\wedge M\bigr)\\
&\quad\le\frac{\mathbf{E}\exp(\alpha\sup_{0\le s\le S}|X(s)|)}{1-\mathbf{E}[\exp(\alpha\sup_{0\le s\le S}|X(s)|);\,T_1^c\le\delta]}\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S-\delta\bigr)\wedge M\bigr).
\end{aligned}$$

Similarly, if $S-2\delta>0$

$$\begin{aligned}
&\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S-\delta\bigr)\wedge M\bigr)\\
&\quad\le\frac{\mathbf{E}\exp(\alpha\sup_{0\le s\le S-\delta}|X(s)|)}{1-\mathbf{E}[\exp(\alpha\sup_{0\le s\le S-\delta}|X(s)|);\,T_1^c\le\delta]}\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S-2\delta\bigr)\wedge M\bigr)\\
&\quad\le\frac{\mathbf{E}\exp(\alpha\sup_{0\le s\le S}|X(s)|)}{1-\mathbf{E}[\exp(\alpha\sup_{0\le s\le S}|X(s)|);\,T_1^c\le\delta]}\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S-2\delta\bigr)\wedge M\bigr).
\end{aligned}$$

Iterating and putting together the above inequalities, we finally obtain

$$\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,S\bigr)\wedge M\bigr) \le \left(\frac{\mathbf{E}\exp(\alpha\sup_{0\le s\le S}|X(s)|)}{1-\mathbf{E}[\exp(\alpha\sup_{0\le s\le S}|X(s)|);\,T_1^c\le\delta]}\right)^{\lfloor S/\delta\rfloor} \times \mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}\bigl(X^c,\delta\bigr)\wedge M\bigr). \tag{4.6}$$

By (4.5) and (4.6), and letting $M\to\infty$ we get $\mathbf{E}\exp(\alpha\,\mathrm{TV}(X^c,S))<+\infty$. Finally, from (4.1) we get

$$\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}^c(X,S)\bigr)<+\infty. \qquad \square$$

Now let (A,ν,γ) be the generating triplet of the process X. By Sato [16], Theorem 28.15, we have

$$\mathbf{E}\exp\Bigl(\alpha\sup_{0\le s\le S}\bigl|X(s)\bigr|\Bigr)<+\infty$$

if and only if

$$\mathbf{E}\exp\bigl(\alpha\bigl|X(1)\bigr|\bigr)<+\infty$$

which, by Sato [16], Corollary 25.8, is equivalent with

$$\int_{|x|>1}\mathrm{e}^{\alpha|x|}\nu(\mathrm{d}x)<+\infty. \tag{4.7}$$

From equivalence of these conditions and Lemma 8 we obtain the following theorem.

Theorem 3. *Let (A,ν,γ) be the generating triplet of the Lévy process X. For any $\alpha>0$ we have*

$$\mathbf{E}\exp\bigl(\alpha\,\mathrm{TV}^c(X,S)\bigr)<+\infty$$

if and only if

$$\int_{|x|>1}\mathrm{e}^{\alpha|x|}\nu(\mathrm{d}x)<+\infty.$$

Theorem 3 may be applied in situations, when the process X satisfies condition (4.7) with some $\alpha>0$ but it is neither Brownian motion nor finite variation process. This holds, for example, for tempered stable process, that is, processes with the Lévy measure given by

$$\nu(\mathrm{d}x)=\frac{c_p}{x^{1+\alpha_p}}\mathrm{e}^{-\lambda_p x}1_{x>0}\,\mathrm{d}x+\frac{c_n}{(-x)^{1+\alpha_n}}\mathrm{e}^{\lambda_n x}1_{x<0}\,\mathrm{d}x,$$

where $\alpha_p,\alpha_n<2$, $\lambda_p,\lambda_n,c_p,c_n>0$. They satisfy (4.7) for any $\alpha<\min(\lambda_p,\lambda_n)$ and have infinite variation when $\alpha_p,\alpha_n\ge 1$. Another example are Meixner processes, used in financial modeling (cf. Kyprianou et al. [7], Chapter I), with Lévy measure given by

$$\nu(\mathrm{d}x)=\delta\frac{\exp(\beta x/\eta)}{x\sinh(\pi x/\eta)}\,\mathrm{d}x,$$

where $\delta, \eta > 0$, $|\beta| < \pi$. They satisfy (4.7) for $\alpha < (\pi - |\beta|)/\eta$.

Acknowledgements

The authors would like to thank the anonymous referee and the Associate Editor for many valuable comments and suggestions, which considerably improved the presentation of the results. The research of the first and the second author was supported by the National Science Center in Poland under decision no. DEC-2011/01/B/ST1/05089.

References

[1] Bednorz, W. (2006). A theorem on majorizing measures. *Ann. Probab.* **34** 1771–1781. MR2271481
[2] Bednorz, W. (2006). On a Sobolev type inequality and its applications. *Studia Math.* **176** 113–137. MR2264359
[3] Bednorz, W. (2010). Majorizing measures on metric spaces. *C. R. Math. Acad. Sci. Paris* **348** 75–78. MR2586748
[4] Fernique, X. (1978). Caractérisation de processus à trajectoires majorées ou continues. In *Séminaire de Probabilités XII* (*Univ. Strasbourg*, *Strasbourg*, 1976/1977). *Lecture Notes in Math.* **649** 691–706. Berlin: Springer. MR0520032
[5] Fernique, X. (1983). Regularité de fonctions aléatoires non gaussiennes. In *Eleventh Saint Flour Probability Summer School* – 1981 (*Saint Flour*, 1981). *Lecture Notes in Math.* **976** 1–74. Berlin: Springer. MR0722982
[6] Kwapień, S. and Rosiński, J. (2004). Sample Hölder continuity of stochastic processes and majorizing measures. In *Seminar on Stochastic Analysis*, *Random Fields and Applications IV*. *Progress in Probability* **58** 155–163. Basel: Birkhäuser. MR2096287
[7] Kyprianou, A. E., Schoutens, W. and Willmott, P. (2005). *Exotic Option Pricing and Advanced Lévy Models*. Chichester: Wiley. MR2343205
[8] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. *Isoperimetry and Processes*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Berlin: Springer. MR1102015
[9] Li, W.V. and Shao, Q.M. (2001). Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes*: *Theory and Methods* 533–597. *Handbook of Statistics* **19**. Amsterdam: North-Holland. MR1861734
[10] Łochowski, R. (2011). Truncated variation, upward truncated variation and downward truncated variation of Brownian motion with drift – their characteristics and applications. *Stochastic Process. Appl.* **121** 378–393. MR2746181
[11] Łochowski, R. (2013). On a generalisation of the Hahn–Jordan decomposition for real càdlàg functions. *Colloq. Math.* **132** 121–138. MR3106093
[12] Łochowski, R. and Miłoś, P. (2013). On truncated variation, upward truncated variation and downward truncated variation for diffusions. *Stochastic Process. Appl.* **123** 446–474. MR3003359
[13] Łochowski, R.M. (2014). Pathwise stochastic integration with respect to semimartingales. *Probab. Math. Statist*. To appear.
[14] Picard, J. (2008). A tree approach to p-variation and to integration. *Ann. Probab.* **36** 2235–2279. MR2478682

[15] Revuz, D. and Yor, M. (2005). *Continuous Martingales and Brownian Motion*. *Grundlehren der Mathematischen Wissenschaften* **293** [*Fundamental Principles of Mathematical Sciences*]. Berlin: Springer. MR1725357

[16] Sato, K.-i. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR1739520

[17] Talagrand, M. (1990). Sample boundedness of stochastic processes under increment conditions. *Ann. Probab.* **18** 1–49. MR1043935

[18] Talagrand, M. (2005). *The Generic Chaining*. *Upper and Lower Bounds of Stochastic Processes*. *Springer Monographs in Mathematics*. Berlin: Springer. MR2133757

Received November 2012 and revised September 2013