Electronic Journal of Probability

Vol. 16 (2011), Paper no. 68, pages 1880–1899.

Journal URL
http://www.math.washington.edu/~ejpecp/

A note on higher dimensional p-variation *

Peter Friz Nicolas Victoir

Abstract

We discuss p-variation regularity of real-valued functions defined on $[0, T]^2$, based on rectangular increments. When $p > 1$, there are two slightly different notions of p-variation; both of which are useful in the context of Gaussian roug paths. Unfortunately, these concepts were blurred in previous works [2, 3]; the purpose of this note is to show that the afore-mentioned notions of p-variations are "ε-close". In particular, all arguments relevant for Gaussian rough paths go through with minor notational changes.

Key words: higher dimensional p-variation, Gaussian rough paths.

AMS 2010 Subject Classification: Primary 60H99.

Submitted to EJP on March 4, 2011, final version accepted July 6, 2011.

*The first author has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement nr. 258237

1 Higher-dimensional p-variation

Let $T > 0$ and $\Delta_T = \{(s,t) : 0 \le s \le t \le T\}$ .We shall regard $((a,b),(c,d)) \in \Delta_T \times \Delta_T$ as (closed) **rectangle** $A \subset [0,T]^2$;

$$A := \begin{pmatrix} a,b \\ c,d \end{pmatrix} := [a,b] \times [c,d];$$

if $a = b$ or $c = d$ we call A degenerate. Two rectangles are called **essentially disjoint** if their intersection is empty or degenerate. A **partition** Π of a rectangle $R \subset [0,T]^2$ is then a a finite set of essentially disjoint rectangles, whose union is R; the family of all such partitions is denoted by $\mathscr{P}(R)$. Recall that **rectangular increments** of a function $f : [0,T]^2 \to \mathbb{R}$ are defined in terms of f evaluated at the four corner points of A,

$$f(A) := f\begin{pmatrix} a,b \\ c,d \end{pmatrix} := f\begin{pmatrix} b \\ d \end{pmatrix} - f\begin{pmatrix} a \\ d \end{pmatrix} - f\begin{pmatrix} b \\ c \end{pmatrix} + f\begin{pmatrix} a \\ c \end{pmatrix}.$$

Let us also say that a **dissection** D of an interval $[a,b] \subset [0,T]$ is of the form $D = \left(a = t_0 \le t_1 \le \cdots \le t_n = b\right)$; we write $\mathscr{D}([a,b])$ for the family of all such dissections.

Definition 1. *Let $p \in [1,\infty)$. A function $f : [0,T]^2 \to \mathbb{R}$ has finite p-**variation** if*

$$V_p\left(f;[s,t]\times[u,v]\right) := \left(\sup_{\substack{D=(t_i)\in\mathscr{D}([s,t]) \\ D'=(t'_j)\in\mathscr{D}([u,v])}} \sum_{i,j} \left| f\begin{pmatrix} t_i, t_{i+1} \\ t'_j, t'_{j+1} \end{pmatrix} \right|^p \right)^{\frac{1}{p}} < \infty;$$

*it has finite **controlled** p-**variation**[1] if*

$$\left|f\right|_{p\text{-var};[s,t]\times[u,v]} := \sup_{\Pi\in\mathscr{P}([s,t]\times[u,v])} \left(\sum_{A\in\Pi} \left|f(A)\right|^p \right)^{1/p} < \infty.$$

The difference is that in the first definition (i.e. of V_p) the sup is taken over **grid-like partitions**,

$$\left\{ \begin{pmatrix} t_i, t_{i+1} \\ t'_j, t'_{j+1} \end{pmatrix} : 1 \le i \le n, 1 \le j \le m \right\},$$

based on D, D' where $D = \left(t_i : 1 \le i \le n\right) \in \mathscr{D}([s,t])$ and $D' = \left(t'_j : 1 \le j \le m\right) \in \mathscr{D}([u,v])$. Clearly, not every partition is grid-like (consider e.g. $[0,2]^2 = [0,1]^2 \cup [1,2] \times [0,1] \cup [0,2] \times [1,2]$) hence

$$V_p(f;R) \le \left|f\right|_{p\text{-var};R}.$$

for every rectangle $R \subset [0,T]^2$.

[1] Our main theorem below will justify this terminology.

Definition 2. *A map $\omega : \Delta_T \times \Delta_T \to [0, \infty)$ is called **2D control** if it is continuous, zero on degenerate rectangles, and **super-additive** in the sense that, for all rectangles $R \subset [0,T]$,*

$$\sum_{i=1}^{n} \omega(R_i) \leq \omega(R), \text{ whenever } \{R_i : 1 \leq i \leq n\} \in \mathscr{P}(R).$$

Our result is

Theorem 1. *(i) For any function $f : [0,T]^2 \to \mathbb{R}$ and any rectangle $R \subset [0,T]$,*

$$\left|f\right|_{1\text{-var};R} = V_1(f;R). \tag{1.1}$$

(ii) Let $p \in [1,\infty)$ and $\varepsilon > 0$. There exists a constant $c = c(p,\varepsilon) \geq 1$ such that, for any function $f : [0,T]^2 \to \mathbb{R}$ and any rectangle $R \subset [0,T]$,

$$\frac{1}{c(p,\varepsilon)} \left|f\right|_{(p+\varepsilon)\text{-var};R} \leq V_p(f;R) \leq \left|f\right|_{p\text{-var};R}. \tag{1.2}$$

(iii) If $f : [0,T]^2 \to \mathbb{R}$ is of finite controlled p-variation, then $R \mapsto \left|f\right|^p_{p\text{-var};R}$ is super-additive.
(iv) If $f : [0,T]^2 \to \mathbb{R}$ is continuous and of finite controlled p-variation, then $R \mapsto \left|f\right|^p_{p\text{-var};R}$ is a 2D control. Thus, in particular, there exists a 2D control ω such that

$$\forall \text{ rectangles } R \subset [0,T] : \left|f(R)\right|^p \leq \omega(R)$$

As will be seen explicitly in the following example, there exist functions f which are of finite p-variation but of infinite controlled p-variation; that is,

$$V_p\left(f;[0,T]^2\right) < \left|f\right|_{p\text{-var};[0,T]^2} = +\infty$$

which also shows that one cannot take $\varepsilon = 0$ in (1.2). In the same example we see that p-variation $R \mapsto V_p(f;R)^p$ can fail to be super-additive[2].

Example 1 (Finite $(1/2H)$-variation of fBM covariance, $H \in (0,1/2]$.). *Let β^H denote fractional Brownian motion with Hurst parameter H; its covariance is given by*

$$C^H(s,t) := \mathbb{E}\left(\beta^H_s \beta^H_t\right) := \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \quad s,t \in [0,T]^2, H \in (0,1/2].$$

We show that C^H has finite $1/(2H)$-variation in 2D sense[3] *and more precisely,*

$$V_{1/(2H)}\left(C^H;[s,t]^2\right) \leq c_H |t-s|^{2H}, \quad \text{for every } s \leq t \text{ in } [0,T].$$

[2]... in contrast to controlled p-variation $R \mapsto \left|f\right|^p_{p\text{-var};R}$ which yields a 2D control, cf part (iv) of the theorem.
[3]This is a minor modification of the argument in [3] where it was assumed that $D = D'$.

(By fractional scaling it would suffice to consider $[s,t]=[0,1]$ but this does not simplify the argument which follows). Consider $D=(t_i), D'=\left(t'_j\right)\in\mathscr{D}[s,t]$. Clearly[4]*,*

$$
\begin{aligned}
3^{1-\frac{1}{2H}}\sum_j\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{t'_j,t'_{j+1}}\right]\right|^{\frac{1}{2H}} &\le 3^{1-\frac{1}{2H}}\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{\cdot}\right]\right|^{\frac{1}{2H}}_{\frac{1}{2H}\text{-}var;[s,t]} \\
&\le \left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{\cdot}\right]\right|^{\frac{1}{2H}}_{\frac{1}{2H}\text{-}var;[s,t_i]} && (1.3)\\
&\quad+\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{\cdot}\right]\right|^{\frac{1}{2H}}_{\frac{1}{2H}\text{-}var;[t_i,t_{i+1}]} && (1.4)\\
&\quad+\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{\cdot}\right]\right|^{\frac{1}{2H}}_{\frac{1}{2H}\text{-}var;[t_{i+1},t]}, && (1.5)
\end{aligned}
$$

by super-additivity of (1D!) controls. The middle term (1.4) is estimated by

$$
\begin{aligned}
\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{\cdot}\right]\right|^{\frac{1}{2H}}_{\frac{1}{2H}\text{-}var;[t_i,t_{i+1}]} &= \sup_{(s_k)\in\mathscr{D}[t_i,t_{i+1}]}\sum_k\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{s_k,s_{k+1}}\right]\right|^{\frac{1}{2H}}\\
&\le c_H\left|t_{i+1}-t_i\right|,
\end{aligned}
$$

where we used that $[s_k,s_{k+1}]\subset[t_i,t_{i+1}]$ implies $\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{s_k,s_{k+1}}\right]\right|\le c_H\left|s_{k+1}-s_k\right|^{2H}$. The first term (1.3) and the last term (1.5) are estimated by exploiting the fact that disjoint increments of fractional Brownian motion have negative correlation when $H<1/2$ (resp. zero correlation in the Brownian case, $H=1/2$); that is, $E\left(\beta^H_{c,d}\beta^H_{a,b}\right)\le 0$ whenever $a\le b\le c\le d$. We can thus estimate (1.3) as follows;

$$
\begin{aligned}
\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{\cdot}\right]\right|^{\frac{1}{2H}}_{\frac{1}{2H}\text{-}var;[s,t_i]} &= \left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{s,t_i}\right]\right|^{\frac{1}{2H}}\\
&\le 2^{\frac{1}{2H}-1}\left(\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{s,t_i}\right]\right|^{\frac{1}{2H}}+E\left[\left|\beta^H_{t_i,t_{i+1}}\right|^2\right]^{\frac{1}{2H}}\right).
\end{aligned}
$$

The covariance of fractional Brownian motion gives immediately $E\left[\left|\beta^H_{t_i,t_{i+1}}\right|^2\right]^{\frac{1}{2H}}=c_H\left(t_{i+1}-t_i\right)$. On the other hand, $[t_i,t_{i+1}]\subset[s,t_{i+1}]$ implies $\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{s,t_i}\right]\right|^{\frac{1}{2H}}\le c_H\left|t_{i+1}-t_i\right|$; hence

$$
\left|E\left[\beta^H_{t_i,t_{i+1}}\beta^H_{\cdot}\right]\right|^{\frac{1}{2H}}_{\frac{1}{2H}\text{-}var;[s,t_i]}\le c_H\left|t_{i+1}-t_i\right|.
$$

As already remarked, the last term is estimated similarly. It only remains to sum up and to take the supremum over all dissections D and D'.

Example 2 (Failure of super-addivity of $(1/2H)$-variation, infinite controlled $(1/2H)$-variation of fBM covariance, $H\in(0,1/2)$.)**.** *We saw above that*

$$
V_{1/(2H)}\left(C^H;[0,T]^2\right)<\infty.
$$

[4]We write $\beta^H_{a,b}\equiv\beta^H_b-\beta^H_a$.

When $H = 1/2$ we deal with Brownian motion and see that its covariance has finite 1-variation, which, by (i),(iv) of Theorem 1, constitues a 2D control for $C^{1/2}$. In contrast, we claim that, for $H < 1/2$, there does not exist a 2D control for the $1/(2H)$-variation of C^H. In fact, the sheer existence of a super-additive map ω (in the sense of definition 2) such that

$$\forall \text{ rectangles } R \subset [0,T] : \left|C^H(R)\right|^{1/(2H)} \leq \omega(R)$$

leads to a contradiction as follows: assume that such a ω exists. By super-addivity,

$$\bar{\omega}(R) := \left|C^H\right|^{1/(2H)}_{1/(2H)\text{-}var;R} \leq \omega(R) < \infty$$

and $\bar{\omega}$ is super-additive (in fact, a 2D control) thanks to part (iv) of the theorem. On the other hand, by fractional scaling there exists C such that

$$\forall (s,t) \in \Delta_T : \bar{\omega}\left([s,t]^2\right) = C|t-s|.$$

Let us consider the case $T = 2$ and the partition

$$[0,2]^2 = [0,1]^2 \cup [1,2]^2 \cup R \cup R'$$

with $R = [0,1] \times [1,2]$, $R' = [1,2] \times [0,1]$. Super-addivitiy of $\bar{\omega}$ gives

$$\begin{aligned} \bar{\omega}\left([0,1]^2\right) + \bar{\omega}\left([1,2]^2\right) + \bar{\omega}(R) + \bar{\omega}\left(R'\right) &\leq \bar{\omega}\left([0,2]^2\right), \\ C(1-0) + C(2-1) + \bar{\omega}(R) + \bar{\omega}\left(R'\right) &\leq 2C, \end{aligned}$$

hence $\bar{\omega}(R) = \bar{\omega}\left(R'\right) = 0$, and thus also

$$C^H(R) = \mathbb{E}\left[\left(B^H_1 - B^H_0\right)\left(B^H_2 - B^H_1\right)\right] = 0;$$

which is false for $H \neq 1/2$ and hence the desired contradiction. En passant, we see that we must have

$$\left|C^H\right|_{1/(2H)\text{-}var;[0,T]^2} = +\infty;$$

for otherwise part (iv) of Theorem 1 would yield a 2D control for the $1/(2H)$-variation of C^H. This also shows that, with $f = C^H$ and $p = 1/(2H)$ one has

$$V_p\left(f;[0,T]^2\right) < |f|_{p\text{-}var;[0,T]^2} = +\infty.$$

Remark 1. *The previous examples clearly show the need for Theorem 1; variational regularity of C^H can be controlled upon considering $[(1/2H)+\varepsilon]$-variation rather than $1/(2H)$-variation. In applications, this distinction never matters. Existence for Gaussian rough paths for instance, requires $1/(2H) < 2$ and one can always insert a small enough ε. It should also be point out that, by fractional scaling,*

$$\left|C^H\right|_{[1/(2H)+\varepsilon]\text{-}var;[s,t]^2} \propto |t-s|^{2H};$$

hence, even in estimates that involve directly that variational regularity of C^H, no ε loss is felt.

Remark 2. *The previous examples dealt with $H \leq 1/2$ and reader may wonder about the case $H > 1/2$. In this case $1/(2H) < 1$ and clearly the (non-trivial) covariance function of fBM with Hurst parameter H will not be of finite $1/(2H)$-variation. Indeed, any continuous function $f : [0,T]^2 \to \mathbb{R}$, with $f(0,\cdot) \equiv f(\cdot,0) \equiv 0$, and finite p-variation for $p \in (0,1)$, is necessarily constant (and then equal to zero).*

Acknowledgement 1. *The authors are indebted to Bruce Driver for pointing out, in the most constructive and gentle way, that $R \mapsto V_p(f;R)^p$ is not, in general, super-additive. P. Friz received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement nr. 258237.*

2 Proof of (i)

We claim the *controlled* 1-variation is exactly equal to its 1-variation. More precisely, for all rectangles $R \subset [0,T]^2$ we have

$$\left|f\right|_{1\text{-var};R} = V_1\left(f;R\right).$$

Proof. Trivially $V_1\left(f;R\right) \leq \left|f\right|_{1\text{-var};R}$. For the other inequality, assume Π is a partition of R. It is obvious that one can find a grid-like partition $\tilde{\Pi}$, based on $D \times D'$, for sufficiently fine dissections D, D', which **refines** Π in the sense that every $A \in \Pi$ can be expressed as

$$A = \cup_i A_i \text{ (essentially disjoint)}, A_i \in \tilde{\Pi}.$$

From the very definition of rectangular increments, we have $f\left(A\right) = \sum_i f\left(A_i\right)$ and it follows that $\left|f\left(A\right)\right| \leq \sum_i \left|f\left(A_i\right)\right|$. (If $|\cdot|$ is replaced by $|\cdot|^p, p > 1$, this estimate is false.[5]) Hence

$$\sum_{A\in\Pi}\left|f\left(A\right)\right| \leq \sum_{A\in\tilde{\Pi}}\left|f\left(A\right)\right| \leq \left|f\right|_{1\text{-var};R}.$$

It now suffices to take the supremum over all such Π to see that $\left|f\right|_{1\text{-var};R} \leq V_1\left(f;R\right)$. □

3 Proof of (ii)

The second inequality $V_p\left(f;R\right) \leq \left|f\right|_{p\text{-var};R}$ is trivial. Furthermore, if $V_p\left(f;R\right) = +\infty$ there is nothing to show so we may assume $V_p\left(f;R\right) < +\infty$. We claim that, for all rectangle $R \subset [0,T]^2$,

$$\left|f\right|_{p+\varepsilon\text{-var};R} \leq c\left(p,\varepsilon\right)V_p\left(f;R\right).$$

For the proof we note first that there is no loss in generality in taking $R = [0,T]^2$; an affine reparametrization of each axis will transform R into $[0,T]^2$, while leaving all rectangular increments invariant. The plan is to show, for an arbitrary partition $\left(Q_k\right) \in \mathscr{P}\left([0,T]^2\right)$, the estimate

$$\left(\sum_k\left|f\left(Q_k\right)\right|^{p+\varepsilon}\right)^{\frac{1}{p+\varepsilon}} \leq c\left(p,\varepsilon\right)V_p\left(f;[0,T]^2\right).$$

where c depends only on p, ε for any partition $\left(Q_k\right) \in \mathscr{P}\left([0,T]^2\right)$. The key observation is that for a suitable choice of $y, x, D = \left(t_i\right), D' = \left(t'_j\right)$ we have

$$\begin{aligned}\sum_k\left|f\left(Q_k\right)\right|^{p+\varepsilon} &= \sum_k\left|f\left(Q_k\right)\right|^{p+\varepsilon-1} sgn\left(f\left(Q_k\right)\right) f\left(Q_k\right) \\ &= \sum_i\sum_j y\begin{pmatrix} t_i \\ t'_j\end{pmatrix} x\begin{pmatrix} t_{i-1}, t_i \\ t'_{j-1}, t'_j\end{pmatrix} \\ &=: \int_{D\times D'} y\, dx.\end{aligned} \tag{3.1}$$

[5]One has $\left|\sum_{i=1}^m a_i\right|^p \leq \left|\sum_{i=1}^m |a_i|\right|^p \leq m^{p-1}\left(\sum_{i=1}^m |a_i|^p\right)$ and this is sharp as seen by taking $a_i \equiv 1$.

Indeed, we may take (as in the proof of part (i)) sufficiently fine dissections $D=\left(t_i\right), D'=\left(t'_j\right)\in\mathscr{D}[0,T]$ such that the grid-like partition based on $D\times D'$ refines $\left(Q_k\right)$; followed by setting[6]

$$\begin{aligned} x &:= f \\ y &:= \sum_k \left|f\left(Q_k\right)\right|^{p-1+\varepsilon} sgn\left(f\left(Q_k\right)\right)\mathbb{I}_{\hat{Q}_k} \end{aligned}$$

where $\hat{Q}_k$ is the of the form $(a,b]\times(c,d]$ whenever $Q_k=[a,b]\times[c,d]$. Lemma 1 below, applied with $p+\varepsilon$ instead of p, says

$$V_q\left(y;[0,T]^2\right)\le 4\left|\sum_k\left|x\left(Q_k\right)\right|^{p+\varepsilon}\right|^{\frac{1}{q}}$$

where $q:=1/\left(1-1/\left(p+\varepsilon\right)\right)$ denotes the Hölder conjugate of $p+\varepsilon$. Since

$$\frac{1}{p}+\frac{1}{q}=1+\left(\frac{1}{p}-\frac{1}{p+\varepsilon}\right)>1,$$

noting also that $y(0,\cdot)=y(\cdot,0)=0$, we can use **Young-Towghi's maximal inequality** [4, Thm 2.1.], included for the reader's convenience as Theorem 3 in the appendix, to obtain the estimate

$$\begin{aligned} \sum_k\left|f\left(Q_k\right)\right|^{p+\varepsilon} &\le c\left(p,\varepsilon\right)V_q\left(y;[0,T]^2\right)V_p\left(x;[0,T]^2\right) \\ &\le 4c\left(p,\varepsilon\right)\left|\sum_k\left|x\left(Q_k\right)\right|^{p+\varepsilon}\right|^{\frac{1}{q}}V_p\left(x;[0,T]^2\right) \end{aligned}$$

Since $1-\frac{1}{q}=\frac{1}{p+\varepsilon}$ and $x=f$ we see that

$$\left(\sum_k\left|f\left(Q_k\right)\right|^{p+\varepsilon}\right)^{\frac{1}{p+\varepsilon}}\le 4c\left(p,\varepsilon\right)V_p\left(f;[0,T]^2\right)$$

and conclude by taking the supremum over all partitions $\left(Q_k\right)\in\mathscr{P}\left([0,T]^2\right)$.

Lemma 1. *Fix $p\ge 1$ and write p' for the Hölder conjugate i.e. $1/p'+1/p=1$. Let $\left(Q_j\right)\in\mathscr{P}\left([0,T]^2\right)$ and $y=\sum_j\left|x\left(Q_j\right)\right|^{p-1}sgn\left(x\left(Q_j\right)\right)\mathbb{I}_{\hat{Q}_j}$. Then*

$$V_{p'}\left(y,[0,T]^2\right)\le|y|_{p'\text{-}var;[0,T]^2}\le 4\left(\sum_i\left|x\left(Q_i\right)\right|^p\right)^{1/p'}.$$

[6] The "right-closed" form of $\hat{Q}_k$ in the definition of y is tied to our definition of $\int_{D\times D'}y\,dx$ which imposes "right-end-point-evaluation" of y. Recall also that Q_k is *really* a point in $((a,b),(c,d))\in\Delta_T\times\Delta_T$; viewing it as closed rectangle is pure convention.

Proof. Only the second inequality requires a proof. By definition, $\left(Q_j\right)$ forms a partition of $[0,T]^2$ into essentially disjoint rectangles and we note that $y(.,0) = y(0,.) = 0$. Consider now another partition $(R_i) \in \mathscr{P}\left([0,T]^2\right)$. The rectangular increments of y over R_i spells out as "$+--+$ sum" of y evaluated at the corner points of R_i. Recall that on each set $\hat{Q}_j$ the function y takes the consant value

$$c_j := \left|x\left(Q_j\right)\right|^{p-1} sgn\left(x\left(Q_j\right)\right).$$

Since the corner points of R_i are elements of $Q_{j_1} \cup Q_{j_2} \cup Q_{j_3} \cup Q_{j_4}$ for suitable (not necessarily distinct) indices $j_1, \ldots, j_4$ we clearly have the (crude) estimate

$$\left|y\left(R_i\right)\right| \leq \sum_{j \in \{j_1,j_2,j_3,j_4\}} \left|c_j\right| \tag{3.2}$$

and, trivially, any $j \notin \{j_1, j_2, j_3, j_4\}$ is not required in estimating $\left|y\left(R_i\right)\right|$. Let us distinguish a few cases where we can do better than in 3.2.

Case 1: There exists j such that all four corner points of R_i are elements of Q_j (equivalently: $\exists j : R_i \subset \hat{Q}_j$). In this case

$$y\left(R_i\right) = c_j - c_j - c_j + c_j = 0.$$

In particular, such an index j is not required to estimate $\left|y\left(R_i\right)\right|$.

Case 2: There exists j such that precisely two corner points[7] of R_i are elements of Q_j. It follows that the corner points of R_i are elements of $Q_{j_1} \cup Q_{j_2} \cup Q_j$ for suitable (not necessarily distinct) indices j_1, j_2. Note however that $j \notin \{j_1, j_2\}$. In this case

$$y\left(R_i\right) = c_{j_1} - c_{j_2} - c_j + c_j = c_{j_1} - c_{j_2}.$$

In general, this quantity is non-zero (although it is zero when $j_1 = j_2$, which is tantamount to say that $R_i \subset Q_{j_1} \cup Q_j$). Even so, we note that

$$\left|y\left(R_i\right)\right| \leq \left|c_{j_1}\right| + \left|c_{j_2}\right|$$

and again the index j is not required in order to estimate $\left|y\left(R_i\right)\right|$.

Case 3: There exists j such that precisely one corner point of R_i is an element of Q_j. In this case, for suitable (not necessarily distinct) indices j_1, j_2, j_3 with $j \notin \{j_1, j_2, j_3\}$

$$\left|y\left(R_i\right)\right| = \left|c_{j_1} - c_{j_2} - c_{j_3} + c_j\right| \leq \left|c_{j_1} - c_{j_2} - c_{j_3}\right| + \left|c_j\right|.$$

In this case, the index j is required to estimate $\left|y\left(R_i\right)\right|$. (There is still the possibily for cancellation between the other terms. If $j_2 = j_3$ for instance, then $\left|y\left(R_i\right)\right| \leq \left|c_{j_1}\right| + \left|c_j\right|$ and indices j_2, j_3 are not required; this corresponds precisely to case 2 applied to Q_{j_2}. Another possiblility is that $\{j_1, j_2, j_3\}$ are all distinct in which case $\left|y\left(R_i\right)\right| \leq \left|c_{j_1}\right| + \left|c_{j_2}\right| + \left|c_{j_3}\right| + \left|c_j\right|$ is the best estimate and all four indices j_1, j_2, j_3, j are needed in the estimate.

The moral of this case-by-case consideration is that only those $j \in \phi(i)$ where

$$\phi(i) := \left\{j : \text{ precisely one corner point of } R_i \text{ is an element of } Q_j\right\}$$

[7] The case that three corner points of R_i are elements of Q_j already implies (rectangles!) that all four corner points of R_i are elements of Q_j. This is covered by Case 1.

are required in estimating $\left|y\left(R_i\right)\right|$; more precisely,

$$\left|y\left(R_i\right)\right| \leq \sum_{j \in \phi(i)} \left|c_j\right|.$$

Since rectangles (here: R_i) have four corner points it is clear that $\#\phi\left(i\right) \leq 4$ where $\#$ denotes the cardinality of a set. Hence

$$\left|y\left(R_i\right)\right|^{p'} \leq 4^{p'-1} \sum_{j \in \phi(i)} \left|c_j\right|^{p'} \equiv 4^{p'-1} \sum_j \phi_{i,j} \left|c_j\right|^{p'}$$

where we introdudced the matrix $\phi_{i,j}$ with value 1 if $j \in \phi\left(i\right)$ and zero else. This allows us to write

$$\begin{aligned} \sum_i \left|y\left(R_i\right)\right|^{p'} &\leq 4^{p'-1} \sum_i \sum_j \phi_{i,j} \left|c_j\right|^{p'} \\ &= 4^{p'-1} \sum_j \left|c_j\right|^{p'} \sum_i \phi_{i,j}. \end{aligned}$$

Consider now, for fixed j, the number of rectangles R_i which have precisely one corner point inside Q_j. Obviously, there can be a most 4 rectangles with this property. Hence

$$\sum_i \phi_{i,j} = \#\left\{i : j \in \phi\left(i\right)\right\} \leq 4.$$

It follows that

$$\sum_i \left|y\left(R_i\right)\right|^{p'} \leq 4^{p'} \sum_j \left|c_j\right|^{p'} = 4^{p'} \sum_j \left|x\left(Q_j\right)\right|^{(p-1)p'} = 4^{p'} \sum_j \left|x\left(Q_j\right)\right|^{p},$$

where we used that $\left(p-1\right)p' = p$. Since $\left(R_i\right)$ was an arbitrary partition of $[0,T]^2$ we obtain

$$\left|y\right|^{p'}_{p'\text{-var};[0,T]^2} \leq 4^{p'} \sum_i \left|x\left(Q_i\right)\right|^{p},$$

as desired. The proof is finished. □

4 Proof of (iii)

The claim is super-additivity of

$$R \mapsto \sup_{\Pi \in \mathscr{P}(R)} \sum_{A \in \Pi} \left|f\left(A\right)\right|^p.$$

Assume $\left\{R_i : 1 \leq i \leq n\right\}$ constitutes a partition of R. Assume also that Π_i is a partition of R_i for every $1 \leq i \leq n$. Clearly, $\Pi := \cup_{i=1}^n \Pi_i$ is a partition of R and hence

$$\sum_{i=1}^n \sum_{A \in \Pi_i} \left|f\left(A\right)\right|^p = \sum_{A \in \Pi} \left|f\left(A\right)\right|^p \leq \omega\left(R\right)$$

Now taking the supremum over each of the Π_i gives the desired result.

5 Proof of (iv)

The assumption is that $f:[0,T]^2\to\mathbb{R}$ is continuous and of finite controlled p-variation. From (iii),

$$\omega(R):=\left|f\right|_{p\text{-var};R}^{p}$$

is super-additive as function of R. It is also clear that ω is zero on degenerate rectangles. It remains to be seen that $\omega:\Delta_T\times\Delta_T\to[0,\infty)$ is continuous.

Lemma 2. *Consider the two (adjacent) rectangles $[a,b]\times[s,t]$ and $[a,b]\times[t,u]$ in $[0,T]^2$.Then,*

$$\begin{aligned}\omega\begin{pmatrix}a,b\\ s,u\end{pmatrix} &\leq \omega\begin{pmatrix}a,b\\ s,t\end{pmatrix}+\omega\begin{pmatrix}a,b\\ t,u\end{pmatrix}\\ &\quad+p2^{p-1}\omega\begin{pmatrix}a,b\\ s,u\end{pmatrix}^{1-1/p}\min\left\{\omega\begin{pmatrix}a,b\\ t,u\end{pmatrix},\omega\begin{pmatrix}a,b\\ s,t\end{pmatrix}\right\}^{1/p}.\end{aligned}$$

Proof. From the very definition of $\omega([a,b]\times[s,u])$, it follows that for every fixed $\varepsilon>0$, there exists a rectangular (not necessarily grid-like) partition of $[a,b]\times[s,u]$, say $\Pi\in\mathscr{P}([a,b]\times[s,u])$, such that

$$\sum_{R\in\Pi}\left|f(R)\right|^p>\omega\begin{pmatrix}a,b\\ s,u\end{pmatrix}-\varepsilon.$$

Let us divide Π in $\Pi_l\cup\Pi_m\cup\Pi_r$ where Π_l contains all $R\in\Pi$ such that $R\subset[a,b]\times[s,t]$, Π_r contains all $R\in\Pi:R\subset[a,b]\times[t,u]$ and Π_m contains all remaining rectangles of Π (i.e. the one such that their interior intersect with the line $[a,b]\times[t,t]$. It follows that

$$\sum_{R\in\Pi_l}\left|f(R)\right|^p+\sum_{R\in\Pi_m}\left|f(R)\right|^p+\sum_{R\in\Pi_r}\left|f(R)\right|^p>\omega\begin{pmatrix}a,b\\ s,u\end{pmatrix}-\varepsilon$$

Every $R\in\Pi_m$ can be split into (essentially disjoint) rectangles $R_1\subset[a,b]\times[s,t]$ and $R_2\subset[a,b]\times[t,u]$. Set $\Pi_m^1=\{R_1:R_1\in\Pi_m\}$ and Π_m^2 similarly. Note that $\Pi_l\cup\Pi_m^1\in\mathscr{P}([a,b]\times[s,t])$ and $\Pi_m^2\cup\Pi_r\in\mathscr{P}([a,b]\times[t,u])$. Then, with

$$\Delta:=\sum_{R\in\Pi_m}\left[\left|f(R)\right|^p-\left|f(R_1)\right|^p-\left|f(R_2)\right|^p\right]$$

we have

$$\sum_{R\in\Pi_l\cup\Pi_m^1}\left|f(R)\right|^p+\sum_{R\in\Pi_m^2\cup\Pi_r}\left|f(R)\right|^p+\Delta>\omega([a,b]\times[s,u])-\varepsilon$$

and hence ,we have

$$\omega\begin{pmatrix}a,b\\ s,t\end{pmatrix}+\omega\begin{pmatrix}a,b\\ t,u\end{pmatrix}+\Delta>\omega\begin{pmatrix}a,b\\ s,u\end{pmatrix}-\varepsilon.$$

We now bound Δ. As $f(R)=f\left(R_1\right)+f\left(R_2\right)$,

$$\begin{aligned}
\Delta &= \sum_{R^j\in\Pi_m}\left|f\left(R_1^j\right)+f\left(R_2^j\right)\right|^p-\left|f\left(R_1^j\right)\right|^p-\left|f\left(R_2^j\right)\right|^p \\
&\leq \sum_{R\in\Pi_m}\left(\left|f\left(R_1^j\right)\right|+\left|f\left(R_2^j\right)\right|\right)^p-\left|f\left(R_1^j\right)\right|^p-\left|f\left(R_2^j\right)\right|^p. \\
&\leq \sum_{R\in\Pi_m}\left(\left|f\left(R_1^j\right)\right|+\left|f\left(R_2^j\right)\right|\right)^p-\left|f\left(R_1^j\right)\right|^p
\end{aligned}$$

If $R^j=[\tau_j,\tau_{j+1}]\times[c,d]$, define $R_3^j=[\tau_j,\tau_{j+1}]\times[s,u]$. Then, quite obviously, we have $\left|f\left(R_1^j\right)\right|^p\leq\omega\left(R_3^j\right)$ and $\left|f\left(R_2^j\right)\right|^p\leq\omega\left(R_3^j\right)$. By the mean value theorem, there exists $\theta\in[0,1]$ such that

$$\begin{aligned}
&\left(\left|f\left(R_1^j\right)\right|+\left|f\left(R_2^j\right)\right|\right)^p-\left|f\left(R_1^j\right)\right|^p \\
&= p\left(\left|f\left(R_1^j\right)\right|+\theta\left|f\left(R_2^j\right)\right|\right)^{p-1}\left|f\left(R_2^j\right)\right| \\
&\leq p2^{p-1}\omega\left(R_3^j\right)^{1-1/p}\left|f\left(R_2^j\right)\right| \\
&\leq p2^{p-1}\omega\begin{pmatrix}\tau_j,\tau_{j+1}\\ s,u\end{pmatrix}^{1-1/p}\omega\begin{pmatrix}\tau_j,\tau_{j+1}\\ t,u\end{pmatrix}^{1/p}.
\end{aligned}$$

Hence, summing over j, and using Hölder inequality

$$\begin{aligned}
\Delta &\leq p2^{p-1}\sum_j\omega\begin{pmatrix}\tau_j,\tau_{j+1}\\ s,u\end{pmatrix}^{p-1}\omega\begin{pmatrix}\tau_j,\tau_{j+1}\\ t,u\end{pmatrix} \\
&\leq p2^{p-1}\left(\sum_j\omega\begin{pmatrix}\tau_j,\tau_{j+1}\\ s,u\end{pmatrix}\right)^{1-1/p}\left(\sum_j\omega\begin{pmatrix}\tau_j,\tau_{j+1}\\ t,u\end{pmatrix}\right)^{1/p} \\
&\leq p2^{p-1}\omega\begin{pmatrix}a,b\\ s,u\end{pmatrix}^{1-1/p}\omega\begin{pmatrix}a,b\\ t,u\end{pmatrix}^{1/p}
\end{aligned}$$

Interchanging the roles of $R_1$and R_2, we also obtain that

$$\Delta\leq p2^{p-1}\omega\begin{pmatrix}a,b\\ s,u\end{pmatrix}^{1-1/p}\omega\begin{pmatrix}a,b\\ t,u\end{pmatrix}^{1/p},$$

which concludes the proof. □

Continuity: ω is a map from $\Delta_T\times\Delta_T\to[0,\infty)$; the identification of points $\left(\left(a_1,a_2\right),\left(a_3,a_4\right)\right)\in\Delta_T\times\Delta_T$ with rectangles in $[0,T]^2$ of the form $A=\begin{pmatrix}a_1,a_2\\ a_3,a_4\end{pmatrix}=\left[a_1,a_2\right]\times\left[a_3,a_4\right]$ is pure convention. If A is non-degenerate (i.e. $a_1 < a_2, a_3 < a_4$) and $|h| = \max_{i=1}^4 |h_i|$ sufficiently small then

$$A^h := \begin{pmatrix} (a_1+h_1) \vee 0, (a_2+h_2) \wedge T \\ (a_3+h_3) \vee 0, (a_4+h_4) \wedge T \end{pmatrix}$$

is again a non-degenerate rectangle in $[0,T]^2$. We can then set for $r > 0$, sufficiently small,

$$A^{\circ;r} := A^{(r,-r,r,-r)}, \ \bar{A}^r := A^{(-r,r,-r,r)}$$

and note that, whenever $|h|$ is small enough to have $A^{\circ;|h|}$ well-defined,

$$\begin{aligned} A^{\circ;|h|} &\subset A \subset \bar{A}^{|h|}, && (5.1)\\ A^{\circ;|h|} &\subset A^h \subset \bar{A}^{|h|}. && (5.2)\end{aligned}$$

The above definition of A^h (and $A^{\circ;r}, \bar{A}^r$) is easily extended to degenerate A, such that the inclusions (5.1),(5.2) remain valid: For instance, in the case $a_1 = a_2$ we would replace the first line in the definition of A^h by

$$\begin{gathered} (a_1+h_1) \vee 0, (a_2+h_2) \wedge T \text{ if } h_1 \le 0 \le h_2 \\ (a_1+h_1) \vee 0, a_2 \text{ if } h_1, h_2 \le 0 \\ a_1, (a_2+h_2) \wedge T \text{ if } h_1, h_2 \ge 0 \\ a_1, a_2 \text{ if } h_1 \ge 0 \ge h_2 \end{gathered}$$

and similarly in the case $a_3 = a_4$. We will prove that, for any rectangle $A \subset [0,T]^2$,

$$\omega\left(A^h\right) \to \omega(A) \text{ as } |h| \downarrow 0.$$

This end we can and will consider $|h|$ is small enough to have $A^{\circ;|h|}$ (and thus $A^h, \bar{A}^{|h|}$) well-defined. By monotonicity of ω, it follows that

$$\omega\left(A^{\circ;|h|}\right) \le \omega\left(A^h\right) \le \omega\left(\bar{A}^{|h|}\right)$$

and the limits,

$$\begin{aligned} \omega^\circ(A) &:= \lim_{r\downarrow 0} \omega\left(A^{\circ;r}\right) \le \omega(A), && (5.3)\\ \bar{\omega}(A) &:= \lim_{r\downarrow 0} \omega\left(\bar{A}^r\right) \ge \omega(A), \end{aligned}$$

exist since $\omega(A^{\circ;r})$ [resp. $\omega(\bar{A}^r)$] are bounded from above [resp. below] and increasing [resp. decreasing] as $r \downarrow 0$. It follows that

$$\omega^\circ(A) \le \varliminf_{|h|\downarrow 0} \omega\left(A^h\right) \le \varlimsup_{|h|\downarrow 0} \omega\left(A^h\right) \le \bar{\omega}(A).$$

The goal is now to show that $\omega^\circ(A) = \omega(A)$ ("inner continuity") and $\bar{\omega}(A) = \omega(A)$ ("outer continuity") since this implies that $\varliminf \omega\left(A^h\right) = \varlimsup \omega\left(A^h\right) = \omega(A)$, which is what we want.
Inner continuity: We first show that ω° is super-additive in the sense of definition 2. To this end, consider $\{R_i\} \in \mathscr{P}(R)$, some rectangle $R \subset [0,T]^2$. For r small enough, the rectangles

$$\left\{R_i^{0,r}\right\}$$

are well-defined and essentially disjoint. They can be completed to a partition of $R^{0,r}$ and hence, by super-additivity of ω,

$$\sum_i \omega\left(R_i^{0,r}\right) \leq \omega\left(R^{0,r}\right);$$

sending $r \downarrow 0$ yields the desired super-addivity of ω°;

$$\sum_i \omega^\circ\left(R_i\right) \leq \omega^\circ(R).$$

On the other hand, continuity of f on $[0,T]^2$ implies

$$\begin{array}{rcl} \left|f(A)\right|^p & \leq & \left|f\left(A^{\circ,r}\right)\right|^p + o(1) \\ & \leq & \omega\left(A^{\circ,r}\right) + o(1) \text{ as } r \downarrow 0 \end{array}$$

and hence $\left|f(A)\right|^p \leq \omega^\circ(A)$, for any rectangle $A \subset [0,T]^2$. Using super-additivity of ω° immediately gives

$$\omega(R) \overset{\text{by def.}}{=} \sup_{\Pi \in \mathscr{P}(R)} \sum_{A \in \Pi} \left|f(A)\right|^p \leq \omega^\circ(R);$$

together with (5.3) we thus have $\omega(R) = \omega^\circ(R)$. Since R was an arbitrary rectangle in $[0,T]^2$ inner continuity is proved.

<u>Outer continuity:</u> We assume $A \subset (0,T)^2$ (i.e. $0 < a_1 \leq a_2 < T, 0 < a_3 \leq a_4 < T$) and take $r > 0$ small enough so that

$$\bar{A}^r = \begin{pmatrix} a_1 - r, a_2 + r \\ a_3 - r, a_4 + r \end{pmatrix};$$

the general case $A \subset [0,T]^2$ is handled by a (trivial) adaption of the argument for the remaining cases (i.e. $a_1 = 0$ or $a_2 = T$ or $a_3 = 0$ or $a_4 = T$). We first note that

$$\begin{array}{rcl} \omega\left(\bar{A}^r\right) - \omega(A) & = & \omega\begin{pmatrix} a_1 - r, a_2 + r \\ a_3 - r, a_4 + r \end{pmatrix} - \omega\begin{pmatrix} a_1, a_2 \\ a_3, a_4 \end{pmatrix} \\ & \leq & \left|\omega\begin{pmatrix} a_1 - r, a_2 + r \\ a_3 - r, a_4 + r \end{pmatrix} - \omega\begin{pmatrix} a_1 - r, a_2 \\ a_3 - r, a_4 + r \end{pmatrix}\right| \\ & & + \left|\omega\begin{pmatrix} a_1 - r, a_2 \\ a_3 - r, a_4 + r \end{pmatrix} - \omega\begin{pmatrix} a_1, a_2 \\ a_3 - r, a_4 + r \end{pmatrix}\right| \\ & & + \left|\omega\begin{pmatrix} a_1, a_2 \\ a_3 - r, a_4 + r \end{pmatrix} - \omega\begin{pmatrix} a_1, a_2 \\ a_3, a_4 + r \end{pmatrix}\right| \\ & & + \left|\omega\begin{pmatrix} a_1, a_2 \\ a_3, a_4 + r \end{pmatrix} - \omega\begin{pmatrix} a_1, a_2 \\ a_3, a_4 \end{pmatrix}\right| \end{array}$$

Now we use lemma 2; with

$$\Delta := \left|\omega\begin{pmatrix} a_1 - r, a_2 + r \\ a_3 - r, a_4 + r \end{pmatrix} - \omega\begin{pmatrix} a_1 - r, a_2 \\ a_3 - r, a_4 + r \end{pmatrix}\right|$$

we have

$$\begin{aligned}\Delta &\leq \omega\begin{pmatrix} a_2, a_2+r \\ a_3-r, a_4+r \end{pmatrix} + c\omega\left([0,T]^2\right)^{1-1/p}\omega\begin{pmatrix} a_2, a_2+r \\ a_3-r, a_4+r \end{pmatrix}^{1/p} \\ &\leq \omega\begin{pmatrix} a_2, a_2+r \\ 0, T \end{pmatrix} + c\omega\left([0,T]^2\right)^{1-1/p}\omega\begin{pmatrix} a_2, a_2+r \\ 0, T \end{pmatrix}^{1/p},\end{aligned}$$

and similar inequalities for the other three terms in our upper estimate on $\omega\left(\bar{A}^r\right) - \omega(A)$ above. So it only remains to prove that for $a \in (0,T)$

$$\omega\begin{pmatrix} a, a+r \\ 0, T \end{pmatrix}, \ \omega\begin{pmatrix} a-r, a \\ 0, T \end{pmatrix}, \ \omega\begin{pmatrix} 0, T \\ a, a+r \end{pmatrix}, \text{ and } \omega\begin{pmatrix} 0, T \\ a-r, a \end{pmatrix}$$

converge to 0 when r tends to 0.But this is easy; using super-addivity of ω and inner-continuity we see that

$$\begin{aligned}\omega\begin{pmatrix} a, a+r \\ 0, T \end{pmatrix} &\leq \omega\begin{pmatrix} a, T \\ 0, T \end{pmatrix} - \omega\begin{pmatrix} a+r, T \\ 0, T \end{pmatrix} \\ &\to 0 \text{ as } r \downarrow 0.\end{aligned}$$

Other expressions are handled similarly and our proof of outer continuity is finished.

6 Appendix

6.1 Young and Young-Towghi discrete inequalities

6.1.1 One dimensional case.

Consider a dissection $D = \left(0 = t_0, ..., t_n = T\right) \in \mathscr{D}([0,T])$. We define the "discrete integral" between $x, y : [0,T] \to \mathbb{R}$ as

$$I^D = \int_D y\,dx = \sum_{i=1}^{n} y_{t_i} x_{t_{i-1},t_i}.$$

Lemma 3. *Let $p, q \geq 1$, assume that $\theta = 1/p + 1/q > 1$. Assume $x, y : [0,T] \to \mathbb{R}$ are finite p- resp. q-variation. Then there exists $t_{i_0} \in D \backslash \{0,T\}$ (equivalently: $i_0 \in \{1, \ldots, n-1\}$) such that*

$$\left| \int_D y\,dx - \int_{D \backslash \left\{t_{i_0}\right\}} y\,dx \right| \leq \frac{1}{(n-1)^\theta} |x|_{p\text{-var},[0,T]} \left|y\right|_{q\text{-var},[0,T]}$$

Iterated removal of points in the dissection, using the above lemma, leads immediately to Young's maximal inequality which is the heart of the Young's integral construction.

Theorem 2 (Young's Maximal Inequality). *Let $p, q \geq 1$, assume that $\theta = 1/p + 1/q > 1$, and consider two paths x, y from $[0, T]$ into $\mathbb{R}$ of finite p-variation and q-variation, with $y_0 = 0$. Then*

$$\left| \int_D y\,dx \right| \leq (1 + \zeta(\theta)) |x|_{p\text{-}var;[0,T]} \left| y \right|_{q\text{-}var;[0,T]}$$

and this estimate is uniform over all $D \in \mathscr{D}([0, T])$.

Proof. Iterative removal of "i_0" gives, thanks to lemma 3,

$$\begin{aligned} \left| \int_D y\,dx - \int_{\{0,T\}} y\,dx \right| &\leq \sum_{n \geq 2} \frac{1}{(n-1)^\theta} |x|_{p\text{-var},[0,T]} \left| y \right|_{q\text{-var},[0,T]} \\ &\leq \zeta(\theta) |x|_{p\text{-var},[0,T]} \left| y \right|_{q\text{-var},[0,T]} \end{aligned}$$

Finally, $\int_{\{0,T\}} y\,dx = y_T x_{0,T} = y_{0,T} x_{0,T}$ since $y_{0,T} = y_T - y_0$ and $y_0 = 0$ and hence

$$\left| \int_{\{0,T\}} y\,dx \right| = \left| y_{0,T} x_{0,T} \right| \leq |x|_{p\text{-var},[0,T]} \left| y \right|_{q\text{-var},[0,T]}$$

and we conclude with the triangle inequality. □

Proof. (**Lemma** 3) Observe that, for any $t_i \in D \backslash \{0, T\}$ with $1 \leq i \leq n - 1$

$$I^D - I^{D \backslash \{t_i\}} = y_{t_i, t_{i+1}} x_{t_{i-1}, t_i}$$

We pick t_{i_0} to make this difference as small as possible:

$$\left| I^D - I^{D \backslash \{t_{i_0}\}} \right| \leq \left| I^D - I^{D \backslash \{t_i\}} \right| \text{ for all } i \in \{1, \dots, n-1\}$$

As an elementary consequence, we have

$$\left| I^D - I^{D \backslash \{t_{i_0}\}} \right|^{\frac{1}{\theta}} \leq \frac{1}{n-1} \sum_{i=1}^{n-1} \left| I^D - I^{D \backslash \{t_i\}} \right|^{1/\theta}.$$

The plan is to get an estimate on $\sum_{i=1}^{n-1} \left| I^D - I^{D \backslash \{t_i\}} \right|^{1/\theta}$ independent of n. In fact, we shall see that

$$\sum_{i=1}^{n-1} \left| I^D - I^{D \backslash \{t_i\}} \right|^{1/\theta} \leq |x|_{p\text{-var},[0,T]}^{1/\theta} \left| y \right|_{q\text{-var},[0,T]}^{1/\theta} \tag{6.1}$$

and the desired estimate

$$\left| I^D - I^{D \backslash \{t_{i_0}\}} \right| \leq \left(\frac{1}{n-1} \right)^\theta |x|_{p\text{-var},[0,T]} \left| y \right|_{q\text{-var},[0,T]}$$

follows. It remains to establish (6.1); thanks to Hölder's inequality, using $1/\left(q\theta\right)+1/\left(p\theta\right)=1$,

$$\begin{aligned}\sum_{i=1}^{n-1}\left|I^{D}-I^{D\setminus\{t_i\}}\right|^{1/\theta} &= \left(\sum_{i=1}^{n-1}\left|y_{t_i,t_{i+1}}\right|^{1/\theta}\left|x_{t_{i-1},t_i}\right|^{1/\theta}\right)^{\theta}\\ &\le \left(\sum_{i=1}^{m-1}\left|y_{t_i,t_{i+1}}\right|^{q}\right)^{\frac{1}{q\theta}}\left(\sum_{i=1}^{n-1}\left|x_{t_{i-1},t_i}\right|^{p}\right)^{\frac{1}{p\theta}}\\ &\le |x|^{1/\theta}_{p\text{-var},[0,T]}\left|y\right|^{1/\theta}_{q\text{-var},[0,T]}.\end{aligned}$$

and we are done. □

6.1.2 Young-Towghi maximal inequality (2D)

We now consider the two-dimensional case. To this end, fix two dissections $D=\left(0=t_0,...,t_n=T\right)$ and $D'=\left(0=t'_0,...,t'_m=T\right)$,and define the discrete integral between $x,y:[0,T]^2\to\mathbb{R}$ as

$$I^{D,D'}=\int_{D\times D'}ydx:=\sum_i\sum_j y\begin{pmatrix}t_i\\t'_j\end{pmatrix}x\begin{pmatrix}t_{i-1},t_i\\t'_{j-1},t'_j\end{pmatrix}. \tag{6.2}$$

Lemma 4. *Let $p,q\ge1$, assume that $\theta=1/p+1/q>1$. Assume $x,y:[0,T]^2\to\mathbb{R}$ are finite p- resp. q-variation. Then there exists $t_{i_0}\in D\setminus\{0,T\}$ (equivalently: $i_0\in\{1,\dots,n-1\}$ such that for every $\alpha\in(1,\theta)$,*

$$\left|\int_{D\times D'}dx-\int_{D\setminus\left\{t_{i_0}\right\}\times D'}ydx\right|\le\left(\frac{1}{n-1}\right)^{\alpha}\left(1+\zeta\left(\frac{\theta}{\alpha}\right)\right)^{\alpha}V_p\left(x;[0,T]^2\right)V_q\left(y;[0,T]^2\right)$$

Iterative removal of "i_0" leads to Young-Towghi's maximal inequality.

Theorem 3 (Young-Towghi Maximal Inequality)**.** *Let $p,q\ge1$, assume that $\theta=1/p+1/q>1$, and consider $x,y:[0,T]^2\to\mathbb{R}$ of finite p- resp. q-variation and $y(0,\cdot)=y(\cdot,0)=0$. Then, for every $\alpha\in(1,\theta)$,*

$$\left|\int_{D\times D'}ydx\right|\le\left[\left(1+\zeta\left(\frac{\theta}{\alpha}\right)\right)^{\alpha}\zeta(\alpha)+(1+\zeta(\theta))\right]V_p\left(x;[0,T]^2\right)V_q\left(y;[0,T]^2\right)$$

and this estimate is uniform over all $D,D'\in\mathscr{D}([0,T])$

Proof. Iterative removal of "i_0" gives

$$\begin{aligned}\left|\int_{D\times D'}ydx-\int_{\{0,T\}\times D'}ydx\right| &\le \sum_{n\ge2}\left(\frac{1}{n-1}\right)^{\alpha}\left(1+\zeta\left(\frac{\theta}{\alpha}\right)\right)^{\alpha}V_p\left(x;[0,T]^2\right)V_q\left(y;[0,T]^2\right)\\ &\le \zeta(\alpha)\left(1+\zeta\left(\frac{\theta}{\alpha}\right)\right)^{\alpha}V_p\left(x;[0,T]^2\right)V_q\left(y;[0,T]^2\right).\end{aligned}$$

It only remains to bound

$$\int_{\{0,T\}\times D'} y\,dx = \sum_j y\begin{pmatrix} T \\ t'_j \end{pmatrix} x\begin{pmatrix} 0,T \\ t'_{j-1},t'_j \end{pmatrix} = \int_{D'} y\begin{pmatrix} 0,T \\ \cdot \end{pmatrix} dx\begin{pmatrix} 0,T \\ \cdot \end{pmatrix}$$

where we used $y\begin{pmatrix} 0 \\ \cdot \end{pmatrix} = 0$ in the last equality. From Young's 1D maximal inequality, we have

$$\begin{aligned} \left|\int_{\{0,T\}\times D'} y\,dx\right| &\leq (1+\zeta(\theta))\left|y\begin{pmatrix} 0,T \\ 0,. \end{pmatrix}\right|_{q\text{-var},[0,T]}\left|x\begin{pmatrix} 0,T \\ 0,. \end{pmatrix}\right|_{p\text{-var},[0,T]} \\ &\leq (1+\zeta(\theta))V_p\left(x;[0,T]^2\right)V_q\left(y;[0,T]^2\right) \end{aligned}$$

The triangle inequality allows us to conclude. □

Proof. (**Lemma** 4) Observe that, for any $t_i \in D\backslash\{0,T\}$

$$\begin{aligned} I^{D,D'} - I^{D\backslash\{t_i\},D'} &= \int_{D'} y\begin{pmatrix} t_i,t_{i+1} \\ \cdot \end{pmatrix} x\begin{pmatrix} t_{i-1},t_i \\ \cdot \end{pmatrix} \\ &= \int_{D'} y\begin{pmatrix} t_i,t_{i+1} \\ 0,\cdot \end{pmatrix} x\begin{pmatrix} t_{i-1},t_i \\ \cdot \end{pmatrix} \end{aligned}$$

where we used $y\begin{pmatrix} \cdot \\ 0 \end{pmatrix} = 0$. We pick t_{i_0} to make this difference as small as possible:

$$\left|I^{D,D'} - I^{D\backslash\{t_{i_0}\},D'}\right| \leq \left|I^{D,D'} - I^{D\backslash\{t_i\},D'}\right| \text{ for all } i \in \{1,\dots,n-1\}$$

As an elementary consequence,

$$\left|I^{D,D'} - I^{D\backslash\{t_{i_0}\},D'}\right|^{1/\alpha} \leq \frac{1}{n-1}\sum_{i=1}^{n-1}\left|I^{D,D'} - I^{D\backslash\{t_i\},D'}\right|^{1/\alpha}. \tag{6.3}$$

The plan is to get an estimate on $\sum_{i=1}^{n-1}\left|I^{D,D'} - I^{D\backslash\{t_i\},D'}\right|^{1/\alpha}$ independent of n and uniformly in $D' \in \mathscr{D}([0,T])$; in fact, we shall see that

$$\Delta^{D,D'} := \sum_{i=1}^{n-1}\left|I^{D,D'} - I^{D\backslash\{t_i\},D'}\right|^{1/\alpha} \leq cV_p\left(x;[0,T]^2\right)^{1/\alpha}V_q\left(y;[0,T]^2\right)^{1/\alpha} \tag{6.4}$$

with $c = 1+\zeta\left(\frac{\theta}{\alpha}\right)$ and the desired estimate

$$\left|I^D - I^{D\backslash\{t_{i_0}\}}\right| \leq \left(\frac{c}{n-1}\right)^{\alpha} V_p\left(x;[0,T]^2\right)V_q\left(y;[0,T]^2\right)$$

follows. It remains to establish (6.4); to this end we consider the removal of $t'_j \in D'\backslash\{0,T\}$ from D' and note that

$$\left(I^{D,D'} - I^{D\backslash\{t_i\},D'}\right) - \left(I^{D,D'\backslash\{t'_j\}} - I^{D\backslash\{t_i\},D'\backslash\{t'_j\}}\right) = y\begin{pmatrix} t_i,t_{i+1} \\ t'_j,t'_{j+1} \end{pmatrix} x\begin{pmatrix} t_{i-1},t_i \\ t'_{j-1},t'_j \end{pmatrix}$$

Using the elementary inequality $|a|^{1/\alpha}-|b|^{1/\alpha}\le|a-b|^{1/\alpha}$ valid for $a,b\in\mathbb{R}$ and $\alpha\ge 1$ we have

$$\begin{aligned}&\left|I^{D,D'}-I^{D\setminus\{t_i\},D'}\right|^{1/\alpha}-\left|I^{D,D'\setminus\{t'_j\}}-I^{D\setminus\{t_i\},D'\setminus\{t'_j\}}\right|^{1/\alpha}\\ &\le\left|\left(I^{D,D'}-I^{D\setminus\{t_i\},D'}\right)-\left(I^{D,D'\setminus\{t'_j\}}-I^{D\setminus\{t_i\},D'\setminus\{t'_j\}}\right)\right|^{1/\alpha}.\end{aligned}$$

Hence, summing over i, we get

$$\begin{aligned}&\Delta^{D,D'}-\Delta^{D,D'\setminus\{t'_j\}}\\ \le\ &\sum_{i=1}^{n-1}\left|\left(I^{D,D'}-I^{D\setminus\{t_i\},D'}\right)-\left(I^{D,D'\setminus\{t'_j\}}-I^{D\setminus\{t_i\},D'\setminus\{t'_j\}}\right)\right|^{1/\alpha}\\ =\ &\sum_{i=1}^{n-1}\left|y\begin{pmatrix}t_i,t_{i+1}\\ t'_j,t'_{j+1}\end{pmatrix}\right|^{1/\alpha}\left|x\begin{pmatrix}t_{i-1},t_i\\ t'_{j-1},t'_j\end{pmatrix}\right|^{1/\alpha} \qquad (6.5)\\ \le\ &\left(\sum_{i=1}^{n-1}\left|y\begin{pmatrix}t_i,t_{i+1}\\ t'_j,t'_{j+1}\end{pmatrix}\right|^{\theta q/\alpha}\right)^{\frac{1}{\theta q}}\left(\sum_{i=1}^{n-1}\left|x\begin{pmatrix}t_{i-1},t_i\\ t'_{j-1},t'_j\end{pmatrix}\right|^{\theta p/\alpha}\right)^{\frac{1}{\theta p}}\\ \le\ &\left(\sum_{i=1}^{n-1}\left|y\begin{pmatrix}t_i,t_{i+1}\\ t'_j,t'_{j+1}\end{pmatrix}\right|^{q}\right)^{\frac{1}{\alpha q}}\left(\sum_{i=1}^{n-1}\left|x\begin{pmatrix}t_{i-1},t_i\\ t'_{j-1},t'_j\end{pmatrix}\right|^{p}\right)^{\frac{1}{\alpha p}};\end{aligned}$$

in the last step we used that the $\ell^{\theta p/\alpha}$ norm on $\mathbb{R}^{n-1}$ is dominated by the ℓ^p norm (because $\theta p/\alpha>p$). It follows that

$$\Delta^{D,D'}-\Delta^{D,D'\setminus\{t'_j\}}\le Y_j^{1/\alpha}X_j^{1/\alpha} \qquad (6.6)$$

where

$$Y_j:=\left(\sum_{i=1}^{n-1}\left|y\begin{pmatrix}t_i,t_{i+1}\\ t'_j,t'_{j+1}\end{pmatrix}\right|^{q}\right)^{\frac{1}{q}},\ X_j:=\left(\sum_{i=1}^{n-1}\left|x\begin{pmatrix}t_{i-1},t_i\\ t'_{j-1},t'_j\end{pmatrix}\right|^{p}\right)^{\frac{1}{p}}$$

We pick $t'_{j_0}\in D'\setminus\{0,T\}$ (i.e. $1\le j_0\le m-1$) to make this difference as small as possible,

$$\Delta^{D,D'}-\Delta^{D,D'\setminus\{t'_{j_0}\}}\le\Delta^{D,D'}-\Delta^{D,D'\setminus\{t'_j\}}\text{ for all }j\in\{1,\dots,m-1\};$$

we shall see below that

$$\left|\Delta^{D,D'}-\Delta^{D,D'\setminus\{t'_{j_0}\}}\right|\le\left(\frac{1}{m-1}\right)^{\frac{\theta}{\alpha}}V_p\left(x;[0,T]^2\right)^{1/\alpha}V_q\left(y;[0,T]^2\right)^{1/\alpha}; \qquad (6.7)$$

iterated removal of "j_0" yields

$$\Delta^{D,D'}\le\Delta^{D,\{0,T\}}+\zeta\left(\frac{\theta}{\alpha}\right)V_p\left(x,[0,T]^2\right)^{1/\alpha}V_q\left(y,[0,T]^2\right)^{1/\alpha};$$

as in (6.5) we estimate

$$\Delta^{D,\{0,T\}}=\sum_{i=1}^{n-1}\left|y\begin{pmatrix}t_i,t_{i+1}\\ 0,T\end{pmatrix}x\begin{pmatrix}t_{i-1},t_i\\ 0,T\end{pmatrix}\right|^{1/\alpha}\le\dots\le V_p\left(x,[0,T]^2\right)^{1/\alpha}V_q\left(y,[0,T]^2\right)^{1/\alpha}$$

and (6.4) follows, as desired. The only thing left is to establish (6.7). Using (6.6) we can write

$$\begin{aligned}
\Delta^{D,D'} - \Delta^{D,D'\setminus\left\{t'_{j_0}\right\}} &\le \left(\prod_{j=1}^{m-1} \Delta^{D,D'} - \Delta^{D,D'\setminus\left\{t'_j\right\}}\right)^{\frac{1}{m-1}} \\
&\le \left(\prod_{j=1}^{m-1} X_j^{1/.\alpha} Y_j^{1/\alpha}\right)^{\frac{1}{m-1}} \\
&= \left(\prod_{j=1}^{m-1} X_j^{p}\right)^{\frac{1}{m-1}\frac{1}{\alpha p}} \left(\prod_{j=1}^{m-1} Y_j^{q}\right)^{\frac{1}{m-1}\frac{1}{\alpha q}}.
\end{aligned}$$

Using the geometric/arithmetic inequality, we obtain

$$\begin{aligned}
\left(\prod_{j=1}^{m-1} X_j^{p}\right)^{\frac{1}{m-1}\frac{1}{\alpha p}} &\le \left(\frac{1}{m-1}\sum_{j=1}^{m-1} X_j^{p}\right)^{\frac{1}{\alpha p}} \\
&\le \left(\frac{1}{m-1}\right)^{\frac{1}{\alpha p}} \left(\sum_{j=1}^{m-1}\sum_{i=1}^{n-1}\left| x\begin{pmatrix} t_{i-1}, t_i \\ t'_{j-1}, t'_j \end{pmatrix}\right|^p\right)^{\frac{1}{\alpha p}} \\
&\le \left(\frac{1}{m-1}\right)^{\frac{1}{\alpha p}} V_p\left(x, [0,T]^2\right)^{1/\alpha}.
\end{aligned}$$

and, similarly,

$$\left(\prod_{j=1}^{m-1} Y_j^{q}\right)^{\frac{1}{m-1}\frac{1}{\alpha q}} \le \left(\frac{1}{m-1}\right)^{\frac{1}{\alpha q}} V_q\left(y, [0,T]^2\right)^{1/\alpha}.$$

Using $\frac{1}{\alpha p} + \frac{1}{\alpha q} = \frac{\theta}{\alpha}$, we thus arrive at

$$\Delta^{D,D'} - \Delta^{D,D'\setminus\left\{t'_{j_0}\right\}} \le \left(\frac{1}{m-1}\right)^{\frac{\theta}{\alpha}} V_p\left(x, [0,T]^2\right)^{1/\alpha} V_q\left(y, [0,T]^2\right)^{1/\alpha}$$

which is precisely the claimed estimate (6.7). □

References

[1] Coutin, L., Qian, Z. : Stochastic analysis, rough path analysis and fractional Brownian motion, Probab. Theory Related Fields 122 (2002), 108–140 MR1883719

[2] Friz, P., Victoir, N.: Differential equations driven by Gaussian signals. Ann. Inst. Henri Poincaré Probab. Stat. 46 (2010), no. 2, 369–41 MR2667703

[3] Friz, P., Victoir, N.: Multidimensional stochastic processes as rough paths. Theory and applications. Cambridge University Press, Cambridge, 2010. (Errata available via CUP's webpages.) MR2604669

[4] Towghi: Multidimensional extension of L.C.Young's Inequality, J. of Inequalities in Pure and Applied Mathematics, Vol 3, 2-22, 2002 MR1906391