

The roughness exponent and its application in finance

by

Xiyue Han

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Actuarial Science

Waterloo, Ontario, Canada, 2024

© Xiyue Han 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Mathieu Rosenbaum Professor, Centre de Mathématiques Appliquées École Polytechnique
Supervisor:	Alexander Schied Professor, Department of Statistics and Actuarial Science University of Waterloo
Internal Members:	Ruodu Wang Professor, Department of Statistics and Actuarial Science University of Waterloo Tony Wirjanto Professor, Department of Statistics and Actuarial Science University of Waterloo
Internal-External Member:	Yuying Li Professor, Cheriton School of Computer Science University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis is based on the following preprints and published works:

(a) Chapter 2 is based on

- (i) [45]: Xiyue Han and Alexander Schied. The roughness exponent and its model-free estimation. *arXiv:2111.10301*, 2021
- (ii) [44]: Xiyue Han. A Gladyshev theorem for trifractional Brownian motion and n -th order fractional Brownian motion. *Electronic Communications in Probability*, 26, 1-12, 2021
- (iii) [48]: Xiyue Han and Alexander Schied. On laws absolutely continuous with respect to fractional Brownian motion. *arXiv:2306.11824*, 2023

(b) Chapter 3 is based on

- (i) [47]: Xiyue Han and Alexander Scheid. Estimating the roughness exponent of stochastic volatility from discrete observations of the realized variance. *arXiv:2307.02582*, 2023

(c) Chapter 4 is based on

- (i) [46]: Xiyue Han and Alexander Scheid. Robust Faber–Schauder approximation based on discrete observations of an antiderivative. *arXiv:2211.11907*, 2022

(d) Chapter 5 is based on

- (i) [49]: Xiyue Han, Alexander Schied and Zhenyuan Zhang. A probabilistic approach to the Φ -variation of classical fractal functions with critical roughness. *Statistics & Probability Letters*, 168, 108920, 2021
- (ii) [50]: Xiyue Han, Alexander Schied and Zhenyuan Zhang. A limit theorem for Bernoulli convolutions and the Φ -variation of functions in the Takagi class. *Journal of Theoretical Probability*, 35, 2853–2878, 2022

Abstract

Rough phenomena and trajectories are prevalent across mathematics, engineering, physics, and the natural sciences. In quantitative finance, sparked by the observation that the historical realized volatility is rougher than a Brownian martingale, rough stochastic volatility models were recently proposed and studied extensively. Unlike classical diffusion volatility models, the volatility process in a rough stochastic volatility model is driven by an irregular process. The roughness of the volatility process plays a pivotal role in governing the behavior of such models.

This thesis aims to explore the concept of roughness and estimate the roughness of financial time series in a strictly pathwise manner. To this end, we introduce the notion of the roughness exponent, which is a pathwise measure of the degree of roughness. A substantial portion of this thesis focuses on the model-free estimation of this roughness exponent for both price and volatility processes. Towards the end of this thesis, we study the Wiener–Young Φ -variation of classical fractal functions, which can be regarded as a finer characterization than the roughness exponent.

Chapter 2 introduces the roughness exponent and establishes a model-free estimator for the roughness exponent based on direct observations. We say that a continuous real-valued function x admits the roughness exponent R if the p^{th} variation of x converges to zero for $p > 1/R$ and to infinity for $p < 1/R$. The main result of this chapter provides a mild condition on the Faber–Schauder coefficients of x under which the roughness exponent exists and is given as the limit of the classical Gladyshev estimates $\widehat{R}_n(x)$. This result can be viewed as a strong consistency result for the Gladyshev estimator $\widehat{R}_n$ in an entirely model-free setting because no assumption whatsoever is made on the possible dynamics of the function x . Nonetheless, we show that the condition of our main result is satisfied for the typical sample paths of fractional Brownian motion with drift, and we provide almost-sure convergence rates for the corresponding Gladyshev estimates. In this chapter, we also discuss the connections between our roughness exponent and Besov regularity and weighted quadratic variation. Since the Gladyshev estimators are not scale-invariant, we construct several scale-invariant modifications of our estimator. Finally, we extend our results to the case where the p^{th} variation of x is defined over a sequence of unequally spaced partitions.

Chapter 3 considers the problem of estimating the roughness exponent of the volatility process in a stochastic volatility model that arises as a nonlinear function of a fractional Brownian motion with drift. To this end, we establish a new estimator $\widehat{\mathcal{R}}_n^*$ based on the Gladyshev estimator $\widehat{R}_n$ that estimates the roughness exponent of a continuous function x , but based on the observations of its antiderivative y . We identify conditions on the underlying trajectory x under which our estimates $\widehat{\mathcal{R}}_n^*(y)$ converge in a strictly pathwise

sense. Then, we verify that these conditions are satisfied by almost every sample path of fractional Brownian motion with drift. As a consequence, we obtain strong consistency of our estimator in the context of a large class of rough volatility models. Numerical simulations are implemented to show that our estimation procedure performs well after passing to a scale-invariant modification of our estimator.

Chapter 4 highlights the rationale of constructing the estimator $\widehat{\mathcal{R}}_n^*$ from the Gladyshev estimator $\widehat{R}_n$. In this chapter, we study the problem of reconstructing the Faber–Schauder coefficients of a continuous function from discrete observations of its antiderivative. This problem arises in the task of estimating the roughness exponent of the volatility process of a financial assets but is also independent interest. Our approach starts with formulating this problem through piecewise quadratic spline interpolation. We then provide a closed-form solution and an in-depth error analysis between the actual and approximated Faber–Schauder coefficients. These results lead to some surprising observations, which also throw new light on the classical topic of quadratic spline interpolation itself: They show that the well-known instabilities of this method can be located exclusively within the final generation of estimated Faber–Schauder coefficients, which suffer from non-locality and strong dependence on the initial value and the given data. By contrast, all other Faber–Schauder coefficients depend only locally on the data, are independent of the initial value, and admit uniform error bounds. We thus conclude that a robust and well-behaved estimator for our problem can be obtained by simply dropping the final-generation coefficients from the estimated Faber–Schauder coefficients.

Chapter 5 studies the Wiener-Young Φ -variation of classical fractal functions with a critical degree of roughness. In this case, the functions have vanishing p^{th} variation for all $p > q$ but are also of infinite p^{th} variation for $p < q$ for some $q \geq 1$. We partially resolve this apparent puzzle by showing that these functions have finite, nonzero, and linear Wiener–Young Φ -variation along the sequence of certain partitions. For instance, functions of bounded variation admit vanishing p^{th} variation for any $p > 1$. On the other hand, Weierstraß and Takagi-van der Waerden functions have vanishing p^{th} variation for $p > 1$ but are also nowhere differentiable and hence not of bounded variation. As a result, power variation and the roughness exponent fail to distinguish the difference of degree of roughness for these functions. However, we can individuate these functions by showing that the Weierstraß and Takagi-van der Waerden functions admit a nontrivial and linear Φ -variation along the sequence of b -adic partitions, where $\Phi(x) = x/\sqrt{-\log x}$. Moreover, for $q > 1$, we further develop a probabilistic approach so as to identify functions in the Takagi class that have linear and nontrivial Φ_q -variation for a prescribed Φ_q . Furthermore, for each fixed $q > 1$, the collection of functions Φ_q forms a wide class of increasing functions that are regularly varying at zero with an index of regular variation q .

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Alexander Schied. All the endeavour presented in this thesis originates from an interesting question that he asked me seven years ago in our first meeting. Since that very first day, I became his student, firstly as a master student and then as a doctoral student. We met every week over the last seven years, sharing the happiness of proving something and the sadness of failing to prove something on a weekly basis. It was his relentless support and sincere kindness that guided me through all stages of my doctoral study. Without him, this thesis would not be possible.

I extend my sincere thanks to Professors Yuying Li, Mathieu Rosenbaum, Ruodu Wang and Tony Wirjanto for their time in reviewing my thesis and providing valuable comments that greatly improve this thesis. I want to thank Professors Peijun Sang, Qinglong Tian, Bin Li, Chengguo Weng, Yi Shen, Mu Zhu, Greg Rice, Zelalem Negeri, Lisa Gao and Fangda Liu, who offered me valuable suggestions in all aspects of my life and research.

I am also deeply grateful to my co-author, Zhenyuan Zhang, with whom I shared countless inspiring discussions and, more importantly, uncountable tearful and joyful moments.

My heartfelt appreciation goes to Mary Lou Dufton for her unwavering support throughout my time at Waterloo. She has always been the first person to congratulate me whenever I made any achievements and the first person to talk with me during difficult times.

I wish to thank Zhiyi Shen, Zachary Van Oosten, Bryn Crandles, Austin Riis-Due, Jean-Gabriel Lauzier, Christopher Blier-Wong, Yimiao Zhao, Qiuqi Wang, Liyuan Lin, Hongdu Hu, Yuanyuan Zhang, Jiayue Zhang and Wenyuan Li for enlightening discussions on research and life. In particular, I want to thank Zachary, Byrn and Austin for bringing me one of the happiest days of my life. I would also like to thank Professor Bin Li, Qianyu Chen, Chengzhi Wang and Jenny Tieu for their generous help when I became an instructor for the first time.

I wish to express my gratitude to my friends in the department, Zhaohan Sun, Yechao Meng, Kecheng Li, Muqiao Huang, Xipeng Huang, Jingyue Huang, Xintong Li, Liyuan Zheng, Chi Zhang and Chi-Kuang Yeh, for every joyful moment we spent together.

I am also deeply indebted to Haodi Liang, Zhuoxuan Wu, Hexiao Huang, Ziwei Song, Yuliang Xu, Fangya Mao, Boxi Lin, Ce Yang, Mingren Yin, Tat Tin Yeung, Bowen Chen, Zhiyi Shen and Shan He for their everlasting friendship and timely encouragement despite the distance. My unique thanks go to the best roommate in the world, Yiming Xu, whose companionship on many late nights witnessed the completion of this thesis.

Finally, I would like to thank my mom for her braveness, fortitude and love.

Dedication

In memory of my grandparents and to my father

with hope and confidence

Table of Contents

1	Introduction	1
1.1	Results in Chapter 2	3
1.2	Results in Chapter 3	5
1.3	Results in Chapter 4	8
1.4	Results in Chapter 5	9
2	The roughness exponent and its model-free estimation	12
2.1	The roughness exponent: definition and existence	13
2.2	A martingale hidden in the Faber–Schauder expansion of a continuous function	19
2.3	The connection with weighted quadratic variation	26
2.4	The connection with self-similarity	31
2.5	The connection with Besov regularity	34
2.6	Application to fractional Brownian motion with drift	37
2.7	The roughness exponent without the reverse Jensen condition	40
2.8	Model-free estimation of the roughness exponent	44
2.9	The roughness exponent for unequally spaced partitions	51
2.10	Proofs for the results in Chapter 2	54
2.10.1	Proofs for the results in Section 2.4	54
2.10.2	Proofs for the results in Section 2.6	64
2.10.3	Proofs for the results in Section 2.9	74

3	Estimating the roughness exponent of stochastic volatility from discrete observations of the realized variance	78
3.1	Problem formulation and the estimator $\hat{\mathcal{R}}_n^*$	79
3.2	Strong consistency theorems	80
3.3	A scale-invariant estimator	83
3.4	Simulation study	85
3.5	Pathwise estimation	87
3.5.1	The rationale behind the estimator $\hat{\mathcal{R}}_n^*$	88
3.5.2	Pathwise consistency of $\hat{\mathcal{R}}_n^*(y)$	92
3.6	Proofs for the results in Chapter 3	101
3.6.1	Proof of Proposition 3.18	102
3.6.2	Proof of Theorem 3.7	108
4	Robust Faber–Schauder approximation based on discrete observations of an antiderivative	117
4.1	Problem formulation and an explicit formula	117
4.2	Error analysis for the Faber–Schauder coefficients	122
4.3	Error analysis for the approximating functions	125
4.4	Matrix analysis of the approximation problem	127
4.5	Proofs for the results in Chapter 4	144
5	The Φ-variation of classical fractal functions	154
5.1	Φ -variation of Takagi–van der Waerden and Weierstraß functions	155
5.2	Central limit theorem and Rademacher functions associated with fractal functions	157
5.3	A limit theorem for Bernoulli-type convolutions	163
5.4	Φ -variation of functions in the Takagi class	168
5.5	Proofs for the results in Chapter 5	174

6	Future Work	186
6.1	Dynamic detection of the roughness exponent	186
6.2	Estimating the roughness exponent of the volatility process from realized volatility	190
	References	198

Chapter 1

Introduction

Rough phenomena and rough trajectories are widely observed in mathematics and in nature. From a mathematical perspective, there are abundant sources of rough trajectories provided by classical fractal functions, such as the Weierstraß function and the Takagi function. Furthermore, many well-known stochastic processes generate sample paths with various degrees of roughness, such as the Brownian motion and the fractional Brownian motion.

Varying roughness behavior can also be widely observed in nature. For instance, the density of the carbon dioxide recorded in the ice core in Antarctica gives rise to rough paths [22]. In quantitative finance, the power-law decay of volatility skew in the options markets becomes a stylized fact since the very first report [11]. Subsequent investigations by Alòs et al. [3] and Fukasawa [36, 37] show that conventional diffusion volatility models fail to generate an implied volatility surface consistent with the power-law of volatility skew. Instead, the volatility process must contain a rough nature to fit the market observation. Furthermore, in the seminal work by Gatheral, Jaisson and Rosenbaum [40], it was discovered that the historic realized volatility of major financial indices is rougher than the sample paths of Brownian martingales. Furthermore, Jaisson and Rosenbaum [56, 57] also examined the formulation of the rough volatility from the microscopic structure. All the above observations motivate the recent development of rough volatility models in which the volatility process is modelled by a stochastic process with rougher trajectories.

Modelling these ‘irregular’ paths requires a quantitative description by rougher trajectories than sample paths of Brownian motions. Typically, such rough models are based on nonlinear functions of fractional Brownian motion with drift. The degree of roughness is determined by the Hurst parameter of the fractional Brownian motion, which was initially introduced to measure the long-range dependence of a time series. While the link between autocorrelation and roughness is well-established for stochastic processes such as fractional Brownian motion, it is by no means a universal truth. In general, the long-range behavior

and the degree of roughness bifurcate. Hence, we aim to study the concept of roughness with pathwise measures, particularly those without any probabilistic assumptions.

Our preference towards a non-parametric characterization and a model-free estimation of the roughness of trajectories, other than the well-known Hurst parameter scheme, is motivated by the following views:

- The roughness of a trajectory should be treated as a local and deterministic property, but not a global and probabilistic property.
- Data availability has significantly increased in recent decades. A model-independent characteristic offers broader applicability compared to a parametric one.
- Due to its greater robustness, a model-free estimation adapts to diverse data types and provides accurate results even under extreme scenarios.

With the above beliefs in mind, this thesis aims to investigate the concept of roughness and estimate the degree of roughness of financial time series in a strictly pathwise manner. To this end, we introduce the notion of a roughness exponent in Chapter 2. Furthermore, we discuss the connection between the roughness exponent and other regularity concepts. Several strongly consistent estimators are developed from the Gladyshev estimator $\hat{R}_n$. However, the above techniques cannot be directly applied to estimating the roughness exponent of volatility from the observations of realized volatility. In Chapter 3, we identify the challenges associated with this specific problem. As a partial response, we construct a new estimator $\hat{\mathcal{R}}_n^*$ that estimates the roughness exponent from the observation of the antiderivative of a function. This novel estimation scheme allows us to estimate the roughness exponent of the volatility process from discrete observations of the integrated volatility. Chapter 4 considers the problem of approximating the Faber–Schauder coefficients of a continuous function from the discrete observations of its antiderivative, which supports the validity of the construction of our estimator in Chapter 3. Chapter 5 studies the Wiener–Young Φ -variation of classical fractal functions, which can be regarded as a reinforced measure of the degree of roughness. Chapter 6 explores potential future research directions based on the findings of this thesis. In particular, we discuss methods to accommodate the proxy error between realized and integrated volatility so as to apply our estimator $\hat{\mathcal{R}}_n^*$ to estimate the roughness exponent of the volatility process directly from the realized volatility. Now, we will present each chapter of this thesis in greater detail.

1.1 Results in Chapter 2

The Hurst parameter was originally defined by Hurst [53] as a measure of the autocorrelation of a time series. But it is well known that it can also determine the degree of ‘roughness’ of a trajectory, e.g., in terms of the fractal dimension of its graph. Both notions are often used interchangeably in the literature. While the link between autocorrelation and roughness is well-established for stochastic processes such as fractional Brownian motion, it is by no means a universal truth. Indeed, Gneiting and Schlather [42] constructed a class of stationary Gaussian processes for which the Hurst parameter and fractal dimension decouple completely; see also Cheridito [12] and Bennedsen et al. [7]. It is therefore necessary to distinguish between the classical, autocorrelation-based Hurst parameter and a suitable index for the roughness of a trajectory. In this chapter, we study such a roughness index, which is based on the p^{th} variation of a continuous real-valued function.

Recall that, if $x : [0, 1] \rightarrow \mathbb{R}$ is a continuous function and $p > 0$, the p^{th} variation of x along the sequence of dyadic partitions is defined as the limit of

$$\langle x \rangle_n^{(p)} := \sum_{k=0}^{2^n-1} |x((k+1)2^{-n}) - x(k2^{-n})|^p,$$

provided that this limit exists. We will say that x admits the *roughness exponent* $R \in [0, 1]$ if

$$\lim_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \begin{cases} 0 & \text{for } p > 1/R, \\ \infty & \text{for } p < 1/R. \end{cases} \quad (1.1)$$

Intuitively, the smaller R is, the rougher the trajectory x will look. For instance, if x is continuously differentiable, then (1.1) holds with $R = 1$, and this is also the largest possible value for R unless x is constant. If x is a typical sample path of a continuous semimartingale such as Brownian motion, then (1.1) holds with $R = 1/2$. If x is a typical sample path of fractional Brownian motion, then R reduces to its classical Hurst parameter.

When measuring the roughness of a trajectory, there are strong reasons for favouring the p^{th} variation over other measures such as fractal dimension. First, the p^{th} variation can be measured in a straightforward manner. Second, other measures may lead to diverging results, and there is no canonical choice. For instance, the graph of a trajectory may have different Hausdorff and box dimensions. Third, and probably most importantly, the p^{th} variation plays a crucial role in the extension of Itô calculus to rough trajectories. A pathwise and strictly model-free version of such an extension for integrators with an arbitrary p^{th} variation was recently established by Cont and Perkowski [20]. Itô calculus for rough trajectories also plays an important role in applications, for instance, to rough volatility models. These models are

based on the observation by Gatheral et al. [40] that the Hurst parameter of the realized volatility of many financial time series is rather small, which makes realized volatility much rougher than the sample paths of a continuous semimartingale.

In Theorem 2.4, our main result, we establish the existence of the roughness exponent under a mild condition on the Faber–Schauder coefficients $\theta_{m,k}$ of x . We call this condition the reverse Jensen condition. Even more important than the existence of the roughness exponent is the fact that the reverse Jensen condition guarantees that the roughness exponent of x can be obtained as the limit of the classical Gladyshev estimates

$$\widehat{R}_n(x) = 1 - \frac{1}{n} \log_2 \sqrt{\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k}^2}. \quad (1.2)$$

This part of Theorem 2.4 can be viewed as a strong consistency result for the Gladyshev estimators in an entirely model-free setting, because no assumption whatsoever is made on the possible dynamics of the function x . In particular, we do not assume that x is a sample path of some stochastic process, even though this is clearly the typical application that we have in mind. While the statement of Theorem 2.4 appears to be deterministic, its proof is probabilistic. It relies on martingale techniques that are applied to a martingale hidden in the Faber–Schauder expansion of x . The corresponding theory is developed in Section 2.2.

The Gladyshev estimator has traditionally been used to estimate the Hurst parameter of stochastic processes such as fractional Brownian motion. So one conceptual contribution of Theorem 2.4 is the observation that the Gladyshev estimator actually estimates the roughness exponent and that its limit only delivers the classical Hurst parameter if and only if the latter coincides with the former. This is the case for fractional Brownian motion, which is the subject of Section 2.6. Our main result in this section shows in particular that the typical sample paths of fractional Brownian motion with drift satisfy the reverse Jensen condition. We also provide almost-sure convergence rates for the corresponding Gladyshev estimates. As an illustrating example, results in Section 2.6 apply to the fractional Ornstein–Uhlenbeck process.

The Gladyshev estimator was originally derived from Gladyshev’s theorem [41], which is an almost-sure limit theorem for the weighted quadratic variation of certain Gaussian processes. In Section 2.3, we systematically investigate the relations between the weighted quadratic variation and the roughness exponent. In particular, we show that for any continuous function, the existence of the weighted quadratic variation is strictly stronger than the convergence of the Gladyshev estimates. However, we also provide an example of a function for which the parameter derived from the weighted quadratic variation differs from

the roughness exponent. Despite that, we obtain a new estimator

$$\widehat{R}_n^* := 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \theta_{n,k}^2} \quad (1.3)$$

of the roughness exponent, which is the cornerstone in constructing the estimator for the roughness exponent of the volatility process in Chapter 3.

In Section 2.4, we also analyze trifractional Brownian motion to elucidate the differences between the self-similarity index and the roughness exponent. In addition, we derive limit theorems for the weighted quadratic variation of the trifractional Brownian motion and the n^{th} -order fractional Brownian motion to show that neither of the aforementioned notions can be obtained from the weighted quadratic variation.

Due to a result by Rosenbaum [85], the existence of the weighted quadratic variation is closely related to the concept of Besov regularity. In Section 2.5, we systematically analyze relations between variants of this concept and the existence of a roughness parameter. First, we observe that Eyink's formulation of Besov regularity [29] leads to a roughness concept that is generally different from ours. Then, we examine a more restrictive requirement, which is sometimes proposed in the rough volatility context. Our main result in this section, Theorem 2.28, links the corresponding roughness measure to p^{th} variations. However, Example 2.29 shows that even this concept of Besov regularity does not imply the existence of a roughness exponent.

In Section 2.8, we carefully discuss the problem of estimating the roughness exponent from discrete observations of a given function x . Here, the fact that the Gladyshev estimator is not scale-invariant can pose an issue. We therefore provide several families of scale-invariant estimators that are derived from the sequence $(\widehat{R}_n)_{n \in \mathbb{N}}$ without deteriorating the corresponding rate of convergence. We also relate the estimator used in [40] to those families. In Section 2.9, we provide an extension of our main results to the case in which the p^{th} variation is defined over a sequence of unequally spaced partitions.

1.2 Results in Chapter 3

In this chapter, we study the problem of estimating the roughness exponent of the volatility process $(\sigma_t)_{t \geq 0}$ in a stochastic volatility model. To this end, let us consider a stochastic volatility model whose price process satisfies

$$dS_t = \sigma_t S_t dB_t, \quad S_0 = s_0 > 0, \quad (1.4)$$

where B is a standard Brownian motion and σ_t is a progressively measurable stochastic process. A specific example suggested in [40] is to model the log volatility by a fractional Ornstein–Uhlenbeck process. Such a model is also known as the (rough) fractional stochastic volatility model. In this model, the ‘roughness’ of the volatility process is governed by the Hurst parameter of the fractional Ornstein–Uhlenbeck process, and it was pointed out in [40] that rather small values of Hurst parameter appear to be most adequate for capturing the stylized facts of empirical volatility of the financial time series. Since the publication of [40], many alternative rough volatility models have been proposed, e.g., the rough Heston model [27, 28] and the rough Bergomi model [6, 35, 55].

In Chapter 2, we have demonstrated that the roughness exponent can be estimated by the Gladyshev estimator $\hat{R}_n$ or the estimator $\hat{R}_n^*$. However, several difficulties hinder the direct application of the estimators $\hat{R}_n$ or $\hat{R}_n^*$ in estimating the roughness exponent of the volatility process.

The first difficulty is that in reality the volatility process σ_t cannot be observed directly; only the asset prices S are known. Thus, one typically computes the quadratic variation of the logarithmic stock prices,

$$\langle \log S \rangle_t = \int_0^t \sigma_s^2 ds, \quad (1.5)$$

which is also called the realized variance or the integrated volatility, and then performs numerical differentiation to estimate proxies $\hat{\sigma}_t$ for the actual values of σ_t . The roughness estimation is then based on those proxy values $\hat{\sigma}_t$. For instance, this two-step procedure is underlying the statistical analysis for empirical volatilities in [40], where roughness estimates were based on proxy values $\hat{\sigma}_t$ taken from the Oxford-Man Institute of Quantitative Finance Realized Library. A problem with that approach is that estimation errors in the proxy values $\hat{\sigma}_t$ might substantially distort the outcomes of the final roughness estimation; see Fukasawa et al. [38] and Cont and Das [19].

As a matter of fact, the quadratic variation (1.5) is usually approximated by a finite sum of the form $\sum_i (\log S_{t_i} - \log S_{t_{i-1}})^2$ based on discrete observations S_{t_i} of the price process. The bias caused by this error is emphasized in [38], where it is assumed that the approximation errors are log-normally distributed and independent of the Brownian motion B in (1.4), and a Whittle-type estimator for the Hurst parameter is developed based on quasi-likelihood. Another attempt to tackle this measurement error is made by Bolko et al. [10], where in a similar framework, the proposed estimator is based on the generalized method of moments approach. Chong et al. [14, 15] substantially extend the previous results by alleviating the assumption on proxy errors and basing the volatility model on a semi-parametric setup, in which, with the exception of the Hurst parameter of the underlying fractional Brownian motion, all components are fully non-parametric. One of the conclusions from [10, 38, 14]

is that the error arising from approximating the quadratic variation (1.5) with finite sum $\sum_i (\log S_{t_i} - \log S_{t_{i-1}})^2$ can be negligible when properly controlled. For this reason, we do not consider that error source in this chapter. A short discussion on how to accommodate this error source into our estimation scheme will be provided in Chapter 6.

Here, we construct a new estimator $\widehat{\mathcal{R}}_n^*$ for the roughness exponent of the volatility process σ_t that is based directly on discrete observations of the quadratic variation (1.5). Our estimator has a very simple form and can be computed with great efficiency on large data sets. It is not derived from distributional assumptions, as most other estimators in the literature, but from strictly pathwise considerations that were developed in Chapter 2 and Chapter 4. We believe that the fact that our estimator is built on a strictly pathwise approach makes it very versatile and applicable also in situations in which trajectories are not based on fractional Brownian motion. As a matter of fact, our Examples 3.10 and 3.13 illustrate that our estimation procedure can work very well for certain deterministic fractal functions.

The pathwise approach to constructing the estimator $\widehat{\mathcal{R}}_n^*$ can be outlined as follows. In Chapter 2, we state conditions under which a given trajectory $x \in C[0, 1]$ admits a roughness exponent R and we provide several estimators that approximate R , based on the Faber–Schauder expansion of x . In Chapter 4, we derive a robust method for estimating the Faber–Schauder coefficients of x for the situation in which only the antiderivative $y(t) = \int_0^t x(s) ds$, and not x itself, is observed on a discrete time grid. As explained in greater detail in Section 3.5.1, that method, when combined with the estimator $\widehat{R}_n^*$ as in (1.3), gives rise to the specific form of the estimator

$$\widehat{\mathcal{R}}_n^* := 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \vartheta_{n,k}^2},$$

where

$$\vartheta_{n,k} := 2^{3n/2+3} \left(y\left(\frac{4k}{2^{n+2}}\right) - 2y\left(\frac{4k+1}{2^{n+2}}\right) + 2y\left(\frac{4k+3}{2^{n+2}}\right) - y\left(\frac{4k+4}{2^{n+2}}\right) \right).$$

In Section 3.5.2, we formulate conditions on the trajectory x under which $\widehat{\mathcal{R}}_n^*(v)$ converges to the roughness exponent of x , resting on discrete observations of the function $v(t) = \int_0^t g(x(s)) ds$, where g is a generic, strictly monotone C^2 -function. In Section 3.6, we then verify that the aforementioned conditions on the trajectory x are satisfied by almost every sample path of fractional Brownian motion (with drift). This verification immediately yields the strong consistency of our estimator for the case in which the stochastic volatility is a nonlinear function of a fractional Brownian motion with drift. This includes, in particular,

the (rough) fractional stochastic volatility model.

Inherited from the estimator $\hat{R}_n^*$, one disadvantage of our original estimator $\hat{\mathcal{R}}_n^*$ is that it is not scale invariant. Using the idea from Chapter 2, we thus propose a scale-invariant modification of $\hat{\mathcal{R}}_n^*$ in Section 3.3. The subsequent Section 3.4 contains a simulation study illustrating the performance of our estimators. This study demonstrates that passing to the scale-invariant estimator can significantly improve the estimation accuracy in practice.

1.3 Results in Chapter 4

Suppose that we have discrete observations $\{F(t) : t \in \mathbb{T}_{n+1}\}$ of a function $F \in C^1[0, 1]$ for $\mathbb{T}_{n+1} := \{k2^{-n-1} : k = 0, \dots, 2^{n+1}\}$ and some $n \in \mathbb{N}$. We are interested in a robust reconstruction of Faber–Schauder coefficients of the derivative f from these observations. This problem naturally arises from the task of estimating the roughness exponent of the volatility process of a financial asset. However, we believe that this problem is of independent interest and has a large number of possible applications. In volatility estimation, the function f would be the square of the volatility process, but only an antiderivative of F is observable in the form of the realized quadratic variation of the asset. The quadratic variation is itself only observed at discrete, but typically equidistant, observation times.

Chapter 2 shows that the roughness exponent of f can be estimated by the estimator $\hat{R}_n^*$ as in (1.3), which requires the knowledge of the Faber–Schauder coefficients $\theta_{m,k}$ of f . To obtain these Faber–Schauder coefficients, the trajectory f needs to be observed directly. However, this is not the case in the context of estimating the roughness exponent of the volatility process. For this reason, we study here robust methods for estimating the Faber–Schauder coefficients of f directly from the observations $\{F(t) : t \in \mathbb{T}_{n+1}\}$. This method allows us to establish the estimator $\hat{\mathcal{R}}_n^*$ based on $\hat{R}_n^*$ to estimate the roughness exponent of the volatility process by replacing the coefficients $\theta_{n,k}$ with the coefficients $\vartheta_{n,k}$ in the corresponding estimators as in Chapter 3.

Suppose that we know the initial value $f(0)$ of the function f . Then estimating all Faber–Schauder coefficients of f up to and including generation n (i.e., all $\theta_{m,k}$ for $m = 0, \dots, n$ and $k = 0, \dots, 2^m - 1$) is equivalent to determining a continuous approximation $\hat{f}_n$ of f that is piecewise linear with supporting grid $\mathbb{T}_{n+1}$. It gives, in turn, rise to a piecewise quadratic antiderivative $\hat{F}_n$ and vice versa. This is the starting point for our analysis: we take a quadratic spline interpolation $\hat{F}_n$ of F with a given initial value $\hat{f}_0 = \hat{F}_n'(0)$ and a supporting grid $\mathbb{T}_{n+1}$, let $\hat{f}_n := \hat{F}_n'$, and then compute its Faber–Schauder coefficients $\vartheta_{m,k}^{(n)}$ for $m \leq n$. Our first result, Theorem 4.1, provides explicit formulas for all $\vartheta_{m,k}^{(n)}$ in terms of

the observed values of F . From these formulas and our subsequent error analysis, we can make three surprising observations.

First, only the coefficients in the final generation, i.e., the numbers $\vartheta_{n,0}^{(n)}, \dots, \vartheta_{n,2^n-1}^{(n)}$, depend on the initial value $\hat{f}_0$. Moreover, their sensitivity with respect to $\hat{f}_0$ grows by the factor $2^{n/2+2}$. All other Faber–Schauder coefficients are independent of $\hat{f}_0$. Thus, the infamous instability of the quadratic spline interpolation with respect to the initial value $\hat{f}_0$ arises exclusively from the coefficients in the final generation.

Second, each Faber–Schauder coefficient $\vartheta_{n,k}^{(n)}$ in the final generation depends in a highly non-local manner on all observations in the interval $[0, (k+1)2^{-n}]$. By contrast, all other coefficients depend only on observations in a local neighborhood.

The third observation is a consequence of our analysis of the approximation error between the estimated Faber–Schauder coefficients of $\hat{f}_n$ and the true Faber–Schauder coefficients of f . Our main results provide sharp upper and lower bounds of that error. These results imply in particular that F can always be chosen in such a way that the error in generation n exceeds the error up to generation $n-1$ by a factor of size $\mathcal{O}(2^n)$ for the ℓ_2 - and ℓ_1 -norms and by a factor of size $\mathcal{O}(2^{n/2})$ for the ℓ_∞ -norm.

These observations shed light on the classical topic of quadratic spline interpolation, as they locate the cause for the well-known instability of this method exclusively within the final generation of the Faber–Schauder coefficients. They also allow us to construct a robust estimator for the Faber–Schauder coefficients of f by simply dropping the final generation of coefficients. This truncated estimator is independent of the initial value $\hat{f}_0$ and avoids the counterintuitive non-local dependence explained above. Moreover, our error analysis yields robust bounds on the approximation error between the estimated and actual Faber–Schauder coefficients, which validates the construction of the estimator $\hat{\mathcal{R}}_n^*$ in Chapter 3.

1.4 Results in Chapter 5

The roughness exponent discussed in the previous chapters only provides a sketchy characteristic of the roughness of trajectories. To wit, the roughness exponent R gives information on the p^{th} variation for all p but not for the critical case $p = 1/R$. For a function with the roughness exponent R , there is no universal conclusion on p^{th} variation for $p = 1/R$. In the following, we study the roughness of fractal functions of the following form

$$f(t) := \sum_{m=0}^{\infty} \alpha_m \varphi(b^m t), \quad t \in [0, 1], \quad (1.6)$$

where $\{\alpha_m\} \in \ell_1$, $b \in \{2, 3, \dots\}$ and $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is $\mathbb{Z}$ periodic and Lipschitz continuous. Then, the series on the right-hand side converges absolutely and uniformly in $t \in [0, 1]$ so that f is indeed a well-defined continuous function. In this chapter, we study the Φ -variation over a sequence of refining partitions for functions at critical degrees of roughness.

Two separate cases are studied in this chapter; first, we consider $\alpha_m = \alpha^m$ for some $\alpha \in (-1, 1)$ in (1.6), and the case of critical roughness $\alpha = \pm 1/b$ is studied. For this case, it was shown by Schied and Zhang [89] that f has vanishing p^{th} variation for all $p > 1$ and $t \in [0, 1]$, and therefore $R = 1$. This, however, does not imply that f is of bounded variation. For instance, if φ is the tent map, $b = 2$, and $\alpha = 1/2$, then f is the classical Takagi function [92], which is nowhere differentiable and hence cannot be of bounded variation; a very short proof of this fact was given by de Rham [24] and later rediscovered by Billingsley [8]. For the Weierstraß function, the proof of nowhere differentiability for all $\alpha \in [1/b, 1)$ is more difficult. Starting from Weierstraß's original work [95], it attracted numerous authors until a definite result was given by Hardy [51].

It is therefore apparent that, in the critical case $|\alpha| = 1/b$, power variation is insufficient to capture the exact degree of roughness of the function f . To give a precise result on the roughness of the function f in the critical case, we search for a strictly increasing function $\Phi : [0, 1) \rightarrow [0, \infty)$ such that the limit

$$\langle f \rangle_t^\Phi := \lim_{n \uparrow \infty} \sum_{k=0}^{\lfloor tb^n \rfloor} \Phi(|f((k+1)b^{-n}) - f(kb^{-n})|)$$

exists and is non-trivial for any $t > 0$. This limit can also be regarded as the Wiener–Young Φ -variation of f (see, e.g., [4]), restricted to the sequence of b -adic partitions. Our results show that Takagi–van der Waerden and Weierstraß functions admit non-trivial and linear Φ -variation for $\Phi(x) = x/\sqrt{-\log x}$.

Furthermore, if $b = 2$ and φ is the tent map, f defines a function in the Takagi class. We consider the case where $\alpha_m = 2^{-m(1-q)}\sqrt{g(m)}$ for $q \in (0, 1)$ and some regularly varying function g . Following from [89], the function f admits the roughness exponent $R = 1 - q$. We demonstrate that functions of the above form have finite, nonzero and linear Φ_q -variation along the sequence of dyadic partitions where

$$\Phi_q(x) = x^{1/(\frac{1}{1-q})} g\left(\frac{-\log_2 x}{1-q}\right)^{-\frac{1}{2(1-q)}}.$$

Moreover, the limiting behavior of the regularly varying function g determines the $(1/R)^{\text{th}}$ variation of function f . For instance, if $g(n)$ converges to infinity, then f admits infinite $(1/R)^{\text{th}}$ variation; if $g(n)$ decays to zero, then f admits zero $(1/R)^{\text{th}}$ variation; finally, if g

is asymptotically constant, then f admits a nonzero $(1/R)^{\text{th}}$ variation. The collection of Φ_q taking the above representation forms a large class of increasing functions that are regularly varying at zero with index q . As a byproduct, we also prove a limiting theorem for certain sums of Bernoulli random variables.

Chapter 2

The roughness exponent and its model-free estimation

This chapter is based on [44, 45, 48] and is structured as follows. In Section 2.1, we introduce the roughness exponent for a continuous function x and discuss the existence of this notion. Theorem 2.4, our main result of this section, demonstrates that the Gladyshev estimates $\widehat{R}_n(x)$ will converge to the roughness exponent under the reverse Jensen condition. The proof of this statement is based on a martingale hidden in the Faber–Schauder expansion of x , which will be highlighted in Section 2.2. In Sections 2.3, 2.4 and 2.5, we systematically investigate the relations between the weighted quadratic variation, self-similarity, Besov regularity and the roughness exponent. Our results demonstrate that the roughness exponent differs from these existing measures. In Section 2.6, we show that the typical sample paths of fractional Brownian motion satisfy the reverse Jensen condition. In addition, we obtain the almost-sure convergence rate of the Gladyshev estimates for fractional Brownian motion with drift. As a byproduct, we provide a simple criterion on the drift coefficient under which the law of the fractional Brownian motion with drift is automatically equivalent to that of fractional Brownian motion. However, the Gladyshev estimator is not scale-invariant. As a result, multiplying x with a constant scaling factor can lead to negative or unreasonably large estimates for the roughness exponent. We resolve this issue in Section 2.8 by constructing several scale-invariant estimators, all of which are persistent to the scaling effect. Section 2.7 collects several additional results that are also valid if the reverse Jensen condition is not assumed. Finally, in Section 2.9, we extend our results to the case where the p^{th} variation is defined over a sequence of unequally spaced partitions. The proofs for the results in Sections 2.4, 2.6 and 2.9 are deferred to Section 2.10.

2.1 The roughness exponent: definition and existence

Let $x : [0, 1] \rightarrow \mathbb{R}$ be any fixed continuous function that is directly observable. This function can be a natural or economic time series, a typical sample path of a stochastic process, or a fractal function. What these phenomena have in common is that the corresponding trajectories are not smooth but exhibit a certain degree of ‘roughness’. In the sequel, our goal is to quantify and measure the degree of that roughness. To this end, we will henceforth exclude the trivial case of a constant function x .

For $p > 0$ and $n \in \mathbb{N}$, we define

$$\langle x \rangle_n^{(p)} := \sum_{k=0}^{2^n-1} |x((k+1)2^{-n}) - x(k2^{-n})|^p, \quad (2.1)$$

which can be regarded as the p^{th} variation of the function x sampled along the dyadic partition $\{k2^{-n} : k = 0, \dots, 2^n\}$. Suppose that there exists q such that

$$\lim_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \begin{cases} 0 & \text{for } p > q, \\ \infty & \text{for } p < q. \end{cases} \quad (2.2)$$

Intuitively, the larger q , the rougher the trajectory x will look. For instance, if x is continuously differentiable, then (2.2) holds with $q = 1$, and this is also the smallest possible value for q , because x is nonconstant by assumption. Moreover, it is easy to see that (2.2) holds if $\lim_n \langle x \rangle_n^{(p)}$ exists in $(0, \infty)$ (see, e.g., the final step in the proof of Theorem 2.1 in [77]). In particular, if x is a typical sample path of a continuous semimartingale such as Brownian motion, then (2.2) holds with $q = 2$. More generally, if x is a typical sample path of a fractional Brownian motion with Hurst parameter $H \in (0, 1)$, then q is equal to $1/H$; see Theorem 2.30 and the subsequent paragraph. The extreme case $q = \infty$ implies that $\langle x \rangle_n^{(p)}$ diverges to infinity for any $p \geq 1$; an example is provided in Example 2.9.

Definition 2.1. Suppose that there exists $q \in [1, \infty]$ such that (2.2) holds. Then $R := 1/q$ is called the *roughness exponent* of $x \in C[0, 1]$.

A concept closely related to the roughness exponent was used by Gatheral et al. [40] to quantify the roughness of empirical volatility time series. Moreover, in Chapter 5, the roughness exponent was computed for certain families of fractal functions. On the other hand, the standard Hurst parameter is defined via the autocorrelation of the time series and measures the amount of long-range dependence in the data. For many stochastic processes, including fractional Brownian motion, the standard Hurst parameter coincides with the

roughness exponent of the sample paths. In general, however, these two parameters may be different; for a discussion, see Gneiting and Schlather [42], where roughness is measured in terms of fractal dimension. Later, Bennedsen et al. [7] adopted this idea to study the bifurcation of the short- and long-term behavior of stochastic volatility with the goal of proposing a new class of stochastic volatility models that can simultaneously incorporate roughness and slowly decaying autocorrelation. Finally, the concept of Besov regularity, introduced by Eyink [29] within the framework of Besov spaces, is somewhat related to our roughness exponent. The corresponding connections are explored in Section 2.5.

Note that if x admits the roughness exponent R , then (2.2) does not make any assertion on the existence of $\lim_n \langle x \rangle_n^{(p)}$ for $p := 1/R$. This limit always exists for $p = 1$ and is equal to the total variation of x . For $p > 1$, however, the limit may or may not exist (see Example 2.18). If it does exist, it can be interpreted as the p^{th} variation of the continuous function x along the sequence of dyadic partitions of $[0, 1]$. This p^{th} variation can be used so as to prove a strictly pathwise version of Itô calculus, an approach that was pioneered by Föllmer [31] for $p \leq 2$ and recently extended to $p > 2$ by Cont and Perkowski [20]. Just as this notion of p^{th} variation, the roughness exponent of a trajectory will typically depend on the underlying partition sequence. As noted above, we have chosen the dyadic partitions here because equally-spaced partitions are natural for many data sources and also lead to the simplest and most elegant mathematical statements. An extension to unequally spaced partitions is provided in Section 2.9.

There exist continuous functions that do not admit a roughness exponent; see Example 2.29. For this reason, we will first establish criteria for the existence of R . Let us start by defining

$$q_* := \sup \left\{ p > 0 : \liminf_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \infty \right\} \quad \text{and} \quad q^* := \sup \left\{ p > 0 : \limsup_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \infty \right\}. \quad (2.3)$$

The proof of Proposition 2.2 will show that the following alternative formulas hold for q_* and q^* ,

$$q_* = \inf \left\{ p \geq 1 : \liminf_{n \uparrow \infty} \langle x \rangle_n^{(p)} = 0 \right\} \quad \text{and} \quad q^* = \inf \left\{ p \geq 1 : \limsup_{n \uparrow \infty} \langle x \rangle_n^{(p)} = 0 \right\}. \quad (2.4)$$

Proposition 2.2. *The roughness exponent of $x \in C[0, 1]$ exists if and only if $q^* = q_*$. In this case, R is given by $R = 1/q^* = 1/q_*$.*

Proof. Clearly, the roughness exponent R is nonnegative by definition. Moreover, since the function x is non-constant by assumption, we have $\langle x \rangle_n^{(p)} \rightarrow \infty$ for all $p < 1$. Hence, we must have $R \leq 1$.

The remaining assertion follows easily from the alternative representation (2.4) of q_* and q^* , so it is sufficient to establish (2.4). To prove this result for the case $q_* < \infty$, let $q_0 := \inf\{p \geq 1 : \liminf_{n \uparrow \infty} \langle x \rangle_n^{(p)} = 0\}$. Let n_0 be such that $\max_k |x((k+1)2^{-n}) - x(k2^{-n})| \leq 1$ for $n \geq n_0$. Then $p \mapsto \langle x \rangle_n^{(p)}$ is non-increasing for $n \geq n_0$, and so we must have $q_* \leq q_0$. To show the converse inequality, we show that $\liminf_n \langle x \rangle_n^{(p)} = 0$ for every $p > q_*$. To this end, we assume by way of contradiction that there exists $p > q_*$ such that $v := \liminf_n \langle x \rangle_n^{(p)} > 0$. The definition of q_* implies that we must have $v < \infty$. Next, for any $\varepsilon > 0$, there exists $n_1 \in \mathbb{N}$ such that $\max_k |x((k+1)2^{-n}) - x(k2^{-n})| \leq \varepsilon$ for $n \geq n_1$. Thus, for $q \in (q_*, p)$ and $n \geq n_1$,

$$\langle x \rangle_n^{(p)} \leq \max_k |x((k+1)2^{-n}) - x(k2^{-n})|^{p-q} \cdot \langle x \rangle_n^{(q)} \leq \varepsilon^{p-q} \langle x \rangle_n^{(q)}.$$

When sending $\varepsilon \downarrow 0$, we see that $0 < v \leq \varepsilon^{p-q} \liminf_n \langle x \rangle_n^{(q)}$ for all $\varepsilon > 0$. This implies that $\liminf_n \langle x \rangle_n^{(q)} = \infty$, which, in view of $q > q_*$, contradicts the definition of q_* . This establishes $q_0 = q_*$. If $q_* = \infty$, then for each $p > 0$, we must have $\liminf_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \infty$. Thus, $\{p \geq 1 : \liminf_{n \uparrow \infty} \langle x \rangle_n^{(p)} = 0\} = \emptyset$, which establishes the first identity in (2.4) for the case $q_* = \infty$. The second identity in (2.4) is proved in the same manner. $\square$

Our analysis of the roughness exponent is based on the Faber–Schauder wavelet expansion of continuous functions. Recall that the *Faber–Schauder functions* are defined as

$$e_\emptyset(t) := t, \quad e_{0,0}(t) := (\min\{t, 1-t\})^+, \quad e_{m,k}(t) := 2^{-m/2} e_{0,0}(2^m t - k) \quad (2.5)$$

for $t \in \mathbb{R}$, $m \in \mathbb{N}$ and $k \in \mathbb{Z}$. It is well known that the restriction of the Faber–Schauder functions to $[0, 1]$ form a Schauder basis for $C[0, 1]$. More precisely, every function $x \in C[0, 1]$ can be uniquely represented by the following uniformly convergent series,

$$x = x(0) + (x(1) - x(0)) e_\emptyset + \sum_{m=0}^{\infty} \sum_{k=0}^{2^m-1} \theta_{m,k} e_{m,k}, \quad (2.6)$$

where the *Faber–Schauder coefficients* $\theta_{m,k}$ are given by

$$\theta_{m,k} = 2^{m/2} \left(2x\left(\frac{2k+1}{2^{m+1}}\right) - x\left(\frac{k}{2^m}\right) - x\left(\frac{k+1}{2^m}\right) \right).$$

We furthermore denote for $n \in \mathbb{N}$,

$$s_n := \sqrt{\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k}^2}, \quad \text{and} \quad \widehat{R}_n(x) := 1 - \frac{1}{n} \log_2 s_n. \quad (2.7)$$

Our main thesis in this chapter is that the roughness exponent of x should be equal to the limit of $\widehat{R}_n(x)$, provided this limit exists. Theorem 2.4 will rigorously establish this result under the following mild condition on the Faber–Schauder coefficients of x . To formulate this and related conditions, we will say that a function $\varrho : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is *subexponential* if $\frac{1}{t} \log \varrho(t) \rightarrow 0$ as $t \uparrow \infty$.

Definition 2.3. The Faber–Schauder coefficients of x satisfy the *reverse Jensen condition* if for each $p \geq 1$ there exist $n_p \in \mathbb{N}$ and a non-decreasing subexponential function $\varrho_p : \mathbb{N} \rightarrow [1, \infty)$ such that

$$\frac{1}{\varrho_p(n)} s_n^p \leq \frac{1}{2^{n-1}} \sum_{k=0}^{2^{n-1}-1} \left(\sum_{m=0}^{n-1} 2^m \theta_{m, [2^{m-n+1}k]}^2 \right)^{p/2} \leq \varrho_p(n) s_n^p \quad \text{for } n \geq n_p. \quad (2.8)$$

If $p \geq 2$, then Jensen’s inequality implies that the left-hand inequality in (2.8) holds with $\varrho_p = 1$. Likewise, the upper bound holds with $\varrho_p = 1$ if $p \leq 2$. So only the other cases are nontrivial, which explains our terminology “reverse Jensen condition”. This terminology will become even more apparent in Remark 2.12, where an equivalent probabilistic formulation is given. As we will see in Theorem 2.30, the reverse Jensen condition is satisfied for the typical sample paths of fractional Brownian motion with arbitrary Hurst parameter. Moreover, two alternative conditions that are easier to verify, but are also stronger, will be provided in Proposition 2.15. Now we can state our first main result, which establishes the existence of the roughness exponent and its relation to the limit of the sequence $\widehat{R}_n(x)$ under the assumption that the reverse Jensen condition holds. The second part of Theorem 2.4 states that condition (2.8) at $p = 1/R$ is necessary for the convergence of $\widehat{R}_n(x)$ to the roughness exponent R .

Theorem 2.4. *Suppose that the Faber–Schauder coefficients of x satisfy the reverse Jensen condition. Then the function x admits a roughness exponent R if and only if the finite limit $\lim_n \widehat{R}_n(x)$ exists, and in this case, we have $\widehat{R}_n(x) \rightarrow R$. Conversely, suppose that $\widehat{R}_n(x)$ converges to the roughness exponent R of x . Then (2.8) holds at $p = 1/R$.*

Theorem 2.4 will be proved in the subsequent Section 2.2. It can be regarded as a model-free consistency result for the estimator $\widehat{R}_n$ of the roughness exponent of the continuous function x . It has the following two aspects.

First, Theorem 2.4 is model-free in the sense that no assumptions are made on x except the reverse Jensen condition. In particular, we do not assume that x is a realization of a stochastic process with prescribed dynamics, even though this is clearly the typical application we have in mind. This line of thought is similar in spirit to the model-free approach to

stochastic analysis, which was pioneered by Föllmer [31] and Lyons [70]. In this approach, one derives results for deterministic trajectories satisfying certain sample path properties, thus avoiding other implicit assumptions (e.g., a Lévy-type modulus of continuity or a law of the iterated logarithm) usually imposed through fixing a probabilistic model. The results can then be applied to all probabilistic models whose trajectories satisfy the assumed sample path properties almost surely. In our analysis of fractional Brownian motion with drift (Section 2.6), we follow precisely this recipe. We show that almost all of the sample paths satisfy the reverse Jensen condition, and we identify the Hurst parameter H as the common roughness exponent of the trajectories. We also provide almost sure convergence rates for the convergence of $\widehat{R}_n$ to H .

Second, based on the work by Gladyshev [41], it has long been known that $\widehat{R}_n$, which is sometimes called the *Gladyshev estimator*, estimates the Hurst parameter for fractional Brownian motion and other Gaussian processes; see, e.g., [65] for an overview. Theorem 2.4 now clarifies that $\widehat{R}_n$ is actually an estimator for the roughness exponent; it estimates the classical Hurst parameter H only if H coincides with the roughness exponent. Our theorem thus demonstrates that the Gladyshev estimator has, in fact, a much wider scope of possible applications, provided that it is correctly understood as an estimator of the roughness exponent and not of the Hurst parameter. The same is true for certain variants of the Gladyshev estimator, such as the scale-invariant modifications discussed in Section 2.8.

The statement of Theorem 2.4 may become false if the reverse Jensen condition is not satisfied; see Proposition 2.19 for a counterexample. However, as shown by the next result, an exception occurs for functions x that display diffusive behavior in the sense that they admit the roughness exponent $R = 1/2$. This includes, in particular, the typical sample paths of any continuous semimartingale with nontrivial martingale component.

Corollary 2.5. *Let $x \in C[0, 1]$ be any function with roughness exponent $1/2$. Then $\lim_n \widehat{R}_n(x) = 1/2$.*

The preceding corollary will follow immediately from Proposition 2.35. The following remark outlines what can be shown if the reverse Jensen condition holds but $\widehat{R}_n(x)$ does not converge to a limit.

Remark 2.6. Suppose that the reverse Jensen condition holds and let us denote

$$r^+ := \limsup_{n \uparrow \infty} \widehat{R}_n(x) \quad \text{and} \quad r^- := \liminf_{n \uparrow \infty} \widehat{R}_n(x). \quad (2.9)$$

One can show as in the proof of Theorem 2.4 that

$$\limsup_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \begin{cases} 0 & \text{if } p > 1/r^-, \\ \infty & \text{if } p < 1/r^-, \end{cases} \quad \liminf_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \begin{cases} 0 & \text{if } p > 1/r^+, \\ \infty & \text{if } p < 1/r^+. \end{cases}$$

To wit, under the reverse Jensen condition, we have $q^* = 1/r^-$ and $q_* = 1/r^+$. By Proposition 2.2, the roughness exponent is well defined if and only if $r^+ = r^-$, and this special case is discussed in Theorem 2.4; on the other hand, if $r^+ > r^-$, the roughness exponent does not exist, and the sequence $(\langle x \rangle_n^{(p)})_{n \in \mathbb{N}_0}$ will fluctuate between zero and infinity for any $p \in (1/r^+, 1/r^-)$.

Remark 2.7. Suppose that $x \in C[0, 1]$ is such that its Faber–Schauder coefficients $\theta_{m,k}$ are the same for all k . Then, the reverse Jensen condition is trivially satisfied. These functions form the so-called Takagi class [52], and, beginning with the subsequent Examples 2.8 and 2.9, several of our examples and counterexamples will belong to this class. The two Examples 2.8 and 2.9 will in particular present for every $R \in [0, 1]$ a function x that satisfies the reverse Jensen condition and has roughness exponent R .

Example 2.8. For $\alpha \in (-\sqrt{2}, \sqrt{2})$, consider the function with Faber–Schauder coefficients $\theta_{m,k} = \alpha^m$. These functions are often called the Takagi–Landsberg functions. For $|\alpha| > 1/\sqrt{2}$, it follows from [77, Theorem 2.1] that their roughness exponent is given by $R = \frac{1}{2} - \log_2 |\alpha|$. Our Theorem 2.4 provides a very short proof of this fact: since

$$s_n^2 = \sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \alpha^{2m} = \frac{(2\alpha^2)^n - 1}{2\alpha^2 - 1},$$

we have $\frac{1}{n} \log_2 s_n \rightarrow \frac{1}{2} \log_2(2\alpha^2) = \frac{1}{2} + \log_2 |\alpha|$, which yields the claim. For $|\alpha| < 1/\sqrt{2}$, we get in the same way that $s_n^2 \rightarrow (1 - 2\alpha^2)^{-1}$. It will be shown in Proposition 2.37 (a) that x must hence be of bounded variation. Finally, in the case $|\alpha| = 1/\sqrt{2}$, we have $s_n^2 = n$, which also leads to $R = 1$. However, in this case, x is the classical Takagi function, which is nowhere differentiable and hence not of bounded variation; see Remark 2.2 (i) in [89] for details.

On the other extreme end of the spectrum are functions with roughness exponent $R = 0$. Such a function is constructed in the following example.

Example 2.9. Consider the function with Faber–Schauder coefficients $\theta_{m,k} = 2^{m/2}(m+1)^{-2}$. One can easily check that this choice gives rise to a well-defined continuous function. It is easy to show that $\langle x \rangle_n^{(p)} \rightarrow \infty$ for all $p \geq 1$. It follows that q_* and q^* , as defined in (2.3), are both infinite, and so Proposition 2.2 yields $R = 0$.

In the following Section 2.2, we present the proof of Theorem 2.4. It is based on a martingale that is hidden in the Faber–Schauder expansion of a continuous function x .

2.2 A martingale hidden in the Faber–Schauder expansion of a continuous function

In this section, we use martingale techniques to develop the probabilistic tools for our analysis of the roughness exponent, and we then prove Theorem 2.4. The idea of using a probabilistic approach to the p^{th} variation of functions goes back to [77, 89], where it was used for certain fractal functions. For a general function $x \in C[0, 1]$, however, the tools from [77, 89] are no longer sufficient; see Remark 2.11.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting an i.i.d. sequence $(U_n)_{n \in \mathbb{N}}$ of $\{0, 1\}$ -valued random variables with symmetric Bernoulli distribution. Furthermore, we define the stochastic processes

$$V_0 := 0, \quad V_m := \sum_{k=1}^m 2^{-k} U_k \quad \text{and} \quad S_m := 2^m \theta_{m, 2^m V_m}^2, \quad m \in \mathbb{N}_0. \quad (2.10)$$

Note that V_m is uniformly distributed on $\{k2^{-m} : k = 0, \dots, 2^m - 1\}$.

We next define the filtration $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_n := \sigma(U_1, \dots, U_n)$ for $n \in \mathbb{N}$. Since $U_1, U_2, \dots, U_n$ can be uniquely recovered from V_n , we also have $\mathcal{F}_n = \sigma(V_n)$. Let $Y_n := 2^{n/2} \theta_{n, 2^n V_n} \cdot (1 - 2U_{n+1})$. As $(1 - 2U_m)$ defines an i.i.d. sequence of centered random variables, and $\theta_{m, 2^m V_m}$ is $\mathcal{F}_m$ -measurable, the sequence

$$X_n := \sum_{m=0}^{n-1} Y_m = \sum_{m=0}^{n-1} 2^{m/2} \theta_{m, 2^m V_m} \cdot (1 - 2U_{m+1}), \quad n \in \mathbb{N},$$

is an $(\mathcal{F}_n)$ -martingale. The link between this martingale and the quantities $\langle x \rangle_n^{(p)}$ is established in the following proposition, which provides the key to the entire analysis in this section.

Proposition 2.10. *Suppose that $x(0) = x(1) = 0$. Then, for $p \geq 1$ and $n \in \mathbb{N}$,*

$$\langle x \rangle_n^{(p)} = 2^{n(1-p)} \mathbb{E}[|X_n|^p]. \quad (2.11)$$

Moreover, there exist constants $0 < A_p \leq B_p < \infty$ depending only on p but not on x , such

that for all $n \in \mathbb{N}$,

$$A_p 2^{n(1-p)} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \leq \langle x \rangle_n^{(p)} \leq B_p 2^{n(1-p)} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right]. \quad (2.12)$$

Furthermore, in the special case $p = 2$, we have

$$\langle x \rangle_n^{(2)} = 2^{-n} s_n^2. \quad (2.13)$$

The formula (2.13) was first obtained by Gantert [39] via a different proof. It can be used to simplify the computation of s_n and $\hat{R}_n(x)$.

Remark 2.11. Suppose that x belongs to the Takagi class [52], and $\theta_{m,k} = \alpha_m$ for a sequence of real numbers (α_m) for which the series $\sum_{m=0}^{\infty} 2^{m/2} \alpha_m$ converges absolutely. In particular, if $\alpha_m = \alpha^m$ for some $\alpha \in (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, the function x is referred to as the Takagi–Landsberg function. The roughness exponent of the Takagi–Landsberg function was studied in [77, 89]. Those results are later extended to functions in the Takagi class in our subsequent Chapter 5. In this case, $Y_n = 2^{n/2} \alpha_n (1 - 2U_{n+1})$ defines a sequence of independent centered random variables, and the sequence (X_n) remains an $(\mathcal{F}_n)$ -martingale. The independence of (Y_n) allows the application of Khintchine’s inequality, which directly gives

$$\begin{aligned} A_p 2^{n(1-p/2)} \left(\langle x \rangle_n^{(2)} \right)^{p/2} &= A_p 2^{n(1-p)} (s_n^2)^{p/2} \leq \langle x \rangle_n^{(p)} \\ &\leq B_p 2^{n(1-p)} (s_n^2)^{p/2} = B_p 2^{n(1-p/2)} \left(\langle x \rangle_n^{(2)} \right)^{p/2}, \end{aligned}$$

where A_p and B_p are constants with $0 < A_p \leq B_p < \infty$, depending only on p but not on x . The analysis in Chapter 5 will then proceed from the above inequality. However, for a general function $x \in C[0, 1]$, the random variables (Y_n) may no longer be independent. Thus, the techniques in Chapter 5 fail to apply to the scope of this chapter.

Remark 2.12. In this remark, we collect several alternative formulations of the reverse Jensen condition. Note that

$$\mathbb{E} \left[\sum_{m=0}^{n-1} S_m \right] = s_n^2 \quad \text{and} \quad \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right] = \frac{1}{2^{n-1}} \sum_{k=0}^{2^{n-1}-1} \left(\sum_{m=0}^{n-1} 2^m \theta_{m, \lfloor 2^{m-n+1} k \rfloor}^2 \right)^{p/2}. \quad (2.14)$$

We can thus rephrase the reverse Jensen condition as follows: there exist $n_p \in \mathbb{N}$ and a non-decreasing subexponential function $\varrho_p : \mathbb{N}_0 \rightarrow [1, \infty)$ such that

$$\frac{1}{\varrho_p(n)} \mathbb{E} \left[\sum_{m=0}^{n-1} S_m \right]^{p/2} \leq \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \leq \varrho_p(n) \mathbb{E} \left[\sum_{m=0}^{n-1} S_m \right]^{p/2} \quad \text{for } n \geq n_p. \quad (2.15)$$

Moreover, Proposition 2.10 shows that the reverse Jensen condition can be reformulated by means of p^{th} variations: for each $p \geq 1$, there exist an $n_p \in \mathbb{N}$ and a non-decreasing subexponential function $\varrho_p : \mathbb{N}_0 \rightarrow [1, \infty)$ such that

$$\frac{1}{\varrho_p(n)} 2^{n(1-p/2)} \left(\langle x \rangle_n^{(2)} \right)^{p/2} \leq \langle x \rangle_n^{(p)} \leq \varrho_p(n) 2^{n(1-p/2)} \left(\langle x \rangle_n^{(2)} \right)^{p/2} \quad \text{for } n \geq n_p. \quad (2.16)$$

Comparing (2.16) with the same inequality for a different value $q \geq 1$ gives the following more general equivalent statement of the reverse Jensen condition: for every $1 \leq p, q < \infty$, there exists $n_{p,q} \in \mathbb{N}$ and a non-decreasing subexponential function $\varrho_{p,q} : \mathbb{N}_0 \rightarrow [1, \infty)$ such that

$$\frac{1}{\varrho_{p,q}(n)} 2^{n(1-p/q)} \left(\langle x \rangle_n^{(q)} \right)^{p/q} \leq \langle x \rangle_n^{(p)} \leq \varrho_{p,q}(n) 2^{n(1-p/q)} \left(\langle x \rangle_n^{(q)} \right)^{p/q} \quad \text{for } n \geq n_{p,q}. \quad (2.17)$$

If $p \geq q$, the left-hand inequality in (2.17) holds with $\varrho_{p,q} = 1$. Likewise, if $p \leq q$, then the right-hand inequality in (2.17) holds with $\varrho_{p,q} = 1$, which can also be understood as the well-known inequality for generalized means.

We need the following lemma for the proof of Proposition 2.10.

Lemma 2.13. *For $n \in \mathbb{N}$ and $m \leq n - 1$, if t is a dyadic rational number in $[0, 1)$ with binary expansion $t = \sum_{j=1}^n 2^{-j} a_j$ for $a_j \in \{0, 1\}$, we have $e_{m,k}(t + 2^{-n}) - e_{m,k}(t) = 2^{m/2-n} (1 - 2a_{m+1}) \mathbb{1}_{\{\lfloor 2^m t \rfloor = k\}}$.*

Proof. For $m \leq n - 1$ and $\psi := \mathbb{1}_{[0, 1/2)} - \mathbb{1}_{[1/2, 1)}$, we have

$$\begin{aligned} e_{m,k}(t + 2^{-n}) - e_{m,k}(t) &= 2^{-m/2} (e_{0,0}(2^m t + 2^{m-n} - k) - e_{0,0}(2^m t - k)) \\ &= 2^{m/2-n} \psi(2^m t - k), \end{aligned} \quad (2.18)$$

Furthermore, it follows that

$$\begin{aligned} \psi(2^m t - k) &= \psi \left(\sum_{j=1}^n 2^{m-j} a_j - k \right) = \psi \left(\sum_{j=m+1}^n 2^{m-j} a_j - \left(k - 2^m \sum_{j=1}^m 2^{-j} a_j \right) \right) \\ &= \psi \left(\frac{1}{2} a_{m+1} + \sum_{j=m+2}^n 2^{m-j} a_j - (k - \lfloor 2^m t \rfloor) \right). \end{aligned}$$

As $k - \lfloor 2^m t \rfloor \in \mathbb{Z}$ and $\sum_{j=m+1}^n 2^{m-j} a_j \in [0, 1)$, we have $\psi(2^m t - k) \neq 0$ if and only if $\lfloor 2^m t \rfloor = k$. Furthermore, since $\sum_{j=m+2}^n 2^{m-j} a_j \in [0, 1/2)$ and $a_{m+1}/2 \in \{0, 1/2\}$, we must

have

$$\psi \left(\frac{1}{2}a_{m+1} + \sum_{j=m+2}^n 2^{m-j}a_j \right) = \psi \left(\frac{1}{2}a_{m+1} \right) = 1 - 2a_{m+1}.$$

Hence, $\psi(2^m t - k) = (1 - 2a_{m+1}) \mathbb{1}_{\{k=\lfloor 2^m t \rfloor\}}$. Substituting this identity back into (2.18) completes the proof. $\square$

Proof of Proposition 2.10. For $n \in \mathbb{N}_0$, let the n^{th} truncation x_n of x be given by

$$x_n(t) := \sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k} e_{m,k}(t). \quad (2.19)$$

Since the Faber–Schauder functions $e_{m,k}$ vanish on $\{k2^{-n} : k = 0, \dots, 2^n\}$ for $m \geq n$, we have $x(k2^{-n}) = x_n(k2^{-n})$ for $k = 0, \dots, 2^n$. Using the fact that V_n is uniformly distributed on $\{k2^{-n} : k = 0, \dots, 2^n - 1\}$ and Lemma 2.13, we get

$$\begin{aligned} \langle x \rangle_n^{(p)} &= 2^n \mathbb{E} [|x_n(V_n + 2^{-n}) - x_n(V_n)|^p] \\ &= 2^n \mathbb{E} \left[\left| \sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k} (e_{m,k}(V_n + 2^{-n}) - e_{m,k}(V_n)) \right|^p \right] \\ &= 2^n \mathbb{E} \left[\left| \sum_{m=0}^{n-1} 2^{m/2-n} \sum_{k=0}^{2^m-1} \theta_{m,k} (1 - 2U_{m+1}) \mathbb{1}_{\{k=2^m V_m\}} \right|^p \right] \\ &= 2^{n(1-p)} \mathbb{E} \left[\left| \sum_{m=0}^{n-1} 2^{m/2} \theta_{m,2^m V_m} (1 - 2U_{m+1}) \right|^p \right] = 2^{n(1-p)} \mathbb{E} [|X_n|^p], \end{aligned} \quad (2.20)$$

which yields (2.11). Therefore, the Burkholder inequality implies the existence of constants $0 < A_p \leq B_p < \infty$ depending only on p such that

$$A_p 2^{n(1-p)} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} Y_m^2 \right)^{p/2} \right] \leq \langle x \rangle_n^{(p)} \leq B_p 2^{n(1-p)} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} Y_m^2 \right)^{p/2} \right]. \quad (2.21)$$

Moreover, as $|1 - 2U_{m+1}| = 1$, we have

$$Y_m^2 = \left(2^{m/2} \theta_{m,2^m V_m} \right)^2 (1 - 2U_{m+1})^2 = 2^m \theta_{m,2^m V_m}^2 = S_m. \quad (2.22)$$

This completes the proof of (2.12) for arbitrary $p > 1$. In the case $p = 2$, the martingale differences (Y_m) clearly satisfy $\mathbb{E}[\sum_{m=0}^{n-1} Y_m^2] = \mathbb{E}[(\sum_{m=0}^{n-1} Y_m)^2]$, which yields (2.22) by way of

(2.20), (2.21), and (2.22). This completes the proof. $\square$

Now, we move toward the proof of Theorem 2.4. The following auxiliary lemma will be needed in that proof.

Lemma 2.14. *If the finite limit $\lim_n \widehat{R}_n(x)$ exists, then it belongs to $[0, 1]$.*

Proof. Let $\xi_n := 1 - \widehat{R}_n(x) = \frac{1}{n} \log_2 s_n$. It suffices to show that $\xi := \lim_n \xi_n = 1 - \lim_n \widehat{R}_n(x)$ belongs to $[0, 1]$. First, let us assume by way of contradiction that $\xi < 0$. Thus, there exists $n_1 \in \mathbb{N}$ such that for $n \geq n_1$, we have $\xi_n \leq -n\xi/2$ and $s_n \leq 2^{-n\xi/2}$. However, this contradicts the fact that $(s_n)_{n \in \mathbb{N}}$ is a non-negative and non-decreasing sequence. Therefore, we must have $\xi \geq 0$.

Next, we assume by way of contradiction that $\xi \in (1, \infty)$. For each $n \in \mathbb{N}$, we take $\lambda(n) := 2^{-2n\xi} s_n^2$. Then λ is a subexponential function, and $\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k}^2 = \lambda(n) 2^{2n\xi}$. It then follows that $\sum_{k=0}^{2^n-1} \theta_{n,k}^2 = \lambda(n+1) 2^{2\xi(n+1)} - \lambda(n) 2^{2n\xi}$, and $\sup_{0 \leq k \leq 2^n-1} |\theta_{n,k}| \geq 2^{(\xi-1/2)n} \sqrt{\lambda(n+1) 2^{2\xi} - \lambda(n)}$. As in (2.19), let x_n denote the n^{th} truncation of x , then

$$\max_{0 \leq t \leq 1} |x_{n+1} - x_n| = 2^{-n/2} \sup_{0 \leq k \leq 2^n-1} |\theta_{n,k}| \geq 2^{(\xi-1)n} \sqrt{\lambda(n+1) 2^{2\xi} - \lambda(n)},$$

Since λ is a subexponential function, the rightmost term tends to infinity as $n \uparrow \infty$. Thus, the series of truncated functions $(x_n)_{n \in \mathbb{N}}$ will not converge. This contradicts the uniform convergence of x_n to x . Thus, we must have $\xi \leq 1$, and this completes the proof. $\square$

Proof of Theorem 2.4. By the argument used in the proof of Lemma 3.1 of [77], we may assume without loss of generality that $x(0) = x(1) = 0$. We start by proving the “only if” direction in the first part of the assertion and assume that the function x admits the roughness exponent R . For simplicity, we use the short-hand notion $r_n := \widehat{R}_n(x)$ in this and the following proofs of this chapter. We must prove that $r_n \rightarrow R$. For any $p > 1/R \geq 1$, Proposition 2.10 and condition (2.15) imply that there exists $A_p > 0$ such that

$$\begin{aligned} A_p 2^{n(1-pr_n - \log_2 \varrho_p(n)/n)} &= A_p \varrho_p^{-1}(n) 2^{n(1-p)} s_n^p \\ &= A_p \varrho_p^{-1}(n) 2^{n(1-p)} \mathbb{E} \left[\sum_{m=0}^{n-1} S_m \right]^{p/2} \leq \langle x \rangle_n^{(p)}. \end{aligned} \tag{2.23}$$

Since $\limsup_n \langle x \rangle_n^{(p)} = 0$, we must thus have $\liminf_n r_n > 1/p$. Sending $p \downarrow 1/R$ now gives $\liminf_n r_n \geq R$. In the same way, one proves that $\limsup_n r_n \leq R$.

To prove the “if” direction, let us assume that the finite limit $\lim_n r_n = r$ exists. Then $r \in [0, 1]$ by Lemma 2.14, and this yields

$$\lim_{n \uparrow \infty} n \left(1 - pr_n \pm \frac{1}{n} \log_2 \varrho_p(n) \right) = \begin{cases} +\infty & \text{if } p < 1/r, \\ -\infty & \text{if } p > 1/r. \end{cases}$$

Applying condition (2.15) to (2.12) and taking $n \uparrow \infty$ implies that $\lim_n \langle x \rangle_n^{(p)} = \infty$ for $p < 1/r$ and $\lim_n \langle x \rangle_n^{(p)} = 0$ for $p > 1/r$; Therefore, we must have $R = r$. This completes the proof.

Let us now prove the converse assertion of the theorem by verifying its contrapositive statement. To this end, we first assume that condition (2.8) does not hold with $p = 1/r > 2$, which implies that there exist $\varepsilon > 0$ and a subsequence $(n_k)_{k \in \mathbb{N}}$ such that

$$\langle x \rangle_{n_k}^{(p)} \geq 2^{n_k(1-p/2+\varepsilon)} \left(\langle x \rangle_{n_k}^{(2)} \right)^{p/2} = 2^{n_k(1-p+\varepsilon)} \left(s_{n_k}^2 \right)^{p/2} = 2^{n_k(1-pr_{n_k}+\varepsilon)}.$$

As $r_n \rightarrow r = 1/p$, then $pr_{n_k} \leq 1 + \varepsilon/2$ for a sufficiently large k . Thus, there exists $k_0 \in \mathbb{N}$ such that for $k \geq k_0$, we get $\langle x \rangle_{n_k}^{(p)} \geq 2^{\varepsilon n_k/2}$. Let $\delta > 0$ be such that $\delta/(p + \delta) < \varepsilon/4$, then it follows from (2.17) that

$$\langle x \rangle_{n_k}^{(p+\delta)} \geq \left(2^{n_k \left(\frac{p}{p+\delta} - 1 \right)} \langle x \rangle_{n_k}^{(p)} \right)^{\frac{p+\delta}{p}} \geq 2^{\frac{p+\delta}{p} \left(\frac{\varepsilon}{2} - \frac{\delta}{p+\delta} \right) n_k} \geq 2^{\frac{(p+\delta)\varepsilon}{4p} n_k}.$$

The above inequality implies that $\limsup_n \langle x \rangle_n^{(p+\delta)} = \infty$, and Proposition 2.2 indicates that x must admit the roughness exponent $R \leq (p + \delta)^{-1} < r$. The proof for the case $p = 1/r < 2$ is analogous, and condition (2.8) holds trivially at $p = 2$. This completes the proof. $\square$

The following proposition provides two alternative conditions on the Faber–Schauder coefficients of an arbitrary continuous function x that may be easier to verify than the reverse Jensen condition, but they are also slightly stronger conditions.

Proposition 2.15. *Consider the following two conditions on the Faber–Schauder coefficients of x .*

- (a) *There exists a non-decreasing subexponential function γ_1 such that for sufficiently large m ,*

$$\max_{k=0, \dots, 2^m-1} |\theta_{m,k}| \leq \gamma_1(m) \min_{k=0, \dots, 2^m-1} |\theta_{m,k}|.$$

- (b) *There exist $\nu \in \mathbb{N}_0$ and a non-decreasing subexponential function $\gamma_2 : \mathbb{N} \rightarrow \mathbb{R}$ such*

that for sufficiently large $m \geq \nu$,

$$\max_{k=0, \dots, 2^{m-\nu}-1} \sum_{j=2^\nu k}^{2^\nu(k+1)-1} \theta_{m,j}^2 \leq \gamma_2(m) \min_{k=0, \dots, 2^{m-\nu}-1} \sum_{j=2^\nu k}^{2^\nu(k+1)-1} \theta_{m,j}^2.$$

Then (a) implies (b), and (b) implies the reverse Jensen condition (2.8).

Proof of Proposition 2.15. (a) $\Rightarrow$ (b): If condition (a) holds, then (b) holds with $\nu = 0$ and $\gamma_2 = \gamma_1^2$.

(b) $\Rightarrow$ (2.8): For the ease of notation, we prove this implication only for the case that $\nu = 1$; the case that $\nu \geq 2$ can be proved analogously. Recall that the upper bound in the reverse Jensen condition holds with $\varrho_p \equiv 1$ if $q := p/2 \leq 1$. So it is sufficient to deduce this upper bound for $q > 1$. We will use the probabilistic formalism introduced in (2.10). Note first that

$$\frac{1}{2^{n-1}} \sum_{k=0}^{2^{n-1}-1} \left(\sum_{m=0}^{n-1} 2^m \theta_{m, \lfloor 2^{m-n} k \rfloor}^2 \right)^{p/2} = \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^q \right].$$

Let $a_0, \dots, a_{n-1} \geq 0$ and note that $(\sum_i a_i)^q \leq n^{q-1} \sum_i a_i^q$ by Jensen's inequality. Hence,

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^q \right] &\leq n^{q-1} \mathbb{E} \left[\sum_{m=0}^{n-1} S_m^q \right] = n^{q-1} \sum_{m=0}^{n-1} 2^{-m} \sum_{k=0}^{2^m-1} 2^{mq} |\theta_{m,k}|^{2q} \\ &\leq 2^{q-1} n^{q-1} \left(|\theta_{0,0}|^{2q} + \sum_{m=1}^{n-1} 2^{-m} \sum_{k=0}^{2^{m-1}-1} \left(2^m \theta_{m,2k}^2 + 2^m \theta_{m,2k+1}^2 \right)^q \right) \\ &\leq 2^{q-1} n^{q-1} \left(|\theta_{0,0}|^{2q} + \sum_{m=1}^{n-1} \frac{1}{2} \left(\max_k (2^m \theta_{m,2k}^2 + 2^m \theta_{m,2k+1}^2) \right)^q \right) \\ &\leq 2^{q-1} n^{q-1} (\gamma_2(n))^q \left(|\theta_{0,0}|^{2q} + \sum_{m=1}^{n-1} \frac{1}{2} \left(\min_k (2^m \theta_{m,2k}^2 + 2^m \theta_{m,2k+1}^2) \right)^q \right) \\ &\leq 2^{q-1} n^{q-1} (\gamma_2(n))^q \sum_{m=0}^{n-1} \mathbb{E}[S_m]^q \leq 2^{q-1} n^{q-1} (\gamma_2(n))^q \left(\sum_{m=0}^{n-1} \mathbb{E}[S_m] \right)^q. \end{aligned}$$

Selecting $\varrho_p(n) := 2^{p/2-1} n^{p/2-1} (\gamma_2(n))^{p/2}$ and applying (2.14) on the above long inequality then yields an upper bound in (2.8). For $0 < q < 1$, an analogous reasoning as above yields the following result

$$\mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^q \right] \geq n^{q-1} 2^{q-1} (\gamma_2(n))^{-q} \sum_{m=0}^{n-1} \mathbb{E}[S_m]^q \geq \frac{1}{\varrho_p(n)} s_n^p,$$

where $\varrho_p(n) := n^{1-p/2}2^{1-p/2}(\gamma_2(n))^{p/2}$. This establishes the lower bound in the reverse Jensen condition, and we complete our proof. $\square$

2.3 The connection with weighted quadratic variation

The Gladyshev estimator has its origins in Gladyshev's theorem [41], which is a limit theorem for the weighted quadratic variation of fractional Brownian motion. This section discusses the relations between $\widehat{R}_n(x)$ and the weighted quadratic variation of an arbitrary function $x \in C[0, 1]$. We start with the following proposition, which states general implications. Its part (b) introduces the estimator $\widehat{R}_n^*$, whose simplified form can sometimes make the analysis easier.

Proposition 2.16. *Let $x \in C[0, 1]$ have the Faber–Schauder coefficients $\theta_{m,k}$.*

(a) *The following conditions are equivalent.*

- (i) *There exists $r_1 \in (0, 1)$ such that the weighted quadratic variation $2^{n(2r_1-1)}\langle x \rangle_n^{(2)}$ converges to the limit $\ell_1 \in (0, \infty)$.*
- (ii) *There exists $r_2 \in (0, 1)$ such that $2^{n(2r_2-2)}s_n^2$ converges to the limit $\ell_2 \in (0, \infty)$.*
- (iii) *There exists $r_3 \in (0, 1)$ such that $2^{n(2r_3-2)}(s_{n+1}^2 - s_n^2)$ converges to the limit $\ell_3 \in (0, \infty)$.*

In this case, we have $r := r_1 = r_2 = r_3$ and $\ell_1 = \ell_2 = (2^{2-2r} - 1)^{-1}\ell_3$.

(b) *The equivalent conditions in part (a) imply that $\lim_n \widehat{R}_n^*(x) = r$, where*

$$\widehat{R}_n^*(x) := 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \theta_{n,k}^2}. \quad (2.24)$$

(c) *If $\widehat{R}_n^*(x)$ converges to the finite limit r , then also $\lim_n \widehat{R}_n(x) = r$.*

Proof. (a): To see the equivalence of (i) and (ii) as well as $r_1 = r_2$ and $\ell_1 = \ell_2$, one argues first as in the proof of [76, Proposition 2.1] that one can assume without loss of generality that $x(0) = x(1) = 0$. Then the claim follows from (2.13). To see that (ii) implies (iii) with $r_3 = r_2$ and $\ell_2 = (2^{2-2r_2} - 1)^{-1}\ell_3$, we let $b_n = 2^{n(2-2r_2)}$. Then $s_n^2/b_n \rightarrow \ell_2$ by assumption. It follows from [76, Lemma 3.1] that

$$\frac{2^{(2r_2-2)n}}{2^{2-2r_2} - 1} (s_{n+1}^2 - s_n^2) = \frac{s_{n+1}^2 - s_n^2}{b_{n+1} - b_n} \longrightarrow \ell_2 \quad \text{as } n \uparrow \infty. \quad (2.25)$$

For the proof that (iii) implies (ii) with $\ell_2 = (2^{2-2r_2} - 1)^{-1}\ell_3$ and $r_2 = r_3$, it suffices to note that the convergence (2.25) implies that $s_n^2/b_n \rightarrow \ell_2$ by means of the Stolz–Cesaro theorem in the form of [78, Theorem 1.22].

(b): Using (iii) yields that

$$r - \widehat{R}_n^*(x) = (r - 1) + \frac{1}{2n} \log_2(s_{n+1}^2 - s_n^2) = \frac{1}{2n} \log_2(2^{n(2r-2)}(s_{n+1}^2 - s_n^2)) \rightarrow 0.$$

(c): Let $\varepsilon \in (0, 2 - 2r)$ be given. By part (b), there is a constant $c > 0$ such that $c^{-1}2^{(2-2r-\varepsilon)n} \leq (s_{n+1}^2 - s_n^2) \leq c2^{(2-2r+\varepsilon)n}$ for all $n \in \mathbb{N}$. It follows that

$$\frac{2^{(2-2r-\varepsilon)n} - 1}{c(2^{2-2r-\varepsilon} - 1)} \leq \sum_{k=0}^{n-1} (s_{k+1}^2 - s_k^2) \leq \frac{c(2^{(2-2r+\varepsilon)n} - 1)}{(2^{2-2r+\varepsilon} - 1)}.$$

Cancelling the terms in the telescopic sum, taking logarithms, dividing by $2n$, passing to the limit $n \uparrow \infty$, and sending ε to zero now yields the result. $\square$

The following corollary follows immediately from Theorem 2.4 and Proposition 2.16.

Corollary 2.17. *If $x \in C[0, 1]$ satisfies the reverse Jensen condition, and one of the conditions in (a), (b) or (c) in Proposition 2.16 holds for some $r \in (0, 1)$, then x admits the roughness exponent r .*

The implications (c) $\Rightarrow$ (b) and (b) $\Rightarrow$ (a) in Proposition 2.16 are generally false. This is illustrated in the following example.

Example 2.18. For given $R \in (0, 1)$, we consider $x \in C[0, 1]$ defined through the Faber–Schauder coefficients $\theta_{m,k} = 2^{m(1/2-R)}$ if m is even and $\theta_{m,k} = 0$ if m is odd. Recall from Remark 2.7 that x satisfies the reverse Jensen condition. A straightforward computation gives

$$\widehat{R}_n(x) = 1 - \frac{1}{2n} \log_2 \frac{2^{4+4(1-R)\lfloor \frac{n-1}{2} \rfloor} - 2^{4R}}{2^4 - 2^{4R}} \longrightarrow R.$$

Theorem 2.4 hence implies that x has the roughness exponent R . However, $\widehat{R}_n^*(x)$ clearly does not converge. It follows that also none of the quantities in part (a) of Proposition 2.16 can converge.

Now take again $R \in (0, 1)$ and let $\theta_{m,k} = 2^{m(1/2-R)} \sqrt{(2^{2(1-R)} - 1)g(m)}$, where the function g is strictly positive and regularly varying with index ϱ . It will be shown in Corollaries 5.12 and 5.13 that the corresponding function x has the roughness exponent R and that the $(1/R)^{\text{th}}$ variation of x is zero for $\varrho < 0$ and infinite for $\varrho > 0$. If g is slowly

varying, so that $\varrho = 0$, then $\liminf_n \langle x \rangle_n^{(1/R)}$ and $\limsup_n \langle x \rangle_n^{(1/R)}$ can take arbitrary values in $[0, \infty]$. We clearly have $\widehat{R}_n^*(x) = R - \frac{1}{2^n} \log_2((2^{2-2R} - 1)g(n)) \rightarrow R$. However, $2^{n(2R-2)}(s_{n+1}^2 - s_n^2) = (2^{2-2R} - 1)g(n)$ does not converge to a finite limit unless $g(n)$ does. This shows that the implication (b) $\Rightarrow$ (a) in Proposition 2.16 does not hold.

As announced above, we now present an example of a continuous function x for which the weighted quadratic variation $2^{n(2\gamma-1)} \langle x \rangle_n^{(2)}$ converges to a finite limit for some $\gamma \in (0, 1)$, but which nevertheless admits a different roughness exponent R . Note that Proposition 2.16 implies that $\widehat{R}_n(x) \rightarrow \gamma$. In other words, the limit of $\widehat{R}_n(x)$ is different from the function's roughness exponent. Hence, x cannot satisfy the reverse Jensen condition. The function x therefore also highlights the importance of the reverse Jensen condition in Theorem 2.4.

Proposition 2.19. *Let $\alpha \in (0, 1]$ and $R \in (0, 1)$. Then the function*

$$x = \sum_{m=0}^{\infty} 2^{(\frac{1}{2}-\alpha R)m} \sum_{k=0}^{\lfloor 2^{\alpha m} \rfloor - 1} e_{m,k}$$

admits the roughness exponent R . On the other hand, let $\gamma := \alpha R + (1 - \alpha)/2$, then

$$\lim_{n \uparrow \infty} 2^{n(2\gamma-1)} \langle x \rangle_n^{(2)} = (2^{2-2\gamma} - 1)^{-1} \quad \text{and} \quad \lim_{n \uparrow \infty} \widehat{R}_n(x) = \gamma. \quad (2.26)$$

Proof. We begin by proving (2.26). To this end, we verify condition (iii) in Proposition 2.16 (a):

$$\lim_{n \uparrow \infty} 2^{n(2\gamma-2)} \sum_{k=0}^{2^n-1} \theta_{n,k}^2 = \lim_{n \uparrow \infty} 2^{n((2\gamma-2)+(1-2\alpha R))} \lfloor 2^{\alpha n} \rfloor = 1.$$

Proposition 2.16 then yields that $\lim_n \widehat{R}_n(x) = \gamma$.

Next, we aim to show that x admits the roughness exponent R . For a fixed $\alpha \in (0, 1]$, let $n_0 \in \mathbb{N}$ be such that $\lfloor 2^{\alpha n} \rfloor / 2 \leq \lfloor 2^{\alpha(n-1)} \rfloor - 1/2$ for $n \geq n_0$. Let

$$\tilde{x} := \sum_{m=n_0}^{\infty} 2^{(\frac{1}{2}-\alpha R)m} \sum_{k=0}^{\lfloor 2^{\alpha m} \rfloor - 1} e_{m,k}.$$

Based on the argument presented in the proof of Lemma 3.1 in [77], the functions x and $\tilde{x}$ share the same roughness exponent. Hence, it will be sufficient to henceforth consider the function $\tilde{x}$. Recalling the definition of V_m and S_m in (2.10), it is clear that for $m \geq n_0$ and $2^m V_m \geq \lfloor 2^{\alpha m} \rfloor$, we have $S_m = 2^{m(2-2\alpha R)}$, and otherwise zero. Clearly, $2^{m-1} V_{m-1} \leq 2^{m-1} V_m$. Therefore, for $m \geq n_0$, if $2^m V_m \leq \lfloor 2^{\alpha m} \rfloor - 1$, then

$$2^{m-1} V_{m-1} \leq \frac{2^m V_m}{2} \leq \frac{1}{2} (\lfloor 2^{\alpha m} \rfloor - 1) \leq \lfloor 2^{\alpha(m-1)} \rfloor - 1. \quad (2.27)$$

In other words, if $S_m = 2^{m(2-2\alpha R)}$, then $S_{m-1} = 2^{(m-1)(2-2\alpha R)}$. On the other hand, for $m \geq n_0$, if $2^m V_m \geq \lfloor 2^{\alpha m} \rfloor$, then

$$2^{m+1} V_{m+1} \geq 2^{m+1} V_m \geq 2 \lfloor 2^{\alpha m} \rfloor \geq 2 \lfloor 2^{\alpha m} \rfloor - 1 \geq \lfloor 2^{\alpha(m+1)} \rfloor. \quad (2.28)$$

Equivalently speaking, if $S_m = 0$, then $S_{m+1} = 0$. For $n \geq n_0 + 3$, it then follows from (2.27) and (2.28) that if there exists $n_0 + 1 \leq m \leq n - 2$ such that $2^m V_m \leq \lfloor 2^{\alpha m} \rfloor - 1$ and $2^{m+1} V_{m+1} \geq \lfloor 2^{\alpha(m+1)} \rfloor$, we have $S_k = 2^{k(2-2\alpha R)}$ for $n_0 \leq k \leq m$ and $S_k = 0$ for $m + 1 \leq k \leq n - 2$. Therefore, for $n \geq n_0 + 3$, the distribution of the random variable $\sum_{m=0}^{n-1} S_m$ is as follows,

$$\mathbb{P}\left(\sum_{m=0}^{n-1} S_m = \sum_{m=n_0}^{n-1} 2^{m(2-2\alpha R)}\right) = \frac{\lfloor 2^{\alpha(n-1)} \rfloor}{2^{n-1}}, \quad \mathbb{P}\left(\sum_{m=0}^{n-1} S_m = 0\right) = 1 - \frac{\lfloor 2^{\alpha n_0} \rfloor}{2^{n_0}},$$

and

$$\mathbb{P}\left(\sum_{m=0}^{n-1} S_m = \sum_{m=n_0}^{k-1} 2^{m(2-2\alpha R)}\right) = \frac{\lfloor 2^{\alpha(k-1)} \rfloor}{2^{k-1}} - \frac{\lfloor 2^{\alpha k} \rfloor}{2^k} \quad \text{for } n_0 + 1 \leq k \leq n - 1.$$

Therefore, for $n \geq n_0 + 3$ and $p \geq 1$,

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{m=0}^{n-1} S_m\right)^{p/2}\right] &= \frac{\lfloor 2^{\alpha(n-1)} \rfloor}{2^{n-1}} \left(\sum_{m=n_0}^{n-1} 2^{m(2-2\alpha R)}\right)^{p/2} \\ &\quad + \sum_{k=n_0+1}^{n-1} \left(\frac{\lfloor 2^{\alpha(k-1)} \rfloor}{2^{k-1}} - \frac{\lfloor 2^{\alpha k} \rfloor}{2^k}\right) \left(\sum_{m=n_0}^{k-1} 2^{m(2-2\alpha R)}\right)^{p/2}. \end{aligned}$$

In the next step, we seek suitable upper and lower bounds for this expression. To obtain a lower bound, we use the following facts,

$$\frac{\lfloor 2^{\alpha(n-1)} \rfloor}{2^{n-1}} \geq 2^{(\alpha-1)(n-1)} - 2^{-(n-1)} \quad \text{and} \quad \sum_{m=n_0}^{n-1} 2^{m(2-2\alpha R)} \sim \frac{2^{n(2-2\alpha R)}}{2^{2-2\alpha R} - 1} \quad \text{as } n \uparrow \infty.$$

Hence, there exists a positive constant c_1 such that for sufficiently large n ,

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{m=0}^{n-1} S_m\right)^{p/2}\right] &\geq \left(2^{(\alpha-1)(n-1)} - 2^{-(n-1)}\right) \left(\sum_{m=n_0}^{n-1} 2^{m(2-2\alpha R)}\right)^{p/2} \\ &\geq c_1 2^{n(p(1-\alpha R) - (1-\alpha))}. \end{aligned} \quad (2.29)$$

To obtain the upper bound, we apply the inequality $\lfloor 2^{\alpha(n-1)} \rfloor / 2^{n-1} \leq 2^{(\alpha-1)(n-1)}$ and Jensen's

inequality,

$$\mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \leq (n^{p/2-1} \vee 1) \sum_{k=n_0+1}^n 2^{(\alpha-1)(k-1)} \left(\sum_{m=n_0}^{k-1} 2^{mp(1-\alpha R)} \right).$$

As $p(1-\alpha R) > 1-\alpha \geq 0$, the following asymptotics hold as $n \uparrow \infty$,

$$\sum_{k=n_0+1}^n 2^{(\alpha-1)(k-1)} \sum_{m=n_0}^{k-1} 2^{mp(1-\alpha R)} = O(2^{n(p(1-\alpha R)-(1-\alpha))}).$$

Thus, for each $p \geq 1$, there exists a positive constant c_2 such that for sufficiently large n ,

$$\mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \leq c_2 (n^{p/2-1} \vee 1) 2^{n(p(1-\alpha R)-(1-\alpha))}. \quad (2.30)$$

It then follows from (2.29) and (2.30) that there exists a positive constant c_p depending on p such that for sufficiently large n ,

$$\frac{1}{c_p} 2^{n(p(1-\alpha R)-(1-\alpha))} \leq \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \leq c_p (n^{p/2-1} \vee 1) 2^{n(p(1-\alpha R)-(1-\alpha))}.$$

Proposition 2.10 now yields that for each $p \geq 1$, there exists $n_0 \in \mathbb{N}$, such that for $n \geq n_0$,

$$\frac{A_p}{c_p} 2^{\alpha n(1-pR)} \leq \langle \tilde{x} \rangle_n^{(p)} \leq c_p B_p 2^{\alpha n(1-pR)}.$$

Taking $n \uparrow \infty$ on both sides of the above inequalities gives

$$\lim_{n \uparrow \infty} \langle \tilde{x} \rangle_n^{(p)} = \begin{cases} 0 & \text{for } p > \frac{1}{R}, \\ \infty & \text{for } p < \frac{1}{R}. \end{cases}$$

Thus, $\tilde{x}$ admits the roughness exponent R , which completes the proof. $\square$

It is worthwhile to point out that $\gamma = \alpha R + (1-\alpha)/2$ is the convex combination of R and $1/2$. If $R > 1/2$, we have $R \geq \gamma > 1/2$ and vice versa. In fact, such a relation holds universally as later shown in Proposition 2.35.

2.4 The connection with self-similarity

Another concept that admits a potential connection with the roughness exponent is self-similarity. Self-similar processes are stochastic processes that are invariant in distribution under suitable scaling of time and space. We say that a stochastic process $(X_t)_{t \geq 0}$ is γ -self-similar, if there exists $\gamma > 0$ such that for any $a > 0$, $(X_{at})_{t \geq 0} \stackrel{d}{=} (a^\gamma X_t)_{t \geq 0}$. Here, the constant γ is called the self-similarity index, which dictates the nature of the self-similar phenomenon. For instance, the fractional Brownian motion W^H is an H -self-similar process. An important question is whether the self-similarity index is generally equivalent to the roughness exponent, as observed in the case of the fractional Brownian motion.

Our findings in this section refute the above assertion. In the first result, we study the weighted quadratic variation of a trifractional Brownian motion, indicating that self-similar behavior is not correlated to the roughness exponent or the weighted quadratic variation. A trifractional Brownian motion was first introduced by Lei and Nualart [67]. In this paper, the authors showed a bifractional Brownian motion could be decomposed into the sum of a fractional Brownian motion and a Gaussian process with absolutely continuous trajectories. This process was subsequently rediscovered and studied by Ma [71], who referred to it as the trifractional Brownian motion. According to Ma [71], a *trifractional Brownian motion* $(Z_t^{H,K})_{t \geq 0}$ with parameters $H \in (0, 1)$ and $K \in (0, 1)$ is a centered Gaussian process with covariance function $C_{H,K}$,

$$C_{H,K}(s, t) = \mathbb{E} \left[Z_s^{H,K} Z_t^{H,K} \right] = t^{2HK} + s^{2HK} - (t^{2H} + s^{2H})^K, \quad s, t \geq 0$$

on a given probability space $(\tilde{\Omega}, \mathcal{G}, \mathbb{Q})$.

Clearly, the trifractional Brownian motion $Z^{H,K}$ is an HK -self-similar process. However, the trajectories of trifractional Brownian motion are absolutely continuous and, thus, of bounded variation. Hence, the trifractional Brownian motion $Z^{H,K}$ admits the roughness exponent $R = 1$ almost surely. The following result proves a limit theorem on the weighted quadratic variation of the trifractional Brownian motion $Z^{H,K}$.

Theorem 2.20. *For $H \in (0, 1)$ and $K \in (0, 1)$, the trifractional Brownian motion $Z^{H,K}$ has the following limiting behavior,*

(a) *If $HK < \frac{1}{2}$, we have*

$$\lim_{n \uparrow \infty} 2^{\alpha n} \langle Z^{H,K} \rangle_n^{(2)} = \begin{cases} 0 & \text{for } \alpha < 2HK, \\ \infty & \text{for } \alpha > 2HK, \end{cases} \quad (2.31)$$

$$(2.32)$$

almost surely.

(b) If $HK \geq \frac{1}{2}$, we have

$$\lim_{n \uparrow \infty} 2^{\alpha n} \langle Z^{H,K} \rangle_n^{(2)} = \begin{cases} 0 & \text{for } \alpha < 1, \\ \infty & \text{for } \alpha > 1, \end{cases} \quad (2.33)$$

$$(2.34)$$

almost surely.

Applying (2.13) to the above asymptotic results and arguing as in the proof of Theorem 2.4 yields our next result. As a probabilistic complement to the deterministic counterexample in Proposition 2.19, the following corollary demonstrates that the Gladyshev estimates fail to converge to either the self-similarity index or the roughness exponent.

Corollary 2.21. *For $H \in (0, 1)$ and $K \in (0, 1)$, we have*

$$\lim_{n \uparrow \infty} \widehat{R}_n(Z^{H,K}) = \begin{cases} HK + \frac{1}{2} & \text{for } HK < \frac{1}{2}, \\ 1 & \text{for } HK \geq \frac{1}{2}, \end{cases} \quad \mathbb{Q}\text{-a.s.}$$

In addition to the trifractional Brownian motion, the following limit theorems for the n^{th} order fractional Brownian motion also reveal that the roughness exponent bifurcates from the self-similarity index. The behavior of the well-known fractional Brownian motion is governed by the Hurst parameter H between 0 and 1. However, many observations reveal that the Hurst parameter could be larger than one in theory and in real life. For instance, a slight modification of the Weyl representation of the fractional Brownian motion allows the Hurst parameter to be larger than one; see, e.g., [83]. Motivated by this fact, Perrin et al. [83] introduced the n^{th} -order fractional Brownian motion as an extension to the fractional Brownian motion. An n^{th} -order fractional Brownian motion $W^{H,n} := (W_t^{H,n})_{t \geq 0}$ defined on the probability space $(\tilde{\Omega}, \mathcal{G}, \mathbb{Q})$ is a centered Gaussian process with the following covariance function,

$$\begin{aligned} G_{H,n}(s, t) &:= \mathbb{E} [W_s^{H,n} W_t^{H,n}] \\ &= (-1)^n \frac{C_H^n}{2} \left(|t - s|^{2H} - \sum_{j=0}^{n-1} (-1)^j \binom{2H}{j} \left(\left(\frac{t}{s} \right)^j s^{2H} + \left(\frac{s}{t} \right)^j t^{2H} \right) \right), \quad s, t \geq 0, \end{aligned}$$

for $n \in \mathbb{N}$, $H \in (n-1, n)$ and $C_H^n := (\Gamma(2H+1)|\sin(\pi H)|)^{-1}$. The special case where $n = 1$ recovers the standard fractional Brownian motion, i.e., $W^{H,1} = W^H$. The class of n^{th} -order fractional Brownian motions allows a wider range of the Hurst parameter H . In other words, the class of n^{th} -order fractional Brownian motions extends the Hurst parameter beyond the constraint $H \in (0, 1)$ and includes the case of fractional Brownian motion for $n = 1$.

Nevertheless, an n^{th} -order fractional Brownian motion remains H -self-similar. Furthermore, following [91, Remark 3.2], $W^{H,m}$ is $(m - 1)$ times differentiable, and thus, admits the roughness exponent $R = 1$, which differs from its self-similarity index.

Now we state the limit theorems for the n^{th} -order fractional Brownian motion.

Theorem 2.22. *For $m \geq 2$ and $H \in (m - 1, m)$, the m^{th} -order fractional Brownian motion $W^{H,m}$ has the following limiting behavior,*

$$\lim_{n \uparrow \infty} 2^{\alpha n} \langle W^{H,m} \rangle_n^{(2)} = 0 \quad \text{for } \alpha < 1, \quad (2.35)$$

almost surely. Moreover, if $m \geq 3$, we have

$$\lim_{n \uparrow \infty} 2^{\alpha n} \langle W^{H,m} \rangle_n^{(2)} = \infty \quad \text{for } \alpha > 1, \quad (2.36)$$

almost surely.

Applying (2.13) to Theorem 2.22 and arguing as in the proof of Theorem 2.4 yields the following result. In contrast to Corollary 2.21, the following result shows that the roughness exponent of the n^{th} -order fractional Brownian motion can be estimated by the Gladyshev estimator $\widehat{R}_n$, or equivalently, by means of a weighted quadratic variation.

Corollary 2.23. *For $m \geq 3$ and $H \in (m - 1, m)$, we have*

$$\lim_{n \uparrow \infty} \widehat{R}_n(W^{H,m}) = 1 \quad \mathbb{Q}\text{-a.s.}$$

The results in Theorem 2.20 and Theorem 2.22 are non-trivial in the sense that they cannot be obtained by applying the Gladyshev theorem [41] or its complement [81]. This point is demonstrated by the following limit theorems at the critical cases for the trifractional Brownian motion and the n^{th} -order fractional Brownian motion, see Remark 2.26.

Theorem 2.24. *For $H \in (0, 1)$ and $K \in (0, 1)$, the trifractional Brownian motion $Z^{H,K}$ has the following limiting behavior,*

- (a) *If $HK < 1/2$, the weighted quadratic variation $2^{2HKn} \langle Z^{H,K} \rangle_n^{(2)}$ does not converge in probability as $n \uparrow \infty$.*
- (b) *If $HK > 1/2$, the weighted quadratic variation $2^n \langle Z^{H,K} \rangle_n^{(2)}$ converges to some non-constant random variable in L^p for all $p \in (1, \infty)$ as $n \uparrow \infty$.*

Theorem 2.25. *For $m \geq 2$ and $H \in (m - 1, m)$, the weighted quadratic variation of the m^{th} -order fractional Brownian motion $2^n \langle W^{H,m} \rangle_n^{(2)}$ converges to some non-constant random variable in L^p for all $p \in (1, \infty)$ as $n \uparrow \infty$.*

Remark 2.26. As observed in Theorem 2.24 and Theorem 2.25, the weighted quadratic variation of the trifractional Brownian motion and the n^{th} -order fractional Brownian motion behaves essentially different to the Gaussian processes that Gladyshev's theorem apply to. To be more specific, if a Gaussian process fulfills conditions in Gladyshev [41] or Norvaiša [81], its weighted quadratic sum at the critical case converges to some positive constant at the critical case almost surely. However, for the trifractional Brownian motion $Z^{H,K}$, its weighted quadratic variation converges to some non-constant random variable in L^p and in probability for $HK > 1/2$ and does not converge in probability for $HK < 1/2$ at each critical case. In addition, as in Theorem 2.25, the weighted sum of the n^{th} -order fractional Brownian motion converges to some non-constant random variable at the critical case. These facts further refute the applicability of the existing literature [41, 81] to the trifractional Brownian motion and the n^{th} -order fractional Brownian motion.

2.5 The connection with Besov regularity

This section establishes a connection between the roughness exponent and Besov regularity. We refer to [93, Section 2.3] for the definition of the Besov space $\mathcal{B}_{p,q}^r[0, 1]$. Ciesielski et al. [16] provided a norm expressed in terms of Faber–Schauder coefficients, which is equivalent to the traditional Besov norm. Rosenbaum [85] defined the following equivalent norm on $\mathcal{B}_{p,q}^r[0, 1]$ for $p, q \in [1, \infty)$ and $r \in (1/p, 1)$,

$$\|x\|_{r,p,q} := |x(0)| \vee \|(a_n)_{n \in \mathbb{N}_0}\|_{\ell_q}, \quad \text{where} \quad a_n := \left(2^{n(pr-1)} \langle x \rangle_n^{(p)}\right)^{1/p}. \quad (2.37)$$

Note that for $p = 2$ the relation to weighted quadratic variation as studied in Section 2.3 becomes apparent. Measuring the degree of roughness of a function $x \in [0, 1]$ through Besov spaces was initially proposed by Eyink [29] through the condition

$$x \in \bigcap_{\varepsilon > 0} \mathcal{B}_{p,\infty}^{r-\varepsilon}[0, 1] \setminus \bigcup_{\varepsilon > 0} \mathcal{B}_{p,\infty}^{r+\varepsilon}[0, 1]. \quad (2.38)$$

Example 2.27. Consider the function x defined in Proposition 2.19 with $\alpha \in (0, 1)$ and $R > 1/2$. The first identity in (2.26) implies that $\|x\|_{\gamma,2,\infty} < \infty$ and that $\|x\|_{\gamma+\varepsilon,2,\infty} = \infty$ for any $\varepsilon > 0$. Hence, condition (2.38) holds for $r = \gamma$, $p = 2$, and $q = \infty$. Nevertheless, as observed in Proposition 2.19, the roughness exponent of x is different from γ , which shows that Eyink's concept of Besov regularity is generally different from our concept of roughness.

In the rough volatility context [40], the parameter r in (2.38) is often assumed to be the

same for all $p \geq 1$, which leads to the requirement

$$x \in \bigcap_{\varepsilon > 0, p \geq 1} \mathcal{B}_{p,\infty}^{r-\varepsilon}[0, 1] \setminus \bigcup_{\varepsilon > 0, p \geq 1} \mathcal{B}_{p,\infty}^{r+\varepsilon}[0, 1]. \quad (2.39)$$

In the following theorem, we explore the connection between this concept of Besov regularity and the roughness exponent. Let q^* be as in (2.39) and $R^- := 1/q^*$. Our result shows that the condition (2.39) implies that $r = R^-$. As a matter of fact, the condition (2.40) used in Theorem 2.28 is actually less restrictive than (2.39). Note moreover that Proposition 2.19 shows that (2.39) or (2.40) cannot be replaced with Eyink's condition (2.38). Indeed, the latter condition holds in Proposition 2.19 for $r = \gamma$, while the function x constructed there admits an arbitrary roughness exponent $R = R^-$.

Theorem 2.28. *Suppose that there exists $r \in (0, 1)$ such that*

$$x \in \bigcap_{\varepsilon > 0, p \geq 1} \mathcal{B}_{p,\infty}^{r-\varepsilon}[0, 1] \setminus \bigcup_{\varepsilon > 0, \frac{1}{r+\varepsilon} < p < \frac{1}{r}} \mathcal{B}_{p,\infty}^{r+\varepsilon}[0, 1]. \quad (2.40)$$

Then $R^- = r$.

Proof of Theorem 2.28. We must show that $q^* = 1/r$, and we start by proving $q^* \leq 1/r$. For $\gamma > 0$ and $0 < \varepsilon < r$, we choose $p := (1 + \gamma)/(r - \varepsilon)$. Then it is clear that $p(r - \varepsilon) = 1 + \gamma > 1$. Hence, $(\sup_n 2^{\gamma n} \langle x \rangle_n^{(p)})^{1/p} = \|x\|_{r-\varepsilon, p, \infty} < \infty$. Since $\gamma > 0$, we get $\limsup_n \langle x \rangle_n^{(p)} < \infty$, which implies that $q^* \leq p = (1 + \gamma)/(r - \varepsilon)$. Taking $\gamma, \varepsilon \downarrow 0$ yields $q^* \leq 1/r$.

Next, let us show that $q^* \geq 1/r$. To this end, we assume by way of contradiction that $q^* < 1/r$. Let us denote $q = \frac{1}{2}(q^* + 1/r)$. Then there exists $\lambda > 0$ such that $q = 1/(r + \lambda)$ as $q < 1/r$. Next, we set

$$\varepsilon := \frac{r}{8} - \frac{r^2 q}{8} = \frac{r}{8} \left(1 - \frac{r}{r + \lambda} \right) = \frac{r \lambda}{8(r + \lambda)} < \lambda.$$

As $q^* < q$, we have $\limsup_{n \uparrow \infty} \langle x \rangle_n^{(q)} = 0$. Let $\mathcal{H}^\alpha[0, 1]$ denote the space of α -Hölder continuous functions on $[0, 1]$. It follows from a fundamental embedding result [93, Section 2.7.1] that $\mathcal{B}_{p,\infty}^r[0, 1] \subset \mathcal{H}^{r-1/p}[0, 1]$, and this yields

$$\bigcap_{\varepsilon > 0, p \geq 1} \mathcal{B}_{p,\infty}^{r-\varepsilon}[0, 1] \subset \bigcap_{\varepsilon > 0} \mathcal{H}^{r-\varepsilon}[0, 1].$$

Next, we take $p := 1/(r + \varepsilon/2)$. Then $p > q$ as $\varepsilon < \lambda$. Applying the above Hölder condition gives $\langle x \rangle_n^{(p)} \leq \kappa 2^{(q-p)(r-\varepsilon)n} \langle x \rangle_n^{(q)}$ for some $\kappa > 0$. Thus, for $n \in \mathbb{N}$, we get $2^{n((r+\varepsilon)p-1)} \langle x \rangle_n^{(p)} \leq$

$\kappa 2^{n((r+\varepsilon)p-1+(q-p)(r-\varepsilon))} \langle x \rangle_n^{(q)}$. One checks that $(r+\varepsilon)p-1+(q-p)(r-\varepsilon) < 0$. Hence,

$$\limsup_{n \uparrow \infty} 2^{n((r+\varepsilon)p-1)} \langle x \rangle_n^{(p)} \leq \kappa \left(\limsup_{n \uparrow \infty} 2^{n((r+\varepsilon)p-1+(q-p)(r-\varepsilon))} \right) \left(\limsup_{n \uparrow \infty} \langle x \rangle_n^{(q)} \right) = 0.$$

Therefore $\|x\|_{r+\varepsilon, p, \infty} < \infty$ for $p = 1/(r+\varepsilon/2)$, which contradicts the underlying assumption that $x \notin \mathcal{B}_{p, \infty}^{r+\varepsilon}[0, 1]$, and thus $q^* \geq 1/r$. $\square$

In our next example, we construct a function x for which condition (2.40) of Besov regularity is satisfied but where $\widehat{R}_n(x)$ oscillates between a given number $R \in (0, 1)$ and 1. In this sense, the function exhibits an entire spectrum of roughness, which is a phenomenon akin to multifractality. Das [21] proposed to use R^- as a “roughness index” for measuring the roughness of a trajectory $x \in C[0, 1]$. Based on the discussion in the present section, it is clear that both the condition of Besov regularity and the “roughness index” of Das [21] capture only one side of the roughness spectrum $[R^-, R^+]$ for $R^+ := 1/q_*$, whereas the existence of the roughness exponent is tantamount to $R^- = R^+$.

Example 2.29. Consider the function x with Faber–Schauder coefficients $\theta_{n,k} = 2^{(1/2-R)n}$ if $n = m!$ for some $m \in \mathbb{N}$ and $\theta_{n,k} = 0$ otherwise. The idea of constructing such a function is based on Faber [30]. It is easy to see that

$$2^{(2-2R)n!} \leq s_{n!+1}^2 \leq n 2^{(2-2R)n!} \quad \text{and} \quad 2^{(2-2R)(n-1)!} \leq s_{n!}^2 \leq (n-1) 2^{(2-2R)(n-1)!},$$

and hence $r^+ = 1$ and $r^- = R$, where $r^\pm$ is as in (2.9). For each $m \in \mathbb{N}$, the Faber–Schauder coefficient $\theta_{m,k}$ does not depend on k , and so x satisfies the reverse Jensen condition by Remark 2.7. Thus, we have $q^* = 1/r^- = 1/R$ and $q_* = 1/r^+ = 1$, and Proposition 2.2 hence implies that x does not admit a roughness exponent. Let us now show that

$$x \in \bigcap_{\varepsilon > 0, p \geq 1} \mathcal{B}_{p, \infty}^{R-\varepsilon}[0, 1] \setminus \bigcup_{\varepsilon > 0, \frac{1}{R+\varepsilon} < p < \frac{1}{R}} \mathcal{B}_{p, \infty}^{R+\varepsilon}[0, 1]. \quad (2.41)$$

We prove first that x belongs to the intersection on the left-hand side. For every $p \geq 1$, it follows from (2.23) that there exist $0 < A_p \leq B_p < \infty$ and a subexponential function ϱ_p such that

$$A_p \varrho_p^{-1}(n) 2^{n(1-p)} s_n^p \leq \langle x \rangle_n^{(p)} \leq B_p \varrho_p(n) 2^{n(1-p)} s_n^p.$$

Since $R \in (0, 1)$, we have that for any $\varepsilon > 0$ and $p \geq 1$,

$$\begin{aligned}
\|x\|_{R-\varepsilon, p, \infty} &= \sup_n \left(2^{n(p(R-\varepsilon)-1)} \langle x \rangle_n^{(p)} \right)^{1/p} \leq \sup_n \left(2^{np(R-\varepsilon-1)} B_p \varrho_p(n) s_n^p \right)^{1/p} \\
&= \sup_n \left(2^{(n!+1)p(R-\varepsilon-1)} B_p \varrho_p(n!+1) s_{n!+1}^p \right)^{1/p} \\
&\leq \sup_n \left(2^{(n!+1)p(R-\varepsilon-1)} B_p \varrho_p(n!+1) n 2^{p(1-R)n!} \right)^{1/p} \\
&= 2^{R-\varepsilon-1} B_p^{1/p} \sup_n \left(2^{-p\varepsilon n!} n \varrho_p(n!+1) \right)^{1/p} < \infty.
\end{aligned}$$

The final inequality holds because $\varrho_p(n)$ is a sub-exponential function. In a next step, we are going to show that for any $\varepsilon > 0$ and $1/R > p > 1/(R + \varepsilon)$, we have $x \notin \mathcal{B}_{p, \infty}^{R+\varepsilon}[0, 1]$. Since $p > 1/(R + \varepsilon)$,

$$\begin{aligned}
\|x\|_{R+\varepsilon, p, \infty} &= \sup_n \left(2^{n(p(R+\varepsilon)-1)} \langle x \rangle_n^{(p)} \right)^{1/p} \geq \sup_n \left(2^{(n!+1)p(R+\varepsilon-1)} A_p \varrho_p^{-1}(n!+1) s_{n!+1}^p \right)^{1/p} \\
&\geq \sup_n \left(2^{(n!+1)p(R+\varepsilon-1)} A_p \varrho_p^{-1}(n!+1) 2^{p(1-R)n!} \right)^{1/p} \\
&= 2^{R+\varepsilon-1} A_p^{1/p} \sup_n \left(2^{p\varepsilon n!} \varrho_p^{-1}(n!+1) \right)^{1/p},
\end{aligned}$$

which is infinite, because ϱ_p is sub-exponential.

2.6 Application to fractional Brownian motion with drift

Fractional Brownian motion with Hurst parameter $H \in (0, 1)$ is a centered Gaussian stochastic process W^H with continuous sample paths and covariance function

$$\mathbb{E}[W_s^H W_t^H] = \frac{1}{2} (|s|^{2H} + |t|^{2H} - |t-s|^{2H}), \quad s, t \geq 0,$$

on a given probability space $(\tilde{\Omega}, \mathcal{G}, \mathbb{Q})$. For $H = 1/2$, we recover a standard Brownian motion $W = W^{1/2}$ as a special case.

The following theorem shows that almost all sample paths of the fractional Brownian motion admit roughness exponent H . In addition, we also obtain the convergence rate of the Gladyshev estimator $\hat{R}_n$ over the fractional Brownian motion.

Theorem 2.30. *For all $H \in (0, 1)$, almost all sample paths of the process W^H satisfy the reverse Jensen condition and admit the roughness exponent H . Moreover, the estimator $\widehat{R}_n$ is strongly consistent and there exists a positive constant κ such that*

$$\limsup_{n \uparrow \infty} \frac{|\widehat{R}_n(W^H) - H|}{\varepsilon_n} \leq \kappa \quad \mathbb{Q}\text{-a.s.}, \quad (2.42)$$

where

$$\varepsilon_n = \begin{cases} 2^{-n/2} n^{-1} \sqrt{\log n} & \text{if } H \in (0, \frac{1}{2}), \\ 2^{-n/2} n^{-1/2} \sqrt{\log n} & \text{if } H = \frac{1}{2}, \\ 2^{(H-1)n} n^{-1} \sqrt{\log n} & \text{if } H \in (\frac{1}{2}, 1). \end{cases} \quad (2.43)$$

For fractional Brownian motion with Hurst parameter $H \leq 1/2$, it follows from Theorem 1.2 in [72] that $\lim_n \langle W^H \rangle_n^{(1/H)} = \pi^{-1/2} 2^{1/(2H)} \Gamma(\frac{H+1}{2H})$ $\mathbb{Q}$ -a.s. As discussed at the beginning of Section 2.1, this implies that W^H admits $\mathbb{Q}$ -a.s. the roughness exponent H . So this part of the assertion of Theorem 2.30 was known beforehand. For the case $H > 1/2$, however, we were only able to find references in the literature that yield convergence of $\langle W^H \rangle_n^{(1/H)}$ in probability (see, e.g., Section 1.18 in [75] and the references therein), which is not quite sufficient to conclude that almost all trajectories of W^H admit the roughness exponent H .

The rate of convergence for $\widehat{R}_n(W^H)$ was initially studied in [64, Theorem 1] for $H > 1/2$. However, as pointed out in our communication [63] with the authors, the convergence rate stated in [64] contains a typo. By modifying and streamlining the original proof idea from [64], Theorem 2.30 provides the correct rate, states a stronger convergence result and extends it to the cases of $H = 1/2$ and $H < 1/2$.

Let us now extend the results in Theorem 2.30 to a fractional Brownian motion with drift. To this end, let X be given by

$$X_t := x_0 + W_t^H + \int_0^t \zeta_s ds, \quad 0 \leq t \leq T, \quad (2.44)$$

where ζ is progressively measurable with respect to the natural filtration generated by W^H . The extension of Theorem 2.30 relies on the following theorem, which provides sufficient conditions that guarantee the law of the process X is absolutely continuous with respect to the law of the original fractional Brownian motion.

Theorem 2.31. *Let X be given by (2.44) and suppose that the following assumption is satisfied.*

- *If $H < 1/2$, we assume that $t \mapsto \xi_t$ is $\mathbb{Q}$ -a.s. bounded in the sense that there exists a finite random variable C such that $\xi_t(\omega) \leq C(\omega)$ for a.e. t and $\mathbb{Q}$ -a.s. ω .*

- If $H > 1/2$, we assume that $t \mapsto \xi_t$ is $\mathbb{Q}$ -a.s. Hölder continuous with some exponent $\alpha > 2H - 1$.

Then the law of $(X_t)_{t \in [0, T]}$ is absolutely continuous with respect to the law of $(x_0 + W_t^H)_{t \in [0, T]}$.

In addition, if the drift ξ is of the form $\xi_t = b(X_t)$ for some function $b : \mathbb{R} \rightarrow \mathbb{R}$, we are able to establish sufficient conditions for the equivalence of the law of X to the one of $x_0 + W^H$. An important special case is the fractional Ornstein–Uhlenbeck process, which was proposed by Gatheral et al. [40] as a suitable candidate for log-volatility. The fractional Ornstein–Uhlenbeck process is defined as the pathwise solution of the following integral equation

$$X_t^H = x_0 + \rho \int_0^t (\mu - X_s^H) ds + W_t^H, \quad t \geq 0, \quad (2.45)$$

where ρ and μ are real parameters; see [13].

Theorem 2.32. *Suppose that X solves the following integral equation,*

$$X_t = x_0 + W_t^H + \int_0^t b(X_s) ds, \quad 0 \leq t \leq T, \quad (2.46)$$

for some locally bounded and measurable function $b : \mathbb{R} \rightarrow \mathbb{R}$. For $H > 1/2$, we assume in addition that b is locally Hölder continuous with some exponent $\alpha > 2 - 1/H$. Then the law of $(X_t)_{t \in [0, T]}$ is equivalent to the law of $(x_0 + W_t^H)_{t \in [0, T]}$.

Clearly, the above applies in particular to the fractional Ornstein–Uhlenbeck process X^H , where $b(x) = \rho(\mu - x)$. Furthermore, as a direct corollary of Theorem 2.31, we are able to show that the fractional Brownian motion with an absolutely continuous drift admits the roughness exponent H . Furthermore, the estimator $\widehat{R}_n$ remains strongly consistent with the same convergence rate as in (2.43).

Corollary 2.33. *Suppose that X is as in (2.44) and satisfies all the assumptions in Theorem 2.31. For all $H \in (0, 1)$, almost all sample paths of the process X satisfy the reverse Jensen condition and admit the roughness exponent H . Moreover, the estimator $\widehat{R}_n$ is strongly consistent and there exists a positive constant κ such that*

$$\limsup_{n \uparrow \infty} \frac{|\widehat{R}_n(X) - H|}{\varepsilon_n} \leq \kappa \quad \mathbb{Q}\text{-a.s.},$$

where

$$\varepsilon_n = \begin{cases} 2^{-n/2} n^{-1} \sqrt{\log n} & \text{if } H \in (0, \frac{1}{2}), \\ 2^{-n/2} n^{-1/2} \sqrt{\log n} & \text{if } H = \frac{1}{2}, \\ 2^{(H-1)n} n^{-1} \sqrt{\log n} & \text{if } H \in (\frac{1}{2}, 1). \end{cases}$$

Note that the Gladyshev estimates $\widehat{R}_n(X)$ converge at the rate $\mathcal{O}(2^{(H-1)n}n^{-1}\sqrt{\log n})$ for $H > 1/2$ almost surely. Consequently, as the Hurst parameter approaches one, the convergence speed of the Gladyshev estimator significantly deteriorates. However, this issue can be solved by utilizing the estimator $\widehat{R}_n^*$ as in (2.24). The following theorem provides the rate of convergence of the estimator $\widehat{R}_n^*$ for the fractional Brownian motion with drift.

Theorem 2.34. *Suppose that X is as in (2.44) and satisfies all the assumptions in Theorem 2.31. The estimator $\widehat{R}_n^*$ is strongly consistent, and there exists a positive constant κ such that*

$$\limsup_{n \uparrow \infty} \frac{|\widehat{R}_n^*(X) + n^{-1} \log_2 \alpha_H - H|}{\varepsilon_n} \leq \kappa \quad \mathbb{Q}\text{-a.s.},$$

where

$$\varepsilon_n = 2^{-n/2}n^{-1}\sqrt{\log n} \quad \text{and} \quad \alpha_H = \sqrt{2^{2-2H} - 1}.$$

As illustrated in the preceding theorem, applying the estimator $\widehat{R}_n^*$ yields an enhanced convergence rate, improving from $\mathcal{O}(2^{(H-1)n}n^{-1}\sqrt{\log n})$ to $\mathcal{O}(2^{-n/2}n^{-1}\sqrt{\log n})$. Nonetheless, in contrast to the Gladyshev estimator $\widehat{R}_n$, the estimator $\widehat{R}_n^*$ entails an extra bias of size $n^{-1} \log_2 \alpha_H$. This bias impedes the direct application of the estimator $\widehat{R}_n^*$ to estimate the Hurst parameter. However, in Section 2.8, we will show that this bias can be removed without deteriorating the corresponding rate of convergence by constructing the scale-invariant modifications of the estimator $\widehat{R}_n^*$.

2.7 The roughness exponent without the reverse Jensen condition

In this section, we collect some general mathematical properties of the roughness exponent, which hold even if the reverse Jensen condition is not satisfied. As in Section 2.1, we fix an arbitrary function $x \in C[0, 1]$ with Faber–Schauder coefficients $(\theta_{m,k})$ and $\widehat{R}_n(x)$ as defined in (2.7). Moreover, $r^+ = \limsup_n \widehat{R}_n(x)$ and $r^- = \liminf_n \widehat{R}_n(x)$ are as in (2.9). We start with the following a priori estimates linking the roughness exponent R with r^+ and r^- .

Proposition 2.35. *Suppose that x admits the roughness exponent R . Then the following assertions hold:*

- (a) *If $r^+ < 1/2$, then $R \leq 1/2$; if $r^- > 1/2$, then $R \geq 1/2$.*
- (b) *If $R \leq 1/2$, then $R \leq r^-$; if $R \geq 1/2$, then $R \geq r^+$.*

Proof of Proposition 2.35. Recall the short-hand notion $r_n := \widehat{R}_n(x)$. To prove (a), we suppose first that $r^+ < 1/2$. Since (2.13) states that $\langle x \rangle_n^{(2)} = 2^{n(1-2r_n)}$ and $\liminf_n \langle x \rangle_n^{(2)} = \infty$, we must have $R \leq 1/2$. In the same way, we get $R \geq 1/2$ if $r^- > 1/2$. To prove (b), we suppose first that $R \leq 1/2$ and take $p > 1/R \geq 2$. Applying Jensen's inequality to (2.21) gives

$$\langle x \rangle_n^{(p)} \geq A_p 2^{n(1-p)} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \geq A_p 2^{n(1-p)} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right) \right]^{p/2} = A_p 2^{n(1-pr_n)}. \quad (2.47)$$

Since $\limsup_n \langle x \rangle_n^{(p)} = 0$, we must have $r^- > 1/p$. Taking $1/p \uparrow R$ gives $R \leq r^-$. The assertion for $R \geq 1/2$ can be proved analogously. $\square$

The preceding proposition provides one-sided bounds on the roughness exponent in terms of $r := \lim_n \widehat{R}_n(x)$. Our next result gives universal two-sided bounds for R . To this end, for $m \in \mathbb{N}$, we denote by

$$\widehat{F}_m(t) := 2^{-m} \left(\sum_{k=0}^{2^m-1} \mathbb{1}_{\{\theta_{m,k}^2 \leq t 2^{-m}\}} \right)$$

the empirical distribution of the m^{th} generation Faber–Schauder coefficients and let $\widehat{F}_m^{-1}$ be a corresponding quantile function.

Proposition 2.36. *Suppose that x admits the roughness exponent R . We define for $\nu \in \mathbb{N}$ and $n \in \mathbb{N}$,*

$$\begin{aligned} r_{\nu,n}^+ &:= 1 - \frac{1}{2n} \log_2 \sum_{m=0}^{n-1} \frac{1}{m^\nu} \sum_{k=1}^{m^\nu} \widehat{F}_m^{-1} \left(2^{-m} \left\lfloor \frac{2^m(k-1)}{m^\nu} \right\rfloor \right), \\ r_{\nu,n}^- &:= 1 - \frac{1}{2n} \log_2 \sum_{m=0}^{n-1} \frac{1}{m^\nu} \sum_{k=1}^{m^\nu} \widehat{F}_m^{-1} \left(2^{-m} \left\lceil \frac{2^m k}{m^\nu} \right\rceil \right). \end{aligned}$$

If the limits $r_\nu^\pm := \lim_{n \uparrow \infty} r_{\nu,n}^\pm$ exist for some $\nu \in \mathbb{N}$, then $r_\nu^- \leq R \leq r_\nu^+$.

Proof of Proposition 2.36. Let us first consider the case $p \geq 2$. It follows from (2.47) that

$$\langle x \rangle_n^{(p)} \geq A_p 2^{n(1-p)} \mathbb{E} \left[\sum_{m=0}^{n-1} S_m \right]^{p/2} = A_p 2^{n(1-p)} \left(\sum_{m=0}^{n-1} \int_0^1 \widehat{F}_m^{-1}(t) dt \right)^{p/2}.$$

Furthermore, we have

$$\int_0^1 \widehat{F}_m^{-1}(t) dt = \sum_{k=1}^{m^\nu} \int_{(k-1)/m^\nu}^{k/m^\nu} \widehat{F}_m^{-1}(t) dt.$$

Note that $2^{-m} \lfloor 2^m(k-1)/m^\nu \rfloor \leq (k-1)/m^\nu \leq k/m^\nu \leq 2^{-m} \lceil 2^m k/m^\nu \rceil$. Moreover, both $2^{-m} \lfloor 2^m(k-1)/m^\nu \rfloor$ and $2^{-m} \lceil 2^m k/m^\nu \rceil$ belong to $\{k2^{-m} : k = 0, \dots, 2^m\}$. Hence,

$$\widehat{F}_m^{-1}\left(2^{-m} \left\lceil \frac{2^m k}{m^\nu} \right\rceil\right) \geq \widehat{F}_m^{-1}(t) \geq \widehat{F}_m^{-1}\left(2^{-m} \left\lfloor \frac{2^m(k-1)}{m^\nu} \right\rfloor\right) \quad (2.48)$$

for all $t \in [(k-1)/m^\nu, k/m^\nu]$. This gives

$$\langle x \rangle_n^{(p)} \geq A_p 2^{n(1-p)} \left[\sum_{m=0}^{n-1} \frac{1}{m^\nu} \sum_{k=1}^{m^\nu} \widehat{F}_m^{-1}\left(2^{-m} \left\lfloor \frac{2^m(k-1)}{m^\nu} \right\rfloor\right) \right]^{p/2} = A_p 2^{n(1-pr_{\nu,n}^+)}.$$

Moreover, applying Jensen's inequality to (2.12), we get

$$\langle x \rangle_n^{(p)} \leq B_p 2^{n(1-p)} n^{p/2} \mathbb{E} \left[\left(\frac{1}{n} \sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \leq B_p 2^{n(1-p)} n^{p/2-1} \sum_{m=0}^{n-1} \mathbb{E}[S_m^{p/2}].$$

We once again apply (2.48) to this inequality and obtain

$$\begin{aligned} \langle x \rangle_n^{(p)} &\leq B_p 2^{n(1-p)} n^{p/2-1} \left(\sum_{m=0}^{n-1} \frac{1}{m^\nu} \sum_{k=1}^{m^\nu} \left[\widehat{F}_m^{-1}\left(2^{-m} \left\lceil \frac{2^m k}{m^\nu} \right\rceil\right) \right]^{p/2} \right) \\ &\leq B_p 2^{n(1-p)} n^{(\nu+1)(p/2-1)} \left[\sum_{m=0}^{n-1} \frac{1}{m^\nu} \sum_{k=1}^{m^\nu} \widehat{F}_m^{-1}\left(2^{-m} \left\lceil \frac{2^m k}{m^\nu} \right\rceil\right) \right]^{p/2} \\ &= B_p 2^{n(1-pr_{\nu,n}^- + (\nu+1)(p/2-1) \log_2 n/n)}. \end{aligned}$$

Moreover, a similar inequality can be obtained for $1 \leq p \leq 2$,

$$A_p 2^{n(1-pr_{\nu,n}^+ + (\nu+1)(p/2-1) \log_2 n/n)} \leq \langle x \rangle_n^{(p)} \leq B_p 2^{n(1-pr_{\nu,n}^-)}.$$

In both cases ($p \geq 2$ or $1 \leq p \leq 2$), the exponents inside the brackets converge to $1 - pr_\nu^\pm$ as $n \uparrow \infty$. Therefore, $\lim_n \langle x \rangle_n^{(p)} = \infty$ for $p < 1/r_\nu^+$ and $\lim_n \langle x \rangle_n^{(p)} = 0$ for $p > 1/r_\nu^-$. This leads to $r_\nu^- \leq R \leq r_\nu^+$, and concludes the proof. $\square$

The next result looks at the special case of functions x of bounded variation. Such functions clearly have the roughness exponent $R = 1$. The converse statement, however, is not true: there exist functions with $R = 1$ that are nowhere differentiable and hence not of bounded variation; see Example 2.8.

Proposition 2.37. *The following assertions hold:*

- (a) *If $\sup_n s_n < \infty$, then the function x is of bounded variation.*

(b) If the function x is of bounded variation, then $\sup_n 2^{-n/2} \sum_{k=0}^{2^n-1} |\theta_{n,k}| < \infty$.

Proof of Proposition 2.37. (a): It follows from Proposition 2.10 that there exists $B_1 > 0$ such that

$$\langle x \rangle_n^{(1)} \leq B_1 \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{1/2} \right] \leq B_1 \left(\sum_{m=0}^{n-1} \mathbb{E}[S_m] \right)^{1/2} = B_1 s_n.$$

Therefore, $\sup_n \langle x \rangle_n^{(1)} \leq B_1 \sup_n s_n < \infty$, and $\sup_n \langle x \rangle_n^{(1)}$ coincides with the total variation of the continuous function x (see, e.g., Theorem 2 in §5 of Chapter VIII in [79]).

(b): Following from Proposition 2.10, we have

$$A_1 \left(2^{-n/2} \sum_{k=0}^{2^n-1} |\theta_{n,k}| \right) = A_1 \mathbb{E}[S_n^{1/2}] \leq A_1 \mathbb{E} \left[\left(\sum_{m=0}^{n-1} S_m \right)^{1/2} \right] \leq \langle x \rangle_n^{(1)},$$

for some $A_1 > 0$. Taking suprema on both sides of the above inequality completes our proof. $\square$

Our final result in this section concerns the value of the roughness exponent of the sum of two continuous functions.

Lemma 2.38. *Suppose that $x, y \in C[0, 1]$ admit respectively the roughness exponents R_x and R_y such that $R_x < R_y$. Then $x + y$ admits the roughness exponent R_x .*

Proof. For simplicity, for any continuous function $f \in C[0, 1]$, we introduce the following generic notation

$$\Delta \mathbf{f}_n = \left(f \left(\frac{1}{2^n} \right) - f(0), f \left(\frac{2}{2^n} \right) - f \left(\frac{1}{2^n} \right), \dots, f(1) - f \left(\frac{2^n-1}{2^n} \right) \right) \in \mathbb{R}^{2^n}.$$

Then, for $n \in \mathbb{N}$ and $p \geq 1$, we have $\langle f \rangle_n^{(p)} = \|\Delta \mathbf{f}\|_{\ell_p}^p$. Moreover, the triangular inequality yields that

$$\left(\left| \|\Delta \mathbf{x}_n\|_{\ell_p} - \|\Delta \mathbf{y}_n\|_{\ell_p} \right| \right)^p \leq \langle x + y \rangle_n^{(p)} = \|\Delta \mathbf{x}_n + \Delta \mathbf{y}_n\|_{\ell_p}^p \leq \left(\|\Delta \mathbf{x}_n\|_{\ell_p} + \|\Delta \mathbf{y}_n\|_{\ell_p} \right)^p.$$

By definition, for any $p > \frac{1}{R_x} > \frac{1}{R_y}$, it gives

$$\lim_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \lim_{n \uparrow \infty} \langle y \rangle_n^{(p)} = 0 \quad (2.49)$$

Thus, we have

$$\lim_{n \uparrow \infty} \|\Delta \mathbf{x}_n\|_{\ell_p} = \lim_{n \uparrow \infty} \|\Delta \mathbf{y}_n\|_{\ell_p} = 0.$$

Thus, for $p > \frac{1}{R_x}$, we have

$$\lim_{n \uparrow \infty} \langle x + y \rangle_n^{(p)} \leq \lim_{n \uparrow \infty} \left(\|\Delta \mathbf{x}_n\|_{\ell_p} + \|\Delta \mathbf{y}_n\|_{\ell_p} \right)^p = 0.$$

Next, for $\frac{1}{R_y} < p < \frac{1}{R_x}$, we have

$$\lim_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \lim_{n \uparrow \infty} \|\Delta \mathbf{x}_n\|_{\ell_p}^p = \infty \quad \text{and} \quad \lim_{n \uparrow \infty} \langle y \rangle_n^{(p)} = \lim_{n \uparrow \infty} \|\Delta \mathbf{y}_n\|_{\ell_p}^p = 0.$$

Hence, we have

$$\lim_{n \uparrow \infty} \langle x + y \rangle_n^{(p)} \geq \lim_{n \uparrow \infty} \left(\left| \|\Delta \mathbf{x}_n\|_{\ell_p} - \|\Delta \mathbf{y}_n\|_{\ell_p} \right| \right)^p = \infty. \quad (2.50)$$

As the quantity $\langle x + y \rangle_n^{(p)}$ is strictly decreasing in p for a sufficiently large n , therefore, the above identity naturally extends to $p < \frac{1}{R_x}$. Combining the identities in (2.49) and (2.50) completes the proof. $\square$

2.8 Model-free estimation of the roughness exponent

In this section, we will introduce and analyze model-free estimators for the roughness exponent. To this end, recall that $\widehat{R}_n(x) = 1 - \frac{1}{n} \log_2 s_n$ is a consistent estimator for the roughness exponent of any function $x \in C[0, 1]$ whose Faber–Schauder coefficients $(\theta_{m,k})$ satisfy the reverse Jensen condition. Figure 2.1 illustrates that the estimator performs very well for sample paths of fractional Brownian motion and yields accurate estimates of the corresponding Hurst parameter.

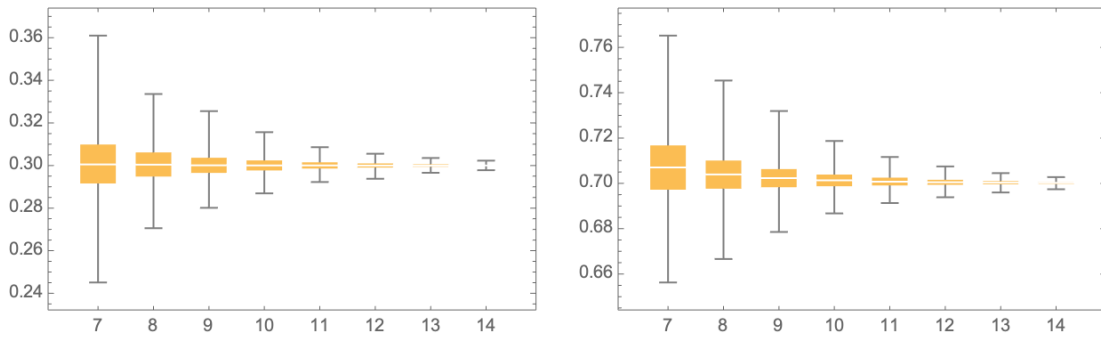


Figure 2.1: Box plots of $\widehat{R}_n$ for $n = 7, \dots, 14$ based on 10,000 sample paths of fractional Brownian motion with $R = 0.3$ (left) and $R = 0.7$ (right).

But the estimator $\widehat{R}_n$ also has a disadvantage: it is not scale-invariant. Indeed, we clearly

have

$$\widehat{R}_n(\lambda x) - \widehat{R}_n(x) = -\frac{\log_2 |\lambda|}{n} \quad \text{for } \lambda \neq 0. \quad (2.51)$$

As a consequence, multiplying x with a constant factor can lead to negative or unreasonably large estimates for R and substantially slow down or speed up the convergence $\widehat{R}_n(x) \rightarrow R$ (note that both the roughness exponent itself and the reverse Jensen condition are scale-invariant). Naive normalization of the data does not resolve this issue, as illustrated in Figure 2.2. In the next subsection, we are therefore going to derive intrinsically defined optimal scaling factors and the corresponding improved estimators for the roughness exponent.

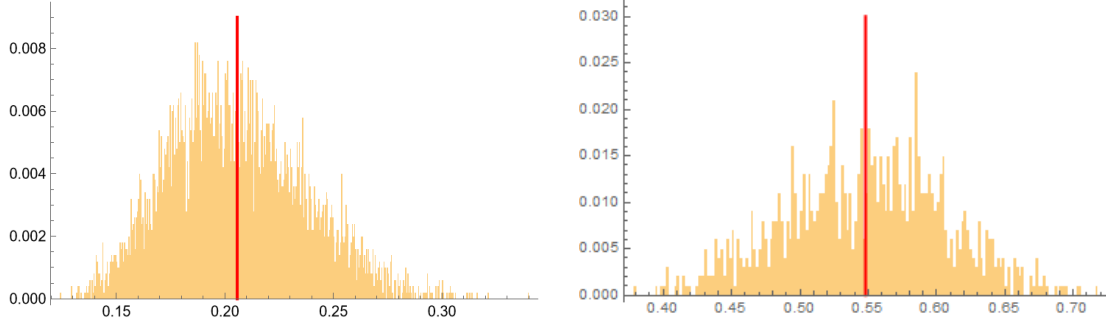


Figure 2.2: Histograms of $\widehat{R}_{12}$ applied to 5000 standardized sample paths of fractional Brownian motion with Hurst parameters $H = 0.3$ (left) and $H = 0.7$ (right). Here, standardization means that each sample path is affinely scaled to have empirical mean zero and variance one. The red vertical lines represent the empirical means of the corresponding estimates.

When presented with a generic time series, we face significant uncertainty about possible data scaling that, as discussed above, can lead to significant underperformance of the estimator $\widehat{R}_n$. In this section, we will therefore analyze two ideas on constructing scale-invariant estimators for the roughness exponent out of the sequence $(\widehat{R}_n)$. Both ideas will lead to new estimators that are linear combinations of $\widehat{R}_n, \dots, \widehat{R}_{n-m}$ for a number m .

The first idea consists in looking for a scaling factor λ_n that optimizes a certain criterion over all possible parameters $\lambda > 0$ and then to take $\widehat{R}_n(\lambda_n x)$ as a new estimator. In particular, we will look at the following two choices. A third choice will be given in Definition 2.40.

Definition 2.39. Fix $m \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_m \geq 0$ with $\alpha_0 > 0$.

- (a) For $n > m$, the *sequential scaling factor* λ_n^s and the *sequential scale estimate* $R_n^s(x)$ are defined as follows,

$$\lambda_n^s := \arg \min_{\lambda > 0} \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{R}_k(\lambda x) - \widehat{R}_{k-1}(\lambda x) \right)^2 \quad \text{and} \quad R_n^s(x) := \widehat{R}_n(\lambda_n^s x). \quad (2.52)$$

The corresponding mapping $R_n^s : C[0, 1] \rightarrow \mathbb{R}$ will be called the *sequential scale estimator*.

- (b) The *terminal scaling factor* λ_n^t and the *terminal scale estimate* $R_n^t(x)$ are defined as follows,

$$\lambda_n^t := \arg \min_{\lambda > 0} \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{R}_n(\lambda x) - \widehat{R}_{k-1}(\lambda x) \right)^2 \quad \text{and} \quad R_n^t(x) := \widehat{R}_n(\lambda_n^t x). \quad (2.53)$$

The corresponding mapping $R_n^t : C[0, 1] \rightarrow \mathbb{R}$ will be called the *terminal scale estimator*.

We will see in Proposition 2.41 that Definition 2.39 is well posed, because the minimization problems in (2.52) and (2.53) admit unique solutions for every function $x \in C[0, 1]$. The intuition for sequential scaling is fairly simple. The sequential scaling factor λ_n^s minimizes the weighted mean-squared differences $\widehat{R}_k(\lambda x) - \widehat{R}_{k-1}(\lambda x)$ for $k = m + 1, \dots, n$. Since $\widehat{R}_n$ is a consistent estimator, the faster the sequence $\widehat{R}_k(\lambda x)$ converges, the more accurate the final estimate $\widehat{R}_n(\lambda x)$ will be. The idea for terminal scaling is also straightforward: Since the estimator $\widehat{R}_n$ is consistent, the last estimate $\widehat{R}_n(x)$ should be more precise than any other previous value $\widehat{R}_k(x)$ with $k < n$. Therefore, we aim at minimizing the weighted mean-squared differences $\widehat{R}_k(\lambda x)$ and $\widehat{R}_n(\lambda x)$.

Our second idea for constructing a scale-invariant estimator is as follows. It follows from (2.51) that for $\lambda > 0$

$$n\widehat{R}_n(x) = n\widehat{R}_n(\lambda x) + \log_2 \lambda.$$

For an ideal scaling factor λ , the right-hand side will be close to $nR + \log_2 |\lambda|$. Therefore, the idea is to perform linear regression on the data points $(n-m)\widehat{R}_{n-m}(x), (n-m-1)\widehat{R}_{n-m-1}(x), \dots, n\widehat{R}_n(x)$ and to take the corresponding slope as an estimator for R .

Definition 2.40. Let $m \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_m \geq 0$ be such that $\sum_{k=0}^m \alpha_k = 1$. Then, for $n > m$, the *regression scale estimate* $R_n^r(x)$ and the *regression scaling factor* λ_n^r are given by

$$(R_n^r(x), \lambda_n^r) = \arg \min_{h \in \mathbb{R}, \lambda > 0} \sum_{k=0}^m \alpha_k \left((n-k)(\widehat{R}_{n-k}(x) - h) - \log_2 \lambda \right)^2. \quad (2.54)$$

The corresponding mapping $R_n^r : C[0, 1] \rightarrow \mathbb{R}$ will be called the *regression scale estimator*.

The following proposition shows in particular that all three estimators can be represented by linear combinations of the estimators $\widehat{R}_k$ for $m \leq k \leq n$ and that they are scale-invariant estimators of the roughness exponent R .

Proposition 2.41. *Consider the context of Definitions 2.39 and 2.40 with fixed $m \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_m \geq 0$ such that $\alpha_0 > 0$.*

- (a) *The three optimization problems in (2.52), (2.53), and (2.54) admit unique solutions for every function $x \in C[0, 1]$. In particular, all objects in Definitions 2.39 and 2.40 are well defined.*
- (b) *The sequential and terminal scale estimators can be represented by the following respective linear combinations of the estimators $\hat{R}_k$,*

$$R_n^s = \beta_{n,n} \hat{R}_n + \dots + \beta_{n,n-m-1} \hat{R}_{n-m-1} \quad \text{and} \quad R_n^t = \gamma_{n,n} \hat{R}_n + \dots + \gamma_{n,n-m-1} \hat{R}_{n-m-1},$$

where

$$\beta_{n,k} = \begin{cases} 1 + \frac{\alpha_0}{c_n^s n^2 (n-1)} & \text{if } k = n, \\ \frac{1}{c_n^s n k} \left(\frac{\alpha_{n-k}}{k-1} - \frac{\alpha_{n-k-1}}{k+1} \right) & \text{if } n-m \leq k \leq n-1, \\ \frac{-\alpha_m}{c_n^s n(n-m)(n-m-1)} & \text{if } k = n-m-1, \end{cases} \quad \text{for } c_n^s := \sum_{k=n-m}^n \frac{\alpha_{n-k}}{k^2 (k-1)^2},$$

and

$$\gamma_{n,k} = \begin{cases} 1 + \frac{1}{c_n^t n^2} \sum_{j=n-m}^n \alpha_{n-j} \frac{n-j+1}{j-1} & \text{if } k = n, \\ \frac{k-n}{c_n^t n^2 k} \alpha_{n-k-1} & \text{otherwise,} \end{cases} \quad \text{for } c_n^t = \sum_{k=n-m}^n \alpha_{n-k} \left(\frac{n-k+1}{n(k-1)} \right)^2.$$

If in addition $\sum_k \alpha_k = 1$, then the regression scale estimator is given by

$$R_n^r = \frac{1}{c^r} \sum_{k=0}^m \alpha_k (n-k)(k-a) \hat{R}_{n-k}, \quad (2.55)$$

where

$$a = \sum_{k=0}^m \alpha_k k \quad \text{and} \quad c^r = a^2 - \sum_{k=0}^m \alpha_k k^2.$$

- (c) *The sequential, terminal, and regression scale estimators are scale-invariant. That is, for $n > m$, $x \in C[0, 1]$, and $\lambda \neq 0$, we have $R_n^t(\lambda x) = R_n^t(x)$, $R_n^s(\lambda x) = R_n^s(x)$ and $R_n^r(\lambda x) = R_n^r(x)$.*
- (d) *If $x \in C[0, 1]$ is such that there exists $\lambda \neq 0$ for which $|\hat{R}_n(\lambda x) - R| = O(a_n)$ as $n \uparrow \infty$ for some sequence (a_n) with $a_n = o(1/n)$, then $|R_n^s(x) - R|$, $|R_n^t(x) - R|$, and $|R_n^r(x) - R|$ are all of the order $O(na_n)$.*

Proof of Proposition 2.41. We prove (a) and (b) together. Fix $x \in C[0, 1]$. For $\lambda > 0$ we write $\phi := \log_2 \lambda$ and we will minimize (2.52) and (2.53) over ϕ rather than over λ . Recall from (2.51) that

$$\widehat{R}_k(2^\phi x) = \widehat{R}_k(x) - \frac{\phi}{k}. \quad (2.56)$$

One therefore sees that the objective functions in (2.52) and (2.53) are strictly convex in ϕ . Thus, minimizers can be computed as the unique zeros of the corresponding derivatives. Differentiating the objective function in (2.52) and summing by parts gives

$$\begin{aligned} & \frac{1}{2} \frac{\partial}{\partial \phi} \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{R}_k(2^\phi x) - \widehat{R}_{k-1}(2^\phi x) \right)^2 \\ &= c_n^s \phi + \frac{\alpha_0 \widehat{R}_n(x)}{n(n-1)} + \sum_{k=n-m}^{n-1} \widehat{R}_k(x) \left(\frac{\alpha_{n-k}}{k(k-1)} - \frac{\alpha_{n-k-1}}{(k+1)k} \right) - \frac{\alpha_m \widehat{R}_{n-m-1}(x)}{(n-m)(n-m-1)}. \end{aligned} \quad (2.57)$$

Setting this expression equal to zero yields the global minimizer ϕ_n^s ,

$$\phi_n^s = -\frac{1}{c_n^s} \left(\frac{\widehat{R}_n(x)}{n} \cdot \frac{\alpha_0}{n-1} + \sum_{k=n-m}^{n-1} \frac{\widehat{R}_k(x)}{k} \left(\frac{\alpha_{n-k}}{k-1} - \frac{\alpha_{n-k-1}}{k+1} \right) - \frac{\widehat{R}_{n-m-1}(x)}{n-m-1} \cdot \frac{\alpha_m}{n-m} \right).$$

Applying (2.56) to $\widehat{R}_n(2^{\phi_n^s} x)$ yields the asserted expression for the sequential scale estimator. The proof for the terminal scale estimator is analogous. The assertion for the regression scale estimator follows by standard computations for linear regression.

(c) For $\eta > 0$, we have

$$\arg \min_{\lambda > 0} \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{R}_k(\lambda \eta x) - \widehat{R}_{k-1}(\lambda \eta x) \right)^2 = \frac{1}{\eta} \arg \min_{\lambda > 0} \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{R}_k(\lambda x) - \widehat{R}_{k-1}(\lambda x) \right)^2 = \frac{\lambda_n^s}{\eta},$$

where λ_n^s is the sequential scaling factor. Therefore, $R_n^s(\eta x) = \widehat{R}_n(\lambda_n^s x) = R_n^s(x)$. The same argument yields the scale invariance of the terminal and regression scale estimators.

(d) Clearly, we can assume without loss of generality that $\lambda = 1$. First, we consider the sequential scale estimator. Using (2.57) one sees that $\sum_k \beta_{n,k} = 1$ so that

$$R_n^s - R = \beta_{n,n}(\widehat{R}_n - R) + \cdots + \beta_{n,n-m-1}(\widehat{R}_{n-m-1} - R).$$

Since $\beta_{n,n-k} = O(n)$ for each k , the assertion follows. The proof for the terminal scale estimator is analogous. For the regression scale estimator, one checks that $\frac{1}{c^r} \sum_{k=0}^m \alpha_k(n-k)(k-a) = 1$ and proceeds as before. $\square$

Part (d) of Proposition 2.41 implies in particular the consistency of the sequential, ter-

minal, and regression scale estimators if $x \in C[0, 1]$ satisfies $|\widehat{R}_n(\lambda x) - R| = o(1/n)$ for some $\lambda \neq 0$. It also enables us to obtain the convergence rate of the estimators R_n^s , R_n^t and R_n^r for the sample paths of fractional Brownian motion with drift.

Corollary 2.42. *Suppose that X is a fractional Brownian motion with drift as in (2.44) and satisfies all the assumptions in Theorem 2.31. Then the following almost sure rates of convergence hold for the scale-invariant estimator R_n^p with $p \in \{s, t, r\}$,*

$$|R_n^p(X) - H| = \begin{cases} \mathcal{O}\left(2^{-n/2}\sqrt{\log n}\right) & \text{if } H \in (0, \frac{1}{2}), \\ \mathcal{O}\left(2^{-n/2}n^{1/2}\sqrt{\log n}\right) & \text{if } H = \frac{1}{2}, \\ \mathcal{O}\left(2^{(H-1)n}\sqrt{\log n}\right) & \text{if } H \in (\frac{1}{2}, 1). \end{cases} \quad (2.58)$$

Remark 2.43. A general scale-invariant estimator $\widetilde{R}_n$ can be constructed by solving an optimization problem of the form

$$\widetilde{\lambda} := \arg \min_{\lambda > 0} \sum_{j > k}^n \alpha_{j,k} |\widehat{R}_j(\lambda x) - \widehat{R}_k(\lambda x)|^2,$$

where $\alpha_{j,k}$ are given coefficients, and by setting $\widetilde{R}_n(x) := \widehat{R}_n(\widetilde{\lambda}x)$. If the global minimizer $\widetilde{\lambda}$ exists, the estimate $\widetilde{R}_n(x)$ will be a linear combination of the estimates $\widehat{R}_k(x)$.

The sequential, terminal, and regression scale estimators can be analogously derived based on the estimator $\widehat{R}_n^*$, following the construction outlined in Definitions 2.39 and 2.40. We denote those scale-invariant estimators by $R^{*,s}$, $R^{*,t}$ and $R^{*,r}$ respectively. Moreover, all properties established in Proposition 2.41 hold true for the scale-invariant estimators $R^{*,s}$, $R^{*,t}$ and $R^{*,r}$. Applying assertion (d) of Proposition 2.41 to Theorem 2.34 directly yields the following corollary. It demonstrates that the bias in Theorem 2.34 will be removed without deteriorating the corresponding rate of convergence.

Corollary 2.44. *Suppose that X is a fractional Brownian motion with drift as in (2.44) and satisfies all the assumptions in Theorem 2.31. Then the following almost sure rates of convergence hold for the scale-invariant estimator $R_n^{*,p}$ with $p \in \{s, t, r\}$,*

$$|R_n^{*,p}(X) - H| = \mathcal{O}\left(2^{-n/2}\sqrt{\log n}\right). \quad (2.59)$$

Let us now investigate the relations between our estimators and the one used in [40]. To

this end, we let for $x \in C[0, 1]$, $n, k \in \mathbb{N}$ and $q \in \mathbb{R}^-$,

$$m(q, k, n) := \left\lfloor \frac{k}{2^n} \right\rfloor^{-1} \sum_{j=1}^{\lfloor 2^n/k \rfloor} |x(kj2^{-n}) - x(k(j-1)2^{-n})|^q.$$

Following Rosenbaum [85], Gatheral et al. [40] assume that there exists a positive constant b_q such that

$$(k2^{-n})^{-qR} m(q, k, n) \rightarrow b_q, \quad \text{for any fixed } k \in \mathbb{N}, \text{ as } n \uparrow \infty,$$

From here, the estimator R_n^v from [40] is computed by way of a linear regression. More precisely, let $\mathcal{K}$ be a finite collection of positive integers and $\mathcal{Q}$ be a finite collection of positive real numbers. For $q \in \mathcal{Q}$, the estimate $R_{n,q}^v(x)$ of the roughness exponent of the continuous function x using the q^{th} variation is obtained by regressing $\log_2 m(q, k, n)$ with respect to $\log_2 k$, i.e.,

$$(R_{n,q}^v(x), b_q) = \arg \min_{h \in \mathbb{R}, b_q > 0} \sum_{k \in \mathcal{K}} \left(\log_2 m(q, k, n) - \log_2 b_q - qh(\log_2 k - n) \right)^2. \quad (2.60)$$

Standard computations for linear regression then yield that

$$R_{n,q}^v(x) = \frac{\sum_{k \in \mathcal{K}} (\log_2 k - \overline{\log_2 \mathcal{K}}) (\log_2 m(q, k, n) - \overline{\log_2 m(q, \mathcal{K}, n)})}{q \cdot \sum_{k \in \mathcal{K}} (\log_2 k - \overline{\log_2 \mathcal{K}})},$$

where $|\cdot|$ denotes the cardinality of a set and

$$\overline{\log_2 \mathcal{K}} = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \log_2 k \quad \text{and} \quad \overline{\log_2 m(q, \mathcal{K}, n)} = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \log_2 m(q, k, n).$$

The *simple regression estimate* $R_n^v(x)$ is then the sample average of $R_{n,q}^v(x)$ for all $q \in \mathcal{Q}$. In other words, we have

$$R_n^v(x) := \frac{1}{|\mathcal{Q}|} \sum_{q \in \mathcal{Q}} R_{n,q}^v(x).$$

The corresponding mapping $R_n^v : C[0, 1] \rightarrow \mathbb{R}$ will be called the *simple regression estimator*, which is a formal description of the estimator used in [40]. It is also clear that the simple regression estimator R_n^v is scale-invariant, i.e., for any $n \in \mathbb{N}$ and $x \in C[0, 1]$, we have $R_n^v(\lambda x) = R_n^v(x)$ for any $\lambda \neq 0$. Our next result states that the simple regression estimator is a particular case of our regression scale estimator under a certain choice of parameters.

Proposition 2.45. *Let $\mathcal{Q} = \{2\}$, $\mathcal{K} = \{1, 2, \dots, 2^m\}$ and $\alpha_k = \alpha$ for all k and some $\alpha > 0$. Then, the regression scale estimator R_n^r coincides with the simple regression estimator R_n^v*

for all $n > m$.

Proof of Proposition 2.45. Since the minimization problems in (2.54) and (2.60) admit unique solutions for every function $x \in C[0, 1]$, it suffices to show (2.54) is equivalent to (2.60) when taking the above parameters. Let $\lambda = 1/\sqrt{b_2}$, and we rewrite optimization problem in (2.60) into

$$\begin{aligned} & \arg \min_{h \in \mathbb{R}, b_q > 0} \sum_{k \in \mathcal{K}} \left(\log_2 m(2, k, n) - \log_2 b_2 - 2h(\log_2 k - n) \right)^2 \\ &= \arg \min_{h \in \mathbb{R}, \lambda > 0} \sum_{j=0}^m \left(2(j - n) + 2\log_2 s_{n-j} + 2\log_2 \lambda - 2h(j - n) \right)^2 \\ &= \arg \min_{h \in \mathbb{R}, b_q > 0} \sum_{j=0}^m \alpha \cdot \left((n - j)(\hat{R}_{n-j}(x) - h) - \log_2 \lambda \right)^2, \end{aligned}$$

where the second identity follows from [76, Proposition 2.1]. $\square$

Unfortunately, there is no rule of thumb in choosing the weighting parameters (α_m) and the type of scale-invariant estimators. As can be seen from the following figures, the performance of the scale-invariant estimators depends significantly on the selection of weighting parameters (α_m) and the actual underlying roughness exponent.

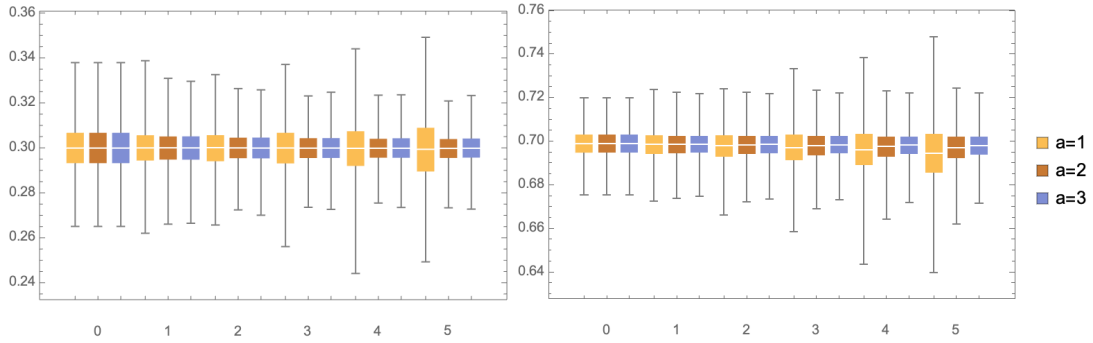


Figure 2.3: Box plots of R_n^s with $\alpha_k = a^{-k}$ applied to 1000 sample paths of fractional Brownian motion with Hurst parameters $H = 0.3$ (left) and $H = 0.7$ (right) with $n = 14$ and $m = 0, 1, \dots, 5$.

2.9 The roughness exponent for unequally spaced partitions

In this section, we will relax our assumption that the data points for sampling the p^{th} variation $\langle x \rangle_n^{(p)}$ of x are located at the n^{th} dyadic partition of the time interval $[0, 1]$. To this

end, let us first recall that a *partition* of the interval $[0, 1]$ is a finite set $\mathbb{T} = \{t_0, \dots, t_m\}$ such that $0 = t_0 < t_1 < \dots < t_m = 1$. Its *mesh* is defined as the maximal distance between two neighboring partition points. A sequence $(\mathbb{T}_n)_{n \in \mathbb{N}_0}$ of partitions of $[0, 1]$ is called a *refining sequence of partitions of $[0, 1]$* if $\mathbb{T}_0 \subset \mathbb{T}_1 \subset \dots$ and the mesh of $\mathbb{T}_n$ tends to 0 as $n \uparrow \infty$. In this case, we will generically write $\mathbb{T}_n = \{t_0^n, t_1^n, \dots, t_{m_n}^n\}$. For simplicity, we will henceforth assume that $\mathbb{T}_0 = \{0, 1\}$ and that, when passing from $\mathbb{T}_n$ to $\mathbb{T}_{n+1}$, exactly one new partition point is added between any two neighboring partition points of $\mathbb{T}_n$, i.e., $|(t_{k-1}^n, t_k^n) \cap \mathbb{T}_{n+1}| = 1$. This assumption can always be satisfied by re-arranging the partitions; in this case, we have $m_n = 2^n$. Furthermore, under this assumption, it follows $t_k^n = t_{2^k}^{n+1}$.

In Section 2.1, the definition of the roughness exponent of a continuous function x was based on the Faber–Schauder expansion of x . The Faber–Schauder functions are clearly tied to the sequence of dyadic partitions, but, similar to Cont and Das [17], the definition of the Faber–Schauder functions can be extended to our present setup as follows. First, let $b_0(t) := t$. Next, for $n \in \mathbb{N}_0$ and $k \in \{0, \dots, 2^n - 1\}$, we define

$$b_{n,k}(t) := \begin{cases} \frac{t - t_{2^k}^{n+1}}{t_{2^{k+1}}^{n+1} - t_{2^k}^{n+1}} & \text{if } t \in [t_{2^k}^{n+1}, t_{2^{k+1}}^{n+1}), \\ \frac{t_{2^{k+2}}^{n+1} - t}{t_{2^{k+2}}^{n+1} - t_{2^{k+1}}^{n+1}} & \text{if } t \in [t_{2^{k+1}}^{n+1}, t_{2^{k+2}}^{n+1}), \\ 0 & \text{otherwise.} \end{cases}$$

It is shown in [17, Section 3.3] that any function $x \in C[0, 1]$ can be represented by the following uniformly convergent series,

$$x(t) = f(0) + (x(1) - x(0))t + \sum_{m=0}^{\infty} \sum_{k=0}^{2^m-1} \theta_{m,k} b_{m,k}(t), \quad t \in [0, 1], \quad (2.61)$$

where

$$\theta_{n,k} := x(t_{2^{k+1}}^{n+1}) - \frac{t_{2^{k+2}}^{n+1} - t_{2^{k+1}}^{n+1}}{t_{2^{k+2}}^{n+1} - t_{2^k}^{n+1}} x(t_{2^k}^{n+1}) - \frac{t_{2^{k+1}}^{n+1} - t_{2^k}^{n+1}}{t_{2^{k+2}}^{n+1} - t_{2^k}^{n+1}} x(t_{2^{k+2}}^{n+1}) \quad (2.62)$$

will be called the *generalized Faber–Schauder coefficients of x* . For $p > 0$ and $n \in \mathbb{N}$, we define the p^{th} variation of the function x along $\mathbb{T}_n$ by

$$\langle x \rangle_n^{(p)} := \sum_{k=0}^{2^n-1} |x(t_{k+1}^n) - x(t_k^n)|^p.$$

If there exists $q \in [1, \infty]$ such that

$$\lim_{n \uparrow \infty} \langle x \rangle_n^{(p)} = \begin{cases} 0 & p > q, \\ \infty & p < q, \end{cases}$$

the *roughness exponent* of x (with respect to $(\mathbb{T}_n)_{n \in \mathbb{N}_0}$) is defined as $R = 1/q$. The following definition formulates the reverse Jensen condition for the irregular Faber–Schauder coefficients defined in (2.62).

Definition 2.46. We say that the coefficients $(\theta_{m,k})$ of x satisfy the *reverse Jensen condition* if for each $p \geq 1$, there exists a non-decreasing subexponential function $\varrho_p : \mathbb{N} \rightarrow [1, \infty)$ such that

$$\begin{aligned} \frac{1}{\varrho_p(n)} \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^{m-1}-1} \frac{\theta_{m,k}^2}{t_{2k+1}^m - t_{2k}^m} + \frac{\theta_{m,k}^2}{t_{2k+2}^m - t_{2k+1}^m} \right)^{p/2} &\leq \left(\sum_{k=0}^{2^n-1} \sum_{m=0}^{n-1} \frac{\theta_{m, \lfloor 2^{m-n+1}k \rfloor}^2}{t_{\lfloor 2^{m-n}k \rfloor + 1}^m - t_{\lfloor 2^{m-n}k \rfloor}^m} \right)^{p/2} \\ &\leq \varrho_p(n) \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^{m-1}-1} \frac{\theta_{m,k}^2}{t_{2k+1}^m - t_{2k}^m} + \frac{\theta_{m,k}^2}{t_{2k+2}^m - t_{2k+1}^m} \right)^{p/2}, \quad n \in \mathbb{N}. \end{aligned}$$

To formulate our extension of Theorem 2.4 to unequally spaced partitions, we need an additional condition on our partitions sequence. Let us denote

$$\bar{\pi}_n := \sup_k |t_{k+1}^n - t_k^n| \quad \text{and} \quad \underline{\pi}_n := \inf_k |t_{k+1}^n - t_k^n|.$$

We will say that the sequence $(\mathbb{T}_n)_{n \in \mathbb{N}_0}$ is *well-balanced* if there exists a non-decreasing subexponential function κ such that

$$\sup_{k \leq n} \frac{\bar{\pi}_k}{\underline{\pi}_k} \leq \kappa(n), \quad n \in \mathbb{N}_0.$$

The similar but slightly different concept of a *balanced* partition sequence was introduced earlier by Cont and Das [18, 17]. Furthermore, for $n \in \mathbb{N}$, we now denote likewise

$$s_n := \sqrt{\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} 2^m \theta_{m,k}^2} \quad \text{and} \quad \hat{R}_n(x) = 1 - \frac{1}{n} \log_2 s_n.$$

The next theorem shows that the roughness exponent defined over a sequence of well-balanced partitions, its value can be obtained from the sequence $(s_n)_{n \in \mathbb{N}}$.

Theorem 2.47. *Suppose that the generalized Faber–Schauder coefficients of x satisfy the reverse Jensen condition and the partition sequence $(\mathbb{T}_n)_{n \in \mathbb{N}_0}$ is well-balanced. Then, the*

function x admits the roughness exponent R if and only if the finite limit $r := \lim_n \widehat{R}_n(x)$ exists, and in that case, we have $R = r$.

Finally, let us show that the main result of dyadic partitions can be directly recovered from Theorem 2.47. Suppose $(\mathbb{T}_n)_{n \in \mathbb{N}_0}$ is the sequence of dyadic partitions and the function $x \in C[0, 1]$ admits the Faber–Schauder coefficients $(\mu_{m,k})$. Then it follows that

$$\mu_{m,k} = 2^{m/2+1} \theta_{m,k} \quad \text{and} \quad s_n = \sqrt{\frac{1}{2} \sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \mu_{m,k}^2},$$

which then reproduces Theorem 2.4.

2.10 Proofs for the results in Chapter 2

2.10.1 Proofs for the results in Section 2.4

For the convenience of exposition in subsequent proofs, we introduce the following notation used in this section exclusively. For a centered Gaussian process X , we denote

$$\phi_{j,k}^{(m,n)} := \mathbb{E} \left[\left(X_{\frac{j}{2^m}} - X_{\frac{j-1}{2^m}} \right) \left(X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}} \right) \right], \quad (2.63)$$

for $m, n \in \mathbb{N}$, $1 \leq j \leq 2^m$ and $1 \leq k \leq 2^n$. In particular, when the arguments $m = n$, we denote $\phi_{j,k}^{(n)} := \phi_{j,k}^{(n,n)}$ in short. Moreover, we write

$$\psi_X(s, t) := \mathbb{E}((X_t - X_s)^2), \quad s, t \geq 0.$$

In this section, we assign the Gaussian process X to be either the trifractional Brownian motion $Z^{H,K}$ or the n^{th} -order fractional Brownian motion $W^{H,m}$.

The proofs of the above theorems intensively rely on proper upper bounds of $\phi_{j,k}^{(m,n)}$ for both processes. The following lemma derives upper bounds of $\phi_{j,k}^{(m,n)}$ for the trifractional Brownian motion, which are needed in the proofs of Theorem 2.20.

Lemma 2.48. *For $H \in (0, 1)$ and $K \in (0, 1)$, we denote*

$$\phi_{j,k}^{(m,n)} := \mathbb{E} \left[\left(Z_{\frac{j}{2^m}}^{H,K} - Z_{\frac{j-1}{2^m}}^{H,K} \right) \left(Z_{\frac{k}{2^n}}^{H,K} - Z_{\frac{k-1}{2^n}}^{H,K} \right) \right].$$

The following inequalities hold:

(a) For $1 < j \leq 2^m$ and $1 < k \leq 2^n$,

$$\phi_{j,k}^{(m,n)} \leq L_1 2^{-(n+m)HK} (k-1)^{HK-1} (j-1)^{HK-1}, \quad (2.64)$$

where $L_1 = 2^K(1-K)KH^2 > 0$.

(b) For $1 < j \leq 2^m$, $1 < k \leq 2^n$ and $H \geq 1/2$,

$$\phi_{j,k}^{(m,n)} \leq L_2 (jk)^{2H-1} \min \left(\frac{(j-1)^{2HK-4H}}{2^{2Hn+(2HK-2H)m}}, \frac{(k-1)^{2HK-4H}}{2^{2Hm+(2HK-2H)n}} \right), \quad (2.65)$$

and for $1 < j \leq 2^m$, $1 < k \leq 2^n$ and $H < 1/2$,

$$\phi_{j,k}^{(m,n)} \leq L_2 [(j-1)(k-1)]^{2H-1} \min \left(\frac{(j-1)^{2HK-4H}}{2^{2Hn+(2HK-2H)m}}, \frac{(k-1)^{2HK-4H}}{2^{2Hm+(2HK-2H)n}} \right), \quad (2.66)$$

where $L_2 = 4K(1-K)H^2 > 0$.

(c) For $m, n \in \mathbb{N}$ and $1 < k \leq 2^n$,

$$\phi_{1,k}^{(m,n)} \leq L_3 (k-1)^{2HK-1-2H} 2^{-2Hm+(2H-2HK)n}, \quad (2.67)$$

where $L_3 = 2HK(1-K) > 0$.

(d) For $m, n \in \mathbb{N}$ and $1 < k \leq 2^n$,

$$\phi_{1,k}^{(m,n)} \leq 2HK(k-1)^{2HK-1} 2^{-2HKn}. \quad (2.68)$$

Proof of Lemma 2.48. First, we prove inequality (2.64). For $s, t > 0$, we get

$$\frac{\partial^2 C_{H,K}}{\partial s \partial t}(s, t) = 4K(1-K)H^2(t^{2H} + s^{2H})^{K-2}(st)^{2H-1} \leq 2^K K(1-K)H^2(st)^{HK-1}. \quad (2.69)$$

For $1 < j \leq 2^m$ and $1 < k \leq 2^n$, applying (2.69) to the following integral representation leads to

$$\begin{aligned} \phi_{j,k}^{(m,n)} &= C_{H,K}\left(\frac{k}{2^n}, \frac{j}{2^m}\right) + C_{H,K}\left(\frac{k-1}{2^n}, \frac{j-1}{2^m}\right) - C_{H,K}\left(\frac{k}{2^n}, \frac{j-1}{2^m}\right) - C_{H,K}\left(\frac{k-1}{2^n}, \frac{j}{2^m}\right) \\ &= \int_{(k-1)/2^n}^{k/2^n} \int_{(j-1)/2^m}^{j/2^m} \frac{\partial^2 C_{H,K}}{\partial s \partial t}(s, t) ds dt \leq 2^{-(m+n)} \sup_{\substack{k-1 \leq 2^n s \leq k \\ j-1 \leq 2^m t \leq j}} \frac{\partial^2 C_{H,K}}{\partial s \partial t}(s, t) \\ &\leq 2^{-(m+n)} L_1 \left(\frac{(j-1)(k-1)}{2^{m+n}} \right)^{HK-1} = L_1 2^{-(m+n)HK} (k-1)^{HK-1} (j-1)^{HK-1}. \end{aligned}$$

The last inequality holds as $HK < 1$. Next, we prove (2.65) for $H \geq 1/2$. As $K-2 < 0$,

one has

$$\frac{\partial^2 C_{H,K}}{\partial s \partial t}(s, t) \leq L_2 \min [s^{2H(K-2)}, t^{2H(K-2)}] (st)^{2H-1}.$$

Therefore,

$$\begin{aligned} \phi_{j,k}^{(m,n)} &\leq L_2 2^{-(m+n)} \sup_{\substack{k-1 \leq 2^n s \leq k \\ j-1 \leq 2^m t \leq j}} \left(\min [s^{2H(K-2)}, t^{2H(K-2)}] (st)^{2H-1} \right) \\ &= L_2 2^{-2H(m+n)} \min \left[\left(\frac{j-1}{2^m} \right)^{2H(K-2)}, \left(\frac{k-1}{2^n} \right)^{2H(K-2)} \right] \left(\frac{jk}{2^{m+n}} \right)^{2H-1}. \end{aligned}$$

Rearranging this inequality gives (2.65), and the proof of (2.66) is analogous. Let us now prove (2.67). To this end, for each fixed $m \in \mathbb{N}$, we take $g(t) := C_{H,K}(1/2^m, t)$. Clearly, g is continuous over $[0, 1]$ and differentiable over $(0, 1)$. Furthermore, since $K < 1$, we have $[1 + (2^{-m}/t)^{2H}]^{K-1} \geq 1 + (K-1)(2^{-m}/t)^{2H}$. Thus,

$$0 < g'(t) = 2HKt^{2HK-1} \left[1 - \left[1 + \left(\frac{2^{-m}}{t} \right)^{2H} \right]^{K-1} \right] \leq L_3 t^{2HK-1-2H} 2^{-2Hm}.$$

By the mean value theorem, there exists $\tau \in [(k-1)/2^n, k/2^n]$ such that

$$\begin{aligned} \phi_{1,k}^{(m,n)} &= C_{H,K}\left(\frac{1}{2^m}, \frac{k}{2^m}\right) - C_{H,K}\left(\frac{1}{2^m}, \frac{k-1}{2^m}\right) = g'(\tau) 2^{-n} \\ &\leq L_3 (k-1)^{2HK-1-2H} 2^{-2Hm+(2H-2HK)n}, \end{aligned}$$

where the last inequality holds as $2HK-1-2H < 0$. This completes the proof of (2.67). Finally, as $g'(t) \leq 2HKt^{2HK-1}$, it leads to

$$\phi_{1,k}^{(m,n)} \leq 2HK \left(\frac{k-1}{2^n} \right)^{(2HK-1)n} 2^{-n} = 2HK (k-1)^{2HK-1} 2^{-2HKn}.$$

Thus, we complete the proof of (2.68). $\square$

Proof of Theorem 2.20. For $n \in \mathbb{N}$ and $\alpha \in \mathbb{R}$, we let $S_n^\alpha := 2^{\alpha n} \langle Z^{H,K} \rangle_n^{(2)}$ in short. As the trifractional Brownian motion is a centered Gaussian process, it then follows from Isserlis' theorem

$$\mathbb{E}(S_n^\alpha) = 2^{\alpha n} \sum_{k=1}^{2^n} \phi_{k,k}^{(n)} \quad \text{and} \quad \text{var}(S_n^\alpha) = 2^{2\alpha n+1} \sum_{j,k=1}^{2^n} (\phi_{j,k}^{(n)})^2. \quad (2.70)$$

In the sequel, we start by analyzing the limits of $\mathbb{E}(S_n^\alpha)$ and $\text{var}(S_n^\alpha)$ as $n \uparrow \infty$. Subsequently, we discuss case by case to prove (2.31) and (2.33). Finally, we prove (2.32) and (2.34). First,

applying (2.64) gives

$$\mathbb{E}(S_n^\alpha) = 2^{\alpha n} \left(\phi_{1,1}^{(n)} + \sum_{k=2}^{2^n} \phi_{k,k}^{(n)} \right) \leq 2^{(\alpha-2HK)n} \left((2-2^K) + L_1 \sum_{k=1}^{2^n-1} k^{2HK-2} \right). \quad (2.71)$$

Seeing that for $2HK-2 < -1$, one has $\sum_{k=1}^{\infty} k^{2HK-2} < \infty$. This implies that if $HK < 1/2$, we have $\mathbb{E}(S_n^\alpha) \downarrow 0$ for $\alpha < 2HK$ as $n \uparrow \infty$. For $HK > 1/2$, one has

$$\sum_{k=1}^{2^n-1} k^{2HK-2} \leq \int_0^{2^n} t^{2HK-2} dt = (2HK-1)^{-1} 2^{(2HK-1)n}. \quad (2.72)$$

Inequalities (2.71) and (2.72) then imply that for $HK > 1/2$, $\mathbb{E}(S_n^\alpha) \downarrow 0$ for $\alpha < 1$ as $n \uparrow \infty$. Last, for $HK = 1/2$, we have

$$\mathbb{E}(S_n^\alpha) \leq 2^{(\alpha-1)n} \left((2-2^K) + L_1 \sum_{k=1}^{2^n} \frac{1}{k} \right) = 2^{(\alpha-1)n} \left((2-2^K) + L_1 h_{2^n} \right),$$

where h_{2^n} denotes the 2^n -th *Harmonic number*, and $h_{2^n} \sim n \log 2$ as $n \uparrow \infty$. Thus, we have $\mathbb{E}(S_n^\alpha) \downarrow 0$ as $n \uparrow \infty$ for $\alpha < 1$ and $HK = 1/2$. Next, we discuss the upper bounds for $\text{var}(S_n^\alpha)$ case by case. By the Cauchy-Schwarz inequality, one has $(\phi_{1,k}^{(n)})^2 \leq \phi_{1,1}^{(n)} \phi_{k,k}^{(n)}$. Then,

$$\text{var}(S_n^\alpha) \leq 2^{2\alpha n+1} \left(\sum_{j,k=2}^{2^n} (\phi_{j,k}^{(n)})^2 + 2 \sum_{k=1}^{2^n} (\phi_{1,k}^{(n)})^2 \right) \leq 2^{2\alpha n+1} \left(\sum_{j,k=2}^{2^n} (\phi_{j,k}^{(n)})^2 + 2 \sum_{k=1}^{2^n} \phi_{1,1}^{(n)} \phi_{k,k}^{(n)} \right),$$

From (2.71), one get

$$\begin{aligned} 2^{2\alpha n} \sum_{k=1}^{2^n} \phi_{1,1}^{(n)} \phi_{k,k}^{(n)} &\leq 2^{(2\alpha-4HK)n} (2-2^K) \left((2-2^K) + L_1 \sum_{k=1}^{2^n-1} k^{2HK-2} \right) \\ &= \begin{cases} \mathcal{O}(2^{2(\alpha-2HK)n}) & \text{for } HK \in (0, 1/2), \\ \mathcal{O}(2^{2(\alpha-1)n}) & \text{for } HK = 1/2, \\ \mathcal{O}(2^{(2\alpha-2HK-1)n}) & \text{for } HK \in (1/2, 1), \end{cases} \end{aligned} \quad (2.73)$$

as $n \uparrow \infty$. Moreover, following (2.64), we have

$$\begin{aligned}
2^{2\alpha n} \sum_{j,k=2}^{2^n} (\phi_{j,k}^{(n)})^2 &\leq L_1^2 \sum_{j,k=1}^{2^n} 2^{(2\alpha-4HK)n} (jk)^{4HK-2} = \left(L_1 2^{(\alpha-2HK)n} \sum_{k=1}^{2^n} k^{2HK-1} \right)^2 \\
&= \begin{cases} \mathcal{O}(2^{2(\alpha-2HK)n}) & \text{for } HK \in (0, 1/2), \\ \mathcal{O}(2^{2(\alpha-1)n} n^2) & \text{for } HK = 1/2, \\ \mathcal{O}(2^{2(\alpha-1)n}) & \text{for } HK \in (1/2, 1), \end{cases} \quad (2.74)
\end{aligned}$$

as $n \uparrow \infty$. Subsequently, the results (2.31) and (2.33) follow directly from the Borel-Cantelli lemma. By virtue of the Borel-Cantelli lemma, the fast L^2 -convergence, i.e., $\sum_{n=0}^{\infty} \text{var}(S_n^\alpha) < \infty$, guarantees the almost sure convergence. For instance, inequalities (2.71), (2.73) and (2.74) imply that $\mathbb{E}(S_n^\alpha) \downarrow 0$ and $\sum_{n=0}^{\infty} \text{var}(S_n^\alpha) < \infty$ for $\alpha < 2HK$ and $HK < 1/2$. Therefore, S_n^α converges to zero almost surely. The proofs for $HK = 1/2$ and $HK \geq 1/2$ also follow from (2.71), (2.73) and (2.74) analogously. Thus, we complete the proof of (2.31) and (2.33).

Now, let us proceed to prove (2.32). There exists a (dependent) sequence of standard normally distributed random variables $\{Y_n, n \geq 1\}$, such that

$$Z_{\frac{1}{2^n}}^{H,K} = \sqrt{2-2^K} 2^{-HKn} Y_n. \quad (2.75)$$

Therefore, for $HK < \frac{1}{2}$, $\alpha > 2HK$ and any given $M > 0$, we get

$$\begin{aligned}
\mathbb{P}(S_n^\alpha \leq M) &\leq \mathbb{P}\left(2^{\alpha n} \left(Z_{\frac{1}{2^n}}^{H,K}\right)^2 \leq M\right) \\
&= \mathbb{P}\left(Y_n^2 \leq \frac{2^{(2HK-\alpha)n} M}{(2-2^K)}\right) = \frac{\gamma\left(1/2, \frac{2^{(2HK-\alpha)n} M}{2(2-2^K)}\right)}{\Gamma(1/2)},
\end{aligned}$$

where $\gamma(k, x) = \int_0^x t^{k-1} e^{-t} dt$ is the *lower incomplete gamma function*. By change of variables and Taylor expansion, we have

$$\frac{\gamma\left(1/2, \frac{2^{(2HK-\alpha)n} M}{2(2-2^K)}\right)}{\Gamma(1/2)} = \frac{2}{\sqrt{\pi}} \int_0^{\sqrt{\frac{2^{(2HK-\alpha)n} M}{2(2-2^K)}}} e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \sqrt{\frac{2^{(2HK-\alpha)n} M}{2(2-2^K)}} + \mathcal{O}(2^{(3HK-\frac{3}{2}\alpha)n}),$$

as $n \uparrow \infty$. As $2HK - \alpha < 0$, the Borel-Cantelli lemma implies that S_n^α converges to infinity almost surely. This completes the proof of (2.33).

Before proving equation (2.34), let us recall the definition of the Lei-Nualart process

$(G_t^K)_{t \geq 0}$. For $K \in (0, 1)$, the Lei-Nualart process is defined as

$$G_t^K = \int_0^\infty (1 - e^{-st}) s^{-\frac{1+K}{2}} dB_s, \quad (2.76)$$

where B is a standard Brownian motion. Lei and Nualart [67] show that the process G^K has trajectories which are infinitely differentiable on $(0, \infty)$ and absolutely continuous on $[0, \infty)$. As a result of the following construction in [67]

$$Z_t^{H,K} = \sqrt{\frac{K}{\Gamma(1-K)}} G_{t^{2H}}^K,$$

the trifractional Brownian motion $Z_t^{H,K}$ also admits infinitely differentiable trajectories on $(0, \infty)$. With slight abuse of notation, let us denote the first- and second-order derivatives of $Z_t^{H,K}$ by $(Z_t^{H,K})'$ and $(Z_t^{H,K})''$ respectively. For any $n \in \mathbb{N}$ and $k \geq 2$, we have

$$Z_{\frac{k}{2^n}}^{H,K} = Z_{\frac{k-1}{2^n}}^{H,K} + 2^{-n} \left(Z_{\frac{k-1}{2^n}}^{H,K} \right)' + 2^{-2n-1} \left(Z_{\frac{u_{k,n}}{2^n}}^{H,K} \right)'',$$

where $u_{k,n} \in (k-1, k)$. Hence,

$$\left(Z_{\frac{k}{2^n}}^{H,K} - Z_{\frac{k-1}{2^n}}^{H,K} \right)^2 = 2^{-2n} \left(\left(Z_{\frac{k-1}{2^n}}^{H,K} \right)' \right)^2 + \mathcal{O}(2^{-3n}).$$

As $(Z_t^{H,K})'$ is still continuously differentiable on $(0, \infty)$ by composition, we have for any $\varepsilon > 0$,

$$\begin{aligned} \lim_{n \uparrow \infty} \sum_{k=\lfloor 2^n \varepsilon \rfloor}^{2^n} 2^n \left(Z_{\frac{k}{2^n}}^{H,K} - Z_{\frac{k-1}{2^n}}^{H,K} \right)^2 &= \lim_{n \uparrow \infty} \sum_{k=\lfloor 2^n \varepsilon \rfloor}^{2^n} 2^{-n} \left(\left(Z_{\frac{k-1}{2^n}}^{H,K} \right)' \right)^2 + \mathcal{O}(2^{-n}) \\ &= \int_\varepsilon^1 \left(\left(Z_t^{H,K} \right)' \right)^2 dt. \end{aligned}$$

As $Z_t^{H,K}$ is non-constant almost surely, $\int_\varepsilon^1 \left(\left(Z_t^{H,K} \right)' \right)^2 dt$ is positive almost surely. As a

result, for $\alpha > 1$, we have

$$\begin{aligned} \lim_{n \uparrow \infty} S_n^\alpha &\geq \lim_{n \uparrow \infty} 2^{(\alpha-1)n} \sum_{k=\lfloor 2^n \epsilon \rfloor}^{2^n} 2^n \left(Z_{\frac{k}{2^n}}^{H,K} - Z_{\frac{k-1}{2^n}}^{H,K} \right)^2 \\ &= \lim_{n \uparrow \infty} 2^{(\alpha-1)n} \int_{\epsilon}^1 \left(\left(Z_t^{H,K} \right)' \right)^2 dt = \infty, \quad \mathbb{Q}\text{-a.s.} \end{aligned}$$

This completes the proof of the result (2.34). Hence, we complete the proof of Theorem 2.20. $\square$

Proof of Theorem 2.22. Throughout the following proof, we adopt the notations in (2.63) with the n -th order fractional Brownian motion. Now, we let $S_n^\alpha = 2^{\alpha n} \langle W^{H,m} \rangle$. Applying Taylor expansion with $z = 1/t$ yields

$$\begin{aligned} \psi_{W^{H,m}}(t, t-1) &= (-1)^m C_H^m \left(\sum_{j=0}^{m-1} (-1)^j \binom{2H}{j} \left(t^j (t-1)^{2H-j} + (t-1)^j t^{2H-j} - t^{2H} - (t-1)^{2H} \right) - 1 \right) \\ &= (-1)^m C_H^m \left(\sum_{j=0}^{m-1} (-1)^j \binom{2H}{j} z^{-2H} \left((1-z)^{2H-j} + (1-z)^j - 1 - (1-z)^{2H} \right) - 1 \right) \\ &= (-1)^m C_H^m \left(\sum_{j=0}^{m-1} (-1)^j \binom{2H}{j} z^{-2H} \left(j(j-2H)z^2 + \mathcal{O}(z^3) \right) - 1 \right) \\ &= (-1)^m C_H^m \left(2H(2H-1) \sum_{j=0}^{m-2} (-1)^j \binom{2H-2}{j} t^{2H-2} - 1 \right) + \mathcal{O}(t^{2H-3}) \\ &= C_{H-1}^{m-1} \binom{2H-3}{m-2} t^{2H-2} + (-1)^{m+1} C_H^m + \mathcal{O}(t^{2H-3}). \end{aligned}$$

As $m \geq 2$ and $H \in (m-1, m)$, we have $2H-2 > 0$, and this gives

$$\psi_{W^{H,m}}(t, t-1) \sim C_{H-1}^{m-1} \binom{2H-3}{m-2} t^{2H-2} \quad \text{as } t \uparrow \infty.$$

As the m^{th} order fractional Brownian motion is H -self-similar, then

$$\sum_{k=1}^{2^n} \phi_{k,k}^{(n)} = 2^{-2Hn} \sum_{k=1}^{2^n} \psi_{W^{H,m}}(k-1, k) = \mathcal{O}(2^{-n}) \quad \text{as } n \uparrow \infty. \quad (2.77)$$

Therefore, $\mathbb{E}(S_n^\alpha) = \mathcal{O}(2^{(\alpha-1)n})$ as $n \uparrow \infty$. Now, let us recall that the Cauchy-Schwarz

inequality gives $(\phi_{j,k}^{(n)})^2 \leq \phi_{j,j}^{(n)} \phi_{k,k}^{(n)}$, which leads to

$$\text{var}(S_n^\alpha) = 2^{2\alpha n+1} \sum_{j,k=1}^{2^n} (\phi_{j,k}^{(n)})^2 \leq 2^{2\alpha n+1} \left(\sum_{k=1}^{2^n} \phi_{k,k}^{(n)} \right)^2 = \mathcal{O}(2^{2(\alpha-1)n}).$$

This then gives the fast L^2 -convergence for the case $\alpha < 1$, which guarantees the almost sure convergence. Hence, the result in (2.35) is then verified.

The proof of (2.36) follows analogously as that of (2.34) in Theorem 2.20. Following [91, Remark 2.3], $W^{H,m}$ is $(m-1)$ times differentiable. Thus, for $m \geq 3$, the arguments in the proof Theorem 2.20 can be applied to the m^{th} order fractional Brownian motion. Hence, we complete the proof of Theorem 2.22. $\square$

The following proposition is needed to study limit theorems at the critical cases, i.e., $\alpha = 1$ in part (b) of Theorem 2.20 and $\alpha = 1$ in Theorem 2.22.

Proposition 2.49. *Let X be a centered Gaussian process. If*

$$\sup_{n \in \mathbb{N}} 2^n \mathbb{E} \left[\langle X \rangle_n^{(2)} \right] < \infty, \quad (2.78)$$

the weighted quadratic variation $2^n \langle X \rangle_n^{(2)}$ converges to some non-constant random variable in L^p as $n \uparrow \infty$ for all $p \in (1, \infty)$.

Proof. According to [94, Theorem 2.1], it suffices to show that the double limit

$$\lim_{n, m \uparrow \infty} 2^{n+m} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2$$

exists. For simplicity, let $a_{m,n} := 2^{n+m} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2$. The Cauchy-Schwarz inequality leads to

$$(\phi_{j,k}^{(m,n)})^2 \leq \mathbb{E} \left[\left(X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}} \right)^2 \right] \mathbb{E} \left[\left(X_{\frac{j}{2^m}} - X_{\frac{j-1}{2^m}} \right)^2 \right] = \phi_{j,j}^{(m)} \phi_{k,k}^{(n)}.$$

Therefore,

$$a_{m,n} \leq 2^{n+m} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} \phi_{k,k}^{(n)} \phi_{j,j}^{(m)} = \left(2^n \sum_{k=1}^{2^n} \phi_{k,k}^{(n)} \right) \left(2^m \sum_{j=1}^{2^m} \phi_{j,j}^{(m)} \right).$$

By the condition (2.78), the double sequence $(a_{m,n})_{m,n \in \mathbb{N}}$ is uniformly bounded. Thus, there exists $M > 0$ such that $\sup_{m,n} a_{m,n} \leq M$. Next, for any $m, n \in \mathbb{N}$, we have

$$(\phi_{j,k}^{(m,n)})^2 = \left(\phi_{j,2k}^{(m,n+1)} + \phi_{j,2k-1}^{(m,n+1)} \right)^2 \leq 2 \left((\phi_{j,2k}^{(m,n+1)})^2 + (\phi_{j,2k-1}^{(m,n+1)})^2 \right).$$

Therefore, one has

$$a_{m,n} = 2^{n+m} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2 \leq 2^{n+m+1} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} \left((\phi_{j,2k}^{(m,n+1)})^2 + (\phi_{j,2k-1}^{(m,n+1)})^2 \right) = a_{m,n+1}.$$

By an analogous argument, one also gets $a_{m,n} \leq a_{m+1,n}$. Therefore, for each fixed $m \in \mathbb{N}$, there exists a positive non-decreasing sequence (b_m) such that $\lim_{n \uparrow \infty} a_{m,n} = \lim_{n \uparrow \infty} a_{n,m} = b_m$. As $b_m \leq M$ for all $m \in \mathbb{N}$, there also exists a positive constant b such that $0 < a_{0,0} \leq b \leq M$ and $b_m \uparrow b$ as $m \uparrow \infty$.

Next, we will show the double limit $\lim_{m,n \uparrow \infty} a_{m,n}$ exists and is equal to b . For any $\epsilon > 0$, there exists $m_1(\epsilon)$ such that for all $m \geq m_1(\epsilon)$, we have $0 \leq b - b_m \leq \epsilon/2$. Moreover, let us select $n_1(\epsilon)$ such that for all $n \geq n_1(\epsilon)$, $0 \leq a_{m_1(\epsilon),n} - b_{m_1(\epsilon)} \leq \epsilon/2$. Taking $N(\epsilon) := \max[n_1(\epsilon), m_1(\epsilon)]$, one has for $m, n \geq N(\epsilon)$,

$$b - a_{m,n} \leq b - a_{m_1(\epsilon),N(\epsilon)} \leq (b - b_{m_1(\epsilon)}) + (b_{m_1(\epsilon)} - a_{m_1(\epsilon),N(\epsilon)}) \leq \epsilon.$$

The first inequality holds as the double sequence $a_{m,n}$ is increasing in each argument. Therefore, the double limit $\lim_{m,n \uparrow \infty} a_{m,n}$ exists and $\lim_{m,n \uparrow \infty} a_{m,n} = b > 0$. Thus, by virtue of [94, Theorem 2.1], the weighted sum $2^n \langle X \rangle_n^{(2)}$ converges to some non-constant random variable in L^p for all $p \in (1, \infty)$. This completes the proof. $\square$

Proof of Theorem 2.24. First, we prove assertion (a) for the case $2H < 1$, and the proof for the case $2H \geq 1$ is analogous. According to [94, Theorem 2.1], it suffices to show that the double limit

$$\lim_{n,m \uparrow \infty} 2^{2HK(n+m)} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2 \quad (2.79)$$

fails to exist. To this end, for $n, m \in \mathbb{N}$, we denote $a_{m,n} := 2^{2HK(n+m)} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2$. Next, we show the limit $\lim_{n \uparrow \infty} a_{n,n}$ exists and is strictly positive. By the Cauchy-Schwarz inequality, we have $(\phi_{j,k}^{(m,n)})^2 \leq \phi_{j,j}^{(m)} \phi_{k,k}^{(n)}$, then

$$a_{n,m} \leq 2^{2HK(n+m)} \sum_{k=1}^{2^n} \sum_{j=1}^{2^m} \phi_{k,k}^{(n)} \phi_{j,j}^{(m)} = \left(2^{2HKn} \sum_{k=1}^{2^n} \phi_{k,k}^{(n)} \right) \left(2^{2HKm} \sum_{j=1}^{2^m} \phi_{j,j}^{(m)} \right).$$

From (2.71), if $HK < 1/2$, the sequence $(\sum_{j=1}^{2^m} \phi_{j,j}^{(m)})_{m \in \mathbb{N}}$ is uniformly bounded. Hence, there exists $M > 0$, such that $\sup_{m,n} a_{m,n} \leq M$. As the trifractional Brownian motion is

HK -self-similar, then

$$a_{n,n} = \sum_{k=1}^{2^n} \sum_{j=1}^{2^n} (2^{2HKn} \phi_{j,k}^{(n)})^2 = \sum_{k=1}^{2^n} \sum_{j=1}^{2^n} \mathbb{E} \left[\left(Z_j^{H,K} - Z_{j-1}^{H,K} \right) \left(Z_k^{H,K} - Z_{k-1}^{H,K} \right) \right]^2.$$

Therefore, $(a_{n,n})_{n \in \mathbb{N}}$ forms a non-decreasing uniformly bounded sequence. Hence, the limit $\lim_{n \uparrow \infty} a_{n,n}$ exists and $\lim_{n \uparrow \infty} a_{n,n} \geq a_{0,0} = \mathbb{E}([Z^{H,K}(1)]^2) = (2 - 2^K) > 0$.

Now, we show that for each fixed $n \in \mathbb{N}$, $\lim_{m \uparrow \infty} a_{m,n} = \lim_{m \uparrow \infty} a_{n,m} = 0$. For $k \geq 2$, inequality (2.66) yields

$$2^{2HK(m+n)} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2 \leq (k-1)^{4HK-4H-2} 2^{(4H-2HK)n} 2^{(2HK-4H)m} \left(L_2^2 \sum_{j=1}^{2^m} j^{4H-2} + L_3^2 \right).$$

If $H < 1/4$, the sequence $(j^{4H-2})_{j \in \mathbb{N}} \in \ell_1$, and $2^{2HK(m+n)} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2$ converges to zero as $m \uparrow \infty$. If $H > 1/4$, one has

$$2^{(2HK-4H)m} \sum_{j=1}^{2^m} j^{4H-2} \leq 2^{(2HK-4H)m} \int_0^{2^m} t^{4H-2} dt = (4H-1)^{-1} 2^{(2HK-1)m}.$$

Therefore, $2^{2HK(m+n)} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2 = \mathcal{O}(2^{(2HK-1)m})$ converges to zero as $m \uparrow \infty$. Last, if $H = 1/4$, then $\sum_{j=1}^{2^m} j^{-1} = h_{2^m} \sim m \log 2$, where h_{2^m} is the 2^m -th Harmonic number. It then implies that $2^{2HK(m+n)} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2$ converges to zero. For $k = 1$, it follows from (2.68) that

$$\begin{aligned} 2^{2HK(m+n)} \sum_{j=2}^{2^m} (\phi_{j,1}^{(m,n)})^2 &\leq 2^{2HK(n-m)+2} H^2 K^2 \sum_{j=2}^{2^m} (j-1)^{4HK-2} \\ &= \begin{cases} \mathcal{O}(2^{-2HKm}) & \text{for } HK < 1/4, \\ \mathcal{O}(2^{-2HKm} m) & \text{for } HK = 1/4, \\ \mathcal{O}(2^{(2HK-1)m}) & \text{for } HK > 1/4, \end{cases} \end{aligned} \quad (2.80)$$

as $m \uparrow \infty$. Finally, for each $n \in \mathbb{N}$, we have

$$0 < 2^{2HK(m+n)} (\phi_{1,1}^{(m,n)})^2 \leq 2^{2HK(n-m)}. \quad (2.81)$$

The above inequalities in together imply that for each $n \in \mathbb{N}$ and $1 \leq k \leq 2^n$, we have

$$\lim_{m \uparrow \infty} 2^{HK(m+n)} \sum_{j=1}^{2^m} (\phi_{j,k}^{(m,n)})^2 = 0.$$

Hence, we have $\lim_{m \uparrow \infty} a_{m,n} = \lim_{m \uparrow \infty} a_{n,m} = 0$. However, as previously proved, $\lim_{n \uparrow \infty} a_{n,n} > 0$. Therefore, the double limit (2.79) does not exist, and we then complete the proof of assertion (a).

We now prove assertion (b). Following (2.71) and (2.72), we have

$$\sup_{n \in \mathbb{N}} 2^n \sum_{k=1}^{2^n} \phi_{k,k}^{(n)} \leq (2 - 2^K) + \frac{L_1}{2HK - 1}.$$

Then Proposition 2.49 implies the assertion (b). $\square$

Proof of Theorem 2.25. This theorem is a direct corollary of Proposition 2.49 and (2.77). $\square$

2.10.2 Proofs for the results in Section 2.6

To prove Theorem 2.30, we start with the following proposition, which provides the speed of convergence in Gladyshev's theorem [41] and is of possible independent interest.

Proposition 2.50. *For ε_n as in (2.43) and any $H \in (0, 1)$, there exists a constant c such that*

$$\limsup_{n \uparrow \infty} \frac{|2^{n(2H-1)} \langle W^H \rangle_n^{(2)} - 1|}{n\varepsilon_n} \leq c \quad \mathbb{Q}\text{-a.s.}$$

Proof. Let Z_n be the 2^n -dimensional centered Gaussian vector with components

$$Z_n^k = 2^{n(H-1/2)} (W_{k2^{-n}}^H - W_{(k-1)2^{-n}}^H).$$

Then $\|Z_n\|_{\ell_2}^2 = 2^{n(2H-1)} \langle W^H \rangle_n^{(2)}$. Denote by Γ_n the covariance matrix of Z_n and by T_n its trace. According to the following concentration inequality from [5, Lemma 3.1], there exists a constant λ such that for all n and each $\delta > 0$,

$$\mathbb{Q} \left[\left| \|Z_n\|_{\ell_2} - T_n \right| \geq \delta \right] \leq \lambda \exp \left(- \frac{\delta^2}{4\|\Gamma_n\|_2} \right),$$

where $\|\Gamma_n\|_2$ is the spectral norm of Γ_n . It was shown in [60, Equation (15)] that $\text{trace } \Gamma_n = 1$, and $\|\Gamma_n\|_2 \leq \gamma b_n$, where γ is a positive constant and $b_n = n^2 \varepsilon_n^2 / \log n$. Taking $\delta_n := 2\sqrt{2\gamma b_n \log n} = n\varepsilon_n \sqrt{8\gamma}$ yields $\mathbb{Q}[\|\|Z_n\|_{\ell_2} - 1| \geq \delta_n] \leq \lambda/n^2$, which is summable. The Borel–Cantelli lemma hence implies that $\|\|Z_n\|_{\ell_2} - 1| \leq \delta_n$ eventually with probability one, and so

$$\limsup_{n \uparrow \infty} \frac{|\sqrt{2^{n(2H-1)} \langle W^H \rangle_n^{(2)}} - 1|}{\delta_n} \leq 1 \quad \mathbb{Q}\text{-a.s.}$$

Using the fact that $\sqrt{1+x} - 1 \sim x/2$ for $x \rightarrow 0$ now yields the result. $\square$

Proof of Theorem 2.30. We start by establishing the reverse Jensen condition. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting the random variables (V_m) from (2.10). Then $(S_m)_{m \in \mathbb{N}} = (2^m \theta_{m, 2^m V_m}^2)_{m \in \mathbb{N}}$ is a sequence of random variables defined on the product space $(\Omega \times \tilde{\Omega}, \mathcal{F} \otimes \mathcal{G}, \mathbb{P} \otimes \mathbb{Q})$. According to (2.15), we must show that for every $p \geq 1$, with $\mathbb{Q}$ -probability one, there exists a subexponential function ϱ_p and $n_p \in \mathbb{N}$ such that for $n \geq n_p$,

$$\begin{aligned} \frac{1}{\varrho_p(n)} \mathbb{E} \left[2^{2n(H-1)} \sum_{m=0}^{n-1} S_m \right]^{p/2} &\leq \mathbb{E} \left[\left(2^{2n(H-1)} \sum_{m=0}^{n-1} S_m \right)^{p/2} \right] \\ &\leq \varrho_p(n) \mathbb{E} \left[2^{2n(H-1)} \sum_{m=0}^{n-1} S_m \right]^{p/2}, \end{aligned} \quad (2.82)$$

where the notation $\mathbb{E}$ denotes as usual the expectation with respect to $\mathbb{P}$. To this end, we note that $\mathbb{Q}$ -a.s. as $n \uparrow \infty$,

$$\begin{aligned} \mathbb{E} \left[2^{2n(H-1)} \sum_{m=0}^{n-1} S_m \right]^{p/2} &= \left(2^{n(2H-1)} 2^{-n} \sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k}^2 \right)^{p/2} \\ &= (2^{n(2H-1)} \langle W^H \rangle_n^{(2)})^{p/2} \longrightarrow 1, \end{aligned} \quad (2.83)$$

where the second identity follows from (2.13) and the convergence from Proposition 2.50. The self-similarity of fractional Brownian motion implies that

$$\begin{aligned} Y_n &:= 2^{n(pH-1)} \langle W^H \rangle_n^{(p)} = 2^{n(pH-1)} \sum_{k=0}^{2^n-1} |W_{(k+1)2^{-n}}^H - W_{k2^{-n}}^H|^p \\ &\sim 2^{-n} \sum_{k=0}^{2^n-1} |W_{k+1}^H - W_k^H|^p, \end{aligned} \quad (2.84)$$

where $\sim$ means equality in law. Since the increments $W_{k+1}^H - W_k^H$ are standard normally distributed, the $\mathbb{Q}$ -expectation of Y_n is bounded from above by a finite constant c , uniformly in n . It follows that $\mathbb{Q}[Y_n > n^2] \leq c/n^2$, and a Borel–Cantelli argument yields that $Y_n \leq n^2$ for sufficiently large n , $\mathbb{Q}$ -a.s. In conjunction with (2.12) and (2.83), this establishes the right-hand inequality in (2.82) with $\varrho_p(n) = 2n^2/A_p$. For the left-hand inequality, we take $q > p$ and note that by (2.84) and Jensen’s inequality,

$$Y_n^{-q} \sim \left(2^{-n} \sum_{k=0}^{2^n-1} |W_{k+1}^H - W_k^H|^p \right)^{-q} \leq 2^{-n} \sum_{k=0}^{2^n-1} |W_{k+1}^H - W_k^H|^{-p/q}.$$

Since the increments $W_{k+1}^H - W_k^H$ are standard normally distributed, the $\mathbb{Q}$ -expectations of Y_n^{-q} are also bounded from above, uniformly in n , as $-p/q > -1$. A similar Borel–Cantelli argument as above then yields that $Y_n \geq 1/n^{2q}$ for sufficiently large n , $\mathbb{Q}$ -a.s., and we obtain the left-hand inequality in (2.82) with $\varrho_p(n) = 2B_p n^{2q}$. Thus, both inequalities in (2.82) hold, e.g., with $\varrho_p(n) := 2(A_p^{-1} \vee B_p)n^{2(p+1)}$. This establishes the reverse Jensen condition.

Next, it follows from Propositions 2.50 and 2.16 that $\widehat{R}_n(W^H) \rightarrow H$ $\mathbb{Q}$ -a.s., and so Theorem 2.4 yields that the sample paths admit the roughness exponent H with probability one.

Finally, we prove the rate of convergence (2.43). Using (2.13), one checks that

$$H - \widehat{R}_n(W^H) = \frac{1}{2n} \log_2 2^{n(2H-1)} \langle W^H \rangle_n^{(2)}.$$

Now the assertion follows from Proposition 2.50 and the fact that $\log_2(1+x) \sim x/\log 2$ for $x \rightarrow 0$. $\square$

To prove Theorem 2.31 and Theorem 2.32, we assume without loss of generality that $x_0 = 0$. Furthermore, we may assume without loss of generality that $\xi_0 = 0$. To see why, we rewrite (2.44) as

$$X_t = \left(W_t^H + \xi_0 t\right) + \int_0^t (\xi_s - \xi_0) ds. \quad (2.85)$$

Theorem 4.1 in [80] states that the law of W^H is equivalent to the one of $(W_t^H + \xi_0 t)_{0 \leq t \leq T}$. Since absolute continuity is a transitive relation, our claim follows.

Our approach to proving Theorem 2.31 combines the Kadota–Shepp Theorem [59] with the transformation theory developed by Norros et al. [80]. The proof of Theorem 2.32 further requires the result concerning the law equivalence of Brownian motion with drift by Yershov [96].

We briefly summarize the results in Kadota and Shepp [59] and Yershov [96]. The proofs of the above results can be found in [68, Chapter 7].

Theorem 2.51 (Kadota and Shepp). *Let X be an Itô process of the form*

$$X_t = W_t + \int_0^t \xi_s ds, \quad 0 \leq t \leq T, \quad (2.86)$$

for some progressively measurable process ξ . If $\int_0^T \xi_t^2 dt < \infty$ $\mathbb{Q}$ -a.s., then the law of $(X_t)_{0 \leq t \leq T}$ is absolutely continuous with respect to the law of $(W_t)_{0 \leq t \leq T}$.

Theorem 2.52 (Yershov). *Suppose that X is a diffusion process satisfying the stochastic integral equation*

$$X_t = W_t + \int_0^t \gamma(s, X) ds, \quad 0 \leq t \leq T, \quad (2.87)$$

where γ is progressively measurable with respect to the natural filtration of W . If $\int_0^T \gamma(t, X)^2 dt$ and $\int_0^T \gamma(t, W)^2 dt$ are both $\mathbb{Q}$ -a.s. finite, then the law of $(X_t)_{0 \leq t \leq T}$ is equivalent to the one of $(W_t)_{0 \leq t \leq T}$

Let us now present the results in [80] in a nutshell. Following the notion therein, we define the following constants,

$$c_H = \sqrt{\frac{2H\Gamma(3/2 - H)}{\Gamma(H + 1/2)\Gamma(2 - 2H)}}, \quad c_1 = \frac{1}{2H\Gamma(\frac{3}{2} - H)\Gamma(H + \frac{1}{2})}, \quad c_2 = \frac{c_H}{2H(2 - 2H)^{1/2}}.$$

We also let

$$w(t, s) = c_1 s^{1/2-H} (t - s)^{1/2-H}, \quad 0 \leq s \leq t \leq T.$$

It is shown in [80] that the following integrals exist $\mathbb{Q}$ -a.s. in a pathwise sense,

$$Y_t = \int_0^t s^{1/2-H} dW_s^H \quad \text{and} \quad M_t = \int_0^t w(t, s) dW_s^H. \quad (2.88)$$

Moreover, Theorem 3.1 in [80] states that M is a continuous Gaussian martingale with quadratic variation $\langle M \rangle_t = c_2^2 t^{2-2H}$. In Theorem 3.2 and Remark 3.1 of [80], it is furthermore proved that Y and M satisfy the following relations,

$$M_t = c_1 \int_0^t (t - s)^{1/2-H} dY_s \quad \text{and} \quad Y_t = 2H \int_0^t (t - s)^{H-1/2} dM_s. \quad (2.89)$$

It follows that the process

$$B_t := \frac{2H}{c_H} \int_0^t s^{H-1/2} dM_s \quad (2.90)$$

is a standard Brownian motion. Theorem 5.2 in [80] states that W^H can be recovered from B as follows,

$$W_t^H = \int_0^t \zeta(t, s) dB_s, \quad (2.91)$$

where for $0 \leq s < t \leq T$,

$$\zeta(t, s) = c_H \left[\left(\frac{t}{s} \right)^{H-1/2} (t - s)^{H-1/2} - (H - 1/2) s^{1/2-H} \int_s^t u^{H-3/2} (u - s)^{H-1/2} du \right]. \quad (2.92)$$

It follows that there is a one-to-one correspondence between W^H and B . In particular, the respective natural filtrations coincide up to sets of measure zero. Our proof is based on extending these arguments to the process X by employing analogous transforms on its drift term $\int_0^t \xi_s ds$.

In the sequel, we shall work in a strictly pathwise manner for a fixed sample path ξ that satisfies our assumptions in Theorem 2.31. Under these assumptions, we may define in analogy to (2.88),

$$\eta_t = \int_0^t s^{1/2-H} \xi_s ds \quad \text{and} \quad \mu_t = \int_0^t w(t, s) \xi_s ds. \quad (2.93)$$

The function η is clearly absolutely continuous with derivative $\eta'_s = s^{1/2-H} \xi_s$, and so we have

$$\mu_t = c_1 \int_0^t (t-s)^{1/2-H} s^{1/2-H} \xi_s ds = c_1 \int_0^t (t-s)^{1/2-H} \eta'_s ds, \quad (2.94)$$

which is the analogue of the first relation in (2.89). An analogue of the second relation in (2.89) is proved in the following lemma.

Lemma 2.53. *Under the assumptions in Theorem 2.31 and the additional assumption $\xi_0 = 0$, the function $t \mapsto \mu_t$ is absolutely continuous with bounded derivative μ'_t and we have*

$$\eta_t = 2H \int_0^t (t-s)^{H-1/2} \mu'_s ds. \quad (2.95)$$

The proof of Lemma 2.53 needs the following short lemma.

Lemma 2.54. *If $f : [0, T] \rightarrow \mathbb{R}$ is Hölder continuous with exponent $\beta \in (0, 1]$, satisfies $f(0) = 0$, and $\alpha \in (0, \beta)$, then $\tilde{f}(t) := t^{-\alpha} f(t)$ is Hölder continuous with exponent $\beta - \alpha$.*

Proof. Let $c \geq 0$ be such that $|f(t) - f(s)| \leq c|t - s|^\beta$ for all $s, t \in [0, 1]$. Let $0 \leq s < t \leq T$ be given. If $s = 0$, then $\tilde{f}(s) = 0$ and $|\tilde{f}(t) - \tilde{f}(s)| = t^{-\alpha} |f(t)| \leq ct^{\beta-\alpha}$. Now suppose that $s > 0$. Using the inequalities $|f(t) - f(s)| \leq c|t - s|^{\beta-\alpha} |t - s|^\alpha \leq c|t - s|^{\beta-\alpha} t^\alpha$ and $|f(s)| = |f(s) - f(0)| \leq cs^\beta$ yields

$$|\tilde{f}(t) - \tilde{f}(s)| \leq t^{-\alpha} |f(t) - f(s)| + (t^{-\alpha} - s^{-\alpha}) |f(s)| \leq c|t - s|^{\beta-\alpha} + c|t^{-\alpha} - s^{-\alpha}| s^\beta.$$

For the rightmost term, we have

$$\begin{aligned} |t^{-\alpha} - s^{-\alpha}|s^\beta &= \alpha \int_s^t s^\beta r^{-\beta} r^{\beta-\alpha-1} dr \leq \alpha \int_s^t r^{\beta-\alpha-1} dr \\ &= \frac{\alpha}{\beta-\alpha} (t^{\beta-\alpha} - s^{\beta-\alpha}) \leq \frac{\alpha}{\beta-\alpha} |t-s|^{\beta-\alpha}. \end{aligned}$$

This completes the proof. $\square$

Proof of Lemma 2.53. We start with the case $H < 1/2$, in which the derivative η' is bounded. Equation (2.94), our assumption $H < 1/2$, and the boundedness of η'_t imply that $t \mapsto \mu_t$ is also absolutely continuous with bounded derivative

$$\mu'_t = c_1(1/2 - H) \int_0^t (t-s)^{-H-1/2} \eta'_s ds.$$

Thus,

$$\begin{aligned} \int_0^t (t-s)^{H-1/2} \mu'_s ds &= c_1(1/2 - H) \int_0^t (t-s)^{H-1/2} \int_0^s (s-r)^{-H-1/2} \eta'_r dr ds \\ &= c_1(1/2 - H) \int_0^t \eta'_r \int_r^t (t-s)^{H-1/2} (s-r)^{-H-1/2} ds dr. \end{aligned}$$

After the change of variables $s = r + u(t-r)$, the inner integral becomes

$$(t-r)^{1+H-1/2-H-1/2} \int_0^1 (1-u)^{H-1/2} u^{-H-1/2} du = \Gamma(1/2 + H) \Gamma(1/2 - H).$$

This gives (2.95) in case $H < 1/2$.

For $H > 1/2$, recall that we assume that ξ_t is Hölder continuous with some exponent $\alpha > 2H - 1$. Let I_{0+}^β denote the fractional Riemann–Liouville integral operator of order $\beta > 0$, i.e.,

$$I_{0+}^\beta f(t) = \frac{1}{\Gamma(\beta)} \int_0^t \frac{f(s)}{(t-s)^{1-\beta}} ds, \quad t \in (0, T], \quad f \in L^1[0, T].$$

Furthermore, the fractional Riemann–Liouville derivative of order $\beta \in (0, 1)$ is given by

$$\mathcal{D}_{0+}^\beta f(t) = \frac{1}{\Gamma(1-\beta)} \frac{d}{dt} \int_0^t \frac{f(s)}{(t-s)^\beta} ds, \quad 0 < t < T,$$

provided the expression on the right makes sense. When defining μ_t as in (2.94) and letting $\tilde{\xi}_t := t^{1/2-H} \xi_t$, we can write

$$\mu_t = c_1 \Gamma(3/2 - H) I_{0+}^{3/2-H} \tilde{\xi}(t). \quad (2.96)$$

By Lemma 2.54, $\tilde{\xi}_t$ is Hölder continuous with exponent $\lambda := \alpha + 1/2 - H > H - 1/2$. It follows from [86, Lemma 13.2] that $\tilde{\xi} \in I_{0+}^{H-1/2}(L^1[0, T])$. In addition, (2.96) and [86, Theorem 2.3] yield that μ_t is absolutely continuous. Its derivative is given by

$$\mu'_t = c_1 \frac{d}{dt} \int_0^t \frac{\tilde{\xi}_s}{(t-s)^{H-1/2}} ds = c_1 \Gamma(3/2 - H) \mathcal{D}_{0+}^{H-1/2} \tilde{\xi}(t). \quad (2.97)$$

To see that μ' is bounded, we use [86, Lemma 13.1] to conclude that there exists a function ψ that is Hölder continuous with exponent $\lambda - (H - 1/2) > 0$ such that

$$\mu'_t = c_1 \Gamma(3/2 - H) \mathcal{D}_{0+}^{H-1/2} \tilde{\xi}(t) = \frac{c_1 \tilde{\xi}_t}{t^{H-1/2}} + \psi(t) = c_1 \xi_t + \psi(t).$$

Hence μ'_t is Hölder continuous and in particular bounded.

By [86, Theorem 2.4], we can recover $\tilde{\xi}$ from μ' as follows,

$$c_1 \Gamma(3/2 - H) \tilde{\xi}_t = I_{0+}^{H-1/2} \mu'(t) = \frac{1}{\Gamma(H - 1/2)} \int_0^t \frac{\mu'_s}{(t-s)^{3/2-H}} ds. \quad (2.98)$$

Integrating with respect to t and using Fubini's theorem yields,

$$\begin{aligned} \eta_t &= \int_0^t \tilde{\xi}_s ds = \frac{1}{c_1 \Gamma(3/2 - H) \Gamma(H - 1/2)} \int_0^t \int_0^r \frac{\mu'_s}{(r-s)^{3/2-H}} ds dr \\ &= \frac{1}{c_1 \Gamma(3/2 - H) \Gamma(H - 1/2)} \cdot \frac{1}{H - 1/2} \int_0^t (t-s)^{H-1/2} \mu'_s ds. \end{aligned}$$

This is (2.95) in our case $H > 1/2$, because one checks that the constant in front of the integral is equal to $2H$. $\square$

Proof of Theorem 2.31. Note that Lemma 2.53 allows us to make the following definition, which is analogous to (2.90),

$$\beta_t := \frac{2H}{c_H} \int_0^t s^{H-1/2} \mu'_s ds. \quad (2.99)$$

By using the two identities (2.94) and (2.95) we can argue exactly as in the proof of [80, Theorem 5.2] to get that

$$\int_0^t \xi_s ds = \int_0^t \zeta(t, s) \beta'_s ds. \quad (2.100)$$

From (2.89) and Lemma 2.53 we now get the pathwise identity

$$Y_t^X := \int_0^t s^{1/2-H} dX_s = Y_t + \eta_t.$$

which shows that Y^X is a semimartingale with decomposition $Y_t^X = Y_t + \eta_t$. Furthermore, we have

$$M_t^X := \int_0^t w(t, s) dY_s^X = M_t + \mu_t.$$

Finally, we let

$$B_t^X := \frac{2H}{c_H} \int_0^t s^{H-1/2} dM_s^X = B_t + \beta_t. \quad (2.101)$$

By (2.91) and (2.100), we have

$$X_t = \int_0^t \zeta(t, s) dB_s^X, \quad (2.102)$$

and so X can be recovered in a pathwise manner from B^X .

Next, we have

$$\int_0^T (\beta'_t)^2 dt = \frac{4H^2}{c_H^2} \int_0^T t^{2H-1} (\mu'_t)^2 dt < \infty, \quad \mathbb{Q}\text{-a.s.}, \quad (2.103)$$

because μ' is bounded by Lemma 2.53. Hence, Theorem 2.51 yields that the law of B^X is absolutely continuous with respect to the law of B . Therefore, the assertion of Theorem 2.31 follows from (2.91) and (2.102). $\square$

Proof of Theorem 2.52. For $H \leq 1/2$ it is obvious from Theorem 2.51 that the law of X is absolutely continuous with respect to that of W^H . For $H > 1/2$, we use the fact that the sample paths of W^H , and in turn those of X , are $\mathbb{Q}$ -a.s. Hölder continuous with exponent β for every $\beta < H$; see, e.g., [75, Section 1.15]. Therefore, the trajectories $t \mapsto b(X_t)$ are Hölder continuous with exponent $\alpha\beta$. By assumption, $\alpha H > 2H - 1$, and hence there exists $\beta < H$ such that $\alpha\beta > 2H - 1$ as well. Hence, Theorem 2.51 gives absolute continuity.

To prove that the law of $(x_0 + W_t^H)_{0 \leq t \leq T}$ is absolutely continuous with respect to the law of $(X_t)_{0 \leq t \leq T}$, we may once again assume without loss of generality that $x_0 = 0$; otherwise we replace X_t with $X_t - x_0$ and b with $b(\cdot + x_0)$. Moreover, as in (2.85), we may assume without loss of generality that $b(X_0) = b(x_0) = b(0) = 0$.

Our goal is to apply Theorem 2.52. We have seen in (2.101) that $B_t^X = B_t + \beta_t$. Using

in addition (2.99) and (2.102), we see that

$$B_t^X = B_t + \int_0^t \gamma(s, B^X) ds$$

where

$$\gamma(t, B^X) = \frac{2Hc_1}{c_H} t^{H-1/2} \frac{d}{dt} \int_0^t \frac{s^{1/2-H}}{(t-s)^{H-1/2}} b\left(\int_0^s \zeta(s, r) dB_r^X\right) ds. \quad (2.104)$$

We conclude from (2.91) that

$$\gamma(t, B) = \frac{2Hc_1}{c_H} t^{H-1/2} \frac{d}{dt} \int_0^t \frac{s^{1/2-H}}{(t-s)^{H-1/2}} b(W_s^H) ds.$$

In the proof of Theorem 2.31, all we needed to conclude that the quantities on the right-hand side of (2.104) are well defined and that $t \mapsto \beta'_t = \gamma(t, B^X)$ is $\mathbb{Q}$ -a.s. bounded were our conditions on b and the $\mathbb{Q}$ -a.s. Hölder continuity of the sample paths of X . Since the trajectories of W^H possess the same Hölder regularity, we conclude that also $t \mapsto \gamma(t, B)$ is $\mathbb{Q}$ -a.s. well defined and bounded. In particular, $\int_0^T \gamma(t, B)^2 dt < \infty$ $\mathbb{Q}$ -a.s. It follows that the conditions of Theorem 2.52 are satisfied so that the equivalence of the two laws is established. $\square$

Proof of Theorem 2.34. It is shown in Theorem 2.31 that the law of X is absolutely continuous with respect to the law of $W^H + x_0$. Hence, it suffices to prove this assertion for the fractional Brownian motion W^H . We denote the Faber–Schauder coefficients of W^H by $(\theta_{m,k})$. Now, let us consider the 2^n -dimensional centered Gaussian vector

$$Z_n := 2^{n(H-1)} (\theta_{n,0}, \theta_{n,1}, \dots, \theta_{n,2^n-1})^\top,$$

Rearranging terms in (2.24) yields that

$$\widehat{R}_n^*(W^H) = H - \frac{1}{n} \log_2 \|Z_n\|_{\ell_2}.$$

Denote by $\Gamma_n = (\gamma_{i,j}^{(n)})_{0 \leq i,j \leq 2^n-1}$ the covariance matrix of Z_n and by T_n its trace. In particular, the covariance matrix Γ_n is a Toeplitz matrix as the fractional Brownian motion W^H admits stationary increments. A simple computation yields that

$$\begin{aligned} \gamma_{i,j}^{(n)} &= 2^{(2H-1)n} \text{cov} \left(2W_{\frac{2i+1}{2^{n+1}}}^H - W_{\frac{i}{2^n}}^H - W_{\frac{i+1}{2^n}}^H, 2W_{\frac{2j+1}{2^{n+1}}}^H - W_{\frac{j}{2^n}}^H - W_{\frac{j+1}{2^n}}^H \right) \\ &= 2^{-n-2H-1} \left(-|2i-2j-2|^{2H} + 4|2i-2j-1|^{2H} - 6|2i-2j|^{2H} + 4|2i-2j+1|^{2H} - |2i-2j+2|^{2H} \right). \end{aligned}$$

Then, we get

$$T_n = \sum_{i=0}^{2^n-1} \gamma_{i,i}^{(n)} = 2^{2-2H} - 1 = \alpha_H^2.$$

Let i, j be fixed and take

$$f(t) = 2|2i - 2j + t|^{2H} - |2i - 2j + t - 1|^{2H} - |2i - 2j + t + 1|^{2H},$$

and we get

$$\gamma_{i,j}^{(n)} = 2^{-n-2H-1} (f(-1) - 2f(0) + f(1)) \leq 2^{-n} \sup_{t \in [-1,1]} |f''(t)|.$$

In a similar manner, we have

$$\begin{aligned} f''(t) &= 2H(2H-1) \left(2|2i - 2j + t|^{2H-2} - |2i - 2j + t - 1|^{2H-2} - |2i - 2j + t + 1|^{2H-2} \right) \\ &\leq \left| \prod_{i=0}^3 (2H-i) \right| \sup_{s \in [-1,1]} |2i - 2j + t + s|^{2H-4}. \end{aligned}$$

Combining the above inequalities yields that for $|i - j| \geq 2$, we get

$$\gamma_{i,j}^{(n)} \leq 2^{-n} \left| \prod_{i=0}^3 (2H-i) \right| (|2i - 2j - 2| \wedge |2i - 2j + 2|)^{2H-4}. \quad (2.105)$$

Let $\|\cdot\|_p$ denote the ℓ_p -induced operator norm. As Γ_n is a Toeplitz matrix and so $\gamma_{i,j}^{(n)} = \gamma_{j,i}^{(n)} = \gamma_{0,|j-i|}^{(n)}$. Hence, we have $\|\Gamma_n\|_1 = \|\Gamma_n\|_\infty$, and this gives

$$\|\Gamma_n\|_2 \leq \sqrt{\|\Gamma_n\|_1 \|\Gamma_n\|_\infty} = \|\Gamma_n\|_1 \leq 2 \sum_{i=0}^{2^n-1} \gamma_{0,i}^{(n)} \leq \gamma \cdot 2^{-n},$$

for some $\gamma > 0$, and the last inequality follows from (2.105). The concentration inequality from [5, Lemma 3.1] yields that there exists a constant λ such that for all n and $\delta > 0$,

$$\mathbb{Q} \left[\left| \|Z_n\|_{\ell_2} - \sqrt{T_n} \right| \geq \delta \right] = \mathbb{P} \left[\left| \|Z_n\|_{\ell_2} - \alpha_H \right| \geq \delta \right] \leq \lambda \exp \left(-\frac{\delta^2}{4\|\Gamma_n\|_2} \right)$$

Taking $\delta_n := 2^{-n/2} \sqrt{8\gamma \log n} = n\varepsilon_n \sqrt{8\gamma}$ yields $\mathbb{Q}[\|\|Z_n\|_{\ell_2} - \alpha_H\| \geq \delta_n] \leq \lambda/n^2$, then the Borel–Cantelli lemma yields that

$$\limsup_{n \uparrow \infty} \delta_n^{-1} \left| \frac{\|Z_n\|_{\ell_2}}{\alpha_H} - 1 \right| \leq \frac{1}{\alpha_H} \quad \mathbb{Q}\text{-a.s.}$$

Applying the identity

$$\frac{1}{n} \log_2 \frac{\|Z_n\|_{\ell_2}}{\alpha_H} = H - \widehat{R}_n^*(W^H) - \frac{1}{n} \log_2 \alpha_H,$$

and using the fact $\log_2(1+x) \sim x/\log 2$ as $x \rightarrow 0$ complete the proof. $\square$

2.10.3 Proofs for the results in Section 2.9

To prepare for the proof of Theorem 2.47, we define for $n \in \mathbb{N}$,

$$V_n := \sum_{k=0}^{2^n-1} \left(\frac{1}{t_{2k+1}^{n+1} - t_{2k}^{n+1}} \mathbb{1}_{[t_{2k}^{n+1}, t_{2k+1}^{n+1})} - \frac{1}{t_{2k+2}^{n+1} - t_{2k+1}^{n+1}} \mathbb{1}_{[t_{2k+1}^{n+1}, t_{2k+2}^{n+1})} \right). \quad (2.106)$$

Then V_n is equal to the right-hand derivative of the function $\sum_{k=0}^{2^n-1} b_{n,k}$. We consider each V_n as a random variable on the probability space $([0, 1], \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is the Borel σ -field of $[0, 1]$ and $\mathbb{P}$ is the Lebesgue measure on $\mathcal{F}$. When defining the σ -field

$$\mathcal{F}_n := \sigma([t_k^n, t_{k+1}^n) : k = 0, \dots, 2^n - 1),$$

one sees that V_n is $\mathcal{F}_{n+1}$ -measurable. Let us moreover define random variables $\vartheta_n : [0, 1] \rightarrow \mathbb{R}$ by

$$\vartheta_n := \sum_{k=0}^{2^n-1} \theta_{n,k} \mathbb{1}_{[t_k^n, t_{k+1}^n)}. \quad (2.107)$$

Clearly, the random variable ϑ_n is $\mathcal{F}_n$ -measurable. For each $n \in \mathbb{N}$, let x_n denote the n^{th} truncation of x , defined in analogy to (2.19). Then

$$x_n(t) = \int_0^t \sum_{m=0}^{n-1} \vartheta_m(s) V_m(s) ds.$$

Since ϑ_m and V_m are constant on intervals of the form $[t_k^n, t_{k+1}^n)$ for $n \geq m$, we have

$$x(t_{k+1}^n) - x(t_k^n) = x_n(t_{k+1}^n) - x_n(t_k^n) = (t_{k+1}^n - t_k^n) \sum_{m=0}^{n-1} \vartheta_m(t) V_m(t). \quad (2.108)$$

Hence, for $p \geq 1$,

$$\begin{aligned} \langle x \rangle_n^{(p)} &= \sum_{k=0}^{2^n-1} |x(t_{k+1}^n) - x(t_k^n)|^p = \sum_{k=0}^{2^n-1} \left| (t_{k+1}^n - t_k^n) \sum_{m=1}^n \vartheta_m(t_k^n) V_m(t_k^n) \right|^p \\ &= \sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n) \cdot \left| \gamma_{n,p}(t_k) \sum_{m=0}^{n-1} \vartheta_m(t_k) V_m(t_k) \right|^p = \mathbb{E} \left[\left| \gamma_{n,p} \sum_{m=0}^{n-1} \vartheta_m V_m \right|^p \right], \end{aligned}$$

where the expectation is taken under the Lebesgue measure $\mathbb{P}$ and

$$\gamma_{n,p}(t) := \sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{\frac{p-1}{p}} \mathbb{1}_{[t_{k+1}^n, t_k^n)}(t).$$

By definition, $\gamma_{n,p}$ is constant on all intervals of the form $[t_k^n, t_{k+1}^n)$ and hence $\mathcal{F}_n$ -measurable. Moreover, Hölder's inequality yields that

$$\begin{aligned} \mathbb{E}[\gamma_{n,p}^{-p}]^{-1} \cdot \mathbb{E} \left[\left| \sum_{m=0}^{n-1} \vartheta_m V_m \right|^{p/2} \right]^2 &\leq \mathbb{E} \left[\left| \gamma_{n,p} \sum_{m=0}^{n-1} \vartheta_m V_m \right|^p \right] \\ &\leq \mathbb{E}[\gamma_{n,p}^{2p}]^{1/2} \cdot \mathbb{E} \left[\left| \sum_{m=0}^{n-1} \vartheta_m V_m \right|^{2p} \right]^{1/2}. \end{aligned} \tag{2.109}$$

Now take $M_0 := 0$ and

$$M_n := \sum_{m=0}^{n-1} \vartheta_m V_m = \sum_{m=0}^{n-1} \vartheta_m (N_{m+1} - N_m), \quad n = 1, 2, \dots$$

Since for each m , the random variable ϑ_m is bounded and $\mathcal{F}_m$ -measurable, $(M_n)_{n=0,1,\dots}$ is a martingale transform of the martingale $(N_n)_{n=0,1,\dots}$ and hence itself a martingale. We first compute the expectations of the random variable $\gamma_{n,p}$ occurring on both bounds,

$$\begin{aligned} \mathbb{E}[\gamma_{n,p}^{2p}]^{1/2} &= \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{2(p-1)} \cdot (t_{k+1}^n - t_k^n) \right)^{1/2} = \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{2p-1} \right)^{1/2}. \\ \mathbb{E}[\gamma_{n,p}^{-p}]^{-1} &= \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{-(p-1)} \cdot (t_{k+1}^n - t_k^n) \right)^{-1} = \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{2-p} \right)^{-1}. \end{aligned}$$

Furthermore, the Burkholder inequality yields constants $0 < A_p \leq B_p < \infty$, depending only on p but not on x , such that

$$A_p \mathbb{E}[[M]_n^p] \leq \mathbb{E}[|M_n|^{2p}] \leq B_p \mathbb{E}[[M]_n^p],$$

where

$$[M]_n = \sum_{m=1}^n (M_m - M_{m-1})^2 = \sum_{m=0}^{n-1} \vartheta_m^2 V_m^2.$$

Note that we can also rephrase the reverse Jensen condition in Definition 2.46 in the same way as in (2.15):

$$\frac{1}{\varrho_p(n)} \mathbb{E} \left[\left(\sum_{m=0}^{n-1} V_m^2 \vartheta_m^2 \right)^{p/2} \right] \leq \mathbb{E} \left[\left(\sum_{m=0}^{n-1} V_m^2 \vartheta_m^2 \right)^{p/2} \right] \leq \varrho_p(n) \mathbb{E} \left[\left(\sum_{m=0}^{n-1} V_m^2 \vartheta_m^2 \right)^{p/2} \right].$$

It then follows from the reverse Jensen condition that

$$\frac{1}{\varrho_p(n)} \mathbb{E} [[M]_n]^p \leq \mathbb{E} [[M]_n^p] \leq \varrho_p(n) \mathbb{E} [[M]_n]^p.$$

Now, we get

$$\mathbb{E} [[M]_n] = \mathbb{E} \left[\sum_{m=0}^{n-1} \vartheta_m^2 V_m^2 \right] = \sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \frac{\theta_{m,k}^2}{t_{2k+1}^{m+1} - t_{2k}^{m+1}} + \frac{\theta_{m,k}^2}{t_{2k+2}^{m+1} - t_{2k+1}^{m+1}}$$

Combining the previous inequalities and the above equality yields

$$\langle x \rangle_n^{(p)} \leq C_p \varrho_p(n) \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \frac{\theta_{m,k}^2}{t_{2k+1}^{m+1} - t_{2k}^{m+1}} + \frac{\theta_{m,k}^2}{t_{2k+2}^{m+1} - t_{2k+1}^{m+1}} \right)^{p/2} \cdot \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{2p-1} \right)^{1/2}.$$

By deriving a corresponding lower bound in the same way, we obtain the following proposition. Note that we have not yet used the assumption that our partition sequence is well-balanced.

Proposition 2.55. *Let $x \in C[0, 1]$ satisfies the reverse Jensen condition. Then for $p \geq 1$, there exist constants $0 < A_p \leq B_p < \infty$ and subexponential function ϱ_p depending only on p but not on x , such that for all $n \in \mathbb{N}$,*

$$\begin{aligned} \langle x \rangle_n^{(p)} &\leq B_p \varrho_p(n) \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \frac{\theta_{m,k}^2}{t_{2k+1}^{m+1} - t_{2k}^{m+1}} + \frac{\theta_{m,k}^2}{t_{2k+2}^{m+1} - t_{2k+1}^{m+1}} \right)^{p/2} \cdot \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{2p-1} \right)^{1/2}, \\ \langle x \rangle_n^{(p)} &\geq \frac{A_p}{\varrho_p(n)} \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \frac{\theta_{m,k}^2}{t_{2k+1}^{m+1} - t_{2k}^{m+1}} + \frac{\theta_{m,k}^2}{t_{2k+2}^{m+1} - t_{2k+1}^{m+1}} \right)^{p/2} \cdot \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{2-p} \right)^{-1}. \end{aligned}$$

Proof of Theorem 2.47. Since $(\mathbb{T}_n)_{n \in \mathbb{N}}$ is a sequence of well-balanced partitions, then for $n \in \mathbb{N}$, we have

$$\kappa^{-1}(n) 2^{-n} \leq t_{k+1}^n - t_k^n \leq \kappa(n) 2^{-n}.$$

Applying this relation to Proposition 2.55 gives

$$\begin{aligned}
\langle x \rangle_n^{(p)} &\leq B_p \varrho_p(n) \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \frac{\theta_{m,k}^2}{t_{2k+1}^{m+1} - t_{2k}^{m+1}} + \frac{\theta_{m,k}^2}{t_{2k+2}^{m+1} - t_{2k+1}^{m+1}} \right)^{p/2} \cdot \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^{2p-1} \right)^{1/2} \\
&\leq B_p \varrho_p(n) \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} 2^m \kappa(m) \theta_{m,k}^2 \right)^{p/2} \cdot \left(\sum_{k=0}^{2^n-1} (2^{-n})^{2p-1} \kappa(n)^{|2p-1|} \right)^{1/2} \\
&\leq B_p \varrho_p(n) \kappa(n)^{1+|2p-1|/2} 2^{n(1-p)} \left(\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} 2^m \theta_{m,k}^2 \right)^{p/2} = B_p \varrho_p(n) \kappa(n)^{1+|2p-1|/2} 2^{n(1-p)} s_n^p.
\end{aligned}$$

Similarly, for the lower bound

$$\langle x \rangle_n^{(p)} \geq \frac{A_p}{\varrho_p(n) \kappa(n)^{p/2+|p-2|}} 2^{n(1-p)} s_n^p.$$

Taking logarithm on both sides of the above two inequalities and letting $n \uparrow \infty$ give the result. This completes the proof. $\square$

Chapter 3

Estimating the roughness exponent of stochastic volatility from discrete observations of the realized variance

This chapter is based on [47] and is arranged as follows. Section 3.1 formulates the problem of estimating the roughness exponent of the volatility process under a strictly pathwise setting. In this section, we also present the exact form of our estimator $\hat{\mathcal{R}}_n^*$ that estimates the roughness exponent of the volatility process based on the discrete observations of the realized variance. In Section 3.2, we demonstrate the strong consistency of our estimator $\hat{\mathcal{R}}_n^*$ when applied to a typical trajectory of fractional Brownian with possible drift. This result, in particular, shows that the estimator $\hat{\mathcal{R}}_n^*$ is a consistent estimator for certain rough stochastic volatility models, including the rough fractional stochastic volatility model [40]. However, one drawback of our estimator $\hat{\mathcal{R}}_n^*$ is that it is not scale-invariant. As a response, Section 3.3 constructed a scale-invariant estimator based on $\hat{\mathcal{R}}_n^*$. Moreover, the simulation studies in Section 3.4 show that our estimation procedure performs well on the fractional stochastic volatility models. Section 3.5.1 highlights the underlying rationale of the construction of our estimator $\hat{\mathcal{R}}_n^*$, and the pathwise consistency of $\hat{\mathcal{R}}_n^*$ is elucidated in Section 3.5.2. In this section, we formulate conditions on a single trajectory $x \in C[0, 1]$ and its antiderivative $y(t) = \int_0^t x(s) ds$ under which the estimates $\hat{\mathcal{R}}_n^*(y)$ converge to the roughness exponent of x . In Section 3.6, we will verify that these conditions in Section 3.5.2 are satisfied for the typical sample paths of fractional Brownian motion to obtain the strong consistency results in Section 3.2. Some technical proofs are deferred to Section 3.6.

3.1 Problem formulation and the estimator $\widehat{\mathcal{R}}_n^*$

Consider a stochastic volatility model whose price process satisfies

$$dS_t = \sigma_t S_t dB_t, \quad S_0 = s_0 > 0, \quad (3.1)$$

where B is a standard Brownian motion and σ_t is a progressively measurable stochastic process. Since the seminal work by Gatheral, Jaisson and Rosenbaum [40], it is now widely acknowledged that the sample paths of σ_t typically display much rougher characteristics than diffusive behavior. A prominent example to accommodate the rough behavior of the volatility process is to model the log-volatility by a fractional Ornstein–Uhlenbeck process X^H as in (2.45). That is,

$$\sigma_t = \exp(X_t^H), \quad (3.2)$$

where X^H solves the following integral equation

$$X_t^H = x_0 + \rho \int_0^t (\mu - X_s^H) ds + W_t^H, \quad t \geq 0,$$

for a fractional Brownian motion W^H with Hurst parameter $H \in (0, 1)$.

As explained in the introduction, our goal in this chapter is to estimate the roughness of the trajectories $t \mapsto \sigma_t$ directly from discrete, equidistant observations of the realized variance,

$$\langle \log S \rangle_t = \int_0^t \sigma_s^2 ds, \quad (3.3)$$

without having first to compute proxy values for σ_t via numerical differentiation of $t \mapsto \langle \log S \rangle_t$. This is important because, in reality, the volatility σ_t is not directly observable and numerical errors in the computation of its proxy values might distort the roughness estimate (see, e.g., [19]).

While our main results in Section 3.2 are concerned with rough stochastic volatility models based on fractional Brownian motion, e.g., the rough fractional stochastic volatility (3.2). A significant portion of our approach actually works completely trajectorial-wise and indeed estimates the roughness exponent; see Section 3.5. The roughness exponent was introduced in Chapter 2, and an estimation procedure of the roughness exponent based on direct observations is provided therein. However, the problem of estimating the roughness exponent for a trajectory of stochastic volatility is more complex because volatility cannot be measured directly; only asset prices and their realized variance (3.3) can be observed. In

our current pathwise setting, this corresponds to making discrete observations of

$$y(t) = \int_0^t g(x(s)) ds, \quad 0 \leq t \leq 1, \quad (3.4)$$

where $g : \mathbb{R} \rightarrow \mathbb{R}$ is sufficiently regular. For instance, in the rough fractional stochastic volatility model (3.2), where the log-volatility is given by a fractional Ornstein–Uhlenbeck process, we will take x as a trajectory of the fractional Ornstein–Uhlenbeck process and $g(t) = (e^t)^2 = e^{2t}$.

Let us now introduce our estimator. Suppose that for some given $n \in \mathbb{N}$ we have the discrete observations $\{y(k2^{-n-2}) : k = 0, \dots, 2^{n+2}\}$ of the function y in (3.4). Based on these data points, we introduce the coefficients

$$\vartheta_{n,k} := 2^{3n/2+3} \left(y\left(\frac{4k}{2^{n+2}}\right) - 2y\left(\frac{4k+1}{2^{n+2}}\right) + 2y\left(\frac{4k+3}{2^{n+2}}\right) - y\left(\frac{4k+4}{2^{n+2}}\right) \right), \quad (3.5)$$

for $0 \leq k \leq 2^n - 1$. Our estimator for the roughness exponent of the trajectory $g \circ x$ is now given by

$$\widehat{\mathcal{R}}_n^*(y) := 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \vartheta_{n,k}^2}. \quad (3.6)$$

As stated in the introduction, the estimator $\widehat{\mathcal{R}}_n^*$ is obtained by applying the robust approximation of the Faber–Schauder coefficients in Chapter 4 to the estimator

$$\widehat{R}_n^* = 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \theta_{n,k}^2}.$$

In Section 3.5.1, we provide a detailed explanation of the rationale behind the estimator $\widehat{\mathcal{R}}_n^*$ and how it relates to the results in Chapter 2 and Chapter 4.

3.2 Strong consistency theorems

We can now state our main results, which show the strong consistency of $\widehat{\mathcal{R}}_n^*$ when it is applied to the situation in which x is a typical trajectory of fractional Brownian motion with possible drift. In the sequel, $W^H = (W_t^H)_{0 \leq t \leq 1}$ will denote a fractional Brownian motion with Hurst parameter H , defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Theorem 3.1. *For $H \in (0, 1)$ and a strictly monotone function $g \in C^2(\mathbb{R})$, let $X_t := g(W_t^H)$*

and

$$Y_t := \int_0^t X_s ds = \int_0^t g(W_s^H) ds, \quad 0 \leq t \leq 1.$$

Then, with probability one, X admits the roughness exponent H and we have $\lim_n \widehat{\mathcal{R}}_n^*(Y) = H$.

The preceding theorem solves our problem of consistently estimating the roughness exponent for a rough volatility model with $\sigma_t^2 = g(W_t^H)$. However, empirical volatility is mean-reverting, and that effect is not captured by this model. Therefore, it is desirable to replace the fractional Brownian motion W^H with a mean-reverting process such as the fractional Ornstein–Uhlenbeck process. This process was first introduced in [13] as the solution of the integral equation

$$X_t^H = x_0 + \rho \int_0^t (\mu - X_s^H) ds + W_t^H, \quad t \in [0, 1], \quad (3.7)$$

where $x_0, \rho, m \in \mathbb{R}$ are given parameters. The integral equation (3.7) can be uniquely solved in a pathwise manner. The fractional Ornstein–Uhlenbeck process was suggested by Gatheral et al. [40] as a suitable model for log-volatility, i.e., $\sigma_t = e^{X_t^H}$. In our context, this model choice implies that we are making discrete observations of the process

$$\int_0^t \sigma_s^2 ds = \int_0^t e^{2X_s^H} ds, \quad 0 \leq t \leq 1.$$

The fractional Ornstein–Uhlenbeck process can simply be regarded as a fractional Brownian motion with starting point x_0 and adapted and absolutely continuous drift $\rho(\mu - X_s^H)$, and so it falls into the class of stochastic processes considered in Theorem 2.31, which we are quoting from Section 2.6 for the convenience of the reader.

Theorem 3.2. *Let X be given by*

$$X_t := x_0 + W_t^H + \int_0^t \xi_s ds, \quad 0 \leq t \leq 1, \quad (3.8)$$

where ξ is progressively measurable with respect to the natural filtration of W^H and satisfies the following additional assumption.

- If $H < 1/2$, we assume that $t \mapsto \xi_t$ is $\mathbb{P}$ -a.s. bounded in the sense that there exists a finite random variable C such that $\xi_t(\omega) \leq C(\omega)$ for a.e. t and $\mathbb{P}$ -a.e. $\omega \in \Omega$.
- If $H > 1/2$, we assume that $\xi_0 = 0$ and that $t \mapsto \xi_t$ is $\mathbb{P}$ -a.s. Hölder continuous with some exponent $\alpha > 2H - 1$.

Then the law of $(X_t)_{t \in [0,1]}$ is absolutely continuous with respect to the law of $(x_0 + W_t^H)_{t \in [0,1]}$.

More specifically, if X is a solution of the fractional integral equation

$$X_t = x_0 + \int_0^t b(X_s) ds + W_t^H, \quad 0 \leq t \leq 1,$$

where $b : \mathbb{R} \rightarrow \mathbb{R}$ is locally bounded and, for $H > 1/2$, locally Hölder continuous with some exponent $\alpha > 2 - 1/H$, it is further stated in Theorem 2.32 that the law of $(X_t)_{t \in [0,1]}$ is equivalent to the law of $(x_0 + W_t^H)_{t \in [0,1]}$. This applies in particular to the fractional Ornstein–Uhlenbeck process X^H defined in (3.7), where $b(x) = \rho(\mu - x)$.

The main result of this chapter is now an immediate corollary of Theorems 3.1 and 3.2.

Corollary 3.3. *Suppose that X is as in Theorem 3.2 and $g \in C^2(\mathbb{R})$ is strictly monotone. Then the stochastic process*

$$g(X_t) = g\left(x_0 + \int_0^t \xi_s ds + W_t^H\right), \quad 0 \leq t \leq 1,$$

admits $\mathbb{P}$ -a.s. the roughness exponent H , and for $Y_t = \int_0^t g(X_s) ds$ we have $\lim_n \widehat{\mathcal{R}}_n^*(Y) = H$ $\mathbb{P}$ -a.s.

By Theorem 3.2, adding a drift to fractional Brownian motion can also be regarded as changing the underlying probability law. Corollary 3.3 can therefore also be stated as follows: The strong consistency of $\widehat{\mathcal{R}}_n^*$ observed in Theorem 3.1 remains true after replacing the law of W^H with a law that arises in the context of Theorem 3.2. This invariance can be seen as robustness of $\widehat{\mathcal{R}}_n^*$ with respect to model misspecification. In addition, the strong consistency of our estimator is unaffected by changes of the nonlinear scale function g , which is yet another indication of the estimator’s robustness and versatility.

As another illustrative example for the robustness and versatility of $\widehat{\mathcal{R}}_n^*$, let us show that our estimator is also consistent for the following class of rough stochastic volatility models. To this end, recall that the Riemann–Liouville process $(Z_t^H)_{t \geq 0}$ is defined as

$$Z_t^H = \frac{1}{\Gamma(H - \frac{1}{2})} \int_0^t (t - s)^{H - \frac{1}{2}} dW_s, \quad t \geq 0.$$

Moreover, for given $\varepsilon \geq 0$, let us consider the following time-shifted Riemann–Liouville process $Z_t^{H,\varepsilon} := Z_{t+\varepsilon}^H - Z_\varepsilon^H$ for $t \geq 0$. The class of stochastic volatility models considered hereupon consists of volatility processes σ_t that admit the following dynamics

$$\sigma_t = \exp(V_t^{H,\varepsilon}) \quad \text{and} \quad V_t^{H,\varepsilon} = \gamma Z_t^{H,\varepsilon} - \frac{\gamma^2}{2} t^{2H}, \quad t \in [0, 1]. \quad (3.9)$$

In particular, taking $\varepsilon = 0$ recovers the rough Bergomi model in the form discussed, e.g., in [35, Equation (28)]. For the sake of reference, we would refer to the above stochastic volatility model (3.9) as the *time-shifted rough Bergomi model*.

The following result obtains the roughness exponent of the volatility process under the time-shifted rough Bergomi model and establishes the strong consistency of the estimator $\widehat{\mathcal{R}}_n^*$ for the time-shifted rough Bergomi model.

Theorem 3.4. *Suppose that $V^{H,\varepsilon}$ is as in (3.9) and $g \in C^2(\mathbb{R})$ is strictly monotone. Then, for any $\varepsilon > 0$, the stochastic process $g \circ V^{H,\varepsilon}$ admits the roughness exponent H almost surely. Moreover, for $Y_t^{H,\varepsilon} := \int_0^t g(V_s^{H,\varepsilon}) ds$, we have $\lim_n \widehat{\mathcal{R}}_n^*(Y^{H,\varepsilon}) = H$ with probability one for any $\varepsilon > 0$.*

3.3 A scale-invariant estimator

By definition, the roughness exponent is scale-invariant, but our estimator is not. To wit, for every trajectory $y \in C[0, 1]$ we have

$$\widehat{\mathcal{R}}_n^*(\lambda y) - \widehat{\mathcal{R}}_n^*(y) = -\frac{\log_2 |\lambda|}{n} \quad \text{for } \lambda \neq 0.$$

Consequently, a scaling factor λ may either remove or introduce a bias into an estimate, and it can notably slow down or speed up the convergence of $\widehat{\mathcal{R}}_n^*(y)$. This will be illustrated by the simulation studies provided in Section 3.4.

A number of scale-invariant modifications of $\widehat{\mathcal{R}}_n^*$ can be constructed in a manner completely analogous to Definitions 2.39 and 2.40 in Chapter 2. Here, we carry this out for the analogue of sequential scaling proposed in Definition 2.39. The underlying idea is fairly simple: We choose $m < n$ and then search for that scaling factor λ that minimizes the weighted mean-squared differences $\widehat{\mathcal{R}}_k^*(\lambda y) - \widehat{\mathcal{R}}_{k-1}^*(\lambda y)$ for $k = m+1, \dots, n$. The intuition is that such an optimal scaling factor λ enforces the convergence of the estimates $\widehat{\mathcal{R}}_k^*(\lambda y)$.

Definition 3.5. Fix $m \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_m \geq 0$ with $\alpha_0 > 0$. For $n > m$, the *sequential scaling factor* λ_n^s and the *sequential scale estimate* $\mathcal{R}_n^{*,s}(y)$ are defined as follows,

$$\lambda_n^s := \arg \min_{\lambda > 0} \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{\mathcal{R}}_k^*(\lambda y) - \widehat{\mathcal{R}}_{k-1}^*(\lambda y) \right)^2 \quad \text{and} \quad \mathcal{R}_n^{*,s}(y) := \widehat{\mathcal{R}}_n^*(\lambda_n^s y). \quad (3.10)$$

The corresponding mapping $\mathcal{R}_n^{*,s} : C^1[0, 1] \rightarrow \mathbb{R}$ will be called the *sequential scale estimator*.

Following the arguments in the proof of Proposition 2.41, one can prove the following result.

Proposition 3.6. Consider the context of Definition 3.5 with fixed $m \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_m \geq 0$ such that $\alpha_0 > 0$.

- (a) The optimization problem (3.10) admits a unique solutions for every function $y \in C[0, 1]$. In particular, all objects in Definition 3.5 are well defined.
- (b) The sequential scale estimator $\mathcal{R}_n^{*,s}$ can be represented as follows as a linear combination of $\widehat{\mathcal{R}}_{n-m-1}^*, \dots, \widehat{\mathcal{R}}_n^*$,

$$\mathcal{R}_n^{*,s} = \beta_{n,n} \widehat{\mathcal{R}}_n^* + \beta_{n,n-1} \widehat{\mathcal{R}}_{n-1}^* + \dots + \beta_{n,n-m-1} \widehat{\mathcal{R}}_{n-m-1}^*,$$

where

$$\beta_{n,k} = \begin{cases} 1 + \frac{\alpha_0}{c_n^s n^2 (n-1)} & \text{if } k = n, \\ \frac{1}{c_n^s n k} \left(\frac{\alpha_{n-k}}{k-1} - \frac{\alpha_{n-k-1}}{k+1} \right) & \text{if } n-m \leq k \leq n-1, \\ \frac{-\alpha_m}{c_n^s n (n-m)(n-m-1)} & \text{if } k = n-m-1, \end{cases} \quad \text{for } c_n^s := \sum_{k=n-m}^n \frac{\alpha_{n-k}}{k^2 (k-1)^2}.$$

- (c) The sequential scale estimator is scale-invariant. That is, for $n > m$, $y \in C^1[0, 1]$, and $\lambda \neq 0$, we have $\mathcal{R}_n^{*,s}(\lambda y) = \mathcal{R}_n^{*,s}(y)$.
- (d) If $y \in C^1[0, 1]$ and $R \in [0, 1]$ are such that there exists $\lambda \neq 0$ for which $|\widehat{\mathcal{R}}_n^*(\lambda y) - R| = O(a_n)$ as $n \uparrow \infty$ for some sequence (a_n) with $a_n = o(1/n)$, then $|\mathcal{R}_n^{*,s}(y) - R| = O(na_n)$.

In particular, part (d) of Proposition 3.6 establishes the consistency criteria and obtains the rate of convergence of the sequential scale estimate. By applying the above result, our next theorem demonstrates the strong consistency and provides the rate of convergence of $\mathcal{R}_n^{*,s}$ for the fractional Brownian motion with an absolutely continuous drift as in (3.8), provided that the nonlinear scale function is chosen to be an identity map.

Theorem 3.7. Let $m \in \mathbb{N}$, $\alpha_0 > 0$ and $\alpha_1, \dots, \alpha_m \geq 0$ be fixed. Let X be as in (3.8), and we take

$$Y_t = \int_0^t X_s ds, \quad t \in [0, 1].$$

Then the following almost sure rate of convergence holds for the sequential scale estimator $\mathcal{R}_n^{*,s}$,

$$|\mathcal{R}_n^{*,s}(Y) - H| = \mathcal{O}(2^{-n/2} \sqrt{\log n}).$$

We now examine the impact of the integration operations on the rate of convergence for sequential scale estimators. Let $R_n^{*,s}$ be the sequence scale estimator constructed based on

$\widehat{R}_n^*$. The convergence speed of the estimator $R_n^{*,s}$ was derived in Corollary 2.44, which we are quoting from Section 2.8 for the convenience of the reader.

Theorem 3.8. *Let $m \in \mathbb{N}$, $\alpha_0 > 0$ and $\alpha_1, \dots, \alpha_m \geq 0$ be fixed. Let X be as in (3.8), and the following almost sure rate of convergence holds for the sequential scale estimator $R_n^{*,s}$,*

$$|R_n^{*,s}(X) - H| = \mathcal{O}(2^{-n/2} \sqrt{\log n}).$$

To employ estimators $R_n^{*,s}$ and $\mathcal{R}_n^{*,s}$, we need to obtain the discrete observations of the process X and its antiderivative Y over the dyadic partition $\mathbb{T}_{n+1}$ and $\mathbb{T}_{n+2}$ respectively. By comparing the rates of convergence in Theorem 3.7 and Theorem 3.8, it becomes evident that employing the sequential scale estimator $R_n^{*,s}$ upon the process X and the sequential scale estimator $\mathcal{R}_n^{*,s}$ upon its antiderivative Y yield the same rates of convergence. Consequently, the smoothing effect induced by integration can be mitigated by transitioning from estimator $R_n^{*,s}$ to $\mathcal{R}_n^{*,s}$ without sacrificing convergence speed, albeit at the expense of doubling the number of observations.

3.4 Simulation study

In this section, we illustrate the practical application of Theorem 3.1, Corollary 3.3, and Proposition 3.6 by means of simulations. We will see that the estimation performance can be significantly boosted by replacing $\widehat{\mathcal{R}}_n^*$ with the sequential scale estimator $\mathcal{R}_n^{*,s}$.

We start by illustrating Theorem 3.1 for the simple choice $g(x) = x$. Recall from (3.5) and (3.6) that for given $n \in \mathbb{N}$, the computation of $\widehat{\mathcal{R}}_n^*(y)$ requires observations of the trajectory y at all values of the time grid $\mathbb{T}_{n+2} := \{k2^{-n-2} : k = 0, 1, \dots, 2^{n+2}\}$. When using for y the antiderivative of a sample path of fractional Brownian motion W^H , we generate the values of W^H on the finer grid $\mathbb{T}_N$ with $N = n + 6$. Then we put

$$Y_{k2^{-n-2}} := 2^{-N} \sum_{j=1}^{2^{N-n-2}k} W_{j2^{-N}}^H, \quad k = 0, 1, \dots, 2^{n+2}, \quad (3.11)$$

which is an approximation of $\int_0^t W_s^H ds$ by Riemann sums. Our corresponding simulation results are displayed in Figure 3.1.

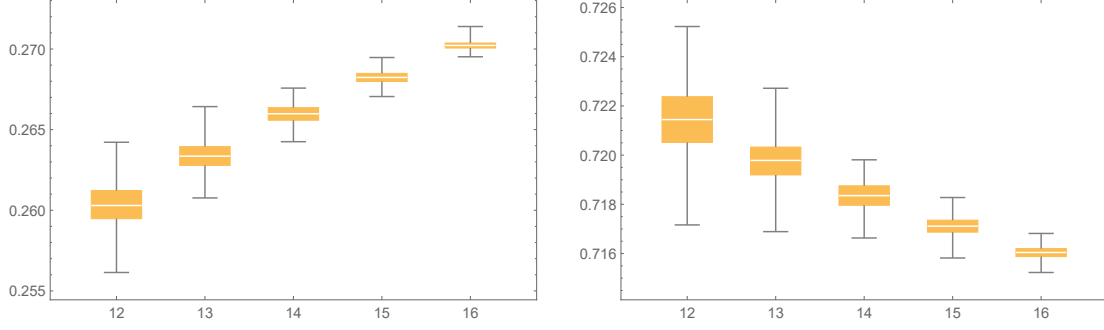


Figure 3.1: Box plots of the estimates $\widehat{\mathcal{R}}_n^*(Y)$ for $n = 12, \dots, 16$, based on 1,000 sample paths of fractional Brownian motion with $H = 0.3$ (left), $H = 0.7$ (right), and Y as in (3.11).

As one can see from Figure 3.1, the estimator $\widehat{\mathcal{R}}_n^*$ performs relatively well but exhibits a certain bias. This bias can be completely removed by passing to the scale-invariant estimator $\mathcal{R}_n^{*,s}$; see Figure 3.2.

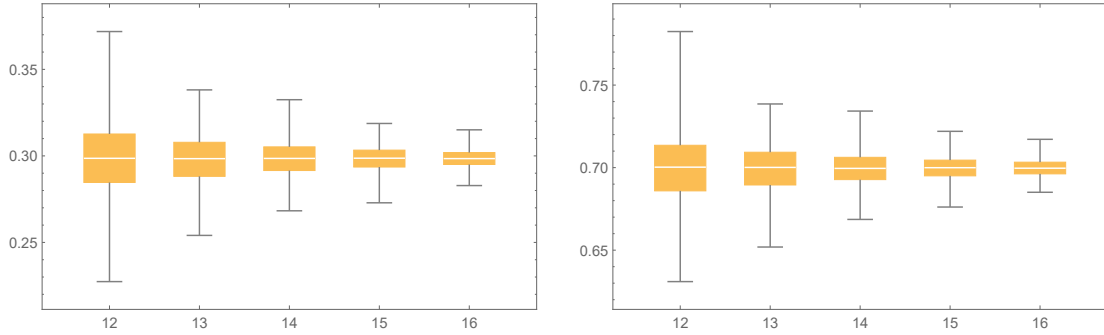


Figure 3.2: Box plots of the sequential scale estimates $\mathcal{R}_n^{*,s}(Y)$ for $n = 12, \dots, 16$, based on 1,000 sample paths of fractional Brownian motion with $H = 0.3$ (left), $H = 0.7$ (right), and Y as in (3.11). The other parameters are chosen to be $m = 3$ and $\alpha_k = 1$ for $k = 0, 1, 2, 3$.

Now we apply our estimator $\widehat{\mathcal{R}}_n^*$ to a model in which log-volatility, $\log \sigma_t$, is given by a fractional Ornstein–Uhlenbeck process X^H of the form

$$X_t^H = x_0 + \rho \int_0^t (\mu - X_s^H) ds + W_t^H, \quad t \in [0, 1],$$

and we make discrete observations of the process

$$\int_0^t \sigma_s^2 ds = \int_0^t e^{2X_s^H} ds, \quad 0 \leq t \leq 1.$$

To this end, we take again $N = n + 6$ and simulate the values $X_{k2^{-N}}^H$ ($k = 0, \dots, 2^N$) by

means of an Euler scheme. Then we put

$$Y_{k2^{-n-2}}^\sigma := 2^{-N} \sum_{j=1}^{2^{N-n-2k}} \exp\left(2X_{j2^{-N}}^H\right), \quad k = 0, 1, \dots, 2^{n+2}, \quad (3.12)$$

which is an approximation of $\int_0^t e^{2X_s^H} ds$ by Riemann sums. As can be seen from Figure 3.3, the original estimator $\widehat{\mathcal{R}}_n^*$ performs poorly in this case, while the sequential scale estimator $\mathcal{R}_n^{*,s}$ performs almost as well as for the simple case $Y_t = \int_0^t W_s^H ds$. This is due to the fact that the function $g(t) = e^{2t}$ used in (3.12) distorts substantially the scale of the underlying process, but this distortion can be remedied by using the sequential scale estimator.

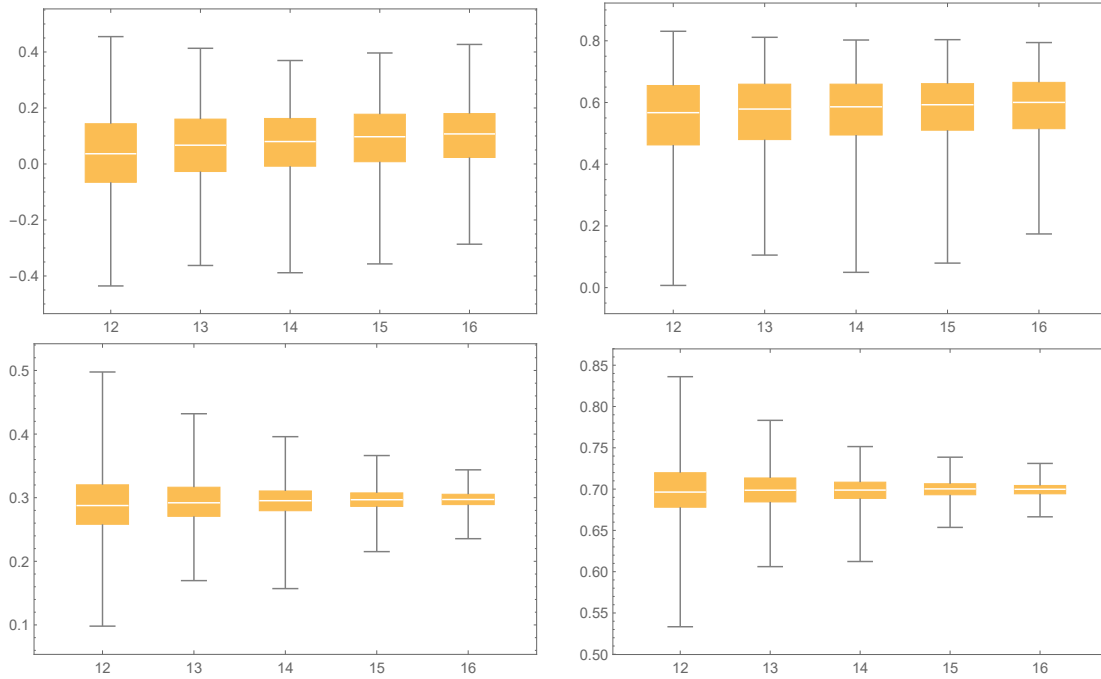


Figure 3.3: Box plots of the original estimates $\widehat{\mathcal{R}}_n^*(Y^\sigma)$ (top) and the sequential scale estimates $\mathcal{R}_n^{*,s}(Y^\sigma)$ (bottom) for $n = 12, \dots, 16$ based on 1,000 simulations of the antiderivative of the exponential Ornstein–Uhlenbeck process (3.12) with $H = 0.3$ (left) and $H = 0.7$ (right). The other parameters are chosen as $x_0 = 0$, $\rho = 0.2$, $\mu = 2$, $m = 3$ and $\alpha_k = 1$ for $k = 0, 1, \dots, 3$.

3.5 Pathwise estimation

In this section, we formulate conditions on a single trajectory $x \in C[0, 1]$ and its antiderivative $y(t) = \int_0^t x(s) ds$ under which the estimates $\widehat{\mathcal{R}}_n^*(y)$ converge to the roughness exponent of x . In Section 3.6, we will verify that these conditions are satisfied for the typical sample

paths of fractional Brownian motion. The results in the present section are hence of independent interest in situations in which it is not clear whether a given trajectory x arises from fractional Brownian motion. We start by summarizing some key results and concepts from Chapters 2 and 4 and also outline our rationale behind the specific form of the estimator $\widehat{\mathcal{R}}_n^*$.

3.5.1 The rationale behind the estimator $\widehat{\mathcal{R}}_n^*$

Recall that the Faber–Schauder functions are defined as

$$e_{-1,0}(t) := t, \quad e_{0,0}(t) := (\min\{t, 1-t\})^+, \quad e_{m,k}(t) := 2^{-m/2} e_{0,0}(2^m t - k)$$

for $t \in \mathbb{R}$, $m \in \mathbb{N}$ and $k \in \mathbb{Z}$. It is well known that the restrictions of the Faber–Schauder functions to $[0, 1]$ form a Schauder basis for $C[0, 1]$. More precisely, our function $x \in C[0, 1]$ can be uniquely represented as the uniform limit $x = \lim_n x_n$, where

$$x_n = x(0) + (x(1) - x(0)) e_{-1,0} + \sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k} e_{m,k},$$

and the Faber–Schauder coefficients $\theta_{m,k}$ are given by

$$\theta_{m,k} = 2^{m/2} \left(2x\left(\frac{2k+1}{2^{m+1}}\right) - x\left(\frac{k}{2^m}\right) - x\left(\frac{k+1}{2^m}\right) \right). \quad (3.13)$$

As a matter of fact, it is easy to see that the function x_n is simply a linear interpolation of x based on the supporting grid $\mathbb{T}_n = \{k2^{-n} : k = 0, \dots, 2^n\}$.

In Chapter 2, we have derived conditions under which the trajectory x admits a roughness exponent $R \in [0, 1]$ and also suggested a way in which the roughness exponent can be estimated from discrete observations of x . Specifically, it follows from Theorem 2.4 and Proposition 2.16 that, if the Faber–Schauder coefficients satisfy the so-called reverse Jensen condition (see Definition 2.3) and the sequence

$$\widehat{R}_n^*(x) := 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \theta_{n,k}^2} \quad (3.14)$$

converges to a finite limit R , then x admits the roughness exponent R .

Note that it is assumed in Chapter 2 that the trajectory x can be observed directly. This, however, is not the case in the context of our case, where x is the (squared) volatility in a

stochastic volatility model. So let us suppose now that we can only observe the values the antiderivative $y(t) = \int_0^t x(s) ds$ takes on the supporting grid $\mathbb{T}_{n+2}$. If we can interpolate the data points $\{y(t) : t \in \mathbb{T}_{n+2}\}$ by means of a piecewise quadratic function $y_{n+2} \in C^1(\mathbb{R})$, then its derivative y'_{n+2} will be a continuous and piecewise linear function with supporting grid $\mathbb{T}_{n+1}$ and hence representable in the form

$$y'_{n+2} = \hat{x}_0 + \hat{\theta}_{-1,0}e_{-1,0} + \sum_{m=0}^{n+1} \sum_{k=0}^{2^m-1} \hat{\theta}_{m,k}e_{m,k} \quad (3.15)$$

for some initial value $\hat{x}_0$ and certain coefficients $\hat{\theta}_{m,k}$. Such a piecewise quadratic C^1 -interpolation y_{n+2} exists in the form of the standard quadratic spline interpolation. Unfortunately, though, it is well known that quadratic spline interpolation suffers some serious drawbacks:

- the initial value $\hat{x}_0$ is not uniquely determined by the given data $\{y(t) : t \in \mathbb{T}_{n+2}\}$;
- the values $y_{n+2}(t)$ depend in a highly sensitive manner on the choice of $\hat{x}_0$;
- the values $y_{n+2}(s)$ depend in a nonlocal way on the given data $\{y(t) : t \in \mathbb{T}_{n+2}\}$, i.e., altering one data point $y(t)$ may affect the value $y_{n+2}(s)$ also if s is located far away from t .

In Chapter 4, we will investigate the analytical properties of the estimated Faber–Schauder coefficients $\hat{\theta}_{m,k}$ defined in (3.15). When looking at quadratic spline interpolation through the lens of these coefficients, we obtain refreshingly new insights into this tool. As will be shown in Theorem 4.1, the Faber–Schauder coefficients of y'_{n+2} take the following explicit forms: For $-1 \leq m \leq n$, we have

$$\begin{aligned} \hat{\theta}_{m,k} = 2^{n+m/2+3} \sum_{j=1}^{2^{n+1-m}} (-1)^j & \left(y\left(\frac{k}{2^m} + \frac{j}{2^{n+2}}\right) - y\left(\frac{k}{2^m} + \frac{j-1}{2^{n+2}}\right) \right. \\ & \left. + y\left(\frac{k+1}{2^m} - \frac{j-1}{2^{n+2}}\right) - y\left(\frac{k+1}{2^m} - \frac{j}{2^{n+2}}\right) \right). \end{aligned} \quad (3.16)$$

For $m = n+1$, we have

$$\begin{aligned} \hat{\theta}_{n+1,k} = -2^{(n+1)/2+2} \hat{x}_0 - 2^{3(n+1)/2+4} \sum_{j=1}^{2k} (-1)^j & \left(y\left(\frac{j}{2^{n+2}}\right) - y\left(\frac{j-1}{2^{n+2}}\right) \right) \\ + 3 \cdot 2^{3(n+1)/2+2} & \left(y\left(\frac{2k+1}{2^{n+2}}\right) - y\left(\frac{2k}{2^{n+2}}\right) \right) - 2^{3(n+1)/2+2} \left(y\left(\frac{2k+2}{2^{n+2}}\right) - y\left(\frac{2k+1}{2^{n+2}}\right) \right). \end{aligned} \quad (3.17)$$

As one can see immediately from those formulas, the coefficients in generations $m = 0, \dots, n$ are independent of $\hat{x}_0$, whereas the coefficients in generation $n + 1$ contain the additive term $-2^{(n+1)/2+2}\hat{x}_0$, which translates any error made in estimating $\hat{x}_0$ into an $2^{(n+1)/2+2}$ -fold error for each final-generation coefficient. Moreover, for $m = 0, \dots, n$, each $\hat{\theta}_{m,k}$ depends only on those data points $y(t)$ for which t belongs to the closure of the support of the corresponding wavelet function $e_{m,k}$. Thus, the entire nonlocality of the function y_{n+2} arises from the coefficients in generation $n + 1$, while the coefficients of all lower generations depend on locally on the given data. We refer to Figure 4.2 for an illustration.

The main results in Chapter 4 concern error bounds for the estimated Faber–Schauder coefficients $\hat{\theta}_{m,k}$. Specifically, we found that the ℓ_2 -norm of the combined errors in generations $m = 0, \dots, n$ is typically benign, whereas the error in the final generation $m = n + 1$ can be larger than a factor of size $\mathcal{O}(2^n)$ times the error of all previous generations combined. While the exact error bounds from Chapter 4 will not be needed in this chapter, the proof of Lemma 3.11 will rely on an algebraic representation of the error terms obtained in Lemma 4.13 and stated in (3.23) below.

The above-mentioned facts make it clear that the coefficients in generations $m = 0, \dots, n$ provide robust estimates for the corresponding true coefficients, while the estimates $\hat{\theta}_{n+1,k}$ are highly non-robust and should be discarded. It is now obvious that in estimating the roughness exponent of x from the data $\{y(t) : t \in \mathbb{T}_{n+2}\}$, we should replace the true coefficients $\theta_{n,k}$ in our formula (3.14) for $\hat{R}_n^*(x)$ with their estimates $\hat{\theta}_{n,k}$. It remains to note that $\hat{\theta}_{n,k}$ is in fact equal to $\vartheta_{n,k}$ defined in (3.5), so that we finally arrive at the rationale behind our estimator $\hat{\mathcal{R}}_n^*$.

Remark 3.9. The above rationale also supports the validity of constructing an estimator $\hat{\mathcal{R}}_n$ from the Gladyshev estimator $\hat{R}_n$ by replacing the actual Faber–Schauder coefficients $\theta_{m,k}$ by $\hat{\theta}_{m,k}$ in (3.16) for all $0 \leq m \leq n - 1$. The estimator $\hat{\mathcal{R}}_n$ takes the following form

$$\hat{\mathcal{R}}_n = 1 - \frac{1}{n} \log_2 \sqrt{\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} (\hat{\theta}_{m,k})^2}.$$

Following the exact same techniques developed in the subsequent sections, one can also show that the estimator $\hat{\mathcal{R}}_n$ is a strongly consistent estimator for the fractional Brownian motion with drift. However, to evaluate the estimate $\hat{\mathcal{R}}_n(y)$ needs to compute all the approximated Faber–Schauder coefficients $\hat{\theta}_{m,k}$. Consequently, numerical computation of the estimated roughness exponent through the estimator $\hat{\mathcal{R}}_n$ will be very costly. Hence, we limit our discussion to the estimator $\hat{\mathcal{R}}_n^*$ in this thesis.

The following example provides a concrete instance where choosing the final-generation

coefficients $\widehat{\theta}_{n+1,k}$ instead of $\vartheta_{n,k} = \widehat{\theta}_{n,k}$ leads to an estimate that is non-robust and also otherwise inferior.

Example 3.10. For $R \in (0, 1]$, let $x^R \in C[0, 1]$ be the function with Faber–Schauder coefficients $\theta_{n,k} = 2^{n(1/2-R)}$. These functions belong to the well-studied class of fractal Takagi–Landsberg functions. It was shown in [77, Theorem 2.1] that x^R has the roughness exponent R . Moreover, for $y^R(t) = \int_0^t x^R(s) ds$, it will be shown in Example 4.2 that the robust approximation (3.5) based on discrete observations of y^R recovers exactly the Faber–Schauder coefficients of x^R . That is, for $n \in \mathbb{N}$ and $0 \leq k \leq 2^n - 1$, we have

$$\vartheta_{n,k} = \theta_{n,k} = 2^{n(\frac{1}{2}-R)}.$$

It follows that

$$\widehat{\mathcal{R}}_n^*(y^R) = 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} 2^{(1-2R)n}} = 1 - \frac{1}{n} \log_2 2^{(1-R)n} = R.$$

Hence, the estimator $\widehat{\mathcal{R}}_n^*$ is not only consistent but also exact in the sense that it gives the correct value R for every finite n .

Now we replace $\vartheta_{n,k} = \widehat{\theta}_{n,k}$ with the final-generation estimates $\widehat{\theta}_{n+1,k}$ as defined in (3.17). Note that this requires the choice of an initial value $\hat{x}_0$. The corresponding estimator is given by

$$\widetilde{\mathcal{R}}_n(y^R) := 1 - \frac{1}{n+1} \log_2 \sqrt{\sum_{k=0}^{2^{n+1}-1} \widehat{\theta}_{n+1,k}^2}.$$

We will get from the subsequent Example 4.2 that for $R < 1/2$,

$$\widehat{\theta}_{n+1,k} = -2^{(n+1)/2+2} \hat{x}_0 + \sum_{m=n+1}^{\infty} 2^{m(\frac{1}{2}-R)} = -2^{(n+1)/2+2} \hat{x}_0 + \frac{2^{(n+1)(\frac{1}{2}-R)}}{1 - 2^{\frac{1}{2}-R}}.$$

Hence,

$$\sqrt{\sum_{k=0}^{2^{n+1}-1} \widehat{\theta}_{n+1,k}^2} = 2^{n+1} \left| -4\hat{x}_0 + \frac{2^{-(n+1)R}}{1 - 2^{\frac{1}{2}-R}} \right|.$$

It follows that

$$\lim_{n \uparrow \infty} \widetilde{\mathcal{R}}_n(y^R) = \begin{cases} 0 & \text{if } \hat{x}_0 \neq 0, \\ R & \text{if } \hat{x}_0 = 0. \end{cases}$$

This shows that the estimator $\widetilde{\mathcal{R}}_n$ is extremely sensitive with respect to the estimate $\hat{x}_0$ of the exact initial value $x(0)$, which in typical applications will be unknown. Even in the case

that $x(0)$ is known, the correct value R is only obtained asymptotically, whereas $\widehat{\mathcal{R}}_n^*(y^R) = R$ for all finite n . These observations illustrate once again why we deliberately discard the final generation $\widehat{\theta}_{n+1,k}$ of estimated Faber–Schauder coefficients.

3.5.2 Pathwise consistency of $\widehat{\mathcal{R}}_n^*(y)$

Let us fix $x \in C[0, 1]$ and denote by $\theta_{m,k}$ its Faber–Schauder coefficients (3.13). As before, we denote by $y(t) = \int_0^t x(s) ds$ the antiderivative of x and by $\vartheta_{n,k}$ the coefficients defined in (3.5). For simplicity, we introduce the following vector notation,

$$\bar{\theta}_n := (\theta_{n,0}, \theta_{n,1}, \dots, \theta_{n,2^n})^\top \in \mathbb{R}^{2^n} \quad \text{and} \quad \bar{\vartheta}_n = (\vartheta_{n,0}, \vartheta_{n,1}, \dots, \vartheta_{n,2^n-1})^\top \in \mathbb{R}^{2^n}, \quad (3.18)$$

Then the estimators $\widehat{R}_n^*$ and $\widehat{\mathcal{R}}_n^*$ defined in (3.14) and (3.6) can be written as

$$\widehat{R}_n^*(x) = 1 - \frac{1}{n} \log_2 \|\bar{\theta}_n\|_{\ell_2} \quad \text{and} \quad \widehat{\mathcal{R}}_n^*(y) = 1 - \frac{1}{n} \log_2 \|\bar{\vartheta}_n\|_{\ell_2}. \quad (3.19)$$

Moreover, we introduce the column vector $\mathbf{z}_n := (z_i^{(n)})_{1 \leq i \leq 2^n}$ with components

$$z_i^{(n)} = 2^{3n/2} \sum_{m=n}^{\infty} 2^{-3m/2} \sum_{k=0}^{2^{m-n}-1} \theta_{m,k+2^{m-n}(i-1)} \quad \text{for } 1 \leq i \leq 2^n. \quad (3.20)$$

For $1 \leq i, j \leq 2^n$, we let furthermore

$$\boldsymbol{\eta}_{i,j} = \begin{cases} \mathbf{r} & \text{for } 1 \leq i = j \leq 2^n, \\ \mathbf{0}_{1 \times 4} & \text{for } 1 \leq i \neq j \leq 2^n, \end{cases} \quad (3.21)$$

where $\mathbf{r} := \frac{1}{4}(-1, +1, +1, -1)$, and $\mathbf{0}_{m \times n}$ denotes the $m \times n$ -dimensional zero matrix. Moreover, we denote

$$Q_n := \begin{bmatrix} \boldsymbol{\eta}_{1,1} & \boldsymbol{\eta}_{1,2} & \cdots & \boldsymbol{\eta}_{1,2^n-1} & \boldsymbol{\eta}_{1,2^n} \\ \boldsymbol{\eta}_{2,1} & \boldsymbol{\eta}_{2,2} & \cdots & \boldsymbol{\eta}_{2,2^n-1} & \boldsymbol{\eta}_{2,2^n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{\eta}_{2^n,1} & \boldsymbol{\eta}_{2^n,2} & \cdots & \boldsymbol{\eta}_{2^n,2^n-1} & \boldsymbol{\eta}_{2^n,2^n} \end{bmatrix} \in \mathbb{R}^{2^n \times 2^{n+2}}. \quad (3.22)$$

It will be shown in Lemma 4.13 that the error between the true and estimated Faber–Schauder coefficients can be represented as follows,

$$\bar{\vartheta}_n - \bar{\theta}_n = \mathbf{w}_n, \quad \text{where } \mathbf{w}_n = Q_n \mathbf{z}_{n+2} \in \mathbb{R}^{2^n}. \quad (3.23)$$

Consider the following condition:

$$\text{There exists } \kappa \in \mathbb{R} \setminus \{1\} \text{ such that } \frac{\|\mathbf{w}_n\|_{\ell_2}}{\|\bar{\boldsymbol{\theta}}_n\|_{\ell_2}} \longrightarrow \kappa \text{ as } n \uparrow \infty. \quad (3.24)$$

We will see in Proposition 3.18 that condition (3.24) is $\mathbb{P}$ -a.s. satisfied for fractional Brownian motion.

Lemma 3.11. *Under condition (3.24), there exist $n_0 \in \mathbb{N}$ and constants $0 < \kappa_- \leq \kappa_+ < \infty$ such that*

$$\kappa_- \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2} \leq \|\bar{\boldsymbol{\vartheta}}_n\|_{\ell_2} \leq \kappa_+ \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2} \quad \text{for all } n \geq n_0. \quad (3.25)$$

Proof. Let κ be as in (3.24). Then, for any $\varepsilon < |\kappa - 1|/2$, there exists $n_\varepsilon \in \mathbb{N}$ such that for $n \geq n_\varepsilon$, we have $\|\bar{\boldsymbol{\theta}}_n\|_{\ell_2} (\kappa - \varepsilon) < \|\mathbf{w}_n\|_{\ell_2} < \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2} (\kappa + \varepsilon)$. Using the representation (3.23) and applying the triangle inequality gives

$$\|\bar{\boldsymbol{\vartheta}}_n\|_{\ell_2} = \|\bar{\boldsymbol{\theta}}_n - \mathbf{w}_n\|_{\ell_2} \leq \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2} + \|\mathbf{w}_n\|_{\ell_2} \leq (\kappa + \varepsilon + 1) \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2}.$$

On the other hand, we have

$$\|\bar{\boldsymbol{\vartheta}}_n\|_{\ell_2} = \|\bar{\boldsymbol{\theta}}_n - \mathbf{w}_n\|_{\ell_2} \geq \left| \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2} - \|\mathbf{w}_n\|_{\ell_2} \right| \geq (|1 - \kappa - \varepsilon| \wedge |1 - \kappa + \varepsilon|) \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2}.$$

This completes the proof. $\square$

By taking logarithms in (3.25), Lemma 3.11 immediately yields the following result.

Proposition 3.12. *Under condition (3.24), the limit $\lim_n \widehat{\mathcal{R}}_n^*(y)$ exists if and only if $\lim_n \widehat{R}_n^*(x)$ exists. Moreover, in this case, $\lim_n \widehat{\mathcal{R}}_n^*(y) = \lim_n \widehat{R}_n^*(x)$.*

Example 3.13. In the situation of Example 3.10, we have seen that $\vartheta_{n,k} = \theta_{n,k} = 2^{n(\frac{1}{2}-R)}$. Applying the representation (3.23) yields that $\mathbf{w}_n = \mathbf{0}_{2^n \times 1}$. This implies $\lim_n \|\mathbf{w}_n\|_{\ell_2} / \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2} = 0$. That is, x^R satisfies condition (3.24). Hence, Proposition 3.12 applies, which gives an additional proof of the previously observed fact that $\lim_n \widehat{\mathcal{R}}_n^*(y) = R$.

Next, we consider the following question: Under which conditions on x and g does $u := g \circ x$ admit the roughness exponent R ? To answer this question, we fix the following notation throughout the remainder of this section,

$$u(t) = g(x(t)) \quad \text{and} \quad v(t) = \int_0^t u(s) ds = \int_0^t g(x(s)) ds. \quad (3.26)$$

Proposition 3.14. *If x admits the roughness exponent R , g belongs to $C^1(\mathbb{R})$, and g' is nonzero on the range $x([0, 1])$ of x , then $u = g \circ x$ also admits the roughness exponent R .*

Proof. For any $p > 0$, the mean value theorem and the intermediate value theorem yield numbers $\tau_{n,k} \in [k2^{-n}, (k+1)2^{-n}]$ such that

$$\begin{aligned} \langle u \rangle_n^{(p)} &= \sum_{k=0}^{2^n-1} \left| g'(x(\tau_{n,k})) \left(x\left(\frac{k+1}{2^n}\right) - x\left(\frac{k}{2^n}\right) \right) \right|^p \\ &= \sum_{k=0}^{2^n-1} |g'(x(\tau_{n,k}))|^p \left| x\left(\frac{k+1}{2^n}\right) - x\left(\frac{k}{2^n}\right) \right|^p, \end{aligned} \quad (3.27)$$

where the notation $\langle u \rangle_n^{(p)}$ was introduced in (2.1). Since g' is continuous and nonzero, there are constants $0 < c_- < c_+ < \infty$ such that $c_- \leq |g'(x(t))| \leq c_+$ for all $t \in [0, 1]$. Hence, $c_-^p \langle x \rangle_n^{(p)} \leq \langle u \rangle_n^{(p)} \leq c_+^p \langle x \rangle_n^{(p)}$ holds for all n . Passing to the limit $n \uparrow \infty$ for $p > 1/R$ and $p < 1/R$ yields the result. $\square$

Now we turn to the following question: Under which conditions do we have $\widehat{\mathcal{R}}_n^*(v) \rightarrow R$, where v is as in (3.26)? The conditions we will introduce for answering this question are relatively strong. Nevertheless, they hold for the sample paths of fractional Brownian motion.

Proposition 3.15. *Suppose that there exists $R \in (0, 1)$ such that the following conditions hold.*

(a) *We have*

$$0 < \liminf_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \vartheta_{n,k}^2 \leq \limsup_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \vartheta_{n,k}^2 < \infty. \quad (3.28)$$

(b) *The function x is Hölder continuous with exponent $\alpha \in (2R/5, 1]$.*

Then, if $g \in C^2(\mathbb{R})$ is strictly monotone, we have $\lim_n \widehat{\mathcal{R}}_n^*(v) = R$.

Proof. In this proof, we will work with the actual and estimated Faber–Schauder coefficients of the various functions x , y , u , and v . For this reason, we will temporarily use a superscript to indicate from which function the Faber–Schauder coefficients will be computed. That is, for any function f , we write

$$\begin{aligned} \theta_{n,k}^f &= 2^{n/2} \left(2f\left(\frac{2k+1}{2^{n+1}}\right) - f\left(\frac{k}{2^n}\right) - f\left(\frac{k+1}{2^n}\right) \right), \\ \vartheta_{n,k}^f &= 2^{3n/2+3} \left(f\left(\frac{4k}{2^{n+2}}\right) - 2f\left(\frac{4k+1}{2^{n+2}}\right) + 2f\left(\frac{4k+3}{2^{n+2}}\right) - f\left(\frac{4k+4}{2^{n+2}}\right) \right). \end{aligned} \quad (3.29)$$

With this notation, the coefficients $\vartheta_{n,k}$ in (3.5) should be re-written as $\vartheta_{n,k}^y$. In particular, (3.28) refers to the coefficients $\vartheta_{n,k}^y$. Our goal in this proof is to show that (3.28) carries over to the coefficients $\vartheta_{n,k}^v$. That is,

$$0 < \liminf_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} (\vartheta_{n,k}^v)^2 \leq \limsup_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} (\vartheta_{n,k}^v)^2 < \infty. \quad (3.30)$$

Taking logarithms, dividing by $2n$, and passing to the limit will then yield $R - \widehat{\mathcal{R}}_n^*(v) \rightarrow 0$, which is the assertion.

It remains to establish (3.30). Rewriting the second line in (3.29) gives after a short computation that

$$\vartheta_{n,k}^f = 2^{n+5/2} \left(\theta_{n+1,2k+1}^f - \theta_{n+1,2k}^f \right). \quad (3.31)$$

Let us introduce the notation $\theta_{m,k}^f(s) := \theta_{m,k}^{f(s+\cdot)}$. That is, $\theta_{m,k}^f(s)$ are the Faber–Schauder coefficients of the function $t \mapsto f(s+t)$ for given $s \geq 0$. One can avoid undefined arguments of functions in case $s+t > 1$ by assuming without loss of generality that all occurring functions on $[0,1]$ are defined on all of $[0,\infty)$. With this notation, we get from (3.31) that for $f \in C^1[0,\infty)$,

$$\vartheta_{n,k}^f = 2^{n+5/2} \int_0^{2^{-n-1}} \theta_{n+1,2k}^{f'}(s) ds. \quad (3.32)$$

Applying the mean-value theorem and the intermediate value theorem yields certain intermediate times $\tau_{n+2,k}(s) \in [2^{-n-2}k + s, 2^{-n-2}(k+1) + s]$ such that for $s \in [0, 2^{-n-1}]$,

$$\begin{aligned} \theta_{n+1,2k}^u(s) &= 2^{(n+1)/2} \left(2u\left(\frac{4k+1}{2^{n+2}} + s\right) - u\left(\frac{4k}{2^{n+2}} + s\right) - u\left(\frac{4k+2}{2^{n+2}} + s\right) \right) \\ &= 2^{(n+1)/2} g'(x(\tau_{n+2,4k}(s))) \left(x\left(\frac{4k+1}{2^{n+2}} + s\right) - x\left(\frac{4k}{2^{n+2}} + s\right) \right) \\ &\quad + 2^{(n+1)/2} g'(x(\tau_{n+2,4k+1}(s))) \left(x\left(\frac{4k+1}{2^{n+2}} + s\right) - x\left(\frac{4k+2}{2^{n+2}} + s\right) \right), \\ &= 2^{(n+1)/2} \left(\frac{g'(x(\tau_{n+2,4k}(s))) + g'(x(\tau_{n+2,4k+1}(s)))}{2} \left(2x\left(\frac{4k+1}{2^{n+2}} + s\right) - x\left(\frac{4k}{2^{n+2}} + s\right) - x\left(\frac{4k+2}{2^{n+2}} + s\right) \right) \right) \\ &\quad + 2^{(n+1)/2} \left(\frac{g'(x(\tau_{n+2,4k}(s))) - g'(x(\tau_{n+2,4k+1}(s)))}{2} \left(x\left(\frac{4k+2}{2^{n+2}} + s\right) - x\left(\frac{4k}{2^{n+2}} + s\right) \right) \right). \end{aligned}$$

The intermediate value theorem and the mean-value theorem also imply that there are

intermediate times $\tau_{n+1,2k}^\sharp(s), \tau_{n+1,2k}^\flat(s) \in [2^{-n-1}2k + s, 2^{-n-1}(2k+1) + s]$ such that

$$\begin{aligned} \frac{1}{2} \left(g'(x(\tau_{n+2,4k}(s))) + g'(x(\tau_{n+2,4k+1}(s))) \right) &= g'(\tau_{n+1,2k}^\sharp(s)), \\ \frac{1}{2} \left(g'(x(\tau_{n+2,4k}(s))) - g'(x(\tau_{n+2,4k+1}(s))) \right) &= \frac{1}{2} g''(\tau_{n+1,2k}^\flat(s)) (x(\tau_{n+2,4k}(s)) - x(\tau_{n+2,4k+1}(s))). \end{aligned}$$

With the shorthand notation

$$\zeta_{n+1,2k}^x(s) := 2^{(n+1)/2} \left(x\left(\frac{4k+2}{2^{n+2}} + s\right) - x\left(\frac{4k}{2^{n+2}} + s\right) \right) (x(\tau_{n+2,4k}(s)) - x(\tau_{n+2,4k+1}(s))),$$

we then have

$$\theta_{n+1,2k}^u(s) = g'(\tau_{n+1,2k}^\sharp(s)) \theta_{n+1,2k}^x(s) + g''(\tau_{n+1,2k}^\flat(s)) \zeta_{n+1,2k}^x(s).$$

Plugging the preceding equation into (3.32) and applying the mean value theorem for integrals yields intermediate times $\tau_{n+1,k}^\sharp, \tau_{n+1,k}^\flat \in [2^{-n-1}k, 2^{-n-1}(k+1)]$ that are independent of s such that

$$\begin{aligned} \vartheta_{n,k}^v &= 2^{n+5/2} g'(x(\tau_{n+1,2k}^\sharp)) \int_0^{2^{-n-1}} \theta_{n+1,2k}^x(s) ds + 2^{n+5/2} g''(x(\tau_{n+1,2k}^\flat)) \int_0^{2^{-n-1}} \zeta_{n+1,2k}^x(s) ds \\ &= g'(x(\tau_{n+1,2k}^\sharp)) \vartheta_{n,k}^y + 2^{n+5/2} g''(x(\tau_{n+1,2k}^\flat)) \int_0^{2^{-n-1}} \zeta_{n+1,2k}^x(s) ds. \end{aligned}$$

Introducing the shorthand notation

$$\tilde{\zeta}_{n+1,2k}^x := 2^{n+5/2} \int_0^{2^{-n-1}} \zeta_{n+1,2k}^x(s) ds,$$

and let us write

$$\begin{aligned} (\vartheta_{n,k}^v)^2 &= (g'(x(\tau_{n+1,2k}^\sharp)))^2 (\vartheta_{n,k}^y)^2 + (g''(x(\tau_{n+1,2k}^\flat)))^2 (\tilde{\zeta}_{n+1,2k}^x)^2 \\ &\quad + 2g'(x(\tau_{n+1,2k}^\sharp))g''(x(\tau_{n+1,2k}^\flat))\vartheta_{n,k}^y\tilde{\zeta}_{n+1,2k}^x. \end{aligned} \tag{3.33}$$

For each of the three terms on the right, we will now analyze its contribution to the quantities in (3.30). The main contribution comes from the first term on the right. Indeed, our assumptions on g imply that there are constants $0 < c_- \leq c_+ < \infty$ such that $c_- < (g'(x(t)))^2 < c_+$ for all $t \in [0, 1]$, and so

$$c_- 2^{n(2R-2)} \sum_{k=0}^{2^n-1} (\vartheta_{n,k}^y)^2 \leq 2^{n(2R-2)} \sum_{k=0}^{2^n-1} (\vartheta_{n,k}^v)^2 \leq c_+ 2^{n(2R-2)} \sum_{k=0}^{2^n-1} (\vartheta_{n,k}^y)^2.$$

This will establish (3.30) as soon as we have shown that the contributions of the two remaining terms in (3.33) are asymptotically negligible. For the second term, we use the Hölder continuity of x to get a constant c_x for which

$$|x(\tau_{n+2,4k}(s)) - x(\tau_{n+2,4k+1}(s))| \leq c_x |\tau_{n+2,4k}(s) - \tau_{n+2,4k+1}(s)|^\alpha \leq c_x 2^{-\alpha n}.$$

Furthermore, there exists $\kappa_x > 0$ such that $32(g''(x(s)))^2 \leq \kappa_x$ for all $s \in [0, 1]$. Then,

$$\begin{aligned} & 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \left(g''(x(\tau_{n+1,2k}^b)) \right)^2 \left(\tilde{\zeta}_{n+1,2k} \right)^2 \\ &= 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \left(g''(x(\tau_{n+1,2k}^b)) \right)^2 \left(2^{n+5/2} \int_0^{2^{-n-1}} \zeta_{n+1,2k}^x(s) ds \right)^2 \\ &\leq \kappa_x 2^{2Rn} \sum_{k=0}^{2^n-1} \left(\int_0^{2^{-n-1}} \zeta_{n+1,2k}^x(s) ds \right)^2 \leq \kappa_x 2^{2Rn} \sum_{k=0}^{2^n-1} 2^{-n-1} \int_0^{2^{-n-1}} \left(\zeta_{n+1,2k}^x(s) \right)^2 ds \\ &\leq \kappa_x 2^{(2R-1)n} \int_0^{2^{-n-1}} \sum_{k=0}^{2^n-1} \left(x\left(\frac{4k+2}{2^{n+2}} + s\right) - x\left(\frac{4k}{2^{n+2}} + s\right) \right)^2 \left(x(\tau_{n+2,4k}(s)) - x(\tau_{n+2,4k+1}(s)) \right)^2 ds \\ &= \kappa_x 2^{\alpha n} \int_0^{2^{-n-1}} 2^{(2R-1-\alpha)n} \sum_{k=0}^{2^n-1} \left(x\left(\frac{2k+1}{2^{n+1}} + s\right) - x\left(\frac{2k}{2^{n+1}} + s\right) \right)^2 (c_x^2 2^{-2\alpha n}) ds \\ &\leq \kappa_x c_x^2 \int_0^{2^{-n-1}} \sup_{s \in [0, 2^{-n-1}]} \left(2^{(2R-1-3\alpha)n} \sum_{k=0}^{2^{n+1}-1} \left(x\left(\frac{2k+1}{2^{n+1}} + s\right) - x\left(\frac{2k}{2^n} + s\right) \right)^2 \right) ds. \end{aligned}$$

Moreover, (b) implies the integrand in the final term converges to zero:

$$\lim_{n \uparrow \infty} \sup_{0 \leq s \leq 2^{-n-1}} 2^{(2R-1-3\alpha)n} \sum_{k=0}^{2^{n+1}-1} \left(x\left(\frac{2k+1}{2^{n+1}} + s\right) - x\left(\frac{2k}{2^n} + s\right) \right)^2 = 0. \quad (3.34)$$

Indeed, by the Hölder continuity of x , we can again use the constant c_x to get

$$2^{(2R-1-3\alpha)n} \sum_{k=0}^{2^{n+1}-1} \left(x\left(\frac{2k+1}{2^{n+1}} + s\right) - x\left(\frac{2k}{2^n} + s\right) \right)^2 \leq 2^{(2R-1-3\alpha)n} \cdot 2^{n+1} \cdot c_x^2 2^{-2(n+1)\alpha};$$

the right-hand side is equal to $c_x^2 \cdot 2^{1-2\alpha} \cdot 2^{(2R-5\alpha)n}$, which converges to zero as $n \uparrow \infty$. Altogether, this shows that the contribution of the second term on the right-hand side of (3.33) is negligible.

For the cross-product term on the rightmost side of (3.33), we get from the Cauchy–

Schwarz inequality,

$$\begin{aligned} & \lim_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} g'(x(\tau_{n+1,2k}^\sharp)) g''(x(\tau_{n+1,2k}^\flat)) \vartheta_{n,k} \tilde{\zeta}_{n+1,2k}^x \\ & \leq \sqrt{\lim_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \left(g'(x(\tau_{n+1,2k}^\sharp)) \right)^2 (\vartheta_{n,k})^2} \sqrt{\lim_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \left(g''(x(\tau_{n+1,2k}^\flat)) \right)^2 (\tilde{\zeta}_{n+1,2k}^x)^2} = 0. \end{aligned}$$

Altogether, (3.30) follows. $\square$

The next theorem describes the asymptotic behavior of the estimator $\widehat{\mathcal{R}}_n^*$ for the antiderivative of the sum of continuous functions. To this end, let $x, u \in C[0, 1]$, and we denote the antiderivatives of x and u by y and v respectively. To present this result, we continue to use a superscript to indicate from which function the Faber–Schauder coefficients will be computed. Furthermore, we remain to use the notation $\theta_{m,k}^f(s) = \theta_{m,k}^{f(s+\cdot)}$.

The following lemma presents a possible sufficient condition, under which, the consistency of the estimator $\widehat{\mathcal{R}}_n^*$ remains to hold for the sum of continuous functions.

Lemma 3.16. *Suppose that there exists $R \in (0, 1)$ such that the following conditions hold.*

(a) *We have*

$$0 < \liminf_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^y \right)^2 \leq \limsup_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^y \right)^2 < \infty. \quad (3.35)$$

(b) *The function $u \in C^2(0, 1)$ admits the following properties:*

(i) *$|u''(t)|$ is strictly decreasing over $(0, 1)$.*

(ii) $\lim_{t \downarrow 0} t^{5-2R} (u''(t))^2 = 0$.

(iii) $\lim_{t \downarrow 0} t^{1/2-R} \cdot \sup_{0 \leq s \leq t} u(s) = 0$.

Then, we have $\lim_n \widehat{\mathcal{R}}_n^*(y + v) = R$.

Proof. To prove the above result, our goal is to show that (3.35) carries over to the coefficients $\vartheta_{n,k}^{y+v}$. That is

$$0 < \liminf_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^{y+v} \right)^2 \leq \limsup_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^{y+v} \right)^2 < \infty. \quad (3.36)$$

If so, taking logarithms, dividing by $2n$, and passing to the limit upon (3.36) will yield $\lim_n \widehat{\mathcal{R}}_n^*(y+v) = R$. Moreover, following the same arguments in the proof of Proposition 3.15, it suffices to establish the following identity

$$\limsup_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} (\vartheta_{n,k}^v)^2 = 0. \quad (3.37)$$

To this end, from (3.31), we get

$$\vartheta_{n,k}^{y+v} = \vartheta_{n,k}^y + \vartheta_{n,k}^v = \vartheta_{n,k}^y + 2^{n+5/2} \int_0^{2^{-n-1}} \theta_{n+1,2k}^u(s) ds.$$

By the mean-value theorem, there exist intermediate times $\tau_{n,k}(s) \in [2^{-n}k, 2^{-n}(k+1)]$ such that

$$\begin{aligned} \theta_{n+1,2k}^u(s) &= 2^{(n+1)/2} \left(2u \left(\frac{4k+1}{2^{n+2}} + s \right) - u \left(\frac{4k}{2^{n+2}} + s \right) - u \left(\frac{4k+2}{2^{n+2}} + s \right) \right) \\ &= 2^{(n+1)/2-2n-3} u''(\tau_{n,k}(s)) = 2^{-3n/2-5/2} u''(\tau_{n,k}(s)). \end{aligned}$$

Applying the Hölder inequality to the above equation yields the existence of an intermediate time $\tau_{n,k} \in [2^{-n}k, 2^{-n}(k+1)]$ such that

$$\begin{aligned} (\vartheta_{n,k}^u)^2 &= 2^{2n+5} \left(\int_0^{2^{-n-1}} \theta_{n+1,2k}^u(s) ds \right)^2 = 2^{-n} \left(\int_0^{2^{-n-1}} u''(\tau_{n,k}(s)) ds \right)^2 \\ &\leq 2^{-2n-1} \int_0^{2^{-n-1}} (u''(\tau_{n,k}(s)))^2 ds = 2^{-3n-2} (u''(\tau_{n,k}))^2, \end{aligned}$$

where the last step follows again from the mean-value theorem. Furthermore, since $|u''|$ is a decreasing function, we have

$$2^{-3n-2} \sum_{k=2}^{2^n-1} (u''(\tau_{n,k}))^2 \leq 2^{-3n-2} \sum_{k=2}^{2^n-1} (u''(2^{-n}k))^2 \leq 2^{-2n-2} \int_{2^{-n}}^1 (u''(s))^2 ds.$$

Applying Cauchy's inequality yields that

$$(\vartheta_{n,0}^v)^2 + (\vartheta_{n,1}^v)^2 \leq 8 \sup_{0 \leq s \leq 2^{-n-1}} \left(\theta_{n+1,0}^u(s) \right)^2 + \left(\theta_{n+1,1}^u(s) \right)^2 \leq 2^{n+9} \sup_{0 \leq s \leq 2^{-n}} u^2(s).$$

Therefore, combining the above inequalities gives

$$2^{n(2R-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^v \right)^2 \leq 2^{(2R-4)n-2} \int_{2^{-n}}^1 \left(u''(s) \right)^2 ds + 2^{(2R-1)n} \sup_{0 \leq s \leq 2^{-n}} u^2(s).$$

In particular, it follows from L'Hôpital's rule that

$$\lim_{n \uparrow \infty} 2^{(2R-4)n} \int_{2^{-n}}^1 \left(u''(s) \right)^2 ds = \lim_{t \downarrow 0} \frac{\int_t^1 \left(u''(s) \right)^2 ds}{t^{2R-4}} = (4-2R) \lim_{t \downarrow 0} \frac{\left(u''(t) \right)^2}{t^{2R-5}} = 0.$$

Furthermore, we have

$$\lim_{n \uparrow \infty} 2^{(2R-1)n} \sup_{0 \leq s \leq 2^{-n}} u^2(s) = \left(\lim_{t \downarrow 0} t^{1/2-R} \cdot \sup_{0 \leq s \leq t} u(s) \right)^2 = 0.$$

Finally, the above identities establish (3.37) and complete the proof. $\square$

To conclude this section, we state and prove a lemma, which will be needed for the proof of Proposition 3.18. For possible future reference, we include it in our present pathwise context. For $n, k \in \mathbb{N}$, let us consider the vector $\mathbf{z}_{(n,k)} = (z_i^{(n,k)}) \in \mathbb{R}^{2^n}$, where

$$z_i^{(n,k)} = 2^{3n/2} \sum_{m=n}^{n+k} 2^{-3m/2} \sum_{j=0}^{2^{m-n}-1} \theta_{m,j+2^{m-n}(i-1)} \quad \text{for } 1 \leq i \leq 2^n. \quad (3.38)$$

It is clear that the vector $\mathbf{z}_{(n,k)}$ is a truncated version of the vector $\mathbf{z}_n$ defined in (3.20). Since each Faber–Schauder coefficient $\theta_{m,k}$ is a linear combination of the values $x(j2^{-n-k-1})$, each $z_i^{(n,k)}$ must admit the following representation,

$$z_i^{(n,k)} = \sum_{j=0}^{2^{n+k+1}} \xi_j^{(n,k,i)} x\left(\frac{j}{2^{n+k+1}}\right), \quad (3.39)$$

for certain coefficients $\xi_j^{(n,k,i)}$. The following lemma computes the values of these coefficients.

Lemma 3.17. *We have*

$$\xi_j^{(n,k,i)} = \begin{cases} 0 & \text{if } j \leq 2^{k+1}(i-1) - 1 \quad \text{or} \quad j \geq 2^{k+1}i + 1, \\ 2^{n/2}(2^{-k} - 2) & \text{if } j = 2^{k+1}(i-1) \quad \text{or} \quad j = 2^{k+1}i, \\ 2^{1-k+n/2} & \text{if } 2^{k+1}(i-1) + 1 \leq j \leq 2^k i - 1. \end{cases} \quad (3.40)$$

Proof of Lemma 3.17. We fix $n \in \mathbb{N}$ and $1 \leq i \leq 2^n$ and proceed by induction on $k \in \mathbb{N}$.

First, let us establish the base case $k = 0$. Then

$$z_i^{(n,0)} = \theta_{n,i-1} = 2^{n/2+1}x\left(\frac{2i-1}{2^{n+1}}\right) - 2^{n/2}x\left(\frac{2i}{2^{n+1}}\right) - 2^{n/2}x\left(\frac{2i-2}{2^{n+1}}\right). \quad (3.41)$$

Moreover, plugging $k = 0$ into (3.40) yields that $\xi_{2i-2}^{(n,0,i)} = \xi_{2i}^{(n,0,i)} = -2^{n/2}$, $\xi_{2i-1}^{(n,0,i)} = 2^{n/2+1}$ and $\xi_j^{(n,0,i)} = 0$ otherwise. It is clear that those coefficients coincide with the corresponding ones in (3.41), which proves our induction for the initial step $k = 0$.

Next, let us assume that (3.40) holds for $k = m$ and subsequently prove that this identity also holds for $k = m + 1$. It follows from (3.38) that

$$z_i^{(n,m+1)} = z_i^{(n,m)} + 2^{n/2-m-1} \sum_{j=2^{m+1}(i-1)}^{2^{m+1}i-1} \left(2x\left(\frac{2j+1}{2^{n+m+2}}\right) - x\left(\frac{j}{2^{n+m+1}}\right) - x\left(\frac{j+1}{2^{n+m+1}}\right) \right). \quad (3.42)$$

For $2^{m+1}(i-1) < j < 2^{m+1}i-1$, the point $2^{-n-m-2}(2j+1)$ cannot be written in the form $\ell 2^{-n-m-1}$ for some $\ell \in \mathbb{N}_0$. Hence

$$\xi_{2j+1}^{(n,m+1,i)} = 2 \cdot 2^{n/2-m-1} = 2^{n/2-m},$$

as the term $x(2^{-n-m-2}(2j+1))$ does not appear in the linear combination (3.39) for $k = m$. Next, for $2^{k-n}(i-1) < j < 2^{k-n}i$, the point $2^{-n-m-2}2j = 2^{-n-m-1}j$ can be written in the form $\ell 2^{-n-m-1}$ for some $\ell \in \mathbb{N}_0$. It thus follows from (3.39) and (3.40) that

$$\xi_{2j}^{(n,m+1,i)} = \xi_j^{(n,m,i)} - 2 \cdot 2^{n/2-m-1} = 2^{n/2-m},$$

as the term $x(2^{-n-m-1}j)$ contributes to the representation of $z_i^{(n,m)}$ with $\xi_j^{(n,m,i)} = 2^{n/2-m+1}$. Moreover, for $j = 2^{k-n}(i-1)$ or $j = 2^{k-n}i$, we have

$$\xi_j^{(n,m+1,i)} = \xi_j^{(n,m,i)} - 2^{n/2-m-1} = 2^{n/2}(2^{-m} - 2) - 2^{n/2-m-1} = 2^{n/2}(2^{-m-1} - 2).$$

Last, for $j \leq 2^{m+2}(i-1) - 1$ or $j \geq 2^{m+2}i + 1$, the term $x(2^{-n-m-2})$ does not appear on right-hand side of (3.42). Thus, we have $\xi_j^{(n,m+1,i)} = 0$. Comparing the above identities with (3.40) proves the case for $k = m + 1$. $\square$

3.6 Proofs for the results in Chapter 3

We begin this section with the proof of Theorem 3.1.

Proof of Theorem 3.1. It was shown in Theorem 2.30 that W^H admits $\mathbb{P}$ -a.s. the roughness

exponent H . It now follows from Proposition 3.14 that the sample paths of $X = g(W^H)$ also admit the roughness exponent H .

Now we prove that, with probability one, $\widehat{\mathcal{R}}_n^*(Y) \rightarrow H$ for $Y_t = \int_0^t X_s ds$. To this end, we use the following result by Gladyshev [41] on the convergence of the weighted quadratic variation of W^H ,

$$2^{(2H-1)n} \langle W^H \rangle_n^{(2)} \longrightarrow 1 \quad \mathbb{P}\text{-a.s.}$$

Hence, if $\bar{\theta}_n = (\theta_{n,k})$ are the Faber–Schauder coefficients of the sample paths of W^H , then Proposition 2.16 yields that

$$2^{(2H-2)n} \|\bar{\theta}_n\|_{\ell_2}^2 = 2^{(2H-2)n} \sum_{k=0}^{2^n-1} \theta_{n,k}^2 \longrightarrow 2^{2-2H} - 1 \quad \mathbb{P}\text{-a.s.} \quad (3.43)$$

Lemma 3.11 now implies that condition (a) of Proposition 3.15 is satisfied. Condition (b) of that proposition is also satisfied, because it is well known that the sample paths of W^H are $\mathbb{P}$ -a.s. Hölder continuous for every exponent $\alpha < H$; see, e.g., [75, Section 1.16]. Hence, we may apply Proposition 3.15 and so $\widehat{\mathcal{R}}_n^*(Y) \rightarrow H$ follows. $\square$

For completing the proof of Theorem 3.1, it remains to establish (3.24). This is achieved in the following proposition.

Proposition 3.18. *With probability one, the sample paths of fractional Brownian motion W^H satisfy condition (3.24).*

3.6.1 Proof of Proposition 3.18

To prove Proposition 3.18, we need to obtain the asymptotic behavior of the $\|\mathbf{w}_n\|_{\ell_2}$ associated with a fractional Brownian motion W^H . Let $\bar{\mathbf{v}}_n$, $\mathbf{z}_{(n,k)}$, and $\mathbf{z}_n$ be defined as in (3.5), (3.38), and (3.20) for the sample paths of W^H . It is clear that $\mathbf{z}_n$ is well defined since the sample paths of W^H satisfy a Hölder condition. Moreover, all three are Gaussian random vectors. Our next lemma characterizes the covariance structure of the Gaussian vector $\mathbf{z}_n$. To this end, consider the function $g_H := h_1 + h_2 + h_3$, where the $h_i : \mathbb{N}_0 \rightarrow \mathbb{R}$ are defined as follows,

$$\begin{aligned} h_1(\varsigma) &= -2 \left(2\varsigma^{2H} + |\varsigma - 1|^{2H} + |\varsigma + 1|^{2H} \right) \\ h_2(\varsigma) &= \begin{cases} \frac{8}{2H+1} \left((\varsigma + 1)^{2H+1} - (\varsigma - 1)^{2H+1} \right) & \text{for } \varsigma \geq 1, \\ \frac{16}{2H+1} & \text{for } \varsigma = 0, \end{cases} \\ h_3(\varsigma) &= -\frac{8}{(2H+2)(2H+1)} \left(|\varsigma + 1|^{2H+2} - 2\varsigma^{2H+2} + |\varsigma - 1|^{2H+2} \right). \end{aligned} \quad (3.44)$$

Furthermore, we introduce the Toeplitz matrix $G_n := (g_H(|i - j|))_{1 \leq i, j \leq 2^n}$.

Lemma 3.19. *For each $n \in \mathbb{N}$, the random vector $\mathbf{z}_n$ is a well-defined zero-mean Gaussian vector with covariance matrix*

$$\Gamma_n = (\gamma_{i,j}^{(n)})_{1 \leq i, j \leq 2^n} = 2^{(1-2H)n} G_n.$$

Proof. For $n, k \in \mathbb{N}$, let us denote

$$\Gamma_{(n,k)} = (\gamma_{i,j}^{(n,k)})_{1 \leq i, j \leq 2^n} := \mathbb{E} \left[\mathbf{z}_{(n,k)} \mathbf{z}_{(n,k)}^\top \right].$$

It suffices to show that the components $\gamma_{i,j}^{(n,k)}$ converges to $\gamma_{i,j}^{(n)}$ as $k \uparrow \infty$. Moreover, by symmetry, it suffices to consider the case $j \geq i$. Lemma 3.17 yields

$$\gamma_{i,j}^{(n,k)} = \mathbb{E} [z_i^{(n,k)} z_j^{(n,k)}] = \sum_{\tau_1=0}^{2^{n+k+1}} \sum_{\tau_2=0}^{2^{n+k+1}} \xi_{\tau_1}^{(n,k,i)} \xi_{\tau_2}^{(n,k,j)} \mathbb{E} \left[W_{\frac{\tau_1}{2^{n+k+1}}}^H \cdot W_{\frac{\tau_2}{2^{n+k+1}}}^H \right].$$

We also get from Lemma 3.17 that $\sum_{\tau=0}^{2^{n+k+1}} \xi_\tau^{(n,k,i)} = 0$ and $\xi_j^{(n,k,i)} = 0$ for $j \leq 2^{k+1}(i-1) - 1$ or $j \geq 2^{k+1}i + 1$. Hence, for $\varsigma := j - i \geq 0$,

$$\gamma_{i,j}^{(n,k)} = - \sum_{\tau_1=0}^{2^{n+k+1}} \sum_{\tau_2=0}^{2^{n+k+1}} \frac{\xi_{\tau_1}^{(n,k,i)} \xi_{\tau_2}^{(n,k,j)}}{2} \left| \frac{\tau_1 - \tau_2}{2^{n+k+1}} \right|^{2H} = -2^{-2Hn} \sum_{\tau_1=0}^{2^{k+1}} \sum_{\tau_2=0}^{2^{k+1}} \frac{\xi_{\tau_1}^{(n,k,i)} \xi_{\tau_2}^{(n,k,j)}}{2} \left| \frac{\tau_1 - \tau_2}{2^{k+1}} + \varsigma \right|^{2H}.$$

Using once again (3.40) yields that

$$\gamma_{i,j}^{(n,k)} = 2^{(1-2H)n} (h_{1,k}(\varsigma) + h_{2,k}(\varsigma) + h_{3,k}(\varsigma)), \quad (3.45)$$

where functions $h_{i,k}$ are defined as follows,

$$\begin{aligned} h_{1,k}(\varsigma) &= -\frac{(2^{-k} - 2)^2}{2} \left(2|\varsigma|^{2H} + |\varsigma - 1|^{2H} + |\varsigma + 1|^{2H} \right), \\ h_{2,k}(\varsigma) &= 2^{-k}(2 - 2^{-k}) \sum_{\tau=1}^{2^{k+1}-1} \left(\left| \frac{\tau}{2^{k+1}} + \varsigma \right|^{2H} + \left| \frac{\tau}{2^{k+1}} + \varsigma - 1 \right|^{2H} + \left| \frac{-\tau}{2^{k+1}} + \varsigma \right|^{2H} + \left| \frac{-\tau}{2^{k+1}} + \varsigma + 1 \right|^{2H} \right), \\ h_{3,k}(\varsigma) &= -2^{1-2k} \sum_{\tau_1=1}^{2^{k+1}-1} \sum_{\tau_2=1}^{2^{k+1}-1} \left| \frac{\tau_1 - \tau_2}{2^{k+1}} + \varsigma \right|^{2H}. \end{aligned}$$

Let us first consider the case $\varsigma \geq 1$. Then,

$$\lim_{k \uparrow \infty} h_{1,k}(\varsigma) = -2 \left(2\varsigma^{2H} + (\varsigma - 1)^{2H} + (\varsigma + 1)^{2H} \right). \quad (3.46)$$

Furthermore,

$$\begin{aligned}
\lim_{k \uparrow \infty} h_{2,k}(\varsigma) &= \lim_{k \uparrow \infty} 2^{1-k} (2 - 2^{-k}) \sum_{\tau=1}^{2^{k+1}-1} \left(\left| \frac{\tau}{2^{k+1}} + \varsigma \right|^{2H} + \left| \frac{\tau}{2^{k+1}} + \varsigma - 1 \right|^{2H} \right) \\
&= 8 \lim_{k \uparrow \infty} 2^{-k-1} \sum_{\tau=1}^{2^{k+1}-1} \left(\left| \frac{\tau}{2^{k+1}} + \varsigma \right|^{2H} + \left| \frac{\tau}{2^{k+1}} + \varsigma - 1 \right|^{2H} \right) \\
&= 8 \left(\int_{\varsigma}^{\varsigma+1} t^{2H} dt + \int_{\varsigma-1}^{\varsigma} t^{2H} dt \right) = \frac{8((\varsigma+1)^{2H+1} - (\varsigma-1)^{2H+1})}{2H+1}
\end{aligned}$$

We also get in a similar way that

$$\begin{aligned}
\lim_{k \uparrow \infty} h_{3,k}(\varsigma) &= -8 \lim_{k \uparrow \infty} 2^{-2-2k} \sum_{\tau_1=1}^{2^{k+1}-1} \sum_{\tau_2=1}^{2^{k+1}-1} \left(\frac{\tau_1 - \tau_2}{2^{k+1}} + \varsigma \right)^{2H} = -8 \int_0^1 \int_0^1 (t - s + \varsigma)^{2H} ds dt \\
&= -\frac{8((\varsigma+1)^{2H+2} - 2\varsigma^{2H+2} + (\varsigma-1)^{2H+2})}{(2H+2)(2H+1)}.
\end{aligned}$$

For the case $\varsigma = 0$, $\lim_{k \uparrow \infty} h_{1,k}(0) = h_1(0)$ as in (3.46). Next, we have

$$\lim_{k \uparrow \infty} h_{2,k}(0) = \lim_{k \uparrow \infty} 2^{2-k} (2 - 2^{-k}) \sum_{\tau=1}^{2^{k+1}-1} \left| \frac{\tau}{2^{k+1}} \right|^{2H} = 16 \int_0^1 t^{2H} dt = \frac{16}{2H+1}.$$

Finally,

$$\begin{aligned}
\lim_{k \uparrow \infty} h_{3,k}(0) &= \lim_{k \uparrow \infty} -2^{1-2k} \sum_{\tau_1=1}^{2^{k+1}-1} \sum_{\tau_2=1}^{2^{k+1}-1} \left| \frac{\tau_1 - \tau_2}{2^{k+1}} \right|^{2H} = -8 \int_0^1 \int_0^1 |t - s|^{2H} ds dt \\
&= 16 \int_0^1 \int_0^t (t - s)^{2H} ds dt = \frac{16}{(2H+1)(2H+2)}.
\end{aligned}$$

Comparing the above equations with (3.44) completes the proof. $\square$

Our next lemma investigates the limit of $2^{n(H-1)} \|\mathbf{z}_n\|_{\ell_2}$ as $n \uparrow \infty$ by applying a concentration inequality from [5, Lemma 3.1]. In the form needed here, it states that if $\mathbf{Z}$ is a centered Gaussian random vector with covariance matrix C , $T := \sqrt{\text{trace } C}$, and $\gamma := \sqrt{\|C\|_2}$, then there exists a universal constant κ independent of C such that

$$\mathbb{P} \left[\left| \|\mathbf{Z}\|_{\ell_2} - T \right| \geq t \right] \leq \kappa \exp \left(-\frac{t^2}{4\gamma^2} \right) \quad \text{for all } t > 0. \quad (3.47)$$

Lemma 3.20. *With probability one,*

$$\lim_{n \uparrow \infty} 2^{n(H-1)} \|\mathbf{z}_n\|_{\ell_2} = 2\sqrt{\frac{1-H}{H+1}}.$$

Proof. It follows from (3.44) that

$$\begin{aligned} \sqrt{\text{trace } \Gamma_n} &= \sqrt{\sum_{i=1}^{2^n} \gamma_{i,i}^{(n)}} = 2^{(1-H)n} \sqrt{g_H(0)} = 2^{(1-H)n} \sqrt{-4 + \frac{16}{2H+1} - \frac{16}{(2H+1)(2H+2)}} \\ &= 2^{(1-H)n+1} \sqrt{\frac{1-H}{H+1}}. \end{aligned} \quad (3.48)$$

Let $\|\cdot\|_p$ denote the ℓ_p -induced operator norm. As shown in Lemma 3.19, the covariance matrix Γ_n is a symmetric Toeplitz matrix and so $\gamma_{i,j}^{(n)} = \gamma_{j,i}^{(n)} = \gamma_{1,|j-i+1|}^{(n)}$. Hence, we have $\|\Gamma_n\|_1 = \|\Gamma_n\|_\infty$, and this gives

$$\begin{aligned} \|\Gamma_n\|_2 &\leq \sqrt{\|\Gamma_n\|_1 \|\Gamma_n\|_\infty} = \|\Gamma_n\|_1 = \max_{1 \leq j \leq 2^n} \sum_{i=1}^{2^n} |\gamma_{i,j}^{(n)}| = \max_{1 \leq j \leq 2^n} \sum_{i=1}^{2^n} |\gamma_{1,|j-i+1|}^{(n)}| \\ &\leq 2 \sum_{i=1}^{2^n} |\gamma_{1,i}^{(n)}| \leq 2^{(1-2H)n+1} \sum_{\varsigma=0}^{2^n} |g_H(\varsigma)|, \end{aligned} \quad (3.49)$$

where the first inequality is a well-known bound for the spectral norm of a matrix; see, e.g., [94, proof of Theorem 2.3]. In the next step, we will show that $g_H(\varsigma) = \mathcal{O}(\varsigma^{2H-4})$ as $\varsigma \uparrow \infty$. For $\varsigma \geq 3$, Taylor expansion yields $u_1 \in (\varsigma-1, \varsigma)$ and $u_2 \in (\varsigma, \varsigma+1)$ such that

$$\begin{aligned} (\varsigma-1)^{2H} &= \varsigma^{2H} + \sum_{i=1}^3 \frac{(-1)^i \prod_{j=1}^i (2H-j+1)}{i!} \varsigma^{2H-i} + \frac{\prod_{j=1}^4 (2H-i+1)}{4!} u_1^{2H-4}, \\ (\varsigma+1)^{2H} &= \varsigma^{2H} + \sum_{i=1}^3 \frac{\prod_{j=1}^i (2H-j+1)}{i!} \varsigma^{2H-i} + \frac{\prod_{j=1}^4 (2H-i+1)}{4!} u_2^{2H-4}. \end{aligned}$$

Note that

$$\sum_{i=1}^3 \frac{((-1)^i + 1) \prod_{j=1}^i (2H-j+1)}{i!} \varsigma^{2H-i} = 2H(2H-1) \varsigma^{2H-2},$$

and therefore, we have

$$h_1(\varsigma) = -2 \left(4\varsigma^{2H} + 2H(2H-1) \varsigma^{2H-2} + \frac{\prod_{j=1}^4 (2H-j+1)}{4!} (u_1^{2H-4} + u_2^{2H-4}) \right). \quad (3.50)$$

In the same way, we obtain

$$h_2(\varsigma) = 8 \left(2\varsigma^{2H} + \frac{2(2H)(2H-1)}{3!} \varsigma^{2H-2} + \frac{\prod_{j=1}^4 (2H-j+1)}{5!} (u_3^{2H-4} + u_4^{2H-4}) \right), \quad (3.51)$$

$$h_3(\varsigma) = -8 \left(\frac{2}{2!} \varsigma^{2H} + \frac{2(2H)(2H-1)}{4!} \varsigma^{2H-2} + \frac{\prod_{j=1}^4 (2H-j+1)}{6!} (u_5^{2H-4} + u_6^{2H-4}) \right), \quad (3.52)$$

for some $u_3, u_5 \in (\varsigma-1, \varsigma)$ and $u_4, u_6 \in (\varsigma, \varsigma+1)$. Since $u_i \geq \varsigma-1$, we get $u_i^{2H-4} \leq (\varsigma-1)^{2H-4}$. Summing up (3.50), (3.51) and (3.52) yields that $g_H(\varsigma) = \mathcal{O}(\varsigma^{2H-4})$ as $\varsigma \uparrow \infty$, which with (3.49) implies that

$$\|\Gamma_n\|_2 \leq 2^{(1-2H)n} \left(g_H(0) + g_H(1) + \sum_{\varsigma=2}^{2^n} g_H(\varsigma) \right) = \mathcal{O}(2^{(1-2H)n}) \quad \text{for } H \in (0, 1). \quad (3.53)$$

Therefore, for each $H \in (0, 1)$, there exist $c_H \geq 0$ and $n_{c,H} \in \mathbb{N}$ such that for $n \geq n_{c,H}$, we have $\|\Gamma_n\|_2 \leq c_H$. Thus, for $n \geq n_{c,H}$ and any given $\varepsilon > 0$, the concentration inequality (3.47) gives

$$\begin{aligned} \mathbb{P} \left(\left| 2^{n(H-1)} \|\mathbf{z}_n\|_{\ell_2} - 2\sqrt{\frac{1-H}{H+1}} \right| \geq \varepsilon \right) &= \mathbb{P} \left(\left| \|\mathbf{z}_n\|_{\ell_2} - 2^{(1-H)n+1} \sqrt{\frac{1-H}{H+1}} \right| \geq 2^{(1-H)n} \varepsilon \right) \\ &\leq \kappa \exp \left(-\frac{2^{n(2-2H)} \varepsilon^2}{\|\Gamma_n\|_2} \right) = \kappa \exp(-c_H^{-1} 2^n \varepsilon^2). \end{aligned}$$

The latter expression is summable in n for every $\varepsilon > 0$, and so a Borel–Cantelli argument yields that $2^{n(H-1)} \|\mathbf{z}_n\|_{\ell_2} \rightarrow 2\sqrt{\frac{1-H}{H+1}}$ with probability one as $n \uparrow \infty$. $\square$

In the following lemma, we will derive the asymptotic behavior of the norms of $\mathbf{w}_n$ defined in (3.23).

Lemma 3.21. *With probability one, we have*

$$\lim_{n \uparrow \infty} 2^{n(H-1)} \|\mathbf{w}_n\|_{\ell_2} = 2^{-2H} \sqrt{\alpha(H)},$$

where $\alpha(H) = g_H(0) - \frac{1}{2}g_H(1) - g_H(2) + \frac{1}{2}g_H(3)$.

Proof. Let us denote the covariance matrix of $\mathbf{w}_n$ by $\Phi_n := (\phi_{i,j}^{(n)})_{i,j=1}^{2^n} = Q_n \Gamma_{n+2} Q_n^\top$. We first show that

$$\text{trace } \Phi_n = 2^{(2-2H)n-4H} \alpha(H).$$

For the fixed $n \in \mathbb{N}$, consider the following partition of the covariance matrix Γ_{n+2} ,

$$\Gamma_{n+2} = \begin{bmatrix} \Gamma_{1,1}^* & \Gamma_{1,2}^* & \cdots & \Gamma_{1,2^n}^* \\ \Gamma_{2,1}^* & \Gamma_{2,2}^* & \cdots & \Gamma_{2,2^n}^* \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{2^n,1}^* & \Gamma_{2^n,2}^* & \cdots & \Gamma_{2^n,2^n}^* \end{bmatrix}, \quad (3.54)$$

where $\Gamma_{i,j}^*$ are 4×4 -dimensional matrices. In particular, for $1 \leq i \leq 2^n$, the diagonal partitioned matrices $\Gamma_{i,i}^*$ are of the form:

$$\Gamma_{i,i}^* = 2^{(1-2H)(n+2)} G_2 = 2^{(1-2H)(n+2)} \begin{bmatrix} g_H(0) & g_H(1) & g_H(2) & g_H(3) \\ g_H(1) & g_H(0) & g_H(1) & g_H(2) \\ g_H(2) & g_H(1) & g_H(0) & g_H(1) \\ g_H(3) & g_H(2) & g_H(1) & g_H(0) \end{bmatrix}.$$

Recall the definition of $\boldsymbol{\eta}_{i,j}$ from (3.21), we get

$$\begin{aligned} \phi_{i,i}^{(n)} &= (\boldsymbol{\eta}_{i,1}, \boldsymbol{\eta}_{i,2}, \dots, \boldsymbol{\eta}_{i,2^n}) \Gamma_{n+1} (\boldsymbol{\eta}_{i,1}, \boldsymbol{\eta}_{i,2}, \dots, \boldsymbol{\eta}_{i,2^n})^\top \\ &= (\mathbf{0}_{1 \times 4}, \dots, \boldsymbol{\eta}_{i,i}, \dots, \mathbf{0}_{1 \times 4}) \Gamma_{n+1} (\mathbf{0}_{1 \times 4}, \dots, \boldsymbol{\eta}_{i,i}, \dots, \mathbf{0}_{1 \times 4})^\top \\ &= \mathbf{r} \Gamma_{i,i}^* \mathbf{r}^\top = 2^{(1-2H)(n+2)} \mathbf{r} G_2 \mathbf{r}^\top. \end{aligned}$$

To evaluate the last argument in the above equation, we have

$$\mathbf{r} G_2 \mathbf{r}^\top = \frac{1}{16} \mathbf{1}_{1 \times 4} \begin{bmatrix} g_H(0) & -g_H(1) & -g_H(2) & g_H(3) \\ -g_H(1) & g_H(0) & g_H(1) & -g_H(2) \\ -g_H(2) & g_H(1) & g_H(0) & -g_H(1) \\ g_H(3) & -g_H(2) & -g_H(1) & g_H(0) \end{bmatrix} \mathbf{1}_{4 \times 1} = \frac{\alpha(H)}{4}.$$

Therefore, we have $\phi_{i,i}^{(n)} = 2^{(1-2H)(n+2)-2} \alpha(H)$ for every $1 \leq i \leq 2^n$, and

$$\text{trace } \Phi_n = \sum_{i=1}^{2^n} \phi_{i,i}^{(n)} = 2^{(2-2H)n-4H} \alpha(H).$$

In our next step, we shall show that $2^{2n(H-1)} \|\mathbf{w}_n\|_{\ell_2}$ converges to $2^{-2H} \sqrt{\alpha(H)}$. It will be shown in Theorem 4.14 that $\|Q_n\|_2 = 1/4$, and due to (3.53), there exists a constant $c_H > 0$ such that

$$\|\Phi_n\|_2 \leq \|Q_n\|_2^2 \|\Gamma_{n+2}\|_2 \leq c_H 2^{(1-2H)n}.$$

For any given $\varepsilon > 0$, the concentration inequality (3.47) yields that

$$\begin{aligned} \mathbb{P} \left(\left| 2^{n(H-1)} \|\mathbf{w}_n\|_{\ell_2} - 2^{n(H-1)} \sqrt{\text{trace } \Phi_n} \right| \geq \varepsilon \right) &= \mathbb{P} \left(\left| 2^{n(H-1)} \|\mathbf{w}_n\|_{\ell_2} - \sqrt{2^{2-4H} \alpha(H)} \right| \geq \varepsilon \right) \\ &\leq \kappa \exp \left(-\frac{2^{n(2-2H)} \varepsilon^2}{\|Q_n\|_2} \right) = \kappa \exp(-c_H^{-1} 2^n \varepsilon^2). \end{aligned}$$

From here, the Borel–Cantelli lemma yields the assertion. $\square$

Proof of Proposition 3.18. By (3.43) and Lemma 3.21,

$$\lim_{n \uparrow \infty} \frac{\|\mathbf{w}_n\|_{\ell_2}}{\|\bar{\boldsymbol{\theta}}_n\|_{\ell_2}} = \lim_{n \uparrow \infty} \sqrt{\frac{2^{n(2H-2)} \|\mathbf{w}_n\|_{\ell_2}^2}{2^{n(2H-2)} \|\bar{\boldsymbol{\theta}}_n\|_{\ell_2}^2}} = \sqrt{\frac{2^{-4H} \alpha(H)}{2^{2-2H} - 1}} = \sqrt{\frac{\alpha(H)}{2^{2+2H} - 2^{4H}}} < 1.$$

See Figure 3.4 for an illustration of the latter inequality. $\square$

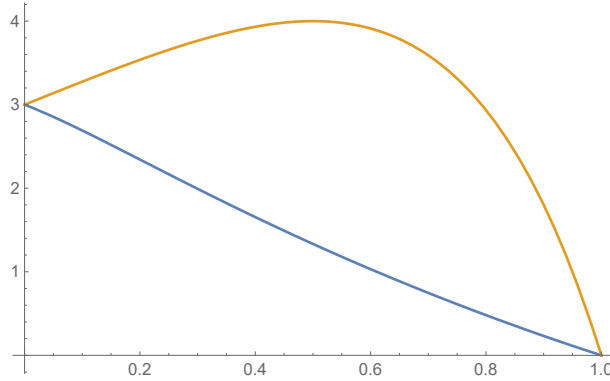


Figure 3.4: Plot of functions $\alpha(H)$ (blue) and $\beta(H) := 2^{2+2H} - 2^{4H}$ (orange) as functions of $H \in (0, 1)$.

3.6.2 Proof of Theorem 3.7

The proof of Theorem 3.7 begins with the following lemma, which provides an upper bound for the covariance between Faber–Schauder coefficients of the fractional Brownian motion at different generations. Such a bound is obtained à la [41, Theorem 1].

Lemma 3.22. *Let $n < m$, $0 \leq k \leq 2^n - 1$ and $0 \leq j \leq 2^m - 1$ be given such that*

$$|2^n j - 2^m k| > 2^{m \vee n}.$$

Then there exists a positive constant ρ_H such that

$$|\text{cov}(\theta_{n,k}, \theta_{m,j})| \leq \rho_H \cdot 2^{-3(m+n)/2} \left(\left| \frac{k+1}{2^n} - \frac{j}{2^m} \right| \wedge \left| \frac{k}{2^n} - \frac{j+1}{2^m} \right| \right)^{2H-4}. \quad (3.55)$$

Proof. For fixed $n, m \in \mathbb{N}$ and $s, t \in \mathbb{R}$, we denote

$$\begin{aligned} f_{n,m}(s, t) &:= \text{cov} \left(W_{s+\frac{1}{2^{n+1}}}^H - W_s^H, W_{t+\frac{1}{2^{m+1}}}^H - W_t^H \right) \\ &= \frac{1}{2} \left(\left| s - t + \frac{1}{2^{n+1}} \right|^{2H} + \left| s - t - \frac{1}{2^{m+1}} \right|^{2H} - |s - t|^{2H} - \left| s - t + \frac{1}{2^{n+1}} - \frac{1}{2^{m+1}} \right|^{2H} \right). \end{aligned} \quad (3.56)$$

Arguing as in the proof of Theorem 2.34 yields

$$\begin{aligned} \text{cov}(\theta_{n,k}, \theta_{m,j}) &= 2^{\frac{n+m}{2}} \int_{\frac{2k}{2^{n+1}}}^{\frac{2k+1}{2^{n+1}}} \int_{\frac{2j}{2^{m+1}}}^{\frac{2j+1}{2^{m+1}}} \frac{\partial^2 f_{n,m}(s, t)}{\partial s \partial t} ds dt \\ &\leq 2^{-\frac{m+n}{2}-2} \sup_{(s,t) \in \mathcal{S}_{\frac{k}{2^n}, \frac{j}{2^m}}^{(n+1, m+1)}} \left| \frac{\partial^2 f_{n,m}(s, t)}{\partial s \partial t} \right|, \end{aligned} \quad (3.57)$$

where for fixed $m, n \in \mathbb{N}$, $\mathcal{S}_{a,b}^{(m,n)}$ denotes the triangular area of the following form

$$\mathcal{S}_{a,b}^{(m,n)} := \left[a, a + \frac{1}{2^n} \right] \times \left[b, b + \frac{1}{2^m} \right] \quad \text{for } a, b \in \mathbb{R}.$$

To provide an upper bound for

$$\sup_{(s,t) \in \mathcal{S}_{\frac{k}{2^n}, \frac{j}{2^m}}^{(n+1, m+1)}} \left| \frac{\partial^2 f_{n,m}(s, t)}{\partial s \partial t} \right|,$$

note that

$$\begin{aligned} \left| \frac{\partial^2 f_{n,m}(s, t)}{\partial s \partial t} \right| &= H|2H-1| \left| |t-s|^{2H-2} + \left| s-t + \frac{1}{2^{n+1}} - \frac{1}{2^{m+1}} \right|^{2H-2} - \left| s-t + \frac{1}{2^{n+1}} \right|^{2H-2} - \left| s-t - \frac{1}{2^{m+1}} \right|^{2H-2} \right| \\ &\leq \frac{1}{2} \prod_{i=0}^3 |2H-i| \int_s^{s+\frac{1}{2^{n+1}}} \int_t^{t+\frac{1}{2^{m+1}}} |u-v|^{2H-4} du dv \\ &\leq 2^{-(n+m)-3} \prod_{i=0}^3 |2H-i| \left(\sup_{(u,v) \in \mathcal{S}_{s,t}^{(n+1, m+1)}} |u-v|^{2H-4} \right). \end{aligned}$$

Applying the above inequality to (3.57) yields

$$|\text{cov}(\theta_{n,k}, \theta_{m,j})| \leq 2^{-\frac{3(m+n)}{2}-5} \prod_{i=0}^3 |2H-i| \sup_{(s,t) \in \mathcal{S}_{\frac{k}{2^n}, \frac{j}{2^m}}^{(n+1, m+1)}} \sup_{(u,v) \in \mathcal{S}_{s,t}^{(n+1, m+1)}} |u-v|^{2H-4}. \quad (3.58)$$

Since $|2^n j - 2^m k| > 2^{m \vee n}$, we have $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \cap \left[\frac{j}{2^m}, \frac{j+1}{2^m}\right] = \emptyset$. As a result, we have

$$\begin{aligned} \sup_{(s,t) \in \mathcal{S}_{\frac{k}{2^n}, \frac{j}{2^m}}^{(n+1, m+1)}} \sup_{(u,v) \in \mathcal{S}_{s,t}^{(n+1, m+1)}} |u-v|^{2H-4} &\leq \sup_{(s,t) \in \mathcal{S}_{\frac{k}{2^n}, \frac{j}{2^m}}^{(n, m)}} |s-t|^{2H-4} \\ &\leq \left(\left| \frac{k}{2^n} - \frac{j+1}{2^m} \right| \wedge \left| \frac{k+1}{2^n} - \frac{j}{2^m} \right| \right)^{2H-4}. \end{aligned}$$

Inserting the above inequality to (3.58) and selecting $\rho_H := 2^{-5} \prod_{i=0}^3 |2H-i|$ completes the proof. $\square$

Recall that $\mathbf{w}_n = Q_n \mathbf{z}_{n+2}$, and let us denote

$$\mathbf{w}_n = (w_{n,0}, w_{n,1}, \dots, w_{n,2^n-1})^\top.$$

By verifying the representation of Q_n in (3.22), we have

$$w_{n,i} = \mathbf{r}^\top \left(z_{4i+1}^{(n+2)}, z_{4i+2}^{(n+2)}, z_{4i+3}^{(n+2)}, z_{4i}^{(n+2)} \right)^\top = \frac{1}{4} \left(-z_{4i+1}^{(n+2)} + z_{4i+2}^{(n+2)} + z_{4i+3}^{(n+2)} - z_{4i+4}^{(n+2)} \right). \quad (3.59)$$

Applying the bound in Lemma 3.22, we can obtain the following result, which concerns the upper bound between the Faber–Schauder coefficient $\theta_{n,k}$ and Gaussian random variable $w_{n,i}$.

Lemma 3.23. *For $n \in \mathbb{N}$, let $0 \leq i, k \leq 2^n - 1$ be given such that $|i - k| > 1$. Then we have*

$$|\text{cov}(\theta_{n,k}, w_{n,i})| \leq \rho_H 2^{(1-2H)n} (|k-i-1| \wedge |k-i+1|)^{2H-4}. \quad (3.60)$$

Proof. We first prove for the case $i > k + 1$. In this context, we have for $1 \leq \delta \leq 4$,

$$\begin{aligned}
\left| \text{cov}(\theta_{n,k}, z_{4i+\delta}^{(n+2)}) \right| &= 2^{3(n+2)/2} \left| \text{cov} \left(\theta_{n,k}, \sum_{m=n+2}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^{m-(n+2)}-1} \theta_{m,j+2^{m-(n+2)}(4i+\delta-1)} \right) \right| \\
&\leq 2^{3(n+2)/2} \sum_{m=n+2}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^{m-(n+2)}-1} \left| \text{cov}(\theta_{n,k}, \theta_{m,j+2^{m-(n+2)}(4i+\delta-1)}) \right| \\
&\leq \rho_H \cdot 2^{3(n+2)/2} \sum_{m=n+2}^{\infty} 2^{-3m-3n/2} \sum_{j=0}^{2^{m-(n+2)}-1} \left(\frac{2^{m-(n+2)}((4i+\delta-1)+j)}{2^m} - \frac{k+1}{2^n} \right)^{2H-4} \\
&= \rho_H \cdot 2^{3(n+2)/2} \sum_{m=n+2}^{\infty} 2^{-3m-3n/2} \sum_{j=0}^{2^{m-(n+2)}-1} \left(\frac{4i+\delta+j-5-4k}{2^{n+2}} \right)^{2H-4}.
\end{aligned}$$

As $j \geq 0$ and $i > k + 1$, we have

$$\begin{aligned}
\left| \text{cov}(\theta_{n,k}, z_{4i+\delta}^{(n+2)}) \right| &\leq \rho_H \cdot 2^{3(n+2)/2} \sum_{m=n+2}^{\infty} 2^{-2m-(n+2)-3n/2} \sup_{0 \leq j \leq 2^{m-(n+2)}-1} \left(\frac{4i+\delta+j-5-4k}{2^{n+2}} \right)^{2H-4} \\
&\leq \rho_H \cdot 2^{-n+1} \sum_{m=n+2}^{\infty} 2^{-2m} \left(\frac{i-k-1}{2^n} \right)^{2H-4} \\
&\leq \rho_H \cdot 2^{-n+1-2(n+2)} \left(\frac{i-k-1}{2^n} \right)^{2H-4} \leq \rho_H \cdot 2^{(1-2H)n} (i-k-1)^{2H-4}.
\end{aligned}$$

Let us now prove for the case $i < k - 1$. Under this circumstance, we now have

$$\begin{aligned}
\left| \text{cov}(\theta_{n,k}, \theta_{m,j+2^{m-(n+2)}(4i+\delta-1)}) \right| &= 2^{-3(n+m)/2} \left(\frac{k}{2^n} - \frac{2^{m-(n+2)}((4i+\delta-1)+j)+1}{2^m} \right)^{2H-4} \\
&= 2^{-3(n+m)/2} \left(\frac{4k-4i-\delta+1-j}{2^{n+2}} - \frac{1}{2^m} \right)^{2H-4}.
\end{aligned}$$

As $\delta \leq 4$, $m \geq n + 2$ and $j \leq 2^{m-(n+2)} - 1$, we must have

$$\frac{4k-4i-\delta+1-j}{2^{n+2}} - \frac{1}{2^m} \geq \frac{k-i-1}{2^{n+2}}$$

Thus, arguing as above yields that

$$\begin{aligned}
\left| \text{cov}(\theta_{n,k}, z_{4i+\delta}^{(n+2)}) \right| &\leq \rho_H \cdot 2^{3(n+2)/2} \sum_{m=n+2}^{\infty} 2^{-2m-(n+2)-3n/2} \sup_{0 \leq j \leq 2^{m-(n+2)}-1} \left(\frac{4k-4i-\delta+1-j}{2^{n+2}} - \frac{1}{2^m} \right)^{2H-4} \\
&\leq \rho_H \cdot 2^{-n+1} \sum_{m=n+2}^{\infty} 2^{-2m} \left(\frac{k-i-1}{2^n} \right)^{2H-4} \leq \rho_H \cdot 2^{(1-2H)n} (k-i-1)^{2H-4}.
\end{aligned}$$

Combining the above bounds for different cases, for $|i - k| > 1$, we have

$$\left| \text{cov}(\theta_{n,k}, z_{4i+\delta}^{(n+2)}) \right| \leq \rho_H \cdot 2^{(1-2H)n} (|k - i - 1| \wedge |k - i + 1|)^{2H-4}.$$

Finally, applying (3.59) gives

$$\left| \text{cov}(\theta_{n,k}, w_{n,i}) \right| \leq \frac{1}{4} \sum_{\delta=1}^4 \left| \text{cov}(\theta_{n,k}, z_{4i+\delta}^{(n+2)}) \right| \leq \rho_H \cdot 2^{(1-2H)n} (|k - i - 1| \wedge |k - i + 1|)^{2H-4}.$$

This completes the proof. $\square$

Given the demonstration in Theorem 3.2, we can show that the distribution $(X_t)_{t \in [0,1]}$ is absolutely continuous with respect to that of $(W_t^H)_{t \in [0,1]}$. Thus, it suffices to establish Theorem 3.7 in the context of fractional Brownian motion. To this end, let $\bar{\theta}_n$, $\bar{\vartheta}_n$, and w_n be as defined in (3.18) and (3.23), respectively, associated with the fractional Brownian motion. Additionally, denote the covariance matrix of $\bar{\vartheta}_n$ by $\Psi_n := (\psi_{i,j}^{(n)})_{i,j=0}^{2^n-1}$. The subsequent lemma explores the behavior of the traces of the covariance matrices $(\Psi_n)_{n \in \mathbb{N}}$.

Lemma 3.24. *For $n \in \mathbb{N}$, we have*

$$\text{trace } \Psi_{n+1} = 2^{2-2H} \text{trace } \Psi_n.$$

Proof. Let us first of all show that the covariance of the Faber–Schauder coefficients of a fractional Brownian motion admits the following properties:

- (a) $\text{cov}(\theta_{n,k}, \theta_{m,j}) = 2^{2H-1} \text{cov}(\theta_{n+1,k}, \theta_{m+1,j})$.
- (b) $\text{cov}(\theta_{n,k}, \theta_{m,j}) = \text{cov}(\theta_{n,k+i}, \theta_{m,j+2^{m-n}i})$.

We show that assertion (a) holds. To this end, recall that $f_{n,m}$ as in (3.56) is given as

$$\begin{aligned} f_{n,m}(s, t) &:= \text{cov} \left(W_{s+\frac{1}{2^{n+1}}}^H - W_s^H, W_{t+\frac{1}{2^{m+1}}}^H - W_t^H \right) \\ &= \frac{1}{2} \left(\left| s - t + \frac{1}{2^{n+1}} \right|^{2H} + \left| s - t - \frac{1}{2^{m+1}} \right|^{2H} - |s - t|^{2H} - \left| s - t + \frac{1}{2^{n+1}} - \frac{1}{2^{m+1}} \right|^{2H} \right). \end{aligned}$$

Clearly, for $s, t \geq 0$, we have

$$f_{n,m}(s, t) = 2^{2H} f_{n+1,m+1}(s/2, t/2).$$

Therefore, we have

$$\begin{aligned}
\text{cov}(\theta_{n+1,k}, \theta_{m+1,j}) &= 2^{\frac{n+m+2}{2}} \left(f_{n+1,m+1} \left(\frac{2k+1}{2^{n+2}}, \frac{2j+1}{2^{m+2}} \right) - f_{n+1,m+1} \left(\frac{2k+1}{2^{n+2}}, \frac{2j}{2^{m+2}} \right) \right. \\
&\quad \left. - f_{n+1,m+1} \left(\frac{2k}{2^{n+2}}, \frac{2j+1}{2^{m+2}} \right) + f_{n+1,m+1} \left(\frac{2k}{2^{n+2}}, \frac{2j}{2^{m+2}} \right) \right) \\
&= 2^{\frac{n+m+2}{2}-2H} \left(f_{n,m} \left(\frac{2k+1}{2^{n+1}}, \frac{2j+1}{2^{m+1}} \right) - f_{n,m} \left(\frac{2k+1}{2^{n+1}}, \frac{2j}{2^{m+1}} \right) \right. \\
&\quad \left. - f_{n,m} \left(\frac{2k}{2^{n+1}}, \frac{2j+1}{2^{m+1}} \right) + f_{n,m} \left(\frac{2k}{2^{n+1}}, \frac{2j}{2^{m+1}} \right) \right) \\
&= 2^{1-2H} \text{cov}(\theta_{n,k}, \theta_{m,j}).
\end{aligned}$$

To prove the second assertion, note that for any $s, t \geq 0$ and any $a \in \mathbb{R}$, we have

$$f_{n,m}(s, t) = f_{n,m}(s+a, t+a).$$

Thus, take $k' = k + i$ and $j' = j + 2^{m-n}i$, we get

$$\frac{k}{2^n} - \frac{j}{2^m} = \frac{k'}{2^n} - \frac{j'}{2^m}.$$

This then yields that

$$\begin{aligned}
\text{cov}(\theta_{n,k}, \theta_{m,j}) &= 2^{\frac{n+m}{2}} \left(f_{n,m} \left(\frac{2k+1}{2^{n+1}}, \frac{2j+1}{2^{m+1}} \right) - f_{n,m} \left(\frac{2k+1}{2^{n+1}}, \frac{2j}{2^{m+1}} \right) \right. \\
&\quad \left. - f_{n,m} \left(\frac{2k}{2^{n+1}}, \frac{2j+1}{2^{m+1}} \right) + f_{n,m} \left(\frac{2k}{2^{n+1}}, \frac{2j}{2^{m+1}} \right) \right) \\
&= 2^{\frac{n+m}{2}} \left(f_{n,m} \left(\frac{2k'+1}{2^{n+1}}, \frac{2j'+1}{2^{m+1}} \right) - f_{n,m} \left(\frac{2k'+1}{2^{n+1}}, \frac{2j'}{2^{m+1}} \right) \right. \\
&\quad \left. - f_{n,m} \left(\frac{2k'}{2^{n+1}}, \frac{2j'+1}{2^{m+1}} \right) + f_{n,m} \left(\frac{2k'}{2^{n+1}}, \frac{2j'}{2^{m+1}} \right) \right) \\
&= \text{cov}(\theta_{n,k+i}, \theta_{m,j+2^{m-n}i}).
\end{aligned}$$

As an immediate corollary of the above assertions, we have must that $\mathbb{E}[\theta_{n+1,k}^2] = 2^{1-2H} \mathbb{E}[\theta_{n,0}^2]$. Moreover, we can simply migrate the above properties from the Faber-Schauder coefficients to the covariance between $\theta_{n,k}$ and $w_{n,k}$.

$$(a) \quad \text{cov}(\theta_{n,k}, w_{n,i}) = 2^{2H-1} \text{cov}(\theta_{n+1,k}, w_{n+1,i}).$$

$$(b) \quad \text{cov}(\theta_{n,k}, w_{n,i}) = \text{cov}(\theta_{n,k+n_0}, w_{n,i+n_0}).$$

Again, let us prove assertion (a) first. To this end, for $1 \leq \delta \leq 4$, we get

$$\begin{aligned}
\text{cov}(\theta_{n+1,k}, z_{4i+\delta}^{(n+3)}) &= 2^{3(n+3)/2} \text{cov} \left(\theta_{n+1,k}, \sum_{m=n+3}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^{m-(n+3)}-1} \theta_{m,j+2^{m-(n+3)}(4i+\delta-1)} \right) \\
&= 2^{3(n+3)/2} \text{cov} \left(\theta_{n+1,k}, \sum_{m=0}^{\infty} 2^{-3(m+n+3)/2} \sum_{j=0}^{2^m-1} \theta_{m+(n+3),j+2^m(4i+\delta-1)} \right) \\
&= \text{cov} \left(\theta_{n+1,k}, \sum_{m=0}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^m-1} \theta_{m+(n+3),j+2^m(4i+\delta-1)} \right) \\
&= \sum_{m=0}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^m-1} \text{cov}(\theta_{n+1,k}, \theta_{m+(n+3),j+2^m(4i+\delta-1)}) \\
&= 2^{1-2H} \text{cov} \left(\theta_{n,k}, \sum_{m=0}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^m-1} \theta_{m+(n+2),j+2^m(4i+\delta-1)} \right) = 2^{1-2H} (\theta_{n,k}, z_{4i+\delta}^{(n+2)}).
\end{aligned}$$

Assertion (a) holds therein, since $w_{n,i}$ is a linear combination of the Gaussian variables $\{z_{4i+\delta}^{(n+2)} : \delta = 1, 2, 3, 4\}$ as in (3.59). To prove assertion (b), note that

$$\begin{aligned}
\text{cov}(\theta_{n,k+n_0}, z_{4(i+n_0)+\delta}^{(n+2)}) &= \sum_{m=0}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^m-1} \text{cov}(\theta_{n,k+n_0}, \theta_{m+(n+2),j+2^m(4i+\delta-1)+2^{m+2}n_0}) \\
&= \sum_{m=0}^{\infty} 2^{-3m/2} \sum_{j=0}^{2^m-1} \text{cov}(\theta_{n,k}, \theta_{m+(n+2),j+2^m(4i+\delta-1)}) = \text{cov}(\theta_{n,k}, z_{4i+\delta}^{(n+2)}).
\end{aligned}$$

Once again, by the as $w_{n,k}$ and $w_{n,k+n_0}$ are constructed as the linear combinations of the $\{z_{4i+\delta}^{(n+2)} : \delta = 1, 2, 3, 4\}$ and $\{z_{4(i+n_0)+\delta}^{(n+2)} : \delta = 1, 2, 3, 4\}$ in the format presented in (3.59), we complete the prove for assertion (b).

Finally, applying the above result, we have

$$\begin{aligned}
\text{trace } \Psi_{n+1} &= \sum_{k=0}^{2^{n+1}-1} \psi_{k,k}^{(n+1)} = \sum_{k=0}^{2^{n+1}-1} \left(\mathbb{E}[\theta_{n+1,k}^2] + \mathbb{E}[w_{n+1,k}^2] + 2\mathbb{E}[\theta_{n+1,k}w_{n+1,k}] \right) \\
&= 2^{n+1+(1-2H)} \left(\mathbb{E}[\theta_{n,0}^2] + \mathbb{E}[w_{n,0}^2] + 2\mathbb{E}[\theta_{n,0}w_{n,0}] \right) = 2^{2-2H} \text{trace } \Psi_n.
\end{aligned}$$

This completes the proof. $\square$

To prove Theorem 3.7, it suffices to prove the following proposition. The results in Theorem 3.7 then simply follow from assertion (d) in Proposition 3.6.

Proposition 3.25. For $H \in (0, 1)$, we take $\varepsilon_n = 2^{-n/2} n^{-1} \sqrt{\log n}$. Then there exist $\tau_H > 0$ and a positive constant c , such that

$$\limsup_{n \uparrow \infty} \varepsilon_n^{-1} \left| H - \widehat{\mathcal{R}}_n^* \left(\frac{W^H}{\sqrt{\tau_H}} \right) \right| \leq c \quad \mathbb{P}\text{-a.s.}$$

Proof. Let us denote $\tau_H = \text{trace } \Psi_0$, it is clear that $\text{trace } \Psi_n = 2^{(2-2H)n} \tau_H$. Furthermore, as shown in Proposition 3.18, as

$$\frac{\mathbb{E}[w_{n,0}^2]}{\mathbb{E}[\theta_{n,0}^2]} = \frac{\alpha(H)}{2^{2+2H} - 2^{4H}} < 1.$$

This then implies that $\tau_H > 0$. Furthermore, following from Lemma 3.19, Lemma 3.21 and Lemma 3.24, the covariance matrix Ψ_n is a symmetric Toeplitz matrix and so $\psi_{i,j}^{(n)} = \psi_{j,i}^{(n)} = \psi_{0,|j-i|}^{(n)}$. Hence, we have $\|\Psi_n\|_1 = \|\Psi_n\|_\infty$, and applying inequalities (3.53), (3.55) and (3.60) yields that there exists a positive constant $\kappa_H > 0$ such that

$$\begin{aligned} \|\Psi_n\|_2 &\leq \sqrt{\|\Psi_n\|_1 \|\Psi_n\|_\infty} = \|\Psi_n\|_1 = \max_{0 \leq j \leq 2^n-1} \sum_{i=0}^{2^n-1} |\psi_{i,j}^{(n)}| \leq 2 \sum_{i=0}^{2^n-1} |\psi_{0,i}^{(n)}| \\ &\leq 2 \left(|\psi_{0,0}^{(n)}| + |\psi_{0,1}^{(n)}| + \sum_{i=2}^{2^n-1} |\psi_{0,i}^{(n)}| \right) \leq \kappa_H 2^{n(1-2H)}, \end{aligned}$$

where the final inequality holds as it follows from the same inequalities that both $2^{n(2H-1)} |\mathbb{E}[\theta_{n,0} \theta_{n,i}]|$, $2^{n(2H-1)} |\mathbb{E}[w_{n,0} w_{n,i}]|$ and $2^{n(2H-1)} |\mathbb{E}[\theta_{n,0} w_{n,i}]|$ are of order $\mathcal{O}(i^{2H-4})$ as $i \uparrow \infty$.

Now, the concentration inequality (3.47) yields that there exists a constant $\lambda > 0$ such that for all $n \in \mathbb{N}$ and $\delta > 0$,

$$\begin{aligned} \mathbb{P} \left[\left| 2^{n(H-1)} \left\| \frac{\bar{\boldsymbol{\vartheta}}_n}{\sqrt{\tau_H}} \right\|_{\ell_2} - 1 \right| \geq \delta \right] &= \mathbb{P} \left[\left| 2^{n(H-1)} \|\bar{\boldsymbol{\vartheta}}_n\|_{\ell_2} - \sqrt{\tau_H} \right| \geq \delta \sqrt{\tau_H} \right] \\ &= \mathbb{P} \left[\left| \|\bar{\boldsymbol{\vartheta}}_n\|_{\ell_2} - \sqrt{\text{trace } \Psi_n} \right| \geq 2^{n(1-H)} \delta \tau_H \right] \\ &\leq \lambda \exp \left(-\frac{2^{n(2-2H)} \delta^2 \tau_H^2}{4 \|\Psi_n\|_2} \right). \end{aligned}$$

Taking $\delta_n := 2^{-n/2} \tau_H^{-1} \sqrt{8\kappa_H \log n}$ yields that

$$\mathbb{P} \left[\left| 2^{n(H-1)} \left\| \frac{\bar{\vartheta}_n}{\sqrt{\tau_H}} \right\|_{\ell_2} - 1 \right| \geq \delta \right] \leq \frac{\lambda}{n^2}.$$

As the expression over the right-hand side of the above inequality is absolutely summable, the Borel–Cantelli lemma yields that

$$\limsup_{n \uparrow \infty} \delta_n^{-1} \left| 2^{n(H-1)} \left\| \frac{\bar{\vartheta}_n}{\sqrt{\tau_H}} \right\|_{\ell_2} - 1 \right| \leq 1 \quad \mathbb{P}\text{-a.s.}$$

Note that

$$H - \widehat{\mathcal{R}}_n^* \left(\frac{Y^H}{\sqrt{\tau_H}} \right) = (H - 1) + \frac{1}{n} \log_2 \left\| \frac{\bar{\vartheta}_n}{\sqrt{\tau_H}} \right\|_{\ell_2} = \frac{1}{n} \log_2 \left(2^{n(H-1)} \left\| \frac{\bar{\vartheta}_n}{\sqrt{\tau_H}} \right\|_{\ell_2} \right),$$

and now using the fact $\sqrt{1+x} - 1 \sim x/2$ as $x \rightarrow 0$ yields the result. This completes the proof. $\square$

Proof of Theorem 3.4. Let us denote the sample path of the time-shifted Riemann–Liouville process $Z^{H,\varepsilon}$ by z . It follows from [84, Theorem 5.8] that the laws of time-shifted Riemann–Liouville process $Z^{H,\varepsilon}$ and fractional Brownian motion W^H are equivalent for any $\varepsilon > 0$. Therefore, the law equivalence between W^H and $Z^{H,\varepsilon}$, together with Proposition 3.18, yields that

$$0 < \liminf_{n \uparrow \infty} 2^{n(2H-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^z \right)^2 \leq \limsup_{n \uparrow \infty} 2^{n(2H-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^z \right)^2 < \infty,$$

with probability one. Furthermore, the sample path v of $V^{H,\varepsilon}$ can be replicated pathwisely through

$$v = \gamma z - \frac{\gamma^2}{2} t^{2H}. \quad t \in [0, 1].$$

It is clear that $t \mapsto \frac{\gamma^2}{2} t^{2H}$ satisfies the conditions in part (b) of Lemma 3.16. Therefore, we have

$$0 < \liminf_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^v \right)^2 \leq \limsup_{n \uparrow \infty} 2^{n(2R-2)} \sum_{k=0}^{2^n-1} \left(\vartheta_{n,k}^v \right)^2 < \infty,$$

with probability one. Finally, the assertion follows from Proposition 3.15. This completes the proof. $\square$

Chapter 4

Robust Faber–Schauder approximation based on discrete observations of an antiderivative

This chapter is based on [46] and studies the problem of reconstructing the Faber–Schauder coefficients of f based on the discrete observations of its antiderivative F . Chapter 4 is organized as follows. In Section 4.1, we formulate our problem and state an explicit solution for each estimated Faber–Schauder coefficient. In Section 4.2, we state our bounds on the approximation error between the estimated and the true Faber–Schauder coefficients. These bounds can be translated into error estimates between the estimated functions $\hat{f}_n$ and $\hat{F}_n$ and the true functions f and F . These results are stated in Section 4.3. The proofs of our main results rely on an analysis of the approximation problem for F in terms of the wavelets formed by the integrated Faber–Schauder functions. The corresponding analysis is presented in Section 4.4. Most proofs for the results stated in Section 4.1 are deferred to Section 4.5. By virtue of the analytical nature of this chapter, we use the notation f to represent a function and F to represent the antiderivative of f .

4.1 Problem formulation and an explicit formula

Recall that the *Faber–Schauder functions* are defined as

$$e_{-1,0}(t) = t, \quad e_{0,0}(t) = (\min\{t, 1 - t\})^+, \quad \text{and} \quad e_{m,k}(t) := 2^{-m/2} e_{0,0}(2^m t - k),$$

for $t \in \mathbb{R}$, $m \in \mathbb{N}$ and $k \in \mathbb{Z}$. It is well known that the restriction of the Faber–Schauder functions to $[0,1]$ forms a Schauder basis of $C[0,1]$. Indeed, for a given function $f \in C[0,1]$,

the function

$$f_n := f(0) + \theta_{-1,0}e_{-1,0} + \sum_{m=0}^n \sum_{k=0}^{2^m-1} \theta_{m,k}e_{m,k}, \quad (4.1)$$

with coefficients $\theta_{-1,0} = f(1) - f(0)$ and

$$\theta_{m,k} = 2^{m/2} \left(2f\left(\frac{2k+1}{2^{m+1}}\right) - f\left(\frac{k}{2^m}\right) - f\left(\frac{k+1}{2^m}\right) \right) \quad \text{for } 0 \leq m \leq n \text{ and } 0 \leq k \leq 2^m - 1, \quad (4.2)$$

is just a piecewise linear interpolation of f with supporting grid $\mathbb{T}_{n+1} := \{k2^{-(n+1)} : k = 0, \dots, 2^{n+1}\}$. In this chapter, we study the question of finding a robust approximation to the Faber–Schauder coefficients of f , if only the values $\{F(t) : t \in \mathbb{T}_{n+1}\}$ of an antiderivative, F , of f are observed.

Our starting point to this question is the following interpolation problem. Based on the given data $\{F(t) : t \in \mathbb{T}_{n+1}\}$, estimate Faber–Schauder coefficients $\vartheta_{m,k}^{(n)}$ and an initial value $\hat{f}_0$ such that

$$\begin{aligned} \hat{F}_n(t) &= F(t) \quad \text{for all } t \in \mathbb{T}_{n+1}, \text{ where } \hat{F}_n(t) = F(0) + \int_0^t \hat{f}_n(s) ds \text{ for} \\ \hat{f}_n &= \hat{f}_0 + \vartheta_{-1,0}^{(n)}e_{-1,0} + \sum_{m=0}^n \sum_{k=0}^{2^m-1} \vartheta_{m,k}^{(n)}e_{m,k}. \end{aligned} \quad (4.3)$$

Since the function $\hat{f}_n$ is continuous and linear on each interval $[k2^{-n-1}, (k+1)2^{-n-1}]$, its antiderivative, $\hat{F}_n$, must be a piecewise quadratic C^1 -interpolation of F with a supporting grid $\mathbb{T}_{n+1}$. The properties of such a quadratic spline interpolation of F have been studied before; see, e.g., [54, 74]. But it is also known that the quadratic spline interpolation suffers from two issues,

- (a) the initial value $\hat{f}_0$ needs to be given so that the coefficients $\vartheta_{m,k}^{(n)}$ are uniquely determined from the data $\{F(t) : t \in \mathbb{T}_{n+1}\}$;
- (b) the interpolating function is notoriously unstable; see Figure 4.1 for an illustration.

In this chapter, we will revisit the classical problem (4.3) within the context of the Faber–Schauder coefficients of $\hat{f}_n$, and we will show that, surprisingly, the source for both issues, (a) and (b), can be identified. It is located in the final generation of the Faber–Schauder coefficients of $\hat{f}_n$, i.e., in the numbers $\vartheta_{n,0}^{(n)}, \dots, \vartheta_{n,2^n-1}^{(n)}$. Some first clues to this observation can be seen from the following theorem, which provides a closed-form solution to the problem (4.3).

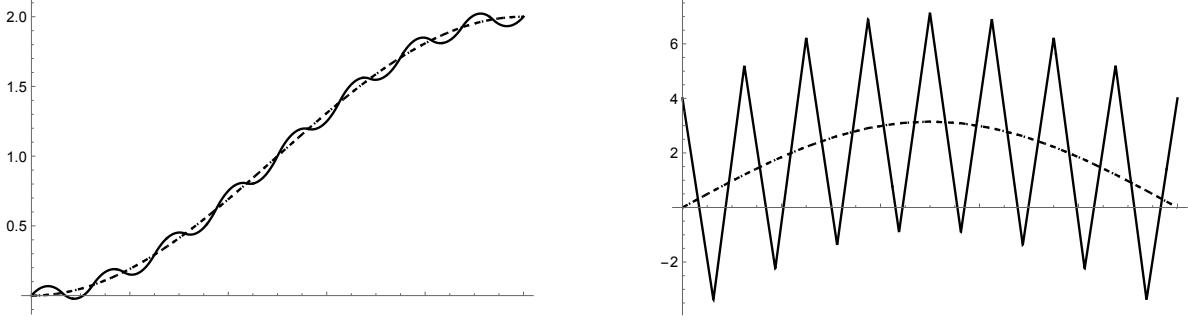


Figure 4.1: Graphical illustration of the instability of the solution to problem (4.3) with $F(t) = 1 - \cos \pi t$ and $n = 3$. While F and $\hat{F}_n$ (left, dotted and dashed respectively) and f and $\hat{f}_n$ (right, dotted and dashed respectively) are basically indistinguishable when taking $\hat{f}_0 = 0$, the value $\hat{f}_0 = 4$ yields wildly oscillating functions $\hat{F}_n$ and $\hat{f}_n$ (solid lines).

Theorem 4.1. *For given $\hat{f}_0$, the coefficients $\{\vartheta_{m,k}^{(n)}\}$ solving problem (4.3) are given as follows. For $m = -1$, we have*

$$\vartheta_{-1,0}^{(n)} = 2^{n+2} \sum_{j=1}^{2^{n+1}} (-1)^j \left(F\left(\frac{j}{2^{n+1}}\right) - F\left(\frac{j-1}{2^{n+1}}\right) \right).$$

For $0 \leq m \leq n-1$ and $0 \leq k \leq 2^m - 1$, we have

$$\vartheta_{m,k}^{(n)} = 2^{n+m/2+2} \sum_{j=1}^{2^{n-m}} (-1)^j \left(F\left(\frac{k}{2^m} + \frac{j}{2^{n+1}}\right) - F\left(\frac{k}{2^m} + \frac{j-1}{2^{n+1}}\right) + F\left(\frac{k+1}{2^m} - \frac{j-1}{2^{n+1}}\right) - F\left(\frac{k+1}{2^m} - \frac{j}{2^{n+1}}\right) \right), \quad (4.4)$$

Finally, for $m = n$ and $0 \leq k \leq 2^n - 1$, we have

$$\begin{aligned} \vartheta_{n,k}^{(n)} = & -2^{n/2+2} \hat{f}_0 - 2^{3n/2+4} \sum_{j=1}^{2k} (-1)^j \left(F\left(\frac{j}{2^{n+1}}\right) - F\left(\frac{j-1}{2^{n+1}}\right) \right) \\ & + 3 \cdot 2^{3n/2+2} \left(F\left(\frac{2k+1}{2^{n+1}}\right) - F\left(\frac{2k}{2^{n+1}}\right) \right) - 2^{3n/2+2} \left(F\left(\frac{2k+2}{2^{n+1}}\right) - F\left(\frac{2k+1}{2^{n+1}}\right) \right). \end{aligned}$$

The first observation we can make from Theorem 4.1 is that only the coefficients of the final generation, i.e., the numbers $\vartheta_{n,0}^{(n)}, \dots, \vartheta_{n,2^n-1}^{(n)}$, depend on the initial value $\hat{f}_0$. Moreover, the formula for the final-generation coefficients $\vartheta_{n,0}^{(n)}, \dots, \vartheta_{n,2^n-1}^{(n)}$ contains the additive term $-2^{n/2+2} \hat{f}_0$, which implies that any error made in estimating $\hat{f}_0$ translates into an additional $2^{n/2+2}$ -fold error for each final-generation coefficient. In particular, the approximation errors $|\vartheta_{n,k}^{(n)} - \theta_{n,k}|$ can be made arbitrarily large by varying the initial condition $\hat{f}_0$. By contrast, all other Faber–Schauder coefficients, i.e., $\vartheta_{m,k}^{(n)}$ with $m \leq n-1$, are independent of $\hat{f}_0$. Thus, the instability with respect to $\hat{f}_0$ observed in Figure 4.1 arises exclusively from the coefficients $\vartheta_{n,0}^{(n)}, \dots, \vartheta_{n,2^n-1}^{(n)}$ of the final generation.

Our second observation is the following counterintuitive phenomenon, which arises independently of the initial value $\hat{f}_0$. On the one hand, each estimated coefficient $\vartheta_{m,k}^{(n)}$ with $m \leq n-1$ is a linear combination of the values of F over the grid $\mathbb{T}_{n+1}$ within the interval $[2^{-m \vee 0}k, 2^{-m \vee 0}(k+1)]$, where $a \vee b$ is shorthand for $\max\{a, b\}$. This interval coincides with the support of the corresponding wavelet function $e_{m,k}$. By contrast, for $m = n$, the estimated coefficient $\vartheta_{n,k}^{(n)}$ involves all observations of F within the interval $[0, 2^{-n}(k+1)]$. This means that when approximating the derivative f by the piecewise linear function $\hat{f}_n$, its behavior within $[2^{-n}k, 2^{-n}(k+1)]$ is affected by values of F potentially far away from that region. A graphical illustration of this phenomenon is presented in Figure 4.2, and a concrete computation will be given in Example 4.2. This nonlocal dependence of the coefficients $\vartheta_{n,k}^{(n)}$ on the observed data is counterintuitive and another cause for the instability of the quadratic spline interpolation.

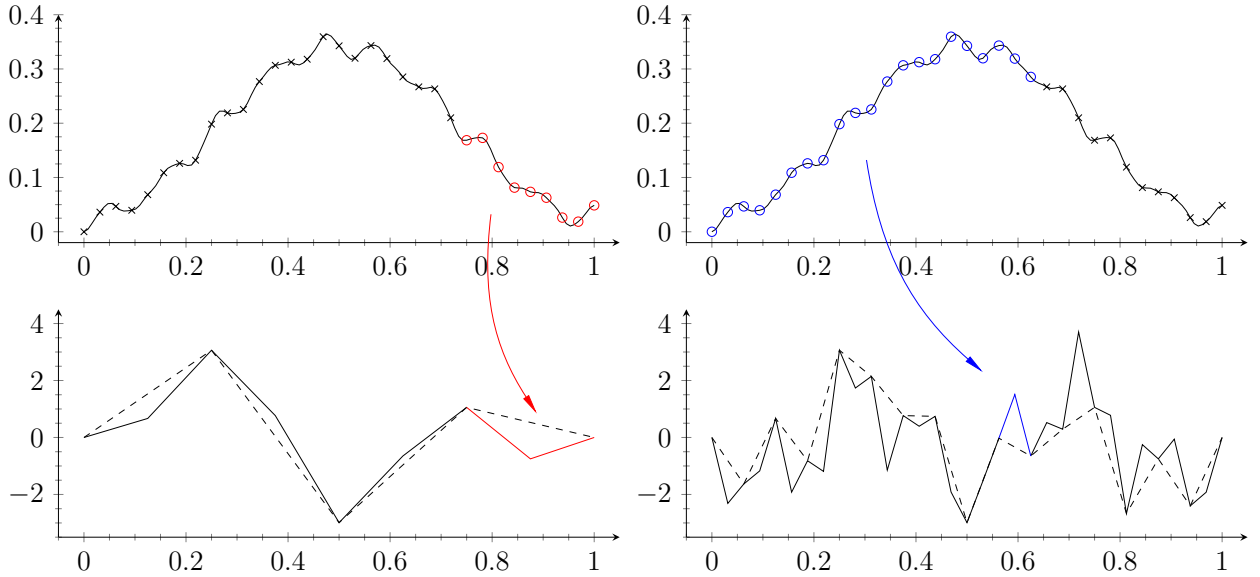


Figure 4.2: Graphical illustrations of the data dependence of the coefficients $\vartheta_{m,k}^{(n)}$. In the bottom-left graph, we plot the approximating functions $\hat{f}_{4,1}$ and $\hat{f}_{4,2}$ in dotted and solid lines respectively, and the contribution $\vartheta_{2,3}^{(4)}e_{2,3}$ is highlighted in red. In the bottom-right graph, we plot the approximating functions $\hat{f}_{4,3}$ and $\hat{f}_{4,4}$ in dotted and solid lines respectively, and the contribution $\vartheta_{4,9}^{(4)}e_{4,9}$ is highlighted in blue. Respective data points of the function F that contribute to the estimation of $\vartheta_{2,3}^{(4)}$ and $\vartheta_{4,9}^{(4)}$ are highlighted by circles in the corresponding upper graphs. All data points that do not contribute to the computation of the coefficients in question are represented by cross marks.

In the following section, we will discover yet another issue with the final-generation Faber–Schauder coefficients of $\hat{f}_n$, which arises even if the true initial value $f(0) = F'(0)$ is known, and $\hat{f}_0$ is set to that value. In this context, we will derive sharp upper and

lower bounds on the approximation errors for the estimated Faber–Schauder coefficients up to generation $n - 1$ and at generation n . That is, we will be looking at the Euclidean ℓ_p -distances between the Faber–Schauder coefficients of $\hat{f}_n$ and the true Faber–Schauder coefficients of f for $p \in \{1, 2, \infty\}$ and under the assumption that $f(0) = F'(0)$ is known. On the one hand, we will discover robust upper error bounds for the estimated Faber–Schauder coefficients up to generation $n - 1$, which hold universally for each function $F \in C^1[0, 1]$. On the other hand, we will see that $F \in C^1[0, 1]$ can always be chosen in such a way that the error in generation n exceeds the error up to generation $n - 1$ by a factor of size $\mathcal{O}(2^n)$ for the ℓ_2 - and ℓ_1 -norms and by a factor of size $\mathcal{O}(2^{n/2})$ for the ℓ_∞ -norm. In Example 4.2, we will encounter an extreme case in which the approximation errors up to and including generation $n - 1$ are all zero, while the error in generation n can go to infinity even if the correct initial value $f(0)$ is known.

Conclusion. Our findings can informally be summarized as follows.

- (a) The estimated coefficients $\vartheta_{m,k}^{(n)}$ for $m \leq n - 1$ are independent of the estimated initial value $\hat{f}_0$, and they admit robust error bounds for their convergence to the corresponding true Faber–Schauder coefficients of $f = F'$.
- (b) By contrast, the final-generation coefficients $\vartheta_{n,0}^{(n)}, \dots, \vartheta_{n,2^n-1}^{(n)}$ are strongly dependent on the estimated initial value $\hat{f}_0$. Even if the true initial value is known, they depend on F in a nonlocal manner. Furthermore, the errors for their convergence to the corresponding true Faber–Schauder coefficients of $f = F'$ are highly sensitive with respect to the function F .

We conclude this section with the following, computationally explicit example, in which the formulas from Theorem 4.1 recover the exact Faber–Schauder coefficients of f up to and including generation $n - 1$. In generation n , however, the estimated Faber–Schauder coefficients will differ from the true ones to the extent that their approximation errors may go to infinity even if the correct initial value $f(0)$ is known.

Example 4.2. Consider the Takagi class, which was introduced by Hata and Yamaguti [52] and motivated by the celebrated, continuous but nowhere differentiable Takagi function. It consists of all functions of the following form:

$$f(t) := \sum_{m=0}^{\infty} c_m \phi(2^m t) = \sum_{m=0}^{\infty} 2^{m/2} c_m \sum_{k=0}^{2^m-1} e_{m,k}(t), \quad \text{for } t \in [0, 1], \quad (4.5)$$

where $\{c_m\}_{m \in \mathbb{N}}$ is an absolutely summable sequence and $\phi(t) := \min_{z \in \mathbb{Z}} |t - z|$ denotes the tent map. The Takagi class has been studied intensively over the past decades; see, e.g., the

surveys [2, 66]. Letting $F(t) := \int_0^t f(s) ds$ and $\hat{f}_0 := f(0) = 0$, the coefficients solving (4.3) are given as follows,

$$\vartheta_{m,k}^{(n)} = \begin{cases} 2^{m/2} c_m & \text{for } 0 \leq m \leq n-1 \text{ and } 0 \leq k \leq 2^m - 1, \\ 2^{n/2} \sum_{m=n}^{\infty} c_m & \text{for } m = n \text{ and } 0 \leq k \leq 2^n - 1. \end{cases}$$

The proof for these formulas is deferred to Section 4.5. It follows that up to and including generation $n-1$, the approximation error between all true and estimated Faber–Schauder coefficients is zero. For generation n , however, we have $|\vartheta_{n,k}^{(n)} - \theta_{n,k}| = 2^{n/2} |\sum_{m=n+1}^{\infty} c_m|$. In particular, if the decay of $|\sum_{m=n+1}^{\infty} c_m|$ is slower than $O(2^{-n/2})$, then the approximation errors $|\vartheta_{n,k}^{(n)} - \theta_{n,k}|$ go to infinity as $n \uparrow \infty$.

4.2 Error analysis for the Faber–Schauder coefficients

In this section, we consider again functions $F \in C^1[0, 1]$ with derivative $F' = f$ and we suppose that $f(0)$ is known so that we can set $\hat{f}_0 := f(0)$ in our problem (4.3). As a matter of fact, we may assume without loss of generality that $f(0) = \hat{f}_0 = 0$ and $F(0) = 0$; otherwise we consider the function $F(t) - F(0) - f(0)t$, whose derivative has the same Faber–Schauder coefficients as f . We refer to [54] for a method for estimating $f(0)$ from the data $\{F(t) : t \in \mathbb{T}_{n+1}\}$.

For studying the Euclidean ℓ_p -norms of the approximation errors, it will be convenient to denote the Faber–Schauder coefficients of f at or up to generation $m \geq -1$ by the respective column vectors $\bar{\theta}_m$ and θ_m , i.e.,

$$\bar{\theta}_m = (\theta_{m,0}, \theta_{m,1}, \dots, \theta_{m,2^{m \vee 0}-1})^\top \in \mathbb{R}^{2^{m \vee 0}} \quad \text{and} \quad \theta_m = (\bar{\theta}_{-1}^\top, \bar{\theta}_0^\top, \dots, \bar{\theta}_m^\top)^\top \in \mathbb{R}^{2^{m+1}}.$$

An analogous notation will be used for the coefficients in Theorem 4.1 solving (4.3). That is, for $-1 \leq m \leq n$, we set

$$\bar{\vartheta}_m^{(n)} = (\vartheta_{m,0}^{(n)}, \vartheta_{m,1}^{(n)}, \dots, \vartheta_{m,2^{m \vee 0}-1}^{(n)})^\top \in \mathbb{R}^{2^{m \vee 0}} \quad \text{and} \quad \vartheta_m^{(n)} = ((\bar{\vartheta}_{-1}^{(n)})^\top, (\bar{\vartheta}_0^{(n)})^\top, \dots, (\bar{\vartheta}_m^{(n)})^\top)^\top \in \mathbb{R}^{2^{m+1}}.$$

Note that in particular $\vartheta_n^{(n)} = \vartheta^{(n)}$, and both notions will be used interchangeably in this chapter. In this section, we study the approximation errors $\theta_m - \vartheta_m^{(n)}$ and $\bar{\theta}_m - \bar{\vartheta}_m^{(n)}$ by giving upper and lower bounds on their Euclidean ℓ_p -norms for $p \in \{1, 2, \infty\}$. These bounds will be given in terms of the ℓ_p -norms of the 2^{n+1} -dimensional column vector $\mathbf{z}_{n+1} :=$

$(z_1^{(n+1)}, \dots, z_{2^{n+1}}^{(n+1)})^\top$, where

$$z_i^{(n+1)} = 2^{3(n+1)/2} \sum_{m=n+1}^{\infty} 2^{-3m/2} \sum_{k=2^{m-n-1}(i-1)}^{2^{m-n-1}i-1} \theta_{m,k} \quad \text{for } 1 \leq i \leq 2^{n+1}. \quad (4.6)$$

The following criterion guarantees that the ℓ_p -norm of $\mathbf{z}_{n+1}$ converges to zero as $n \uparrow \infty$. Combined with our subsequent error bounds, this proposition will yield a sufficient criterion for the convergence of the estimated Faber–Schauder coefficients to their true values.

Proposition 4.3. *Suppose that f is continuously differentiable and that its derivative is Hölder continuous with exponent $\alpha \in (0, 1]$. Then*

$$\|\mathbf{z}_{n+1}\|_{\ell_p} = \begin{cases} \mathcal{O}(2^{-(n+1)(\alpha-\frac{1}{2})}) & \text{for } p = 1, \\ \mathcal{O}(2^{-(n+1)\alpha}) & \text{for } p = 2, \\ \mathcal{O}(2^{-(n+1)(\alpha+\frac{1}{2})}) & \text{for } p = \infty. \end{cases} \quad (4.7)$$

The following theorem gives sharp upper and lower bounds for the ℓ_2 -norm of the approximation error $\boldsymbol{\vartheta}_{n-1}^{(n)} - \boldsymbol{\theta}_{n-1}$ of the estimated Faber–Schauder coefficients up to and including generation $n - 1$.

Theorem 4.4. *The following assertions hold.*

(a) *For arbitrary choice of F , we have*

$$\|\boldsymbol{\vartheta}_{n-1}^{(n)} - \boldsymbol{\theta}_{n-1}\|_{\ell_2} \leq \frac{1}{2} \|\mathbf{z}_{n+1}\|_{\ell_2}. \quad (4.8)$$

(b) *For each $\varepsilon > 0$ and $-1 \leq m \leq n - 1$, the function F can be chosen in such a way that*

$$\|\boldsymbol{\vartheta}_{n-1}^{(n)} - \boldsymbol{\theta}_{n-1}\|_{\ell_2} \geq \|\bar{\boldsymbol{\vartheta}}_m^{(n)} - \bar{\boldsymbol{\theta}}_m\|_{\ell_2} \geq (1 - \varepsilon) \frac{1}{2} \|\mathbf{z}_{n+1}\|_{\ell_2}.$$

Our next results look specifically at the error term $\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_2}$ for the Faber–Schauder coefficients of the n^{th} generation.

Theorem 4.5. *For $n \geq 2$, the following assertions hold:*

(a) *For arbitrary choice of F , we have*

$$\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_2} \leq \sqrt{2 \left(1 - \cos \frac{\pi}{2^{n+1}}\right)^{-1} + \frac{3}{4}} \cdot \|\mathbf{z}_{n+1}\|_{\ell_2}. \quad (4.9)$$

(b) For each $\varepsilon > 0$, the function F can be chosen in such a way that

$$\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_2} \geq (1 - \varepsilon) \sqrt{2 \left(1 - \cos \frac{\pi}{2^{n+1}}\right)^{-1} - \frac{3}{4}} \cdot \|\mathbf{z}_{n+1}\|_{\ell_2}. \quad (4.10)$$

Since¹ $1 - \cos x \sim x^2/2$ as $x \downarrow 0$, the coefficients in (4.9) and (4.10) behave asymptotically as

$$\sqrt{2 \left(1 - \cos \frac{\pi}{2^{n+1}}\right)^{-1} \pm \frac{3}{4}} \sim \frac{2^{n+1}}{\pi} \quad \text{for } n \uparrow \infty. \quad (4.11)$$

As a matter of fact, by combining the elementary inequality $1 - \cos x \leq x^2/2$ with Theorem 4.4 (a) and Theorem 4.5 (b), we obtain the following corollary. As mentioned above, it states that the ℓ_2 -norm of the error in the final generation of the Faber–Schauder coefficients can be substantially larger than a factor of size $\mathcal{O}(2^n)$ times the ℓ_2 -norm of the error of all previous generations combined. Moreover, it is worthwhile to point out that both vectors $\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n$ and $\boldsymbol{\vartheta}_{n-1}^{(n)} - \boldsymbol{\theta}_{n-1}$ are of length 2^n . Thus, the relatively significant error in the final-generation Faber–Schauder coefficients purely results from the quadratic spline interpolation.

Corollary 4.6. *For all $n \geq 2$ and $\varepsilon > 0$, the function F can be chosen in such a way that*

$$\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_2} \geq 2(1 - \varepsilon) \sqrt{\frac{2^{2n+2}}{\pi^2} - \frac{3}{4}} \|\boldsymbol{\vartheta}_{n-1}^{(n)} - \boldsymbol{\theta}_{n-1}\|_{\ell_2}.$$

Our bounds for the ℓ_1 - and ℓ_∞ -norms, as stated in the following theorem, are even stronger than (4.9) and (4.10), because the constants for the upper and lower bounds are the same.

Theorem 4.7. *We fix $n \geq 2$ and let for $m = -1, 0, \dots, n-1$ and $p \in \{1, \infty\}$,*

$$\bar{c}_{p,m} := \begin{cases} 2^{\frac{m \vee 0 - n - 3}{2}} & \text{if } p = 1, \\ 2^{\frac{n - m \vee 0 - 1}{2}} & \text{if } p = \infty, \end{cases} \quad c_{p,m} := \begin{cases} 2^{-\frac{n+3}{2}} \left(\frac{2^{\frac{m+1}{2}} - 1}{2^{1/2} - 1} \mathbb{1}_{\{m \geq 0\}} + 1 \right) & \text{if } p = 1, \\ 2^{\frac{n-1}{2}} & \text{if } p = \infty, \end{cases}$$

$$\bar{c}_{p,n} := \begin{cases} 2^{n+\frac{1}{2}} - 2^{-3/2} & \text{if } p = 1, \\ 2^{n+\frac{3}{2}} - \sqrt{2} & \text{if } p = \infty. \end{cases}$$

Then the following assertions hold for $m = -1, 0, \dots, n-1$ and $p \in \{1, \infty\}$.

- (a) *For arbitrary choice of F , we have $\|\bar{\boldsymbol{\vartheta}}_m^{(n)} - \bar{\boldsymbol{\theta}}_m\|_{\ell_p} \leq \bar{c}_{p,m} \|\mathbf{z}_{n+1}\|_{\ell_p}$, and for any given $\varepsilon > 0$, the function F can be chosen in such a way that $\|\bar{\boldsymbol{\vartheta}}_m^{(n)} - \bar{\boldsymbol{\theta}}_m\|_{\ell_p} > (1 - \varepsilon) \bar{c}_{p,m} \|\mathbf{z}_{n+1}\|_{\ell_p}$.*

¹For functions $g(x)$ and $h(x)$, we say $g \sim h$ as $x \rightarrow x_0$ if $g(x)/h(x) \rightarrow 1$ for $x \rightarrow x_0$.

- (b) For arbitrary choice of F , we have $\|\boldsymbol{\vartheta}_m^{(n)} - \boldsymbol{\theta}_m\|_{\ell_p} \leq c_{p,m} \|\mathbf{z}_{n+1}\|_{\ell_p}$, and for any given $\varepsilon > 0$, the function F can be chosen in such a way that $\|\boldsymbol{\vartheta}_m^{(n)} - \boldsymbol{\theta}_m\|_{\ell_p} > (1 - \varepsilon) c_{p,m} \|\mathbf{z}_{n+1}\|_{\ell_p}$.
- (c) For arbitrary choice of F , we have $\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_p} \leq \bar{c}_{p,n} \|\mathbf{z}_{n+1}\|_{\ell_p}$, and for any given $\varepsilon > 0$, the function F can be chosen in such a way that $\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_p} > (1 - \varepsilon) \bar{c}_{p,n} \|\mathbf{z}_{n+1}\|_{\ell_p}$.
- (d) For all $n \geq 2$ and $\varepsilon > 0$, the function F can be chosen in such a way that $\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_p} \geq (1 - \varepsilon) \bar{c}_{p,n} c_{p,n-1}^{-1} \|\boldsymbol{\vartheta}_{n-1}^{(n)} - \boldsymbol{\theta}_{n-1}\|_{\ell_p}$.

Note that part (d) in the preceding theorem plays the role of Corollary 4.6 for the ℓ_1 - and ℓ_∞ -norms and that the coefficients in that part satisfy

$$\frac{\bar{c}_{1,n}}{c_{1,n-1}} = \frac{(2^{n+\frac{1}{2}} - 2^{-3/2})}{2^{-\frac{n+3}{2}} \left(\frac{2^{n/2}-1}{2^{1/2}-1} + 1 \right)} \sim (\sqrt{2} - 1) 2^{n+2} \quad \text{as } n \uparrow \infty,$$

and

$$\frac{\bar{c}_{\infty,n}}{c_{\infty,n-1}} = \frac{2^{n+\frac{3}{2}} - \sqrt{2}}{2^{\frac{n-1}{2}}} \sim 2^{\frac{n}{2}+2} \quad \text{as } n \uparrow \infty.$$

Thus, for all $p \in \{1, 2, \infty\}$, the ratio between $\|\bar{\boldsymbol{\vartheta}}_n^{(n)} - \bar{\boldsymbol{\theta}}_n\|_{\ell_p}$ and $\|\boldsymbol{\vartheta}_{n-1}^{(n)} - \boldsymbol{\theta}_{n-1}\|_{\ell_p}$ can become arbitrarily large as n goes to infinity. This underscores our thesis that $\boldsymbol{\vartheta}_{n-1}^{(n)}$ is a more robust estimate for the true Faber–Schauder coefficients of f than $\boldsymbol{\vartheta}_n^{(n)}$.

Theorems 4.4, 4.5, and 4.7 are actually corollaries to our results stated in Section 4.4. More specifically, they follow immediately from the equations (4.21) and (4.22) and Theorem 4.14.

4.3 Error analysis for the approximating functions

As discussed above, our analysis of the estimated Faber–Schauder coefficients suggests that dropping the final generation $\bar{\boldsymbol{\vartheta}}_n^{(n)} = (\vartheta_{n,0}^{(n)}, \dots, \vartheta_{n,2^n-1}^{(n)})^\top$ of coefficients leads to a more robust estimate $\boldsymbol{\vartheta}_{n-1}^{(n)}$ of the true Faber–Schauder coefficients of f . A Faber–Schauder expansion based on these remaining coefficients should in turn lead to more robust approximations for the unknown functions f and F . There is, however, a price to be paid for this robustness: the new approximating functions will no longer be interpolations of f and F . That is, they may no longer coincide with f and F on the observation grid $\mathbb{T}_{n+1}$.

To analyze these induced approximating functions, we let again $F \in C^1[0, 1]$ with $f = F'$

and assume as in Section 4.2 that $F(0) = f(0) = 0$. For $m = -1, 0, \dots, n$, we then define

$$\hat{f}_{n,m} = \sum_{i=-1}^m \sum_{k=0}^{2^{i \vee 0} - 1} \vartheta_{i,k}^{(n)} e_{i,k} \quad \text{and} \quad \hat{F}_{n,m} = \sum_{i=-1}^m \sum_{k=0}^{2^{i \vee 0} - 1} \vartheta_{i,k}^{(n)} \psi_{i,k}, \quad (4.12)$$

where the coefficients $\vartheta_{i,k}^{(n)}$ are as above. Then both $\hat{f}_{n,m}$ and the function f_m defined in (4.1) will be piecewise linear with supporting grid $\mathbb{T}_{m+1}$.

Error analysis for quadratic spline interpolation has been well studied in the existing literature. Typically, results such as those in [74] are based on the mean-value theorem and provide error bounds for the $L_\infty[0, 1]$ -norm in terms of the first-order modulus of continuity of f . By contrast, our results presented here are based on the theorems in Section 4.2. They provide error bounds with respect to the $L_p[0, 1]$ -norms for $p \in \{1, 2, \infty\}$ and in terms of the second-order modulus of continuity of f ,

$$\omega_2(f, t) := \sup_{\substack{\tau_1, \tau_2 \in [0, 1] \\ |\tau_1 - \tau_2| \leq t}} \left\{ \left| f(\tau_1) + f(\tau_2) - 2f\left(\frac{\tau_1 + \tau_2}{2}\right) \right| \right\}.$$

Our results will involve the following constants, which we define for fixed $n \geq 2$,

$$\begin{aligned} A_p &:= \begin{cases} 2^{-(n+3)/2} & \text{for } p = 1, \\ 1/\sqrt{3} & \text{for } p = 2, \\ 2^{(n+1)/2} & \text{for } p = \infty, \end{cases} & B_p &:= \begin{cases} 2^n & \text{for } p = 1, \\ 2^{n+1}/\sqrt{3} & \text{for } p = 2, \\ 2^{n+1} & \text{for } p = \infty, \end{cases} \\ C_p &:= \begin{cases} 2^{-(n+1)/2} & \text{for } p = 1, \\ \gamma_2 & \text{for } p = 2, \\ 2^{(n+3)/2} & \text{for } p = \infty, \end{cases} & D_p &:= \begin{cases} 2^{n+1} & \text{for } p = 1, \\ 2^{n+1}\gamma_2 & \text{for } p = 2, \\ 2^{n+2} & \text{for } p = \infty, \end{cases} \end{aligned} \quad (4.13)$$

where

$$\gamma_2 := \frac{1}{\sqrt{3}} \left(1 + 2^{-n-1} \sqrt{2 \left(1 - \cos \frac{\pi}{2^{n+1}} \right)^{-1} + \frac{3}{4}} \right).$$

Theorem 4.8. *The following inequalities hold for fixed $n \geq 2$, $0 \leq m \leq n-1$, and $p \in \{1, 2, \infty\}$,*

$$\begin{aligned} \|f_m - \hat{f}_{n,m}\|_{L_p[0,1]} &\leq A_p \|z_{n+1}\|_{\ell_p} \leq B_p \sum_{k=n+1}^{\infty} \omega_2(f, 2^{-k}), \\ \|f_n - \hat{f}_{n,n}\|_{L_p[0,1]} &\leq C_p \|z_{n+1}\|_{\ell_p} \leq D_p \sum_{k=n+1}^{\infty} \omega_2(f, 2^{-k}) \end{aligned}$$

Remark 4.9. Recall from (4.3) that $\|z_{n+1}\|_{\ell_p} \rightarrow 0$ as $n \uparrow \infty$ if f is continuously differentiable with Hölder continuous derivative. Thus, in this case, Theorem 4.8 yields the result that $\|f_{m_n} - \hat{f}_{n,m_n}\|_{L_p[0,1]} \rightarrow 0$ if (m_n) is any sequence with $m_n \leq n$.

Remark 4.10. Theorem 4.8 gives error bounds between the truncated function f_m and the approximation $\hat{f}_{n,m}$. Error bounds for $\|f - \hat{f}_{n,m}\|_{L_p[0,1]}$ can be obtained as follows. For instance, for $p = 1$, we observe that $\|e_{m,k}\|_{L_1[0,1]} = 2^{-3m/2-2}$ and $|\theta_{m,k}| \leq 2^{m/2}\omega_2(f, 2^{-m})$. From here, it is easy to show that

$$\|f - f_m\|_{L_1[0,1]} \leq \frac{1}{4} \sum_{k=m+1}^{\infty} \omega_2(f, 2^{-k}).$$

The desired error bound for $\|f - \hat{f}_{n,m}\|_{L_1[0,1]}$ can now be obtained in a straightforward manner from the triangle inequality and Theorem 4.8. Similar arguments also work in the cases of the L_2 - and L_∞ -norms, where we can use that $\|e_{m,k}\|_{L_2[0,1]} = 2^{-m-1}/\sqrt{3}$ and $\|e_{m,k}\|_{L_\infty[0,1]} = 2^{-m/2-1}$. In the L^∞ -case, one can alternatively use the fact that $\|f - f_m\|_{L_\infty[0,1]} \leq 4\omega_2(f, 2^{-m-1})$, which was established in [73, Theorem 1].

By combining the bounds explained in the preceding remark with the next theorem, one obtains bounds for the respective error $|F - \hat{F}_{n,n}|$ in various L_p -norms.

Theorem 4.11. *For $n \geq 2$, and $p \in \{1, 2, \infty\}$, we have*

$$\|F - \hat{F}_{n,n}\|_{L_p[0,1]} \leq a_p \|f - \hat{f}_{n,n}\|_{L_p[0,1]},$$

where $a_1 = 2^{-n-1}$, $a_2 = 2^{-n}/\pi$, and $a_\infty = 2^{-n-2}$.

4.4 Matrix analysis of the approximation problem

In this section, we present our approach to the proofs of the main results in this chapter. It consists of a formulation of the problem (4.3) by means of linear algebra. To this end, consider the following wavelet system. We define $\psi_{-1,0}(t) = t^2/2$,

$$\psi_{0,0}(t) = \begin{cases} 0 & t \in (-\infty, 0), \\ \frac{1}{2}t^2 & t \in [0, \frac{1}{2}), \\ \frac{1}{4} - \frac{1}{2}(1-t)^2 & t \in [\frac{1}{2}, 1), \\ \frac{1}{4} & t \in [1, \infty), \end{cases} \quad \text{and} \quad \psi_{m,k}(t) = 2^{-3m/2}\psi_{0,0}(2^m t - k) \quad (4.14)$$

for $m \in \mathbb{N}$ and $0 \leq k \leq 2^m - 1$.

Remark 4.12. One easily verifies that $\int_0^t e_{m,k}(s) ds = \psi_{m,k}(t)$ for $m \geq -1$, $0 \leq k \leq (2^m - 1) \vee 0$, and $t \in \mathbb{R}$. Therefore, the fact that the Faber–Schauder functions are a Schauder basis of $C[0, 1]$ implies that the collection of all functions $\psi_{m,k}$ augmented with $\{1, t\}$ forms a Schauder basis for the Banach space $C^1[0, 1]$.

Using the functions $\{\psi_{m,k}\}$, we can transform the problem (4.3) into a standard linear equation. To this end, we assume again that $f(0) = 0$ is known, that $\hat{f}_0 = 0$, and that $F(0) = 0$; see the beginning of Section 4.2. We denote for $m \geq -1$ and $k \in \mathbb{N}$,

$$Q_{(m,k)} := \begin{bmatrix} \psi_{m,0}\left(\frac{1}{2^k}\right) & \psi_{m,1}\left(\frac{1}{2^k}\right) & \cdots & \psi_{m,2^{m \vee 0}-1}\left(\frac{1}{2^k}\right) \\ \psi_{m,0}\left(\frac{2}{2^k}\right) & \psi_{m,1}\left(\frac{2}{2^k}\right) & \cdots & \psi_{m,2^{m \vee 0}-1}\left(\frac{2}{2^k}\right) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{m,0}(1) & \psi_{m,1}(1) & \cdots & \psi_{m,2^{m \vee 0}-1}(1) \end{bmatrix} \in \mathbb{R}^{2^k \times 2^{m \vee 0}}. \quad (4.15)$$

Next, for $m, k \in \mathbb{N}$, we write

$$\Psi_{(m,k)} = \left[Q_{(-1,k)}, \dots, Q_{(m,k)} \right] \in \mathbb{R}^{2^k \times 2^{m+1}}. \quad (4.16)$$

Remark 4.12 implies that $\Psi_{(m,k)}$ has full rank. In particular, $\Psi_{n+1} := \Psi_{(n,n+1)} \in \mathbb{R}^{2^{n+1} \times 2^{n+1}}$ is an invertible square matrix. We also use the shorthand notation

$$\mathbf{y}_{n+1} := \left(F\left(\frac{1}{2^{n+1}}\right), F\left(\frac{2}{2^{n+1}}\right), \dots, F(1) \right)^\top \in \mathbb{R}^{2^{n+1}}$$

for the values of F over $\mathbb{T}_{n+1}$, our given data points. Using our assumptions $F(0) = 0$ and $\hat{f}_0 = 0$, the problem (4.3) then becomes equivalent to the 2^{n+1} -dimensional linear equation

$$\boldsymbol{\vartheta}^{(n)} = \Psi_{n+1} \mathbf{y}_{n+1}, \quad (4.17)$$

in which the solution $\boldsymbol{\vartheta}^{(n)}$ describes the Faber–Schauder coefficients of $\hat{f}_n$ or, equivalently, the coefficients in the expansion of $\hat{F}_n$ in terms of the basis functions $\psi_{m,k}$.

It is a natural guess that the linear system (4.17) can be solved by a simple numerical inversion of the corresponding matrix Ψ_{n+1} . In practice, however, the numerical inversion of the matrix Ψ_n becomes extremely unstable as the number of observation points increases; see Table 4.1. The main reason for this instability is due to the fact that matrix columns corresponding to neighboring wavelet functions become asymptotically co-linear. Fortunately, it is possible to algebraically invert the matrix Ψ_n . This inversion formula gives rise to an

Numerical values for $\log_{10} \det(\Psi_n)$ for several values of n					
n	2	3	4	5	6
$\log_{10} \det(\Psi_n)$	-4.97	-13.4	-33.9	-82.03	-192.81

Table 4.1: Numerical values for $\log_{10} \det(\Psi_n)$ for several values of n . Already for small values of n , the determinant of Ψ_n is extremely small, which makes the problem (4.3) ill-posed from a numerical point of view.

explicit solution of (4.17), which was stated in Theorem 4.1. The insight gained into the structure of Ψ_n by deriving our inversion formula also forms a basis for our analysis of the approximation error. Our first lemma relates the two vectors $\vartheta^{(n)} - \theta_n$ and z_{n+1} , where z_{n+1} was defined in (4.6), and θ_n denotes as before the true Faber-Schauder coefficients of f up to and including generation n .

Lemma 4.13. *For $n \in \mathbb{N}$, we set $P_n := \Psi_{n+1}^{-1} Q_{(n+1, n+1)}$. Then $\vartheta^{(n)} - \theta_n = P_n z_{n+1}$.*

Proof. Remark 4.12 implies that $y_{n+1} = \lim_{m \uparrow \infty} \Psi_{(m, n+1)} \theta_m$. Thus, we have

$$\begin{aligned}
\vartheta^{(n)} &= \Psi_{(n, n+1)}^{-1} y_{n+1} = \Psi_{(n, n+1)}^{-1} \lim_{m \uparrow \infty} \Psi_{(m, n+1)} \theta_m \\
&= \lim_{m \uparrow \infty} \Psi_{(n, n+1)}^{-1} \left[\Psi_{(n, n+1)}, Q_{(n+1, n+1)}, \dots, Q_{(m, n+1)} \right] \cdot \left[\theta_n^\top, \bar{\theta}_{n+1}^\top, \dots, \bar{\theta}_m^\top \right]^\top \\
&= \theta_n + \Psi_{(n, n+1)}^{-1} \sum_{m=n+1}^{\infty} Q_{(m, n+1)} \bar{\theta}_m,
\end{aligned} \tag{4.18}$$

where the third equality follows from (4.16). For $n \in \mathbb{N}$ and $m \geq n+1$, we define the $2^{n+1} \times 2^m$ dimensional matrix

$$U_{(m, n+1)} := 2^{(3n+3-3m)/2} \cdot \begin{bmatrix} 11 \dots 1 & 00 \dots 0 & \dots & 00 \dots 0 \\ 00 \dots 0 & 11 \dots 1 & \dots & 00 \dots 0 \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{00 \dots 0}_{2^{m-n-1} \text{ times}} & \underbrace{00 \dots 0}_{2^{m-n-1} \text{ times}} & \dots & \underbrace{11 \dots 1}_{2^{m-n-1} \text{ times}} \end{bmatrix},$$

and by a simple matrix calculation, one can easily verify that

$$z_{n+1} = \sum_{m=n+1}^{\infty} U_{(m, n+1)} \bar{\theta}_m. \tag{4.19}$$

Moreover, for $m \geq n+1$, we have $\psi_{m,k}(2^{-(n+1)}j) = 2^{-3m/2} \psi_{0,0}(2^{m-n-1}j - k) = 0$ for $k \leq 2^{m-n-1}j$ and otherwise $2^{-3m/2-2}$. Thus, for $m \geq n+1$, the matrix $Q_{(m, n+1)}$ is of the following

form:

$$Q_{(m,n+1)} = 2^{-3m/2-2} \cdot \begin{bmatrix} 11 \cdots 1 & 00 \cdots 0 & \cdots & 00 \cdots 0 \\ 11 \cdots 1 & 11 \cdots 1 & \cdots & 00 \cdots 0 \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{11 \cdots 1}_{2^{m-n-1} \text{ times}} & \underbrace{11 \cdots 1}_{2^{m-n-1} \text{ times}} & \cdots & \underbrace{11 \cdots 1}_{2^{m-n-1} \text{ times}} \end{bmatrix}. \quad (4.20)$$

In particular, the matrix $Q_{(n+1,n+1)}$ is a lower triangular matrix that every non-zero entries are equal to one after a proper scaling. Therefore, we get $Q_{(m,n+1)} = Q_{(n+1,n+1)} \cdot U_{(m,n+1)}$, this is because $Q_{(n+1,n+1)}$ can be regarded as a row operation acting on the matrix $U_{(m,n+1)}$ as a partial sum of rows. Applying this identity and (4.19) to (4.18) gives

$$\boldsymbol{\vartheta}^{(n)} - \boldsymbol{\theta}_n = \Psi_{(n,n+1)}^{-1} \sum_{m=n+1}^{\infty} Q_{(m,n+1)} \bar{\boldsymbol{\theta}}_m = P_n \sum_{m=n+1}^{\infty} U_{(m,n+1)} \bar{\boldsymbol{\theta}}_m = P_n \mathbf{z}_{n+1}. \quad \square$$

Next, for any $-1 \leq m \leq n$, the vectors $\boldsymbol{\vartheta}_m^{(n)}$ and $\bar{\boldsymbol{\vartheta}}_m^{(n)}$ are subvectors of $\boldsymbol{\vartheta}^{(n)}$, and we can represent them via simple linear transformations of $\boldsymbol{\vartheta}^{(n)}$. To this end, we denote for $-1 \leq m \leq n$,

$$R_m^{(n)} := [\mathbf{I}_{2^{m+1} \times 2^{m+1}}, \mathbf{0}_{(2^{n+1}-2^{m+1}) \times 2^{m+1}}] \quad \text{and} \quad \bar{R}_m^{(n)} := [\mathbf{0}_{(2^{n+1}-2^m \vee 0) \times 2^m \vee 0}, \mathbf{I}_{2^m \vee 0 \times 2^m \vee 0}, \mathbf{0}_{(2^{n+1}-2^{m+1}) \times 2^m \vee 0}],$$

where $\mathbf{I}_{m \times m}$ denotes the $m \times m$ dimensional identity matrix, and $\mathbf{0}_{m \times k}$ denotes the $m \times k$ dimensional zero matrix. Thus, $R_m^{(n)}$ is a $2^{m+1} \times 2^{n+1}$ dimensional matrix, and $\bar{R}_m^{(n)}$ is a $2^{m \vee 0} \times 2^{n+1}$ dimensional matrix. Furthermore, one sees by means of a simple matrix manipulation that

$$\boldsymbol{\vartheta}_m^{(n)} - \boldsymbol{\theta}_m = R_m^{(n)} P_n \mathbf{z}_{n+1} \quad \text{and} \quad \bar{\boldsymbol{\vartheta}}_m^{(n)} - \bar{\boldsymbol{\theta}}_m = \bar{R}_m^{(n)} P_n \mathbf{z}_{n+1}.$$

It follows that for $p \in [1, \infty]$,

$$\|\boldsymbol{\vartheta}_m^{(n)} - \boldsymbol{\theta}_m\|_{\ell_p} \leq \|R_m^{(n)} P_n\|_p \|\mathbf{z}_{n+1}\|_{\ell_p}, \quad (4.21)$$

where $\|\cdot\|_p$ denotes the ℓ_p -induced operator norm for a matrix. Moreover, and for any $\varepsilon > 0$, we can choose a vector $\mathbf{z}_{n+1} \in \mathbb{R}^{2^{n+1}}$ such that

$$\|\boldsymbol{\vartheta}_m^{(n)} - \boldsymbol{\theta}_m\|_{\ell_p} \geq (1 - \varepsilon) \|R_m^{(n)} P_n\|_p \|\mathbf{z}_{n+1}\|_{\ell_p}. \quad (4.22)$$

Indeed, for any given vector $\mathbf{z}_{n+1}$, one can find Faber–Schauder coefficients $\theta_{m,k}$ that yield

back $\mathbf{z}_{n+1}$ via (4.6); for instance we can take $\theta_{m,k}$ arbitrary for $m \leq n$, $\theta_{m,k} = 0$ for $m \geq n+2$, and $\theta_{n+1,k} = z_{k+1}^{(n+1)}$ for $k = 0, \dots, 2^{n+1} - 1$. Thus, the Theorems 4.4 through 4.7 will follow immediately from corresponding bounds on the operator norm $\|R_m^{(n)} P_n\|_p$.

For the cases $p = 1$ and $p = \infty$, the following result states exact expressions for this operator norm. For $p = 2$, we have upper and lower bounds that are slightly different from each other.

Theorem 4.14. *For $n \in \mathbb{N}$, $-1 \leq m \leq n$, and $p \in \{1, \infty\}$, $\|\bar{R}_m^{(n)} P_n\|_p = \bar{c}_{p,m}$ and $\|R_m^{(n)} P_n\|_p = c_{p,m}$, where the constants $\bar{c}_{p,m}$ and $c_{p,m}$ are as in Theorem 4.7. For $p = 2$, we have $\|\bar{R}_m^{(n)} P_n\|_2 = \|R_m^{(n)} P_n\|_1 = \frac{1}{2}$ if $-1 \leq m \leq n-1$, and for $m = n$ we have*

$$\sqrt{2 \left(1 - \cos \frac{\pi}{2^{n+1}}\right)^{-1} - \frac{3}{4}} \leq \|\bar{R}_n^{(n)} P_n\|_2 \leq \sqrt{2 \left(1 - \cos \frac{\pi}{2^{n+1}}\right)^{-1} + \frac{3}{4}}.$$

To prove results stated in Theorem 4.4, we first need an explicit representation of the matrix P_n . This is the content of the next lemma, whose proof will, in turn, require another lemma. Before stating it, though, we need to introduce some notation. For $m \geq 0$, we denote the column Rademacher vectors $\mathbf{r}_m^+, \mathbf{r}_m^- \in \mathbb{R}^{2^m}$ by

$$\mathbf{r}_m^+ = \underbrace{(-1, +1, -1, +1, \dots, -1, +1)}_{2^m \text{ times}}^\top \quad \text{and} \quad \mathbf{r}_m^- = \underbrace{(+1, -1, +1, -1, \dots, +1, -1)}_{2^m \text{ times}}^\top. \quad (4.23)$$

For given $n \in \mathbb{N}$ and $0 \leq m \leq n-1$, we define the row vectors

$$\boldsymbol{\eta}_{i,j}^{(m)} = \begin{cases} 2^{\frac{m-n-3}{2}} ((\mathbf{r}_{n-m}^+)^{\top}, (\mathbf{r}_{n-m}^-)^{\top}) & \text{for } 1 \leq i = j \leq 2^m, \\ \mathbf{0}_{1 \times 2^{n+1-m}} & \text{for } 1 \leq i \neq j \leq 2^m. \end{cases}$$

We also define the column vectors $\mathbf{v} = (\frac{3}{2\sqrt{2}}, \frac{-1}{2\sqrt{2}})$ and $\mathbf{u} = (\sqrt{2}, -\sqrt{2})$. Then we define $\mathcal{C}_{-1}, \mathcal{C}_0, \dots, \mathcal{C}_n$ as follows. For $0 \leq m \leq n-1$, we let

$$\mathcal{C}_m := \begin{bmatrix} \boldsymbol{\eta}_{1,1}^{(m)} & \boldsymbol{\eta}_{1,2}^{(m)} & \cdots & \boldsymbol{\eta}_{1,2^m-1}^{(m)} & \boldsymbol{\eta}_{1,2^m}^{(m)} \\ \boldsymbol{\eta}_{2,1}^{(m)} & \boldsymbol{\eta}_{2,2}^{(m)} & \cdots & \boldsymbol{\eta}_{2,2^m-1}^{(m)} & \boldsymbol{\eta}_{2,2^m}^{(m)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{\eta}_{2^m,1}^{(m)} & \boldsymbol{\eta}_{2^m,2}^{(m)} & \cdots & \boldsymbol{\eta}_{2^m,2^m-1}^{(m)} & \boldsymbol{\eta}_{2^m,2^m}^{(m)} \end{bmatrix}. \quad (4.24)$$

For $m = -1$ and $m = n$, we let

$$\mathcal{C}_{-1} = 2^{-\frac{n+3}{2}}(\mathbf{r}_{n+1}^+)^{\top} \quad \text{and} \quad \mathcal{C}_n = \underbrace{\begin{bmatrix} \mathbf{v} & \mathbf{0}_{1 \times 2} & \cdots & \mathbf{0}_{1 \times 2} & \mathbf{0}_{1 \times 2} \\ \mathbf{u} & \mathbf{v} & \cdots & \mathbf{0}_{1 \times 2} & \mathbf{0}_{1 \times 2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{u} & \mathbf{u} & \cdots & \mathbf{v} & \mathbf{0}_{1 \times 2} \\ \mathbf{u} & \mathbf{u} & \cdots & \mathbf{u} & \mathbf{v} \end{bmatrix}}_{2^n \text{ times}},$$

so that $\mathcal{C}_m \in \mathbb{R}^{2^{m \vee 0} \times 2^{n+1}}$ for $m = -1, \dots, n$.

Lemma 4.15. *The matrix P_n is of the following form,*

$$P_n = \left((\mathcal{C}_{-1})^{\top}, (\mathcal{C}_0)^{\top}, \dots, (\mathcal{C}_n)^{\top} \right)^{\top}. \quad (4.25)$$

To prove Lemma 4.15, the following lemma is needed. For $n \in \mathbb{N}$ and $-1 \leq m \leq n-1$, we define the column vector $\mathbf{w}_{(m,n)}^+ \in \mathbb{R}^{2^{n-m}}$ and its reverse vector $\mathbf{w}_{(m,n)}^- \in \mathbb{R}^{2^{n-m}}$,

$$\mathbf{w}_{(m,n)}^+ = \left(\frac{2k-1}{2^{(n-1-m \vee 0)/2}} \right)_{k=1,2,\dots,2^{n-m}} \quad \text{and} \quad \mathbf{w}_{(m,n)}^- = \left(\frac{2k-1}{2^{(n-1-m \vee 0)/2}} \right)_{k=2^{n-m},\dots,2,1}. \quad (4.26)$$

Lemma 4.16. *For given $n \in \mathbb{N}$ and $0 \leq m \leq n$, we define the column vectors*

$$\alpha_{i,j}^{(m)} = \begin{cases} ((\mathbf{w}_{(m,n)}^+)^{\top}, (\mathbf{w}_{(m,n)}^-)^{\top})^{\top} & \text{for } 1 \leq i = j \leq 2^m, \\ \mathbf{0}_{2^{n-m+1} \times 1} & \text{for } 1 \leq i \neq j \leq 2^m. \end{cases}$$

Then we define $\mathcal{A}_{-1}, \mathcal{A}_0, \dots, \mathcal{A}_n$ as follows: For $0 \leq m \leq n$, we let

$$\mathcal{A}_m := \begin{bmatrix} \alpha_{1,1}^{(m)} & \alpha_{1,2}^{(m)} & \cdots & \alpha_{1,2^m-1}^{(m)} & \alpha_{1,2^m}^{(m)} \\ \alpha_{2,1}^{(m)} & \alpha_{2,2}^{(m)} & \cdots & \alpha_{2,2^m-1}^{(m)} & \alpha_{2,2^m}^{(m)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{2^m,1}^{(m)} & \alpha_{2^m,2}^{(m)} & \cdots & \alpha_{2^m,2^m-1}^{(m)} & \alpha_{2^m,2^m}^{(m)} \end{bmatrix} \in \mathbb{R}^{2^{n+1} \times 2^m}. \quad (4.27)$$

For $m = -1$, we set $\mathcal{A}_{-1} = \omega_{(-1,n)}^+$, so that $\mathcal{A}_m \in \mathbb{R}^{2^{m \vee 0} \times 2^{n+1}}$ for $m = -1, \dots, n$. Then the matrix $A_n := Q_{(n+1,n+1)}^{-1} \cdot \Psi_{n+1}$ is well-defined and of the following form

$$A_n = (\mathcal{A}_{-1}, \mathcal{A}_0, \dots, \mathcal{A}_n).$$

Proof. The proof consists of two steps. First, we claim and show that $Q_{(n+1,n+1)}$ is invertible

and

$$Q_{(n+1,n+1)}^{-1} = 2^{\frac{3(n+1)}{2}+2} \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}. \quad (4.28)$$

Taking $m = n + 1$ in (4.20) implies that $Q_{(n+1,n+1)}$ is of the form

$$Q_{(n+1,n+1)} = 2^{-\frac{3(n+1)}{2}-2} \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix},$$

and thus, we only need to show that

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix} = \mathbf{I}_{2^{n+1} \times 2^{n+1}},$$

The above identity indeed holds for the following reason: The first matrix can be regarded as a row operation acting on the subsequent lower triangular matrix by subtracting preceding rows. To be more specific, for $n \in \mathbb{N}$, the n^{th} row of the second matrix consists of ones at its first n positions and zeros at the remaining position. The difference of the n^{th} and the $(n-1)^{\text{th}}$ row then only has a one at the n^{th} position and zeros elsewhere. This then leads to

$$A_n = Q_{(n+1,n+1)}^{-1} \Psi_{(n,n+1)} = \left(Q_{(n+1,n+1)}^{-1} Q_{(-1,n+1)}, Q_{(n+1,n+1)}^{-1} Q_{(0,n+1)} \cdots, Q_{(n+1,n+1)}^{-1} Q_{(n,n+1)} \right).$$

Again, the matrix $Q_{(n+1,n+1)}^{-1}$ acts on each matrix $Q_{(m,n+1)}$ by subtracting preceding rows after proper scaling. For instance, the first column of the matrix $Q_{(n+1,n+1)}^{-1} Q_{(m,n+1)}$ becomes

$$2^{3(n+1)/2+2} \left(\psi_{m,0} \left(\frac{1}{2^{n+1}} \right), \cdots, \psi_{m,0}(1) - \psi_{m,0} \left(\frac{2^{n+1}-1}{2^{n+1}} \right) \right)^{\top}. \quad (4.29)$$

Note that

$$\begin{aligned} \psi_{m,0}\left(\frac{k}{2^{n+1}}\right) - \psi_{m,0}\left(\frac{k-1}{2^{n+1}}\right) &= 2^{-3m/2} \left(\psi_{0,0}\left(\frac{k}{2^{n-m+1}}\right) - \psi_{0,0}\left(\frac{k-1}{2^{n-m+1}}\right) \right) \\ &= \begin{cases} 2^{m/2-2n-3}(2k-1) & \text{for } 1 \leq k \leq 2^{n-m}, \\ 2^{m/2-2n-3}(2^{n-m+2} - (2k-1)) & \text{for } 2^{n-m} + 1 \leq k \leq 2^{n-m+1}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (4.30)$$

Hence, (4.29) and (4.30) demonstrate that the first column fits into the form (4.27). Furthermore, by translation, this assertion carries over to all columns of the matrix $Q_{(n+1,n+1)}^{-1} Q_{(m,n+1)}$ for $m \geq 0$. For the case $m = -1$, we have

$$Q_{(n+1,n+1)}^{-1} Q_{(-1,n+1)} = 2^{3(n+1)/2+1} \left(\left(\frac{1}{2^{n+1}} \right), \left(\frac{2}{2^{n+1}} \right)^2 - \left(\frac{1}{2^{n+1}} \right)^2, \dots, 1 - \left(\frac{2^{n+1}-1}{2^{n+1}} \right)^2 \right)^\top.$$

Since for any $1 \leq k \leq 2^{n+1}$, we have

$$2^{3(n+1)/2+1} \left(\frac{(k-1)^2}{2^{2n+2}} - \frac{k^2}{2^{2n+2}} \right) = 2^{-n/2+1/2}(2k-1),$$

thus, $Q_{(n+1,n+1)}^{-1} Q_{(-1,n+1)} = \omega_{(-1,n)}^+$. This completes the proof. $\square$

Now we can proceed to proving Lemma 4.15.

Proof of Lemma 4.15. For $n \in \mathbb{N}$, let A_n be defined as in (4.27). Since A_n is of full-rank, it suffices to show that $L_n := P_n A_n$ is equal to $\mathbf{I}_{2^{n+1} \times 2^{n+1}}$. We have

$$L_n = \begin{bmatrix} \mathcal{C}_{-1} \mathcal{A}_{-1} & \mathcal{C}_{-1} \mathcal{A}_0 & \cdots & \mathcal{C}_{-1} \mathcal{A}_n \\ \mathcal{C}_0 \mathcal{A}_{-1} & \mathcal{C}_0 \mathcal{A}_0 & \cdots & \mathcal{C}_0 \mathcal{A}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}_n \mathcal{A}_{-1} & \mathcal{C}_n \mathcal{A}_0 & \cdots & \mathcal{C}_n \mathcal{A}_n \end{bmatrix} =: \begin{bmatrix} \mathcal{L}_{(-1,-1)} & \mathcal{L}_{(-1,0)} & \cdots & \mathcal{L}_{(-1,n)} \\ \mathcal{L}_{(0,-1)} & \mathcal{L}_{(0,0)} & \cdots & \mathcal{L}_{(0,n)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{(n,-1)} & \mathcal{L}_{(n,0)} & \cdots & \mathcal{L}_{(n,n)} \end{bmatrix},$$

where $\mathcal{L}_{(m,k)} := \mathcal{C}_m \mathcal{A}_k \in \mathbb{R}^{2^{m \vee 0} \times 2^{k \vee 0}}$ for $-1 \leq m, k \leq n$. Now, we only need to show that $\mathcal{L}_{(m,k)} = \mathbf{0}_{2^{m \vee 0} \times 2^{k \vee 0}}$ and $\mathcal{L}_{(m,m)} = \mathbf{I}_{2^{m \vee 0} \times 2^{m \vee 0}}$ for $-1 \leq m \neq k \leq n$.

First, we show that that $\mathcal{L}_{(m,m)} = \mathbf{I}_{2^m \times 2^m}$ for $0 \leq m \leq n-1$. To this end,

$$\begin{aligned} \mathcal{L}_{(m,m)} &= \begin{bmatrix} \boldsymbol{\eta}_{1,1}^{(m)} & \boldsymbol{\eta}_{1,2}^{(m)} & \cdots & \boldsymbol{\eta}_{1,2^m}^{(m)} \\ \boldsymbol{\eta}_{2,1}^{(m)} & \boldsymbol{\eta}_{2,2}^{(m)} & \cdots & \boldsymbol{\eta}_{2,2^m}^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{2^m,1}^{(m)} & \boldsymbol{\eta}_{2^m,2}^{(m)} & \cdots & \boldsymbol{\eta}_{2^m,2^m}^{(m)} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{1,1}^{(m)} & \boldsymbol{\alpha}_{1,2}^{(m)} & \cdots & \boldsymbol{\alpha}_{1,2^m}^{(m)} \\ \boldsymbol{\alpha}_{2,1}^{(m)} & \boldsymbol{\alpha}_{2,2}^{(m)} & \cdots & \boldsymbol{\alpha}_{2,2^m}^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\alpha}_{2^m,1}^{(m)} & \boldsymbol{\alpha}_{2^m,2}^{(m)} & \cdots & \boldsymbol{\alpha}_{2^m,2^m}^{(m)} \end{bmatrix} \\ &=: \begin{bmatrix} \ell_{1,1}^{(m)} & \ell_{1,2}^{(m)} & \cdots & \ell_{1,2^m}^{(m)} \\ \ell_{2,1}^{(m)} & \ell_{2,2}^{(m)} & \cdots & \ell_{2,2^m}^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{2^m,1}^{(m)} & \ell_{2^m,2}^{(m)} & \cdots & \ell_{2^m,2^m}^{(m)} \end{bmatrix}, \end{aligned}$$

where $\ell_{i,j}^{(m)} = \sum_{k=1}^{2^m} \boldsymbol{\eta}_{i,k}^{(m)} \boldsymbol{\alpha}_{k,j}^{(m)}$. Since $\boldsymbol{\eta}_{i,j}^{(m)} = \mathbf{0}_{1 \times 2^{n+1-m}}$ and $\boldsymbol{\alpha}_{i,j}^{(m)} = \mathbf{0}_{2^{n+1-m} \times 1}$ for $i \neq j$, therefore $\ell_{i,j}^{(m)} = 0$ for $i \neq j$. Moreover, we get in the same way that

$$\begin{aligned} \ell_{i,i}^{(m)} &= \boldsymbol{\eta}_{i,i}^{(m)} \boldsymbol{\alpha}_{i,i}^{(m)} = 2^{(m-n-3)/2} \left((\mathbf{r}_{n-m}^+)^{\top} \mathbf{w}_{(m,n)}^+ + (\mathbf{r}_{n-m}^-)^{\top} \mathbf{w}_{(m,n)}^- \right) \\ &= 2^{m-n-1} \left(\sum_{k=1}^{2^{n-m}} (-1)^k (2k-1) + \sum_{k=2^{n-m}+1}^{2^{n-m+1}} (-1)^{k+1} (2^{n-m+2} - (2k-1)) \right) \\ &= 2^{m-n} \sum_{k=1}^{2^{n-m}} (-1)^k (2k-1) = 2^{m-n} \sum_{k=1}^{2^{n-m}-1} 2 = 1. \end{aligned}$$

Next, for the case $m = -1$, we prove analogously that

$$\mathcal{C}_{-1} \mathcal{A}_{-1} = 2^{-\frac{n+3}{2}} (\mathbf{r}_{n+1}^+)^{\top} \mathbf{w}_{(-1,n)}^+ = 2^{-n-1} \left(\sum_{k=1}^{2^{n+1}} (-1)^k (2k-1) \right) = 1.$$

Last, it follows that

$$\mathcal{C}_n \mathcal{A}_n = 2^{1/2} \begin{bmatrix} \frac{3}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & 0 & \cdots & 0 & 0 \\ \sqrt{2} & -\sqrt{2} & \frac{3}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sqrt{2} & -\sqrt{2} & \sqrt{2} & -\sqrt{2} & \cdots & \frac{3}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}_{2^n \times 2^n}.$$

The above identity holds because the second matrix can be regarded as a column operation

acting on the first matrix by adding subsequent columns. To be more specific, for $1 \leq m \leq 2^n$, the m^{th} column of the matrix $\mathcal{L}_{(n,n)}$ equals the sum of the $(2m-1)^{\text{th}}$ and the $(2m)^{\text{th}}$ columns of $\mathcal{C}_n$. Prior to the m^{th} entry, each entry of the $(2m-1)^{\text{th}}$ and the $(2m)^{\text{th}}$ column of $\mathcal{C}_n$ equals zero and all entries are equal to $\pm\sqrt{2}$ afterwards respectively. Thus, it is clear that $\mathcal{C}_n \mathcal{A}_n = \mathbf{I}_{2^n \times 2^n}$.

Next, we are going to show that $\mathcal{L}_{(m,k)} = \mathbf{0}_{2^m \times 2^k}$ for $0 \leq m \neq k \leq n-1$, and we first show this holds with $m < k$. Under this assumption, we have

$$\begin{aligned} \mathcal{C}_m \mathcal{A}_k &= \begin{bmatrix} \boldsymbol{\eta}_{1,1}^{(m)} & \mathbf{0}_{1 \times 2^{n+1-m}} & \cdots & \mathbf{0}_{1 \times 2^{n+1-m}} \\ \mathbf{0}_{1 \times 2^{n+1-m}} & \boldsymbol{\eta}_{2,2}^{(m)} & \cdots & \mathbf{0}_{1 \times 2^{n+1-m}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1 \times 2^{n+1-m}} & \mathbf{0}_{1 \times 2^{n+1-m}} & \cdots & \boldsymbol{\eta}_{2^m, 2^m}^{(m)} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{1,1}^{(k)} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \mathbf{0}_{2^{n+1-k} \times 1} & \boldsymbol{\alpha}_{2,2}^{(k)} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2^{n+1-k} \times 1} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & \boldsymbol{\alpha}_{2^k, 2^k}^{(k)} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\eta}_{1,1}^{(m)} \boldsymbol{\alpha}_{1,1}^{(m,k)} & \mathbf{0}_{1 \times 2^{k-m}} & \cdots & \mathbf{0}_{1 \times 2^{k-m}} \\ \mathbf{0}_{1 \times 2^{k-m}} & \boldsymbol{\eta}_{2,2}^{(m)} \boldsymbol{\alpha}_{2,2}^{(m,k)} & \cdots & \mathbf{0}_{1 \times 2^{k-m}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1 \times 2^{k-m}} & \mathbf{0}_{1 \times 2^{k-m}} & \cdots & \boldsymbol{\eta}_{2^m, 2^m}^{(m)} \boldsymbol{\alpha}_{2^m, 2^m}^{(m,k)} \end{bmatrix}, \end{aligned}$$

where the $2^{n+m-1} \times 2^{k-m}$ dimensional matrix $\boldsymbol{\alpha}_{i,i}^{(m,k)}$ is defined as

$$\boldsymbol{\alpha}_{i,i}^{(m,k)} := \begin{bmatrix} \boldsymbol{\alpha}_{1,1}^{(k)} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \mathbf{0}_{2^{n+1-k} \times 1} & \boldsymbol{\alpha}_{2,2}^{(k)} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2^{n+1-k} \times 1} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & \boldsymbol{\alpha}_{2^{k-m}, 2^{k-m}}^{(k)} \end{bmatrix},$$

for any $1 \leq i \leq 2^m$. Further calculation yields that for each $1 \leq i \leq 2^m$,

$$\begin{aligned}
\boldsymbol{\eta}_{i,i}^{(m)} \boldsymbol{\alpha}_{i,i}^{(m,k)} &= 2^{(m-n-3)/2} \left((\mathbf{r}_{n-m}^+)^{\top}, (\mathbf{r}_{n-m}^-)^{\top} \right) \begin{bmatrix} \boldsymbol{\alpha}_{1,1}^{(k)} & \mathbf{0}_{2^{n-k+1} \times 1} & \cdots & \mathbf{0}_{2^{n-k+1} \times 1} \\ \mathbf{0}_{2^{n-k+1} \times 1} & \boldsymbol{\alpha}_{2,2}^{(k)} & \cdots & \mathbf{0}_{2^{n-k+1} \times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2^{n-k+1} \times 1} & \mathbf{0}_{2^{n-k+1} \times 1} & \cdots & \boldsymbol{\alpha}_{2^{k-m}, 2^{k-m}}^{(k)} \end{bmatrix} \\
&= 2^{(m-n-3)/2} \left(\underbrace{(\mathbf{r}_{n-k}^+)^{\top}, \dots, (\mathbf{r}_{n-k}^+)^{\top}}_{2^{k-m} \text{ times}}, \underbrace{(\mathbf{r}_{n-k}^-)^{\top}, \dots, (\mathbf{r}_{n-k}^-)^{\top}}_{2^{k-m} \text{ times}} \right) \begin{bmatrix} \mathbf{w}_{(k,n)}^+ & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \mathbf{w}_{(k,n)}^- & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{w}_{(k,n)}^+ & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{w}_{(k,n)}^- & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{w}_{(k,n)}^+ \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{w}_{(k,n)}^- \end{bmatrix} \\
&= 2^{(m-n-3)/2} \left(\underbrace{(\mathbf{r}_{n-k}^+)^{\top} (\mathbf{w}_{(k,n)}^+ + \mathbf{w}_{(k,n)}^-), \dots}_{2^{k-m-1} \text{ times}}, \underbrace{(\mathbf{r}_{n-k}^-)^{\top} (\mathbf{w}_{(k,n)}^+ + \mathbf{w}_{(k,n)}^-), \dots}_{2^{k-m-1} \text{ times}} \right).
\end{aligned}$$

Since $\mathbf{w}_{(k,n)}^+ + \mathbf{w}_{(k,n)}^- = 2^{\frac{m-n+1}{2}} ((2^{n-m+1} - 1) + 1) \cdot \mathbf{1}_{2^{n-k} \times 1} = 2^{\frac{n-m+3}{2}} \cdot \mathbf{1}_{2^{n-k} \times 1}$, it follows that

$$(\mathbf{r}_{n-k}^+)^{\top} (\mathbf{w}_{(k,n)}^+ + \mathbf{w}_{(k,n)}^-) = (\mathbf{r}_{n-k}^-)^{\top} (\mathbf{w}_{(k,n)}^+ + \mathbf{w}_{(k,n)}^-) = 0. \quad (4.31)$$

Thus, we have $\boldsymbol{\eta}_{i,i}^{(m)} \boldsymbol{\alpha}_{i,i}^{(m,k)} = \mathbf{0}_{1 \times 2^{k-m}}$, and hence, $\mathcal{L}_{(m,k)} = \mathbf{0}_{2^m \times 2^k}$. Since the matrices A_n and P_n are symmetric by construction, we have $\mathcal{L}_{(m,k)} = \mathcal{L}_{(k,m)}^{\top} = \mathbf{0}_{2^k \times 2^m}$ for $m > k$. This now proves our claim $\mathcal{L}_{(m,k)} = \mathbf{0}_{2^m \times 2^k}$ for $0 \leq m \neq k \leq n-1$.

For the case $m = -1$, we get in the same manner that for $-1 \leq k \leq n-1$,

$$\mathcal{C}_{-1} \mathcal{A}_k = 2^{-\frac{n+3}{2}} \left(\underbrace{(\mathbf{r}_{n-k}^+)^{\top} (\mathbf{w}_{(k,n)}^+ + \mathbf{w}_{(k,n)}^-), \dots, (\mathbf{r}_{n-k}^-)^{\top} (\mathbf{w}_{(k,n)}^+ + \mathbf{w}_{(k,n)}^-)}_{2^k \text{ times}} \right) = \mathbf{0}_{1 \times 2^k}.$$

Finally, for $-1 \leq m \leq n-1$, recall that we denote the (i, j) -entry of $\mathcal{L}_{(n,m)}$ by $\ell_{i,j}^{(n,m)}$. A simple matrix calculation implies that

$$\ell_{i,j}^{(n,m)} = \left(\underbrace{\mathbf{u}, \dots, \mathbf{u}}_{i \text{ times}}, \underbrace{\mathbf{v}, \mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{2^{n-i-1} \text{ times}}, \underbrace{\mathbf{0}_{1 \times 2^{n-m}}, \dots}_{2j \text{ times}}, \underbrace{(\mathbf{w}_{(m,n)}^+)^{\top}, (\mathbf{w}_{(m,n)}^-)^{\top}, \dots, \mathbf{0}_{1 \times 2^{n-m}}, \dots}_{2^{m+1}-2j-2 \text{ times}} \right)^{\top}.$$

The fact that

$$\underbrace{(\mathbf{u}, \dots, \mathbf{u})}_{2^{n-m-1} \text{ times}} \mathbf{w}_{(m,n)}^{\pm} = \sqrt{2} \cdot (\mathbf{r}_{n-m}^+)^{\top} \mathbf{w}_{(m,n)}^{\pm} = 0.$$

directly yields that $\ell_{i,j}^{(n,m)} = 0$ for $i \geq 2^{n-m}(j+1)$. Otherwise, for $I \leq 2^{n-m}j - 1$, we clearly have $\ell_{i,j}^{(n,m)} = 0$, as every non-zero vector must be multiplied by a zero vector. For $2^{n-m}j \leq i \leq 2^{n-m-1}(2j+1) - 1$, we denote $\tau = i - 2^{n-m}j$. Then a matrix calculation gives

$$\begin{aligned}\ell_{i,j}^{(n,m)} &= \frac{\sqrt{2}}{2^{(n+3-m)/2}} \left(\sum_{k=1}^{\tau} (-1)^k (2k-1) + \frac{3}{4}(2\tau+1) - \frac{1}{4}(2\tau+3) \right) \\ &= \frac{\sqrt{2}}{2^{2n-m/2-1}} \left(\tau - \frac{3}{4}(2\tau+1) - \frac{1}{4}(2\tau+3) \right) = 0.\end{aligned}$$

The case $2^{n-m-1}(2j+1) \leq i \leq 2^{n-m}(j+1) - 1$ can be proved analogously using the symmetry property of $\mathbf{w}_{(m,n)}^{\pm}$ for $m \geq 0$. For the case $m = -1$, one can easily apply the same method to obtain $\mathcal{L}_{(n,-1)} = \mathbf{0}_{2^{n+1} \times 1}$. This completes the proof. $\square$

Next, we shall apply Lemma 4.15 to obtain the matrix norms of $\bar{R}_m^{(n)} P_n$ and $R_m^{(n)} P_n$ as stated in Theorem 4.14. We give separate proofs for the cases $p = \infty$, $p = 1$, and $p = 2$. Recall that for an $m \times n$ matrix $A = (a_{ij})$, we have $\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|$, which is simply the maximum absolute column sum of the matrix, and $\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|$ is the maximum absolute row sum of the matrix. For the special case $p = 2$, $\|A\|_2 = \sigma_{\max}(A)$ is the largest singular value of the matrix.

Proof of Theorem 4.14 for $p = \infty$. We start with computing the norm $\left\| \bar{R}_m^{(n)} P_n \right\|_{\infty}$ for $-1 \leq m \leq n-1$. Note that $\bar{R}_m^{(n)} P_n = \mathcal{C}_m$, and since $\boldsymbol{\eta}_{i,j}^{(m)}$ is a zero vector for $i \neq j$, then

$$\begin{aligned}\|\mathcal{C}_m\|_{\infty} &= \max_{1 \leq i \leq 2^m} \left\| (\boldsymbol{\eta}_{i,1}^{(m)}, \boldsymbol{\eta}_{i,2}^{(m)}, \dots, \boldsymbol{\eta}_{i,2^{m-1}}^{(m)}, \boldsymbol{\eta}_{i,2^m}^{(m)}) \right\|_{\infty} \\ &= \max_{1 \leq i \leq 2^m} \left\| (\boldsymbol{\eta}_{i,1}^{(m)}, \boldsymbol{\eta}_{i,2}^{(m)}, \dots, \boldsymbol{\eta}_{i,2^{m-1}}^{(m)}, \boldsymbol{\eta}_{i,2^m}^{(m)})^{\top} \right\|_{\ell_1} = \|(\boldsymbol{\eta}_{i,i}^{(m)})^{\top}\|_{\ell_1} \\ &= \begin{cases} 2^{\frac{m-n-3}{2}} \|(\mathbf{r}_{n-m}^+, \mathbf{r}_{n-m}^-)\|_{\ell_1} & \text{for } m \geq 0 \\ 2^{\frac{-n-3}{2}} \|\mathbf{r}_{n+1}^+\|_{\ell_1} & \text{for } m = -1 \end{cases} \\ &= 2^{\frac{m \vee 0 - n - 3}{2}} \cdot 2^{n+1-m \vee 0} = 2^{\frac{n-1-m \vee 0}{2}}.\end{aligned}$$

Furthermore, for $-1 \leq m \leq n-1$, we get

$$\left\| R_m^{(n)} P_n \right\|_{\infty} = \|(\mathcal{C}_{-1}^{\top}, \mathcal{C}_0^{\top}, \dots, \mathcal{C}_m^{\top})^{\top}\|_{\infty} = \max_{-1 \leq i \leq m} \|\mathcal{C}_i\|_{\infty} = 2^{\frac{n-1}{2}}.$$

Last, for $m = n$, we have

$$\begin{aligned}\left\|\bar{R}_n^{(n)}P_n\right\|_\infty &= (2^n - 1)\|\mathbf{u}\|_\infty + \|\mathbf{v}\|_\infty = (2^n - 1)\|\mathbf{u}\|_{\ell_1} + \|\mathbf{v}\|_{\ell_1} \\ &= (2^n - 1) \cdot 2\sqrt{2} + \frac{1}{2\sqrt{2}} + \frac{3}{2\sqrt{2}} = 2^{n+\frac{3}{2}} - \sqrt{2}.\end{aligned}$$

This concludes the proof of Theorem 4.14 for $p = \infty$. $\square$

Proof of Theorem 4.14 for $p = 1$. By utilizing the identity $\bar{R}_m^{(n)}P_n = \mathcal{C}_m$, we get

$$\|\mathcal{C}_m\|_1 = \max_{1 \leq i \leq m} \left\| \left((\boldsymbol{\eta}_{1,i}^{(m)})^\top, (\boldsymbol{\eta}_{2,i}^{(m)})^\top, \dots, (\boldsymbol{\eta}_{2^m-1,i}^{(m)})^\top, (\boldsymbol{\eta}_{2^m,i}^{(m)})^\top \right)^\top \right\|_1 = \max_{1 \leq i \leq m} \left\| \boldsymbol{\eta}_{i,i}^{(m)} \right\|_1.$$

Since, for every $1 \leq i \leq 2^m$, the entries of the vector $\boldsymbol{\eta}_{i,i}^{(m)}$ are either $2^{(m \vee 0 - n - 3)/2}$ or $-2^{(m \vee 0 - n - 3)/2}$, the vector $\boldsymbol{\eta}_{i,i}^{(m)}$, regarded as a linear functional, has the ℓ_1 -induced norm $\left\| \boldsymbol{\eta}_{i,i}^{(m)} \right\|_1 = 2^{(m \vee 0 - n - 3)/2}$. Next, we have

$$\begin{aligned}\left\|R_m^{(n)}P_n\right\|_1 &= \left\| \left(\mathcal{C}_{-1}^\top, \mathcal{C}_0^\top, \dots, \mathcal{C}_m^\top \right)^\top \right\|_1 = \sum_{i=-1}^m \|\mathcal{C}_i\|_1 \\ &= \sum_{i=-1}^m 2^{(i \vee 0 - n - 3)/2} = \begin{cases} 2^{-\frac{n+3}{2}} & \text{for } m = -1, \\ 2^{-\frac{n+3}{2}} \left(\frac{2^{\frac{m+1}{2}} - 1}{\sqrt{2} - 1} + 1 \right) & \text{for } 0 \leq m \leq n - 1, \end{cases} \\ &= 2^{-\frac{n+3}{2}} \left(1 + \frac{2^{\frac{m+1}{2}} - 1}{\sqrt{2} - 1} \mathbb{1}_{\{m \geq 0\}} \right).\end{aligned}$$

Last, for the case $m = n$, it follows that

$$\left\|\bar{R}_n^{(n)}P_n\right\|_1 = (2^n - 1)\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 = (2^n - 1)\sqrt{2} + \frac{3}{2\sqrt{2}} = 2^{n+\frac{1}{2}} - \frac{\sqrt{2}}{4}.$$

This completes the proof. $\square$

The proof of Theorem 4.14 for $p = 2$ is more complicated than the ones for $p = 1$ and $p = \infty$ and requires additional preparation. We start with a lemma concerning the exact form of product $\mathcal{C}_m \mathcal{C}_k^\top$ for $-1 \leq m, k \leq n - 1$.

Lemma 4.17. *For $-1 \leq m \leq n$, let $\mathcal{C}_m$ be as in the context of Lemma 4.15. Then,*

$$\mathcal{C}_m \mathcal{C}_k^\top = \begin{cases} \frac{1}{4} \mathbf{I}_{2^{m \vee 0} \times 2^{m \vee 0}} & \text{for } 1 \leq m = k \leq n - 1, \\ \mathbf{0}_{2^{m \vee 0} \times 2^{k \vee 0}} & \text{for } 1 \leq m \neq k \leq n - 1. \end{cases}$$

Proof. Let us denote $\Phi_{(m,k)} = (\phi_{i,j}^{(m,k)}) := \mathcal{C}_m \mathcal{C}_k^\top$ and prove the case for $m = k$. Since

$$\Phi_{(m,m)} = \begin{bmatrix} \boldsymbol{\eta}_{1,1}^{(m)} & \boldsymbol{\eta}_{1,2}^{(m)} & \cdots & \boldsymbol{\eta}_{1,2^m}^{(m)} \\ \boldsymbol{\eta}_{2,1}^{(m)} & \boldsymbol{\eta}_{2,2}^{(m)} & \cdots & \boldsymbol{\eta}_{2,2^m}^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\eta}_{2^m,1}^{(m)} & \boldsymbol{\eta}_{2^m,2}^{(m)} & \cdots & \boldsymbol{\eta}_{2^m,2^m}^{(m)} \end{bmatrix} \begin{bmatrix} (\boldsymbol{\eta}_{1,1}^{(m)})^\top & (\boldsymbol{\eta}_{2,1}^{(m)})^\top & \cdots & (\boldsymbol{\eta}_{2^m,1}^{(m)})^\top \\ (\boldsymbol{\eta}_{1,2}^{(m)})^\top & (\boldsymbol{\eta}_{2,2}^{(m)})^\top & \cdots & (\boldsymbol{\eta}_{2^m,2}^{(m)})^\top \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{\eta}_{1,2^m}^{(m)})^\top & (\boldsymbol{\eta}_{2,2^m}^{(m)})^\top & \cdots & (\boldsymbol{\eta}_{2^m,2^m}^{(m)})^\top \end{bmatrix},$$

then $\phi_{i,j}^{(m,m)} = \sum_{k=1}^{2^m} \boldsymbol{\eta}_{i,k}^{(m)} (\boldsymbol{\eta}_{j,k}^{(m)})^\top$. As $\boldsymbol{\eta}_{i,k}^{(m)} = \mathbf{0}_{1 \times 2^{n-m+1}}$ for $i \neq k$, then $\phi_{i,j}^{(m,m)} = 0$ for all $i \neq j$. Furthermore, it again follows from Lemma 4.15 that for $1 \leq i \leq 2^m$,

$$\phi_{i,i}^{(m,m)} = \boldsymbol{\eta}_{i,i}^{(m)} (\boldsymbol{\eta}_{i,i}^{(m)})^\top = 2^{m \vee 0 - n - 3} \left(\|\mathbf{r}_{n-m}^+\|_{\ell_2}^2 + \|\mathbf{r}_{n-m}^-\|_{\ell_2}^2 \right) = 2^{m \vee 0 - n - 3} \cdot 2^{n+1-m \vee 0} = \frac{1}{4}.$$

This shows that $\Phi_{(m,m)} = \frac{1}{4} \mathbf{I}_{2^{m \vee 0} \times 2^{m \vee 0}}$. Next, for $0 \leq m < k \leq n-1$, we have

$$\begin{aligned} \mathcal{C}_m \mathcal{C}_k^\top &= \begin{bmatrix} \boldsymbol{\eta}_{1,1}^{(m)} & \mathbf{0}_{1 \times 2^{n+1-m}} & \cdots & \mathbf{0}_{1 \times 2^{n+1-m}} \\ \mathbf{0}_{1 \times 2^{n+1-m}} & \boldsymbol{\eta}_{2,2}^{(m)} & \cdots & \mathbf{0}_{1 \times 2^{n+1-m}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1 \times 2^{n+1-m}} & \mathbf{0}_{1 \times 2^{n+1-m}} & \cdots & \boldsymbol{\eta}_{2^m,2^m}^{(m)} \end{bmatrix} \begin{bmatrix} (\boldsymbol{\eta}_{1,1}^{(k)})^\top & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \mathbf{0}_{2^{n+1-k} \times 1} & (\boldsymbol{\eta}_{2,2}^{(k)})^\top & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2^{n+1-k} \times 1} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & (\boldsymbol{\eta}_{2^k,2^k}^{(k)})^\top \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\eta}_{1,1}^{(m)} \boldsymbol{\eta}_{1,1}^{(m,k)} & \mathbf{0}_{1 \times 2^{k-m}} & \cdots & \mathbf{0}_{1 \times 2^{k-m}} \\ \mathbf{0}_{1 \times 2^{k-m}} & \boldsymbol{\eta}_{2,2}^{(m)} \boldsymbol{\eta}_{2,2}^{(m,k)} & \cdots & \mathbf{0}_{1 \times 2^{k-m}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1 \times 2^{k-m}} & \mathbf{0}_{1 \times 2^{k-m}} & \cdots & \boldsymbol{\eta}_{2^m,2^m}^{(m)} \boldsymbol{\eta}_{2^m,2^m}^{(m,k)} \end{bmatrix}, \end{aligned}$$

where

$$\boldsymbol{\eta}_{i,i}^{(m,k)} := \begin{bmatrix} (\boldsymbol{\eta}_{1,1}^{(k)})^\top & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \mathbf{0}_{2^{n+1-k} \times 1} & (\boldsymbol{\eta}_{2,2}^{(k)})^\top & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2^{n+1-k} \times 1} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & (\boldsymbol{\eta}_{2^k,2^k}^{(k)})^\top \end{bmatrix},$$

for any $1 \leq i \leq 2^m$. Note that for $0 \leq m \leq n-1$ and $1 \leq i \leq 2^m$,

$$\begin{aligned}
\boldsymbol{\eta}_{i,i}^{(m)} \boldsymbol{\eta}_{i,i}^{(m,k)} &= 2^{(m-n-3)/2} \left((\mathbf{r}_{n-m}^+)^{\top}, (\mathbf{r}_{n-m}^-)^{\top} \right) \begin{bmatrix} (\boldsymbol{\eta}_{1,1}^{(k)})^{\top} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \mathbf{0}_{2^{n+1-k} \times 1} & (\boldsymbol{\eta}_{2,2}^{(k)})^{\top} & \cdots & \mathbf{0}_{2^{n+1-k} \times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2^{n+1-k} \times 1} & \mathbf{0}_{2^{n+1-k} \times 1} & \cdots & (\boldsymbol{\eta}_{2^{k-m}, 2^{k-m}}^{(k)})^{\top} \end{bmatrix} \\
&= 2^{m-n-3} \left(\underbrace{(\mathbf{r}_{n-k}^+)^{\top}, \dots, (\mathbf{r}_{n-k}^+)^{\top}}_{2^{k-m} \text{ times}}, \underbrace{(\mathbf{r}_{n-k}^-)^{\top}, \dots, (\mathbf{r}_{n-k}^-)^{\top}}_{2^{k-m} \text{ times}} \right) \begin{bmatrix} \mathbf{r}_{n-k}^+ & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \mathbf{r}_{n-k}^- & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{r}_{n-k}^+ & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{r}_{n-k}^- & \cdots & \mathbf{0}_{2^{n-k} \times 1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{r}_{n-k}^+ \\ \mathbf{0}_{2^{n-k} \times 1} & \mathbf{0}_{2^{n-k} \times 1} & \cdots & \mathbf{r}_{n-k}^- \end{bmatrix} \\
&= 2^{m-n-3} \left(\underbrace{(\mathbf{r}_{n-k}^+)^{\top} (\mathbf{r}_{n-k}^+ + \mathbf{r}_{n-k}^-), \dots}_{2^{k-m-1} \text{ times}}, \underbrace{(\mathbf{r}_{n-k}^-)^{\top} (\mathbf{r}_{n-k}^+ + \mathbf{r}_{n-k}^-), \dots}_{2^{k-m-1} \text{ times}} \right) = \mathbf{0}_{1 \times 2^{k-m}}.
\end{aligned}$$

Hence, $\mathcal{C}_m \mathcal{C}_k^{\top} = \mathbf{0}_{2^m \times 2^k}$ for $0 \leq m < k \leq n-1$, and for $0 \leq k < m \leq n-1$, one has $\mathcal{C}_m \mathcal{C}_k^{\top} = (\mathcal{C}_k \mathcal{C}_m^{\top})^{\top} = \mathbf{0}_{2^k \times 2^m}^{\top} = \mathbf{0}_{2^m \times 2^k}$. Therefore, for $0 \leq m \neq k \leq n-1$, we have $\mathcal{C}_m \mathcal{C}_k^{\top} = \mathbf{0}_{2^m \times 2^k}$. Lastly, for the special case of $m = -1$ and $0 \leq k \leq n-1$, we have

$$\mathcal{C}_{-1} \mathcal{C}_k = 2^{-n-3+k/2} \left(\underbrace{(\mathbf{r}_{n-k}^+)^{\top} (\mathbf{r}_{n-k}^+ + \mathbf{r}_{n-k}^-), \dots, (\mathbf{r}_{n-k}^+)^{\top} (\mathbf{r}_{n-k}^+ + \mathbf{r}_{n-k}^-)}_{2^k \text{ times}} \right) = \mathbf{0}_{1 \times 2^k},$$

and by a similar argument as above, we get $\mathcal{C}_k \mathcal{C}_{-1} = \mathbf{0}_{2^k \times 1}$. This completes the proof. $\square$

Let us next consider the following $n \times n$ dimensional matrix D_n , where the $(i, j)^{\text{th}}$ -entry $d_{i,j} = i \vee j - 1/2$ for $1 \leq i, j \leq n$. The next lemma computes the ℓ_2 -induced norm of D_n .

Lemma 4.18. *For $n \geq 3$, we have $\|D_n\|_2 = \frac{1}{2}(1 - \cos \frac{\pi}{2n})^{-1}$.*

Proof. In the first step, we first derive the inverse of D_n and then obtain $\|D_n\|_2$. To this end, let us consider the $n \times n$ dimensional tridiagonal matrix E_n which is defined as

$$E_n := \begin{bmatrix} 3 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \ddots \\ 0 & -1 & 2 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & -1 \\ 0 & \ddots & 0 & -1 & 1 \end{bmatrix},$$

and we are about to demonstrate $E_n = D_n^{-1}$ for $n \geq 3$. Since D_n and E_n are symmetric, it suffices to show that the product matrix $F_n := E_n D_n = \mathbf{I}_{n \times n}$. Denote the i^{th} row of the matrix D_n by $\mathbf{d}_i$ for $1 \leq i \leq n$, and $D_n = (\mathbf{d}_1^\top, \mathbf{d}_2^\top, \dots, \mathbf{d}_n^\top)^\top$. Writing the i^{th} row of the matrix F_n by $\mathbf{f}_i$ and regarding the matrix as a row operation acting on the matrix D_n leads to

$$F_n = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_{n-1} \\ \mathbf{f}_n \end{pmatrix} = E_n \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_{n-1} \\ \mathbf{d}_n \end{pmatrix} = \begin{pmatrix} 3\mathbf{d}_1 - \mathbf{d}_2 \\ 2\mathbf{d}_2 - \mathbf{d}_1 - \mathbf{d}_3 \\ \vdots \\ 2\mathbf{d}_{n-1} - \mathbf{d}_{n-2} - \mathbf{d}_n \\ \mathbf{d}_n - \mathbf{d}_{n-1} \end{pmatrix}$$

It remains to verify the form of $\mathbf{f}_i$ term by term. First, we have

$$\mathbf{f}_1 = 3\mathbf{d}_1 - \mathbf{d}_2 = 3 \cdot \left(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2} \right) - \left(\frac{1}{2}, \frac{3}{2}, \dots, \frac{3}{2} \right) = (1, 0, \dots, 0).$$

Next, for $2 \leq i \leq n-1$ and $1 \leq j \leq n$, we have

$$t_{i,j} = 2(i \vee j) - (i-1) \vee j - (i+1) \vee j = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, for $2 \leq i \leq n-1$, we have

$$\mathbf{f}_i = 2\mathbf{d}_i - \mathbf{d}_{i-1} - \mathbf{d}_{i+1} = \underbrace{(0, \dots, 0)}_{i-1 \text{ times}}, 1, \underbrace{(0, \dots, 0)}_{n-1 \text{ times}}.$$

Last, we have

$$\begin{aligned} \mathbf{f}_n &= \mathbf{d}_n - \mathbf{d}_{n-1} = \left(\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots, \frac{2n-3}{2}, \frac{2n-1}{2} \right) - \left(\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots, \frac{2n-3}{2}, \frac{2n-3}{2} \right) \\ &= (0, 0, \dots, 0, 1). \end{aligned}$$

Hence, the above calculation establishes the result that $E_n = D_n^{-1}$.

For the second step of the proof, it follows from [69, Theorem 4] that the determinant of $E_n - \lambda \mathbf{I}_{n \times n}$ admits the following representation:

$$\det(E_n - \lambda \mathbf{I}_{n \times n}) = \prod_{j=0}^{n-1} \left(2 - \lambda - 2 \cos \frac{(j+1)\pi}{2n} \right),$$

and this leads to $\lambda_{\min}(E_n) = 2 - 2 \cos \frac{\pi}{2n}$. Thus, we have

$$\|D_n\|_2 = \lambda_{\max}(D_n) = \frac{1}{\lambda_{\min}(E_n)} = \frac{1}{2} \left(1 - \cos \frac{\pi}{2n}\right)^{-1},$$

and this completes the proof. $\square$

We are now ready to complete the proof of Theorem 4.14.

Proof of Theorem 4.14 for $p = 2$. Let us denote the largest singular value of a matrix M by $\sigma_{\max}(M)$. By the definition of the spectral norm, we get

$$\begin{aligned} \left\| \bar{R}_m^{(n)} P_n \right\|_2 &= \sqrt{\sigma_{\max}(\bar{R}_m^{(n)} P_n P_n^\top (\bar{R}_m^{(n)})^\top)} \\ &= \sqrt{\sigma_{\max}(\mathcal{C}_m \mathcal{C}_m^\top)} = \frac{1}{2} \sqrt{\sigma_{\max}(\mathbf{I}_{2^m \times 2^m})} = \frac{1}{2}, \end{aligned} \quad (4.32)$$

where the second last identity follows from Lemma 4.17. To derive the ℓ_2 -induced norm of matrices $R_m^{(n)} P_n$, note that

$$R_m^{(n)} P_n P_n^\top (R_m^{(n)})^\top = \begin{bmatrix} \mathcal{C}_{-1} \mathcal{C}_{-1}^\top & \mathcal{C}_{-1} \mathcal{C}_0^\top & \cdots & \mathcal{C}_{-1} \mathcal{C}_m^\top \\ \mathcal{C}_0 \mathcal{C}_{-1}^\top & \mathcal{C}_0 \mathcal{C}_0^\top & \cdots & \mathcal{C}_0 \mathcal{C}_m^\top \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}_m \mathcal{C}_{-1}^\top & \mathcal{C}_m \mathcal{C}_0^\top & \cdots & \mathcal{C}_m \mathcal{C}_m^\top \end{bmatrix} = \frac{\mathbf{I}_{2^{m+1} \times 2^{m+1}}}{4}.$$

An analogous argument as in (4.32) yields that $\left\| \bar{R}_m^{(n)} P_n \right\|_2 = \frac{1}{2}$. Lastly, it remains to obtain the spectral norm of $\bar{R}_n^{(n)} P_n$, or equivalently, the largest singular value of $\mathcal{C}_n \mathcal{C}_n^\top$. To this end, let us first derive the exact form of the matrix $G_n = (g_{i,j}^{(n)}) = \mathcal{C}_n \mathcal{C}_n^\top$. Next, we shall claim and show that

$$g_{i,j}^{(n)} = \begin{cases} 4i - \frac{11}{4} & \text{for } 1 \leq i = j \leq 2^n, \\ 4(i \wedge j) - 2 & \text{for } 1 \leq i \neq j \leq 2^n. \end{cases} \quad (4.33)$$

First, one has

$$\begin{aligned} g_{i,i}^{(n)} &= \left\| \left(\underbrace{\mathbf{u}, \dots, \mathbf{u}}_{i-1 \text{ times}}, \mathbf{v}, \underbrace{\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{2^n - i \text{ times}} \right) \right\|_{\ell_2}^2 = 4(i-1) + \left(\left(\frac{3}{2\sqrt{2}} \right)^2 + \left(\frac{1}{2\sqrt{2}} \right)^2 \right) \\ &= \frac{5}{4} + 4(i-1) = 4i - \frac{11}{4}. \end{aligned}$$

On the other hand, we assume that $i \leq j - 1$ without loss of generality, then

$$\begin{aligned} g_{i,j}^{(n)} &= \left(\underbrace{\mathbf{u}, \dots, \mathbf{u}}_{i-1 \text{ times}}, \underbrace{\mathbf{v}, \mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{2^n - i \text{ times}}, \underbrace{\mathbf{u}^\top, \dots, \mathbf{u}^\top}_{j-1 \text{ times}}, \underbrace{\mathbf{v}^\top, \mathbf{0}_{2 \times 1}, \dots, \mathbf{0}_{2 \times 1}}_{2^n - j \text{ times}} \right)^\top \\ &= 4(i-1) + \sqrt{2} \left(\frac{3}{2\sqrt{2}} + \frac{1}{2\sqrt{2}} \right) = 4i - 2. \end{aligned}$$

This demonstrates the exact form of the matrix G_n as shown in (4.33). Since $G_n = 4D_{2^n} - \frac{3}{4}\mathbf{I}_{2^n \times 2^n}$, thus

$$\|G_n\|_2 = \left\| 4D_{2^n} - \frac{3}{4}\mathbf{I}_{2^n \times 2^n} \right\|_2 \leq 4\|D_{2^n}\|_2 + \frac{3}{4} \leq 2 \left(1 - \cos \frac{\pi}{2^{n+1}} \right)^{-1} + \frac{3}{4},$$

where the last inequality follows from Lemma 4.18. Similarly, we have

$$\|G_n\|_2 = \left\| 4D_{2^n} - \frac{3}{4}\mathbf{I}_{2^n \times 2^n} \right\|_2 \geq \left| 4\|D_{2^n}\|_2 - \frac{3}{4} \right| = 2 \left(1 - \cos \frac{\pi}{2^{n+1}} \right)^{-1} - \frac{3}{4}.$$

This completes the proof. $\square$

4.5 Proofs for the results in Chapter 4

Proof of Theorem 4.1. We consider first the case in which $\hat{f}_0 = 0$. Let us consider the 2^{n+1} dimensional column vector $\boldsymbol{\omega}_{n+1} = (\omega_i^{(n+1)})_{i=1}^{2^{n+1}}$, where $\omega_i^{(n+1)} = F(i2^{-n-1}) - F((i-1)2^{-n-1})$. Recall the explicit form of $Q_{(n+1,n+1)}^{-1}$ in (4.28), which can be regarded as a row operation acting on a vector by subtracting preceding rows after properly scaling. Therefore, we have $Q_{(n+1,n+1)}^{-1}\mathbf{y}_{n+1} = 2^{3(n+1)/2+2}\boldsymbol{\omega}_{n+1}$. Thus,

$$\boldsymbol{\vartheta}_{n-1}^{(n)} = R_{n-1}^{(n)} P_n Q_{(n+1,n+1)}^{-1} \mathbf{y}_{n+1} = 2^{3(n+1)/2+2} R_{n-1}^{(n)} P_n \boldsymbol{\omega}_{n+1}.$$

Analogously, for $0 \leq m \leq n-1$, we partition $\boldsymbol{\omega}_{n+1}$ as follows,

$$\boldsymbol{\omega}_{n+1} = \left((\boldsymbol{\omega}_1^{(m,n+1)})^\top, (\boldsymbol{\omega}_2^{(m,n+1)})^\top, \dots, (\boldsymbol{\omega}_{2^m}^{(m,n+1)})^\top \right)^\top,$$

where each vector $\boldsymbol{\omega}_k^{(m,n+1)}$ is a 2^{n+1-m} dimensional column vector. Then, for a similar reason, we have

$$\begin{aligned}\vartheta_{m,k}^{(n)} &= 2^{3(n+1)/2+2} \left(\boldsymbol{\eta}_{k+1,1}^{(m)}, \boldsymbol{\eta}_{k+1,2}^{(m)}, \dots, \boldsymbol{\eta}_{k+1,2^m}^{(m)} \right) \cdot \boldsymbol{\omega}_{n+1} = 2^{3(n+1)/2+2} \boldsymbol{\eta}_{k+1,k+1}^{(m)} \boldsymbol{\omega}_{k+1}^{(m,n+1)} \\ &= 2^{3(n+1)/2+(m-n-3)/2+2} \left((\boldsymbol{r}_{n-m}^+)^{\top}, (\boldsymbol{r}_{n-m}^-)^{\top} \right) \left((\boldsymbol{\omega}_{2k+1}^{(m+1,n+1)})^{\top}, (\boldsymbol{\omega}_{2k+2}^{(m+1,n+1)})^{\top} \right)^{\top} \quad (4.34) \\ &= 2^{n+m/2+2} \left((\boldsymbol{r}_{n-m}^+)^{\top} \boldsymbol{\omega}_{2k+1}^{(m+1,n+1)} + (\boldsymbol{r}_{n-m}^-)^{\top} \boldsymbol{\omega}_{2k+2}^{(m+1,n+1)} \right).\end{aligned}$$

A routine calculation gives that

$$(\boldsymbol{r}_{n-m}^+)^{\top} \boldsymbol{\omega}_{2k+1}^{(m+1,n+1)} = \sum_{j=1}^{2^{n-m}} (-1)^j \omega_{2^{n-m+1}k+j}^{(n+1)} = \sum_{j=1}^{2^{n-m}} (-1)^j \left(F\left(\frac{k}{2^m} + \frac{j}{2^{n+1}}\right) - F\left(\frac{k}{2^m} + \frac{j-1}{2^{n+1}}\right) \right),$$

and we similarly have

$$\begin{aligned}(\boldsymbol{r}_{n-m}^-)^{\top} \boldsymbol{\omega}_{2k+2}^{(m+1,n+1)} &= \sum_{j=1}^{2^{n-m}} (-1)^{j+1} \omega_{2^{n-m}(2k+1)+j}^{(n+1)} = \sum_{j=0}^{2^{n-m}-1} (-1)^{j+1} \omega_{2^{n-m+1}(k+1)-j}^{(n+1)} \\ &= \sum_{j=1}^{2^{n-m}} (-1)^j \omega_{2^{n-m+1}(k+1)-(j-1)}^{(n+1)} = \sum_{j=1}^{2^{n-m}} (-1)^j \left(F\left(\frac{k+1}{2^m} - \frac{j-1}{2^{n+1}}\right) - F\left(\frac{k+1}{2^m} - \frac{j}{2^{n+1}}\right) \right).\end{aligned}$$

Plugging the above two identities back to (4.34) gives (4.4). Last, for $m = -1$, we have

$$\begin{aligned}\vartheta_{-1,0}^{(n)} &= 2^{3(n+1)/2+2} \mathcal{C}_{-1} \boldsymbol{\omega}_{n+1} = 2^{n+2} \boldsymbol{r}_{n+1}^+ \boldsymbol{\omega}_{n+1} = 2^{n+2} \sum_{j=1}^{2^{n+1}} (-1)^j \omega_j^{(n+1)} \\ &= 2^{n+2} \sum_{j=1}^{2^{n+1}} (-1)^j \left(F\left(\frac{j}{2^{n+1}}\right) - F\left(\frac{j-1}{2^{n+1}}\right) \right).\end{aligned}$$

Last, note that $\bar{\vartheta}_n^{(n)} = 2^{3(n+1)/2+2} \bar{R}_n^{(n)} P_n \omega_{n+1} = 2^{3(n+1)/2+2} \mathcal{C}_n \omega_{n+1}$, and therefore,

$$\begin{aligned}
\vartheta_{n,k}^{(n)} &= 2^{3(n+1)/2+2} \left(\underbrace{\mathbf{u}, \dots, \mathbf{u}}_{k \text{ times}}, \mathbf{v}, \underbrace{\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{2^n - k - 1 \text{ times}} \right) \omega_{n+1} \\
&= 2^{3(n+1)/2+2} \left(\underbrace{\mathbf{u}, \dots, \mathbf{u}}_{k \text{ times}}, \mathbf{v}, \underbrace{\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{2^n - k - 1 \text{ times}} \right) \left((\omega_1^{(n,n+1)})^\top, (\omega_2^{(n,n+1)})^\top, \dots, (\omega_{2^n}^{(n,n+1)})^\top \right)^\top \\
&= 2^{3(n+1)/2+2} \left(\sum_{j=1}^k \mathbf{u} \cdot \omega_j^{(n,n+1)} + \mathbf{v} \cdot \omega_{k+1}^{(n,n+1)} \right) \\
&= 2^{3(n+1)/2+2} \left(\sqrt{2} \sum_{j=1}^k (\omega_{2j-1}^{(n+1)} - \omega_{2j}^{(n+1)}) + \frac{3}{2\sqrt{2}} \omega_{2k+1}^{(n+1)} - \frac{1}{2\sqrt{2}} \omega_{2k+2}^{(n+1)} \right) \\
&= 2^{3(n+1)/2+5/2} \sum_{j=1}^k \left(2F\left(\frac{2j-1}{2^{n+1}}\right) - F\left(\frac{2j-2}{2^{n+1}}\right) - F\left(\frac{2j}{2^{n+1}}\right) \right) \\
&\quad + 3 \cdot 2^{3(n+1)/2+1} \left(F\left(\frac{2k+1}{2^{n+1}}\right) - F\left(\frac{2k}{2^{n+1}}\right) \right) - 2^{3(n+1)/2+1} \left(F\left(\frac{2k+2}{2^{n+1}}\right) - F\left(\frac{2k+1}{2^{n+1}}\right) \right).
\end{aligned}$$

This completes the proof for the case $\hat{f}_0 = 0$.

For $\hat{f}_0 \neq 0$, we let $\tilde{\vartheta}_{m,k}^{(n)}$ be the estimated Faber–Schauder coefficients for $\hat{f}_0 = 0$, as in the first part of this proof. Then the functions

$$\tilde{f}_n := \tilde{\vartheta}_{-1,0}^{(n)} e_{-1,0} + \sum_{m=0}^n \sum_{k=0}^{2^m-1} \tilde{\vartheta}_{m,k}^{(n)} e_{m,k} \quad \text{and} \quad \tilde{F}(t) := F(0) + \int_0^t \tilde{f}_n(s) ds$$

are such that $\tilde{F}(t) = F(t)$ for all $t \in \mathbb{T}_{n+1}$. Now we define $\vartheta_{m,k}^{(n)} := \tilde{\vartheta}_{m,k}^{(n)}$ for $m < n$ and $\vartheta_{n,k}^{(n)} := \tilde{\vartheta}_{n,k}^{(n)} - 2^{-n/2+2} \hat{f}_0$. Then we let

$$\hat{f}_n := \hat{f}_0 + \vartheta_{-1,0}^{(n)} e_{-1,0} + \sum_{m=0}^n \sum_{k=0}^{2^m-1} \vartheta_{m,k}^{(n)} e_{m,k}, \quad \text{and} \quad \hat{F}_n(t) := F(0) + \int_0^t \hat{f}_n(s) ds.$$

With these definitions, we have that

$$\hat{F}_n(t) = \tilde{F}(t) + \hat{f}_0 \cdot t - 2^{-n/2+2} \hat{f}_0 \int_0^t \sum_{k=0}^{2^n-1} e_{n,k}(s) ds = \tilde{F}(t) + \hat{f}_0 \cdot t - 2^{-n/2+2} \hat{f}_0 \sum_{k=0}^{2^n-1} \psi_{n,k}(t).$$

By using (4.14) and considering separately the case $t \in \mathbb{T}_n$ and $t \in \mathbb{T}_{n+1} \setminus \mathbb{T}_n$, one easily checks that $2^{-n/2+2} \sum_{k=0}^{2^n-1} \psi_{n,k}(t) = t$ for $t \in \mathbb{T}_{n+1}$. Hence, $\hat{F}(t) = \tilde{F}(t) = F(t)$ for all $t \in \mathbb{T}_{n+1}$, and so the coefficients $\vartheta_{m,k}^{(n)}$ are as desired. To show uniqueness, one argues in a

similar way, transforming any solution to (4.3) back to the case $\hat{f}_0 = 0$ and using the already established uniqueness there. $\square$

Proof of the formulas in Example 4.2. Consider the following partition of $\mathbf{z}_{n+1}$,

$$\mathbf{z}_{n+1} = \left((\mathbf{z}_1^{(m,n+1)})^\top, (\mathbf{z}_2^{(m,n+1)})^\top, \dots, (\mathbf{z}_{2^m}^{(m,n+1)})^\top \right)^\top,$$

where each vector $\mathbf{z}_k^{(m,n+1)}$ is a 2^{n+1-m} dimensional column vector for $0 \leq m \leq n+1$ and $1 \leq k \leq 2^{n+1-m}$. Moreover, for each $1 \leq i \leq 2^{n+1}$, we have

$$\begin{aligned} z_i^{(n+1)} &= 2^{3(n+1)/2} \sum_{m=n+1}^{\infty} 2^{-3m/2} \sum_{k=2^{m-n-1}(i-1)}^{2^{m-n-1}i-1} \theta_{m,k} \\ &= 2^{3(n+1)/2} \sum_{m=n+1}^{\infty} 2^{-3m/2} \sum_{k=2^{m-n-1}(i-1)}^{2^{m-n-1}i-1} 2^{m/2} c_m \\ &= 2^{3(n+1)/2} \sum_{m=n+1}^{\infty} 2^{-3m/2+m-n-1+m/2} c_m = 2^{(n+1)/2} \sum_{m=n+1}^{\infty} c_m =: \kappa_{n+1} < \infty, \end{aligned} \tag{4.35}$$

where the finiteness of κ_{n+1} follows from the absolute summability of the sequence (c_m) . Therefore, for every $1 \leq k \leq 2^m$, we have

$$\mathbf{z}_k^{(m,n+1)} = \underbrace{(\kappa_{n+1}, \kappa_{n+1}, \dots, \kappa_{n+1})}_{2^{n+1-m} \text{ times}}.$$

Furthermore, it follows from Lemma 4.15 that $\boldsymbol{\eta}_{i,j}^{(m)} = \mathbf{0}_{1 \times 2^{n-m+1}}$ for $i \neq j$, then for $0 \leq m \leq n-1$ and $0 \leq k \leq 2^m - 1$

$$\begin{aligned} \vartheta_{m,k}^{(n)} - \theta_{m,k} &= \left(\boldsymbol{\eta}_{k+1,1}^{(m)}, \boldsymbol{\eta}_{k+1,2}^{(m)}, \dots, \boldsymbol{\eta}_{k+1,2^m}^{(m)} \right) \mathbf{z}_{n+1} = \boldsymbol{\eta}_{k+1,k+1}^{(m)} \mathbf{z}_{k+1}^{(m,n+1)} \\ &= 2^{\frac{m-n-3}{2}} \left((\mathbf{r}_{n-m}^+)^{\top}, (\mathbf{r}_{n-m}^-)^{\top} \right) \left((\mathbf{z}_1^{(m,n)})^{\top}, (\mathbf{z}_1^{(m,n)})^{\top} \right)^{\top} \\ &= 2^{\frac{m-n-3}{2}} (\mathbf{r}_{n-m}^+ + \mathbf{r}_{n-m}^-)^{\top} \mathbf{z}_1^{(m,n)} = 0. \end{aligned}$$

The proof for the case $m = -1$ is analogous and we next prove the case $m = n$. Let us use

the shorthand notation $\boldsymbol{\kappa}_{n+1} = (\kappa_{n+1}, \kappa_{n+1})$, and an analogous calculation yields that

$$\begin{aligned}
\vartheta_{n,k}^{(n)} - \theta_{n,k} &= \left(\underbrace{\mathbf{u}, \dots, \mathbf{u}}_{k \text{ times}}, \mathbf{v}, \underbrace{\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{2^n - k - 1 \text{ times}} \right) \mathbf{z}_{n+1} \\
&= \left(\underbrace{\mathbf{u}, \dots, \mathbf{u}}_{k \text{ times}}, \mathbf{v}, \underbrace{\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{2^n - k - 1 \text{ times}} \right) \left(\underbrace{\boldsymbol{\kappa}_{n+1}, \boldsymbol{\kappa}_{n+1}, \dots, \boldsymbol{\kappa}_{n+1}}_{2^n \text{ times}} \right)^\top \\
&= k(\mathbf{u} \cdot \boldsymbol{\kappa}_{n+1}^\top) + \mathbf{v} \cdot \boldsymbol{\kappa}_{n+1}^\top = \left(\frac{3}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}} \right) \cdot (\kappa_{n+1}, \kappa_{n+1})^\top \\
&= \frac{\sqrt{2}}{2} \kappa_{n+1} = 2^{n/2} \sum_{m=n+1}^{\infty} c_m.
\end{aligned}$$

Thus, we have

$$\vartheta_{n,k}^{(n)} = 2^{n/2} c_n + 2^{n/2} \sum_{m=n+1}^{\infty} c_m = 2^{n/2} \sum_{m=n}^{\infty} c_m.$$

This completes the proof. $\square$

Proof of Proposition 4.3. By assumption, there exists $c \in [0, 1]$ such that

$$\begin{aligned}
\omega_2(f, t) &= \sup_{\substack{\tau_1, \tau_2 \in [0, 1] \\ |\tau_1 - \tau_2| \leq t}} \left\{ \left| f(\tau_1) + f(\tau_2) - 2f\left(\frac{\tau_1 + \tau_2}{2}\right) \right| \right\} \\
&\leq |\tau_1 - \tau_2| \sup_{\substack{\tau_1, \tau_2 \in [0, 1] \\ |\tau_1 - \tau_2| \leq t}} \left(\sup_{u_1, u_2 \in [\tau_1, \tau_2]} |f'(u_1) - f'(u_2)| \right) \leq c |\tau_1 - \tau_2|^{1+\alpha} \leq c \cdot t^{1+\alpha}.
\end{aligned} \tag{4.36}$$

Note next that (4.2) implies $|\theta_{m,k}| \leq 2^{m/2} \omega_2(f, 2^{-m})$. Thus,

$$\begin{aligned}
|z_i^{(n+1)}| &= \sum_{m=n+1}^{\infty} 2^{-3(m-n-1)/2} \sum_{k=2^{m-n-1}(i-1)}^{2^{m-n-1}i-1} |\theta_{m,k}| \\
&\leq \sum_{m=n+1}^{\infty} 2^{-3(m-n-1)/2} \cdot \left(2^{3m/2-n-1} \omega_2(f, 2^{-m}) \right) \\
&= 2^{(n+1)/2} \sum_{m=n+1}^{\infty} \omega_2(f, 2^{-m}) \leq c \cdot 2^{(n+1)/2} \sum_{m=n+1}^{\infty} 2^{-m(1+\alpha)} \\
&\leq \frac{c \cdot 2^{-(n+1)(\alpha + \frac{1}{2})}}{1 - 2^{-\alpha-1}}.
\end{aligned} \tag{4.37}$$

From here, the assertion follows easily. $\square$

Proof of Theorem 4.8 We start with the case $p = 1$. For $m < n$, Theorem 4.7 and the fact

that $\|e_{m,k}\|_{L_1[0,1]} = 2^{-3m/2-2}$ yield that

$$\begin{aligned}
\|f_m - \hat{f}_{n,m}\|_{L_1[0,1]} &\leq \left\| \sum_{i=-1}^m \sum_{k=0}^{2^{i \vee 0} - 1} (\vartheta_{i,k} - \theta_{i,k}) e_{i,k} \right\|_{L_1[0,1]} \\
&\leq \sum_{i=-1}^m \left\| \sum_{k=0}^{2^{i \vee 0} - 1} (\vartheta_{i,k} - \theta_{i,k}) e_{i,k} \right\|_{L_1[0,1]} \\
&= \sum_{i=0}^m 2^{-3i/2-2} \|\bar{\vartheta}_i^{(n)} - \bar{\theta}_i\|_{\ell_1} + \frac{1}{2} \|\bar{\vartheta}_{-1}^{(n)} - \bar{\theta}_{-1}\|_{\ell_1} \\
&\leq \sum_{i=0}^m 2^{-i-2-(n+3)/2} \|\mathbf{z}_{n+1}\|_{\ell_1} + 2^{-(n+5)/2} \|\mathbf{z}_{n+1}\|_{\ell_1} \\
&= 2^{-(n+3)/2} \|\mathbf{z}_{n+1}\|_{\ell_1}.
\end{aligned} \tag{4.38}$$

For the case $m = n$, we get in the same manner $\|f_n - \hat{f}_{n,n}\|_{L_1[0,1]} \leq 2^{-(n+1)/2} \|\mathbf{z}_{n+1}\|_{\ell_1}$. To relate these two inequalities with the second-order modulus of continuity, we recall from (4.37) that $|z_i^{(n+1)}| \leq 2^{(n+1)/2} \sum_{m=n+1}^{\infty} \omega_2(f, 2^{-m})$. From here, the proof of the case $p = 1$ is easily completed.

For $p = 2$, we have $\|e_{m,k}\|_{L_2[0,1]} = 2^{-m-1}/\sqrt{3}$ and $e_{m,k} \cdot e_{m,\ell} = 0$ for $k \neq \ell$. Thus,

$$\begin{aligned}
\|f_m - \hat{f}_{n,m}\|_{L_2[0,1]} &\leq \left\| \sum_{i=-1}^m \sum_{k=0}^{2^{i \vee 0} - 1} (\vartheta_{i,k} - \theta_{i,k}) e_{i,k} \right\|_{L_2[0,1]} \\
&\leq \sum_{i=-1}^m \left\| \sum_{k=0}^{2^{i \vee 0} - 1} (\vartheta_{i,k} - \theta_{i,k}) e_{i,k} \right\|_{L_2[0,1]} \\
&= \frac{1}{\sqrt{3}} \sum_{i=-1}^m 2^{-i-1} \|\bar{\vartheta}_i^{(n)} - \bar{\theta}_i\|_{\ell_2} \\
&\leq \frac{1}{\sqrt{3}} \sum_{i=-1}^m 2^{-i-2} \|\mathbf{z}_{n+1}\|_{\ell_2} \leq \frac{\|\mathbf{z}_{n+1}\|_{\ell_2}}{\sqrt{3}}.
\end{aligned} \tag{4.39}$$

Next, we can argue as in the proof for $p = 1$ to get $\|\mathbf{z}_{n+1}\|_{\ell_2} \leq 2^{n+1} \sum_{m=n+1}^{\infty} \omega_2(f, 2^{-m})$, which gives the first inequality. The proof of the second one is analogous.

For $p = \infty$, we have $\|e_{-1,0}\|_{L_{\infty}[0,1]} = 1$ and $\|e_{m,k}\|_{L_{\infty}[0,1]} = 2^{-m/2-1}$ for $m \in \mathbb{N}$ and

$0 \leq k \leq 2^m - 1$. The triangle inequality then yields that for $0 \leq m \leq n - 1$

$$\begin{aligned}
\|f_m - \hat{f}_{n,m}\|_{L_\infty[0,1]} &\leq \left\| \sum_{i=-1}^m \sum_{k=0}^{2^{i \vee 0} - 1} (\vartheta_{i,k} - \theta_{i,k}) e_{i,k} \right\|_{L_\infty[0,1]} \\
&\leq \sum_{i=-1}^m \left\| \sum_{k=0}^{2^{i \vee 0} - 1} (\vartheta_{i,k} - \theta_{i,k}) e_{i,k} \right\|_{L_\infty[0,1]} \\
&= \sum_{i=0}^m 2^{-i/2-1} \|\bar{\vartheta}_i^{(n)} - \bar{\theta}_i\|_{\ell_\infty} + \|\bar{\vartheta}_{-1}^{(n)} - \bar{\theta}_{-1}\|_{\ell_\infty} \\
&\leq \sum_{i=0}^m 2^{-i/2-1+(n-i-1)/2} \|\mathbf{z}_{n+1}\|_{\ell_\infty} + 2^{(n-1)/2} \|\mathbf{z}_{n+1}\|_{\ell_\infty} \\
&= \left(2^{(n-3)/2} \sum_{i=0}^m 2^{-i} + 2^{(n-1)/2} \right) \|\mathbf{z}_{n+1}\|_{\ell_\infty} \\
&\leq 2^{(n+1)/2} \|\mathbf{z}_{n+1}\|_{\ell_\infty},
\end{aligned} \tag{4.40}$$

where we have used Theorem 4.7 in the fourth step. For the case $m = n$, an analogous decomposition leads to

$$\begin{aligned}
\|f_n - \hat{f}_{n,n}\|_{L_\infty[0,1]} &\leq \left\| (f_{n-1} - \hat{f}_{n,n-1}) + \sum_{k=0}^{2^n-1} (\theta_{n,k} - \vartheta_{n,k}^{(n)}) e_{n,k} \right\|_{L_\infty[0,1]} \\
&\leq \|f_{n-1} - \hat{f}_{n,n-1}\|_{L_\infty[0,1]} + 2^{-n/2-1} \|\bar{\vartheta}_n^{(n)} - \bar{\theta}_n\|_{\ell_\infty} \\
&\leq 2^{(n+1)/2} \|\mathbf{z}_{n+1}\|_{\ell_\infty} + \left(\frac{2^{n+\frac{3}{2}} - \sqrt{2}}{2^{n/2+1}} \right) \|\mathbf{z}_{n+1}\|_{\ell_\infty} \\
&\leq 2^{(n+3)/2} \|\mathbf{z}_{n+1}\|_{\ell_\infty}.
\end{aligned} \tag{4.41}$$

It follows from (4.37) that

$$\|\mathbf{z}_{n+1}\|_{\ell_\infty} \leq 2^{(n+1)/2} \sum_{k=n+1}^{\infty} \omega_2(f, 2^{-k}), \tag{4.42}$$

and plugging this inequality back into (4.40) and (4.41) completes the proof. $\square$

The proof of Theorem 4.11 will follow immediately from the next lemma. Indeed, the quadratic spline interpolation $\hat{F}_{n,n}$ coincides with F on the dyadic partition $\mathbb{T}_{n+1}$, and so $\hat{F}_{n,n} - F$ belongs to the linear space $\mathcal{X}_{n+1}$ introduced in (4.43) below.

Lemma 4.19. *Let us consider a vector space $\mathcal{X}_n$ of functions, such that*

$$\mathcal{X}_n := \left\{ F \in C^1[0, 1] \left| F\left(\frac{k}{2^n}\right) = 0 \text{ for } 0 \leq k \leq 2^n \right. \right\} \text{ for } n \in \mathbb{N}. \quad (4.43)$$

Moreover, we also consider the vector space $\mathcal{Y}_n$, where

$$\mathcal{Y}_n := \left\{ f \in C[0, 1] \left| \int_0^1 f(s) ds \in \mathcal{X}_n \right. \right\} \text{ for } n \in \mathbb{N}.$$

Furthermore, by $\mathcal{X}_n^1$, $\mathcal{X}_n^2$ and $\mathcal{X}_n^\infty$, we denote the normed linear space $\mathcal{X}_n$ equipped with the $L_1[0, 1]$, $L_2[0, 1]$ and $L_\infty[0, 1]$ norms respectively, as well as, the analogously defined normed spaces $\mathcal{Y}_n^1$, $\mathcal{Y}_n^2$ and $\mathcal{Y}_n^\infty$. Let $T : \mathcal{Y}_n \rightarrow \mathcal{X}_n$ denote the integral operator, i.e., $Tf = \int_0^1 f(s) ds$. Then, its operator norms are given by $\|T\|_{\mathcal{Y}_n^p \rightarrow \mathcal{X}_n^p} = 2a_p$, where a_p is as in Theorem 4.11.

Proof. First, we consider the case $p = \infty$. Since $F(2^{-n}k) = 0$ for all k ,

$$\begin{aligned} \|F\|_{L_\infty[0,1]} &= \max_{1 \leq k \leq 2^n} \max_{\frac{k-1}{2^n} \leq t \leq \frac{k}{2^n}} |F(t)| \\ &\leq \max_{1 \leq k \leq 2^n} \max_{\frac{k-1}{2^n} \leq t \leq \frac{k}{2^n}} \int_{\frac{k-1}{2^n}}^t |f(s)| ds \leq 2^{-n-1} \|f\|_{L_\infty[0,1]}. \end{aligned} \quad (4.44)$$

This shows that $\|T\|_{\mathcal{Y}_n^\infty \rightarrow \mathcal{X}_n^\infty} \leq 2^{-n-1}$. To get the converse inequality, we take $\varepsilon \in (0, 2^{-n-3})$ and consider the function f_ε that is equal to 1 on intervals of the form $[\frac{k}{2^n} + \varepsilon, \frac{2k+1}{2^{n+1}} - \varepsilon]$, equal to zero on intervals of the form $[\frac{2k+1}{2^{n+1}} + \varepsilon, \frac{k+1}{2^{n+1}} - \varepsilon]$ and linearly interpolated everywhere else. Then $f_\varepsilon \in \mathcal{X}_n$ and we have $\|f_\varepsilon\|_{L_\infty[0,1]} = 1$ and $\|Tf_\varepsilon\|_{L_\infty[0,1]} = 2^{-n-1} - 2\varepsilon$, and the proof for $p = \infty$ is complete.

For the $L_2[0, 1]$ norm, let us take any $f \in \mathcal{Y}_n$ with the corresponding $F := Tf \in \mathcal{X}_n$ and define $F_k(t) = F(2^{-n}(t + k))$ for $t \in [0, 1]$. Then

$$\|F\|_{L_2[0,1]}^2 = \sum_{k=0}^{2^n-1} \int_{2^{-n}k}^{2^{-n}(k+1)} F^2(t) dt = 2^{-n} \sum_{k=0}^{2^n-1} \int_0^1 F_k^2(t) dt = 2^{-n} \sum_{k=0}^{2^n-1} \|F_k\|_{L_2[0,1]}^2. \quad (4.45)$$

Furthermore, let us denote $f_k(t) := F'_k(t)$. Then $f_k(t) = 2^{-n}f(2^{-n}(t + k))$ and so

$$\begin{aligned} \sum_{k=0}^{2^n-1} \|f_k\|_{L_2[0,1]}^2 &= \sum_{k=0}^{2^n-1} \int_0^1 f_k^2(t) dt = 2^{-2n} \sum_{k=0}^{2^n-1} \int_0^1 f^2(2^{-n}(t + k)) dt \\ &= 2^{-n} \int_0^1 f^2(t) dt = 2^{-n} \|f\|_{L_2[0,1]}^2. \end{aligned}$$

Next, for $0 \leq k \leq 2^n - 1$, in the spirit of [43, Problem 188], the Hölder inequality yields that

$$\begin{aligned}
\|F_k\|_{L_2[0,1]}^2 &= \int_0^1 F_k^2(t) dt = \int_0^1 \left(\int_0^t f_k(s) ds \right)^2 dt \\
&= \int_0^1 \left(\int_0^t \sqrt{\cos \frac{\pi s}{2}} \frac{f_k(s)}{\sqrt{\cos \frac{\pi s}{2}}} ds \right)^2 dt \\
&\leq \int_0^1 \left(\int_0^t \cos \frac{\pi s}{2} ds \int_0^t \frac{f_k^2(s)}{\cos \frac{\pi s}{2}} ds \right) dt \\
&= \frac{2}{\pi} \int_0^1 \sin \frac{\pi t}{2} \int_0^t \frac{f_k^2(s)}{\cos \frac{\pi s}{2}} ds dt \\
&= \frac{2}{\pi} \int_0^1 \frac{f_k^2(s)}{\cos \frac{\pi s}{2}} \int_s^1 \sin \frac{\pi t}{2} dt ds \\
&= \frac{4}{\pi^2} \int_0^1 \frac{f_k^2(s)}{\cos \frac{\pi s}{2}} \left(-\cos \frac{\pi}{2} + \cos \frac{\pi s}{2} \right) ds = \frac{4}{\pi^2} \|f_k\|_{L_2[0,1]}^2.
\end{aligned} \tag{4.46}$$

Thus, $\|Tf\|_{L_2[0,1]} \leq \frac{2^{-n+1}}{\pi} \|f\|_{L_2[0,1]}$.

To prove the converse inequality, note that the inequality (4.46) becomes an identity for $f_k(t) := \cos \frac{\pi t}{2}$. To see why, note first that $F_k(t) := Tf_k(t) = \frac{2}{\pi} \sin \frac{\pi t}{2}$ satisfies $\|F_k\|_{L_2[0,1]}^2 = \frac{4}{\pi^2} \int_0^1 (\sin \frac{\pi t}{2})^2 dt = \frac{2}{\pi^2}$ and $\|f_k\|_{L_2[0,1]}^2 = \int_0^1 (\cos \frac{\pi t}{2})^2 dt = \frac{1}{2}$. In such a way, one sees that the function $f(t) = \cos 2^{n-1}\pi t$ belongs to $\mathcal{Y}_n$, and $\|Tf\|_{L_2[0,1]} = \frac{2^{-n+1}}{\pi} \|f\|_{L_2[0,1]}$. This completes the proof for $p = 2$.

Finally, the case for $p = 1$ is similar to the case for $p = 2$. We get as in (4.45) that $\|F\|_{L_1[0,1]} = 2^{-n} \sum_{k=0}^{2^n-1} \|F_k\|_{L_1[0,1]}$, where F_k and $f_k = F'_k$ are as above. In the same way, we get $\sum_{k=0}^{2^n-1} \|f_k\|_{L_1[0,1]} = \|f\|_{L_1[0,1]}$. Moreover, for each $0 \leq k \leq 2^n - 1$,

$$\|F_k\|_{L_1[0,1]} = \int_0^1 |F_k(t)| dt \leq \int_0^1 \int_0^t |f_k(s)| dt \leq \|f_k\|_{L_1[0,1]}.$$

Plugging this inequality back to the above formulae shows that $\|T\|_{\mathcal{Y}_n^1 \rightarrow \mathcal{X}_n^1} \leq 2^{-n}$. To show the reverse inequality, we take $\varepsilon < \frac{1}{2}$ and $m \geq 1$ and define the function $g_{m,\varepsilon} \in C[0,1]$ as follows,

$$g_{m,\varepsilon} = \frac{mt}{\varepsilon^2} \mathbb{1}_{[0, \frac{\varepsilon}{m}]} + \frac{1}{\varepsilon} \mathbb{1}_{(\frac{\varepsilon}{m}, \varepsilon]} - \frac{m}{\varepsilon^2} \left(t - \frac{(m+1)\varepsilon}{m} \right) \mathbb{1}_{(\varepsilon, \frac{(m+1)\varepsilon}{m}]},$$

where $\mathbb{1}_A$ denotes the indicator function of a set A . In other words, the function $g_{m,\varepsilon}$ defines an isosceles Trapezoid with height $1/\varepsilon$. Thus, $\|g_{m,\varepsilon}\|_{L_1[0,1]} = 1$ for any $\varepsilon < 1/2$ and $m \geq 1$.

Clearly, $G_{m,\varepsilon} := Tg_{m,\varepsilon}$ is given by

$$\begin{aligned} G_{m,\varepsilon}(t) &= \frac{mt^2}{2\varepsilon^2} \mathbb{1}_{[0, \frac{\varepsilon}{m}]} + \left(\frac{1}{\varepsilon} \left(t - \frac{\varepsilon}{m} \right) + \frac{1}{2m} \right) \mathbb{1}_{(\frac{\varepsilon}{m}, \varepsilon]} \\ &\quad + \left(\frac{2m-1}{2m} - \frac{m}{2\varepsilon^2} \left(t - \frac{(m+1)\varepsilon}{m} \right)^2 \right) \mathbb{1}_{(\varepsilon, \frac{(m+1)\varepsilon}{m}]} + \mathbb{1}_{(\frac{(m+1)\varepsilon}{m}, 1]}. \end{aligned}$$

Thus, we have $\|G_{m,\varepsilon}\|_{L_1[0,1]} \geq 1 - \frac{(m+1)\varepsilon}{m}$. Note that for any pair of m and ε , we have $g_{m,\varepsilon}(0) = g_{m,\varepsilon}(1) = 0$. Thus, can define a function $f \in C[0, 1]$ by $f(t) := 2^n g_{m,\varepsilon}(2^n t - k)$ for $t \in [2^{-n}k, 2^{-n}(k+1)]$ so that $f_k = g_{m,\varepsilon}$ for all k . Furthermore, we have

$$\|F\|_{L_1[0,1]} = 2^{-n} \sum_{k=0}^{2^n-1} \|F_k\|_{L_1[0,1]} \geq 2^{-n} \left(1 - \frac{(m+1)\varepsilon}{m} \right) = 2^{-n} \left(1 - \frac{(m+1)\varepsilon}{m} \right) \|f\|_{L_1[0,1]}.$$

Simultaneously taking $m \uparrow \infty$ and $\varepsilon \downarrow 0$ shows that $\|T\|_{\mathcal{Y}_n^1 \rightarrow \mathcal{X}_n^1} = 2^{-n}$. This completes the proof. $\square$

Chapter 5

The Φ -variation of classical fractal functions

This chapter is based on [49, 50]. In this chapter, we study the Wiener–Young Φ -variation of some classical fractal functions. These functions admit the following representation:

$$f(t) := \sum_{m=0}^{\infty} \alpha_m \varphi(b^m t), \quad t \in [0, 1],$$

where $(\alpha_m) \in \ell_1$ and $b \in \{2, 3, \dots\}$. Furthermore, we require that the base function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is Lipschitz continuous and periodic with period 1. Then the series on the right-hand side of the above equation converges absolutely and uniformly in $t \in [0, 1]$, so that f is indeed a well-defined continuous function. In particular, if $\alpha_m = \alpha^m$ for some $\alpha \in (-1, 1)$ and $\varphi(t) = \nu \sin(2\pi t) + \rho \cos(2\pi t)$ for real constants ν and ρ , then f is a Weierstraß function. If $\alpha_m = \alpha^m$ for some $\alpha \in (-1, 1)$ and $\varphi(t) = \min_{z \in \mathbb{Z}} |z - t|$ is the tent map, then f is a Takagi–van der Waerden function. Finally, if $b = 2$ and $\varphi(t) = \min_{z \in \mathbb{Z}} |z - t|$ is the tent map, then f belongs to the Takagi class.

This chapter is structured as follows. In Section 5.1, we show that the Weierstraß and Takagi–van der Waerden function admits a nonzero, finite and linear Φ -variation along the b -adic partition for $\Phi = x/\sqrt{-\log x}$. Those results are consequences of suitable central limit theorems for the Rademacher functions in the Faber–Schauder development of f . The development of the corresponding theory will be presented in Section 5.2.

The latter part of this chapter introduces a probabilistic approach to identifying those functions in the Takagi class that have nonzero, finite and linear Φ -variation for a prescribed

function Φ of the form

$$\Phi_q(x) := x^{\frac{1}{1-q}} g\left(\frac{-\log_2 x}{1-q}\right)^{-\frac{1}{2(1-q)}}, \quad 0 < x < 1,$$

where $q \in [0, 1)$ and $g : [0, \infty) \rightarrow (0, \infty)$ is any regularly varying function¹. Our approach to the Φ_q -variation of functions in the Takagi class depends on the parameter q . For $q = 0$, the argument extends the approach developed in Section 5.2. For $q \in (0, 1)$, however, a new limit theorem for Bernoulli-type convolutions is needed. This convergence result is collected in Section 5.3. Our main result on the Φ_q -variation of functions in the Takagi class, Theorem 5.10, is stated in Section 5.4. The proof of Theorem 5.10 is given in Section 5.5.

5.1 Φ -variation of Takagi–van der Waerden and Weierstraß functions

We consider a base function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ that is periodic with period one and Lipschitz continuous. Our aim is to study the function

$$f(t) := \sum_{m=0}^{\infty} \alpha^m \varphi(b^m t), \quad t \in [0, 1], \quad (5.1)$$

where $b \in \{2, 3, \dots\}$ and $\alpha \in (-1, 1)$. Then the series on the right-hand side converges absolutely and uniformly in $t \in [0, 1]$, so that f is indeed a well-defined continuous function. If $\varphi(t) = \nu \sin(2\pi t) + \rho \cos(2\pi t)$ for real constants ν and ρ , then f is a Weierstraß function. If $\varphi(t) = \min_{z \in \mathbb{Z}} |z - t|$ is the tent map, then f is a Takagi–van der Waerden function. It was shown in [89] that, under some mild conditions on φ , the function f is of bounded variation for $|\alpha| < 1/b$, whereas for $|\alpha| > 1/b$ and $p := -\log_{|\alpha|} b$ it has nontrivial and linear p^{th} -variation along the sequence

$$\mathbb{T}_n := \{kb^{-n} : k = 0, \dots, b^n\}, \quad n \in \mathbb{N}, \quad (5.2)$$

¹ Recall that a measurable and strictly positive function g , which is defined on some interval $[a, \infty)$, is called regularly varying (at infinity) with index $\varrho \in \mathbb{R}$ if $g(\lambda x)/g(x) \rightarrow \lambda^\varrho$ as $x \uparrow \infty$ for all $\lambda > 0$. If $\varrho = 0$, the function g is also called slowly varying. Throughout this chapter, we will always assume that all regularly varying functions are defined on $[0, \infty)$. This assumption can be made without loss of generality, as one can always consider the function $x \mapsto g(a \vee x)$.

of b -adic partitions of $[0, 1]$. That is, for all $t \in (0, 1]$,

$$\langle f \rangle_t^{(q)} := \lim_{n \uparrow \infty} \sum_{k=0}^{\lfloor tb^n \rfloor} |f((k+1)b^{-n}) - f(kb^{-n})|^q = \begin{cases} 0 & \text{if } q > p, \\ t \cdot \mathbb{E}[|Z|^q] & \text{if } q = p, \\ +\infty & \text{if } q < p. \end{cases} \quad (5.3)$$

Here, Z is a certain random variable, whose law is known in some special cases. For instance, if φ is the tent map and b is even, then the law of $\alpha b Z$ is the infinite Bernoulli convolution with parameter $1/(|\alpha|b)$ (see also [39, 87, 76, 77] for earlier results and [90] for subsequent extensions in this special setup). Clearly, the parameter $p = -\log_{|\alpha|} b$ can be regarded as a measure for the “roughness” of the function f . In particular, if $b = 2$, the function f admits the roughness exponent $R = 1/p = -\log_2 |\alpha|$.

In this section, we study the case of critical roughness, $\alpha = -1/b$ or $\alpha = 1/b$, in which $p = 1$. For this case, it was shown in [89] that $\langle f \rangle_t^{(q)} = 0$ for all $q > 1$ and $t \in [0, 1]$. This, however, does *not* imply that f is of bounded variation. For instance, if φ is the tent map, $b = 2$, and $\alpha = 1/2$, then f is the classical Takagi function, which is nowhere differentiable and hence cannot be of bounded variation; a very short proof of this fact was given by de Rham [24] and later rediscovered by Billingsley [8]. For the Weierstraß function, the proof of nowhere differentiability for all $\alpha \in [1/b, 1)$ is more difficult. Starting from Weierstraß’s original work, it attracted numerous authors until a definite result was given by Hardy [51].

It is therefore apparent that, in the critical case $|\alpha| = 1/b$, power variation $\langle f \rangle^{(q)}$ is insufficient to capture the exact degree of roughness of the function f . To give a precise result on the roughness of the function f in the critical case, we take a strictly increasing function $\Phi : [0, 1) \rightarrow [0, \infty)$ and investigate the limit

$$\langle f \rangle_t^\Phi := \lim_{n \uparrow \infty} \sum_{k=0}^{\lfloor tb^n \rfloor} \Phi(|f((k+1)b^{-n}) - f(kb^{-n})|),$$

which can be regarded as the Wiener–Young Φ -variation of f (see, e.g., [4]), restricted to the sequence of b -adic partitions (5.2). Our main results will show that the correct choice for Φ is the function

$$\Phi(x) = \frac{x}{\sqrt{-\log x}} \quad \text{for } x \in (0, 1) \quad \text{and} \quad \Phi(0) := 0.$$

For the case $b = 2$, to clearly distinguish the notion $\langle f \rangle_n^{(q)} := \sum_{i=0}^{2^n-1} |f((k+1)2^{-n}) - f(k2^{-n})|^q$ used throughout Chapter 2 and Chapter 3 with $\langle f \rangle_t^{(q)}$ as in (5.2), we use the notion $V_n^{(q)} = \sum_{i=0}^{2^n-1} |f((k+1)2^{-n}) - f(k2^{-n})|^q$ exclusively in this chapter.

We fix this function Φ throughout the remainder of this section. Our first result establishes the Φ -variation of f for the class of Takagi–van der Waerden functions.

Theorem 5.1. *Let $\varphi(t) = \min_{z \in \mathbb{Z}} |z - t|$ be the tent map, $b \in \{2, 3, \dots\}$, and $|\alpha| = 1/b$. Then the Φ -variation of the Takagi–van der Waerden function f exists along $\{\mathbb{T}_n\}_{n \in \mathbb{N}}$. If b is even, then it is given by*

$$\langle f \rangle_t^\Phi = t \cdot \sqrt{\frac{2}{\pi \log b}}, \quad t \in [0, 1].$$

If b is odd, then

$$\langle f \rangle_t^\Phi = t \cdot \sqrt{\frac{2(b + \operatorname{sgn}(\alpha))}{\pi(b - \operatorname{sgn}(\alpha)) \log b}}, \quad t \in [0, 1].$$

The following theorem establishes the Φ -variation of Weierstraß functions.

Theorem 5.2. *Suppose $\varphi(t) = \nu \sin(2\pi t) + \rho \cos(2\pi t)$, $b \in \{2, 3, \dots\}$, and $|\alpha| = 1/b$. Then the Φ -variation of the Weierstraß function f exists along $\{\mathbb{T}_n\}_{n \in \mathbb{N}}$ and is given by*

$$\langle f \rangle_t^\Phi = t \cdot 2 \sqrt{\frac{\pi(\nu^2 + \rho^2)}{\log b}}, \quad t \in [0, 1].$$

5.2 Central limit theorem and Rademacher functions associated with fractal functions

The above results are consequences of suitable central limit theorems. For Theorem 5.1, the case of an even b will be settled by the standard central limit theorem, whereas the case of an odd b will require the use of a central limit theorem for Markov chains. To establish the Φ -variation of the critical Weierstraß functions, as stated in Theorem 5.2, we rely on the martingale central limit theorem. The corresponding theory is developed in the following paragraphs, where the proofs of Theorem 5.1 and Theorem 5.2 are also given.

We first consider only the Φ -variation $\langle f \rangle_t^\Phi$ for $t = 1$. The case $t < 1$ will be discussed at the end of this section, simultaneously for both theorems. We fix $b \in \{2, 3, \dots\}$ and $\alpha \in \{-1/b, +1/b\}$. Following the proof of Proposition 2.10, we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting an independent sequence $U_1, U_2, \dots$ of random variables with a uniform distribution on $\{0, 1, \dots, b-1\}$ and define the stochastic process $R_m := \sum_{i=1}^m U_i b^{i-1}$.

Note that R_m has a uniform distribution on $\{0, \dots, b^m - 1\}$. Therefore, for $n \in \mathbb{N}$ such

that all increments $|f((k+1)b^{-n}) - f(kb^{-n})|$ are less than 1,

$$V_n := \sum_{k=0}^{b^n-1} \Phi(|f((k+1)b^{-n}) - f(kb^{-n})|) = b^n \mathbb{E} \left[\Phi(|f((R_n+1)b^{-n}) - f(R_nb^{-n})|) \right]. \quad (5.4)$$

To simplify the expectation on the right-hand side of (5.4), let the n^{th} truncation of f be given by $f_n(t) = \sum_{m=0}^{n-1} \alpha^m \varphi(b^m t)$. The periodicity of φ implies that

$$\begin{aligned} f((R_n+1)b^{-n}) - f(R_nb^{-n}) &= f_n((R_n+1)b^{-n}) - f_n(R_nb^{-n}) \\ &= b^{-n} \operatorname{sgn}(\alpha)^n \sum_{m=1}^n \operatorname{sgn}(\alpha)^m \frac{\varphi((R_n+1)b^{-m}) - \varphi(R_nb^{-m})}{b^{-m}}. \end{aligned}$$

Moreover, the periodicity of φ implies that for $m \leq n$,

$$\varphi(x + R_nb^{-m}) = \varphi\left(x + \sum_{i=1}^n U_i b^{i-1-m}\right) = \varphi\left(x + \sum_{i=1}^m U_i b^{i-1-m}\right) = \varphi(x + R_mb^{-m}).$$

Therefore,

$$\operatorname{sgn}(\alpha)^m \frac{\varphi((R_n+1)b^{-m}) - \varphi(R_nb^{-m})}{b^{-m}} = \operatorname{sgn}(\alpha)^m \frac{\varphi((R_m+1)b^{-m}) - \varphi(R_mb^{-m})}{b^{-m}} =: Y_m.$$

It follows that

$$V_n = b^n \mathbb{E} \left[\Phi\left(b^{-n} \left| \sum_{m=1}^n Y_m \right| \right) \right]. \quad (5.5)$$

Lemma 5.3. *Suppose that $Z_0, Z_1, Z_2, \dots$ is a sequence of random variables with $Z_0 = 0$ and uniformly bounded increments such that the laws of $\frac{1}{\sqrt{n}}Z_n$ converge weakly to some normal distribution $N(0, \sigma^2)$ with $\sigma^2 > 0$ and that the expression $\frac{1}{n}\mathbb{E}[Z_n^2]$ is bounded in n . Then*

$$b^n \mathbb{E} \left[\Phi(b^{-n} |Z_n|) \right] \longrightarrow \sqrt{\frac{2\sigma^2}{\pi \log b}}.$$

Proof. The fact that $\frac{1}{n}\mathbb{E}[Z_n^2]$ is bounded implies together with standard arguments that for every nondegenerate interval $I \subset [0, \infty)$,

$$\lim_{n \uparrow \infty} \mathbb{E} \left[\mathbb{1}_{\{|\frac{1}{\sqrt{n}}Z_n| \in I\}} \left| \frac{1}{\sqrt{n}}Z_n \right| \right] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\{|z| \in I\}} |z| e^{-z^2/(2\sigma^2)} dz. \quad (5.6)$$

We have

$$b^n \mathbb{E} \left[\Phi(b^{-n} |Z_n|) \right] = \mathbb{E} \left[\frac{|Z_n|}{\sqrt{n \log b - \log |Z_n|}} \mathbb{1}_{\{|Z_n| > 0\}} \right].$$

Let C be an almost sure uniform bound for $|Z_{k+1} - Z_k|$. Hence, for all $\beta \in (0, \log b)$ there exists $n_0 \in \mathbb{N}$ such that $n\beta < n \log b - \log(Cn)$ for all $n \geq n_0$. Hence,

$$\sqrt{n \log b - \log |Z_n|} \geq \sqrt{n\beta} \quad \text{for } n \geq n_0, \quad (5.7)$$

and taking $I := (0, \infty)$ in (5.6) gives

$$\limsup_{n \uparrow \infty} b^n \mathbb{E} \left[\Phi(b^{-n} |Z_n|) \right] \leq \frac{1}{\sqrt{2\pi\sigma^2\beta}} \int_{\{|z| \in I\}} |z| e^{-z^2/(2\sigma^2)} dz = \sqrt{\frac{2\sigma^2}{\pi\beta}}.$$

To get a lower bound, observe that for every $\varepsilon > 0$ and $n \geq 1/\varepsilon^2$,

$$\mathbb{1}_{\{|\frac{1}{\sqrt{n}} Z_n| \geq \varepsilon\}} \sqrt{n \log b - \log |Z_n|} \leq \mathbb{1}_{\{|\frac{1}{\sqrt{n}} Z_n| \geq \varepsilon\}} \sqrt{n \log b}.$$

Hence, we get from (5.6) that

$$\liminf_{n \uparrow \infty} b^n \mathbb{E} \left[\Phi(b^{-n} |Z_n|) \right] \geq \frac{1}{\sqrt{2\pi\sigma^2 \log b}} \int_{\{|z| \geq \varepsilon\}} |z| e^{-z^2/(2\sigma^2)} dz.$$

Sending $\varepsilon \downarrow 0$ and $\beta \uparrow \log b$ gives the result. $\square$

Proof of Theorem 5.1 for $t = 1$. For b even, [89, Proposition 3.2 (a)] states that $Y_1, Y_2, \dots$ is an i.i.d. sequence of symmetric $\{-1, +1\}$ -valued Bernoulli random variables. Therefore, (5.5), the classical central limit theorem, and Lemma 5.3 give $V_n \rightarrow \sqrt{2/(\pi \log b)}$. If b is odd, then [89, Proposition 3.2 (b)] states that the random variables $\text{sgn}(\alpha)^m Y_m$ form a time-homogeneous Markov chain on $\{-1, 0, +1\}$ with an initial distribution $\mu_1 = (\frac{b-1}{2b}, \frac{1}{b}, \frac{b-1}{2b})$ and transition matrix P_+ , where

$$P_{\pm} := \frac{1}{2b} \begin{pmatrix} b \pm 1 & 0 & b \mp 1 \\ b - 1 & 2 & b - 1 \\ b \mp 1 & 0 & b \pm 1 \end{pmatrix}.$$

It follows that $Y_1, Y_2, \dots$ also form a time-homogeneous Markov chain with an initial distribution μ_1 and a transition matrix P_+ for $\alpha > 0$ and P_- for $\alpha < 0$. Since 0 is a transient state, we can clearly consider only the restriction of the Markov chain to its positive recurrent states, -1 and $+1$. Let $\bar{P}_{\pm}$ be the 2×2 -matrix obtained from $P_{\pm}$ by deleting the second row and second column from P , and define $\bar{\mu}_1 = (1/2, 1/2)$. Then $\bar{\mu}_1$ is the unique stationary

distribution for $\overline{P}_\pm$. Moreover,

$$\overline{P}_\pm^n = \frac{1}{2} \begin{pmatrix} 1 + (\pm b)^{-n} & 1 - (\pm b)^{-n} \\ 1 - (\pm b)^{-n} & 1 + (\pm b)^{-n} \end{pmatrix}.$$

For the state-constraint Markov chain $\overline{Y}_1, \overline{Y}_2, \dots$ with an initial distribution $\bar{\mu}_1$ and a transition matrix $\overline{P}_\pm$, we thus have $\text{var}(\overline{Y}_1) = 1$ and

$$\text{cov}(\overline{Y}_1, \overline{Y}_{n+1}) = \sum_{y_1, y_{n+1} \in \{-1, +1\}} \bar{\mu}_1(y_1) \overline{P}_\pm^n(y_1, y_{n+1}) y_1 y_{n+1} = (\pm b)^{-n}.$$

Letting

$$\sigma^2 := \text{var}(\overline{Y}_1) + 2 \sum_{n=1}^{\infty} \text{cov}(\overline{Y}_1, \overline{Y}_{n+1}) = \frac{b \pm 1}{b \mp 1},$$

the central limit theorem for Markov chains (see, e.g., [58]) implies that $\frac{1}{\sqrt{n}} \sum_{k=1}^n \overline{Y}_k$ converges in law to $N(0, \sigma^2)$. Due to the stationarity of the Markov chain, we have moreover

$$\begin{aligned} \mathbb{E} \left[\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \overline{Y}_k \right)^2 \right] &= \frac{1}{n} \sum_{k=1}^n \text{var}(\overline{Y}_k) + \frac{2}{n} \sum_{k=1}^{n-1} \sum_{\ell=k+1}^n \text{cov}(\overline{Y}_k, \overline{Y}_\ell) \\ &= 1 + \frac{2}{n} \sum_{k=1}^{n-1} \sum_{\ell=k+1}^n (\pm b)^{k-\ell} \leq 1 + \frac{2}{n} \cdot \frac{b^{1-n} + bn + b - n}{(b-1)^2}, \end{aligned}$$

which is uniformly bounded in n . Therefore, Lemma 5.3 and (5.5) give

$$V_n \rightarrow \sqrt{2(b \pm 1)/(\pi(b \mp 1) \log b)} \quad n \uparrow \infty.$$

This completes the proof. $\square$

Now we prepare for the proof of Theorem 5.2 for $t = 1$. Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n := \sigma(U_1, \dots, U_n)$ for $n \in \mathbb{N}$. Then each Y_n is $\mathcal{F}_n$ -measurable. Since $U_1, \dots, U_n$ can be recovered from R_n , we have $\mathcal{F}_n = \sigma(R_n)$ for $n \geq 1$. We define $Z_0 := 0$ and $Z_n := \sum_{k=1}^n Y_k$ for $n \in \mathbb{N}$.

Lemma 5.4. *If $\varphi(t) = \nu \sin(2\pi t) + \rho \cos(2\pi t)$, then $\{Z_n\}_{n \in \mathbb{N}_0}$ is a martingale with respect to $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$.*

Proof. We must show that $\mathbb{E}[Y_n | R_{n-1}] = 0$ $\mathbb{P}$ -a.s. for $n \geq 1$. To this end, we use the fact

that $R_n = R_{n-1} + U_n b^{n-1}$, where $R_0 := 0$ and U_n is independent of R_{n-1} . Therefore,

$$\begin{aligned}\mathbb{E}[Y_n | R_{n-1} = r] &= (\text{sgn}(\alpha))^n \mathbb{E} \left[\frac{\varphi((r + U_n b^{n-1} + 1)b^{-n}) - \varphi((r + U_n b^{n-1})b^{-n})}{b^{-n}} \right] \\ &= \frac{(\text{sgn}(\alpha)b)^n}{b} \sum_{k=0}^{b-1} \left(\varphi((r+1)b^{-n} + k/b) - \varphi(rb^{-n} + k/b) \right).\end{aligned}\quad (5.8)$$

If $n = 1$, then r must be zero, and the sum in (5.8) is a telescopic sum with value $\varphi(1) - \varphi(0) = 0$. Now consider the case $n \geq 2$. Then, for all $x \in \mathbb{R}$, $i = \sqrt{-1}$, and $\Re$ denoting the real part of a complex number,

$$\sum_{k=0}^{b-1} \varphi(x + k/b) = \Re \left((\rho - i\nu) \sum_{k=0}^{b-1} e^{2\pi i(x+k/b)} \right) = \Re \left((\rho - i\nu) e^{2\pi i x} \cdot \frac{e^{2\pi i b/b} - 1}{e^{2\pi i/b} - 1} \right) = 0.$$

Therefore, the sum in (5.8) vanishes. $\square$

Lemma 5.5. *With δ_x denoting the Dirac measure in $x \in \mathbb{R}$ and λ denoting the Lebesgue measure on $[0, 1]$, we have $\mathbb{P}$ -a.s., $\frac{1}{n} \sum_{k=1}^n \delta_{b^{-k} R_k} \rightarrow \lambda$ weakly as $n \uparrow \infty$.*

Proof. Without loss of generality, we can extend the sequence $\{U_i\}_{i \in \mathbb{N}}$ to a two-sided sequence $\{U_i\}_{i \in \mathbb{Z}}$ of i.i.d. random variables with a uniform distribution on $\{0, \dots, b-1\}$. Then we define $X_n := \sum_{j=1}^{\infty} U_{n+1-j} b^{-j} = \sum_{j=0}^{\infty} U_{n-j} b^{-(j+1)}$ for $n \in \mathbb{Z}$. Each X_n is uniformly distributed on $[0, 1]$, i.e., has law λ . Moreover, in comparison with X_n , the random variable X_{n+1} is obtained by shifting the sequence $\{U_i\}_{i \in \mathbb{Z}}$ one step to the right. It is well-known that the dynamical system corresponding to such a two-sided Bernoulli shift is mixing and hence ergodic (for a proof, see, e.g., Example 20.26 in [61]). By Birkhoff's ergodic theorem, we thus have $\frac{1}{n} \sum_{k=1}^n f(X_k) \rightarrow \int_0^1 f d\lambda$ $\mathbb{P}$ -a.s. for each bounded Borel-measurable function on $[0, 1]$. Since $|b^{-n} R_n - X_n| \leq b^{-n}$, we hence obtain $\frac{1}{n} \sum_{k=1}^n f(b^{-k} R_k) \rightarrow \int_0^1 f d\lambda$ $\mathbb{P}$ -a.s. for each (uniformly) continuous function on $[0, 1]$. Since $C[0, 1]$ is separable, the result follows. $\square$

Proof of Theorem 5.2 for $t = 1$. Let $\langle Z \rangle_n := \sum_{k=1}^n \mathbb{E}[Y_k^2 | \mathcal{F}_{k-1}]$ be the predictable quadratic variation of the martingale $\{Z_n\}_{n \in \mathbb{N}_0}$. We define $\psi_n(x) := (\varphi(x + b^{-n}) - \varphi(x))/b^{-n}$. Then $\psi_n(x) \rightarrow \varphi'(x)$ uniformly in x . By arguing as in (5.8), we see that

$$\mathbb{E}[Y_k^2 | \mathcal{F}_{k-1}] = \frac{1}{b} \sum_{\ell=0}^{b-1} (\psi_k(b^{-k} R_{k-1} + \ell/b))^2.$$

We therefore conclude from Lemma 5.5 that

$$\frac{1}{n}\langle Z \rangle_n = \frac{1}{b} \sum_{\ell=0}^{b-1} \sum_{k=1}^n (\psi_k(b^{-k}R_{k-1} + \ell/b))^2 \longrightarrow \int_0^1 (\varphi'(t))^2 dt = 2\pi^2(\nu^2 + \rho^2) =: \sigma^2.$$

Analogously, one sees easily that $s_n^2 := \mathbb{E}[\langle Z \rangle_n]$ satisfies $\frac{1}{n}s_n^2 \rightarrow \sigma^2$. Since the increments Y_k are uniformly bounded, the Lindeberg condition,

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E}[Y_k^2 \mathbb{1}_{\{Y_k^2 \geq \varepsilon n\}} | \mathcal{F}_{k-1}] \longrightarrow 0 \quad \mathbb{P}\text{-a.s. for all } \varepsilon > 0,$$

is clearly satisfied. Therefore, the martingale central limit theorem in the form of [25, (7.4) in Chapter 7] yields that the laws of $\frac{1}{\sqrt{n}}Z_n$ converge weakly to $N(0, \sigma^2)$. Lemma 5.3 hence gives

$$V_n \longrightarrow 2\sqrt{\frac{\pi(\nu^2 + \rho^2)}{\log b}}. \quad \square$$

Finally, we show how the preceding results can be extended to the case $0 \leq t < 1$. Writing Z_n for $\sum_{k=1}^n Y_k$, the Φ -variation over the interval $[0, t]$ is equal to

$$\begin{aligned} V_{n,t} &:= \sum_{k=0}^{b^n-1} \Phi(|f((k+1)b^{-n}) - f(kb^{-n})|) \mathbb{1}_{[0,t]}(kb^{-n}) \\ &= b^n \mathbb{E} \left[\Phi(|f((R_n+1)b^{-n}) - f(R_nb^{-n})|) \mathbb{1}_{\{b^{-n}R_n \leq t\}} \right] \\ &= b^n \mathbb{E} \left[\Phi \left(b^{-n} \left| \sum_{m=1}^n Y_m \right| \right) \mathbb{1}_{\{b^{-n}R_n \leq t\}} \right] = \mathbb{E} \left[\frac{|Z_n|}{\sqrt{n \log b - \log |Z_n|}} \mathbb{1}_{\{|Z_n| > 0\}} \mathbb{1}_{\{b^{-n}R_n \leq t\}} \right]. \end{aligned}$$

Let $\delta > 0$ be given and pick $m \in \mathbb{N}$ such that $b^{-m} \leq \delta$. Clearly, $\{b^{-n}R_n \leq t\} \subset \{b^{-n}R_{m,n} \leq t\}$, where $R_{m,n} := R_n - R_{n-m} = \sum_{k=n-m+1}^n U_k b^{k-1}$. In addition, we argue as in the proof of Lemma 5.3 and take $\beta \in (0, \log b)$ and $n_0 \in \mathbb{N}$ such that $n\beta < n \log b - \log(Cn)$ for all $n \geq n_0$ and (5.7) holds. Therefore, for $n \geq m \vee n_0$.

$$V_{n,t} \leq \frac{1}{\sqrt{n\beta}} \mathbb{E}[|Z_n| \mathbb{1}_{\{b^{-n}R_{m,n} \leq t\}}] \leq \frac{1}{\sqrt{n\beta}} \mathbb{E}[|Z_{n-m}| \mathbb{1}_{\{b^{-n}R_{m,n} \leq t\}}] + \frac{1}{\sqrt{n\beta}} \mathbb{E} \left[\left| \sum_{k=n-m+1}^n Y_k \right| \right].$$

Clearly, the rightmost term converges to zero as $n \uparrow \infty$. Moreover, Z_{n-m} and $R_{m,n}$ are independent, and so

$$\limsup_{n \uparrow \infty} V_{n,t} \leq \sqrt{\frac{2\sigma^2}{\pi\beta}} \limsup_{n \uparrow \infty} \mathbb{P}[b^{-n}R_{m,n} \leq t] \leq \sqrt{\frac{2\sigma^2}{\pi\beta}} \limsup_{n \uparrow \infty} \mathbb{P}[b^{-n}R_n \leq t+\delta] = \sqrt{\frac{2\sigma^2}{\pi\beta}}(t+\delta),$$

where the second inequality follows from the fact that $b^{-n}R_{m,n} \geq b^{-n}R_n - \delta$ for $n > m$. Sending $\beta \uparrow \log b$ and $\delta \downarrow 0$ gives the desired upper bound.

To get a corresponding lower bound, we choose $\delta > 0$ and m as in the upper bound. In addition, we choose $\varepsilon > 0$. For $n > m \vee 1/\varepsilon^2$, we then get as in the proof of Lemma 5.3,

$$V_{n,t} \geq \mathbb{E} \left[\frac{|Z_n|}{\sqrt{n \log b}} \mathbb{1}_{\{|\frac{1}{\sqrt{n}}|Z_n| \geq \varepsilon\}} \mathbb{1}_{\{b^{-n}R_n \leq t\}} \right] \geq \mathbb{E} \left[\frac{|Z_n|}{\sqrt{n \log b}} \mathbb{1}_{\{|\frac{1}{\sqrt{n}}|Z_n| \geq \varepsilon\}} \mathbb{1}_{\{b^{-n}R_{m,n} \leq t-\delta\}} \right].$$

Now let C be a uniform upper bound for $|Y_k|$ and choose n_1 such that $mC \leq \varepsilon\sqrt{n_1}$. Then, for $n \geq n_1 \vee m \vee 1/\varepsilon^2$,

$$V_{n,t} \geq \mathbb{E} \left[\frac{|Z_{n-m}|}{\sqrt{n \log b}} \mathbb{1}_{\{|\frac{1}{\sqrt{n}}|Z_{n-m}| \geq 2\varepsilon\}} \mathbb{1}_{\{b^{-n}R_{m,n} \leq t-\delta\}} \right] - \frac{1}{\sqrt{n \log b}} \mathbb{E} \left[\left| \sum_{k=n-m+1}^n Y_k \right| \right].$$

Again, the second expectation on the right converges to zero. Using as before, the independence of Z_{n-m} and $R_{m,n}$ now easily gives the desired lower bound. This concludes the proof of Theorems 5.1 and 5.2 for $0 \leq t < 1$.

5.3 A limit theorem for Bernoulli-type convolutions

From the section, we study the Φ -variation of functions in the so-called Takagi class, which was introduced by Hata and Yamaguti [52] and motivated by Takagi's [92] celebrated example of a continuous nowhere differentiable function. This class was chosen due to its richness and tractability and also because it is closely related to Wiener–Lévy construction of the Brownian motion. The Takagi class consist of all function $f : [0, 1] \rightarrow \mathbb{R}$ that admit an expansion of the form

$$f(t) := \sum_{m=0}^{\infty} \alpha_m \varphi(2^m t), \quad t \in [0, 1],$$

where $\varphi(t) = \min_{z \in \mathbb{Z}} |z - t|$ is the tent map, and (α_m) is a sequence of real numbers for which the series $\sum_{m=0}^{\infty} \alpha_m$ converges absolutely. The special choice $\alpha_m = \alpha^m$ for some $\alpha \in (-1, 1)$ yields the class of so-called Takagi–Landsberg functions, which, for $1/2 < |\alpha| < 1$, have finite nontrivial and linear p^{th} variation for $p = -\log_{|\alpha|} 2$ as shown in [89]. Thus, it admits the roughness exponent $R = -\log_2 |\alpha|$. For $|\alpha| < 1/2$, it is well known that f is of finite total variation and admits the roughness exponent $R = 1$. The borderline case $\alpha = 1/2$ corresponds to the classical Takagi function, which is nowhere differentiable and hence not of finite total variation, even though $p = -\log_{1/2} 2 = 1$ and $R = 1$. Instead, as shown in

Section 5.2, for $|\alpha| = 1/2$, the function f has linear Φ -variation along the dyadic partitions

$$\langle f \rangle_t^\Phi := \lim_{n \uparrow \infty} \sum_{k=0}^{\lfloor t2^n \rfloor} \Phi(|f((k+1)2^{-n}) - f(k2^{-n})|) = \sqrt{\frac{2}{\pi}} \cdot t, \quad 0 \leq t \leq 1, \quad (5.9)$$

where

$$\Phi(x) := \frac{x}{\sqrt{-\log_2 x}}, \quad 0 < x < 1. \quad (5.10)$$

The theme starting from this section is to locate those functions in the Takagi class that have linear Φ -variation for a prescribed function Φ of the form

$$\Phi_q(x) := x^{\frac{1}{1-q}} g\left(\frac{-\log_2 x}{1-q}\right)^{-\frac{1}{2(1-q)}}, \quad 0 < x < 1, \quad (5.11)$$

where $q \in [0, 1)$ and $g : [0, \infty) \rightarrow (0, \infty)$ is any regularly varying function. For instance, the function Φ in (5.10) is obtained by taking $q = 0$ and $g(x) = x$. In this way, we obtain natural trajectories with linear Wiener–Young Φ -variation for a large class of given functions Φ .

As mentioned earlier in this chapter, our approach to the Φ_q -variation of functions f in the Takagi class depends on the parameter q in (5.11). For $q = 0$, the argument extends the approach via the central limit theorem for the Rademacher functions in the Faber–Schauder development of f as presented in Section 5.2. For $q \in (0, 1)$, however, a different limit theorem is needed, which might be of independent interest.

For some $p \in [1, \infty]$, let $(Y_m)_{m \in \mathbb{N}}$ be a sequence of random variables in $L^p := L^p(\Omega, \mathcal{F}, \mathbb{P})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a given probability space. We assume for simplicity that the L^p -norms of the Y_m are uniformly bounded. Let furthermore $\{\beta_m\}_{m \in \mathbb{N}_0}$ be a sequence of real numbers and consider the random variables

$$Z_n := \sum_{m=1}^n \beta_{n-m} Y_m, \quad n \in \mathbb{N}.$$

We are first concerned with the limit of

$$\frac{1}{s_n} Z_n \quad \text{where} \quad s_n^2 := \sum_{m=0}^{n-1} \beta_m^2$$

as $n \uparrow \infty$.

Theorem 5.6. *Suppose that $\beta_m \geq 0$ for all m and that there exist $q > 0$, $b > 1$, and a slowly varying function ℓ such that*

$$\lim_{n \rightarrow \infty} \frac{s_n^2}{2^{2qn} \ell(b^n)} = 1. \quad (5.12)$$

Then as $n \uparrow \infty$,

$$\frac{1}{s_n} Z_n \longrightarrow Z := \sqrt{2^{2q} - 1} \sum_{m=1}^{\infty} 2^{-qm} Y_m \quad \text{in } L^p.$$

Proof of Theorem 5.6. For $\delta \in (0, \delta_0]$, denote

$$\lambda := \sqrt{2^{2q} - 1}, \quad \lambda_{\delta}^+ := \sqrt{2^{2q} \frac{(1+\delta)^2}{1-\delta} - 1}, \quad \text{and} \quad \lambda_{\delta}^- := \sqrt{2^{2q} \frac{(1-\delta)^2}{1+\delta} - 1},$$

where $\delta_0 \in (0, \min\{1, q\})$ is chosen such that the arguments of the square roots are always positive. Thus,

$$2^{2q}(1-\delta)^2 - (1+\delta) > 0 \quad \text{for all } \delta \in (0, \delta_0]. \quad (5.13)$$

Let $\varepsilon > 0$ be given. We first choose $K \in \mathbb{N}$ such that

$$\lambda \sum_{m=K+1}^{\infty} 2^{-(q-\delta_0)m} < \varepsilon. \quad (5.14)$$

Then we choose $\delta_1 \in (0, \delta_0]$ such that

$$1 - 2^{-\delta m} < \varepsilon \quad \text{for } m = 0, \dots, K \text{ and all } \delta \in (0, \delta_1]. \quad (5.15)$$

Finally, we choose $\delta \in (0, \delta_1]$ such that

$$\left| \frac{\lambda_{\delta}^{\pm}}{\sqrt{1 \mp \delta}} - \lambda \right| < \varepsilon. \quad (5.16)$$

Given this δ , we use (5.12) to pick $n_0 \in \mathbb{N}$ such that

$$\ell(b^n) 2^{2qn} (1-\delta) < s_n^2 < \ell(b^n) 2^{2qn} (1+\delta) \quad \text{for all } n > n_0. \quad (5.17)$$

Hence,

$$\begin{aligned} 2^{2q(m+1)} ((1-\delta)\ell(b^{m+1}) - (1+\delta)\ell(b^m)2^{-2q}) &< \beta_m^2 \\ &< 2^{2q(m+1)} ((1+\delta)\ell(b^{m+1}) - (1-\delta)\ell(b^m)2^{-2q}). \end{aligned} \quad (5.18)$$

Since ℓ is slowly varying, we can choose $n_1 \geq n_0$ such that

$$1 - \delta < \frac{\ell(b \cdot b^m)}{\ell(b^m)} < 1 + \delta \quad \text{for all } m \geq n_1.$$

Combining this with (5.13), we see that the left-hand side of (5.18) is strictly positive for

$m \geq n_1$. Hence, we may take square roots to get

$$\begin{aligned} 2^{q(m+1)} \sqrt{(1-\delta)\ell(b^{m+1}) - (1+\delta)\ell(b^m)2^{-2q}} &< \beta_m \\ &< 2^{q(m+1)} \sqrt{(1+\delta)\ell(b^{m+1}) - (1-\delta)\ell(b^m)2^{-2q}} \end{aligned} \quad (5.19)$$

for all $m \geq n_1$. Moreover, with (5.17) we get that for $n > m \geq n_1$,

$$\begin{aligned} \frac{\beta_m}{s_n} &< \frac{2^{q(m+1)} \sqrt{(1+\delta)\ell(b^{m+1}) - (1-\delta)\ell(b^m)2^{-2q}}}{\sqrt{(1-\delta)2^{2qn}\ell(b^n)}} \\ &= 2^{q(m-n)} \sqrt{\frac{\ell(b^m)}{\ell(b^n)} \left(2^{2q} \frac{1+\delta}{1-\delta} \frac{\ell(b^{m+1})}{\ell(b^m)} - 1 \right)} < 2^{q(m-n)} \sqrt{\frac{\ell(b^m)}{\ell(b^n)}} \lambda_\delta^+. \end{aligned}$$

Similarly, we conclude a corresponding lower bound and obtain that for $n > m \geq n_1$,

$$2^{-q(n-m)} \sqrt{\frac{\ell(b^m)}{\ell(b^n)}} \lambda_\delta^- < \frac{\beta_m}{s_n} < 2^{-q(n-m)} \sqrt{\frac{\ell(b^m)}{\ell(b^n)}} \lambda_\delta^+. \quad (5.20)$$

The Potter bounds (see, e.g., Theorem 1.5.6 in [9]) yield that there is $n_2 \geq n_1$ such that for $n \geq m \geq n_2$,

$$(1-\delta)2^{-2\delta(n-m)} = (1-\delta)b^{-2\delta(n-m)\log_b 2} \leq \frac{\ell(b^n)}{\ell(b^m)} \leq (1+\delta)b^{2\delta(n-m)\log_b 2} = (1+\delta)2^{2\delta(n-m)}.$$

With (5.20), we obtain that for $n \geq n_2 + m$,

$$2^{-(q+\delta)m} \frac{\lambda_\delta^-}{\sqrt{1+\delta}} < \frac{\beta_{n-m}}{s_n} < 2^{-(q-\delta)m} \frac{\lambda_\delta^+}{\sqrt{1-\delta}}. \quad (5.21)$$

Note that

$$Z^{(n)} := \lambda \sum_{m=1}^n 2^{-qm} Y_m \longrightarrow Z \quad \text{in } L^p.$$

By Minkowski's inequality and with $c := \sup_m \|Y_m\|_p < \infty$,

$$\begin{aligned} \left\| \frac{Z_n}{s_n} - Z^{(n)} \right\|_p &\leq \sum_{m=1}^n \left| \frac{\beta_{n-m}}{s_n} - \lambda 2^{-qm} \right| \|Y_m\|_p \\ &\leq c \sum_{m=1}^{n-n_2} \left| \frac{\beta_{n-m}}{s_n} - \lambda 2^{-qm} \right| + c \sum_{m=n-n_2+1}^n \left| \frac{\beta_{n-m}}{s_n} - \lambda 2^{-qm} \right|. \end{aligned}$$

To deal with the first sum on the right-hand side, we use (5.16) and (5.21) to get that for

$$m \leq n - n_2,$$

$$\begin{aligned} \frac{\beta_{n-m}}{s_n} - \lambda 2^{-qm} &\leq 2^{-(q-\delta)m} \left(\frac{\lambda_\delta^+}{\sqrt{1-\delta}} - \lambda \right) + \lambda(2^{-(q-\delta)m} - 2^{-qm}) \\ &\leq \varepsilon 2^{-(q-\delta)m} + \lambda 2^{-(q-\delta)m} (1 - 2^{-\delta m}). \end{aligned}$$

In the same way, we get the lower bound

$$-\varepsilon 2^{-(q+\delta)m} - \lambda 2^{-qm} (1 - 2^{-\delta m}) \leq \frac{\beta_{n-m}}{s_n} - \lambda 2^{-qm}.$$

Using (5.15) and (5.14), we thus get the following estimate for the first sum on the right-hand side of (5.3)

$$\sum_{m=1}^{n-n_2} \left| \frac{\beta_{n-m}}{s_n} - \lambda 2^{-qm} \right| \leq \varepsilon \sum_{m=1}^{n-n_2} 2^{-(q-\delta)m} + \lambda \sum_{m=1}^K 2^{-(q-\delta)m} \varepsilon + \lambda \sum_{m=K+1}^{\infty} 2^{-(q-\delta)m} \leq \varepsilon c',$$

where $c' := 2 \sum_{m=0}^{\infty} 2^{-(q-\delta_0)m} + 1$. For the right-most sum in (5.3), we have

$$\begin{aligned} \sum_{m=n-n_2+1}^n \left| \frac{\beta_{n-m}}{s_n} - \lambda 2^{-qm} \right| &\leq \sum_{m=n-n_2+1}^n \frac{\beta_{n-m}}{s_n} + \lambda \sum_{m=n-n_2+1}^n 2^{-qm} \\ &\leq \frac{n_2}{s_n} \max_{i=0, \dots, n_2-1} \beta_i + \lambda 2^{-qn} \frac{2^{qn_2} - 1}{2^q - 1}. \end{aligned} \tag{5.22}$$

We will show below that our assumptions imply that $s_n \rightarrow \infty$. Therefore, the right-hand side of (5.22) can be made smaller than ε by choosing n sufficiently large. Thus, we see that the L^p -distance between Z_n/s_n and $Z^{(n)}$ is smaller than $c(c' + 1)\varepsilon$ if n is sufficiently large.

Finally, we argue that $s_n \rightarrow \infty$ if (5.12) holds and $q > 0$. To this end, let $g(x) := x^{2q}\ell(x)$, so that the denominator in (5.12) is equal to $g(2^n)$. Proposition 1.3.6 in [9] states that $g(x) \rightarrow \infty$, which in view of (5.12) yields our claim. $\square$

Now we briefly discuss the case where $q = 0$ in (5.12). In that case, we are typically in the regime of the central limit theorem. This is suggested, for instance, by [97], where a central limit theorem for reversible Markov chains was obtained under the condition that the variance, σ_n^2 , satisfies $\sigma_n^2 = n\ell(n)$ for a slowly varying function ℓ . Since $g(x) := x\ell(x)$ is regularly varying with index $\varrho = 1$, we see that, in our situation, this condition is a special case of condition (5.26) with $q = 0$. In the following proposition, we formulate implications between stronger and weaker versions of condition (5.12) with $q = 0$.

Now, we will formulate a simple central limit theorem in the form needed in the next section. We believe that this result is well known but were unable to find an exact reference

to the literature. Due to the large number of existing variants of the central limit theorem, we confine ourselves here to the basic case that will actually be needed in Section 5.5.

Lemma 5.7. *Suppose that $Y_1, Y_2, \dots$ is an i.i.d. sequence of symmetric $\{-1, +1\}$ -valued Bernoulli random variables such that $s_{n-1}^2/s_n^2 \rightarrow 1$ as $n \uparrow \infty$. Then, for every $p \in [1, \infty)$, the laws of $\frac{1}{s_n}Z_n$ converge in the L^p -Wasserstein metric to the standard normal distribution.*

Proof. As $s_{n-1}^2/s_n^2 \rightarrow 1$, it then follows that

$$\lim_{n \uparrow \infty} \frac{\beta_n^2}{s_n^2} = \lim_{n \uparrow \infty} \left(1 - \frac{s_{n-1}^2}{s_n^2}\right) = 0.$$

It is hence easy to see that the sequence $\{Z_n\}_{n \in \mathbb{N}_0}$ satisfies Lindeberg's condition, which yields that $\frac{1}{s_n}Z_n$ converges in law to $N(0, 1)$. Thus, to conclude convergence in the L^p -Wasserstein metric, it is sufficient to show that $\mathbb{E}[|s_n^{-1}Z_n|^q]$ is uniformly bounded in n for some $q > p$; this follows, e.g., from Lemma 5.61 and Corollary A.50 in [34]. We actually show a stronger condition, namely that $\mathbb{E}[e^{|\lambda Z_n/s_n|}]$ is uniformly bounded for all $\lambda \in \mathbb{R}$. Since $e^{|x|} \leq e^x + e^{-x}$, we can drop the absolute value in the exponent. Then,

$$\log \mathbb{E}[e^{\lambda Z_n/s_n}] = \log \prod_{k=0}^n \cosh(\lambda \beta_k/s_n) = \sum_{k=0}^n \log \cosh(\lambda \beta_k/s_n) \leq \sum_{k=0}^n \frac{\lambda^2 \beta_k^2}{2s_n^2} = \frac{\lambda^2}{2}.$$

This completes the proof. □

5.4 Φ -variation of functions in the Takagi class

In this section, we will apply the limit theorems from Section 5.3 to the problem of computing the Φ -variation of functions in the Takagi class. To this end, let $\Phi : [0, 1) \rightarrow [0, \infty)$ be an increasing function with $\Phi(0) = 0$. The Φ -variation of a continuous function $f : [0, 1] \rightarrow \mathbb{R}$ along the sequence of dyadic partitions is defined as

$$\langle f \rangle_t^\Phi := \lim_{n \uparrow \infty} \sum_{k=0}^{\lfloor t2^n \rfloor} \Phi(|f((k+1)2^{-n}) - f(k2^{-n})|), \quad 0 \leq t \leq 1,$$

provided that all limits exist. For $\Phi(x) = x$ and $t = 1$, we obtain the total variation of f . For $\Phi(x) = x^2$, we obtain the usual quadratic variation, and for $\Phi(x) = x^p$ with $p \geq 1$ the p^{th} -variation. Here, we are interested in functions for which the correct Φ may *not* be of the form $\Phi(x) = x^p$ but rather has a more complex structure.

Let $\varphi(t) = \min_{z \in \mathbb{Z}} |z - t|$ denote the tent map. The Takagi class, as introduced by Hata and Yamaguti [52], consists of all function $f : [0, 1] \rightarrow \mathbb{R}$ that admit an expansion of the form

$$f(t) := \sum_{m=0}^{\infty} \alpha_m \varphi(2^m t), \quad t \in [0, 1], \quad (5.23)$$

where $\{\alpha_m\}$ is a sequence of real numbers for which the series $\sum_{m=0}^{\infty} \alpha_m$ converges absolutely. Clearly, the series on the right-hand side of (5.23) converges uniformly in $t \in [0, 1]$, so that f is indeed a well-defined continuous function. Typically, the functions in the Takagi class have a fractal structure. For instance, the choice $\alpha_m = 2^{-m}$ corresponds to the classical, nowhere differentiable Takagi function, which was originally introduced in [92] but rediscovered many times. More generally, the choice $\alpha_m = a^m$ for some $a \in (-1, 1)$ yields the class of so-called Takagi–Landsberg functions. Analytical properties of the functions in the Takagi class, such as differentiability, are discussed in [62].

Now we take f as in (5.23) as given. The following result bridges the roughness exponent of functions in the Takagi class with the results in Chapter 2. Rewriting (5.23) yields the following representation of the function f using the Faber–Schauder functions,

$$f(t) = \sum_{m=0}^{\infty} 2^{m/2} \alpha_m \sum_{k=0}^{2^m-1} e_{m,k}(t), \quad t \in [0, 1].$$

In other words, we have $\theta_{n,k} = 2^{n/2} \alpha_n$ for $0 \leq k \leq 2^n - 1$. In this case, the quantity s_n , as in (2.7), takes the following form

$$s_n = \sqrt{\sum_{m=0}^{n-1} \sum_{k=0}^{2^m-1} \theta_{m,k}^2} = \sqrt{\sum_{m=0}^{n-1} \alpha_m^2 4^m}, \quad n \in \mathbb{N}. \quad (5.24)$$

As a result, the Gladyshev estimate $\hat{R}_n(f) = 1 - \frac{1}{n} \log_2 s_n = 1 - \frac{1}{n} \log_2 \sqrt{\sum_{m=0}^{n-1} \alpha_m^2 4^m}$. By verifying conditions in Proposition 2.15, it is clear that f satisfies the reverse Jensen condition, and thus, the Gladyshev estimator is a consistent estimator for the roughness exponent for the functions in the Takagi class. That is, suppose that $\hat{R}_n(f) \rightarrow R$ as $n \uparrow \infty$, then we have

$$\lim_{n \uparrow \infty} V_n^{(p)} = \lim_{n \uparrow \infty} \sum_{k=0}^{2^n-1} |f((k+1)2^{-n}) - f(k2^{-n})|^p = \begin{cases} 0 & p > \frac{1}{R}, \\ \infty & p < \frac{1}{R}. \end{cases}$$

In the manner of Remark 2.6, the following proposition extends the above result and provides

a priori bounds on $V_n^{(p)}$, based on the range of possible growth rates of s_n .

Proposition 5.8. *Let*

$$q_* := \liminf_{n \uparrow \infty} \frac{1}{n} \log_2 s_n \quad \text{and} \quad q^* := \limsup_{n \uparrow \infty} \frac{1}{n} \log_2 s_n.$$

Then, for $p \geq 1$, we have $\liminf_n V_n^{(p)} = 0$ if $p(1 - q_*) > 1$ and $\limsup_n V_n^{(p)} = \infty$ if $p(1 - q^*) < 1$. In particular, if $q^* = q_*$, then the function f admits the roughness exponent $R = 1 - q$.

Example 5.9. Faber [30] considered the function f with

$$\alpha_m = \begin{cases} \frac{1}{10^n} & \text{if } m = n! \text{ for some } n \in \mathbb{N}, \\ 0 & \text{otherwise,} \end{cases}$$

and showed that it is nowhere differentiable. It is clear that this choice leads to $q^* \geq 1$ and $q_* \leq 0$, which implies that $\liminf_n V_n^{(p)} = 0$ and $\limsup_n V_n^{(p)} = \infty$ for all $p \geq 1$.

In the sequel, we shall, in particular, investigate the situation in which $\lim_n V_n^{(p)}$ exists but is either zero or infinity. In this case, p^{th} -variation or the roughness exponent is clearly no longer sufficient for characterizing the exact “roughness” of f .

To state our first main result, we recall that Corollary 1.4.2 in [9] allows us to assume without loss of generality that all regularly varying functions are bounded on compact intervals (see the footnote on page 155 for the definition of a regularly varying function). If g is regularly varying and $q \in [0, 1)$, consider

$$\Phi_q(x) = \Phi_{q,g}(x) := x^{\frac{1}{1-q}} g\left(\frac{-\log_2 x}{1-q}\right)^{-\frac{1}{2(1-q)}} \quad (5.25)$$

for $0 < x < 1$. By Proposition 1.5.7 in [9], $y \mapsto g((\log_2 y)/(1 - q))$ is slowly varying, and so $\psi_q(y) := \Phi_q(1/y)$ is regularly varying with index $-1/(1 - q)$. We conclude that $\Phi_q(x) \rightarrow 0$ as $x \downarrow 0$. It therefore makes sense to extend Φ_q to the domain $[0, 1)$ by setting $\Phi_q(0) := 0$. Note also that Φ_q is *slowly varying at zero* with index $1/(1 - q)$ and that the functions arising in this manner form a very wide class of functions that are regularly varying at zero.

Theorem 5.10. *For f as in (5.23), let s_n^2 be as in (5.24) and suppose that there exist $q \in [0, 1)$ and a regularly varying function g such that*

$$\lim_{n \uparrow \infty} \frac{s_n^2}{2^{2qn} g(n)} = 1. \quad (5.26)$$

Let moreover $\{Y_m\}_{m \in \mathbb{N}}$ be an i.i.d. sequence of symmetric $\{-1, +1\}$ -valued Bernoulli random variables and Φ_q as in (5.25). Then the following hold:

- (a) The function f is of bounded variation if and only if $\lim_n s_n^2 < \infty$ and, in this case, the total variation of f is equal to $\mathbb{E}[|\tilde{Z}|]$, where $\tilde{Z} := \sum_{m=0}^{\infty} \alpha_m 2^m Y_{m+1}$.
- (b) If $s_n^2 \rightarrow \infty$ and $q = 0$, then the Φ_0 -variation of f is given by

$$\langle f \rangle_t^{\Phi_0} = \sqrt{\frac{2}{\pi}} \cdot t. \quad (5.27)$$

- (c) If $q > 0$, then the Φ_q -variation of f is given by

$$\langle f \rangle_t^{\Phi_q} = \mathbb{E}[|Z|^{\frac{1}{1-q}}] \cdot t \quad \text{where} \quad Z := \sqrt{2^{2q} - 1} \sum_{m=1}^{\infty} 2^{-qm} Y_m. \quad (5.28)$$

Remark 5.11. The random variable Z in (5.28) can be represented as

$$Z = \sqrt{1 - 2^{-2q}} \sum_{m=0}^{\infty} 2^{-qm} Y_{m+1}$$

and hence has as its law a scaled version of the infinite Bernoulli convolution with parameter 2^{-q} . These laws exhibit some fascinating properties and have been studied in their own right for many decades; see, e.g., [82] and the references therein.

In the next two results, we consider again the situation of Proposition 5.8 and take a closer look at what happens to the power variation of a function f in the Takagi class satisfying (5.26).

Corollary 5.12. Suppose that (5.26) holds with $q \in [0, 1)$ and let $p := 1/(1 - q)$.

- (a) The function f has infinite r^{th} variation for $1 \leq r < p$ and vanishing r^{th} variation for $r > p$.
- (b) If $\lim_n g(n) = c$ for some number $c \in [0, \infty]$, then the p^{th} -variation of f satisfies $\langle f \rangle_t^{(p)} = c^{p/2} \langle f \rangle_t^{\Phi_q}$ for $t \in [0, 1]$.

If g is regularly varying with index $\varrho \in \mathbb{R}$, then $g(x) \rightarrow 0$ for $\varrho < 0$ and $g(x) \rightarrow \infty$ for $\varrho > 0$ according to Proposition B.1.9 in [23]. If $\varrho = 0$, then g is slowly varying and may or may not converge to a number $c \in [0, \infty]$. In conjunction with Corollary 5.12, we thus get immediately the following corollary.

Corollary 5.13. *Suppose that (5.26) holds with $q \in [0, 1)$ and a function g that is regularly varying with index $\varrho \in \mathbb{R}$. We let $p := 1/(1 - q)$.*

(a) *If $\varrho < 0$, then f has vanishing p^{th} -variation.*

(b) *If $\varrho > 0$, then f has infinite p^{th} -variation.*

Example 5.14. Let us consider the classical Takagi function,

$$f(t) = \sum_{m=0}^{\infty} 2^{-m} \varphi(2^m t),$$

as discussed previously. Then $s_n^2 = n$ and so we can take $q = 0$ and $g(x) := x$. Corollary 5.13 thus immediately yields the well-known fact that f is of unbounded total variation. Moreover, we have $\Phi_0(t) = x/\sqrt{-\log_2 x}$, and so we recover here a special case of Theorem 5.1.

Allaart [1] introduced an extension of the Takagi class by multiplying each replicate of the tent map in the expansion (5.23) with an arbitrary sign. More precisely, he considered the class of functions of the form

$$f(t) := \sum_{m=0}^{\infty} \alpha_m \sigma_m(t) \varphi(2^m t), \quad t \in [0, 1], \quad (5.29)$$

where $\sigma_m : [0, 1] \rightarrow \{-1, +1\}$ is constant on each interval $[(k-1)2^{-m}, k2^{-m})$ for $k = 1, \dots, 2^m$, and $\{\alpha_m\}_{m \in \mathbb{N}_0}$ is a sequence of real numbers for which the series $\sum_m \alpha_m$ converges absolutely. Then f is well-defined and continuous due to the uniform convergence of the series in (5.29). Certain deterministic choices for σ_m yield fractal functions studied in other fields such as information theory; see [1, 2] and the references therein. When the signs of σ_m are chosen in a random manner, the expansion (5.29) becomes closely related to the Wiener–Lévy–Ciesielski expansion of Brownian motion by means of the Faber–Schauder functions, and the functions (5.29) with fixed coefficients $\{\alpha_m\}_{m \in \mathbb{N}_0}$ but random signs form a non-Gaussian stochastic process with rough sample paths. More precisely, for each m , let $\{\xi_{m,k}\}_{k=1, \dots, 2^m}$ be a sequence of $\{-1, +1\}$ -valued Bernoulli random variables and define

$$\sigma_m(t) := \sum_{k=1}^{2^m} \xi_{m,k} \mathbb{1}_{[(k-1)2^{-m}, k2^{-m})}(t) \quad \text{and} \quad X_t := \sum_{m=0}^{\infty} \alpha_m \sigma_m(t) \varphi(2^m t). \quad (5.30)$$

Then the following corollary states that $\{X_t\}_{t \in [0, 1]}$ is a stochastic process whose sample paths have the deterministic linear Φ_q -variation (5.28) if the sequence $\{\alpha_m\}_{m \in \mathbb{N}_0}$ satisfies the conditions stated in Theorem 5.10. An illustration of this stochastic process is provided

in Figure 5.1. For the choice $\alpha_m = a^m$ with $a \in (-1, 1)$, the p^{th} -variation of the functions (5.29) was studied in [77, 87].

Corollary 5.15. *The statements of Theorem 5.10 and of the Corollaries 5.12 and 5.13 remain fully valid for the functions of the form (5.29).*

When the goal is to utilize Theorem 5.10 for the construction of functions with a prescribed Φ_q -variation for a given function g , it can be inconvenient that the condition (5.26) is formulated in terms of the sums $s_n^2 = \sum_{m=0}^{n-1} \alpha_m^2 4^m$ rather than in terms of the coefficients α_m themselves. To deal with this issue, we are going to formulate Proposition 5.17, which will also be used in the proofs of our main results.

Remark 5.16. Let us recall the following facts on slowly and regularly varying functions so as to put the statement of the following Proposition 5.17 into the context of Theorem 5.10.

- (a) If g is a regularly varying function and $b > 1$, then $\ell(x) := g(\log_b x)$ is slowly varying according to Proposition 1.5.7 (ii) in [9]. Moreover, our condition (5.26) implies that

$$\lim_{n \rightarrow \infty} \frac{s_n^2}{2^{2qn} \ell(b^n)} = 1.$$

- (b) Recalling that we assume all regularly (and hence slowly) varying functions to be locally bounded, and thus locally integrable, we may define

$$\ell(x) := \frac{1}{\log b} \int_1^x \frac{1}{t} L(t) dt = \int_0^{\log_b x} L(b^s) ds \quad (5.31)$$

for a given slowly varying function L . Then, according to Proposition 1.5.9 (a) in [9], the function ℓ is slowly varying and satisfies $\ell(x)/L(x) \rightarrow +\infty$ as $x \uparrow \infty$.

Proposition 5.17. *Let $\{\beta_m\}_{m \in \mathbb{N}_0}$ be a sequence of real numbers, $s_n^2 = \sum_{m=0}^{n-1} \beta_m^2$, and $b > 1$.*

- (a) *Suppose that $s_n \rightarrow \infty$ as $n \uparrow \infty$, let L be a slowly varying function, define ℓ as in (5.31), and consider the following two conditions:*

- (i) $\beta_n^2 / L(b^n) \rightarrow 1$ as $n \uparrow \infty$.
- (ii) $s_n^2 / \ell(b^n) \rightarrow 1$ as $n \uparrow \infty$.

Then (i) $\Rightarrow$ (ii).

- (b) *If ℓ is a slowly varying function and $q > 0$, the following are equivalent:*

- (iii) $\beta_n^2/2^{2qn}\ell(b^n) \rightarrow 1$ as $n \uparrow \infty$.
- (iv) $s_n^2/2^{2qn}\ell(b^n) \rightarrow (2^{2q} - 1)^{-1}$ as $n \uparrow \infty$.
- (c) Any of the conditions (i)–(iv) implies $s_{n-1}^2/s_n^2 \rightarrow 2^{2q}$ as $n \uparrow \infty$, where we take $q = 0$ in the context of (a).

The proofs of Proposition 5.8, Theorem 5.10, Corollary 5.15, Corollary 5.12, and Proposition 5.17 are given in Section 5.5. Theorem 5.10 and Proposition 5.17 yield immediately the following result, which provides a simple construction for a function with prescribed Φ_q -variation. When applying it to the stochastic process (5.30), it provides a straightforward way of constructing a stochastic process whose sample paths have deterministic and linear Φ_q -variation; see Figure 5.1 for an illustration.

Corollary 5.18. *Let g be a regularly varying function, $q \in [0, 1)$, and Φ_q as in (5.25).*

- (a) *For $q \in (0, 1)$, the functions f in (5.23) and (5.29) with coefficients*

$$\alpha_m := 2^{-m(1-q)} \sqrt{(2^{2q} - 1)g(m)}$$

have linear Φ_q -variation (5.28).

- (b) *For $q = 0$, we assume in addition that $g(x) \rightarrow \infty$ as $x \uparrow \infty$ and that g is absolutely continuous with regularly varying derivative g' . Then the functions f in (5.23) and (5.29) with coefficients*

$$\alpha_m := 2^{-m} \sqrt{g'(m)}$$

have linear Φ_q -variation (5.27).

Proof. (b) Since the function $L(x) := g'(\log_2 x)$ is slowly varying by Remark 5.16, Proposition 5.17 (a) yields that $s_n^2/\ell(2^n) \rightarrow 1$, where $\ell(x) = \int_0^{\log_2 x} L(2^t) dt = g(\log_2 x) - g(0)$. Hence, the conditions of Theorem 5.10 (b) are satisfied. Part (a) is now obvious. $\square$

5.5 Proofs for the results in Chapter 5

Proof of Proposition 5.17. (i) $\Rightarrow$ (ii): For any given $\varepsilon > 0$, there exists n_0 such that for all $m > n_0$, we have $L(b^m)(1 - \varepsilon) < \beta_m^2 < L(b^m)(1 + \varepsilon)$. Hence, for all $\nu \geq n_0$ and $n > \nu$,

$$\sum_{m=0}^{\nu} \beta_m^2 + (1 - \varepsilon) \sum_{m=\nu+1}^{n-1} L(b^m) < \sum_{m=0}^{n-1} \beta_m^2 < \sum_{m=0}^{\nu} \beta_m^2 + (1 + \varepsilon) \sum_{m=\nu+1}^{n-1} L(b^m). \quad (5.32)$$

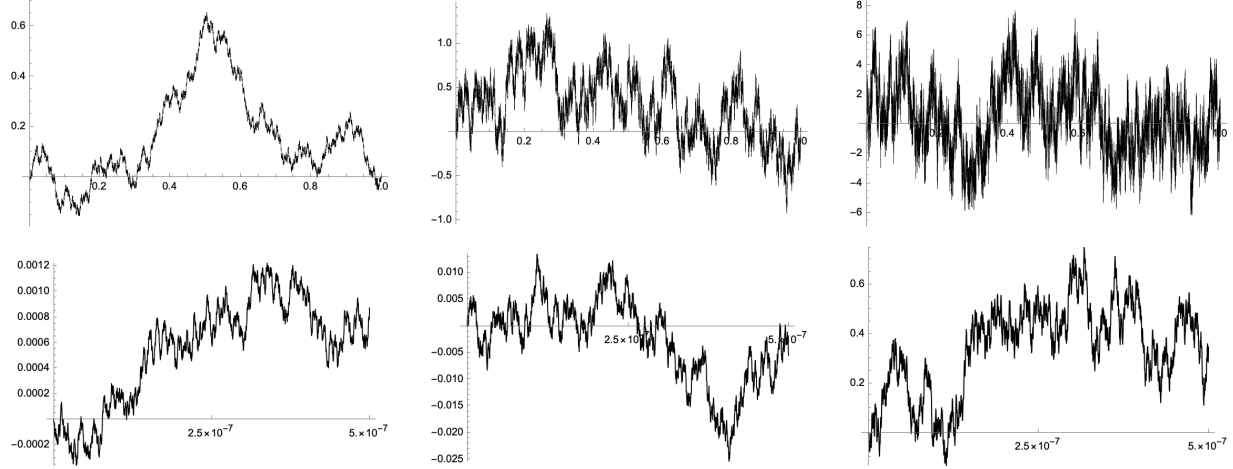


Figure 5.1: Sample paths of the stochastic process $\{X_t\}_{t \in [0,1]}$ defined in (5.30) with $\alpha_m = 2^{-m(1-q)} \sqrt{(2^{2q} - 1)g(m)}$ for $q = 0.7$, $g(x) = (1+x)^e$ with $\rho = -2$ (left), $\rho = 0$ (center), $\rho = 2$ (right), and signs $\{\xi_{m,k}\}_{m \in \mathbb{N}_0, k=1, \dots, 2^m}$ that are i.i.d. $\{-1, +1\}$ -valued Bernoulli random variables. For each ρ , a new realization of $\{\xi_{m,k}\}$ was chosen. One can see from the three top panels that the parameter ρ has a strong influence on the macroscopic fluctuations of the trajectories. Following from Proposition 5.8, all three functions share a common roughness exponent of $R = 1 - q = 0.3$. One would thus expect the fluctuations of the three functions to be similar on a microscopic scale. This effect can indeed be observed in the lower panels below, which show the same three functions after zooming into the interval $[0, \varepsilon]$ for $\varepsilon = 5 \cdot 10^{-7}$.

The Potter bounds (e.g., Theorem 1.5.6 in [9]) yield for every $\delta > 0$ some $n_1 \in \mathbb{N}$ such that

$$(1 - \varepsilon)b^{\delta(m-t)} \leq \frac{L(b^t)}{L(b^m)} \leq (1 + \varepsilon)b^{\delta(t-m)} \quad \text{for all } t, m \geq n_1.$$

We hence get that for $m \geq n_1$,

$$L(b^m)(1 - \varepsilon) \frac{1 - b^{-\delta}}{\delta \log b} \leq \int_m^{m+1} L(b^t) dt \leq L(b^m)(1 + \varepsilon) \frac{b^\delta - 1}{\delta \log b}.$$

Since $\pm(1 - b^{\mp\delta})/(\delta \log b) \rightarrow 1$ as $\delta \downarrow 0$, we can choose δ small enough such that for any $m \geq n_1$,

$$(1 - \varepsilon)^2 L(b^m) \leq \int_m^{m+1} L(b^t) dt \leq (1 + \varepsilon)^2 L(b^m).$$

Using this estimate in (5.32) and dividing both sides of the result by $\int_0^n L(b^t) dt = \ell(b^n)$ gives

$$\frac{\sum_{m=0}^{n_1-1} \beta_m^2}{\int_0^n L(b^t) dt} + \frac{(1 - \varepsilon)^3 \int_{n_1}^n L(b^t) dt}{\int_0^n L(b^t) dt} < \frac{s_n^2}{\ell(b^n)} < \frac{\sum_{m=0}^{n_1-1} \beta_m^2}{\int_0^n L(b^t) dt} + \frac{(1 + \varepsilon)^3 \int_{n_1}^n L(b^t) dt}{\int_0^n L(b^t) dt}. \quad (5.33)$$

Since $s_n \rightarrow \infty$, the inequalities in (5.33) imply that we must also have $\ell(b^n) = \int_0^n L(b^t) dt \rightarrow \infty$ as $n \uparrow \infty$. It follows that the left- and right-hand sides of (5.33) tend to $(1 - \varepsilon)^3$ and $(1 + \varepsilon)^3$ as $n \uparrow \infty$, respectively. Sending $\varepsilon \downarrow 0$ now yields (ii).

(iv) $\Leftrightarrow$ (iii): It follows from the Stolz theorem and its converse given in Lemma 3.1 of [76] that the existence of one of the following limits entails the existence of the other one and that in this case both must be equal:

$$\lim_{n \uparrow \infty} \frac{s_n^2}{2^{2qn} \ell(b^n)} \quad \text{and} \quad \lim_{n \uparrow \infty} \frac{s_{n+1}^2 - s_n^2}{2^{2q(n+1)} \ell(b^{n+1}) - 2^{2qn} \ell(b^n)}.$$

Moreover, if it exists, the limit on the right-hand side is equal to

$$\lim_{n \uparrow \infty} \frac{\beta_n^2}{2^{2qn} \ell(b^n)} \cdot \frac{1}{(2^{2q} \frac{\ell(b \cdot b^n)}{\ell(b^n)} - 1)},$$

and here the second factor converges to $(2^{2q} - 1)^{-1}$.

(c): We have

$$\lim_{n \uparrow \infty} \frac{s_{n-1}^2}{s_n^2} = 2^{2q} \lim_{n \uparrow \infty} \frac{\ell(b^{n-1})}{\ell(b \cdot b^{n-1})} = 2^{2q}.$$

This concludes the proof. $\square$

Now we turn toward proving Theorem 5.10 (and also Corollary 5.15 along the way). Following the probabilistic development of the Rademacher functions in the Faber–Schauder expansion of f in Section 5.2, we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting an independent sequence $U_1, U_2, \dots$ of symmetric $\{0, 1\}$ -valued Bernoulli random variables. Then we define the stochastic process $R_m := \sum_{k=1}^m U_k 2^{k-1}$ and set

$$Y_m := \frac{\varphi((R_m + 1)2^{-m}) - \varphi(R_m 2^{-m})}{2^{-m}}, \quad \beta_m := 2^m \alpha_m, \quad \text{and} \quad s_n^2 := \sum_{m=0}^{n-1} \beta_m^2. \quad (5.34)$$

Proposition 3.2 (a) in [89] states that $Y_1, Y_2, \dots$ is an i.i.d. sequence of symmetric $\{-1, +1\}$ -valued Bernoulli random variables. Moreover, the proof of Theorem 2.1 in [77] gives the same result in the context of Corollary 5.15, so that all subsequent observations are also valid for the functions of the form (5.29).

Note that R_m has a uniform distribution on $\{0, \dots, 2^m - 1\}$. Therefore, for $n \in \mathbb{N}$ such that all increments $|f((k+1)2^{-n}) - f(k2^{-n})|$ are less than 1,

$$V_n := \sum_{k=0}^{2^n-1} \Phi(|f((k+1)2^{-n}) - f(k2^{-n})|) = 2^n \mathbb{E} \left[\Phi(|f((R_n + 1)2^{-n}) - f(R_n 2^{-n})|) \right].$$

To analyze the expectation on the right, let the n^{th} truncation of f be given by $f_n(t) = \sum_{m=0}^{n-1} \alpha_m \varphi(2^m t)$. With the notation introduced in (5.34), we get

$$\begin{aligned} f((R_n + 1)2^{-n}) - f(R_n 2^{-n}) &= f_n((R_n + 1)2^{-n}) - f_n(R_n 2^{-n}) \\ &= 2^{-n} \sum_{m=0}^{n-1} \beta_m \frac{\varphi((R_n + 1)2^{m-n}) - \varphi(R_n 2^{m-n})}{2^{m-n}} \\ &= 2^{-n} \sum_{m=0}^{n-1} \beta_m Y_{n-m}, \end{aligned}$$

where in the last step we have used that $\varphi(x + R_n 2^{-m}) = \varphi(x + R_m 2^{-m})$ due to the periodicity of φ . Hence, if we define

$$Z_n := \sum_{m=0}^{n-1} \beta_m Y_{n-m} = \sum_{m=1}^n \beta_{n-m} Y_m,$$

then

$$V_n = 2^n \mathbb{E}[\Phi(|2^{-n} Z_n|)]. \quad (5.35)$$

Thus, the Φ -variation of f is determined by the limit of the right-hand expectations as $n \uparrow \infty$. It is worthwhile to point out that Proposition 5.8 can be directly obtained from Proposition 2.10. Nevertheless, we provide a different proof of the result from the well-known Khintchine's inequality.

Proof of Proposition 5.8. Taking $\Phi(x) = x^p$, we get from (5.35) that $V_n^{(p)} = 2^{n(1-p)} \mathbb{E}[|Z_n|^p]$. By Khintchine's inequality, there exist constants A_p and B_p only depending on p such that $0 < A_p \leq B_p < \infty$, and

$$A_p 2^{n(1-p+pq_n)} = A_p 2^{n(1-p)} (s_n^2)^{p/2} \leq V_n^{(p)} \leq B_p 2^{n(1-p)} (s_n^2)^{p/2} = B_p 2^{n(1-p+pq_n)},$$

where $q_n := \frac{1}{n} \log_2 s_n$. From here, the assertion is straightforward. $\square$

Now we prove Theorem 5.10. For better accessibility, we have divided the proof into several parts.

Proof of Theorem 5.10 (a). Assuming that $\lim_n s_n^2 < \infty$, we define $\beta_m = \alpha_m 2^m$ and

$$\tilde{Z}_n := \sum_{m=0}^{n-1} \beta_m Y_m.$$

Then the exchangeability of the sequence $\{Y_m\}_{m \in \mathbb{N}_0}$ implies that Z_n and $\tilde{Z}_n$ have the same law. Clearly, $\tilde{Z}_n \rightarrow \tilde{Z} := \sum_{m=0}^{\infty} \beta_m Y_m$ in L^1 . Therefore, (5.35) yields for the choice $\Phi(x) := x$,

$$\sum_{k=0}^{2^n-1} |f((k+1)2^{-n}) - f(k2^{-n})| = 2^n \mathbb{E}[|2^{-n} Z_n|] \rightarrow \mathbb{E}[|\tilde{Z}|].$$

Since f is continuous, this limit must coincide with the total variation of f (see, e.g., Theorem 2 in §5 of Chapter VIII in [79]). Conversely, parts (b) and (c) of Theorem 5.10 will imply that f can only be of finite total variation if $\lim_n s_n^2 < \infty$. $\square$

Proof of Theorem 5.10 (b) for $t = 1$. Here, we identify the Φ_0 -variation $\langle f \rangle_1^{\Phi_0}$ at time 1. The linearity of the Φ_0 -variation $t \mapsto \langle f \rangle_t^{\Phi_0}$ will be proved subsequently. Condition (5.26) for $q = 0$ together with our assumption $s_n^2 \rightarrow \infty$ implies via Proposition 1.5.1 in [9] that g is regularly varying with index $\rho \geq 0$. Moreover, Remark 5.16, Lemma 5.7, and Proposition 5.17 yield that the law and the first moment of Z_n/s_n converge, respectively, to $N(0, 1)$ and to its first moment.

Let us write $\psi(x) := \sqrt{g(x)}$ so that $2^n \Phi_0(2^{-n} |Z_n|) = |Z_n|/\psi(n - \log_2 |Z_n|)$. It follows that

$$2^n \Phi_0(2^{-n} |Z_n|) \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon\}} = \frac{\psi(n)}{\psi(n - \log_2 |Z_n|)} \cdot \frac{|Z_n|}{\psi(n)} \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon\}}. \quad (5.36)$$

To deal with the first factor on the right, we estimate $n - \log_2 |Z_n|$. First, on $\{|Z_n|/\psi(n) \geq \varepsilon\}$, we have $\log_2 |Z_n| \geq \log_2 \varepsilon + \ell(n)$, where $\ell(x) := \log_2 \psi(x)$ is slowly varying by Proposition 1.5.7 in [9] and hence satisfies $\ell(n)/n \rightarrow 0$ according to Proposition 1.3.6 (v) in [9]. Therefore, there exists $n_3 \in \mathbb{N}$ such that

$$n - \log_2 |Z_n| \leq n - \log_2 \varepsilon - \ell(n) \leq (1 + \varepsilon)n \quad \text{on } \{|Z_n|/\psi(n) \geq \varepsilon\} \text{ for all } n \geq n_3. \quad (5.37)$$

Second, to get a lower bound, we use that $|Z_n| \leq \sum_{m=0}^{n-1} |\beta_m|$. Jensen's inequality gives furthermore that

$$|Z_n| \leq \sum_{m=0}^{n-1} |\beta_m| \leq n \cdot \sqrt{\frac{1}{n} \sum_{m=0}^{n-1} \beta_m^2} = \sqrt{n} s_n. \quad (5.38)$$

By (5.26), there is $n_4 \in \mathbb{N}$ such that $s_n \leq (1 + \varepsilon)\psi(n)$ for all $n \geq n_4$. Thus, there is $n_5 \geq n_4$ such that

$$n - \log_2 |Z_n| \geq n - \log_2(1 + \varepsilon) - \frac{1}{2} \log_2 n - \ell(n) \geq (1 - \varepsilon)n \quad \text{for all } n \geq n_5. \quad (5.39)$$

Combining (5.37) and (5.39) now yields

$$\frac{\psi(n)}{\psi(n - \log_2 |Z_n|)} \geq \inf_{1-\varepsilon \leq \lambda \leq 1+\varepsilon} \frac{\psi(n)}{\psi(\lambda n)} \quad \text{for } n \geq n_5 \text{ on } \{|Z_n|/\psi(n) \geq \varepsilon\}. \quad (5.40)$$

According to the uniform convergence theorem for regularly varying functions (Theorem 1.5.2 in [9]) and the fact that ψ is regularly varying with index $\varrho/2 \geq 0$ by Proposition 1.5.7 (i) in [9], we have

$$\inf_{1-\varepsilon \leq \lambda \leq 1+\varepsilon} \frac{\psi(n)}{\psi(\lambda n)} \longrightarrow (1 - \varepsilon)^{\varrho/2}. \quad (5.41)$$

To deal with the second factor on the right-hand side of (5.36), we choose $n_6 \geq n_5$ such that $\psi(n) \leq (1 + \varepsilon)s_n$ for all $n \geq n_6$. Then, for $n \geq n_6$,

$$\frac{|Z_n|}{\psi(n)} \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon\}} \geq \frac{|Z_n|}{(1 + \varepsilon)s_n} \mathbb{1}_{\{|Z_n|/s_n \geq \varepsilon(1+\varepsilon)\}}. \quad (5.42)$$

Altogether, we get

$$\liminf_{n \uparrow \infty} 2^n \mathbb{E} \left[\Phi_0(2^{-n} |Z_n|) \right] \geq \frac{(1 - \varepsilon)^{\varrho/2}}{(1 + \varepsilon)\sqrt{2\pi}} \int_{\{|z| \geq \varepsilon(1+\varepsilon)\}} |z| e^{-z^2/2} dz.$$

With (5.35), we thus get $\liminf_n V_n \geq \sqrt{2/\pi}$ by sending $\varepsilon \downarrow 0$.

To get an upper bound, we recall that $x \mapsto \Phi_0(1/x)$ is regularly varying (at infinity) with index $-1 < 0$. Theorem 1.8.2 in [9] hence yields a function h that is regularly varying with index -1 , strictly decreasing, and satisfies $\Phi_0(1/x) \leq h(x)$ as well as $h(x)/\Phi_0(1/x) \rightarrow 1$ as $x \uparrow \infty$. Thus, the function $\Psi(x) := h(1/x)$ is strictly increasing, regularly varying at zero with index 1 , and satisfies $\Phi_0 \leq \Psi$ as well as $\Psi(x)/\Phi_0(x) \rightarrow 1$ as $x \downarrow 0$. Let $\varepsilon > 0$ be given and choose $x_0 > 0$ such that $\Psi(x) \leq \Phi_0(x)/(1 - \varepsilon)$ for all $x \leq x_0$. By (5.38), there is $n_7 \in \mathbb{N}$ such that $2^{-n}|Z_n| \leq x_0$ for all $n \geq n_7$. For such n , we get

$$\begin{aligned} 2^n \mathbb{E} \left[\Phi_0(2^{-n} |Z_n|) \right] &\leq 2^n \mathbb{E} \left[\Psi(2^{-n} |Z_n|) \right] \\ &\leq 2^n \mathbb{E} \left[\Psi(2^{-n}(\varepsilon + |Z_n|)) \right] \leq \frac{2^n}{1 - \varepsilon} \mathbb{E} \left[\Phi_0(2^{-n}(\varepsilon + |Z_n|)) \right] \\ &= \frac{s_n}{(1 - \varepsilon)\psi(n)} \mathbb{E} \left[\frac{\varepsilon + |Z_n|}{s_n} \cdot \frac{\psi(n)}{\psi(n - \log_2(\varepsilon + |Z_n|))} \right]. \end{aligned}$$

Clearly, the factor $s_n/\psi(n)$ converges to 1. Moreover, by (5.39) there is $n_8 \geq n_5 \vee n_7$ such that

$$(1 - \varepsilon)n \leq n - \log_2(\varepsilon + |Z_n|) \leq n - \log_2 \varepsilon \leq (1 + \varepsilon)n \quad \text{for all } n \geq n_8.$$

As in (5.40) and (5.41), we thus get

$$\frac{\psi(n)}{\psi(n - \log_2(\varepsilon + |Z_n|))} \leq \sup_{1-\varepsilon \leq \lambda \leq 1+\varepsilon} \frac{\psi(n)}{\psi(\lambda n)} \longrightarrow (1 + \varepsilon)^{e/2}.$$

We hence obtain the upper bound

$$\limsup_{n \uparrow \infty} V_n \leq \frac{(1 + \varepsilon)^{e/2}}{(1 - \varepsilon)\sqrt{2\pi}} \int |z| e^{-z^2/2} dz, \quad (5.43)$$

where the right-hand side reduces to $\sqrt{2/\pi}$ as $\varepsilon \downarrow 0$. $\square$

The proof of Theorem 5.10 (b) for $0 < t < 1$ requires that we truncate the summation for V_n in (5.35) at those indices k for which $k2^{-n} \leq t$. This is equivalent to restricting the expectation in (5.35) to the set $\{2^{-n}R_n \leq t\}$. The following proof adapts the arguments of the proof of Theorem 5.10 (b) for $t = 1$ to this restricted case.

Proof of Theorem 5.10 (b) for $0 < t < 1$. Let $t \in (0, 1)$ be given. The Φ_0 -variation of f over the interval $[0, t]$ is equal to $\lim_n V_{n,t}$, where

$$V_{n,t} := \sum_{k=0}^{2^n-1} \Phi_0(|f((k+1)2^{-n}) - f(k2^{-n})|) \mathbb{1}_{[0,t]}(k2^{-n}) = 2^n \mathbb{E} \left[\Phi_0(2^{-n}|Z_n|) \mathbb{1}_{\{2^{-n}R_n \leq t\}} \right].$$

This can be proved similarly as in the derivation of (5.35) as in Section 5.1. We fix $\delta \in (0, t \wedge (1 - t))$ and $m \in \mathbb{N}$ that $2^{-m} \leq \delta$. Recall that $R_m := \sum_{k=1}^m U_k 2^{k-1}$ and let us denote $R_{m,n} := R_n - R_{n-m}$ for short. Then

$$\{2^{-n}R_{m,n} \leq t - \delta\} \subseteq \{2^{-n}R_n \leq t\} \subseteq \{2^{-n}R_{m,n} \leq t\}. \quad (5.44)$$

First, we derive a lower bound for $V_{n,t}$. To this end, we define for $0 \leq m < n$,

$$Z_{m,n} := \sum_{k=1}^{n-m} \beta_{n-k} Y_k = Z_n - \sum_{k=n-m+1}^n \beta_{n-k} Y_k.$$

Then, since $s_n \rightarrow \infty$,

$$\frac{|Z_n - Z_{m,n}|}{s_n} \leq \frac{1}{s_n} \sum_{k=0}^{m-1} |\beta_k| =: a_n \longrightarrow 0,$$

and so the laws of $Z_{m,n}/s_n$ converge to $N(0, 1)$ in the L^1 -Wasserstein distance by Slutsky's theorem. By combining (5.36), (5.40), (5.41), and (5.42), there is $n_9 \geq n_6 \vee m$ such that for

all $n \geq n_9$, with $c_\varepsilon := (1 - \varepsilon)^{1+\varrho/2}/(1 + \varepsilon)$,

$$\begin{aligned} 2^n \Phi_0(2^{-n}|Z_n|) \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon, 2^{-n}R_n \leq t\}} &\geq \frac{c_\varepsilon |Z_n|}{s_n} \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon, 2^{-n}R_n \leq t\}} \\ &\geq \frac{c_\varepsilon |Z_{m,n}|}{s_n} \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon, 2^{-n}R_n \leq t\}} - \frac{c_\varepsilon |Z_n - Z_{m,n}|}{s_n} \\ &\geq \frac{c_\varepsilon |Z_{m,n}|}{s_n} \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon, 2^{-n}R_n \leq t\}} - \frac{c_\varepsilon \sum_{k=0}^{m-1} |\beta_k|}{s_n}, \end{aligned}$$

where the last two inequalities hold due to the triangle inequality. Again, the right-most term decays to zero as $n \uparrow \infty$. Hence, there is $n_{10} \geq n_9$ such that for $n \geq n_{10}$ and $|Z_{m,n}|/s_n \geq 2\varepsilon(1 + \varepsilon)$,

$$\frac{|Z_n|}{s_n} \geq \frac{|Z_{m,n}| - |Z_n - Z_{m,n}|}{s_n} \geq \frac{|Z_{m,n}| - \sum_{k=0}^{m-1} |\beta_k|}{s_n} \geq \varepsilon(1 + \varepsilon).$$

Hence, for those n and m , we have $\{|Z_{m,n}|/s_n \geq 2\varepsilon(1 + \varepsilon)\} \subseteq \{|Z_n|/s_n \geq \varepsilon(1 + \varepsilon)\}$ and in turn

$$\begin{aligned} V_{n,t} + c_\varepsilon a_n &\geq \mathbb{E} \left[\frac{c_\varepsilon |Z_{m,n}|}{s_n} \mathbb{1}_{\{|Z_n|/\psi(n) \geq \varepsilon, 2^{-n}R_n \leq t\}} \right] \geq \mathbb{E} \left[\frac{c_\varepsilon |Z_{m,n}|}{s_n} \mathbb{1}_{\{|Z_{m,n}|/s_n \geq 2\varepsilon(1+\varepsilon), 2^{-n}R_{m,n} \leq t-\delta\}} \right] \\ &\geq \mathbb{E} \left[\frac{c_\varepsilon |Z_{m,n}|}{s_n} \mathbb{1}_{\{|Z_{m,n}|/s_n \geq 2\varepsilon(1+\varepsilon)\}} \right] \cdot \mathbb{P}[2^{-n}R_n \leq t - \delta], \end{aligned}$$

where the last step follows from (5.44) and the independence of $Z_{m,n}$ and $R_{m,n}$. Clearly, $\mathbb{P}[2^{-n}R_n \leq t - \delta] \rightarrow t - \delta$ and

$$\mathbb{E} \left[\frac{c_\varepsilon |Z_{m,n}|}{s_n} \mathbb{1}_{\{|Z_{m,n}|/s_n \geq 2\varepsilon(1+\varepsilon)\}} \right] \longrightarrow \frac{c_\varepsilon}{\sqrt{2\pi}} \int_{\{|z| \geq 2\varepsilon(1+\varepsilon)\}} |z| e^{-z^2/2} dz,$$

since $s_{n-m}/s_n \rightarrow 1$ by Proposition 5.17. Hence

$$\liminf_{n \uparrow \infty} V_{n,t} \geq \frac{c_\varepsilon(t - \delta)}{\sqrt{2\pi}} \int_{\{|z| \geq 2\varepsilon(1+\varepsilon)\}} |z| e^{-z^2/2} dz. \quad (5.45)$$

To obtain a corresponding upper bound, we can argue exactly as in the derivation of (5.43) to get

$$\limsup_{n \uparrow \infty} V_{n,t} \leq (1 + \varepsilon)^{\varrho/2} \limsup_{n \uparrow \infty} \mathbb{E} \left[\frac{\varepsilon + |Z_n|}{s_n} \mathbb{1}_{\{2^{-n}R_{m,n} \leq t\}} \right] = (1 + \varepsilon)^{\varrho/2} \limsup_{n \uparrow \infty} \mathbb{E} \left[\frac{|Z_n|}{s_n} \mathbb{1}_{\{2^{-n}R_{m,n} \leq t\}} \right].$$

Using again the independence of $Z_{m,n}$ and $R_{m,n}$, we find that

$$\mathbb{E}[|Z_n| \mathbb{1}_{\{2^{-n}R_{m,n} \leq t\}}] \leq \mathbb{E}[|Z_{m,n}|] \cdot \mathbb{P}[2^{-n}R_{m,n} \leq t] + \sum_{k=0}^{m-1} |\beta_k|.$$

By (5.44), we have $\mathbb{P}[2^{-n}R_{m,n} \leq t] \leq \mathbb{P}[2^{-n}R_n \leq t+\delta] \rightarrow t+\delta$. Recall that $\frac{1}{s_n} \sum_{k=0}^{m-1} |\beta_k| \rightarrow 0$ as $n \uparrow \infty$. Next, the function $\ell(x) := g(\log_2 x)$ is slowly varying by Remark 5.16, and so

$$\frac{g(n)}{g(n-m)} = \frac{\ell(2^n)}{\ell(2^{-m}2^n)} \rightarrow 1 \quad \text{as } n \uparrow \infty. \quad (5.46)$$

Thus, our condition (5.26) for $q = 0$ implies that

$$\mathbb{E}\left[\frac{|Z_{m,n}|}{s_n}\right] = \frac{s_{n-m}}{s_n} \mathbb{E}\left[\frac{|Z_{m,n}|}{s_{n-m}}\right] \rightarrow \sqrt{\frac{2}{\pi}} \quad \text{as } n \uparrow \infty.$$

Altogether, we conclude that $\limsup_n V_{n,t} \leq (1+\varepsilon)^{q/2}(t+\delta)\sqrt{2/\pi}$. Combining this inequality with (5.45) and sending ε and δ to zero, we conclude the proof of the linearity of the Φ_0 -variation. $\square$

The proof of part (c) of Theorem 5.10 will be prepared with the following two lemmas.

Lemma 5.19. *Under the conditions of Theorem 5.10 (c), there is a constant K such that for all $\omega \in \Omega$ and all sufficiently large $n \in \mathbb{N}$,*

$$2^{-qn}|Z_n(\omega)| \leq K\sqrt{g(n)} \quad \text{and} \quad |Z_n(\omega)| \leq Ks_n.$$

Proof. The function $\ell(x) := g(\log_2 x)$ is slowly varying by Remark 5.16, and so ℓ and the sequence $\{\beta_m\}_{m \in \mathbb{N}_0}$ satisfy the conditions of Theorem 5.6. Recall from (5.19) in the proof of that theorem that there exists $\delta_1 > 0$ such that for every $\delta \in (0, \delta_1)$ there is $n_1 \in \mathbb{N}$ such that for all $m \geq n_1$,

$$2^{qm}\sqrt{(1-\delta)\ell(2^m) - (1+\delta)\ell(2^{m-1})2^{-2q}} < |\beta_m| < 2^{qm}\sqrt{(1+\delta)\ell(2^m) - (1-\delta)\ell(2^{m-1})2^{-2q}},$$

and the arguments of the two square roots are positive. Now let $\delta \in (0, \delta_1)$ be given and take $n_2 \geq n_1$ such that $1-\delta < g(n)/g(n-1) < 1+\delta$ for any $n \geq n_2$; this is possible by (5.46). Since g is regularly varying, we have $2^{2qn}g(n) \rightarrow \infty$ as $n \uparrow \infty$ (see, e.g., Proposition B.1.9 in [23]). Therefore, we may apply the Stolz–Cesàro theorem in its general form so as

to obtain

$$\begin{aligned}
\limsup_{n \uparrow \infty} \frac{\sum_{m=0}^n |\beta_m|}{2^{qn} \sqrt{g(n)}} &\leq \limsup_{n \uparrow \infty} \frac{|\beta_n|}{2^{qn} \sqrt{g(n)} - 2^{q(n-1)} \sqrt{g(n-1)}} \\
&\leq \limsup_{n \uparrow \infty} \frac{\sqrt{(1+\delta)g(n) - (1-\delta)g(n-1)2^{-2q}}}{\sqrt{g(n)} - \sqrt{2^{-2q}g(n-1)}} \\
&\leq \frac{\sqrt{(1+\delta)2^{2q} - (1-\delta)}}{\sqrt{2^{2q}(1-\delta)} - 1}.
\end{aligned}$$

Since $|Z_n| \leq \sum_{m=0}^{n-1} |\beta_m|$, the claim follows. $\square$

Lemma 5.20. *Suppose that $\{z_n\}_{n \in \mathbb{N}_0}$ is a sequence of nonnegative numbers converging to $z > 0$. Then $2^n \Phi_q(2^{-n} s_n z_n) \rightarrow z^{1/(1-q)}$ as $n \uparrow \infty$.*

Proof. We may assume without loss of generality that $z_n > 0$ for all n . Then we may write

$$\begin{aligned}
2^n \Phi_q(2^{-n} s_n z_n) &= 2^n (2^{-n} s_n z_n)^{1/(1-q)} g\left(\frac{n - \log_2 s_n - \log_2 z_n}{1-q}\right)^{-1/(2(1-q))} \\
&= \left(\frac{s_n}{2^{qn} \sqrt{g(n)}}\right)^{1/(1-q)} z_n^{1/(1-q)} \left(\frac{g(n)}{g\left(\frac{n - \log_2 s_n - \log_2 z_n}{1-q}\right)}\right)^{1/(2(1-q))}.
\end{aligned}$$

By (5.26), the first factor on the right converges to 1. The second factor converges to $z^{1/(1-q)}$ by assumption. To deal with the third factor, (5.26) implies that there is $n_0 \in \mathbb{N}$ such that

$$qn - \log_2 \sqrt{g(n)} - 1 \leq \log_2 s_n \leq qn - \log_2 \sqrt{g(n)} + 1 \quad \text{for all } n \geq n_0.$$

Theorem 1.4.1 and Proposition 1.3.6 (i) in [9] give moreover that there exists $\kappa^+, \kappa^- \in \mathbb{R}$ such that $\kappa^- \log_2 n \leq \log_2 \sqrt{g(n)} \leq \kappa^+ \log_2 n$. Thus, if $n \geq n_0$ is sufficiently large such that $z/2 \leq z_n \leq 2z$, then

$$n + \frac{-1 - \log_2(2z) - \kappa^+ \log_2 n}{1-q} \leq \frac{n - \log_2 s_n - \log_2 z_n}{1-q} \leq n + \frac{1 - \log_2(z/2) - \kappa^- \log_2 n}{1-q}. \quad (5.47)$$

For sufficiently large n , the center term can thus be expressed as $n\lambda_n$, where $\lambda_n \in [1/2, 3/2]$ and $\lambda_n \rightarrow 1$. The uniform convergence theorem for regularly varying functions (e.g., Theorem 1.5.2 in [9]) hence implies that

$$\frac{g(n)}{g\left(\frac{n - \log_2 s_n - \log_2 z_n}{1-q}\right)} \rightarrow 1$$

as $n \uparrow \infty$. This concludes the proof. $\square$

Proof of Theorem 5.10 (c). We prove part (c) only for $t = 1$. The extension to $0 < t < 1$ is almost verbatim identical to the one in part (b) and hence left to the reader. We write

$$\widehat{Z}_n := \frac{1}{s_n} \sum_{m=1}^n |\beta_{n-m}| Y_m.$$

Then the law of $2^n \Phi_q(2^{-n} |Z_n|)$ is the same as that of $2^n \Phi_q(2^{-n} s_n |\widehat{Z}_n|)$. By Theorem 5.6, we have $|\widehat{Z}_n| \rightarrow |Z|$ in L^∞ . Moreover, by either Corollary 6.6 or Remark 6.7 in [26], the law of Z has no atoms. In particular, we have $\mathbb{P}[Z = 0] = 0$. Lemma 5.20 hence yields that $\mathbb{P}$ -a.s. $2^n \Phi_q(2^{-n} s_n |\widehat{Z}_n|) \rightarrow |Z|^{1/(1-q)}$. Fatou's lemma thus gives immediately that

$$\liminf_{n \uparrow \infty} V_n = \liminf_{n \uparrow \infty} 2^n \mathbb{E}[\Phi_q(2^{-n} |Z_n|)] = \liminf_{n \uparrow \infty} \mathbb{E}[2^n \Phi_q(2^{-n} s_n |\widehat{Z}_n|)] \geq \mathbb{E}[|Z|^{1/(1-q)}].$$

To get an upper bound, we argue as in the proof of Theorem 5.10 (b) for $t = 1$ that there exists an increasing function Ψ that is regularly varying at zero with index $1/(1-q)$ and satisfies $\Phi_q(x) \leq \Psi(x)$ for all sufficiently small x as well as $\Psi(x)/\Phi_q(x) \rightarrow 1$ as $x \downarrow 0$. In particular, we have $2^n \Psi(2^{-n} s_n z_n) \rightarrow z^{1/(1-q)}$ if $\{z_n\}_{n \in \mathbb{N}_0}$ is a sequence of nonnegative numbers converging to $z > 0$. Let $\varepsilon > 0$ be given. We choose $x_0 > 0$ such that $\Psi(x) \leq \Phi_q(x)/(1-\varepsilon)$ for all $x \leq x_0$. Since the sequence $\{\widehat{Z}_n\}$ is uniformly bounded, there is $n_0 \in \mathbb{N}$ such that $2^{-n} s_n (\varepsilon + |\widehat{Z}_n|) \leq x_0$ $\mathbb{P}$ -a.s. for all $n \geq n_0$. For such n , we hence have

$$\begin{aligned} V_n &= 2^n \mathbb{E}[\Phi_q(2^{-n} s_n |\widehat{Z}_n|)] \leq 2^n \mathbb{E}[\Psi(2^{-n} s_n |\widehat{Z}_n|)] \\ &\leq 2^n \mathbb{E}[\Psi(2^{-n} s_n (\varepsilon + |\widehat{Z}_n|))] \leq \frac{2^n}{1-\varepsilon} \mathbb{E}[\Phi_q(2^{-n} s_n (\varepsilon + |\widehat{Z}_n|))]. \end{aligned}$$

Moreover, we see from (5.47) that there exists $n_1 \geq n_0$ such that the random variables Λ_n defined through

$$\Lambda_n := \frac{n - \log_2 s_n - \log_2 (\varepsilon + |\widehat{Z}_n|)}{n(1-q)}$$

take values in $[1/2, 3/2]$ for $n \geq n_1$ and satisfy $\Lambda_n \rightarrow 1$ $\mathbb{P}$ -a.s. Hence, with ϱ denoting the index of regular variation of g ,

$$0 \leq \frac{g(n)}{g\left(\frac{n - \log_2 s_n - \log_2 (\varepsilon + |\widehat{Z}_n|)}{1-q}\right)} = \frac{g(n)}{g(n\Lambda_n)} \leq \sup_{1/2 \leq \lambda \leq 3/2} \frac{g(n)}{g(n\lambda)} \rightarrow \left(\frac{3}{2}\right)^\varrho$$

according to the uniform convergence theorem for regularly varying functions (e.g., Theorem 1.5.2 in [9]). Hence, Lemma 5.20 and dominated convergence give that

$$\limsup_{n \uparrow \infty} V_n \leq \limsup_{n \uparrow \infty} \frac{1}{1-\varepsilon} 2^n \mathbb{E}[\Phi_q(2^{-n} s_n (\varepsilon + |\widehat{Z}_n|))] = \frac{1}{1-\varepsilon} \mathbb{E}[(\varepsilon + |Z|)^{1/(1-q)}].$$

Sending $\varepsilon \downarrow 0$ gives the desired upper bound. $\square$

Proof of Corollary 5.12. (a) Taking logarithms in (5.26) and using the fact that $\frac{1}{n} \log_2 g(n) \rightarrow 0$, we get $\frac{1}{n} \log_2 s_n \rightarrow q$ as $n \uparrow \infty$. Hence, the result follows from Proposition 5.8.

(b) If $0 < c < \infty$, then (5.26) holds also if g is replaced with $\tilde{g}(x) := c$. Hence, the assertion follows immediately from Theorem 5.10. If $c = 0$, then for any $\varepsilon > 0$ there is $\delta > 0$ such that

$$g\left(\frac{-\log_2 \eta}{1-q}\right) < \varepsilon \quad \text{for all } \eta < \delta.$$

Now it suffices to take $n_0 \in \mathbb{N}$ such that $|f((k+1)2^{-n}) - f(k2^{-n})| < \delta$ for all k and $n \geq n_0$ to obtain that

$$\sum_{k=0}^{\lfloor t2^n \rfloor} |f((k+1)2^{-n}) - f(k2^{-n})|^p \leq \varepsilon^{p/2} \sum_{k=0}^{\lfloor t2^n \rfloor} \Phi_q(|f((k+1)2^{-n}) - f(k2^{-n})|).$$

Sending $\varepsilon \downarrow 0$ gives the result. The case $c = \infty$ is obtained analogously. $\square$

Chapter 6

Future Work

In this chapter, we outline several potential directions for future research. In Section 6.1, we extend our scale-invariant estimators from Definition 2.39 to detect dynamic changes in the roughness exponent of a time series over time. Section 6.2 explores how the estimator $\widehat{\mathcal{R}}_n^*$ can be adapted to estimate the roughness exponent of the volatility process directly from discrete observations of the realized volatility.

6.1 Dynamic detection of the roughness exponent

In many applications, dynamically monitoring the Hurst parameter or the roughness exponent of a time series is important. An example is the ExtremeHurstTM function on the Bloomberg Terminal, which can be used for market timing. For dynamic monitoring, the original time series is divided into several (possibly overlapping) intervals, and the Hurst parameter or the roughness exponent is then measured on each of these intervals.

For a formal description, we fix $x \in C[0, T]$ and an interval shift $0 \leq \tau \leq T - 1$. Then we let $x_\tau(t) := x(t + \tau)$ for $t \in [0, 1]$. For any estimator $\widehat{R}$, the value $\widehat{R}(x_\tau)$ gives an estimate of roughness exponent of x over the interval $[\tau, \tau + 1]$.

Definition 6.1. Let $x \in C[0, T]$, $0 \leq \tau \leq T - 1$, and $\mathbb{T} := \{0 \leq \tau_1 < \tau_2 < \dots < \tau_N \leq T - 1\}$. Fix $m \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_m \geq 0$ with $\alpha_0 > 0$.

- (a) For $n > m$, the $\mathbb{T}$ -adjusted sequential scaling factor $\lambda_n^{s, \mathbb{T}}$ and the $\mathbb{T}$ -adjusted sequential scale estimate $R_n^{s, \mathbb{T}}(x_\tau)$ are defined as follows,

$$\lambda_n^{s, \mathbb{T}} := \arg \min_{\lambda > 0} \sum_{j=1}^N \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{R}_k(\lambda x_{\tau_j}) - \widehat{R}_{k-1}(\lambda x_{\tau_j}) \right)^2 \quad \text{and} \quad R_n^{s, \mathbb{T}}(x) := \widehat{R}_n(\lambda_n^{s, \mathbb{T}} x). \quad (6.1)$$

The corresponding mapping $R_n^{s,\mathbb{T}} : C[0, 1] \rightarrow \mathbb{R}$ will be called the ($\mathbb{T}$ -adjusted) *sequential scale estimator*.

- (b) For $n > m$, the *terminal scaling factor* $\lambda_n^{t,\mathbb{T}}$ and the $\mathbb{T}$ -adjusted *terminal scale estimate* $R_n^{t,\mathbb{T}}(x)$ are defined as follows,

$$\lambda_n^{t,\mathbb{T}} := \arg \min_{\lambda > 0} \sum_{j=1}^N \sum_{k=n-m}^n \alpha_{n-k} \left(\widehat{R}_n(\lambda x_{\tau_j}) - \widehat{R}_{k-1}(\lambda x_{\tau_j}) \right)^2 \quad \text{and} \quad R_n^{t,\mathbb{T}}(x) := \widehat{R}_n(\lambda_n^{t,\mathbb{T}} x). \quad (6.2)$$

The corresponding mapping $R_n^{t,\mathbb{T}} : C[0, 1] \rightarrow \mathbb{R}$ will be called the ($\mathbb{T}$ -adjusted) *terminal scale estimator*.

The estimates $R_n^{s,\mathbb{T}}(x_\tau)$ and $R_n^{t,\mathbb{T}}(x_\tau)$ are the sequential and terminal scale estimates of the roughness exponent of x over the interval $[\tau, \tau + 1]$ with a scaling factor chosen with respect to the collection $\mathbb{T}$. We have the following representation result for our $\mathbb{T}$ -adjusted estimators.

Proposition 6.2. *Consider the context of Definitions 6.1 with fixed $m \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_m \geq 0$ such that $\alpha_0 > 0$.*

- (a) *The two optimization problems in (6.1) and (6.2) admit unique solutions for every function $x \in C[0, 1]$. In particular, all objects in Definitions 6.1 are well defined.*
- (b) *The $\mathbb{T}$ -adjusted sequential scaling factor and scale estimate can be represented as follows:*

$$\log_2 \lambda_n^{s,\mathbb{T}} = \frac{1}{N} \sum_{j=1}^N \log_2 \lambda_n^{s,\{\tau_j\}} \quad \text{and} \quad R_n^{s,\mathbb{T}}(x_\tau) = \widehat{R}_n(x_\tau) + \frac{1}{N} \sum_{j=1}^N \left(R_n^{s,\{\tau_j\}}(x_{\tau_j}) - \widehat{R}_n(x_{\tau_j}) \right),$$

Moreover, analogous representations hold if we replace the sequential factor, estimate and estimator with the terminal factor, estimate and estimator respectively.

- (c) *The sequential and terminal scale estimators are scale-invariant. That is, for $n > m$, $x \in C[0, T]$, $0 \leq \tau \leq T - 1$, and $\lambda \neq 0$, we have $R_n^{t,\mathbb{T}}(\lambda x_\tau) = R_n^{t,\mathbb{T}}(x_\tau)$ and $R_n^{s,\mathbb{T}}(\lambda x_\tau) = R_n^{s,\mathbb{T}}(x_\tau)$.*

The proof of Proposition 6.2 is similar to the one of Proposition 2.41 and left to the reader. The following figure illustrates that applying the $\mathbb{T}$ -adjusted scale estimator significantly reduces the fluctuation of the scale estimates over a sequence of intervals compared with a direct application of corresponding estimators. This also results in a smooth curve of the estimated roughness exponent over time.

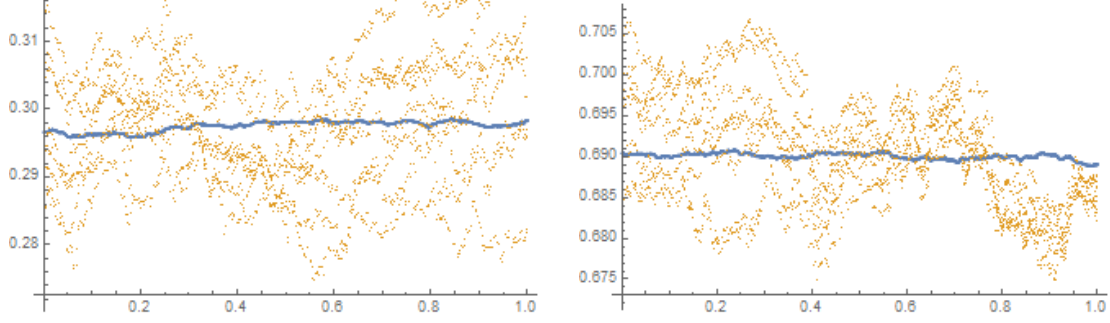


Figure 6.1: Plots of terminal scale estimates $R_n^{t, \mathbb{T}}(x_\tau)$ (in blue) and $R_n^{t, \{\tau\}}(x_\tau)$ (in yellow) for $\tau \in \{k2^{-14} : k = 0, 1, \dots, 2^{14}\}$ applied to a single trajectory of fractional Brownian motion with Hurst parameter $H = 0.3$ (left) and $H = 0.7$ (right). The other parameters are chosen to be $m = 3$, $n = 14$, $\alpha_k = 1$ and $\mathbb{T} = \{k2^{-14} : k = 0, 1, \dots, 2^{14}\}$.

In the sequel, we apply the $\mathbb{T}$ -adjusted terminal scale estimator to estimate the roughness of certain high-frequency financial time series. We implement our estimator with $n = 11$, $m = 1$ and $\alpha_k = 2^{-k}$ based on 5-minute data and rolling estimation intervals, each with previous $2^{11} + 1$ data points, i.e., 7.11 days = $2^{11} \times 5$ minutes. Furthermore, we set $\mathbb{T}$ to be the collection of all data points. Before plotting the estimated roughness exponent over time, let us first test the validity of the reverse Jensen condition for our empirical time series. To this end, we plot the ratio

$$\psi_p(n) := \frac{2^{n-1} s_n^p}{\sum_{k=0}^{2^{n-1}-1} \left(\sum_{m=0}^{n-1} 2^m \theta_{m, \lfloor 2^{m-n+1}k \rfloor}^2 \right)^{p/2}}$$

as a function of n for various p . Figure 6.2 demonstrates that for different choices of p , the ratios $\psi_p(n)$ are approximately linear; thus, the Faber–Schauder coefficients of the log prices satisfy the reverse Jensen condition, and the subexponential functions ϱ_p are approximately linear functions.

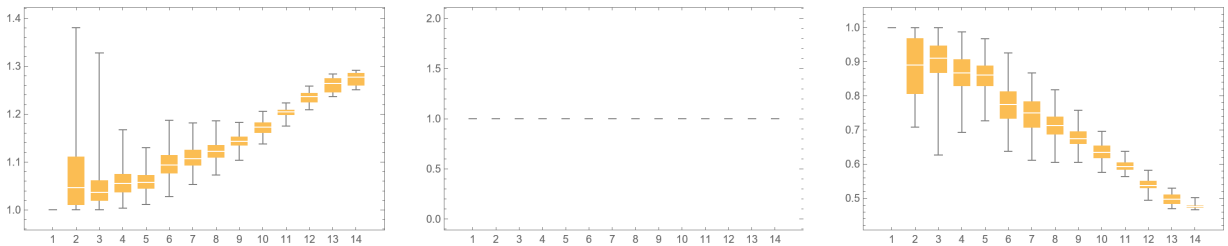


Figure 6.2: Box plots of the ratio $\psi_p(n)$ for log prices of continuous e-mini SPX futures based on 1,000 data intervals at random locations for $p = 1$ (left), $p = 2$ (center) and $p = 3$ (right).

Now, Theorem 2.4 and Proposition 6.2 evince that the $\mathbb{T}$ -adjusted estimator are strongly

consistent, and we plot the estimates of the roughness exponent in blue and the futures price in orange from the year 2006 to the year 2022.

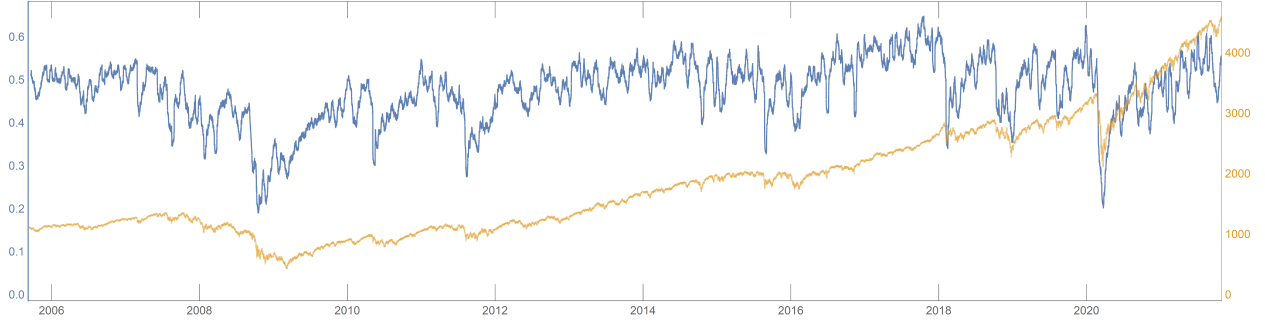


Figure 6.3: Plot of the roughness exponent (blue) for continuous e-mini SPX futures (ES, orange), estimated by $R_n^{t,T}$ with $n = 11$, $m = 1$ and $\alpha_k = 2^{-k}$.

As one can see from Figure 6.3, the roughness exponent stays around 0.5 most of the time. The market thus exhibits diffusive behaviour, which is consistent with the arbitrage requirement of finite and non-vanishing quadratic variation observed by Föllmer [32] (see also [33] and Proposition 2.1 in [88] for extensions of this argument). Exceptions to this behaviour, however, tend to occur during market crises, during which we observe a positive correlation between futures prices and their roughness exponent. In Figure 6.4, we shrink our time span down to two years starting from the year 2008 to gaze at the joint changes of the futures prices and the roughness exponent during the 2008 financial crisis.

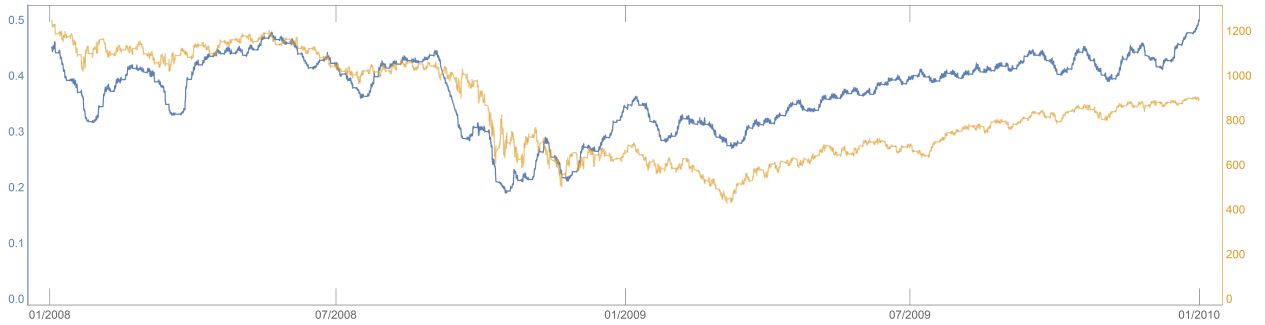


Figure 6.4: Plot of the roughness exponent (blue) for continuous e-mini SPX futures (ES, orange) during 2008 and 2009, estimated by $R_n^{t,T}$ with $n = 11$, $m = 1$ and $\alpha_k = 2^{-k}$.

It is clearly shown that the drop in roughness precedes the drop in stock prices. Moreover, the roughness exponent quickly deviates from 0.5 in the steep drop before the crisis.

Next, we consider the roughness of the realized volatility of the SPX index. Gatheral et al. [40] split the volatility process into halves and estimate the Hurst parameter of each

separated process. Based on their numerical results, they claim that the Hurst parameter does not vary remarkably over time. We proxy the realized volatility by the daily realized volatility estimates from the Oxford-Man Institute of Quantitative Finance Realized Library (from Jan 03, 2000 to Nov 15, 2021) and apply the terminal $\mathbb{T}$ -adjusted scale estimator to estimate the degree of smoothness of the log volatility. We implement our estimation with $m = 1$ and $\alpha_k = 2^{-k}$ based on daily data and rolling estimation intervals, each with previous $2^n + 1$ data points, i.e., 2^n days for $n = 8$ (approx. 1 year) and $n = 10$ (approx. 4 years). Furthermore, we set $\mathbb{T}$ to be the collection of all data points.

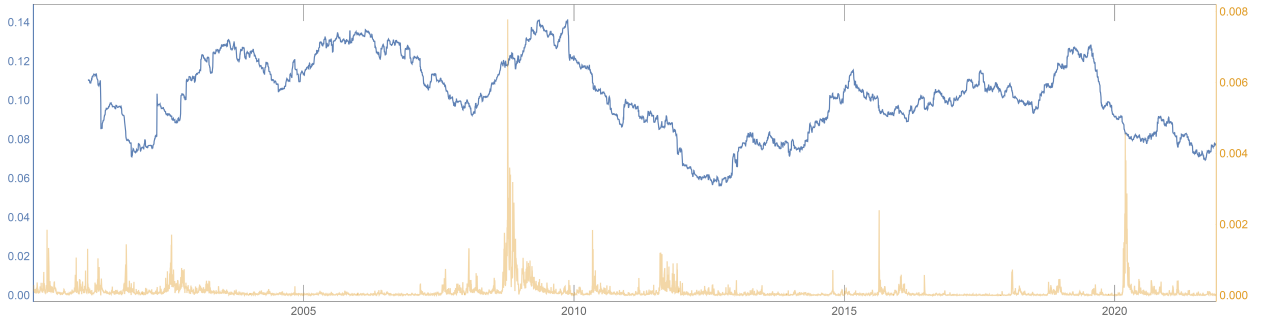


Figure 6.5: Plot of the roughness exponent (blue) for the logarithm of the estimated realized volatility of the SPX index (orange), estimated by $R_n^{t,\mathbb{T}}$ with $n = 8$, $m = 1$ and $\alpha_k = 2^{-k}$.

We can notice from the above plot that even though different estimation intervals will alter the shape of the curve of the roughness exponent, all the estimated roughness exponents are concentrated around 0.1. This appears to confirm the claim in [40] that the degree of roughness is more or less constant over time.

However, one possible drawback of the above method is the assumption that the scale of the time series remains constant over time. Moreover, in reality, the scale of many financial time series often varies. For example, the S&P 500 index has increased from 1,100 to 5,000 since the year 2008. Clearly, assuming a constant scale for the time series may not be appropriate in general. In the future, we aim to incorporate the time-varying effect of the possible scale into the presented framework.

6.2 Estimating the roughness exponent of the volatility process from realized volatility

In this section, we discuss how to accommodate the estimator $\widehat{\mathcal{R}}_n^*$ to estimate the roughness exponent of the volatility process from realized volatility. To this end, let $(B_t)_{t \in [0,1]}$ be a

standard Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. We assume that price process S_t is directly observable and admits the following dynamics:

$$\log S_t = \log S_0 + \int_0^t \sigma_s dB_s, \quad t \in [0, 1],$$

where the volatility process $(\sigma_t)_{t \in [0, 1]}$ is a continuous $\mathcal{F}_t$ -adapted process. Here, we allow the volatility process σ_t and the Brownian motion B_t to be correlated in order to generate the leverage effect. The integrated volatility Y_t is defined as

$$Y_t := \langle \log S \rangle_t = \int_0^t \sigma_s^2 ds, \quad t \in [0, 1]. \quad (6.3)$$

Chapter 3 shows if the integrated volatility is directly observed, the roughness exponent of σ_t can be directly estimated from the integrated volatility Y_t in a strict pathwise manner by the estimator

$$\widehat{\mathcal{R}}_n^*(Y) := 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \vartheta_{n,k}^2},$$

and the coefficients $\vartheta_{n,k}$ are defined as follows:

$$\vartheta_{n,k} = 2^{3n/2+3} \left(Y_{\frac{4k}{2^{n+2}}} - 2Y_{\frac{4k+1}{2^{n+2}}} + 2Y_{\frac{4k+3}{2^{n+2}}} - Y_{\frac{4k+4}{2^{n+2}}} \right). \quad (6.4)$$

However, in reality, the integrated volatility is not directly observed, so it needs to be estimated based on the realized volatility from the price process. To this end, we let $(m_n)_{n \in \mathbb{N}_0}$ be a fixed increasing sequence, and we denote

$$\widehat{Y}_t^{(n)} = \sum_{k=1}^{\lfloor 2^n m_n t \rfloor} \left(\log S_{\frac{k}{2^n m_n}} - \log S_{\frac{k-1}{2^n m_n}} \right)^2. \quad (6.5)$$

The process $\widehat{Y}_t^{(n)}$ is the realized volatility calculated from the price process S_t with mesh size $(2^n m_n)^{-1}$. In a similar manner, let us consider the proxy coefficients $\widetilde{\vartheta}_{n,k}$ of the coefficients $\vartheta_{n,k}$,

$$\widetilde{\vartheta}_{n,k} := 2^{3n/2+3} \left(\widehat{Y}_{\frac{4k}{2^{n+2}}}^{(n+2)} - 2\widehat{Y}_{\frac{4k+1}{2^{n+2}}}^{(n+2)} + 2\widehat{Y}_{\frac{4k+3}{2^{n+2}}}^{(n+2)} - \widehat{Y}_{\frac{4k+4}{2^{n+2}}}^{(n+2)} \right). \quad (6.6)$$

Analogously, we construct the estimator from the realized volatility as

$$\widetilde{\mathcal{R}}_n^*(S) := 1 - \frac{1}{n} \log_2 \sqrt{\sum_{k=0}^{2^n-1} \widetilde{\vartheta}_{n,k}^2},$$

where the notion S specifies price process that $\tilde{\mathcal{R}}_n^*$ estimates from. Clearly, the estimator $\tilde{\mathcal{R}}_n^*$ is constructed by replacing the coefficients $\vartheta_{n,k}$ with $\tilde{\vartheta}_{n,k}$ in the estimator $\hat{\mathcal{R}}_n^*$, and thus, can be directly calculated from the discrete observations of the price process S_t . The first result of this section shows that if the sequence (m_n) is properly chosen, the estimator $\tilde{\mathcal{R}}_n^*$ and $\hat{\mathcal{R}}_n^*$ converge to the same value as if the integrated volatility Y_t is directly observed.

Theorem 6.3. *Let Y_t and $\hat{Y}_t^{(n)}$ be as in (6.3) and (6.5). Suppose that $\sup_{t \in [0,1]} \sigma_t^2 \phi_t^2 \leq M$ $\mathbb{P}$ -a.s., and there exists $R \in (0, 1)$ such that $\hat{\mathcal{R}}_n^*(Y) \rightarrow R$ $\mathbb{P}$ -a.s.. If*

$$\lim_{n \uparrow \infty} \frac{1}{n} \log_2 m_n > 2R,$$

then $\tilde{\mathcal{R}}_n^*(S) \rightarrow R$ as $n \uparrow \infty$ $\mathbb{P}$ -a.s.

In particular, Theorem 6.3 provides a sufficient condition that guarantees the estimates $\tilde{\mathcal{R}}_n^*(S)$ and $\tilde{\mathcal{R}}_n^*(\langle \log S \rangle)$ converge to the same limit as n approaches infinity. As shown in Chapter 3, the estimator $\hat{\mathcal{R}}_n^*$ is robust and versatile, consistently estimating the roughness exponent under various rough volatility models, such as the rough fractional stochastic volatility model. As a result, the above theorem demonstrates that when the mesh used to compute the realized volatility is properly chosen, the estimator $\tilde{\mathcal{R}}_n^*$ is a consistent estimator for the roughness exponent for the actual roughness exponent of the rough fractional stochastic volatility model.

As a result, the estimator $\tilde{\mathcal{R}}_n^*$ addresses all the concerns outlined in Chapter 3 and estimates the roughness exponent in a strictly pathwise manner. In the future, we aim to apply $\tilde{\mathcal{R}}_n^*$ to study the roughness exponent of many high-frequency financial time series.

Before proving the above theorem, let us first sketch the idea of this proof. It follows from (6.5) that

$$\hat{Y}_{\frac{k}{2^n}}^{(n)} - \hat{Y}_{\frac{k-1}{2^n}}^{(n)} = \sum_{j=m_n(k-1)+1}^{m_n k} \left(\log S_{\frac{j}{2^n m_n}} - \log S_{\frac{j-1}{2^n m_n}} \right)^2 = \sum_{j=m_n(k-1)+1}^{m_n k} \left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^2.$$

For short, let us denote

$$\xi_j^{(n)} := \left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^2 - \int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \quad \text{and} \quad Z_k^{(n)} = \sum_{j=m_n(k-1)+1}^{m_n k} \xi_j^{(n)}.$$

We can then obtain the following decomposition

$$\begin{aligned}\widehat{Y}_{\frac{k}{2^n}}^{(n)} - \widehat{Y}_{\frac{k-1}{2^n}}^{(n)} &= \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} \sigma_s^2 ds + \sum_{j=m_n(k-1)+1}^{m_n k} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^2 - \int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right] \\ &= \left(Y_{\frac{k}{2^n}} - Y_{\frac{k-1}{2^n}} \right) + \sum_{j=m_n(k-1)+1}^{m_n k} \xi_j^{(n)} = \left(Y_{\frac{k}{2^n}} - Y_{\frac{k-1}{2^n}} \right) + Z_k^{(n)}.\end{aligned}$$

It is clear that if the quantity $Z_k^{(n)} := \sum_{j=m_n(k-1)+1}^{m_n k} \xi_j^{(n)}$ is relatively small compared with $Y_{\frac{k}{2^n}} - Y_{\frac{k-1}{2^n}}$ as $n \uparrow \infty$, then $\widehat{Y}_{\frac{k}{2^n}}^{(n)} - \widehat{Y}_{\frac{k-1}{2^n}}^{(n)}$ will be sufficiently close to the difference of the integrated volatility $Y_{\frac{k}{2^n}} - Y_{\frac{k-1}{2^n}}$, so are the values of $\vartheta_{n,k}$ and $\widetilde{\vartheta}_{n,k}$. Our next result derives the upper bounds for the second and fourth moment of the random variable $Z_k^{(n)}$, which is the content of the next lemma.

Lemma 6.4. *Suppose that $\sup_{t \in [0,1]} \sigma_t^2 \leq M$ $\mathbb{P}$ -a.s., then there exists a constant $\kappa > 0$ such that for $n \in \mathbb{N}$ and $1 \leq k \leq 2^n$, we have*

$$\mathbb{E} \left[\left(Z_k^{(n)} \right)^2 \right] = \frac{\kappa}{2^{2n} m_n} \quad \text{and} \quad \text{var} \left[\left(Z_k^{(n)} \right)^2 \right] \leq \frac{\kappa}{2^{4n} m_n^2}.$$

Proof. Clearly, we have

$$\left(Z_k^{(n)} \right)^2 = \left(\sum_{j=m_n(k-1)+1}^{m_n k} \xi_j^{(n)} \right)^2 = \sum_{j=m_n(k-1)+1}^{m_n k} \left(\xi_j^{(n)} \right)^2 + 2 \sum_{m_n(k-1)+1 \leq i < j \leq m_n k} \xi_i^{(n)} \xi_j^{(n)}.$$

Note that for $i < j$, we get

$$\mathbb{E} \left[\xi_i^{(n)} \xi_j^{(n)} \right] = \mathbb{E} \left[\mathbb{E} \left[\xi_i^{(n)} \xi_j^{(n)} \middle| \mathcal{F}_{\frac{i}{2^n m_n}} \right] \right] \mathbb{E} \left[\xi_i^{(n)} \mathbb{E} \left[\xi_j^{(n)} \middle| \mathcal{F}_{\frac{i}{2^n m_n}} \right] \right].$$

In particular, we have

$$\begin{aligned}\mathbb{E} \left[\xi_j^{(n)} \middle| \mathcal{F}_{\frac{i}{2^n m_n}} \right] &= \mathbb{E} \left[\int_0^{\frac{j}{2^n m_n}} \sigma_s dB_s - \int_0^{\frac{j-1}{2^n m_n}} \sigma_s dB_s \middle| \mathcal{F}_{\frac{i}{2^n m_n}} \right] \\ &= \int_0^{\frac{i}{2^n m_n}} \sigma_s dB_s - \int_0^{\frac{i}{2^n m_n}} \sigma_s dB_s = 0.\end{aligned} \tag{6.7}$$

Furthermore, Itô's isometry yields

$$\begin{aligned}
\mathbb{E} \left[\left(\xi_j^{(n)} \right)^2 \right] &= \mathbb{E} \left[\left(\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^2 - \int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^2 \right] \\
&= \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^4 \right] - 2 \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^2 \int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right] \\
&\quad + \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^2 \right] \\
&\leq \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^4 \right] + 2 \sqrt{\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^4 \right]} \sqrt{\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^2 \right]} \\
&\quad + \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^2 \right].
\end{aligned}$$

The Burkholder–Davis–Gundy inequality yields that there exists a constant $\kappa_1 > 0$ such that

$$\begin{aligned}
\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^4 \right] &= \mathbb{E} \left[\left(\int_0^{\frac{1}{2^n m_n}} \sigma_{s+\frac{j-1}{2^n m_n}} dB_{s+\frac{j-1}{2^n m_n}} \right)^4 \right] \\
&\leq \kappa_1 \mathbb{E} \left[\left(\int_0^{\frac{1}{2^n m_n}} \sigma_{s+\frac{j-1}{2^n m_n}}^2 ds \right)^2 \right] \\
&= \kappa_1 \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^2 \right].
\end{aligned}$$

Applying the above inequality gives

$$\begin{aligned}
\mathbb{E} \left[\left(\xi_j^{(n)} \right)^2 \right] &\leq (\sqrt{\kappa_1} + 1)^2 \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^2 \right] \\
&\leq (\sqrt{\kappa_1} + 1)^2 \frac{\mathbb{E} \left[\left(\sup_{s \in [0,1]} \sigma_s^2 \right)^2 \right]}{2^{2n} m_n^2} \leq \frac{(\sqrt{\kappa_1} + 1)^2 M}{2^{2n} m_n^2}.
\end{aligned}$$

and hence,

$$\mathbb{E} \left[\left(Z_k^{(n)} \right)^2 \right] = \frac{(\sqrt{\kappa_1} + 1)^2 M}{2^{2n} m_n}.$$

Next, we are going to analyze the fourth moment $\mathbb{E} \left[\left(Z_k^{(n)} \right)^4 \right]$. It follows from the multinomial expansion and a similar argument as in (6.7) that

$$\begin{aligned} \mathbb{E} \left[\left(Z_k^{(n)} \right)^4 \right] &= \sum_{j=m_n(k-1)+1}^{m_n k} \mathbb{E} \left[\left(\xi_j^{(n)} \right)^4 \right] + 6 \sum_{m_n(k-1)+1 \leq i < j \leq m_n k} \mathbb{E} \left[\left(\xi_i^{(n)} \right)^2 \left(\xi_j^{(n)} \right)^2 \right] \\ &\leq \sum_{j=m_n(k-1)+1}^{m_n k} \mathbb{E} \left[\left(\xi_j^{(n)} \right)^4 \right] + 6 \sum_{m_n(k-1)+1 \leq i < j \leq m_n k} \sqrt{\mathbb{E} \left[\left(\xi_i^{(n)} \right)^4 \right] \mathbb{E} \left[\left(\xi_j^{(n)} \right)^4 \right]}. \end{aligned} \quad (6.8)$$

Furthermore, it follows from the Cauchy–Schwarz inequality that

$$\begin{aligned} \mathbb{E} \left[\left(\xi_j^{(n)} \right)^4 \right] &= \mathbb{E} \left[\left(\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^2 - \int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^4 \right] \\ &\leq \sum_{i=0}^4 \binom{4}{i} \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^{8-2i} \left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^i \right] \\ &\leq \sum_{i=0}^4 \binom{4}{i} \sqrt{\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^{16-4i} \right]} \sqrt{\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^{2i} \right]} \end{aligned}$$

The Burkholder–Davis–Gundy inequality yields that there exists $\kappa_2 > 0$ such that for $0 \leq i \leq 4$,

$$\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s dB_s \right)^{16-4i} \right] \leq \kappa_2 \mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^{8-2i} \right],$$

therefore, we get

$$\begin{aligned} \mathbb{E} \left[\left(\xi_j^{(n)} \right)^4 \right] &\leq \sqrt{\kappa_2} \sum_{i=0}^4 \binom{4}{i} \sqrt{\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^{8-2i} \right]} \sqrt{\mathbb{E} \left[\left(\int_{\frac{j-1}{2^n m_n}}^{\frac{j}{2^n m_n}} \sigma_s^2 ds \right)^{2i} \right]} \\ &\leq \sqrt{\kappa_2} \sum_{i=0}^4 \binom{4}{i} \left(\frac{M}{2^n m_n} \right)^{4-i} \left(\frac{M}{2^n m_n} \right)^i = \frac{16M^4}{2^{4n} m_n^4}. \end{aligned}$$

Inserting the above inequality to (6.8) yields that

$$\text{var} \left(\left(Z_k^{(n)} \right)^2 \right) \leq \mathbb{E} \left[\left(Z_k^{(n)} \right)^4 \right] \leq 16 \left(\frac{M^4}{2^{4n} m_n^3} + \frac{6M^4}{2^{4n} m_n^2} \right).$$

This completes the proof. $\square$

Proof of Theorem 6.3. Let $\mathbf{a} = (-1, +1, +1, -1)^\top$. Applying the above equality to (6.6) yields that

$$\begin{aligned}\tilde{\vartheta}_{n,k} &= 2^{3n/2+3} \left(\widehat{Y}_{\frac{4k}{2^{n+2}}}^{(n+2)} - 2\widehat{Y}_{\frac{4k+1}{2^{n+2}}}^{(n+2)} + 2\widehat{Y}_{\frac{4k+3}{2^{n+2}}}^{(n+2)} - \widehat{Y}_{\frac{4k+4}{2^{n+2}}}^{(n+2)} \right) = 2^{3n/2+3} \sum_{i=1}^4 a_i \left(\widehat{Y}_{\frac{4k+i}{2^{n+2}}}^{(n+2)} - \widehat{Y}_{\frac{4k+i-1}{2^{n+2}}}^{(n+2)} \right) \\ &= 2^{3n/2+3} \left(\sum_{i=1}^4 a_i \left(Y_{\frac{4k+i}{2^{n+2}}} - Y_{\frac{4k+i-1}{2^{n+2}}} \right) + \sum_{i=1}^4 a_i Z_{4k+i}^{(n+2)} \right) = \vartheta_{n,k} + 2^{3n/2+3} \sum_{i=1}^4 a_i Z_{4k+i}^{(n+2)}.\end{aligned}$$

For short, let us denote $\zeta_{n,k} = 2^{3n/2+3} \sum_{i=1}^4 a_i Z_{4k+i}^{(n+2)}$. Then, by Jensen's inequality, we get

$$\zeta_{n,k}^2 = 2^{3n+6} \left(\sum_{i=1}^4 a_i Z_{4k+i}^{(n+2)} \right) \leq 2^{3n+8} \sum_{i=1}^4 \left(Z_{4k+i}^{(n+2)} \right)^2.$$

Moreover, we have

$$\begin{aligned}\mathbb{E} \left[2^{(2R-2)n} \sum_{k=0}^{2^n-1} \zeta_{n,k}^2 \right] &= 2^{(2R-2)n} \sum_{k=1}^{2^{n+2}} \mathbb{E} \left[\left(Z_{4k+i}^{(n+2)} \right)^2 \right] \\ &\leq 2^{(2R-2)n+3n+8+(n+2)} \frac{\kappa}{2^{2(n+2)} m_{n+2}} = \frac{2^{2Rn+4} \kappa}{m_{n+2}}.\end{aligned}$$

Furthermore,

$$\begin{aligned}\text{var} \left[2^{(2R-2)n} \sum_{k=0}^{2^n-1} \zeta_{n,k}^2 \right] &= 2^{(4Rn-4)n+6n+16} \sum_{k=1}^{2^{n+2}} \left(\text{var} \left[\zeta_{n,k}^2 \right] + \sum_{1 \leq i \neq j \leq 2^{n+2}} \text{cov} \left(\zeta_{n,k}^2, \zeta_{n,j}^2 \right) \right) \\ &\leq 2^{(4Rn+2)n+12} \sum_{k=1}^{2^{n+2}} \left(\text{var} \left[\zeta_{n,i}^2 \right] + \sum_{1 \leq i \neq j \leq 2^{n+2}} \sqrt{\text{var} \left[\zeta_{n,i}^2 \right] \text{var} \left[\zeta_{n,j}^2 \right]} \right) \\ &\leq 2^{(4Rn+2)n+12} 2^{2(n+2)} \frac{\kappa}{2^{4(n+2)} m_{n+2}^2} = \frac{2^{4Rn+4} \kappa}{m_{n+2}^2}.\end{aligned}$$

As $\lim_{n \uparrow \infty} 2^{-2Rn} m_n = \infty$, we have $\mathbb{E} \left[2^{(2R-2)n} \sum_{k=0}^{2^n-1} \zeta_{n,k}^2 \right]$ and $\text{var} \left[2^{(2R-2)n} \sum_{k=0}^{2^n-1} \zeta_{n,k}^2 \right]$ converge to zero as $n \uparrow \infty$. Thus, we have

$$2^{(2R-2)n} \sum_{k=0}^{2^n-1} \zeta_{n,k}^2 \longrightarrow 0 \quad \mathbb{P}\text{-a.s.},$$

by the fast L_2 -convergence. Finally, note that

$$2^{(2R-2)n} \sum_{k=0}^{2^n-1} \tilde{\vartheta}_{n,k}^2 = 2^{(2R-2)n} \left(\sum_{k=0}^{2^n-1} \vartheta_{n,k}^2 + \sum_{k=0}^{2^n-1} \zeta_{n,k}^2 + 2 \sum_{k=0}^{2^n-1} \vartheta_{n,k} \zeta_{n,k} \right),$$

and it remains to show that the third argument on the right-hand side of the above equation converges to zero as $n \uparrow \infty$. The Cauchy–Schwarz inequality yields

$$\limsup_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \vartheta_{n,k} \zeta_{n,k} \leq \sqrt{\limsup_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \vartheta_{n,k}^2} \sqrt{\lim_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \zeta_{n,k}^2} = 0.$$

Thus, we have

$$\liminf_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \tilde{\vartheta}_{n,k}^2 = \liminf_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \vartheta_{n,k}^2,$$

and

$$\limsup_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \tilde{\vartheta}_{n,k}^2 = \limsup_{n \uparrow \infty} 2^{(2R-2)n} \sum_{k=0}^{2^n-1} \vartheta_{n,k}^2.$$

Taking logarithms on both sides of the above identities, dividing by n and letting $n \uparrow \infty$ complete the proof. $\square$

References

- [1] Pieter C. Allaart. On a flexible class of continuous functions with uniform local structure. *Journal of the Mathematical Society of Japan*, 61(1):237–262, 2009.
- [2] Pieter C. Allaart and Kiko Kawamura. The Takagi function: a survey. *Real Analysis Exchange*, 37(1):1–54, 2011.
- [3] Elisa Alòs, Jorge A León, and Josep Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11(4):571–589, 2007.
- [4] Jürgen Appell, Józef Banaś, and Nelson Merentes. *Bounded variation and around*, volume 17 of *De Gruyter Series in Nonlinear Analysis and Applications*. De Gruyter, Berlin, 2014.
- [5] Fabrice Baudoin and Martin Hairer. A version of Hörmander’s theorem for the fractional Brownian motion. *Probability Theory and Related Fields*, 139(3-4):373–395, 2007.
- [6] Christian Bayer, Peter Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016.
- [7] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. Decoupling the short-and long-term behavior of stochastic volatility. *Journal of Financial Econometrics*, 20(5):961–1006, 2022.
- [8] Patrick Billingsley. Van der Waerden’s continuous nowhere differentiable function. *American Mathematical Monthly*, 89:691, 1982.
- [9] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.
- [10] Anine E Bolko, Kim Christensen, Mikko S Pakkanen, and Bezirgen Veliyev. A GMM approach to estimate the roughness of stochastic volatility. *Journal of Econometrics*, 235(2):745–778, 2023.

- [11] Peter Carr and Liuren Wu. The finite moment log stable process and option pricing. *Journal of Finance*, 58(2):753–777, 2003.
- [12] Patrick Cheridito. *Regularizing fractional Brownian motion with a view towards stock price modelling*. PhD thesis, ETH Zürich, 2001.
- [13] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima. Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability*, 8:1–14, 2003.
- [14] Carsten Chong, Marc Hoffmann, Yanghui Liu, Mathieu Rosenbaum, and Grégoire Szymanski. Statistical inference for rough volatility: Central limit theorems. *arXiv preprint arXiv:2210.01216*, 2022.
- [15] Carsten Chong, Marc Hoffmann, Yanghui Liu, Mathieu Rosenbaum, and Grégoire Szymanski. Statistical inference for rough volatility: Minimax Theory. *arXiv preprint arXiv:2210.01214*, 2022.
- [16] Z. Ciesielski, G. Kerkycharian, and B. Roynette. Quelques espaces fonctionnels associés à des processus gaussiens. *Studia Mathematica*, 107(2):171–204, 1993.
- [17] Rama Cont and Purba Das. Quadratic variation along refining partitions: Constructions and examples. *Journal of Mathematical Analysis and Applications*, 512(2):126173, 2022.
- [18] Rama Cont and Purba Das. Quadratic variation and quadratic roughness. *Bernoulli*, 29(1):496–522, 2023.
- [19] Rama Cont and Purba Das. Rough volatility: fact or artefact? *Sankhya B*, pages 1–33, 2024.
- [20] Rama Cont and Nicolas Perkowski. Pathwise integration and change of variable formulas for continuous paths with arbitrary regularity. *Transactions of the American Mathematical Society. Series B*, 6:161–186, 2019.
- [21] Purba Das. *Roughness properties of paths and signals*. PhD thesis, University of Oxford, 2022.
- [22] Jörn Davidsen and James Griffin. Volatility of unevenly sampled fractional brownian motion: An application to ice core records. *Physical Review E*, 81(1):016107, 2010.
- [23] Laurens de Haan and Ana Ferreira. *Extreme value theory. An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006.
- [24] Georges de Rham. Sur un exemple de fonction continue sans dérivée. *L’Enseignement Mathématique*, 3:71–72, 1957.

- [25] Rick Durrett. *Probability: theory and examples*. Brooks/Cole—Thomson Learning, Belmont, CA, third edition, 2005.
- [26] Dorin Ervin Dutkay and Palle E. T. Jorgensen. Harmonic analysis and dynamics for affine iterated function systems. *Houston Journal of Mathematics*, 33(3):877–905, 2007.
- [27] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38, 2019.
- [28] Omar El Euch and Mathieu Rosenbaum. Perfect hedging in rough Heston models. *The Annals of Applied Probability*, 28(6):3813–3856, 2018.
- [29] Gregory Eyink. Besov spaces and the multifractal hypothesis. *Journal of Statistical Physics*, 78:353–375, 1995.
- [30] G. Faber. Einfaches Beispiel einer stetigen nirgends differenzierbaren Funktion. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 16:538–540, 1907.
- [31] Hans. Föllmer. Calcul d’Itô sans probabilités. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980)*, volume 850 of *Lecture Notes in Math.*, pages 143–150. Springer, Berlin, 1981.
- [32] Hans Föllmer. Probabilistic aspects of financial risk. In *European Congress of Mathematics, Vol. I (Barcelona, 2000)*, volume 201 of *Progr. Math.*, pages 21–36. Birkhäuser, Basel, 2001.
- [33] Hans Föllmer and Alexander Schied. Probabilistic aspects of finance. *Bernoulli*, 19(4):1306–1326, 2013.
- [34] Hans Föllmer and Alexander Schied. *Stochastic finance. An introduction in discrete time*. De Gruyter Graduate. De Gruyter, Berlin, fourth revised and extended edition, 2016.
- [35] Martin Forde, Masaaki Fukasawa, Stefan Gerhold, and Benjamin Smith. The Riemann–Liouville field and its GMC as H^0 , and skew flattening for the rough Bergomi model. *Statistics & Probability Letters*, 181:109265, 2022.
- [36] Masaaki Fukasawa. Short-time at-the-money skew and rough fractional volatility. *Quantitative Finance*, 17(2):189–198, 2017.
- [37] Masaaki Fukasawa. Volatility has to be rough. *Quantitative Finance*, 21(1):1–8, 2021.

- [38] Masaaki Fukasawa, Tetsuya Takabatake, and Rebecca Westphal. Consistent estimation for fractional stochastic volatility model under high-frequency asymptotics. *Mathematical Finance*, 32(4):1086–1132, 2022.
- [39] Nina Gantert. Self-similarity of Brownian motion and a large deviation principle for random fields on a binary tree. *Probability Theory and Related Fields*, 98(1):7–20, 1994.
- [40] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.
- [41] E. G. Gladyshev. A new limit theorem for stochastic processes with Gaussian increments. *Theory of Probability and Its Applications*, 6:57–66, 1961.
- [42] Tilmann Gneiting and Martin Schlather. Stochastic models that separate fractal dimension and the Hurst effect. *SIAM Review*, 46(2):269–282, 2004.
- [43] Paul Richard Halmos. *A Hilbert space problem book*, volume 19. Springer Science & Business Media, 2012.
- [44] Xiyue Han. A Gladyshev theorem for trifractional Brownian motion and n -th order fractional Brownian motion. *Electronic Communications in Probability*, 26:1–12, 2021.
- [45] Xiyue Han and Alexander Schied. The roughness exponent and its model-free estimation. *arXiv:2111.10301*, 2021.
- [46] Xiyue Han and Alexander Schied. Robust Faber–Schauder approximation based on discrete observations of an antiderivative. *arXiv:2211.11907*, 2022.
- [47] Xiyue Han and Alexander Schied. Estimating the roughness exponent of stochastic volatility from discrete observations of the realized variance. *arXiv:2307.02582*, 2023.
- [48] Xiyue Han and Alexander Schied. On laws absolutely continuous with respect to fractional Brownian motion. *arXiv:2306.11824*, 2023.
- [49] Xiyue Han, Alexander Schied, and Zhenyuan Zhang. A probabilistic approach to the Φ -variation of classical fractal functions with critical roughness. *Statistics & Probability Letters*, 168:108920, 2021.
- [50] Xiyue Han, Alexander Schied, and Zhenyuan Zhang. A limit theorem for Bernoulli convolutions and the ϕ -variation of functions in the Takagi class. *Journal of Theoretical Probability*, 35(4):2853–2878, 2022.
- [51] G. H. Hardy. Weierstrass’s non-differentiable function. *Transactions of the American Mathematical Society*, 17(3):301–325, 1916.

- [52] Masayoshi Hata and Masaya Yamaguti. The Takagi function and its generalization. *Japan Journal of Applied Mathematics*, 1(1):183–199, 1984.
- [53] Harold Edwin Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1):770–799, 1951.
- [54] J Ingtem. Minimal-norm-derivative spline function in interpolation and approximation. *Moscow University Computational Mathematics and Cybernetics*, 32(4):201–213, 2008.
- [55] Antoine Jacquier, Claude Martini, and Aitor Muguruza. On vix futures in the rough Bergomi model. *Quantitative Finance*, 18(1):45–61, 2018.
- [56] Thibault Jaisson and Mathieu Rosenbaum. Limit theorems for nearly unstable Hawkes processes. *Annals of Applied Probability*, 25(2):600–631, 2015.
- [57] Thibault Jaisson and Mathieu Rosenbaum. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *The Annals of Applied Probability*, 26(5):2860, 2016.
- [58] Galin L. Jones. On the Markov chain central limit theorem. *Probability Surveys*, 1:299–320, 2004.
- [59] T. T. Kadota and L. A. Shepp. Conditions for absolute continuity between a certain pair of probability measures. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 16:250–260, 1970.
- [60] Ruben Klein and Evarist Giné. On quadratic variation of processes with Gaussian increments. *The Annals of Probability*, 3(4):716–721, 1975.
- [61] Achim Klenke. *Probability theory*. Universitext. Springer, London, second edition, 2014.
- [62] Norio Kôno. On generalized Takagi functions. *Acta Mathematica Hungarica*, 49(3-4):315–324, 1987.
- [63] Kestutis Kubilius. private communication. 2021.
- [64] Kestutis Kubilius and Dmitrij Melichov. On the convergence rates of Gladyshev’s Hurst index estimator. *Nonlinear Analysis. Modelling and Control*, 15(4):445–450, 2010.
- [65] Kestutis Kubilius, Yuliya Mishura, and Kostiantyn Ralchenko. *Parameter estimation in fractional diffusion models*, volume 8. Springer, 2017.

- [66] Jeffrey C. Lagarias. The Takagi function and its properties. In *Functions in number theory and their probabilistic aspects*, RIMS Kôkyûroku Bessatsu, B34, pages 153–189. Res. Inst. Math. Sci. (RIMS), Kyoto, 2012.
- [67] Pedro Lei and David Nualart. A decomposition of the bifractional Brownian motion and some applications. *Statistics & Probability Letters*, 79(5):619–624, 2009.
- [68] Robert S. Liptser and Albert N. Shiryaev. *Statistics of random processes. I*, volume 5 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, expanded edition, 2001. General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.
- [69] Laszlo Losonczi. Eigenvalues and eigenvectors of some tridiagonal matrices. *Acta Mathematica Hungarica*, 60(3-4):309–322, 1992.
- [70] Terry J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.
- [71] Chunsheng Ma. The Schoenberg–Lévy kernel and relationships among fractional Brownian motion, bifractional Brownian motion, and others. *Theory of Probability and Its Applications*, 57(4):619–632, 2013.
- [72] Michael B Marcus and Jay Rosen. p -variation of the local times of symmetric stable processes and of Gaussian processes with stationary increments. *The Annals of Probability*, pages 1685–1713, 1992.
- [73] Viktor Anatolievich Matveev. Series in a Schauder system. *Matematicheskie Zametki*, 2(3):267–278, 1967.
- [74] Holger Mettke, Eckehard Pfeifer, and Edward Neuman. Quadratic spline interpolation with coinciding interpolation and spline grids. *Journal of Computational and Applied Mathematics*, 8(1):57–62, 1982.
- [75] Yuliya Mishura. *Stochastic calculus for fractional Brownian motion and related processes*, volume 1929 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008.
- [76] Yuliya Mishura and Alexander Schied. Constructing functions with prescribed pathwise quadratic variation. *Journal of Mathematical Analysis and Applications*, 442(1):117–137, 2016.
- [77] Yuliya Mishura and Alexander Schied. On (signed) Takagi–Landsberg functions: p th variation, maximum, and modulus of continuity. *Journal of Mathematical Analysis and Applications*, 473(1):258–272, 2019.

- [78] Marian Mureşan. *A concrete approach to classical analysis*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2009.
- [79] I. P. Natanson. *Theory of functions of a real variable*. Frederick Ungar Publishing Co., New York, 1955. Translated by Leo F. Boron with the collaboration of Edwin Hewitt.
- [80] Ilkka Norros, Esko Valkeila, and Jorma Virtamo. An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motions. *Bernoulli*, 5(4):571–587, 1999.
- [81] Rimas Norvaiša. A complement to Gladyshev’s theorem. *Lithuanian Mathematical Journal*, 51(1):26–35, 2011.
- [82] Yuval Peres, Wilhelm Schlag, and Boris Solomyak. Sixty years of Bernoulli convolutions. In *Fractal geometry and stochastics, II (Greifswald/Koserow, 1998)*, volume 46, pages 39–65. Birkhäuser, Basel, 2000.
- [83] Emmanuel Perrin, Rachid Harba, Corinne Berzin-Joseph, Ileana Iribarren, and Aline Bonami. n th-order fractional Brownian motion and fractional Gaussian noises. *IEEE Transactions on Signal Processing*, 49(5):1049–1059, 2001.
- [84] Jean Picard. Representation formulae for the fractional Brownian motion. *Séminaire de probabilités XLIII*, pages 3–70, 2011.
- [85] Mathieu Rosenbaum. First order p -variations and Besov spaces. *Statistics and Probability Letters*, 79(1):55–62, 2009.
- [86] Stefan G. Samko, Anatoly A. Kilbas, and Oleg I. Marichev. *Fractional integrals and derivatives*. Gordon and Breach Science Publishers, Yverdon, 1993.
- [87] Alexander Schied. On a class of generalized Takagi functions with linear pathwise quadratic variation. *Journal of Mathematical Analysis and Applications*, 433:974–990, 2016.
- [88] Alexander Schied and Iryna Voloshchenko. Pathwise no-arbitrage in a class of Delta hedging strategies. *Probability, Uncertainty and Quantitative Risk*, 1, 2016.
- [89] Alexander Schied and Zhenyuan Zhang. On the p -th variation of a class of fractal functions. *Proceedings of the American Mathematical Society*, 148(12):5399–5412, 2020.
- [90] Alexander Schied and Zhenyuan Zhang. Weierstrass bridges. *Transactions of the American Mathematical Society*, 377(04):2947–2989, 2024.

- [91] Tommi Sottinen and Lauri Viitasaari. Transfer principle for n -th order fractional Brownian motion with applications to prediction and equivalence in law. *Theory of Probability and Mathematical Statistics*, 98:199–216, 2019.
- [92] Teiji Takagi. A simple example of the continuous function without derivative. In *Proceedings of the Physico-Mathematical Society of Japan*, volume 1, pages 176–177, 1903.
- [93] Hans Triebel. *Theory of Function Spaces*. Modern Birkhäuser Classics, 1983.
- [94] Lauri Viitasaari. Necessary and sufficient conditions for limit theorems for quadratic variations of Gaussian sequences. *Probability Surveys*, 16:62–98, 2019.
- [95] Karl Weierstraß. Ueber continuirliche Functionen eines reellen Arguments, die für keinen Wert des letzteren einen bestimmten Differentialquotienten besitzen. In *Mathematische Werke. Herausgegeben unter Mitwirkung einer von der Königlich Preussischen Akademie der Wissenschaften eingesetzten Commission. Zweiter Band. Abhandlungen II.*, pages 71–74, Berlin, 1895. Mayer & Müller.
- [96] M.P. Yershov. On absolute continuity of measures corresponding to diffusion type processes. *Theory of Probability and Its Applications*, 17(1):173–178, 1972.
- [97] Ou Zhao, Michael Woodroffe, and Dalibor Volný. A central limit theorem for reversible processes with nonlinear growth of variance. *Journal of Applied Probability*, 47(4):1195–1202, 2010.